Digital soil resource inventories: status and prospects

D.G. Rossiter

Abstract. This article presents an inventory of digital soils data and supporting information available or publicized on the World Wide Web. The amount and quality of digital soil survey information at global, national, regional, and local scales is increasing dramatically. Some is freely available online, some is available on CD-ROM and only publicized, not distributed, on the Web. The world is completely mapped at small scale. Many regions are covered by medium-scale inventories, especially using the SOTER methodology. Large-scale digital data are limited to the USA, Canada, Australia and Europe, but there is discrepancy with respect to their philosophies of public access to foundation data such as soil maps. Remaining problems include the lack of metadata, limited interpretations for professionals who are not soil specialists, and geodesic incompatibility with other digital data. The frequent reorganization of websites leads, at best, to user frustration and, at worst, to the disappearance of information.

Keywords: Digital information, data availability, soil inventories, SOTER methodology, World Wide Web

INTRODUCTION

Information about soil properties and behaviour over tracts of land is vital for making decisions on proper land use and management, environmental protection, and land use planning. This has been the motivation for systematic soil surveys, soil survey interpretations, and maps of soil properties required by empirical or process models.

Soil resource inventories are usually presented as

area-class maps which divide the survey area into polygons, with an accompanying report that describes the soils in groups of polygons known as map units, each of which contains one or more named soil types. Rarely, continuous-field maps of soil properties or types are presented. Another kind of inventory is a database of point observations, usually with analytical data from the laboratory.

Traditional distribution channels for soil surveys included county agents, agricultural extensionists, educational institutions and public libraries. The data-producing organization itself was often a principal user of the data, mainly for farm planning, but also offered interpreting on request for other uses. However, it was soon appreciated that interpreted information would reach a wider user base, and this became part of some organizations’ activities (Klingebiel 1991). In recent years there has been increasing demand for soil geographical information from governmental agencies, national governments officers, consultants, and researchers dealing with a wide range of new applications (e.g. environmental protection, town and country planning, health, resource extraction) as well as traditional applications in agricultural extension and soil conservation. Some users want map unit or point data as collected or interpreted by the survey organization, while others require that a specialist interpret it in a form that is directly useful for their work, so the primary data must be made available to outsiders.

The advent of the World Wide Web (WWW) and inexpensive, user-friendly geographical information systems (GISs) have the potential to revolutionize the distribution and use of soil geographical data. In particular, the ease of combining a digital soils theme with spatially related information such as land use, land ownership, planning units, infrastructure, and surveys of other resources can greatly increase the utility of soil surveys. If these are made available in digital form, with no or liberal-use restrictions, anyone can use the data, perhaps in novel ways. This article reviews the status of digital soil geographical data available online, or via on-line order, from the viewpoint of an informed data user searching for available information.

The ‘Compendium of On-Line Soil Survey Information’

Since 1997, the author has maintained the website ‘A Compendium of On-Line Soil Survey Information’ (Rossiter 1997), to reference all information on soil survey activities, institutions, data sets, research and teaching that is available via the WWW, with downloading if the referred site permits. The items are links to on-line resources, annotated with opinions of their reliability and usefulness, and are divided into eight sections: (1) Soil survey institutes and activities; (2) Learning about soil survey; (3) Soil classification for soil survey; (4) Applications and Models (including soil survey interpretations); (5) Methods and Techniques; (6) Land

Department of Earth Systems Analysis, International Institute for Geo-Information Science and Earth Observation (ITC), PO Box 99, 7500AA Enschede, Netherlands
Fax: +31 (0)53 487 4336. E-mail: rossiter@itc.nl.
Evaluation; (7) Digital Soil Geographic Databases; and (8) Computer programmes for soil survey. The Compendium does not cover general soil science or subdisciplines, nor cartography, remote sensing, landscape analysis or GISs except where these topics are directly relevant to soil survey. The major effort has been to provide centralized access to digital data sets and information that makes them interpretable, such as soil classifications.

The author spends several weeks a year actively searching for new information and updating links as websites are reorganized and soil survey organizations revise distribution policies: a Sisyphean task. Information is often removed altogether when a project finishes, a government agency is closed or merged with another, or a scientist moves to another institution.

