The Historical Role of Base Maps in Soil Geography

Bradley A Miller, Leibniz-Centre for Agricultural Landscape Research (ZALF)
R. J. Schaetzl, Michigan State University

Available at: https://works.bepress.com/bradley_miller/5/
THE HISTORICAL ROLE OF BASE MAPS IN SOIL GEOGRAPHY

B.A. Miller*, R.J. Schaetzl

*Leibniz-Centre for Agricultural Landscape Research (ZALF) e.V., Institute of Soil Landscape Research, Eberswalder Str. 84, 15374 Müncheberg, Germany

Michigan State University, Dept. of Geography, 673 Auditorium Rd., East Lansing, Michigan 48912, USA

Email addresses: miller@zalf.de (B.A. Miller), soils@msu.edu (R.J. Schaetzl).

Abstract

Soil mapping is a major goal of soil science. Soil maps rely upon accurate base maps, both for positional reference and to provide environmental data that can assist in the prediction of soil properties. This paper reviews the historical development of base maps used for soil mapping, and evaluates the dependence of soil mapping on base maps. The availability of geographic technology for producing base maps has both constrained and directed the geographic study of soil. The lack of accurate methods for determining location limited early geographic descriptions of soils to narratives, or to listings of attributes for property-based map units. The first real base maps available for soil mapping were outline maps produced in the late 18th century, fueled by governments’ interests in documenting national boundaries and popular interest in world atlases. These early soil maps primarily used outline maps as a positional reference onto which soil-related thematic detail was added. Eventually, additional spatial information, in the form of topographic maps and later aerial photographs, increased the predictive role of base maps in soil mapping. In the current digital, geospatial revolution, global positioning systems and geographic information systems have nearly replaced the positional reference function of base maps. Today, base maps are more likely to be used as parameters in soil-landscape models for predicting the spatial distribution of soil properties.
and classes. Formerly, as a reference for spatial position, paper base maps controlled the cartographic scale of soil maps. However, this relationship is no longer true in geographic information systems. Today, as parameters for digital soil maps, base maps constitute the library of predictive variables and constrain the supported resolution of the soil map.

**Keywords:** soil mapping; geographic technology; history; environmental correlation; model parameters; history of mapping

### Highlights

1. Soil maps have always depended on base maps
2. The evolution of soil mapping is tied to the development of base maps
3. Paper base maps constrain and control cartographic scale and analysis scale of soil maps
4. Digital base maps only control resolution of the soil map
5. Standardized definitions for describing base map scale are proposed
1. Introduction

Soil geography and mapping are indispensable components of soil science, because soils are inherently spatial (Arnold, 1994; Campbell and Edmonds, 1984; Fridland, 1974; Goryachkin, 2005; Hole, 1953; Hole, 1978; Hudson, 1992). The processes for creating soil maps, with uses ranging from scientific study of soil pattern to applied use and management decisions, are firmly rooted in the concepts and methodologies of geography (Bushnell, 1929; Florinsky, 2012; Helms et al., 2002).

Therefore, it can be beneficial to review the historical development of soil mapping, as it relates to evolving concepts in geography.

The on-line Dictionary of Cartography (www.geography-dictionary.org, accessed 2014) defines a base map as a “map on which information may be placed for purposes of comparison or geographical correlation.” From the beginning of accurate surveying techniques (17th century), until the widespread completion of topographic maps, base maps were often limited to outline maps, which provide only positional reference. Outline maps tended to consist mainly of political borders, because these were of greatest interest to the governments who funded their compilation (Harley, 1988). Early soil maps used these base maps because the mapper needed a positional reference on which to plot observations, and these base maps were the only ones available. When other kinds of base maps - containing information useful beyond positional reference - became available for soil mapping, the additional information in them could be used to predict known soil patterns (Dokuchaev, 1883/1967; Florinsky, 2012; Hole and Campbell, 1985; Milne, 1935). Thus, historically, the base map has served as both a positional reference and a source of soil-landscape model parameters for the production of soil maps.

Although positional reference is a key function of all maps, recent developments in global positioning systems (GPS) and geographic information systems (GIS) have nearly replaced the need for separate maps to provide positional reference. Because of the heavy reliance on base maps for
In the past, these new geographic technologies may lead some to consider the term ‘base map’ as outdated. However, base maps still have use in “geographic correlation.”

