Geoderma 170 (2012) 232–238

Contents lists available at SciVerse ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Soil and informatics science combine to develop S-map: A new generation soil information system for New Zealand

L.R. Lilburne *, A.E. Hewitt, T.W. Webb

Landcare Research, PO Box 40, Lincoln 7640, Canterbury, New Zealand

ARTICLE INFO

Article history:
Received 11 January 2011
Received in revised form 19 October 2011
Accepted 21 November 2011
Available online 24 December 2011

Keywords:
Soil survey
National soil database
Information technology

ABSTRACT

Upgrading from a traditional soil spatial database, consisting of a single GIS layer of polygons and basic attributes, to a modern soil information system has required significant informatics resourcing. This paper describes the design criteria of New Zealand's new soil mapping system, and the IT infrastructure required for its support. The hybrid design incorporates both traditional soil survey techniques and data, and newer digital soil mapping techniques and information, to (eventually) achieve full coverage of New Zealand at 1:50 000 scale.

Full advantage has been taken of recent advances in informatics science in the area of integrated database tools, modelling, remote sensing and web technologies. A number of in-house modules have been developed for functionality including:

1. Data entry, storage and validation of soil data and photos
2. Dynamic generation of spatial data, maps and factsheets highly customised for a range of end-users
3. Automated generation of relevant metadata reports
4. Running pedo-transfer functions (PTFs) and other digital soil mapping operations
5. Managing and simulating soil uncertainty information

1. Introduction

Worldwide there is renewed awareness of the need for good information about soils, to support wise use of such a critical life-sustaining resource (Oldeman and van Engelen, 1993). This information is needed at a variety of scales and for many purposes ranging, for example, from micro management of plots of land to studying global climate change impacts. Many countries are now focusing on updating and modernising their soil databases (Hartemink et al., 2008) and are collaborating to create a new global soil map (Sanchez et al., 2009). In this renewed effort in collecting soil resource information, the role of informatics has become central, not only for storing the data but also for a range of other key tasks. These tasks include manipulating and transforming basic soil data into a variety of quantitative soil information via soil inference systems (McBratney et al., 2002); devising new methods for different users to search, query and access a range of suitably formatted soil information; and maintenance (verification and update) of the soil information and inference system. This paper discusses the significant behind-the-scenes informatics contribution to the development of New Zealand's new soil information system.

New Zealand has a large accumulation of soil data in the form of soil maps acquired over the last 70 years. Most New Zealand soil surveys have been undertaken prior to the advent of computer programs and simulation models, and soil data have largely been collected in the context of interpretation for land use suitability. In this context, a qualitative estimate of such attributes as drainage class or soil depth was considered to be adequate in providing guidance on land-use issues. These data served the agricultural community well last century but modern users of soil information require soil data that are more quantitative, digital and accurately geo-referenced, for a wider array of uses including assessing irrigation demand, or crop-growth and environmental-risk simulation modelling. Further problems with New Zealand's historical soil surveys include the imprecise definition of soil series, the proliferation of soil series (many of which appear to be very similar to one another), inconsistency between survey maps, and lack of data on map unit composition.

The combination of the very limited number and location of soil profiles with measured data and the complex nature of soil distribution processes in New Zealand means that digital soil mapping (DSM) techniques (Lagacherie et al., 2007) are not able to deliver a new soil map of sufficient detail that covers all of New Zealand. There are insufficient resources to fund collection of new comprehensive point data, and useful covariate layers for Scorpan (McBratney et al., 2003) modelling are unavailable in areas with low relief. However,

Abbreviations: DSM, digital soil mapping; pdf, probability distribution functions; PTF, pedo-transfer functions.

* Corresponding author: Tel.: +64 3 321 9850; fax: +64 3 321 9998.
E-mail addresses: lilburnel@landcareresearch.co.nz (L.R. Lilburne), hewitta@landcareresearch.co.nz (A.E. Hewitt), webbt@landcareresearch.co.nz (T.W. Webb).

0016-7061/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2011.11.012

more traditional soil survey mapping methods (TSM) are also not able to produce a new database for the whole country due to the same lack of resources to tackle the necessary manual tasks of upgrading existing maps and mapping new areas with no legacy data. Therefore, we have devised an integrated hybrid approach that combines restructured and updated legacy data for the lowland areas of New Zealand with more modern DSM methods for the more poorly mapped higher-relief areas, to create a flexible, dynamic and innovative soil database that can meet the various needs of modern users of soil information.

This paper describes this approach and the resulting soil information system (S-map). A key part of S-map is the underlying informatics infrastructure, which is also described.

2. Background

2.1. New Zealand

Soil survey in New Zealand from 1938 to 2001 has resulted in a set of soil maps of varying quality at varying scales. By the end of 2001, the whole country was covered by 1:253 440 scale soil maps. In addition, just over 50% of the country is covered by more detailed maps (with scales ranging from 1:126 720 through to 1:10 000). Many soil surveys included a soil bulletin in which the mapped soil series were qualitatively described. Most of these soil maps have been made under guidelines defined in *Soil Survey Method* (Taylor and Pohlen, 1970) and *Soil Description Handbook* (Milne et al., 1995). Simple map units are defined as delineations where at least 85% of the area contains soils conforming to the definition of a single soil type. Compound map units comprise more than one named soil type. Unhappily, this standard of map purity has rarely been attained. Studies have shown that there is considerable variability in soil properties and soil types within soil map units on alluvial plains in New Zealand (Adams and Wilde, 1976; Di and Kemp, 1989; Karageorgis, 1980; Webb et al., 2000). In general, knowledge of soil variability gained in the course of soil surveys has not been captured on the resulting soil maps and reports. Most map units are simply described as being represented by a typical or modal profile, leaving the user with no information on soil variability.

