Data-Standard Guidelines for Improving the Quality of Permanent Plot Data Archived in the National Vegetation Survey Databank — First Approximation

C.L. Newell and A.B. Baldwin

Landcare Research
PO Box 69, Lincoln 8152
New Zealand

Landcare Research Contract Report: LC9900/140

PREPARED FOR:
The Foundation for Research, Science and Technology
PO Box 12 240, Wellington
New Zealand

DATE: August 2000

Contents

1. Summary

1.1 Project and Client

This report outlines a set of initial standard guidelines for the recording and correcting of data from permanent forest plots archived in the National Vegetation Survey (NVS) databank. These guidelines are intended to achieve a consistent quality of data recording and a consistent manner of treating errors. The guidelines have been formulated as part of a quality audit of data archived in the NVS databank. The project was carried out by Landcare Research, Lincoln, between December 1999 and June 2000 and was funded by the Foundation for Research, Science and Technology as part of the Landcare Research's Database Integration Project.

1.2 Objectives

- To develop a set of standard guidelines for data that are transparent, logical and repeatable in the future to improve the quality and ensure consistent quality control of existing data.
- To preserve the integrity of the original archival data as well as improving the quality of the data archived.
- To develop a standardised error-correction method using a Correction-Problem Sheet to record corrections made to the data and inconsistencies that should be resolved with a future remeasurement.
- For each plot, to maximise the compatibility of information between different data collection methods within a remeasurement year and between remeasurement years.
- To apply these principles by data-checking the set of benchmark datasets described by Bellingham et al. (2000).

1.3 Methods and Results

- We outlined nine guidelines that ensure that errors identified in the permanent-forest-plot data archived in the NVS databank are handled in a consistent manner across all datasets.
- We applied these guidelines to a set of benchmark datasets described by Bellingham et al. (2000).
- One of our guidelines was to develop a Correction-Problem Sheet that archives changes that are made to data and itemises queries that typically can only be resolved with additional information gained during a future remeasurement.
- We also developed a lineage for each benchmark survey to understand how datasets within a survey relate to one another and to document the type of data available for each measurement. The lineage summarised the total number of plots measured per each sampling method in a remeasurement and tracked each plot across all measurements.
- We applied the guidelines to four specific data-quality problems that had previously been identified from permanent-forest-plot data. We undertook a close look at specific problems associated with inconsistent plant species codes, tag number duplication, grid references, and other site information, and developed a set of rules to solve the most common problems identified.

1.4 Conclusions

- Our set of standard guidelines provide the foundations for improving the quality of permanent-forest-plot data archived in the NVS databank. These guidelines set out error-checking methods that are consistent, repeatable, and can be tracked by future researchers, and, moreover, preserve the integrity of original data.

1.5 Recommendations

- These guidelines represent an initial set of standards for (a) reducing inconsistencies and errors in existing NVS data from permanent forest plots (see Sections 5.1 and 6.2), (b) minimising errors and inconsistencies when entering data (see Section 6.3), and (c) enhancing the quality of data during the field collection phase (see Section 6.4).
- These guidelines should be trialled, refined, and extended in the future as other data quality issues are addressed.

2. Introduction

This report outlines a set of initial standard guidelines for the recording and correcting of data from permanent forest plots archived in the National Vegetation Survey (NVS) databank. These guidelines are intended to achieve a consistent quality of data recording and a consistent manner of treating errors. The guidelines have been formulated as part of a quality audit of data archived in the NVS databank. The project was carried out by Landcare Research, Lincoln, between December 1999 and June 2000 and was funded by the Foundation for Research, Science and Technology as part of the Landcare Research's Database Integration Project.

3. Background

3.1 National Vegetation Survey Databank

The National Vegetation Survey (NVS) databank is a physical archive and archival computer databank that contains physical data sheets and associated material (e.g. maps, aerial photos) and computerised databases for approximately 45 000 vegetation survey plots. It includes data from c. 14 000 permanent plots and c. 8000 Protected Natural Areas Programme (PNAP) survey records, which, collectively, cover a broad range of habitats and geographic locations across New Zealand.

The NVS databank (or NVS for short) is a unique, nationally–significant collection of vegetation data, which spans over 50 years of data collection. As such it represents many decades of field surveys, data compilation and checking. The databank was developed from surveys conducted by the former New Zealand Forest Service, Department of Lands and Survey, and DSIR Botany Division. In recent years data have steadily been added to NVS from surveys and research conducted by the Department of Conservation (DOC), regional councils, universities, and Landcare Research. Such widely sourced information collated in one databank is part of the value of NVS to New Zealand. At the same time, the interests of data providers are protected through written agreements that determine access rights to specific datasets archived in the databank.

The NVS databank archives two major types of vegetation data (see Allen 1992, 1993; Wiser & Rose 1997; Wiser et al. 1999 for specific details of methods): (1) Point-based compositional (and usually structural) data are available for *general survey plots*. These typically are not permanently marked. Reconnaissance (recce) plots and PNAP plots are examples of general survey data; and (2) NVS also archives approximately 14 000 *permanent plots*. Permanent plots are fixed-area plots or transects where precise vegetation measurements (e.g., tagged trees, species lists, stereophotographs) have been collected using a standard method.

For permanent forest plots, permanently tagged individuals and seedling subplots allow repeat measurements of vegetation composition and structure. Three types of data are typically collected from a permanent forest plot. These include the diameter (hereafter known as *diameter* or *quad* data), measured at breast height (1.35 m), of each stem ∃2.5 cm in diameter, the frequency of individual species found in permanently marked seedling subplots (hereafter known as *understorey* data), and the composition of all vascular species present in the permanent plot, assessed using a reconnaissance plot (hereafter known as *recce* data). Approximately 35% of permanent forest plots have been remeasured at least once since their establishment. When plots are remeasured Landcare Research encourages that data are recorded on preprinted plot sheets that itemise the previous measurements for each tagged tree, so that field staff can compare new information with diameter measurements and species identification from past measurements. A programme for producing the preprinted plot sheet is available in the software package PC-DIAM (Hall 1994).

All hard-copy archival plot sheets and master digital databases are archived in NVS at Landcare Research, Lincoln, where staff administer and oversee data additions, error checking, database maintenance, and liaise with users. Hard-copy plot sheets are archived in a fireproof vault in boxes specific to each survey. These boxes are ordered by their ecological district number. Each archive-box is labelled with its survey name, survey year, survey type (e.g., forest or grassland), sampling method (e.g., recce, understorey, quad), and ecological district number. General information such as the number of plots sampled, a printout from the NVS standard error-checking programme, and other miscellaneous information associated with the survey (e.g., often maps, information on unique methods) are stored in the archive-box. Plot sheets are generally sorted by sampling method in the archive-box.

Digital databases are generally archived according to DOC Conservancies. Each Conservancy has separate subdirectories for each of the major sampling methods. There is also a subdirectory for miscellaneous information, such as non-standard environmental data, text files providing additional information, and data methods for a particular survey. Quality and storage of the digital data is important. Accordingly, all changes and additions to the digital data archived in NVS are directed through one specified NVS administrator.

3.2 Need for data standards

Even the most careful researchers occasionally make mistakes when collecting and processing data. This means that errors can potentially occur even in high-quality databases. The level of errors in a database can be reduced by developing formal guidelines for adding, updating and editing data. It is important to remove errors before even the simplest of analyses take place; left in they can lead to spurious results and misleading conclusions.

One of the main purposes of the NVS is to provide archival data storage of nationally important datasets. Recently Wiser et al. (1999) identified major types of errors that occur in NVS, quantified the extent of these errors, and made recommendations on how these problems might be rectified. This present report follows on by developing a set of proposed guidelines that ensure consistent and repeatable quality-control methods that can be tracked by future NVS data caretakers. Furthermore, these guidelines were set to ensure that the integrity of the original data is preserved. A major aim is to ensure that all corrections made and problems identified for future resolution are adequately documented. We use these guidelines to identify and correct, where possible, errors in a set of benchmark permanent-plot databases that are archived in NVS. We discuss each type of error and the problems associated with it in detail, providing a series of correction rules for dealing with specific variants of an error. Finally, we recommend a set of simple guidelines to help improve the quality of existing NVS databases and data collected in the field.

