TERN NetCDF User Manual

NetCDF Conversions and Reference Tables

Version 2.0

tern
Ecosystem Research Infrastructure

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

Acknowledgment

TERN acknowledges that this document is based on the Integrated Marine Observing System (IMOS) NetCDF User's Manual. It re-uses extensively the NetCDF data format table, definition, structure, and references included in the manual and in the NetCDF User's Guide.

Document History

Version	Date	Authors	Revision History
1.0	Nov 2018	Biswajit Bala, Siddeswara Guru	Draft version of TERN NetCDF user's manual was developed.
2.0	1 November 2021	Anusuriya Devaraju	• Restructured the headings. • Added document history. • Updated the global attributes based on Attribute Convention for Data Discovery 1-3 and excluded redundant attributes. • Updated broken links. • Updated CF variable attributes to include parameter and instrument metadata. • Included data examples from TERN OzFlux NetCDF files.
2.0	16 May 2022	Anusuriya Devaraju	• Updated the requirement level of global attributes (required and recommended) and added the attribute 'statistic_type' in data attribute as requested by TERN EP central team.

Acronyms

Acronym	Definition
ACDD	Attribute Convention for Data Discovery 1-3
CDM	Unidata's Common Data Model (CDM)
CDL	Common Data Language
CF Metadata (conventions)	NetCDF Climate and Forecast (CF) Metadata Conventions
COARDS	Cooperative Ocean/Atmosphere Research Data Service
NetCDF	Network Common Data Form
TDDP	TERN Data Discovery Portal
TDSA	TERN Data Services and Analytics
TERN	Terrestrial Ecosystem Research Network
UDUNITS	Unidata Units

Table of Contents

1 Introduction

This document is intended to provide an overview of NetCDF conventions applied by the TERN data infrastructure. Contributors should provide data in NetCDF file format that conform to the Climate and Forecast (CF) Conventions outlined in this document so that TERN can provide systematic and standardised access and services to the research community. The guidelines for CF 1.8[1] and/or latest and general NetCDF user guidance (NUG)[2] can be used to build this file format. However, for managing and publishing data through TERN data infrastructure, TERN highly encourages the data contributors to represent TERN's metadata within global attributes of the NetCDF file following the Unidata's Attribute Convention for Dataset Discovery (ACDD)[3]. ACDD defines a list of NetCDF global attributes recommended improving the discovery of a NetCDF dataset.

2 Network Common Data Form (NetCDF)

NetCDF is one of many file formats available for storing TERN data. It is a self-describing and portable data file format. It contains information about the data (metadata), such as variables and units. Data in a dataset is represented in a form that can be accessed by machines with different methods of storing integers, characters, and floating-point numbers. It has a strong set of functional libraries that can be used to compress, subset and transform data. The NetCDF software libraries and documentation are available online from Unidata[4]. Many NetCDF manipulation and display software utilities are also available online[5].

NetCDF files contain the data as 'variables', which can be single numbers, vectors, or multi-dimensional arrays. Variables can be of data types such as char (character), byte, short, int (integer), float or real, and double. NetCDF files contain data organised into a collection of named array variables along with named data file attributes. The format is widely applicable to many data types. The Unidata NetCDF User's Guide thoroughly documents the many benefits and few limitations of NetCDF format[6]. Unidata released the NetCDF-4 format in 2008. This format is more flexible than the classic format and offers additional features such as groups, compound types and variable length arrays.

2.1 Common Data Language (CDL)

Common Data Language (CDL) is a human readable text notation that is used to describe NetCDF objects. The NetCDF utility *ncdump* can be used to convert NetCDF object binary to CDL text. The NetCDF utility ncgen creates a NetCDF binary file from a well-formed CDL text file. A CDL description of a NetCDF dataset takes the form:

```
netcdf name { // example of CDL notation
dimensions: .... ;
variables: .... ;
data: .... ; }
```

[1] https://cfconventions.org/documents.html
[2] https://www.unidata.ucar.edu/software/netcdf/documentation/NUG/
[3] https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3
[4] https://www.unidata.ucar.edu/downloa /
[5] https://www.unidata.ucar.edu/software/netcdf/software.html
[6] https://www.unidata.ucar.edu/software/netcdf/docs/faq.html

Where the *name* is used only as a default in constructing file names by the ncgen utility. The CDL description consists of three optional parts, each introduced by the keywords *dimensions*, *variables* and *data*. NetCDF dimension declarations appear after the dimensions keyword, variables and attributes are defined after the variables keyword, and variable data assignments appear after the data keyword. CDL statements are terminated by a semicolon. Spaces, tabs, and new lines can be used freely for readability. Comments in CDL follow the characters '//' on any line. A CDL example which describes TERN flux data collected from and processed by TERN OzFlux and TERN Ecosystem Processes is shown in Figure 1.

