https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

A spatially explicit dataset of agriculture liming across the contiguous United States

Samuel Shou-En Tsao[1], Tim Jesper Surhoff[2,3], Giuseppe Amatulli[1], Maya Almarez[1], Jon Gewirtzman[1], Beck Woollen[3], Eric W. Slessarev[4], James E. Saiers[1], Christopher T. Reinhard[6], Shuang Zhang[5], Noah J. Planavsky[3], Peter A. Raymond[1]

[1] School of the Environment, Yale University, New Haven, CT 06511, USA

[2] Yale Center for Natural Carbon Capture, Yale University, New Haven, CT 06511, USA

[3] Department of Earth and Planetary Sciences, Yale University, New Haven, CT 06511, USA

[4] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06511, USA

[5] Department of Oceanography, Texas A&M University, College Station, TX 77843, USA

[6] School of Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

Correspondence to: Samuel Shou-En Tsao (samuel.tsao@yale.edu)

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

Abstract

Agricultural lime has historically been applied to croplands in the United States to counteract soil acidification and enhance soil fertility, with important consequences for crop productivity and Earth's carbon cycle. Previous work on agricultural liming has largely focused on either region-specific case studies or national-level estimates of total application rates, leaving a major gap in understanding the spatial variability in lime application. This study addresses this gap by presenting the first spatially explicit dataset of agricultural lime application across the contiguous United States. The dataset comprises state-level data for 1930–1950 and a more detailed county-level dataset for 1954–1987, enabling comprehensive spatial-temporal analyses at multiple scales. Counties in the Midwest region exhibited the highest total amounts of lime applied in the latter half of the twentieth century, reflecting intensive agricultural activity. These counties were characterized by higher overall lime application rates (amount of lime applied per unit of limed area each year) but relatively lower liming frequency (ratio of limed area to total agricultural land area each year). In contrast, counties in the southeastern coastal region exhibited lower lime application rates per unit of limed area but more frequent lime applications. We used a machine learning framework, to elucidate key environmental and agricultural drivers of lime application. Our results show that the total amount of lime applied, as well as the application rate and frequency, are strongly associated with regional climatic conditions and soil properties. However, we also found evidence that agricultural management practices (such as crop production, fertilizer use, and soil pH recommendations) played a key role in shaping liming applications. Spatiotemporal integration of the data product results in a revised national estimate of total lime application, with a range of 15-25 million tons (Mt) per year. This study establishes a critical observational baseline for assessing the potential of agricultural lime application as a climate mitigation strategy and highlights the need for further research into its long-term environmental impacts.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

1 Introduction

Agricultural lime has been historically applied to the surface of cropland to enhance soil pH and increase yields (Tennant, 1799; Johnston, 1849). Soils can naturally acidify through a range of processes, including atmospheric deposition of strong acids, nitrification, sulfur oxidation, and release of organic acids by roots, microorganisms, and decomposing organic matter (Goulding, 2016). In croplands, soil acidification is intensified by repeated addition of nitrogen fertilizers, which promotes the release of hydrogen ions (H^+) via nitrification and by removal of base cations in crop biomass (Guo et al., 2010). Approximately 50% of global arable land is considered acidic (pH < 5.5) (von Uexküll & Mutert, 1995). As soil pH decreases, excess hydrogen ions replace base cations such as calcium (Ca^{2+}) and magnesium (Mg^{2+}), causing their leaching and leading to a relative abundance of toxic elements like aluminum (Al^{3+}), which can be detrimental to crop growth (Fageria & Baligar, 2008). In response, where access is available lime application has become a common strategy to mitigate the issue of soil acidification in croplands where access is available. When lime is applied to soil, the Ca^{2+} and Mg^{2+} it releases can displace H^+, effectively reducing soil acidity (Hooda & Alloway, 1996). Repeated nitrogen application and harvest increase soil acidity, prompting periodic lime applications to restore soil pH towards neutral (pH ~7). In addition to counteracting soil pH decline, liming also replenishes key nutrients like Ca^{2+}, Mg^{2+}, and potassium (K) (Han et al., 2019), improves soil fertility and crop yield (Pagani & Mallarino, 2012), and has become an essential practice in agriculture management in the U.S. (Agegnehu et al., 2021).

When lime is applied to soils, it can be weathered by strong acids, such as nitric acid (HNO_3^-) produced from fertilizers, or by the weaker carbonic acid (H_2CO_3), which may originate from the dissolution of atmospheric CO_2 or respired organic carbon (Semhi et al., 2000). The relative contribution of strong acid derived acidity and carbonic acid depends on the pH of the soil (Plummer & Wigley, 1976). Depending on soil pH, the bicarbonate ion produced from weathering also, through equilibration of the carbonic acid system, converts to carbonic acid and leads to CO_2 source to the atmosphere. At higher pH values, however, carbonate weathering can act as a CO_2 sink, even on short timescales. 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). However, West & McBride (2005) estimated that only 38% of lime is dissolved by strong acids. Subsequent studies have revised this estimate to be even lower, with reported fractions ranging from 14% (Oh & Raymond, 2006) to 25% (Hamilton et al., 2007). These findings suggest that most of the lime reacts with carbonic acid (H_2CO_3), forming bicarbonate (HCO_3^-) in soil water. Nonetheless, previous estimates are derived from conceptual models or localized studies, underscoring the need for a spatially explicit dataset on agricultural lime applications to better quantify the fraction of strong acid weathering and the carbon removal

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

potential of lime across diverse regions.

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). Current ERW trials predominantly emphasize silicate rocks, in part because carbonate amendments like lime have historically been viewed as net CO_2 sources. However, carbonates weather significantly faster than silicates and contribute substantially to carbon cycling and sequestration on shorter timescales (<3,000 years) (Gaillardet et al., 1999; Liu et al., 2011). Enhanced carbonate weathering, particularly through large-scale agricultural liming, could offer a promising and more immediate approach to atmospheric CO_2 mitigation given its current widespread use. Using modelling approaches, Zeng et al. (2022) estimated that application of lime to its maximum extent globally could potentially sequester up to ~ 0.9 Gt of carbon per year, and Zhang et al. (2022) shows that rivers have a generally high a capacity to transport this additional sequestered carbon to the ocean (Knapp & Tipper, 2022). Furthermore, the application of lime has been shown to reduce the emission of nitrous oxide (N_2O) and methane (CH_4) emissions from soils (Khaliq et al., 2019; Shaaban et al., 2020), due to changes in the biotic and abiotic environment that are tied to soil pH (Zhang et al., 2022). Since lime can act as both a carbon source and a sink for atmospheric CO_2, 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 spatial understanding of where it can be marketed as a robust carbon removal practice. A more detailed understanding of historical lime applications across regions would serve as a valuable observational baseline to understand its climate impact in different spatial regimes and further explore the potential of enhanced lime application as a climate solution.

West & McBride (2005) provided the first national estimate of agricultural lime applied in the continental U.S., reporting a peak of approximately 35 Tg C/year during the 1970s. However, their data included only a time-series of national total estimates, lacking spatial detail. To address this research gap, this study aims to

1. Compile a spatially explicit liming dataset (state-level for the period 1930–1950 and at the county-level for 1950–1990) for the U.S.
2. Use machine learning approaches to identify key agricultural and environmental factors that explain the spatial variation in lime applications.

We find that liming has historically occurred almost exclusively in the eastern half of the United States, which is characterized by more acidic soils. The spatial variation of lime application has been further shaped by variation in agricultural practices, climate, and soil properties. Our spatially explicit, multi-decadal dataset enables improved assessment of liming's agricultural and climate impacts across the U.S.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

2. Data and Methods

2.1 Liming data

We compiled state-level liming data from 1930 to 1950 (Mehring et al. 1957) and county-level liming and agricultural data from 1954 to 1987 (U.S. Bureau of the Census, 1946, 1956, 1961, 1967, 1977, 1981, 1984, 1989) as tabulated within ICPSR dataset at the University of Michigan (Haines et al. 2018; https://www.icpsr.umich.edu/web/ICPSR/studies/35206). Liming variables, including 'lime (tons)' and 'lime (acres)', were only available for the years 1954, 1959, 1964, 1969, 1974, 1978, 1982, and 1987. In each census years, there are some counties that did not report liming-related data. We will address the treatment of these missing values in the following section on spatiotemporal interpolation. Lime application data was not reported in Census records prior to 1954, and after 1987, it was aggregated as 'area of fertilizer and lime application'. Our dataset nevertheless covers a critical period when national agricultural lime application reached its peak during the 1970s (West & McBride, 2005). We convert all the units of liming variables to metric tons (t) and hectares (ha) in the calculations, visualizations, and final data output.

