Response to Review 1:

We thank the reviewer for scrutinizing our manuscript and providing the insightful comments and constructive suggestions, which greatly improve the quality of the manuscript. Please see our responses to the comments and related changes (highlighted in red) as follows.

Cao et al. present a comprehensive spatial dataset of historical nitrogen fertilizer use in the United States. The authors are to be commended for undertaking an ambitious project and integrating new types of data useful to the agriculture and biogeochemical modeling communities, including fertilizer type and timing of application. While the general methodological approach makes sense, my primary critique is that the rationale for the authors' specific choices is not always clear and the methodological description is confusing in some places.

A major limitation is that it appears fertilizer use by crop is spatialized to a cropland map that is not crop-specific. This needs to be made clear, and the limitations of the resulting spatial patterns discussed. Does state or county cropland area enter into the analysis? This could lead to some substantial biases in the data product.

Reply: Thanks to the reviewer to point out the unclear description about the cropland distribution map we used in this study. The cropland database developed by Yu and Lu (2017) includes annual cropland density maps and crop type distribution maps covering the conterminous US spanning from 1850 to 2016.

We have added the description about the historical land cover dataset developed by Yu and Lu, (2017) to the **2.4 Spatializing state-level crop-specific N fertilizer input to gridded maps** in the manuscript.

Page 7, line 197. For spatial analysis, we downscaled the imputed state-level crop-specific N management data to gridded maps based on 1 km × 1 km historical land

cover data (including crop density and crop type distribution maps) of the contiguous U.S. from 1850 to 2015 developed by Yu and Lu (2017) and Yu et al. (2018). The cropland density maps, by incorporating various sources of inventory data and high spatial resolution satellite images, were reconstructed to represent the area of cropped land each year while excluding summer idle/fallow. The crop type maps were reconstructed using satellite images and the USDA National Agricultural Statistics Service (NASS) survey data, and state-level land area of each crop type in each year is consistent with USDA survey. More details about cropland maps can be found in Yu and Lu (2017) and Yu et al. (2018).

References:

Yu, Z. and Lu, C. Q.: Historical cropland expansion and abandonment in the continental U.S. during 1850 to 2015, Glob. Ecol. Biogeogr., 27, 322-333, 2017, DOI: 10.1111/geb.12697.

Yu, Z., Lu, C., Cao, P., and Tian, H.: Long-term terrestrial carbon dynamics in the Midwestern United States during 1850-2015: Roles of land use and cover change and agricultural management, Glob. Change Biol., 1-18, 2018, DOI: 10.1111/gcb.14074.

Paragraph at 115: There is a reference here to eq. 1 as providing the state N fertilizer use rate of each crop, but eq. 1 is only total N consumption in each state.

Reply: thanks for pointing this out. In our study, equation 1 and equation 2 are not only applied in imputing national crop-specific N use rate, but also used in generating state-level crop-specific N use rate, and state-level consumption of Calcium Nitrate, Diammonium Phosphate, and Monoammonium Phosphate. To avoid the confusion, we reconstructed the equations and rephrased the sentence.

Page 4, line 117.

$$Raw\ data_i = \frac{Referenced\ trend_i}{Referenced\ trend_{i+1}} \times Raw\ data_{i+1} \tag{1}$$

where *Raw data* refers to the raw data that contains missing values, the *Referenced trend* refers to the complete data from which we extracted the changing trend that raw data can refer to, and *i* is the years between 1850 to 1930.

Page 4, line 125.

$$Raw\ data_{i+k} = \frac{Referenced\ trend_{i+k} \times Raw\ data_i}{Referenced\ trend_i} \times \frac{k-i}{j-i} + \frac{Referenced\ trend_{i+k} \times Raw\ data_j}{Referenced\ trend_j} \times \frac{j-k}{j-i}$$

(2)

where *Raw data* refers to the raw data that contains missing values, *Referenced trend* refers to the complete data that provides a reliable changing trend, the years *i* and *j* were the beginning and ending year of the gap, and the year *i+k* was the kth missing year.

Paragraph at 118: It's unclear what benefit is provided by the cubic spline approach, since it shows some strange values for Maryland in Fig. 2. Why bother with two different approaches?

Reply: We agree with the reviewer that it is important to stress on the benefit of adopting these two approaches under different situations. The nitrogen fertilizer use survey was periodically conducted by each state based on their own schedules. In addition, the survey schedules also varied among crop types within the state. Collectively, there are many gaps with irregular length in the century-long N fertilizer use data, and most of these gaps are less than three years. Cubic spline interpolation is advantaged in high efficiency (using "spline" command with "fmm" method in R). More importantly, cubic spline method imputes the missing values smoothly by considering the effects of points at sides of the gap based on the third-order polynomials. Nonetheless, we found that cubic spline often over-fits the N rates when the gaps are more than three years. Whereas the distance-weighted method is capable to reproduce the interannual variation of N use rates. In comparison, when the missing years are less than three years, the distance-weighted method may largely bias the gap-filled values when the trend of referenced rates is opposite to that of original Thus, we use cubic spline to gap-fill N use rates and avoid abrupt biases.

