Supplementary Information (SI):

1. Liming data processing

The county level data for liming, agricultural and environmental variables was joined with the U.S. county-level shapefile in 2012

Notes for Joining the Datasets

Miami-Dade County, Florida: Previously known as Dade County, this county was renamed Miami-Dade County in 1997, and its FIPS code changed from 12025 to 12086. To join with the county-level shapefile, we updated the county FIPS code to the current code (12086).

Broomfield County, Colorado: Created in 2001 from parts of Adams, Boulder, Jefferson, and Weld counties, Broomfield County (FIPS 8014) did not exist during the 1954-1987 period. Therefore, we assumed its liming application to be 0 for this time frame.

Baltimore, MD; St. Louis, MO; and independent cities in Virginia: Independent cities, such as Baltimore, St. Louis, and several cities in Virginia, are administratively independent from any county and are not included in county-level agricultural data from the U.S. Bureau of the Census. Since agricultural data is typically reported at the county level and these independent cities do not have significant agricultural activity, liming application is assumed to be 0 for these regions.

1974 Limed Area: The 1974 data do not report limed area (acres). Therefore, for 1974, we interpolated the limed area (acres) and lime application rate (tons per acre) using data from the previous census years, 1969 and 1978.

Cropland Area: Cropland area values recorded as 4.000000e+10 are treated as errors and marked as missing values.

Table S1. Predictor Variables by Category, Temporal Range, and Source

Category	Variable	Temporal Range	Source
Agriculture	Land in farms (acres), Harvested crop land (acres), Irrigated land (acres), Total Cropland (acres), Cropland used only for pasture (acres), Cropland not harvested/pastured (acres), Approximate land area (acres)	Yearly varying	U.S. Census of Agriculture
	Corn for grain or seed (bu.), Corn for silage, green chop (tons, green), Sorghum for grain or seed (bu.), Wheat for grain (bu.), Cotton (bales), Tobacco (pounds), Soybeans, beans (bu.), Alfalfa & alfalfa mixtures (tons)	Yearly varying	U.S. Census of Agriculture
Nutrient Input	N fertilizer (kg), P fertilizer (kg), N manure (kg), P manure (kg)	Yearly varying	Falcone (2021)
Climate	Mean Precipitation (in.), Mean Temperature (°F)	Static (Decadal mean)	PRISM Climate Group
Soil Properties	Soil pH, pH recommendations, Clay (%), Silt (%), Sand (%), Cation Exchange Capacity (cmol(+)/kg), Soil Organic Matter (kg/m²), $CaCO_3$ stock (kg/m²), Rock Fragments (0–25 cm) %, Bulk Density (g/cm³), Soil Depth (cm), Electrical Conductivity (dS/m), Sodium Adsorption Ratio, Depth to Restrictive Layer (cm), Available Water Storage (cm), Sat. Hyd. Conductivity (Ksat) – Mean	Static	Walkinshaw et al. (2023), UC Davis Soil Properties
Lithology	Carbonate Sedimentary Rocks (%)	Static	Hartmann & Moosdorf (2012);

Category	Variable	Temporal Range	Source
			GLiM database
Other Physical Factors	Wind Erodibility Index (tons/ac/yr)	Static	Walkinshaw et al. (2023)

Fig. S2

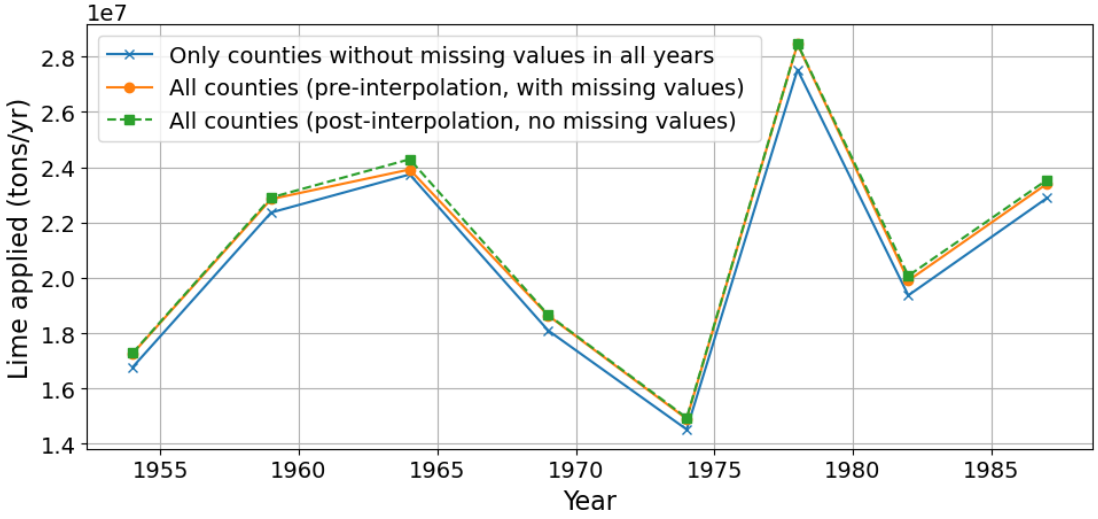

Fig. S1 Comparison of total agricultural lime applied in the U.S. before and after temporal interpolation and data filling. The blue line represents the sum of the counties that have valid data in all years, while the orange line shows the sum of all counties data, before interpolation and data filling. The green dashed line represents the sum of all counties data, after interpolation and data filling. All three lines demonstrate similar trends, with minor differences indicating the effect of missing data, data interpolation and filling.

Fig. S2

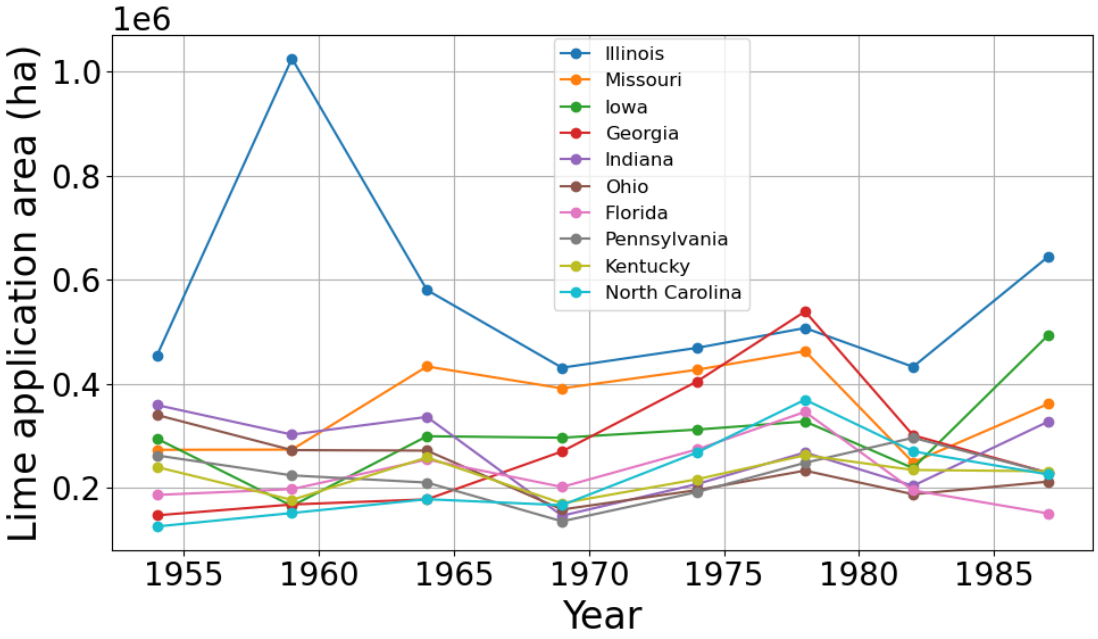

Fig. S2 Temporal variation in the total limed area (hectares, ha) for the top 10 states in the continental U.S. from 1954 to 1987. The states are ranked in descending order based on total limed area across the period, with Illinois having the largest area, followed by Missouri, Iowa, Georgia, Indiana, Ohio, Florida, Pennsylvania, Kentucky, and Mississippi.

Fig. S3

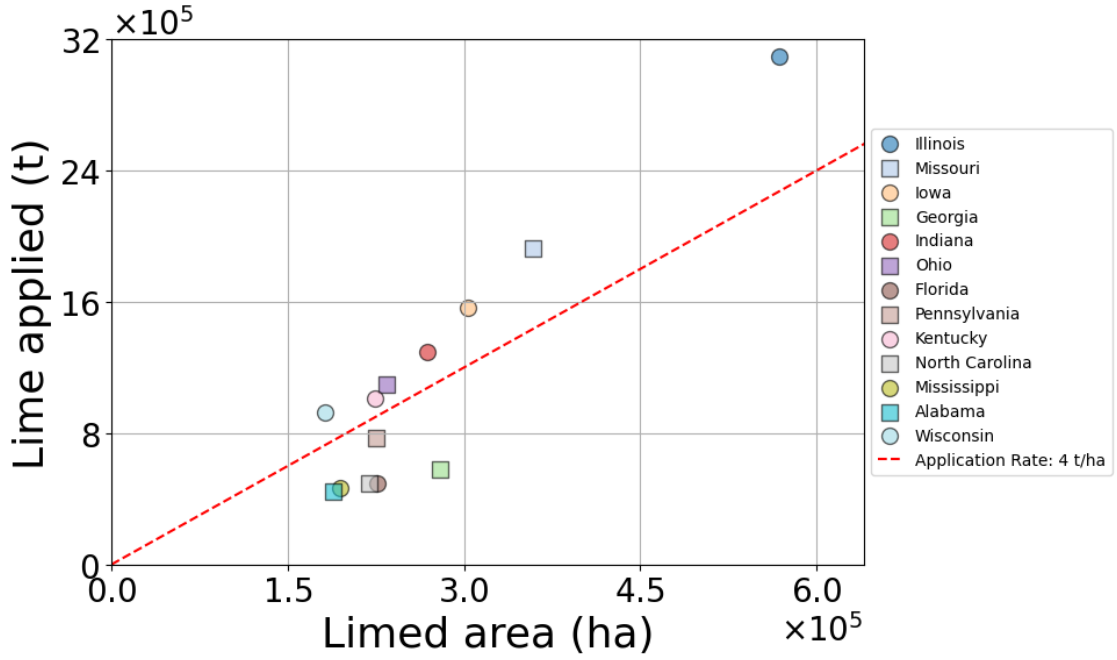

Fig. S3

Total limed weight (metric tons, t) versus total limed area (hectares, ha) for the 13 states with the highest average limed area. The values represent the averages from 1954 to 1987. The red dashed line indicates a lime application rate of 4 mertic tons per hectare.