3.3 Definitions

Archival digital data — the master digital file of a survey measurement archived in NVS.

Archival plot sheet — the hard-copy plot sheet archived in NVS: either the original or a high-quality photocopy of the plot sheet on which the information was recorded in the field during the measurement.

Dataset — digital data associated with one or all measurements of a survey.

Metadata — descriptive information about a survey and/or remeasurement.

Permanent plots — permanently located, fixed-area plots or transects in which precise vegetation measurements are taken.

Remeasurement — a resampling of a set of permanent plots after the initial measurement, generally a full or partial measurement of a survey.

Partial remeasurement — remeasurement that only includes a subset of the original set of plots.

Sampling methods — the three standard types of data collected at a permanent forest plot. These include diameter (or quad) data, understorey (or seedling) data, and composition of all vascular plants assessed using a reconnaissance (or recce) plot.

Survey — a set of permanent plots that have been established in (usually) one geographic region at (usually) one time period. The term “survey” is often misapplied.

Tag — a metal label annotated with a multiple digit number, which is attached at 1.35 m to each woody stem ∃2.5 cm in diameter at breast height (1.35 m).

4. Objectives

- To develop a set of standard guidelines for data that are transparent, logical and repeatable in the future to improve the quality and ensure consistent quality control of existing data.
- To preserve the integrity of the original archival data as well as improving the quality of the data archived.
- To develop a standardised error-correction method using a Correction-Problem Sheet to record corrections made to the data and inconsistencies that should be resolved with a future remeasurement.
- For each plot, to maximise the compatibility of information between different data collection methods within a remeasurement year and between remeasurement years.
- To apply these principles by data-checking the set of benchmark datasets described by Bellingham et al. (2000).

5. Methods and Results

5.1 Standard procedures for data quality

We developed the following set of standard guidelines for permanent forest plots to (1) ensure that any errors identified in NVS data are handled in a consistent manner using procedures that are understandable, logical and repeatable in the future, and (2) ensure that summary information is consistent for all datasets archived in NVS.

5.1.1 Data quality standards

(1) *Information on the physical plot sheets represents the most accurate archival record of data:* Where there are discrepancies between the archival digital file and plot sheets, the latter are considered to contain the most accurate record. This guideline will only be ignored where there is written evidence to suggest that information in the digital data is more accurate, e.g., where a text file accompanying a survey noted that all *Uncinia uncinata* on the plot sheets had been renamed *Uncinia clavata* in the digital file, because this taxon had been consistently misidentified in the field.

(2) *Physical plot sheets will be archived in an original, untouched form, to ensure the integrity of the original data collected:* Original plot sheets will be archived where possible, as some information may be lost on photocopies. If this is not possible high-quality photocopies will be archived. Notes will not be added to the plot sheet after data collection as this can lead to confusion and reduce the clarity of the original data. This is particularly true if an error has been recorrected several times.

(3) *Corrections and queries associated with a plot will not be written on the archival plot sheets but on an accompanying NVS Correction-Problem Sheet (see below) which will be archived with the plot sheet. Where possible, this information will also be stored in an associated ASCII format text file:* The one exception to this is that missing information for the sampling date (year at a minimum) and survey name should be written on the plot sheet to ensure that the metadata is complete.

(4) *An NVS Correction-Problem Sheet will be archived with every plot that has had identified errors corrected and/or itemised to be resolved during a future remeasurement. Where possible, this information will also be stored in an associated ASCII format text file:* The specific nature of each correction and problem will be written on the Correction-Problem Sheet, followed by a small explanation for the correction or query, and signed and dated by the annotator (see Section 5.1.2).

(5) *An individual plot represents the standard archival unit so that it can be accessed as a stand-alone unit, independent of all other plots in a survey:* This is in part necessitated by a growing trend for individual plots to be selected from a number of different surveys for larger-scale analyses (e.g., monitoring carbon stocks, Coomes et al. unpubl. data) and for remeasuring subsets of an original set of permanent plots (e.g., see 5.1.3(2)).

(6) *Provide clear summary information for each survey tracking the lineage of each permanent plot:* We have developed a lineage for each benchmark survey (see Section 5.1.3), summarising the years that each plot was remeasured, the types of sampling methods used and the names of all digital files with information pertaining to each plot.

(7) *Taxonomic names in the archival digital data will reflect those used on the archival plot sheets to preserve the integrity of the original data collected:* Nomenclature will not be updated to reflect taxonomic names in current use. Accordingly, there may be discrepancies between remeasurement years for the taxonomic name of a given individual. Such discrepancies will be noted in an archival ASCII-formatted digital text file.

(8) *Each survey will be accompanied by an ASCII-formatted digital text file:* The digital text file will include notes on methodology, the names of team members involved in each measurement, taxonomic changes between different remeasurement years, and other notes that help clarify methods/usage of the information archived (e.g., plot name changes between remeasurement years, reasons why permanent plots were remeasured on two consecutive years).

(9) *Changes must be logical and clear to someone with little prior knowledge of the data.*

5.1.2 Correction-Problem sheet

In the past, corrections and queries associated with the data have been noted on the plot sheets. This has caused many difficulties. The most common problems are listed below.

(1) Little explanation (if any) for correction(s), as there is no space designated for corrections on the plot sheet. For example, in Appendix 9.1 there is no explanation for the double arrow between tag numbers 7156 and 75 in subplot A.

(2) Inconsistent procedures for recording corrections. Example one: the order of corrections listed on a plot sheet can not be tracked if corrections were not dated. Accordingly, it is not possible to determine whether a change on a plot sheet was made in the field at the plot location, or 10 years later by a researcher in the office who decided that a specific change should be made. Example two: there are four different plot numbers written on each plot sheet. It is impossible to determine which is the appropriate numbering system.

(3) Corrections are often written on a plot sheet in ink of a contrasting colour. However, once the plot sheet is photocopied, these corrections can not be distinguished from the original data. Photocopies of archival plot sheets are often sent out to a client for a forthcoming remeasurement. Similarly, photocopied plot sheets are often archived in NVS. We illustrate this with a comparison of the colour photocopy of Plot 11 in Appendix 9.1 with a black and white photocopy of the same plot in Appendix 9.2.

(4) Information on an archival plot sheet becomes increasingly confusing as corrections are added to the plot sheets, compromising the integrity of the original data collected. The problem is compounded if a correction has been recorrected several times, at times compromising the legibility of the original field data (Appendix 9.1).

(5) Under the past system errors were often repeated in subsequent remeasurements because there was no formal method of passing on updated information to monitoring agencies unless they requested a copy of the archival plot sheets prior to a remeasurement.

We have developed a Correction-Problem Sheet to overcome these problems and to provide a link between the archival plot sheet and digital data. As outlined in our proposed data-quality standards, a Correction-Problem Sheet will accompany each plot sheet where corrections have been made and/or there are problems that need to be resolved. We have developed a plot-specific Correction-Problem Sheet rather than summarising corrections and problems for an entire survey, to ensure that individual plots can function as stand-alone units. The Correction-Problem Sheet consists of two pages, with (1) corrections and (2) problems that need to be to resolved at a future remeasurement itemised on separate pages (Appendices 9.3, 9.4). The survey and plot identifiers, sampling method and date are at the top of each page, enabling both pages to function as separate entities. Each correction and problem is itemised on a separate line.

Correction sheet: The correction sheet is used to note corrections that can be made immediately. Each correction is recorded on the correction sheet (Appendix 9.3), followed by an explanation of the error, and signed and dated by the person who found the error. It is important that the date of the correction is noted as this enables us to track the order of specific changes made to the data over time.

Problem sheet: The problem sheet is used to itemise queries that can not be corrected immediately and typically can only be resolved with additional information gained during a future remeasurement (Appendix 9.4). The person who identified the problem fills out the first two columns, noting the problem and suggested solution, and recording the date and their name. For example, *tree tag number 3455 was recorded as NOTmen in 1978, but NOT fus in 1982: determine correct identification in next remeasurement*. The field team will take a copy of the problem sheet into the field at the next remeasurement to resolve the itemised problems. Once the problem has been resolved the third and fourth columns are filled, noting the solution to the problem, and the name and date of the person who resolved the problem.

