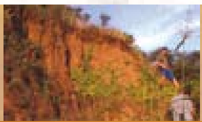
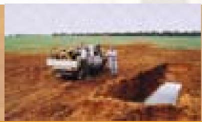
National Action Plan for Salinity and Water Quality

Soil and landscape attributes






Regional Natural Resource Management in Queensland

A report on the creation of a soil and landscape attribute information system for Queensland

July 2006

Queensland Government
action Salinity & Water AUSTRALIA

Soil and landscape attributes

A report on the creation of a soil and landscape attribute information system for Queensland

Project: SA03 Landscape Attributes for Salinity Processes
National Action Plan for Salinity and Water Quality

D.M. Brough, J. Claridge and M.J. Grundy

QNRM06186

The National Action Plan for Salinity and Water Quality (NAPSWQ) is a joint Australian and Queensland Government initiative that encourages governments and regional communities to work together to address salinity and water quality issues in priority catchments throughout Queensland. This document has been produced under the NAPSWQ using Australian and Queensland Government financial support.

The project was coordinated by:
Daniel Brough - Land Resources Officer
Land and Environment Assessment
Resource Condition and Trend
Natural Resource Sciences
Queensland Department of Natural Resources, Mines and Water
Email: Daniel.Brough@nrm.qld.gov.au
Phone: (07) 3896 9333

Acknowledgments

The authors would like to sincerely thank the many Departmental staff who contributed to and laid the foundation for this work over the last five years. Thanks also to David Jacquier, Brian Moore and Andrew Biggs for reviewing this report and for their valuable contributions. The following staff from Queensland Department of Natural Resources, Mines and Water, have made a significant contribution in various capacities over the duration of this project. Nathalie Baxter, Bruce Forster, Andrew Biggs, Tessa Chamberlain, John Loi, Peter Wilson, Ben Harms, Jon Burgess, Lauren Eyre, Ross Bigwood, Nev Christianos, Ian Webb and Ross Searle.

Citation: Brough, D.M., Claridge, J. and Grundy, M.J. (2006). *Soil and landscape attributes: A report on the creation of a soil and landscape information system for Queensland*. Natural Resources, Mines & Water, Brisbane, Australia. QNRM06186.

Contents

List of Tables

List of Figures

List of Maps

Summary

This report details the background information and technical methods used during the creation of soil and landscape attribute information for Queensland. The soil and landscape attribute information system is a combination of all existing data, information and knowledge on Queensland's landscapes, plus newly collected data from a strategic sampling scheme. A description of the current iteration of creating attribute surfaces for Queensland is given and a brief history of previous methods is also included. The methodology for the new surfaces is based on that for the Australian Soil Resource Information System, the emerging national standard for providing interpretations of soil and landscape attributes.

This project was funded by the National Action Plan for Salinity and Water Quality (NAPSWQ), a joint Australian and Queensland Government initiative that encourages governments and regional communities to work together to address salinity and water quality issues in priority catchments throughout Queensland. NAPSWQ program has provided a crucial impetus for the improvements to soil and landscape information by supporting the collection of new soil and regolith data in the NAPSWQ priority catchments across the state. NAPSWQ supported the collection of 1506 new soil and regolith sites, which includes some 110 sites sampled to depths of 3m or more. Most of the sites collected were strategically located to improve our knowledge of the variation in soil and landscape attributes in both environmental and geographic space.

The sampling methodology developed in this project was very effective in achieving a strategic sampling distribution to assist in the production of robust soil attribute surfaces. The analysis carried out in this study provides an understanding of current sampling and assists in devising new sampling programs.

Prior to the compilation of this, and previous versions of the soil and landscape attribute information system, the only complete state wide land resource information was continental scale soil mapping. This mapping was at a scale of 1:2 million, such a scale is inadequate for assessments at scales below the regional level. The version produced during this project comprises information at scales from 1:25,000 to 1:2 million, and was been achieved by combining all the appropriate land resource data for Queensland into a single 'Combined Soils' dataset.

By using much of the existing knowledge available from land resource data stored in the Natural Resources, Mines and Water Soil and Land Information database, the interpretation and estimation of attributes has been greatly enhanced. The ability to utilise expert knowledge has been important in producing surfaces that are reliable, updateable and accurate. The ability to process large amounts of SALI data has provided a system where the attribute values for a polygon can be area weighted to produce more accurate results. The use and storage of non-mapped information will allow users of the data to assess the variability inherent in the mapped units of the State's landscapes.

The soil and landscape attribute system will be an important and on-going resource for the modelling, assessment and interpretation of land resources for Queensland for many years to come, while its ability to be easily updated will ensure that the surfaces will reflect our best knowledge for that point in time.

1 Introduction

This report details the background information and technical methods used during the creation of soil and landscape attribute information for Queensland. The soil and landscape information is a combination of all existing data, information and knowledge on Queensland's landscapes, plus newly collected data from a strategic sampling scheme. The surfaces can contribute significant information into the conceptual understanding of various processes which occur across the State's landscapes. This report describes the current version of the process of generating soil and landscape information at spatial resolutions applicable for a variety of landscape assessment and modelling purposes. Prior approaches to developing reliable and consistent coverages of soil attributes have been described by Smith (2000) and Brough (2001, 2003). These previous versions have used varying processes to achieve similar outcomes, but the current process provides both more detail for the attributes in question and an improved spatial reliability through the use of repeatable processes. The current version is based on the standards defined for the Australian Soil Resource Information System (ASRIS[1]; McKenzie *et al.* 2005).

For Queensland, prior to 2000, the only complete state wide land resource information is from the Atlas of Australian Soils (Northcote *et al.* 1968). The soil and landscape attribute information system, and its predecessors, have improved on the information available from the Atlas. The Atlas was produced at a scale of 1:2 million and is not adequate for assessments at scales below the regional level. The version produced in this project comprises information at scales from 1:25,000 to 1:2 million, this has been achieved by combining all the appropriate land resource data for Queensland into a single 'Combined Soils' dataset.

The National Action Plan for Salinity and Water Quality (NAPSWQ) is a joint Australian and Queensland Government initiative that encourages governments and regional communities to work together to address salinity and water quality issues in priority catchments throughout Queensland. NAPSWQ consists of a number of component projects developed to address specific areas of concern. The landscape attributes for salinity processes project (Project Sa03) focuses on developing spatially reliable datasets of landscape processes. These datasets are used for the improved understanding and modelling of salinity processes in Queensland. The main aims of the project were to:

1) provide an understanding of the spatial distribution of key soil and land attributes,
2) build the capacity to predict that distribution using spatial environmental predictors,
3) build models of spatial processes linked with land management,
4) link with models and modellers at the point and regional levels and
5) develop information packages and strategies to communicate and use this information to inform land management.

The National Action Plan in Queensland focuses on the priority regions of the Burdekin, Fitzroy, Condamine-Balonne, Maranoa, Border Rivers-Moonie, Burnett, Mary, Lockyer-Bremer-Upper Brisbane Catchments. These catchments represent about 30% of the State or 470 000km^2 (Figure 1).

[1] www.asris.csiro.au

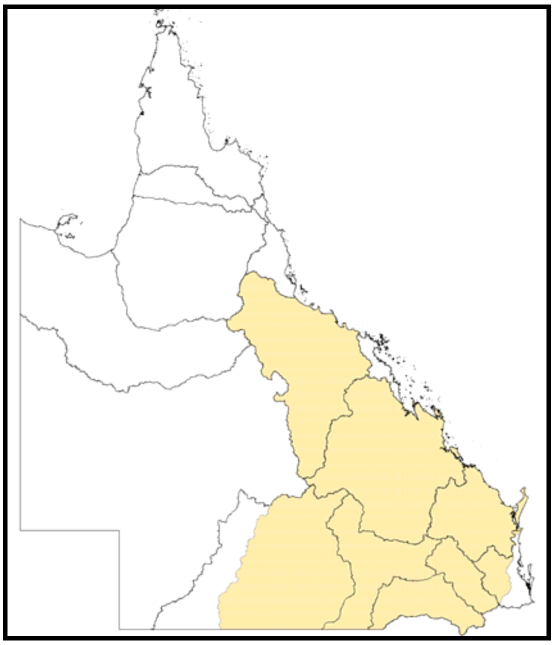


Figure 1. Queensland NAPSWQ priority areas.

The process of creating soil and landscape information as described in this report has utilised several opportunities to significantly improve the reliability, repeatability and updateability of these important datasets. The improvements fall into three broad categories;

1) The million dollar redevelopment by Natural Resources, Mines and Water (NRMW) of the Soil and Land Information (SALI) database system,
2) The development of ASRIS to provide a national framework for the interpretation of soil and landscape attributes and
3) The NAPSWQ Sa03 project, which provided $950,000 of its budgeted $1.76 million to collate and collect land resource data and build the soil and landscape attribute information system

The redevelopment of SALI to provided a single point-of-truth for soil and land information within NRMW and has presented a consistent framework and logical database design in a centralised system for the storage, management and use of SALI data. The consistent framework and centralisation of SALI has enabled the soil attribute information to be generated with ease and allowed for the design of a system to produce the information that is both repeatable and easily updateable. The implementation of SALI in Oracle (2006) and ESRI (2006) products has facilitated the use of industry standard tools and techniques to be used in the creation of the soil and landscape attribute information system.

The development of the ASRIS methodology has enable efficiencies in the production of the information to suit a wide variety of purposes. The ASRIS methodology, which contains recommendations based on Queensland experiences, is a nationally recognised standard and has provided a peer reviewed methodology. Previous examples of deriving soil and landscape information for Queensland have been important precursors to the current iterations of both the Queensland and ASRIS frameworks.

The NAPSWQ Sa03 project has provided the crucial component for the improvements to the soil and landscape attribute information system by supporting the collection of new soil and regolith data in the NAPSWQ priority catchments across the state. The Sa03 project supported the collection of 1506 new soil and regolith sites, which includes some 110 sites sampled to depths of 3m or more. The Sa03 project used two sampling schemes, the first involved the strategic collection of new site information, while the second built a set of reference sites. Most of the sites collected during Sa03 were selected based on the first sampling scheme These sites were strategically located to improve our knowledge of the variation in soil and landscape attributes in both environmental and geographic space. The purposive sampling scheme was designed with the intention of providing a richer set of sites to improve the spatial reliability and accuracy of the soil and landscape attribute surfaces.

The second sampling scheme utilised in Sa03 was used to build a set of Key Reference Sites where soil and landscape processes are known in detail and where the sites have been well characterised over time. Details of the Key Reference Sites and the reasons for selecting each one can be found in Harms and Main (2006).

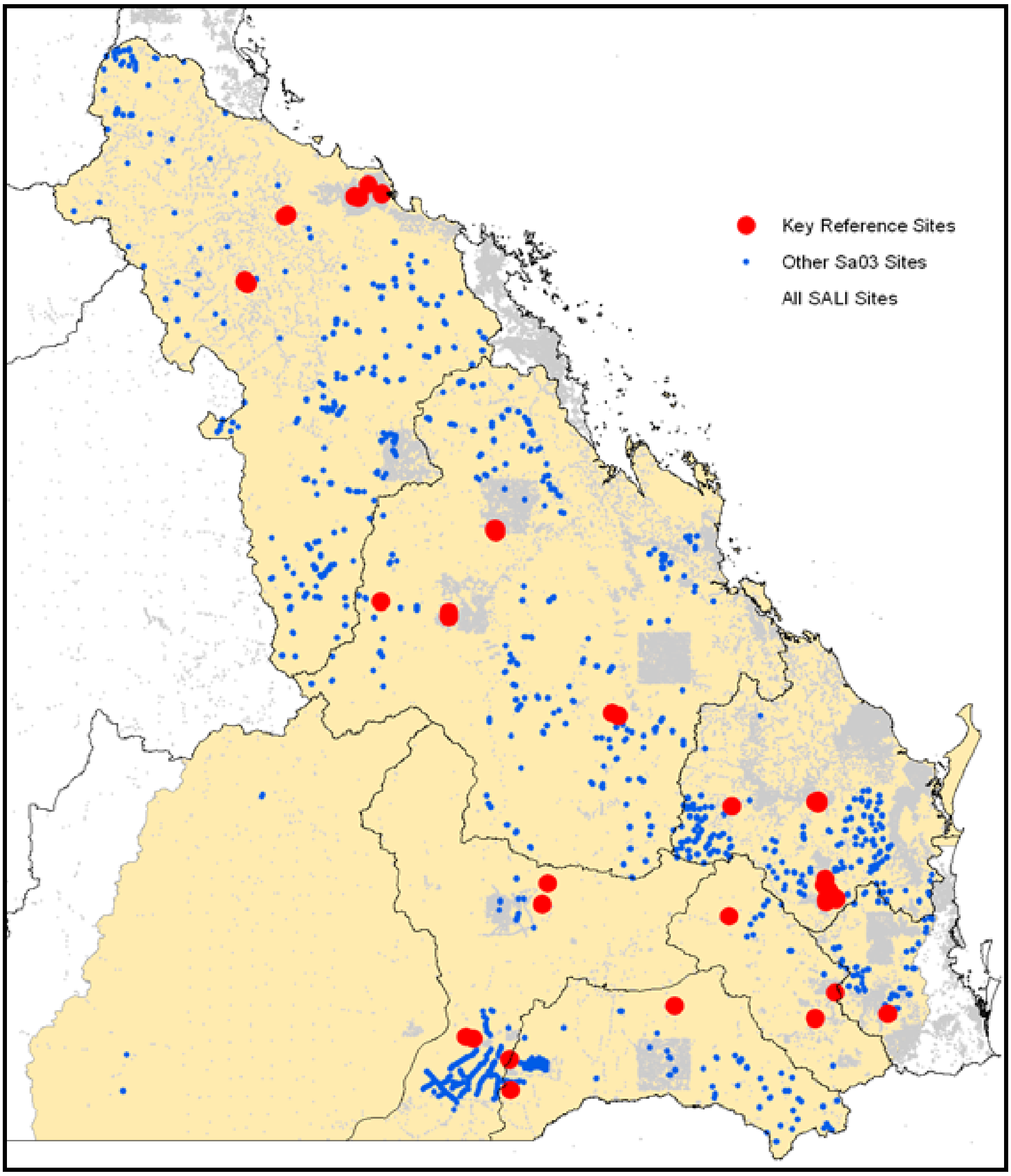


Figure 2. Soil sites collected during the Sa03 project in relation to all existing SALI sites.

2 NAPSWQ landscape attribute project

2.1 Project background

This project was designed to meet the overall aims of the NAPSWQ and broader NRMW policies and priorities. It was also designed to build the capacity of community and landholder groups to recognise land and water resources systems locally. The project aimed to assist in understanding the role that current and potential land management systems play and to integrate this understanding into plans for improved natural resource management and stewardship. There are a number of information requirements which are needed to underpin this capacity. These include those which allow a direct understanding of landscapes and their function and those which feed into a modelled understanding of salinity hazard and risk.

2.2 Project objective

The objective of the Sa03 project is the development of spatially reliable datasets of landscape processes for the improved understanding and modelling of salinity in Queensland. Landscape attribute assessment provides an understanding of the spatial distribution of key soil and land attributes in several key areas

- it builds the capacity to predict that distribution using spatial environmental predictors,
- builds models of spatial processes linked with land management,
- links with models and modellers at the point and regional levels, and
- develops information packages / strategies to communicate

The information generated from landscape attribute assessment is used to inform land management decisions and broader natural resource management issues.

2.3 Project outcomes

The Sa03 project has produced information sets which are spatially reliable. The information can be readily used by a large number of potential uses within NAPSWQ activities and which complement the analytical, modelling and extension projects within the NAPSWQ.

The products and information sets developed by Sa03 are digital elevation models, land use mapping, regolith information and mapping, the description of a group of Key Reference Sites and the soil and landscape attribute surfaces.

The component reports detailing the results from the Sa03 project, are:

1) Modelling the Land Surface - Smith and Brough (2006),
2) Mapping Land Use in Queensland – Witte *et al.* (2006),
3) Regolith and Physiography in Queensland – Chamberlain (2006),
4) Key Reference Sites for Queensland- Harms and Main (2006), and
5) Soil and Landscape Attributes – Brough *et al.* (2006) – this report

3 Site selection strategy

This section was previously published as Claridge and Grundy (2003) for the Burdekin Catchment. The same process was used for all NAPSWQ catchments.

The site selection strategy presented here uses an expert derived pedogeomorphic model for the division of the environment into physiographic domains. The degree to which these domains have been characterised (by morphological and analytical soil sites) can then be assessed. It has been argued that intuitive regionalisations are inflexible and are of limited value because the classification procedure is not explicit and repeatable. This is not always strictly true. Neldner *et al.* (1995) investigated five methods of assessing sampling adequacy and found that all require subjective decisions and rely on the skill and knowledge of the ecologist.

3.1 Background and reasoning

This section contains information on the need for unbiased site data and the move towards soil properties rather than soil mapping. It also discusses some of the issues when considering developing an effective sampling strategy more aligned with modelling than mapping.

From polygon information to site data

The conventional polygon mapping of soil is rarely tested for accuracy. Soil survey organisations have considered that polygons at the ‘soil series’ level (maps with a spatial resolution of 1:50,000 map scale or larger) have a ‘purity’ of 70 – 80%, which is generally regarded as acceptable. However there have been many studies that have shown this is too optimistic; with both soil classes and mapped soil units being extremely variable, purity is often much less than 70% (Burrough *et al.* 1997). In a study by Beckett and Webster (1971), it was found that only 50% of the randomly chosen sites matched the mapped soil series. A Netherlands study quoted by Burrough (1993) on the quality of conventional soil survey found that mapping units tested for purity by independent sampling, ranged from 64 – 70% for 1:50,000 maps, 59 – 62% for 10,000 maps and 7–12% for the soil classification units.

Soil properties and spatial prediction

Conventional soil maps have some major limitations when soil information is used for environmental modelling including low spatial resolution and uniform attribute value within units (Zhu *et al.* 1997). The mapping of soils using choropleth maps with sharply delineated and homogenous polygons being used to represent the spatial distribution of soils has been shown to be impractical and unscientific (Zhu *et al.* 1997). In reality, changes between soils often occur gradually with a diffuse boundary rather than sharp; and soil attributes often vary considerably within mapping units.

The sampling biases that we inherit

The field phase is one of the more subjective and less explicit stages in land resource surveys (McKenzie, Austin 1993; Slater, Grundy 1999; McKenzie *et al.* 2000a). Most broad scale land resource surveys use observations that are irregularly located according to the surveyor's judgement ('Free survey technique'; Gunn *et al. 1988*) for locating and sampling sites. Free surveys tend to place emphasis on checking boundary locations and the validation of units delineated on the basis of air photo patterns and leads to an unknown level of bias in the survey coverage.

Better sampling strategies

There are explicit survey methods currently being used that randomly locate sites within stratified units so as to minimise site bias (e.g. McKenzie, Ryan 1999, O'Connell *et al.* 2000 for forested landscapes). Stratified sampling of a survey area is more effective than random sampling because the varying degree of spatial dependence in soil attributes reduces the efficiency of random sampling (Gessler *et al.* 1995). However, stratified sampling using information about the spatial dependence structure (e.g. geology), increases the information obtained from sample sites. Disadvantages of stratified sampling include the uneven distribution of samples in geographic space (McKenzie, Austin 1993), it may be impracticable and loss of information may occur if an inappropriate stratification was used. Models developed from data collected using a purposive stratified sampling approach only test the stratified environment and not the actual soil-landscape.

Borrowing sampling principals from vegetation surveys

Survey data should cover the full range of environmental space and be sampled proportionally (Margules, Stein 1989) or biased towards the less common types (Austin, Heyligers 1989). The number of samples should be as large as possible within resource constraints (Margules, Stein 1989), with replicates covering the geographic range of the biota studied (Nicholls 1989). From Neldner *et al.* 1995, p2 "These principals probably also hold true for soil sampling with the exception of biasing sampling towards less common soils." Soils range along a continuum and therefore it is usually not as important to capture the extremes as it is to sample rare vegetation types.

The issue of scale

It is generally accepted by pedologists that the pedogenic processes controlling soil distribution are multi-scaled (McSweeney *et al.* 1994; Ryan *et al.* 2000). Investigations of soil–landscapes show that soil-forming relationships are rarely linear across scales (Thwaites, Slater 2000). Therefore optimally the scale for the environmental variables should be comparable to the scale of processes controlling soil formation. In reality it is impossible to accurately measure all variables at the scale of soil forming processes (McKenzie, Ryan 1999). Therefore the choice and scale of the variables used to create the physiographic domains can be controlled by data availability and expert knowledge.