Fig. S4

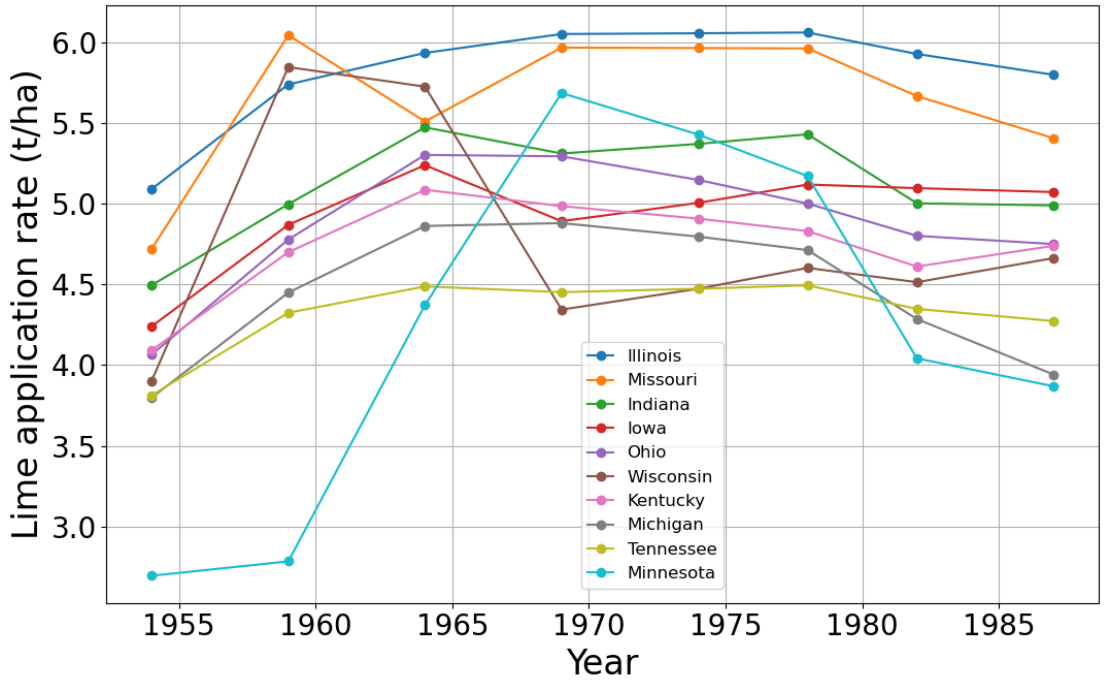

Fig. S4 Temporal variation in lime application rate (metric tons/hectare) for the top 10 states in the continental U.S. from 1954 to 1987. The states are ranked in descending order based on average lime application rate across the period, with Illinois having the largest area, followed by Missouri, Indiana, Iowa, Wisconsin, Kentucky, Michigan, Tennessee, Minnesota.

Random Forest Analysis

2.1 Feature selection

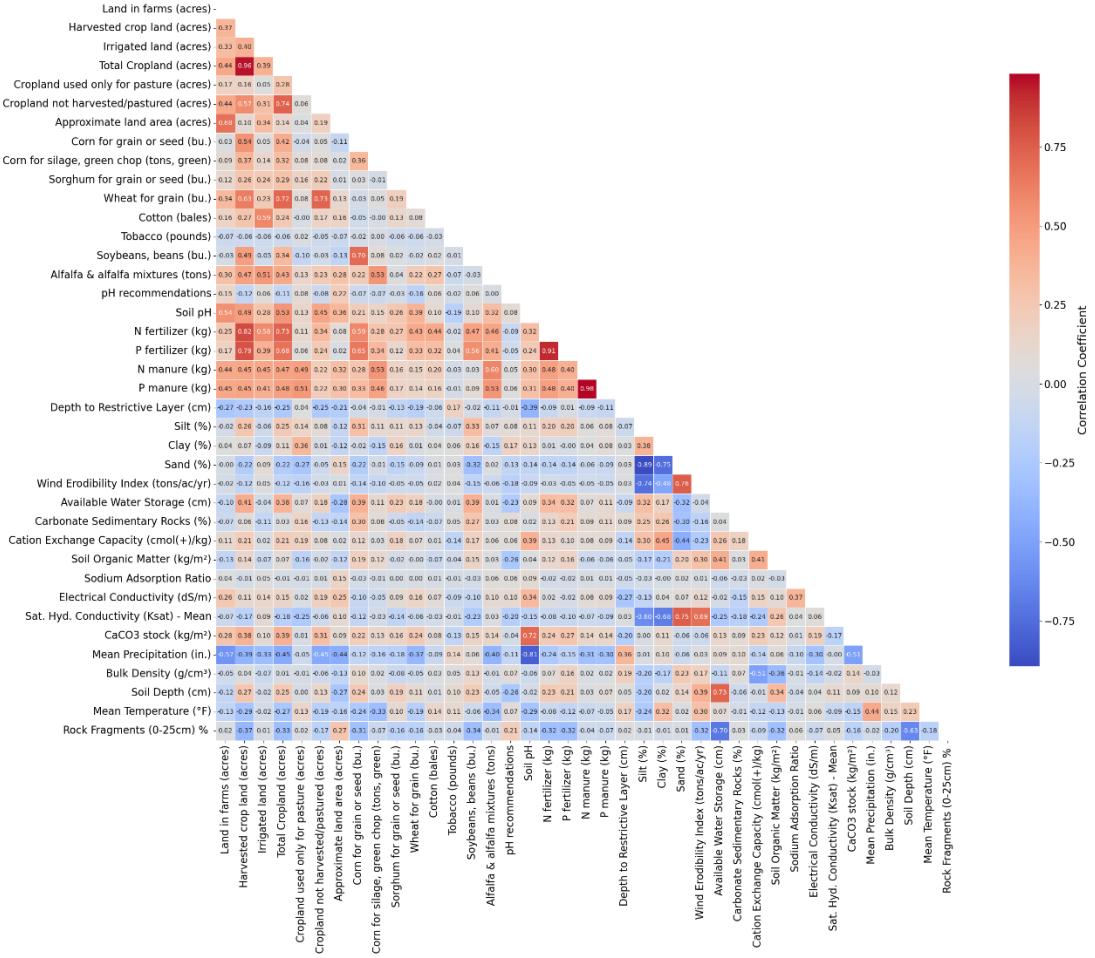

Fig. S5 Correlation matrix of predictors and Lime applied (tons), showing Pearson correlation coefficients. Each cell represents the correlation between two variables, with values closer to 1 (red) indicating a strong positive linear relationship, values closer to -1 (blue) indicating a strong negative linear relationship, and values near 0 (white) indicating no linear correlation. The heatmap highlights significant correlations, such as between N manure and P manure (0.96), and between N fertilizer and P fertilizer (0.91), which suggest potential collinearity among specific predictors.

Table S2 Variance Inflation Factor (VIF) Analysis of final list of predictor variables after removing collinear variables

Variable	VIF
N fertilizer (kg)	7.32
Soil pH	7.19
Sand (%)	7.09
Wind Erodibility Index (tons/ac/yr)	5.56
Corn for grain or seed (bu.)	5.05
Clay (%)	4.80
Mean Precipitation (in.)	4.70
Wheat for grain (bu.)	4.08
Sat. Hyd. Conductivity (Ksat) – Mean	3.86
N manure (kg)	3.69
Irrigated land (acres)	3.33
Cation Exchange Capacity (cmol(+)/kg)	3.21
Mean Temperature (°F)	3.15
Soybeans, beans (bu.)	3.02
Soil Organic Matter (kg/m²)	2.93
Land in farms (acres)	2.92
$CaCO_3$ stock (kg/m²)	2.88
Cropland not harvested/pastured (acres)	2.79
Rock Fragments (0–25 cm) (%)	2.79
Approximate land area (acres)	2.55

Variable	VIF
Alfalfa & alfalfa mixtures (tons)	2.50
Bulk Density (g/cm³)	2.50
Soil Depth (cm)	2.36
Corn for silage, green chop (tons, green)	2.11
Cotton (bales)	2.09
Cropland used only for pasture (acres)	1.96
Electrical Conductivity (dS/m)	1.57
Carbonate Sedimentary Rocks (%)	1.46
pH recommendations	1.37
Sorghum for grain or seed (bu.)	1.36
Depth to Restrictive Layer (cm)	1.35
Sodium Adsorption Ratio	1.20
Tobacco (pounds)	1.12