2.2 Environmental and agronomic variables

We compiled a comprehensive suite of environmental parameters, including agricultural metrics, climate conditions, soil properties, and lithological features, to further understand the potential controls of the spatial variation of agricultural liming (Table S1). Agricultural data on specific crop production (e.g., corn, soybeans, wheat, cotton) and land use (e.g., cropland harvested, cropland pastured, irrigated acreage) were also collected from the Census reports. County-level nutrient data from fertilizers and manure are extracted from Falcone (2021) (https://pubs.usgs.gov/publication/ofr20201153). Averaged precipitation and temperature data (1978-1987) were sourced from the PRISM Climate Group (PRISM Climate Group, Oregon State University, 2014; accessed March 20, 2025). Soil properties (i.e., cation exchange capacity and soil carbonate stocks as $CaCO_3$ equivalents) were obtained from re-gridded SSURGO data obtained from Walkinshaw et al. (2023) (https://casoilresource.lawr.ucdavis.edu/soil-properties/). Lithological data, specifically the occurrence of carbonate sedimentary rocks, were sourced from Hartman & Moosdorf (2012) (https://www.geo.uni-hamburg.de/en/geologie/forschung/aquatische-geochemie/glim.html), as these rocks are assumed to weather more rapidly than silicates and thus influence soil acidity. We adapted the county-level soil pH recommendation from a state-level agronomic extension of pH recommendations based on the most commonly grown crop in each U.S. county using CroplandCROS (https://croplandcros.scinet.usda.gov). Raster-format datasets were spatially aggregated by averaging pixel values within the geographic boundaries of each county. This process assigns a single representative value to each county, ensuring that all raster-based variables are consistently aligned with the county-level analysis. The agricultural data corresponds to the specific years

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

of the lime application records, while the climate, soil, lithology, and pH recommendations lack temporal resolution and are treated as static across all years. To help explain the temporal variability observed across different years, we also extracted information on fertilizer application area and costs from the Census reports, as well as farm expenditures and income.

2.3 Spatial temporal interpolation

For visualization and analysis, the county-level data were joined with the U.S. county-level shapefile (https://www.census.gov/cgi-bin/geo/shapefiles/index.php). We excluded counties outside the contiguous United States (e.g., Alaska, Hawaii). In each census year, there were some counties that did not report data on lime application. The counties with missing data for each year (Fig. 1) were addressed using linear interpolation based on values from the preceding and following census years within the 1954–1987 time range. In cases where data were available only for the preceding year (e.g., the final year in the time series), missing values were extrapolated using the previous value. After applying temporal interpolation, a number of independent cities in the shapefile, including Baltimore, St. Louis, and several independent cities in Virginia, remain 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 agricultural land, it is reasonable to assume that lime was not applied in these areas (see SI for further details on data processing). Notably, most counties missing original lime application data are situated in areas with limited agricultural activity, particularly in the arid and desert regions of the western United States (Fig. 1). Since these missing gaps are not located in major liming regions, the total lime application estimates are only slightly affected by the number of missing counties each year and the interpolation procedure (Fig. S1).

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

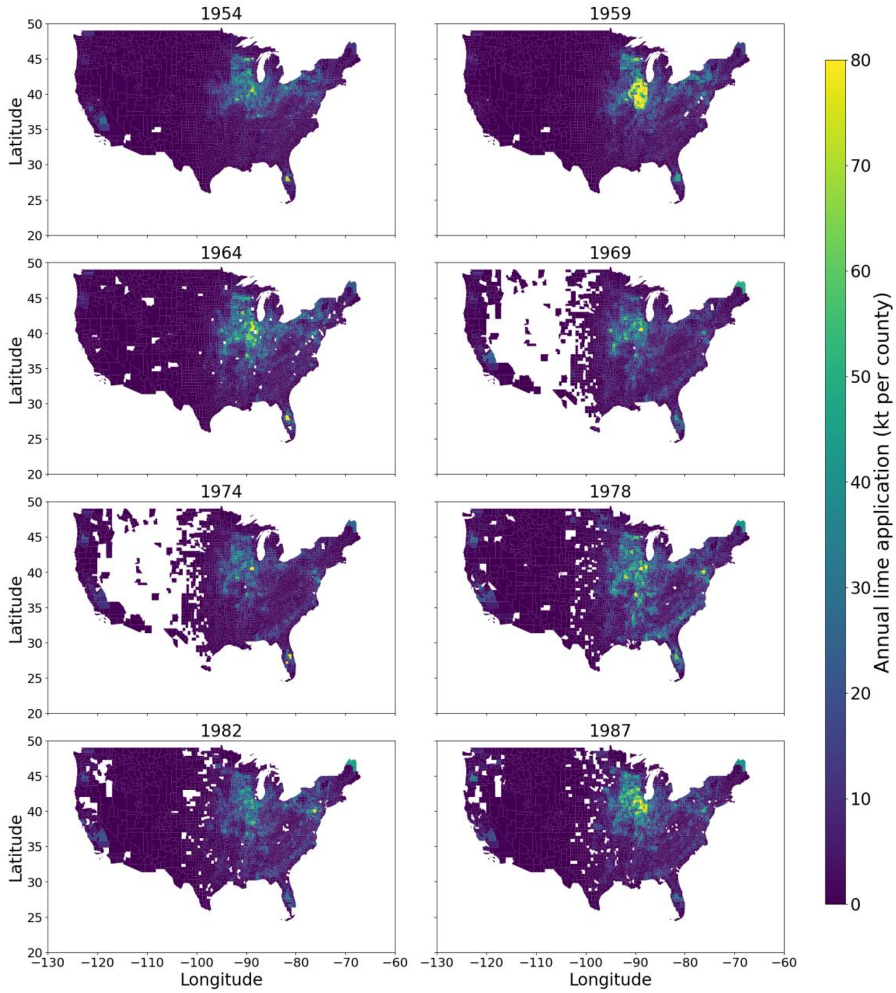

Figure 1 Lime application (kilotons, kt) by county in the continental U.S. from 1954 to 1987. The white gaps indicate counties that originally reported missing data, before temporal interpolation and spatial filling were applied. The eight maps correspond to the census years between 1954 and 1987.

We conducted the data processing mentioned above (temporal interpolation and spatial gap filling) for both lime applied weight (kt) (Fig. 4) and limed area (ha) (Fig. 5) and all the other agricultural metrics (i.e., cropland area, crop production, nutrient input) for each county. This resulted in a complete dataset with no missing value, enabling us to conduct further derivation and statistical analysis. We were then able to derive the lime application rate (lime applied weight/limed area) and liming frequency (limed area/cropland area) and further analyze the main environmental controls on each variable.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

2.4 Statistical analysis

2.4.1 Pearson correlation and variable selection

To investigate the key factors controlling the spatial distribution of lime application, we used the average across three census years (1977, 1982, and 1987), representing the most recent decade with overlapping liming and environmental data. This period offers a more stable and representative snapshot of liming conditions and is closer to modern practices. We began with a comprehensive set of agronomic, soil, and climatic variables (Table S1) and conducted a Pearson correlation analysis to evaluate linear relationships between each predictor and lime application. To mitigate multicollinearity, we also examined pairwise correlations among predictors and removed variables from highly correlated pairs (Pearson's $r > 0.90$), including *Harvested crop land (acres)* vs. *Total Cropland (acres)* ($r = 0.96$), *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, data availability, and predictive strength, we retained one variable from each pair and removed the others. Additionally, 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. The final list of features also excluded: *Harvested crop land (acres)*, *P manure (kg)*, *P fertilizer (kg)*, *Silt (%)*, *Total Cropland (acres)*, 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 (see SI, Table S2). This step ensured a more robust and interpretable analysis of environmental controls.

2.4.2 Random Forest Analysis

We employed a machine learning technique to account for complex relationships between lime applications and the predictor variables. Random Forest (RF) is a commonly used machine learning algorithm that can be used in both classification or regression by constructing an ensemble of decision trees to capture complex interactions and non-linear dependencies between predictor variables and the target variable (Breiman, 2001). This approach is particularly suited for modelling in an environmental or agricultural context characterized by high-dimensional, heterogeneous relationships (Shen et al., 2020; Burdett & Wellen, 2022; Siqueira et al., 2024; Jeong et al., 2016). In the context of our study, the RF model is not intended to extrapolate temporal trends beyond the observed data range. Instead, it is used to identify and quantify the relative importance of key environmental and agricultural factors that may influence lime application patterns.

We implemented the model using the RandomForestRegressor class derived from bootstrapped datasets (Pedregosa, 2011). We performed a grid search across multiple combinations of hyperparameters using 5-fold cross-validation. Specifically, we tested: number of trees (n_estimators) = 100, 300, 500; maximum depth (max_depth) = 10, 20, and unlimited (None); minimum samples to split (min_samples_split) = 2 or

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

5; minimum samples per leaf (min_samples_leaf) = 1 or 2; and number of features considered at each split (max_features) = 'sqrt' or 0.5. Despite testing this range of values, model performance was relatively stable, with test R^2 scores varying modestly across configurations. All models were implemented using scikit-learn (https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html). To validate the model, we randomly split our dataset (approximately 3000 counties in the contiguous USA) into 5 equal subsets, known as *folds*. We then performed 5-fold cross-validation, where each iteration trained the model on 4 of these folds and used the remaining fold as a testing set to evaluate model performance. Each fold served once as the testing set during the 5-fold cross-validation. The best-performing model was selected based on the highest mean cross-validated R^2 and used for further interpretation, as our primary objective was to assess the relative importance of input features.