We have added the comparison of these two approaches into **11 Comparison between cubic spline method and distance_weighted method in interpolating the missing**

values in the supplementary file.

11 Comparison between cubic spline method and distance-weighted method in interpolating the missing values

We adpoted two methods, cubic spline method (for gaps less than three years) and distance-weighted method (Eq. 2, for gaps larger or equal to three years), to gap-fill the missing values in state-level crop-speicific N use rates. It is of high probility that the trend of referenced rates are reverse to the row rates in short periods (less than three years), which may produce abrupt biases while using distance-weighted method (Fig. S2a&b). whereas cubic spline could interpolate the missing values more smoothly due to it adopts the third-order polynomials to interpolate the missing values, in which the function fits through at most four points at side of the gap. In comparison, cubic spline may over-fits the N use rates and fails to capture the long-term interannual variation of the N use rates in long period (larger than three years) (Fig. S1c&d). Whereas the distance-weighted approach, which considers the contribution of the weight of the neatest value and uses the trend of corrected rate (Fig. S3c) as base-line, are capable to restore reproduce the interannual variation of N use rates.
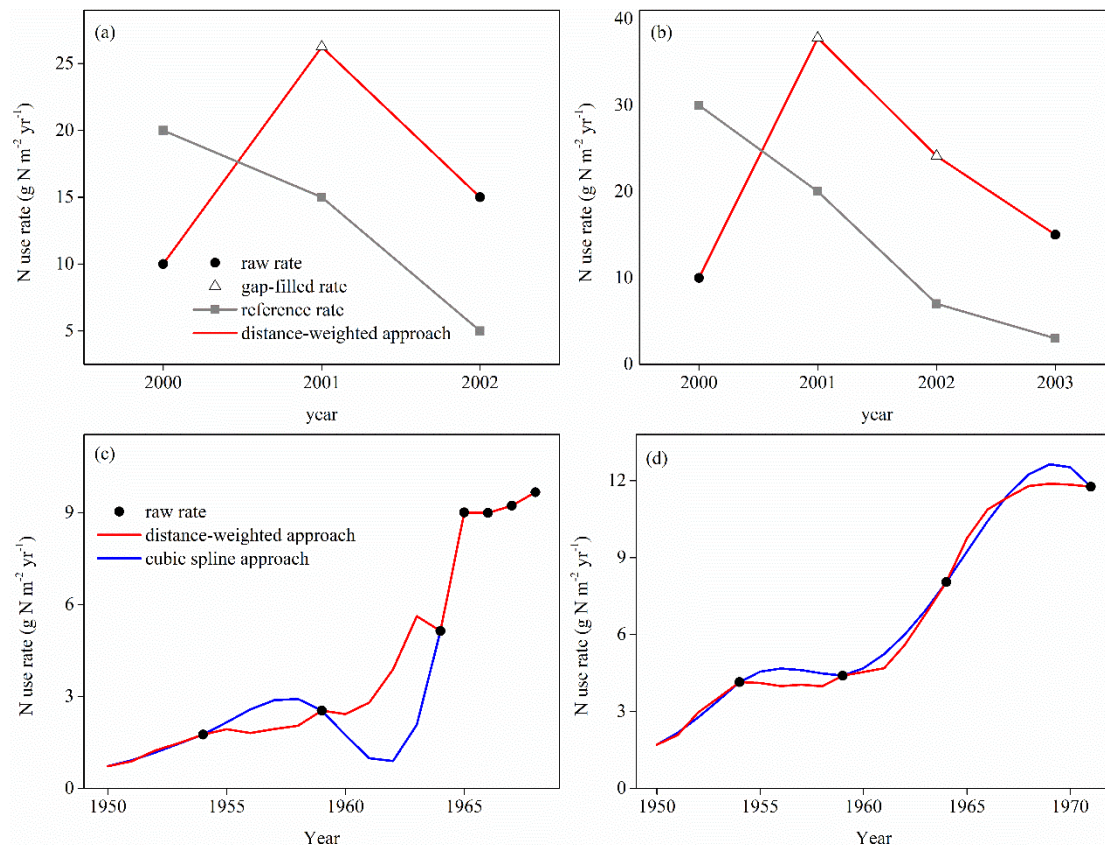
Supplementary figure S2: Difference between distance-weighted imputation and cubic spline interpolation on different length of gaps. (a) One missing year gap-filling based on distance-weighted method, (b) Two missing years gap-filling based on distance-weighted method, (c) Comparison of two gap-filling methods in N use rate of corn in Maryland, (d) Comparison of two gap-filling methods in N use rate of corn in Delaware. Grey line is referenced rate, red line is N fertilizer rate gap-filled by distance-weighted imputation, blue line is N fertilizer rate gap-filled by cubic spline approach, triangle is the gap-filled value, and dots are the raw data.