In summary, the main limitations of these soil survey data are

- Incomplete coverage
- Inconsistency and varying standards between survey maps
- Proliferation of soil series (many of which appear to be very similar to one another)
- No information on map unit variability
- Qualitative information targeted at land use suitability
- Little quantitative information, making the survey data unusable for input into simulation models

The current national soil map is contained within the *New Zealand Land Resource Inventory* (NZLRI) (Newsome, 1992), which was later enhanced by the addition of 15 soil properties collectively known as the fundamental soil layers (FSL) (Wilde et al., 2000). The NZLRI was compiled at a scale of 1:63 360 from a range of soil maps, mostly pre-dating 1979. This required simplification of more-detailed survey polygons; consequently it does not contain the best available linework. In some areas the only source of information is the General Soil Survey maps of New Zealand (1:253 440 scale).

Variability information in the NZLRI is limited to identifying two component soils in some map units. No confidence or reliability information is provided. In contrast, confidence and variability information is supplied with the FSLs but again the best knowledge is not available, as, firstly, the description of each soil property is limited to five (in most cases) predefined ranges or classes. Secondly, variability is simply defined as the number of classes of predefined ranges in each map unit, resulting in large ranges for each soil property for each map unit.

Essentially the current soil map consists of a single spatial layer of polygons or map units and attributes describing 20 soil classes and properties. In common with most other national databases, the attributes represent modal or representative values for each soil type. Attribute information mostly pertains to the whole profile or to just the topsoil, with three attributes for the subsoil layer. Standard GIS tools were used to manage both the polygon and attribute data, which have been static since the completion of the more recent soil attributes in 2002.

2.2. Australia

Johnston et al. (2003) and McKenzie et al. (2008) describe the development of the Australian Soil Resource Information System (ASRIS). As a larger country they have had to reconcile differences in taxonomy, soil survey and soil analysis procedures that were applied by the various agencies in the different states. The size of Australia has also required them to develop a spatial hierarchical approach of seven levels where lower levels contain more spatial detail (McKenzie et al., 2005). Only the top levels have complete coverage. Levels 4–6 including the "facet" level (1:100 000 scale) have 60% coverage of intensive areas and 5% coverage of rangelands in 2007 (ASRIS, 2011). Soils are modelled as having five contiguous layers, with attributes relating to soil thickness, water storage, permeability, fertility, salinity, and erodibility. Uncertainty estimates (spatial and attribute) are provided with most data, along with a method code describing how the estimate was derived. ASRIS is a combination of vector and raster data. DSM techniques are now included in the latest edition of Australia's soil survey guidelines (Bui, 2007).

2.3. Europe

A report details the state of soil mapping and databases in each country in Europe (Jones et al., 2005). The state of the national soil information systems varies considerably from country to country, although minimal funding for new soil survey work over the last few decades is a common theme. Smaller countries such as Belgium, The Netherlands and Denmark are well served by complete coverage of soil survey maps, a comprehensive set of profile data, and the establishment of pedo-transfer functions including geostatistical estimates for a wide range of soil properties. The United Kingdom has complete coverage of soil survey maps at the 1:250 000 scale. They are trialling DSM-type techniques to produce soil maps at 1:50 000 to meet user demand (Thompson et al., 2005). Larger countries such as France do not have complete coverage and are looking to develop new techniques (King et al., 2005).

2.4. North America

The USA established the SSURGO (1:12 000–1:63 360) and STATSGO (1:250 000) soil databases in the 1990s. Coverage of SSURGO is close to complete in the east, less so in the west (NRCS, 2010a). SSURGO contains the range and representative value for each polygon for a range of soil properties (NRCS, 2010b). These ranges and polygons were defined by soil surveyors using available measured profiles and soil–landscape relationships. Map units can be composed of up to three component soils designated as consociations, associations, complexes or undifferentiated groups. Methods for aggregating component information have been developed (NRCS, 2010c). Landform properties form part of the soil definition.

Canada also has multi-scale databases where the Canada Land Inventory is an older land capability database of 1:250 000 scale. Detailed soil surveys vary in scale from 1:10 000 to 1:250 000 (AAFC, 2010). These are legacy surveys and only cover the most significant agricultural areas. In general, there are correlation differences and lack of edge matching between soil surveys.

2.5. Rest of the world

South Korea is another country that has invested in the development of a soil information system. Soil property and crop suitability layers are delivered over the web. They plan to incorporate DSM techniques (Hong et al., 2010). South Africa has developed a web portal to deliver generalised soil maps and some interpretative layers. The Brazilians are actively investigating DSM techniques to map some of their states (Hartemink et al., 2008). Most other countries appear to have either very poor soil data or digitised legacy survey maps (ITC, 2011).