Challenges associated with large areas and low sampling intensity

There are a number of unique challenges that exist for sampling of large areas including:

1) accurate characterisation of the area to devise an effective sampling strategy;
2) the range of soil variation is significantly larger than what can be sampled;
3) difficulties in capturing the critical components of the landscape and in the right proportions,
4) with large areas there are access (often a limited number of roads), and
5) time and other logistical constraints.

These challenges are more formidable when sampling for modelling purposes compared to mapping soil types.

3.2 Assessing sampling adequacy

A sampling distribution can be evaluated by assessing how well it captures either geographic or environmental space. For the development of accurate soil attribute surfaces a sampling distribution must effectively sample both. Additionally, there are at least two broad groups of sites: those that have only morphological data and those that also contain chemical data and are therefore of higher intrinsic value to the process. In order to assess the sampling of environmental space a stratification of the pedogeomorphic environment was developed.

The stratification developed here is an expert conceptual model of the pedogeomorphic environment. It attempts to capture the range of different soil forming environments and the factors governing the distribution of soils so that the developed site network is representative of the environment. A pedo-lithological framework (pedolith) has been developed for this purpose and an example for the Burdekin Catchment can be seen in Table 1. The application of the pedolith framework in the Burdekin catchment used soils mapping as the base layer. The best soils mapping at any location has been combined into the one coverage for the whole of Queensland and has been documented in Brough (2001). The stratification process involved assigning map codes from the best soils coverage to a lithology (Map 1). The decision to assign a map code to a particular class was based on information from the land resource survey map / report and where possible expert knowledge from a relevant person. The advantage of using soils mapping rather than rock geology is that soil distribution should be more accurately characterised in both a conceptual and spatial sense.

There were a few difficulties encountered during this process due to the scale and type of mapping across the catchment. Some of the mapping was very coarse (1:1,000,000) and mapcodes were only a general soil type and therefore were often not consistent with a single geology. This made assigning a mapcode to a pedolith class very difficult in some cases. Edge matching of the different soils coverages was not always seamless due to the different scales of land resource mapping and sometimes due to mapping type. For example a seamless join was not always possible between soils and land systems mapping.

This broad pedolithological definition of the environment was further subdivided according to terrain and a description of this process can be seen in Section 3.4. The splits according to terrain separate the broad soil groups (pedolith classes) into more discrete physiographic domains occurring within these areas.

3.2.1 Sampling of geographic space

Distance and density can be used to describe how well the geographic space has been sampled. Three methods have been used to approximate sampling of geographic space by soil site data:

1) Distance to nearest site,
2) Site density on a fixed radius using a kernel approach and
3) Sampling intensity required for different mapping intensities

The sampling intensity for mapping used a neighbourhood function with an expanding radius until at least one site was found and divided by the search area. A look up table was then used to remap the resultant grid according to Gunn *et al.* (1988). These sampling intensities of geographic space are shown in Maps 2, 3 and 4.

Table 1. Pedo-lithological classes of the Burdekin Catchment.

Classification 1	Subgroup 1	Description	Lithcode	Pedolith number
Acid igneous	Granites (adamellites)		Gn	1
	Granodiorites (tonalities)		Gd	2
	Acid pyroclastics and lavas		Py	3
Intermediate igneous	Andesites, diorites, and associated pyroclastics		An	4
Mafic igneous	Basalts, gabbros	Includes dolerites, amphibolites	Ba	5
Ultramafics	Serpentinite, ultramafic intrusives and extrusives		Sr	6
Sediments (includes low to medium grade metamorphics)	Highly siliceous	Cherts, silcretes, quartzites	Ch	7
	Fine grained sediments	Shales, mudstones, siltstones	Shl	8
	Coarse sediments - quartz rich (sublabile)	Sandstones, greywacke, conglomerate,	Sdq	9
	Coarse sediments - lithic/ feldspathic (labile)		Sdl	10
	Calcareous		Li	11
High grade metamorphics	Highly metamorphosed sequences / gneiss / migmatite (mainly Pre-Cambrian)	Excludes quartzites (Sdq) and amphibolites (Ba)	Mg	12
Residual/ weathered surfaces	Quartzose weathered sediments		Rq	13
	Soft weathered sediments		Rs	14
	Duricrusts	Silcrete, laterite	Rd	15
Unconsolidated	Estuarine	Dominantly saline	Ul	16
	Sands (beach, aeolian)		Us	17
	Lacustrine / lagoonal		Ue	18
	Alluvium / colluvium	Recent	Ur	19
		Old	Uo	20
		Fans	Uf	21

3.2.2 Sampling of environmental space

The simplest way to assess sampling of environmental space is to calculate the area represented by each site within each pedolith class. The further subdivision of each pedolith class into different pedogeomorphic environments using some terrain derivative like slope or compound topographic index was not done in the main analysis due to the low density of the existing sampling distribution. The sampling of these individual physiographic domains was then assessed by comparing the sampling intensity of these areas and is presented in Section 3.3.

The sampling of individual physiographic domains was then to be assessed by looking at the sampling intensity of these areas. This method of creating the physiographic domains was considered inappropriate due to large areas of the catchment being under sampled and the large geographic extent of the catchment. The large extent of the catchment meant that a pedolith class could vary significantly from one end of the catchment to the other (over 700 km for the Burdekin Catchment). For example the basalt flows in the north of the Burdekin

Catchment are of a different age and geochemistry compared to those in the south. The characteristics of alluvium also vary significantly depending on the source material.

3.2.3 Other considerations not incorporated into the analysis

Generally the larger the area the less intensely it needs to be sampled to capture the soil variability. To account for this fact a sampling intensity index (area/site) could be used to adjust the required sampling depending on the area (a logarithmic type function). Some pedolith groups have inherently more variability and will therefore need to be sampled at a relatively higher intensity to account for this variation. To account for the variation between pedolith groups each group could be assigned a weighting factor to adjust the number of samples required. It is also possible to bias the sampling based on an area's importance and where more intensive information is needed. For salinity some pedolith groups will have a higher salinity risk and special attention may be given to these areas. In vegetation surveys it is important to sample heavily fractured areas more intensely than vegetation communities that occur in large contiguous areas. This principal is probably not as important in soils work.

3.2.4 Assessment of geographic and environmental space

To overcome the variation that may occur in a pedolith class due to the large spatial extent a combined approach was used to define the sampling priorities. A search radius was used to calculate sampling intensity within pedolith classes, thus incorporating geographic space and limiting the variability of pedolith classes across the catchment. The new sampling intensity index used a moving window to calculate the number of sites in the same pedolith class using a fixed search radius divided by the area of that pedolith class within the search area. Two different radii were used, 10 km (Map 5) and 50 km. The grey areas are where there were no sites within 10 km in the same pedolith class. This approach captures the concept of both geographic and environmental space.

The final product of the combined assessment was a map (Map 6) that identified areas of five different existing sampling densities and can be seen in Appendix 1. The sampling densities were:

1) No sites in the same pedolith class within 50 km,
2) No sites in the same pedolith class within 10 km,
3) Less than 1 site/100 km^2 in the same pedolith class within a search radius of 50 km ~ 1:500,000,
4) Less than 1 site/25km^2 in the same pedolith class within a search radius of 50 km ~ 1:250,000 and
5) Sampled better than 1:250,000

These densities then led to the development of a sampling strategy aimed at achieving an approximate new density better than 1:500,000 or more than 1 site/100 km^2.

3.3 Assessment of sampling biases

3.3.1 Terrain biases

The existing sites network has not sampled the range of soil forming environments equally and this was not the intention for the original mapping purposes. In order to assess where new samples should be located to capture the environmental space in a more equitable manner terrain variables were used to subdivide each pedolith group into physiographic domains.

Within each pedolith class the terrain variables slope, Compound Topographic Index (CTI) and the Multiresolution Valley Bottom Flatness Index (MRVBF; Gallant, Dowling 2003) were divided into four equal area categories. The sites in each of the equal area terrain categories were counted. The results for each terrain derivative for the Burdekin Catchment are shown in Table 2, Table 3 and Table 4.

In most pedolith classes there was a strong bias toward flatter areas, in the Burdekin Catchment 40% of sites were in the flattest slope class, and in accumulation areas with high CTI values. For MRVBF there was an emphasis towards larger flat areas except for alluvial lithologies where the smaller flats are sampled more intensely. This clearly reflects the agricultural and edaphic considerations, which drove the original projects.

3.3.2 Pedolith biases

A simple analysis of sites and area of the different pedolith classes across the whole of the Burdekin catchment is given in Table 5. The sampling priorities of the pedolith classes across the entire catchment assessed by sampling intensity would be 1) acid pyroclastics and lavas, 2) sublabile sediments, 3) deeply weathered sediments, 4) labile sediments, 5) metamorphics, 6) granites (adamellites), 7) granodiorites (tonalites), 8) fine grained sediments, 9) old alluvial/colluvial material and 10) fans/reworked materials.

Existing sites are distributed unevenly in pedolith groups. Pedolith classes that are well sampled may occur only in well surveyed areas, while those poorly sampled may occur primarily in low intensity survey areas. This is certainly the case for the acid pyroclastics and lavas, which primarily occur in the Townsville sheet.

Examining just the analysed sites the sampling intensity of the pedolith classes can be seen in Table 5. This table indicates that acid pyroclastics and lavas have a paucity of chemical data. It also shows that recent alluvial areas have been very well sampled. Because there are large areas of the catchment that have not been sampled it is most useful to look at these values in absolute terms. It should also only be used as a guide because as the area of lithology increases a smaller number of new sites are required to accurately characterise it. Therefore pedolith classes with large extents have an area/site ratio that is larger than a balanced sampling method would indicate.

Table 2. The number of sites in slope categories using an equal area four class split within each pedolith class for the Burdekin Catchment.

Pedolith class	Slope categories				Total
	1	2	3	4	
Granites (adamellites)	115	72	50	14	251
Granodiorites (tonalites)	81	78	78	15	252
Acid pyroclastics and lavas	23	11	6	1	41
Andesites, diorites	1354	110	39	14	1517
Basalts, gabbros	342	379	108	86	915
Highly siliceous		2			2
Fine grained sediments	49	48	39	26	162
Coarse sediments –quartz rich	142	72	74	39	327
Coarse sediments –feldspathic	44	34	20	11	109
Calcareous	64	59	40	8	171
Highly metamorphosed	78	54	32	20	184
Soft weathered sediments	239	123	114	120	596
Duricrusts		2		1	3
Estuarine	41		1	3	45
Sands (beach, aeolian)	1				1
Alluvium / colluvium - recent	5974	1770	1886	1340	10970
Alluvium / colluvium – old	263	172	154	89	678
Alluvium / colluvium – fans	46	54	44	25	169
Total	6325	1996	2084	1454	16393

Table 3. The number of sites in CTI categories using an equal area four class split within each pedolith class for the Burdekin Catchment.

Pedolith class	CTI categories				Total
	1	2	3	4	
Granites (adamellites)	27	43	71	110	251
Granodiorites (tonalites)	19	86	73	74	252
Acid pyroclastics and lavas	5	7	11	18	41
Andesites, diorites	17	49	137	1314	1517
Basalts, gabbros	110	188	312	305	915
Highly siliceous		1	1		2
Fine grained sediments	29	40	49	44	162
Coarse sediments –quartz rich	41	89	91	106	327
Coarse sediments –feldspathic	14	17	42	36	109
Calcareous	10	44	62	55	171
Highly metamorphosed	22	32	65	65	184
Soft weathered sediments	143	139	151	163	596
Duricrusts		1	1	1	3
Estuarine	5	12	16	12	45
Sands (beach, aeolian)				1	1
Alluvium / colluvium - recent	1474	3051	3067	3378	10970
Alluvium / colluvium – old	104	210	195	169	678
Alluvium / colluvium – fans	33	50	44	42	169
Total	1611	3311	3306	4072	16393

Table 4. The number of sites in MRVBF categories using an equal area four class split within each pedolith class for the Burdekin Catchment.

Pedolith class	*MRVBF categories*				*Total*
	1	*2*	*3*	*4*	
Granites (adamellites)	*36*	*46*	*52*	*117*	*251*
Granodiorites (tonalites)	*53*	*63*	*61*	*75*	*252*
Acid pyroclastics and lavas	*17*		*9*	*15*	*41*
Andesites, diorites	*152*		*321*	*1044*	*1517*
Basalts, gabbros	*184*	*144*	*202*	*385*	*915*
Highly siliceous		*1*		*1*	*2*
Fine grained sediments	*33*	*40*	*40*	*49*	*162*
Coarse sediments –quartz rich	*50*	*82*	*97*	*98*	*327*
Coarse sediments –feldspathic	*17*	*20*	*33*	*39*	*109*
Calcareous	*28*	*43*	*42*	*58*	*171*
Highly metamorphosed	*29*	*39*	*50*	*66*	*184*
Soft weathered sediments	*148*	*117*	*138*	*193*	*596*
Duricrusts	*1*	*2*			*3*
Estuarine	*31*	*13*	*1*		*45*
Sands (beach, aeolian)		*1*			*1*
Alluvium / colluvium - recent	*2475*	*2503*	*5896*	*96*	*10970*
Alluvium / colluvium – old	*159*	*265*	*199*	*55*	*678*
Alluvium / colluvium – fans	*42*	*36*	*45*	*46*	*169*
Total	*2676*	*3306*	*6140*	*197*	*16393*

Table 5. Number of sites within each pedolith class for the Burdekin Catchment.

Pedolith class	Area (km^2)	Number of analysed sites	Analytical sites (km^2/site)	Number of morphological sites	Morphological sites (km^2/site)
Acid pyroclastics and lavas	5715	9	635	41	139.4
Coarse sediments –quartz rich	19670	66	298	327	60.2
Soft weathered sediments	23987	103	232	596	40.2
Coarse sediments –feldspathic	3565	19	187	109	32.7
Highly metamorphosed	5447	51	106	184	29.6
Granites (adamellites)	7015	49	143	251	27.9
Granodiorites (tonalites)	6357	55	115	252	25.2
Sands (beach, aeolian)	25	1	24	1	25.0
Fine grained sediments	3196	12	266	162	19.7
Alluvium / colluvium – old	12958	58	223	678	19.1
Alluvium / colluvium – fans	3137	19	165	169	18.6
Estuarine	767	3	255	45	17.0
Basalts, gabbros	13984	89	157	915	15.3
Andesites, diorites	5999	64	93	1517	4.0
Calcareous	291	5	58	171	1.7
Alluvium / colluvium - recent	16662	559	29	10970	1.5
Duricrusts	1.07			3	0.4
Highly siliceous	0.5			2	0.3

3.4 Development of a supplementary sampling program

The original objective was to develop a stratified random sampling strategy across the whole catchment using detailed information on landscape position and parent material. This objective was found to be inappropriate, due to the extensive areas that had not been sampled, and was subsequently modified. It was decided to focus on these large gaps for further sampling rather than embarking on a supplementary sampling program across the whole area. On examination of the analysis results discrete areas (sampling gaps) were identified and then assigned an importance. From the importance and size of the area, times for field work were allocated for sampling teams after considering the budgetary and logistical constraints.

For the initial field work, sites were randomly located based on the stratification of the area according to geology and topographic position. The actual collection of these sites in the field was found to have a number technical and logistical problems due to the large areas, low sampling intensities and remoteness of the areas.

To overcome some of the problems associated with low intensity sampling of large areas the classical stratified random sampling procedure was modified so that ‘new sites’ were called sampling localities and the environment was only stratified by parent material. The term sampling localities were used because the location just provides a starting point for the sampling of a toposequence. The flexibility of a sampling location means that access restrictions and soil variation can influence actual site positions. The sampling of toposequences captures the catenary/geomorphic environment and allows the incorporation of expert knowledge into the sampling strategy. It is also easier for the surveyor to interpret the landscape compared to using a classical stratified random sampling pattern, which would result in relatively isolated site observations. Because we are only locating sites at a very low sampling intensity it is probably more important to capture representative areas within our stratification or more to the point to avoid strange or localized features. The use of sampling localities makes use of expert knowledge for this purpose.

This method of sampling has been very effective in achieving the desired outcomes. The use of the analysis carried out in this study provides an understanding of the strengths and weaknesses of the current sampling.

4 Soil and landscape attribute surfaces

The method of producing soil and landscape attribute surfaces presented here uses a data driven approach to soil modelling. The processes utilises all of the suitable land resource information in SALI to derive surfaces that provide the best available estimate of an attribute at the best available scale across Queensland. By using all the suitable information from SALI the 50 year catalogue of land resource information has been used. This has relied upon the significant investments made by the Queensland Government and CSIRO.

4.1 Historical approaches and recent developments

Over the past several years there has been an increasing need for reliable and consistent datasets of soil and landscape attributes to support a range of high priority tasks across Queensland. The increase in modelling and the need for broadscale interpretations has led to the development of a series of datasets where all of the available land resource information is combined into a single dataset. The various programs to develop these 'combined' coverages have developed over time to the current incarnation through a series of increasingly complicated methods, which have produced more information and allowed for more detailed interpretations to be developed.

4.1.1 Background and requirements

The objective of the Sa03 project was to develop information on soil and landscape attributes of importance to salinity processes and water quality issues. Sa03 was the only project in the range of NAPSWQ Strategic Investment Projects to produce landscape attribute information, thereby increasing the significance and range of attributes that were developed. The attribute surfaces that have been produced are applicable to a wide variety of landscape assessment tasks as they form a primary data theme for many landscape assessment endeavours. Examples of previous uses:

- Salinity Hazard and Risk Assessments
- Modelling sediment loads and quality of waters flowing onto the Great Barrier Reef
- Soil Erosion and Soil Condition Hazard Assessments for Reef Catchments
- Bioregion mapping
- Modelling pre-clearing vegetation extent

The importance of making regionally consistent interpretations of soil and landscape attributes has increased in the recent past because of programs such as NAPSWQ, the Natural Heritage Trust, the Great Barrier Reef Protection Plan and the Vegetation Management Act. Future proposed programs in the Queensland Government such as the Queensland Landscape Monitoring System and the Rural Leasehold Lands Renewal Strategy will also rely heavily on the soil and landscape attributes produced from this project.

Definition of land resource data types used

Queensland has a wealth of land resource information that has been collected over the last half century. Much of the effort has been provided by the Queensland Government, through the Department of Natural Resources, Mines and Water (and its predecessors). CSIRO provided significant investment in the early broadscale assessments of the States land resources. The information collected over the years falls into two broad categories, observed and interpreted data. The interpreted data can be further categorised into polygonal mapping, soil profile classes (SPC) and land systems information. SPC and land system information are taxonomic classes and conceptual units, they assist in the definition and communication of the landscape. Observed data is collected through the site concept. The following paragraphs contain brief descriptions of the different types of land resource data available for Queensland. See Gunn *et al.* (1988), McDonald *et al.* (1990) and Christian and Stewart (1953, 1968) for detailed definitions of the land resource data types.

Project data

The manner in which land resource assessment is undertaken in Queensland is through projects. A project maybe specifically funded or for a specific purpose. The project record in SALI (Biggs *et al.* 2000) contains high level project information that describes the project in sufficient detail that insight into the purpose and reasoning for a project maybe discovered. SALI stores information on when, where and why a project was undertaken. Other information lists the officers involved in the project and the standards to which data is collected. For the soil attribute system the key information is the status of the project, the type of survey and the scale of the project.

Site data

A site is a small area of land (approximately 30m in diameter) considered to be representative of the landform, vegetation, land surface and other land features associated with the soil observation (McDonald *et al.* 1990). Sites have observations, which describe many of the features of a site that are typically transient or part of dynamic systems, such as vegetation and the soil profile. Both these features of an observation may be altered by human influence or natural process such as ecological succession. Multiple observations occur when a site is sampled multiple times, for example during annual monitoring, or where different observations are taken at the same time, as in the case of the Key Reference Sites. In Queensland, the terms site and observation, from the data perspective, are typically interchangeable as 99.8% of sites in SALI have only one observation[1].

Polygon data

The two types of polygon mapping used in Queensland are map reference or mapcode mapping and the unique map area (UMA) methods. Mapcode mapping lists each unit in the reference and each feature on the map is given that same description. The UMA approach defines each polygon on the map as being unique, even though similar units contain the same soils they can occur in different proportions. In UMA mapping the concept of un-mapped components or entities exists. Entities record detailed information for the different component soils within a polygon. In SALI each polygon has at least one entity, for mapcode mapping only one entity can be recorded with the pertinent details of the reference map unit.

