Harmonization of Legacy Soil Maps in North America: Status, Trends, and Implications for Digital Soil Mapping Efforts

J.A. Thompson, T.W. Nauman, & N.P. Odgers
West Virginia University

Z. Libohova & J.W. Hempel
US Department of Agriculture, Natural Resources Conservation Service

ABSTRACT: Use of soil property data in large regional to global models has highlighted some data artifacts in soil survey data in North America. There are often noticeable boundary differences between soil property value maps of adjacent survey areas, often done on a smaller governmental land-unit basis (e.g., US counties), and along international boundaries. Soil survey specialists have identified differences in map unit composition and inconsistencies in soil component property estimates as some potential causes to these discrepancies. These differences have been examined and even addressed in updates in very limited areas of North America, but there is still no standardized approach to the issue. As efforts to complete the initial national coverage of the Soil Survey Geographic (SSURGO) database are concluding around the US, the USDA-NRCS leadership has identified addressing SSURGO harmonization issues as a priority. Now, as this challenge is being tackled actively, questions arise regarding the goals of harmonization. In the US, SSURGO was not designed to be a soil property map; so as it is being modified to appear more seamless, the questions must be asked; How is the accuracy, precision, and uncertainty addressed in new maps? How can this data be used? Are there better digital soil mapping techniques out there for correcting these problems? Examples of these data discrepancies and possible corrective measures are presented to help provide context for discussion of harmonization.

1 INTRODUCTION

Applications of soil information and queries of soil databases are becoming increasingly larger in extent, covering multi-county, multi-state, and multi-national areas in order to address issues that require regional to continental to global response (Grunwald et al., 2011). The use of legacy soil survey products, which often were created to respond to more localized land use issues, reveals logical inconsistencies when multiple survey areas are used together for modeling or decision-making purposes. Within a country such as the United States, the discrepancies at political boundaries may be attributable to the piecemeal approach to collecting and publishing soil survey information. The spatial and tabular soils information included in Soil Survey Geographic (SSURGO) database (Soil Survey Staff, 2012) was originally compiled as part of approximately 3,000 independent soil surveys. These individual maps and associated tabular data are of different vintages and different scales, they were created using different mapping concepts, and they often use different soil components and different estimated property data to represent the same soil-landscape features. Consequently, there are frequently artificial boundaries in the spatial data associated with political borders. These discrepancies in the maps are caused by inconsistencies in map unit composition and/or in estimated soil property data. These inconsistencies in the map unit characteristics often lead to discontinuities in mapped soil properties and in soil use and management interpretations. These discrepancies are further accentuated at international borders where different entities have been responsible for producing the information, such that national priorities dictate further differences in purpose, scale, intensity, and classification systems.

During the 2011 National Cooperative Soil Survey Conference in Asheville, North Carolina, the Director of the USDA-NRCS Soil Survey Division announced that the National Cooperative Soil Survey (NCSS) was beginning a new, two-phase initiative to improve the quality of soil survey data in the US (Golden, 2011). Phase 1 of this initiative is to harmonize the NCSS soil attribute database. This is to involve targeted projects that examine data voids and gaps for selected soil properties, and address discontinuities at county, state, and national borders (Golden, 2011). While this portion of the initiative focuses primarily on the tabular data in SSURGO, Phase 2 is to focus on enhancement of the spatial data. The intention is to address errors in the spatial representation of map unit delineations (Golden, 2011). The objectives of this paper are to (i) highlight some of the issues related to the observed discrepancies in SSURGO and their possible causes, and (ii) identify some of the potential corrective measures in the context of digital soil mapping.

