Collating and using Australian soil data - a process of aggregation or aggravation?

Peter L. Wilson[A], David Jacquier[A] and Linda Gregory[A]

[A]CSIRO Land and Water, GPO Box 1666, Canberra ACT 2601, Australia. Email peter.wilson@csiro.au

Abstract
Soil and land resource data in Australia is captured and maintained by many jurisdictional agencies and individuals for many purposes. The Australian Collaborative Land Evaluation Program and the Australian Soil Resource Information System have tackled the challenge of collating consistent, standardised national soil data sets largely from legacy data. But how do we best simplify the complexity of soil data and provide appropriate useable data and information products for a wide range of users? Can we specify a minimum set of attributes and a spatio-temporal data model to satisfy more frequent and increasingly complex requests for soil data?

Key Words
Soil mapping, legacy data, aggregation, attributes, information products.

Introduction

Australian soil and land resource data have been captured over the last 70 years through a number of national approaches or regionally specific surveys. Data have been collected for a range of purposes, including scientific research and developing soil-landscape systems understanding, regional land resource inventory and characterisation, and local and site specific studies for particular development or management applications.

Soil and land resource data have been collected by a range of government and non-government organisations within individual jurisdictions and through multi-jurisdictional and nationally collaborative activities. Survey and analytical methods have varied between projects, and have advanced over time. This has resulted in a huge array of different and conceptually inconsistent legacy data. Some of these have been captured in digital information systems, but much remains in original hard copy formats or has been lost to science through inadequate desire, capacity or ability to properly manage historical data.

A strong, though relatively small, community of soil and land resource professionals exists within Australia. This is embodied in organisations such as the Australian Society of Soil Science Inc. (ASSSI) and through coordination mechanisms including the National Committee on Soil and Terrain (NCST). The latter supports the Australian Collaborative Land Evaluation Program (ACLEP) and activities to develop and promote standards such as the *Australian Soil and Land Survey Handbook* (National Committee on Soil and Terrain 2009) and *The Australian Soil Classification* (Isbell 2002). This effort is underpinned by an improving data and information infrastructure, and mechanisms for consistent national data collation, analysis and reporting such as the Australian Soil Resource Information System (ASRIS) (McKenzie *et al*. 2005).

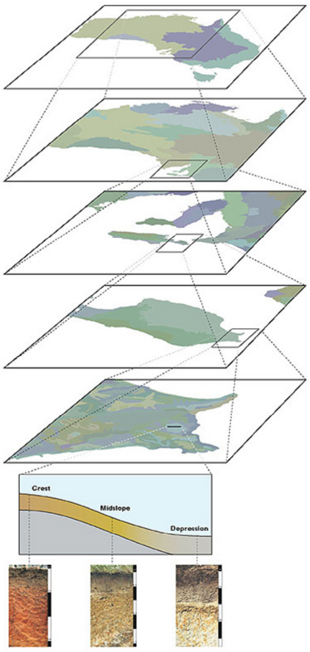
Making soil data useful

Soils ain't soils. The processes of soil formation, interaction and modification are complex. The data collected to understand soil processes, describe their characteristics, map their distribution and model their interactions and alterations over time are complex. The systems developed to capture, manage, store, analyse and disseminate data and information on soils are also, therefore necessarily complex.

"Quick, give me a soil map!" is a common request issued by scientists, modellers, policy and decision makers to the gate keepers of soil information systems. "Sure - but which one?" is the common reply. This should be the start of a healthy conversation, exploring the complexities of soils and the data describing them and then, deciphering the user's specific needs, namely: the explicit soil characteristics required and their spatial and temporal representation. However, the conversation often ends in aggravation and frustration borne by data supplier and user alike due to poor user-needs articulation, and an inability by the custodian to supply the required attributes in the desired timeframe.

Aiming for "Utopia" - or striking a happy medium?

The ASRIS (www.asris.csiro.au) is underpinned by a hierarchy of mapping resolutions from broad national divisions (Level 1) through tract mapping of districts and systems (Level 4 to 5), to detailed site-based property and condition descriptions (Level 7) (Figure1). Adding to the spatial complexity is the data base recording of values for unmapped components within each of the mapped tracts and the use of a five layer functional depth model to accommodate variability of characteristics down the soil profile (Figure 2).



