DEPARTMENT OF CONSERVATION TECHNICAL SERIES No. 6

DATA STANDARDS FOR THE DEPARTMENT OF CONSERVATION

by

Information Services Unit
Science and Research Division
Department of Conservation

PUBLISHED BY DEPARTMENT OF CONSERVATION, WELLINGTON, NEW ZEALAND

ISSN 1172-6873
ISBN 0-478-01631-X

Cataloguing-in-Publication data

New Zealand. Dept. of Conservation. Information Services Unit.
Data standards for the Department of Conservation / by Information Services Unit. Wellington, N.Z. : Dept. of conservation, 1994.
1 v. ; 30 cm. (Department of Conservation technical series, 1172-6873 ; no. 6.)
ISBN 047801631X
1. Information retrieval. 2. Data base management. 3. Data structures (Computer science) I. Series: Department of Conservation technical series ; no. 6.
zbn94-076772

Keywords: information systems, definitions, computers, conservation, databases

CONTENTS

DATA STANDARDS
for the
DEPARTMENT OF CONSERVATION

by
Information Services Unit
Science and Research Division, Department of Conservation,
PO Box 10-420, Tory Street, Wellington.

ABSTRACT

The team responsible for developing conservation information systems has developed a set of standards for the Department. These standards are presented with accepted definitions of how individual data items are to be input, output and/or stored within a database. The data standards are available for use by any individual or organisation with responsibility for conservation information management.

1. INTRODUCTION

1.1 Aims

The aims of this publication are to:

- Document data standards for the Department.
- Encourage discussion and debate on the data standards and thereby facilitate feedback to the systems development team.
- Encourage adherence to the data standards by all groups and individuals involved in application and database development for the Department of Conservation.

1.2 Purpose

To facilitate the consistent handling of conservation data within DOC and, wherever possible, encourage the same in external information systems. This will make it possible for:

- Information systems within the Department to share data and, therefore, move closer towards integration.
- Easier sharing of data with external agencies.

All databases that are developed by the Information Services Unit from now on will adhere to these standards. We encourage all other database developers within the Department to do the same.

1.3 Document Design

This document is designed to allow it to evolve or change to meet new user needs as these are identified. The "list" is not complete and new standards will be added as they are developed. The publication has been produced in a ring binder format to facilitate the update process. Each standard is contained on a separate sheet and will simply lift out and be replaced if updated or added to the back if new. All recipients within the Department of Conservation will be on our mailing list for the updates.

1.4 Field Names

Please note that the field names appearing in these standards can be altered to suit particular circumstances and do not need to be strictly adhered to. There will be times when a particular field such as PHONE _NO will need to be recorded more than once within a record, e.g., when we want to record both a work and home telephone number. Some of the suggested names might be restricted or reserved words in the database managment system you are using, e.g., Date or Time.

In these instances we suggest using a suitable prefix to ensure the name is not duplicated within your system or is not a reserved/restricted word.

1.5 Feedback please!

During the development of the standards there has been little opportunity to consult and we would really appreciate hearing your comments. These standards are not "set in concrete" and will be changed if there are good reasons to do so. Please direct your comments to:

DOC Data Standards Group
Information Services Unit
Science and Research Division
Department of Conservation
PO Box 10-420
WELLINGTON

2. DEFINITION OF TERMS

CHARO Used to describe the format for a data item as a character field. A number inside brackets indicates the length of the field.

Coordinates Two measurements used for fixing a position on the ground with reference to a map projection or grid.

Data Basic elements of information which can be stored, processed, retrieved, or produced by a computer.

Database An organised collection of information or data.

DATE Used to describe the format for a data item as a date field.

Datum Known position which is the basis of a survey system.

dBASE A database management software package for personal computers, released by Borland International.

Field Name Suggested name for the data item being described.

Format Used to describe some general system rules about how a data item is input, output or stored, e.g., CHAR - Character field: NUM -Numeric field, DATE - System date field, etc.