On typical workdays in early 2004, about thirty different Internet users visited the Compendium (Nedstat bv. 2004). Their distribution by Internet domain corresponds to the availability of Internet access: about 50% from the USA (.edu, .com, .gov, .mil), 24% from Europe, 6% from Canada and Australia, 6% from Asia, 4% from Mexico, South and Central America, 1% from Africa, and 9% from international organizations and networks (.org, .net). Over 160 websites include a link to the Compendium: mostly libraries, portals, soil science faculties, and soil survey organizations. Usage is increasing, probably because of a steady increase of these inward links. In mid-February 2004, it was the second site listed in a Google search for the term ‘soil survey’ (Google 2004), following the home page of the US Natural Resources Conservation Service (NRCS).

**SOIL GEOGRAPHICAL DATA AVAILABLE VIA THE WEB**

The following inventory was current as of mid-February 2004. Data are accessible via the Compendium web page.

**World.** Climate modellers have abstracted the printed 1:5 M Soil Map of the World (FAO-UNESCO 1974) to provide a 1° x 1° grid of relevant soil properties, including generalized texture, depth and slope (Zobler 1986; Webb et al. 1993). This resolution is about twice as coarse (linear) at the equator as the minimum legible delineation implied by this scale, 625 km². The Zobler set has been resampled to 30° x 30’ by Oak Ridge National Laboratory, USA, but the resolution of the map was not actually increased: rather the 1° squares were divided into four 30’ squares with the necessary modification of ocean boundaries. The associated soil type information is retained from the Webb et al. data. Other derived properties, such as water-holding capacity and organic carbon, are included in related data sets available from Distributed Active Archive Center (DAAC) at Oak Ridge. FAO has revised the Soil Map of the World: the polygon map and a 5’ x 5’ rasterization of soil properties is available on CD-ROM ($350) but not on-line (FAO 1998a). ISRIC and FAO have used this as the basis of a 30’ x 30’ version which considers the full map unit composition of the 36 component 5’ x 5’ grid cells (WISE) (Batjes & Bridges 1994). Applications of this database include agroecological zoning and studies of global change (Batjes 2002).

**Regional.** The outstanding examples of regional soil data bases are those made with the SOTER methodology, jointly developed by UNEP, IUSS, ISRIC and FAO (Engelen & Wen 1995). These are currently available for Latin America and the Caribbean (SOTERLAC) (FAO 1998b) and Central and Eastern Europe (SOVEUR) (FAO 1998c); both are available on-line from ISRIC and as CD-ROMs in the FAO’s Land and Water Digital Media Series. FAO has also produced a data set for northeast Africa (FAO 1998d) and a soil and physiographical database for North and Central Eurasia (FAO 1999), available only on CD-ROM. The Mekong River Commission (2002) has produced a medium-scale CD-ROM of the soils of the lower Mekong Basin.

The 1:1 M European soils data set is not distributed outside of the European Commission itself and its approved research projects, with the explanation that the copyright of the original provider has not been extended to allow redistribution (and that the Joint Research Centre does not have sufficient data management facilities), even though the data were re-interpreted for this map which could be considered a new product.

**National.** The USA (Natural Resources Conservation Service, NRCS) and Canada (Agriculture and Agri-Food) have complete national coverages at small scale (1:5 M) and sectional coverages at medium scale (in the USA, STATSGO at 1:250 K), as well as many detailed surveys (in the USA, SSURGO at 1:24 K to 1:12 K) available for downloading, along with supporting metadata, interpretive tables, and descriptions of soil types. On-line national coverages at small scale are available for Mexico, Brazil and Kenya (country-level SOTER); these were not produced by their national agencies, but rather by international data centres (UNEP, EROS and ISRIC, respectively). Australia’s national map (1:2 M), part of the Australian Natural Resources Atlas, is also freely available after signing a licence agreement. This is a new digital product, drawing upon a wide variety of surveys. A 1:2.5 M map of Russian soils, complete with metadata, map unit descriptions, and representative profiles, is included in the Land Resources of Russia CD-ROM (Stolbovii et al. 2002) and also on-line from the International Institute for Applied Systems Analysis and the Russian Academy of Science. A 1:1 M map of Italian soil regions is freely available. Digital data from various countries including the Netherlands, England & Wales, Scotland, France, New Zealand, and Austria can be ordered on-line, after licensing and applicable fees.

**States, provinces.** Many of the states in the USA have established their own geospatial data infrastructure, usually in the form of a site that points to data sets held by various agencies: for soils, these are the same coverages that are available from the national databases (STATSGO and SSURGO). Several German states, for example, Lower Saxony, have on-line references to digital soil maps at various scales that are available for purchase, but there is no on-line access.

