REVIEW ARTICLE

Managing biodiversity information: development of New Zealand's National Vegetation Survey databank

Susan K. Wiser, Peter J. Bellingham and Larry E. Burrows
Landcare Research, P.O. Box 69, Lincoln 8152, Canterbury, New Zealand (E-mail: wisers@landcare.cri.nz)

Abstract: The National Vegetation Survey (NVS) databank is designed to safeguard the investment of millions of dollars spent over the last 50 years collecting, computerising and checking New Zealand vegetation data and to optimise the potential knowledge gains from these data. Data such as these can be synthesised across a range of spatial and temporal scales, allow novel ecological questions to be considered, and can underpin land management and legal reporting obligations. The NVS databank builds largely on the base of data collected under the auspices of the New Zealand Forest Service from the 1940s to 1987. In more recent years, it has incorporated data from Protected Natural Area (PNA) surveys and from new and remeasured plots in a range of ecosystems collected by staff of, among others, the Department of Conservation, Landcare Research, regional councils and universities. The databank currently stores data from approximately 14 000 permanent plots, 52 000 reconnaissance descriptions and PNA plots, and 14 000 timber volume plots measured in the 1940s and 1950s. Ecosystems that are best represented are grasslands in montane and alpine areas and indigenous forests. Geographic coverage is widespread but patchy. As the NVS databank continues to develop and grow, a range of data management issues are being addressed. These include (i) developing mechanisms to meet the needs of both data users and data providers and incentives to encourage individuals and organisations to deposit data into the databank, (ii) ensuring that metadata are adequate to allow raw data to be interpreted, and (iii) ensuring that the data stored meet set quality standards. In the future, the databank will take advantage of changing technology to best meet the needs of data users and providers. Further information about the NVS databank can be obtained from www.landcare.cri.nz/science/nvs.

Keywords: biodiversity information; data archive; data management; database; environmental monitoring; information science; metadata; New Zealand; permanent plots; vegetation.

Introduction

Data from a broad range of scales are vital if we are to address many of the issues at the forefront of ecology (Michener *et al.,* 1997). Such issues often require more data than an individual or team of researchers could collect. In New Zealand, syntheses of broad-scale data have been used to formulate and test hypotheses about factors controlling vegetation structure and composition (e.g., Holloway, 1954; McKelvey, 1963; Osawa and Allen, 1993; Leathwick *et al.,* 1998; Bellingham *et al.,* 1999). More recently, such syntheses have allowed national-scale issues to be considered that were not anticipated at the time of data collection, such as how much carbon is stored in indigenous forests (e.g., Hall *et at.,* 2001). The existence of long-term data from permanent plots has allowed consideration of questions about forest dynamics (e.g., Mark *et al.,* 1991; Smale *et al.,* 1995), weed invasion in grasslands and forests (e.g., Scott, 1993; Wiser *et al.,* 1998), and grazing impacts in non-forest ecosystems (e.g., Dickinson *et al.,* 1992).

Worldwide, efforts are ongoing to ensure that vegetation data are well documented, archived and made accessible (Table 1). In New Zealand, such efforts are seen as increasingly important by agencies that fund data collection, or use such data to support policy decisions and assess compliance with legal obligations (Whitehouse, 1998). New Zealand has a range of international reporting requirements as a signatory to the Convention on Biological Diversity and the Framework Convention on Climate Change, and as a participant in the Forest Resource Assessment of the Food and

Table 1. Examples of international efforts to document, archive and increase access to vegetation data. Vegetation data may be the main emphasis in some or included among a range ecological data. Data include both those from one-off surveys and from permanent plots.

Organisation and Project	Scope	Internet address (URL)
Ecological Monitoring and Assessment Network (EMAN, Canada)	Provides a metadata search facility to allow searches for ecological data sets available from around the world.	http://metadata.cciw.ca/search/main_e.html
National Biological Information Infrastructure (NBII), U.S.A.	Electronic gateway to biological data and information maintained by U.S. federal, state and local government agencies and private sector organisations and other parties around the world.	http://www.nbii.gov/
Ecological Archives of the Ecological Society of America	Stores data sets of ecological significance described in, or supplemental to, papers published in *Ecology, Ecological Monographs,* and *Ecological Applications.*	http://esapubs.org/archive
TROPIS - Tree growth and permanent plot information system	Maintains a searchable index of people and institutions worldwide that hold permanent plot data in both plantations and natural forests.	http://www.cifor.cgiar.org/tropis/
U.S. Long Term Ecological Research (L TER) Program	Umbrella organisation for 24 research sites in the U.S.A. Sites independently manage their own long-term data (e.g., from permanent vegetation plots, animal censuses, climate data).	http://ltemet.edu
Environmental Change Network (ECN) (United Kingdom)	The U.K.'s long-term environmental monitoring programme. It collects, stores, analyses and interprets long-term data from a range of terrestrial and freshwater sites across the U.K.	http://www.nmw.ac.uk/ecn/data_info.htm
Center for Tropical Forest Science (CTFS) of the Smithsonian Tropical Research Institute	Provides data from a number of the 50-ha Forest Dynamics Plots. Within each plot, trees are identified, marked, measured and plotted on a map.	http://www.ctfs.si.edu/
Canada's National Forest Inventory (Natural Resources Canada, Canadian Forest Service)	Intends to provide access to data collected from a plot-based inventory system across Canada.	http://www.pfc.cfs.nrcan.gc.ca/monitoring/inventory
National Vegetation Information System (NVIS)	Developing a national framework for compiling and communicatmg information about Australia's vegetation.	http://www.alJa.gov.au/docs/rural_science/nfi/nvis/
French National Forest Inventory (Inventaire Forestier National, IFN), Countryside and Forestry France. Department	Access to cartographic, mensuration and ecological information collected throughout	http://www.ifn.fr/
U.S. Man and the Biosphere (U.S. MAB) Program, in association with the Information Center for the Environment	Developing databases of vascular plant and vertebrate animal occurrences on the world's biosphere reserves and other protected areas.	http://ice.ucdavis.edu/mab/
National Vegetation Map of southern Africa Project	Vegetation survey data from sites across South Africa.	http://www.nbi.ac.za/research/vegmap.htm
Nordic vegetation survey	Aims to coordinate analysis, description and classification of Nordic vegetation. Participants have agreed upon a common concept of data sampling, storage, quality control and analysis.	http://hjem.get2net.dk/lawesson/The%20Nordic %20vegetation%20survey.htm
Natural Heritage Network (U.S.)	Comprises 85 biodiversity data centres throughout the Western Hemisphere (mostly in the U.S.A.). Personnel collect, organise and share data using a common, standards-based methodology. The network helps provide information for land-use decisions and is also consulted for research and education.	http://www.abi.org/

Agriculture Organisation (FAO) and the Montreal Process (both related to sustainable forest management) (see Bellingham *et al.,* 2000). Domestically, government agencies are charged with ensuring compliance with key environmental laws [such as the Conservation Act 1987, Resource Management Act 1991 and the Forests Act 1949 (and 1993 amendments)]. Accessible vegetation data of known quality are required to meet these obligations. In recent years, the resurgence of interest in vegetation monitoring has resulted in a proliferation of data collected and stored locally (e.g., by Department of Conservation conservancies, territorial local authorities and private consultants). However, these data are often not archived or made accessible in ways that allow issues spanning larger temporal or spatial scales to be considered. Adequate documentation and storage of data are especially important in long-term studies based on permanent plots (Brunt, 1994).

In New Zealand, the National Vegetation Survey (NVS) databank stores, manages and provides access to a large portion of the data on vegetation composition and structure collected in this country over the last 50 years. Other important vegetation databases in New Zealand include the South Island high country monitoring data currently held by Knight Frank (NZ) Ltd. (Webster, 1994) and data held by Timberlands West Coast Ltd. A myriad of smaller vegetation data sets are held by individuals at universities, in private consultancies, within national and local government agencies and Crown Research Institutes [see Meurk and Buxton (1991) and Bellingham (1996) for partial listings]. Some important vegetation data are not available electronically [e.g., North Island Ecological Transects: McKelvey and Cameron (1958); data from an extensive survey of Stewart Island: Wilson (1987)].

In this paper we describe the history of standardized collection of vegetation data in New Zealand and the resulting evolution of the NVS databank from paper to the current electronic version. We then characterise the types of data stored in the databank, including the geographic, ecosystem and temporal coverage. Finally, we outline future plans for the databank, including plans for meeting needs of both data users and providers, expanding metadata and quality control, and enhancing flexibility and utility.