TERN uses NetCDF Climate and Forecast (CF) Metadata Conventions v1.8 wherever possible. However, sometimes it was desirable to incorporate attributes or concepts from other conventions, e.g., ACDD. For more information about the metadata properties included, see the column 'Source' in Table 1).

2.2 TERN NetCDF Data Format

The TERN NetCDF data format will enable the inclusion of standard terms for the short names of both coordinates and data variables (measurements). File names should be created using an agreed NetCDF file-naming convention (section 4). Coordinate variables, which describe the dimensions of a data set, are limited to a single set of 3-dimensional axes representing longitude, latitude, and time (X, Y and T). If data cannot be represented in a single time axis, then separate files are created for these datasets. TERN data files will be flexible enough to contain all TERN data variables.

TERN NetCDF file conventions require that:

- Units are compatible with COARDS/UDUNITS
- The time parameter is encoded as recommended by COARDS and CF
- Parameters are given standard names from the CF standard name table
- Where time is specified as an attribute, the ISO8601 standard is used.

For more information on CF, COARDS, NetCDF, Udunits, and ISO8601 see:

- **NetCDF:** https://www.unidata.ucar.edu/software/netcdf/docs/index.html
- **Udunits**: http://www.unidata.ucar.edu/software/udunits/
- **CF:** https://cfconventions.org/
- **COARDS:** https://ferret.pmel.noaa.gov/Ferret/documentation/coards-netcdf-conventions
- **ISO8601**: http://en.wikipedia.org/wiki/ISO_8601

```
dimensions:
    time = 191476;
    latitude = 1;
    longitude = 1;
    name_strlen = 8;
variables:
    double time(time=191476);
      :long_name = "time";
      :standard_name = "time";
      :units = "days since 1800-01-01 00:00:00.0";
      :calendar = "gregorian";

    double latitude(latitude=1);
      :long_name = "latitude";
      :standard_name = "latitude";
      :units = "degree_north";

    double longitude(longitude=1);
      :long_name = "longitude";
      :standard_name = "longitude";
      :units = "degree_east";

    double AH(time=191476, latitude=1, longitude=1);
      :coverage_L3 = "96%";
      :height = "2m";
      :instrument = "HMP45C";
      :long_name = "Absolute humidity";
      :standard_name = "mass_concentration_of_water_vapor_in_air";
      :statistic_type = "average";
      :units = "g/m^3";
      :valid_range = 0.0, 35.0; // double

……

// global attributes:
  :Conventions = "CF-1.8";
  :acknowledgement = "This work used eddy covariance data collected by the TERN-OzFlux facility.
OzFlux\nwould like to acknowledge the financial support of the Australian Federal Government
via\nthe National Collaborative Research Infrastructure Scheme and the Education
Investment\nFund.";
  :altitude = "21m";
  :canopy_height = "3m";
  :comment = "CF metadata, OzFlux standard variable names";
  :contact = "ozfluxtech@adelaide.edu.au";
  :coverage_flux_L3 = "89%";
  :coverage_meteorology_L3 = "95%";
  :coverage_radiation_L3 = "97%";
  :coverage_soil_L3 = "83%";
  :data_link = "http://data.ozflux.org.au/";
  :date_created = "2021-08-06 21:27:53";
  :history = "June 2021 processing";
  :institution = "University of Adelaide";
  :latitude = "-34.0027";
  :license = "https://creativecommons.org/licenses/by/4.0/";
  :license_name = "CC BY 4.0";
  :longitude = "140.5875";
  :metadata_link = "http://http://www.ozflux.org.au/monitoringsites/<site_name>/index.html";
  :nc_nrecs = "191476";
  :ozflux_link = "http://ozflux.org.au/";
  :processing_level = "L3";
  :publisher_name = "TERN Ecosystem Processes,OzFlux";
  :pyfluxpro_version = "PyFluxPro V3.3.0";
  :python_version = "3.7.7 (default, Mar 13 2020, 21:39:43) \n[GCC 9.2.1 20190827 (Red Hat
9.2.1-1)]";
  :site_name = "Calperum";
  :site_pi = "Wayne Meyer";
  :soil = "Reddish brown, yellowish red, loamy sand, free grain structure; low organic matter,
loose consistence; pH 6";
  :source = "20m flux tower plus array of soil sensors";
  :time_coverage_end = "2021-07-01 12:30:00";
  :time_coverage_start = "2010-07-30 11:00:00";
  :time_step = "30";
  :time_zone = "Australia/Adelaide";
  :title = "Flux tower data set from the Calperum site for the calendar year 2021, Li-7500RS";
  :tower_height = "20m";
  :vegetation = "Mallee";
}
```