In the modern definition of base maps, geography is recognizing the continuing role of base maps for spatial association. Indeed, GPS data do not actually replace base maps, because GPS only provides information on location. In contrast, base maps provide more information than location alone; they include spatially associated attributes and context (Abler, 1993). Base maps’ significance to soil mapping has, in fact, grown over the past century, as new technologies have provided new and more data-rich base maps with critical information about the soil environment. In digital soil mapping applications, base maps with these kinds of data have been commonly referred to by terms such as ‘parameters’, ‘covariates’, or ‘input variables’ (e.g. Behrens and Scholten, 2006; Grunwald, 2009; McBratney et al., 2003). These terms have been carried over from non-spatial modelling applications. However, maps used as input variables are a special kind of parameter used in soil-landscape modelling (Dobos et al., 2000; Levi and Rasmussen, 2014), with important spatial characteristics. Therefore, the continuing influence of base map properties on soil geography and soil mapping needs to be clarified, explained and hence, appreciated. Acknowledgment of base maps’ role in GIS-based methods of soil mapping provides a key link between traditional soil mapping endeavors of the past, and the new wave of digital soil mapping (McKenzie and Ryan, 1999). Thus, in this paper we examine the changing role of the base map in the creation of soil maps throughout history, guiding this discussion within the context of geographic technology’s evolution over time.

2. Early Soil Geography

2.1 Background

Ever since early peoples began sowing crops, and perhaps before, humans have had a vested interest in the geography of soil. Dating back to 3,000 - 2,000 BCE in central India, archeologists have
found evidence of farming on the fertile, black soils formed in the Deccan basalt, but not in areas to the north where these soils are absent (Shchetenko, 1968). Around the same time, the people of Mesopotamia were adjusting cropping patterns in response to observed differences in soil fertility (Krupenikov, 1992; Brevik and Hartemink, 2010). In eastern Sweden, where farming systems date back to 500 BCE, patterns of settlements, fields, meadows, and pastures are correlated to soil type (Widgren, 1979). Similarly, archeological sites in the United States (U.S.) have been shown to be preferentially located on well-drained soils (Almy, 1978). In each of these cases, a mental map of where certain soil characteristics existed was likely used in deciding where to live, settle, and plant. Although such information about soils was probably gathered by trial and error, eventually some rational patterning must have emerged so that numerous excavations were no longer required – if at all. Early peoples most likely used information observable from the surface, e.g., plants, slope, wetness, etc., as reference data with which to build their mental map of desired soil properties.

The Greeks of the Hellenistic Period were particularly astute at observation. Among these observations, they were the first to recognize the information that could be gained from examining a soil profile, which led to improved ideas related to pedogenesis (Krupenikov, 1993). Beginning with Parmenides and Eudoxus (ca. 540-470 BCE), they also recognized the connections among similar soils, vegetation, and climates on large scales - in belts on the Earth (Isachenko, 1971; Krupenikov, 1993). Theophrastus (ca. 371-287 BCE) observed that different soils react differently to the vagaries of weather, and that different crops were better suited for different soils (Theophrastus, 1916). Extending the soil geography of Theophrastus to predictive indicators, Marcus Cato (ca. 234-149 BCE) recalled a point made by Diophanes of Bithynia that, “you can judge whether land is fit for cultivation or not, either from the soil itself or from the vegetation growing on it” (p. 205, Cato, 1934). Even though the concepts of spatial and functional association were in place at this time, an actual soil map was still missing.
Many of the techniques for making the kind of paper map we know today began to appear after 600 BCE (Eratosthenes, 2010). However, the preferred method for describing soil geography remained the written narrative. There are two primary reasons for the lack of interest in producing a physical soil map at this point in history: (1) knowledge of soil was minimal and generalized, and (2) base maps of any kind were extremely crude. To be sure, the manual measurement of distance across land was relatively reliable. However, because a large amount of the geographic data was based on the narrative descriptions of explorers, assembling locations together on a two-dimensional format, i.e. the map, was not as reliable. For example, Herodotus (ca. 484-425 BCE) was highly critical of maps, because of their potential to be misleading. He believed that a map showing Persia far away from Lacedaemon (modern Greece) convinced the Spartans not to assist in the Ionian Revolt. In place of maps, he felt that his table of distances provided more accurate information (Eratosthenes, 2010). Of course, the potential for maps to be misleading still exists today (Monmonier, 1996), but the primary problem with maps of any kind during this early time was accurately representing the position of all features consistently.

Due to the difficulty of accurately measuring position on the Earth, map scale was very inconsistent on early maps. Even though early maps essentially did not have a cartographic scale, their relationship between extent and generalization remain consistent with the modern use of map scale terms. Specifically, for a given size of media (i.e. paper), the map could either include more detail about a small extent or it could be more generalized for covering a larger extent.

