P1.8   SERVING THE NODC/OCL WORLD OCEAN DATABASE
VIA THE LIVE ACCESS SERVER (LAS)

Joseph McLean *
JISAO/University of Washington, Seattle, WA

1. INTRODUCTION

Fundamental to the design of the Live Access Server (LAS) is its ability to provide browse services (subsetting and visualization) for databases of in-situ observations (non-gridded data). Over the past two years, we have turned or attention to further developing that capability. Configuring the LAS to access such databases and generate products useful to the in-situ community involves changes to the LAS User Interface, data access, and Product Generation machinery. This document will explain these changes and provide examples of LAS capabilities by exploring the specific customizations used to provide access to the NODC World Ocean Database.

2. THE LIVE ACCESS SERVER

The Live Access Server[1] is a Web-server designed to provide access to geo-referenced scientific data. In its basic form, a user is presented with an interface through a web browser with which he can select the data set(s) of interest, variables, the desired region in longitude-latitude-vertical-time (X-Y-Z-T) and an output product. When the user requests a result LAS then accesses the data (locally or remotely via OPeNDAP) and launches a "back-end" application whose output is returned to the user's browser.

3. THE WORLD OCEAN DATABASE

The National Oceanographic Data Center/Ocean Climate Laboratory (NODC/OCL) World Ocean Database 2001[2] (WOD01), M. E.

* *Corresponding author address*: Joseph McLean, JISAO/PMEL, Bldg #3 7600 Sand Point Way NE, Seattle, WA, 98115; e-mail: joe.mclean@noaa.gov

* *with* Steve Hankin, Jonathan Callahan, Roland Schweitzer, Kevin O'Brien, and Ansley Manke

Conkright, (2002), is a collection of ocean profile data comprising over 7 million stations and hundreds of millions of individual data measurements.

The data include metadata such as Location, Cruise ID, Project affiliation, Institution affiliation, Quality flags, etc. Standard presentation of the data is a set of eight CD-ROMS containing OCL formatted ASCII files representing nearly 14 gb of data.

4. REFORMATTING THE WORLD OCEAN DATABASE

We approached the problem of accessing the OCL formatted data with a requirement that the X-Y-Z-T region must always be specified. To support high speed access of this type we reformatted the data into a 3-tiered hierarchy of NetCDF[3] (Network Common Data Format) files. NetCDF was chosen because of the readily available libraries and interface and the experience of the LAS development team in working with NetCDF.
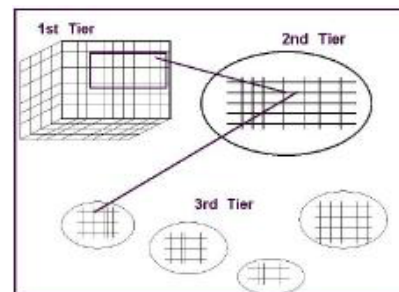

Figure 1 - Concept of reformatted in-situ Database

- The 1st tier is an XYT gridded data set containing pointers (indices) to the first element of the 2nd tier blocks (described below), and the number of profiles in each block.

- The 2nd tier of the hierarchy consists of station or profile metadata in 1-dimensional fields. These fields include the latitude, longitude, time, and other parameters associated with a

specific profile. Each of the 2nd tier blocks consist of the metadata of profiles within 1 WMO geo-spatial 10X10 degree square and within the range of one axis cell from the 1st tier data T-axis. If the 1st tier request is for all profiles lying within 0N, 10N, 130W, 120W, and recorded between 1910 and 1960 the corresponding side-by-side 2nd tier blocks are collected for further filtering on depth.

- The 3rd tier is a collection of NetCDF files containing the depth and values of the data measurements. Each of the 3rd tier blocks consists of the depth and data measurements of one profile. This tier is accessed using pointers (indices) and number of depth levels stored in the 2nd tier, the same way that the 2nd tier metadata is accessed by the 1st tier pointers.

5. LAS DBM ACCESS

Perl modules have been written for the LAS to facilitate access to the in-situ data in the above Database as well as MySQL and JGOFS[4] (Joint Global Ocean Flux Study) Relational Database Management Systems. Besides filtering on XYZT parameters, the LAS can also filter on other profile metadata values available in the WOD, such as Probe Type and Institution Affiliation. The data is collected and stored in a series of arrays, which are then written to a temporary NetCDF file. The file is non-gridded, and it contains the Longitude, Latitude, Time, and Depth coordinates, for each measurement value. It also contains the total number of Profiles, the total number of observations, and a Profile index needed to generate a product.
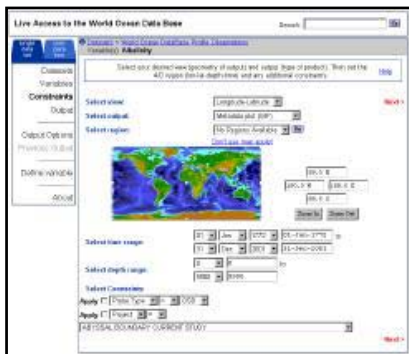
6. LAS IN-SITU USER INTERFACE

Figure 2 - The WODB LAS UI Constraints page.

The LAS User Interface is configurable through XML files. For the WODB LAS, custom views (slices of data in which a user is interested, i.e. XY, YT, XZ, Time Series) and custom products have been added to the UI selection menus.