2.6. Global soil map

Global soil maps have been coordinated by FAO (2010) at 1:5 000 000 scale. More recently a new initiative called the Global Soil Map (GSM) is underway to develop a common database structure for soil raster data at the 90-m-resolution scale. This will provide outputs that are more easily assimilated by models than traditional soil polygon soil maps. In more developed areas, existing databases will be converted and ported into GSM; in other countries, particularly African and Asian, new DSM approaches will be developed. The initial concept is for six attributes to be mapped (GSM, 2010).

3. S-map database principles and approach

S-map is New Zealand's new national soil information system. S-map soil data are primarily polygon-based despite much attention in the literature to modelling of soil attribute surfaces (Burrough et al., 1997; Moore et al., 1993). We recognise the power of statistical interpolation techniques and other DSM techniques, and their importance, and are designing S-map to underpin such techniques. A polygon base is required for the following reasons. First, available point data in New Zealand are limited at a national scale and by a sampling density that will not allow generation of accurate high-resolution surfaces by geostatistical analysis of point data. Useful covariate layers are also limited. Second, there is much polygon-based information within existing hard-copy soil survey reports, and as expert knowledge, and this needs to be made available for digital analysis (Hewitt et al., 2008). Third, completion of a national soil polygon layer will provide a sound starting point for use of legacy soil information in DSM by achieving correlation of soil taxonomic classes, harmonisation of neighbouring soil maps, and registration of soil maps to a common up-to-date topographic base. Fourth, the use of polygons operationally suits the description of many landforms in New Zealand that have relatively well-defined boundaries but contain high soil variability, for example the alluvial fans, terraces, and flood plains. The polygon base information will be an important input for the generation of other raster soil layers where the available information and spatial distribution patterns are more suited to DSM techniques.

S-map has seven founding principles. It:

1. Describes soil (to a depth of 1 m) according to its functional characteristics (the 1-m depth limitation was selected for pragmatic reasons due to limited resources)
2. Correlates soils on a national basis
3. Focuses on data rather than mapping — thus releasing cartographic constraints on map unit depiction
4. Incorporates knowledge of map unit variability and uncertainty
5. Develops a soil information platform suited to modelling
6. Combines traditional and digital soil mapping techniques with the ability to include some expert knowledge
7. Is able to provide readily understood, easily accessed information to a wide range of users.

3.1. Soil description

S-map focuses on soil information. Parent material is included as it is part of the soil classification. Other closely related environmental data, such as climate and topography-related variables (e.g. rainfall, slope), will no longer be part of the soil database, nor will they be used directly to define soil types. Even landform type will no longer be explicitly specified. It is expected S-map will be combined with these other environmental layers, as appropriate, to produce spatial layers of interest. For example, a simple model of water deficit would include rainfall, evapotranspiration and vegetation information as well as soil variables.

The key descriptive criteria for describing soil now focus on its functional characteristics, i.e. the practical performance or utility of the soil for management purposes. This does not preclude soil genetic interpretations from the soil data.

3.2. National soil correlation

The second principle is to develop a national legend through a clear objective definition of soil taxonomic units. Currently, soils with very similar properties have been recognised as unique soil series because of subtle differences in geomorphology, climate, parent materials or geography. The correlation of soils from widely separated geographic locations (such as North Island and South Island) may cause some confusion for local soil map users, as many familiar soil series names will disappear. A strict definition and application of the soil taxonomy criteria will also require the splitting or realigning of many current soil series, where the current series spans the new taxonomic criteria. However, there are great benefits from this process. Well-defined soils enable confidence and consistency in classification and clarity of communication about soil characteristics. It will also better support technology transfer and enable easier pooling of data for application of pedo-transfer functions to unsampled soils. The new correlation criteria are described in Webb and Lilburne (2011), Hewitt (2010) and Schmidt and Hewitt (2004). The taxonomic hierarchy is now Soil Order, Soil Group, Soil Subgroup, Soil Family, Soil Sibling. Each sibling is defined as having up to six Functional Horizons (Webb, 2003), which are defined according to soil morphology (stone content, structure-size, texture and consistency).

3.3. Data focus

In the past, cartographic constraints have dictated the level of spatial and thematic detail, e.g. polygon size and orientation, and label information. For example, long thin horizontal polygons were more acceptable than vertical ones as the former could more easily be labelled. The need for an aesthetically consistent hard-copy product meant that more detailed linework was spatially generalised, such that in some areas the national soil layer did not contain as much information, or was not as accurate, as the source soil survey. As S-map is primarily a digital product, it can now contain the best available soil information at any given location. However, surveys more detailed than 1:15 000 will not be included. The digital rather than paper format also enables full flexibility in what information is conveyed in polygon labels.

3.4. Incorporation of map unit variability and uncertainty

To date little information has been provided on map unit variability or uncertainty in New Zealand. However, pedologists have accumulated much more knowledge of soil variability from considerable fieldwork over several decades. It is important to record this knowledge before it is lost due to retirement of personnel. Important knowledge to record is classification reliability or accuracy, and likely map unit composition (including the component siblings, their proportions and alternatives). This is achieved through qualitative reliability or confidence codes for each component sibling and each of its 17 soil classification properties, and the opportunity for the pedologist to specify alternative classification options. This allows the pedologist to specify for example, that the assigned topsoil stones category of a map unit is reliable but the drainage could be a different category (poorly drained) to that specified (imperfectly drained).

