Testing the representativeness of soil carbon data held in databases underpinning the New Zealand Soil Carbon Monitoring System

Hugh Wilde[1], Murray Davis[2], Kevin Tate[3] and David J. Giltrap[4]

[1]Landcare Research, Private Bag 11052, Palmerston North, New Zealand. www.landcareresearch.co.nz Email wildeh@landcareresearch.co.nz
[2]Forest Research, Private Bag 29237, Christchurch, New Zealand. www.forestresearch.co.nz Email murray.davis@forestresearch.co.nz
[3]Landcare Research, Private Bag 11052, Palmerston North, New Zealand. www.landcareresearch.co.nz Email tatek@landcareresearch.co.nz
[4]Small Office Systems Ltd, P O Box 46024, Lower Hutt, New Zealand. Email sos@actrix.gen.nz

Abstract
The New Zealand Soil Carbon Monitoring System involves stratification of the land area into cells based on climatic, soil and land-cover classes. Cell C values were obtained by averaging data contained in existing soil databases. Soil C data contained in these databases have come from carefully selected representative soils or from forest mensuration or trial plots, and were not randomly selected. There is thus a risk of bias in site selection, and a question about the representativeness of the underpinning data. This study tested the representativeness of underpinning databases by restricted random sampling of one cell (11 sites) using a grid-based system, and comparing results with the database values. The Temperate Volcanic Improved Pasture cell was selected for the test. This cell was already well sampled (29 points to 0.3 m depth) with a well-established mean and variance from the non-random sampling. Mean values for C mass derived from the two methods were similar, the randomly derived values being within 5.2%, 1.6%, and 1.4% of the database values for the 0–0.1 m, 0.1–0.3 m and 0–0.3 m layers respectively. With one exception, the randomly derived means for all three estimates were well within the 95% confidence limits of the database values. Database means for all attributes fell within the confidence limits derived from the random samples. These limits were wider than those of the database samples because we had fewer samples. We conclude that the original database soil C values, derived from non-random sampling, for this cell are representative.

Key Words
Soil carbon, representativeness, databases, underpinning

Introduction
The New Zealand Soil Carbon Monitoring System (CMS) was developed to estimate carbon in New Zealand soils and to quantify soil carbon changes following land-use change. Building the system involved stratifying the land area of New Zealand into soil, climatic, and land-cover classes (Scott *et al.* 2002). The New Zealand Land Resource Inventory (NZLRI) (Water and Soil Division 1979) was used to map the distribution of New Zealand soils at a scale of 1:1 000 000. Generalised soil polygons within the NZLRI were assigned to subgroups of the New Zealand Soil Classification (Hewitt 1998) and subsequently reclassified into six IPCC soil categories (IPCC 1996) based on clay activity, organic matter content, wetness, texture and presence of amorphous constituents (Daly and Wilde 1997). Podzols were added to the six IPCC categories because they are widespread throughout New Zealand. Temperature and moisture stratification was based on the USDA Soil Classification system (Soil Survey Staff 1999). Temperature was stratified into two categories: Cryic (<8 ^{O}C), and Mesic (8–15 ^{O}C), which included the Thermic regime (>15 ^{O}C) that is only of minor extent in New Zealand. Moisture categories were based on water balance, and included five categories. Ten categories of land use/land cover were based on the 1:1 000 000 scale Vegetative Cover Map of New Zealand (Newsome 1987) which recognised 47 vegetative cover classes including six grassland and eight forest classes. The six grassland and eight forest classes were subsequently reduced to two and three classes respectively for soil carbon sequestration work. Overall, 39 combinations of these three factors (soil/climate/land-use cells) were derived that describe 93% of the New Zealand landscape (Table 1).

Georeferenced soil carbon data (carbon concentration and bulk density) contained in existing soil databases was the primary source of data used to estimate average soil carbon for each of the 39 cells. Historically, most soil C data contained in these databases came from carefully selected 'representative' (or modal) soil pedons that were sampled and analysed as a part of soil survey operations (Landcare

Research database) or from forest mensuration or trial plots (Forest Research database). In neither case were sampling sites randomly selected; thus there is a risk of bias in site selection, with the possibility of some sites being preferred over others. A question is therefore posed about the representativeness of the existing database data underpinning the Soil CMS. The object of this study was to test the representativeness of database data underpinning the CMS by resampling one soil/climate/land-use cell using a grid-based sampling protocol, and comparing the results with the original database values.

Table 1: Seven soil classes, 7 combined climate classes and 10 land-use classes, when combined, provide 39 soil-climate-landuse combinations (cells) representing 93% of the New Zealand land area.

