National Soil Database utility: a review and process for development

Envirolink 1072-MLDC73

Landcare Research
Manaaki Whenua

National Soil Database utility: a review and process for development

Allan Hewitt, James Barringer, John Claydon, Stephen Campbell, Sam Carrick, Scott Fraser, Colin Gray, Ian Lynn, David Medyckyj-Scott, Alistair Ritchie, Trevor Webb

Landcare Research

Envirolink Medium Advice Grant 1072-MLDC73

Prepared for:

Marlborough District Council

15 Seymour Street
PO Box 443
Blenheim 7240
New Zealand

June 2012

Landcare Research, Gerald street, PO Box 40, Lincoln 7640, New Zealand, Ph +64 3 321 9999,Fax +64 3 321 9998, www.landcareresearch.co.nz

Reviewed by:	*Approved for release by:*
Sharn Hainsworth Researcher Landcare Research	Alison Collins Portfolio Leader Characterising Land Resources
Landcare Research Contract Report:	LC 1033

Disclaimer

This report has been prepared by Landcare Research for regional councils. If used by other parties, no warranty or representation is given as to its accuracy and no liability is accepted for loss or damage arising directly or indirectly from reliance on the information in it.

ISO 14001

Contents

1 Introduction

1.1 What is the National Soils Database?

The National Soils Database (NSD) is a crucial part of our soil data legacy in New Zealand. Our legacy of soil maps and reports and the new information products delivered in S-map Online and the LRIS Portal[1] are the most visible manifestations of our national soil dataset but the NSD sits in the background, a less visible but a crucial foundation underpinning all other datasets. It comprises the profile data collected in soil pits scattered throughout New Zealand that are the record of almost all we know of the soil chemistry and soil physical properties of our soils. From it we have derived our soil classifications, interpretations of our soil maps, land management models, and our understanding of how soil properties vary with geology, rainfall, vegetation, topography, and land management across the major gradients of the New Zealand landscape. It is the fundamental dataset that underpins our soil knowledge of New Zealand.

The NSD comprises the dataset for the soil at a site, and there are more than 1500 sites in New Zealand (not counting others from New Zealand awaiting data entry, the Pacific Islands, and Antarctica). At each site the soil profile was described from soil exposed in a soil pit (Figure 1) following the methodolgy of either *Soil Survey Method* (Taylor & Pohlen 1979) or *Soil Description Handbook* (Milne et al. 1995). Samples were collected from all or some of the soil horizons, and analysed either in the Landcare Research Environmental Chemistry Laboratory or previously by the DSIR Soil Bureau Laboratory. All profiles have at least soil chemical analyses for a number of soil horizons, and few soils have the complete set of moisture retention, mineralogy, XRF analyses, and particle size (see NSD views in Section 3.3). The NSD itself is a set of read-only tables of data (Figure 2) that have undergone a structured check and authorisisation process. Section 3.3 provides a convenient view of these data that is searchable by site.

1.2 Potted history of the NSD

Sites representative of soil series or soils with special characteristics were described and sampled since the beginnings of soil survey in New Zealand. Data were recorded on data cards and files maintained by the DSIR Soil Bureau.

In the mid-1980s data on cards were entered into computer storage using a VAX Datatrieve database. Earlier cards were not entered into Datatrieve because analytical methods had not been standardised. Only data with code numbers over SB08000 were captured electronically. The analytical lab data was transferred into the database both manually and electronically. The soil description data, however, was transferred using student labour to type in data from

[1] (http://smap.landcareresearch.co.nz/home)

(http://lris.scinfo.org.nz/#/layers/global/oceania/new-zealand/)

file cards. This resulted in incomplete data capture, and transcription errors were introduced that persist to the present day.

Figure 1 Sam Carrick describing the morphology of a soil profile exposed in a pit prior to sampling the soil horizons, at a site in Southland

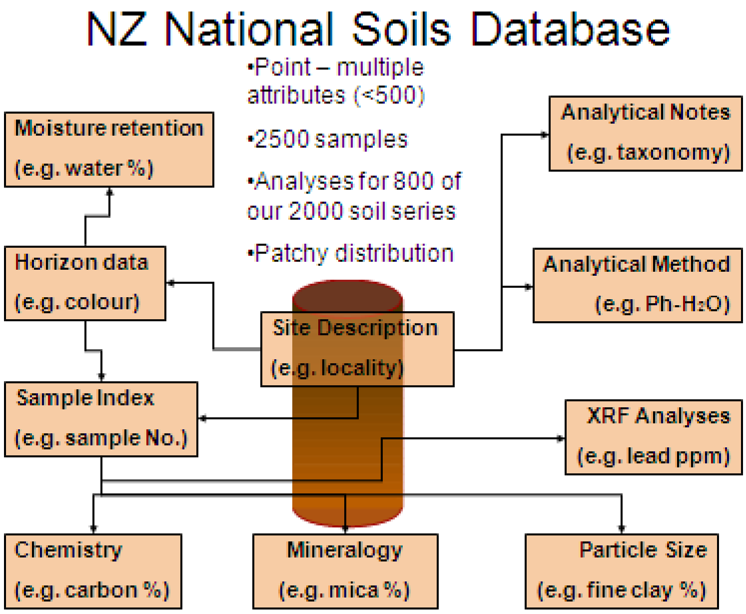


Figure 2 Data tables comprising the NSD (diagram Peter Newsome).

Prior to the 1990s the wider database system included a staging read-write accessible TEMP database where all data was held prior to final transfer to the NSD. The transfer to the NSD occurred after checking and formal approval by the chief soil correlator and the chemistry, physics, mineralogy and X-ray Fluorescence lab managers. The wider database system at that time included data input and export scripts with pre-formatted data reports.

The initial intent for the NSD was that it would only comprise full profiles, with complete descriptions and a full core analytical dataset qualified for entry to the NSD. This policy was revised when sites with part profiles and incomplete analytical data were added.

In the early 1990s the Datatrieve Database tables were migrated to a Paradox database platform. At this stage further errors (e.g. rounding errors) were inadvertently introduced due to the properties of Paradox. The majority of the functions and features of the wider database system, including the approval system, was not transferred nor recreated.

The platform was subsequently upgraded again by transferring the dataset into SQL Server. Around this time many problems with the data were identified and rectified, and there has been ongoing maintenance and repair. From the mid-1980s there was an overall substantial drop in funding for soil science, including funding for the maintenance and development of land resource databases and information. Datasets that were generated by research since the mid-1990s were no longer incorporated into the NSD but remained as accessory temporary datasets. This reflected not only poor resourcing, but also technical problems identified in the NSD that needed to be rectified before adding new data. Again lack of funding stalled progress.

A turning point was the recognition of the Land Resource Information System (LRIS) by government as one of the nationally significant databases that should be maintained. Some funding was reallocated to databases followed this decision, and was further secured by giving the funding long-term status as 'Backbone Funding'. The calls on this funding were large and only a modest sum could be allocated to NSD development.

1.3 Why this report now?

A resurgence of demand for quality soil information is driven by the need to better underpin the land management decisions required to achieve the important national outcomes of water quality, and water quantity, greenhouse gasses, and the better use of the nation's soil, land and water resources to meet production and sustainability goals. These national outcomes are propelled by domestic law, international commitments, the need to secure trade access, and to raise economic performance.

The NSD in its present state cannot provide the needed support and requires development. It is time to reverse the withdrawal of funding experienced in the 1980s and develop a database that can meet the new national imperatives.

2 Objectives

The aims of this project are to:

1. Review the requirements of the National Soils Database by Regional and District Councils including Marlborough District Council – the council that has championed this report

2. Review the current status of the NSD

3. Based on 1 and 2, scope out a programme to update the database that includes a global scan for systems and standards, a review of appropriate data modelling systems, an assessment of the existing database's ability to integrate with other databases, and its platform requirements

3 NSD status

3.1 Current access to the NSD

The primary legacy NSD dataset comprises the yellow cards of field-collected data and white cards of laboratory-derived data. This is held at Landcare Research in Palmerston North. The cards are currently being scanned with a third of them completed. This primary dataset is not generally available in its physical form but the data is accessible by the following routes.

1. The data on yellow and white cards, originally entered into the Datatrieve database, are available through the Landcare Research network as tables that replicate the look of the yellow and white cards. Examples are shown below (Section 3.3) for the Wairau silt loam site SB10084.

2. This same data is accessible via the Landcare Research network from the SQL Server database to query and download datasets for data analysis.

3. The Paradox version of the NSD was provided on a CD and distributed under licence to many users.

4. NSD data is publically available on the following website. It provides a subset of the more commonly used soil attributes.

   http://soils.landcareresearch.co.nz/contents/SoilData_NSD_About.aspx?currentPage=SoilData_NSD&menuItem=SoilData

   It is expected that this portal will be terminated when a new delivery system is implemented.

3.2 The National Soils Archive (NSA) collection

The NSA is a physical collection of samples of nationally significant soils dating from the 1930s. It is housed in Palmerston North in a purpose-designed storage shed, with air-dry 2-mm-sieved soils retained in glass jars. The NSA catalogue is retained by the Lab Manager, Environmental Chemistry Lab, Landcare Research, Palmerston North.

Other physical soil archives associated with soil datasets also exist nationally. Protocols and delegated responsibilities for all archives need to be improved. It is imperative to establish a protocol by which subsamples may be taken from an archived soil sample for research purposes. The option to centralise all archived soils into the Palmerston North NSA should also be considered.

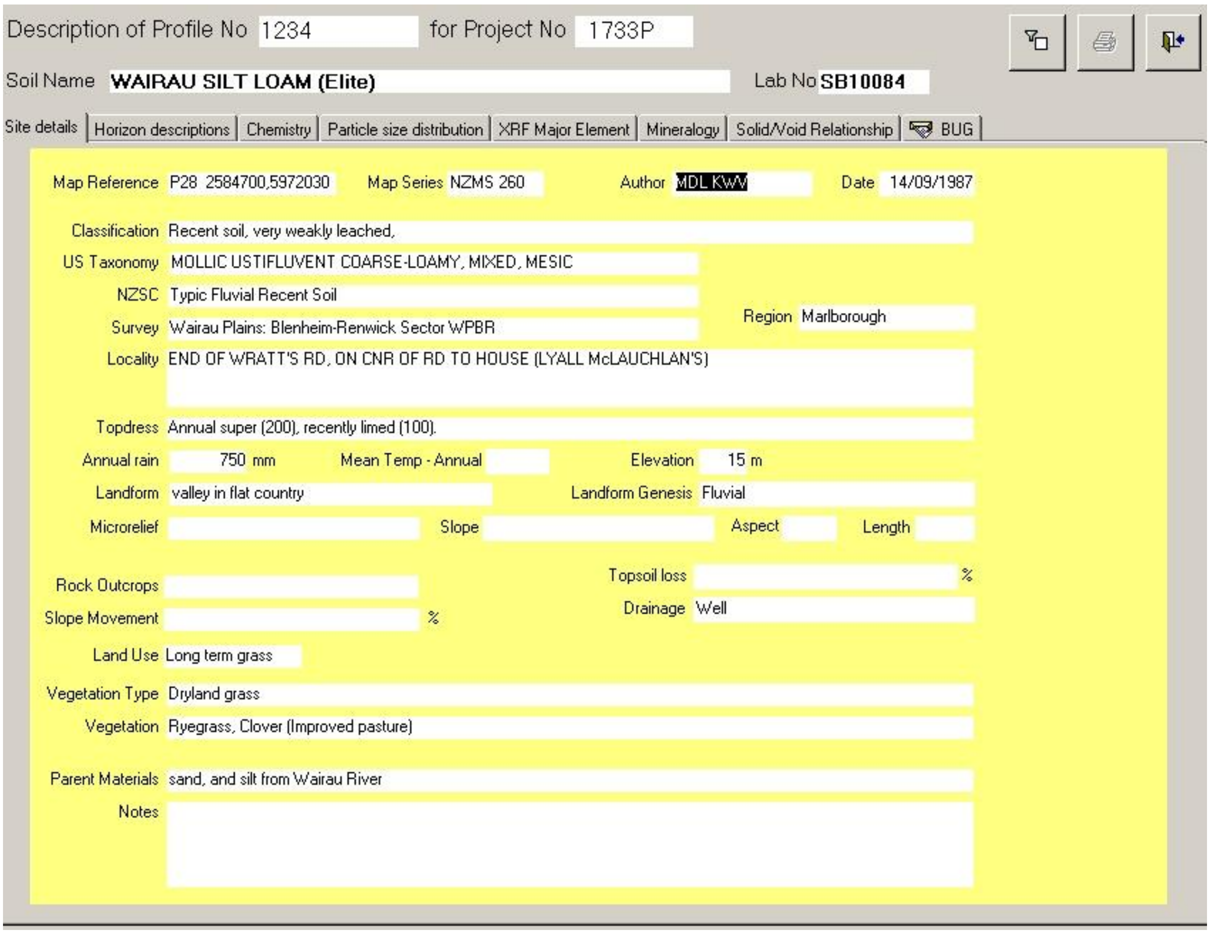
The connection between the database and sample archive can enable powerful analyses. For example in recent research by Pierre Roudier and Carolyn Hedley about 2000 samples were non-destructively sampled and scanned to obtain near-infrared spectra (NIR) to derive an NIR spectral library. This will become an additional resource for the NSD. One proposed application is to build a model to predict soil carbon in the field using a portable NIR scanner.