In addition, it was considered important to quantitatively characterise the variability of key soil properties in a way that reflected best available knowledge — right down to the individual polygon level if appropriate. This necessitates a flexible structure that is not based on predefined ranges or depths. Our new database structure is based on specifying probability distribution functions (pdf) for each of the key soil properties (profile depth, depth to slow layer, rooting depth, horizon thickness, stone, clay and sand content for each horizon). These pdfs describe the variability of the soil property within the map unit. Source-of-estimate codes describe the quality of information used to estimate parameter-values for each pdf.

This structure must be the same for soils for which much is known (lowland soils) as for upland soils where little is known. The latter will be less precisely and accurately defined than the former. Storing information about soil variability and uncertainty will allow environmental modellers to be more aware of the limitations of soil data. S-map will be capable of providing alternative spatial realisations of soil properties of interest that can be used in stochastic uncertainty analysis modelling, e.g. (Hansen et al., 1999; Heuvelink et al., 2010)

3.5. Development of a soil information platform

Soil data need to be provided as a comprehensive platform for modelling of any environmental state, land-use suitability or vulnerability that involves the soil. S-map is designed so that it can be readily integrated with other data sources. A strong link with soil profile data stored in the national soils database (NSD, Landcare Research, 2003) is essential so that future pedological modellers can develop additional spatial layers of specific soil properties based on analyses in the NSD. It is expected that the combination of S-map and the NSD will significantly enhance the value of both databases. Finally, a modelling platform requires knowledge to be stored in formalised fields in a database. Use of textual or descriptive information as found in the NSD or in legacy map legends can be very difficult to interpret automatically, therefore free text fields are avoided in S-map.

3.6. Combination of traditional and digital soil mapping techniques

Very good polygonal data exist for much of New Zealand's lowlands. The rolling, hill and high country of New Zealand are less well mapped. Hillier landscapes are more suited to predictive modelling using terrain, geological and climate attributes as covariates. Consequently an approach whereby the landscape is digitally divided first into soilscapes (Hewitt et al., 2008), then into land elements, using a digital terrain model (Barringer et al., 2008; Schmidt and Hewitt, 2004) and linked by means of limited sampling and expert knowledge to S-map siblings (Barringer et al., 2008), has been used to map large tracts of the eastern dry greywacke high country in the South Island. Other DSM techniques have been tested (Hewitt et al., 2010). The advent of the ALOS PRISM satellite is opening up opportunities to apply DSM techniques on the hilly land in the North Island. This is because the 5-m DEM that can be derived from this satellite is much more suited to the shorter hillslopes found in the North Island.

S-map will eventually comprise both vector and raster layers that either link to siblings, key soil properties (directly recorded or predicted via pedo-transfer functions PTFs), or to individual soil attributes (which might be based on a model of sibling soil properties combined with other layers, or derived from interpolated point data). Initially the focus is on the polygonal (vector) format, but S-map also comprises raster information where this is more appropriate.

3.7. Readily accessible data

S-map includes a number of database views and formatted reports aimed at a variety of users ranging from soil experts to those with little soil knowledge. Spatial soil data can be easily provided to environmental modellers according to their data requirements. These will be made available through web service technologies as well as by means of traditional downloads.

4. Design of S-map and its associated software tools

Advances in information technology are enabling the development of complex soil databases more appropriately termed soil information systems (Lagacherie and McBratney, 2007). S-map is an information system comprising a complex database structure, pedotransfer functions or soil inference system (McBratney et al., 2002), rule-based validation checks, automatic taxonomic correlation, data modification history, dynamic factsheet generation, an image/photo database, and database management. This and further functionality is provided through the development of a number of software modules (Fig. 1) which are now described. Modules preceded by "Smap" are developed in-house specifically for S-map, those preceded by LCR are generic modules developed in-house by Landcare Research. Modules with no shading have yet to be developed.

All data are stored in a Microsoft SQL Server database. These include all the soil attribute and uncertainty data, images of profiles, valid users and their permissions, and all metadata (including metamodels, i.e. metadata for PTFs and other models). Polygons are stored in the Microsoft SQL Server database as well — currently via ESRI's SDE product.

PTFs are coded using the .NET Framework Common Language Runtime (CLR) technology, which has been integrated into MS SQL Server. This makes for a very powerful and extremely flexible programming tool allowing for the development and maintenance of complex PTFs in the form of CLR functions stored in a database "assembly" or function library (SmapPTFLibrary). Spatial SDE views (SmapSpatialViews) can then be established using these CLR functions. This effectively allows dynamic access to all of the soil data including those derived using the PTFs, from a GIS. These spatial views are used to generate layers with any combination of specific soil information as required by users.