Correction-Problem Sheet usage: To date we have mostly used the Correction-Problem Sheet in conjunction with our data quality checks of existing NVS datasets. However, the sheet also has a beneficial role during the initial data entry phase. The Correction-Problem Sheet has been trialled at Landcare Research by Michelle Breach, the NVS database administrator, since March 2000 for this purpose, and it has provided a very useful means of noting corrections and problems that have been identified during the initial data-entry phase. Michelle has mostly used the sheet to record where two individuals in a plot have the same tag number. In this case she has noted that one of the two duplicate tag numbers must be retagged at the next measurement. Copies of the annotated Correction-Problem Sheets, generated during initial data entry and data quality checks, will be sent out to DOC and other vegetation monitoring agencies to ensure that they can keep track of changes made to the archival data held in NVS and make them aware of problems that must be resolved in future remeasurements.

We believe that the Correction-Problem Sheet provides a transparent and logical method of tracking corrections and problems associated with NVS data. We recommend that it should be used by other agencies entering data or undertaking data quality-control checks. Blank Correction-Problem Sheets are available on request.

5.1.3 Lineage

Understanding the lineage of a survey region, how datasets within a region relate to one another, both spatially and temporarily, and what types of data have been sampled for each set of permanent plots is central to NVS. With increasing numbers of remeasurements, and partial remeasurements of specific surveys, each survey needs a detailed lineage to summarise clearly what information is available in the survey and to enable individual plots to be tracked through time between different survey remeasurements. Additionally, a lineage can help overcome the following difficulties that we have encountered.

(1) Although all permanent plots in a survey have been consistently remeasured each year, not all three sampling methods (quad, understorey, recce; see Section 3 for details) may have been used each time.

(2) Partial remeasurements make it difficult to track individual plots over time. This is compounded in cases where one or two plots from specific surveys have been remeasured as part of larger-scale projects (see Appendix 9.5).

(3) The remeasurement survey name and archival digital file name of partial remeasurements does not link intuitively with the original survey and digital file name. For example, in the Hokitika region 'WHITCO' and 'KOKATA' are partial remeasurements of the survey 'HOKITK' (See Appendix 9.5). Some of this information can be tracked through the NVS metadata file.

(4) Similar digital file names for unrelated surveys from a similar geographic region cause confusion. For example, in the Hokitika region 'HOK' and 'HOKITK' have been used for two unrelated surveys (Appendix 9.5).

(5) Changes in plot name between remeasurements make it difficult to track a plot across all remeasurements. This can be complicated by the presence of multiple plot numbers on a plot sheet.

We developed a lineage for each benchmark survey. The lineage of each survey summarises for each measurement (1) the total number of plots measured using each sampling method (quad, recce, understorey) and (2) the digital filename of each remeasurement. Each plot is also itemised by measurement year, listing (3) which sampling methods were used, and (4) the name of the digital file that it has been archived in. For example, in Appendix 9.6 we can see that line 67 plot 1 was established in 1971, remeasured in years 1979, 1979, and 1996, but was not remeasured in 1999. At present we have developed lineages for all benchmark surveys, but intend to generate a lineage for all surveys with permanent plots and remeasurements.

A hard copy of the lineage is archived with the archival plot sheets from each remeasurement. This enables a researcher to quickly determine how the data collected at a particular measurement relates to other years. Each lineage will also be archived digitally. A copy of each lineage will be sent to the relevant DOC Conservancy and other monitoring agencies.

5.2 General data-quality problems in permanent-forest-plot data

Our quality checks were undertaken on a set of benchmark datasets identified by Bellingham et al. (2000). Data were read from ASCII formats into a SAS database (SAS Institute 1998) and programs written by Susan Wiser and Claire Newell were run in SAS to test specific errors associated with the data. Where appropriate, quality checks were made on all three major types of files associated with permanent plots.

Several general problems reoccurred during our quality checks and these are listed below. These vary in magnitude, but all highlight the need for careful data collection methods and archiving. The introduction of the lineage and Correction-Problem Sheet will help reduce some of these problems.

(1) No archival plot sheets for a survey (most likely never deposited in NVS). These are mostly noted on the NVS metadata file.

(2) Difficulties linking archival plot sheets with digital data (due to multiple plot numbers on plot sheet, or some plot sheets missing, plot names changed in digital files for analysis, or digital file name for survey not recorded with plot sheets in archive-boxes).

(3) Difficulties linking plots across remeasurements due to changes in plot name between remeasurement years.

(4) Interpretation of plot sheet is difficult (due to bad photocopy, illegible or untidy writing, dirty plot sheet; see Appendices 9.1, 9.2, 9.7). This is sometimes noted in the NVS metadata file. Photocopied plots sheets often do not capture the 5 mm around the edge of the plot sheet, missing information written near that edge. Note that currently at Landcare Research, photocopies of original plot sheets are individually checked by Michelle Breach to ensure all information is present and legible. Michelle does not return the original plot sheets to the monitoring agency until after data entry and initial error checking are complete.

(5) New data written on photocopy of previous remeasurement (Appendices 9.1, 9.7). In some instances the new data are distinguished from the photocopied data by a different-coloured pen, but such differentiation is lost if the plot sheet is photocopied again.

(6) Important plot information incomplete. For example, date of sampling, the survey, and/or line and plot number are missing. This information is critical for linking a plot with its relevant survey and archive-box.

5.3 Specific data-quality problems in permanent-forest-plot data

5.3.1 Inconsistent plant taxonomic names in diameter data

Wiser et al. (1999) found that 1.4% of trees in the 7564 permanent forest plots that they checked had different species codes recorded in different measurement years. We systematically examined each case where there was a discrepancy in the species code and its associated tag number between different measurement years. Not all the discrepancies have been resolved, due to missing plot sheets. Six types of problems were identified (see Table 1). We discuss these in detail and (where possible) provide a solution to the problem.

Table 1 Types of errors associated with inconsistent species codes found in plant diameter data of benchmark datasets archived in the NVS databank.

Conservancy	Total inconsistences identified	Tag # duplication	Inconsistent identification ⎞Field checks required	Species code corrected	Data not available for checking	Taxonomic name change between measurements	Error in tag number	Other problems
Auckland	0							
Bay of Plenty	26	4		3				19
Hawke's Bay-East Coast	55*							
Northland	0							
Tongariro-Taupo	0							
Waikato	114	27	78	4			2	3
Wanganui	327		75	251		1		
Wellington	416			194	222			
Canterbury	5**		2	3				
NelsonMarlborough	268	4	50	16	5	193		
Otago	86			86				
Southland	826	32	453	234	14	65	28	
West Coast	248	3	125	67	19	4		30

* (Waiting for data from recent remeasurement before checking errors in detail)
** (In addition we are undertaking a detailed plot-by-plot check of one survey)

(1) Plant identification inconsistencies between remeasurements:

Example: Tree tag number 5336 has been identified as Nothofagus solandri in 1971, N. fusca in 1977, 1982, 1990 and 1993 and N. solandri in 1999.

Problem: Tree tag number 5336 has been identified as N. fusca in four out of six measurements. We can not assume that either taxonomic name is "correct" and that the taxonomic identification of an individual has been checked during a remeasurement unless:

(i) the remeasurement used preprinted plot sheets listing species and tag numbers from the previous remeasurement, AND
(ii) a change in identification for a particular tag number was noted on the plot sheet.

If these conditions can not be met the problem must be noted on the problem sheet to ensure that the specific taxonomy of the individual is checked during next remeasurement.

Correction Rules: The taxonomic name of an individual can only change when the "correct identification" has been determined. Changes will be backdated in all digital files and noted on the Correction-Problem Sheets associated with the relevant archival plot sheets (Appendix 9.8).

(2) Nomenclatural change in taxa between remeasurements:

Example: Unambiguous change — tree tag number 5567 was identified as Podocarpus ferrugineus in 1981 and Prumnopitys ferruginea in 1996. Ambiguous change — due to recent taxonomic changes,

taxa identified as Blechnum capense in the past may currently refer to either Blechnum novae zelandiae or Blechnum procerum (Chambers & Farrant 1998).
Problem: As noted earlier, the nomenclature of the digital data must reflect what has been used on the archival plot sheets. Accordingly, there will be discrepancies in the nomenclature between remeasurements.
Correction Rules: Nomenclatural changes between remeasurements will be documented in the digital ASCII text file that accompanies a survey. The correct identification of species with ambiguous taxonomic changes must be resolved during a future remeasurement.