Geodetic Datum 1949 This is the reference datum covering New Zealand. A number of map projections based on geodetic datum 1949 have been used for mapping purposes including: transverse mercator projection for imperial mapping grid; New Zealand map grid for metric mapping; and the standard mercator projection for marine charts.

NUMO Describes the format for a data item as a numeric field. A number inside brackets indicates the length of the field.

Picture Used to provide more graphical information about the format of a data item. "X" - any valid character, "A" -alphanumeric character, "C" - alphabetic character only, "9" - digit, "." a decimal point, "Z" - a digit or blank, "DD" "MM" "YY" "HH" "MM" "SS" - digits to represent day month year hours minutes and seconds.

Resolution Measure of the precision of numeric data.

3. DATA STANDARDS

1. DATE

1.1 Introduction

Where there is a need to store a precise date then a standard date format of DD/MM/YYYY is sufficient. Database systems like dBASE allow dates to be stored in this fashion in special date fields (system dates). However, where partial dates are to be stored, a character field needs to be used in conjunction with a system date field to indicate the date's accuracy. This dual standard is outlined below.

1.2 Format

Database

Field Name	Format	Picture	Comments
DATE	CHAR(8)	DDMMYYYY	Text field for storing date as input
SYS DATE	DATE		Input converted to system date for date manipulation and searches.
DATE AC	NUM(5)		Accuracy indicator for date, i.e., ± number of days

Input/Output

Field Name	Format	Picture	Comments
DAY	CHAR(2)	DD	Day of month
MONTH	CHAR(2)	MM	Month of the year
YEAR	CHAR(4)	YYYY	Year

1.3 General Discussion

The standard has been designed to allow the recording of imprecise, precise and ranges of dates. Two date fields will be stored:

- text field in the format DDMMYYYY and
- system date field.

The text field will contain the actual data as input, and the system date which has to be a complete and valid date will be used to do date manipulation and searches. There is also a date accuracy indicator field which in most applications will record the accuracy of the system date field as a "t number of days".

Although stored as a single text field, to reduce any possible confusion from users on the required format, the date will be split into 3 separate fields for input, i.e., Day (2 characters), Month (2 characters) and Year (4 characters). The fields will be separated by a "/" and only valid numerical values will be accepted, e.g., 01 or 1 to 12 for Month.

2. TIME

2.1 Introduction

Time is treated similarly to date. A text field is to be provided for the storage of imprecise data while all time data is also stored in a system time field if such is available in the system. An accuracy indicator gives a range of accuracy for imprecise time and an am/pm field permits use of a 12 hour clock.

2.2 Format

Database

Field Name	Format	Picture	Comments
TIME	CHAR(6)	HHMMSS	Text field for storing input time
SYS TIME	DATE		Input converted to system time
TIME AC	NUM(3)		Accuracy indicator -:t number of minutes

Input/Output

Field Name	Format	Picture	Comments
IOHRS	CHAR(2)	HH	Hours part of time
IOMINS	CHAR(2)	MM	Minutes (part of time) separated from hours by screen colon
IOSECS	CHAR(2)	SS	Seconds (part of time) separated from minutes by screen colon
IOHRSAC	CHAR(2)	HH	Accuracy of time indicated by * number of hours
IOMINSAC	CHAR(2)	MM	Accuracy of time indicated by t number of minutes
AMPM	CHAR(2)	CC	AM / PM indicator

2.3 General Discussion

Time is to be entered in a screen picture of the form HH:MM:SS. This will be stored as a text string (without the colons) and will also be converted and stored as a system time.

Where an imprecise time is permitted accuracy is to be entered in a screen picture of the form ± HH:MM. This will be converted and stored as a number of minutes.

A two-character am/pm indicator is to be included in all cases where time is not explicitly entered on a 24 hour clock basis.

3. NAME

3.1 Introduction

The name comprises a title (Dr, Ms, Mr ...), either first name or initials and the surname. Surname could also be used to denote businesses, e.g., Department of Conservation.

3.2 Format

Database[InputlOutput

Field Name	Format	Picture	Comments
TITLE	CHAR(8)		Person's title.
NAAME	CHAR(15)		A persons first name(s) or initials.
SURNAME	CHAR(40)		A persons surname or a company/business name.