Random Forest models further provide an assessment of variable importance using permutation importance, which quantifies the relative contribution of each predictor variable by measuring the decrease in model performance when the variable's values are randomly shuffled (Altmann et al. 2010). These scores were averaged across the five folds to ensure a robust estimation of variable influence. In addition, partial dependence analyses offer insights into the marginal effect of each predictor on the target outcome, holding all other variables constant. This allows visualization of non-linear relationships and interaction effects between predictors and target outcome. The combination of linear (Pearson correlation) and non-linear (Random Forest model) approaches provides a more comprehensive view of the environmental factors influencing lime application, revealing the dominant drivers of the spatial variability in lime application. We applied this statistical analysis to three response variables: (1) annual lime applied (metric tons per year, normalized by hectare of land area), (2) annual lime application rate (metric tons per hectare of limed area per year), and (3) liming frequency (ratio of limed area to total cropland area per year) in each county.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

3. Results

3.1 Spatial-temporal pattern of lime application

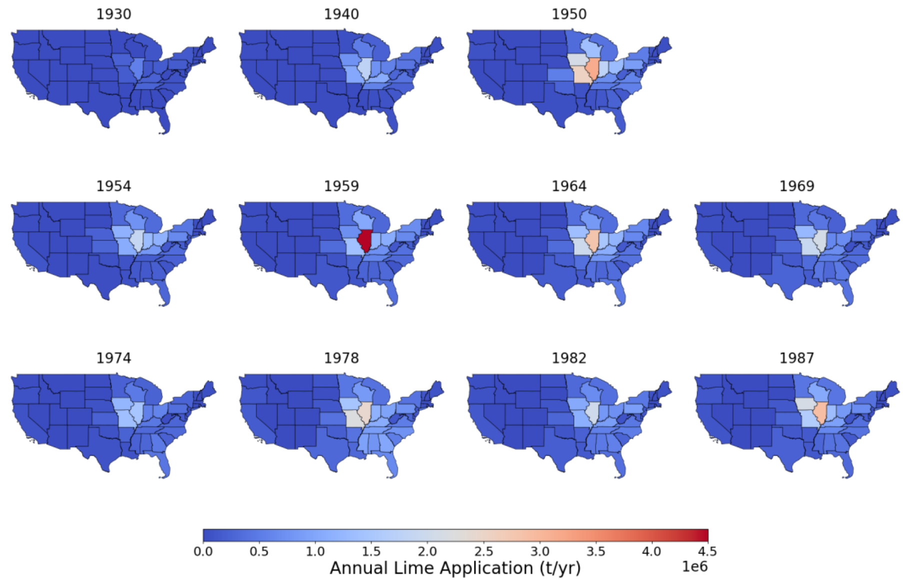

Figure 2 Annual lime application (metric tons per year, t/yr) per state in the contiguous U.S. from 1930 to 1987. The data from 1930 to 1950 is from the state-level data from the USDA report, while the data between 1954-1987 is summed from the county-level data from the Census reports. The data represents the condition during each specific year, not including the period in between.

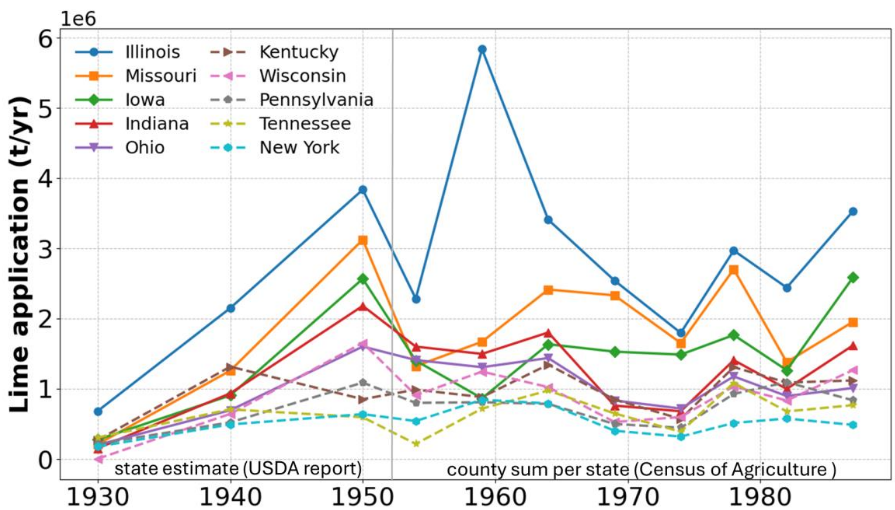

Figure 3 Temporal variations in lime applied (metric tons per year, t/yr) per state in the

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

contiguous U.S. from 1930 to 1987. The data from 1930 to 1950 is from the state-level data from the USDA report, while the data between 1954–1987 is summed from the county-level data from the Census reports. The figure shows the top 10 states ranked by state-level total lime application during 1954–1987, with Illinois applying the most lime, followed by Missouri, Iowa, Indiana, Ohio, Kentucky, Wisconsin, Pennsylvania, Tennessee, and Georgia.

The growth in state-level estimates of lime use from 1930–1987 was disproportionately concentrated in the Midwest (Fig. 1). 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 increasing trend does not persist after 1950, and the total amount of lime applied exhibits considerable temporal variation (Fig. 2). In the early 1950s, most of the lime was applied in the Midwest states. In 1959 an abrupt increase of lime is shown in Illinois (exceeding 6 Mt lime year^{-1}), which could potentially be due to a combination of soil management needs, and government programs or educational initiatives (such as the Illinois Voluntary Limestone Program, https://www.iaap-aggregates.org/agricultural-lime.html). Starting from the 1960s, there is a gradual increase of lime application extending into the southern U.S. and particularly in the southeastern coast, reaching a peak in 1978, which Georgia alone exceeded 1 Mt lime year^{-1} statewide (Fig. 2). In the 1980s lime application slightly declined on the southeastern coast but maintained a similar spatial pattern to the 1960s. Fig. 3 & Fig. 4 show the county-level lime application mass from 1954 to 1987, highlighting lime application being initially concentrated around the Midwest in the 1950s, a gradual decline in the later 60s to early 70s, an increase and expansion to the South and Southeast in the late 70s, and a slight decline in the 80s.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

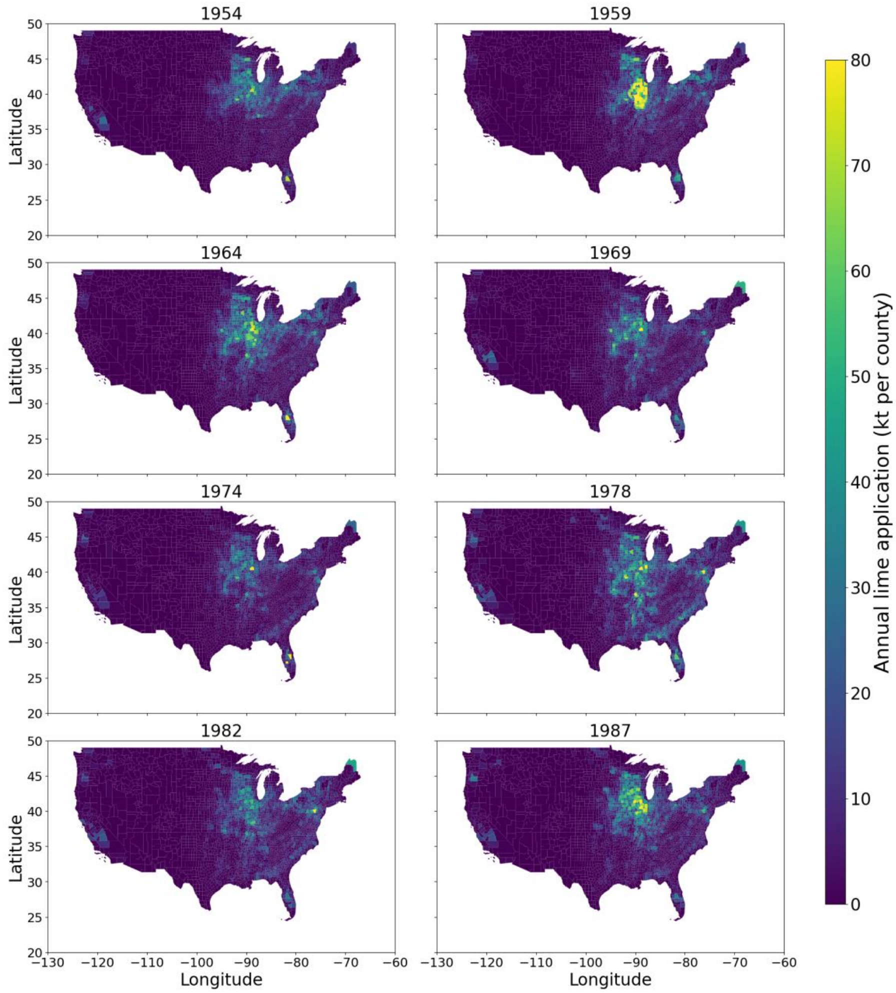

Figure 4 Annual lime application (kilotons, kt) per county in the contiguous U.S. for the eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display the lime application across counties with missing data temporarily interpolated and spatially filled, as described in the Methods section. The data only represents the condition during each specific census year, not including the period in between.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

The temporal-spatial pattern of county-level limed area (Fig. 5) approximately resembles that of limed applied mass, which the counties with the higher area of lime application are also distributed in the midwestern regions. Limed areas were high in the Midwest by the 1950s and gradually extended southward into the eastern and southeastern coasts starting from the 1970s. However, compared to limed weight, the limed area in the southeastern coast shows a more significant signal, suggesting that an equivalent amount of lime is applied over a larger area in the southeast. Compared to the states with the highest lime weight applied being all in the Midwestern regions, the states that have the highest limed area include both midwestern and southeastern states (Fig. S3). The relationship between limed weight and limed area across regions reflects the variation of lime application rates across regions, which may be influenced by local soil and climate properties (Fig. S4) .