Figure 1. A partial view of CDL description of the NetCDF file (Level 3) that contains flux data from the Calperum Chowilla station.

3 NetCDF File Structure

3.1 Global Attributes

3.1.1 Definition

The global attribute section of a NetCDF file contains metadata that describes the overall contents of the file and allows for data discovery. All fields should be human-readable and can be of either 'character' or 'numeric' type. TERN recommends that all listed attributes be used and contain meaningful information unless there are technical reasons rendering this impossible (for example, information that is not available for historical data). Files must at least contain the attributes listed as "mandatory"[7]. Please contact the TERN Data Services and Analytics platform (TDSA) if this is proving difficult. Global attributes can be thought of as conveying five kinds of information:

- What: What are the data in the dataset
- Where: The spatial coverage of the data
- When: The temporal coverage of the data
- Who: Who produced the data
- How: How were the data produced and made available

3.1.2 Global Attributes List

Table 1 lists all the global attributes used to define a TERN dataset. Required fields are marked with (*) whereas highly recommended fields are marked with (*). The attributes are determined based on the Attribute Convention for Data Discovery (ACDD) 1-3[8]. Some of the attributes are TERN-specific; see the 'Source' column in Table 1. The "Type" values are **S** for string and **N** for numeric (byte, short, long, integer, float or double).

Table 1. Global Attributes.

Attribute Name	Type	Example	Definition	Source
WHAT				
**Conventions	S	Conventions = "CF-1.8"	Name of the format convention used by the dataset. If you want to specify CF and ACDD conventions, please use Conventions = "CF-1.8, ACDD-1.3".	CF Conventions
id	S	id =".."	An identifier for the data set, provided by and unique within its naming authority. The combination of the "naming authority" and the "id" should be globally unique, but the id can be globally unique by itself also. IDs can be URLs, URNs, DOIs, meaningful text strings, a local key, or any other unique string of characters. The id should not include white space characters. Use the DOI of the published data collection, e.g., as displayed in the TERN Data Discovery Portal (TDDP).	ACDD 1-3
**title	S	title = "Flux tower data set from the Adelaide	Short title of the dataset.	CF Conventions

[7] https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#_attributes

[8] https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3

		River site for the calendar year 2009"		
summary	S	summary = "....."	The abstract describing the dataset including the type of data contained in the dataset, how the data was created, the creator of the dataset, the project for which the data was created, the geospatial coverage of the data, the temporal coverage of the data.	ACDD 1-3
*keywords	S	keywords = "SHORTWAVE RADIATION, AIR TEMPERATURE, …"	A comma separated list of keywords and phrases. We recommend using the GCMD Science Keywords, see https://gcmd.earthdata.nasa.gov/kms/concepts/concept_scheme/sciencekeywords/ Non-GCMD keywords may be used at your discretion.	ACDD 1-3
keywords_vocabulary	S	keywords_vocabulary = "GCMD:GCMD Science Keywords"	If you are using a controlled vocabulary for the words/phrases in your "keywords" attribute, this is the unique name or identifier of the vocabulary from which keywords are taken. If more than one keyword vocabulary is used, each may be presented with a prefix and a following comma, so that keywords may optionally be prefixed with the controlled vocabulary key. Example: 'GCMD: GCMD Science Keywords, CF: NetCDF COARDS Climate and Forecast Standard Names'.	ACDD 1-3
*date_created	S	date_created = "2019-11-17T22:19:33Z"	The date on which the data was created. See section 3.1.3 on time format below.	ACDD 1-3
date_modified	S	date_modified = "2019-11-17T22:19:33Z"	The date on which the data was modified. See section 3.1.3 on time format below. If this attribute is used for the first time or modified, a new entry needs to be added to the "history" attribute.	ACDD 1-3
comment	S	comment = "....."	Miscellaneous information about the data, not captured elsewhere. This attribute is defined in the CF Conventions. Any free- format text is appropriate.	CF Conventions
references	S	references = ""	Published or web-based references that describe the data or methods used to produce it. Recommend URIs (such as a URL or DOI) for papers or other references. This attribute is defined in the CF conventions. Include a reference to TERN and a project-specific reference if appropriate.	CF Conventions
naming_authority	S	naming_authority = "tern.org.au"	The organization that provides the initial identifier for the dataset. The naming authority should be uniquely specified by this attribute. We recommend using reverse-DNS naming for the naming authority; URIs are also acceptable. This will always be "tern.org.au"	ACDD 1-3
cdm_data_type	S	cdm_data_type = "Station"	The "cdm_data_type" attribute gives the Unidata CDM (Common Data Model) data type used by THREDDS, e.g., "Point", "Trajectory", "Station", "Radial", "Grid", "Swath". For more information, seehttps://www.unidata.ucar.edu/software/tds/current/catalog/InvCatalogSpec.html	ACDD 1-3
metadata_link	S	metadata_link = "https://portal.tern.org.au/adelaide-river-flux-collection/22132"	URL to the metadata record corresponding to the NetCDF file. This can be a link to TERN Geonetwork record, data landing page (TDDP) or OzFlux metadata page.	ACDD 1-3
featureType	S	featureType = "timeSeries"	Required for all Discrete Geometry representations except the orthogonal multidimensional array representation, for which it is highly recommended. The value assigned to the featureType attribute is case-insensitive; see http://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#_features_and_feature_types	CF Conventions