Paragraph at 135: The specification of these regression models is unclear from the text - please show examples using equations.

Reply: The details of regression models and how they were applied in gap-filling state-level crop-specific N use rates have been added to the supplementary material. Please see more details in section **4 Regression models (Table S4)** and section **12 Gap-filling the raw state-level crop-specific N use rates with referenced N use rates by regression model (Fig S3)** in the supplementary file.

4 Regression models

The referenced state-level crop-specific N use rates generated in section 2.1 combined national crop-specific N use rates with state-level N use rates, which can potentially provide the referenced interannual variation and temporal trend for gap-filling the missing values in the raw state-level crop-specific N use rates. But since national crop-specific N use rates and state-level N use rates are the average of N used by a given crop across the nation and by all crops within a given state, their combination, here we named the referenced state-level crop-specific N use rates, mismatchs the amplitidue and changing rate of the state-level crop-specific N use rates more or less (Fig. S2a). So we built regression models (Table S4) between the referenced crop-specific rates and raw state-level crop-specific rates to tune these two data sets to the same amplitude and changing rate (Fig. S2b&c).

Supplementary table S4. Regression models used to gap-fill raw state-level crop-specific N use rates

Model	Equation
Quadratic Model	$y = ax^2 + bx + c$
Cubic Mode	$y = ax^3 + bx^2 + cx + d$
Exponential Model	$y = ae^x + b$
Logarithmic Model	$y = alog(x) + b$

Where y is the raw state-level crop-speicific N use rates, x is the referenced state-level crop-specific N use rates.

12 Gap-filling the raw state-level crop-specific N use rates with referernced N use rates by regression model

We combined the national crop-specific N use rates and state-level N use rates as the reference for gap-filling the raw data of state-level crop-specific N use rates. However, there are discrepancies in the amplitude and the changing rate between the referenced N use rates and raw N use rates (Fig S3a), which produce abrupt changes (outliers) during the gap-filling. Thus, we first built the regression models (Table S4) between referenced rates and raw rates (Fig. S3b) to make the correction on the amplitude and changing rate of referenced rates (Fig. S3c). Then we gap-filled the missing values

based on the corrected reference (Fig. S3d).
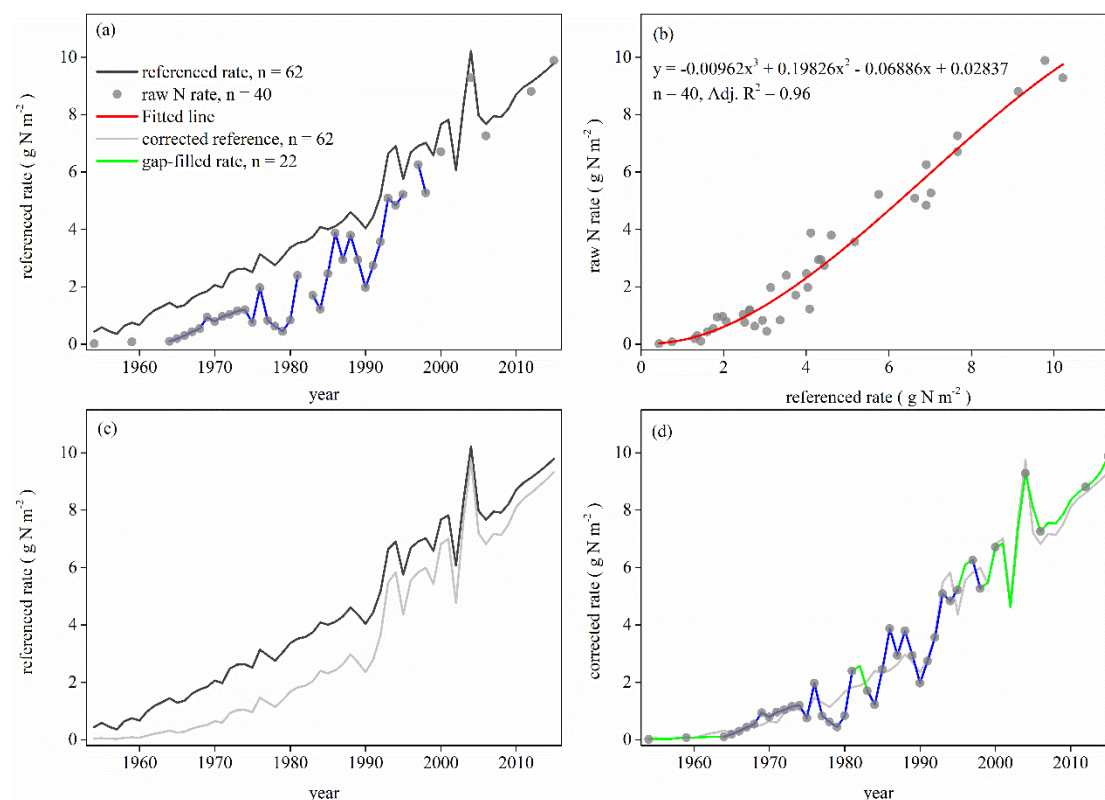
Supplementary figure S3. Gap-fill the missing N use rate of Spring Wheat in South Dakoda Spring wheat during 1954 to 2015. (a) Referenced rate and raw data (blue line means there is no missing years between the data points). (b) Regression model we built between raw data and referenced rate, which was used to correct referenced rate for better magnitude and changing rate (The Cubic Model is the best regression model with high R^2 of 0.96). (c) the referenced rate (black line) and corrected reference (gray line) based on cubic model in fig b. (d) Final product of gap-filled values (green line) by using weighted-distance approach and corrected reference.