Evolution of a New Zealand vegetation databank

History of collection of standardised vegetation data

Leonard Cockayne provided the first comprehensive descriptions of New Zealand plant communities (e.g., Cockayne, 1899; 1928). Later, he extended these descriptions to include changes in plant communities over time, based on observations from permanently marked sites (e.g., Cockayne and Calder, 1932). Formal national surveys of New Zealand's vegetation began in 1923 with the National Forest Inventory, a standardised inventory of the country's forests to assess their potential timber yield (Anon., 1926). The second standardised survey was the National Forest Survey (NFS) of 1946-55, which was primarily a timber inventory but ecological data were also collected (Thomson, 1946; Masters *et al.,* 1957). It mainly covered lowland and mid-altitude forests from which timber could be extracted, with limited coverage of upland forests. In 1956/57 this coverage was extended by the North Island Forest Ecological Survey (Ecosurvey) which provided comprehensive ecological information on forests not surveyed in the NFS (McKelvey, 1995). The NFS and Ecosurvey provided the foundation for a community classification of New Zealand forests (e.g., Nicholls, 1976; McKelvey, 1984).

The increasing focus on the role of natural forest and grassland ecosystems in protecting catchments and the vulnerability of these to the effects of browsing mammals ushered in an era of vegetation monitoring. Standardised methods were developed and later refilled for forests, grasslands and other non-woody ecosystems (Holloway and Wendelken, 1957; McKelvey and Cameron, 1958; Wraight, 1962; Scott, 1965; Atkinson, 1975; Wardle and Guest, 1977; Batcheler and Craib, 1985; Dickinson *et al.,* 1992; Allen, 1993; Wiser and Rose, 1997). Vegetation communities were described in many parts of New Zealand where standardised survey data were scant (e.g., Kelly, 1972). Based on methods in widespread use internationally (e.g., Mueller-Dombois and Ellenberg, 1974), standard methods using reconnaissance descriptions were tailored to New Zealand ecosystems and adopted for general surveys and for data collection in the Protected Natural Areas (PNA) Programme (e.g., Myers *et al.,* 1987; Allen, 1992).

In 1987 the Department of Conservation (DOC) was established. This was during a time of upheaval in the New Zealand civil service (Kelsey, 1997) and staff turnover was high. During the late 1980s and early 1990s, vegetation survey and monitoring was a relatively low priority for the Department (Bellingham, 1996), although some national initiatives continued, notably the PNA Programme. A result was the loss of many skilled staff who had undertaken vegetation surveys. Staff attrition resulted in a loss of institutional memory and a loss of appreciation of the value of major data sets (Bellingham, 1996). This led to some unfortunate losses of irreplaceable vegetation data during this period. Parallel events and loss of data also occurred in some research institutes and other government agencies.

In the early 1990s, standardised data collection continued in a piecemeal fashion by individuals in government agencies, universities, private consultancies and research institutions. Starting in 1997, new management procedures within DOC led to a revival of vegetation survey and monitoring, and the Department began to rebuild the requisite skill base. Standard methods are now being used increasingly within DOC to ensure comparability of results. Regional and local authorities, too, are placing more emphasis on vegetation survey and monitoring to meet requirements of the Resource Management Act 1991.

Development of a physical archive and electronic databank

From the late 1960s, access to mainframe computers enabled the organisation of data collected using standard methods into defined electronic data formats. This made it possible to analyse large amounts of data from throughout New Zealand (e.g., Wardle, 1970). The Forest Research Institute (FRI) and, from the early 1970s to 1986, the New Zealand Forest Service, adopted standard methods of data collection for reconnaissance surveys and permanent plots in forest and grassland. Concurrently, FRI developed standard formats for electronic data entry and storage, and computer packages for data checking and analysis (Allen *et al.,* 1983; Hall and Allen, 1985). From the early 1980s, data were collected and entered using these standard formats for many of the reconnaissance descriptions used in PNA surveys (e.g., Arand and Glenny, 1990). At that time, however, computer files and data sheets were held in offices and storerooms throughout New Zealand.

In the late 1980s, the creation of the NIVS (National Indigenous Vegetation Survey) database formalised the process of obtaining and archiving electronic data, copies of original field data sheets, maps, aerial and plot photographs, ancillary information and reports at FRI in Christchurch (Payton *et al.,* 1988; Forest Research Institute, 1989). The NIVS database also included data from plot types such as variable area forest plots (Batcheler and Craib, 1985) and those collected using the cruciform method (Holloway and Wendelken, 1957). Hard copies of data sheets and ancillary information were organised in a central archive and arranged by ecological region and district to allow ready retrieval. At that time, the electronic database and analysis packages could be accessed (read-only) by anyone linked to the Ministry of Forestry V AX computer system. Later, 14 reports produced for DOC listed all available data sets for each Conservancy (e.g., Hall *et al.,* 1991).

The NIVS database and staff associated with its development and maintenance transferred from FRI to Landcare Research when it was established in 1992. Agreement was reached that copies of NFS data (plot sheets and electronic data) and attendant maps and documentation would form part of NIVS. To adapt to changing technology, data analysis packages were rewritten to allow them to be run from personal computers (e.g., Hall, 1994a, b).

In 1997, the vegetation database was renamed the National Vegetation Survey (NVS) databank and incorporated data from the NIVS and NFS (NFS and Ecosurvey data: Forest Research Institute, 1989) databases and reconnaissance descriptions collected by the PNA Programme. The name reflects the intention to encompass data spanning a wide range of New Zealand's vegetation types including communities where either indigenous or exotic plants dominate. In 1998, the Foundation for Research, Science and Technology accorded the NVS databank the status of a Nationally Important Database, and since 1999 has funded its maintenance.

The NVS databank has two primary functions. The first is to serve as a national archive into which data can be deposited with confidence that future retrieval will be straightforward, and that, with provisos, these data may be made available to others. The second is to achieve as much consistency as possible in the manner in which data are stored and accessed to allow ready analysis of combined data sets that span space and time.

What data are stored in the NVS databank?

The NVS databank pertains largely to vascular plants. Data have been collected from both permanent plots and one-off surveys (e.g., reconnaissance surveys, PNA surveys and the NFS). The databank is not a single agglomerative database; rather, it is composed of individual data sets pertaining to individual surveys. These are mostly groups of plots (range I to c. 1000) within a defined survey area collected over a set time period. Most data from permanent plots in indigenous forests have been collected from 20-m x 20-m plots, within which individual stems are tagged (Allen, 1993). Data also include seedling and sapling counts. For grasslands, permanent plot data consist of frequency measurements of all vascular plant species, stature and density of dominant tussock species, and stereophotographs (Wiser and Rose, 1997). Data from reconnaissance descriptions, including those collected under the auspices of the PNA Programme, include assessments of abundance of each species in a given area. Some reconnaissance descriptions are associated with permanent plots. On all plot types site attributes such as altitude, slope and aspect have usually been recorded. Spatial location coordinate data (recorded to the nearest 100 m) are present for -95% and -65% of the permanent forest and grassland plots respectively, for all NFS and

Ecosurvey plots, and for 80% of the reconnaissance descriptions from reconnaissance and PNA surveys.

Currently the databank contains data from -10 000 permanent plots in forests, of which -6500 are 20-m x 20-m plots, and of these, ~2100 (34%) have been remeasured at least once. There are data from -3800 permanent grassland transects [most follow methods of Wraight (1962)], and of these ~390 (10%) have been remeasured. Data collected under the auspices of NFS and Ecosurvey comprise 14000 plots. The databank also contains data from ~52 000 reconnaissance descriptions (PNA Programme included); of these ~9000 (17%) represent repeat measurements, usually associated with permanent plots. Hard copies of data sheets completed during field surveys and ancillary material such as maps and aerial photographs showing plot locations are stored in the herbarium (CHR) at Landcare Research, Lincoln. Archiving hard copy as well as electronic data is essential; there are many unfortunate stories of the loss of data because of the dependence on electronic media alone (e.g., Michener *et al.,* 1997). Also, hard copies of data sheets normally contain information such as location maps, that is not computerised.

For most data entry, checking, storage, analysis and export, the NVS databank currently uses a computer system designed in the 1980s and refined continually since. Most data are entered via REFLEX® (Borland International, 1989), a database management system. For storage and analysis, data are converted to standardised condensed formats of ASCII text. Data of different types (e.g., tree diameters, sapling and seedling counts, species composition) from the same survey are stored in files having the same name, but different extensions. Such condensed formats were required in the 1980s when electronic storage space was limited. Programs for data checking and analyses are written in FORTRAN (e.g., Hall, 1994a, b) and remain available for a nominal cost from Landcare Research (see URL: www.landcare.cri.nz/ science/nvs). These programs also allow export of summarised data to other formats [e.g., suitable for analysis by statistical packages, and vegetation analysis programs such as CANOCO (ter Braak and Smilauer, 1998) and TWINSPAN (Hill, 1979)]. Programs have been written in SASTM (SAS Institute, 1999) to readily manipulate and combine these ASCII files, and to export and import data to and from a widerrange of file types (e.g., MS-ACCESS, MS-EXCEL, tab-delimited). A Microsoft Windows-based platform for data entry and management in a relational database system is currently being developed. Electronic files in the NVS databank are stored on a Digital Prioris HX 200 MHz Pentium Pro server using Novell Netware version 4.11 at Landcare Research, Lincoln. Back-ups are made every night with a full back-up performed twice a week. Monthly back-ups are stored in a fire- and earthquake-proof room in a separate building.