3.3 Core dataset description

The important data are contained in a series of tables within the NSD. Screen shots of the NSD Viewer are shown on the following pages. These are available for each NSD site through the Landcare Research network.

Sites

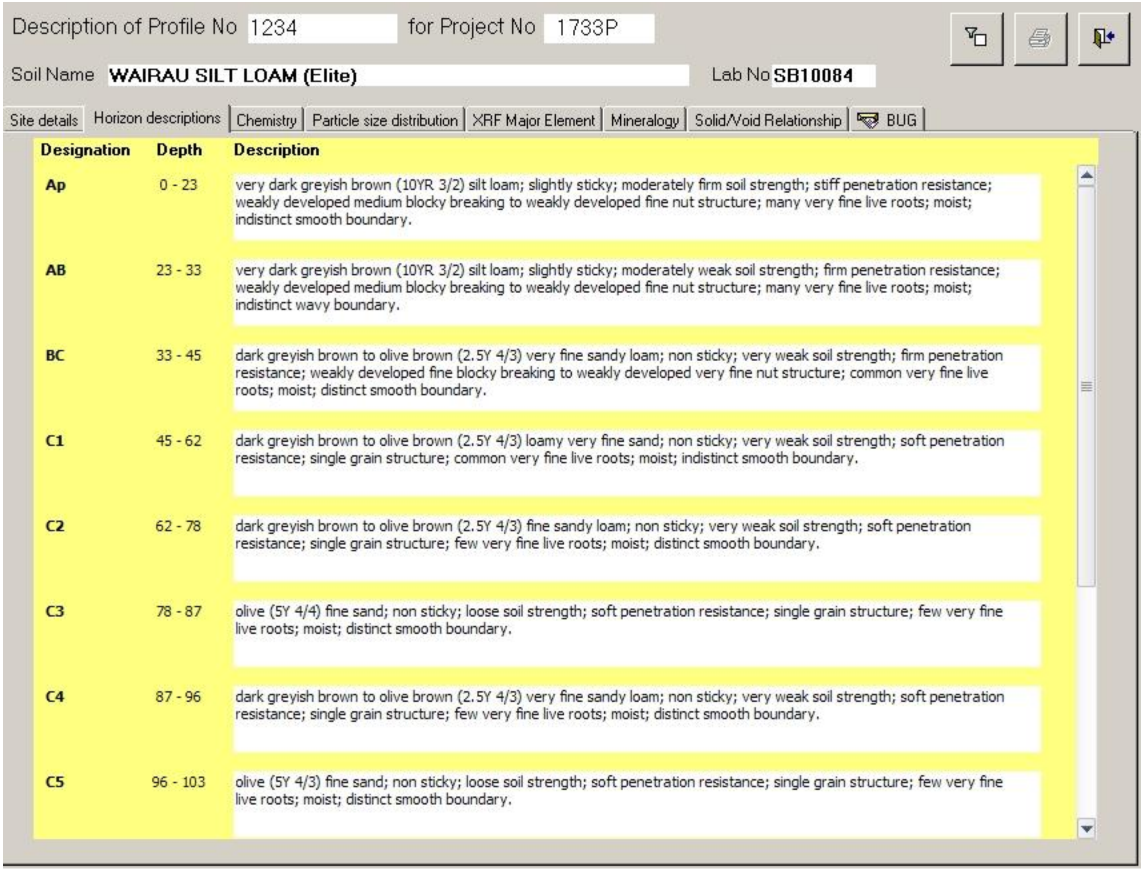
Site data include topography and drainage, parent materials, climate, land management attributes, vegetation cover etc. In the view of this data below, the soil site example is of Wairau silt loam (Lab no. SB10084).


Description of Profile No 1234 for Project No 1733P
Soil Name WAIRAU SILT LOAM (Elite) Lab No SB10084
Site details | Horizon descriptions | Chemistry | Particle size distribution | XRF Major Element | Mineralogy | Solid/Void Relationship | BUG

Map Reference P28 2584700,5972030
Map Series NZMS 260
Author MDL KWV
Date 14/09/1987
Classification Recent soil, very weakly leached,
US Taxonomy MOLLIC USTIFLUVENT COARSE-LOAMY, MIXED, MESIC
NZSC Typic Fluvial Recent Soil
Survey Wairau Plains: Blenheim-Renwick Sector WPBR
Region Marlborough
Locality END OF WRATT'S RD, ON CNR OF RD TO HOUSE (LYALL McLAUCHLAN'S)
Topdress Annual super (200), recently limed (100).
Annual rain 750 mm
Mean Temp - Annual
Elevation 15 m
Landform valley in flat country
Landform Genesis Fluvial
Microrelief
Slope
Aspect
Length
Rock Outcrops
Topsoil loss %
Slope Movement %
Drainage Well
Land Use Long term grass
Vegetation Type Dryland grass
Vegetation Ryegrass, Clover (Improved pasture)
Parent Materials sand, and silt from Wairau River
Notes

Horizons

Data in this view includes soil profile descriptions, horizon depth and designation, colour/mottling, texture, structure, consistence, etc. An important limitation in the current NSD output is that stone content is not described for each horizon (this data does exist but is not currently accessible). The page below extends to another sheet describing horizons C6 103–125 cm, and 2C 125–126 cm.



Description of Profile No 1234 for Project No 1733P

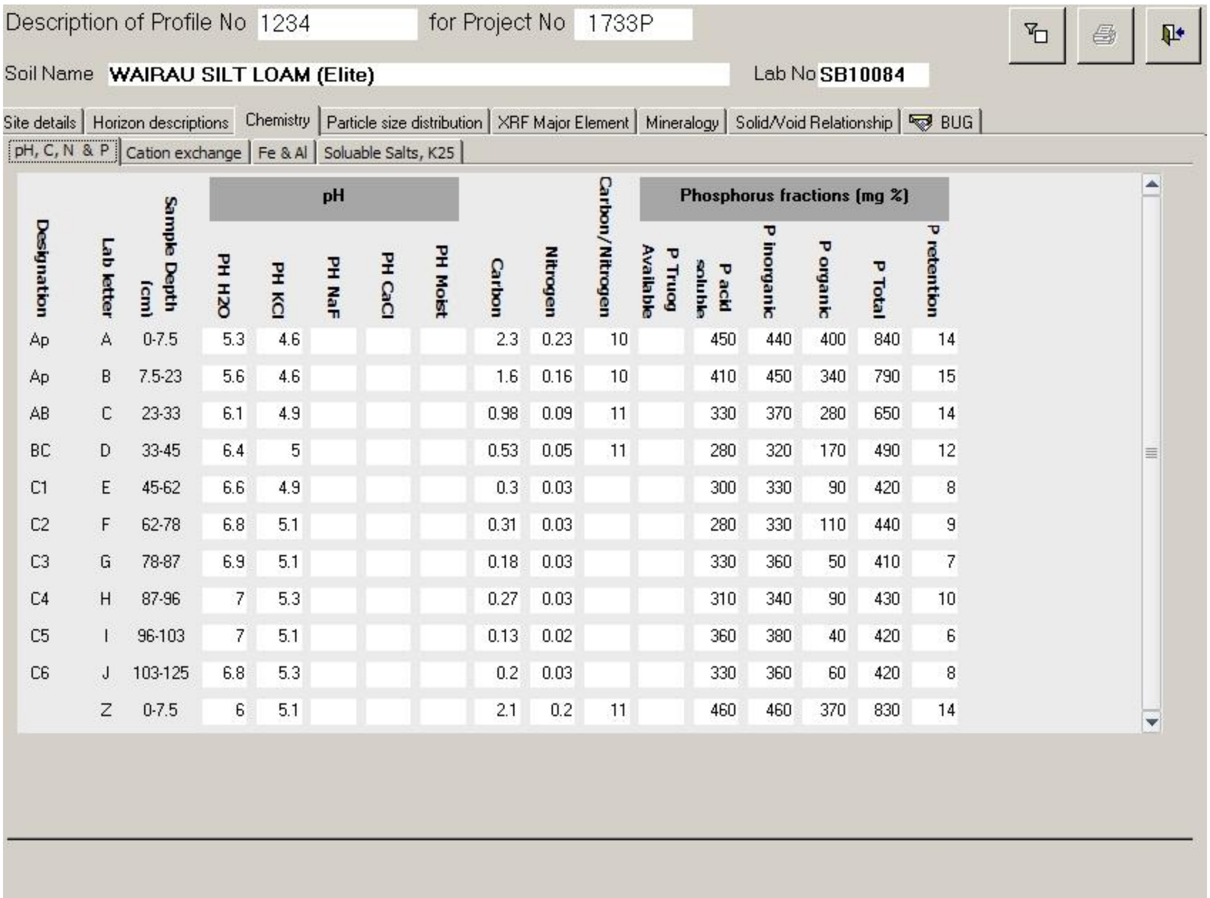
Soil Name WAIRAU SILT LOAM (Elite) Lab No SB10084

Site details | Horizon descriptions | Chemistry | Particle size distribution | XRF Major Element | Mineralogy | Solid/Void Relationship | BUG

Designation	Depth	Description
Ap	0 - 23	very dark greyish brown (10YR 3/2) silt loam; slightly sticky; moderately firm soil strength; stiff penetration resistance; weakly developed medium blocky breaking to weakly developed fine nut structure; many very fine live roots; moist; indistinct smooth boundary.
AB	23 - 33	very dark greyish brown (10YR 3/2) silt loam; slightly sticky; moderately weak soil strength; firm penetration resistance; weakly developed medium blocky breaking to weakly developed fine nut structure; many very fine live roots; moist; indistinct wavy boundary.
BC	33 - 45	dark greyish brown to olive brown (2.5Y 4/3) very fine sandy loam; non sticky; very weak soil strength; firm penetration resistance; weakly developed fine blocky breaking to weakly developed very fine nut structure; common very fine live roots; moist; distinct smooth boundary.
C1	45 - 62	dark greyish brown to olive brown (2.5Y 4/3) loamy very fine sand; non sticky; very weak soil strength; soft penetration resistance; single grain structure; common very fine live roots; moist; indistinct smooth boundary.
C2	62 - 78	dark greyish brown to olive brown (2.5Y 4/3) fine sandy loam; non sticky; very weak soil strength; soft penetration resistance; single grain structure; few very fine live roots; moist; distinct smooth boundary.
C3	78 - 87	olive (5Y 4/4) fine sand; non sticky; loose soil strength; soft penetration resistance; single grain structure; few very fine live roots; moist; distinct smooth boundary.
C4	87 - 96	dark greyish brown to olive brown (2.5Y 4/3) very fine sandy loam; non sticky; very weak soil strength; soft penetration resistance; single grain structure; few very fine live roots; moist; distinct smooth boundary.
C5	96 - 103	olive (5Y 4/3) fine sand; non sticky; loose soil strength; soft penetration resistance; single grain structure; few very fine live roots; moist; distinct smooth boundary.