**Point data.** The Soil Survey Laboratory of the US National Soil Survey Center has made available on-line and on CD-ROM the analytical data for about 20000 pedons of USA soils and about 1100 from other countries, most
with morphological descriptions. A standardized set of geo-referenced profiles for global environmental research produced during the WISE project (Batjes 1995) is available for downloading from ISRIC and DAAC and on CD-ROM.

**RELATED INFORMATION**

*Soil classification systems.* Both the World Reference Base (Decker et al. 1998; FAO 1998e) and Soil Taxonomy (Soil Survey Staff 1999) have their own home pages from which both keys and documents can be downloaded. National systems on-line include the Australian (Isbell 1996) and the 2nd Canadian (Agriculture Canada Expert Committee on Soil Survey 1987), but not the 3rd, which is only available in print. A few other systems are briefly described on-line, but not in useful detail. Among interpretive classifications: the Unified (engineering) System is provided on-line by the United States Army (1992); the FAO Topsoil Classification (Land and Water Development Division 1998), which is partly based on the Fertility Capability Classification (Sánchez et al. 2003), is provided by the FAO; and a summary of the UK Land Capability Classification (Bibby et al. 1991) is provided by the British Society of Soil Science.

*Methods.* The USA makes available the full text of the Soil Survey Manual (Soil Survey Division Staff 1993), the Field Book for Describing and Sampling Soils (Schoneberger et al. 2002), the Soil Survey Laboratory Methods Manual (Soil Survey Staff 1996), and the National Soil Survey Handbook (Soil Survey Staff 1997), which includes soil survey interpretations. The Italian Centro Nazionale di Cartografia Pedologica has also placed its procedure manuals online.

*The digital divide* Developing and transitional countries are almost invisible in the effort to make soils information digital and accessible. Such activities are promised (e.g. in India and the Philippines) but with limited results to date. Almost all the digital information on these countries is at small-to-medium scales and produced through international projects. This is not entirely, or even principally, because of the lack of soil surveys in these areas: in part it reflects the overall digital divide and problems with government institutions.

*CULTURES OF DATA DISTRIBUTION* There are clear differences among soil survey organizations in the way that their information is made available to the public. This is related to the legal status of data within a geospatial data infrastructure, the relation of public agencies to the private sector and public (Kabel 2000), financing mechanisms and, more generally, to philosophies of government and civil society.

At one end of the spectrum stand Canada and the national cooperative soil survey of the United States which consider soils information to be a public good, which implies that digital soils data should be made freely available for download. It is seen as a ‘multiplier’: the private sector and academia can use public information as they see fit, and this is expected to lead to economic efficiency for society as a whole, thus fulfilling one of the roles of government (Locke 1690), namely to provide the greatest public good. ‘Our philosophy has been that we want the customers to use the information and if we can give it away it will get used more’ (T. Calhoun pers. comm. 2002). Not incidentally, this helps build political support for the agency and its mission when budgets are discussed by the legislature.

This philosophy dates from the earliest days of the soil survey and its precursors. When the United States Department of Agriculture (USDA) was authorized in 1862, it was to ‘acquire and diffuse [my italics] among the people of the United States useful information on subjects connected with agriculture [and] rural development... in the most general and comprehensive sense of those terms’ (Soil Survey Staff 1997). US Code Title 44/35 of 1980, amended in 1996, mandated ‘the dissemination of public information on a timely basis, on equitable terms, and in a manner that promotes the utility of the information to the public and makes effective use of information technology’. This was put into operational form by Executive Order 12906 (Clinton 1994) which established a national spatial data infrastructure (NSDI) and directed all agencies of the federal government to make their geospatial data available to the public, documented using standards developed by an inter-agency working group, the Federal Geographic Data Committee (2004). This executive order solved the funding problem by making the provision of geospatial data part of the regular tasks of each agency. In the case of the soils data, the agency is the National Resources Conservation Service. The obvious way to carry out this order was to digitize primary data and interpretations, and make these freely available on the WWW, and for reproduction at-cost on CD-ROM.