The same CLR technology has also been used to develop another library of functions (SmapAggregateLibrary) for reformatting soil property data as required by different users (e.g. different units, numbers vs text strings, concatenated string labels). These functions also include the ability to aggregate the properties of multiple siblings in a polygon. For example, one function determines the dominant texture (or any categorical property) and its proportion of a polygon by taking into account all the component siblings and their proportions. Another can calculate the minimum value, maximum, range or weighted average of any numeric soil property including those derived from PTFs, e.g. the average profile available water (weighted by the proportion of each sibling). This allows uncertainty information to be easily presented to the user.

A .NET application has been written to facilitate data entry by pedologists (SmapDataEntry). It has been designed in a spreadsheet format with built-in assistance (Fig. 2). This tool is used to enter all the family and sibling soil information — qualitative and quantitative, the uncertainty data, the map-unit (polygon) composition data, as

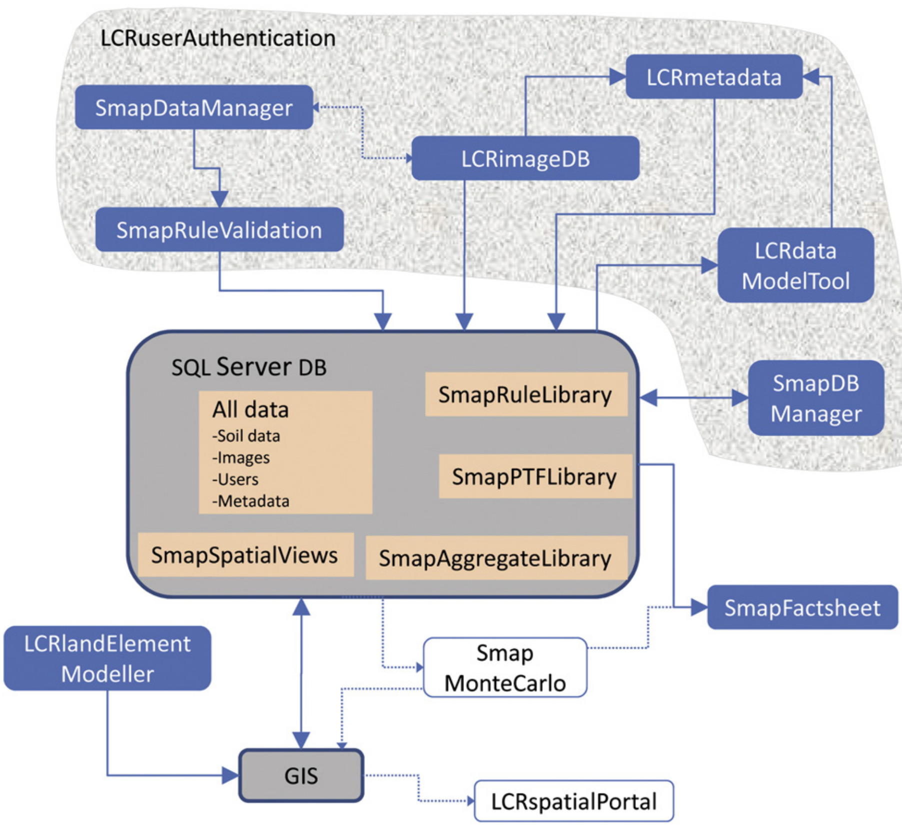

Fig. 1. Diagram of the S-map software infrastructure. An MS SQL Server database contains all of the data. Modules prefixed by "Smap" are S-map specific, and by "LCR" are generic in-house modules. Hollow modules and dashed lines have not yet been completed.

well as correlation history, and factsheet links. The tool performs taxonomic correlation on-the-fly with newly entered data being linked to the appropriate family and sibling. Simple built-in database verification checks are complemented by another .NET application (SmapRuleValidation) that has been written to manage rather more complex rule-based consistency checks of the data. For example, CLR routines (SmapRuleLibrary) have been written to check that the sum of the horizon thicknesses as described by the pdfs and the horizon morphology is consistent with the pdf for rooting depth, and the soil classification of the linked sibling. A peer-review system for the data has been built into the SmapDataEntry module. This allows any exceptions to the rule validation process to be defended by a pedologist, and then either accepted or rejected by another senior pedologist. The explanations are stored for future reference.

Landcare Research (LCR) has developed a generic module (LCRimageDB) to manage its large collection of photos and other images of New Zealand flora and fauna. Soil profile photos have been entered into this database system and linked to S-map siblings.

In addition to the standard SQL Server database management tools, an additional tool (SmapDBManager) was developed to facilitate deleting linked records (e.g. all quantitative and uncertainty data associated with a specific sibling), and some other tasks made more difficult due to the complex database structure. A generic module (LCRuserAuthentication) manages users and their associated permissions. These specify what operations each user can perform (e.g. read-only, data entry, data edit, peer review, rule management, image management, S-map database management).