Chemistry

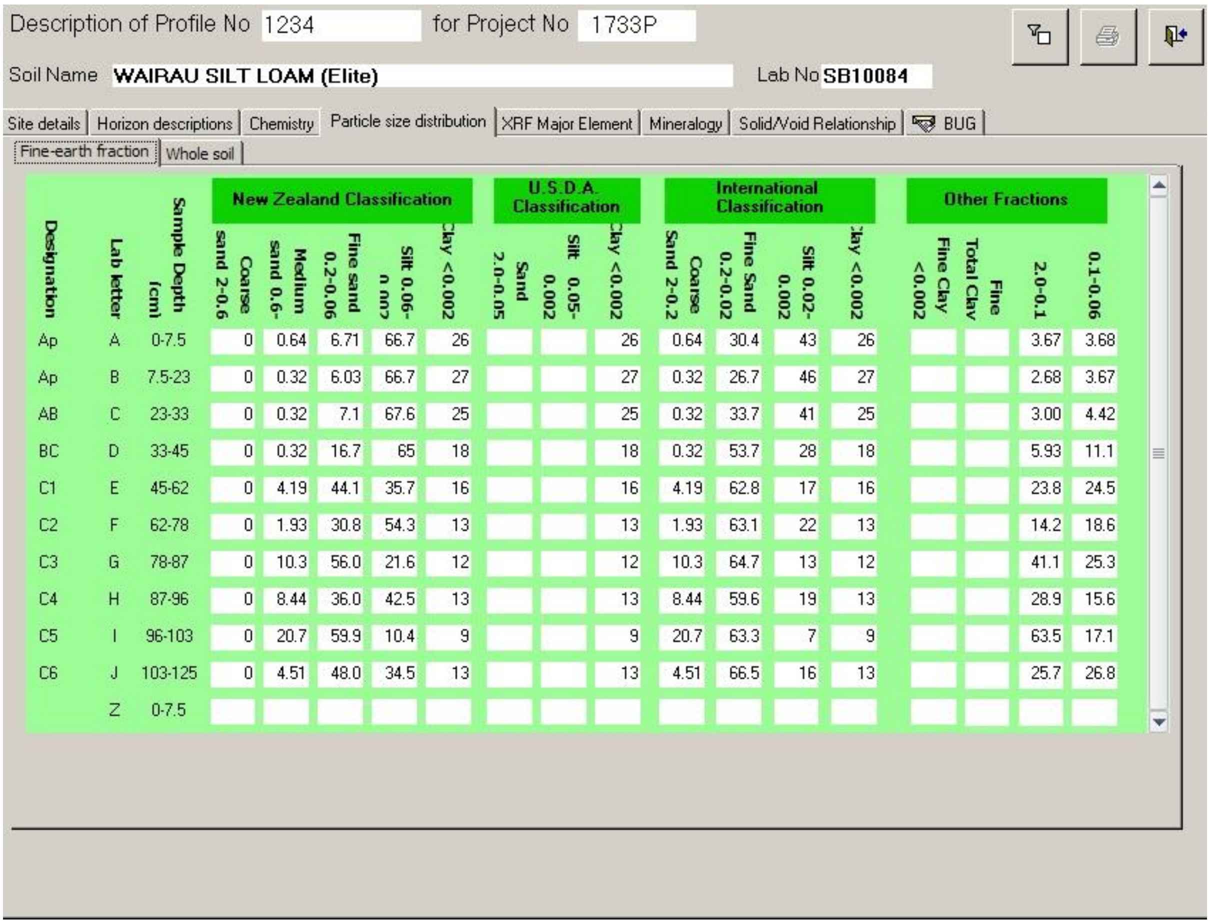
This view is used principally for soil characterisation and classification but with some fertility measurements. Analyses include: organic matter (C, N); pH, cation exchange properties, phosphorus fractions including an available P measurement; P retention, available sulphate & reserves of Mg, and K; Al, Fe, and Si fractions, etc.



			pH									Phosphorus fractions (mg %)					
Designation	Lab letter	Sample Depth (cm)	PH H2O	PH KCl	PH NaF	PH CaCl	PH Moist	Carbon	Nitrogen	Carbon/Nitrogen	P Truog Available	P acid soluble	P inorganic	P organic	P Total	P retention
Ap	A	0-7.5	5.3	4.6				2.3	0.23	10		450	440	400	840	14
Ap	B	7.5-23	5.6	4.6				1.6	0.16	10		410	450	340	790	15
AB	C	23-33	6.1	4.9				0.98	0.09	11		330	370	280	650	14
BC	D	33-45	6.4	5				0.53	0.05	11		280	320	170	490	12
C1	E	45-62	6.6	4.9				0.3	0.03			300	330	90	420	8
C2	F	62-78	6.8	5.1				0.31	0.03			280	330	110	440	9
C3	G	78-87	6.9	5.1				0.18	0.03			330	360	50	410	7
C4	H	87-96	7	5.3				0.27	0.03			310	340	90	430	10
C5	I	96-103	7	5.1				0.13	0.02			360	380	40	420	6
C6	J	103-125	6.8	5.3				0.2	0.03			330	360	60	420	8
	Z	0-7.5	6	5.1				2.1	0.2	11		460	460	370	830	14

Physical analyses

This view comprises particle size distribution.


Description of Profile No 1234 for Project No 1733P

Soil Name WAIRAU SILT LOAM (Elite) Lab No SB10084

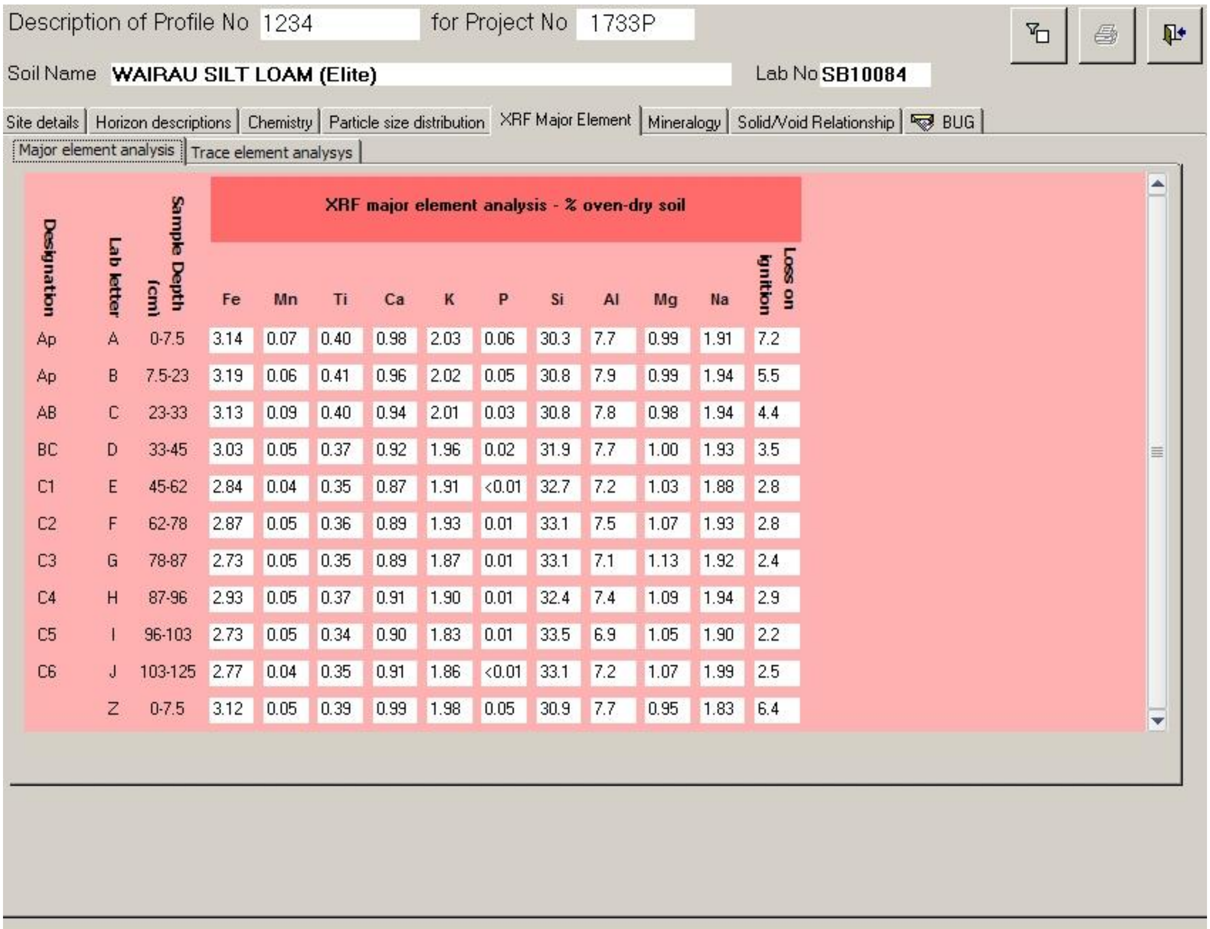
Site details | Horizon descriptions | Chemistry | Particle size distribution | XRF Major Element | Mineralogy | Solid/Void Relationship | BUG

Fine-earth fraction | Whole soil

			New Zealand Classification					U.S.D.A. Classification			International Classification				Other Fractions			
Designation	Lab letter	Sample Depth (cm)	Coarse sand 2-0.6	Medium sand 0.6-	Fine sand 0.2-0.06	Silt 0.06-0.00?	Clay <0.002	Sand 2.0-0.05	Silt 0.05-0.002	Clay <0.002	Coarse Sand 2-0.2	Fine Sand 0.2-0.02	Silt 0.02-0.002	Clay <0.002	Fine Clay <0.002	Fine Total Clay	2.0-0.1	0.1-0.06
Ap	A	0-7.5	0	0.64	6.71	66.7	26			26	0.64	30.4	43	26			3.67	3.68
Ap	B	7.5-23	0	0.32	6.03	66.7	27			27	0.32	26.7	46	27			2.68	3.67
AB	C	23-33	0	0.32	7.1	67.6	25			25	0.32	33.7	41	25			3.00	4.42
BC	D	33-45	0	0.32	16.7	65	18			18	0.32	53.7	28	18			5.93	11.1
C1	E	45-62	0	4.19	44.1	35.7	16			16	4.19	62.8	17	16			23.8	24.5
C2	F	62-78	0	1.93	30.8	54.3	13			13	1.93	63.1	22	13			14.2	18.6
C3	G	78-87	0	10.3	56.0	21.6	12			12	10.3	64.7	13	12			41.1	25.3
C4	H	87-96	0	8.44	36.0	42.5	13			13	8.44	59.6	19	13			28.9	15.6
C5	I	96-103	0	20.7	59.9	10.4	9			9	20.7	63.3	7	9			63.5	17.1
C6	J	103-125	0	4.51	48.0	34.5	13			13	4.51	66.5	16	13			25.7	26.8
	Z	0-7.5																

XRF analyses

Included in this view is analysis of total amounts of major elements (Ca, K, P, Mg, Na, Si, Al, Si, etc.) and minor elements (Mn, Cu, Zn, etc.).


