An Assessment of the Quality of Data Stored in the National Vegetation Survey Database, With Recommendations for Minimising Errors

S.K. Wiser, P.J. Bellingham, D.A. Coomes, L.E. Burrows, and R.F.S. Gordon

Landcare Research
P.O. Box 69, Lincoln
New Zealand

Landcare Research Contract Report: LC9899/139

DATE: July 1999

Contents

1. Summary

1.1 Project and Client

New Public Good Science Funding for the National Vegetation Survey (NVS) database was provided to developing quality assurance protocols that provide for updating and editing of data, and to recognise data quality and origin. Development of quality assurance depends on an assessment of data quality, and recommendations for ensuring high quality of newly collected data, and these are presented in this report. The project was carried out by Landcare Research, Lincoln, between January – June 1999. The report was funded by the Foundation for Research, Science and Technology as part of Landcare Research's Land Resources Information Systems programme, under contract CO9626.

1.2 Objectives

Our objectives were to give an assessment of the extent of errors in NVS data sets, based on automated checks. Time-consuming corrections of erroneous data were not made at this stage. The following checks were made on the NVS data.

a. Identify plots that have identical grid coordinates, where this should not be the case (i.e. exclosures and their controls may have the same coordinates).

b. Check the consistency of plot site data, i.e. data on localities and physical site information. For example, are these consistent (where they should be) across different types of data collected in the same year (e.g. sapling vs. recce file), and across different survey years?

c. Assess the level of taxonomic precision (examining in particular genera or families that often cause difficulties in the field, e.g., grasses and the genera *Uncinia*, *Carex*, and *Hymenophyllum*).

d. Check the extent of changes in tree identifications of tagged trees from census to census (those likely to be errors and those that result from nomenclatural changes).

e. Check a newly collected field data set for reliability of earlier plot locations, site data, and establish the reliability of plot dimensions.

f. Use a newly collected field data set to quantify errors that arise from measuring and marking trees, and to determine the taxonomic reliability of current and earlier plot measurements.

1.3 Methods

We assessed the quality of data in the electronic NVS database, by summarising data from permanent forest plots, and evaluating these data for spatial accuracy, for site information, for errors in species identification across surveys, and for apparent errors in growth increment data for tagged trees. We resurveyed 25 permanent forest plots in montane rain forest in Westland to evaluate errors in a single data set. Here we quantified the reliability of plot locations in the field, of plot dimensions, of site data, and the capacity to relocate permanent seedling subplots. In this data set we quantified errors that arise from tagging and measuring trees, and we evaluated the taxonomic reliability of earlier surveys.

1.4 Results

Spatial location data were available for 96% of 7564 permanent forest plots. Most plots have fairly complete site information, although mean top height has often not been measured. Discrepancies in site information are generally few between seedling plots and tree plots in the same location, except in the case of mean top height values. Taxonomic precision is highest in tree plot data, with 99% of measurements identified to species level. Tree growth data is reliable as 97% of stems grew at credible rates (i.e. between -5 and 5 mm/year trunk diameter growth). In the test data set from Westland, plot locations in the 1999 census were on average within 130 m of previous location data. Site data were generally measured at similar values to earlier measurements from the same plots, except aspect values when slope was <5E. Plot dimensions were very reliable in the field: of 14 plots measured the actual mean plot area was 397 m^2 (cf. intended 400 m^2). Permanent seedling subplots could be relocated readily in most plots: on average 22 of 24 subplots per plot were relocated. Taxonomic reliability of permanent tagged stems ∃2.5 cm trunk diameter was high: only 1.6% of *c*. 2800 stems were incorrectly identified. Taxonomic misidentifications were more prevalent in seedling subplots, especially of sedges, grasses and some ferns.

1.5 Conclusions

Evaluation of the overall database and of a case study dataset from the NVS database gives confidence in the accuracy of data from permanent plots, especially with respect to taxonomy of tree, tree growth rates, and location data. There is scope for improvement in recording of site data for plots, and in taxonomic accuracy for non-woody species, especially ferns, grasses and sedges.

1.6 Recommendations

We recommend that similar evaluations be conducted of other major data sets contained in the NVS database. Improved spatial information can be achieved using GPS. Improved metadata are required in NVS on limits to accuracy and measurement of individual surveys. We recommend several ways in which both existing data and new field measurements can be improved. We recommend that guidelines be prepared in light of this report to encourage improvement of field measurements in future plot remeasurements.

2. Introduction

Even careful researchers make occasional mistakes when collecting and processing data, and such errors occur in even the best of databases. The numbers of new errors entering a database can be reduced by developing formal quality assurance protocols for adding, updating and editing data. There is also a need to screen data for errors before embarking on analyses. The importance of removing such errors almost goes without saying, because they can otherwise lead to spurious results and misleading conclusions.

The purpose of this report is to identify types of error that occur in the National Vegetation Survey (NVS) database, quantify the extent of these errors, and make recommendations on how the problems may be rectified. We concentrate on a subset of the NVS data, i.e. that collected from permanent plots in New Zealand's indigenous forests, although many of the error types examined will be common across most vegetation data. Permanent forest NVS plots were originally intended to sample forest dynamics, particularly for the effects of introduced mammalian herbivores (in this case, the pervasive effects of red deer (*Cervus elaphus*) and brushtail possums (*Trichosurus vulpecula*)). The plots can still address the issues for which they were established (e.g., Bellingham *et al*. 1999), but are increasingly used for additional purposes not originally foreseen. These include reporting on carbon storage in indigenous forests (Hall & Hollinger 1997; Hall *et al.* 1998; Coomes *et al*. 1999), and examination of weed invasion in forests (Wiser *et al*. 1998). It is also anticipated that they will be used to address national strategies in coming years, e.g., the Biodiversity Strategy (Department of Conservation and Ministry for the Environment 1998). At no stage in its history has NVS found so many applications, and consequently the need for quality assurance has become paramount.

Some types of errors can be assessed by performing diagnostics on the existing database, and this is the focus of the first section of the report. For example, when more than one file pertains to the same plot (recce, seedling, sapling, and tree data), the site information should be identical in each, and there should be data for the same categories of site information. Similarly, a tagged tree should have the same species code at every enumeration date, and trees should grow at a plausible rate. The database was checked for errors of this kind, and the percentage of erroneous data is given. Given adequate resources, most of the existing errors can be rectified. We also recommend ways of minimising the chances of such data entering the database in future.

Other types of error are less easy to recognise simply by examining existing data. For example, how often are species misidentified in the surveys, especially those in taxonomically difficult genera? For this reason, a second test of the accuracy and reliability was made by remeasuring a set of permanent plots in montane rain forests. We chose a set of 25 permanent forest plots established in 1972 in the Whitcombe River valley, central Westland (43E05'S, 171E01'E) (James *et al*. 1973), for which all existing information resides in the NVS database. These plots are representative of many of the *c*. 9000 such plots established nationally, in that they representatively sample a catchment with an extensive spatial layout. Problems with errors increase with time between enumerations. We chose the plots in the Whitcombe River valley to test whether a set of plots that had not been visited in 19 years could still be relocated and yield useful data. We chose to test this in an area which might be considered fairly extreme compared with many examples elsewhere in New Zealand, because it is an area of extremely high rainfall (>8 m per year, Griffiths & McSavaney 1983), and has a high disturbance frequency due to flooding and landslides that occur on an unstable schist bedrock, close to the Alpine Fault. We chose a forest system in which substantial changes due to canopy mortality or disturbance events might be expected. We also chose these plots to assess taxonomic reliability of data in NVS collected from permanent plots because the forests are species rich and show high variation over small distances (Reif & Allen 1988).

3. Background

3.1 What is the National Vegetation Survey?

The National Vegetation Survey (NVS) database is a physical archive and computer database containing records from *c.*100,000 vegetation survey plots—including data from over 11 000 permanent plots and *c.* 8000 Protected Natural Areas Programme (PNAP) survey records. NVS provides a time-series record, spanning more than 50 years, for both indigenous and exotic plants in terrestrial ecosystems, from Northland to Stewart Island and the Chatham Islands, with broad habitat coverage, especially indigenous forest and grasslands.