Web-based factsheets can be dynamically generated from the S-map database. Currently the reports are developed using the Crystal report writer software (SmapFactsheet). These access the PTF and reformatting functions in the SmapPTFLibrary and SmapAggregateLibrary, as well as photos from the image database (LCRimageDB). Factsheets can be generated by specifying a soil name or a spatial location (see http://smap.landcareresearch.co.nz). It is planned to move the factsheet generation to a new generic module still under development (LCRfactsheet). A set of factsheet templates is being developed,

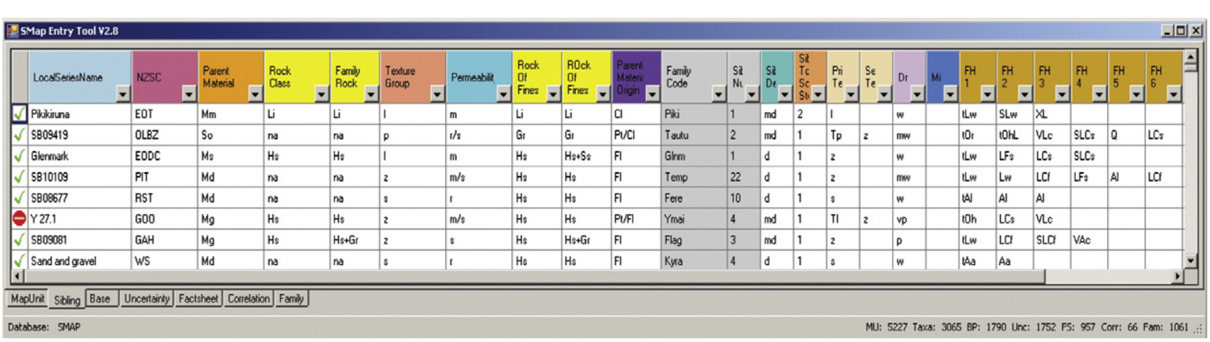

Fig. 2. Screen snapshot of the SmapDataEntry tool.

some with a particular focus on providing soil information in lay terms, others more technical. Factsheets can be easily designed to suit a variety of end-users, e.g. pastoral farmers vs cropping farmers, forestry and other land managers, regulatory agency staff interested in management of risks, people trained in soil science, school children.

The GIS tool is used to export spatial data for users in a variety of standard formats. The GIS/SQL Server combination facilitates customisation of the data export for different users, where specific soil properties can be added, renamed or omitted as required. This requires the use of another generic module (LCRdataModelTool) that can dynamically generate a metadata report that is specific to the exported data. For example, if a data export includes a particular soil property or PTF, then the appropriate attribute information, descriptive text, lookup values, formula, etc. are output into a metadata report that can be sent to the end-user along with the exported data. Both the LCRimageDB and LCRdataModelTool access the LCRmetadata service, which manages the relevant XML schemas and data.

The land element and DSM modelling work are encoded in standalone GIS programs (LCRlandElementModeller).

Yet to be completed are web-based tools (LCRportal) for dynamic spatial dissemination of the soil information. The aim is to provide user access to the information at different levels: the simple easy-to-use point-and-click "what soil data is here" level, predefined data download, and direct modeller access via web services. Another tool that will be developed (SmapMonteCarlo) is an application to create random realisations of soil properties (spatial and aspatial) given the uncertainty and variability information in the database. These can then be used in Monte Carlo uncertainty analysis or in stochastic models.

5. Discussion

The S-map information system is considerably more powerful than its outdated ancestor. This has required the development of a software infrastructure to perform a variety of tasks ranging from managing linked data in separate tables and databases, data verification, reformatting and derivation of further soil information, user permissions, and report generation. The flexibility of the new design has allowed pedologists to enter the best available soil knowledge whether this is detailed expert knowledge of soils in a particular location, or modelled estimates based on terrain and other environmental correlates for much larger areas. Both highly variable soils and uniform soils can be appropriately described, and pedologists are able to highlight any soil properties in any particular area that are less certain than others. Considerably more information is stored on soil variability than in the previous model of dominant soil and a modal profile. Dynamic soil correlation on-the-fly is implicitly resulting in a major restructuring of old soil series names into a much more functionally organised taxonomy.

Coverage of New Zealand in S-map is far from complete, however. Efforts to gain funding for S-map have targeted national and regional government agencies and farm servicing agencies. Each of these has budget constraints, and although they generally express a desire to support the project this is not readily translated into financial support. Funding is slowly building but carries with it two problems. First, each agency has its own interest and few are free to consider S-map development for the general good of New Zealand. Second, funding accumulates in a series of small parcels that carry a high transaction cost to administer.

DSM promises efficiencies, especially in New Zealand where in the young landscape there are close relationships between soils and landscape features. As described above, the high-relief hilly and mountainous terrain is particularly well suited because spatial covariates for modelling are readily available (Schmidt and Hewitt, 2004). DSM in the flat lowlands is presently limited because existing digital terrain models (based on 20-m contours) are unable to be adequately applied as covariate layers. Remote sensing capability has not yet been successfully applied to develop alternative covariates. A challenge is the very high coverage of lush green pasture that masks the soil signal. Gamma radiometrics have potential but currently there is no mining industry able to provide the level of coverage enjoyed by other countries. Electromagnetic radiation is now being used for high-resolution paddock-scale mapping. This will provide useful windows of fine-scale variability within more generalised-scale maps, which can in turn be used to identify areas that have sufficient variability to warrant investment in precision agriculture. A major driver of such investment is water quality and water quantity, requiring high-resolution mapping for the operation of variable-rate irrigation. Our challenge is to incorporate this detailed mapping of small areas into S-map.

