6.7 INTEGRATED IN-SITU OCEAN DATA DELIVERY FOR LOW AND HIGH BANDWIDTH INTERNET CONNECTIONS

L. Charles Sun [1*], Nancy N. Soreide [2], John. R. Osborne [3], and Joe. Sirott [2]
[1]NOAA/NESDIS/NODC, 1315 East-West Highway, Maryland
[2]NOAA/OAR/PMEL, Seattle, Washington
[3]OceanAtlas Software, Vashon, Washington

1. INTRODUCTION

Currently many developing countries in the Pacific Islands Region are facing with inadequate technical capacity to assess to diverse climate data sets; and without adequate dissemination capacity, there is a tremendous gap in the ability of ocean products and services. Furthermore, the ability of a user to work productively with very large and diversified datasets has been limited by difficulties related to the network transfer of significant subsets of the data, and the subsequent search through the data for specific observations or data characteristics of interest. This difficulty is compounded when the user needs to compare or integrate real-time observations with the extensive climatologic data sets. There is no easy way to integrate the most recently updated observations into user's local datasets and there is no convenient way to explore the data in these archives and use them in user's preferred visualization and analysis tools.

This paper describes a strategy and framework for developing and implementing operationally a state-of-the-art data and information portal with capabilities of exploring in-situ data from real-time data streams and integrating the data streams with archives across the Internet through a "one-stop shopping" (single interface) query for users who have a high bandwidth Internet connection. Alternatively, users without high-speed network access can order CD/DVD-ROMs from the National Oceanographic Data Center (NODC) that contain the integrated dataset and then use software over potentially low-bandwidth/high-latency network connection to periodically update the CD/DVD-ROM-based archive with new data.

* *Corresponding author address:* Charles Sun, National Oceanographic Data Center, 1315 East West Highway, Silver Spring, MD 20910; e-mail: Charles.Sun@noaa.gov.

NODC, in cooperation with NOAA's Pacific Environmental Laboratory (PMEL), has developed and implemented a Pacific region ocean data and information portal (PRODIP) for delivering integrated real-time and historical environmental datasets to remote users with high and low bandwidth Internet connections. See http://www.nodc.noaa.gov/PRODIP.

2. TECHNICAL APPROACH

Our technical approach is to leverage from and integrate numerous newly developed, proven technologies and prototypes listed below to build a framework for the PORDIP:

- Global Argo Data Repository System: http://www.nodc.noaa.gov/argo
- Global Temperature-Salinity Profile Program: http://www.nodc.noaa/gov/gtspp
- Interactive Data Access and Retrieval System: http://www.nodc.noaa.gov/idars
- Argo NdEdit: http://www.nodc.noaa.gov/argo/tools/ane/install.htm
- The Climate Data Portal: http://www.epic.noaa.gov/cdp/cdpweb.htm

Our developmental strategy was: Step 1) to fully test a prototype of the system at NODC's developmental servers located inside the NODC firewall, then, Step 2). To implement the final (operational) version of the system in NODC's database server located outside the firewall for public access. End-users access the operational system through specially written applications programs installed at NODC's official Web server.

3. PORDIP COMPONENTS

The PRODIP is designed with the "customers/users" in mind. It can be integrated

easily with the existing operational global ocean observing systems to form an end-to-end, unified search, discovery, and access service of historical and near real-time data. Users will have the choice of accessing the data remotely (without having to download any data), or of maintaining an automatically updating local repository.

Figure 1 illustrates a conceptual design of the PORDIP. It consists of four major components. They are:

- Data acquisition component,
- Data access component,
- Data selection component, and
- Data delivery component.

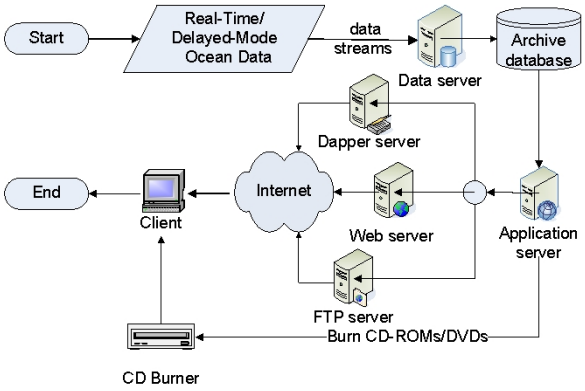

Figure 1. A Schematic diagram shows a conceptual design of an end-to-end ocean data and information.

We shall briefly describe the functionalities of each component in the following sections.

3.1 Data Acquisition Component

As a starting point, the PRODIP distributes near real-time data streams of Argo and Global Temperature-Salinity Profile Program (GTSPP) data in the Pacific Ocean, although the PRODIP can be integrated easily with the existing operational global ocean observing systems. NODC operates the Global Argo Data Repository for managing Argo data and assembles both real-time and delayed-mode GTSPP data into a continuously managed database. Argo and GTSPP are internationally coordinated programs that provide high quality, real-time monitoring of Temperature-Salinity profiles in the global oceans. These programs are major components of the Global Ocean Observing System and the Climate Variability and Predictability program of the World Climate Research Programme.

3.2 Data Access Component

Dapper, an OPeNDAP-enabled in-situ data server developed by Sirott (2004), provides local and distributed access to data. The Dapper Server developed as part of the Climate Data Portal (Soreide et. al., 2001) was modified to include a pluggable framework for easily adding interfaces to new databases and profile datasets. This will provide a common interface for all of the tools that will be used to subset and visualize the data.

3.3 Data Selection Component

NdEdit developed by Osborne and Denbo (2004), an existing interactive tool for selecting and sub-setting large in-situ data collections using spatial/temporal metadata, is used as the primary query interface. NQuery is a program to allow queries based upon the actual measured parameters in a set of data files and to refine that data selection based on the characteristics of the data (Osborne and Denbo, 2005). NQuery reads subset-Argo/GTSPP inventory files produced by Argo/GTSPP NdEdit. NQuery will be extended to ingest Argo and GTSPP netCDF data files at NODC via the Internet and the HTTP netCDF library. NQuery will produce an output file that allows users with access to high-bandwidth network to order NODC data and download the data files directly into desktop visualization tools or users can order NODC data on CD-ROMS.

3.4 Data Delivery Component

NODC published the PORDIP data on optical disc media such as compact-disc (CD) and/or digital versatile disc (DVD). As the disc uses the ISO9660 standard with the RockRidge extension, the platform-independent files can, in principle, be read on all operating systems which are compatible with that format. The disc is designed strictly as a "data-storage medium" with a simple Hypertext Markup Language (HTML) interface to aid in navigating the disc. The HTML was written as generically as possible in order to work with as many browsers and operating systems as possible.

Subsequent monthly updates of the disc are available online for users to download to their local storage devices. Alternatively, users without access to high-speed Internet connections can request that the NODC send them the updates on a hardcopy form. An online version of the disc is available at http://www.nodc.noaa.gov/PRODIP.

4. DATAT INTER-OPERABILITY

Data distributed by the PRODIP are in the ACSII text format and the NetCDF (network common data form) format. We believe that the PRODIP's data formats can be integrated seamlessly with freely available desktop visualization and analysis tools such as:

- ncBrowse (http://www.epic.noaa.gov/java/ncBrowse/)
- Java OceanAtlas (http://odf.ucsd.edu/joa/jsindex.html), and
- Ocean Data View (http://www.awi-bremerhaven.de/GEO/ODV/).

5. OCEAN DATA EXPLORER

NODC has also developed a software package, known as the "Ocean Data Explorer" (ODE) for users to examine oceanographic data stored on the PRODIP CDs/DVDs. The ODE (Figure 2) is a Java-based application that provides interactive graphical exploration and conversion of oceanographic vertical profile data stored on optical disc media.

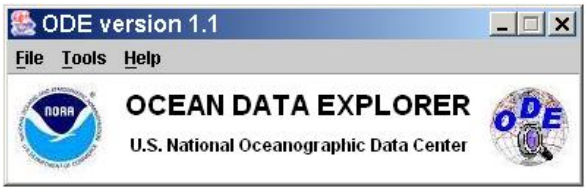

Figure 2. A screen shot of the Ocean Data Explorer (ODE) main window.

5.1 System Requirements

The ODE was developed in Java, an object-oriented programming language. With very few changes to the source code, The ODE can be run on any platform that has a Java Virtual Machine (JRE 1.4.2.x or above) available. The platforms supported currently by ODE include Windows 95, 98, ME, NT, 2000, and XP, Mac OS 8-9.X, Mac OS X, Solaris, Linux, and other flavors of UNIX.

5.2 User Interfaces

The main features of the ODE allow the user:

(1) To explorer Argo profiling floats' metadata (Figure 3) and the "waterfall" plots (Figure 4) of vertical profiles of the measured parameters,

(2) To select and sub-set large in-situ data collections using spatial/temporal metadata (Figure 5),

(3) To convert in-situ station profile data from the NODC NetCDF format to the ASCII text format (Figure 6), and

(4) To decompress ("un-zip") the "win-zip" format files stored on compact-discs.

6. CONCLUSIONS AND FUTURE PLANS

The PRODIP provides efficient data sharing and integration, easy access to and long-term archiving of climate data sets and the capability of synthesis of information in the islands of Hawaii, the Northwestern Hawaii Islands, and the American Flag (AF) Territories. It demonstrates that NOAA products and services are responsive to needs of the Pacific region ocean communities, governments and business.

The current version of the PRODIP serves only ocean vertical profiles of temperature, salinity, conductivity observations collected from the Argo and Global Temperature-Salinity Profile Program projects. Data sets that have coverage in the Pacific Ocean basin will be included in the PRODIP in the future are:

- Coastal Ocean Time Series Data: http://www.nodc.noaa.gov/idars/tsdb/index.html
- NOAA Marine Environmental Buoy Data: http://www.nodc.noaa.gov/BUOY/buoy.html
- World Ocean Circulation Experiment (WOCE) Global Data Version 3.0: http://www.nodc.noaa.gov/woce_v3/index.html

7. REFERENCES

Osborne, J.R., and D.W. Denbo, 2004: NdEdit: A tool for browsing and subsetting 4D netCDF files. Proceedings of the 20th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, 2004 AMS Annual Meeting, Seattle, WA, 12–15 January 2004, paper 10.4.

Osborne, J.R., K.T. McHugh, and D.W. Denbo, 2005: Nquery: A network-enabled data-based query tool for multi-disciplinary earth-science datasets. In The 21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, The 85th AMS Annual Meeting, San Diego, CA, 9–13 January 2005, Paper J2.5 (on CD).

Sirott, J., D.W. Denbo, and W.H. Zhu, 2004: Dapper: An OPeNDAP server for in-situ data. Proceedings of the 20th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, 2004 AMS Annual Meeting, Seattle, WA, 12–15 January 2004, paper 9.12

Soreide, N.N., C.L. Sun, B.J. Kilonsky, D.W. Denbo, and W. Zhu, 2001: A climate data portal. In 17th Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, AMS, Albuquerque, NM, 14–19 January 2001, 191–193.

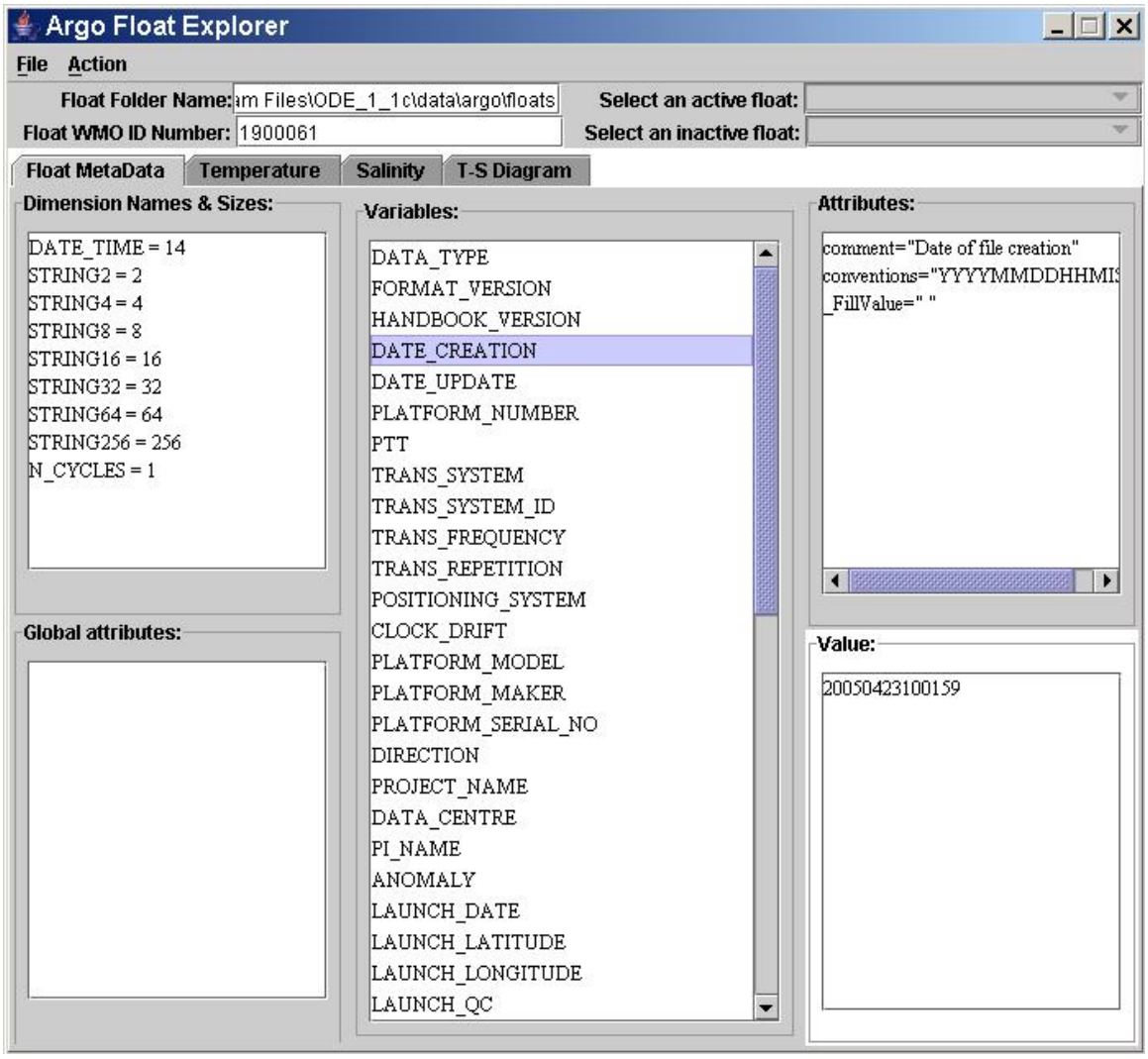

Figure 3. A screen shot of the “Argo Float Explorer (AFE)” application. AFE reads Argo float’s metadata file produced by the Argo Global Data Assembly Center and allows users to view the value of each metadata variable.

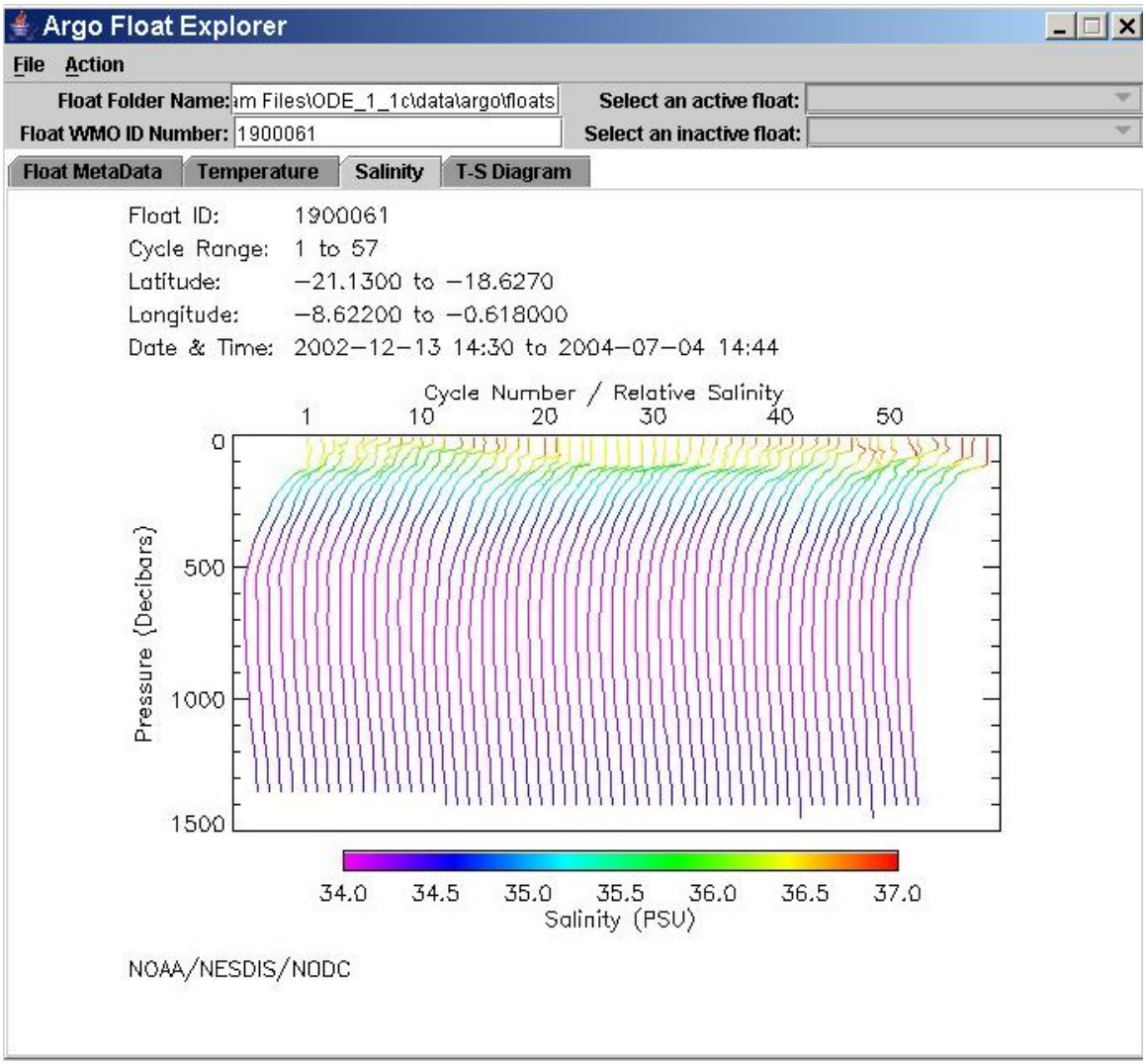

Figure 4. A screen shot of Argo float's waterfall plot showing the salinity observation profiles of Float# 190061.

