Comparison of the Soil Properties in State Geographic (STATSGO) and National Resources Inventory (NRI) Databases

T. Ding1, M. Zhang2 and P. Hendley3
1Department of Agronomy and Range Science, 2Department of Land, Air and Water Resources
University of California, Davis, Davis, CA, 95616
3Zeneca Ag Products, 1200 S. 47th Street, Richmond, CA 94804

Abstract
Soil is an important factor for risk assessment and ecosystem modeling. In our study, we compared the 18 properties of soil surface layers estimated from NRI and STATSGO soil databases in six states, California, Florida, Iowa, New York, Texas, and Washington. Results from regression analyses show that there is no good agreement between estimations from both soil databases. The coefficients of determination (R²) vary from 0.69 to 0.86. R² for soil pH value and organic matter was 0.69 and 0.51 in average, respectively. Statistical tests on the coefficients of regression functions show that the STATSGO database tends to overestimate when soil properties are very high and underestimate when soil properties are in lower values. Due to the disagreements between these soil databases, we suggest to aggregate the point samples of NRI database with spatial coverage of overlays for STATSGO, MLRA (Major Land Resource Area), hydrologic unit and county boundary. This approach can generate aggregated data for more accurate estimation of soil properties for regional risk assessment.

Introduction
Landscape analysis has become one of the important approaches for regional risk assessment. For assessing pesticide and nutrient contamination in groundwater and surface water, soil data are essential. The U.S. Department of Agriculture’s (USDA) Natural Resource Conservation Service (NRCS) has been establishing three digital soil geographic databases: the Soil Survey Geographic (SSURGO) database, the State Soil Geographic (STATSGO) database, and the National Geographic (NATSGO) database (Reyoldt and Telfe, 1989). While the availability of SSURGO is limited, STATSGO, which is primarily designed for planning, management, and monitoring at the regional, state, and river-basin scale, provides complete countrywide coverage with the finest resolution. Because of its digital format and present data availability, scientists and managers have heavily employed STATSGO for risk assessment.

The National Resources Inventory (NRI), the most comprehensive natural resources inventory in the United States, includes soils, water, land cover and use, and other natural resource characteristics on non-Federal land in the United States. NRI is characterized by the extent of coverage in both time and spatial scale. The most recently available data in 1992 contain information of about 800,000 sample points throughout the non-Federal land of the United States (Kafogli et al, 1994, USDA 1994).

Because NRI database links land use/cover, soil properties, anthropogenic variables, and other natural resource characteristics on a point-to-point basis, it supplies rich, detailed, and precise background for risk assessment. However, the spatial location of NRI sample points is not available for public use. Geographical representation of NRI is usually completed by joining the point data with polygon coverage such as STATSGO, MLRA (Major Land Resource Area), hydrologic unit and county boundaries (Figure 2).

The objective of this study is to compare some soil properties from both NRI and STATSGO databases and recommend the use of the aggregated data from NRI and spatial elements from STATSGO for risk assessment in geographic information systems (GIS). By using this approach, we can take the advantage of geographic elements from the GIS map overlays (STATSGO, MLRA, hydrologic unit and counties) with a point-to-point basis rich with detailed NRI data.

Materials and Methods
Six states (California, Florida, Iowa, New York, Texas, and Washington) were selected for the soil property comparisons. Eighteen properties of soil surface layers (Table 1 and Table 2) were considered in the analysis. STATSGO ANALYST, an ARC/VIEW extension written by AVENUE language was used to extract information from the layer tables to the map unit table (Jones and Zhang 1999). For each map unit, the values reported for the surface layer of components were weighted by the percentage area of that component to derive a weighted average value for the entire map unit. The percentage of components that participated in calculation was also accounted.

In order to have a common ground for comparing with STATSGO (Figure 1), the point records of NRI were spatially aggregated into the STATSGO map unit (Figure 2). The arithmetical means of sampling points within STATSGO map units were calculated.

The components of map unit for each soil property where the valid values exist are summarized. For each soil property variable, NRI and STATSGO were compared on those map units. Simple regressions and GIS were used in the analysis. The coefficient of determination (R²) was used as a measure of the agreement of the regression relationship.
Results

• The number of NRI points are linearly related to STATSGO polygon areas where the states have the most agriculture land (Figure 3).

• Except for the Bulk Density in California, all the other soil properties show significant regression relationships between NRI and STATSGO databases in all selected states (Table 3, Iowa as an example).

• The relationship for pH values between NRI and STATSGO have the higher R-square (0.69 for average). The rest of soil properties have rather lower R-square whose averages range from 0.242 to 0.512.

• The number of NRI points are linearly related to STATSGO polygon areas in Illinois.

• STATSGO provides ready to use GIS data and spatial elements with which to link the NRI point data and other point datasets. It is commonly used within many areas of scientific research.

• However, we propose four important advantages of the combination approach of NRI and STATSGO, HUC, MLRA and county FIPS databases:
  
  (1) NRI database not only provides soil properties but also land use, water, anthropogenic variables, and other natural resource characteristics for each sample point.
  
  (2) Within non-federal land, NRI database supplies complete soil property information. STATSGO database frequently displays missing values for some soil properties.
  
  (3) Historical NRI databases are available for trend analysis.
  
  (4) With the overlay of STATSGO, hydrologic units, MLRA and county boundaries, the spatial geographic elements of NRI can be adequately resolved.

• However, NRI at present has few data points on federal land. This might be problematic for regional risk assessment of some regions that are largely federally owned; nevertheless, this should not be a problem for agricultural systems.

Discussion and Conclusion

This study shows the disagreement of NRI and STATSGO databases for the selected soil properties. This result implies the outputs of risk assessment or ecosystem modeling will be influenced by the selection of data source.

Based on the observations about the discrepancies between NRI and STATSGO, we suggest an alternative approach to estimate values of regional soil properties for risk assessment. That is, aggregating the point samples in NRI database and representing the results on the spatial boundaries defined by county, MLRA, hydrologic units, STATSGO map units, or combination of them.

STATSGO provides ready to use GIS data and spatial elements with which to link the NRI point data and other point datasets. It is commonly used within many areas of scientific research.

However, we propose four important advantages of the combination approach of NRI and STATSGO, HUC, MLRA and county FIPS databases:

(1) NRI database not only provides soil properties but also land use, water, anthropogenic variables, and other natural resource characteristics for each sample point.

(2) Within non-federal land, NRI database supplies complete soil property information. STATSGO database frequently displays missing values for some soil properties.

(3) Historical NRI databases are available for trend analysis.

(4) With the overlay of STATSGO, hydrologic units, MLRA and county boundaries, the spatial geographic elements of NRI can be adequately resolved.

However, NRI at present has few data points on federal land. This might be problematic for regional risk assessment of some regions that are largely federally owned; nevertheless, this should not be a problem for agricultural systems.

References


Acknowledgement

Many thanks to Yinyin Guo in helping with graphs and Mary Nett in reviewing the draft.