Description of Profile No 1234 for Project No 1733P

Soil Name WAIRAU SILT LOAM (Elite) Lab No SB10084

Site details | Horizon descriptions | Chemistry | Particle size distribution | XRF Major Element | Mineralogy | Solid/Void Relationship | BUG

Major element analysis | Trace element analysys

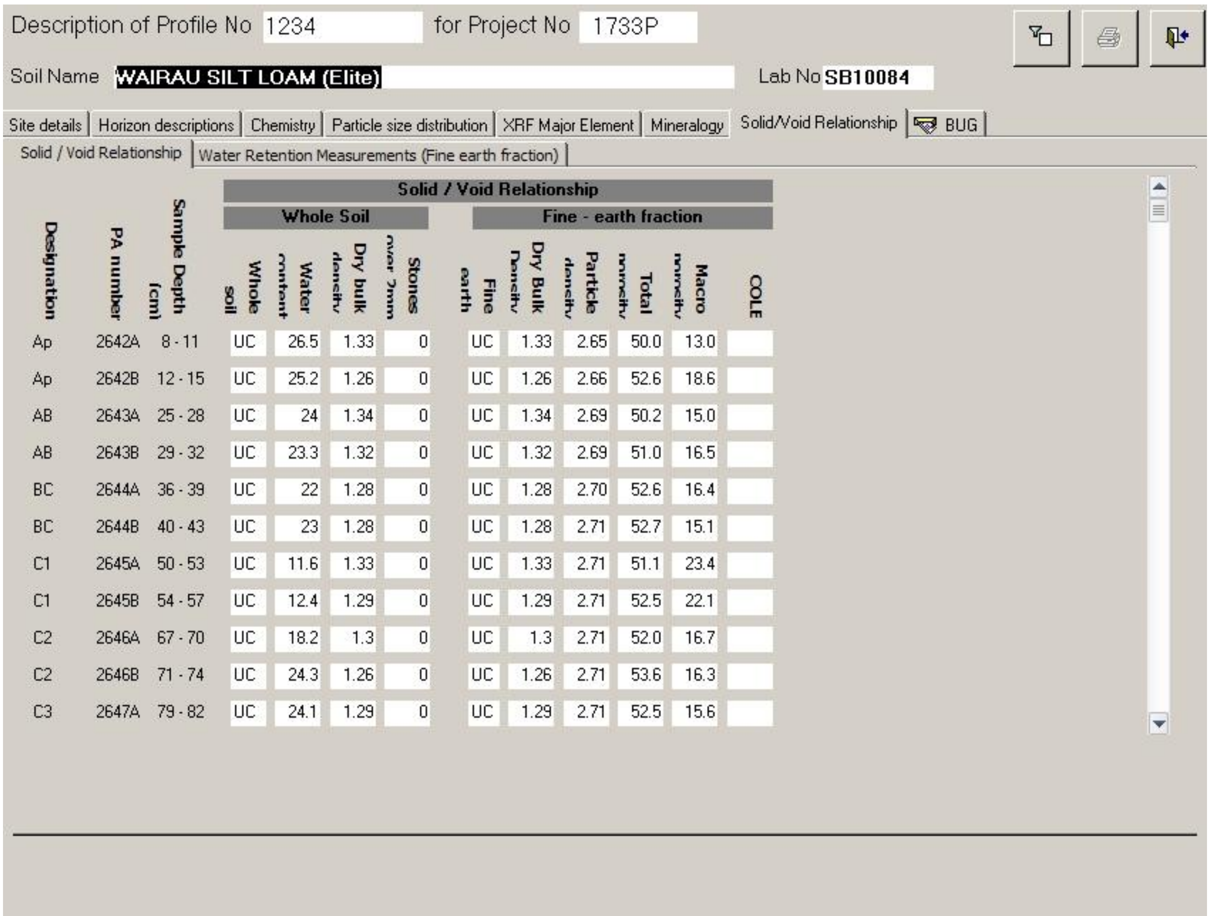
Designation	Lab letter	Sample Depth (cm)	Fe	Mn	Ti	Ca	K	P	Si	Al	Mg	Na	Loss on ignition
			XRF major element analysis - % oven-dry soil										
Ap	A	0-7.5	3.14	0.07	0.40	0.98	2.03	0.06	30.3	7.7	0.99	1.91	7.2
Ap	B	7.5-23	3.19	0.06	0.41	0.96	2.02	0.05	30.8	7.9	0.99	1.94	5.5
AB	C	23-33	3.13	0.09	0.40	0.94	2.01	0.03	30.8	7.8	0.98	1.94	4.4
BC	D	33-45	3.03	0.05	0.37	0.92	1.96	0.02	31.9	7.7	1.00	1.93	3.5
C1	E	45-62	2.84	0.04	0.35	0.87	1.91	<0.01	32.7	7.2	1.03	1.88	2.8
C2	F	62-78	2.87	0.05	0.36	0.89	1.93	0.01	33.1	7.5	1.07	1.93	2.8
C3	G	78-87	2.73	0.05	0.35	0.89	1.87	0.01	33.1	7.1	1.13	1.92	2.4
C4	H	87-96	2.93	0.05	0.37	0.91	1.90	0.01	32.4	7.4	1.09	1.94	2.9
C5	I	96-103	2.73	0.05	0.34	0.90	1.83	0.01	33.5	6.9	1.05	1.90	2.2
C6	J	103-125	2.77	0.04	0.35	0.91	1.86	<0.01	33.1	7.2	1.07	1.99	2.5
	Z	0-7.5	3.12	0.05	0.39	0.99	1.98	0.05	30.9	7.7	0.95	1.83	6.4

Mineralogy

Note that the mineralogy data are not currently available through NSD viewer. These include sand, silt and clay mineralogy; amorphous constituents, and heavy minerals.

Soil/void relationships

Bulk density, total porosity, macro-porosity data is shown in this view.



Description of Profile No 1234 for Project No 1733P

Soil Name WAIRAU SILT LOAM (Elite) Lab No SB10084

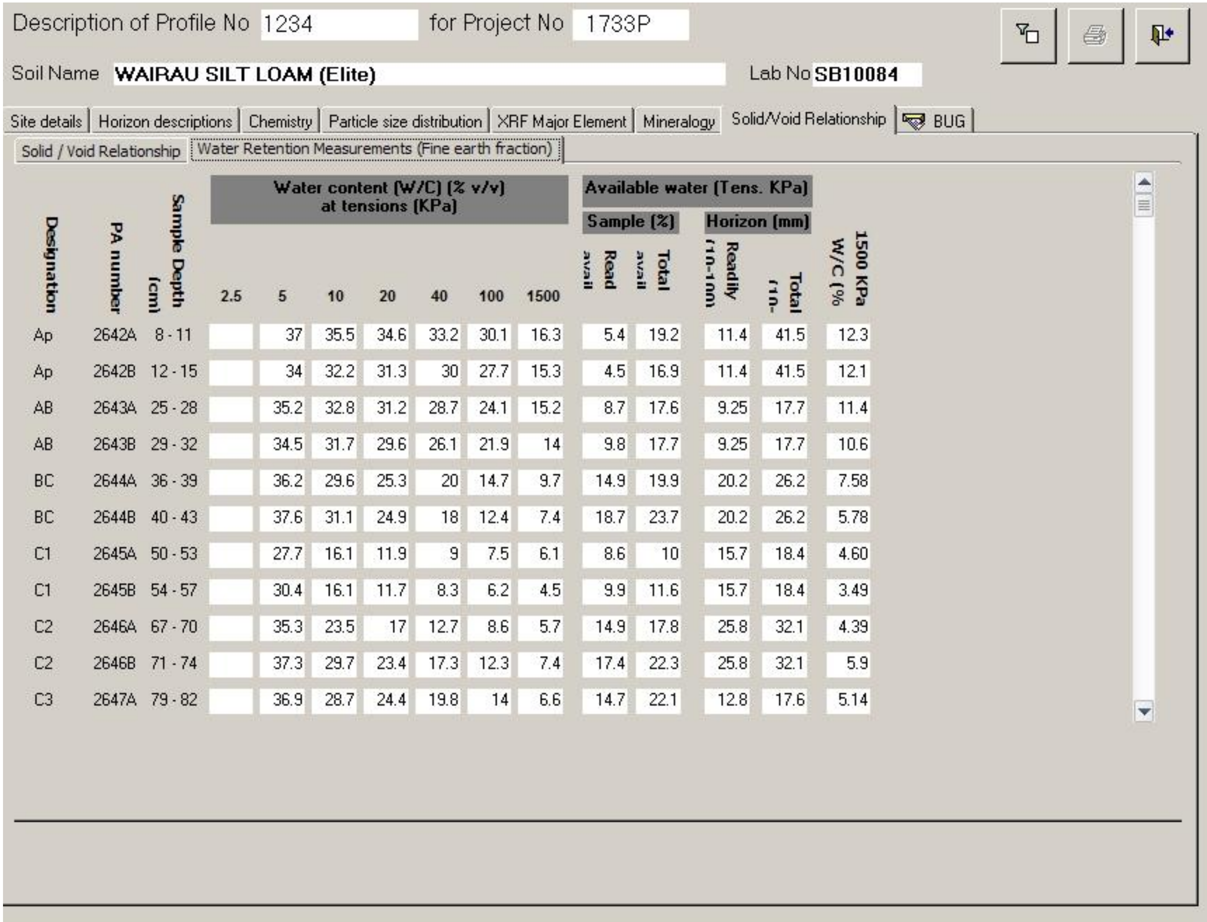
Site details | Horizon descriptions | Chemistry | Particle size distribution | XRF Major Element | Mineralogy | Solid/Void Relationship | BUG

Solid / Void Relationship | Water Retention Measurements (Fine earth fraction)

			Solid / Void Relationship										
			Whole Soil				Fine - earth fraction						
Designation	PA number	Sample Depth (cm)	Whole soil	Water content	Dry bulk density	Stones over 2mm	Fine earth	Dry Bulk Density	Particle density	Total porosity	Macro porosity	COLE
Ap	2642A	8 - 11	UC	26.5	1.33	0	UC	1.33	2.65	50.0	13.0	
Ap	2642B	12 - 15	UC	25.2	1.26	0	UC	1.26	2.66	52.6	18.6	
AB	2643A	25 - 28	UC	24	1.34	0	UC	1.34	2.69	50.2	15.0	
AB	2643B	29 - 32	UC	23.3	1.32	0	UC	1.32	2.69	51.0	16.5	
BC	2644A	36 - 39	UC	22	1.28	0	UC	1.28	2.70	52.6	16.4	
BC	2644B	40 - 43	UC	23	1.28	0	UC	1.28	2.71	52.7	15.1	
C1	2645A	50 - 53	UC	11.6	1.33	0	UC	1.33	2.71	51.1	23.4	
C1	2645B	54 - 57	UC	12.4	1.29	0	UC	1.29	2.71	52.5	22.1	
C2	2646A	67 - 70	UC	18.2	1.3	0	UC	1.3	2.71	52.0	16.7	
C2	2646B	71 - 74	UC	24.3	1.26	0	UC	1.26	2.71	53.6	16.3	
C3	2647A	79 - 82	UC	24.1	1.29	0	UC	1.29	2.71	52.5	15.6	

Water retention measurements

Shown in this view are available and readily available water, etc.