2.2 Random Forest modelling

- Lime applied per year

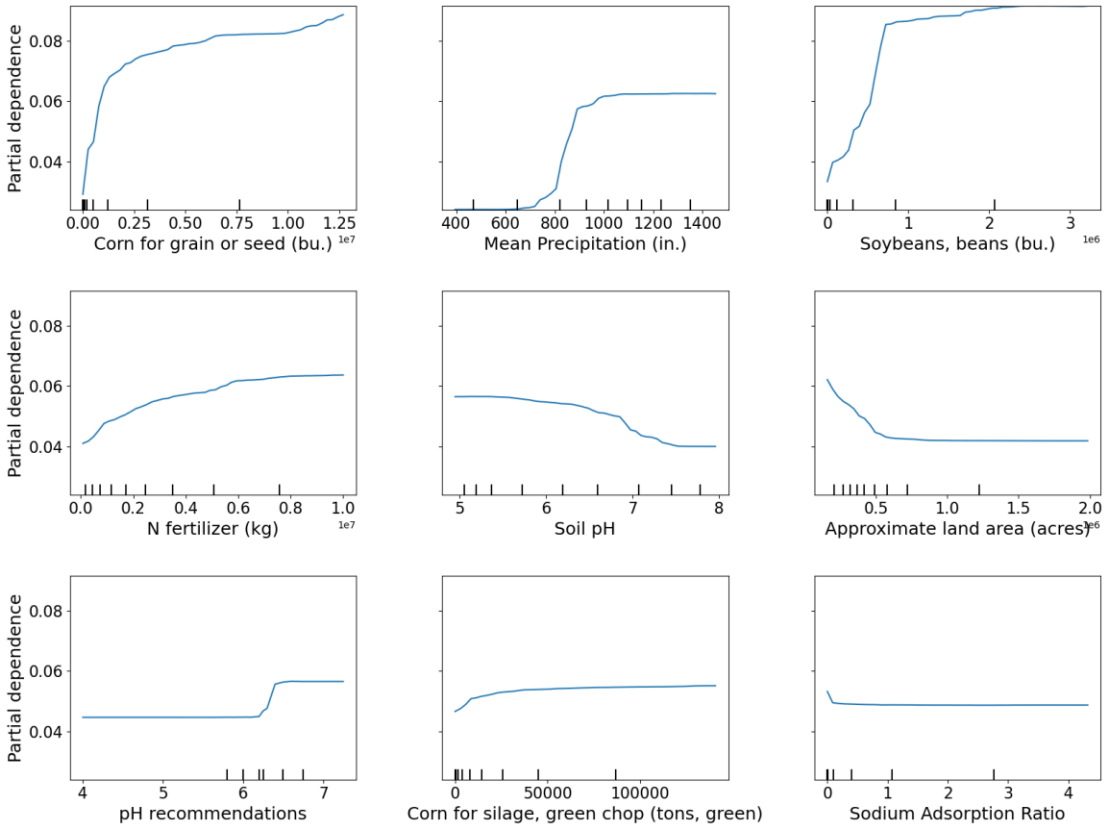

Fig. S6 The partial dependence plots of the variables that have the higher feature importances with annual lime application (metric tons per year) in each county, showing the effect of each variable on the prediction results.

-Lime application rate

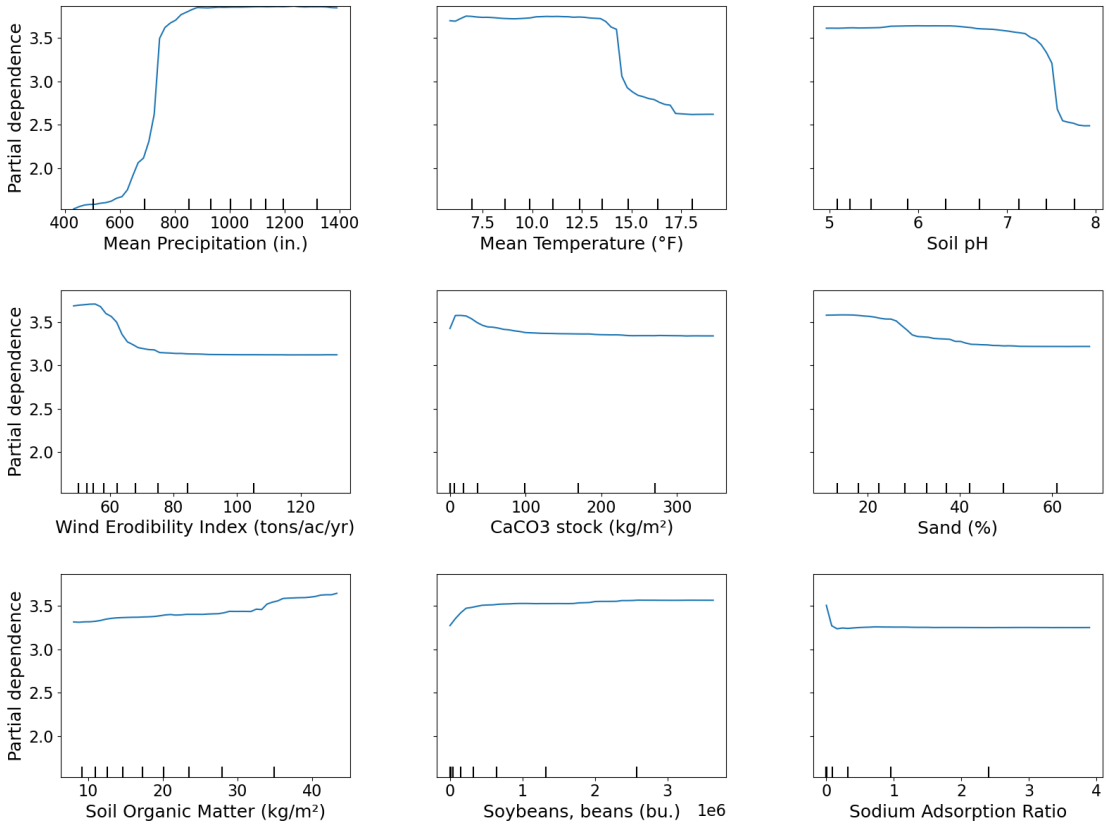

Fig. S7 The partial dependence plots of the variables that have the higher feature importances with lime application rate (limed applied per limed area; t/ha) in each county, showing the effect of each variable on the prediction results.

-Liming Frequency (limed area/cropland area)

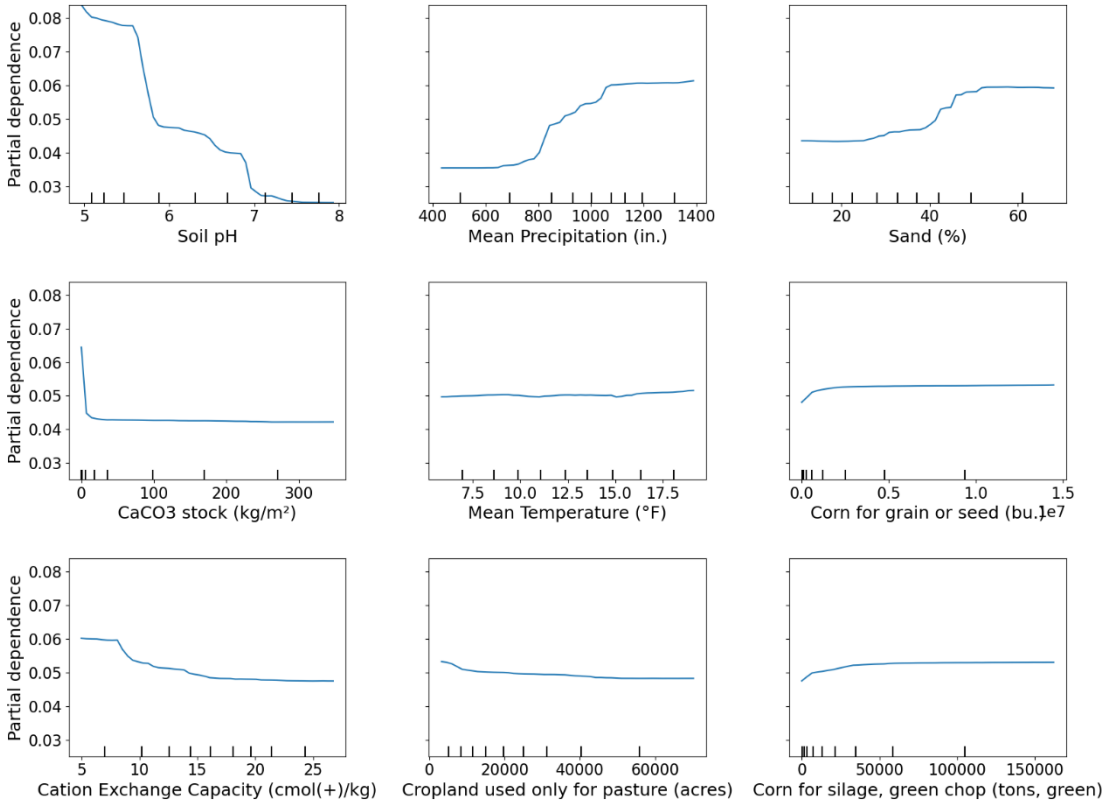

Fig. S8 The partial dependence plots of the variables that have the higher feature importances with liming frequency (limed area/cropland area) in each county, showing the effect of each variable on the prediction results.