Geographic and ecosystem coverage

Forests and grasslands dominated by indigenous species are well represented by permanent-plot data in the NVS databank. Indigenous ecosystems that are either poorly or not represented by permanent plots are subalpine and successional shrublands, freshwater and estuarine wetlands and turfs. Across the two main islands, geographic coverage of permanent plots in indigenous forests is widespread but patchy, with some areas well represented (e.g., Fiordland, southern North Island forests) and others poorly represented (notably Northland, inland Taranaki and north Westland) (Fig. 1). In general, upland areas are better represented by permanent plots than are lowland areas. The NVS databank also includes data from plots in forests of the Chatham, Kermadec and Stewart islands. Grasslands in montane and alpine areas are well represented and coverage is strongest in wetter regions and on land managed by DOC. Grasslands in drier regions of New Zealand, especially the induced grasslands in the eastern South Island, are poorly represented, and no permanent plot data are held for indigenous lowland and coastal grasslands.

Data from one-off surveys (e.g., reconnaissance, PNA, NFS) are more comprehensive in geographic coverage and the range of vegetation types sampled. One-off surveys include data from coastal turfs to high-altitude grasslands, and both woody and non-woody vegetation. Reconnaissance descriptions are concentrated on land administered by DOC, and in terms of absolute numbers there is geographic bias towards some areas (e.g., South Westland).

Temporal coverage

Most permanent plots in grasslands were established in the 1960s and 1970s and in forests in the 1970s and 1980s (Fig. 2). With time, these data comprised an increasing proportion of plot remeasurement data versus establishment data. Most reconnaissance survey data were collected in the 1970s and 1980s, and PNA data in the 1980s. Fewer plots are represented in the databank by data from the I 990s, reflecting the lower level of data collection during that decade (Fig. 2).

Permanent plots, especially those with a history of measurement, can provide benchmarks against which to assess long-term change in ecosystems (Bakker *et al.,* 1996). TheNVS databank contains data from some especially notable permanent-plot surveys from forests. These have an average measurement span of 21 years (Table 2). Most used similar methods, with plots systematically spaced along randomly located transects. As such, they record the average' dynamics and stand structure of the catchment. The databank also holds data from other notable permanent forest plot networks, including data sets from the Orongorongo Valley (Campbell, 1990) and the Hunua Ranges. These too have outstanding histories of measurement, but were based on different sampling regimes.

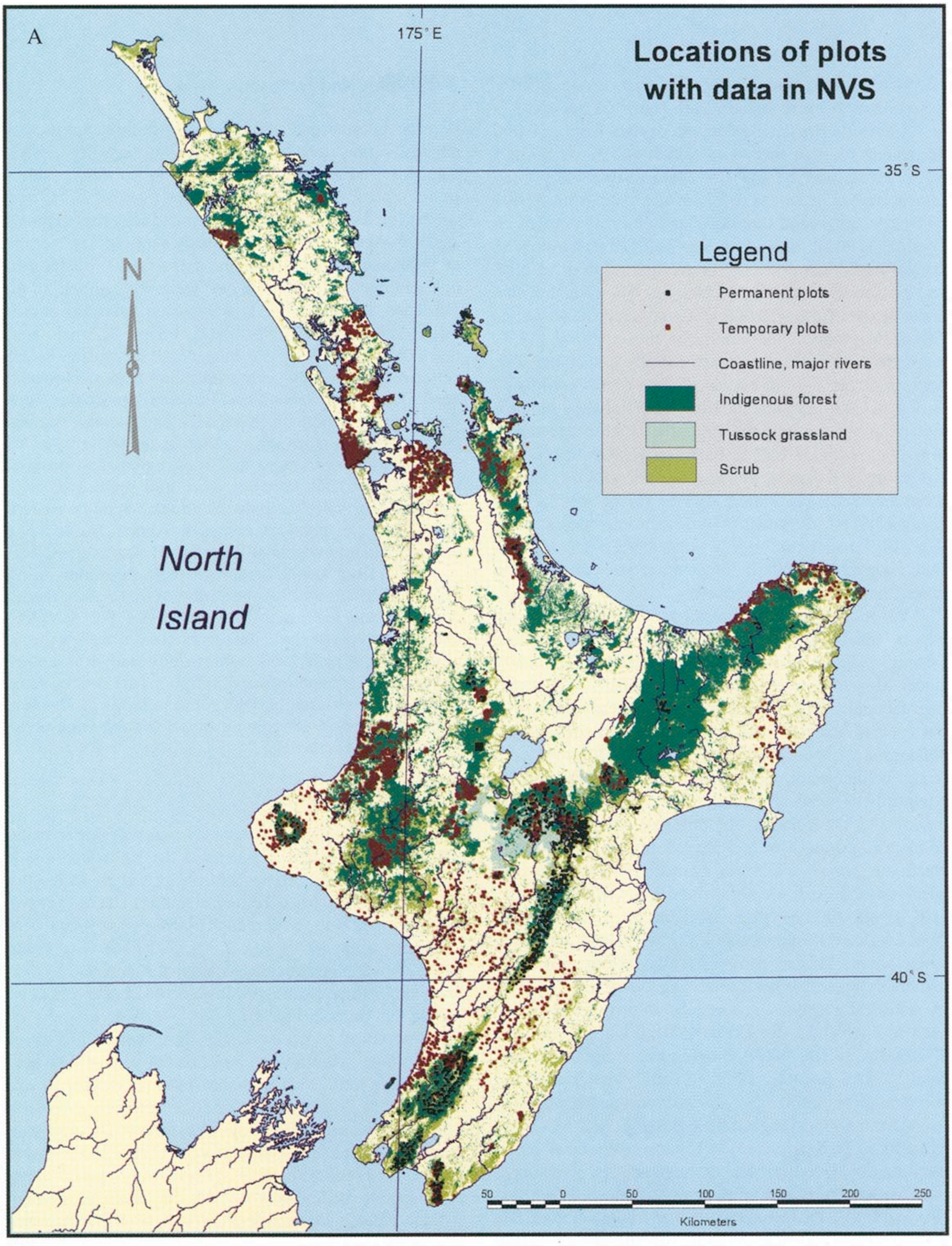
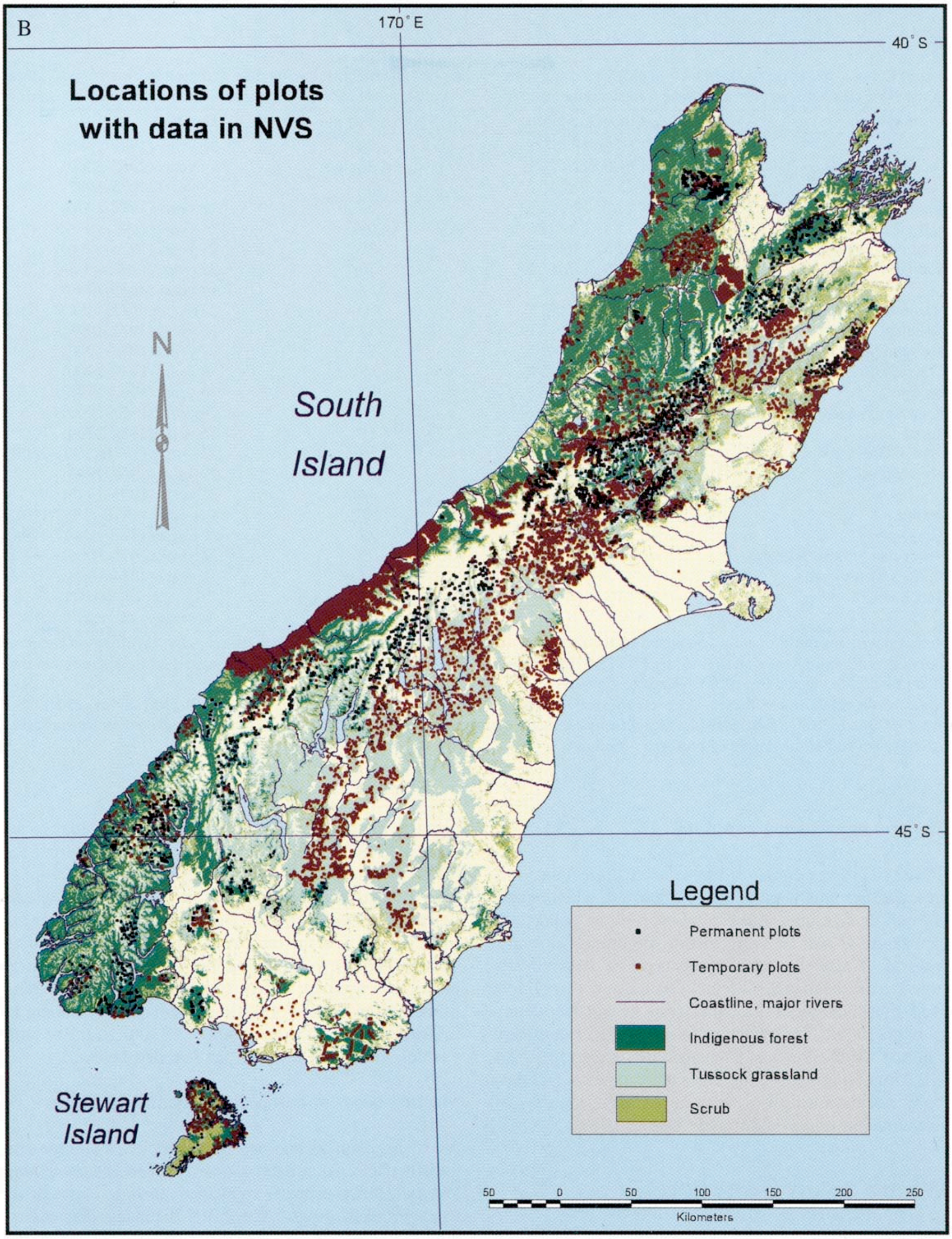