The database is a unique, nationally important collection of vegetation data and includes plot sheets, maps, and photographs from the surveys. As such it represents many years of field surveys, data compilation and checking, as well as reformatting to ensure ready accessibility, reliability and comprehensive coverage. Software has been specifically prepared for statistical analysis and is an integral part of NVS.

The NVS database has been built up from surveys conducted by the New Zealand Forest Service, Department of Lands & Survey, and the DSIR Botany Division. Data are constantly being added to NVS from surveys and research conducted by the Department of Conservation, Regional Councils, Universities and Landcare Research. The value of NVS to the Nation is enhanced by making such widely sourced information available through modern computing and communication facilities. But at the same time, the interests of data providers are protected through written agreements that determine access rights to specific data sets within NVS.

3.2 The value of NVS Information

NVS data support reporting requirements for the *Biodiversity Convention*, *Framework Convention on Climate Change*, *Resource Management Act*, *State of Environment*, and the *Montreal Process*. They also assist resource management. The historical nature of NVS information has great significance in enabling New Zealand to address issues unforeseen at the time of data collection, such as the impacts of climate change on indigenous ecosystems and the storage of carbon in indigenous ecosystems. Recent examples of NVS information use include the following:

- the assessment of carbon storage in New Zealand's indigenous forests (Hall & Hollinger 1997);
- possible changes in forest tree distribution following global warming (Leathwick 1995);
- the elevational distribution of conifer-broadleaved hardwood forests on the South Island (Allen *et al.* 1991);
- the historical assessment of damage to the Kaweka forest area by deer (Allen & Allan 1997);
- the dynamics of kanuka forest (Smale *et al.* 1995);

- setting priorities for controlling possums in the Nelson-Marlborough and Westland areas (Rose 1996);
- measurement of changes caused by possum and deer browsing, natural disturbance (wind, fire, earthquake), or human disturbance;
- predicting how weed invasion in an indigenous forest changes over time (Wiser *et al*. 1998);
- assessing the regeneration of snow tussocks in alpine grasslands (Rose & Platt 1990).

3.3 What kinds of vegetation data does NVS contain?

1. *General survey data* from locations that have not been permanently marked suitable for understanding species distributions and for studies needing only coarse measurement of changes in vegetation. In excess of 100 000 survey plot assessments in NVS comprise:
    - point-based compositional (and usually) structural description of vegetation
    - relative abundance in fixed structural tiers (usually included)
    - locational information (>50% have NZMS grid references).

    Reconnaissance plots, Protected Natural Areas programme plots are typical of these data.
    Examples of Potential Uses:
    Quantification of local and widespread plant species richness, hence use in recommendations for designating protected areas, planning, etc. These data underpin the formulation of environmental domain maps.

2. *Permanent plot data* where fixed area plots or transects have been established, and the vegetation has been measured precisely (e.g., tagged trees, species lists, stereophotographs). Assessments of *c*.11 000 permanent plots in NVS are ideal for monitoring vegetation changes and the effects of management.
    - Nearly all follow standard methods, where all trees within a standard area (usually 400 m^2) are permanently tagged to allow repeat measurements
    - Most contain permanently marked seedling subplots to determine changes in seedling and herbaceous composition with time.
    - Most are along randomly located transects.
    - More than 75% have NZMS grid references.

    Examples of potential uses:
    - Forest Plots: absolute changes in abundance, growth, mortality, by species.
    - Grassland and Shrubland Plots: change in structure and composition.

3.4 Software Packages for Analysis of NVS Data

Analytical software has been tailored to NVS data for typical vegetation assessment purposes (Hall 1992, 1994*a*,*b*, 1996). All packages run on stand-alone PCs, using data entered in a standard ASCII text file format.

For analysing vegetation composition and vegetation communities in the landscape (usually from vegetation survey data).

PC-RECCE Analyses reconnaissance (recce) plots, where vegetation information includes measures of plant species composition, and usually of relative abundance. Data are often collected from vertical tiers, and relative abundance of species quantified within tiers in abundance or cover classes. Data typically analysed using this package include standard Recce plots (methods: see Recce survey manual, Allen 1992), and Protected Natural Areas Programme survey plots.

For long-term or short-term vegetation monitoring, using permanent fixed-area plots or transects.
(a) Forest plots
PC-DIAM Analyses vegetation composition, based on absolute abundance or basal area of woody species from measurements of stem diameters. If stems have been tagged for repeated measurements (methods: see forest plot manual, Allen 1993), changes in growth, mortality and recruitment with time can be quantified and contrasted among species (e.g., of tree species palatable to possums and those not palatable).

PC-USTOREY Analyses changes in abundance and growth of saplings and seedlings with time from data collected from permanent plots (methods: see forest plot manual, Allen 1993). Comparisons can be made among species (e.g., useful in assessing browsing by mammals – frequently used in contrasting differences between exclosure plots and adjacent control pots).

(b) Grassland plots (also useful for shrublands, and non-woody communities, e.g., salt marshes, cushion communities, wetlands, etc.)

PC-TRANSECT Analyses changes in abundance of species and changes in vertical structure from permanent plots or transects using permanent fixed area plots or height-frequency transects (methods: grasslands survey manual, Wiser & Rose 1997). Comparisons can be made among species (e.g., in response to changing management, or invasion by exotic species).

All packages are capable of converting data to formats for widely used vegetation classification and ordination programmes (e.g., CANOCO, ter Braak 1987), and the classification package TWINSPAN (Hill 1979) is included in all packages.

4. Objectives

Our objectives were to give an assessment of the extent of errors in NVS data sets, based on automated checks. Time-consuming corrections of erroneous data were not made at this stage. The following checks were made on the NVS data.

a. Identify plots that have identical grid coordinates, where this should not be the case (i.e. exclosures and their controls may have the same coordinates).
b. Check the consistency of plot site data, i.e. data on localities and physical site information. For example, are these consistent (where they should be) across different types of data collected in the same year (e.g. sapling vs. recce file), and across different survey years?
c. Assess the level of taxonomic precision (examining in particular genera or families that often cause difficulties in the field, e.g., grasses and the genera *Uncinia*, *Carex*, and *Hymenophyllum*).
d. Check the extent of changes in tree identifications of tagged trees from census to census (those likely to be errors and those that result from nomenclatural changes).

e. Check a newly collected field data set for reliability of earlier plot locations, site data, and establish the reliability of plot dimensions.

f. Use a newly collected field data set to quantify errors that arise from measuring and marking trees, and to determine the taxonomic reliability of current and earlier plot measurements.

5. Methods

5.1 Quality assessment of electronic data

Data from NVS forest plots are currently being applied in a number of national and regional studies, so our analyses place an emphasis on this data (including diameter-, understorey-, and recce- data) . Grassland and PNA data have not been assessed, as no computer programs have yet been written that can access more than one data file simultaneously.

Data were read from ASCII formats into a SAS database (SAS Institute 1989), and programs were written by one of us (SKW) to test for each of the error types outlined below. For broad-scale comparisons, results of quality assessment are summarised by Department of Conservation Conservancy. We recognise that summaries could also be done at the level of individual data files; and this information would be useful to add to the REFLEX file that holds the metadata associated with the data files.

5.2 Remeasurement of permanent forest plots

The plots in the Whitcombe River valley are in montane conifer/broadleaved hardwood rain forests. These plots, when established in 1972, were intended to serve as a comparison with nearby plots in the Kokatahi River valley (James *et al*. 1973). Both the Whitcombe and Kokatahi valleys are in the Hokitika River catchment, and when the Kokatahi valley plots were established in 1972, extensive mortality of canopy trees was underway. In contrast, forest canopies were largely intact in the Whitcombe River valley. Plots in the two valleys were thus intended to allow comparisons between forest dynamics in a valley where canopy mortality (believed to be due to browsing by possums) was advanced (Kokatahi) and one where canopy mortality was not apparent. A remeasurement of the plots in the Whitcombe River valley 27 years after their establishment should allow evaluation of whether forest dynamics still differed substantially between the two catchments.