2. Spatial Pattern of agricultural and environmental parameters.

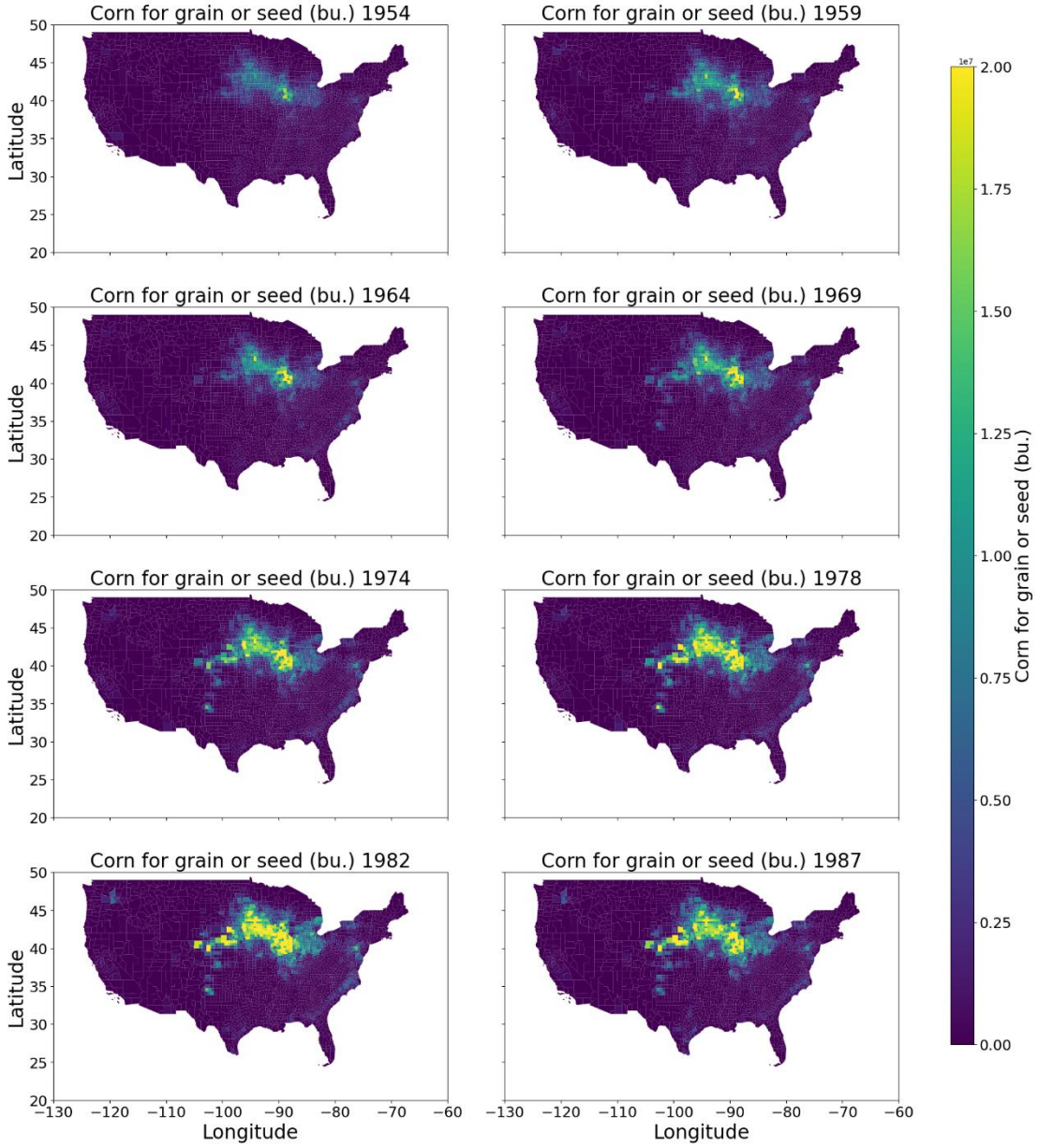

Fig. S9 Corn production for grain or seed (bu.) by county in the continental United States for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display the spatial distribution of corn production. Missing data is not shown and may be addressed through interpolation or spatial filling methods as described in the Methods section.

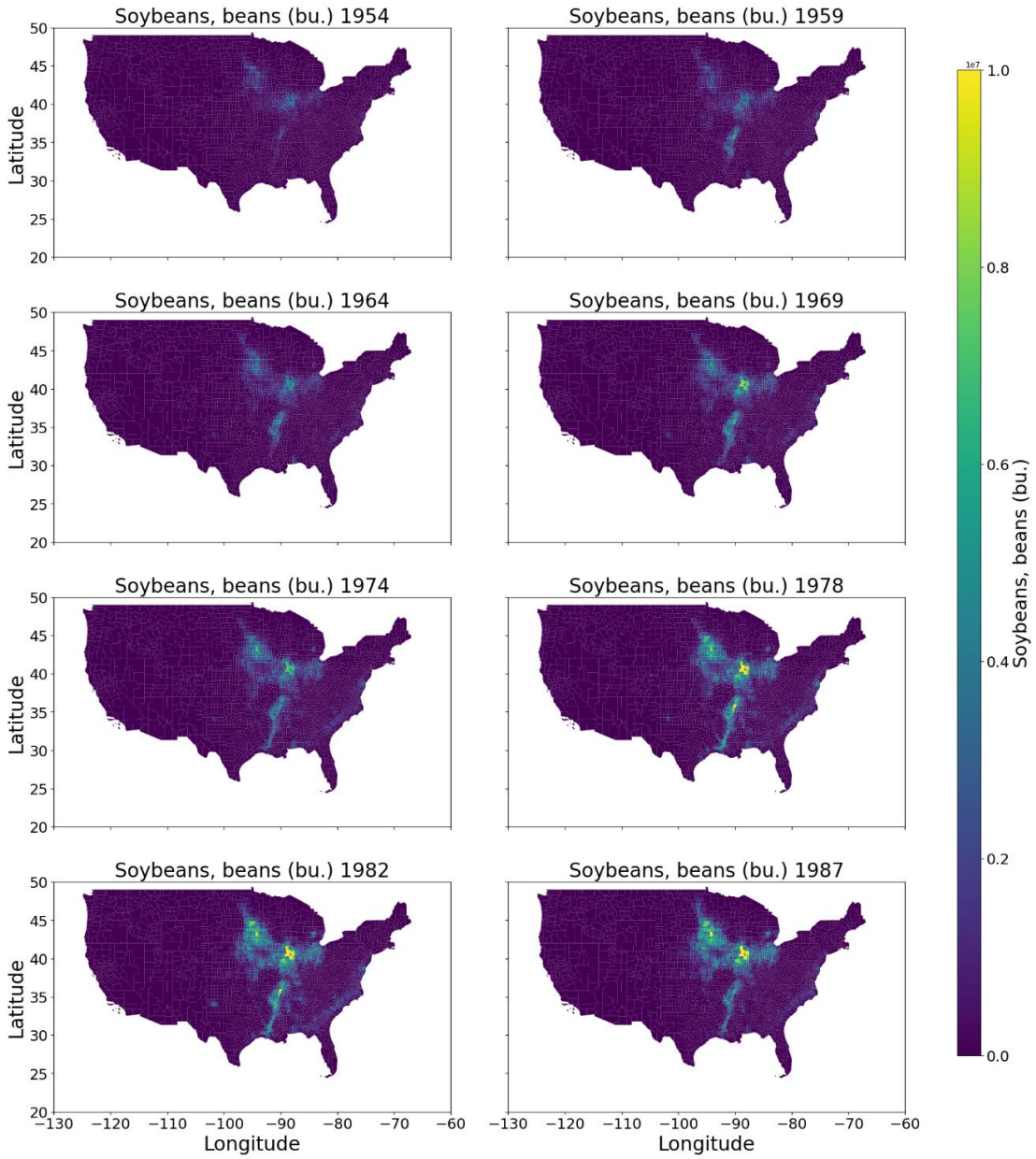

Fig. S10 Soybean production (bu.) by county in the continental United States for eight census years from 1954 to 1987, as reported by the Census Bureau. Missing data is not shown and may be addressed through interpolation or spatial filling methods, as described in the Methods section."

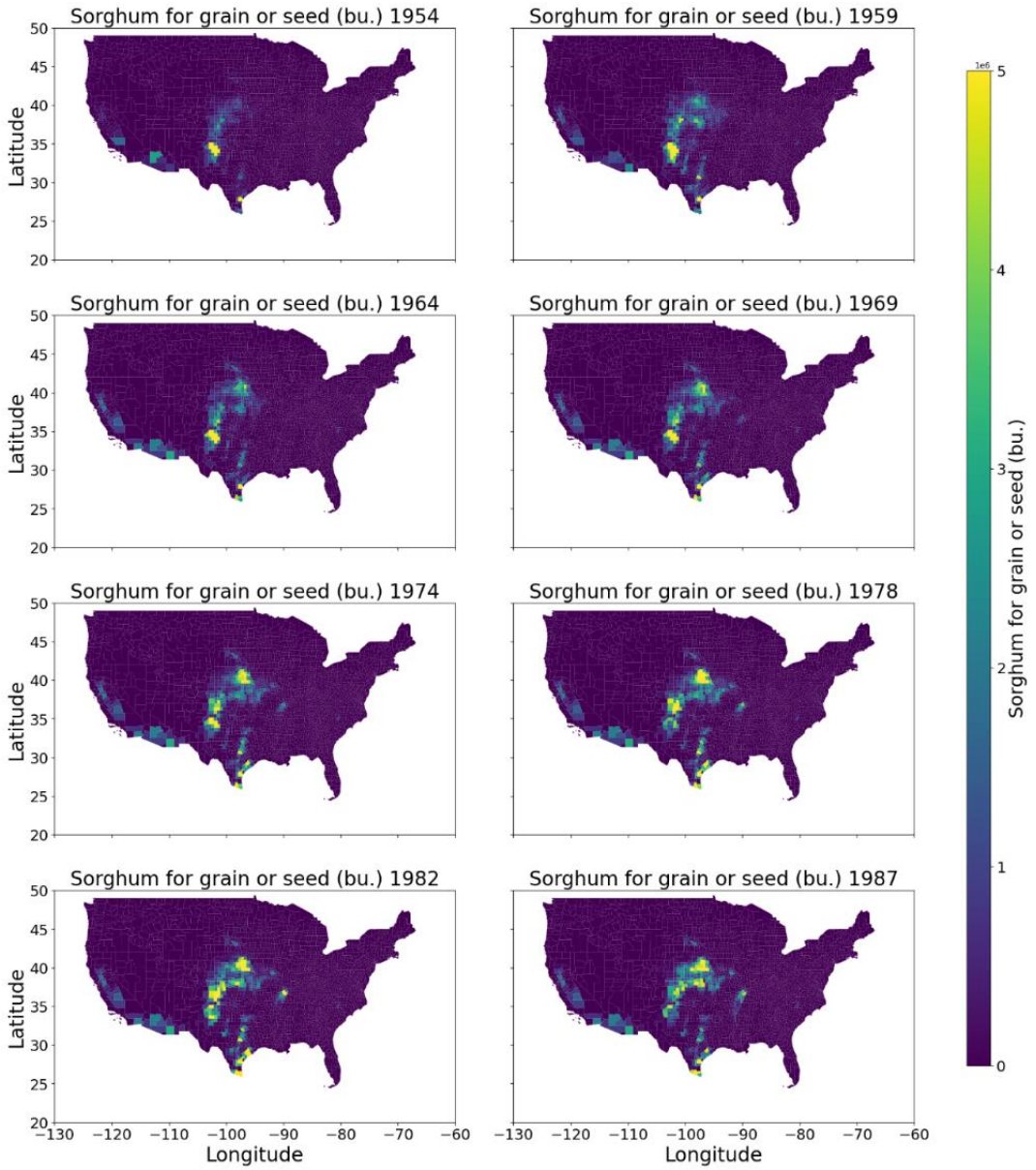

Fig. S11 Sorghum production for grain or seed (bu.) by county in the continental United States for eight census years from 1954 to 1987, as reported by the Census Bureau. Missing data is temporarily interpolated and spatially filled, as described in the Methods section.