Designation	PA number	Sample Depth (cm)	Water content (W/C) (% v/v) at tensions (KPa)							Available water (Tens. KPa) Sample (%)		Available water (Tens. KPa) Horizon (mm)		1500 KPa W/C (%
			2.5	5	10	20	40	100	1500	Read avail	Total avail	Readily (10-100	Total (10-	
Ap	2642A	8 - 11		37	35.5	34.6	33.2	30.1	16.3	5.4	19.2	11.4	41.5	12.3
Ap	2642B	12 - 15		34	32.2	31.3	30	27.7	15.3	4.5	16.9	11.4	41.5	12.1
AB	2643A	25 - 28		35.2	32.8	31.2	28.7	24.1	15.2	8.7	17.6	9.25	17.7	11.4
AB	2643B	29 - 32		34.5	31.7	29.6	26.1	21.9	14	9.8	17.7	9.25	17.7	10.6
BC	2644A	36 - 39		36.2	29.6	25.3	20	14.7	9.7	14.9	19.9	20.2	26.2	7.58
BC	2644B	40 - 43		37.6	31.1	24.9	18	12.4	7.4	18.7	23.7	20.2	26.2	5.78
C1	2645A	50 - 53		27.7	16.1	11.9	9	7.5	6.1	8.6	10	15.7	18.4	4.60
C1	2645B	54 - 57		30.4	16.1	11.7	8.3	6.2	4.5	9.9	11.6	15.7	18.4	3.49
C2	2646A	67 - 70		35.3	23.5	17	12.7	8.6	5.7	14.9	17.8	25.8	32.1	4.39
C2	2646B	71 - 74		37.3	29.7	23.4	17.3	12.3	7.4	17.4	22.3	25.8	32.1	5.9
C3	2647A	79 - 82		36.9	28.7	24.4	19.8	14	6.6	14.7	22.1	12.8	17.6	5.14

4 Methods

The analyst's name, method of analyses, and date are provided.

Soil chemical analyses from 1966 to mid-1987 follow the procedures in Blakemore et al (1972, 1977, 1981). From the mid-1980s the procedures follow those described by Blakemore et al. (1987). Individual method codes were assigned to values.

Mineral analyses and interpretation of instrument data since 1980 follow Whitton & Churchman (1987). Method codes are given in parentheses. For mineral analyses carried out before 1980, Claridge (1969) was generally followed.

Chemical, physical and mineralogical information for the whole collection of recorded profiles are variable. Soils collected and analysed since the early 1980s generally have a more complete set of analytical data compiled for them compared with earlier collections. Before 1980 the emphasis was on soil chemistry and mineralogy. During the 1980s soil physical data became more valued and the analytical work carried out on New Zealand soils began to reflect this. There was only a short period when complete characterisation was done routinely before funding declined, and as a result the comprehensiveness of NSD physical analyses was limited. It is important to consider changes in methods over time when interpreting data. Analytical data from soils collected before about 1960 should be treated with added caution.

Soil profile descriptions use the method of Taylor & Pohlen from 1962 onwards, with the following modifications: FAO horizon designations for mineral horizons (FAO-UNESCO 1974) and the Canadian system of nomenclature for organic horizons (Canadian Soil Survey Committee 1978) have been in use since 1981. Slope classes follow the system used in the New Zealand Land Resource Inventory (NWASCO 1975–1979, 1979) and have been used in the NSD since 1981. Consistence terminology follows earlier drafts of US Department of Agriculture (1993) and has been used in the NSD since 1983. Penetration resistance classes are based on those of Griffiths (1985). Horizon designations since 1988 follow the system of Clayden & Hewitt (1989). From 1994 on, the method of soil profile description changed from the methods of Taylor and Pohlen (1962) to the methods of Milne et al. (1995).

4.1 Data coverage

Willoughby et al. (2001) made an audit of the NSD, and the NSD is little changed since that date. There were 1449 soil sites with data in the NSD (Figure 3). Nearly all of these have site and soil morphology descriptions. Samples are analysed for significant horizons usually not deeper than 1 m. Soil chemical analyses are the most represented and comprehensive with 1395 sites with at least soil pH data. On a horizons basis these 1395 sites represent 7453 samples with pH data. Soil physics analyses are poorly represented with only 423 sites with bulk density, 747 with sand, silt, and clay, and 425 sites with water retention data. Only 239 sites have soil mineralogy data. Apart from the SWAMP dataset (not yet incorporated in the NSD), there are no soil water dynamics analyses in the NSD. (SWAMP was the Soil Water Assessment and Measurement Programme which measured dynamic soil water characteristics for a set of national benchmark soils).

Sites are clustered around former areas of soil survey activity, and research sites. Of the approximately 2000 soil series that have been recognised in New Zealand less than half have representative analyses – and then frequently only at one site.

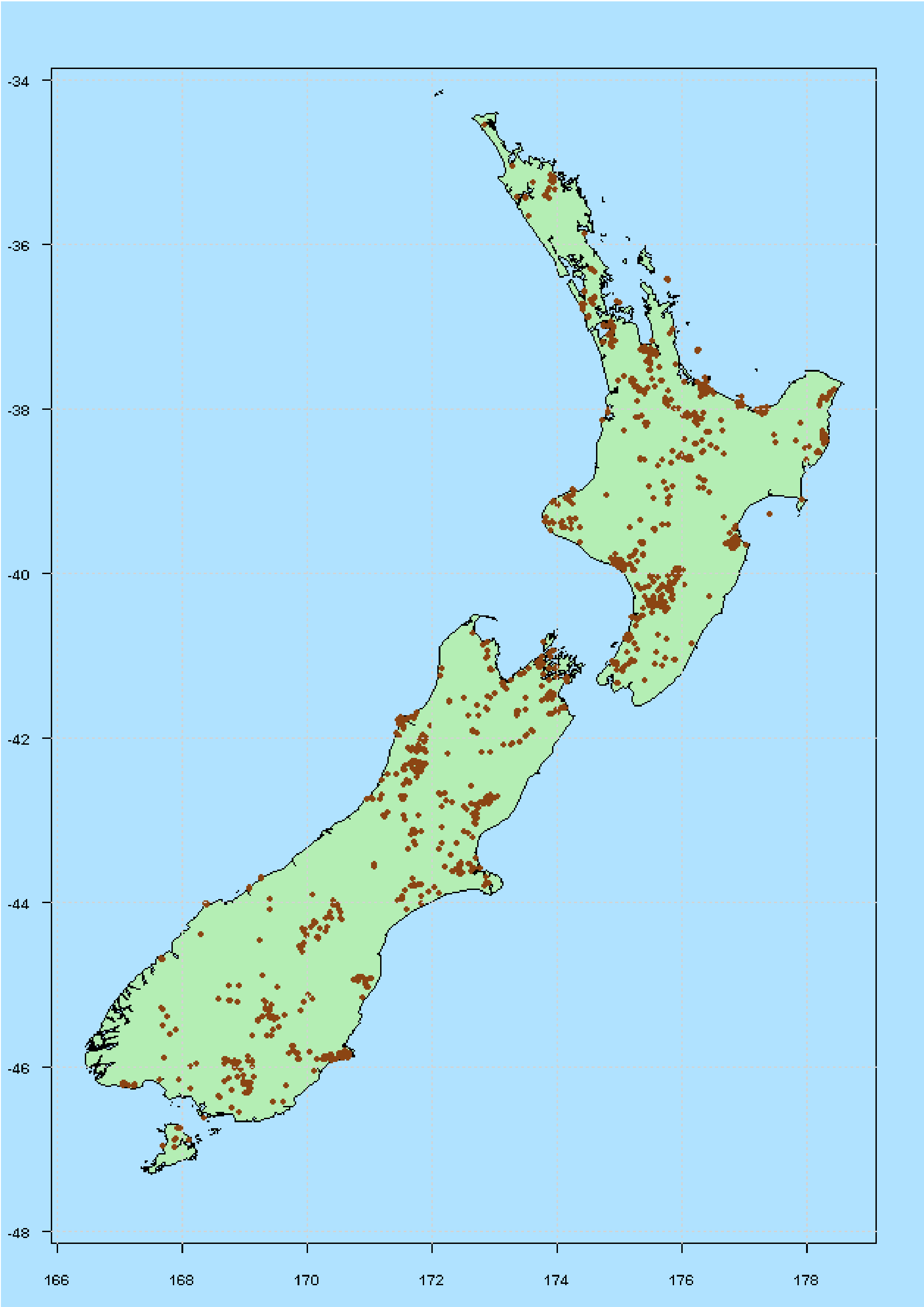


Figure 3 Sample site locations of the current core NSD.

4.2 A Southland case study

The Southland Region provides an example of the importance of providing access to available soil data through the NSD, and points to the potential to be gained by addressing the points made in this report. There are 64 NSD sites in Southland in clusters related to previous soil survey areas, and research sites. Some soils are well represented and many not represented. Of the 64 NSD sites 21 are located in the hill country areas, so have limited applicability for issues in the intensively farmed lowlands.

From 1998 to 2001 there was a region-wide soil survey of 800 000 ha of the lowlands by a community-funded project, Topoclimate South. Across the lowlands 166 soil series were mapped and 563 representative profiles described with basic chemical and physical analysis. Ten years after completion of the Topoclimate soil survey, none of this data has made it into the NSD, and has had very limited use, except as input into interpretative generalised information sheets for each soil type.

The limitations of the NSD came to the fore in 2011 when Environment Southland convened a scientific working party to advise on nitrate and phosphorus sources in the catchment of Waituna Lagoon, which has seen a rapid decline in water quality over the last decade. The Topoclimate South survey mapped 25 different soil series in the catchment; however, the NSD was of limited use, because it contains no sites within the catchment, with only three relevant sites nearby. In comparison the Topoclimate database has 18 sites located within the catchment, and another 19 sites on the same soil series in the immediate area of the catchment.

The Overseer model was used to predict the likely nitrogen and phosphorus losses from different soil types in the catchment, with phosphorus loss per hectare estimated to be 2–3 times greater on Podzol and Organic soils than Gley and Brown soils (Robson et al. 2011). The high estimated phosphorus leaching from the Podzol soils was attributed to the poor anion storage capacity of these soils affecting their ability to retain phosphorus. However, the Overseer model uses NSD data to predict leaching from different soil types, but the lack of spatial representativeness in the NSD data limits the quality of the model predictions. Overseer predictions of phosphorus loss from Podzols are based on the low anion retention capacity of West Coast Podzols – very different to those in the Waituna catchment. The Topoclimate dataset has nine podzol sites in the Waituna catchment, with moderate to high topsoil P-retention, similar to the Brown soils in the catchment. If the Topoclimate data was available to the Overseer model we would expect the Waituna Podzols would have estimated phosphorus leaching similar to the Brown soils.