processing_level	S	processing_level = “L3”	A textual description of the processing (or quality control) level of the data.	ACDD 1-3
product_version	S	product_version = “v1”	Version identifier of the data file or product as assigned by the data creator.	ACDD 1-3
WHERE				
**latitude	N	latitude = -36.6732;	The latitude of the point.	TERN
**longitude	N	longitude = 145.0294;	The longitude of the point.	TERN
geospatial_vertical_min	N	geospatial_vertical_min = 10	Minimum depth for measurements in meters.	ACDD 1-3
geospatial_vertical_max	N	geospatial_vertical_max = 2000	Maximum depth for measurements in meters.	ACDD 1-3
WHEN				
*time_coverage_start	S	time_coverage_start = “2011-12-01T08:35:00Z”	Start date of the data in UTC. See section 3.1.3 on time format below	ACDD 1-3
*time_coverage_end	S	time_coverage_end = “2019-11-16T08:35:00Z”	Final date of the data in UTC. See chapter 3.1.3 on time format below	ACDD 1-3
time_coverage_resolution	S		Describes the targeted period between each value in the data set. Use ISO 8601:2004 duration format, preferably the extended format as recommended in the Attribute Content Guidance section in ACDD.	ACDD 1-3
time_zone	S	time_zone = “Australia/Darwin”	Local time zone. See chapter 2.1.3 on time format below. If local time does not fall into one zone for the full dataset, do not use this attribute.	TERN
**time_step	S	time_step = "30"	Time interval (minutes)	TERN
WHO				
*publisher_name	S	publisher_name= “TERN”	The name of the person (or other entity specified by the publisher_type attribute) responsible for publishing the data file or product to users, with its current metadata and format.	ACDD 1-3
publisher_type	S	publisher_type = “institution”	Specifies type of publisher with one of the following: 'person', 'group', 'institution', or 'position'. If this attribute is not specified, the publisher is assumed to be a person.	ACDD 1-3
publisher_email	S	publisher_email = “esupport@tern.org.au”	The email address of the person (or other entity specified by the publisher_type attribute) responsible for publishing the data file or product to users, with its current metadata and format.	ACDD 1-3
publisher_url		publisher_url= “https://www.tern.org.au”	The URL of the publisher responsible for publishing the data file or product to users, with its current metadata and format.	ACDD 1-3
contact	S	contact = “jason.beringer@monash.edu”	NetCDF file author contact e- mail address	TERN
site_pi	S	site_pi = “Jason Beringer”	Name of the person responsible for the creation of the dataset	TERN
*institution	S	institution = “Monash University”	The name of the institution principally responsible for originating this data. This attribute is recommended by the CF convention. Multiple institutions should be separated with a comma.	CF convention
project	S	Project = “TERN”	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by comma.	TERN
HOW				
citation	S		The citation to be used in publications using the dataset should follow the DataCite citation format: Creator (PublicationYear): Title. Version. Publisher. (resourceTypeGeneral). Identifier	TERN
acknowledgement	S	acknowledgement = “Data was sourced from the Terrestrial Ecosystem Research Network (TERN) - an initiative of the Australian Government	Any users (including re-packagers) of TERN data are required to clearly acknowledge the source of the material in this format. Add additional acknowledgment if required, for example to OzFlux and any other organisations contributes to the data creation and publication.	ACDD 1-3