Soil Profile Class data

A Soil Profile Class (SPC) is a group of similar soil profiles, defined at any level of generalisation, that form a local taxonomic classification of the soils found within an area or region (Isbell 1988). An SPC has a defined range of values for an attribute, the variation of the attribute within each class typically increases as more generalisation is applied to the class or more soil individuals are included. The variation within a class is generally less than

[1] As at May 2006.

the variation between classes. A SPC is used to define the typical soil description for a named taxonomic class and is an important communication tool for the land resource assessment community. An SPC may have variants or phases and may also be correlated to related classes or taxonomic groupings of soil profiles.

Land system data

The concept of a land system as defined by Christian and Stewart (1953, 1968) are those used in Queensland rather than those of Speight (1988). A land system as defined by Christian and Stewart is "*an area or groups of areas throughout which there is a recurring pattern of topography, soils and vegetation*". A land system is an amalgamation of one or more land units which are the individual components of the land systems. A unit may occur in several systems with the relative area of the unit varying between systems. Land units are relatively uniform in terms of landform, geology, soils and vegetation. In Queensland, two types of land system mapping have been used. The first is mapcode mapping of land systems where the same land system and land unit concepts are applied to all polygons of that particular type. The second type of mapping is similar to the UMA approach where individual land units are recorded for the mapped land systems and a land unit may occur in several land systems.

4.1.2 History of previous methods

In Queensland there is a history of providing interpretations of land resource data for the broadscale assessment of the State's natural resources. A variety of different methods have been used to produce interpreted products over the years ranging from basic interpretations through increasingly complicated methods. The increase in computing power and development of integrated information systems, such as SALI, has promoted the development of quantitative approaches to producing soil and landscape attribute surfaces. This section provides a history of the development of interpreted soil and landscape information for Queensland.

Previous broadscale interpretations

Until the year 2000, state and regional assessments of landscape attributes relied on very broadscale mapping, such as the Atlas of Australian Soils (Atlas; Northcote *et al.* 1968). While the Atlas was and still is a significant dataset for Queensland, the 1:2 million scale of capture is its main limitation. Previous assessments, such as the map "pH of Surface Soils, Queensland" (Ahern *et al.* 1992), rely on expert interpretations of the broadscale mapping to produce statewide information. McKenzie *et al.* (2000b) developed methods for data driven interpretations of soil types from the Factual Key (Northcote 1979) to be applied to Atlas mapping units. McKenzie *et al.* (2000b) lists several caveats on the use of these interpretations. The major points from the caveats are that reconnaissance scale soil-landscape maps usually have a low predictive capability and very large variation within each map unit. While not discounting the value and importance of these previous assessments, it is clearly evident from the aforementioned caveats the results from these projects are coarse estimates of the attributes and their variability[1].

[1] The authors of this report would like to note that the Atlas is the best available mapping across parts of Queensland. This fact should not escape the attention of the users of these attribute surfaces. In some areas, where the Atlas is the best available mapping land use conflicts and environmental pressures are increasing, in these areas much care should be taken in assessing and using the attribute surfaces.

Interpretations based on combined datasets

Smith (2000) developed an approach to predict pre-clearing soil organic carbon for the Interim Biogeographic Regionalisation of Australia (IBRA; Environment Australia 2000) areas of Queensland. Estimations of soil carbon were made within IBRA regions for national consistency and in accordance with the guidelines from the Australian Greenhouse Office, who commissioned the study. Complete accounts of the methods used are described in Smith (2000). An excerpt of the methodology is listed below.

> *The approach used the available site data as the most reliable estimate of pre-clearing soil carbon at the site and as one of a population of observations which could be used to characterise polygons dominated by similar soils within IBRA regions. Since most soil mapping has been conducted to determine the spatial variation of soil types and a limited number of easily observable soil features, they are likely to have limitations as predictors of soil carbon variation. That limitation is likely to increase as scale broadens – so the analysis used the finest scale polygon data available for any part of the state.*

> *The spatial estimation was derived by overlay of the site coverage over the polygons. Where one or more sites fell within a polygon, the mean value was ascribed to the polygon. For other polygons, the mean values of soil carbon observed in polygons of similar dominant soil type (Australian Soil Classification Order; Isbell 2002) within the same IBRA region was used.*

Brough (2001) developed a revised methodology for the prediction of soil and landscape attributes of importance to salinity processes for Queensland. Estimations for available water, bulk density, clay percent, horizon thickness, saturated hydraulic conductivity, nutrients, plant available water capacity (PAWC), rooting depth, solum thickness and water capacity were produced for Queensland. Many of these attribute interpretations were included in the NAPSWQ Salinity Hazard Mapping Program (Project Sa01; Grundy *et al.* In preparation). The updated approach was improved from previous estimates of soil attributes by obtaining more site and polygon data and included data from the recently developed Soil Profile Class module for the SALI database.

By using a similar process to that of Smith (2000), combined with the use of various 'lookup' tables based on the approach of McKenzie *et al.* (2000b) produced a dataset with superior spatial estimates of a greater range of soil attributes. Complete accounts of the methods used are described by Brough (2001). An excerpt of the methodology is listed below.

> *The approach used the available site (point), polygonal and soil profile class (SPC or taxonomic unit) data from within the SALI database system as the basis for the creation of the soil attribute surfaces. The approaches used in this study follow those of Smith (2000) for the analysis of pre-clearing soil carbon levels and McKenzie et al. (2000b) for the estimation of soil properties from Principle Profile Forms (PPF; Northcote 1979). These processes were used in this project to assign attribute levels to polygons from a lookup table of values for the PPF in question.*

> *Since most soil mapping is used to determine the variation of soil types from limited observations and from easily observable features of the soil profile, limitations to*

> *the prediction of soil attributes exist and increase with scale (Smith 2000). The mapping scales of the data used in this project can be seen in Figure 1. The use of soil taxonomic classes for the prediction of individual soil properties also provides limitations to the usefulness of soil attribute mapping (McKenzie et al. 2000) and will depend on a soil attributes correlation with a mapped entity.*

In 2003, Brough produced a series of updated surfaces using the same basic methodology by including new data and provided estimations for exchangeable sodium percentage, Universal Soil Loss Equation K-factor (Littleboy 1997), air dry moisture content, 15 bar moisture content, sand percentage, drainage, permeability and salt content. Slight improvements to the interpretation process were included in the new version of the soil and landscape attribute surfaces.

4.1.3 Recent developments

The current method of interpretation in Queensland is based on the ASRIS approach for the definition and description of soil and landscape attributes. By following the specifications of ASRIS, the current surfaces have been developed within a nationally consistent framework. The methods outlined in the ASRIS Technical Specifications (McKenzie *et al.* 2005) provide a richer set of information than was available by using any of the previous approaches as described in prior sections. The true power of the current methodology is derived not from the new interpretation methods but from the recent improvements to the SALI database. SALI is now a centralised system that provides a single point-of-truth to site, polygon, soil profile class (SPC) and land system information for Queensland. The new surfaces are derived by directly accessing SALI and building the interpretations by the querying and manipulation of all the stored data.

Soil and Land Information System (SALI)

Since 2003, SALI has undergone a million dollar redevelopment to become a centralised, single point-of-truth system that holds all site, polygon, SPC and land system data. Prior to this, SALI existed as a central data library that contained most completed[1] site data and some polygon data in various formats. All of the working data was stored in Access databases spread throughout the Department. There was no SALI module for SPC or land system data available. The redeveloped SALI stores all land resource data in a consistent and logical framework that simplifies the creation of the attribute surfaces. During the redevelopment process SALI was spatially enabled. Spatial objects, such as polygons, are now stored and managed in the one physical database. The use of a consistent framework, along with the centralisation, has made the process of generating the attribute surfaces easier.

The design of SALI has promoted the logical and consistent design principles that are a requirement of this data intensive attribute surfacing work. SALI has five basic modules, Projects, Sites, Polygons, Soil Profile Class (SPC) and Land Components.

---

[1] In SALI, completed projects are referred to as Master. Master projects are set to a read-only status within the system. Only certain users have the access privilege to alter the status of a project, if a change is made to a status a reason must be recorded within the system and is associated with the project record.

The Project, Site and Polygon modules of SALI have been spatialised, Figure 3 depicts a diagrammatic representation of the interaction between the different modules. The design of the SPC and Land Component modules allows for single features to be related to multiple projects, for example a single SPC can occur in multiple projects within the same region. Another key design feature of SPC and Land Component modules is the concept of versions. Over time, our knowledge of the various landscapes and our ability to characterise them improves. This improved knowledge and the subsequent alteration to the concept for a SPC is managed in SALI by creating versions of that particular soil type. The land component module stores information relating to land systems and land units.

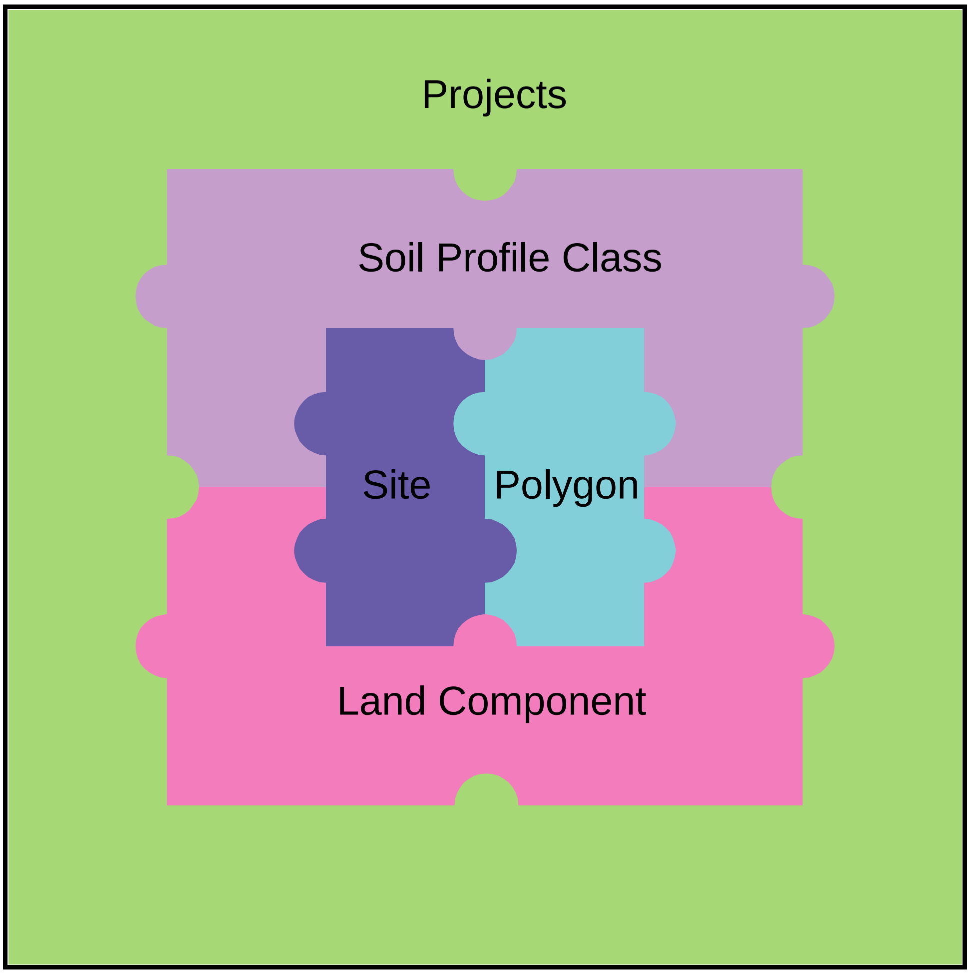


Figure 3. SALI data jigsaw, the relationship between the different SALI modules.

The redeveloped SALI is housed in a single physical Oracle 10g database (Oracle 2006), with the spatial data stored in the same physical database and managed by the ArcSDE 9.1 application (ESRI 2006). The storage and management structure has been simplified from the pre-redevelopment arrangements of an Ingres database to store completed data and Microsoft Access databases for all project data. The previous system was difficult to manage, relied on staff sending their completed Access databases to Brisbane for loading into the Ingres database, which acted as a library. Access to the data was achieved through haphazard means and building a set of all the suitable data for inclusion in any large assessment process was both difficult and time consuming. The spatial data for SALI, which historically was just the polygon mapping, was also stored in multiple formats and at multiple locations within the Department.

The use of industry standard tools, techniques and technologies has improved the useability and reliability of SALI for the production of a range of outputs. By utilising standard technologies that have appropriate levels of scalability and a high level of industry support the future expandability and management of SALI, and by association the attribute surfaces, are assured for at least the medium term. Oracle and ESRI products are the standard software packages, within NRMW, for relational databases and geographic information systems, thereby further enhancing the level of support available for SALI.

In concert with the recent improvements to the SALI database, there has also been a significant investment in the capture of existing land resource information into SALI. The important task of acquiring, this often non-digital, information has proved to be a major benefit to the task of creating the soil and landscape attribute surfaces. The storage of a large number of existing soil profile classes within SALI has enhanced the surfaces in those areas of the state that are covered by detailed land resource assessment projects. The development and deployment of the Land Component module to SALI has allowed reliable interpretations to be achieved for large tracts of land across Queensland. The ability to make consistent estimations of attributes from broadscale mapping is a major benefit of this new process, something that was not achievable until very recently. While much 'historical' data capture has been accomplished in the past few years a large part of the job still remains to be done It is with this in mind that the new process of estimating soil and landscape attributes has been designed to be easily repeatable and updatable.

Australian Soil Resource Information System (ASRIS)

In 2001, the Australian Soil Resource Information System was initially released to provide primary inputs for a broad range of simulation modelling studies supported by the National Land and Water Resources Audit (NLWRA). The 2001 release of ASRIS, initiated in 1999 by the NLWRA (NLWRA 2002, Henderson *et al.* 2001), covered much of continent with very broadscale data. ASRIS 2001 included some detailed data where it was available and able to be modelled. For a more detailed description of the history of ASRIS please refer to McKenzie *et al.* (2005). An excerpt of the history is listed below.

> *The ASRIS 2001 team achieved a great deal given the short time available and daunting nature of the task (see Johnston et al. 2003). During the project, the core team and the National Committee on Soil and Terrain Information (which acted as the Steering Committee) identified a series of deficiencies in the land resource information base for Australia. They also identified a logical pathway for overcoming these problems to ensure a greatly improved system for providing information to support natural resource management in Australia. The task was recognized to be long-term, and requiring a permanent project.*
>
> *With this background, the Australian Collaborative Land Evaluation Program was commissioned to provide with online access to soil and land resource information, and assessments of land suitability. The information is to be available at a range of scales, and in a consistent and easy-to-use format across Australia. The activity must also provide a scientific framework for assessing and monitoring the extent and condition of Australia's soil and land resources.*

The current ASRIS framework is based on three main principles. The first is the definition and delineation of the continent in a series of mapping hierarchies. The second is to produce estimates of attributes from site data from identified layers (or control sections) through a series of rules. Thirdly the provision of spatial estimates are derived by a list of methods with decreasing spatial reliability. The use of the ASRIS framework has allowed the significant investment made by Queensland, through the NAPSWQ program, for the production of spatially reliable soil and landscape attributes to be compatible and consistent with all other national assessments of a broad range attributes. The production of a framework based on a set of nationally agreed and consistent standards is a major achievement for the ASRIS development team and the National Committee on Soil and Terrain (as the Steering Committee). The peer reviewed ASRIS framework has been utilised for the production of soil and landscape information as it is a nationally accepted standard. Efficiencies are achieved in producing a single dataset while meeting several Departmental obligations at once.

The methodology used in ASRIS provides more detail for each attribute that was achieved through any of previous interpretation methods (see Section 4.1.2). Previous methods provided estimates of the attributes as a broad grouping of the A or B horizons (Brough 2001, 2003) as averages. The new methods involve the prediction of an attribute at one of five possible layers (control sections) for each site. Each profile is represented by, (up to) five contiguous soil layers that discriminate the soil materials in terms of their function in relation to water and gas movement, nutrient supply, plant growth, and physical behaviour more generally. By providing an estimate of the attribute in a more detailed model a greater reliability in the prediction of the attribute can be achieved, it also allows for the summarisation of attributes back to horizon or profile weighted averages. The methodology also allows for of both attribute and landscape uncertainties to be recorded for each attribute and control section combination.

4.2 Current methodology

This section outlines the current process of deriving soil and landscape attribute information for Queensland. The soil and landscape attribute information system is a combination of all existing data, information and knowledge on Queensland's landscapes. The estimates have been generated at spatial resolutions applicable to a variety of landscape assessments and modelling purposes. The current process provides both more detail for the attributes in question and an improved spatial reliability through the use of repeatable processes. The processes undertaken to provide spatially reliable estimates of attributes from the interpretation of the suitable site, polygon, SPC and land system data from SALI are outlined in the sections below.

4.2.1 Attributes estimated

A number of soil and landscape attribute surfaces are required to be developed to suit a wide variety of uses. Most of the attributes generated are those required by ASRIS, while other attributes have been estimated for Queensland's purposes. Currently, estimates of 28 attributes have been completed, while another 16 attributes are planned to be completed over the coming months. Some attributes requested by ASRIS may never be completed for Queensland due to a current lack of data. In the future, if adequate pedotransfer functions are developed these attributes maybe completed. The new attribute surfacing system has been designed to provide flexibility, much of the code to derive the attributes is of a modular design allowing for easy re-use. Table 6 lists both the current and proposed attributes currently designated within the system, Appendix 2 contains a complete description of each attribute.

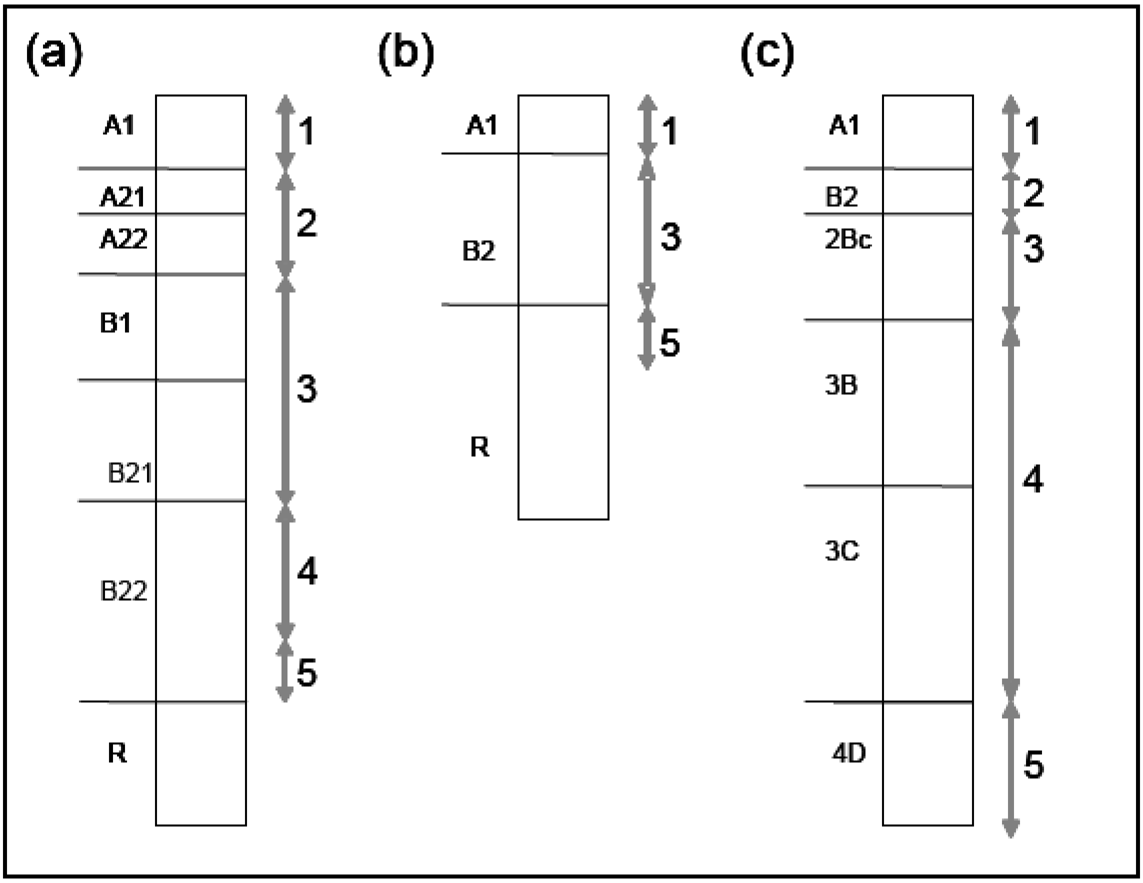


Figure 4. Examples of horizon sequences and allocation to the five-layer model used to describe idealised soil profiles in ASRIS. Example (a) is a common sequence. In example (b), Layers 2 and 4 are recorded as missing because the profile is shallow and has only a few horizons. Example (c) is a complex profile and Layers are specified according to their influence on plant growth and water movement. From McKenzie *et al.* (2005).