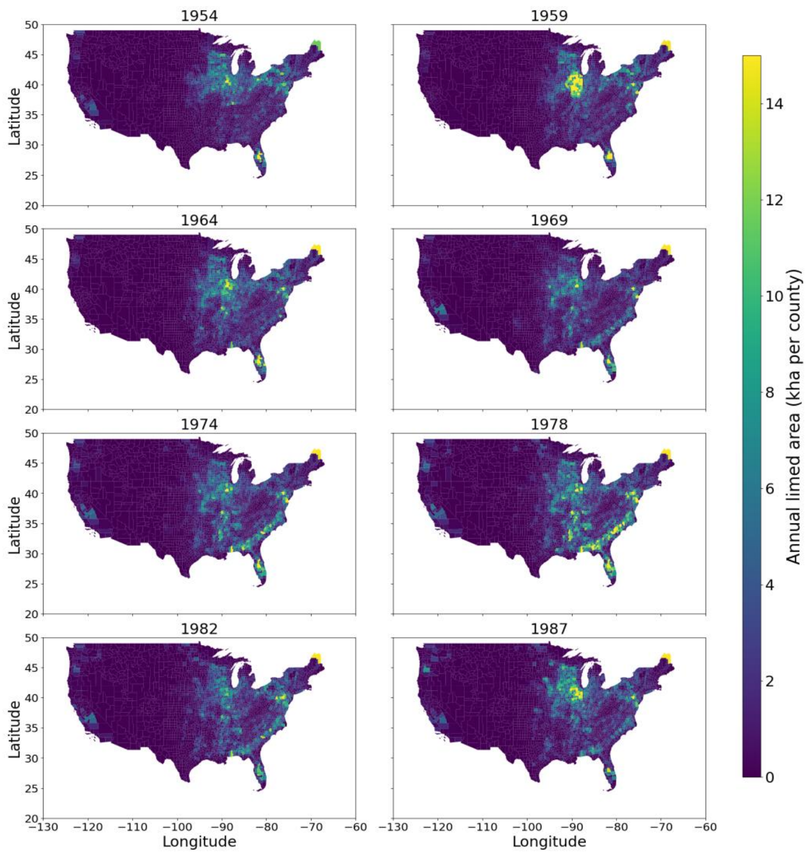

Figure 5 Annual limed area (kilohectares, kha) per county in the contiguous U.S. for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display lime application across counties, with missing data temporarily interpolated and spatially filled, as described in the Methods section.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

The lime application rate for each county (Fig. 6), calculated as the ratio of applied mass to area of lime application, shows a sharp transition between the western and eastern parts of the U.S. Most of the lime was applied only in the eastern half of the U.S., while only a few coastal regions received lime in the western half of the US. Lime application rates were relatively lower during the 1950s but reached a similar level from the 1960s to the 1980s. Fig. S4 shows that the midwestern states of the U.S. had the highest lime application rate (> 4 t ha^{-1} yr^{-1}). Notably, Michigan and Minnesota, though not ranked as the top 10 states of total amount applied (Fig. 2), have a high lime application rate ranking (8th and 10th respectively). In comparison with the mid-west region, the eastern, southeastern, and southern regions of the U.S, have lower lime application rates (~2 t ha^{-1} yr^{-1}).

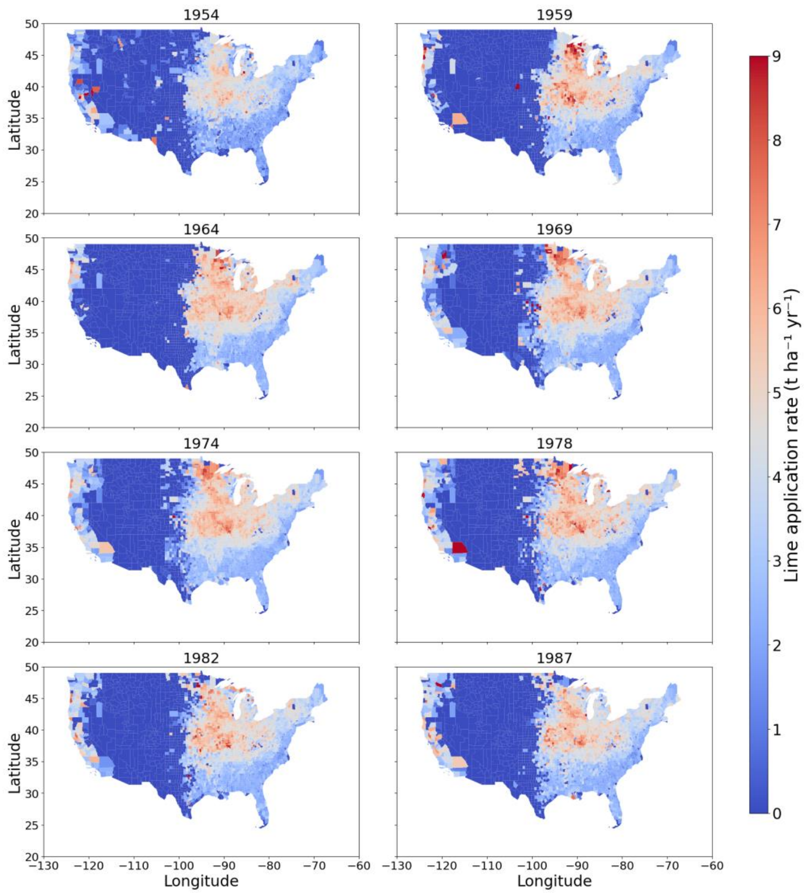

Figure 6. Lime application rate (t ha^{-1} yr^{-1}) for each county in the contiguous U.S. from 1954 to 1987. Rates were calculated as the total lime applied (t) divided by the limed area (ha) for each census year. The eight maps represent the eight census years that reported data during this period.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

The liming frequency, calculated as the ratio of agricultural land limed to total agricultural land each year, reveals a distinct Midwest–Southeast contrast in the eastern half of the US. (Fig. 7). The Southeast generally shows a higher frequency compared to the Midwest. In the 1950s, most regions had a low ratio (<10% of agricultural lands limed per year), with only Florida and a few areas along the eastern seaboard exhibiting higher values (~20%). A marked increase in this ratio occurred throughout the Southeast beginning in the 1960s, peaking at ~30% in 1970. In contrast, the Midwest consistently maintained much lower values, with only ~5% of cropland limed annually. Assuming 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. This contrasting spatial pattern in both Fig. 6 (lime application rate) and Fig. 7 (liming frequency) highlights the spatial differences in lime management, likely driven by regional variations in climate and soil properties.

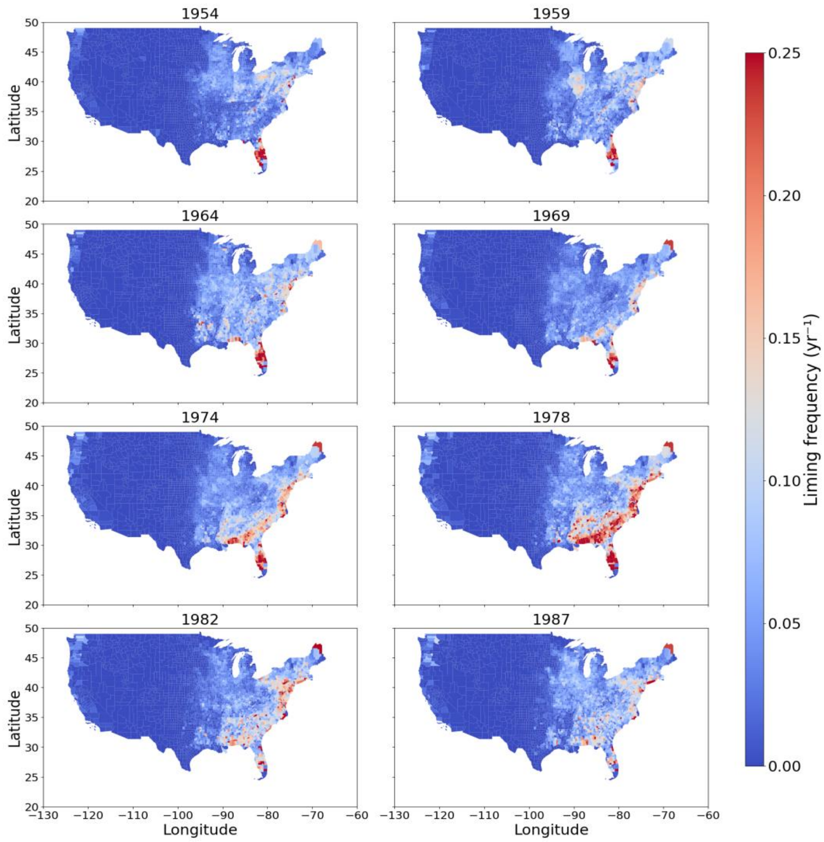

Figure 7. Liming frequency (yr^{-1}) for each county in the contiguous U.S. from 1954 to 1987, calculated as the ratio of limed area to total cropland area. The eight maps represent the eight census years, showing the fraction of cropland treated with lime annually.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

3.2 Statistical analysis of environmental controls

We examined potential environmental predictors of lime-application patterns in terms of three separate outcome variables: (a) total lime applied per county, (b) lime application rate, and (c) liming frequency. After removing the highly correlated predictors from the initial environmental parameters (Fig. S5, Table S2), our Random Forest models demonstrated strong explanatory power, with R^2 values ranging from 0.83 to 0.86 on test data (20% of data) (Fig. 8), while the R^2 on the training set (80% of data) averaged 0.97 to 0.98, indicating strong predictive performance with limited overfitting.

