Field to Market – 02/12/2021

Data Processing for National Indicators Report and Benchmarks

Data source: USDA, NRCS, 2017 National Resources Inventory, September 2020

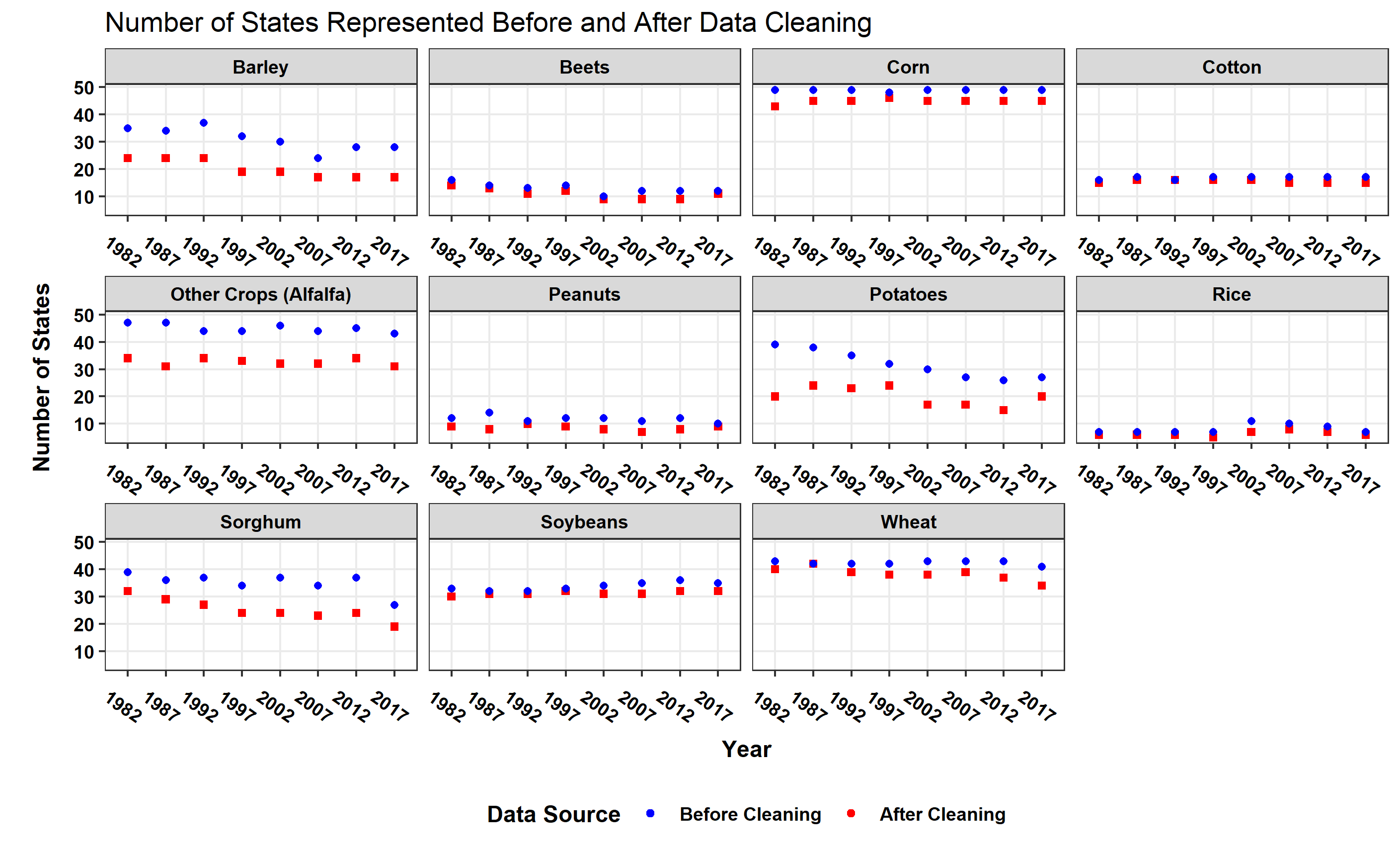
**Background**

Data sent by USDA to Field to Market consist of soil erosion rates and rates margin of errors for crops in Field to Market’s program. The data consist of 11 crops by year and state, starting on year 1982 and then every five years.