(3) Typing errors:
Problem: These are identified when there are discrepancies between the species codes associated with a tag number in the archival plot sheets and digital files. In some cases this has also lead to tag replication (see below).
Correction Rules: *Information on hard copy plot sheets represents the most accurate record of the data. A discrepancy is corrected in the digital file but does not need to be noted on the Correction-Problem Sheet as it is a typing error.*

(4) Tag number replication: See Section 5.3.2 for details.

(5) Zero versus blank problem:
Problems: Differentiation between zeros and blanks has not been consistently used for tag numbers. There are some instances where a particular tag number has been recorded with a zero in front of the tag number in one measurement year (e.g. 0567), but entered without the zero in a subsequent remeasurement (e.g 567).

This problem can also cause tag duplications (see below) where five-column numbers have been reduced to four digits causing a duplication with another existing tag number. For example, tag C0678 is entered as 0678 and therefore can not be distinguished from tag 678 because of inconsistent use of zero and blank.

The blank versus zero problem partly arises because most field teams are not aware that only the last four digits of a tag number can be entered into the digital file, and that zeros and blanks can not be distinguished. We need to ensure that tags beginning with zero are consistently entered as such and check to see that this is consistent in digital files from previous remeasurements.

(6) Retagged individuals are tagged with a new tag number: We identified a few cases where an individual tree had been retagged with a new number, rather than retagging with the original number. This can cause many problems, including duplicating tag numbers within a plot.
Correction Rules: *Always retag an individual with its original tag number.*

5.3.2 Tag number duplications in diameter data

A search for duplicate tag numbers found that <1% of plots in the benchmark datasets had tag duplications. However, additional duplications were identified during our taxonomic names check. Some of these were also encountered by chance whilst checking for other errors. The types of tag replication errors that we have encountered are itemised below.

(1) Tag number used twice for different individuals in a plot: Data-checking routines in PC-DIAM (Hall 1994) identify plots with duplicate tag numbers. In the past, if this problem was identified at the initial data entry stage, one of the duplicate tags was typically “digitally” retagged in the digital file. Where possible, a “0” was added in front of the tag number to distinguish it from its

duplicate. This, however, can cause problems, as noted in the previous section. In some surveys duplicate tag numbers have been digitally retagged with the following series; "0001", "0002" etc. However, we are not sure how widespread this practice has been. Duplicate tags are now noted on the Correction-Problem Sheet.

In many cases such digital tag number reassignments have not been archived. Accordingly, it is difficult to link information between the digital and archival plot sheets. The problem can be compounded during the next remeasurement if photocopies of previous measurements are not taken into the field (also next error itemised).
Suggested Correction Rule: *Where possible, one individual with a duplicate tag number will be "digitally" retagged in the digital file using a number that has not been previously used in the plot. The new tag number will begin with an "*" followed by the duplicate tag number if it is three or fewer digits. Duplicate tag numbers that are four digits will be retagged as "***1", "***2" etc. This change will be backdated through all digital files and noted on the Correction-Problem Sheet, to ensure that the archive plot sheet is consistent with the digital files. The digitally retagged individual must be physically retagged during the next remeasurement and then backdated through all digital files*.

(2) Inconsistent reassignment of tag numbers in different remeasurement years: In one survey tag duplications have been treated differently in two different measurement years. In our example there are two trees with tag 516 in a plot. In 1978 one of the duplicates was assigned tag 0001. However, in 1988 one duplicate was assigned the tag 0516. As a result there is no link between these individuals in the two measurements.
Suggested Correction Rule: *Where possible, one individual with a duplicate tag number will be "digitally" retagged in the digital file using a number that has not been previously used in the plot. The new tag number will begin with an "*" followed by the duplicate tag number if it is three or fewer digits. Duplicate tag numbers that are four digits will be retagged as "***1", "***2" etc. This change will be backdated through all digital files and noted on the Correction-Problem Sheet, to ensure that the archive plot sheet is consistent with the digital files. The digitally retagged individual must be physically retagged during the next remeasurement and then backdated through all digital files.*

(3) Tag from a dead individual reused for a new individual: Sometimes tags from dead individuals have been reused to tag new individuals as a way of economising tag use. However, this is not recommended. This problem is difficult to identify. We have mostly identified cases where the dead and new individual had a different species code quite by chance during species code inconsistency checks. However, detecting cases where the dead and new individuals have the same species code will be more difficult. It is particularly important to solve the latter case as analyses including data from more than one measurement would treat the dead and new individual as one individual. This would produce inaccurate results.
Correction Rule: *Digitally retag the dead individual with a new tag number beginning with "***" using a number that has not previously been used in the plot, e.g., "***1", "***2", etc. Backdate the change to the dead individual in all relevant digital files and note the new tag number for the dead individual on all relevant correction sheets.*

(4) Duplicate tag numbers for two individuals in a plot, but one individual died prior to latest remeasurement, or both individuals have died:
Correction Rule: *Digitally retag the dead individual with a new tag number beginning with "***" using a number that has not previously been used in the plot, e.g., "***1", "***2", etc. Backdate the*

change to the dead individual in all relevant digital files and note the new tag number for the dead individual on all relevant correction sheets.

(5) Tag number duplication because the full tag number was not written on the plot sheet:
Example: In some instances only the first tag number in a sequence will be written in full on a plot sheet and all subsequent tag numbers will only list the last one or two digits. Whilst this may have been clear to the field team as the time, such abbreviations are not always clear to others trying to interpret what has been written down at data entry and later checking stages. In some cases tag abbreviations have caused tag duplications in the digital data. Appendix 9.1 shows several examples of this problem. For example, in the first column, tag number 58 and 57 follow tag number K7155, and most likely represent tag numbers K7158 and K7157, respectively. However, the full sequence of tag numbers can not always be understood. For example, half way down the second column, it is not obvious whether tag numbers 62 and 61 should follow the sequence associated with tag numbers L1257 or L1305.
Correction Rule: *Note duplicate tag numbers on problem sheet for clarification during next remeasurement. Always write the full tag number on the plot sheet. The use of preprinted plot sheets in remeasurements will help reduce this problem if it does not occur in the original measurement.*

(6) Tag duplication because full tag number could not be entered in the digital data:
Problem: The current file formats only provide four digits for a tag number. Accordingly, tag numbers with more than four digits are truncated and may appear to duplicate another tag number in the plot. This prevents, for example, tag number C5467 from being distinguished from tag number D5467 as both would be entered as 5467.
Correction Rule: *Note duplicate tag numbers on problem sheet for clarification during next remeasurement. Always write the full tag number on the plot sheet.*

5.3.3 Grid references

Grid references are used in our analyses of data-quality checking as part of the unique identifier for a plot. Because there are duplicate numbering systems between surveys, each plot is identified by plot name (line number (if present) and plot number) and grid reference. For this reason alone it is important to reduce the errors associated with grid references. Wiser et al. (1999) discussed some of the reasons for lack of precision for grid references. We identified six types of errors.

(1) Low-resolution grid references:
Problem: Low-resolution codes that have been referenced as “-00-00” rather than “.--.--“; for example, “5..4..” was recorded as “500400”. However, “-00-00” numbers represent an actual point in space, denoting a higher level of accuracy than has been measured. This problem is not uncommon throughout NVS, although only 30 plots in the benchmark surveys were found to have true low-resolution (“.--.--“) grid references. These were checked and most were found to be an accurate record of what was recorded on the plot sheet.
Correction rule: *Low-resolution codes will be assigned as “.--.--“* (e.g. 5--4--).
Resolution: Note on problem sheet that the grid reference must be accurately measured during a future remeasurement.

(2) Changes in the grid references of a plot between remeasurement: We determined the consistency of grid references between quad, recce and understorey files for a particular plot between measurement years. Twelve plots had inconsistent grid references between remeasurements. These were mostly understorey files.