		~~being conducted~~ as part of the National Collaborative Research Infrastructure Strategy."		
license	S	license = "https://creativecommons.org/licenses/by/4.0/";	Provide the URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	ACDD 1-3
license_name	S	license_name = "CC BY 4.0";	The name of the license.	TERN
data_link	S	data_url = " http://dap.ozflux.org.au/thredds/fileServer/ozflux/sites/AdelaideRiver/L3/default/AdelaideRiver_L3.nc"	The URL to access the published data files, e.g., from THREDDs server.	TERN
**site_name	S	site_name= " Adelaide River"	The name of the flux monitoring station.	TERN
platform	S	platform = "..."	Name of the platform(s) that supported the sensor data used to create this data set or product. Platforms can be of any type, including satellite, ship, station, aircraft or other.	ACDD 1-3
history	S	history = "....."	Provides an audit trail for modifications to the original data. It should contain a separate line for each modification, with each line beginning with a timestamp and including username, modification name and modification arguments.	ACDD 1-3
source	S	source = "21m flux tower plus array of soil sensors"	The method of production of the original data. If it was model-generated, the source should name the model and its version. If it is observational, the source should characterise it. This attribute is defined in the CF Conventions. Examples: 'temperature from CTD #1234'; 'world model v.0.1'.	CF Conventions

3.1.3 Time Formats

Time consists of three components:

- The actual time as hours, minutes, and seconds
- The time zone of the location at which the measurement was made
- The reference point against which the time is measured

3.1.3.1 Actual Time

All time will be recorded in hours, minutes, seconds, and decimal fractions of seconds relative to UTC that is Universal Time Coordinate or the old GMT. Whenever time information is given in the global attributes, it ought to be a string of the format: "YYYY- MM-DDThh:mm:ssZ" (i.e., year – month – day T hour : minute : second Z). If higher resolution than seconds is needed, any number of decimal digits (".s") for the seconds is acceptable: "YYYY-MM-DDThh:mm:ss.sZ"

Time should be recorded using the international standard ISO 8601 (International Organisation for Standardisation, 2009). Examples are shown below.

- Complete date:
  - 2013-07-16
- Complete date plus hours and minutes with optional "Z":
  - 2005-10-24T08:00:00Z
  - 2008-01-01T22:50:02.03Z

3.1.3.2 Time Zone

The time zone value gives the local time which is important in considering many biological processes and phenomena. For example, we may measure time as hours, minutes, and seconds from a reference such as UTC but to be able to relate that to local diurnal processes we also need to know the local time zone. The local time zone will be recorded as the hours plus or minus from the longitude meridian. Examples of the time format are shown below:

- 2008-10-24T08:00:00Z (UTC)
- 2008-10-24T18:00:00+10 (Local)

The time zone information should be included in the global attribute *time_zone*. Corrections from local to UTC time for Australian time zones can be found online at: http://www.timeanddate.com/worldclock/timezone.html?n=396

3.1.3.3 The Reference Time

This value represents the reference point against which the time is measured. This value will be used in the next chapter and particularly in the attributes representing the coordinate variables. TERN suggests that all the TERN data should use the Unix reference time (epoch) of 1st January of 1970 as it is widely used in file formats and operating systems. The value will be stored as the number of days since this reference time.

3.1.4 User-specified Global Attributes

The global attributes listed in Table 1 are most important to define a dataset as clearly as possible. However, this list will not in all cases be exhaustive and TERN requests that other meaningful global attributes be used where necessary. It is possible to add global attributes to meet specific facility needs. New attributes will need to be self-defined, including a description and an example of how it is used. User-defined global attributes should be added to the existing list in the next version of the TERN NetCDF User's manual.

3.2 Data File Dimensions

NetCDF file dimensions provide information on the number and size of the data variables. TERN allows a single variable for each of the data dimensions, i.e., *time*, *latitude*, and *longitude*. There may only be one unlimited dimension, i.e., as many instances of this variable as needed, for a limited number of coinciding variables. Other dimensions may be greater than 1 but must be defined and may not be unlimited. Coordinate types other than latitude, longitude, depth, and time are allowed.