Because of the limitations in geographic technology, maps before 1500 CE were either theoretical maps about the Earth with a small cartographic scale, or maps created for civil purposes with a large cartographic scale. Small cartographic scale maps were popular with natural philosophers for the purpose of illustrating conceptualizations about the known world (Figure 1). Large cartographic scale maps were generally produced for individual cities and for more practical purposes. For example, the Romans surveyed grid systems of “centuries” as a base map for city planning and levying of taxes.
Advances by Ptolemy (ca. 100-178 CE) in map projections and coordinate systems (including the prototype for latitude and longitude), greatly improved the utility of maps. However, it would take technological developments, namely the printing press, in the Early Modern period (ca. 1500-1800 CE) to popularize the use of paper maps (Brown, 1979).

2.2. Early Maps with Soil Attributes

The earliest known spatial representations, i.e., maps, of soil properties were tied to the assessment of land valuation at the analysis scale of parcels or fields. In general, the valuation of land - from which the levying of taxes was based - was assessed by an index of soil productivity (Kain and Baigent, 1992). Therefore this soil attribute was associated with documents of land ownership. Cadastral recordings of soil quality were fairly common for levying taxes in the feudal system of western Europe. From the 6th to the 15th century, cadastrals in the Arabian Caliphatates used soil productivity as a factor (with crop type) to levy taxes (Krupenikov, 1993). From approximately 300 to 1951 CE, China compiled data on land quality by crop fields, as well as with other physiographic properties, into special geographic descriptions called difanchzhi (Lee, 1921; Zaichikov, 1955). Although not a soil map per se, the linking of soil data with the spatial information of property boundaries was an advancement toward the creation of a real soil map.

Although cadastral maps advanced the geographic collection of soil data, they lacked a fundamental theory of prediction. Also, geographic information was limited, because soil boundaries often were simply field boundaries. It is likely that land assessors had some type of mental model for interpreting the fields they viewed, but like the ancients, they lacked a systematic approach for testing those relationships. More importantly, the cadastral maps summarized soil information with an analysis scale based on field and property delineations, instead of as a natural body or continuous surface. As this example shows, the approach to soil science was, at least in part, directed by the available base maps.
3. The Emergence of Topographic Maps

3.1. Surveying

Before thematic maps (i.e. geologic and soil maps) would be considered viable endeavors, mappers needed more reliable base maps of the landscape. Among the key ingredients for producing a reliable base map was the accurate determination of position. Early maps, with small cartographic scales, were so speculative that a basic outline of land masses was generally sufficient. Nonetheless, during ancient times even the configuration of the continents was known only generally. The scientific basis of continent outlines steadily improved during the Early Modern period (ca. 16th to 19th centuries). In 1554, Gerardus Mercator improved Ptolemy’s estimation of the Mediterranean Sea’s length from 62° to 53° longitude, which was corrected to 41° longitude by Guillaume Delisle in 1700 (Brown, 1979). The technology to geographically plot observations and examine generalized spatial patterns became more available later in the 18th century. Although the same technology was needed to produce accurate maps with large cartographic scales, it would take much longer to map large areas of extent at the associated level of detail. For this reason, thematic maps used for scientific study still tended to use small cartographic scales, up to and through the 19th century.

Although the principles for determining position were mostly worked out by the ancients, it would take inventions of the 17th century to produce instruments accurate enough to determine position within a few meters (Bennett, 1987). The development of this technology was hastened by the Age of Exploration, which began in the 15th century (Martin and Martin, 2005). At that time, to calculate latitude one needed only the angular measurement of a celestial body at its maximum height in the sky and the assistance of mathematical tables based on previous observations. However, until devices measuring fractions of angular seconds were available, the measurement of latitude commonly had an error range of a couple of degrees (equivalent to approximately 100 km per degree of latitude, in the mid-latitudes). Important innovations for increasing both the accuracy and precision of measuring celestial angles were filar micrometers and the replacement of alidades with
telescopes fitted with cross-hairs (Bennett, 1987). After the invention of accurate devices for measuring angles, the determination of longitude on land only required the establishment of a celestial clock (Brown, 1979). By 1676, sufficient data had been collected on the eclipses of Jupiter’s satellites to be used as a consistent timepiece anywhere in the world.

During much of the history of modern cartography, the production of base maps using the improved methods of location determination required the patronage of national governments. In 1668, France became the first country in Europe to sponsor a systematic survey of its lands using accurate methods. Its motivation was to improve infrastructure for the purpose of increasing trade and tax revenue (Brown, 1979). After establishing a series of points in a line using astronomic observations, detailed topographic surveying could be done, extending out from the established (meridian) line. Because of the time it took to produce such an accurate survey, the project focused on surveying France’s border first. In 1739, 70 years after the meridian line was established, the first survey of France’s border was completed. The topographic survey of France then ensued, requiring monies from private investors and local governments through time of war. After four generations of work by the Cassini family on this project, the first general topographic map of France was completed in 1815 (Konvitz, 1987) (Figure 2).