Resolution: We checked the grid references of plots that were marked on a topographic map to resolved this problem. Discrepancies should only remain if (i) plot locations marked on maps in different measurement years disagree, or (ii) units (imperial versus metric) differ. Such cases must be noted on the problem sheet to ensure that the grid reference is accurately measured during a future remeasurement.

(3) Duplicate grid references for different plots: We identified 555 cases, across all three sampling methods, where there were duplicate grid references for different plots. Approximately 40% of these were associated each with recce and quad files. This problem also includes several cases where all permanent plots on a continuous transect have the same grid reference. In such cases a note on the precision of the grid reference information should be added to the digital text file. Where possible, we recommend that continuous transects have grid references calculated for each permanent plot in future remeasurements.

(4) Grid reference consistently defaulted from initial or early measurement: Some surveys use the grid reference from the first measurement, or the initial measurement with a grid reference, as the default for all subsequent (and previous) measurements. We can not determine the accuracy of this information.

(5) Grid references often only recorded in the digital files: This prevents checks with archival plot sheets. In some cases this relates to work by Martin Fastier in the late 1980s where grid references were generated by assuming plot locations were placed equidistantly along plot lines of known location, or interpolated between plots with a known location. Where relevant, the ASCII text file should note that grid references were calculated using this method.

(6) No grid reference: This was not a major problem in our data quality work due to efforts in the late 1990s by Landcare Research staff to calculate coordinates for permanent forest plots lacking grid references. Staff were able to calculate the grid reference of a plot if it was located on a topographic map. We list this problem to record the method for generating the grid reference for an existing plot.
Correction Rule: *Grid references can be generated for a plot if the plot location is marked on a topographic map and/or if plot altitude and plot line position are known.*

5.3.4 Site information other than grid references

We produced a list of plots with incomplete information on header files (also see Wiser et al. 1999) and checked the information on the archival plot sheets. Discrepancies between different sampling methods were checked by referencing header files from different plot methods. In most cases information in the digital files was consistent with the archival plot sheets. The following problems were encountered.

(1) Site information defaulted from one remeasurement to another: In some cases altitude, aspect, slope, mean top height and ground cover percentages have been defaulted from one measurement to another. This was identified by comparing missing information on the plot sheet with information in the digital file that related to information found on a subsequent or previous remeasurement. These site variables can change over time and should be remeasured with each remeasurement. This will, however, produce another set of inconsistencies which will need addressing.

(2) Site information backdated from later measurement: We also found cases where site information collected at a second (or later) measurement was backdated to header lines in the digital files of earlier measurements.

(3) Discrepancies in site information between measurements: The majority of the plots had small discrepancies in altitude, slope and aspect between remeasurements. In most instances we expect that these relate to measurement errors or difficulties rather than changes in site information (also see Wiser et al. 1999). Variation in altitude most likely relates to differences between field-derived and map-derived measurements and variation in barometric pressure during different measurements. Similarly, accurate measurements of slope and aspect can be difficult to obtain on plots with complex topography.

(4) Missing values for site information: In some cases missing values for altitude, aspect and slope were recorded as zero. This is misleading as zero represents an actual measurement. In the past missing values have sometimes been entered as “99” or “999”. However, it is Landcare Research policy to record a blank value for missing site information rather than add any numerical data.

6. Recommendations

This report provides a set of initial standards for recording and correcting data from permanent forest plots archived in the NVS databank. We recommend that these standards be trialled, refined, and extended in the future as other data-quality issues are addressed to enable a comprehensive set of guidelines to be developed.

6.1 General recommendations

- Claire Newell will meet with Landcare Research staff who manage the NVS databank to discuss whether to accept the suggested correction rules itemised in this report.
- Reconfigure file formats to increase the flexibility of NVS digital files, in particular to allow tag numbers with >4 digits and grid references with >3 digits for eastings and northings.
- Continued financial support from the Foundation of Research, Science and Technology to fund development of data-standard guidelines for permanent plot data archived in the NVS databank.
- Conduct checks similar to those undertaken in this report on other permanent-plot databases archived in NVS.
- Ensure that standard blank plot sheets have enough blank space around the edge to prevent loss of information during photocopying.

6.2 Reducing inconsistencies and errors in existing NVS permanent-forest-plot data

- Follow proposed standard data-quality procedures outlined in 5.1.
- Any changes made to digital files must also be written on a correction sheet.
- Obtain hard copies of all missing plot sheets, if possible.
- Replace low-quality photocopied plot sheets archived in NVS with originals (where possible) or high-quality photocopies of original plot sheets.
- Ensure that no information is lost from original plot sheets on photocopies.

- Ensure that all archival plot sheets have the sampling date, survey name, line and plot number written on them.
- Develop a lineage for all permanent plot surveys in NVS.
- Develop a lineage for a survey before starting any error checking.
- Determine the precision and consistency of grid references between sampling methods within a measurement and between years in other permanent plot data.
- Determine where site information has been defaulted from original or subsequent remeasurement and note these in the digital text file accompanying a survey.
- Determine the consistency of tag numbers beginning with zero and blank between measurements.
- Check zero values for site information to determine how many represent missing data.
- Grid-reference discrepancies should be resolved by geo-referencing plot locations with a global positioning system (GPS).
- Allow >3 digits (measured with a GPS) for eastings and northings to record grid references of closely grouped permanent plots and continuous transects.
- Obtain more accurate altitude measurements (e.g., using a GPS or regularly corrected altimeter).
- For each plot archive on a Corrections-Problems Sheet all corrections made to the data and queries to be resolved during the next remeasurement.
- Send a copy of each lineage to the relevant DOC Conservancy and other monitoring agencies.
- Send Corrections-Problems Sheet to DOC and other vegetation monitoring agencies.
- Develop a method to ensure that monitoring staff take problem sheets into the field.
- Itemise all one-to-one nomenclatural changes to update users on current nomenclature and its relationship with archival data.

6.3 Minimising errors and inconsistencies during data entry phase

- Undertake a rigorous check of digital data entered, by either fully comparing digital data entered with data on plot sheets, or enter the data twice and compare the two digital versions.
- All changes made to the digital data must be written on a Correction-Problems Sheet and noted in an ASCII text file.
- Follow up on species codes corrected in the field and, where appropriate, backdate changes in digital files and on a Correction-Problems Sheet to accompany the archival plot sheets.

6.4 Minimising errors and improving the quality of data during field collection phase

- Contact NVS staff to obtain up-to-date problem sheets and nomenclatural changes associated with the plots to be remeasured.
- Take preprinted quad plot sheet of previous remeasurements into field and enter new data on this sheet.
- Also take a photocopy of the last remeasurement for additional information, but DO NOT write new data on this.
- Take any relevant problem sheets into the field.
- Tell field teams that only the last four numbers of a tag can be entered into the digital file at present, but to record the whole tag number.

- Tell field teams that zeros and blanks are not consistently distinguished from each other in the digital data.
- Ensure plot sheets are legible and understandable in future to those who are not familiar with the plot.
- Measure altitude, aspect, slope, mean top height and ground cover percentages at each remeasurement, as these site variables can change over time.
- Write "DEAD" on plot sheet where a tree has died, DO NOT remeasure and put an asterix by measurement (as suggested in manual) as the asterix may not be seen if the plot sheet gets dirty.
- Write the whole tag number.
- Tag trees systematically from subplot to subplot.
- Where possible, do not tag new individuals with random tag numbers.
- Do not retag new individuals with tag numbers from dead trees.
- (If possible) tag new individuals with a tag number series that has not previously been used on the plot.
- Replace lost/broken tags with same tag number.
- Species code corrections written on the preprinted plot sheet must be accompanied by a short explanatory note in the space provided.
- Once the plot has been remeasured, check to ensure that all categories have been completed.

7. Acknowledgements

Thanks to Rob Allen, Peter Bellingham, Michelle Breach, Larry Burrows, Ian Payton and Susan Wiser for discussions on data quality checking. Peter Bellingham and Susan Wiser had close involvement in the development of the guidelines outlined in this report and, along with Larry Burrows and Michelle Breach, made many helpful comments on an earlier draft of this manuscript.