The plots were established in 1972 along randomly located compass lines, and established at regular altitudinal intervals (*c*. every 200 m above sea level) from the valley floor until timberline or subalpine shrubland was reached. All but one of the plots were remeasured in 1980, and at that time permanent seedling subplots were also established in each plots (methods follow Allen 1993). In 1999, transect origins were remarked with red permanent marking, and lines remarked. All transects and all plots along transects were relocated without great difficulty. Markers denoting the centres of animal pellet counts that were established in 1971 were at 1 chain (*c*. 20 m) intervals along transects and these greatly facilitated plot relocation. One transect (line 22) was poorly marked through its middle sections, as the transect crossed unstable terrain where few large trees could be marked. Plots were re-established by searching for corner markers of the 20 x 20 m plots, then 5 x 5 m subplots were established by locating seedling pegs that lie along the edges of the 5x5 m subplots (see Allen 1993), so that as close as possible to an exact reconstruction of the 1980 plot layout was achieved. In 14 of 25 plots, measurements of outer perimeters were made to the nearest 0.1 m. Where corner pegs could not be relocated, new corner pegs were placed 20 m from an existing corner peg. Exact plot areas were calculated subsequently from the area of a quadrilateral, using exactly measured distances along the 4 outer perimeters (to the nearest 0.1 m).

Remeasurement of 25 permanent plots took place over the summer of 1998 – 1999. Six plots along two lines (9 and 11) were remeasured in November 1998, and the remaining 19 plots along six lines (7, 8, 10, 17, 19, and 22) were remeasured in January – February 1999. One plot on another line (plot 20/1) in Vincent Creek was not measured; this was established in 1972, but not remeasured in 1980.

GPS determined plot locations were recorded for 12 of 25 plots (i.e, plots 7/1, 7/3, 8/1 (nearby), 8/2, 8/3, 10/2, 17/1, 19/1, 19/2, 19/4, 22/2, and 22/4). Map references were derived in the field for each plot from NZMS 260 K34. Plot altitudes were measured by altimeters. Slopes were measured with Suunto clinometers. Aspects were measured as uncorrected magnetic values. For each plot, a shelter index was measured (angle to horizon, measured by clinometer from plot centre at 45E angles around compass, McNab 1993). Mean top height was assessed as mean height over entire plot where canopy was of variable heights (e.g., included treefalls).

Methods for remeasurement of permanent 20 x 20 m plots followed Allen (1993), i.e.:

- all stems ∃2.5 cm dbh were tagged and measured (many earlier tagged stems, especially those in 1972 establishment tagged very low, near the base of trees rather than the standard 1.4 m – this usually noted on plot sheets). New data were recorded for all tree ferns (predominantly *Cyathea smithii*) – all were tagged, at 1.4 m, dbh recorded and height recorded to nearest 0.1 m;
- a full enumeration was conducted of saplings (>135 cm tall and <2.5 cm dbh) per 5 x 5 m subplots;
- all seedling subplots established in 1980 were relocated where possible, and 0.75 m^2 subplots were remeasured. New seedling subplots were established when thorough searches with metal detectors failed to locate the original seedling pegs. Exact height measurements were usually recorded of *Podocarpus hallii*, *Prumnopitys ferruginea*, *Libocedrus bidwillii*, *Metrosideros umbellata* and *Weinmannia racemosa* seedlings >15 cm tall in these plots. Seedlings were permanently tagged for future remeasurement in seedling plots in plots 9/1 and 9/2 (seedlings of *Podocarpus hallii*, *Pseudowintera colorata*, *Quintinia acutifolia* and *Weinmannia racemosa*).

Since the 1980 measurement of the plots, two plots (8/3 and 10/1) are now overlapped in part by new plots established by L. Batcheler and D. Craib in 1985: the upper half of 8/3 is overlapped by a larger 40 x 40 m plot centred above, and 10/1 is overlapped by 20 x 20 m plot, in the north-east corner. In both cases the result is that many trees in these two plots (especially 8/3) bear 2 tags – one from the original 1972 plots remeasured during this survey, and a newer tag from the plots established in 1985. To avoid more "double-tagged" trees for the future, stems that were new since the 1980 remeasurement of the original plots, but which were tagged when new plots were established in 1985, were regarded as new recruits in the 1999 measurement, and the tags used in 1985 were recorded.

Recce plots were conducted within a defined search area of the 20 x 20 m permanent plots; therefore data is unlikely to be directly comparable with 1972 and 1980 recces, although the centres of those recce plots should be the same as the 1999 measurement. Recces were conducted according to methods of Allen (1992). Voucher specimens were collected of plants that could not be identified in the field for later determination, using keys and reference collections in the CHR herbarium. Voucher specimens collected from the plots are lodged in the CHR herbarium.

6. Results

6.1 Quality assessment of electronic data

6.1.1 Spatial location information

Complete grid coordinates (to the nearest 100 m northing and easting) have been recorded on 96 % of the permanent forest plots that have diameter data (Table 1). Where there are no maps indicating plot locations for a survey, grid coordinates may never have been entered in the data or may be incomplete. For some of these surveys, plot locations may be marked on aerial photos, and these could be used as a basis for determining grid coordinates.

Lack of precision in the recorded grid coordinates means that plots that are within 100 m of each other will have the same coordinate, even though they do not occur at the exact same locality; this affects 2.6 % of the plots (Table 1). Examples of this include exclosure plots and their controls, subplots of the large plot at Orongorongo and subplots of the Hunua transects.

Duplicate grid coordinates for different plots may also indicate true errors (Table 1) of three main types. The first is a mistake in determining a plot's grid coordinate. The second is that a plot has not been named consistently from year to year so that it is unclear that a remeasured plots is actually the same as the original (e.g. F 31 vs. F31). The third is that plot header information was incorrectly combined with plot data information at the time of data entry. These true errors affect only 30 plots or 0.4% of the total, and these are currently being corrected.

Considerable efforts have been made to complete and correct the grid coordinates on the diameter data from permanent forest plots. A similar assessment should be made on the other forms of data.

6.1.2 Site information other than spatial location

Table 2 provides a summary of the site information missing from the files containing diameter data, understorey data and total composition data; data in the first two categories has usually been collected only from permanent plots; the latter may be from either permanent plots or point-in-time surveys.

There are very plots for which the year is missing; this is relatively straightforward to add. In most instances, altitude, slope and aspect is recorded on the plot records of a given data type for at least 95% of the plots in a Conservancy; but there are some exceptions (e.g. over 15% of the Recce plots surveyed in Northland and Hawke's Bay-East Coast have no altitude entered; over 10% of the diameter plots in Northland , Wanganui, Wellington and West Coast lack slope data). Site records are more commonly incomplete for physiography, drainage, ground cover assessments and mean top height.

Table 1 Summary of grid coordinate information on permanent forest plots. The number of plots in each category is provided. A total of 7564 plots were determined from this analysis

Conservancy	Coordinates missing	Coordinates incomplete	Different plot numbers for same coordinate		Coordinates complete and unique
			Likely error	Precision (e.g., exclosures & controls	
Auckland	0	0	0	80	35
Bay of Plenty	0	23	6	25	107
Hawke's Bay - East Coast	6	36	58	114	578
Northland	0	0	0	8	0
Tongariro-Taupo	4	10	0	16	216
Waikato	0	4	0	24	119
Wanganui	0	0	0	412	171
Wellington	0	2	2	83	472
Canterbury	28	31	10	2	1513
Nelson-Marlborough	64	0	2	18	1596
Otago	0	2	0	0	340
Southland	14	0	6	97	1103
West Coast	107	2	24	88	766
TOTAL	223	110	30	185	7016
Percentage	3.2	1.6	0.42	2.6	92

In many cases this information exists, but was not copied over from data from previous censuses or was not entered in the headers for all of the data categories of a census (i.e. entered on the diameter records and not the understorey records).

Table 3 shows where there is inconsistency in site information among records of files of different types (e.g., diameter vs. understorey vs. recce) that was collected from the same permanent plots at the same time. The rate of this type of error is relatively low, typically affecting fewer than 5% of the plot measurements in a Conservancy. The error rate is highest in the records for the Canterbury and West Coast Conservancies, where inconsistencies in individual fields may be as high as 13 – 15% respectively, and for the field mean top height across all conservancies.

This inconsistency could stem from two sources. First, as described above, the data may have not been entered on all file types for a survey. This is relatively easy to correct and is likely the reason for most of the mismatches. More problematic is where different values have been entered in files of different types. Here, original data sheets will need to be checked to ascertain the true values.