Soon after the successful beginnings of France’s mapping endeavor, other countries began initiating their own topographic surveys. By 1730, topographic surveys were underway in Bohemia, Moravia, Silesia, Lombardy, Transylvania, Austria, Hungary, and Germany (Brown, 1979). Astronomic and triangulation survey in the U.S. began with Meriwether Lewis and William Clark’s expedition in 1803, and continued sporadically with subsequent military reconnaissance parties. However, a systematic topographic survey of the U.S. did not start until after 1885 (Wheeler, 1885).

The transition between simple outline maps and topographic maps has considerable relevance to the field of soil geography. Topographic maps quickly evolved from being outline maps -usually the survey of a political boundary- as a starting point, to containing additional information about the
Although topographic maps usually contain information about topography, they also regularly include information on cultural features, water bodies, and vegetation. As more information beyond the outline of political boundaries came to be included in these types of base maps, they provided the soil mapper with data that were useful beyond locational reference alone. This added information provided parameters useful to the soil mappers' mental model of spatial association. And yet, the thematic maps that were to come would prove to be a key ingredient to better soil maps.

3.2. The Rise of Thematic Maps

From their inception, thematic maps have relied on base maps for positional reference, because of the special equipment and time-consuming work of accurately determining location on the Earth’s surface. In other words, a thematic map could have been created from astronomically referenced positions, but it was far more efficient to utilize the previously made positional references in an existing base map. Base maps were purchased by publishing companies, who printed compilations of them as atlases. These atlases were widely popular, because they presented new views of the world and reported the discoveries of European explorers. By the 1600s, hand-coloring of these maps was a fashionable activity for the recreational study of geography (Brown, 1979). Base maps were used like popular coloring books, which could be either educational or used to add whatever theme one wished.

During the Age of Exploration, atlases grew from being collections of base maps from around the world to also including thematic maps of those places. In 1890, Titov summarized the work of several mid to late 17th century Russian geographers, who described the distributions of climate, vegetation, and soils (Krupenikov, 1993). Also in typical atlas style of the time, the French Constantin-François Volney published View of the Soils and Climatic Pictures of the United States of America in 1804. This early physiographic description and map of North America compared Canada with Siberia and the prairies with the Russian steppes (Volney, 1804).
The reliance of thematic maps on base maps focused the production of thematic maps at the cartographic scale of the available base maps. The primary sponsors of base map production were national governments interested in outlining the borders of their country (Harley, 1988). Therefore, the first available, accurate, base maps were at the cartographic scale best suited for fitting the area of a country on the largest piece of paper available. However, in order to add detail to the maps, their cartographic scales needed to be increased and their extents decreased. To expand the map extent again required drawing the maps on adjoining sections of paper (e.g. Figure 3). In the 18th through 20th centuries, these physical limitations set the cartographic scale of maps based on the desired extent.

3.3. Use of Topographic Maps for Soil Mapping

When agrogeologists met the newly available, accurate, topographic maps, a new synergy was found. For the first time, a resource existed that allowed them to record and visualize spatial information with a reasonable amount of efficiency and accuracy. Scientists could simply plot their observations on the base map. Further, they could connect their observations on the map based on experience alone or guided by additional information in the base map (e.g. elevation). For example, William Smith’s 1815 geology and soil map of England was built on printings of John Cary’s 1812 topographic map (Figure 3). The geological attributes were added by hand, using water colors (University of New Hampshire, 2013). The linking of soil and geology attributes was based on the mapper’s experience of soil characteristics corresponding to the spatial distribution of geology as mapped at that cartographic scale (Boud, 1975; Brevik and Hartemink, 2010). Stanislaw Staszic’s 1815 geology map of Eastern Europe used the same approach - coloring topographic maps that were reproduced from an earlier cartographer’s efforts (Grigelis et al., 2011).

The famous map produced by V.V. Dokuchaev in 1883, showing the distribution of Chernozems in European Russia, is another example of the reliance of early soils maps on base maps (Figure 4). It can be assumed that Dokuchaev did not determine the location of the cities, rivers, and land...
boundaries included on his map. Instead, he plotted observations of humus content on a pre-existing base map. This followed the style of geographic study of the natural environment popularized by Alexander von Humboldt (Hartshorne, 1958; Robinson and Wallis, 1967). Hence, the cartographic scale in Dokuchaev’s and other thematic maps of the natural world was controlled by the available base map. Consequently, their analysis scale was also dictated by the level of detail that could be added to those maps.