The example in the table below allows for measurements at an unlimited number of time steps, at one latitude, longitude. Requirements are described further in the section on coordinate variables (3.3.1).

Name	Example	Comment
Time	time = unlimited	Number of time steps
Latitude	latitude = 1	Dimension of the latitude coordinate variable
Longitude	longitude = 1	Dimension of the longitude variable

3.3 Variables

NetCDF variables include data measured by instruments, parameters derived from the primary measurements and coordinate variables, which may be nominal values such as values for depth for instruments that do not directly record depth. Each variable has a specific set of attributes, some of which are mandatory. This section will be divided into 5 different subsections as listed below:

- The coordinate variables
- The data variables
- Quality control 'sets'
- Defining uncertainties in data measurement
- An example of a NetCDF header.

3.3.1 Coordinate Variables

The coordinate variables orient the data in time and space. For this purpose, they have an "axis" attribute defining that they point in X, Y and T dimensions. The X and Y of course represent horizontal space while T is time. The use of a common set of spatial and temporal units and measures is the basic requirement to be able to integrate the various data collected by the TERN project.

3.3.1.1 Time

All time will be recorded in hours, minutes, seconds, and decimal fractions of seconds relative to UTC that is Universal Time Coordinate or the old Greenwich Mean Time. Time consists of the following components:

- The actual time as hours, minutes, and seconds
- The time zone of the location at which the measurement was made
- The reference point against which the time is measured. This component is used to define the reference point in the attribute named "units".

The following Table 2 presents the different attributes used to represent the time variable. Mandatory fields are marked with an asterisk (*). The "Type" values are **S** for string and **N** for numeric (byte, short, long, integer, float or double). To identify the time variable, three parameters are used: Type, Name and Dimension.

Table 2. List of attributes defining the TIME variable for NetCDF files.

Attributes	Type	Example	Comment
*standard_name	S	standard_name = "time"	A standard name that references a description of a variable's content in the CF standard name table
*long_name	S	long_name = "time "	A descriptive name that indicates a variable's content. This name is not standardised.
*units	S	units = "days since 1950-01-01T00:00:00Z"	Units of a variable's content. The acceptable units for time are listed in the udunits.dat file. The reference time string may include date alone; date and time; or date, time, and time zone.
calendar	S	calendar = "gregorian"	Calendar used for encoding time axes, see 4.4 at http://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html

Example:

```
double time(time) ;
      time:long_name = "time" ;
      time:standard_name = "time" ;
      time:units = "days since 1800-01-01 00:00:00.0" ;
      time:calendar = "gregorian" ;
```

3.3.1.2 Location (X-Y horizontal Space)

With the advent of GPS systems, it is now possible to measure position in space easily and accurately. The location will be measured by GPS or equivalent as Latitude/Longitude in Decimal degrees using the WGS84 projection with locations south of the equator as negative values and values west of zero degrees of Longitude being negative. Table 3 and Table 4 present the different attributes used to define the variables Latitude and Longitude.

Table 3. List of attributes that define the Latitude variable for TERN NetCDF files.

Attributes	Type	Example	Comment
*standard_name	S	standard_name = “latitude”	A standard name that references a description of a variable’s content in the standard name table. Latitude of the measurements.
*long_name	S	long_name = “latitude”	A descriptive name that indicates a variable’s content. This name is not standardised.
*units	S	units = “degrees_north”	Units: degrees north; southern latitudes are negative Example: -44.4991 for 44 29’ 56.76” S
axis	S	axis = “Y”	Identifies Y axes

Table 4. List of attributes to define Longitude variable for TERN NetCDF files.

Attributes	Type	Example	Comment
*standard_name	S	standard_name = “longitude ”	A standard name that references a description of a variable’s content in the standard name table. Longitude of the measurements.
*long_name	S	long_name = “longitude”	A descriptive name that indicates a variable’s content. This name is not standardised.
*units	S	units = “degrees_east”	Units: degrees east; western longitudes are negative Example: -16.7222 for 16 43’ 19.92” W
axis	S	axis =“X”	Identifies X axes

Example:

```
double latitude(latitude) ;
      latitude:long_name = "latitude" ;
      latitude:standard_name = "latitude" ;
      latitude:units = "degree_north" ;
double longitude(longitude) ;
      longitude:long_name = "longitude" ;
      longitude:standard_name = "longitude" ;
      longitude:units = "degree_east" ;
```

3.3.2 Data Variables

The data variables names should begin with a letter and be composed of letters, digits, and underscores. To identify a variable, three parameters are used: Type, Name and Dimension. In data variable definitions, the dimensions (time, latitude, longitude) must be used in this order (if present) as they appear in CDL. When "extra" dimensions are used, such as with model runs, they should appear to the left of the standard dimensions in a variable definition.