As previously stated, attribute surfaces have been generated for Queensland specific purposes. The list of attributes from Table 6 is by no means complete[1]. Previous experience tells us that once a system has been developed requests will be made for the inclusion of new features; requests have already been made for new attributes to be included in the system. The process of including new attributes will be a trade-off between client needs and the ability of the system to generate meaningful estimations.

4.2.2 System structure

A brief discussion of system structure follows in this section, the structure of the soil and landscape attribute system is similar to that proposed in the ASRIS Technical Specifications (McKenzie *et al.* 2005). There are three basic components to the information system, the relational data tables, the spatial datasets and the system code. Appendix 2 contains more system documentation including entity relationship diagrams, table definitions and dimension table information as extracted from the attribute information system.

[1] To obtain the current list of attributes completed for Queensland please contact the NRSc Data Coordinator. Contact details are listed in Section 4.4.

Table 6. Attributes estimated and their completion status.

Parameter	Parameter Description	Status
ASC_ORD	ASC Order	Completed
AGG_STAG	Aggregate Stability	Not Started
ADMC	Air Dry Moisture Content	Completed
AWC	Available Water Capacity	Not Started
BULK_DENSITY	Bulk Density	Not Started
CEC	Cation Exchange Capacity	Completed
CHLORIDE	Chloride Content	Completed
CLAY	Clay Content	Completed
COARSE_FRAGMENTS	Coarse Fragments	Completed
SAND_COARSE	Coarse Sand Content	Completed
DEPTH_A	Depth of A horizon	Not Started
DEPTH_A1	Depth of A1 horizon	Not Started
DEPTH_REGOLITH	Depth of Regolith	Not Started
DEPTH_B2	Depth to B2 horizon	Not Started
DEPTH_ACP_IMP	Depth to Impeding Layer - Annual Crops and Pastures	Not Started
DEPTH_PNV_IMP	Depth to Impeding Layer - Perennial Native Vegetation	Not Started
DEPTH_PP_IMP	Depth to Impeding Layer - Perennial Pastures	Not Started
DRAINAGE	Drainage	Completed
EC	Electrical Conductivity	Completed
EX_CA	Exchangeable Calcium	Completed
EX_MG	Exchangeable Magnesium	Completed
EX_K	Exchangeable Potassium	Completed
EX_NA	Exchangeable Sodium	Completed
ESP	Exchangeable Sodium Percentage	Completed
TEXTURE_CODE	Field Texture	Completed
TEXTURE_GRADE	Field Texture Grade	Completed
SAND_FINE	Fine Sand Content	Completed
IMP_TYPE	Impeding Layer Type	Not Started
KS	Ksat	Not Started
DEPTH_LOWER	Lower Depth of Control Section	Completed
BAR_15	Moisture Content - 15 Bar	Completed
ORGANIC_CARBON	Organic Carbon	Completed
PED_SIZE	Pedality Size	Completed
PED_TYPE	Pedality Type	Completed
PERMEABILITY	Permeability	Completed
PROFILE_AWC	Profile Available Water Capacity	Not Started
REG_TYPE	Regolith Type	Not Started
ROCK_OUTCROP	Rock Outcrop Abundance	Not Started
SILT	Silt Content	Completed
REG_KS	Substrate Ksat	Not Started
SCF_2-60	Surface Coarse Fragments 0.002-0.06m	Not Started
SCF_60-200	Surface Coarse Fragments 0.06-0.2m	Not Started
SCF_200-600	Surface Coarse Fragments 0.2-0.6m	Not Started
SCF_600-2M	Surface Coarse Fragments 0.6-2m	Not Started
SCF_2M	Surface Coarse Fragments >2m	Not Started
THICKNESS	Thickness of Control Section	Not Started
TOTAL_N	Total Nitrogen - Kjeldahl, automated	Completed
TOTAL_P	Total Phosphorous - XRF	Completed
DEPTH_UPPER	Upper Depth of Control Section	Completed
WATER_REPEL	Water Repellence	Not Started
PH	pH	Completed

Relational data tables

The design and structure of the relational data tables follows data warehousing principles. The data tables are multidimensional data structures that are designed for query and analysis rather than for transaction processing. SALI is a transactional database and is designed and optimised for fast and reliable transaction handling where most interactions involve a small number of rows from a large group of tables. Table 7 contains some brief contrasts between SALI and the soil attribute system. It is beyond the scope of this document to discuss data warehousing – for more information of a general nature refer to *The Data Warehouse Toolkit* (Kimball and Ross 2002).

The system structure has been defined to hold all site, polygon, SPC and land system information in the same series of tables, compared to SALI where all of this information is held in different data structures with linkages between them. The system contains three main 'fact' tables, which are the features and results tables and a table that links each feature to other features, or the feature composition table. Each of these tables contains a column to define the feature type, the feature type is one of the key pieces of information for each record. An example of the feature type identifier is SI for site data. Other unique identifiers within the system are the project code and individual feature identifiers such as site and observation numbers for site data and the polygon and entity numbers for polygon information. Default project identifiers of SPC_DATA and LS_DATA exist for SPC and land system data respectively. The feature composition table provides linkages between each of the features described within the system, for example the table records the linkages between sites and polygons, soil profile classes and land systems.

The data in the feature and feature compositions tables is provided by direct population from each of the SALI component modules. The results table is populated through a querying and interpretation process from the raw SALI data. There are also a number of codes or dimension tables within the system, these tables provide information and descriptions of much of the other information recorded within the fact tables.

Table 7. Contrasting SALI and the Soil Attribute System Relational Database Environments.

Item	SALI	Soil Attribute System
Data Structures	Complex	Multidimensional
Query Span	Tens of records	Thousands to millions of records
Indexes	Few	Many
Joins	Many	Few
Derived Data and Aggregates	Rare	Common
Duplicated Data	Normalised database	Denormalised database

Spatial datasets

The spatial datasets for the attribute information system are built as ArcSDE feature datasets, the major spatial data table holds the polygons used to defined the attribute surfaces. The polygons are a spatial interpretation of all the polygons from suitable projects in SALI that have been combined within the spatial hierarchy. Key identifiers of polygon data in the table are the project code and polygon number as defined in SALI. The process of building the spatial dataset is described in Section 4.2.3.

The soil and landscape attribute information system uses the base polygon information from the combined polygons dataset to produce a series of views of the data. Spatially enabled views have been implemented to reduce redundancy incurred by using the same polygon data multiple times to illustrate different combinations of the attributes and control sections. Spatially enabled views operate the same as views or queries within any other relational database management system but they include the spatial data table which allows combination of spatial and textual attribute information to be displayed in a Geographic Information System.

System code

As with any relational database management system, which includes a data warehouse, a large amount of customised system commands, or code, must be developed. The soil and landscape attribute information system is no exception to this. The code for the attribute system has been written to manage all steps in the attribution, interpretation and spatialisation stages of the soil and landscape attribute information system. The language used to interact with an Oracle database is the Structured Query Language (SQL).

SQL, pronounced either “sequel” or “S-Q-L”, is a very capable tool. Using it does not require any programming experience. The SQL language has structure, just as the English language has structure. It has rules of grammar and syntax, but they are basically the normal rules of careful English speech and can be readily understood. Using SQL you request which information you want to select, insert, update or delete. These four verbs are the primary words in the SQL language.

The attribute information system code has been developed in the PL/SQL language. PL/SQL is Oracle’s procedural language (PL) superset of the Structured Query Language (SQL). PL/SQL allows the codification of business rules through the creation of stored procedures and packages. It also adds programming logic to the execution of SQL commands. PL/SQL code is grouped into structures called blocks. Many of the blocks of code are given names. A block of PL/SQL code contains three sections - declarations, executable commands and exception handling. Table 8 describes these three elements of a PL/SQL block.

The bulk of the code developed for the attribute system has been compiled as PL/SQL packages with numerous functions and procedures. To date some 12,000 lines of code have been written to manage all parts of the system. A small percentage of the code has been developed to interact with the spatial area of the system. Some of this code is used with the ArcSDE application and operates as scripts and the command line interface. The remainder of the spatial code is written in the ArcGIS 9.1 application programming interface (ESRI 2006), using either Python or the ArcGIS ModelBuilder, to manage the creation of the spatial datasets to complete the attribute system.

The benefits of developing the attribute system in a language such as PL/SQL is that the modularity and re-use options for individual functions and procedures are significantly enhanced. Without the ability to re-use, or call functions from different parts of the code system, the amount of code needed to run the system would exponentially increase. Examples of this code modularity will become evident in Section 4.2.4.

It is beyond the scope of this document to fully describe the packages and procedures compiled for the attribute information system, it is hoped that this short introduction to the methods used are suitable to satisfy the requirements of most readers.

Table 8. Elements of a PL/SQL code block.

Section	Description
Declarations	Defines and initialises the variables used in the block
Executable Commands	Uses flow-control commands (such as loops and if statements) to execute commands and assign values to the declared variables
Exception Handling	Provides customised handling of error conditions

4.2.3 Producing the spatial hierarchy

The production of the spatial hierarchy of the polygon data from SALI is one of the critical steps in the construction of spatially accurate and reliable soil and landscape attribute surfaces. To complete the task of producing attribute surfaces, each project with polygonal data from SALI was overlaid to obtain the best spatial polygon information at any point across Queensland. Each project, and therefore polygon, is assigned to a hierarchal level within the soil and landscape attribute information system. ASRIS has defined several levels to its mapping hierarchy (McKenzie *et al.* 2005 Table 3). The hierarchal arrangement all the data produced by the soil and landscape attribute information system is mapped onto Levels 4 and 5 of the ASRIS hierarchy.

The attribute surfaces developed for Queensland fit below the ASRIS mapping hiatus, between Levels 3 and 4, as they are aggregations of land resource surveying data. While some data exists for Queensland that is appropriate to Level 6 of the ASRIS hierarchy it has not been included in the attribute information system at this point in time. At some future point this data may be included in the soil and landscape attribute system.

Each of the mapping projects from SALI has two key pieces of information to assist in the defining its level with the spatial hierarchy, these are the scale and type of survey[1]. For Queensland we have identified that there are thirteen individual levels of the hierarchy to produce information for Levels 4 and 5. There are multiple planes within each of the ASRIS levels to account for projects that have spatial overlaps in their extent. By eliminating the spatial overlaps each level in the hierarchy is topologically correct. This will assist in any future interpretations. Table 9 lists the levels defined within our hierarchy as well as the dominant scale and survey type of the projects assigned to each level. Figure 5 and Figure 5 depict the spatial extent of each of the thirteen levels within the hierarchy and the combined hierarchy image to gives an indication of the extent and interaction of the different levels.

The hierarchal numbering system operates in relatively straight-forward, as the hierarchal number increases so does the perceived accuracy of the information those polygons convey. The numbering system is based on that used for ASRIS (e.g. Level 4) - with intermediate levels are determined by the scale and survey type (e.g. Level 4.2). Within intermediate levels, sub-levels are used where the spatial arrangement of projects creates an overlap within the intermediate level (e.g. Level 4.51). Projects are assigned to the various sub-

[1] Typical survey types for Queensland at Level 4 include land system and land resource area surveys, at Level 5 survey type is typically soil or soil and land suitability surveys.

levels based on expert opinion. There is no difference in projects between related sub-levels, they are assigned to simplify the procedure to creating the combined polygon coverage and to ensure topological correctness.

Table 9. Hierarchal levels of the system, including dominant scale and survey type.

Hierarchy	Tract	Dominant Scale	Dominant Survey Type
1	Division		
2	Province		
3	Zone		
4	District		
4.1	District - Intermediate 1	2,000,000	Soil Survey
4.2	District - Intermediate 2	500,000	Land System Survey
4.3	District - Intermediate 3	1,000,000	Soil Survey
4.4	District - Intermediate 4	250,000	Land System Survey
4.51	District - Intermediate 5.1	250,000	Land Resource Area Survey
4.52	District - Intermediate 5.2	250,000	Land Resource Area Survey
4.6	District - Intermediate 6	250,000	Soil Survey
5	System		
5.11	System - Intermediate 1.1	100,000	Soil Survey
5.12	System - Intermediate 1.2	100,000	Soil Survey
5.13	System - Intermediate 1.3	100,000	Soil Survey
5.21	System - Intermediate 2.1	50,000	Soil Survey
5.22	System - Intermediate 2.2	50,000	Soil Survey
5.3	System - Intermediate 3	25,000	Soil Survey
6	Facet		

Processing the spatial hierarchy

Building the polygonal spatial coverage from the hierarchy occurs in an ArcGIS environment. A Python script was created to construct the coverage using the best polygons as defined from the spatial hierarchy data. The coverage consists of the polygons from higher in the hierarchy being 'cookie cut' into those from a lower hierarchy. The process used in this iteration of the combined dataset is conceptually the same as previous versions, as described by Smith (2000) and Brough (2001, 2003). In practice the re-development of SALI and the spatialisation of polygon data combined with the use of the SDE environment have significantly simplified the process.

The polygonal data from each level of the hierarchy is extracted from the soil and landscape attribute information system, or SALI as the case maybe. Each layer of the hierarchy is unioned in ArcGIS starting with the layers at the highest level, e.g. smallest scale, and working down through the hierarchy. The final unioned coverage is dissolved on the project code and polygon number from the highest polygon to produce a spatial dataset containing some 115,000 individual polygons[1]. This spatial dataset of the combined polygons is used to convey all of the information for the soil and landscape attribute surfaces.

Another product from the creation of the spatial hierarchy coverage is a table defining the spatial relationship of polygons between the different levels of the hierarchy. This is built as part of the unioning process prior to the coverage being dissolved. For each polygon in the hierarchy a list of the polygons it intersects from the levels below is produced. This table is

[1] As at May 2006.

used in the production of outputs from the system to account for polygons where no information is able to be derived.

4.2.4 Attribute data processing

There are a number of data processing steps that are required to produce the soil and landscape attribute surfaces. These steps can be broken into the processing of site data, SPC data, land system data and polygon data. The goal of processing site, SPC and land system data is to define the attribute values that are utilised in the population of the polygons. There is much similarity between the processing of the base data types, this will become apparent in the following sections. All of the data processing steps from the attribute interpretation system are very similar to those described in the ASRIS Technical Specifications (2005).

Processing site data

The first step in the construction of the soil and landscape attribute surfaces is to process the base data through the interpretation of the 84,000 sites in SALI[1]. The data processed from SALI is the observation data for each site. The observation data is utilised through the feature compositions table with its link to their respective site. Each of the observations in SALI passes through a series of rules to compute the horizon that is used as a placeholder for the derivation of attribute values. Each attribute that is derived has a series of horizon rules that must be followed and the rules for each horizon may have a number of sub-rules. Many of the rules are the same for a number of attributes, for example the computed horizons for texture grade are the same as those for coarse fragments.

The rules for the calculation of a horizon are used in a predefined order, if a rule returns a horizon number to the calling function, subsequent processing of the horizon definition rules is halted. There are three possible results of the execution of the horizon calculation process 1) a valid horizon is returned, 2) the horizon is defined as missing and the Missing code is returned or 3) a valid horizon was not able to be determined and the Not Recorded code is returned. The values for Missing records are -9999 and NA for numeric and character type attribute respectively, the value for Not Recorded (not determined) records are -1234 and NR for numeric and character type attribute respectively. The values for Missing and Not Recorded are defined in the ASRIS Technical Specifications.

In the attribute system there are currently 20 individual rules for the definition of horizons, these rules form 22 unique combinations for the calculation of horizon details for each control section. The 22 combinations are utilised on 254 different attribute and control section combinations. This number of combinations of rules and attributes for each of the control sections is where the modularity of the PL/SQL code is used, each of the 20 rules only had to be written once as a function, this allows the function to be called multiple times to define the horizon for the rule combination in question. Each rule combination is completed for the observation data and the horizon return is stored in the database results table. This permits all attributes to be calculated from the observation data in one intensive data processing step.

[1] As at May 2006.

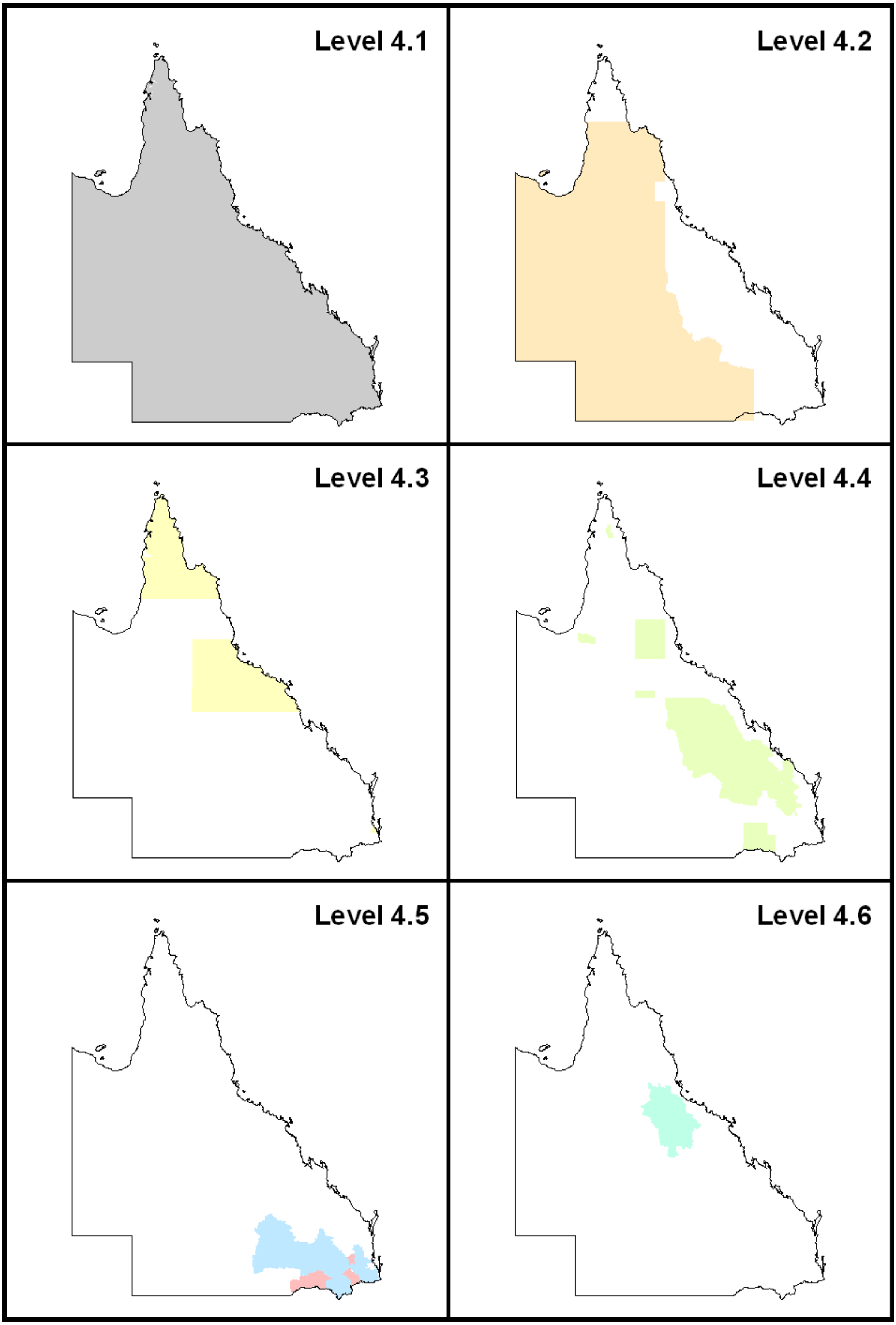


Figure 5. Area covered by levels 4.1 to 4.6 of the spatial hierarchy.