4.3 Datasets related to the NSD

Datasets exist that should be either part of, or linked to, the NSD. Some of these conform to the NSD structure and data format, and could be entered into the NSD with little work needed. Others may require significant investments to align them to the NSD. In a current independent study, staff within Landcare Research were canvassed to identify datasets (Table 1) that may be entered into, or linked to, the NSD. A linked dataset would be one that did not conform to the NSD so that entry into the NSD is precluded but could be made interoperable with the NSD.

Table 1 Datasets existing outside the current NSD that could potentially be entered into, or linked to, the NSD. The list includes major datasets and is not exhaustive

Dataset	Date/lineage	Comment	Medium	Extent
SINDI	1995–present	Pooled 50-m transects	Digital	national
SWAMP (including McLeod additions)	1980–present	Physics	Digital	National
Topoclimate South	1997–2001	Chemistry physics	Digital	Southland
Forest samples (Parfitt)		Chemistry physics	Digital	National
Carbon Monitoring System		Chemistry	Digital	National
Canterbury Plains hydraulic data (Webb)	1990s	Chemistry physics	Digital	Canterbury
All Physical Analysis Field books, Sample Registers, Raw Data, Calculation records, Results Cards 1976–1993	NSD precursor data	Required to trace NSD physics pre analysis	Paper	National
All Hamilton Physics Lab Field books, Sample Registers 1993–2012	NSD precursor data	As above	Paper	National
All Physical Analysis electronic raw data and results from Taita Analytical Lab 1986–1993	NSD precursor data	As above	Digital	National
All Hamilton Physics Lab raw data and results 1993–2012	NSD precursor data	As above	Digital	National
Taxonomy Samples (TI)	1970s–1980s	Chemistry Physics	Digital	National
Investigation Samples (IS)	1970s–1990s	Various	Digital	National
NSD re-sampling data	Old site, new data	C,N&P data	Digital	National
NSD re-sampling GPS	Updated GPS	Add to NSD	Digital	National
NSD site/profile photos	Various sources	Link to new NSD	Image	National
Antarctic soils database		NSD compatible	Digital	National
SPG1 Profile descriptions on the vax	1970s–1980s	Morphology	digital	National
Waikato Uni, 25 paired pits	2012	comprehensive	digital	Waikato

4.4 NSD documentation

NSD documentation (ordered chronologically) includes field definitions, identification of computed fields, missing value assignments, and data precision definitions.

System and field manuals

1979 [Taylor NH, Pohlen IJ] Soil survey method. A New Zealand handbook for the field study of soils. New Zealand Soil Bureau *25*. 242 p. (1962 version reprinted with amendments)

1984 [McArther, Giltrap] SPG1 Data Base System manual (V1.1) NZ Soil Bureau Laboratory Report SS4.

1988 [McDonald, Giltrap, McArther] Revised SPG1 Data Base System manual (V1.2) NZ Soil Bureau Laboratory Report SS16.

1994 [Willoughby EJ, Baker IR] Standard input/output procedures for the National Soils Database using Paradox and Windows. Unpublished Landcare Research Report. 21 p + appendix.

1995 [Milne JDG, Clayden B, Singleton PL, Wilson AD] Soil description handbook. Lincoln, Canterbury, Manaaki Whenua Press

1996 [Wilde RH, Ross CW] New Zealand Reference Soil Collection and the National Soils Database. New Zealand Soil News 44: 224–227.

2003 [Wilde RH ed.] Manual for National Soils Database. Unpublished Landcare Research report.

Analyses documentation

1972 Soil Bureau laboratory methods. Soil Bureau Scientific Report 10.

1987 [Whitton & Churchman] Standard methods for mineral analysis of soil survey samples for characterisation and classification in NZ Soil Bureau. NZ Soil Bureau Scientific Report 79.

1987 [Blakemore, Searle, Daly] Methods for chemical analysis of soils. NZ Soil Bureau Scientific Report 80.

1988 [McQueen] Glossary of soil water terms. Soil Bureau Laboratory Report EP31.

1989 Determination of particle-size distribution in fine-grained soil – pipette method. Division of Land and Soil Sciences Technical Record LH5.

4.5 Core dataset problems

Users, data managers, and domain experts, past and present, were interviewed to compile an inventory of issues that need to be addressed. The major problems are listed below:

- Related to design or need for better metadata, e.g. integration of data based on variable depth genetic horizons or fixed depths, hidden dependencies, storage of temporal data, methods code explanation, improved method descriptions, methods for site location in temporal sampling
- Data gaps, e.g. bulk density, stone content
- Errors, e.g. confusion between air capacity and macroporosity, insertion of false precision by Paradox database, inconsistencies in morphology descriptions, soil classification assignment
- Character values in mineralogy mixed with numerical values, e.g. changes in mineralogy notation in time
- Unsatisfactory data approval system
- Site and horizons do not conform to original yellow-card data

Many of these problems resulted from database platform changes, reduced funding, and staff losses.

Initially there was a good quality control system while the system was implemented in Datatrieve. It used a Temporary Soils database to hold data until it had been approved by analytical lab leaders and the chief soil correlator. Clearance allowed transfer to NSD. The database structure, quality control procedures, search and reporting capabilities built into both the Temporary and NSD resulted in the database being both a repository of historical information, as well as giving a centralised, live, storage facility for laboratories, and access to well-managed data at that time.

When the NSD was transferred to a PC platform (Borland Paradox) in the early 1990s, the Temporary Soils database was not transferred to Paradox. The previous data input procedures and system were not recreated. This resulted in the centralised live-data-storage functionality of the system being lost, and the individual laboratories reverted to localised storage of data. The laboratory data storage still now remains localised.

5 User requirements

5.1 Regional and unitary councils' needs

Users from five regional and unitary councils provided feedback on the NSD.

Their verbatum comments:

- 'My view is there's useful information for regional councils in NSD. So I'd support any proposal that (at modest cost) can make access to NSD more user-friendly.'
- 'I would be generally supportive to improve access and usability of the NSD soil data information.'
- '…interested in future use and particularly of information that may help … improve management of the region, especially under increasing intensification of land use, and potential use in tools or by the resource consents team.'
- '…available infiltration rate data and structural vulnerability class data incorporated into the dataset.'
- '…important that there is a process developed for the updating of the database which would include tagging updated data with a date, and informing users of updates when they occur.'
- 'There has been quite a lot of reports, files and theses conducted that have not made their way into the FSL database.'
- '…there is a big grey area across the MUL [metropolitan urban land] that has no soil information. I have been asked about what soil information we have for the MUL, particularly from a stormwater perspective that are interested in a soil's drainage properties and infiltration rates.'

From this feedback, other discussions, and observations of regional council use of soil data, we summarise regional and unitary council needs in the following:

- Convenient and user-friendly access to NSD data
- New data to increase NSD coverage
- Clear information on updating
- Input data for management tools, and to support resource consents teams
- Linkage between soil quality data held in SINDI and NSD
- Storage of soil profile and site descriptions of new soil quality monitoring sites, and data from other investigations
- Reference locations to NSD site for resampling and observing type sites
- Specific data on infiltration, soil drainage properties, and structure vulnerability

Essentially, the NSD provides six functions, all of which underpin regional and unitary council operations.

1. Data to develop models of several kinds for prediction of complex information (e.g. infiltration rate) from fundamental soil data, and construction and improvement of management models such as Overseer

2. Input data to run models. Model accuracy is highly dependent on good data for the soil of interest

3. Underpinning of S-map soil classification and to provide accurate soil information products delivered by soil factsheets (see S-map Online)

4. Access to further analytical data on profiles represenatative of soils identified in field investigations or by S-map Online

5. Storage and national curation of data collected in various field investigations

6. A fundamental resource for research into spatial and temporal variability of New Zealand soils, and their management

5.2 Science needs

Science needs are highly varied as shown by the following sample of science applications. In all of these the NSD is essential and indicates the high level of reuse of the data.

A survey of science users yielded the following applications.

- Model building (e.g. Overseer)
- Model application (e.g. SPASMO application)
- Soil process understanding (e.g. sources and processes soil N by correlations with C, Al, Fe, clay mineralogy, and air dry water content)
- Data dumps (e.g. MfE LUCAS carbon monitoring system)
- Sample archive analysis (e.g. analysis of Bomb C14 and Cs137 fallout by analysis of pre-bomb C14 and Cs137 in samples of various ages – an example of an 'out of left field' application that would not have been anticipated)
- Non-destructive scanning of the sample archive (e.g. scanning to produce a near-infrared (NIR) scan library of approximately 2000 samples, to be used to develop models to predict soil C in the field using a portable NIR scanner)
- Resampling to monitor temporal change (e.g.160 NSD pasture sites resampled for carbon to estimate change over a period of 20–30 years in relation to predominant land use)
- Taxonomy development (e.g. development of the New Zealand Soil Classification system)
- Definition and quantification of soil natural capital and modelling of soil services (e.g. research in SLURI programme)
- Variability studies of New Zealand soil classes (e.g. a study of New Zealand stony soils, their variation and soil services)

- Studies of soil genesis (e.g. short-range-order mineralogy in Podzols)
- Interpretation of soil quality target levels (e.g. SINDI use of NSD ranges as a comparative standard for soil quality indicator interpretation)
- National-scale digital soil mapping and land environments mapping. (e.g.national 25-m soil carbon map, and carbon/nitrogen ratio map)

Most of these regional council and science needs will continue but more intensive land use will reqiure better models to manage land with greater precision. This will require a finer network of NSD sites and and wider range of analyses, particularly of dynamic soil water properties (e.g. infiltration rate). The inference system of S-map will be able to deliver a wider range of soil properties with greater accuracy but this will be dependant on NSD improvements.

6 Vision for the National Soils Database (NSD)

The NSD will be a database of soil attributes sampled mainly in the New Zealand landscape, that:

1. **Has staged development** towards an end product where the NSD acts as a single virtual database. In Stage 1, start with core datasets that can most readily be incorporated so the database can be used and prove itself. The core datasets are the original-core plus Landcare Research temporary datasets that can be joined without extensive revision. The temporary datasets have similar format and standard to the existing NSD (i.e. datasets with profile and analytical characterisation, e.g. Topoclimate, Parfitt forest dataset). Stage 1 has three components: 1. Developing the data model; 2. Expanding the core dataset; 3. Basic access/delivery systems. We need to get the core data in as early as possible, then build usability and visibility. The priority is to serve the needs of model developers (including pedotransfer functions, management processes, soil services) so that better underpinning by the NSD will flow through model application to high-priority applications. In Stage 2, investigate other datasets for inclusion or linking to the NSD (e.g. other CRI datasets, soil quality, Z data, fertiliser company datasets).

2. **Has levels of accessibility**. The principle is ease of use for most, but with the facility to make more sophisticated SQL queries for more acomplished users. A diagram of the main functional components of the NSD is shown in Figure 4 and shows four access types.