Table 2 Proportions of records missing specific types of site information

Conservancy and File Type	No. Records	Year	Altitude	Aspect	Slope	Physio-graphy	Drainage	Ground cover	Mean Top Height
Auckland									
Diameter	192	0	<1	1	1	14	68	77	74
Recce	205	0	<1	1	1	3	<1	0	3
Understorey	53	0	11	8	8	16	4	17	17
Bay of Plenty									
Diameter	233	0	2	3	2	<1	18	34	8
Recce	605	0	<1	2	<1	0	0	<1	77
Understorey	187	0	2	4	2	1	15	35	35
Hawke's Bay - East Coast									
Diameter	949	0	<1	13	1	20	29	26	2
Recce	2771	0	17	3	5	5	6	4	52
Understorey	601	0	2	5	16	16	29	28	19
Northland									
Diameter	8	0	0	0	25	25	0	0	0
Recce	392	0	0	<1	1	<1	<1	0	<1
Understorey	24	0	0	0	25	13	0	0	0
Tongariro-Taupo									
Diameter	304	0	<1	5	4	3	4	8	10
Recce	1298	0	11	4	2	3	1	1	2
Understorey	228	0	5	8	4	3	4	7	4

Conservancy and File Type	No. Records	Year	Altitude	Aspect	Slope	Physio-graphy	Drainage	Ground cover	Mean Top Height
Waikato									
Diameter	248	0	2	4	3	4	24	48	3
Recce	1219	0	10	6	6	20	32	49	31
Understorey	293	<1	8	5	4	6	10	30	10
Wanganui									
Diameter	627	0	0	12	12	7	<1	3	2
Recce	1876	0	9	2	4	4	<1	3	4
Understorey	279	2	<1	5	5	15	<1	3	15
Wellington									
Diameter	1126	0	<1	13	27	24	15	24	10
Recce	2170	0	18	6	12	10	3	9	6
Understorey	924	0	<1	4	18	15	6	17	5
Canterbury									
Diameter	4486	0	<1	5	2	9	34	34	36
Recce	8380	0	1	2	2	4	15	12	7
Understorey	4074	0	<1	6	2	8	33	37	35
Nelson-Marlborough									
Diameter	2162	0	<1	3	1	1	4	10	3
Recce	5689	0	1	3	4	8	6	7	15
Understorey	2221	0	<1	3	2	3	6	12	7
Otago									
Diameter	544	0	<1	3	3	2	1	1	<1

Conservancy and File Type	No. Records	Year	Altitude	Aspect	Slope	Physio-graphy	Drainage	Ground cover	Mean Top Height
Recce	809	0	<1	1	1	<1	<1	6	4
Understorey	548	<1	2	3	3	2	2	2	1
Southland									
Diameter	1787	0	3	9	8	8	11	13	10
Recce	6531	<1	11	3	2	5	3	10	6
Understorey	1888	<1	3	7	7	7	10	15	13
West Coast									
Diameter	1492	0	7	9	10	16	26	42	24
Recce	14399	0	<1	2	2	2	10	5	13
Understorey	924	0	<1	3	3	4	14	28	18

Table 3 Percentage of mismatches by field among site records of data of different types (e.g., diameter vs. sapling vs. recce) for the same survey

Conservancy	Number of Records	Year	Altitude	Aspect	Slope	Physiography	Drainage	Ground cover	Mean Top Height
Auckland	49	0	2	0	0	2	0	0	23
Bay of Plenty	216	0	<1	0	<1	0	<1	<1	38
Hawke's Bay - East Coast	822	<1	<1	<1	<1	1	<1	<1	18
Northland	12	0	0	0	0	0	0	0	2
Tongariro-Taupo	315	2	1	5	4	3	2	4	5
Waikato	291	0	<1	0	0	8	<1	<1	8
Wanganui	274	0	<1	<1	2	<1	<1	<1	10
Wellington	1116	0	<1	<1	<1	2	<1	<1	4
Canterbury	4191	2	13	12	12	8	2	8	5
Nelson-Marlborough	2212	<1	9	1	<1	<1	<1	<1	5
Otago	547	0	<1	0	<1	0	0	<1	4
Southland	1783	<1	6	2	1	3	<1	1	3
West Coast	981	7	9	8	15	13	12	11	12

Specific checks were run on selected site fields. When a plot is sampled on a flat surface there can be no aspect. Missing values for aspect have been indicated by either entering '999' or leaving the space blank. Across all diameter, recce and understorey surveys there are a total of 5073 header records with slopes of 0E. Of these, 99 have the aspect entered as '999', 365 have the aspect column left blank, 3420 have aspect entered as '0', and 1189 have a value for aspect.

Slopes greater than 70^{o} are unlikely to have been sampled. Plot sheets were examined where such high slopes were recorded in headers for diameter files and entries corrected where necessary. This check has yet to be made on headers in recces, understorey or grassland data.

6.1.3 Precision in species identifications

The precision of identifications on plot records is uneven across surveys, and depends on the botanical skills of the people who made the measurements, and the maturity of the plant material observed in the field (see Section 6.2.6). We assessed the taxonomic precision on all plots. The proportion of stems and recorded presences in subplots of taxa identified no further than to genus was determined for recces, diameter plots and understorey plots (sapling and seedling) respectively. Results are summarised by Conservancy and file type in Table 4. Taxonomic precision is very high for diameter and sapling data; within conservancies generally plots average having 99% of the measurements identified to species level. The sapling and diameter measurements on permanent plots have the advantage of accrued knowledge over time, and taxonomic difficulties are less pronounced in trees than in other groups (e.g., ferns, graminoids). Trees are permanently tagged, so identities can be checked and improved in precision. Saplings are often long-lived, so although not tagged, identities can be rechecked at the same location. Recce data and data from seedling subplots is less precise. Still, on average, plots in most conservancies have more than 90% of their species measurements identified to species level.

Examining the range of taxonomic precision observed across plots highlights instances where the taxonomic precision is low. This is especially pronounced when the number of taxa encountered in a plot were low, and these were not all identified to species.

Table 4 Precision of species identification. The percentage of the total species in a plot identified to genus only (Recces), or occurrences (Diameter) or presences in subplots (Sapling, Seedling) was determined, and averaged across all plots in a Conservancy. The range in this figure is also shown

	Recce			Diameter			Sapling			Seedling		
Conservancy	mean	range	no. plots	mean	range	no. plots	mean	range	no. plots	mean	range	no. plots
Auckland	0.02	0-0.21	205	<0.01	0-0.20	192	<0.01	0-1.00	49	0.03	0-0.23	53
Bay of Plenty	0.05	0-0.31	604	<0.01	0-0.13	233	<0.01	0-0.14	181	0.19	0-0.81	187
Hawke's Bay - East Coast	0.07	0-0.53	2684	0.01	0-0.95	947	0.04	0-0.87	559	0.18	0-0.75	586
Northland	<0.01	0-0.14	392	0.01	0-0.03	8	<0.01	0-0.06	15	0.01	0-0.08	24
Tongariro-Taupo	0.09	0-0.89	1292	<0.01	0-0.18	304	0.01	0-0.38	140	0.13	0-0.53	228
Waikato	0.03	0-0.22	1218	<0.01	0-0.25	248	0.01	0-0.44	291	0.03	0-0.45	293
Wanganui	0.06	0-0.50	1876	0.02	0-0.56	627	<0.01	0-0.50	195	0.05	0-0.48	278
Wellington	0.09	0-0.74	2014	<0.01	0-0.23	1126	0.01	0-0.45	859	0.22	0-0.76	1083
Canterbury	0.04	0-1.00	8365	<0.01	0.13	4484	<0.01	0-1.0	1542	0.05	0-1.0	4024
Nelson-Marlborough	0.05	0-1.00	5684	<0.01	0.18	2162	0.01	0-1.0	1501	0.12	0-0.79	2162
Otago	0.05	0-0.48	809	<0.01	0.30	544	<0.01	0-0.67	399	0.18	0-0.8	533
Southland	0.04	0-0.57	6529	<0.01	0.19	1787	<0.01	0-0.23	1474	0.16	0-0.83	1881

6.1.4 Growth Increments

Trees can increase in stem diameter by up to 20 mm per year under ideal conditions, but few trees in native forests are expected to grow more than 5 mm per year. It is not uncommon for trees to shrink between measurements, in part because of random error in measuring tree diameters, but also because of natural phenomena such as the flaking off of bark. The data set was scanned for trees that increased or decreased in size by unrealistic amounts. Table 5 provides data on tagged trees that grew or shrunk > 50 mm/year. In general there are very few errors of this kind (0.8% of tree diameter records). The entire database could be corrected for this type of error.