Additional constraint menus, in which a user can request a Probe Type or Institution Affiliation filter are also available. The XML configuration of the LAS sets up the User Interface, and the custom DB-access Perl modules instruct the LAS how to do the filtering. An Options page for adding additional parameters to the data request is also available and configurable.

7. LAS PRODUCT GENERATION

The default back-end visualization and analysis application for the LAS is Ferret[5]. Although Ferret, Hankin, S. (1996), is designed to work with gridded data, some effort has been put into creating scripts which handle the formatted NetCDF files which the Perl DB access code has generated. Various plots are generated on-the-fly and returned to the user's browser.

- The Metadata plot shows the locations of Profiles within the X-Y-T range selected by a user. The access speed of the reformatted database enables the collection and plot generation of thousands of profiles in seconds. The plot below was returned to the browser in less than 10 seconds after the request and shows the locations of over 49,000 profiles.
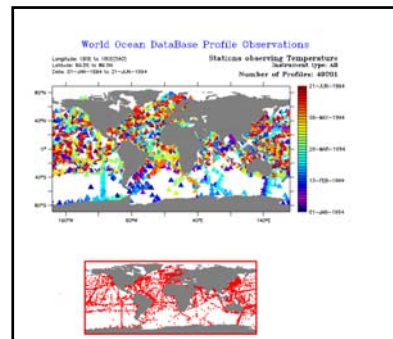

Figure 3 - The in-situ Metadata plot showing locations of over 49,000 Profiles.

- The Gaussian filled plot performs interpolation operations and returns an anlayzed view of the data.
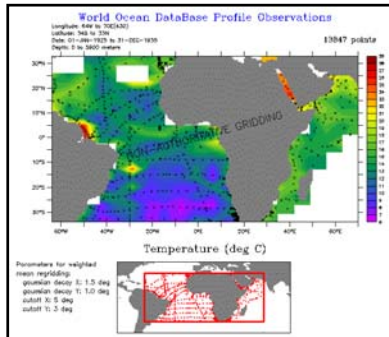

Figure 4 - The in-situ Gaussian-filled plot

- Many of our customers using in-situ data are interested in observational statistics. The Time Series view offers a Number of Profiles per Month plot.
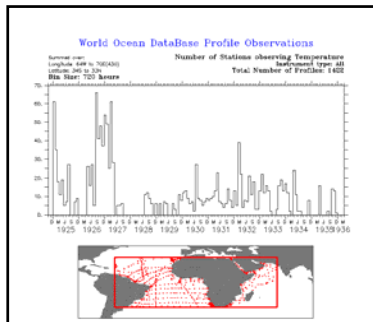

Figure 5 - The Number of Profiles per Month plot

The configurable Options page enables the user to choose a bin size other than Month.

8. SINGLE PROFILE ACCESS

The LAS WODB LAS is designed to present a familiar User Interface in which in-situ data subsets are chosen in much the same way that gridded data subsets are chosen. A region in time and space is requested and the LAS pulls that subset from the database before doing any work with the data. If additional Profile constraints are requested, then only those data are pulled from the database.

We are currently experimenting with interfaces that allow the end-user to access single profiles. Initially, a subset is described as above by the user. Then an intermediate interface is generated on-the-fly which provides access to single profiles. This can either be a hypertext list of profiles or a clickable map.

9. CONCLUSION

The development of the WODB LAS has shown the usefulness of LAS as an access provider for large collections of in-situ data. The continuing development of in-situ LAS capabilities must focus on an all-encompassing picture of the needs of the scientific community. Among these are the inclusion of more filtering request parameters such as Quality Flags and Cruise ID's. The ability to integrate data access for in-situ and gridded data will provide a much needed tool for comparing the two types of data and validation of model data. The recent push for a better understanding of oceanographic observational statistics demands more practical access to large collections of in-situ data. The WODB LAS is a tool which demonstrates practicality and speed needed for that access.

10. ACKNOWLEDGEMENTS

11. REFERENCES

M. E. Conkright, J.I. Antonov, O. Baranova, T. P. Boyer, H.E. Garcia, R. Gelfeld, D. Johnson, R.A. Locarnini, P.P. Murphy, T.D. O'Brien, I. Smolyar, C. Stephens, (2002):
World Ocean Database 2001, Volume 1: Introduction. Ed: Sydney Levitus, NOAA Atlas NESDIS 42, U.S. Government Printing Office, Washington, D.C., 167 pp.

Hankin, S., Sirott, J., Callahan, J., O'Brien, K., Manke, A. (2002):
Live Access Server: A Toll for Web Access to In Situ Data Collections, Too

Hankin, S., D.E. Harrison, J. Osborne, J. Davison, and K. O'Brien (1996): A strategy and a tool, Ferret, for closely integrated visualization and analysis. Journal of Visualization and Computer Animation, 7, 149–157.

---

[1] Information about the **LAS** can be found at
http://www.ferret.noaa.gov/Ferret/LAS/

[2] Information about the **WOD01** can be found at
http://www.nodc.noaa.gov/OC5/WOD01/

[3] Information about **NetCDF** can be found at
http://www.unidata.ucar.edu/packages/netcdf/

[4] Information about **JGOFS** and the JGOFS LAS
can be found at
http://www.ferret.noaa.gov/jgofs

[5] Information about **Ferret** can be found at
http://www.ferret.noaa.gov/Ferret