Example of variable:

```
double AH(time, latitude, longitude) ;
      AH:long_name = "Absolute humidity";
      AH:standard_name = "mass_concentration_of_water_vapor_in_air";
      AH:statistic_type = "average";
      AH:units = "g/m^3";
      AH:valid_range = 0., 35.;
```

The following table presents the different attributes used to represent a specific variable. Highly recommended fields are marked with an asterisk (*).

Table 5. List of data variables attributes for TERN NetCDF files.

Attributes	Type	Example	Comment
*standard_name	S	standard_name = "sea_surface_temperature"	A standard name that references a description of a variable's content in the CF standard name table.
*units	S	units = "degC"	UNIDATA's Udunits.
*_FillValue	D	_FillValue = 99999	A value used to represent missing or undefined data. The example uses '9999' to represent the missing value in the data. If only one missing value is needed for a variable, then we recommend that this value be specified using the _FillValue attribute. The _FillValue should be outside the range specified by the attribute valid_range (if specified) for a variable.
*long_name	S	long_name = "Surface temperature in degree Celsius"	A descriptive name that indicates a variable's content. This name is not standardised and maybe used for labelling plots.
valid_min	N	valid_min = -2.0	Minimum value for valid data
valid_max	N	valid_max = 40	Maximum value for valid data
valid_range			A vector of two numbers specifying the minimum and maximum valid values for this variable, equivalent to specifying values for both valid_min and valid_max attributes. Any of these attributes define the valid range. The attribute valid_range must not be defined if either valid_min or valid_max is defined.
ancillary_variables	S	ancillary_variables = "TEMP_quality_control"	Identifies a variable that contains closely associated data, e.g., the measurement uncertainties of instrument data. See section 3.3.3.
*instrument	S	instrument = "Open Path CO2/H2O Gas Analyzer - LI-7500"	This attribute is defined by TERN based on ACDD 1-3. Use controlled vocabularies to describe the instrument, e.g., see TERN instrument vocabulary scheme http://linked.data.gov.au/def/tern-cv/a3088b5c-622d-4e25-8a75-4c4961b0dfe8
statistic_type	S	statistic_type = "average"	This attribute has been added as requested by TERN EP central team to use different interpolation methods for sums, e.g., rainfall.

3.3.3 Ancillary Variables

When one data variable provides metadata about the individual values of another data variable it may be desirable to express this association by providing a link between the variables. For example, instrument data may have associated measures of uncertainty; data points may have associated quality control flags. The attribute “ancillary_variables” is used to express these types of relationships.

Example:

```
float TEMP(TIME,  LATITUDE, LONGITUDE) ;
      TEMP:long_name = "Water Temperature in degrees C" ;
      TEMP:units = "degC" ;
      TEMP:standard_name = "sea_water_temperature" ;
      TEMP:_FillValue = 99999 ;
      TEMP:ancillary_variables = "TEMP_qc TEMP_uncertainty";

byte TEMP_qc(TIME,  LATITUDE, LONGITUDE) ;
      TEMP_qc:standard_name = "sea_water_temperature status_flag" ;
      TEMP_qc:convention = "CF-1.8";
      TEMP_qc:_FillValue = -10 ;
      TEMP_qc:flag_values = 0, 1, 2, 3, 4, 5;
      TEMP_qc:flag_meanings = "no_qc_performed good_data
probably_good_data bad_data_that_are_potentially_correctable
bad_data missing_value";
```

3.3.4 User-specified Data Variable Attributes

User-defined data variables will be included in future versions of the NetCDF document. This document will be revised regularly, and new user-defined data variable attributes incorporated into Table 5.

3.3.5 Quality Control (QC)

Quality control involves some sort of assessment of the data to identify data points or even data sets which have errors that limit their use. The basic approach used by TERN is to keep all the data but to flag data or data sets that do not meet the quality assessment standards of data collectors / principal investigators. Quality Control is a complex area and one that will be implemented in a more sophisticated manner as the project progresses. We present here the basic requirements for QC of TERN data. Most TERN platforms are currently using a quality control procedure. TDSA will not proscribe changes to procedures that are already in use. On the other hand, it is appropriate that all TERN facilities use the same convention to qualify their quality control procedures: for example, the naming of the QC variables or the naming of possible variable attributes.