When the U.S. Soil Survey began in 1899, its primary goal was to create maps specific enough to provide guidance for crop selection (Whitney, 1900; Whitney, 1909). However, the base maps available prevented the use of larger cartographic scales and limited the amount of detail that could be shown (Simonson, 1952). Furthermore, by this time, the U.S. Geological Survey had produced topographic maps for only a small portion of the U.S., generally focusing on areas of interest for mineral mining, irrigation, and land reclamation (Brown, 1979). As a result, the typical kit for the first U.S. soil surveyors included a six foot soil auger with extensions, shovel, compass, protractor and scale, accompanied by “a copy of the usually inadequate county or other available base map” (Lapham, 1949, p. 12). Soil boundaries were determined by frequent borings. In situations where a topographic map wasn’t available, soil boundaries would be sketched on a blank plat book. The plat books were usually based on the General Land Office surveys, which commonly had many errors (Lapham, 1949). Thus, the endeavor of detailed soil mapping in the early 1900s was limited by the inadequacy of base maps.

4. Introduction of Aerial Photography

The ability to locate features in soil surveys accurately and with greater detail was markedly enhanced by the use of aerial photography in the 1930’s (Soil Survey Staff, 1951). Many advances during World War I had improved the utility of aerial photographs for mapmaking in general (Smith, 1985). Soon thereafter, the U.S. Geological Survey began using aerial photography for making topographic maps (Davey, 1935). Aerial photography for soil survey work in the U.S. had been
proposed as early as 1923 (Cobb, 1923), but the U.S. Bureau of Chemistry and Soils showed little interest in taking on the expense (Simonson, 1989). Bushnell (1929) compared soil mapping with and without aerial photography, and made four main points on the advantages of using aerial photography. Aerial photographs made soil mapping (1) more efficient, (2) of higher quality, (3) with improved consistency, and (4) with reduced cartographic demands for the determination of location.

The use of aerial photography in soil survey operations in the U.S. was gradually adopted as commercial aerial photography (Bushnell, 1932), or state level aerial photographs, became more widely available. Michigan was the first state to report use of a mirror stereoscope in soil mapping, allowing for soil boundary determinations over areas that were thick in vegetation or otherwise difficult to traverse (Millar, 1932). In 1951, standards for interpreting aerial photographs were added to the U.S. Soil Survey Manual (Soil Survey Staff, 1951). By the end of the 20th century it would be common practice for the U.S. Soil Survey to use aerial photographs taken with a 153 mm lens, resulting in a 23 cm square image with a scale between 1:20,000 and 1:80,000 (Simonson, 1989; Soil Survey Staff, 1993). An example of a soil map that used an aerial photograph as both a source of positional reference and spatial association is provided in Figure 5.

The paradigm shift caused by the implementation of aerial photography has yielded lessons about the use of base maps. For example, Brown (2006) criticized the U.S. Soil Survey for limiting the efficiency of soil mapping because it continued to use extensive ground observations and field sampling of every delineation. However, when Pomerening and Cline (1953) compared results from soil mapping done using aerial photo interpretation alone versus that done with field mapping, they found that field reconnaissance was still important for maintaining map accuracy. In their study of soil mapping procedures, the mean for soil series mapped correctly was only 48% using aerial photograph interpretation alone, but increased to 84% when photo interpretation was combined with field reconnaissance. A similar increase in mapping accuracy was obtained for other soil attributes such as drainage class, parent material, and land use capability. In summary, although
high-quality aerial photography improved the U.S. Soil Survey map products by facilitating greater detail and precision, it did not replace the need to understand the contextual information in aerial photographs through field observation.

By the late 1900’s, even though aerial photography had greatly increased the level of detail that could be included in a soil map, limits still existed. The minimum delineation size on paper maps is relative to the legibility of identifying symbols and colors, as well as levels of spatial accuracy (Hupy et al., 2004; Arnold, 2006). For example, errors of location and internal composition increase as delineation size decreases. Therefore, the decision of cartographic scale and associated minimum delineation size for a map depended on the purpose, degree of boundary precision possible, and the maximum amount of acceptable error. In the U.S. Soil Survey program, those limits are expressed by definitions of map orders (Schoeneberger et al., 2012).