2 ISSUES

The nature of the disharmony in the SSURGO database is illustrated when viewing the spatial distribution of soil properties across multi-state regions that encompass multiple soil survey areas (Fig. 1). Most soil property data in SSURGO are estimates that were developed by aggregating information from field and laboratory analyses of multiple pedons. A derived soil property that is commonly calculated using the SSURGO database is soil organic carbon (SOC) stock. We followed the methods of Bliss et al. (1995) to compute SOC storage within the upper 100 cm of soil on an areal basis (kg/m^2) using the representative values for soil properties contained in SSURGO. Soil organic matter (OM) content values for each horizon were converted to SOC values by dividing by 1.724. The SOC stock of each horizon (to a depth of 100 cm) was calculated using SOC content, bulk density, thickness, and rock fragment content data of the horizon. The SOC content of each horizon was summed over the 100-cm depth to determine the SOC content of each soil in the survey area. The SOC content of each map unit was calculated as the weighted average of all the soils represented in each map unit.

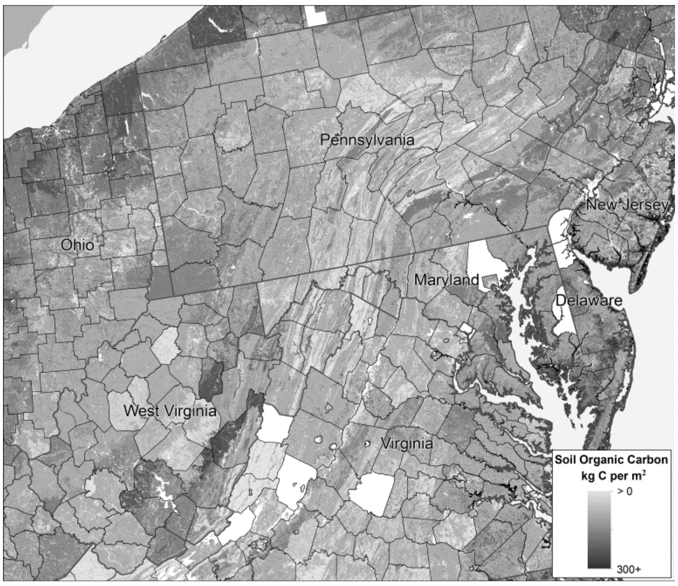

Figure 1. Soil organic carbon stocks across a multi-state region of the northeastern United States.

To understand the possible causes of this disharmony, we further examined the data on the currency, scale, and density of the mapping among individual soil survey areas in SSURGO. We also examined the distribution of map unit types within soil survey areas.

2.1 *Data Currency*

Quality assurance and quality control of soil survey products concludes with the final correlation of the survey data. The correlation process, when completed, ensures consistency and accuracy in the mapping, data, and interpretations within and among soil survey products, and requires that data entered into the database meets national standards. The earliest correlation date associated with a soil survey area in the SSURGO database is for Jackson County, Alabama, which is from 1943. This is one of 11 survey areas in the US that were correlated in the 1940s (Table 1). The median year of correlation is 1983, such that half of the SSURGO data is approximately 30 years old or older. Approximately half of the survey areas in SSURGO were correlated in the 1970s and 1980s. These survey areas represent a little less than 48% of the land area in the US.

Table 1. Summary of the currency of soil survey information in SSURGO.

Publication date	Number of survey areas	Areal extent (km^2)	(%)
1940s	11	17,813	0.3
1950s	68	113,429	1.6
1960s	428	929,586	13.1
1970s	704	1,588,418	22.4
1980s	725	1,808,290	25.5
1990s	461	1,270,875	17.9
2000s	480	1,060,921	15.0
Unspecified	70	304,238	4.3
Total	2947	7,093,570	100

2.2 *Map Scale*

The SSURGO database includes data from soil maps from individual survey areas that were created at scales from 1:12,000 to 1:125,000 (Table 2), with 13 distinct map scales represented (although only six of these map scales have more than 10 survey areas). The most common map scale is 1:24,000, encompassing almost 45% of the published soil survey area and 55% of the mapped land area of the US.