It is interesting to compare the direction of evolution of soil databases in the different countries. The development of S-map has paralleled that of ASRIS in Australia due to the need to maximise the use of good legacy data. Thus polygonal soil survey data still form the backbone of both the ASRIS and S-map information systems but both are moving to incorporate DSM techniques and raster data. Explicit descriptions of uncertainty are important to both. They have quite different approaches to scale due to the difference in size of the two countries. The lack of any need to integrate data from multiple states and organisations as in Australia, Europe, and the USA, has meant that New Zealand could put more focus on a national correlation effort. A good database of geo-referenced point data (soil pits and augur data) has enabled some countries to adopt more of a DSM approach. Drivers or issues such as carbon accounting, soil quality monitoring, land use change, and mining exploration can all provide information that can influence the evolution of each nation's soil information, as can the existence of a strong local pedometrics research capability.

6. Conclusion

Just as the advent of GIS technology allowed a huge step forward in the storage and provision of soil maps in the 1970s and 80s, recent advances in database, modelling and web technology now permit another significant step forward. This new technology is allowing the development of powerful soil information systems capable of the dynamic extension and provision of both spatial and aspatial soil information. Tools can be developed to automate some of the core data-entry, data-validation, inference engine and modelling, dissemination and documentation tasks that in the previous incarnation of New Zealand's soil information system were manual or one-off tasks. These tools form the basis of the S-map infrastructure, and are essential to the creation of a soil information system better able to serve modern end-users. This infrastructure will also be important in supporting the input of New Zealand soil information into the global soil mapping project.

Use of more objective taxonomic criteria with greater weight given to soil functionality than soil genesis has involved a substantial restructuring and renaming of New Zealand's soils. This revision of New Zealand soils and the new infrastructure supporting S-map will considerably enhance our ability to link to measured profile data, and to undertake DSM and other predictive modelling. Similarly, a stronger focus on recording information about soil variability and uncertainty also enhances modelling capability. Once complete coverage is achieved, S-map will be well placed to serve the varying needs of current and future end-users.

Acknowledgements

We are grateful for the soil, modelling and programming contributions to the S-map project of our colleagues, including Sam Carrick, Ian Lynn, James Barringer, Malcolm McLeod, Bavo de Pauw, Nick

Spencer and Aaron Wilton. Two anonymous reviewers, Malcolm McLeod and Pierre Roudier kindly reviewed the paper making many useful suggestions. This work was supported by the New Zealand Foundation for Research, Science and Technology, contract number C09X0704.

References

AAFC, 2010. The National Soil DataBase (NSDB). http://sis.agr.gc.ca/cansis/nsdb/intro.html. Accessed: 5 August.

Adams, J.A., Wilde, R.H., 1976. Variability within a soil mapping unit mapped at the soil type level in the Wanganui district. New Zealand Journal of Agricultural Research 19, 165–176.

ASRIS, 2011. Methods. http://www.asris.csiro.au/methods.html. Accessed: 17 Oct 2011.

Barringer, J.R.F., Hewitt, A.E., Lynn, I.H., Schmidt, J., 2008. National mapping of landform elements in support of S-Map, a New Zealand soils database. In: Zhou, Q., Lees, B., Tang, G. (Eds.), Advances in digital terrain analysis. Lecture notes in geoinformation and cartography, Springer, Berlin, pp. 443–458.

Bui, E., 2007. A review of digital soil mapping in Australia. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Developments in Soil Science, vol. 31. Elsevier B. V., Amsterdam, pp. 23–37.

Burrough, P.A., van Gaans, P.F.N., Hootsman, R., 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. Geoderma 77, 115–135.

Di, H.J., Kemp, R.A., 1989. Variation in soil physical properties between and within morphologically defined series taxonomic units. Australian Journal of Soil Research 27, 259–273.

FAOFAO, 2010. Natural Resources: soils and maps. http://www.fao.org/nr/land/soils/it/. Accessed: 4 Dec 2010.

GSM, 2010. *GlobalSoilMap.net* meetings in Rome May 2010. http://www.globalsoilmap.net/. Accessed: 5th August, 2010.

Hansen, S., Thorsen, M., Pebesma, E.J., Kleeschulte, S., Svendsen, H., 1999. Uncertainty in simulated nitrate leaching due to uncertainty in input data. A case study: Soil Use and Management, 15, pp. 167–175.

Hartemink, A.E., McBratney, A., Mendonça-Santos, M.L., 2008. Digital soil mapping with limited data. Springer.

Heuvelink, G.B.M., Burgers, S.L.G.E., Tiktak, A., Den Berg, F.V., 2010. Uncertainty and stochastic sensitivity analysis of the GeoPEARL pesticide leaching model. Geoderma 155 (3–4), 186–192.

Hewitt, A.E., 2010. New Zealand soil classification. Landcare Research science series no. 1. Manaaki Whenua Press, Lincoln, Canterbury, New Zealand. 136 pp.

Hewitt, A.E., Barringer, J.R.F., Forrester, G.J., McNeill, S.J., 2008. Soilscapes basis for digital soil mapping in New Zealand. In: Hartemink, A.E., McBratney, A., Botenger, J. (Eds.), Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. Third Global Workshop on Digital Soil Mapping (DSM 2008). Springer, Logan, Utah, USA.