Overall, it should be clearer if and how the derived rates are (or are not) consistent with state and national consumption estimates.

Reply: The development of state-level crop-specific nitrogen (N) use rates in this study was controlled by the state agricultural N consumption (excluding nonfarm use and permanent pasture use, supplementary table S2 and fig. S1), which means sum of state N consumption from the derived rates should theoretically equals to the state agricultural N consumption provided by the USDA. We first gap-filled the N use rates

of nine major crops (corn, soybean, winter wheat, spring wheat, cotton, sorghum, rice, barley, and durum wheat) of each state, then we derived state-level N use rates of cropland pasture from 1945 to 2015. We assume the rest of N fertilizer in each state was applied to other crops, which vary widely among states. By keeping the sum of crop-specific fertilizer use consistent to state total, we ensure our data is consistent with survey-based state and national agricultural fertilizer consumption in magnitude over time (excluding non-farm use and permanent pasture use).

Pages 5-6: Is the Agronomy Guide only used to populate Table 1, whereas the timing maps are coming from the survey? This could be clearer. Also, is it fair to say the Agronomy Guide is providing recommended timing or is this based on any survey data? The table legend should be more explicit.

Reply: Nitrogen fertilizer application timing (Table S4) and application timings for 11 N fertilizer types (Table 1 and Table S7) are different. Nitrogen fertilizer application timing is state-level crop-specific timing information, which is based on the survey conducted by USDA-ERS. This timing information was used for splitting annual state-level crop-specific N use rates into four timings (Timing maps, Fig. 5), but there is no information for NH_4^+-N/NO_3^--N ratio. The application timing for 11 N fertilizer types was derived from the Agronomy Guide and a survey in Minnesota (Bierman et al., 2012). It provides the reference to split annual consumption of each N fertilizer type (e.g. Anhydrous Ammonia, Ammonium Nitrate, Urea, etc.) into four timings. We adopted it to further divide state-level crop-specific N use at each timing to NH_4^+-N and NO_3^--N use rates. Such information can be found in the methodology and the workflow chart in the manuscript. Here we provide a flowchart to help illustrate these processes.
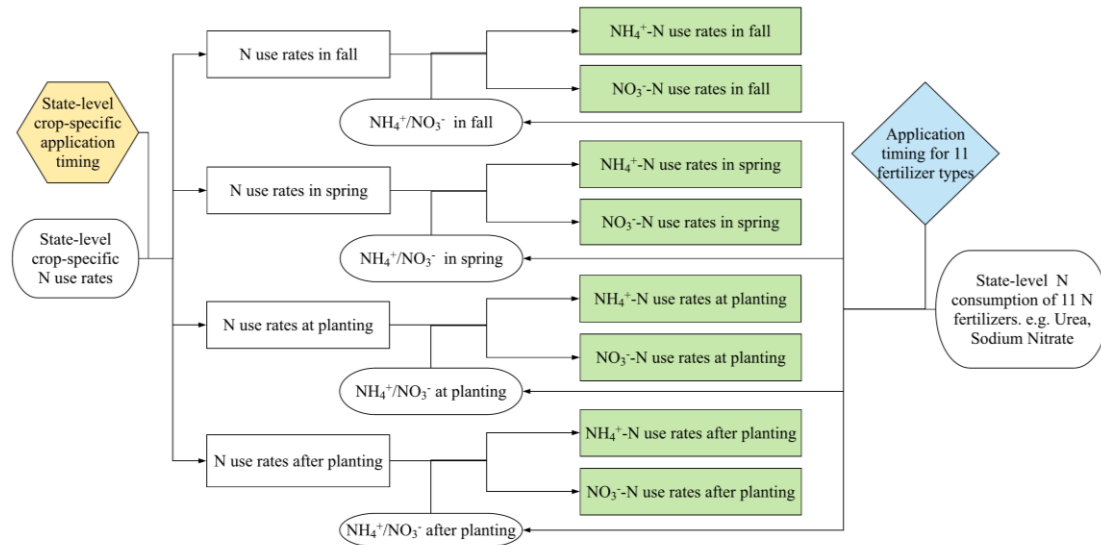
Section 2.4 - Please provide more background about the historical land cover dataset, and its coverage for different land cover and crop types (see earlier comment).