Figure 1. Locations of plots for which data are stored in the NVS databank, and location data is present, as of December 2000. These are overlain on areas mapped as either forest, scrub or tussock grassland by the Landcover Database (derived from a classification of SPOT satellite imagery acquired in the summer of 1996/97); (a) North Island, (b) South and Stewart islands. Variable area plots, NFS and Ecosurvey plots are excluded.


B
170° E
40° S
45° S
Locations of plots
with data in NVS
N
South
Island
Stewart
Island
Legend
Permanent plots
Temporary plots
Coastline, major rivers
Indigenous forest
Tussock grassland
Scrub
50 0 50 100 150 200 250
Kilometers

-

Table 2. Local networks of 20-m x 20-m plots sampling indigenous forest catchments. having data spanning at least 14 years stored in the NVS databank and the most recent measurement since 1992. Localities of plot networks are ordered from north to south.

Locality	Latitude, Longitude	Number of plots	Duration
Pirongia	37°59'S, 175°02'E	20	20 years (1979-1999)
Okataina	38°08'S, 176°27'E	36	17 years (1983-2000)
North Pureora	38°23'S, 175°35'E	28	18 years (1975-1993)
Kaimanawa	39°12'S, 175°59'E	30	18 years (1982-2000)
Kaweka	39°15'S, 176°25'E	140	14+ years (1981-(1995-1999))
Ruahine (Pohangina Valley)	40º03'S, 176°02'E	10	21 years (1975-1996)
Southern Tararua[1]	41°01 'S, 175°12'E	10	22+ years (1975-(1996-1999))
		5	14 years (1985-1999)
Hurunui	42°43'S, 172°0 I'E	102	25+ years (1975-(2000-2001))
Kokatahi	42°57'S, 171°12'E	22	23 years (1972-1995)
Whitcombe[2]	43°05'S, 171°01'E	23	27 years (1972-1999)
Craigieburn[2]	43°10'S, 171°35'E	250	30 years (1970-2000)
Okarito[3]	43°13'S, 170º16'E	32	14 years (1983-1997)
Arawata (Waipara)	44°15'S, 168°41'E	4	29 years (1971-2000)
Kaipo	44°26'S, 167°53'E	10	15 years (1984-1999)
Caples-Greenstone	44°55'S, 168°14'E	38	22+ years (1976-(1997-2000))
Murchisons	45°18'S, 167°38'E	5	29 years (1969-1998)
		34	23 years (1975-1998)
Waitutu	46°12'S, 167°04'E	107	19+years (1978-(1997/1998))
Longwood[4]	46°I3'S, 167°50'E	42	20+ years (1977-(1997/1998))
North east Stewart Island	46°47'S, 167°59'E	23	24 years (1976-2000)
North Stewart Island	46°47'S, 168°oo'E	47	18 years (1981-1999)
Bench Island	46°54'S, 168°15'E	5	20 years (1979-1999)
		L = 984	Average = 21 years

Remeasurements most recently conducted by the Department of Conservation (or its contractors) except by: [1]Wellington Regional Council; [2]Landcare Research (funded by FRST); [3]Timberlands West Coast (plots located on a grid within the catchment rather than along randomly located transects); [4]Waikato University.
Note that the Ruahine and Tararua data sets are small subsets, remeasured in the 1990s, of very large (> 100 plot) surveys established in the 1970s and 1980s. Likewise, the Murchisons data set is a subset of a much larger North Fiordland original data set from a survey covering several of the ranges of northern Fiordland.

Progress in archiving new data

Efforts to procure copies of important historical data sets are ongoing. This is particularly important when people retire or change jobs. At such times, lifetime collections of data are at risk of being lost or forgotten. Currently, archival of new data focuses on types of data already stored in the databank. NVS lacks data from other widely used methods, notably data from height-frequency transects (Scott, 1965; Dickinson *et al.,* 1992) and forest transects (McKelvey and Cameron, 1958). Future efforts will focus on incorporating these types of data, and data from permanent plots with a notable history of measurement (e.g., Calder and Wardle, 1969).

Where to go from here

The overall goal of the NVS databank is to safeguard millions of dollars worth of past investment in data and thus facilitate knowledge gains from these data. The unique time-series record from permanent plots and one-off vegetation records from the past are irreplaceable and become more valuable with time. For this goal to be achieved, and to become a truly national resource, the NVS databank must be seen as the logical place for long-term storage of vegetation data and the first port-of-call when such data are sought (e.g., for design of monitoring programmes and for information on vegetation status). Current barriers to achieving this goal, some real, others perceived, include issues surrounding rights of data users and providers, provision of adequate metadata to interpret raw data, assurance of data quality and technological issues. These are detailed below.

Meeting needs of both data users and providers

As advances in technology have simplified storage and transfer of electronic data, issues of data access, ownership and intellectual property rights have emerged worldwide (e.g., Frankel, 1999). The NVS databank data-access policy has attempted to strike a balance between making data freely available and protecting the rights of data providers. The databank does not 'own' data; rather it serves as an intermediary between data providers and data users (cf. Nash, 1993). Data providers can set conditions of use. Data users agree to a set of obligations that govern use of data (e.g., concerning citation, provision of data to third parties etc.; see Appendix 1). Worldwide, such agreements are becoming standard, particularly for large databanks.

Much of the data stored in the NVS databank (i.e., data designated as nationally important) lies in the public domain. This includes most data collected before 1987 (when DOC was formed). The policy of Land care Research regarding access to these data is aligned with the policy for national databases and collections owned by Crown Research Institutes. That policy was developed by the Crown Company Monitoring Advisory Unit in 1996/97 and is designed to provide access to these data for public good or personal use, except where the access is clearly not to the benefit of New Zealand (Whitehouse, 1998). Requests for public domain data can be met on the basis of cost of supply (e.g., costs of downloading electronic archives, determining any restrictions on distribution of data, photocopying original data sheets). Costs may range from nominal for simple queries to significant where considerable manipulation of data is required. For private good or commercial use other restrictions and costings may apply; these are handled on a case-by-case basis.

The NVS databank also stores data to which access is restricted by the data provider (Appendix 1). There are two levels of restriction. The first, and most common, is that access is contingent on permission from the data provider. Data providers may be individual researchers or institutions (e.g., DOC). This restriction protects the proprietary rights of data providers and is in accord with the recommendations of Nash (1993) that the generating researcher(s) or institution(s) should control access to their data. In most cases this is formalised via a memorandum of understanding with the data provider. Access restrictions have been put in place because without them many providers will not agree to store their data in the NVS databank. The second level of restriction is reserved for confidential or commercially sensitive data, where the NVS databank functions as a data archive only. For individual data sets, data access levels are periodically reviewed. With time, and agreement of their owners, it is expected that many currently protected data sets will move from restricted access into the public domain.

Preparing a data set for deposit requires some effort to ensure it is properly documented, hard copies or ancillary information are available and the data are properly organised. Clearly, there are advantages to being a data user; less clear are the advantages to data providers. This is a problem faced by databank projects worldwide, and the solution is to have tangible rewards for data providers (Porter and Callahan, 1994). These could include provision of resources by funding agencies or databank managers to support technical services, such as data entry and quality assurance, allowing data providers preferential access to a databank and ensuring that data providers receive adequate recognition for their efforts (Porter and Callahan, 1994). Recognition may include acknowledgement in publications, collaboration and co-authorship of any publications based on their data (or the right to publish a disclaimer), and acknowledgement on the databank website. In response to the recognised lack of incentive for data providers, the Ecological Society of America has adopted a policy to encourage publication of 'data papers'. Such papers emphasise the "collection, organisation, synthesis and thorough documentation of data sets of ecological value" (Peet, 1998). The data will be stored in Ecological Archives (Table I). Providing better incentives to data providers is an area that needs to be pursued to promote archival of New Zealand vegetation data.