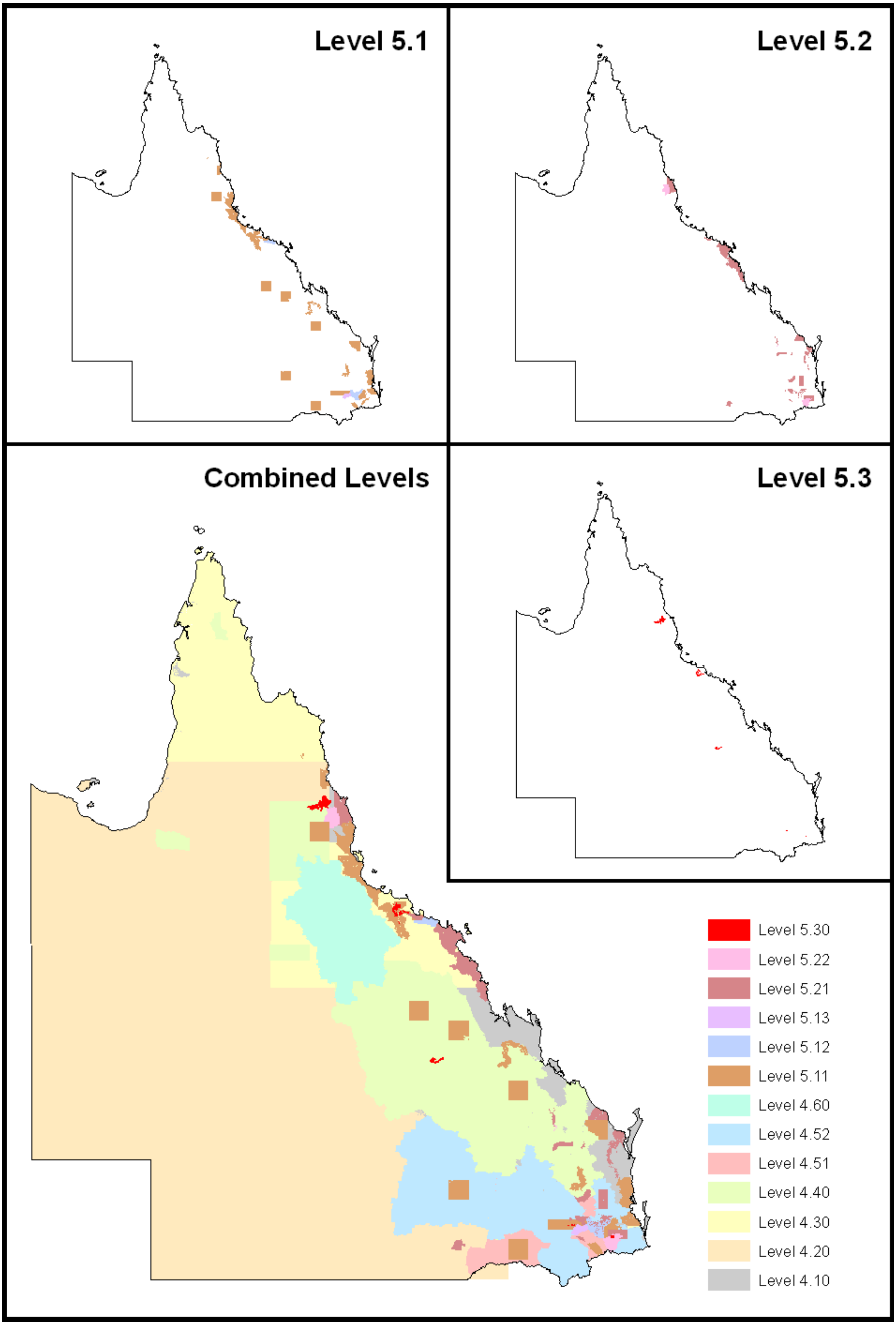


Figure 6. Area covered by levels 5.1 to 5.3 of the hierarchy and the combined hierarchy.

For example, if an observation has a profile consisting of A1 (non-peat texture), A2, B21, B22 (thickness 0.3 m) and B23 (thickness 0.8 m, horizon upper depth 0.9 m) horizons. Using the horizon definition rules from Table 10, control section 1 will be defined as the A1 horizon from Rule 2. Control sections 2, 3 and 4 will be derived as the A2, B21 and B23 horizons from Rule 1 for their respective control sections. Control section 5 will be recorded as missing.

Table 10. Control section rules for the calculation of the horizons for texture grade.

Rule	Description
Control Section 1	
Rule 1	If the surface layer is a peat, refers to the 7 classes of organic materials defined by McDonald and Isbell (1990).
Rule 2	If there is a single A1 horizon without subdivisions (e.g. A11, A12 etc.), then refers to the A1 horizon. Ignores any possible overlying surface horizons.
Rule 3	If there are subdivisions within the A1 horizon, refers to the thickest A horizon layer within the top 0.20 m of the soil profile (the upper layer is used if thicknesses are equal). If the surface layer is an O horizon, refers to an A horizon as defined accord with the above criteria.
Rule 4	If the surface layer is an O horizon and there is no underlying A1 horizon, refers to the thickest layer in the 0.20 m directly beneath the O horizon.
Control Section 2	
Rule 1	If the surface layer is not an A1 horizon (e.g. O horizon) and there is no underlying A horizon, the attribute is recorded as missing, otherwise refers to the lower portion of the A horizon that is below control section 1, if not below control section 1 then record as missing
Control Section 3	
Rule 1	If a B horizon is present, refers to the uppermost B horizon (usually B1 or B21).
Rule 2	If no B horizon is present and the sequence consists of an AC profile, refers to the major part of the materials in the 0.20m directly below the A horizon.
Control Section 4	
Rule 1	Refers to the thickest B horizon below the control section 3 that has an upper depth above 1.5m
Control Section 5	
Rule 1	Refers to the lower 0.1m of the horizon above 2m, or above a pan or above a C or R horizon, and is below the horizon defined for control section 4, otherwise recorded as missing.

Processing Soil Profile Class data

The processing of SPC data is the second step in the interpretation process to produce surfaces of soil and landscape attribute information. This step involves the interpretation of information for the 3,500 soil profile classes described in SALI[1]. Each SPC passes through a two stage interpretation process, the first is closely aligned to the interpretation of site data while the second is similar to the process of estimating attributes for polygons and entities.

As with sites, each SPC passes through a series of rules to compute the horizon that is used as a placeholder for the derivation of attribute values. These rules are exactly the same as those for observation data. The difference for an SPC is that the horizon for a rule combination is processed up to three times. As an SPC is a conceptual unit that may record a range of attribute values, each is assessed with respect to its minimum, modal and maximum attribute ranges. The ability to define a range of attributes for a SPC is an important conceptual point within the attribute system. Currently, the data derived for the

[1] As at May 2006.

minimum and maximum range of an attribute is not utilised for any further processing within the system. In the future, this information maybe used in the calculation of the range of possible values for modelling purposes, or the attribute values estimated for a polygon maybe displayed as falling within a range of potential values.

The first step in processing SPC data does not interpret data for the SPC it uses only information recorded for that SPC. For example texture data is ascribed to the SPC but no laboratory data is estimated. Within SALI, no laboratory information is stored for a SPC.

For attributes not defined during the first processing step, a process similar to the attribute estimation for polygon entities is used. The estimation methods use site data that is either representative or has recorded the SPC as its soil taxonomic unit. If the attribute is unable to be estimated from site data it is estimated from soil classification information.

Due to nuances in the coding of the soil and landscape attribute information system each control section is first checked against the result value recorded for field texture. If the field texture is recorded as missing the new attribute result value is recorded as missing also. This precursory examination of the control sections prevents some SPC control sections existing as missing for one attribute and valid values for another.

For a complete description of the process of estimating attributes, based on the polygon approach please refer to the Processing polygon data section.

Processing land component data

The processing of land system data in the soil and landscape attribute system is the third step in the production of the attribute surfaces. This step involves the interpretation of attributes for the 415 land systems and 1300 land units currently described in SALI[1]. There are two components to land system data, these are the land units and land systems. The estimation of attributes for land units and land systems components are processed differently though both are based on the polygon estimation methods.

As for the attribute estimation of SPCs, the texture code result is checked for each land component during a precursory examination of the land component.

Land unit data

The land unit is the first data type to be processed within the land systems concept. SALI holds the descriptions of land units which include information on the SPC and sites that are representative of that land unit.

The process of estimating attributes for a land unit is the same as for the estimation of entities. The attributes are estimated from representative sites for the land unit or from the SPC recorded for the land unit. If the attribute is not able to be estimated from site data it will then be estimated from the soil classification information.

[1] As at May 2006.

For a complete description of the process of estimating attributes, based on the polygon approach please refer to the Processing polygon data section below.

Land system data

Following the estimation of attributes for land units, the attributes for a land system are estimated. SALI holds the descriptions of land systems which include their component land units and the percentage of the system in which they occur.

The process of estimating attributes for a land system is the same as for the estimation of polygons. The first attempt at attribute estimation is to prepare an area weighted average of the attribute if it is of a numeric type (e.g. pH), if the attribute is of a character type (e.g. texture code) then the dominant value is used. If the first attempt at estimation results in no data being able to be determined then the available site, SPC and soil classification data is used to create an attribute value.

For a complete description of the process of estimating attributes, based on the polygon approach please refer to the Processing polygon data section below.

Processing polygon data

The final step in the attribute estimation process is to calculate attribute values for the polygons from SALI. There are two components to polygon information in SALI. These are the polygon and entity records, of which 151,000 and 217,000 have been described respectively[1]. The calculation of polygon attributes depends on the estimation of the site, SPC and land system records from SALI. The estimation of attributes for the polygons and entities are processed differently as mentioned in the previous sections.

Entity data

The estimation of attributes for entities is the penultimate step in the creation of the soil and landscape attribute surfaces. The estimation processes follow those listed in the ASRIS Technical Specifications. The estimation rules define how an attribute value is estimated by using the best available data. The rules are used in a predefined order, if a rule returns a result for a feature (in this case an entity) then subsequent processing of that feature is halted. There are four possible results from the estimation process 1) a valid result is returned, 2) the layer is defined as missing in the base data and the Missing code is returned, 3) the layer is defined as not determined in the base data and the Not Recorded code is returned, 4) no valid value is able to be determined and the Not Recorded code is returned. The values for Missing and Not Recorded are the same as defined those for site data and the ASRIS Technical Specifications.

In the soil and landscape attribute information system there are currently nine individual rules for the estimation of attributes. These rules form 189 combinations of parameters, control sections and estimation rules. The number of combinations of estimation rules for each of the attributes is where the modularity of the PL/SQL code is useful, each of the nine rules only had to be written once as a function, this allows the function to be called multiple times

[1] As at May 2006.

to estimate the result for the rule combination in question. Examples of estimation rule re-use are shown in Table 11.

Table 11. Estimation method rules defined for field texture grade and clay percentage.

Rule No	Estimation Method	Description
Field Texture		
1	Texture 1	Estimate based on an replicated and representative measurements in the land unit tract
2	Texture 2	Estimate based on an un-replicated and representative measurement in the land unit tract
3	Texture 3	Estimate based on direct measurements of similar soils in the same land unit type (e.g., modal profiles)
4	Texture 4	Estimate based on direct measurements of similar soils in the region or project area, includes sites that are located within the land unit tract
5	Texture 5	Estimate based on experience with similar soils (e.g., same taxa in the Australian Soil Classification but from other regions)
Clay Percentage		
1	Texture 1	Estimate based on an replicated and representative measurements in the land unit tract
2	Texture 2	Estimate based on an un-replicated and representative measurement in the land unit tract
3	Texture 3	Estimate based on direct measurements of similar soils in the same land unit type (e.g., modal profiles)
4	Clay 4	Estimate is based on field textures from representative soil profiles in the land-unit tract (estimate of field texture from Texture 1 rule)
5	Texture 4	Estimate based on direct measurements of similar soils in the region or project area, includes sites that are located within the land unit tract
6	Clay 6	Estimate is based on field textures from soil profiles in the land-unit tract (estimate of field texture from Texture 3 rule)
7	Clay 7	Estimate is based on field textures from similar soils in the project area (estimate of field texture from Texture 4 rule)
8	Texture 5	Estimate based on experience with similar soils (e.g., same taxa in the Australian Soil Classification but from other regions)

The typical sequence of estimation rules for the soil and landscape attribute information system is the use of representative sites within the entity. If multiple sites exist, this defines the first or second estimation method. In SALI, sites maybe recorded as either representative or not for a feature. The third estimation method uses the attribute value from the local classification type of the entity. This may be either site, SPC, land unit or land system information. Much of this information is derived from the feature composition table. The fourth estimation method utilises sites that are recorded in SALI as being within the polygon, these sites maybe non-representative. If a polygon has only one entity, then any site that is spatially within the polygon is used as a second sub-step to this method. The final estimation method utilised uses the soil classification data recorded for the entity, SPC or sites within the polygon to derive an attribute value. The features are used in the order they are listed. The classification methods used are the Australian Soil Classification (Isbell 2002) and the Factual Key (Northcote 1979) respectively. If any of the methods estimates a value for the entity then further processing is halted.

A different estimation method is used for determining drainage and permeability attributes where expert interpretation has been applied to the entity. The ability to record attributes such as drainage and permeability for entities is catered for in SALI. Drainage and permeability are routinely recorded during detailed UMA style surveys. The expert interpretation rule is used prior to other estimation rules. In the future other attributes may also use this method after careful consideration of the likelihood of the attributes representativeness for an entity.

Due to nuances in the coding of the soil and landscape attribute information system each control section is first checked against the result value recorded for field texture. If the field texture is recorded as missing the new attribute result value is recorded as missing also. This precursory examination of the control sections prevents some entity control sections existing as missing for one attribute and valid values for another.

The purpose of using a complicated series of estimation methods, when compared to previous attempts at interpret attributes for a combined dataset, is to provide more reliable estimates of soil and landscape attributes. Using the estimation methods defined above the earlier an attribute is derived in the estimation process the greater the accuracy of the attribute.

UMA and mapcode mapping data

The estimation of attribute results for polygons is the final step in creation of the soil and landscape attribute surfaces. The estimate for a polygon is derived as an area weighted average of its entities. In SALI, each polygon has at least one entity record, each entity is recorded as either a percentage or proportion of the parent polygon. If an entity only has proportions recorded, these are converted to percentages during the population of the features and feature composition tables. Each proportion is converted to its maximum percentage, summed for the polygon and scaled to 100%, see Table 12 for an example. Any values of missing or not recorded are ignored during the first step of polygon attribution.

The area weighting of results occurs only for attributes with a numerical data type, e.g. pH. A geometric mean is used for attributes that are log normally distributed, e.g. electrical conductivity. If an attribute has a character data type, for example field texture, the unique result with the largest summed percentage of area for the entities is used. All values of Not Recoded (result values of NR or -1234) are ignored in the area weighting process.

To enable attribute results to be created for each polygon a secondary processing step is used. Where the value for a polygon is not determined and the result value for texture grade is missing, the polygon attribute result is recorded as missing. If a result value is still not determined it is finally recorded as not recorded.

Table 12. Example of proportion to percentage calculations.

Entity	Proportion	Percentage			
		Actual	Minimum	Maximum	Scaled
1	7 (61-70%)	70	60	70	64
2	2 (11-20%)	15	10	20	18
3	2 (11-20%)	15	10	20	18
Total		**100**	**80**	**110**	**100**

Table 13. Example of area weighting calculations.

Entity	Percentage of Polygon	pH	Area Weighted pH	Field Texture Grade
1	70%	6.5	4.55	Clay Loam
2	15%	5.5	0.82	Sandy Loam
3	15%	8.5	1.28	Medium Clay
Polygon Result	**100%**		**6.65**	**Clay Loam**

Miscellaneous processing steps

This section describes several other important processing steps that take place in the in the creation of the soil and landscape attribute surfaces, these processes are not directly aligned with those outlined in the previous sections.

Spatially locating sites in polygons

To spatially locate and record in the soil and landscape attribute information system the arrangement of sites in polygons an ArcGIS Python script is used. The script is included as part of the spatial hierarchy process. The purpose of this process is to improve the estimation of the polygons where no other information exists, typically this is achieved in estimation method 4 (from Table 11).

Following the completion of the ArcGIS process, the table containing the sites and polygon data is inserted into the feature composition table. A secondary process of altering the polygon record is achieved by 1) if the polygon has one entity, the site is recorded for that entity or 2) if the polygon has multiple entities the taxonomic classification of the site and entities are compared, if a match is found then the site is assigned to that entity. Of the 150,000 polygons recorded in SALI, approximately 64% have only one entity record[1].

Creating the grouped result information

Some of the attributes used as part of the information system are grouped from a set of other attributes, for example pH is defined by four different analytical methods and a fifth overall pH result is also recorded. The grouping of attributes allows the most accurate result to be used from the most reliable estimation method. The attributes that are grouped to improve the accuracy of their estimation are pH, Electrical Conductivity and the cation measurements[2].

[1] As at May 2006.
[2] Cations include Calcium (Ca), Magnesium (Mg), Sodium (Na), Potassium (K), Cation Exchange Capacity (CEC) and Effective Cation Exchange Capacity (ECEC).

Using pH as the example from Table 14, it is estimated from four different analytical methods. The degree of measurement accuracy decreases as the order of use increases. To derive an estimate of the most accurate pH for a feature, within each estimation method the measurement with the lowest 'order of use' is recorded as the attribute result.

Creating the polygon results table

As part of the process of creating more usable outputs from the system a new polygon results table is created from the results. The polygon results table contains two results fields, one each for numeric and character results to simplify the data display process in many GIS packages. Two fields capturing hierarchy data are included in the table, they record the hierarchy of the spatial polygon and the hierarchy of the polygon the result is derived from.

The polygon results table contains records for only those polygons in the combined polygon coverage as described in Section 4.2.3. If a polygon has an attribute result estimated as Not Recorded (NR or -1234) then the results from polygons lower in the hierarchy that it intersects with are used.

Many of the character attributes used in the soil and landscape attribute information system have a numerical equivalent defined in the codes table. For example, the numerical codes associated with field texture code are produced as a sequential number based on the increasing clay content of the field texture. This numbering sequence does not have any particular meaning but allows many of the character based attributes to be used more efficiently in a GIS environment for display and data manipulation purposes.

Table 14. Grouped attributes and their order of grouping.

Order of Use	Measurement Method		
	pH	Electrical Conductivity	Cations
1	Laboratory	Laboratory	Laboratory – Alcoholic Method
2	Field – 1:5 soil/water	Field – Laboratory Method	Laboratory – Aqueous Method
3	Field – Electrode Probe	Field – 1:5 soil/water	
4	Field – Raupach & Tucker		

4.2.5 Data update mechanisms and frequency

The soil and landscape attribute information system has been developed to be dynamic and updateable. It is envisaged that the system will become the standard method of viewing and querying interpreted attributes for Queensland. It is with this in mind that the entire system has been produced to be easily updatable, compared to a normal data warehouse where the information is generally static.

There is currently no planned regular update frequency. Given the sporadic nature of the updates or new data collection made to SALI, the estimated update period is approximately six months.

The data update mechanism is relatively simple, the soil and landscape attribute information system will scan SALI for features that have been created or updated, based on the auditing fields in SALI. Any of the created or updated features have been updated will be re-processed by the attribute information system. Any feature related to the those that are updated will also be updated. For example, a site receives new analytical data, the site is also representative of an entity and a SPC. The site (including observation), SPC and any polygon related to the site or SPC will be re-processed by the attribute information system. The updated results will then be immediately available through the data delivery mechanisms of the attribute system.

4.2.6 Data delivery mechanisms

The ability to effectively and efficiently query, display and interpret the soil and landscape attributes from the information system is a key factor in its use by a wide variety of stakeholders. Currently data display mechanisms are aimed at the provision of basic information from the system. It is envisaged that the type and format of information from the system will be improved over time as the system gains a wider audience and feedback is received with respect to outputs available. This section describes the manner in which outputs are currently provided from the attribute system. There are two major components to the output of information from the system, these are the spatially enabled views, which manage the spatial display of data, and the export formats, which enable the use of information outside of the information system.

Spatially enabled views and data display

The ability of Oracle and the ArcSDE applications to provide spatially enabled views, to combine the spatial and textual data, has been widely used to output data from the attribute system. The use of spatially enabled views allows a textual dataset, the polygon results table, to be joined to a spatial dataset, the combined polygons coverage, and spatially displayed. The ability to use one spatial dataset to produce hundreds of data layers significantly reduces the overheads of storage space and management of the spatial section to the attribute information system.

The disadvantage of spatially enabled views compared to standard spatial datasets is the performance decrease in the ability of the system to read data. With more efficient indexing, both of the textual and spatial data, this limitation maybe overcome or further consultation with system users may result in an opinion that this is not a limiting factor of the system. The ASRIS team have decided against the use of spatially enable views because of significant performance decreases caused by increased processing times (D Jacquier *pers comm*).

Current export formats

The export format of soil and landscape attribute data from the system maybe handled in a variety ways. The current preferred method is to export the combined polygons dataset as an ESRI shapefile, each record in the shapefile contains a unique identifier to simplify joins with the attribute data files. The attribute information is exported as delimited text files, each attribute and control section combination is exported into separate files to improve their manageability, the data in these files is selected from the polygon results table with the inclusion of the unique identifier for each polygon. A series of summary files are exported from the system. These summary files contain the results only for each of the attributes, they do not contain the uncertainty or hierarchy level details.