Level and tract name	Mapping window	Main attributes used for mapping	Typical uses for the information
Level 1 Division	30 km	Broad physiography (slope and relief)	Broad geographic context
Level 2 Province	10 km	Water balance, physiography	National natural resource policy
Level 3 Zone	3 km	Substrate lithology, water balance, physiography	Regional natural resource policy
Level 4 District	1 km	Groupings of geomorphically related systems	Catchment planning, location of new industries
Level 5 System	300 m	Local climate, relief, slope, lithology, drainage network, soil profile class	Catchment management, hydrological modelling, land conservation, infrastructure planning
Level 6 Facet	30 m	Slope, aspect, land curvature, soil profile class	Farm management, land-use planning, on-ground works
Level 7 Site	10 m	Soil properties, surface condition, microrelief	Precision agriculture, site development

Figure 1. ASRIS seven level spatial hierarchy (McKenzie *et al*. 2005)

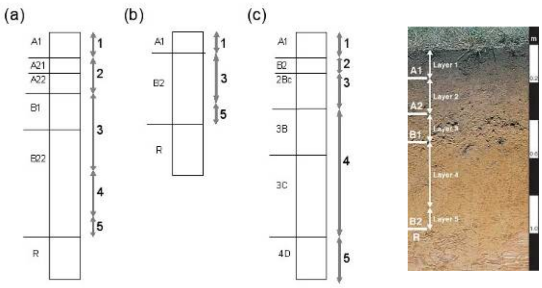


Figure 2. ASRIS five layer depth model (McKenzie *et al.* 2005)

A seemingly simple request for a soil map therefore, may result in a seemingly simple soil classification product, based on the dominant occurrence in non-homogeneous mapped entities, using conceptual or reference profile data which may not adequately or accurately represent the spatial, depth or temporal variation of soil characteristics at any actual point on the ground. Over simplification or over generalisation of the complex soils data can result in meaningless products or misrepresentation of the characteristics and the variability that may be essential to a particular modelling or management decision process. Striking a happy medium and producing a generic soil map can result in something not much more useful than a decorative picture.

On the other hand, the spatial and temporal complexity of soil characteristics cannot easily be summarised or visually represented. The specific data requirements for a particular modelling or decision making process need to be clearly and explicitly stated. Often this could result in a number of complex soil inputs that may not be able to be easily understood or assimilated into particular modelling or decision making processes. Aiming for "Utopia" therefore does not always satisfy users who often just want a simplified representation of the soil complexity, albeit at spatial and temporal resolutions that fulfil their specific needs. Users only want the attributes important to them, not necessarily the full complexity of the soil landscape system.

ASRIS aggregation products

The ASRIS web mapping interface provides a window into the complexity of data comprising Australia's national soil data collection. However, the interface does not provide access to all collated data and does not allow downloads of the data due to licensing limitations by some data contributors. In an effort to make at least some national soil data more readily accessible, the interface does provide visualisation and query of all available tract mapping at the various ASRIS levels. It also has aggregated representations of attribute values for a number of the commonly required soil properties for each of the five depth layers (Figure 3). However, many users of ASRIS are critical that they cannot find or use the data they want, and cannot easily understand the level/layer framework of ASRIS or the products presented - even in this simplified form.

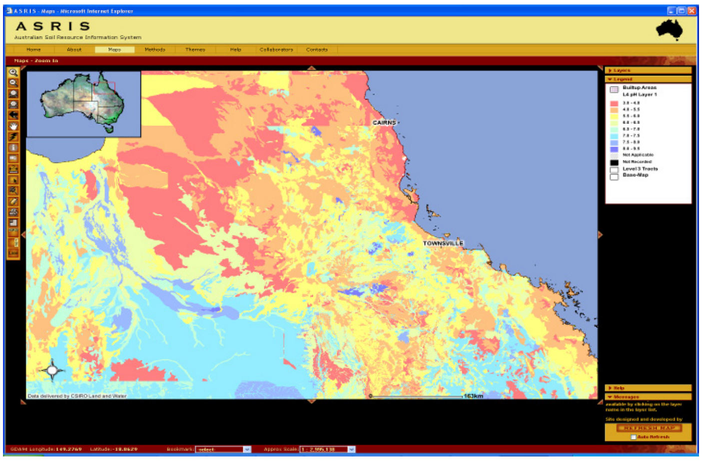


Figure 3. ASRIS Level 4 mapping showing dominant pH for layer 1

There are many methods by which component and depth data can be aggregated (Soil Data Viewer 5.2 User Guide - United States Department of Agriculture http://soildataviewer.nrcs.usda.gov/default.aspx). Generic ASRIS data views use dominant component or weighted average approaches. However many different views of the same data could be created using different aggregation approaches, and each may be more appropriate for a particular application. Aggregation of complex data using large national collations (such as ASRIS) can be computationally intensive, and may not lend themselves to on-the-fly, on-demand supply of requested data products. Additionally, regular updates or improvements to underlying data would necessitate lengthy re-creation of stored aggregation data sets.

New aggregated data products have recently been completed to support specific project requirements for carbon and acidification monitoring and prioritisation (Baldock *et al*. in prep, Baldock *et al*. 2009, Wilson *et al*. 2009a). The data have been aggregated both spatially and with depth to create 0-10 cm and 0-30 cm depth-slices using weighted averages. These products have been generated for parts of Australia using a 250 m cell sized representation to allow analytical computation within realistic project timeframes (Figure 4). These products provide another view of ASRIS data, which may or may not be useful inputs to other analyses.