Metadata

Metadata are the descriptive information about the data. Comprehensive metadata should describe what data are stored, why and how they were collected, their quality, their structure and storage medium and how they can be accessed (Michener *et al.,* 1997; Michener, 1998). Metadata are essential for two primary reasons (Conley and Brunt, 1991; Stafford, 1993; Hale, 2000). First, metadata provide the information required for long-term use of a data set (Colwell, 1995). The importance of metadata to the NVS databank has become increasingly apparent with the uneven financial support for databases, loss of personnel and loss of institutional memory as a result of the restructuring of New Zealand science over the last 20 years. Even without such events, good documentation is required because of the difficulty of remembering details about a research project that was completed years ago (Fig. 3). Secondly, metadata allows users to ensure their use of the data is not beyond the bounds of the questions that the data can answer (Chrisman, 1994). This is especially important when a user is attempting to scale up point data to regional or national spatial scales.

The types of metadata required for ecological databases have been reviewed in numerous articles (e.g., Colwell, 1995; Michener *et al,* 1997; Hale, 2000). International standards exist for geospatial data [e.g., the U.S. Spatial Data Transfer Standard; National Institute of Standards and Technology (1992)] and taxonomic names (e.g., Bisby, 1995). The recently produced Biological Data Profile (FGDC Biological Data Working Group and USGS Biological Resources Division, 1999) incorporates these standards and provides standards for other types of metadata associated with biological data, such as data collection methods and electronic data field content. Michener *et al.* (1997) suggested the major categories of 'other' information, as a minimum, should include data set descriptors (e.g., originator of the study, research objectives, location), research origin descriptors (e.g., site description, sampling design, personnel), data set status and accessibility (e.g., data quality assessment, contact person, copyright restrictions), data structural descriptors (e.g., format and storage mode, descriptions of variables), and supplemental descriptors (location of data sheets and related materials such as maps, history of data set usage).

Currently the NVS databank includes a database that provides metadata for all individual data sets held (summarised in Table 3). These metadata were initially published as a series of reports (e.g., Hall *et al.,* 1991). Since then, the electronic database has been updated continually. The storage of information in fields allows easy searching of this database and ensures adequate documentation of each data set. In 1999 a standard metadata form for data providers *to* complete was produced to ensure that adequate documentation accompanies all data deposited in the NVS databank. Additional information about individual data sets resides in associated text files. These files also include records of corrections and changes made *to* electronic data files. Metadata for the actual variables stored in data sets [i.e., data structural descriptors *sensu* Michener *et al.* (1997)] are provided in manuals that describe the different standard data formats used (e.g., Hall, 1994a, b). Metadata about individual plots (e.g., grid reference, date sampled, altitude) are included as part of the data itself or in an associated text file if *non*-standard methods were used.

Several key types of metadata have *not* traditionally been stored in the NVS databank, but should be in future developments. These include methods used for determining values of site variables collected (e.g., whether grid references were determined from a Global Positioning System or read off a map), a record of the personnel who collected the data, full references for publications and reports based on the data, and dates when electronic files have been updated. Improving the quality and breadth of metadata, particularly for older data sets, is a priority for the NVS databank.

Quality control

The number of new errors entering a database can be reduced by developing formal quality-control procedures for adding, updating and editing data. It is essential *to* remove errors before analysis *to* prevent spurious results and misleading conclusions. Currently, some electronic data in the NVS databank have not received the level of checking desirable and this provides an ongoing challenge for database managers and users.

When data are entered, efforts are made *to* ensure that the data are as error-free as possible. Authority tables are used *to* ensure that data, such as 6-letter codes for species, are valid (e.g., Hall, 1994a). Validation protocols are used *to* ensure that data values fall within reasonable limits (e.g., that tree diameters are *not* excessively large, that the aspect for a plot does not exceed 360°). For the most part, data stored is 'raw', i.e., exactly as recorded on the data sheets. However, past efforts to standardise data

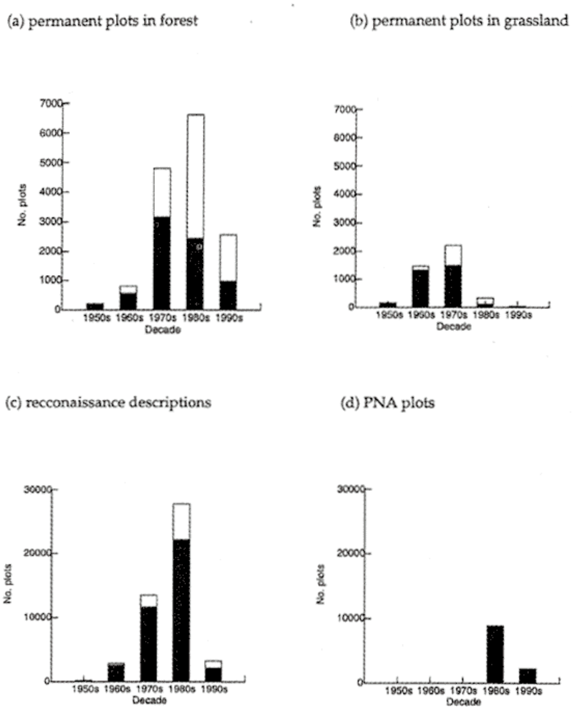


Figure 2. The number of plot measurements in each decade for which data are stored in the NVS databank. NFS and Ecosurvey plots are not shown. Summaries for different types *of* data are given in a-d. The filled portion *of* the bar indicates plots that were first measured in that decade; the open portion indicates remeasurement of existing plots. Note different scales on the vertical axes.

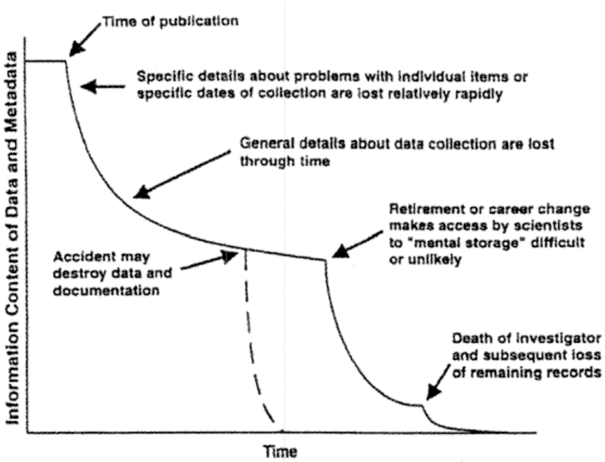


Figure 3. How the information content of data and associated metadata decays over time. Reprinted with permission from Michener *et al.* (1997).

resulted in units being converted from imperial to metric before the data were stored. In some cases this gives the impression of higher levels of precision (e.g., "500 feet" has become "152 metres"). Diagnostic tests have been performed on different subsets of the electronic data to screen for errors. For example, tree diameter data were recently examined for 7564 permanent plots in indigenous forests to determine whether tagged trees had plausible rates of diameter growth.

Other types of errors cannot be recognised simply by examining existing electronic data. For this reason we conducted a test of data accuracy by remeasuring a set of 25 permanent plots in relatively species-rich montane rainforests of the Whitcombe Valley, central Westland (James *et at.,* 1973). As these plots had not been visited in 19 years and occur in an area where high rainfall causes frequent flooding and landslides, we anticipated that relocation of plots and subplots would be difficult. In fact, all plots were successfully relocated. On average, the locations as recorded on metric NZMS 260 series maps (1 :50 000) were 130 horizontal metres from the location originally recorded on imperial NZMS 1 series maps (1:63 000). Altitude, aspect and slope data were similar to those recorded in the past, although measurements of aspect when slopes were < 5° proved unrepeatable. In most plots, permanent markers for seedling subplots were relocated readily using metal detectors; on average 22 of 24 markers per plot were relocated. For permanently tagged stems having diameters ≥2.5 cm, species identifications were highly accurate; only 1.6% of *c.* 2800 stems had been incorrectly identified during the original survey. Taxonomic problems were more common in seedling subplots, especially for sedges, grasses and some ferns. Some problems resulted from changes in taxonomic concepts since the last remeasurement and taxa such as *Hymenophyllum* and *Uncinia* only being identified to genus level in the past. Some previous identifications were suspect but could not be verified because we could not find that taxon on the plot. Other errors arose because original tree tags had been replaced with tags having different numbers; current data formats do not distinguish retagged trees from newly tagged ones.

Efforts are underway to improve quality control procedures for the NVS databank. Automated procedures to allow longitudinal checks (Le., comparisons with data collected in the past) on permanent plot data at the time of data entry are being developed. Discrepancies are much easier to resolve when the people who collected the data can still remember what they did! For permanent plots, discrepancies that can only be resolved on-site at the time of next remeasurement are now recorded in text files which are retained with the original data sheets (copies of which are given to data providers); these will be given to the next remeasurement team. To assure data integrity, write-access is currently restricted to a highly trained database administrator who has more than 10 years experience working with the NVS databank and its precursors. Errors found by data users and those who curate the databank are corrected on the electronic files by the database administrator. Currently, data users tend to find errors and correct them on a copy of the file that they are using; only rarely are these corrections fed back into the master copy of the data held by the NVS databank. Conversely, these users do not reap the benefits of ongoing updates of the electronic data.