Across all conservancies, 97.2% of trees were found to grow within the range -5 to + 5 mm/year. Figure 1 shows that most of the remaining 2.8% appeared to grow (or shrink) by less than 5–10 mm/year. The number of trees growing 5 – 10 mm/year is rather high, and probably contains some erroneous data, but these trees are impossible to distinguish from those that are genuinely fast growing. This is a type of random error – sometimes a larger diameter is recorded than the actual size, while sometimes a smaller diameter is recorded, but on average the measurements are unbiased. The only method of removing the effects of random errors of this kind is to reenumerate the plots several times, allowing long-term growth rates to be untangled from measurement error.

Table 5 Percentage of tagged trees that appeared to shrink or grow by at least 49 mm/year. There was a total of 333 312 trees in the analysis

Conservancy	<–49 mm/year	>49 mm/year
Auckland	0.00	0.00
Bay of Plenty	0.02	0.05
Canterbury	0.00	0.00
Hawke’s Bay-East Coast	0.06	0.11
Nelson-Marlborough	0.04	0.00
Otago	0.13	0.08
Southland	0.13	0.08
Tongariro-Taupo	0.30	0.17
Waikato	0.01	0.04
Wanganui	0.05	0.02
Wellington	0.03	0.15
West Coast	0.11	0.10
Total	0.88	0.80

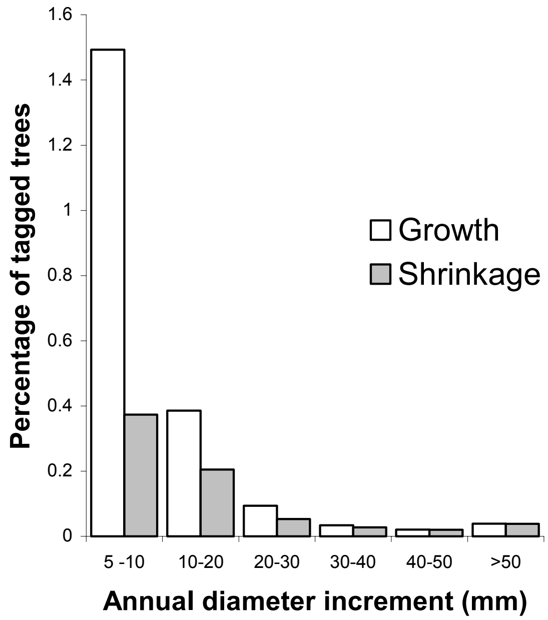

Fig. 1. The percentage of tagged trees that appeared to grow, or shrink, by at least 5 mm/year. Note that most trees (97.2%) grew within the -5 to + 5 mm/year range, which is not shown.

6.1.5 Mismatches of tree species codes

Table 6 provides data on the number of trees that have different species codes on two enumeration dates, and is based on all trees that were measured at least twice (34.5%). Over all, 1.4% of trees changed their species codes. These errors could occur for one of three reasons: (a) nomenclature changes (e.g., many of the changed names of podocarps, Connor & Edgar 1987), (b) the tree was incorrectly identified at the first enumeration, but the correct species code was not backdated, (c) the wrong species codes was entered because there are ambiguities in the codes (e.g., *Pseudowintera colorata* and *Pseudopanax colensoi* both abbreviate to PSECOL, when the latter should be entered as NEOCOL).

6.2 Quality in remeasurement of permanent forest plots

6.2.1 Reliability of plot locations

Differences in grid references in the Whitcombe River valley, Westland, between 1980 references (also derived in the field) were compared with those made in 1999. Exact comparisons were difficult because 1980 references were derived from NZMS 1 series maps (imperial 1:63 000) while those in 1999 used metric NZMS 260 series maps (1:50 000). For comparisons, imperial 1980 grid references were converted to NZMS 260 grid references using Grid reference conversion program version 2.0.4 (Ross Pickard, DoC Science & Research, 1998). Because of differences in scale, some conversion errors were inevitable (e.g., one plot was placed on the opposite bank of the river from its true location). Only one plot had exactly the same NZMS 260 grid reference in 1980 as in 1999. The mean planimetric distance between 1980 grid references and 1999 grid references was 130 ± 20.5 (standard error of the mean) m (range 0 – 460 m).

Table 6 Number of trees that have different species codes between enumerations (‘mismatches), compared with the number that have consistent codes.

Conservancy	No. mismatches	No. correct matches	% mismatched
Auckland	46	430	10.7
Bay of Plenty	68	6645	1.0
Canterbury	349	165 571	0.2
Hawke’s Bay-East Coast	889	8058	11.0
Nelson-Marlborough	389	43 115	0.9
Otago	186	11 816	1.6
Southland	1349	33 889	4.0
Tongariro-Taupo	149	2980	5.0
Waikato	120	9405	1.3
Wanganui	343	4111	8.3
Wellington	637	27 398	2.3
West Coast	286	23 573	1.2
Complete data set	4811	165 571	2.9

6.2.2 Reliability of site data

Data quantifying site variables in 1980 was apparently applied retrospectively to 1972 data, as the data are identical. New measurements of site variables were made in 1999.

- Altitude, measured by altimeters in 1999, were 20 ± 5.0 m lower than on plot records of 1980 (range from 80 m lower than 1980 records to 40 m higher).
- Slope, measured by Suunto clinometers in 1999 from plot centres, were 1±0.8E greater than 1980 measurements (range from 6E less than 1980 measurements to 9E greater).
- Aspect was measured as uncorrected magnetic bearing of the predominant physiography of the plot from its centre in 1999. Five of 25 plots had the same aspect as that recorded in 1980. On five plots with little slope (#5E), measured aspects were sometimes nearly diametrically the opposite of those measured in 1980 (e.g., 285E *vs* 90E, 270E *vs* 105E); the average difference between measurements on these five plots was 95±30E. Thus use of aspect values from plots with little slope should be interpreted with caution or not used in analysis, though in combination with slope, calculation of potential solar radiation (after Frank & Lee 1966) would be less affected by the very different values for aspect. On the 20 plots of greater (>5E) slope, differences between aspects measured in 1980 *vs* 1999 were 20±3.5E (range: 0-55E).

Mean top height: In 1999 this was estimated over the entire 20 H 20 m plot, including treefall gaps. In 1980 it is not clear how mean top height was estimated; however it was probably assessed similarly but within an unbounded reconnaissance plot that overlapped to some extent with the 20 H 20 m plot. Thus estimates may not be strictly comparable. Moreover, in contrast to site variables, biotic variables such as mean top height will change as young vigorous stands increase in height, or as older stands die. The mean top height estimated in 1999 was 4±1 m lower among the 25 plots than that estimated in 1980 (range from 1 m taller than 1980 estimates to 14 m less than 1980 estimates).

6.2.3 Reliability of plot dimensions

Plot dimensions were checked in 14 of 25 permanent plots. All four corner pegs were located in six of these 14 plots, a single peg was missing in four plots, and two corner pegs were not relocated in another four plots. The mean plot area of plots in which all four corner pegs were relocated was 397 ± 12.7 m^2 (Table 7), and that of all 14 plots (including those in which at most two corner pegs were replaced) was 397 ± 5.79 m^2 (compared with the intended 400 m^2).

Table 7 Plot dimensions of permanent forest plots in montane rain forest in the Whitcombe River valley, Westland

Number of corner pegs found	Number of plots	Mean plot area (m^2) ± SE
4	6	397 ± 12.7
3 – remaining peg substituted exactly 20 m from nearest original pegs	4	387 ± 2.25
2 – remaining 2 pegs substituted exactly 20 m from nearest original pegs	4	408 ± 2.13
Overall	14	397 ± 5.79

6.2.4 Capacity to relocate permanent seedling subplots

Twenty-four seedling subplots were established in 1980 in all but one of 23, 20 × 20 m plots (21 subplots were established in one plot (10/3) where a stream course bisects the plot). In both 20 × 10 m plots, 10 seedling subplots were established in 1980. In the 1999- remeasurement, 2 ± 0.7 seedling plots per plot (or 11 ± 2.9% of seedling subplots per plot) were lost between 1980 and 1999 and were replaced. Of these, 1 ± 0.3 seedling plots per plot (or 2 ± 0.1% of seedling subplots per plot) were lost due to major disturbance such as landslides or flooding that would probably have eliminated the seedling flora in the subplot.