3.3 General Discussion

Using mixed case in names is preferable, creating a more readable output when viewing or printing. Initials should be uppercase and do not use punctuation.

4. ADDRESS

4.1 Introduction

Address has been separated into the postal address for PO Box numbers and street address as a location indicator, e.g., 5th Floor, Macarthy Trust Building, 58 Tory Street, WELLINGTON.

4.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
ATTENTION	CHAR(32)		Name of contact person within an organisation
POST1	CHAR(30)		Postal address line 1 .
POST2	CHAR(30)		Postal address line 2.
POST3	CHAR(30)		Postal address line 3
POST4	CHAR(30)		Postal address line 4
POST5	CHAR(30)		Postal address line 5
ST1	CHAR(30)		Street address line 1
ST2	CHAR(30)		Street address line 2
ST3	CHAR(30)		Street address line 3
ST4	CHAR(30)		Street address line 4
ST5	CHAR(30)		Street address line 5

4.3 General Discussion

New Zealand Post standards recommend that the only item of an address that is capitalised should be the post centre, e.g., WELLINGTON. All other fields should be lowercase. Abbreviations and punctuation should be avoided to facilitate searching (with the exception of the standard abreviation for Post Office Box, i.e., PO Box).

5. PHONE and FAX

5.1 Introduction

Below is the recommended standard for storing international and local telephone and facsimile numbers.

5.2 Format

Database/Input/output

Field Name	Format	Picture	Comments
PH ISTD	CHAR(8)		Country/Area/Access Codes
PHONE NO	CHAR (10)		Telephone number
FAX ISTD	CHAR(8)		Country/Area/Access Codes
FAX_NO	CHAR(10)		Fax number

5.3 General Discussion

The telephone and facsimile numbers can be broken down in four parts:

- Access code - which is specific to the country or city you are calling from. (For NZ the access code for international calls is "00")
- Country code - identifies the "country" being called. 1 - 4 digits in length.
- Area code - subdivision of "country". 0 - 4 digits in length.
- Telephone or fax number - The maximum length is currently 7 digits but this is to be increased to 10 digits on 1 January 1996.

We have simplified the above by combining Access, Country and Area codes into one field to make life easier for our users. It is not always easy to identify where a country code ends and area code begins.

The actual phone or fax portion of the number has been set to a maximum of 10 digits to allow for the new international standard which will be effective from 1 January 1996.

6. METRIC GRID - Geographic Reference

6.1 Introduction

This standard refers to storage of metric grid for the New Zealand Map Grid (NZMG). It allows storage to a resolution of 1 metre (7 figure grid reference) and includes a resolution indicator for less precise grid references.

6.2 Format

Database

Field Name	Format	Picture	Comments
METMAP	CHAR(3)	C99	Metric map number
METE	NUM(7)	9999999	Metric Easting
METN	NUM(7)	9999999	Metric Northing
METRES	NUM(2)		Resolution indicator for the metric grid reference. Derived from input.

Input/Output

Field Name	Format	Picture	Comments
METMAP	CHAR(3)	C99	Metric map number
METE	CHAR(7)	999ZZZZ	Metric Easting - Input as character field to allow recording of less than 7 digits. "9" = compulsory digits and "Z" = a digit or blank.
METN	CHAR(7)	999ZZZZ	Metric Northing - Input as character field to allow recording of less than 7 digits. "9" = compulsory digits and "Z" = a digit or blank.

6.3 General Discussion

The standard has been set to allow the recording of up to a 7 digit easting and northing i. e., a resolution of 1 metre. Input and output will be via character fields as this will allow users to input less than seven digits. Character fields will accept trailing blanks where a numeric field will zero fill these spaces.

The easting and northing will be stored within the database as a numeric field so valid mathematical functions can be performed. A resolution indicator will also be stored to record the number of digits actually input.

For example, a metric grid reference entered as R27 577 E 918 N will be stored as R27 2657700 5991800 with a resolution = 5 indicating a resolution to the nearest 100 metres.