Reply: We have added the description about the historical land cover dataset developed by Yu and Lu, (2017) to the **2.4 Spatializing state-level crop-specific N fertilizer input to gridded maps** in the manuscript. Please see the response to earlier comment.

Paragraph at line 216: Histograms across total land area or agricultural area could be used to show graphically the change in distribution.

Reply: We agree with the reviewer's suggestion for showing the N fertilizer change in distribution. The national N fertilizer use rates of the continental U.S. has been continuously increasing since 1850. This general increase, however, may vary among different regions due to the difference in crop cultivation. Thus, we based on the N fertilizer use rates (per land area) of 1960, 1980, and 2015 to discuss the change of the rates distribution across the continental U.S. We didn't include the year 2000 because the rates increased very slowly after the 2000s. We counted the numbers of grid cells at every 0.1 g N m^{-2} yr^{-2} interval from 0 to 21 g N m^{-2} yr^{-1} and divided the numbers by total grid cells to generate the frequency in percentage. We extended the text in **3.1**

Nitrogen fertilizer application rate.

Page 7, line 210. In 1960, most agricultural area received low N fertilizer input (less than 6 g N m^{-2} yr^{-1}), while due to expansion of cropland and rapid increase in N fertilizer use from 1960 to 1980, the area receiving larger N input (2-8 g N m^{-2} yr^{-1}) greatly increased in 1980, and frequency peaked at a few high N input values, which may be caused by intensive fertilizer use in certain crop types (e.g., maize and rice). In 2015, we found less areas were characterized by low to medium-level fertilizer input (< 8 g N m^{-2} yr^{-1}), and high-level N fertilizer application (8-18 g N m^{-2} yr^{-1}) was more widespread compared to 1960 and 1980 (insert figure in Fig. 2a).
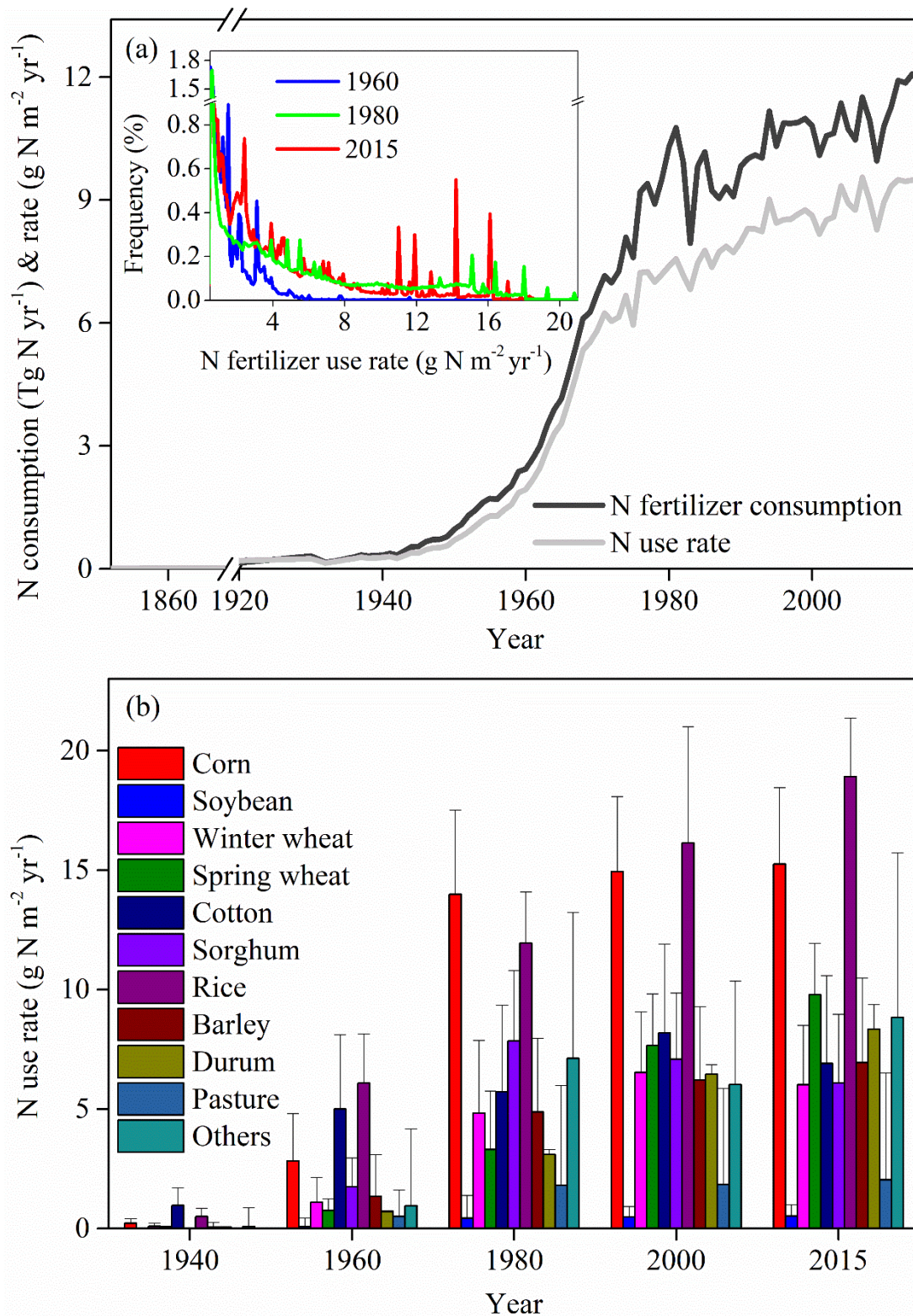
Figure 2: Time series of N fertilizer use in the United States: (a) national commercial N fertilizer consumption and use rate of the continental U.S. during 1850-2015 derived from this study (figure inset provides the N fertilizer rates distribution across the continental U.S. in the year 1960, 1980, and 2015); (b) national crop-specific N fertilizer application rate (error bars indicate standard error across states in one year).