Metadata is available for the polygon coverage and is available with the data or upon request. Metadata for individual export files is not available but a standard metadata document is available. Departmental standards have been used to create metadata for the soil and landscape attribute information system and is compliant with ANZLIC standards.

4.3 Validation process

Several validation processes have been undertaken during the development of the attribute information system, both for the system itself and the methodology used. The validation processes have included

1) The validation of the ASRIS methodology by the National Committee on Soil and Terrain Information and representatives of the various government agencies across Australia,
2) The validation of the system code to ensure the horizon definition processes of the base attribute data are occurring as the rules stated,
3) The validation of the system code to ensure the estimation processes of attribute data are occurring as the methods stated and
4) The validation of the final products by the regional Land Resource Officers from NRMW to ensure that the attribute surfaces match their knowledge of Queensland's landscapes.

Validation processes 2 and 3 are part of the ongoing commitment to system enhancements. Whenever new rules, estimation methods or other system changes are made, the appropriate data processing steps are checked for consistency against the rules defined.

The validation of final products by NRMW officers to ensure the attribute surfaces matched their knowledge of the Queensland's landscapes was an important step to allow the new interpretation methods to gain a wide acceptance as the preferred methodology within Queensland.

4.4 Data storage

All datasets, interpretations and system code for the soil and landscape attribute information system are stored in the NRMW Spatial Information Resource (SIR) environment. A subset of the attribute surfaces will be available via ASRIS (www.asris.csiro.au). All SALI data will continue to be maintained in the SALI system.

To obtain a copy of the soil and landscape attribute surfaces or SALI data from NRMW please contact the NRSc Data Coordinator. The contact details for the Data Coordinator are

NRSc Data Coordinator
Data Delivery Team
Ph: (07) 3896 9862
Email: NRScDataCoordinator@nrm.qld.gov.au

4.5 Future improvements

A number of potential future improvements to the attribute information system have already been identified. These range from the collection of new data to improvements in the interpretation process itself.

The data collection activities to improve the soil and landscape attribute information system include the collection of more paper based information into SALI. This is especially true for land systems data where only a small proportion of the available data is recorded in SALI. The collection of new site, polygon and SPC data will continue to improve the system. A focus on capturing data in areas where either information and knowledge is poor or through a targeted strategic sampling program, as described in Section 3, will continue to improve the accuracy and reliability of the attribute estimations.

Within the attribute system, future improvements will include the investigation of stratifying the use of soil classification data into broad physiographic zones to improve the attribute estimation based on the Australian Soil Classification and the Factual Key. Currently it is unknown how the broad effects of climate and geology affect the estimation of attributes from soil classification data when compared to a single statewide assessment. Other improvements that will be investigated include the use of correlation data from SALI for both SPC and Land Systems records. The use of data from a similar SPC may provide greater certainty in the prediction process, when compared to the use of soil classification data.

5 Conclusions

This section contains the conclusions to the report, and has been split into the two major sections of this report, the site selection strategy and the soil and landscape attribute surfaces.

5.1 Site selection strategy

When developing a sampling system it is important to consider why the data is being collected and how it will be used. Most of the existing soil sites in Queensland have been sampled for the purpose of soil mapping with projects having a clear agricultural and edaphic bias. An appreciation of why existing sites were collected enables an understanding of the types of biases that exist in current sampling distributions. The analysis clearly shows that particular lithologies are sampled more intensively and there is a strong bias towards flatter areas that occupy the lower parts of the landscape.

The current purpose for sampling is to use the sites to produce soil attribute surfaces for use in environmental process modelling. This creates the need to capture the range of soil attribute values in the proportion that they occupy. To do this two assumptions are made 1) that a soil attribute will vary with soil type, and 2) that soils will vary with environmental variables relevant to soil formation. These assumptions are very difficult to validate, but are thought to be accurate for many situations and are better than any alternatives.

The type of analysis that was done in this study has effectively highlighted the bias that exists in the current data and could be used to develop a systematic sampling program to fill these gaps. Terrain was not used to subdivide the environment into physiographic units finer than lithology for the prioritisation of sampling because of the large gaps that needed filling first.

The sampling methodology that has been developed in this project has been very effective in achieving a strategic sampling distribution to assist in the production of robust soil attribute surfaces. The use of the analysis carried out in this study provides an understanding of the current sampling and assists in devising new sampling programs.

This site selection strategy, outlined in this report, to assist in the creation a range of reliable and accurate soil and landscape attributes highlights important considerations for devising sampling strategies. It provides a framework for further work and the continued development of a sampling distribution that captures the environmental variation of soil in a representative and proportional manner.

5.2 Soil and landscape attribute surfaces

The newly created soil and landscape attribute information system has provided a rich resource for the modelling and broadscale assessment of Queensland's natural resources. The method of producing soil and landscape attribute surfaces presented in this report uses a data driven approach to soil modelling. By using all suitable information from SALI, we have relied on the significant investment, of the Queensland Government and CSIRO, in describing and mapping Queensland's landscapes over the past 50 years.

The soil and landscape attribute information system has improved on prior iterations of broadscale interpretations of soil information and combined analyses of existing data. This system has improved on earlier attempts because of its use of a great range of data, a strong rule base to attribute derivation and estimation and creates a system which combines the best aspects of newly available technologies with the lessons learnt from the past.

The improvements of the current version of the attribute system have been underpinned by the great leap forward of SALI and the development of a nationally consistent framework for the estimation of soil and landscape attributes. Without the centralisation of SALI, and the associated effort to capture existing land resource information much of the effort in developing this system would have been wasted. With ASRIS providing a nationally consistent framework the development of attribute surfaces for Queensland will easily fit into national and interstate modelling tasks.

By using much of the existing knowledge available from the land resource data in SALI, the interpretation and estimation of attributes has been greatly enhanced. The ability to utilise expert knowledge has been important in producing surfaces that are reliable, updateable and accurate.

The ability to process large amounts of SALI data through the attribute system has provided a system where the attribute values for a polygon can be area weighted to produce more accurate results, while the use and storage of non-mapped information will allow users of the data to assess the variability inherent in the mapped units of the State's landscapes.

The soil and landscape attribute information system will be an important and on-going resource for the modelling, assessment and interpretation of land resources for Queensland for many years to come while its ability to be easily updated will ensure that the surfaces will reflect our best knowledge for that point in time.

6 References

Ahern, C.R., Weinand, M.M.G. & Isbell, R.F. 1992. *pH of Surface Soils, Queensland*, Department of Primary Industries, Brisbane, Australia.

Austin, M.P. & Heyligers, P.C. 1989. "Vegetation survey design for conservation: gradsect sampling of forests in north-eastern New South Wales", *Biological Conservation*, vol. 50, no. 1-4, pp. 13-32.

Beckett, P.H.T. & Webster, R. 1971. The development of a system of terrain evaluation over large areas. *R. Eng. J.* 85**:** 243-258.

Biggs, A.J.W., Kidston, E.M., Searle, R.D., Wilson, P.L., Slater, B.K., Grundy, M.J., Heiner, I.J., Smith, C.D. & Sweeney, C. 2000. "SALI - The soil and land information system for Queensland". In *Proceedings of the NZSSS/ASSSI Soil 2000 Conference.* Canterbury, New Zealand. NZSSS/ASSSI.

Brough, D.M. 2001. *Creation of Soil Attribute Surfaces for Landscape Salinity Hazard Assessment*, Queensland Department of Natural Resources and Mines, Brisbane.

Brough, D.M. 2003, *Creation of Soil Attribute Surfaces for Landscape Salinity Hazard Assessment - Version 2*, Queensland Department of Natural Resources and Mines, Brisbane.

Burrough, P.A. 1993. Soil variability: a late 20th century view. *Soils and fertilizers* 56(5): 529-562.

Burrough, P.A., Van Gaans, P.F.M. & Hootsmans, R. 1997. "Continuous classification in soil survey: Spatial correlation, confusion and boundaries", *Geoderma*, vol. 77, no. 2-4, pp. 115-135.

Chamberlain, T. 2006. Regolith and Physiography in Queensland: Technical report on regolith and physiographic mapping for the Queensland NAPSWQ regions. Queensland Department of Natural Resources, Mines and Water. Brisbane, Australia. QNRM06187.

Christian, C.S. & Stewart, G.A. 1953. *General Report on Survey of Katherine - Darwin Region, 1946,* CSIRO, Australia.

Christian, C.S. & Stewart, G.A. 1968. "Methodology of Integrated Surveys", *Aerial Surveys and Integrated Surveys, Proceesings Toulouse Conference 1964.* UNESCO, pp. 233.

Claridge, J. & Grundy, M.J. 2003. *Development of a Second Stage Sampling Strategy by Assessing the Representiveness of the Current Site Distribution*, Natural Resources & Mines, Brisbane, Australia.

Environment Australia 2000. *Revision of the Interim Biogeographic Regionalisation of Australia (IBRA) and the Development of Version 5.1. - Summary Report*. Department of Environment and Heritage, Canberra.

ESRI 2006. 19 May 2006-last update*, ESRI* [Homepage of ESRI], [Online]. Available: www.esri.com [2006, 05/26] .

Gallant, J.C. & Dowling, T.I. 2003. "A multiresolution index of valley bottom flatness for mapping depositional areas", *Water Resources Research,* vol. 39, no. 12.

Gessler, P.E., Moore, I.D., McKenzie, N.J. & Ryan, P.J. 1995. "Soil-landscape modelling and spatial prediction of soil attributes", *International Journal of Geographical Information Systems*, vol. 9, no. 4, pp. 421-432.

Grundy, M., Moss, J., Gordon, I., Pearce, B., Biggs, A., Searle, R. & Brough, D. In preparation. *Salinity Hazard Mapping in Queensland – A report on the development of the Salinity Hazard Map released for the priority action catchments of the National Action Plan for Salinity and Water Quality*. Queensland Department of Natural Resources, Mines and Water. Brisbane, Australia.

Gunn, R.H., Beattie, R.E., Reid, R.E. & van de Graff, R.H.M. 1988. *Australian Soil and Land Survey Handbook: Guidelines for Conducting Surveys,* Inkata Press, Melbourne, Australia.

Harms, B. & Main, A., 2006. *Key Reference Sites for Queensland – a compendium of soil and regolith reference sites*. Queensland Department of Natural Resources, Mines and Water. Brisbane, Australia. QNRM06197.

Henderson, B.L., Bui, E.N., Moran, C.J., Simon, D.A.P., Carlile, P. 2001. *'ASRIS: Continental-scale soil property predictions from point data'*, CSIRO Land and Water, Canberra, Australia.

Isbell, R.F. 1988. "Soil Classification" in *Australian Soil and Land Survey Handbook: Guidelines for Conducting Surveys*, eds. R.H. Gunn, R.E. Beattie, R.E. Reid & R.H.M. van de Graff, Inkata Press, Melbourne, Australia, pp. 20-37.

Isbell, R.F. 2002. *The Australian Soil Classification,* Revised edn, CSIRO Publishing, Collingwood, Victoria, Australia.

Johnston, R.M., Barry, S.J., Bleys, E., Bui, E.N., Moran, K.C.J., Simon, D.A.P., Carlile, P., McKenzie, N.J., Henderson, B.L., Chapman, G., Imhoff, M., Maschmedt, D., Howe, D., Grose, C., Schoknecht, N., Powell, B. & Grundy, M. 2003. "ASRIS: The database", *Australian Journal of Soil Research,* vol. 41, no. 6, pp. 1021-1036.

Kimball, R. & Ross, M. 2002. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd edn, Wiley.

Littleboy, M. 1997. *Spatial generalisation of biophysical simulation models for quantitative land evaluation: a case study for dryland wheat growning areas of Queensland*. PhD. Thesis. The University of Queensland, Brisbane, Australia.

Margules, C.R. & Stein, J.L. 1989. "Patterns in the distributions of species and the selection of nature reserves: an example from Eucalyptus forests in south-eastern New South Wales", *Biological Conservation*, vol. 50, no. 1-4, pp. 219-238.

McDonald, R.C. & Isbell, R.F. 1990. "Soil Profile" in *Australian Soil and Land Survey Field Handbook*, eds. R.C. McDonald, R.F. Isbell, J.G. Speight, J. Walker & M.S. Hopkins, 2nd edn, Inkata Press, Melbourne, Australia, pp. 103-152.

McDonald, R.C., Isbell, R.F., Speight, J.G., Walker, J. & Hopkins, M.S. 1990. *Australian Soil and Land Survey Field Handbook,* 2nd edn, Inkata Press, Melbourne, Australia.

McKenzie, N.J. & Austin, M.P. 1993. "A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation", *Geoderma*, vol. 57, no. 4, pp. 329-355.

McKenzie, N.J. & Ryan, P.J. 1999. "Spatial prediction of soil properties using environmental correlation", *Geoderma*, vol. 89, no. 1-2, pp. 67-94.

McKenzie, N.J., Cresswell, H.P., Ryan, P.J. & Grundy, M. 2000a. "Contemporary land resource survey requires improvements in direct soil measurement", *Communications in Soil Science and Plant Analysis,* vol. 31, no. 11-14, pp. 1553-1569.

McKenzie, N.J., Jacquier, D.W., Ashton, L.J. & Cresswell, H.P. 2000b. *Estimation of Soil Properties Using the Atlas of Australian Soils*, CSIRO Land and Water, Canberra, Australia.

McKenzie, N.J., Jacquier, D.W., Maschmedt, D.J., Griffin, E.A. & Brough, D.M. 2005. *Australian Soil Resource Information System Technical Specifications, Version 1.5,* Australian Collaborative Land Evaluation Program, Canberra.

McSweeney, K., Slater, B.K., Hammer, R.D., Bell, J.C., Gessler, P.E., Petersen, G.W., Amundson, R., Harden, J. & Singer, M. 1994. Towards a new framework for modeling the soil-landscape continuum. *Factors of soil formation: a fiftieth anniversary retrospective.* Madison, SSSA Special Publication No. 33.

Neldner, V.J., Crossley, D.C. & Cofinas, M. 1995. Using geographic information systems (GIS) to determine the adequacy of sampling in vegetation surveys. *Biological Conservation* 73: 1-17

Nicholls, A.O. 1989. "How to make biological surveys go further with generalised linear models", *Biological Conservation*, vol. 50, no. 1-4, pp. 51-75.

NLWRA. 2002. '*Natural resource economics*.' National Land and Water Resources Audit, Canberra.

Northcote, K.H. 1979. *A Factual Key for the Recognition of Australian Soils,* 4th edn, Relim, Adelaide, Australia.

Northcote, K.H., Beckmann, G.G., Bettenay, E., Churchwood, H.M., van Dijk, D.C., Dimmock, G.M., Hubble, G.D., Isbell, R.F., McArthur, W.M., Murtha, G.G., Nicolls, K.D., Paton, T.R., Thompson, C.H., Webb, A.A. & Wright, M.J. 1968. *Atlas of Australian Soils, Sheets 1 to 10, with explanatory data*, CSIRO and Melbourne University Press, Melbourne, Australia.

O'Connell, D.A., Ryan, P.J., McKenzie, N.J. & Ringrose-Voase, A.J. 2000. Quantitative site and soil descriptors to improve the utility of forest soil surveys. *Forest Ecology and Management* 138: 107-122.

Oracle 2006. 26 May 2006-last update, *Oracle Corporation* [Homepage of Oracle Corporation], [Online]. Available: www.oracle.com [2006, 05/26] .

Ryan, P.J., McKenzie, N.J., O'Connell, D., Loughhead, A.N., Leppert, P.M., Jacquier, D. & Ashton, L. 2000. "Integrating forest soils information across scales: Spatial prediction of soil properties under Australian forests", *Forest Ecology and Management*, vol. 138, no. 1-3, pp. 139-157.

Slater, B.K. & Grundy, M.J. 1999. Enhanced Resource Assessment: Integration of Innovative Technologies by Agency Soil Scientists in Australia. *The Second Approximation International Conference on Soil Resources: Their Inventory, Analysis and Interpretation in the 21st Century*, Minneapolis, Minnesota USA.

Smith, D.M. 2000. *Analysis of soil carbon levels in the IBRA regions of Qld - Technical Report*, Queensland Department of Natural Resources, Brisbane, Australia.

Smith, D.M. & Brough, D.M. 2006, *Modelling the land surface: Technical report on the creation of digital elevation models for Queensland*. Queensland Department of Natural Resources, Mines and Water. Brisbane, Australia. QNRM06196.

Speight, J.G. 1988. "Land Classification" in *Australian Soil and Land Survey Handbook: Guidelines for Conducting Surveys*, eds. R.H. Gunn, R.E. Beattie, R.E. Reid & R.H.M. van de Graff, Inkata Press, Melbourne, Australia, pp. 38-59.

Thwaites, R.N. & Slater, B.K. 2000. "Soil-landscape resource assessment for plantations - A conceptual framework towards an explicit multi-scale approach", *Forest Ecology and Management*, vol. 138, no. 1-3, pp. 123-138.

Witte, C., Van den Berg, D., Rowland, T., O'Donnell, T., Denham, R., Pitt, G. & Simpson, J., 2006. Mapping Land Use in Queensland: Technical report on the 1999 land use map for Queensland. Queensland Department of Natural Resources, Mines and Water. Brisbane, Australia. QNRM06185.

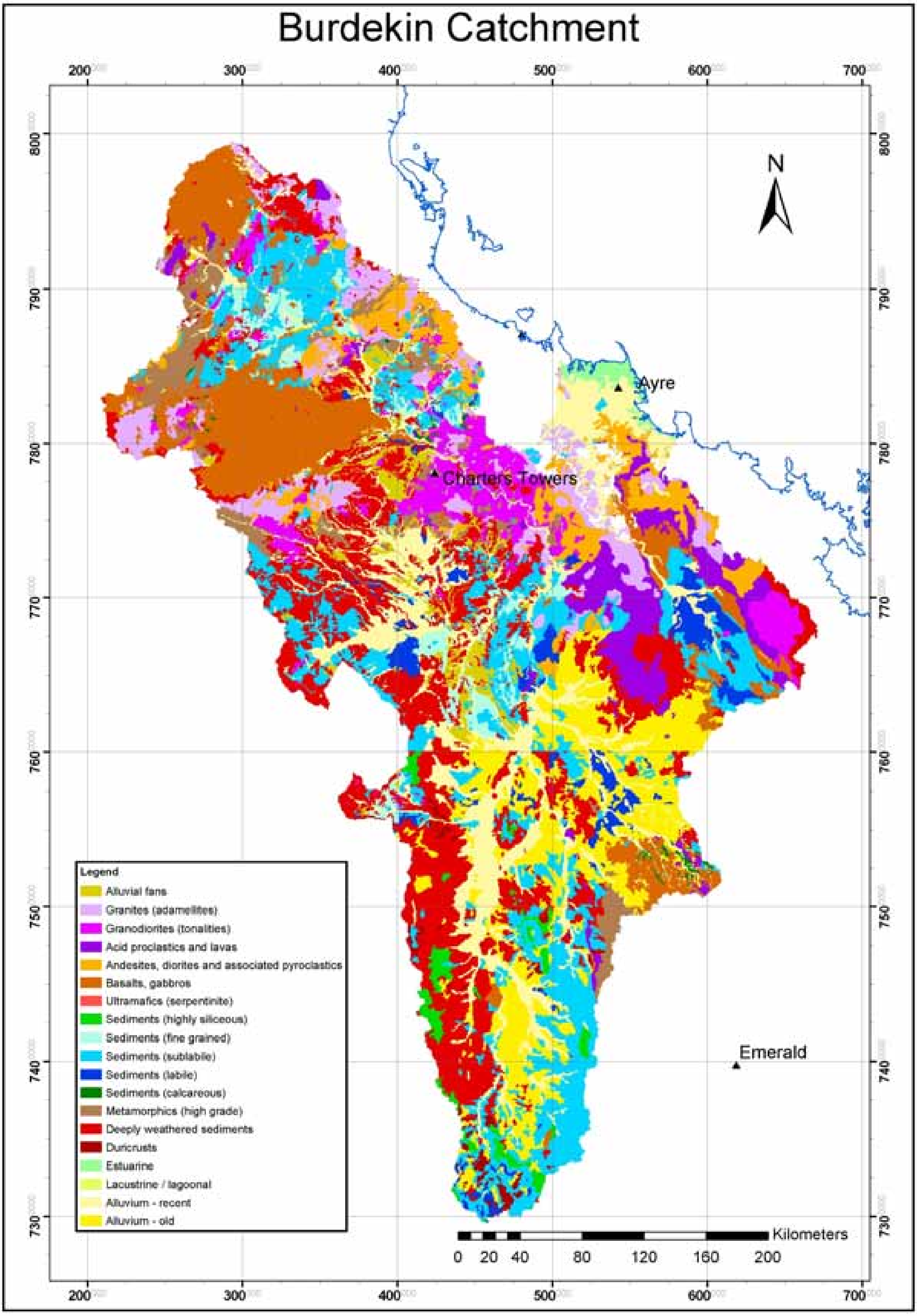
Zhu, A.-., Band, L., Vertessy, R. & Dutton, B. 1997. "Derivation of soil properties using a soil land inference model (SoLIM)", *Soil Science Society of America Journal*, vol. 61, no. 2, pp. 523-533.