7. IMPERIAL GRID - Geographic Reference

7.1 Introduction

This standard refers to storage of imperial grid references for the New Zealand Imperial Grid on the NZMS 1 map series. It allows storage to a resolution of 1 yard (6 figure grid reference) and includes a resolution indicator for less precise grid references.

7.2 Format

Database

Field Name	Format	Picture	Comments
ISLE	CHAR(1)	C	Island
IMPMAP	CHAR(3)	ZZZ	Imperial map number
IMPE	NUM(6)	999999	Imperial Easting
IMPN	NUM(6)	999999	Imperial Northing
IMP RES	NUM(2)	99	Resolution indicator for the imperial grid reference. Derived from input.

Input/Output

Field Name	Format	Picture	Comments
ISLE	CHAR(1)	C	Island. e.g."N" North Island, "S" South Island
IMPMAP	CHAR(3)	ZZZ	Imperial map number
IMPE	CHAR(6)	9ZZZZZ	Imperial Easting - Input as character field to allow recording of less than 7 digits. "9" = compulsory digits and "Z" = a digit or blank.
IMPN	CHAR(6)	9ZZZZZ	Imperial Northing - Input as character field to allow recording of less than 7 digits. "9" = compulsory digits and "Z" = a digit or blank.

7.3 General Discussion

The standard has been set to allow the recording of up to a 6 digit easting and northing, i.e., a resolution of 1 yard. Input and output will be via character fields as this will allow users to input less than six digits. Character fields will accept trailing blanks where a numeric field will zero fill these spaces.

The easting and northing will be stored within the database as a numeric field so valid mathematical functions can be performed. A resolution indicator will also be stored to record the number of digits actually input.

For example, an imperial grid reference entered as N164 2324 E 1241 will be stored as N164 232400 124100 with a resolution = 4 indicating a resolution to the nearest 100 yards.

8. LATITUDE - Geographic Reference

8.1 Introduction

This standard refers to the storage of latitude for coordinate references using latitude and longitude. Latitude may be stored to a resolution of 0.01 seconds, equivalent to about 0.2 metres at Wellington. The standard refers to the geodetic datum 1949. A resolution indicator records precision of the original entry.

8.2 Format

Database

Field Name	Format	Picture	Comments
LATITUDE	-NUM(7)	-99.99999	Latitude in decimalised degrees, stored as a signed number field. The sign indicates direction from the equator., i.e., "+" = North and"-" = South.
LATRES	NUM(2)		Resolution indicator for latitude. Records the number of digits input.

Input/Output

Field Name	Format	Picture	Comments
LATD	CHAR(2)	9Z	Latitude degrees - Input as a character field to allow blanks. "9" = compulsory digits and "Z" = a digit or blank allowed.
LATM	CHAR(2)	ZZ	Latitude minutes (see above)
TATS	CHAR(5)	ZZ.ZZ	Latitude decimal seconds (see above)
LATDIR	CHAR(t)	C	Direction of latitude from the equator. i.e. "N" = North, "S" = South

8.3 General Discussion

The standard has been set to allow the recording of latitude to a resolution of 0.01 seconds, equivalent to about 0.2 metres at Wellington. Input and output will be via character fields as this will allow users to input only the required digits. Character fields will accept trailing blanks where a numeric field will zero fill these spaces.

Latitude will be stored within the database as decimalised degrees within a signed numeric field so valid mathematical functions can be performed. A resolution indicator will also be stored to record the number of digits actually input. For example, a latitude entered as 41 016'07"S will be stored as -41.26860 with a resolution = 6 indicating a resolution to the nearest second.

NOTE: It is particularly important that users of global positioning systems convert any coordinate references to the geodetic datum 1949 before storage. Significant differences can occur if a different reference datum is used.

9. LONGITUDE - Geographic Reference

9.1 Introduction

This standard refers to the storage of longitude for coordinate references using latitude and longitude. Coordinates may be stored to a resolution of 0.01 seconds, equivalent to about 0.2 metres in Wellington. The standard refers to the geodetic datum 1949. A resolution indicator records precision of the original entry.