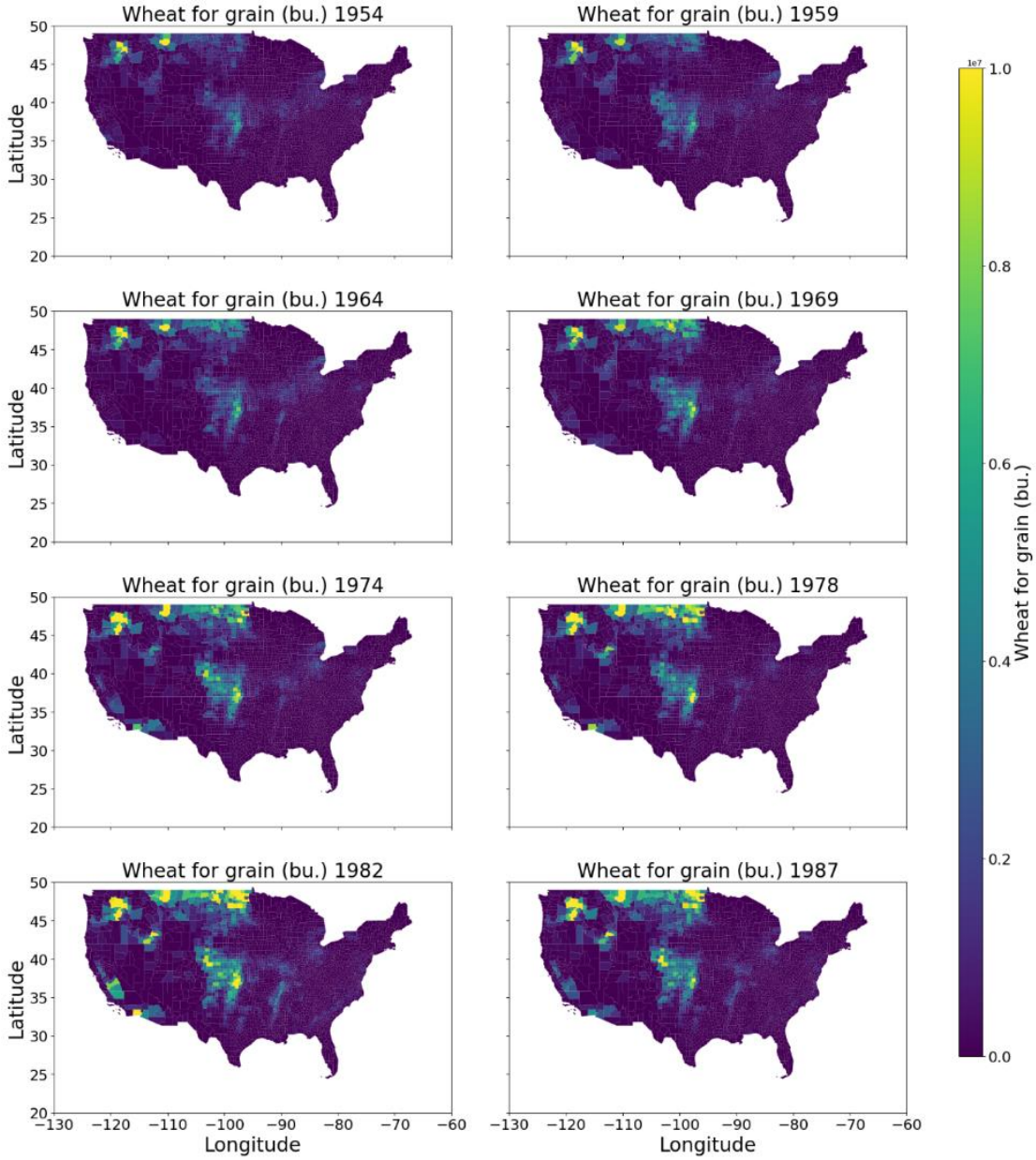

Fig. S12 Wheat production for grain (bu.) by county in the continental United States for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps depict the spatial distribution of wheat production. Missing data is temporarily interpolated and spatially filled, as described in the Methods section.

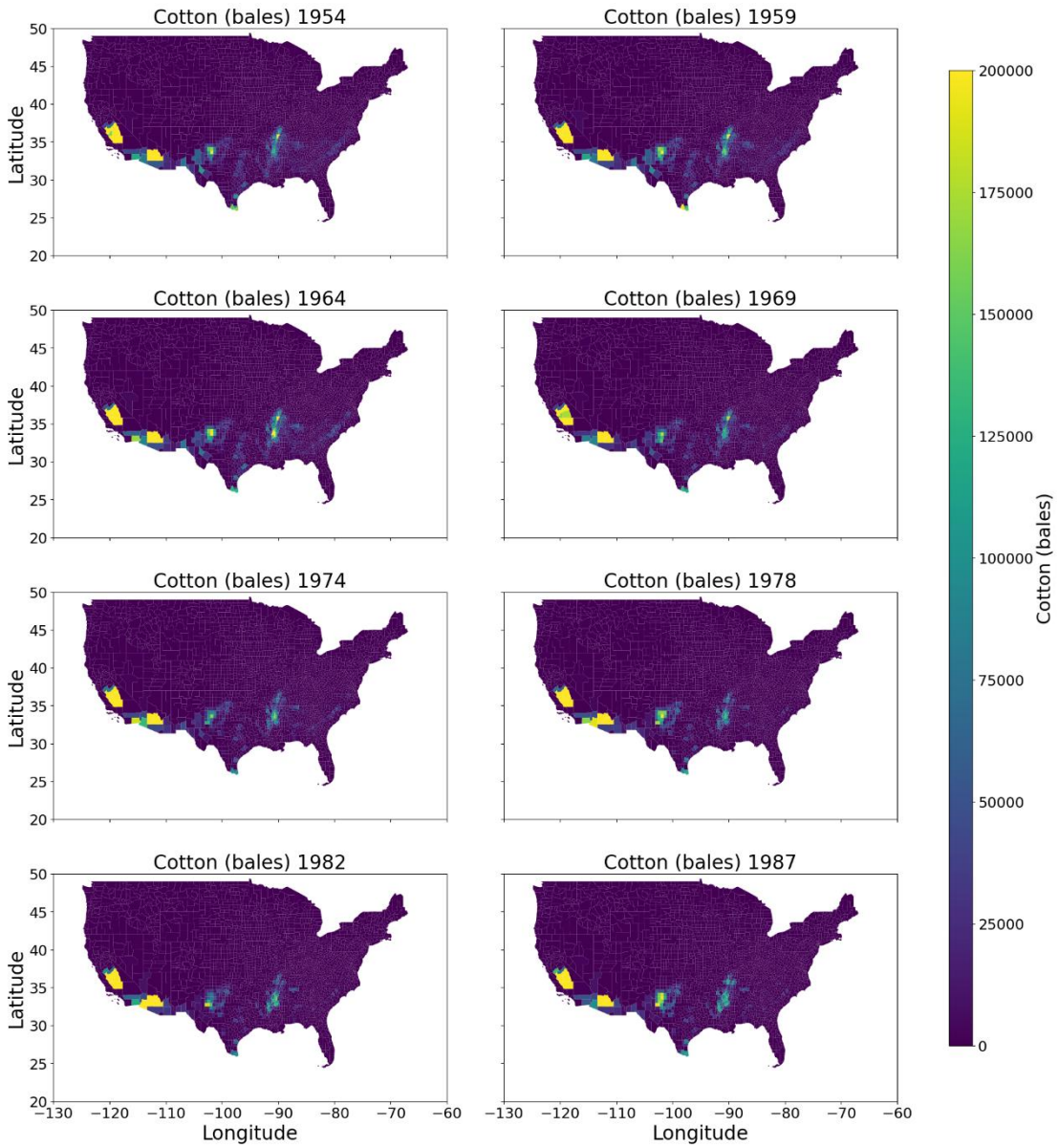

Fig. S13 Cotton production (bales) by county in the continental U.S. for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display the spatial variation of cotton production, with missing data temporarily interpolated and spatially filled, as described in the Methods section.