Soil classes	Climate classes			
	Temperature	Moisture	Combined	Land use classes
Organic	Cryic	Udic	Boreal	Horticulture
Aquic	Cryic	Perudic	Humid Boreal	Arable crops
High Clay Activity	Mesic	Aridic	Very dry Temperate	Improved pasture
Podzols	Mesic	Xeric	Dry Temperate	Unimproved pasture
Volcanic	Mesic	Udic	Moist Temperate	Shrubland
Low Clay Activity	Mesic	Perudic	Humid Temperate	Indigenous forest (Mixed)
Sandy	Mesic	Aquic	Aquic	Indigenous forest (Broadleafed)
				Exotic forest
				Wetlands
				Alpine

Methods

To test the representativeness of the existing database, the 'Temperate Volcanic soils under improved pasture' cell was selected. This cell was already well sampled (29 data points to 0.3 m depth from the underpinning database) with a well-established mean and variance. It also had a restricted national area (central North Island and Taranaki), making it more cost effective to sample than a cell of wide national distribution.

A 5 km x 5 km grid was used to select an adequate number of sample points. The grid was laid over the entire national area of Temperate Volcanic soils. Grid squares containing 30% or more of the target cell were numbered, and from these, 50 grid squares were randomly selected for sampling, and numbered in order of selection. The centre of the grid square was selected as the sampling point. In some cases this was not available because the site could not be sampled (water body, structure, road or non-pasture land cover) or was mapped as a different cell. In these cases the sampling point was moved to the centre point of the line bounding the north side of the grid square, and clockwise around the compass 90^{o} intervals until an available sampling point was obtained. The sampling point was initially located on a 1:250 000 NZMS 262 topographical map, and then transferred to a 1:50 000 NZMS 260 topographical map for locating the sample point in the field. The fifty randomly-selected grid squares were initially provided to ensure a sufficient supply of sample plots.

In the field it was found that a number of points had soils or vegetation that were not as mapped. In these cases, this point was noted, and then moved a random distance (selected using a calculator random number generator) between 100– and 500 m north of the original point. If this point was in turn not as mapped it was moved clockwise around the compass in 90^{o} intervals until a point with the correct soil and land-use classes was located. Sample sites and their selection details are shown in Table 2.

The soil sampling method used is described by Davis *et al.* (2004). A 20 m x 20 m plot was laid out and divided into four quarters and a point was randomly located in each quarter. For determination of soil C concentration , eight 25 mm diameter cores were taken in 0.1 m intervals to 0.3 m deep, at 0.25 m intervals, four on either side of the random point. Soil mass was determined using single, 98 mm diameter x 100 mm deep cores to 0.3 m depth at each random point. A soil profile description was made for the plot based on profiles exposed at each of the random points (i.e. four profiles per plot).

Published on CDROM. Website www.regional.org.au/au/asssi/

Table 2. Sample site locations and their selection details.

Site	Location	Sample point selection[1]	Point not as mapped for soil or for vegetation
1	Hawera 1	Grid centre not improved pasture but exotic woodlot. Point moved north	X
2 rejected	Edgecumbe	Whole grid square is now forest. Not sampled.	X
3	Broadlands	OK	
4	Otorohanga	Grid centre was pond. N, E & S points on adjacent property. Point moved west.	X
5	Kinloch	Grid centre not improved pasture but exotic woodlot. Point moved north.	X
6	Pukeokahu	Grid centre was road. N point not volcanic soil. Point moved east.	X
7	Moawhango	Grid centre not volcanic soil. Moved north.	X
8	Patea	OK	
9	Kapuni	Grid centre occupied by building. Point moved north.	X
10	Hawera 2	OK	
11	Taumarunui	OK	
12	Waiotapu	OK	

[1] See text for procedure used for moving selected sample points.

Results

Soil analysis

Mean values for C mass derived from the two methods were similar, the randomly derived values being within 5.2%, 1.6%, and 1.4% of the database values for the 0–0.1 m, 0.1–0.3 m and 0–0.3 m layers respectively (Table 3). The randomly derived means for the three carbon estimates were well within the 95% confidence limits of the database values. With one exception (bulk density of the 0.1–0.3 m layer), the random means for bulk density and C concentration were also within the 95% confidence limits of the database values. Because of fewer sample numbers, confidence limits of the random samples were wider than for the database samples and database means for all attributes fell within the confidence limits derived from the random samples.

Bulk density was much less variable than C concentration, as indicated by the lower coefficients of variation for bulk density (Table 3). This was true despite the fact that the random sample C concentration was derived from many more cores per plot (32) than bulk density (4). Comparison of the coefficients of variation also show that variability of all attributes was greater in the 0.1–0.3 m layer than the 0–0.1 m layer.

Table 3. Soil bulk density, C concentration and C mass values for the Temperate Volcanic cell as determined using the existing soil CMS database, and from random sampling in 2001.