This section will present the different QC procedures currently used within the TERN project. It will also present the different attributes and the variables available to define a quality control procedure.

3.3.5.1 Definition of the variables and attributes

Highly recommended attributes for variables as marked by * sign in Table 5 must be provided by data contributors.

3.3.5.2 Quality Flag

The coordinate variables (TIME, LATITUDE and LONGITUDE) utilise the same quality control variables as the data variables. The quality of the data in a variable is described by the ancillary variable <VARIABLE_qc>. To identify a QC variable, three parameters are used: Type, Name and Dimension.

Example for the TIME, LATITUDE, LONGITUDE and DEPTH variables:

```
Byte TIME_qc(TIME);
Byte LATITUDE_qc(LATITUDE)
Byte LONGITUDE_qc(LONGITUDE);
Byte DEPTH_qc(DEPTH)
```

The attributes flag_values, flag_masks and flag_meanings are intended to make variables that contain flag values self-describing. For more information about these attributes, see NetCDF Climate and Forecast (CF) Metadata Conventions 1.8 (section 3.5. Flags).

4 TERN NetCDF File Naming Convention

This section specifies the format of filenames that will be used to distribute TERN data in a NetCDF format. The file name extension of each NetCDF file must be **".nc".** Filenames can be up to 255 characters in length and are composed of up to 10 fields separated by **'_'** (underscore) characters. Characters which can be used within fields are letters (A to Z) and whole numbers (0 to 9). The hyphen character (-) may also be used within fields.
TERN recommends the following NetCDF file naming convention:

```
<SiteName>_<ProcessingLevel>_<StartDate><EndDate>.nc
```

These fields are mandatory and must conform to the following content guidelines.

- <SiteName>: The short, unique name of the flux monitoring site.
- <ProcessingLevel>: Level of processing such as L3,L4,L5 and L6.
- <StartDate>: start date and time of the measurement, not of file creation. The date and time should be in YYYYMMDD format, and the date would be UTC date.
- <EndDate>: end date and time of the measurement. The date format is the same as the start date.

The following fields are optional. In a case where the file cannot be uniquely identified by those mandatory fields, the following optional names are desired after those mandatory fields.

- <FileVersion>: value representing the version of the file
- <FileType>: Whether data is Summary, Monthly, Daily, Cumulative, Annual

Fields | Example	WombatStateForest_L6_20100120_20200120.nc	WombatStateForest_L6_20100120_20200120_Annual.nc
SiteName	WombatStateForest	WombatStateForest
ProcessingLevel	L6	L6
Fromdate	20100120	20100120
ToDate	20200120	20200120
Type (If applicable)	--	Annual

Related Resources

1. Eaton, B., Gregory, J., Drach, B., Taylor, K, Hnakin, S., Blower, J., Caron, J., Signell, R., Bentley, P. and Rappa, G., Hoeck, H., Pamment, A., Juckes, M., Raspaud, M., Horne, R., Whiteaker, T., Blodgett, D., Zender, C., Lee, D. NetCDF Climate and Forecast (CF) Metadata Conventions, v1.8, URL: http://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html
2. Mancini, S., Tattersall, K., Proctor, R. (2009), IMOS User's Manual, NetCDF Conventions and Reference Tables , v1.2, URL: http://imos.org.au/fileadmin/user_upload/shared/emii/IMOS_netCDF_usermanual_v1.2.pdf
3. TERN and CSIRO AusCover (2008). NetCDF File Conventions using Climate and Forecast metadata (CF-1.6) and Attribute Conventions Data Discovery metadata (ACDD 1.3), v1, URL: https://docs.google.com/document/d/1bVWf9SIRhkeEbtLYiclIxf_87ibgMwIFasyfOxh6fi0/edit
4. UniData NetCDF User's Guide v1.1, URL: https://www.unidata.ucar.edu/software/netcdf/documentation/NUG/

We at TERN acknowledge the Traditional Owners and Custodians throughout Australia, New Zealand and all nations. We honour their profound connections to land, water, biodiversity and culture and pay our respects to their Elders past, present and emerging.

TERN is enabled by NCRIS. Our work is a result of collaborative partnerships with many universities and institutions.

To find out more please go to **tern.org.au.**

tern
Ecosystem Research Infrastructure

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative