TUTORIAL

NetCDF Data Extraction at the ARM Archive Using NCVweb

NCVweb Overview

Like many scientific organizations, the ACRF Data Archive stores and distributes atmospheric data from the ACRF sites in Network Common Data Format, or NetCDF. This file format is storage efficient, machine independent, and self describing. However, to work with the data files directly requires some programming skill and use of the NetCDF programming libraries, or API.

NCVweb was developed to overcome these shortcomings with the data format. The application reduces the need of and problems associated with downloading large volumes of data, installing and configuring visualization software, or writing custom data exploration software. Since it operates by querying each file for the details necessary for visualization, little or no work goes into preparing the tool for new data streams. Since the tool is Web-based, ARM researchers and instrument mentors can visualize large and complex data sets without needing to be skilled in NetCDF programming.

The tool supports a wide number of visualizations as outlined in the following table:

NCVweb Plot Types

Plot Type	Example
X-Y	Time vs. Irradiance or Temperature vs. Altitude
X-Y-Y2	Ambient and Calibration Temperatures Vs. Time
Multi-Facility	Soil Temperature at several different Facilities Vs. Time
X-Y-Z Color Image	Temperature Profile (over height and time)
Horizontal Slices	Plots of Temperature vs. Time at several Altitudes
Vertical Slices	Plots of Temperature vs. Altitude at several, specific instances during the day

General plotting features

- Plots data across multiple files
- Detects and does not plot data points designated as "missing"
- Plots any compatible variable against any other
- Allows arbitrary zoom or pan to any data region
- Detects and does not connect data across significant time gaps
- Provides useful time conversions for the time axis
- Allows for various symbol and/or line combinations
- Supports multiple plot sizes to accommodate various screen sizes
- Auto-generates appropriate plot labels based on the data being viewed
- X/Y Axis Flip

Additional features

- Choice of color table for false color three-dimensional visualizations
- Day/Night cycle background color coding (yellow/blue)
- Computes statistics for the aggregate data set
- Displays textual data values of the current zoom region
- Will convert NetCDF variables to ASCII formats for use with other tools
- Generates custom subsetted NetCDF files for more efficient downloads.
- Highlights data that have failed associated quality control flags
- Displays details of file contents, variables, and dimensions
- Online help and directions

NetCDF Data Extraction

A new version of NCVweb is available at the ARM Archive to allow users to create custom NetCDF files from ARM data files.

Features:
- Works with single or multi-dimensioned variables
- Attributes of selected variables are copied into the custom NetCDF file
- All datastream global attributes are copied
- Parentage of data documented in new file
- Works across multiple files from the same datastream
- Greatly reduces size of overall data download

Typically, the archive user will receive an email notification when their data request is ready. As part of that notification will be a URL that launches NCVweb and takes the user directly to their data request.

File selection is performed exactly the same way as if the user was using NCVweb to plot or otherwise inspect a file.

As an example, let's say the user would like a custom NetCDF file containing only a time series of the first cloud base height field from the SGP ceilometer. Order the data, and click on the NCVweb URL received as part of the data request fulfillment notification. A list of files to choose from will be displayed as shown in Figure 1.
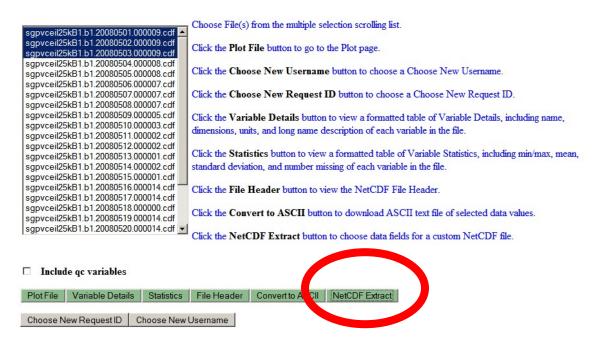
Choose File(s)

Figure 1. File Selection

Multiple files are selected by shift clicking and/or ctrl clicking depending on your browser. After files are selected, press the 'NetCDF Extract' button to proceed to the field selection form. Depending on the size and quantities of the files selected, there may be a delay before the next form appears. Currently, you may select up to 1GB of data at a time. This restriction will be removed in a future release. A status message will indicate how much data is in the process of being accumulated. Figure 2 shows the example of selecting only first_cbh (First cloud base height) along with base_time and time_offset.

Checking size of 3 files...Please wait while I combine 18.70 MB of data...Thanks, I'm done.

Select Variables to Extract from File: sgpvceil25kB1.b1.20080501..20080503.cdf

Pick	Index	Variable Name	Dimensions	Units	Variable Long Name
☑	0	base_time	[scalar]	seconds since 1970-1-1 0:00:00 0:00	Base time in Epoch
☑	1	time_offset	[time=17280]	seconds since 2008-05-01 00:00:09 0:00	Time offset from base_time
☐	2	time	[time=17280]	seconds since 2008-05-01 00:00:00 0:00	Time offset from midnight
☐	4	range	[range=252]	m	Distance to the center of the corresponding range bin.
☐	5	detection_status	[time=17280]	unitless	Detection status.
☐	6	status_flag	[time=17280]	unitless	Ceilometer status indicator.
☑	7	first_cbh	[time=17280]	m	Lowest cloud base height detected.
☐	9	vertical_visibility	[time=17280]	m	Vertical visibility
☐	11	second_cbh	[time=17280]	m	Second lowest cloud base height
☐	13	alt_highest_signal	[time=17280]	m	Altitude of highest signal
☐	15	third_cbh	[time=17280]	m	Third cloud base height
☐	17	laser_pulse_energy	[time=17280]	%	Laser pulse energy
☐	19	laser_temperature	[time=17280]	C	Laser temperature
☐	21	receiver_sensitivity	[time=17280]	%	Receiver sensitivity
☐	22	window_contamination	[time=17280]	mV	Window contamination
☐	23	tilt_angle	[time=17280]	deg	Tilt angle
☐	25	background_light	[time=17280]	mV	Background light
☐	26	sum_backscatter	[time=17280]	1/srad	SUM of detected and normalized backscatter
☐	27	backscatter	[range=252,time=17280]	1/(srad*km*10000.0)	Backscatter
☐	28	measurement_parameters	[string_len=8,time=17280]	unitless	6 character string describing instrument measurement parameters
☐	29	status_string	[string_len=8,time=17280]	unitless	Warning and alarm status bits
☐	30	lat	[scalar]	degrees	north latitude
☐	31	lon	[scalar]	degrees	east longitude
☐	32	alt	[scalar]	meters above Mean Sea Level	altitude

[Extract and Download NetCDF] [Choose New File] [Choose New Request ID] [Choose New Username]

Figure 2. Field Selection

To start the process, click on 'Extract and Download NetCDF'. After a brief delay, the extraction is performed and the new custom file is made available for download. By using the context menu of your browser (usually right-click) on the new file name, so can choose to save the link (or target) to your computer. In this case, the resulting custom file was only 0.2 MB in size, saving the user from downloading an additional 18.47 MB (see Figure 3). The NetCDF file header for the resulting custom file is shown in Figure 4.

Please wait while I create your custom file from a 18.67 MB data chunk...I finished!

Here is your custom NetCDF subset file

Right-click on your file to download it to your computer

sgpvceil25kB1.b1.20080501..20080503.subset.cdf (file size: 0.20 MB)

Using this tool, you saved yourself from downloading an extra 18.47 MB!

| Choose New File | Choose New Request ID | Choose New Username |

Figure 3. Download the Custom File

File Header for sgpvceil25kB1.b1.20080501..20080503.subset.cdf

```
netcdf sgpvceil25kB1.b1.20080501..20080503.subset {

dimensions:
        time = UNLIMITED ; // (17280 currently)
variables:
        int base_time ;
                base_time:string = "1-May-2008,0:00:09 GMT" ;
                base_time:long_name = "Base time in Epoch" ;
                base_time:units = "seconds since 1970-1-1 0:00:00 0:00" ;
        double time_offset(time) ;
                time_offset:long_name = "Time offset from base_time" ;
                time_offset:units = "seconds since 2008-05-01 00:00:09 0:00" ;
        float first_cbh(time) ;
                first_cbh:long_name = "Lowest cloud base height detected." ;
                first_cbh:units = "m" ;
                first_cbh:valid_min = 0.f ;
                first_cbh:valid_max = 7620.f ;
                first_cbh:missing_value = -9999.f ;
                first_cbh:values = "  If detection_status = 1, 2 or 3,\n",
    " lowest cloud base height measured.\n",
    " If detection_status = 4 this value\n",
    " will be stored as -9999 and a value\n",
    " for vertical_visibility will be stored.\n",
    " If detection_status = 0 or 5, -9999 will \n",
    " be stored for this value and vertical_visibility.\n",
    " (5 slashes were read from field)." ;

// global attributes:
                :ingest_version = "process-ingest-vceil_ingest-8.1-0" ;
                :libingest_version = "ds-dsutil-ingest_lib-1.7-0" ;
                :libdslibc_version = "ds-dslib-c_lib-1.2-0" ;
                :libdsdb_version = "database-dsdb-c_lib-1.2-0" ;
                :ingest_software = " vceil_ingest.c,v 8.2 2006/09/08 22:53:57 ermold process-ingest-vceil_ingest-
8.1-0 $" ;
                :proc_level = "b1" ;
                :input_source = "/data/collection/sgp/sgpvceil25kB1.00/08050100.DAT" ;
                :site_id = "sgp" ;
                :facility_id = "B1: Hillsboro, Kansas" ;
                :sample_int = "15 seconds" ;
                :averaging_int = "11.5 seconds" ;
                :serial_number = "R085002" ;
                :comment = "None" ;
                :resolution_description = "The resolution field attributes refer to the number of significant\n",
    "digits relative to the decimal point that should be used in\n",
    "calculations.  Using fewer digits might result in greater uncertainty;\n",
    "using a larger number of digits should have no effect and thus is\n",
    "unnecessary.  However, analyses based on differences in values with\n",
    "a larger number of significant digits than indicated could lead to\n",
    "erroneous results or misleading scientific conclusions.\n",
    "\n",
    "resolution for lat= 0.001\n",
    "resolution for lon = 0.001\n",
    "resolution for alt = 1" ;
                :get_algorithm = "Instrument response to GET ALGORITHM query.\n",
    "noise_scale              = 1.7\n",
    "minimum_scale            = 30\n",
    "minimum_extinction_coef = 6" ;
```

```
                    :get_factory = "Instrument response to GET FACTORY query.\n",
"in_power                 = 200\n",
"out_power                = 1500 \n",
"coarse_compensation      = 20 \n",
"fine_compensation        = 255 \n",
"receiver_test_value      = 100 \n",
"clean_window             = 100 mV" ;
                    :get_message = "Instrument response to GET MESSAGE query\n",
"angle_correction         = OFF\n",
"height_offset            = 0\n",
"noise_h2_compensation    = ON\n",
"profile_scale            = 100%" ;
                    :unit_num = "C" ;
                    :software_level = "20" ;
                    :qc_method = "Standard Mentor QC" ;
                    :Mentor_QC_Field_Information = "For each qc_<field> interpret the values as follows:\n",
"\n",
"Basic mentor QC checks (bit values):\n",
"========================================\n",
"0x0 = value is within the specified range\n",
"0x1 = value is equal to \'missing_value\'\n",
"0x2 = value is less than the \'valid_min\'\n",
"0x4 = value is greater than the \'valid_max\'\n",
"0x8 = value failed the \'valid_delta\' check\n",
"\n",
"If the value is a \'missing_value\' no min, max, or delta checks are performed.\n",
"\n",
"The delta checks are done by comparing each data value to the one just\n",
"prior to it in time. If a previous data value does not exist or is a\n",
"\'missing_value\' the delta check will not be performed.\n",
"\n",
"Note that the delta computation for multi-dimensioned data compares the\n",
"absolute value between points in the same spatial location to the previous\n",
"point in time.\n",
"\n",
"If the associated non-QC field does not contain any mentor-specified minimum,\n",
"maximum, or delta information a qc_field is not generated.\n",
"" ;
                    :zeb_platform = "sgpvceil25kB1.b1" ;
                    :history = "Mon Jun 30 17:14:46 2008: /var/www/html/ncvweb/ncrcat -O -p
/http/ftp/armguest_user/moore/112788 sgpvceil25kB1.b1.20080501.000009.cdf sgpvceil25kB1.b1.20080502.000009.cdf
sgpvceil25kB1.b1.20080503.000009.cdf /var/ftp/tmp/ncvweb/tmp/1214860386.787/sgpvceil25kB1.b1.20080501..20080503.cdf\n",
"created by user dsmgr on machine ruby at 1-May-2008,3:10:04, using $State: ds-zebra-zeblib-4.16-0 $" ;
                    :NCV_extraction_from_file =
"/var/ftp/tmp/ncvweb/tmp/1214860386.787/././sgpvceil25kB1.b1.20080501..20080503.cdf" ;
                    :NCV_extraction_timestamp = "Mon Jun 30 17:14:57 2008" ;
                    :NCV_extraction_software = "$Id: ncv_anders.pro,v 1.4 2007/12/05 05:03:29 moore Exp $" ;
}
```

Figure 4. Example Custom File Header