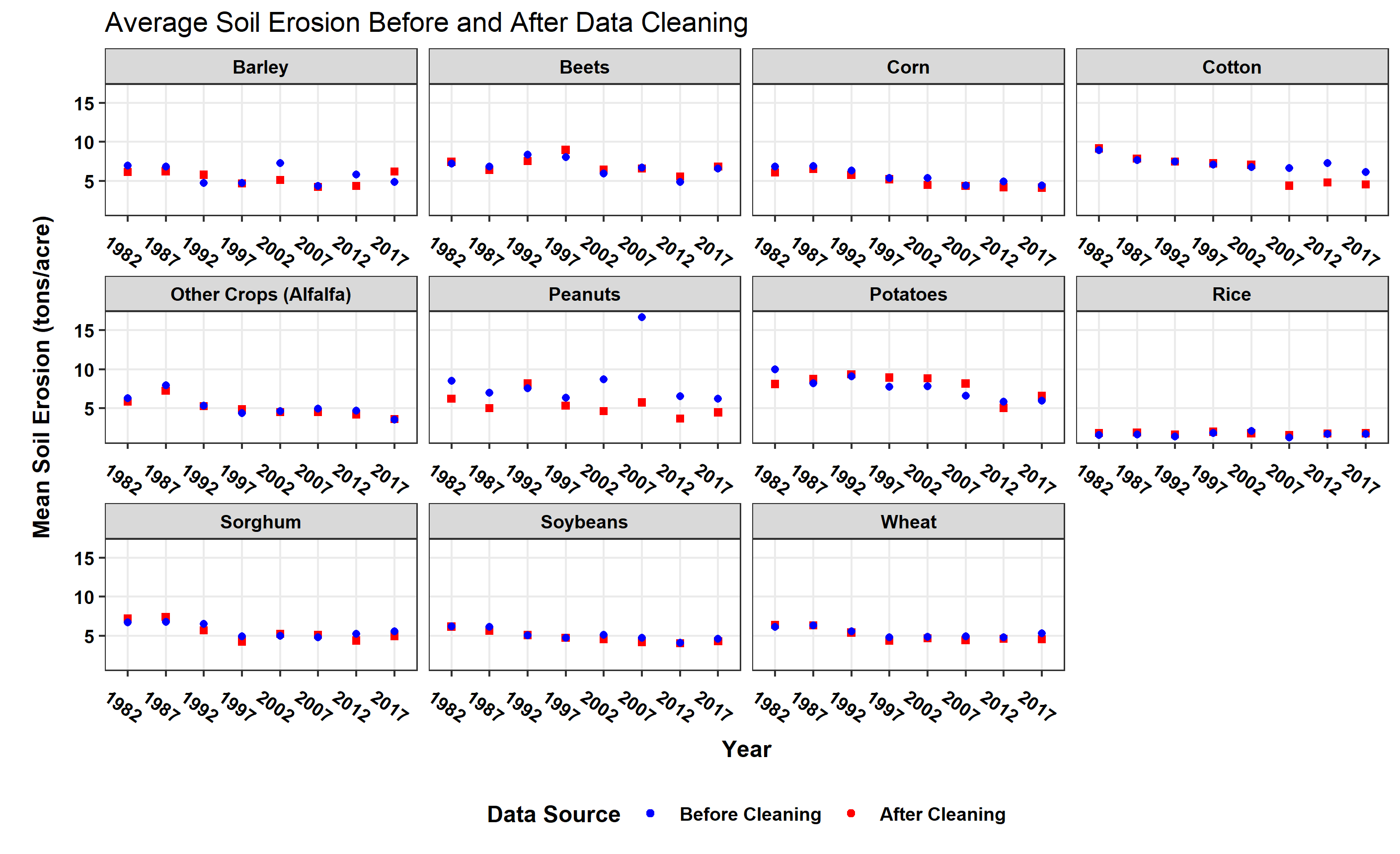
Many data points appear to be extreme outliers and/or with high margins of error. The following data cleaning steps were applied:

1. Removed data from the Caribbean and the National values.
2. Removed values where the lower bound of the erosion rate estimate were negative.
3. Removed extreme outliers after grouping by crop and year. Outliers were classified as those that exceeded the interquartile range by a factor of three.

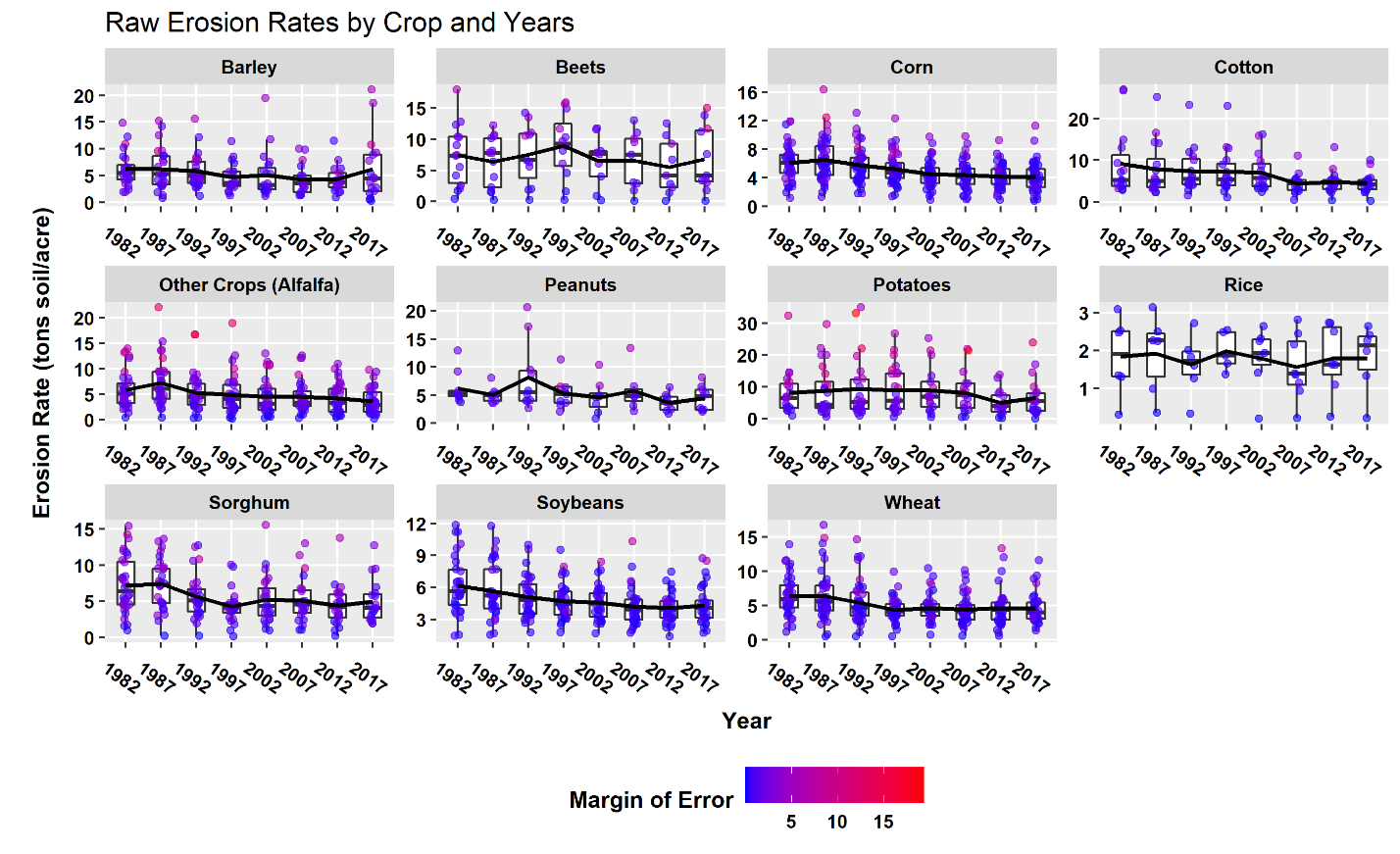
The above steps excluded approx. 24% of original data points. Summary figures are shown below.



**Figure 1. Number of States represented in the data set before and after applying cleaning process. Data are shown by crop and year. The crops with most data removed were Barley, Other Crops (Alfalfa), Potatoes, and Sorghum.**



**Figure 2. Mean soil erosion rate before and after applying cleaning process by crop and year. The largest erosion rate differences included the last three sampling periods for Cotton, and sampling periods for Barley, Peanuts, and Potatoes.**



**Figure 3. Raw erosion rates by crop and year using the data after applying the cleaning process. Graph includes boxplots, jittered points colored by margin of errors, and a line with mean erosion rates over time.**