2.3 *Polygon Density*

Commensurate with these differences in the map scale, there are also differences in the intensity of the mapping. We quantified these differences by calculating the number of soil polygons per hectare (Table 3). Unlike map scale, there is a fairly even distribution of survey areas across the range of mapping densities. However, the areal extent of these surveys with low-density mapping is great, with over 35% of the mapped land area at densities of less than 0.003 features per hectare. (Table 3). The median density of polygons per hectare is approximately 0.008.

Table 2. Summary of scale of soil survey information in SSURGO.

Map Scale	Number of survey areas	Areal extent (km^2)	Areal extent (%)
1:12000	250	366,283	5.16
1:15000	1	652	0.01
1:15840	467	630,220	8.88
1:20000	831	1,756,289	24.76
1:24000	1315	3,905,147	55.05
1:25000	14	24,318	0.34
1:31680	50	234,839	3.31
1:40000	1	4,673	0.07
1:42240	1	968	0.01
1:48000	6	48,051	0.68
1:62500	1	4,677	0.07
1:63000	1	5,347	0.08
1:63360	8	105,293	1.48
1:125000	1	7,465	0.11
Total	2947	7,093,570	100

Table 3. Summary of density of soil survey information in SSURGO.

Polygons per hectare	Number of survey areas	Areal extent (km^2)	Areal extent (%)
Less than 0.0015	290	1,659,628	23.4
0.0015 - 0.0030	293	907,563	12.8
0.0030 - 0.0045	296	808,610	11.4
0.0045 - 0.0060	311	741,221	10.4
0.0060 - 0.0080	285	595,074	8.4
0.0080 - 0.010	315	625,301	8.8
0.010 - 0.015	278	485,218	6.8
0.015 - 0.020	301	476,247	6.7
0.020 - 0.025	297	416,299	5.9
0.025 - 0.080	281	378,411	5.3
Total	2947	7,093,570	100

The survey area with the highest intensity of mapping is Winnebago County, Iowa (correlation date of 1984, map scale of 1:15,840). The average size of map unit delineations in this county, which is located in the intensive agricultural area of the Midwestern US, is 2 ha (Fig. 2). Contrast this with the SSURGO soil map from the Mojave Desert Area, California (correlation date of 2007, map scale of 1:24,000). Despite being listed as mapped at the same scale, this survey area, which is in a remote desert environment, has the lowest intensity of mapping and an average size of map unit delineations of approximately 3,800 ha (38 km^2) (Fig. 3).

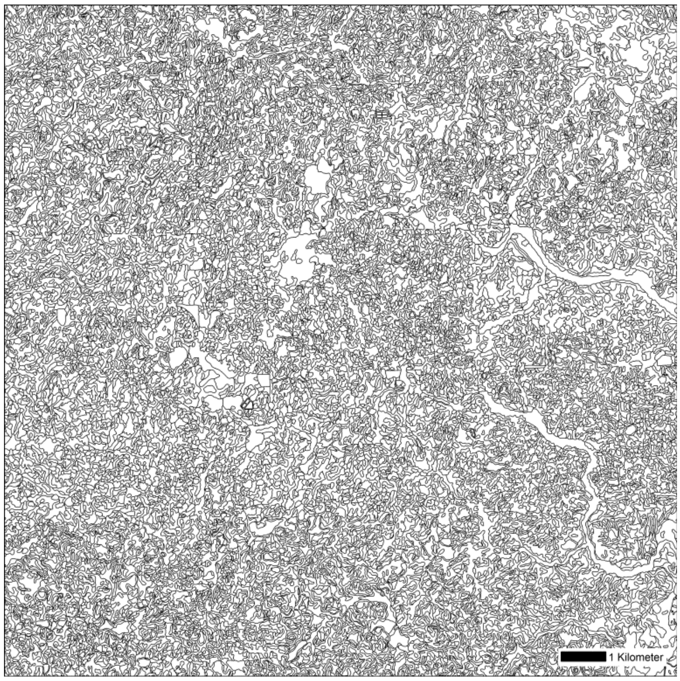

Figure 2. Density of map units in Winnebago County, Iowa, survey area. The extent of the data shown is approximately 15 by 15 km.

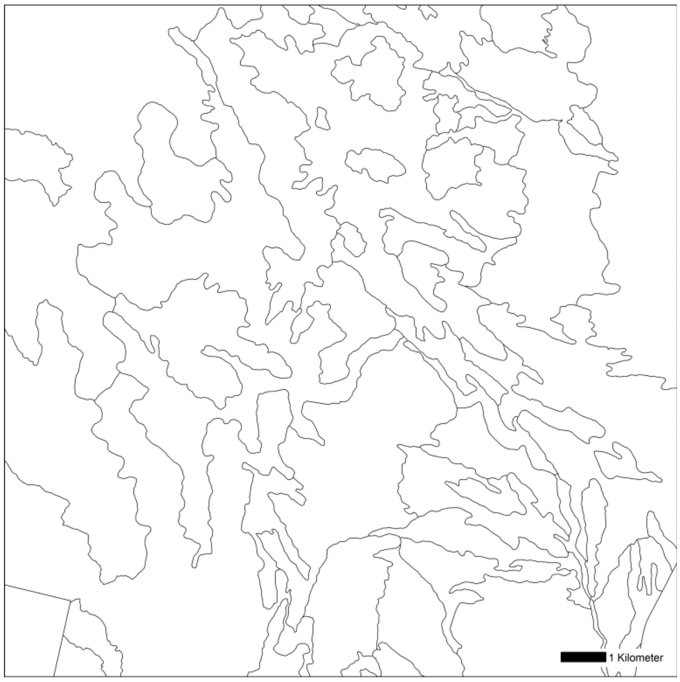

Figure 3. Density of map units in Mojave Desert Area, California, survey area. The extent of the data shown is approximately 15 by 15 km.

2.4 *Mapping Concepts*

Four different kinds of soil map units may be used in SSURGO, including simple map units, which represent one named soil type, and combined map units, which represent two or more named soil types. The simple map units (consociations) are normally used when mapping is more detailed (larger scale maps). The combined map units (complexes, associations, undifferentiated groups) are used when individual soil types cannot be separated on the map because of the intricate spatial arrangement of the soil types, the scale of the map, or overriding characteristics that render differences in soil properties of limited importance. When adjoining soil survey areas have differing amounts and different kinds of map units, this implies that the concept of map unit design and the map unit composition differed between the soil scientists that created the maps. In some areas of the US, the difference in the distribution of the different kinds of map units can be striking (Fig. 4).

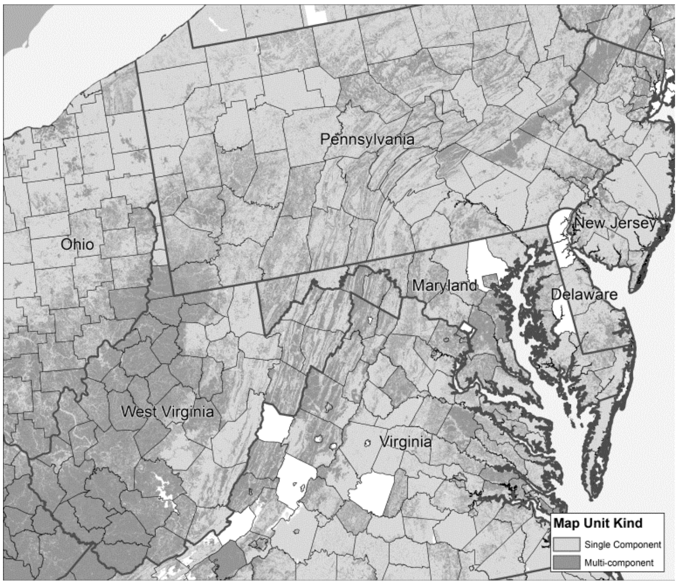

Figure 4. Map unit kinds across a multi-state region of the northeastern United States. Consociations are shown in light gray, while combined map units (associations, complexes, and undifferentiated groups) are represented by dark gray.