6.2.5 Errors that arise from marking and measuring of adult tree stems

Two kinds of errors result from previous measurements of adult stems in plots. The first is that stems ∃2.5 cm dbh were tagged, but not measured, in either the earliest measurement (1972) or in both earlier measurements (1972 and 1980). Among 25 permanent plots, stems tagged but not measured in 1972, and first measured in 1980 represented 3.4 ± 0.63% (range: 0 – 10.8%) of total stems per plot in 1980. Stems tagged but not measured in 1972 or 1980, and first measured in 1999 represented 0.3 ± 0.1% (range: 0 – 2.3%) of total stems per plot in 1999.

The second kind of error that results is when a new tree tag was attached to a stem in the 1980 measurement that already bore a tag from the 1972 measurement. This occurred when stem diameter increment had caused the tag applied in 1972 to be engulfed. When this occurred, this was usually noted on the field sheets in 1980, but because the electronic data format in NVS does not currently allow any such notes to be recorded, the effect was that the 1972 tag number received no data in 1980, and this is normally surmised to mean mortality of that stem (and the computer program PC-DIAM written for NVS data treats lack of current diameter data to mean mortality). Conversely, the new

1980 tag is treated as a new stem, i.e. the result of diameter growth of a stem too small to be tagged in 1972, but which grew to become of a size able to be tagged in 1980. As a result, without scrutiny of the original field sheets, it would be possible substantially to overestimate mortality rates and recruitment rates per plot, and this would be even more true for individual species which do have rapid diameter growth rates. In the 1999 measurement, a metal tape was carried in the field that allowed replacement of the original tag with the original number. Where possible the earliest (i.e. 1972) tag was reapplied, and electronic data were amended in both 1972 and 1980 data sets to show a constant tag number for all diameter measurements. Of the total stems per plot in 1980, 1.6 ± 0.57% (range: 0 – 6.8%) received a new tag in 1980 on a stem already tagged in 1972. Calculation of tree mortality rates (*sensu* Sheil *et al*. 1995) in all plots between 1972 and 1980 was 1.5%/yr using uncorrected data, and was 1.4%/yr using corrected data. Calculation of tree recruitment rates (*sensu* McCune & Cottam 1985) in all plots between 1972 and 1980 was 1.9%/yr using uncorrected data, and was 1.8%/yr using corrected data. Although these discrepancies are small among all plots, at an individual plot level, discrepancies resulting from retagging of trees were much greater (e.g., mortality rates of 2.5%/yr (uncorrected) *vs*. 1.2%/yr (corrected) and 2.2%/yr *vs*. 1.8%/yr; recruitment rates of 1.9%/yr *vs*. 0.9%/yr and 1.4%/yr *vs*. 1.1%/yr in plots 10/2 and 22/1 respectively).

6.2.6 Taxonomic reliability of data

Stems ∃2.5 cm dbh: Dubious or incorrect identifications accounted for 45 stems (1.6%) of 2770 stems tagged in 1972, and 33 stems (1.1%) of 2973 stems tagged in 1980. Details of these are stems identified in field sheets as:

- *Coprosma parviflora*. This species only occurs north of 36ES (Wilson & Galloway 1993). Stems identified as this species are mostly misidentifications, in both 1972 and 1980, of an entity not recognised in 1972, now known by the informal name of *Coprosma* sp. "t", (but also likely to include some misidentified *Coprosma ciliata*). Those stems determined in earlier measurements as *C. parviflora* and still alive in 1999 were all determined as *Coprosma* sp. "t". In the 1999 census every effort was made to distinguish *Coprosma* sp. "t" from *Coprosma ciliata*, with which it is often confused (see Appendix by D.S. Glenny in Coomes *et al*. 1999).
- *Pittosporum eugenioides*. While the distribution of *P. eugenioides* includes the study area, none were seen during the survey. Three of four stems determined in 1972 and 1980 as were redetermined in 1999 as *Pittosporum tenuifolium* (includes *P. colensoi*).
- *Coprosma linariifolia*. Six stems were determined as this species in 1972, of which 4 remained alive in 1980. This species is mainly found east of the Main Divide in the South Island (Wilson & Galloway 1993). Of the four stems alive in 1980, all were identified as *Coprosma pseudocuneata* in 1999.
- *Raukaua edgerleyi*. While the distribution of *R. edgerleyi* includes the study area, none were seen during the survey. The single stem determined as *R. edgerleyi* in 1972 and 1980 was not alive in 1999.

One of 2973 stems tagged in 1980 was determined to genus only (*Gaultheria* sp.) – determination to species level should have been possible even with vegetative material; the stem was not located in the 1999 measurement. Another stem tagged in 1980 was not determined to genus or species, and was not located in the 1999 measurement.

Some stems were affected by taxonomic changes between their earlier measurements and 1999. Of stems tagged in 1972, such changes affected 219 stems (7.9% of 2770 stems) of six species (13% of

48 species), and of the total number of tagged stems in 1980, such changes affected 216 stems (7.3% of 2973 stems) of six species (12% of 51 species). The species affected by such changes were:

- *Coprosma banksii*, now included in *Coprosma colensoi* (Wilson & Galloway 1993);
- *Dacrydium biforme*, for which *Halcarpus biformis* is now the recommended name (Connor & Edgar 1987);
- Various species previously included in the genus *Pseudopanax*. Taxonomic revisions in the New Zealand Araliaceae by Mitchell *et al*. (1997) reinstate the genus *Raukaua*, hence *Raukaua anomalus* replaces *Pseudopanax anomalus*, *Raukaua edgerleyi* replaces *Pseudopanax edgerleyi* and *Raukaua simplex* replaces *Pseudopanax simplex*.
- *Podocarpus dacrydioides* for which *Dacrycarpus dacrydioides* is now the recommended name (Connor & Edgar 1987).

Seedlings and other species in seedling subplots: In 1980, 136 taxa were recorded on seedling subplots. Of these:

- Taxonomic changes affect 17 taxa, but do not introduce problems in interpreting field data (e.g., *Todea superba* becomes *Leptopteris superba*). Another taxon, *Coprosma banksii*, is synonymous with *Coprosma colensoi* and the two taxa should be merged in comparisons with 1999 data.
- One taxonomic change (taxonomic revision of the *Blechnum* "capense" group, Chambers & Farrant 1998) means that 1980 data may not be comparable with data collected in 1999 (as many as three taxa are recognised from this group in the area in 1999, as well as hybrids).
- Two non-vascular species were recorded in 1980 (i.e., *Dendroligotrichum dendroides* and "moss") – it is not convention to record non-vascular species on plots, and none were recorded in 1999.
- Twelve taxa were listed only as genus in 1980, and some were common (e.g., *Hymenophyllum* sp. recorded from 22.4% of all seedling subplots). Of the taxa identified to genus level only, it is likely that lack of field identification skills, or lack of later determination from vouchered herbarium records, precluded determination to species level in 11 of the 12 taxa. Although later remeasurements from the same seedling plots may allow educated guesses to be made about the identity of species identified to genus only at earlier surveys, this is risky because many of the plants may have short life spans (e.g., herbs, grasses and some sedges) and it is possible that a species in a genus recorded at an earlier census may have died and been replaced by another species in the same genus at a later census.

  Of note among taxa identified to genus level only are *Uncinia* spp. Only two *Uncinia* taxa were recognised in 1980 (*U. uncinata* and *U*. 'fine'), and in nine plots, taxa on seedling subplots were identified only as *Uncinia* sp. In contrast in 1999, seven *Uncinia* taxa were recognised on seedling plots (all supported by voucher specimens), and the overall frequency of *Uncinia* species increased from being recorded on 119 seedling plots in 1980 to 191 seedling plots in 1999 (a 61% increase). While it is conceivable that there has been an increase in both diversity and abundance of *Uncinia* species over the 19 years between measurements, it is more likely they were overlooked in the 1980 measurement because of taxonomic uncertainty. In an appendix in Coomes *et al*. (1999), Landcare Research plant systematist Dr David Glenny also made pertinent observations about difficulties in determining *Uncinia* specimens collected from permanent plots: "This is clearly the worst genus for field workers and for botanists doing identifications alike. I compiled a 15- character interactive key to *Uncinia* when I saw how many specimens there were to be done, but it was of limited value because the published descriptions are quite poor. That is, maxima and minima of leaf widths are not at all comprehensive."