Enhancing flexibility and utility

Although the system meets most needs of the current NVS databank users, it is now desirable to take advantage of evolving technology to improve its capability and flexibility. The technical solutions devised in the 1980s anticipated many future developments. With new developments, the underlying strategy is to advance in small steps that are driven by the needs of data providers and users while retaining the flexibility to allow future developments that cannot be anticipated. To determine optimal solutions to enhance the NVS databank, a thorough review of how vegetation and permanent plot databases have been designed and implemented elsewhere in the world is underway. Each development of the NVS databank will be preceded by a pilot project to ensure that the functionality of the database is improved.

How is flexibility of a database best enhanced by new developments? In contrast to business data, scientific data are often less structured and less formally organised, and the needs of users are less predictable (Hale, 1999). Software used for databanks should provide maximum flexibility to import and export data and to allow access via different platforms [e.g., IBM, MacIntosh, Sun workstations; Porter (1998); Burley (1998)]. Many ecological data archives require data to be stored as plain ASCII text (sometimes called 'flat files') with clearly defined formats. The advantages of this were summarised by Colwell (1995) as: (i) ASCII is platform-independent; (ii) ASCII text can be read and written by all proprietary software (e.g., relational database management systems, spreadsheets, statistical packages), whereas directly reading and writing between software systems is often problematic; and (Hi) ASCII offers the maximum flexibility for structuring data. For some ecological databanks, data entry and quality control work is done using a relational database management system, but data are stored in ASCII text (e.g., Stafford, 1993).

Currently the NVS databank is going through a major upgrade to increase its flexibility, improve accessibility of data and integrate it with other Landcare Research databases. The first goal is to facilitate storage of data that do not fit into the data structures currently supported. This will require tailoring structures to accommodate other widely used survey methods (e.g.,

-

Table 3. Information stored in the metadata file in the NVS databank that describes individual data sets (modified from Hall *et al,* 1991). Descriptor classes follow Michener *et at.* (1997).

Descriptor	Definition/explanation
Class I. Data set descriptors	
Survey name	Survey name and year data were collected.
Class II. Research origin descriptors- overall and specific subproject descriptions	
Organiser, organisation	The survey organizer or principal investigator and their organisation, government department, institution etc.
Department	The department or conservancy that initiated the survey.
Aims	The rationale for the survey.
Vegetation	Type of vegetation sampled, e.g., forest, scrub, grassland.
Location	Includes general location, specific location (catchment, hill, forest), Ecological Region, Ecological District, Ecological Code, DOC Conservancy, topographical map code of the imperial NZMS I series or the metric NZMS 260 series.
Survey method	Describes which of a set of standard sampling methods was used.
Modified	Describes non-standard sampling methods or how standard methods were modified.
Remeasure	Whether the survey remeasures previously surveyed areas or plots. Previous measurement years listed where relevant.
No. lines, plot numbers	Number of lines and plots measured.
Exclosure	Indicates whether the survey includes plots from animal exclosures.
Species information	Indicates what information was recorded about species (e.g., occurrence in tiers, cover classes, stem density).
Site information	Indicates what site information was recorded (e.g., altitude, aspect, slope).
Plot coordinates	Indicates whether grid coordinates were recorded on data sheets.
Class III. Data set status and accessibility	
Contact person	Whoever knows most about the data at present.
Access, access address and phone	Proprietary restrictions on use of data; contact details for permission for access.
Class IV. Data structural descriptors	
File name and directory	Name of computer file and directory where file resides. Size
Size (kb)	of the computer files in kilobytes.
Data entry	Data entry operator or data source if imported electronically.
Class V. Supplemental descriptors	
Data location	Agency where the original plot sheets or copies reside.
Box number	Box number where data sheets reside in the NVS archive at CHR.
Missing plot sheets	Number of plot sheets not held at Landcare Research.
Photocopy	Describes whether data sheets in the NVS archive are photocopies of originals, and whether the quality is adequate.
File errors, species errors, missing data, warnings	Summaries of results from quality control checks.
Corrections	Corrections needed to reduce the number of errors found during quality control checks.
Aerial photos, slides, soil records, bird records, animal census, browse records, maps, location diagrams on plot sheets	Indicates presence/absence of this type of information.
Reports/refs	The author and year of any published or unpublished material generated from the data.
File distribution	Describes who has been provided copies of computer files or data sheets and when this was done.
Notes	Any miscellaneous information about the survey.

height-frequency data). For plots of standard types already supported, there is a need to accommodate ancillary information such as additional site information (e.g., soil chemistry data, GPS coordinates, topographic variables), and more attributes of individual plants measured on plots [e.g., indices of browsing by" introduced animals, (Payton *et al.,* 1999); individual tree heights, spatial location of trees within a plot, presence of flowers, fruits and parasitic plants]. To allow storage of vegetation data that do not fit into a standard format, an approach such as that of Conley and Brunt (1991) is being adopted. They designed a generalised data structure (stored as ASCII text) that contains both the data and full documentation in one file that stands alone. The data can be extracted from such files with any text editor, then read into the software system of choice (e.g., spreadsheet, statistical package, graphics package).

The second goal of the current databank upgrade is to enable data users to readily query the NVS databank using software systems of their choice, including Geographic Information Systems (GIS). Over the next year, the data are being restructured to enable just that. Users will be able to access data with conventional software such as database, spreadsheet, statistical, or graphics packages, with computer packages designed specifically for the NVS databank (e.g., Hall, 1994a, b), or specialised vegetation analysis software such as PC-ORD (McCune and Mefford, 1999), CANOCO or TWINSPAN. Conversion of spatial location data to a range of forms is underway to increase the utility of the data and facilitate interrogation with GIS.

The third goal is development of an internet site to facilitate access to information about the NVS databank and to data stored there. Currently the site includes general information about the databank, copies of data formats used, copies of data collection manuals, data request and deposit forms, and maps of plot locations. In the near future the metadata will be available for querying. Eventually, we hope to have the plot data available as well.

The NVS databank is an invaluable source of point data on vegetation composition and structure, and there are wide-ranging knowledge gains to be made by integrating these data with other New Zealand databases and sources of information. To date, this has been done to a limited extent. GIS has been used to overlay plot and animal distribution data to determine areas most susceptible to damage by exotic animals (e.g., Rose *et al.,* 1994) and to model distributions of species in relation to climate (e.g., Leathwick *et al.,* 1998). NVS data have been linked to species-attribute information in the Taxonomic Names Database held by Landcare Research, to summarise point vegetation data in terms of plant family membership and exotic or native status. Georeferenced plot data can also be used to verify other data sources. Data from forest canopy gaps have been used to verify canopy gap locations generated from digital canopy-elevation models derived from aerial photographs (Betts *et al.,* 2000). Vegetation data from which carbon storage has been calculated have been linked to satellite images to allow ground-truthing for carbon monitoring (Pairman *et al.,* 1999).

Concluding comment

Data not only provide the foundation for science, they will increasingly provide the basis for many of our management decisions. As data accumulate, there will be a critical need to standardise, integrate and disseminate biodiversity information - we are at the beginning of a revolution (Burley, 1998). Vehicles such as the NVS databank can be used to ensure that the substantial investment of time spent collecting, entering, correcting and managing biodiversity information is safeguarded for the future. For more information about the NVS databank see the internet site www.Landcare.cri.nz/science/nvs. Queries about NVS can be sent to nvs@landcare.cri.nz.

Acknowledgements

We thank Terry Savage for producing the maps, and Rob Allen, Ian Payton, Cathy Allan, Ian Whitehouse, John Leathwick, David Coomes, Michelle Breach and Christine Bezar who all provided valuable comments on an earlier version of the manuscript. The NVS databank owes its existence to many individuals whose efforts over the years have allowed it to become what it is today. These efforts include, but are not limited to, the work done to standardise methods, collate and collect hard copies of plot sheets and ancillary data, computerise data, develop data checking and analysis routines, and document datasets. Some of the individuals who played a major role are: John Holloway, John Nicholls, John Wardle, Graeme Hall, Ian Payton, Martin Fastier, Shane Andreasend, Rob Allen, Graeme Evans, Mac McLennan, John Leathwick, Shirley Chee, Ngaire Breuton and Michelle Breach. Special thanks go to all those people who wore out their boots collecting data and those who have saved and provided data to the databank. Financial support for the preparation of this paper was provided by the New Zealand Foundation for Research, Science and Technology as part of contract CO9806.

References

Allen, R.B. 1992. RECCE - *an inventory method for describing New Zealand vegetation.* Forest Research Institute Bulletin 176. Forest Research Institute, Christchurch, N.Z.

Allen, R.B. 1993. *A permanent plot method for monitoring changes in indigenousforests.* Manaaki Whenua - Landcare Research New Zealand Ltd, Christchurch, N.Z.