3 CAUSES

As described previously, the inconsistencies in the estimated SOC stock derived from SSURGO are evident, especially at county and state political boundaries (Fig. 1). The previous discussion summarizes how the original data that was used to create SSURGO is of different vintages and different scales, and created using different map unit design and composition concepts. In addition, the map units that straddle political boundaries may use different soil components and different estimated property data to represent the same soil-landscape features. Here we illustrate two cases where the inconsistencies in map unit characteristics have lead to discontinuities in mapped soil properties (in this case, SOC). Both examples are from the border between the states of New York and Pennsylvania in the northeastern US (shown in top center of Figs. 1 and 4). In each case, the spatial join between the two map units that meet at the border is perfectly aligned. However, these adjoining map units do not have the same estimated SOC stock in the upper 100 cm.

3.1 *Example 1: Same Series, Different Properties*

Chautauqua County, New York: Volusia channery silt loam, 3 to 8 percent slopes—SOC = 9.5 kg m^{-2}

Erie County, Pennsylvania: Volusia gravelly silt loam, 3 to 8 percent slopes—SOC = 7.4 kg m^{-2}

In this example, both map units are consociations of the same series. However, the amount of SOC in the map unit in New York is 2.1 kg m^{-2} higher than the map unit in Pennsylvania. The primary cause of these differing SOC stocks is the estimated property data associated with the Volusia soil (Table 4). The surface horizon for both soils has the same OM content (4.5%). However, in Chautauqua County, New York, which reports more SOC in the Volusia map unit, this surface horizon is 8 cm thicker. Additionally, the OM content of the second horizon of the Volusia soil in New York was 1% higher. There are negligible differences in the bulk density and rock fragment estimates between the two soils. As a result of these differences, the SOC for the Volusia soil in New York is greater than the SOC for the same soil in Pennsylvania.

Table 4. Summary of soil property data from the SSURGO data for the Volusia series of Chautauqua County, New York, and Erie County, Pennsylvania.

Chautauqua Co., NY		Erie County, PA	
Depth (cm)	OM (%)	Depth (cm)	OM (%)
0-23	4.5	0-15	4.5
23-38	1.5	15-43	0.5
38-107	0.5	43-152	0.5

3.2 *Example 2: Different Series & Map Unit Composition*

Chautauqua County, New York: Schuyler silt loam, 25 to 35 percent slopes—SOC = 25.0 kg m^{-2}

Erie County, Pennsylvania: Langford and Erie silt loams, 25 to 50 percent slopes—SOC = 8.9 kg m^{-2}

Table 5. Summary of soil property data from the SSURGO data for the Schuyler series from of Chautauqua County, New York, and the Langford and Erie series in Erie County, Pennsylvania.

Chautauqua Co., NY			Erie County, PA		
Series	Depth (cm)	OM (%)	Series	Depth (cm)	OM (%)
Schuyler	0-8	90	Langford	0-23	6
	8-30	5.5		23-76	0.25
	30-81	1.5		76-137	0.25
	81-183	0.5			
			Erie	0-15	5
				15-56	0.25
				56-109	0.25

In this example there are several discrepancies between the two adjoining map units. The map unit in Chautauqua County, New York, is a consociation, while the map unit in Erie County, Pennsylvania, is an undifferentiated group with two named series. However, the major difference between soils in these two map units, and the reason the SOC on the New York side of the border is almost three times greater, is that an organic surface horizon was described for the Schuyler soil, while no such organic horizon was described for the two soils in Erie County, Pennsylvania. Below the organic surface, the OM content (and the other soil properties) are similar among these soils.

These two examples illustrate the differences in calculated SOC stock due to differences in OM content, either because of differing estimated amounts of OM, differences in horizon thickness, or differences in the horizonation between soil series (e.g., presence of an organic surface horizon). Discrepancies in SOC stock across political borders could occur for other reasons, such as differences in rock fragment content, differences in bulk density, or the presence of minor components with distinctly different amounts of SOC.