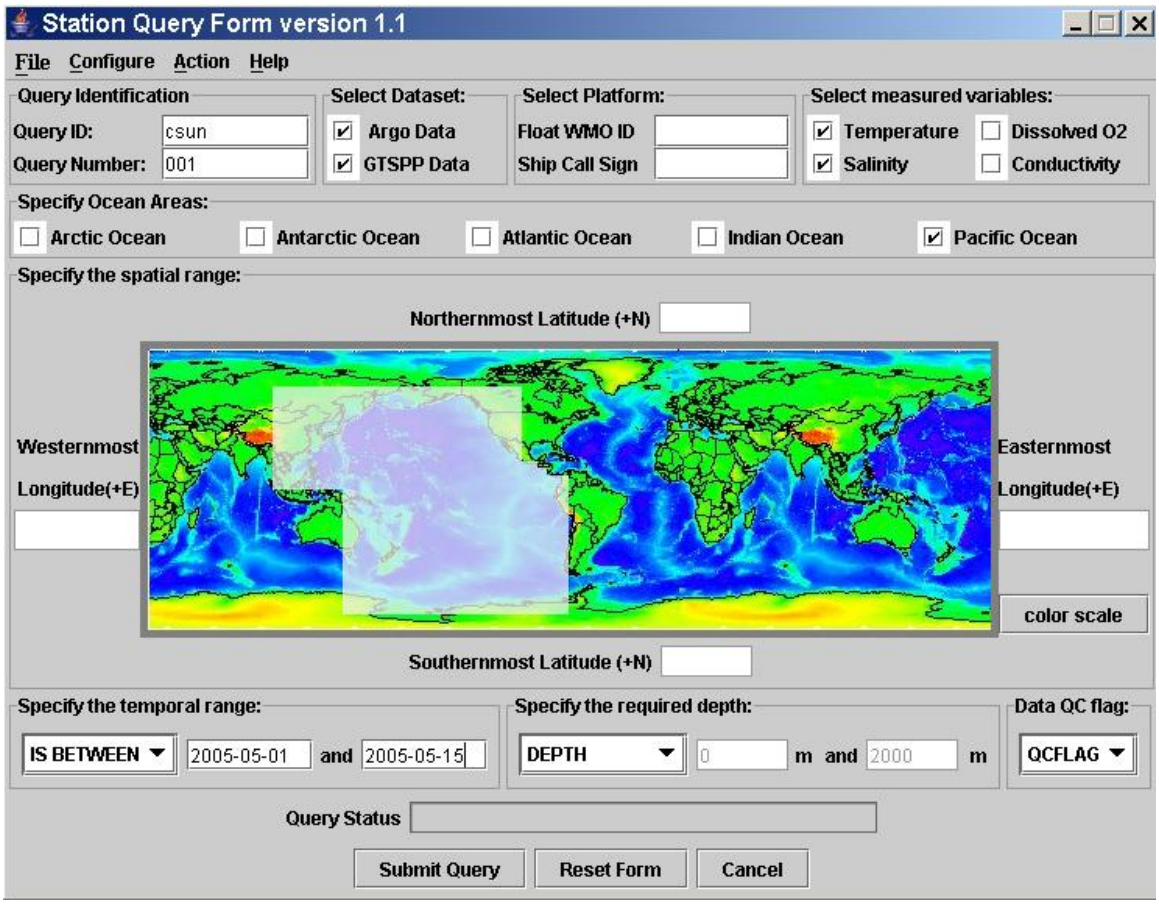

Figure 5. The Main Station Query Form of the Ocean Data Explorer.

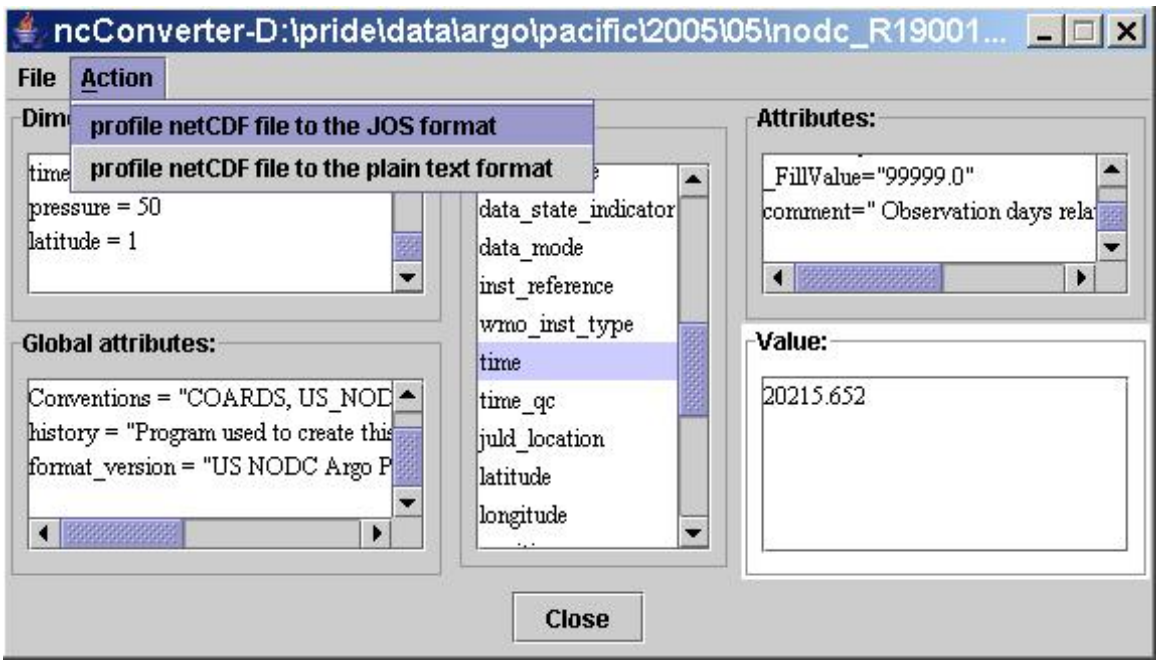

Figure 6. Screen shot of the “ncConverter” program for converting data from the netCDF format to the plain text format.