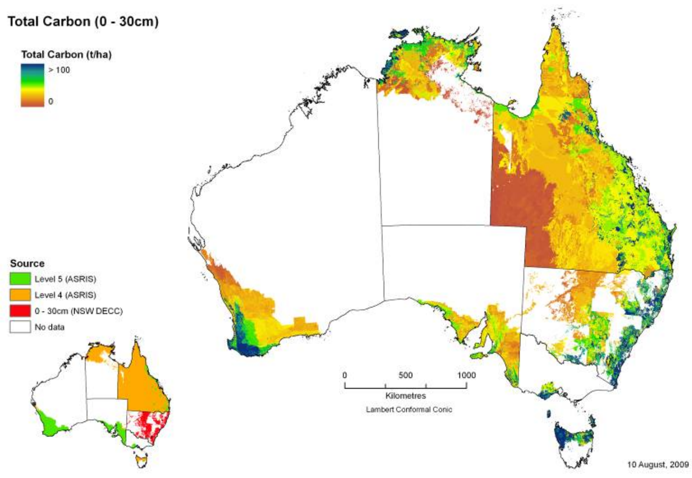


Figure 4. Spatial and layer aggregation of ASRIS data (Baldock *et al.* in prep)

Discussion

ASRIS supports the vision of making the best available soils data and information readily accessible and useable to a wide range of users. Current national data collations are now extensive and comprehensive enough to be useful inputs to many national and multi-jurisdictional analyses. A number of generic views of ASRIS data and project-specific aggregations, such as depth-weighted limitation values for agricultural land suitability assessment of northern Australia (Wilson *et al*. 2009b) have now been created, but the utility of these for other projects remains untested. The specific soil data needs of different analyses need to be well considered so that the right information products can be developed. Much angst can be generated if users do not understand the complexity of soils data, and cannot explicitly define the data inputs they require. Different aggregation processes and analytical methods (such as calculating values for individual un-mapped components then averaging, versus averaging components then calculating) can produce very different data products and therefore very different analytical results. The use of easily obtainable, over-simplified soil inputs will not result in appropriate representation of soil characteristic variability in modelling or decision-making processes, and therefore outputs and the decisions that ensue may be compromised.

Current discussions are focussed on creation of fine resolution data sets (e.g. 90 m continental grids, see www.globalsoilmap.net) with comprehensive depth models (e.g. attribute depth splines, see Bishop *et al*. 1999). While data sets in this format may be logical for better representation of the spatial, depth and temporal variability of soils, they may actually exacerbate analytical difficulties for different users due to the data volumes that will result. It is also questionable whether current computing capacity and software will allow ready use of such data models by many users. A short term solution will be to develop an agreed approach to grid representations of key soil attributes with explicit estimates of uncertainty and data quality. There is a need for soil scientists and information managers to work closely with soil data users to define the data requirements, and to build tools to assist flexible generation of products in a timely manner. The ACLEP will assist with this task, and also the promotion of information on soil and soil data complexities to the wider user community.

Conclusion

Users of soil data and information need to be adequately aware of the complexity of soil characteristics and their landscape interactions over space and time. Complex data required for soil description must be simplified to provide easily assimilated inputs into modelling and decision-making processes without losing the essential components of variability. Potentially, a set of standardised, aggregated soil characteristic data may be defined that could support a wide range of users needs. However, many users will require non-generic, project-specific inputs. Tools allowing flexible query and timely generation of appropriate data must be developed and made freely available. ACLEP will progress ongoing development of ASRIS national data and access tools to fulfil this need.

References

Baldock JA, Grundy MJ, Griffin EA, Webb MJ, Wong MTF, Broos K (in prep.) Building a foundation for soil condition assessment. CSIRO Land and Water Science Report.

Baldock J, Grundy M, Wilson P, Jacquier D, Griffin T, Chapman G, Hall J, Maschmedt D, Crawford D, Hill J, Kidd D (2009) Identification of areas within Australia with the potential to enhance soil carbon content. CSIRO Land and Water Science Report.

Bishop TFA, McBratney AB, Laslett GM (1999). Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma* **91,** 27–45

Isbell RF (2002) The Australian Soil Classification, Revised Ed. CSIRO Publishing: Collingwood, Australia.

McKenzie NJ, Jacquier DW, Maschmedt DJ, Griffin EA, Brough DM (2005) Australian Soil Resource Information System, Technical specifications Version 1.5 October 2005. National Committee on Soil and Terrain Information.

National Committee on Soil and Terrain (2009) Australian Soil and Land Survey Field Handbook, Third edition. CSIRO Publishing: Collingwood, Australia.

Wilson PL, Baldock J, Grundy M, Jacquier D, Griffin T, Moody P, Chapman G, Hall J, Maschmedt D, Crawford D, Hill J, Kidd D (2009a) Identification of land with a risk of acidification. CSIRO Land and Water Science Report.

Wilson PL, Ringrose-Voase A, Jacquier D, Gregory L, Webb M, Wong MTF, Powell B, Brough D, Hill J, Lynch B, Schoknecht N, Griffin T (2009b) Land and soil resources in Northern Australia Land and Water Taskforce *full report* October 2009. CSIRO Sustainable Ecosystems Science Report.