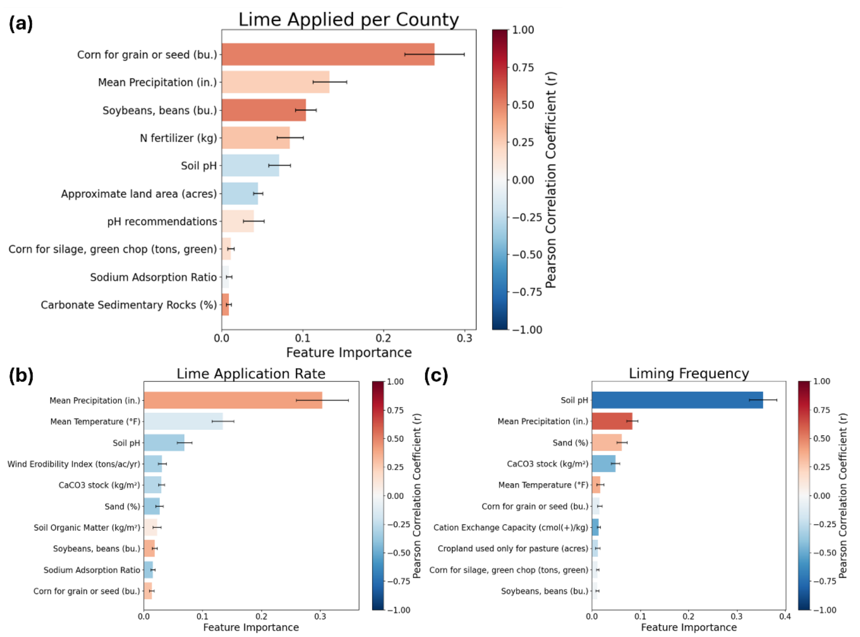

Figure 8 Statistical and machine learning analysis of average annual (a) total lime applied per county per year, normalized by county land area (t ha^{-1} yr^{-1}), (b) lime application rate (t ha^{-1} yr^{-1}), and (c) liming frequency (yr^{-1}) per county across the three census years from 1978 to 1987. For (b) and (c) we ran the analysis only on the counties that have greater than 10% cropland by area, to minimize the influence of counties that have minimal agriculture. The bar length represents the permutation importance of all variables in the random forest model, with the error bar showing the standard error for the permutation importance across the 5-folds of cross validation training. Only the top ten variables with the highest permutation importance are shown here. The colors represent the Pearson correlation coefficients with the target variable.

For the total lime applied per county, the most influential predictors derived from the Random Forest model were corn grain production, mean precipitation, and soybean production, all of which showed strong positive associations. These were

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

followed by N fertilizer use, soil pH, approximate land area, and pH recommendations. Partial dependence plots (Fig. S6) indicate that counties with greater crop production, higher precipitation, lower soil pH elevated fertilizer inputs, and higher pH targets tend to apply more lime. With respect to the lime application rate, the most important variable was mean precipitation, positively correlated with the response, followed by mean temperature, soil pH, wind erodibility index, and carbonate mineral stocks, all showing negative correlations. Additional predictors included sand percentage (negative correlation), soil organic matter (positive correlation) (Fig. 8b). Partial dependence analyses reveal increased lime application rates in wetter and colder regions with lower soil pH, wind erodibility, carbonate mineral stocks, sand percentage, and higher soil organic matter (Fig. S7). For liming frequency, soil pH emerged as the dominant predictor, exhibiting a strong negative correlation. Other important predictors included mean precipitation and sand percentage (positive correlations), and carbonate mineral stocks, mean temperature (negative correlations). Counties with lower soil pH, lower carbonate minerals, higher sand content, and greater rainfall exhibited more frequent liming practices (Fig. 8c, Fig. S8). It is worth noting that the relative ranking of these features may change slightly with different hyperparameter settings in the RF model, the important features in the model reveal important information of the underlying relationships. Overall, the average quantity of lime applied in counties is largely influenced by agricultural management practices, including crop production, fertilizer application, and soil pH recommendations, as well as climate. In contrast, the application rate and frequency are predominantly governed by regional climatic conditions and soil properties.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

4. Discussion

4.1 Spatial and temporal change of county-level lime application

From our results we showed that most agricultural lime was applied in the Midwest region of the continental U.S. during the years 1930-1987. This pattern likely reflects higher precipitation and lower soil pH in the east compared to western U.S. (Hoffmeister, 1947; Wieczorek, 2019). Precipitation leaches base cations (calcium, magnesium) and depletes buffering agents such as $CaCO_3$, causing soil pH to drop nonlinearly as rainfall increases (Slessarev et al. 2016). Consistent with this observation, we found that the average soil pH (from SSURGO) has a strong negative correlation with average precipitation (from PRISM) (Pearson's r=-0.80; Fig. S5).

An extensive amount of lime is applied in the Corn Belt in the American Midwest, which has more than 70 million ha in corn and soybean rotation and is one of the most intensively managed agricultural regions globally (Green et al., 2018). Corn grain production emerged as the strongest predictor of total lime applied per county in the permutation-based model (Fig. 8a), followed by mean precipitation, soybean production, nitrogen fertilizer use, and soil pH. We further show that there is a gradual expansion of the lime application area (Fig. 5), correlating with the expansion, in particular a southward extension, of corn and soybean production area (Fig. S9, S10), as there was a slight decline in tobacco and cotton production (Fig. S13, S14). There is also an increase in lime application on the southeastern coast starting from the 1960s and 1970s, matching the increase of agriculture in the region. However, lime application along the southeastern coast declined in the 1980s, potentially due to a decrease of agricultural land-use resulting from urbanization and conservation programs (Napton et al., 2010. Additionally, the growth of the regional lumber industry during this period may have further reduced agricultural acreage, contributing to the observed decline in lime application (Prestemon & Abt, 2002).

The higher importance of corn and soybeans in determining the amount of lime applied compared to other crop types (e.g., hay, wheat, cotton, tobacco) (S11-S15), may be primarily 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 higher intensity (Annan et al., 2024). The intensive agriculture, including frequent fertilizer inputs, acidifies the soil, necessitating lime application. Additionally, soybeans are also a nitrogen fixing species, which can promote nitrification and increase soil acidification (Nyantsanga & Pierre, 1973). Although other crops like alfalfa may be more pH-sensitive, their production is less fertilizer intensive and may result in a weaker apparent relationship with lime application (Mallarino et al., 2011). As for non-alfalfa hay systems, relatively low profit margins may deter investment in fertilizer and lime applications.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Furthermore, counties in the Midwest that predominantly grow corn and soybeans tend to have a higher proportion of cropland relative to total land area (Fig. S16), which adds to the association between production of these crops and the amount of lime applied relative to total county area. Taken together, these results suggest that while broad patterns such as the association between lime use and corn-soybean agriculture are evident, the drivers of lime application are complex, regionally dynamic, and shaped by interacting environmental, agronomic, and economic factors. Our machine learning framework highlights just one of the perspectives to these complex relationships, which we have shown is more dynamic at the regional scale than commonly recognized.

4.2 Spatial distribution of lime application rate and frequency

We show a spatial pattern of higher liming application rates (> 4 t ha^{-1} yr^{-1}) and longer liming intervals (~20 years) in the Midwest and lower application rates with shorter intervals (~5 years) near the southeastern coastal regions (Fig. 6-7). Permutation importance analysis (Fig. 8b) shows that mean precipitation is the most important predictor of lime application rate, followed by mean temperature, soil pH, wind erodibility index, carbonate mineral stock, sand content, and soil organic matter. In contrast, liming frequency is most strongly explained by soil pH, followed by mean precipitation, sand content, and carbonate mineral stock (Fig. 8c). Together, these findings suggest that lime application practices are governed by a complex interplay between acidification drivers and soil buffering responses.

The required lime application rate and liming frequency is heavily dependent on the soil acidification rate and pH buffering capacity (Xu et al., 2022; Kanzaki et al. 2024). Higher precipitation leads to faster leaching of base cations, while regions of warmer temperatures lead to faster decomposition and organic acid release, with both processes leading to faster soil acidification (Ulrich, 1986). Both processes promote faster soil acidification. However, precipitation appears to be a necessary condition for acidification to manifest at scale, while temperature acts both as a direct driver and an indirect proxy for soil characteristics. Specifically, warmer regions tend to have more weathered soils with low buffering capacity, while cooler regions, particularly in the Midwest, often retain glacially derived carbonates and clay minerals that confer greater buffering potential. This helps explain the somewhat counterintuitive negative correlation between temperature and application rate: although warmer regions acidify faster, their low-buffering soils allow for smaller lime doses per application, whereas cooler, more buffered soils require larger doses when they are eventually limed. This pattern is further supported by the partial dependence plots (Fig. S7), which reveal a distinct threshold-like response: lime application rate increases steeply with precipitation beyond approximately 750–800 mm/year, and declines sharply with

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

temperature above 12–13°C. These non-linear relationships imply that precipitation is a key threshold variable enabling acidification, while temperature's response reflects both its role in driving acid production and its spatial correlation with regional differences in soil buffering capacity.

4.3 Regional variation in buffering capacity and liming strategy

Resisting the higher acidification rates associated with warmer, wetter conditions in the southeastern USA likely entail more frequent lime additions. However, these regions typically do not require higher application rates per event, as their soils exhibit relatively low pH buffering capacity. The pH buffering capacity of soils varies systematically between the Midwest and the Southeast, which is another main cause in the spatial patterns in lime application rate and frequency. Soil buffering capacity is influenced by naturally occurring carbonate minerals ($CaCO_3$ and $CaMg(CO_3)_2$), which can directly neutralize acidity when it dissolves. The influence of carbonate minerals is particularly relevant in some Midwestern soils, where residual carbonate from glacial parent materials helps maintain higher soil pH values. In contrast, Southeastern soil typically lacks carbonate minerals, making them more vulnerable to acidification (Fig. S17). In addition to carbonates, cation exchange capacity (CEC) helps to resist swings in pH by capturing or releasing adsorbed base cations and protons (Brady & Weil, 2016). CEC also varies systematically between the southeastern and midwestern USA. Several factors contribute to the lower CEC observed in highly-weathered southeastern soils (Fig. S18), which are dominated by low activity clay minerals, coarser particle sizes, elevated concentrations of detrital quartz (Woodruff et al., 2009), and lower organic matter content (Solly et al., 2020). Conversely, soils in the upper Midwest are derived from glacial parent materials and loess (Borker et al. 2018), which are often characterized by higher organic matter contents, smaller soil particle sizes, and higher activity clay minerals (Fig. S19, S20), providing more CEC and better pH buffering capacity. These finer-textured soils are also generally less erodible by wind, consistent with the observed negative correlation between wind erodibility and lime application rate.