   - Profile locations displayed in S-map with data summarised in a factsheet
   - NSD viewer similar to the present viewer (see Section 3.3) with the core dataset publicly available (Stage 1)
   - SQL query access as standard templates or specialist free format (Stage 1)
   - Web data feeds to user tools or models

   It is possible that there may be a demand for storage of proprietary datasets in the database with password access. These could be regional/unitary authorities, industries etc. where information privacy would not allow public access. The National Inventory of Vegetation Surveys provides a similar service (Stage 2)

3. **Is comprehensive.** It incorporates or links to all available national datasets (Landcare Research, regional councils, universities, and CRIs) that can be accommodated by the data model.

4. Is fed by a **gaps sampling** programme to ensure representation of all NZ Soil Groups, and environments. Initially some work is required to correct the NZSC classification assignments of existing NSD sites to fit modern NZSC and S-map requirements.

5. Has **quality** levels driven by metadata in a form readily used in querying. For example, a query for analyses by lab (for Landcare Research, lab needs to distinguish the code – e.g. SB or IS etc.) and date, or a query for field morphology described according to Taylor & Pohlen (1979), and institutional origin. The database will need to store profile description data that is unaccompanied by analytical data.

6. Has a **dedicated database manager** with soil and informatics knowledge (or advisers at hand). Responsibilites would include quality control, stakeholders, maintenance of data systems/platform/ development, training, and supervising gap-filling(Stage 1). Advisors would be needed for each data domain to adjudicate on technical issues.

7. Has a **data entry facility** that includes electronic field entry, and integrates this with lab data (Stage 1), and includes an ability to link new data to old data at the same site in temporal-change studies. A query for new data should direct the inquirer to any old data at the same site. It must also provide for description uncertainty information associated with temporal sampling. Site and profile photos should also be stored (Stage 2).

8. **Integrates with S-map** by showing: (1) NSD sites in S-map with point and click – 'Soil properties here', (2) NSD storage of profile descriptions that can be part of 'Soil properties here'. In contrast to S-Map, the NSD holds fundamental attributes, calculated attributes, and pedotransfer functions that estimate missing attributes (e.g. bulk density). S-Map, however, holds complex derived attributes and information products. It should be possible to see the data model and NSD inputs used to derive the complex derived attributes seen in S-Map.

9. Maintains **soil quality** data in the SINDI database. Fertiliser testing data could be linked to SINDI, and NSD z samples. Origin, and methods would be identified by the metadata.

10. Able to **communicate** with other environmental datasets nationally and internationally, by using international standards and communication protocols.

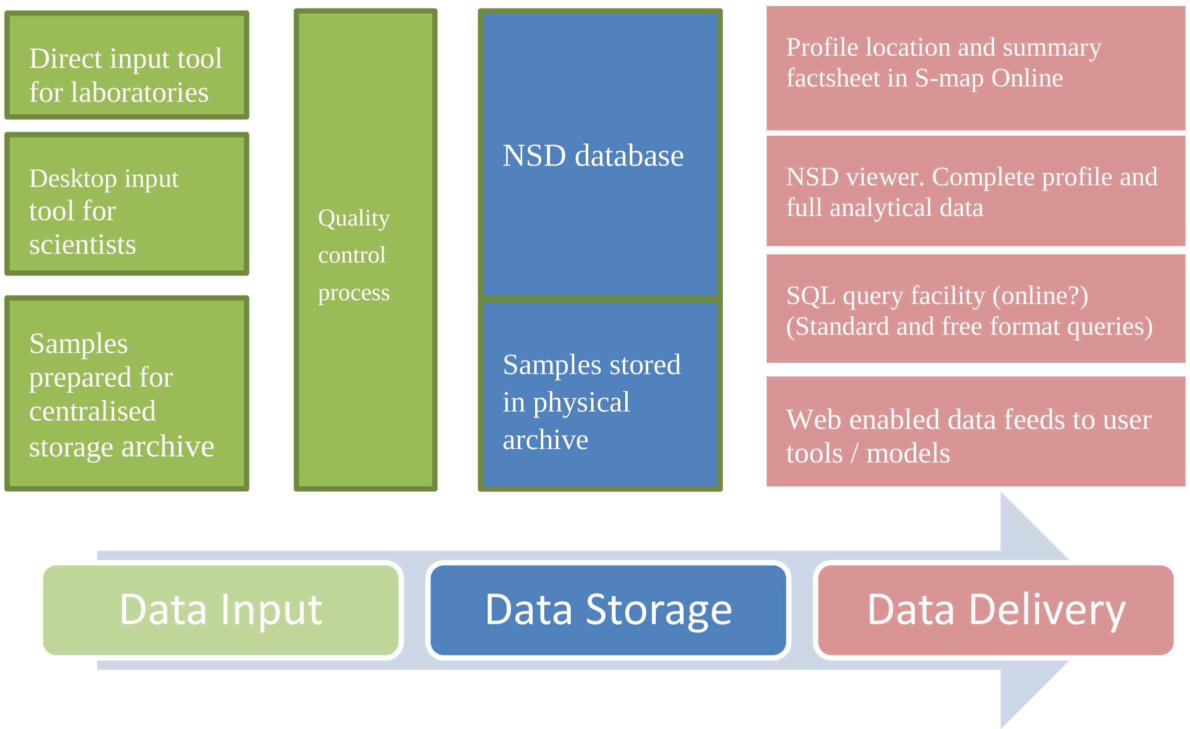


Figure 4 Schematic overview of the re-developed NSD.

7 Development plan

Development of the NSD must consider the informatics design, data quality issues and filling data gaps.

7.1 Data quality

All datasets, both the data within the current core NSD and datasets that may be either merged or linked to the NSD, need to be assessed for data quality issues reviewed in the NSD status section of this report (Section 3). A triage is then needed to determine dataset clean-up priorities. Suggested triage criteria are:

1. Investment needed to address issues
2. Level of urgency for use of the data
3. Any funding tagged to recovery of the data

7.2 Filling data gaps

Prioritise gaps based on:

1. Important soil attributes
2. Land use types
3. Geographic distribution
4. Soil families
5. Convenient sampling in conjunction with other work
6. Cost

7.3 NSD access options

A schematic diagram shows a representation of the deployment and data access options (Figure 5). The input of data to the NSD presupposes field and lab methods and procedures, a data submission and approval system, and a data model as the basis of data validation and accommodation in the NSD.

The numbers refer to proposed NSD informatics projects. There is an implicit order of deployment in the numbering system, but it may change according to the needs of stakeholders. Note too that the steps may be revised as the breadth of content of the database grows.

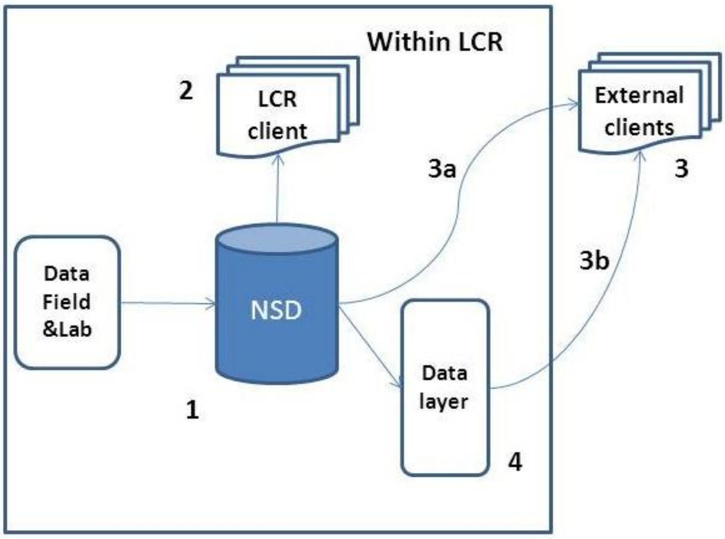


Figure 5 Schematic representation of deployment and data access options. LCR = Landcare Research.

In Figure 5:

1 is the development of the database platform and stored data.

2 is data provision to portals developed and maintained by Landcare Research, e.g. S-map, Our Environment, or a future NSD Portal, and any other Landcare Research applications where data is drawn dir ectly from the NSD. The data may be unmodified and read directly from the NSD, or may be processed in some way (e.g. by a pedotransfer function).

3 is provision of access to external clients. In 3a, controlled access to the NSD is provided to to external clients via a secure interface allowing SQL-style querying of the NSD data model.The results can then be downloaded. In 3b, client access is mediated by an NSD data delivery layer (see 4). 3a is an interium option, while 3b/4 is the long-term preferred approach.

4 is the deployment of the NSD data delivery layer (see Section 7.4, NSD Data Delivery).

The SND components diagram Figure 5 helps identify development priorities. Clearly all components could ideally be developed, but cut-down options are needed to match funding levels. The following options are suggested (Table 3).

Table 3 NSD development options. The numbers relate to Figure 5.

Option	Components developed	Explanation
A	1 only	The NSD is developed but access is not provided. Data would need to be provided by few qualified individuals who would take a request and send the results of required queries to users. [Note, users would tend to retain their copies resulting in versioning problems]
B	1+2	Landcare Research research-orientated service
C	1+4	The ‘data streaming’ option (equivalent to the NIWA meteorological data service)
C	1+2+4	Landcare Research client service + data streaming elsewhere
D	1+3a	A possible interim service but option E is preferred
E	1+4+3b	Preferred option for external clients
F	1+2+4+3b	Comprehensive service

7.4 NSD data delivery

Consideration must be given to ensuring that NSD data is delivered in a robust and stable manner. Robustness refers to suitable availability of secure data services, while stability refers to the provision of data using a consistent structure and set of formats. The latter is the basis of what is commonly referred to as ‘interoperability’ as it allows a network of heterogeneous systems to interoperate, or work together to exchange data.

An interoperable network of systems could be organised in several ways: along client–server lines – where a potentially unlimited number of client applications can connect to a single, standardised data source; in a distributed system where the clients connect to what is effectively a virtual data-store composed of multiple servers, each exposed using the exact same standards; or using a combination of both. The flow of data within the network may be unidirectional, where clients simply consume the available data, or bidirectional, where clients not only consume data, but also submit new or changed data to the appropriate repository.

The standards involved define the communication protocols used by nodes in the network to request and transmit data, and standards that define the structure and content of the data. The two are related, but are not necessarily tightly coupled. For the purposes of this report, it is assumed that the communication will occur across the Internet, facilitated by web services (see below).

A full analysis of available protocols and standards is beyond the scope of this report, but consideration should be given for an appropriate means of exchanging data between participating systems. Criteria for the evaluation of data-exchange systems should cover:

1. The ability to provide secure access to data, by restricting:

    a) Access to a subset of the features available in the data (e.g. excluding confidential information), or whole suite of features or properties

   b) Rights to manipulate the data – e.g. allow read-only access, or the ability to create, update or delete existing data

   c) A combination of the above

2. The technical feasibility of deployment according to the current capability of information systems.

3. The likely cost of deployment and, more importantly, ongoing maintenance and support.

4. The ability to adapt to known trends for data exchange, and accommodate increasing volumes of data.

5. Stable data abstraction layers that decouple storage mechanisms from data exchange mechanisms (allowing organisations to hide and therefore change their information systems as needed, without exposing those changes to the broader network).

6. Well-thought-out information models – including clearly defined content – governed by active scientific communities (careful consideration should be given to supporting integration with other environmental datasets).

Care must be taken to separate the development of information models that inform the structure and content of the data being exchanged from the technology used to store, transfer and analyse the data. It is envisaged the protocols and formats used to exchange data will evolve at a more rapid rate than the conceptual basis of the soil science constraining the information model.