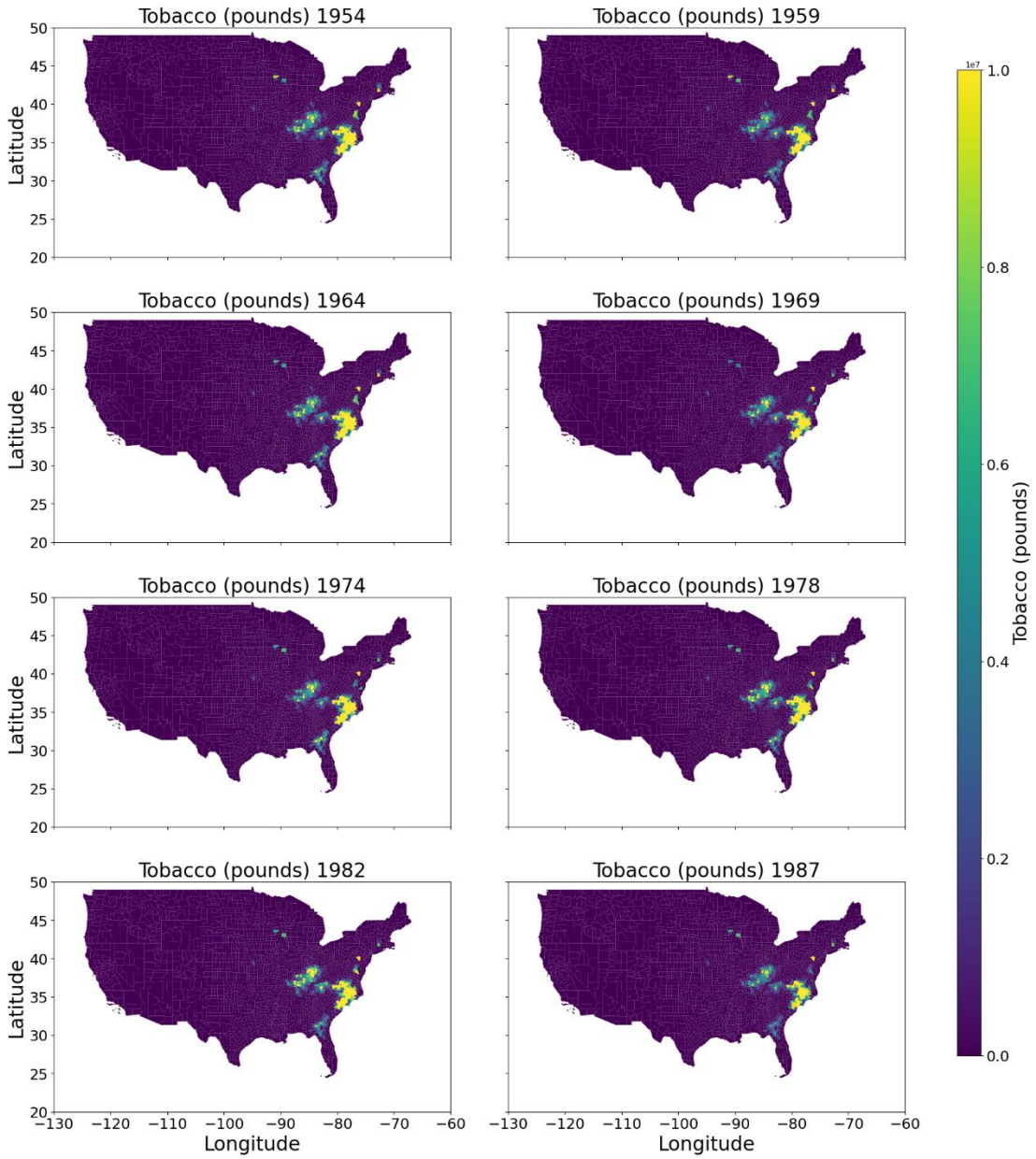

Fig. S14 Tobacco production (pounds) by county in the continental U.S. for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display the spatial variation of tobacco production across counties, with missing data temporarily interpolated and spatially filled, as described in the Methods section.

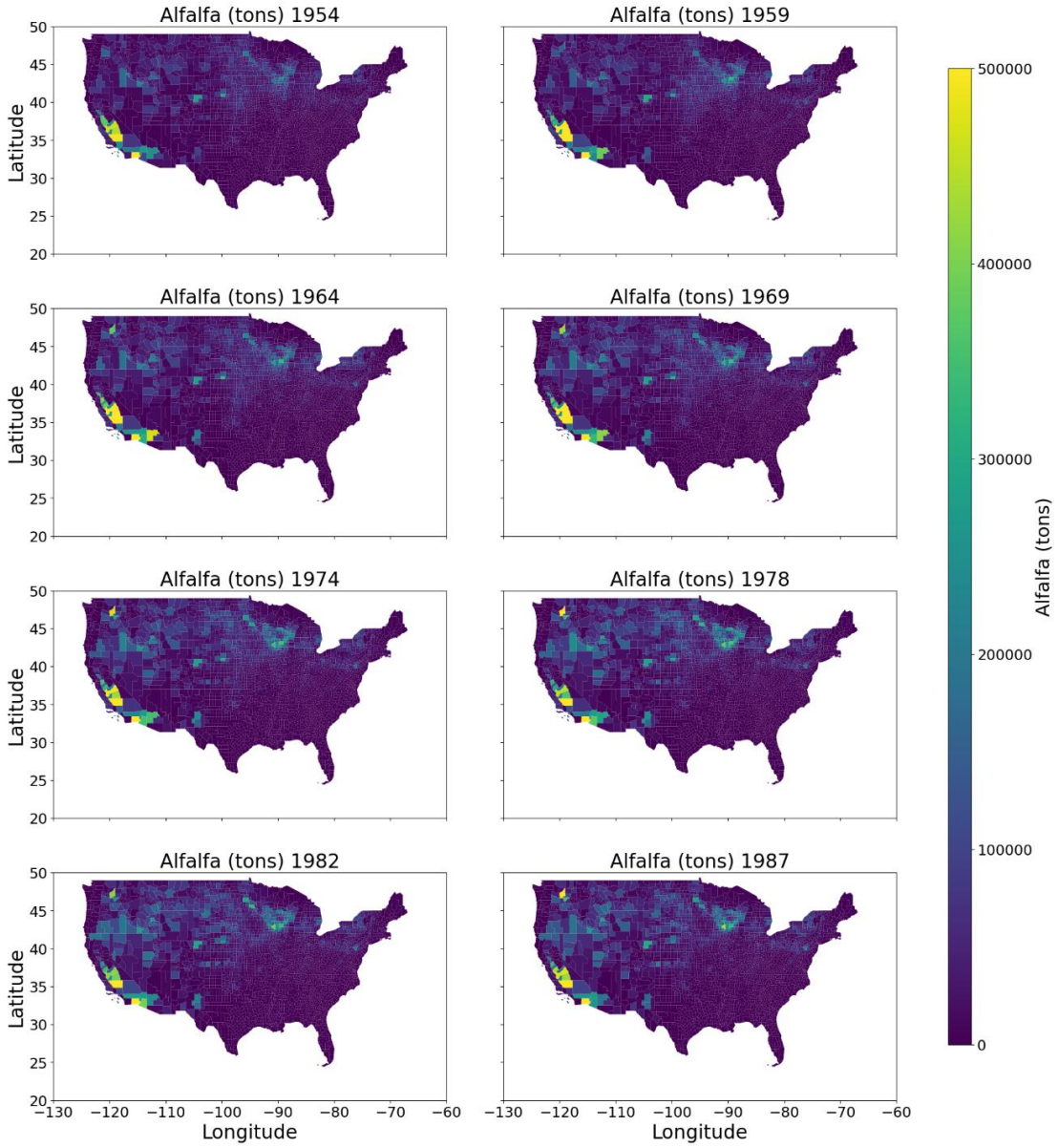

Fig. S15 Alfalfa production (tons) by county in the continental U.S. for eight census years from 1954 to 1987, as reported by the Census Bureau. The maps display the spatial variation of alfalfa production across counties, with missing data temporarily interpolated and spatially filled, as described in the Methods section.

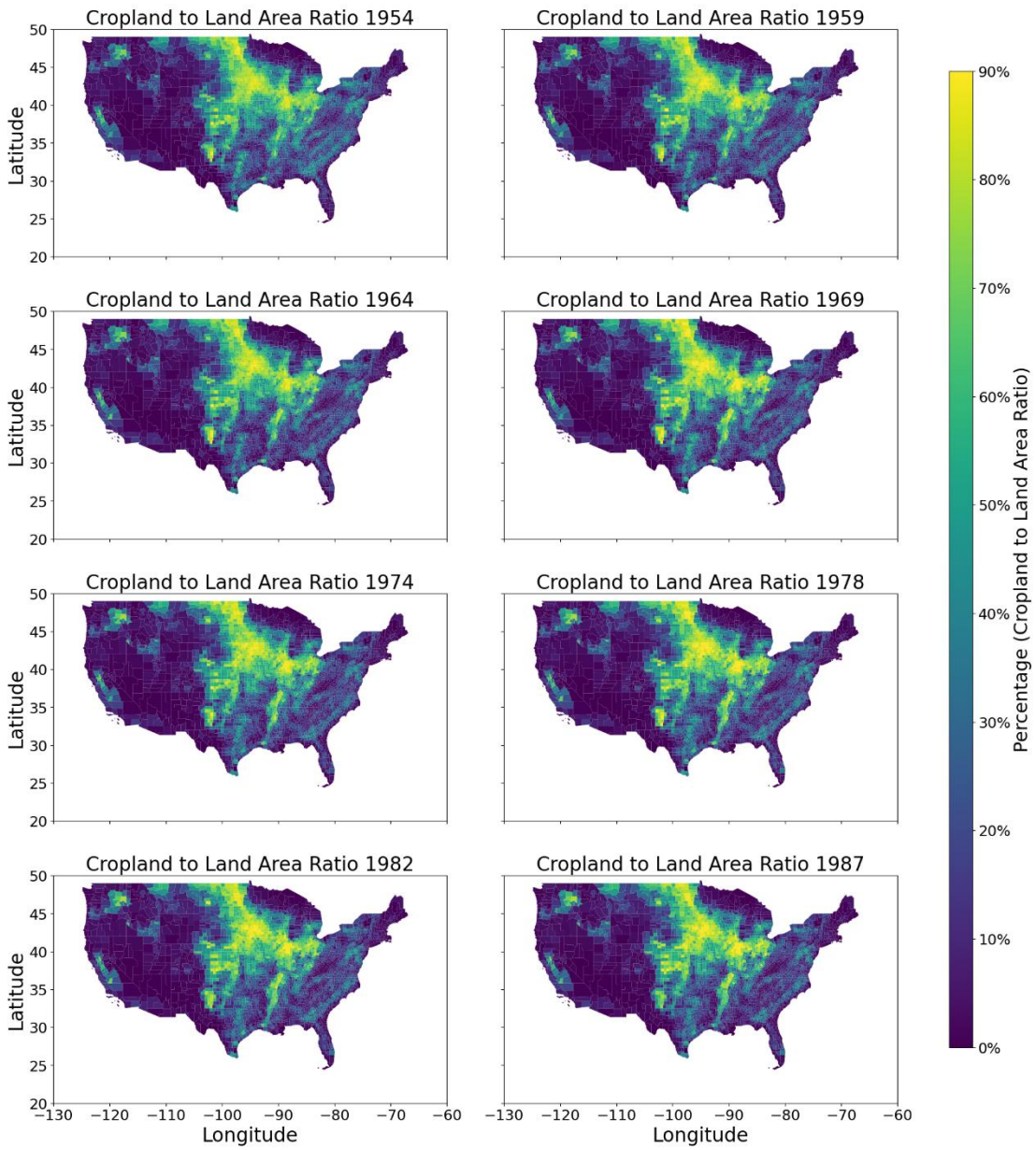

Fig. S16 Ratio of total cropland divided by the approximate land area per county for eight census years from 1954 to 1987, as reported by the Census Bureau. The approximate land area is calculated based on the shapefile of each county.