In 1980 some distinctions in the genus *Uncinia* were made in the field in identification as tag names, but these tag names were not supported by voucher specimens. One of these can be resolved with some confidence in remeasurements, i.e. in 1980 in many higher altitude (>700 m a.s.l.) seedling plots a taxon was identified as *Uncinia* 'fine', which almost certainly corresponds with *Uncinia filiformis*, determined and supported by voucher specimens in the 1999 remeasurement. However, in two high altitude plots in which *Uncinia filiformis* was common in seedling plots in 1999, no *Uncinia* species were recorded at all in seedling plots in 1980. It is possible that the species has colonised these two plots in the intervening 19 years, but it is more likely they were overlooked because of taxonomic uncertainty. For other *Uncinia* species recorded in 1980 at genus level only, it is not possible to resolve which species these are likely to have been.

In the 1999 measurement, identification of some material collected in November was not possible because the material lacked key characters. This points to the need to measure plots at seasons when material can be most readily identified.

- As noted above, grasses, especially those of short life span, are prone to misidentification. The comparative abundance of *Poa annua* in seedling subplot records in 1980 (14 records) *vs* none in seedling subplots in 1999 (although it was recorded in a reconnaissance plot in one plot) may be due to spurious accuracy in 1980 records, i.e. any unknown grass of *Poa*-like appearance was referred to this species. Similarly, the absence of grasses recorded in 1999 (e.g., *Deschampsia tenella*, seven records) may be due to it either being overlooked or aggregated with other species in the 1980 measurement. For future surveys, voucher specimens of most grasses should be collected.
- Identification of one taxon (*Alseuosmia macrophylla*) in 1980 is almost certainly incorrect: this is most likely to be misidentified *Alseuosmia pusilla*, collected from plots and supported by a voucher specimen in 1999. It is probable that identification of another taxon (*Hydrocotyle moschata*) in 1980 was incorrect; voucher specimens from the same seedling plot in 1999 were determined as *Hydrocotyle microphylla*. Identification of a taxon as *Phormium tenax* in a high altitude (845 m a.s.l.) plot in 1980 is suspect: in the 1999 census only *Phormium cookianum* was recorded from plots, but no seedlings of any species of *Phormium* were recorded from the plot in which it was recorded in 1980, so it is not possible to be sure this was an error.
- Distinctions made in 1980 between two taxa (*Parsonsia capsularis* and *Parsonsia heterophylla*) are not possible in seedlings (the two species can only be separated on floral characters) and the two should be condensed to a single entity (no distinction was made in 1999 and data were recorded as *Parsonsia* sp. only).

In 1999, 142 taxa were recorded on seedling subplots, only a small increase (4%) over total species richness in seedling subplots in 1980. Of these the taxa recorded on seedling subplots in 1999, eight were identified to genus only, including:

- sterile material (e.g., of *Pterostylis* spp. (on nine subplots), and of *Uncinia* sp. (on one subplot));
- seedlings were too small to be identified with certainty, i.e., of *Coprosma* spp. (on five subplots), *Astelia* spp. (on two subplots), *Blechnum* spp. (on one subplot) and *Clematis* spp. (on one subplot));

- taxonomically undescribed material (an undescribed species of *Gentiana* on one seedling subplot; a voucher specimen from near this subplot is being used in a current PGSF-funded revision of this genus);
- where specific distinction is not possible on juvenile material. As described earlier, distinctions between seedlings of the two *Parsonsia* species is not possible without flowering material, and *Parsonsia* seedlings occurred on 25 seedling subplots.

Voucher specimens in the field were collected and later determined. Of 65 specimens lodged as voucher records in the CHR herbarium, three could not be determined with certainty to species. Database links are possible from the CHRP database to individual NVS plots because among details entered from each voucher record (i.e. on both herbarium labels and the electronic database) are the permanent plot number and its NZMS 260 grid reference. A likely benefit is that the site of defined search area (400 m2) can be revisited with some confidence that living material could be found at the same site in future. Likewise a detailed ecological interpretation could be derived from the description of the site that could aid identification of suitable habitat for a given species. Similar links are possible between the NVS database and the CHRP database because a personal collection number (e.g., PJB 962) is denoted alongside specimens for which voucher specimens were collected on the hard copy field sheets. At present there is no data field in the NVS electronic database into which to record either personal collection numbers or CHR numbers for voucher specimens.

7. Recommendations

7.1 General recommendations

- Checks similar to those conducted in this report need to be made on other databases included in NVS (e.g., grassland data).
- Additional guidelines for plot remeasurements should be prepared based on the findings in this report, to be inserted into field manuals, and provided to all end-users of NVS data desiring to remeasure permanent plots.
- Increased precision of grid coordinates for plots should be strongly encouraged, as well as use of GPS in field surveys.
- The results of electronic quality checks, such as those in this report, should be included into metadata associated with individual surveys.

7.2 Reducing errors in stem measurements in permanent forest plots

7.2.1 Future surveys

Every effort should be made in the field to reconcile stems which have received different tags in different surveys. Since radial diameter growth of stems (rapid for some species) results in tags being occluded, a "Dymo" metal tape that permits reapplications of original tags should be mandatory equipment for field crews. Every effort should be made to reinstitute the earliest tag number applied to a stem, and notes should be made on field sheets describing how tag numbers for the same stem may have changed between censuses.

7.2.2 Existing data

An additional data column might be added to show a different tag number for the same stem where these have changed between surveys, and in preparation of field sheets for new surveys to highlight the need to scrutinise stem identification.

At present, the data structure does not allow two stems in the same plot to bear the same tag number. However, on repeat measurements, it is quite likely that a new series of tags could include tag numbers already used in a plot. The current error checking routine that does not permit the same tag number to occur twice in a plot should be retained to allow warnings. Additionally, there would be use in allowing more than the current maximum of 4 digits for a tag number and to allow an alpha/numeric option for tags as well. Often tags have up to 6 digit or 6 character sequences, and inclusion of these additional data would often redress apparent duplicate tag numbers, or avoid apparent data ambiguity.

Relocation of stems in the field would be assisted if pro-forma data sheets incorporating earlier measurements also included the 5 H 5 m subplots in which stems were last recorded. At present, no data field exists to nominate subplots.

7.3 Improving the accuracy of species identification

7.3.1 Future surveys

In resurveys of permanent plots, reconnaissance plots should be delimited to sample the same area as the permanent plots, i.e. a complete floristic coverage within a fixed area.

Improved taxonomic reliability from permanent plots can be achieved by the following steps (partly adapted from David Glenny's comments in an appendix to Coomes *et al*. 1999):

- Each voucher should have a jeweller's tag with the plot number on it.
- Supply to plant systematists working on voucher specimens a list of plots and where they are, as to make identifications some idea about habitat and altitude is important. All interactive keys have province as a character, sometimes the best character available for a specimen.
- Small "field guides" should be compiled for field workers. For example, for *Chionochloa*, it would help if field workers had a laminated photocopy of the key in Connor (1991) revision, as the key is easy to use in the field. This is also clearly required for traditionally problematic genera such as *Hymenophyllum* and *Uncinia*.
- Spurious accuracy can be applied to juveniles and seedlings (e.g., in distinguishing *Parsonsia* species, as above). Another example is of seedlings of trees in the genus *Raukaua*, in which seedlings of *R. simplex* and *R. edgerleyi* cannot be distinguished (P.B. Heenan, pers. comm.), so while in some cases field workers may be able to guess their identity from adults in the vicinity, in other cases this may not be possible.

7.3.2 Existing databases

There is a need:

- to ensure cross-referencing between the various data sources collected from a permanent plot. All species recorded on data sheets from stems ∃2.5 cm dbh, from saplings and from seedling subplots should be cross-referenced to ensure that they appear on the reconnaissance plots, and that consistent names have been used across data sheets as different data sets are often collected by different members of a field team.