5. Base Maps in the Age of Geographic Information Systems

The recent introduction of geographic information technologies, such as GPS and GIS, has instigated another revolution in soil geography. Similar to aerial photographs, and topographic maps before them, these new technologies have created opportunities in soil mapping by providing critical spatial information and new methods to analyze that data. In short, modern geospatial technologies are resetting the soil mapping paradigm yet again. New technologies beckon the reevaluation of current procedures and the development of new techniques for the advancement science. For this reason, the current geospatial revolution has caused soil science to spawn a new sub-discipline, digital soil mapping (Scull et al., 2003). Many aspects of reality have remained the same, however, justifying the continued use of proven concepts from the previous paradigm. New geospatial technologies are shifting the soil mapping paradigm in three primary ways: (1) by increasing the types of spatial information available as base maps, (2) by replacing base maps with GPS as the primary source of positional referencing, and (3) by decoupling the relationships between different aspects of map scale.
Remote sensing technologies, beginning in the 20\textsuperscript{th} century and continuing today, have greatly improved the quality and variety of information that can be used as base maps (Lillesand et al., 2008). The process of remote sensing is generally based on recording the soil’s or other related environmental properties’ interaction with the electromagnetic spectrum. This interaction includes the reflectance/absorption of radiation from the sun or man-made emitters. That information is then analyzed for single time frames and bands, or processed together, to produce unique perspectives of the environment (Pohl and van Genderen, 1998). New remote sensing technologies range from proximal to aerial to satellite based sensors, each with unique capabilities for measuring soil properties directly, or for providing indirect evidence of soil properties by spatial association (Doolittle, 1987; Doolittle and Brevik, 2014; Lobell et al., 2010; Mulder et al., 2011). In addition, remote sensing can interpret the time required for a signal to travel between the emitter, a landscape feature, and back to a sensor, and then use this information to provide data about landscape morphology (e.g. LiDAR (Wehr and Lohr, 1999)). The spectral and spatial resolution, plus accuracy, of these data continue to improve, which regularly adds to the library of base maps available to soil-landscape modelers.

Base maps, such as aerial photographs, have also provided important contextual information for the soil mapper. From an aerial photograph, interpretations can be made about geomorphology, landscape position, vegetation, soil wetness, erosion status, land use, among other parameters useful for predicting soil properties (Goosen, 1968; Soil Survey Staff, 1993). Spatial analysis methods in GIS can quantify context much more efficiently than traditional methods, but also offer new methods of analysis. For example, multiple remote sensing bands can be analyzed together to produce indicators such as the normalized difference vegetation index (Rouse et al., 1973; Kriegler et al., 1969).
Although the list of spatial analysis methods available today is large, one of the most commonly used categories of analysis to produce base maps for soil mapping is land-surface derivatives (Florinsky et al., 2002; Moore et al., 1993). Each of these scale-dependent products of digital terrain analysis (Albani et al., 2004; Hupy et al., 2004; Roecker and Thompson, 2010, Wood, 1996) provides different types of information about the soil environment (Wilson and Gallant, 2000). The fact that land-surface derivatives can be calculated at many analysis scales (Figure 6), and new methods of analysis are being regularly proposed, serves only to multiply the number of base maps of this type that are available to soil-landscape modelers.

5.2. GPS becomes primary source of spatial referencing

GPS provides an efficient means for accurately determining location by triangulating for the distance between a GPS receiver and the known locations of satellites in orbit around the Earth. The distance between the receiver and the satellites is calculated using the time it takes a radio signal to travel from the respective satellites to the receiver (Hofmann-Wellenhof et al., 1993). By this method, the location of the surface-based receiver can be determined rapidly, accurately, and at relatively low cost. Thus, the preferred method for spatial referencing field observations is now with a GPS receiver (Schaetzl et al., 2002). Similarly, GPS enables the spatial referencing of remote sensing data as they are collected, with only minor post-processing adjustments to improve accuracy (Lillesand et al., 2008; Mostafa and Schwarz, 2000).

Despite the efficiency of GPS, the need for base maps for positional reference has not been completely replaced. Data collected without GPS available need to be georeferenced before they can be used in a GIS. The majority of data requiring georeferencing are legacy data, created before the availability of GPS and GIS technologies (e.g. Pásztor et al., 2010). Such legacy maps can only be incorporated into a GIS by georeferencing points viewable in both the legacy map and a base map. Another example is the importing of points referenced by distance from landmarks, such as political
boundaries (Veenstra and Burras, 2012). Although the role of base maps has been clearly altered by
GPS, base maps are still useful resources for positional reference when a GPS is unavailable or fails.

5.3. Challenging definitions of scale

A common problem for disciplines studying spatial phenomena is the wide range of interpretations
and definitions of scale (Withers and Meentemeyer, 1999; Dungan et al., 2002). Now that
geographic technologies have evolved to the point where scale can be examined in a variety of ways,
it is important to distinguish and standardize definitions for the different types of scale. This notion
is especially relevant to the consideration of how base maps are used in soil mapping. The
decoupling of the relationship between cartographic scale, analysis scale, extent, and resolution has
been a subtle change that has challenged our definitions and understanding of scale (Goodchild and
Proctor, 1997).