Hewitt, A., et al., 2010. Aligning New Zealand digital soil mapping with the global soil mapping project. World Congress Soil Science, Brisbane, Australia.

Hong, S.Y., et al., 2010. Development of Soil Information System and Its Application in Korea. http://www.agnet.org/library/eb/612/eb612.pdf. Accessed: 17 Oct 201.

ITC, 2011. A compendium of on-line soil survey information: soil geographic databases. http://www.itc.nl/~rossiter/research/rsrch_ss_digital.html. Accessed: 18 Oct 2011.

Johnston, R.M., et al., 2003. ASRIS: the database. Australian Journal of Soil Research 41 (6), 1021–1036.

Jones, R.J.A., Houšková, B., Bullock, P., Montanarella, L., 2005. Soil Resources of Europe, second edition. Office for Official Publications of the European Communities, Luxembourg.

Karageorgis, D., 1980. Soil variability and related crop productivity within a sample area of the Templeton soil mapping unit. MSc Thesis, University of Canterbury, Christchurch, New Zealand.

King, D., Stengel, P., Jamagne, M., Le Bas, C., Arrouays, D., 2005. Soil mapping and soil monitoring: state of progress and use in France. Soil Resources of Europe In: Jones, R.J.A., Houšková, B., Bullock, P., Montanarella, L. (Eds.), European Soil Bureau Research Report No.9, EUR 20559 EN, second edition. Office for Official Publications of the European Communities, Luxembourg, p. 420.

Lagacherie, P., McBratney, A.B., 2007. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Developments in Soil Science, vol. 31. Elsevier B. V., Amsterdam.

Lagacherie, P., McBratney, A., Voltz, M. (Eds.), 2007. Digital soil mapping — an introductory perspective. : Developments in Soil Science. Elsevier, Amsterdam.

Landcare Research, 2003. Manual for National Soils Database. In: Wilde, R.H. (Ed.), Landcare Research New Zealand Ltd, Palmerston North.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer functions to soil inference systems. Geoderma 109 (1–2), 41–73.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.

McKenzie, N.J., Jacquier, D.W., Maschmedt, D.J., Griffin, E.A., Brough, D.M., 2005. ASRIS Technical Specifications. CSIRO.

McKenzie, N.J., Jacquier, D.W., Gregory, L.J., 2008. Online soil information systems — recent Australian experience. In: Hartemink, A.E., McBratney, A., Mendonça-Santos, M.L. (Eds.), Digital soil mapping with limited data. Springer.

Milne, J.D.G., Clayden, B., Singleton, P.L., Wilson, A.D., 1995. Soil description handbook. Manaaki Whenua Press, Lincoln, Canterbury, New Zealand.

Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57, 443–452.

Newsome, P.F.J., 1992. New Zealand land resource inventory Arc/Info data manual. Landcare Research.

NRCS, 2010a. Available soil survey data, Wednesday August 04, 2010. http://soildatamart.nrcs.usda.gov/SDMDB/SSURGSDM.pdf. Accessed.

NRCS, 2010b. Soil Survey Geographic (SSURGO) Database Description. http://soils.usda.gov/survey/geography/ssurgo/description.html. Accessed: 6 Aug, 2010.

NRCS, 2010c. SSURGO meta-data. http://soildatamart.nrcs.usda.gov/SSURGOMetadata.aspx. Accessed: 5 August, 2010.

Oldeman, L.R., van Engelen, V.W.P., 1993. A world soils and terrain digital database (SOTER) — an improved assessment of land resources. Geoderma 60 (1–4), 309–325.

Sanchez, P.A., et al., 2009. Digital soil map of the world. Science 325 (5941), 680–681 Washington.

Schmidt, J., Hewitt, A., 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. Geoderma 121, 243–256.

Taylor, N.H., Pohlen, I.J., 1970. Soil survey method: a New Zealand handbook for the field study of soils. NZ Soil Bureau Bulletin 25 242 pp.

Thompson, T.R.E., Bradley, R.I., Hollis, J.M., Bellamy, P.H., Dufour, M.J.D., 2005. Soil information and its application in the United Kingdom: an update. Soil Resources of Europe In: Jones, R.J.A., Houšková, B., Bullock, P., Montanarella, L. (Eds.), European Soil Bureau Research Report No.9, EUR 20559 EN, second edition. Office for Official Publications of the European Communities, Luxembourg, p. 420.

Webb, T.H., 2003. Identification of functional horizons to predict physical properties for soils from alluvium in Canterbury, New Zealand. Australian Journal of Soil Research 41, 1005–1019.

Webb, T.H., Lilburne, L.R., 2011. Criteria for defining the soil family and soil sibling: the fourth and fifth categories of the New Zealand Soil Classification Landcare Research Science Series no. 3. Manaaki Whenua Press, Lincoln, New Zealand.

Webb, T.H., Claydon, J.J., Harris, S.R., 2000. Quantifying variability of soil physical properties within map units to address modern land-use issues on the Canterbury Plains, New Zealand. Australian Journal of Soil Research 38, 1115–1129.

Wilde, R.H., Willoughby, E.J., Hewitt, A.E., 2000. Data manual for the national soils database spatial extension. Landcare Research.