Appendix 1

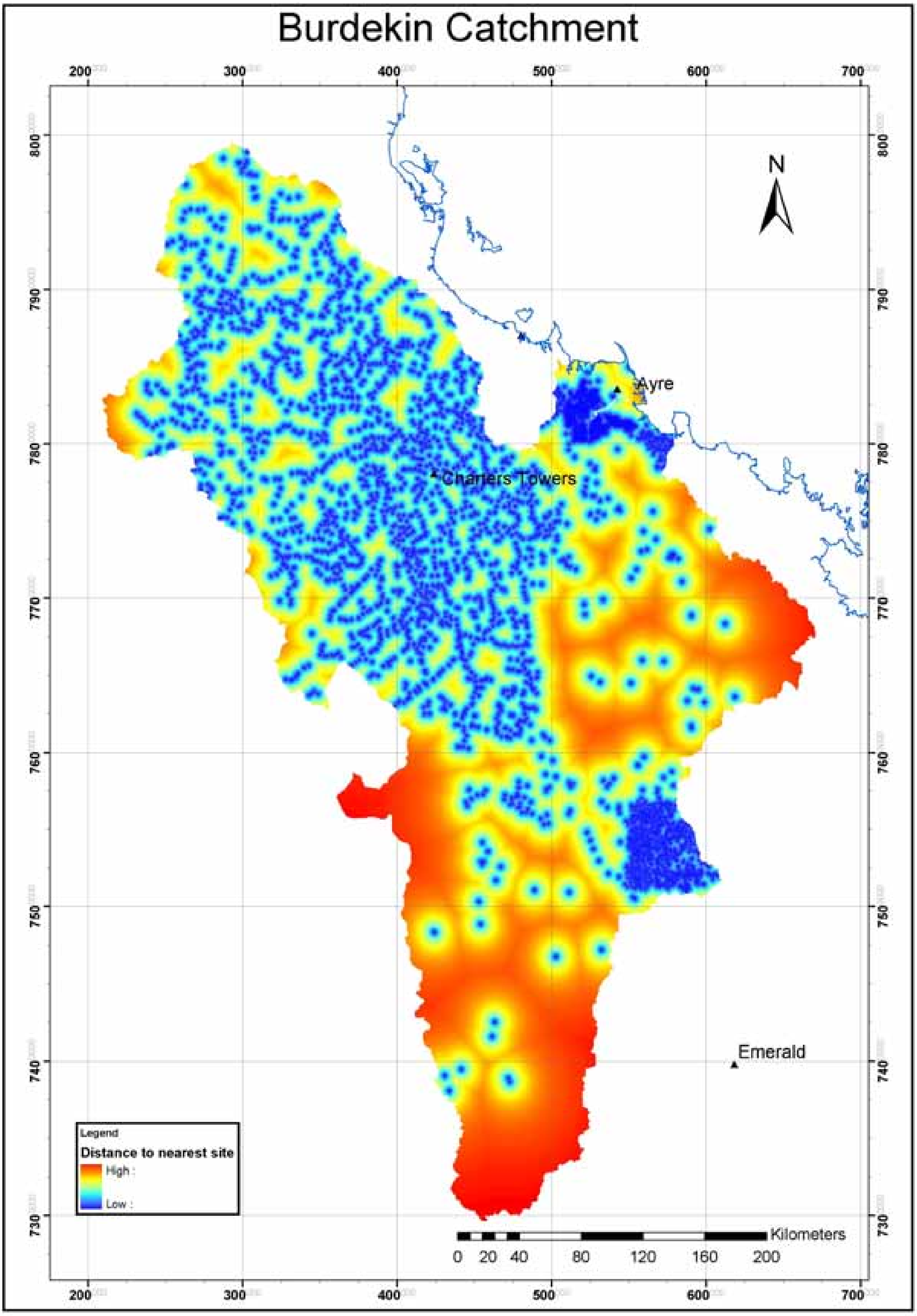
Example Maps from the site selection strategy

Map 1. Pedolith classes in the Burdekin Catchment.
Map 2. Site intensity in geographic space using distance to nearest site.
Map 3. Site intensity in geographic space using a site density function.
Map 4. Approximate mapping scale according to sampling intensity.
Map 5. Sampling intensity calculated on the number of sites within a particular pedolith class with a 10km search radius.
Map 6. Sampling priorities developed according to sampling of geographic and environmental space.

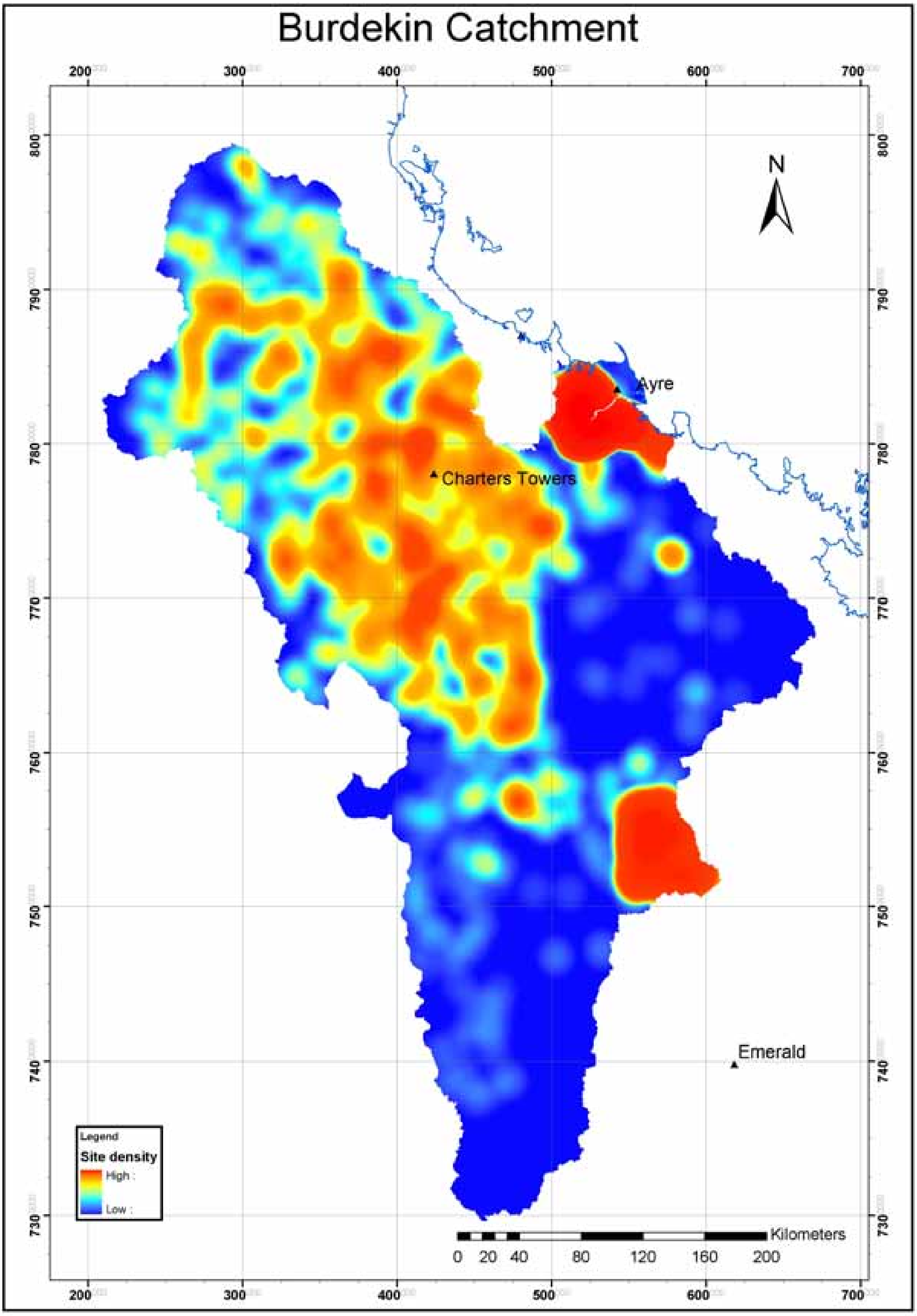
Map 1. Pedolith classes in the Burdekin Catchment.



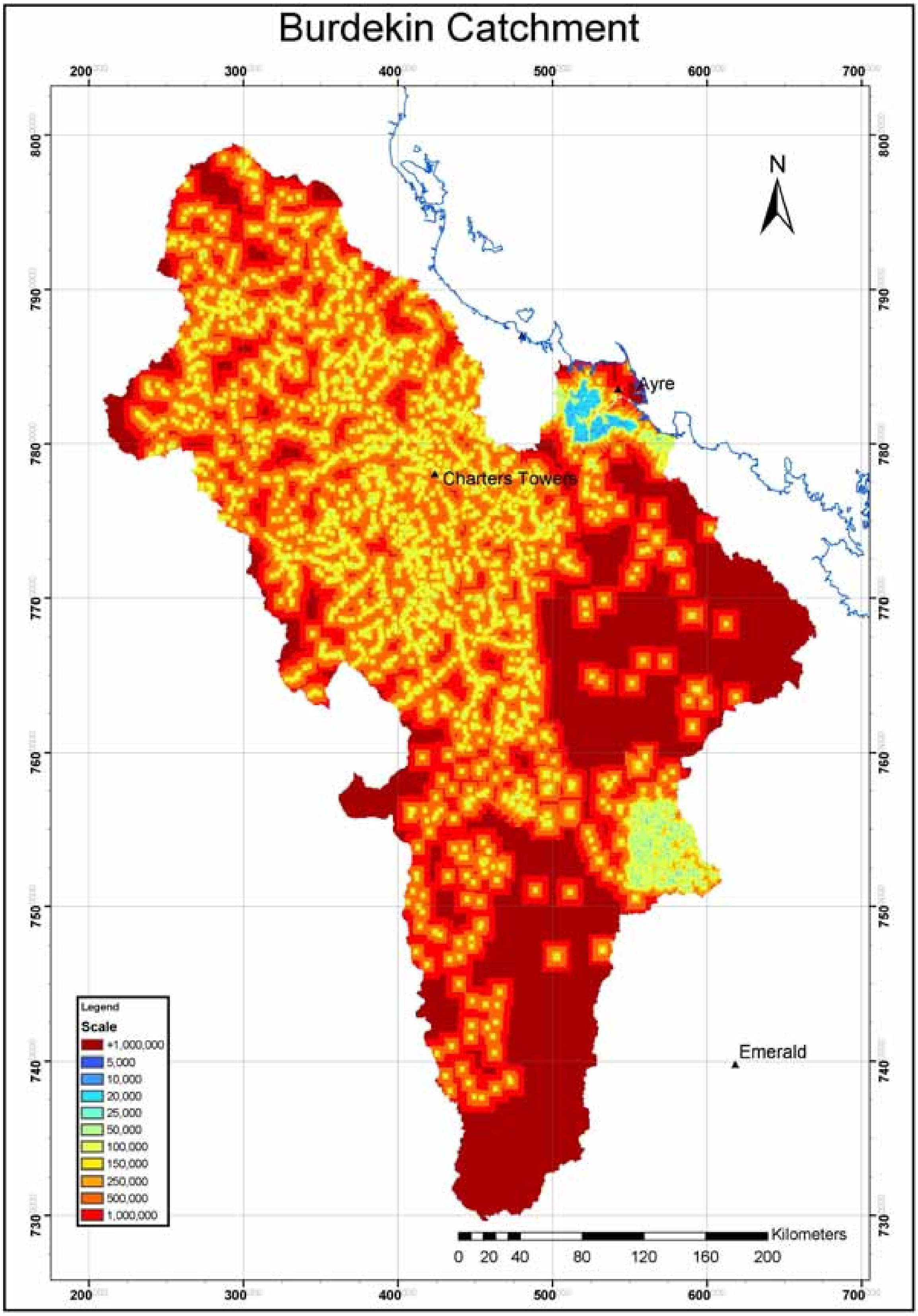
Map 2. Site intensity in geographic space using distance to nearest site.



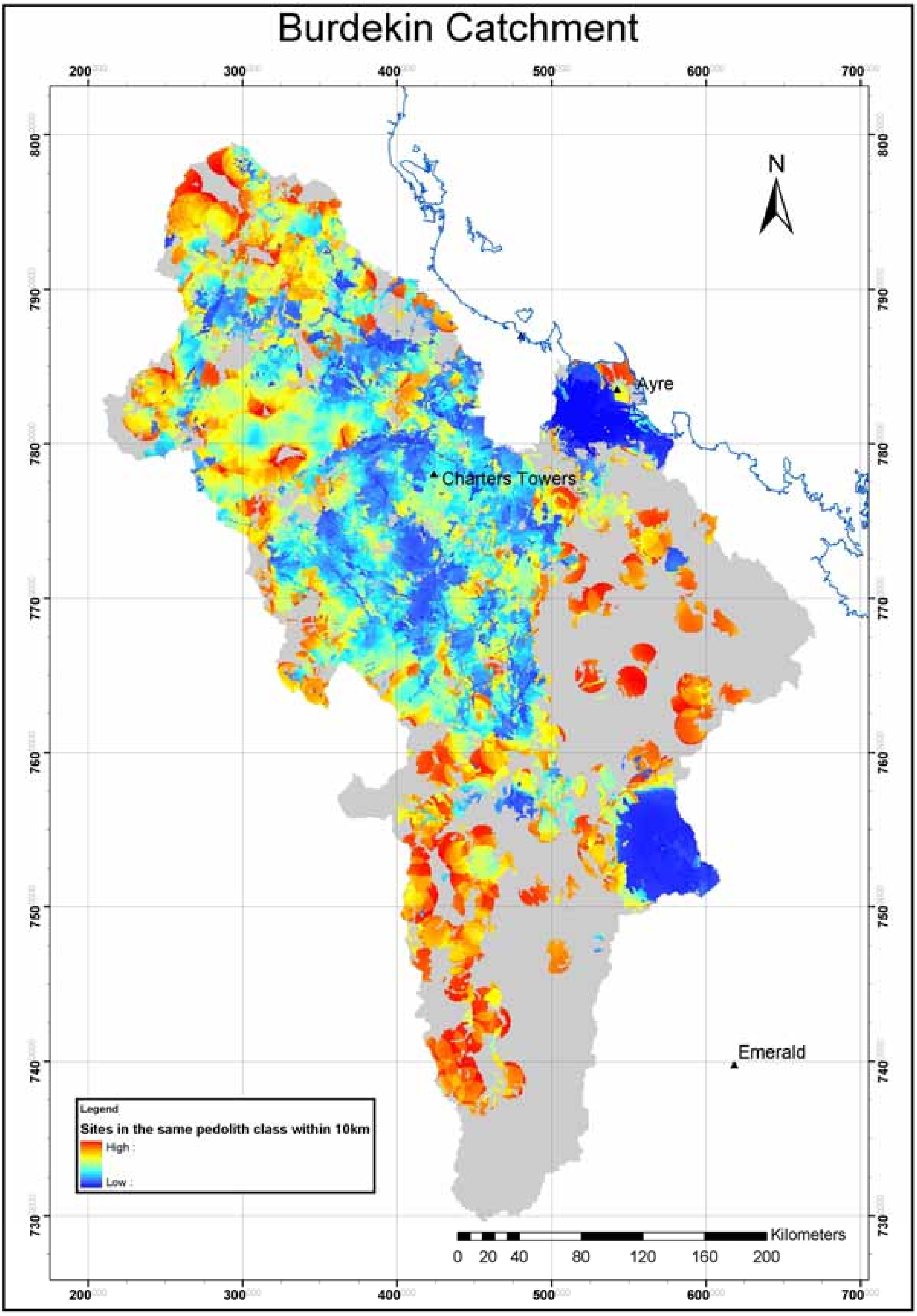
Map 3. Site intensity in geographic space using a site density function.



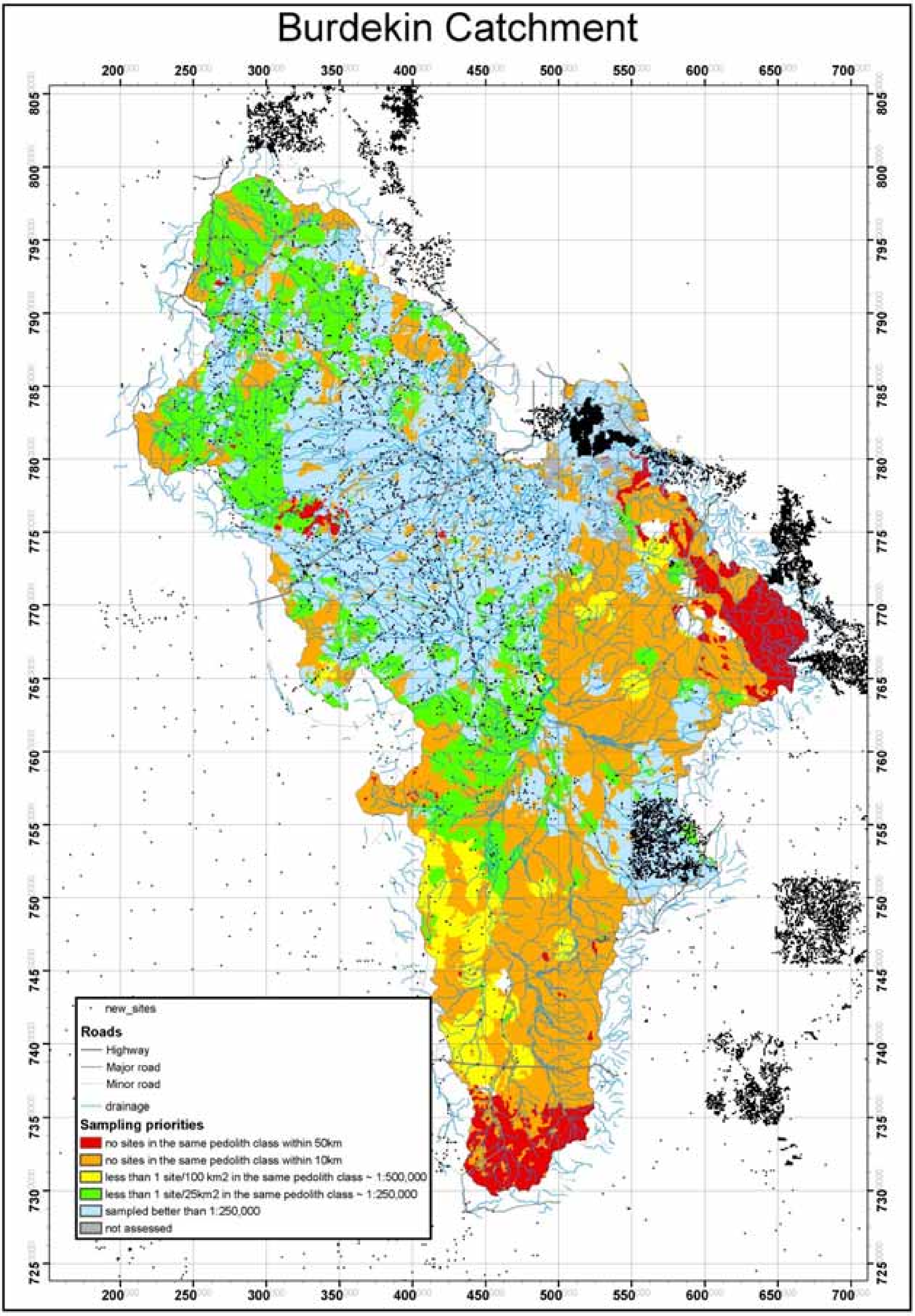
Map 4. Approximate mapping scale according to sampling intensity.



Map 5. Sampling intensity calculated on the number of sites within a particular pedolith class with a 10km search radius.



Map 6. Sampling priorities developed according to sampling of geographic and environmental space.



Appendix 2

Soil and landscape attribute information system documentation.