Although CEC and carbonate content are key conceptual components of buffering capacity, they did not emerge as top predictors in the permutation importance analysis. This is likely due to collinearity with soil pH, which reflects the integrated outcome of both acidification pressure and buffering history. In this sense, soil pH serves as a practical proxy for underlying soil properties that govern liming frequency. The spatial correlation is particularly relevant in the southeastern U.S., where low soil pH tends to co-occur with both low CEC and minimal carbonate content, rendering pH a strong indicator of inherently low buffering capacity. A similar rationale applies temperature's high variable importance in predicting lime

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

application rate, which as discussed earlier, is likely to reflect its spatial co-correlation with the degree of soil weathering and associated buffering capacity.

In summary, both climatic drivers (temperature precipitation) and soil mineralogy (including soil pH, carbonate presence, wind erodibility, soil organic matter, and particle size distribution) together explain the spatial variation of lime application rate and frequency. The Midwestern region requires less frequent but larger amounts of lime application to maintain soil pH, while the southeastern region requires more frequent but smaller amounts of lime application to effectively maintain optimal pH levels (Ross et al., 1964).

4.4 Temporal variability of total agricultural lime application.

We further hypothesize that economics may have played an important role in shaping the temporal variability of lime use. As shown in Fig. S21, the late 1960s and early 1970s witnessed a surge in fertilizer expenditures and a drop in fertilizer cost, likely driven by technological advances, leading farmers to prioritize fertilizer over liming. This behavior may have contributed to the notable decline in lime application during this period. The immediate yield gains from increased fertilizer use could have overshadowed the longer-term implications of reduced liming, particularly given that soil acidification tends to accumulate gradually over multi-year or decadal timescales. In 1974, fertilizer prices rebounded sharply in conjunction with the global fuel crisis (Brunelle et al., 2015), aligning with the most pronounced dip in lime application rate observed that year. In the late 1970s to early 1980s, farm incomes peaked, and input costs stabilized (Fig. S22), creating conditions conducive to greater focus on long-term soil health. Consequently, this period saw a resurgence in lime application. It is also likely that this period marked growing recognition of the soil acidification problem, as several studies on liming and soil pH management were published during this time (Nyborg & Hoyt, 1978; Bache, 1980). By the late 1980s, declining crop sales and lower input costs suggest a period of adjustment and input optimization, potentially explaining the modest reduction in total lime applied (Fig. S22)

4.5 Comparison of total lime application estimate with West & McBride (2005)

Fig. 9 shows a comparison between the total lime application in our dataset and the national estimate of lime application in West & McBride (2005), which to our knowledge is the first national estimate of agriculture lime applied in the US. Our study and West & McBride both show that lime application increases very slowly starting from the 1910s but started to rapidly increase in ~1930 and reach a high range of values, exceeding 20 Mt/yr around 1950. However, there are some discrepancies from 1950 through 1980. West & McBride estimated the total lime applied to continue to increase, particularly during the 1970s, which exceeded 30 Mt per year, while our dataset shows the total lime applied to be more consistent in the range of

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

15-25 Mt per year. Oh & Raymond (2006) conducted a regional study based on the county-level liming dataset from the Census reports and first noticed disagreements in the trends of the regional sum of the county-level data and the national estimate from West & McBride. However, Oh and Raymond's study was done over a small region of the US, and a spatially explicit dataset of lime application across the U.S. was not yet available at that time to address the uncertainty in the total lime estimates.

Here we revisited the discrepancy between the national estimate of lime application and that of West & McBride (2005). We suggest that the difference is primarily due to uncertainties in the methodology of West & McBride, which relies on data from the Mineral Yearbook (U.S. Bureau of Mines, 1927–1996; U.S. Geological Survey, 1906–1927, 1997–2002). The Mineral Yearbook classifies crushed stones by their intended use, with agricultural limestone being a sub-category of "agricultural uses". We noticed roughly half of crushed stones are classified under an "unspecified end-use" category. West & McBride estimated total agricultural lime by combining explicitly labeled agricultural limestone with an inferred fraction (approximately 2–3%) from the unspecified category, assuming proportional distribution across all categories.

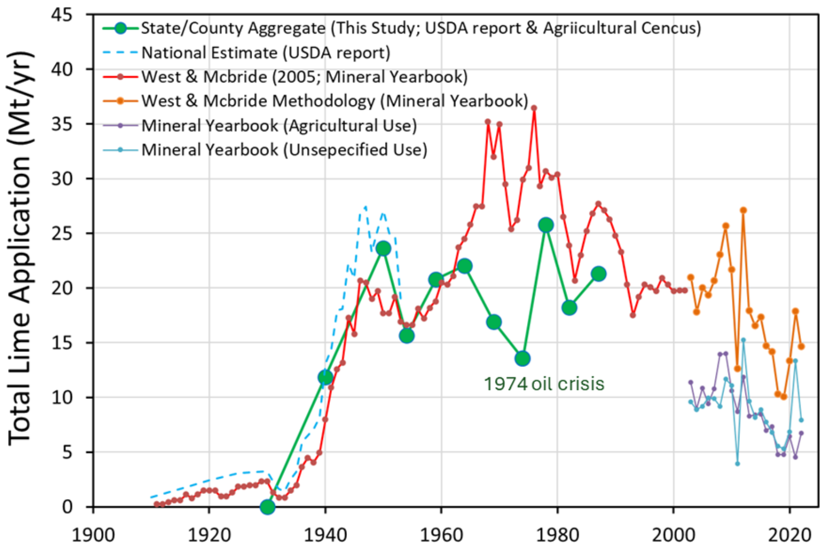

Figure 9 Comparison of national estimates of the total weight of lime applied (tons) over time. Our study's compilation (state-level aggregate (1930–1950) from the USDA report and county-level aggregate (1954–1987) from the Agriculture Census) are shown as green dots. The green dashed line represents the national estimate reported annually in the USDA report (1910-1950). The red line represents national estimates from West & McBride (2005). The orange dots indicate the extended total lime applied calculation based on the Mineral Yearbook using methods from West & McBride (2005). The purple line represents the fraction of lime applied for agricultural use, and the blue line reflects the unspecified end-use fraction, scaled by the ratio of agricultural use to total specified end-use, representing the

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

potential lime used for agriculture.

We further extended the methodology of West & McBride (2005) until 2022 (orange line) (data extracted from https://www.usgs.gov/centers/national-minerals-information-center/crushed-stone-statistics-and-information. We aggregated the number of crushed stones specified as “agricultural limestone” with an estimated percentage of agricultural lime within the “unspecified” end-use category. We show that the fraction of data in the unspecified end-use category (blue line; assumed to be used for agriculture) is similar in magnitude as those specified to be used for agriculture (purple line). This suggests that ~50% of the estimate in West & McBride is sensitive to how “unspecified” use is interpreted, which leads to high uncertainty and potential biases in total estimate. Given that agricultural limestone is such a small fraction of the total unspecified crushed rocks (2–3%), small variations in this percentage would lead to a large impact on the total estimated agricultural lime. Therefore, we believe that our compiled dataset here offers a better representation of the total amount of lime applied in the US, and that 15-25 Tg per year could be a more accurate range instead of the 25–35 Tg per year suggested in West & McBride.

Based on our extended analysis with West & McBride (2005)’s methodology, we also show that the total annual applied lime in the last two decades varied between 10–25 Tg per year, but did not surpass the peak in the 1970s. A similar trend is shown in the total nutrient input via fertilizers, where there is no significant increase of N and P input since 1980 (Fig. S23). This suggests that the amount of agricultural lime applied likely did not increase since then, providing evidence that the temporal range of our dataset in this study covered the period when lime application reached its full extent in the USA.

5. Data Availability

The dataset presented in this study will be made publicly available on Zenodo at https://doi.org/10.5281/zenodo.15758275 (Tsao, S. S.-E., Surhoff, T. J., Amatulli, G., & Raymond, P. (2025). Agricultural lime application across the contiguous United States, 1930–1987 [Data set]. Zenodo.). Reviewers can access the dataset during peer review via the private preview link: https://zenodo.org/records/15758275?preview=1&token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjdiN2VkNDkwLTE1YjctNGNiZC1iMjY2LTBlOGMzZjkzYTZmZSIsImRhdGEiOnt9LCJyYW5kb20iOiJjZTFmNTA4ZDliZGU4NDQzOGNlMDkzMTYxNjJkMDQxNSJ9.2yChJg8txya_OZzIo6nmgFAjVj-haakBfMtFcLrDukLXaZxKVM19i-DrUZ3kKuI443l4z1KjREqlKiOlHbJ42g

This dataset provides spatially explicit estimates of agricultural lime application across the contiguous United States, including: 1. State-level data from 1930 to 1950, and 2. County-level data from 1954 to 1987.