8. References

Allen, R.B. 1992: RECCE. An inventory method for describing New Zealand vegetation. *Forest Research Institute Bulletin 176*. New Zealand. 25 p.

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

Bellingham, P.B.; Wiser, S.K.; Coomes, D.A. 2000: Permanent forest plots as reference sites for long-term monitoring. *Science for Conservation 151*. Department of Conservation, Wellington.

Chambers, T.C.; Farrant, P.A. 1998: The *Blechnum procerum* ('capense') (Blechnaceae) complex in New Zealand. *New Zealand Journal of Botany 36*: 1–19.

Hall, G.H.J. 1994: PC-DIAM: stem diameter data analysis. Landcare Research, New Zealand. 88 p.

SAS Institute, 1998: SAS Version 6.12. SAS Institute Inc., Cary, North Carolina, USA.

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

Wiser, S.K.; Bellingham, P.J.; Coomes, D.A.; Burrows, L.E.; Gordon, R.F.S. 1999: An assessment of the quality of data stored in the National Vegetation Survey Database, with recommendations for minimising errors. Landcare Research Contract Report LC9899/139 (unpublished), 32 p.

9. Appendices

Appendix 9.1 Colour photocopy of a messy plot sheet. Note the incomplete tag numbers and a multitude of corrections on plot sheet. Appendix 9.2 represents a black and white photocopy of the same plot sheet.

20-1-80

Fs. B80

STEM DIAMETER RECORD SHEET

PROTECTION FORESTRY DIVISION
FOREST RESEARCH INSTITUTE
P.O. BOX 31-011
CHRISTCHURCH

SURVEY: Taramakau 1979
AREA: Mid Taramakau
LINE/PLOT NO.: Plot 11
MEASURED BY: Andy.
RECORDED BY: Brian.

QUARTER.	SPECIES	TAG NO.	CODE	DIAMETER	QUARTER.	SPECIES	TAG NO.	CODE	DIAMETER
A	Gri Lit	K7153		8.6	E	Lib bid	75		25.0
	" "	7154		34.5			77		18.5
	" "	7156		21.5 ✓		Pod hal	L1261		4.5
	" "	7155		10.3		Cop Pse	78		7.7
	Pod hal	58		4.3	<>	~~Phy Alp~~		~~(1)~~	
	Phy Alp	57		19.0		Cop Pse	76		3.9 D
	Cop Pse	60		3.8	F	Pod hal	81		12.2
	" "	59	Dead	4.4*		Gri Lit	82		25.6
	Cop Pse	96		3.7		Myr Div	84		25.9
	" "	L12 64		3.5		Cop Pse	83		*
<>	~~Cop Pse (4)~~					Pse lin	79		5.4 *
	~~Myr Div (1)~~								
					G	Gri Lit	85		19.4
B	Cop Pse	7164		6.8 T	Pod hal.	~~Met~~ Umb	86		62.1
	" "	7165		6.1		Myr div	87		6.5
<>	~~Pod hal (1)~~					" "	63		7.6
	~~Cop Cil (1)~~				<>	~~Pod hal (1)~~			
	~~Cop Pse (1)~~								
					H	Pod hal	L1257		4.0
C	Myr Div	K219		2.0		" "	L1305		3.3
		K7166		12.5*		Met Umb	62	retagged as 26199	25.0
	Gri Lit	67		14.0		Myr Div	61		5.4
	Myr Div	68		13.4					
<>	~~" " (1)~~				I	Cop Pse	89		7.4
	~~Cop Pse (1)~~					" "	88		8.3 D
	~~Pse Col (1)~~					" "	26198	new tag	5.6
						Gri Lit	90		9.2
D	Phy Alp	71	Antler thread	25.3		" "	91		17.1
	Gri Lit	72		27.8		Phy Alp	92		33.3
	Pod hal	70A		6.8					
	Gri Lit	69		54.0	J	Phy Alp	7216		29.1
	Phy Alp	26200	new tag	16.8		Pod hal	7217		29.0
	Gri Lit	73		12.3		Gri Lit	26197	new tag	12.7
	Lib bid	74		28.3					
	Met Umb	L1263		2.9					
	Myr Div / GRI lit	L1260 / 70		2.9 / 18.2 ✓		GRI lit			

QUARTER	SPECIES	TAG NO.	CODE	DIAMETER	QUARTER	SPECIES	TAG NO.	CODE	DIAMETER
K	Met Umb	7218	New tag 26196	27.0					
	Gri Lit	15		17.6					
	Pse Sim	14		24.1					
	Gri Lit	26195	new tag	7.1					
	Myr Div	01		9.5					
	Gri Lit	13		23.7					
	Pod hal	12		13.8					
	Met Umb	7180		9.8					
	⟷								
<>	Met Umb	(1)							
L	Gri Lit	19		12.7					
	" "	11		19.6					
	Myr Div	09		10.3					
<>	Phy Alp	(1)							
	Pse Col	(2)							
	Pod hal	(1)							
M	Gri Lit	07		23.5					
	" "	08		7.7					
	Gri Lit	06		14.7					
	" "	05		12.8					
	" "	04		40.3					
	Cop Pse	10		6.7					
N	Pod hal	02		12.4					
	Gri Lit	03		38.0					
	Myr Div	26193	new tags	2.9					
	" "	26194		3.0					
<>	Myr Div	(2)							
O	Cop Cil	7200		5.2 *					
	Myr Div	7199		5.8					
	Cop Cil	98	new tag 26192	8.6					
	Cop Cil	97		9.8 *					
<>	Myr Div	(2)							
	Gri Lit	96		23.1					
	" "	94	new tag 26191	43.7					
	Myr Div	93		6.2					
		95	not found						
E		[illegible]							

20312A—5000/10/78 G

Appendix 9.2 A black-and-white photocopy of the plot sheet in Appendix 9.1. Note that the information written in coloured ink in Appendix 9.1 does not stand out so clearly on the photocopy.

20-1-80

Fs. B80

STEM DIAMETER RECORD SHEET

PROTECTION FORESTRY DIVISION
FOREST RESEARCH INSTITUTE
P.O. BOX 31-011
CHRISTCHURCH

SURVEY: Taramakau 1979
AREA: Mid Taramakau
LINE/PLOT NO.: Plot 11
MEASURED BY: Andy.
RECORDED BY: Brian.

QUARTER.	SPECIES	TAG NO.	CODE	DIAMETER	QUARTER.	SPECIES	TAG NO.	CODE	DIAMETER
A	Gri Lit	K7153		8.6	E	Lib bid	75		25.0
	" "	7154		34.5			77		18.5
	" "	7156		21.5✓		Pod hal	L1261		4.5
	" "	7155		10.3		Cop Pse	78		7.7
	Pod hal	58		4.3	< >	[redacted]			-
	Phy Alp	57		19.0		Cop pse	76		3.9 D
	Cop Pse	60		3.8	F	Pod hal	81		12.2
	" "	59	Dead	4.4*		Gri Lit	82		25.6
	Cop Pse	96		3.7		Myr Div	84		25.9
	" "	L1264		3.5		Cop Pse	83		*
< >	[redacted]					Pse tin	79		5.4 *
	[redacted]								
					G	Gri Lit	85		19.4
B	Cop Pse	7164	38	6.8	Pod hal.	Met Umb	86		32.1
	" "	7165		6.1		Myr Div	87		6.5
< >	[redacted]					" "	63		7.6
	[redacted]				< >	[redacted]			
	[redacted]								
					H	Pod hal	L1257		4.0
C	Myr Div	K219		2.0		" "	L1305		3.3
		K7166		12.5*		Met Umb	62	Re-tagged as 26199	25.0
	Gri Lit	67		14.0		Myr Div	61		5.4
	Myr Div	68		13.4					
< >	[redacted]				I	Cop Pse	89		7.4
	[redacted]					" "	88		8.3 D
	[redacted]					" "	26198	new tag	5.6
						Gri Lit	90		9.2
D	Phy Alp	71	Antler thrash	25.3		" "	91		17.1
	Gri Lit	72		27.8		Phy Alp	92		33.3
	Pod hal	70A		6.8					
	Gri Lit	69		54.0	J	Phy Alp	7216		29.1
	Phy Alp	26200	new tag	16.8		Pod hal	7217		29.0
	Gri Lit	73		12.3		Gri Lit	26197	New tag	12.7
	Lib bid	74		28.3					
	Met Umb	L1263		2.9					
	Myr Div / SRI lit	L1262 / 70		2.9 / 18.2✓		SRI lit			

Appendix 9.3 An example of the Correction Sheet showing errors identified and corrected. Note that each error is annotated with the name of the person who identified the error and the date that the correction was made.