Since many numbers are reported as rates over all land area, the results are not immediately accessible to those who may be interested in how cropland application rates have changed over time (although these numbers could be backed out of the dataset). I think the authors could demonstrate the richness of the dataset by highlighting a case study, e.g. maize in Iowa, and then show how the rates (per maize area) have changed.

Reply: We appreciate the suggestion for highlighting a case study of rate change of a crop within state. However, our study aims at spatializing the state-level crop-specific survey data of N fertilizer use rates (per crop area) based on cropland area and distribution maps. The raw data can be downloaded from USDA-ERS (https://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx), from which the changes of fertilizer uses can be directly estimated. We also agree there is large spatial heterogeneity in N fertilizer use rate even for one crop type in one state, but sub-state crop-specific data is unavailable so far. We expect to have more information in the future, to improve our data set and to inform modeling assessment.

Finally, it would be helpful to show more comparisons with other N fertilizer maps, including global products which may rely upon coarser land use maps. The authors have their own map that should be compared against, for example. For this, it would be good to put the datasets on a common grid and also include difference maps.

Reply: We appreciate the reviewer's constructive suggestion in improving the manuscript. Three more N fertilizer use data sets have been added to compare with our data set. One is county-level N fertilizer use in the continental U.S. developed by IPNI (NuGIS), another two are global N fertilizer use derived from IFA and FAO, respectively.

We have add the comparison of other three datasets to our data in **4.1 Comparison**

with existing N fertilizer maps in the manuscript.

Page 9, line 310. In order to examine the validity of our data, we compared our product with other studies. The temporal variations and spatial patterns of N fertilizer use were compared among our study, one regional data set from IPNI (2018), and two global data sets from Lu and Tian (2017) and Nishina et al. (2017) (Fig. 8 and Fig. 9). IPNI (2018) developed the N fertilizer use data set of the U.S. spanning from 1987 to 2012 by adopting method in Ruddy et al. (2006). It is based on county-level N consumption data derived from state-level N fertilizer sale and agricultural land acreage. Lu and Tian (2017) and Nishana et al. (2017) derived the N fertilizer use data from the International Fertilizer Association (IFA) and the Food and Agriculture Organization of the United Nations (FAO), respectively. Four data sets showed the same inter-annual variation of national N fertilizer consumption pattern from 1961 to 2014 but small difference in magnitude (Fig. 8). The N fertilizer consumption in our study is a little smaller than the value from IPNI (2018) except the year 2001 because our study excluded the permanent pasture. The historical N fertilizer use amount in our study is slightly lower than Lu and Tian (2017). This is because the N fertilizer data from Lu and Tian (2017) contains permanent pasture and nonfarm N fertilizer use and also covers territories of the U.S. rather than continental US. Whereas our data is marginally higher than Nishina et al. (2017) in most years. The overall spatial pattern of N fertilizer use rate in the continental U.S. in 2010 from our study is similar to other three maps (Fig. 9), indicating the hotspots of N fertilizer use occurred in the Midwest, West and East Coast, and some parts of the Northwest, the Southern Great Plains. Nevertheless, discrepancies were also found in the magnitude of N use rates between our data and the other studies. There were only a few states in Lu and Tian (2017) and no state in Nishina et al. (2017) applied over 7 g N m^{-2} yr^{-1} in the year 2010 (Fig. 9c&d). In comparison, two global data sets showed many regions received fertilizer at 2-4 g N m^{-2} yr^{-1}, in which very little N fertilizer application was indicated in this study and IPNI (2018). It is because the global cropland maps used in these two global N fertilizer data sets were based on HYDE 3.2 (Klein Gildewijk, 2006) in Lu and Tian (2017) and HYDE 3.1 (Hurrt et al., 2011) in Nishina et al. (2017), respectively. The HYDE data overestimated cropland percentage