The situation in Europe is clouded by an EC directive (European Community 1996b) which grants database providers, including public agencies, a so-called ‘extraction right’ that allows them to restrict access. The data are considered proprietary, no matter that the taxpayer has already paid for it to be collected, and this is coupled with mandates from governments that their agencies should be fully or partly self-supporting. So, for example, data are available from the National Soil Resources Institute of England & Wales only under temporary licence from the data provider. Without entering into a specific licence agreement with the provider, no one may provide the data to a third-party, even if interpreted and value-added. Royalties are, in principle, payable for any application, although these may be waived. This is the logical outcome of privatization of public services: if a third party is making money from soils data, the data provider should share in the benefit. The Dutch and several other European nations fall between these extremes. They are charged with diffusing soil survey information to the public but, at the same time, with providing some of their funding. They have a four-level pricing policy: (1) Alterra (own organization); (2) other agencies in the same Ministry; (3) educational organizations sponsored by this Ministry; (4) unsponsored universities and the private sector. In any case an agreement...
for each user must be reached, specifying the terms of use; prices range from approximately €0.05–3 km$^{-2}$ of digital product, depending on the client, intended use and map scale. A major difference from England & Wales is that agreements are permanent and non-revocable, with only a one-time payment, as long as the client abides by the terms of the agreement.

At the European Union level, national policies constrain the international policy. The introduction to the technical report of the 1:1 M European soils database (Jones & Buckley 1996) states ‘data holdings are now viewed as valuable sources of income’. Yet the authors point out the increasing demand for soils data and the wide variety of users. The contradiction between maximizing income to the data producer and maximizing economic value to the European people is not addressed.

Some countries consider geo-data (including soils information) to be state secrets: a relic of colonial times (Shrikanthia 1999) or a totalitarian government. Agencies, in fact, try to work around this by under-the-table compensation and off-the-record agreements. In many developing and transitional countries, the laws governing distribution are unclear, contradictory, or unworkable; and data distribution relies on personal contacts. This situation provides only small benefits to both the state and public.

At the international level, the FAO has a clear policy of making all of its data and publications easily available, if not always downloadable, and ISRIC-World Soil Information is placing data sets co-produced with its partners in the public domain via their website, using its well-known diplomatic skills to make access as free as possible.

BEYOND DOWNLOADING: ADEQUACY AND CORRECT USE

Simply having digital data available does not ensure that it is adequate for use, or that it is used appropriately. An important distinction in this regard is between digital products that are direct conversions of analogue maps and reports, and true soil geographical databases (SGDB), or soil information systems (SIS) (Burrough 1991), which are designed from the start as a digital product. Many projects that begin as simple conversions grade into SGDB as the providers begin to appreciate the intrinsic differences between analogue and digital products.

A major issue as soil survey organizations move to widely-distributed digital products is their fitness for use or adequacy (Forbes et al. 1982). Surveys originally made for agricultural land evaluation define soil classes and mapping units that may not best meet other needs. This is one reason why a SGDB is preferable to a simple conversion. At the very least, new interpretations linked to the existing spatial entities are required. Also, the geodesy is often quite poor for maps made on unrectified or semi-controlled airphoto bases, let alone on old topographic maps, which makes integration into a multi-theme GIS difficult. Experience in the USA shows that it is necessary to recompile boundaries on to a reliable topographic or ortho-photo base rather than simply to digitize them (D’Avelo & McLeese 1998). Expensive re-compilation rather than simple conversion thus seems to be indicated for both thematic and geodesic reasons. In the case of new surveys made on stereo vertical airphotos, it is possible to convert photo interpretation directly to geometrically-correct and geo-referenced maps in a low-cost desktop GIS (Rossiter & Hengl 2001).

A major concern for anonymous provision of data is that the WWW or CD-ROM user may misuse the data, either wilfully or more likely through lack of understanding of the data’s origins, lineage (how the data set was produced, including its source materials and working procedures) and limitations. One approach to this problem is the provision of metadata explaining these issues. The US Content Standard for Digital Geospatial Metadata (Federal Geographic Data Committee 1998) includes a subsection on ‘Use Constraints’, which has been used to good effect in the SSURGO and STATSGO metadata to inform the user about the mapping scale and consequent minimum delineation size, and the purposes for which the survey was made. An entire section on ‘Data Quality’ includes subsections on lineage and accuracy (attribute and positional). Metadata standards are now being internationalized via the International Organization for Standardization (2003).