- for cross-referencing between NVS and herbarium databases (e.g., CHRP). In CHRP database a field for plot identifier is required as well as grid reference, and in NVS a field is required for an herbarium voucher record number.
- to establish metadata files that contain information about changes in nomenclature, making specific references to survey that are affected.
- to back-date corrections of species codes to all previous enumerations, noting the changes on the original data sheet, but expunging incorrect information from the database.
- to allow storage of complete species names and ready links to databases providing nomenclature changes.
- to flag the data set for ambiguous species codes, e.g., PSECOL.

7.4 Minimising errors in numerical data

7.4.1 Future surveys

When remeasuring established plots:

- produce a preprinted data sheet for recces with a complete list of species determined from last measurement.
- verify site parameters (altitude, slope, aspect, grid coordinates (with GPS))
- test estimated mean top heights vs actual measured with Vertex measuring device. Compare with previous assessments. For some forest types and development stages it will be clear if previous estimates were incorrect.
- provide the survey team with metadata describing anomalies in earlier surveys.

7.4.2 Existing databases

- An interface should be written so that trees that grow by dubiously large amounts are flagged *at the time of data entry,* allowing data to be checked immediately for errors. Existing software only provides checks on data formatted as an NVS ASCII file.
- The whole database should be systematically checked for trees that appear to grow or shrink by 20 mm/year or more. There are only around 1000 trees in this category.
- A provision is required to denote whether a stem has died or is a new recruit.
- Aspect should be corrected where values are 0E.
- All files pertaining to a given plot should be cross-referenced so that site data are consistent.

8. Acknowledgements

We thank the Department of Conservation's West Coast Conservancy for issuing permits to collect voucher material, and for assistance, especially from Phil Knightbridge in remeasuring plots. We also thank Phil Suisted, Deb Zanders, Kate Orwin, and Simon Burrows for assistance in remeasuring permanent plots. We thank Michelle Breach for assistance in providing data, and for entering new data from field sheets.

9. References

Allen, R. B. 1992: RECCE: an inventory method for describing New Zealand's vegetative cover. *Forest Research Institute Bulletin 181*. 25 p.

Allen, R. B. 1993: A permanent plot method for monitoring changes in indigenous forests. Landcare Research, Christchurch. 35 p.

Allen, R. B.; Allan, C. 1997: Mountain beech forest dynamics in the Kaweka Range and the influence of browsing animals. *Science for Conservation 44*. 23 p.

Allen, R. B.; Reif, A.; Hall, G. M. J. 1991: Elevational distribution of conifer-broadleaved hardwood forests on South Island, New Zealand. *Journal of Vegetation Science 2*: 323–330.

Bellingham, P .J.; Wiser, S. K.; Hall, G. M. J.; Alley, J. C.; Allen, R. B.; Suisted, P. A. 1999: Impacts of possum browsing on the long-term maintenance of forest biodiversity. *Science for Conservation 103*. 59 p.

Chambers, T. C.; Farrant, P. A. 1998: The *Blechnum procerum* ("capense") complex in New Zealand. *New Zealand Journal of Botany 36:* 1–19.

Connor, H. E. 1991: *Chionochloa* Zotov (Gramineae) in New Zealand. *New Zealand Journal of Botany 29*: 219–283.

Connor, H. E.; Edgar, E. A. 1987: Name changes in the indigenous New Zealand flora, 1960–1986 and Nomina Nova IV, 1983–1986. *New Zealand Journal of Botany 25*:5–170.

Coomes, D.; Allen, R.; Partridge, T.; Scott, N.; Glenny, D.; Wiser, S.; Savage, T.; Buxton, R.; Beets, P.; Nordmeyer, A.; Goulding, C.; Robertson, K. 1999: Monitoring carbon in indigenous forests and scrublands. Task 2 report: a test of methods and logistics along a south-island transect. Landcare Research/New Zealand Forest Research Institute contract report (JNIxxx/xxx) for the Ministry for the Environment.

Department of Conservation and Ministry for the Environment. 1998: New Zealand's biodiversity strategy: our chance to turn the tide. Wellington, Department of Conservation and Ministry for the Environment. 142 p.

Frank, E. C.; Lee, R. 1966: Potential solar beam irradiation on slopes: tables for 30E to 50E latitude. U.S. Forest Service Rocky Mountain Forest Experiment Station General Technical Report RM-131. Fort Collins, Colorado.

Griffiths, G. A.; McSaveney, M. J. 1983: Distribution of mean annual precipitation across some steepland regions of New Zealand. *New Zealand Journal of Science 26*: 197–209.

Hall, G. M. J. 1992: PC-RECCE. Vegetation inventory data analysis. *Forest Research Institute Bulletin 177.* 108 p.

Hall, G. M. J. 1994a: PC-DIAM: Stem diameter analysis. Auckland, Landcare Research. 88 p.

Hall, G. M. J. 1994b: PC-USTOREY: Seedling and sapling data analysis. Auckland, Landcare Research. 44 p.

Hall, G. M. J. 1996: PC-TRANSECT: Vegetation transect data analysis. Lincoln, Landcare Research. 74 p.

Hall, G. M. J.; Hollinger, D.Y. 1997: Do the indigenous forests affect the net CO2 emission policy of New Zealand? *New Zealand Forestry* 41: 24–31.

Hall, G.; Wiser, S.; Allen, R.; Moore, T.; Beets, P.; Goulding, C. 1998: Estimate of the carbon stored in New Zealand's indigenous forest and scrub vegetation for 1990. Landcare Research/New Zealand Forest Research Institute contract report (JNT 9798/147) for the Ministry for the Environment. 57 pp.

Hill, M.O. 1979: TWINSPAN. A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Section of Ecology and Systematics. Ithaca, New York, Cornell University. 79 pp.

James, I. L.; Jane, G.; Barr, C. 1973: The forests and sub-alpine shrublands of the Hokitika catchment. *Forest & Range Experiment Station, Protection Forestry Division Report No. 116.*

Leathwick, J. R. 1995: Climatic relationships of some New Zealand forest tree species. *Journal of Vegetation Science 6*: 237–248.

McCune, B.; Cottam, G. 1985: The successional status of a southern Wisconsin oak woods. *Ecology 66*:1270–1278.

McNab, W. H. 1993: A topographic index to quantify the effect of mesoscale landform on site productivity. *Canadian Journal of Forest Research* 231:100–1107.

Mitchell, A. D.; Frodin, D. G.; Heads, M. J. 1997: Reinstatement of *Raukaua*, a genus of the Araliaceae centred in New Zealand. *New Zealand Journal of Botany 35*: 309–315.

Rose, A. 1996: Possum impacts and threats to conservation values in Nelson/Marlborough Conservancy. *Proceedings of the New Zealand Ecological Society Conference, July 1996.*

Rose, A. B.; Platt, K.H. 1990: Age-states, population structure, and seedling regeneration of *Chionochloa pallens* in Canterbury alpine grasslands. *Journal of Vegetation Science* 1: 89–96.

Reif, A.; Allen, R. B. 1988: Plant communities of the steepland conifer-broadleaved hardwood forests of central Westland, South Island, New Zealand. *Phytocoenologia 16*: 145–224.

SAS Institute 1989.: SAS/STAT user's guide. Version 6, 4th edition, vols 1 and 2. Cary, North Carolina, SAS Institute.

Sheil, D.; Burslem, D. F. R. P.; Alder, D. 1995: The interpretation and misinterpretation of mortality rate measures. *Journal of Ecology 83*: 331–33.

Smale, M. C.; Hall, G. M. J.; Gardner, R. O. 1995: Dynamics of kanuka (*Kunzea ericoides*) forest on South Kaipara Spit, New Zealand, and the impact of fallow deer (*Dama dama*). *New Zealand Journal of Ecology 19*: 131–141.

ter Braak, C. J. F. 1987: CANOCO – a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). Wageningen, The Netherlands, TNO Institute of Applied Computer Science.

Wilson, H.D. & Galloway, T. 1993. Small-leaved shrubs of New Zealand. Manuka Press, Christchurch. 307 p.

Wiser, S.K., Allen, R.B., Clinton, P.W. & Platt, K.H. 1998. Community structure and forest invasion by an exotic herb over 23 years. *Ecology 79:* 2071-2081.

Wiser, S.K. & Rose, A.B. 1997. Two permanent plot methods for monitoring changes in grasslands: a field manual. Landcare Research, Christchurch. 51p.