4 HARMONIZATION

The soil science community in Europe, faced with continental-scale social and environmental issues that required soil information (e.g., Eckelmann et al, 2006), made efforts to harmonize soil survey information among the countries of Europe (Finke et al., 1998). This work may provide valuable guidance for harmonization efforts in the US. Harmonization can occur at several different levels within the structure of such complex databases. Some harmonization may occur at the level of the map legend, correlating soil types and aggregating them to recognize few distinct soil types. Harmonization may involve correlating soil map unit definitions to create consistent application of mapping concepts to distinct geologic or geomorphic landscape features. It may also be required to harmonize at the level of the soil type or soil horizon by consistently assigning physical, chemical, or morphological properties to soil materials, such that every occurrence in the database of a particular soil type will have the same suite of soil properties associated with it.

Potential also exists to harmonize SSURGO using digital soil mapping applications to help provide new maps with consistency across multiple soil survey areas. By using *scorpan* variables (McBratney et al., 2003) that are consistent across multiple survey areas, SSURGO data could be analyzed using scorpan variables to help disaggregate map-unit constructions into their component 'common denominators' (i.e. soil series). This would enable creation of seamless soil series maps to help solve issues with map unit design discrepancies. These procedures could be modeled in a raster framework, and released both as a raster map for researchers and as a re-aggregated SSURGO-equivalent vector format to maintain end-user interface continuity.

5 CONCLUSIONS

Soil surveys in the US have been conducted to meet user needs by creating an inventory of local soil resources. While the primary objective of soil survey efforts in the US has been to develop soil maps that delineate uniform management areas across the landscape, newer and more intensive uses of soil spatial databases have given further emphasis to the role of soil maps for providing users with information on soil properties instead of generalized soil use and management interpretations.

So while SSURGO was not designed to be a soil property map, efforts are now underway to modify the database to create a more seamless, consistent product. Moving forward, it is necessary to ask questions such as: How are accuracy, precision, and uncertainty addressed in new maps? How can this data be used? Are there available digital soil mapping techniques for addressing these concerns?

6 REFERENCES

Bliss, N.B., S.W. Waltman, and G.W. Petersen. 1995. Preparing a soil carbon inventory for the United States using geographic information systems. p. 275–295. In R. Lal et al. (ed.) Soils and global change. Adv. Soil Sci. CRC Press, Boca Raton, FL.

Eckelmann, W., Baritz, R., Bialousz, S., Bielek, P., Carre, F., Houšková, B., Jones, R.J.A., Kibblewhite, M.G., Kozak, J., Le Bas, C., Tóth, G., Tóth, T., Várallyay, G., Yli Halla, M. & Zupan, M. (2006). Common Criteria for Risk Area Identification according to Soil Threats. European Soil Bureau Research Report No.20, EUR 22185 EN, 94pp. Office for Official Publications of the European Communities, Luxembourg.

Finke, P.A., Hartwich, R., Dudal, R., Ibàñez, J.J., Jamagne, M., King, D., Montanarella, L. and Yassoglou, N. (1998): Georefernced Soil Database for Europe, Manual of Procedures, Version 1.0. - European Soil Bureau Research report No. 5, EUR 18092 EN, 170pp. Office for Official Publications of the European Communities, Luxembourg

Golden, M.L. 2011 Future Priorities in a Digital World for the New Soil Survey. Available online at http://soils.usda.gov/partnerships/ncss/conferences/2011_national/agenda.html. (Accessed February, 2012)

Grunwald S., J.A. Thompson, and J.L. Boettinger. 2011. Digital soil mapping and modeling at continental scales – finding solutions for global issues. Soil Sci. Soc. Am. J. (SSSA 75th Anniversary Special Paper) 75(4): 1201-1213.

McBratney AB, Santos MLM, and Minasny B 2003. On digital soil mapping. Geoderma 117(1-2), 3-52.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at http://soildatamart.nrcs.usda.gov. Accessed January, 2012.