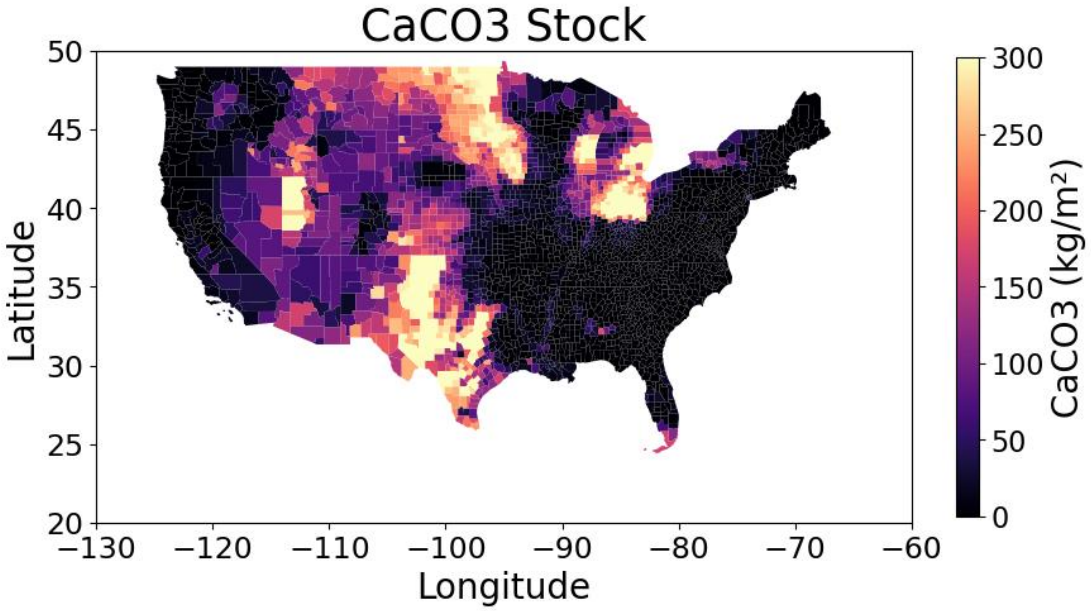

Fig. S17 Spatial distribution of $CaCO_3$ stock (kg/m²) across counties in the continental United States. The map illustrates areas with varying levels of $CaCO_3$ stock, with higher concentrations predominantly located in the central and western regions.

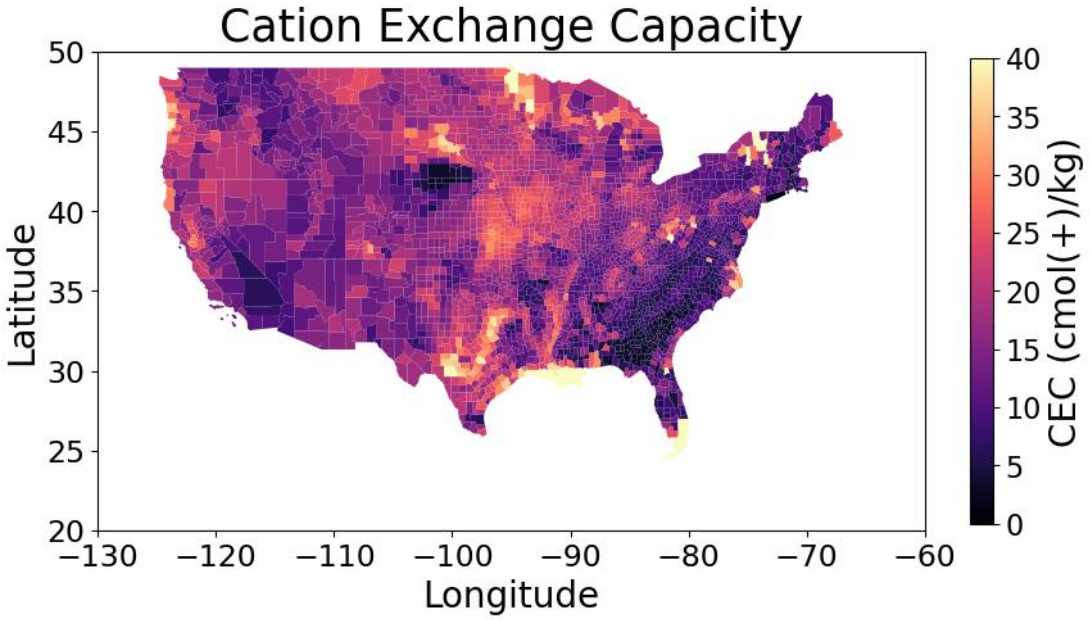

Fig. S18 Spatial distribution of Cation Exchange Capacity (CEC) (cmol(+)/kg) across counties in the continental United States. The map highlights regional variations in soil's ability to retain and exchange cations,

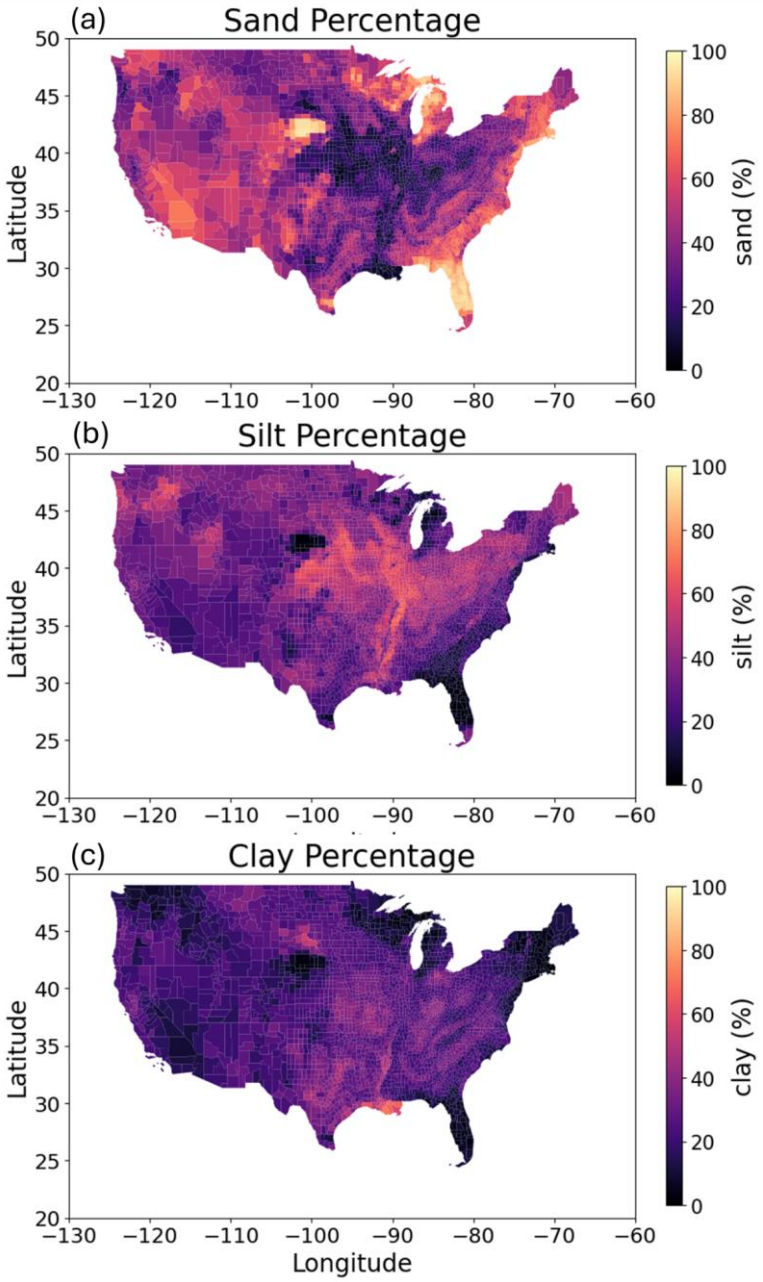

Fig. S19 Spatial distribution of percentage of (a) sand, (b) silt, (c) clay across counties in the continental United States. The map highlights regional variations in soil particle sizes.

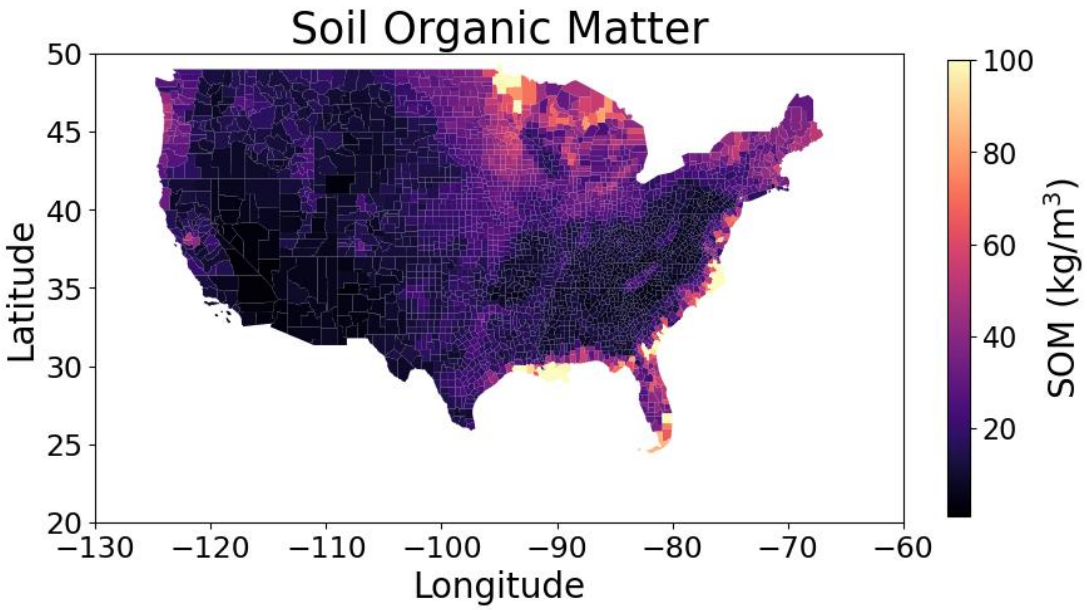

Fig. S20 Spatial distribution of soil organic matter (kg/m^3) across counties in the continental United States.