National Vegetation Survey (NVS) Database

Manaaki Whenua
Landcare Research

Correction sheet

Record of corrections, amendments or comments to data for the following plot:

Survey	North Fiordland
Area	
Line/plot	Opp 8
Sampling date	1970 1976 and 1985
Plot type	Qd diam

Corrections made (state problem, reason for change, surname of annotator and date).

	Surname	Date
Tree # s 6372, 6373, 6374, 6375, 6378, 6434, 6435, 6436 and 6450 were identified as ARISER in the 1970 survey. This identification was corrected to HOHGLA in the 1985 survey. The species code has been changed in the 1970 digital file to be consistent with this correction.	Baldwin	11 July 2000

Additional notes

Send corrections to:
Michelle Breach
National Vegetation Survey (NVS) database
Landcare research Lincoln
PO Box 69 Lincoln

(P.T.O)

Appendix 9.4 An example of the Problem Sheet showing queries that can not be resolved until the next remeasurement.

National Vegetation Survey (NVS) Database

Manaaki Whenua
Landcare Research

Problems to resolve at next remeasurement by field teams

Take a photocopy of this sheet and the last remeasurement into the field to reconcile problems for the following plot.

Survey	North Fiordland
Area	
Line/plot	OPP 8
Sampling date	1970 and 1975
Plot type	Qd diam

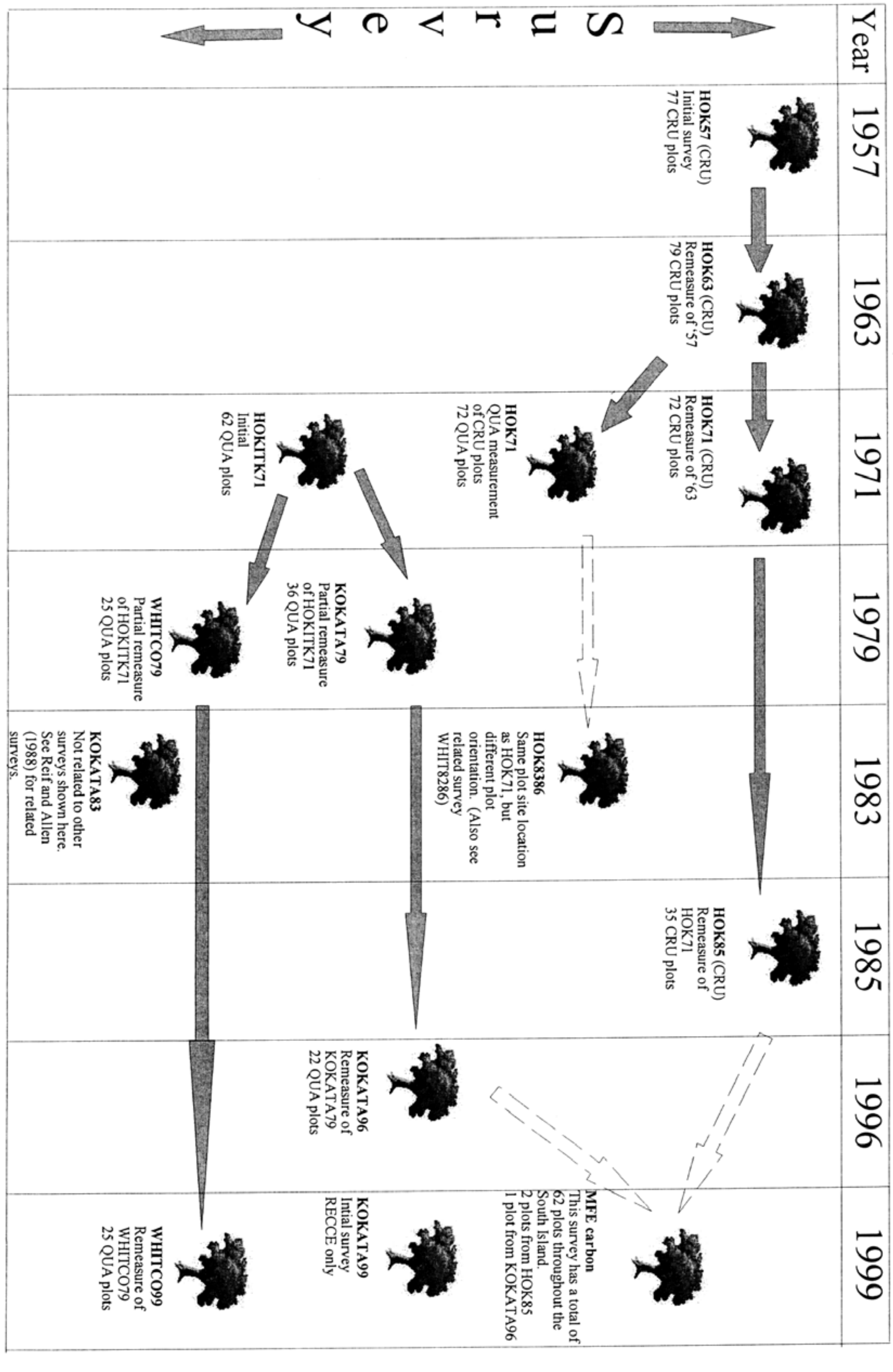
Date and surname of annotator	Description of problem and suggested solution	Description of solution	Resolution date and surname of resolver
Baldwin 7 July 2000	Tree # 6450 was identified as CARSER in 1970 and as HOHGLA in 1975. →Confirm identity at next remeasurement		

Send corrections to:
Michelle Breach
National Vegetation Survey (NVS) database
Landcare research Lincoln
PO Box 69 Lincoln

(P.T.O)

Appendix 9.5 Diagram showing the relationship between surveys in the Hokitika River valley. This tracks the lineage of permanent plots in two major surveys (HOK) and (HOKITK) that were established in the same geographic region and shows the relationship with subsequent surveys in the same area. Also see Appendix 9.6 for the lineage.

Diagram showing relationships between surveys in the Hokitika area



Appendix 9.6 Lineage of the HOKITK survey, established in 1971. The lineage summarises plots per remeasurement year by sampling method, lists the digital file names associated with each measurement, and itemises for each plot the years sampled and the sampling methods used. Here we only show the lineage information for plots. See Appendix 9.5 to see how the measurements relate to another.

HOKITIKA
J:\NVS\WESTCOA0\RECCE\HOKITK etc.
K*= plots filed and stored under the survey name KOKATA
W*= plots filed and stored under the survey name WHITCO
MFE = plots used in the MFE carbon transect: J:\NVS\Mfe Carbon transect\1998 99\NVS formatted\Mfedia99.qua etc.