9.2 Format

Database

Field Name	Format	Picture	Comments
LONGITUDE	-NUM(8)	-999.99999	Longitude in decimalised degrees, stored as a signed number field. The sign indicates direction from the prime meridian, i.e., East and -~ = West.
LONGRES	NUM(2)		Resolution indicator for longitude. Records the number of digits input.

Input/Output

Field Name	Format	Picture	Comments
LONGD	CHAR(3)	9ZZ	Longitude degrees - Input as a character field to allow blanks. "9" = compulsory digits and "Z" = a digit or blank allowed.
LONGM	CHAR(2)	ZZ	Longitude minutes (see above)
LONGS	CHAR(5)	ZZ.ZZ	Longitude decimal seconds (see above)
LONGDIR	CHAR(1)	C	Direction of longitude from the prime meridian, i.e., "E" = East, "W" = West

9.3 General Discussion

The standard has been set to allow the recording of longitude to a resolution of 0.01 seconds, equivalent to about 0.2 metres at Wellington. Input and output will be via character fields as this will allow users to input only the required digits. Character fields will accept trailing blanks where a numeric field will zero fill these spaces.

Longitude will be stored within the database as a signed numeric field so valid mathematical functions can be performed. A resolution indicator will also be stored to record the number of digits actually input. For example, a longitude entered as 174°45'46"E will be stored as +174.76270 with a resolution = 7 indicating a resolution to the nearest second.

NOTE: It is particularly important that users of global positioning systems convert any coordinate references to the geodetic datum 1949 before storage. Significant differences can occur if a different reference datum is used.

10. ALTITUDE - Geographic Reference

10.1 Introduction

This standard refers to the storage of altitude in metres or feet.

10.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
MET ALT	NUM(6)	999999	Altitude recorded in whole metres
METALTRES	NUM(6)	999999	Resolution indicated by t no. of metres
IMP ALT	NUM(6)	999999	Altitude recorded in whole feet
IMPALTRES	NUM(6)	999999	Resolution indicated by t no. of feet

10.3 General Discussion

The standard has been set to allow the recording of altitude to the nearest metre or foot. A resolution indicator records the precision of the original entry. This is of particular importance for converted data, where conversions can give spurious impressions of precision such as 500 feet converting to 158.28 metres.

11. AREA - Geographic Reference

11.1 Introduction

This standard refers to the storage of area in hectares.

11.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
AREA	NUM(12)	99999999.9999	Area given in hectares. The four decimal places allow recording down to 1 m .
AREA-RES	NUM(12)	99999999.9999	Resolution of the area as ± number of hectares.

11.3 General Discussion

Data presented in imperial units will be converted and stored as hectares. Area can be stored to an accuracy of 1 metre .

12. GEOGRAPHIC REFERENCE INDICATOR

12.1 Introduction

This standard refers to an indicator for geographical coordinate data, recording the original basis on which coordinate was originally presented. This is of particular importance for converted data, where for example an imperial grid reference was arithmetically converted to a metric or lat/long coordinate.

12.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
GEOREFIND	CHAR(t)	C	Indicates how the geographical reference was originally recorded, i.e., M - Metric, I - Imperial, L- Latitude & longitude

12.3 General Discussion

The geographic reference indicator is used to determine the original method for recording a particular coordinate reference. This may be needed when editing coordinates or reconciling differences in location that can arise during the conversion process between coordinate systems.

13. SPECIES and TAXONOMY

13.1 Introduction

Species information can be stored in a number of ways. The outline below includes a taxonomic hierarchy based on the most commonly used taxonomic levels. A similar format can be used to include less commonly used groupings, such as infra-orders.