Usage problems with digital soils data occur if the user attempts to use the data in a GIS at a scale other than that of the original survey. In addition to the purely cartographic generalization found in any theme, soil map unit type and composition change as scale is reduced. Users must understand that, at any scale, map units are not pure and may contain a specified proportion of even strongly-contrasting soils (Young et al. 1997). Medium-scale databases such as STATSGO explicitly mention several soil types per map unit; large-scale databases such as SSURGO make reference to map purity standards, but unless the user is already familiar with this concept, it is easy to miss.

One way to encourage correct usage is to encourage or require data users to register before downloading, and maintain contact with the data provider by e-mail. The above-cited metadata have a subsection on ‘Point of Contact’, and most web pages from which data can be downloaded have at least an e-mail address for follow-up questions.

FUTURE DIRECTIONS

It is to be hoped that more information is made freely available on the Web or on low-cost CD-ROMs, matched with metadata and user documentation to ensure its correct use. There is a dichotomy between those soil survey organizations, professional societies, and educational institutions that are actively making available as much comprehensive information as possible, and those that are continuing with traditional non-distribution. As we strive to increase public awareness of the value of soils information, it seems logical that easy access for both primary and interpreted information should be a top priority. Luke (11: 33) writes ‘No man, when he hath lighted a candle, putteth it in a secret place, neither under a bushel, but on a candlestick, that they which come in may see the light.’ Unfortunately, experience from that day to this shows that this is not always true.
Existing digital soils data for developing countries are provided almost exclusively by development cooperation projects or international organizations. The technological barriers to locally developed and served databases are relatively low: the institutional, legal and cultural barriers are much higher. The development of soils databases will probably mirror development of national data infrastructures in these countries.

ACKNOWLEDGEMENTS

Thanks to Sharon Waltman and Tom Calhoun of the USDA-NRCS for help with the legal status of soil surveys in the USA, and to Niels Batjes of ISRIC for additional information on world soils data sets.

REFERENCES


Jones RJA & Buckley B 1996. European soil database information access and data distribution procedures. EUR 17265 EN, Joint Research Centre of the European Commission, European Soil Bureau Via E. Fermi 1 1-21020 Ispra (VA) Italy.


Locke J 1690. The second treatise of government (an essay concerning the true original, extent and end of civil government). ed JW Gough (1966), Blackwell Oxford UK.


BOOK REVIEW


Scale and the movement between scales are at the heart of most of soil science. Traditionally we study certain phenomenon at specific scales, and then try to infer results for soil behaviour at other scales. Sometimes the movement is to smaller scales, as we attempt to infer molecular level behaviour from ped size behaviour. At other times it is to larger scales, as we try to infer the behaviour of whole landscapes from single soil profiles. Almost every result we obtain is constrained by the levels at which we can measure and observe, and these are almost inevitably different from the scales we want to study. Soil science is thus always addressing the issue of scale, either consciously or unconsciously.

So, it is timely that we have a new book which discusses the developing field of looking at the movement between scales – upscaling (generalizing to bigger scales) and downscaling (inferring smaller scales).

The overall impression is that this is now an active area of research, pursuing all sorts of avenues. In this, soil science is of course not alone, sharing common problems with most other earth sciences (perhaps most notably hydrology) that are also addressing the same issue. So the tools developed here overlap parallel studies in other disciplines, in a way that is truly multidisciplinary.

The editors have assembled a collection of 20 essays, each by eminent practitioners in the field who address the central theme of scaling. That they can do so without significant overlap is itself a reflection of the vigour, breadth and vitality of the field. The contributions range from the downscaling view of soil structure that is given in a a pore-size fractal model, which sees more and more detail as it moves to smaller and smaller scales, to the opposite extreme, which integrates soil science with remote sensing data to give results at a whole catchment scale.

The end result is an enormously valuable book, which encompasses the main threads of an exciting area of work. There is something here for every soil scientist to ponder, even those who find the sometimes-dense diet of mathematics a little indigestible. For those who work in this area of research, it will be an invaluable starting point, opening a veritable Pandora’s box for them to explore.

Inevitably, in a rapidly developing field, no single book will ever be a standard work of reference. Only after the methods have been developed will there be a need for the definitive textbook. But for the time being, this will be a major starting point for those wishing to explore this fascinating area of soil science. Equally, it can never be totally complete, and it is perhaps a pity that none of the authors considers the issue of scale in the other major dimension that affects soil science, namely, time. The editors, authors and publishers are to be congratulated on producing such a fine volume, which I recommend to all soil scientists.

Adrian Armstrong

DOI: 10.1079/SUM2004244

© British Society of Soil Science 2004