Summary created March 2000

Surveys used in this lineage: NO. OF PLOTS BY SAMPLING METHOD

	QUA	REC	UND
HOKITK71	62	560	
KOKATA79	36	36	36
WHITCO79	25	25	25
KOKATA96	22	22	22
WHITCO99	25	25	25
Mfe	1	1	1

There are other surveys, not mentioned here, that have been carried out in the Hokitika area. Information on these surveys can be found in text files in the folder: J\NVS\WESTCOA0\TXT

Line and plot	1971	1979	1996	1999
890	REC			
9 1	QUA	QUA REC UND W*		QUA REC UND W*
9 2	QUA REC	QUA REC UND W*		QUA REC UND W*
9 3	QUA	QUA REC UND W*		QUA REC UND W*
910	REC			
920	REC			
980	REC			
10 1	QUA	QUA REC UND W*		QUA REC UND W*
10 2	QUA REC	QUA REC UND W*		QUA REC UND W*
10 3	QUA	QUA REC UND W*		QUA REC UND W*
1010	REC			
6620	REC			
67 1	QUA	QUA REC UND K*	QUA UND REC K*	
67 2	QUA	QUA REC UND K*	QUA UND REC K*	
67 3	QUA	QUA REC UND K*	QUA UND REC K*	
68 1	QUA	QUA REC UND K*	QUA UND REC K*	
68 2	QUA REC	QUA REC UND K*	QUA UND REC K*	
6810	REC			
6820	REC			
6830	REC			
6840	REC			
69 1	QUA	QUA REC UND K*	QUA UND REC K*	
69 2	QUA REC	QUA REC UND K*	QUA UND REC K*	
69 3	QUA	QUA REC UND K*	QUA UND REC K*	
69 4	QUA	QUA REC UND K*	QUA UND REC K*	
6910	REC			
71110	REC			
72 1	QUA	QUA REC UND K*		QUA REC UND (MFE plot C8)
72 2	QUA REC	QUA REC UND K*		
72 3	QUA	QUA REC UND K*		
72 4	QUA	QUA REC UND K*		
72 5	QUA	QUA REC UND K*		
72 6	QUA	QUA REC UND K*		

Appendix 9.7 This plot sheet illustrates several different problems. The photocopy is of low quality and is an example where new data have been written on a photocopy of the last remeasurement. The information is confusing, with three years of data written on the plot sheet, (one year photocopied and two additional years data written on that). The date has only been noted for one of the two additional remeasurements. Also note that the names of the field team have not been written in.

STEM DIAMETER RECORD SHEET

FOREST & RANGE EXPERIMENT STATION
P. O. BOX 106
RANGIORA

SURVEY: Taramakau
AREA: Mid Taramakau
LINE/PLOT NO.: Plot 8
MEASURED BY:
RECORDED BY:

DYING (ROT SET IN)

QUARTER		SPECIES	TAG NO.	DIAMETER ~~CODE~~ 1969	DIAMETER
C	✓	OLE ave	01	13.0	
"	✓	" "	02	9.4	8.3
"	✓	" "	03	14.5	
"	✓	" "	04	10.7	
"	✓	" "	05	10.9	
"	✓	" "	06	5.6	
"	✓	DRA tra	07	25.1	
"	✓	OLE ave	08	12.2	
"	✓	" "	09	6.4	
"	✓	" "	10	10.7	
"	✓	" "	11	8.6	
"	✓	" "	12	8.1	
"	✓	" "	13	7.1	7.9
"	✓	" "	14	9.7	
"	✓	" "	15	6.6	
"	✓	" "	16	6.4	
C	✓	" "	17	13.0	
D	✓	MYR div	18	3.6	
"	✓	GRI lit	19	6.1	
"	✓	" "	20	3.6	
"	✓	" "	21	6.9	
"	✓	PHY alp	22	15.7	
"	✓	GRI lit	23	35.8	
"	✓	OLE ave	24	17.8	
"	✓	GRI lit	25	12.2	
"	✓	PHY alp	26	19.8	
"	✓	GRI lit	27	6.4	
"	✓	OLE ave	28	11.9	
"	✓	MYR div	29	3.0	
D	✓	GRI lit	30	11.7	
B	✓	DRA tra	31	13.0	
B		OLE ave	32	15.7	15.3
D	✓	" "	33	11.7	
D	✓	" "	34	11.4	
B	✓	POD hal	35	32.5	
"	✓	" "	36	27.9	
"	✓	OLE ave	37	14.5	
"		" "	38	23.1	
"	✓	MYR div	39	4.8	
"	✓	DRA tra	40	17.3	
"	✓	GRI lit	41	20.1	
"	✓	COP pse	42	5.6	
"	✓	ARC tra	43	18.3	
"	✓	GRI lit	44	28.4	
"	✓	" "	45	12.2	
B	✓	" "	46	28.2	

QUARTER		SPECIES	TAG NO.	DIAMETER ~~CODE~~ 1969	DIAMETER
B	✓	GRI lit	47	11.9	
B	✓	PHY alp	48	11.7	
A		Dead	49		
"	✓	GRI lit	50	10.4	
"	✓	" "	51	6.9	
"	✓	" "	52	5.1	
"	✓	OLE ave	53	8.1	
"	✓	" "	54	6.1	
"	✓	" "	55	6.4	
"	✓	" "	56	6.4	
"	✓	" "	57	7.9	
"	✓	~~CARM. spp.~~	58	10.4	
A	✓	COP cil ~~par~~	59	7.6	
C	✓	MYR div	60	6.9	
A	✓	MYR div	61	3.8	
D	✓	MYR div	1314		3.2
		OLE ave			

Appendix 9.8 An example of a reprinted plot sheet that has been written neatly and filled out properly. The species codes of two individuals have been deliberately changed on this plot and these can be back-corrected in previous measurements.

SURVEY : CAPLES/GREENSTONE 1989 SURVEY-PAGE(27) PLOT-PAGE(
MEASURED BY : P.W.B. RECORDED BY : M.H. DATE : 14-11-89.
Altitude: 825metres, Aspect: 35deg. Slope:15deg. Physiography:Face

PLOT	S	SPECIES	TAG		76	82	Diameter	Plot-notes
8 2	I	NOTCLI	894		31.9	32.0	-.-	dead
.	~~I~~ H	.	895		15.4	16.5	18.4	
.	A	.	896		13.5	13.8	13.8	
.	A	NOTMEN	897		4.8	4.8	-.-	dead
.	A	NOTCLI	898		6.7	6.8	7.1	
.	A	.	899		6.8	7.0	7.4	
.	B	NOTFUS	900		5.1	5.7	6.1	
.	B	NOT~~MEN~~ CLI	901		3.2	3.7	3.4	correct ID NOT cli
.	B	NOTCLI	902		8.4	8.9	9.1	
.	B	.	903		19.5	20.4	21.5	
.	B	.	904		35.0	36.7	37.7	
.	B	.	905		25.4	26.3	28.1	
.	C	.	906		25.8	27.2	27.3	
.	D	NOT~~MEN~~ CLI	907		32.0	33.2	33.6	correct ID NOT cli
.	[illegible]	NOT MEN	908		13.5	13.8	13.2	NOT men
.	[illegible]	PHYALP	909		11.5	11.7	11.9	
.	D	NOTCLI	910		19.1	19.2	19.4	
.	E	.	911		9.2	9.2	9.3	
.	E	.	912		19.4	21.0	21.7	
.	F	.	913		42.0	42.7	44.0	
.	G	.	914		10.3	9.6	9.6	
.	G	.	915		29.0	29.7	29.5	
.	G	NOTMEN	916		5.9	6.5	1.0	
.	H	NOTCLI	917		6.5	6.9	7.5	
.	C	NOTMEN	918		5.9	5.9	5.9	
.	H	NOTCLI	919		5.5	5.6	5.7	
.	H	.	920		33.0	34.0	34.7	
.	H	.	921		3.8	3.9	-.-	clead
.	H	.	922		3.3	3.3	-.-	dead
.	H	.	923		3.0	3.9	4.4	
.	H	NOTMEN	924		3.8	4.2	4.4	
.	H	.	925		3.2	3.8	3.9	
.	H	.	926		7.2	7.5	7.4	
.	I	.	927		6.0	6.1	5.6	
.	I	NOT.cli	928		9.1	9.4	9.7	correct ID NOT cli
.	J	NOT.men	929		3.0	3.0	2.7	NOT men
.	J	.	930		8.0	8.4	9.1	
.	J	.	931		13.5	14.5	16.1	
.	H	NOTCLI	933		19.5	19.7	18.9	
.	J	NOTMEN	934		3.9	4.5	4.6	
.	K	.	935		2.8	3.1	2.8	
.	H	NOTCLI	936		5.5	6.5	6.6	
.	J	.	937		24.7	25.0	25.8	
.	J	.	938		17.1	18.3	20.3	
.	J	.	939		28.2	28.8	30.7	
.	K	.	940		10.1	10.5	11.0	
.	K	.	941		12.1	12.3	12.5	
.	M	.	942		27.7	28.6	28.7	
.	G	NOTCLI	K739				4.2	
.	-	------	----				-.-	
.	-	------	----				-.-	
.	-	------	----				-.-	
.	-	------	----				-.-	
.	-	------	----				-.-	