in low crop-coverage states while underestimating crop percentage in the US Corn Belt (Yu and Lu, 2017). This explains the different spatial patterns revealed by these four N fertilizer data sets that although their total consumptions are very close nationally. By comparison, our data considers cross-crop divergence in N fertilizer use, covers longer period, and shows more details with finer resolution.
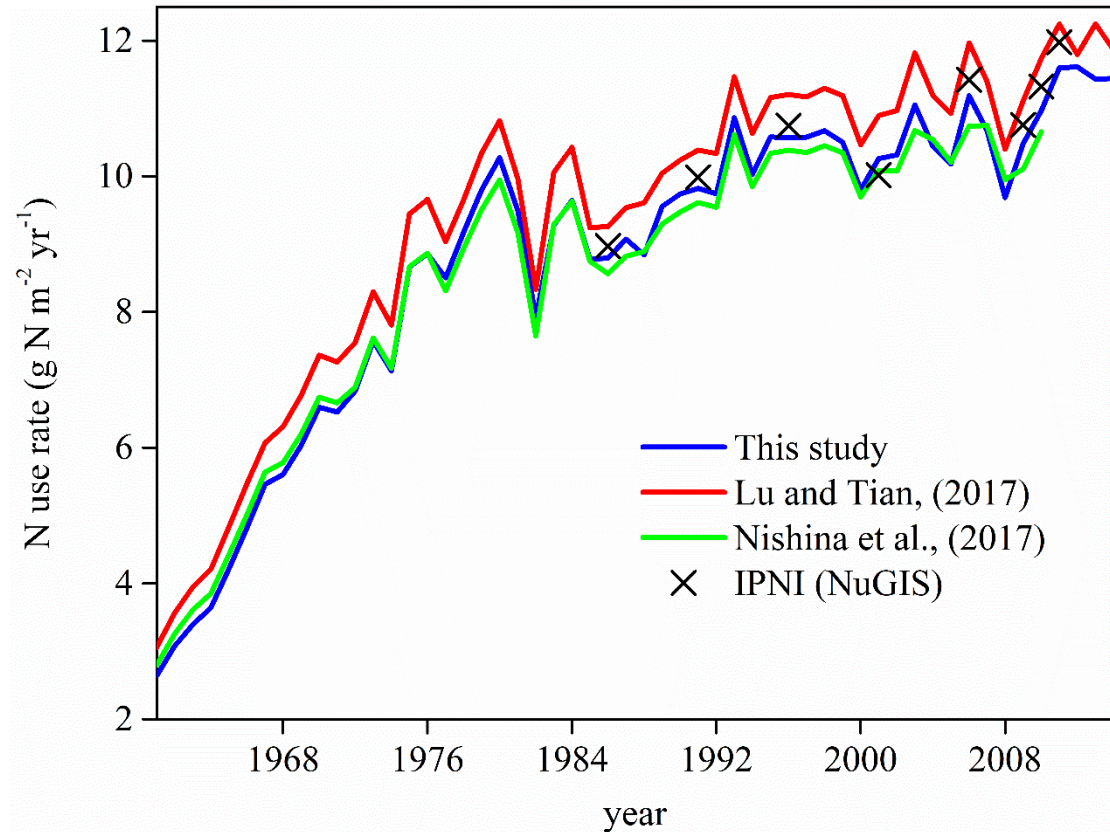