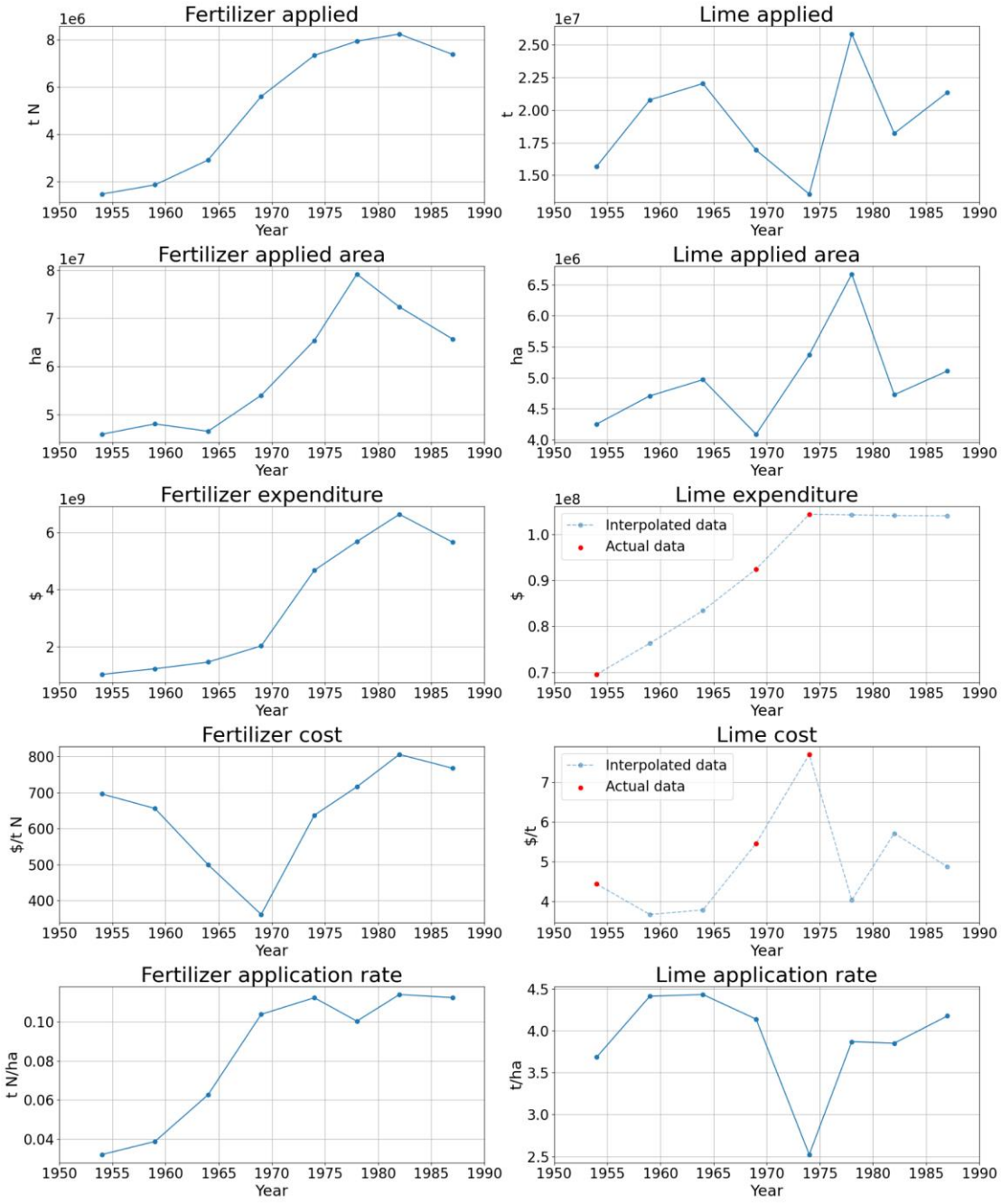

Fig. S21 Trends in total fertilizer and lime use, cost, area, and application rates in the United States from 1950 to 1990. Fertilizer applied (weight of N) is from the county-level data of Falcone (2021), while all the other variables are from the Census reports. Cost for fertilizer and lime is calculated by dividing the total expenditure to the amount of application. Application rate is calculated from dividing the total weight of application to the applied area. For lime expenditure and cost, red dots represent available actual data, and dotted blue lines represent interpolated estimates used to complete missing years.

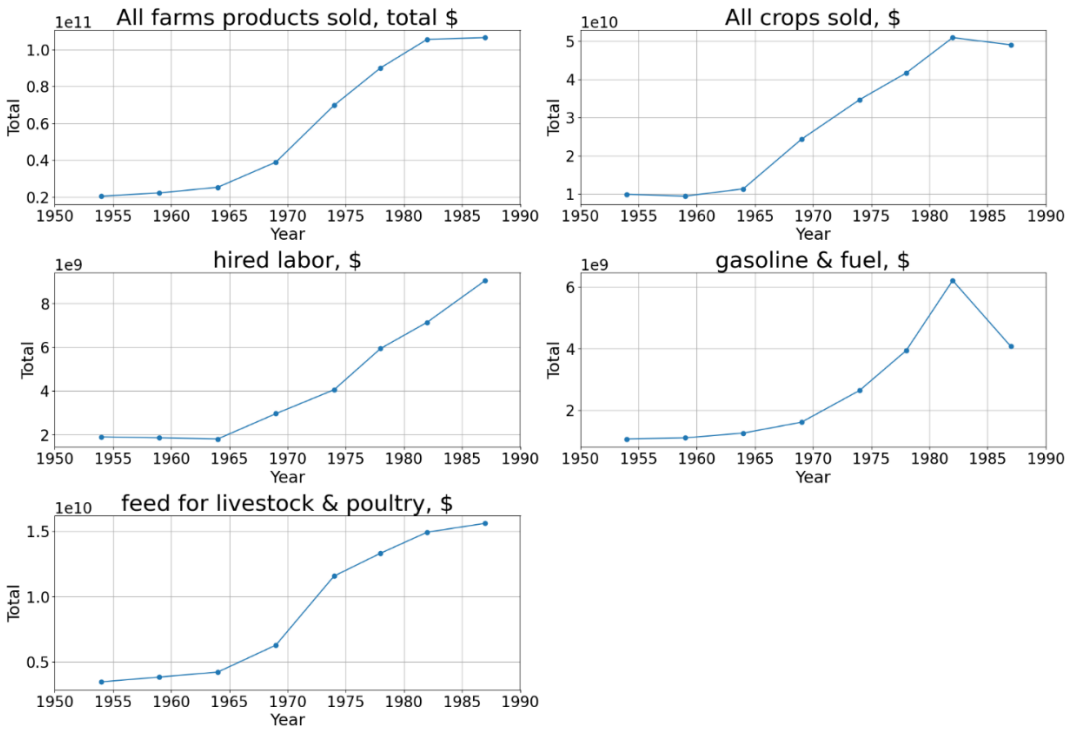

Fig. S22 Trends in total sales and input expenditures in U.S. agriculture from 1950 to 1990. This figure presents the inflation-agnostic trends (in nominal USD). While overall sales of agricultural products and crops steadily increased across the decades, input costs—especially labor, fuel, and feed—also rose, reflecting growing capital and operational intensities in U.S. farming.

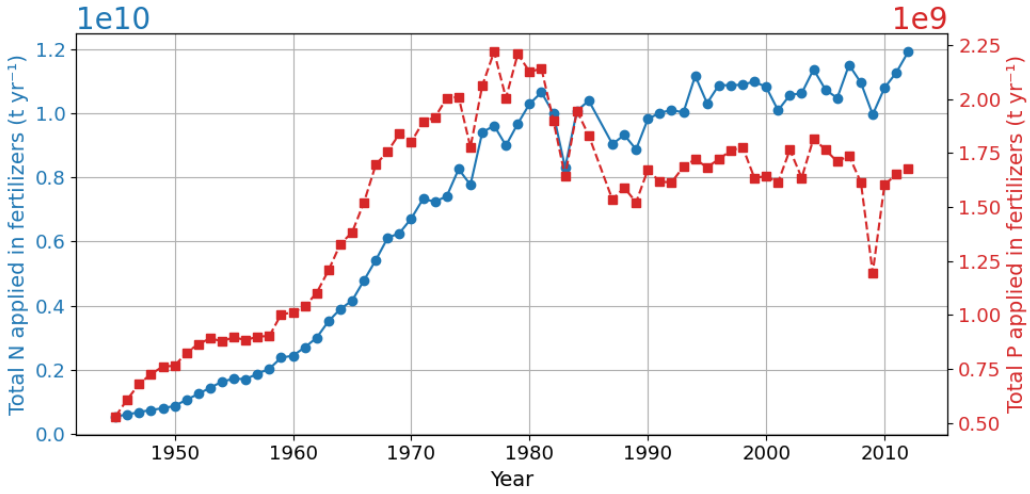

Fig. S23 Annual total N and P input via fertilizers (metric tons per year) within 1945-2012, based on the county-level total data from Falcone (2021).