13.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
KINGDOM	CHAR(50)		Kingdom name
PHYLUM	CHAR(50)		Phylum name
SUBPHYLUM	CHAR(50)		Sub-phylum name
CLASS	CHAR(50)		Class name
SUB CLASS	CHAR(50)		Sub-class name
TAXORDER	CHAR(50)		Order name
SUB ORDER	CHAR(50)		Sub-order name
FAMILY	CHAR(50)		Family name
SUBFAMILY	CHAR(50)		Sub-family name
GENUS	CHAR(30)		Genus name
SPECIES	CHAR(40)		Species/subspecies name
SP CODE	CHAR(6)		Unique code for scientific name
COM NAME	CHAR(60)		Common/Maori name for species

13.3 General Discussion

Normally the common name will be in a separate file to the other taxonomic identifiers, linked by species code. This allows multiple common names to be stored for each scientific name with a minimum of storage space. The species code should therefore be unique for each genus-and-species (or genus-and-species/subspecies) combination. Although the genus and species are treated as a single binomial name for outputting, etc., they are in fact separate taxonomic levels, which is why they are stored as such. This also provides much greater flexibility in a relational database environment than if they were stored as a single field ("scientific name").

Common name can also be associated with higher taxonomic levels such as family, e.g., "Hominidae" has a common name "Humans".

14. CONSERVANCY and FIELD CENTRES

14.1 Introduction

There are currently 14 conservancies and 78 associated field centres in New Zealand.

14.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
DOCCODE	CHAR(5)	99.99	Code for conservancy or field centre
DOCDIST	CHAR(30)		Name of the conservancy or field centre

14.3 General Discussion

The standard allows for the entry of conservancies and/or field centres, with the last two digits of the code identifying the type of office. A code with "00" after the decimal point indicates a conservancy, while at present, field centres are identified numerically, starting with ".01 ", with the ability to enter up to 100 codes, i.e., ".99". Only the 5 digit code will be stored in any data file, with an associated table of codes and descriptions from which to select the conservancy or field centre name.

15. ECOLOGICAL REGION / DISTRICT

15.1 Introduction

There are currently 85 Ecological Regions and 268 Ecological Districts in New Zealand (Ecological Regions and Districts of New Zealand, 3rd edition, 1988).

15.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
ECOLCODE	CHAR(5)	99.99	Code for ecological region or district
ECOLDIST	CHAR(30)		Ecological region or district name

15.3 General Discussion

Although the code for ecological district is defined as a character field it actually contains two digits a decimal point seperator and a further two digits. The first two digits indicate the ecological region, and the second two digits indicate the district within that region, i.e., "11.04" denotes the Hamilton ecological district within the Waikato ecological region.

16. COMMENT or DESCRIPTION

16.1 Introduction

The comment or description field is a text field for the descriptive kind of information commonly required in a data record.

16.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
COMMENT	CHAR(255)		Comment or description field.

16.3 General Discussion

This is a text field of 255 characters, the maximum permissible length for a normal character field in many database systems. Guidelines for the use of these fields are mixed case, avoid abbreviations and unnecessary punctuation.

17. SITE / LOCALITY

17.1 Introduction

The name of a site can be a specific reference to a point on a map, or a general description of a location.

17.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
SITE NAME	CHAR(70)		Specific name of site
LOCALITY	CHAR(255)		General location of site

17.3 General Discussion

A site name should refer to a specific point on a map, e.g., Acacia Gully *Pond;* whereas a locality should give the user an indication of the general location of the site, and should reference another site name if possible, e.g., Coastal cliff approximately *1.5 km north of Woody Head near Mount Kariod, Raglan.*

Guidelines for these fields are: mixed case, avoid abbreviations and unnecessary punctuation.

18. LOCAL AUTHORITIES

18.1 Introduction

There are currently 73 local authorities in New Zealand, including district and city boundaries.

18.2 Format

Database/Input/Output

Field Name	Format	Picture	Comments
LAUTCODE	CHAR(2)	CC	Code for local authority
LAUTDIST	CHAR(30)		Name of the local authority

18.3 General Discussion

Each local authority boundary is identified by a 2-character alphabetic code. This includes district and city boundaries. Local authority names should be stored as mixed case with no abbreviations or punctuation.

4. ALPHABETIC LIST OF STANDARDS