Soil depth (m)	Attribute	Mean		Standard deviation		95% confidence limit		Coefficient of variation	
		Data-base	Random	Data-base	Random	Data-base	Random	Data-base	Random
0–0.1	BD (g/cm^3)	0.81	0.81	0.158	0.094	0.060	0.063	19.4	11.6
	C (%)	7.79	7.73	2.829	1.732	1.076	1.163	36.3	22.4
	C (t/ha)	60.0	63.1	15.205	17.51	5.784	11.77	25.3	27.7
0.1–0.3	BD (g/cm^3)	0.81	0.9	0.197	0.157	0.074	0.105	24.2	17.5
	C (%)	4.69	3.92	2.125	1.556	0.808	1.045	45.3	39.7
	C (t/ha)	71.50	70.3	28.407	28.69	10.806	19.28	39.7	40.8
0–0.3	C (t/ha)	131.5	133.4	41.56	45.61	15.508	30.64	31.6	34.0

Cell map accuracy

Sixty percent of the sampling sites, initially selected from a grid overlay, needed to be moved (Table 2). One sample point (Site 2) was totally rejected because although the cell map showed the site and the total 1 km^2 grid square to be pasture, it was in fact forest (Table 2). Two additional sites (1 & 5) were rejected but re-sited within the 1 km grid because field inspection showed the sample points to be woodlots

Published on CDROM. Website www.regional.org.au/au/asssi/

instead of pasture. Two more sites (6 & 7) were also rejected, but re-sited within the 1 km grid, because soil profile inspection showed their soils to be High Clay Activity soils rather than Volcanic soils in terms of the IPCC soil classification. Thus five of twelve sites (42%) were wrongly mapped. This incorrect mapping would have occurred partly because of relatively recent conversion from pasture to forestry, and the forests would either not have been present at the time of mapping, or were very young. Incorrect mapping of these sites is therefore considered to be relatively unimportant from the soil C sequestration standpoint. A further two sites (4 & 9) were rejected because one was a pond and the other a building. These sites were also re-sited within the 1 km grid.

Incorrect soil mapping was probably a result of the scale of soil maps used to produce the cell map. Errors in soil type will have serious implications if the soil C concentration of the soil type encountered in field sampling differs substantially from that of the mapped cell. In the present study the rejected 'soil-type' sites were both Moist Temperate, High Clay Activity Soils. These soils have substantially less C than Moist Temperate Volcanic Soils (cell means are 93 and 131 t C/ha respectively for the 0–0.3 m layer). However, one of the rejected sites occurred in a swampy depression where soil C would be higher than average, and if this site had been sampled it could have balanced the lower value of the other site.

Conclusion

Means of soil C mass from 11 randomly-sampled sites within the Moist Temperate Volcanic soils under improved pasture were found to fall within 5.2%, 1.7% and 1.4% of the database means for the 0–0.1 m, 0.1–0.3 m and 0–0.3 m layers respectively. All soil C mass values also fell within the 95% confidence limits of the database values. Consequently, the initial database soil C values derived from non-random sampling, and used to predict C content of this cell, are considered to be representative.

References

Daly, BK, Wilde RH 1997: Contribution of soil carbon to New Zealand's CO_2 emissions. I. Reclassification of New Zealand soil series to IPCC categories. 9 p. and one map. (Landcare Research Contract Report LC9697/096).

Davis MR, Wilde RH, Garrett L, Oliver G (2004). New Zealand Carbon Monitoring System Soil Data Collection Manual. 56 p. Prepared for: (New Zealand Climate Change Office Ministry for the Environment, PO Box 10-362, Wellington).

Hewitt AE 1998: New Zealand Soil Classification. Landcare Research Science Series No. 1. p 133. (Manaaki WhenuaPress, Landcare Research, Lincoln, New Zealand).

IPCC (1996). Land-use change and forestry. In *'Intergovernmental Panel for Climate Change. Revised 1996 guidelines for national greenhouse gas inventories: reference manual'* pp. 5.6–5.75. (Intergovernmental Panel for Climate Change, Bracknell, UK).

Newsome PFN1987: Vegetative Cover of New Zealand. Water & Soil Miscellaneous Publication No. 112. 153 p. and 2 map sheets. (National Water and Soil Conservation Authority, Wellington, New Zealand).

Scott NA, Tate KR, Giltrap DJ, Smith CT, Wilde RH, Newsome PJF, Davis MR (2002). Monitoring land-use change effects on soil carbon in New Zealand: quantifying baseline soil carbon stocks. *Environmental Pollution* ***116***, S167–S186.

Soil Survey Staff (1999). 'Soil Taxonomy. A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Second Edition'. p 869. (United States Department of Agriculture Natural Resources Conservation Service. Agriculture Handbook No. 436).

Water and Soil Division 1979: Our Land Resources. (Water and Soil Division, Ministry of Works and Development, Wellington).