Allen, R.B.; Rose, A.B.; Evans, G.R. 1983. *Grassland survey manual: A permanent plot method.* Forest Research Institute Bulletin 43. Forest Research Institute, Christchurch, N.Z.

Anon., 1926. *The New Zealand official year-book.* Government Printer, Wellington, N.Z.

Arand, J.; Glenny, D. 1990. *Mathias and Mt Hutt Ecological Districts.* Protected Natural Areas Programme Survey Report No. 12. Department of Conservation, Wellington, N.Z.

Atkinson, I.A.E. 1975. A method for permanent transects in vegetation. *Tuatara 21*: 81-91.

Bakker, J.; Olff, H.; Willems, J.H.; Zobel, M. 1996. Why do we need permanent plots in the study of longterm vegetation dynamics? *Journal of Vegetation Science* 7: 147-156.

Batcheler, C.L; Craib, D.G. 1985. A variable area plot method for assessment of forest condition and trend. *New Zealand Journal of Ecology* 8: 83-96.

Bellingham, P.J. 1996. *Surveys and monitoring of vegetation.* Landcare Research Contract Report LC9596/026 for the Department of Conservation. Landcare Research, Lincoln, N.Z.

Bellingham, P.J.; Stewart, G.H.; Allen, R.B. 1999. Tree species richness and turnover throughout New Zealand forests. *Journal of Vegetation Science 10:* 825-832.

Bellingham, P.J.; Wiser, S.K.; Coomes, D.A.; Dunningham, A. 2000. *A review of permanent plots for long-term monitoring of New Zealand's indigenous forests.* Science for Conservation 151. Department of Conservation, Wellington, N.Z.

Betts, H.; Brown, L; Stewart, G.H. 2000. Mapping Canopy gaps in beech forest from a digital Elevation model. *In:III Southern Connection Congress, Programme and Abstracts,* p. 21. Lincoln University, Canterbury, NZ.

Bisby, F. 1995. *Plant names* in *botanical databases.* Plant Taxonomic Database Standards No.3. Hunt Institute for Botanical Documentation, Pittsburgh, U.S.A.

Borland International Inc. 1989. *REFLEX® 2.0.* Borland International Inc. Scotts Valley, California, U.S.A.

Brunt, J. W. 1994. Research data management in ecology: A practical approach for long-term projects. *In:* French, J.C.; Hinterberger, H. (Editors), *Seventh international working conference on scientific and statistical database management,* pp. 272-275. IEEE Computer Society Press, Washington, D.C., U.S.A.

Burley, J. 1998. *Joining the revolution: A strategyfor the standardization, integration and dissemination of biodiversity information - as a prospective model for the management of other kinds of environmental information.* Manaaki Whenua - Landcare Research, Lincoln, N.Z.

Calder, J.W.; Wardle, P. 1969. Succession in subalpine vegetation at Arthur's Pass, New Zealand. *Proceedings of the New Zealand Ecological Society* 16: 36-47.

Campbell, D.J. 1990. Changes in structure and composition of a New Zealand lowland forest inhabited by brushtail possums. *Pacific Science* 44: 277-296.

Chrisman, N.R. 1994. Metadata required to determine the fitness of spatial data for use in environmental analysis. *In:* Michener, W.K.; Brunt, J.W.; Stafford, S.G. (Editors), *Environmental information management and analysis: Ecosystems to global scales,* pp. 177-190. Taylor & Francis, London, U.K.

Cockayne, L. 1899. A sketch of the plant geography of the Waimakariri River basin, considered chiefly from an oecological point of view. *Transactions and Proceedings of the New Zealand Institute* 32: 95-136.

Cockayne, L. 1928. *The vegetation of New Zealand,* Second edition. Engelmann, Leipzig, Germany.

Cockayne, L; Calder, J.W. 1932. The present vegetation of Arthur's Pass (New Zealand) as compared with that of thirty-four years ago. *Journal of Ecology 20:* 270-283.

Colwell, R.K. 1995. Ecological Society of America special committee on ESA communications in the electronic age. *Bulletin of the Ecological Society of America* 76: 120-131.

Conley, W.; Brunt, J.W. 1991. An institute for theoretical ecology? Part V: Practical data management for cross-site analysis and synthesis of ecological information. *Coenoses* 6: 173-180.

Dickinson, KJ.M.; Mark, A.F.; Lee, W.G. 1992. Long-term monitoring of non-forest communities for biological conservation. *New Zealand Journal of Botany 30:* 163-179.

FGDCBiologicalData Working Group and USGS Biological Resources Division. 1999. *Content standardfordigital geospatial metadata - biological data profile, FGDCSTD-OO* 1.1 - 1 999. Federal Geographic Data Committee. Washington, D.C., U.S.A.

Forest Research Institute. 1989. *Databases for New Zealand's indigenous vegetation.* What's New in Forest Research 175. Forest Research Institute, Rotorua, N .Z.

Frankel, M.S. 1999. Public access to data. *Science* 283:1114.

Hale, *S.S.* 1999. How to manage data badly (part 1). *Bulletin of the Ecological Society of America 80:* 265-268.

Hale, *S.S.* 2000. How to manage data badly (part 2). *Bulletin of the Ecological Society of America* 81:101-103.

Hall, G.MJ. 1994a. *PC-DIAM: Stem diameter analysis.* Landcare Research, Auckland, N.Z.

Hall, G .MJ. 1994b. *PC- USTOREY: Seedling and saplingdata analysis.* Landcare Research, Auckland, N.Z.

Hall, G.; Allen, R. 1985. *Reconnaissance vegetation survey programs.* Forest Research Institute Bulletin 88. Forest Research Institute, Christchurch, N.Z.

Hall, G.M.J.; Payton, I; Burrows, L; Fastier, M.; Andreasend, S. 1991. *National Indigenous Vegetation Survey database: Database directories.* Forest Research Institute contract report FWE 91/ 16, prepared for the Department. of Conservation. Forest Research Institute, Christchurch, N.Z.

Hall, G.M.J.; Wiser, SoK.; Allen, RoB.; Beets, PoN.; Goulding, CJ. 2001. Strategies to estimate national carbon biomass from forest inventory data: The 1990 New Zealand baseline. *Global Change Biology* 7: 389-403.

Hill, M.O. 1979. *TWINSPAN. A FORTRAN programfor arranging multivariate data in an ordered two-way table by classification of the individuals and attributes.* Section of Ecology and Systematics, Cornell University, Ithaca, N.Y., U.S.A.

Holloway, J.T. 1954. Forests and climate in the South Island of New Zealand. *Transactions of the Royal Society of New Zealand* 82: 329-410.

Holloway, J.T.; Wendelken, W.J. 1957. Some unusual features of sample plot design. *New Zealand Journal of Forestry* 7: 77-83.

James, I.L.; Jane, G.; Barr, C. 1973. *Theforests and sub-alpine shrublands of the Hokitika catchment.* Forest and Range Experiment Station, Protection Forestry Division Report 116. Forest Range and Experiment Station, Rangiora, N.Z.

Kelly, G.C. 1972. *Scenic reserves of Canterbury.* Biological Survey of Reserves Report 2. Botany Division, Department of Scientific and Industrial Research, Wellington, N.Z.

Kelsey, J. 1997. *The New Zealand experiment - a world model for structural adjustment.* Auckland University Press, Auckland, N.Z.

Leathwick, J.R.; Bums, B.R.; Clarkson, B.D. 1998. Environmental correlates of tree alpha-diversity in New Zealand primary forests. *Ecography* 31: 235-246.

Mark, A.F; Baylis, G.T.S.; Dickinson, K.J.M. 1991. Monitoring the impacts of deer on vegetation condition of Secretary Island, Fiordland National Park, New Zealand: A clear case for deer control and ecological restoration. *Journal of the Royal Society of New Zealand 21:* 43-54.

Masters, S.E.; Holloway, J.T.; McKelvey, P.I. 1957. *The national forest survey of New Zealand,* 1955. New Zealand Forest Service, Wellington, N.Z.

McCune, B.; Mefford, M.J. 1999. *PC-ORD. Multivariate analysis of ecological data, Version* 4. MjM Software Design, Gleneden Beach, Oregon, U.S.A

McKelvey, P.I. 1963. The synecology of the West Taupo indigenous forests. *New Zealand Science Bulletin* 14. Government Printer, Wellington, N.Z.

McKelvey, P.J. 1984. Provisional classification of South Island virgin indigenous forests. *New Zealand Journal of Forestry Science* 14: 151-178.

McKelvey, P.I. 1995. *Steepland forests.* Canterbury University Press, Christchurch, N.Z.

McKelvey, P.I.; Cameron, R.I. 1958. Design for a forest study. *New Zealand Journal of Forestry* 7: 116-122

Meurk, C.D.; Buxton, R.P. 1991. *A New Zealand register of permanent vegetation plots.* Department of Scientific and Industrial Research Land Resources, Christchurch, contract report 91/35, prepared for the Department of Conservation. Department of Scientific and Industrial Research, Christchurch, N.Z.