Key variables include total lime applied (metric tons), area limed (hectares),

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

application rate (t ha^{-1} yr^{-1}), liming intensity (t ha^{-1} yr^{-1}), and liming frequency (yr^{-1}), which are visualized in Figures 4–7.

The county-level shapefile contains the following columns:

-Year: Year of observation

-GEOID: Unique identifier for each U.S. county

-LimeTons: Total lime applied (metric tons)

-LimeArea: Area limed (hectares)

-CropArea: Total cropland area (hectares)

-LimeRate: Lime application rate (t ha^{-1} yr^{-1})

-LimeIntsty: Liming intensity (t ha^{-1} yr^{-1})

-LimeFreq: Liming frequency (yr^{-1})

All files are provided in open formats and are accompanied by metadata and documentation to ensure transparency and reproducibility.

6. Code Availability

The code used for data processing, interpolation, and machine learning analysis will be made publicly available upon publication of the final version of this manuscript. A link to the repository will be provided at that time to ensure full reproducibility.

7. Conclusion

In this study, we compiled and presented the first spatially explicit data set of agricultural lime application in the U.S. from 1930–1987. We showed that the Midwest region has the most amount of lime applied, while the south and southeastern coastal region of the U.S. has gradually increased lime application. We also showed that the Midwest has a much longer liming interval and higher lime application rate than the Southeastern and eastern US, while the liming frequency is higher in the Southeastern and eastern regions than in the Midwest. Using machine learning algorithms, we showed that on the county level, the total amount of lime applied is largely influenced by the agricultural management practices (soybean and corn production, fertilizers, soil pH recommendation) alongside natural factors such as soil pH and climate. In contrast lime application rate and frequency are predominantly governed by natural regional climate (e.g., precipitation, temperature) and soil characteristics (e.g., soil pH, carbonate minerals, particle size distribution, soil texture, organic matter). We further show that the temporal variability in total lime application may be influenced by economic factors, including competing input decisions such as fertilizer use, as well as fluctuations in fuel prices and farm income and expenditure.

Our study synthesizes historical lime application practices across both temporal and

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

spatial dimensions, offering a critical observational baseline for future modeling efforts. As liming gains attention for its potential role in carbon removal and climate mitigation, understanding historical application patterns is essential for accurately assessing its long-term carbon sequestration potential and contribution to greenhouse gas reductions. Future research should build on these trends by integrating spatial and temporal dynamics with variations in soil types, climate conditions, and other environmental factors to comprehensively assess the viability of liming as a nature-based climate solution.

8. Author Contributions

Conceptualization: MA, TJS, GA, CTR, NJP, PAR
Methodology: SST, TJS, GA, JG, EWS, CTR, SZ, PAR,
Software: GA
Validation: SST, TJS
Formal Analysis: SST, TJS, EWS
Investigation: SST, TJS, GA,
Resources: PAR
Data Curation: TJS, GA, SST
Writing – Original Draft: SST
Writing – Review & Editing: TJS, GA, BW, MA, EWS, JES, CTR, SZ, NJP, PAR
Visualization: SST
Supervision: PAR
Project Administration: PAR
Funding Acquisition: PAR

9. Competing Interest

The authors declare that they have no conflict of interest.

10. Acknowledgements.

10. This research was supported by the U.S. Department of Energy Earthshot Initiative (grant no. DE-SC0024709). Additional support was provided by the Yale Center for Natural Carbon Capture and the Yale School of the Environment. We thank Rachel Sperling (Yale University Library) for assistance with data mining and the Spatial Ecology summer workshop team for training and help in accelerating the data-processing workflow. We are also grateful to all members of the Raymond Lab and Saiers Lab for their constructive feedback during the preparation of this manuscript.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

References

Agegnehu, G., Amede, T., Erkossa, T., Yirga, C., Henry, C., Tyler, R., Nosworthy, M. G., Beyene, S., and Sileshi, G. W.: Extent and management of acid soils for sustainable crop production system in the tropical agroecosystems: A review, Acta Agric. Scand. Sect. B Soil Plant Sci., 71, 852–869, https://doi.org/10.1080/09064710.2021.1954239, 2021.

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T.: Permutation importance: a corrected feature importance measure, Bioinformatics, 26, 1340–1347, 2010.
Annan, K., Fausti, S. W., Van der Sluis, E., and Kolady, D. E.: Corn Acreage Intensification Levels in US Corn Belt States, J. Agric. Appl. Econ., 56, 353–372, 2024.

Bache, B. W.: The acidification of soils, in: Effects of acid precipitation on terrestrial ecosystems, edited by: Dochinger, L. S. and Seliga, T. A., Springer US, Boston, MA, 183–202, https://doi.org/10.1007/978-1-4613-3083-1_11, 1980.

Beerling, D. J., Kantzas, E. P., Lomas, M. R., Wade, P., Eufrasio, R. M., Renforth, P., Sarkar, B., Andrews, M. G., James, R. H., Pearce, C. R., Mercure, J.-F., Pollitt, H., Holden, P. B., Edwards, N. R., Khanna, M., Koh, L., Quegan, S., Pidgeon, N. F., Janssens, I. A., and Banwart, S. A.: Potential for large-scale CO2 removal via enhanced rock weathering with croplands, Nature, 583, https://doi.org/10.1038/s41586-020-2448-9, 2020.

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, 2001.

Brunelle, T., Dumas, P., Souty, F., Dorin, B., and Nadaud, F.: Evaluating the impact of rising fertilizer prices on crop yields, Agric. Econ., 46, 653–666, 2015.

Burdett, H. and Wellen, C.: Statistical and machine learning methods for crop yield prediction in the context of precision agriculture, Precis. Agric., 23, 1553–1574, https://doi.org/10.1007/s11119-022-09897-0, 2022.

De Klein, C. et al.: Chapter 11 - N2O Emissions from managed Soils and CO2 Emissions From Lime and Urea Application, IPCC Guidelines for Natl. Greenh. Gas Invent., Vol. 4, 11.1–11.54, 2006.

Falcone, J. A.: Estimates of county-level nitrogen and phosphorus from fertilizer and manure from 1950 through 2017 in the conterminous United States, USGS Open-File

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Open Access Earth System Science Data Discussions

Report, https://doi.org/10.3133/ofr20201153, 2021.

Fageria, N. K. and Baligar, V. C.: Ameliorating soil acidity of tropical Oxisols by liming for sustainable crop production, Adv. Agron., 99, 345–399, https://doi.org/10.1016/S0065-2113(08)00407-0, 2008.

Goulding, K. W. T.: Soil acidification and the importance of liming agricultural soils with particular reference to the United Kingdom, Soil Use Manage., 32, 390–399, https://doi.org/10.1111/sum.12270, 2016.

Green, T. R., Kipka, H., David, O., and McMaster, G. S.: Where is the USA Corn Belt, and how is it changing?, Sci. Total Environ., 618, 1613–1618, 2018.

Guo, J. H., Liu, X. J., Zhang, Y., Shen, J. L., Han, W. X., Zhang, W. F., Christie, P., Goulding, K. W. T., Vitousek, P. M., and Zhang, F. S.: Significant acidification in major Chinese croplands, Science, 327, 1008–1010, https://doi.org/10.1126/science.1182570, 2010.

Haines, M., Fishback, P., and Rhode, P.: United States Agriculture Data, 1840–2012, Inter-university Consortium for Political and Social Research, https://doi.org/10.3886/ICPSR35206.v4, 2018.

Hamilton, S. K., Kurzman, A. L., Arango, C., Jin, L., and Robertson, G. P.: Evidence for carbon sequestration by agricultural liming, Glob. Biogeochem. Cycles, 21, https://doi.org/10.1029/2006GB002738, 2007.

Han, T., Cai, A., Liu, K., Huang, J., Wang, B., Li, D., Qaswar, M., Feng, G., and Zhang, H.: The links between potassium availability and soil exchangeable calcium, magnesium, and aluminum are mediated by lime in acidic soil, J. Soils Sediments, 19, 1382–1392, https://doi.org/10.1007/s11368-018-2145-6, 2019.

Hartmann, J. and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochem. Geophys. Geosyst., 13, Q12004, https://doi.org/10.1029/2012GC004370, 2012.

Hoffmeister, H. A.: Alkali problem of western United States, Econ. Geogr., 23, 1–9, 1947.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Hooda, P. S. and Alloway, B. J.: The effect of liming on heavy metal concentrations in wheat, carrots and spinach grown on previously sludge-applied soils, J. Agric. Sci., 127, 289–294, https://doi.org/10.1017/S0021859600078448, 1996.

Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K.-M., Gerber, J. S., Reddy, V. R., and Kim, S.-H.: Random Forests for global and regional crop yield predictions, PLoS ONE, 11, e0156571, https://doi.org/10.1371/journal.pone.0156571, 2016.

Johnston, J. F. W.: On the use of lime in agriculture, 1849.

Kanzaki, Y., Planavsky, N., Zhang, S., Jordan, J., Suhrhoff, T. J., and Reinhard, C. T.: Soil cation storage as a key control on the timescales of carbon dioxide removal through enhanced weathering, Authorea Preprints, 2024.

Khaliq, M. A., Tarin, M. W. K., Jingxia, G., Yanhui, C., and Guo, W.: Soil liming effects on CH4, N2O emission and Cd, Pb accumulation in upland and paddy rice, Environ. Pollut., 248, 408–420, 2019.

Knapp, W. J. and Tipper, E. T.: The efficacy of enhancing carbonate weathering for carbon dioxide sequestration, Front. Clim., 4, https://doi.org/10.3389/fclim.2022.928215, 2022.