Rules and estimation methods

Entity relationship diagrams

Table 15. Derivation rule combination and descriptions.

Combination Name	Description
text_cs1	Texture control section 1 rule combination
text_cs2	Texture control section 2 rule combination
text_cs3	Texture control section 3 rule combination
text_cs4	Texture control section 4 rule combination
text_cs5	Texture control section 5 rule combination
bulk_cs3	Bulk density control section 3 combination
ph_cs1	pH control section 1 combination
depth_a1	Depth of A1 horizon
depth_a_horizon	Total thickness of A horizon
depth_to_b2	Depth to B2 horizon
depth_regolith	Depth to base of Regolith
ksat_cs1	Ksat control section 1 combination

Table 16. Derivation rules and rule combinations.

Combination Name	Rule No			
	1	2	3	4
bulk_cs3	bulk_cs3_rule1	bulk_cs3_rule2		
depth_a_horizon	depth_a_horizon			
depth_a1	depth_a1			
depth_regolith	depth_regolith			
depth_to_b2	depth_to_b2			
ksat_cs1	ksat_cs1_rule1	text_cs1_rule1	ksat_cs1_rule3	ksat_cs1_rule4
ph_cs1	text_cs1_rule1	ph_cs1_rule2	ph_cs1_rule3	
text_cs1	text_cs1_rule1	text_cs1_rule2	text_cs1_rule3	text_cs1_rule4
text_cs2	text_cs2_rule1			
text_cs3	text_cs3_rule1	text_cs3_rule2		
text_cs4	text_cs4_rule1			
text_cs5	text_cs5_rule1			

Table 17. Parameters and derivation rule combinations.

Parameter	Control Section				
	1	2	3	4	5
15_BAR	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
ADMC	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
AGG_STAB	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
AWC_FC	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
AWC_LOWER_DEPTH	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
AWC_UPPER_DEPTH	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
AWC_WP	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
BULK_DENSITY	text_cs1	text_cs2	bulk_cs3	text_cs4	text_cs5
CEC	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
CEC_ALC	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
CEC_EFFECTIVE	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
CHLORIDE	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
CLAY	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
COARSE_FRAGMENTS	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
DEPTH_A	depth_a_horizon				
DEPTH_A1	depth_a1				
DEPTH_B2	depth_to_b2				
DEPTH_LOWER	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
DEPTH_REGOLITH	depth_regolith				
DEPTH_UPPER	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
DRAINAGE	drain_perm				
EC	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
ESP	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
EX_CA	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
EX_K	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
EX_MG	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
EX_NA	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
KS	ksat_cs1	text_cs2	text_cs3	text_cs4	text_cs5
KS_CLASS	ksat_cs1	text_cs2	text_cs3	text_cs4	text_cs5
ORGANIC_CARBON	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
PEDALITY_SIZE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
PEDALITY_TYPE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
PERMEABILITY	drain_perm				
PH	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
SAND_COARSE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
SAND_FINE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
SILT	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
TEXTURE_CODE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
TEXTURE_GRADE	text_cs1	text_cs2	text_cs3	text_cs4	text_cs5
TOTAL_N	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5
TOTAL_P	ph_cs1	text_cs2	text_cs3	text_cs4	text_cs5

Table 18. Derivation rules and descriptions.

Function Name	Description
text_cs1_rule1	If the surface layer is a peat, refers to the 7 classes of organic materials defined by McDonald and Isbell (1990).
text_cs1_rule2	If there is a single A1 horizon without subdivisions (e.g. A11, A12 etc.), then refers to the A1 horizon. Ignores any possible overlying surface horizons.
text_cs1_rule3	If there are subdivisions within the A1 horizon, refers to the thickest A horizon layer within the top 0.20 m of the soil profile (the upper layer is used if thicknesses are equal). If the surface layer is an O horizon, refers to an A horizon as defined accord with the above criteria.
text_cs1_rule4	If the surface layer is an O horizon and there is no underlying A1 horizon, refers to the thickest layer in the 0.20 m directly beneath the O horizon.
text_cs2_rule1	If the surface layer is not an A1 horizon (e.g. O horizon) and there is no underlying A horizon, the attribute is recorded as missing, otherwise usually refers to the lower portion of the A horizon that is below layer 1, if not below layer 1 then record as missing
text_cs3_rule1	If a B horizon is present, refers to the uppermost B horizon (usually B1 or B21).
text_cs3_rule2	If no B horizon is present and the sequence consists of an AC profile, refers to the major part of the materials in the 0.20m directly below the A horizon.
text_cs4_rule1	Refers to the B horizon below the Layer 3 horizon with the maximum field texture if muliple horizons have the same texture it uses the uppermost horizon, otherwise recorded as missing
text_cs5_rule1	Refers to the lower 0.1m of the horizon above 2m, or above a pan or above a C or R horizon, and is below the layer 4 horizon, otherwise recorded as missing.
bulk_cs3_rule1	If a B2 horizon is present, refers to the max density in the upper 0.2m of the B2 horizon, or the major part of the B2 if less than 0.2m thick.
bulk_cs3_rule2	If no B horizon is present and the sequence consists of an AC profile, refers to the max value of the materials in the 0.20m directly below the A horizon.
ph_cs1_rule2	Refers to the upper 0.05m of the surface A1 horizon or the whole A1 if less than 0.05m thick.
ph_cs1_rule3	If the surface layer is an O horizon refers to the 0.05m directly beneath the O horizon.
depth_a1	The depth of the A1 horizon, if A1 is not present then recorded as missing.
depth_a_horizon	Total thickness of the A horizon, if A horizon not present then recorded as missing.
depth_to_b2	Depth from the land surface to the top of the B2 horizon, if there is no B2 horizon then recorded as missing.
depth_regolith	Depth to the base of the regolith.
ksat_cs1_rule1	If the soil has a surface crust or surface flake, the estimate is for the upper 0.01m of the surface horizon
ksat_cs1_rule3	Refers to the major part of the A1 horizon within the top 0.2m, or the whole A1 horizon if thinner than 0.05m, where there is no O horizon.
ksat_cs1_rule4	If the surface is an O horizon then estimate is from the upper 0.05m of the underlying material, normally an A1 horizon.

Table 19. Parameters and estimation rules.

Parameter	Estimation Rule Method								
	0	1	2	3	4	5	6	7	8
15_BAR		text_est1	text_est2	text_est3	text_est4	text_est5			
ADMC		text_est1	text_est2	text_est3	text_est4	text_est5			
ASC_ORD	expert value	text_est1	text_est2	text_est3	text_est4				
CEC		text_est1	text_est2	text_est3	text_est4	text_est5			
CHLORIDE		text_est1	text_est2	text_est3	text_est4	text_est5			
CLAY		text_est1	text_est2	text_est3	clay_est4	text_est4	clay_est6	clay_est7	text_est5
COARSE_FRAGMENTS		text_est1	text_est2	text_est3	text_est4	text_est5			
DEPTH_LOWER		text_est1	text_est2	text_est3	text_est4	text_est5			
DEPTH_UPPER		text_est1	text_est2	text_est3	text_est4	text_est5			
DRAINAGE	expert value	text_est1	text_est2	text_est3	text_est4	text_est5			
EC		text_est1	text_est2	text_est3	text_est4	text_est5			
ESP		text_est1	text_est2	text_est3	text_est4	text_est5			
EX_CA		text_est1	text_est2	text_est3	text_est4	text_est5			
EX_K		text_est1	text_est2	text_est3	text_est4	text_est5			
EX_MG		text_est1	text_est2	text_est3	text_est4	text_est5			
EX_NA		text_est1	text_est2	text_est3	text_est4	text_est5			
ORGANIC_CARBON		text_est1	text_est2	text_est3	text_est4	text_est5			
PEDALITY_SIZE		text_est1	text_est2	text_est3	text_est4	text_est5			
PEDALITY_TYPE		text_est1	text_est2	text_est3	text_est4	text_est5			
PERMEABILITY	expert value	text_est1	text_est2	text_est3	text_est4	text_est5			
PH		text_est1	text_est2	text_est3	text_est4	text_est5			
SAND_COARSE		text_est1	text_est2	text_est3	clay_est4	text_est4	clay_est6	clay_est7	text_est5
SAND_FINE		text_est1	text_est2	text_est3	clay_est4	text_est4	clay_est6	clay_est7	text_est5
SILT		text_est1	text_est2	text_est3	clay_est4	text_est4	clay_est6	clay_est7	text_est5
TEXTURE_CODE		text_est1	text_est2	text_est3	text_est4	text_est5			
TEXTURE_GRADE		text_est1	text_est2	text_est3	text_est4	text_est5			
TOTAL_N		text_est1	text_est2	text_est3	text_est4	text_est5			
TOTAL_P		text_est1	text_est2	text_est3	text_est4	text_est5			

Table 20. Estimation rules.

Estimation Rule Name	Description
text_est1	Estimate based on an replicated measurement in the land unit tract
text_est2	Estimate based on an un-replicated measurement in the land unit tract
text_est3	Estimate based on direct measurements of similar soils in the same land unit type (e.g., modal profiles)
text_est4	Estimate based on direct measurements of similar soils in the region or project area
text_est5	Estimate based on experience with similar soils (e.g., same taxa in the Australian Soil Classification but from other regions)
clay_est4	Estimate is based on field textures from representative soil profiles in the land-unit tract
clay_est6	Estimate is based on field textures from soil profiles in the land-unit tract
clay_est7	Estimate is based on field textures from similar soils in the project area
expert value	Estimate is based on actual value recorded for the feature

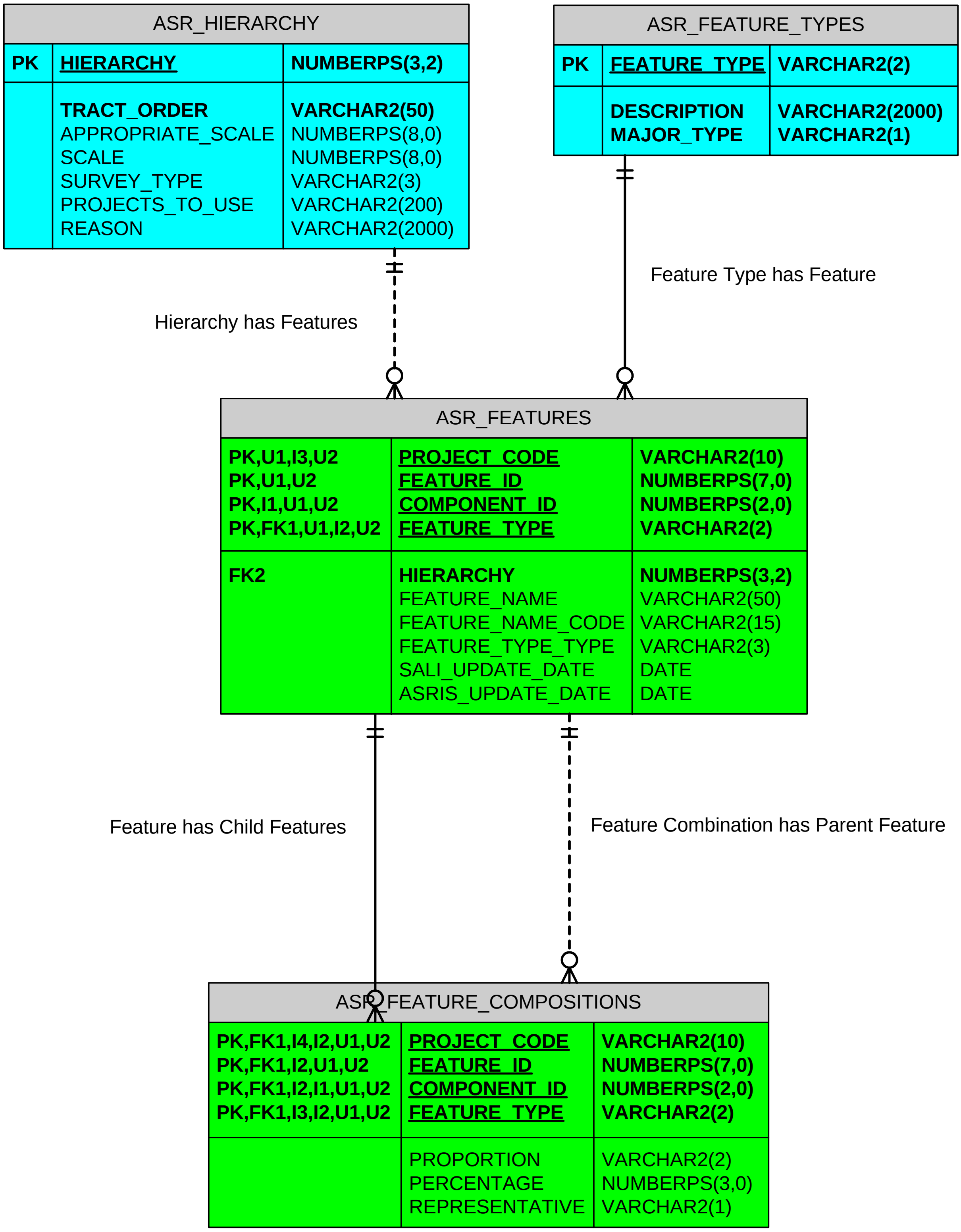


Figure 7. Entity relationship (ER) diagram for the feature tables.

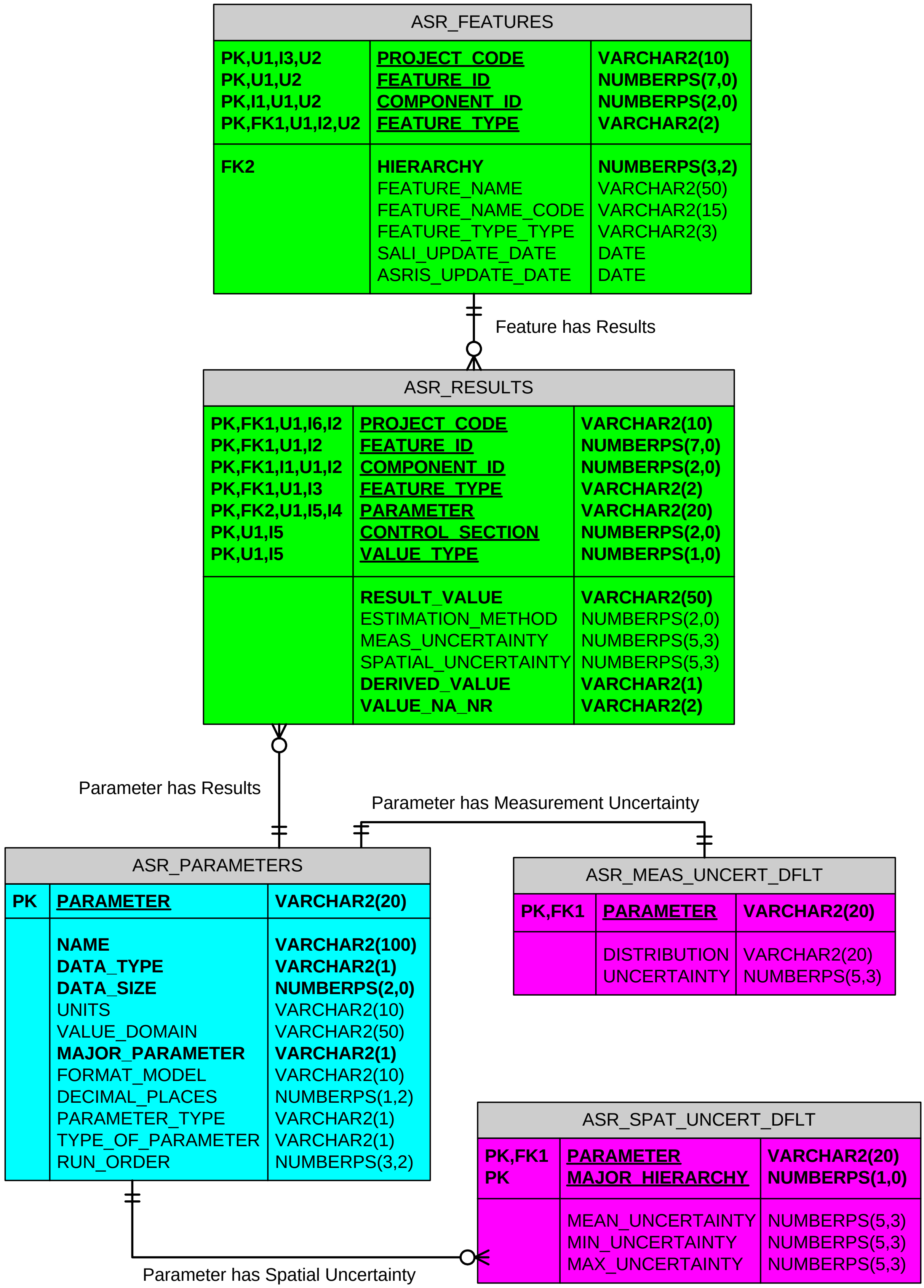


Figure 8. Entity relationship (ER) diagram for the result and parameter tables.

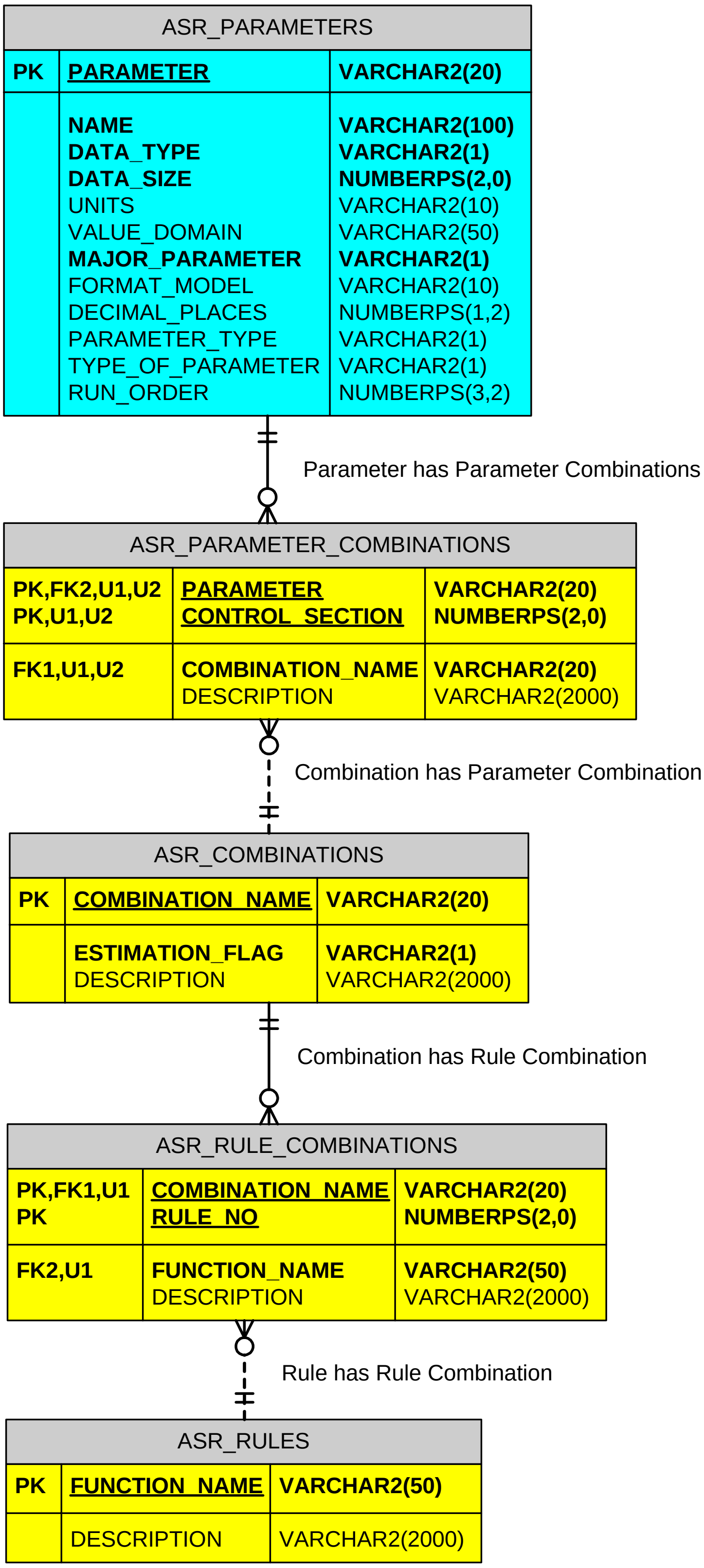


Figure 9. Entity relationship (ER) diagram for the rule tables.