7.5 Candidate service infrastructures

A variety of service infrastructures may be used to deliver data; however, this is a fundamental decision that is relevant to *all* information provided by the environmental and earth science community, not just soil scientists and their stakeholders. As such it is recommended that NSD data delivery aligns with policies and practices being defined by larger communities: e.g. Land Information New Zealand (LINZ) has defined a New Zealand Spatial Data Infrastructure (SDI) based on the specifications of the Open Geospatial Consortium (OGC). The NZ SDI is not unusual as OGC-based SDIs have been adopted or defined elsewhere (e.g. the FGDC[2], AuScope GRID[3], US-GIN[4], GIN-RIES[5] and INSPIRE[6]).

---

[2] The Federal Geographic Data Committee (http://www.fgdc.gov/)

[3] Australian national earth science infrastructure program (http://auscope.org.au/site/grid.php)

[4] US Geoscience Information Network (http://usgin.org/)

[5] Canadian Groundwater Information Network (http://gw-info.net/)

[6] Infrastructure for Spatial Information in the European Community (http://inspire.jrc.ec.europa.eu/)

The OGC itself is 'an international industry consortium of 456 companies, government agencies and universities participating in a consensus process to develop publicly available interface standards'[7]. It has, in close collaboration with the International Standards Organisation Technical Committee 211 (ISO/TC211), defined standards for a suite of geospatially oriented web services to support the delivery of maps, features (raw vector data), coverages (gridded data) and catalogues (that enable the publication of the other services). The consortium has also implemented ISO/TC211 standards to define practices for the modelling and encoding of geospatial information – see below. This breadth of coverage of the necessary standards, including strong geospatial support, combined with increasing implementation in the environmental sciences, makes the OGC suite of standards a strong contender for an initial infrastructure.

The provision of NSD data should not necessarily be restricted to a single standard and consideration should be given, with time, to supporting other infrastructures as well – particularly the tightly bound standards of Representational State Transfer (REST; Fielding 2000) and the Semantic Web (W3C 2012). This is not necessarily at odds with an OGC implementation, as the consortium is actively addressing alignment with these standards.

7.6 Candidate information models

An advantage of adopting the OGC standards is access to a wide variety of conceptual and physical modelling tools. As stated, they build on the published ISO/TC211 standards for methods, tools and services required for geographic information. These introduce the General Feature Model as a basis for describing real-world features, and define a methodology for applying it to the description of domain-specific features. The resulting application schema (so-called because they apply the general model to a particular domain of interest) can be encoded as XML and transmitted using OGC web services.

Communities have formed to define models for domains that include the description of observations and sampling practices (Cox 2010), geology (CGI 2012), climate (NERC 2011) and water (OGC 2012). The International Union of Soil Sciences (IUSS) has formed a working group to define the Soil Mark-up Language (SoilML), a model for soil information exchange (Montanarella et al. 2010) and it will reconcile a number of existing models, including the European INSPIRE soil conceptual model (INSPIRE 2011), the Soil and Terrain Mark-up Language (Pourabdollah et al. 2012) and the Australian OzSoilML.

These community-driven modelling activities are the real strength of the OGC, bringing together experts from a variety of sectors to develop robust models, based on scientific or real-world principles (in that they describe the actual entity of interest, not how a computer system may interact with, or present, data about the entity). It is recommended that the soil-related information models be reviewed as candidates for the data-exchange models. Better still, Landcare Research should participate in the development and testing of the information models.

---

[7] http://www.opengeospatial.org/ogc

It must be noted that some of the needs of the NSD may not be met by these models, either because more data are captured within the NSD than the community recognises, or there are specific data products developed to meet the needs of a subset of end-users (perhaps by post-processing the core NSD data). In these circumstances it is permissible to develop an NSD-specific data model, as a last resort, preferably by extending the community model, but perhaps as a standalone one.

That this is a last resort cannot be understated; the development of a comprehensive, scientific information model is a laborious task. As an example, the North American Conceptual Data Model for Geologic Maps (NADM 2004) was developed by a group of nine geoscientists from six states and national geological surveys over 5 years. Making the most of existing thinking, and only focusing on the gaps, is a necessity.

This applies to the modelling of the NSD database as well. The need to redesign the database according to scientific requirements has been cited when interviewing Landcare Research staff, and modelling activity from OGC-aligned scientific communities could well inform design efforts. These standards, principally Observations and Measurements and the IUSS soil modelling work, should be consulted and reused wherever possible.

8 Recommendations

We conclude by recommending priority development tasks. The NSD vision statement in this report suggests a two-stage development process with a Stage 1 goal of doing the essentials needed to provide good basic access to a core dataset comprising the least problematic data. The following steps are suggested that would contribute to Stage 1 and provide a good interim service that will allow the NSD to prove itself.

1. Appoint a database manager to plan development tasks and revise and implement database procedures.

2. Develop a soil data model that will facilitate database development and enable the NSD to conform to international data exchange standards.

3. Review datasets to identify those with potential for addition to the NSD.

4. Establish a new-NSD by copying in sound datasets and retaining the old-NSD as a legecy dataset.

5. Develop access routes by (1) maintaining the NSD Viewer and if posssible opening it to external view, (2) provideing a client-service query system for both standard template queries and free-format SQL queries. The goal will be to first serve the needs of model developers (including pedotransfer functions, management processes, soil services) so that the NSD can better play its role in underpinning high-national-priority applications.

6. Provide NSD storage for regional council datasets that meets the database standards and proceedures – by submitting data through via database manager.

7. Set up a gaps-sampling programme cooperatively with other funded projects.

8. Develop a data-entry facility in conjunction with a field-data-entry system for S-map fieldwork.

These recommendations are not yet costed. Much of the developent involves informatics investment which must be carefully planned and budgeted. The recommendations are therefore conditional on a thorough informatics plan, which was outside the scope of this report. Where possible, development costs need to be leaveraged on other related projects. It is recommended that application be made for an Envirolink Tools grant to gain initial momentum in development.

9 Acknowledgements

We thank Colin Gray (Marlborough District Council), Jon Roygard and Malcolm Todd (Horizons Regional Council), Fiona Curran-Cournane (Auckland City), Andrew Burton (Tasman District Council), Dave Cameron and John Dewry (Greater Wellington Regional Council) for their input into this project. The work was funded by the Ministry of Science and Innovation through an Envirolink Medium Advice Grant to Marlborough District Council

10 References

Blakemore LC, Searle PL, Daly BK 1972. Methods for chemical analysis of soils. Part A. Soil Bureau Laboratory Methods. New Zealand Soil Bureau Scientific Report 10 [with subsequent editions in 1977 and 1981].

Blakemore LC, Searle PL, Daly BK 1987. Methods for chemical analysis of soils. New Zealand Soil Bureau Scientific Report 80. 103 p.

Canada Soil Survey Subcommittee, Subcommittee on Soil Classification 1978. The Canadian system of soil classification. Canadian Department of Agriculture Publication 1646. Ontario, Canada, Supply and Services Ottawa. 164 p.

CGI 2012. GeoScience Mark-up Language. http://www.cgi-iugs.org/tech_collaboration/geosciml.html, accessed 6 June 2012.

Claridge GGC 1969. Physico-chemical methods for soil analysis. New Zealand Soil Bureau Information Series 23. 39 p.

Clayden B, Hewitt AE 1989. Horizon notation for New Zealand soils. DSIR Division of Land and Soil Sciences Scientific Report 1. 30 p.

Cox SD ed. 2010. Geographic Information: Observations and Measurements. OGC Abstract Specification Topic 20 https://portal.opengeospatial.org/files/?artifact_id=41579, accessed 6 June 2012.

FAO-UNESCO 1974. Soil map of the World. 1:5 000 000. Vol. 1 Legend. Paris, UNESCO. 59 p.

Fielding RT 2000. Architectural styles and the design of network-based software architectures. Unpublished PhD thesis, University of California, Irvine. http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm, accessed 6 June 2012.

Griffiths E 1985. Interpretation of soil morphology for assessing moisture movement and storage. New Zealand Soil Bureau Scientific Report 74. 20 p.

INSPIRE Thematic Working Group SOIL 2011. D2.8.III.3 INSPIRE Data specification on SOIL – draft guidelines. http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_SO_v2.0.pdf [Accessed on 2012/06/06]

McArther, Giltrap 1984. SPG1 Data Base System Manual (V1.1) NZ Soil Bureau Laboratory Report SS4.

McDonald, Giltrap, McArther 1988. Revised SPG1 Data Base System Manual (V1.2) NZ Soil Bureau Laboratory Report SS16.

McQueen 1988. Glossary of soil water terms. Soil Bureau Laboratory Report EP31.

Milne JDG, Clayden B, Singleton PL, Wilson AD 1995. Soil description handbook. Lincoln, Canterbury, Manaaki Whenua Press.

Montanarella L, Wilson P, Cox S, McBratney A, Ahamed S, MacMillan B, Jacquier D, Fortner J 2010. Developing SoilML as a global standard for the collation and transfer of soil data and information. Geophysical Research Abstracts 12. http://eusoils.jrc.ec.europa.eu/esdb_archive/eusoils_docs/Poster/montanarella_EGU2010_XML.pdf, accessed 6 June 2012.

NADM Data Model Design Team 2004. NADM Conceptual Model 1.0 – A conceptual model for geologic map information. U.S. Geological Survey Open-File Report 2004-1334.

NERC 2011. Climate science modelling language v3.0. http://csml.badc.rl.ac.uk/, accessed 6 June 2012.

NWASCO 1975–1979. New Zealand Land Resource Inventory Survey, 1:63 360. Wellington, National Water and Soil Conservation Organisation.

NWASCO 1979. Our land resources. Wellington, National Water and Soil Conservation Organisation. 79 p.

OGC 2012. WaterML 2.0 Standards Working Group. http://www.opengeospatial.org/projects/groups/waterml2.0swg, accessed 6 June 2012.

Pourabdollah A, Leibovici DG, Simms DM, Tempel P, Hallett SH, Jackson MJ 2012. Towards a standard for soil and terrain data exchange: SoTerML Computers and Geosciences. http://dx.doi.org/10.1016/j.cageo.2011.11.026.

Robson M, Monaghan R, McDowall R 2011. Potential nitrogen and phosphorus losses from example farms in the Waituna catchment: sources and mitigation. AgResearch report for Environment Southland. 23 p.

Taylor NH, Pohlen IJ (1962) and (1979). Soil survey method. A New Zealand handbook for the field study of soils. New Zealand Soil Bureau 25. 242 p. [1962 version reprinted with amendments]

United States Department of Agriculture 1993. Soil survey manual. Agriculture Handbook No. 18. 437 p.

Whitton JG, Churchman JG 1987. Standard methods for mineral analysis of soil survey samples for characterisation and classification in New Zealand Soil Bureau. New Zealand Soil Bureau Scientific Report 79. 27 p.

Wilde RH ed. 2003. Manual for National Soils Database. Unpublished Landcare Research report.

Wilde RH, Ross CW 1996. New Zealand Reference Soil Collection and the National Soils Database. New Zealand Soil News 44: 224–227.

Willoughby EJ, Baker IR 1994. Standard input/output procedures for the National Soils Database using Paradox and Windows. Unpublished Landcare Research Report. 21 p + appendix.

Willoughby EJ, Wilde RH, McLeod M, Hewitt AE, Webb TW 2001. National Soils Database audit: A document describing Landcare Research soils data. Unpublished Landcare Research report.

W3C 2012. Semantic Web http://www.w3.org/standards/semanticweb/, accessed 20 February 2012.