Figure 8: Time series of N fertilizer use in the United States: national commercial N fertilizer consumption of the U.S. during 1961-2014 derived from four data sets. Red line: this study, Green line: Lu and Tian, (2017), Blue line: Nishina et al. (2017), and ×: IPNI (2018). The N fertilizer consumption derived from Lu and Tian (2017) covers all states in the U.S., whereas other three data sets only cover the continental U.S.
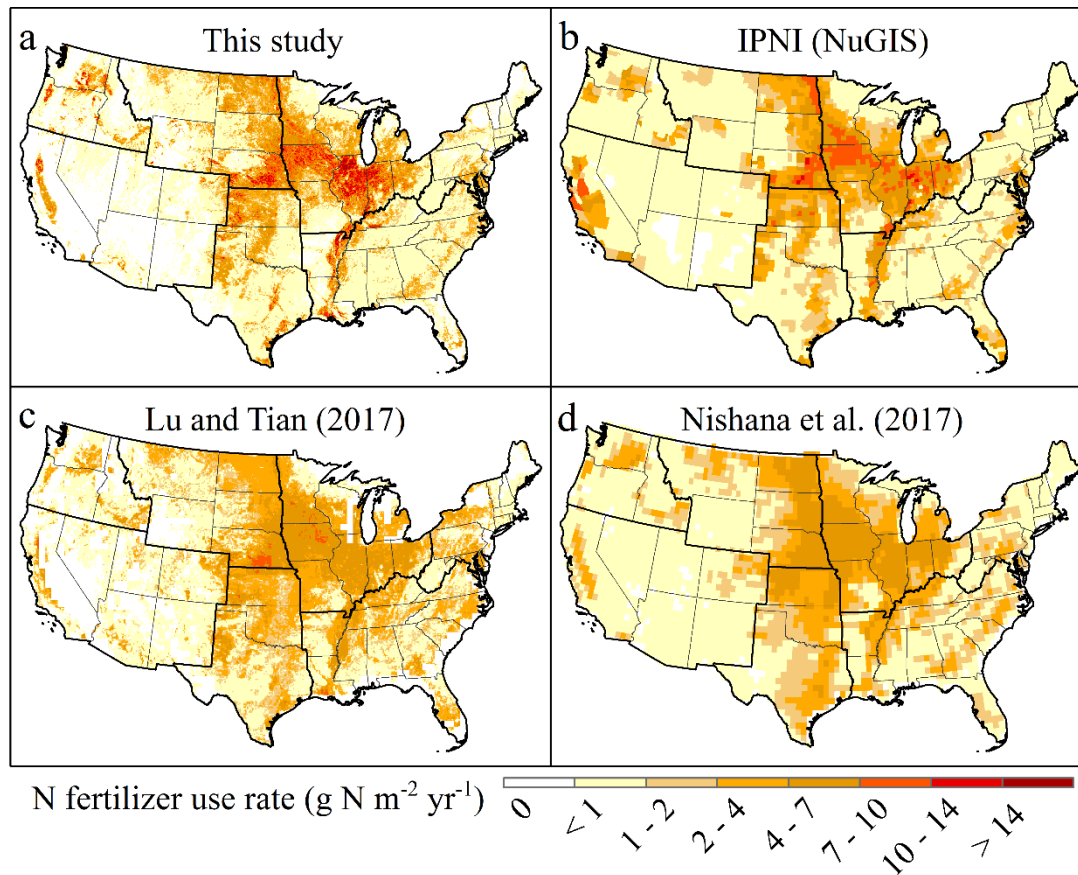
Figure 9: Comparison of N fertilizer use rate across the continental U.S. in 2010. (a) This study; (b) IPNI (2018); (c) Lu and Tian (2017); (d) Nishana et al. (2017).

References:

Hurtt, G., Chini, L. P., Frolking, S., Betts, R., Feddema, J., Fischer, G., Fisk, J., Hibbard, K., Houghton, R., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., Wang, Y. P.: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109, 117–161, 2011.

IPNI (International Plant Nutrition Institute): A Nutrient Use Information System (NuGIS) for the U.S. Norcross, GA. November 1, 2011, available at: www.ipni.net/nugis, last access: 1 Marth 2018, 2018.

Klein Goldewijk, K.: A historical land use data set for the Holocene; HYDE 3.2, DANS, doi:10.17026/dans-znk-cfy3, 2016.

Lu, C., and Tian, H.: Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance, Earth Syst. Sci. Data, 9, 181-192, https://doi.org/10.5194/essd-9-181-2017, 2017.

Nishina, K., Ito, A., Hanasaki, N., and Hayashi, S.: Reconstruction of spatially detailed global map of NH4$^+$ and NO3$^-$ application in synthetic nitrogen fertilizer, Earth Syst. Sci. Data, 9, 149-162, https://doi.org/10.5194/essd-9-149-2017, 2017.

Line 10: "essential implications" - awkward wording

Line 208: "consumptions" -> "consumption"

Line 349: "maximizing crop production" -> "maximizing profit" would be more accurate

Reply: We thank the reviewer for these words correction and have corrected them.