Liu, Z., Dreybrodt, W., and Liu, H.: Atmospheric CO2 sink: Silicate weathering or carbonate weathering?, Appl. Geochem., 26, S292–S294, https://doi.org/10.1016/j.apgeochem.2011.03.085, 2011.

Mallarino, A., Pagani, A., and Sawyer, J.: Corn and soybean response to soil pH level and liming, Iowa State Univ. Ext., https://dr.lib.iastate.edu/handle/20.500.12876/43941, 2011.

Mehring, A. L., Adams, J. R., and Jacob, K. D.: Statistics on fertilizers and liming materials in the United States, USDA Soil and Water Conservation Research Branch, 191, 1957.

Napton, D. E., Auch, R. F., Headley, R., and Taylor, J. L.: Land changes and their driving forces in the Southeastern United States, Reg. Environ. Change, 10, 37–53, 2010.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Nyatsanga, T. and Pierre, W. H.: Effect of nitrogen fixation by legumes on soil acidity, Agron. J., 65, 936–940, 1973.

Nyborg, M. and Hoyt, P. B.: Effects of soil acidity and liming on mineralization of soil nitrogen, Can. J. Soil Sci., 58, 331–338, 1978.

Oh, N. and Raymond, P. A.: Contribution of agricultural liming to riverine bicarbonate export and CO2 sequestration in the Ohio River basin, Glob. Biogeochem. Cycles, 20, https://doi.org/10.1029/2005GB002565, 2006.

Pagani, A. and Mallarino, A. P.: Soil pH and crop grain yield as affected by the source and rate of lime, Soil Sci. Soc. Am. J., 76, 1877–1886, https://doi.org/10.2136/sssaj2012.0119, 2012.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É.: Scikit-learn: Machine learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Plummer, L. N. and Wigley, T. M. L.: The dissolution of calcite in CO2-saturated solutions at 25 °C and 1 atmosphere total pressure, Geochim. Cosmochim. Acta, 40, 191–202, 1976.

Prestemon, J. P. and Abt, R. C.: Southern forest resource assessment highlights: The southern timber market to 2040, J. For., 100, 16–22, 2002.

Robertson, G. P., Paul, E. A., and Harwood, R. R.: Greenhouse gases in intensive agriculture: Contributions of individual gases to the radiative forcing of the atmosphere, Science, 289, 1922–1925, https://doi.org/10.1126/science.289.5486.1922, 2000.

Ross, G. J., Lawton, K., and Ellis, B. G.: Lime requirement related to physical and chemical properties of nine Michigan soils, Soil Sci. Soc. Am. J., 28, 209–212, 1964.
Semhi, K., Amiotte Suchet, P., Clauer, N., and Probst, J.-L.: Impact of nitrogen fertilizers on the natural weathering-erosion processes and fluvial transport in the Garonne basin, Appl. Geochem., 15, 865–878, https://doi.org/10.1016/S0883-2927(99)00076-1, 2000.

Shaaban, M., Wu, Y., Wu, L., Hu, R., Younas, A., Nunez-Delgado, A., Xu, P., Sun,

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

Z., Lin, S., Xu, X., and Jiang, Y.: The effects of pH change through liming on soil N2O emissions, Processes, 8, 702, https://doi.org/10.3390/pr8060702, 2020.
Slessarev, E. W., Lin, Y., Bingham, N. L., Johnson, J. E., Dai, Y., Schimel, J. P., and Chadwick, O. A.: Water balance creates a threshold in soil pH at the global scale, Nature, 540, 567–569, 2016.

Smil, V.: Enriching the Earth: Fritz Haber, Carl Bosch, and the transformation of world food production, MIT Press, Cambridge, MA, 2004.

Tennant, S.: On different sorts of lime used in agriculture, Philos. Mag., 5, 209–216, 1799.

Tsao, S. S.-E., Surhoff, T. J., Amatulli, G., & Raymond, P. (2025). Agricultural lime application across the contiguous United States, 1930–1987 [Data set]. Zenodo. https://doi.org/10.5281/zenodo.15758275

Solly, E. F., Weber, V., Zimmermann, S., Walthert, L., Hagedorn, F., and Schmidt, M. W. I.: A critical evaluation of the relationship between the effective cation exchange capacity and soil organic carbon content in Swiss forest soils, Front. For. Glob. Change, 3, https://doi.org/10.3389/ffgc.2020.00098, 2020.

Siqueira, R. G., Moquedace, C. M., Fernandes-Filho, E. I., Schaefer, C. E. G. R., Francelino, M. R., Sacramento, I. F., and Michel, R. F. M.: Modelling and prediction of major soil chemical properties with Random Forest: Machine learning as a tool to understand soil-environment relationships in Antarctica, Catena, 235, 107677, 2024.

Ulrich, B.: Natural and anthropogenic components of soil acidification, Z. Pflanzenernähr. Bodenkd., 149, 702–717, https://doi.org/10.1002/jpln.19861490607, 1986.

U.S. Bureau of Mines: Mineral Resources of the United States, U.S. Department of the Interior, Bureau of Mines, Washington, DC, 1927–1996.

U.S. Bureau of the Census: Census of Agriculture: 1954, U.S. Govt. Print. Off., Washington, D.C., 1956.
U.S. Bureau of the Census: Census of Agriculture: 1959, U.S. Govt. Print. Off., Washington, D.C., 1961.
U.S. Bureau of the Census: Census of Agriculture: 1964, U.S. Govt. Print. Off., Washington, D.C., 1967.

https://doi.org/10.5194/essd-2025-411
Preprint. Discussion started: 1 September 2025

U.S. Bureau of the Census: Census of Agriculture: 1969, U.S. Govt. Print. Off., Washington, D.C., 1972.

U.S. Bureau of the Census: Census of Agriculture: 1974, U.S. Govt. Print. Off., Washington, D.C., 1977.

U.S. Bureau of the Census: Census of Agriculture: 1978, U.S. Govt. Print. Off., Washington, D.C., 1981.

U.S. Bureau of the Census: Census of Agriculture: 1982, U.S. Govt. Print. Off., Washington, D.C., 1984.

U.S. Bureau of the Census: Census of Agriculture: 1987, U.S. Govt. Print. Off., Washington, D.C., 1989.

U.S. Geological Survey: Mineral Resources of the United States, U.S. Department of the Interior, U.S. Geological Survey, Washington, DC, 1906–1927.

U.S. Geological Survey: Minerals Yearbook, vol. 1, U.S. Department of the Interior, U.S. Geological Survey, Washington, DC, 1997–2002.

Walkinshaw, M., O'Geen, A. T., and Beaudette, D. E.: Soil Properties, California Soil Resource Lab, https://casoilresource.lawr.ucdavis.edu/soil-properties/, 2023.

Weil, R. R., Brady, N. C., and Weil, R. R.: The nature and properties of soils, Pearson, Vol. 1104, 2017.

West, T. O. and McBride, A. C.: The contribution of agricultural lime to carbon dioxide emissions in the United States: Dissolution, transport, and net emissions, Agric. Ecosyst. Environ., 108, 145–154, https://doi.org/10.1016/j.agee.2005.01.002, 2005.

Wieczorek, M. E.: Area and depth-weighted average of soil pH from STATSGO2 for the conterminous United States and District of Columbia [Dataset], U.S. Geological Survey, https://doi.org/10.5066/P95YOSFQ, 2019.

Woodruff, L. G., Cannon, W. F., Eberl, D. D., Smith, D. B., Kilburn, J. E., Horton, J. D., Garrett, R. G., and Klassen, R. A.: Continental-scale patterns in soil geochemistry and mineralogy: Results from two transects across the United States and Canada, Appl. Geochem., 24, 1369–1381, https://doi.org/10.1016/j.apgeochem.2009.04.009, 2009.

Xu, D., Zhu, Q., Ros, G., Cai, Z., Wen, S., Xu, M., Zhang, F., and de Vries, W.:

Calculation of spatially explicit amounts and intervals of agricultural lime applications at county-level in China, Sci. Total Environ., 806, 150955, https://doi.org/10.1016/j.scitotenv.2021.150955, 2022.

Zeng, S., Liu, Z., and Groves, C.: Large-scale CO2 removal by enhanced carbonate weathering from changes in land-use practices, Earth-Sci. Rev., 225, https://doi.org/10.1016/j.earscirev.2021.103915, 2022.

Zhang, H.-M., Liang, Z., Li, Y., Chen, Z.-X., Zhang, J.-B., Cai, Z.-C., Elsgaard, L., Cheng, Y., van Groenigen, K. J., and Abalos, D.: Liming modifies greenhouse gas fluxes from soils: A meta-analysis of biological drivers, Agric. Ecosyst. Environ., 340, 108182, https://doi.org/10.1016/j.agee.2022.108182, 2022.

Zhang, S., Planavsky, N. J., Katchinoff, J., Raymond, P. A., Kanzaki, Y., Reershemius, T., and Reinhard, C. T.: River chemistry constraints on the carbon capture potential of surficial enhanced rock weathering, Limnol. Oceanogr., 67, https://doi.org/10.1002/lno.12244, 2022.

Zulauf, C., Schnitkey, G., Paulson, N., and Colussi, J.: Concentration of US crops in corn and soybeans: Importance to increasing US production of grains and oilseeds, Farmdoc Daily, 13, https://farmdocdaily.illinois.edu/2023/09/concentration-of-us-crops-in-corn-and-soybeans-importance-to-increasing-us-production-of-grains-and-oilseeds.html, 2023.