One of the most fundamental differences between traditional soil mapping and digital soil mapping
is that the cartographic scale—often referred to as map scale—of base maps no longer constrains the
cartographic scale of the resulting soil map. For practical reasons, such as legibility, the amount of
information that can be placed on a paper map is limited by a finite space, i.e. the piece of paper.
Put another way, the level of detail on a paper map is limited by cartographic scale. Because base
maps provide both positional reference and parameters for spatial association, a paper base map
controls the cartographic scale and the resolution simultaneously. Thus the resolution, or the spatial
density of information, has a direct relationship with the cartographic scale. Resolution can also be
equated to the minimum delineation size described in protocols for traditional soil mapping (e.g.,
Schoeneberger et al., 2012), although other factors are also involved in setting the lower limit of
minimum delineation size. For example, in figure 5, the aerial photograph clearly shows darker
(muck) areas that the mapper knew existed, but could not delineate them due to time and legibility
issues. In contrast, the user-interface of a GIS allows the display size of a map (i.e. cartographic scale)
to be freely changed by the viewer, without affecting the density of information (i.e. resolution)
stored in the spatial database. Because of this characteristic, GIS has eliminated the issue of legibility as a constraining or limiting factor in the determination of the minimum delineation size. However, it has not eliminated issues of spatial accuracy and resolution (Frank, 2009). Therefore, base maps no longer control cartographic scale. However, the resolution of the base map is still the lower limit of resolution possible for soil maps, without interpolation or downsampling. In order to properly convert lessons from the past and utilize opportunities of the new digital medium, it is essential to recognize the changing relationships between these classic geographic properties.

Further adding to confusion in the new world of GIS, there are multiple meanings for the term ‘scale’ used in popular and scientific literature. Geography distinguishes between cartographic scale, and the concept of analysis scale as a tool for identifying phenomenon scale (Montello, 2001). Analysis scale differs from resolution by being able to summarize an area containing multiple data points (Miller, 2014). A common example of resolution is the size of raster cells, which provides a single value for what has been observed or estimated about the space represented by each cell. In this way, the resolution describes the level of spatial detail available. The analysis scale can be left as equivalent to the resolution, or it can consider a greater neighborhood of data. Traditionally, changing the analysis scale used different zones, such as soil delineations or political delineations, to summarize neighborhoods of data (Gehlke and Biehl, 1934). However, with modern raster analysis, the context of a cell can be quantified by summarizing or otherwise analyzing an area larger than the cell without altering the spatial resolution. The change in results due to the change in analysis scale has been termed the modifiable areal unit problem, where this effect is divided into the zoning problem and the scale problem (Jelinski and Wu, 1996; Openshaw, 1983). When the spatial patterns of an analysis scale coincide with the spatial patterns of a target variable, the analysis scale is described as detecting the scale of the phenomenon.

To illustrate this difference between resolution and analysis scale, consider a cell in a digital elevation model (DEM). The DEM resolution is the cell size. For the original grid of elevation values,
the analysis scale is equivalent to resolution. If the mean elevation within soil delineations is calculated from that elevation grid, the analysis scale is changed to a level equivalent to the soil delineations, even though the information is based on data collected at the original DEM’s resolution. The DEM also can be analyzed for slope, but must use a neighborhood of cells in order to determine a change in elevation over a distance. The size of that neighborhood is the analysis scale. The cell is the center point and the neighborhood is shifted respectively for each of the other cells. Thus, the resolution is unchanged, despite a larger analysis scale being used to calculate new values for each cell location.

Although the definition of scale has long been a point of confusion, using the definitions of the three types of scale set by geography offer a basis for clarifying ideas, definitions, and meanings, as we transition to a new soil mapping paradigm. Cartographic scale, which was once paramount for describing the characteristics of a map, is now only a description of how the map is being displayed. Analysis scale is not necessarily viewable in the map, but is a key part in the process of creating the map. And finally, phenomenon scale is the scale at which spatial patterns occur in reality, which can be discovered or best represented by appropriately matching the scale at which the data are analyzed.

Other aspects of the digital map are best described by terms other than scale. Resolution is the spatial detail of the data, without dependencies on any type of scale. Further, extent describes the real-world area encompassed by the map, which in a GIS is also independent of any type of scale or the spatial data’s resolution.

We recommend the use of these definitions to disambiguate the various uses of the term ‘scale’ in the context of describing base maps and the process of using them to create soil maps.

6. Conclusions
Soil maps have always relied on base maps for information needed to predict the spatial distribution of soil properties. Over time, the spatial information garnered from base maps has shifted in emphasis between positional reference in early maps, to attributes that are more reliably associated with soil properties. The only possible exceptions to the reliance on base maps are spatial interpolation techniques that solely rely on spatial autocorrelation. However, recent spatial interpolation methods, such as co-kriging, reiterate the utility of base maps for making accurate soil maps. Even before paper maps, it seems likely that a mental map of above ground observations informed the creation of a mental soil map.