Michener, W.K.; Brunt, J.W.; Helly, J.J.; Kirchner, T.B.; Stafford, S.G. 1997. Nongeospatial metadata for the ecological sciences. *Ecological Applications* 7: 330-342.

Michener, W.K. 1998. Ecological metadata. *In:* Michener, W.K.; Porter, J.H.; Stafford, S.G. (Editors). *Data and information management in the ecological sciences: A resource guide,* pp. 47-51. Long Term Ecological Research Network Office, University of New Mexico, Albuquerque, New Mexico, U.S.A.

Mueller-Dombois, D.; Ellenberg, H. 1974. *Aims and methods of vegetation ecology.* John Wiley & Sons, New York, U.S.A.

Myers, S.C.; Park, G.N.; Overmars, F.B. 1987. The *New Zealand Protected Natural Areas Programme - a guidebookfor the rapid ecological survey of natural areas.* New Zealand Biological Resources Centre Publication No.6. Department of Conservation, Wellington, N.Z.

Nash, J.E. 1993. Ownership and outreach: A model for administration of shared data. *Annals of the Missouri Botanical Garden 80:* 304-308.

National Institute of Standards and Technology. 1992. *Spatial data transfer standard (Federal Information Processing Standard* 173). National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.

Nicholls, J.L. 1976. A revised classification of the North Island indigenous forests. *New Zealand Journal of Forestry* 21: 105-132.

Osawa, A; Allen, R.B. 1993. Allometric theory explains self-thinning relationships of mountain beech and red pine. *Ecology* 74: 1020-1032.

Pairman, D.;McNeill,S.; Scott, N.; Bellis, S. 1999. Vegetation identification and biomass estimation using AlRSAR data. *Geocarto International* 14: 67-75.

Payton, 1.1.; Hall, G.; Burrows, L.; Hunt, J. 1988. *Establishment of a vegetation survey database - progress report,* 1987/88. Forest Research Institute report prepared for the Department of Conservation. Forest Research Institute, Christchurch, N.Z.

Payton, 1.1.; Pekelharing, C.J.; Frampton, C.M. 1999. *Foliar browse index: A method for monitoring possum* (Trichosurus vulpecula) *damage to plant species and forest communities.* Manaaki Whenua - Landcare Research, Lincoln, N.Z.

Peet, R.K. 1998. ESA jour'!.als: Evolution and revolution. *Bulletin of the Ecological Society of America* 79: 177-181.

Porter, J.H. 1998. Scientific databases for environmental research. *In:* Michener, W.K.; Porter, J.H.; Stafford, S.G. (Editors), *Data and information management in the ecological sciences: A resource guide,* pp. 4146. Long Term Ecological Research Network Office, University of New Mexico, Albuquerque, New Mexico, U.S.A

Porter, J.H.; Callahan, J.T. 1994. Circumventing a dilemma: Historical approaches to data sharing in ecological research. *In:* Michener, W.K.; Brunt, J.W.; Stafford, S.G. (Editors), *Environmental information management and analysis: Ecosystems to global scales,* pp. 193-202. Taylor & Francis, London, U.K.

Rose, A.B.; Wiser, S.K.; Platt, K.H. 1994. Forest susceptibility to browsing by possums. *In:* Rose, A.B. (Editor), *A review of possums and possum-vulnerable species in Nelson/Marlborough Conservancy,* pp. 8-21. Landcare Research Contract Report LC9394/119. Landcare Research, Blenheim, N.Z.

SAS Institute, 1999. The SAS system for Windows, release 8.00. SAS Institute, Cary, North Carolina, U.S.A.

Scott, D. 1965. A height frequency method for sampling tussock and shrub vegetation. *New Zealand Journal of Botany* 3: 253-260.

Scott, D. 1993. Time segment analysis of permanent quadrat data: Changes in *Hieracium* cover in the Waimakariri in 35 years. *New Zealand Journal of Ecology* 17: 53-57.

Smale, M.C.; Hall, G .MJ.; Gardner, R.O. 1995. Dynamics ofkanuka *(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.

Stafford, S.G. 1993. Data, data everywhere but not a byte to read: Managing monitoring information. *Environmental Monitoring and Assessment* 26: 1125-141.

ter Braak, C.J.F.; Smilauer, P. 1998. *CANOCO reference manual and user's guide to Canoco for Windows: Software for Canonical Community Ordination (version* 4). Microcomputer Power, Ithaca, New York, U.S.A.

Thomson, A.P. 1946. Design for a forest survey. *New Zealand Journal of Forestry* 5: 191-199.

Wardle, J.A. 1970. The ecology of *Nothofagus solandri.* 2. The associations. *New Zealand Journal of Botany* 8: 532-570.

Wardle, J.; Guest, R. 1977. Forests of the Waitaki and Lake Hawea catchment. *New Zealand Journal of Forestry Science* 7: 44-67.

Webster, B. 1994. Vegetation monitoring - Landcorp Property Ltd programme and data. *In: Proceedings of the* 1994 *New Zealand Conference on Sustainable Land Management,* pp 192-198. Lincoln University, Canterbury, N.Z.

Whitehouse, I. 1998. *Science database and collection issues: Oceans of data, vulnerable collections, and terabytes of power.* Ministry of Research, Science and Technology, Wellington, N.Z.

Wilson, H.D. 1987. Vegetation of Stewart Island, New Zealand. *New Zealand Journal of Botany Supplement:* 1-80.

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.* Manaaki Whenua Press, Lincoln, N.Z.

Wraight, MJ. 1962. Methods of measurement of alpine grassland. *In: Methods of measuring plant communities: Report of a seminar held at Cass on 16-18 October [1962J under the auspices of the New Zealand Institute of Agricultural Science with the co-operation of the Botany Department, University of Canterbury,* pp 15-23. New Zealand Institute of Agricultural Science, Wellington, N.Z.

Apendix 1.

NVS Protocol

Purpose of the National Vegetation Survey (NVS) Databank

The goal is to develop NVS as New Zealand's prime repository for ecological data on vegetation structure and composition, and specifically:

1. to enhance archival data storage of nationally important datasets;
2. to enhance availability of archived data to users, while protecting the interests of data suppliers;
3 and to encourage users of stored data to provide some benefit which enhances NVS as an in-kind contribution.

Protocol for Data Deposition and Storage in NVS

4. Hard copies of data, electronic copies where available, and documentation about the data should be provided.
5. No costs will be associated with data deposit, storage and retrieval by the provider.
6. Landcare Research will not normally purchase data for inclusion in NVS.

Issues of ownership of, and access to, data are of concern to data providers. Specific conditions regarding issues of ownership and access will be clearly defined in a Memorandum of Understanding between Landcare Research (as curators and custodians of NVS data) and providers (as per the attached Agreement on Confidential Disclosure of Information and Memorandum of Understanding). All data sets provided will be assigned one of the access levels listed below in consultation with Landcare Research.

Levels of Proprietary Ownership

Level I (**Open Datasets**) No limitation on availability of data. The provider puts no conditions on use of the data;
Level 2 (**Conditional Datasets**) The existence ofthese data will be shown on data listings, but use is restricted by the provider. Written approval must be obtained from the provider before data will be supplied;
Level 3 (**Reserved Datasets**) Confidential or commercially sensitive (the existence of data will not be advertised; they will be archived in NVS predominantly for data security).

Protocol for Data Use

8. Data contained in NVS hardcopy andlorelectronic files are copyright and subject to Licence Agreements where used by any party.
Licensed users of NVS data may not use the data for any purpose other than the purpose specified in the Licence Agreement, Or subsequently agreed in writing between Landcare Research and the Licensee.
9. Licensed users of NVS data may not pass this information to any other party in any form unless this use is specifically provided
for in the Licence Agreement, or subsequently agreed in writing between Landcare Research and the Licensee.
10. Data are provided on a single-use basis unless otherwise negotiated.
11. Modification or addition of ancillary data does not confer ownership of the original data to the user.

Cost of Data Retrieval

12. Costs of data handling (e.g., retrieval, copying, analysis) must be met by the user.

Acknowledgements

13. A clear acknowledgement of NVS as a data source must appear in any products (e.g., publications, unpublished reports) in the following terms:
*We <**or User/Institution name**> acknowledge the use of data <**or other information**> drawn from the National Vegetation Survey Databank (NVS).*
Additional acknowledgement of the original collector or organisation may also be necessary as a condition of use.

Data Accuracy

14. Landcare Research attempts to hold the most up-to-date and complete copies of data in NVS, but does not guarantee that all data are error-free. Users are encouraged to furnish copies of updated or corrected data or plot remeasurement data within a sensible time frame for the purpose of updating records.

User Lists

15. Landcare Research will maintain a log of data users for reporting purposes (e.g., to Public Good Science Fund). Information on other users of requested data may be provided at cost and to the extent allowed by Licence Agreements.

Last updated 14 Jun 2000