Base maps determine many of the characteristics of the soil map, because base maps are the primary ingredients in the production of the soil map. When soil maps were produced on paper, the cartographic scale of the soil map had to match the cartographic scale of the base map, because the base maps were the most practical source for positional referencing. When base maps became available with information about spatially associated variables, such as geology, vegetation, or elevation, soil mappers were able to use that information to improve the quality of the soil map. However, unless the soil mappers made additional observations, the resolution of the soil map could not reliably exceed the resolution of information available in the base maps. Also because of the practical limits in detail that can be included on a paper map, as determined by cartographic scale, base maps also strongly influenced the analysis scale of soil maps. In contrast, digital maps are free from many of the constraints of paper maps. However, as the source of spatial data, base maps are still the limiting factor in the density of spatial information available.

As the sophistication and format of base maps have evolved over time, so have soil maps and the geographical concepts used to understand them. Modern definitions of scale in geography help clarify concepts that were once inextricably linked. Multiple map characteristics that at one time could be summarily described with the scale of the map now need separate definitions to accommodate the new format and tools of GIS. Cartographic scale, analysis scale, map extent, and
map resolution, must now be treated as independent concepts. In contrast to paper maps, the only link between these concepts in a GIS is the limit of computing power for handling large file sizes. In this case, map resolution and map extent could be related, but the relationship is weak and somewhat arbitrary. Therefore, the description of base maps used in the scientific study of soil mapping should always include map extent, map resolution, and analysis scale when it differs from the resolution.

Acknowledgements

We appreciate the helpful comments on this manuscript from the anonymous reviewers.
References


Photogrammetric Engineering. 1, 3.


Dobos, E., Micheli, E., Baumgardner, M.F., Biehl, L., Helt, T., 2000. Use of combined digital elevation
model and satellite radiometric data for regional soil mapping. Geoderma. 97, 367-391. doi:
10.1016/S0016-7061(00)00046-X.

Dokuchaev, V.V., 1967. The Russian Chernozem (Kaner, N., Trans.). Israel Program for Scientific
Translations, Jerusalem. (Original work published in 1883)


Dungan, J.L., Perry, J.N., Dale, M.R.T., Legendre, P., Citron-Pousty, S., Fortin, M.J., Jakomulska, A.,
Ecography. 25, 626-640.

EHESS, 2014. Des villages de Cassini aux communes d’aujourd’hui. Available online:

Jersey.

Florinsky, I.V., 2012. The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the


Library Foundation, Buffalo and Erie County Public Library, 2013. A Delineation of the Strata of England and Wales with part of Scotland: Exhibiting the Collieries and Mines, the Marshes and Fen Lands Originally Overflowed by the Sea, and the Varieties of Soil According to Variations in the Substrata, Illustrated by the Most Descriptive Names. Available online at: 


Figure 1. A map that predates the use of cartographic scale – Eratosthenes’s map of the known world, ca. 194 BCE (Bunbury, 1883).

Figure 2. A portion of the first systematically surveyed, topographic map of France published between 1756 and 1815. Although this map was a milestone for accurate positioning of features, note the lack of elevation information. The cartographic scale was 1:86,400. (EHESS, 2014)
Figure 3. The first geological/soil map of Britain, produced by water coloring a pre-existing topographic map. Published by William Smith in 1815. In order to increase the cartographic scale, the base map needed to be printed on multiple pieces of adjoining paper. The cartographic scale was approximately 1:316,000. (Library Foundation, Buffalo and Erie County Public Library, 2013)
Figure 4. Isohumus belts identified by Dokuchaev based on quantitative point observations, plotted on a topographic base map. The cartographic scale was 1:4.2 million. (Dokuchaev, 1883/1967)
Figure 5. A portion of the Lincoln County, Wisconsin soil survey, conducted by the U.S. Natural Resources Conservation Service, which typically uses aerial photograph base maps (Mitchell, 1996). Even though this is a forested, and not a bare-soil, image, the tone, texture, and value of the photograph provide clear clues about the location of soil map unit boundaries. Line placement on the photograph shows how the aerial photograph provides positional reference. The assignment of the same map unit to areas of similar tone is an example of the use of the aerial photograph for spatial association. The original cartographic scale was 1:20,000.

Figure 6. The same portion of Lincoln County, Wisconsin as shown in figure 5, analyzed for slope gradient. The analysis was conducted on a 3 m resolution DEM (USDA-NRCS, 2014). The calculation of slope gradient was conducted at two different analysis scales: (a) 9 m and (b) 90 m. Although there is a gradient between analysis scales, the different analysis scales provide different base maps that could be used for soil mapping.