13.2 THE INTEGRATED DATA VIEWER – A WEB-ENABLED APPLICATION FOR SCIENTIFIC ANALYSIS AND VISUALIZATION

Don Murray*, Jeff McWhirter, Stuart Wier, Steve Emmerson
Unidata Program Center, Boulder, Colorado

1. INTRODUCTION

In September 2002, the Unidata Program Center released the first version of a new, freely available visualization and analysis framework called the Integrated Data Viewer (IDV). (http://my.unidata.ucar.edu/software/metapps) The IDV distribution includes a reference application that combines the framework components to produce an application that is tailored to the atmospheric science community. This application allows a user to display and work with satellite imagery, gridded data (primarily from model output), surface, upper air and radar data within a unified interface. (Figure 1)

Built entirely using the Java language and on top of the VisAD analysis and visualization package (Hibbard, 1998), the IDV reference application provides many of the standard 2D data displays that atmospheric scientists are used to, as well as 3-D views of the atmosphere. It allows users to interactively slice and probe data to create cross-sections, profiles, animations and value readouts of multi-dimensional data sets. Using the power of the data model in VisAD, IDV allows users to perform calculations on disparate data sets without having to convert from one format to another.

While it can be used as a stand alone application for viewing local datasets, many features have been built in to allow it to work better in a networked environment. These include client/server access to data, distributed configuration files, an integrated HTML viewer and the use of Java Web Start™. (Figure 2)

2. CLIENT/SERVER DATA ACCESS

To take advantage of network access, on of the most important features is the ability to locate and work with geo-referenced datasets located on remote servers. As an increasing number of datasets are made accessible to the atmospheric science community (e.g., data from new sensors,

Corresponding author address: Don Murray, Unidata/UCAR, PO Box 3000, Boulder, CO 80307; e-mail <dmurray@unidata.ucar.edu>

local meso-scale models, higher spatial and temporal resolution data), it is becoming impractical to distribute these datasets in real-time using conventional methods like Unidata's Internet Data Distribution (IDD) system (Baltuch, 1997)

Digital libraries are indexing references to Earth System data resources, including real-time and archived data sets. Through the framework developed by the NSDL-funded Thematic Real-time Environmental Distributed Data Services (THREDDS) project, the IDV can access and display these data sets (Domenico, 2002). THREDDS embodies a network of thematic servers that provide educators and students coherent access to a large collection of distributed data (data web). At the heart of the THREDDS framework are metadata catalogs that are used by data providers to describe and publish their datasets. These catalogs provide discovery and usage metadata and can be indexed by a digital library like the Digital Library for Earth System Education (DLESE). Using the metadata catalogs, the IDV provides users with options of selecting entire data sets, or just small subsets of data (spatial and temporal). The integrated nature of the IDV allows users to bring in data from multiple servers through a variety of protocols (e.g. FTP, DODS, ADDE, HTTP) and combine them into a single display. It provides a powerful mechanism for using real-time and archived data in a classroom setting.

3. XML CONFIGURATION AND PERSISTENCE

Much of the IDV, including the user interface, is configured through a set of eXtensible Markup Language (XML) files. Developers and users can easily tailor the IDV to provide custom applications based on the underlying framework by using their own set of configuration files. Everything from map projections to color tables to the available display components is configurable. These configurations can be local files on the user's system or located on a remote Web server and loaded in when the application is started.

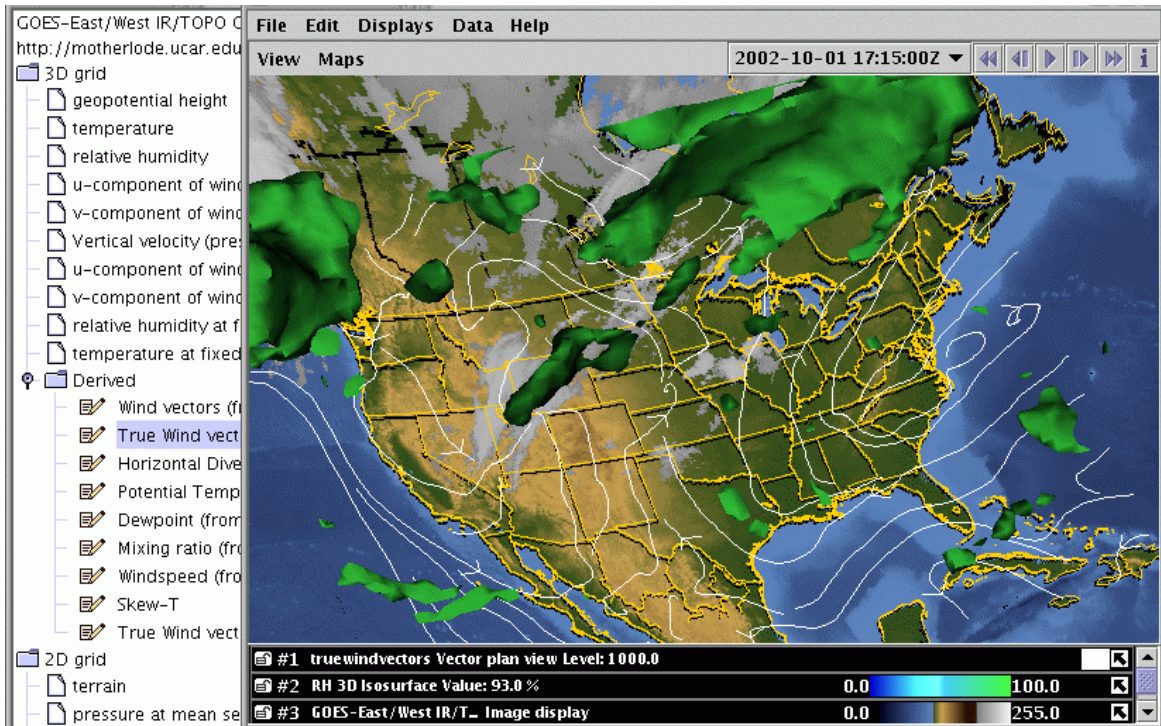The IDV also provides an XML-based mechanism for saving the state of the

Figure 1. IDV reference application showing displays of satellite imagery, wind streamlines and isosurface of relative humidity
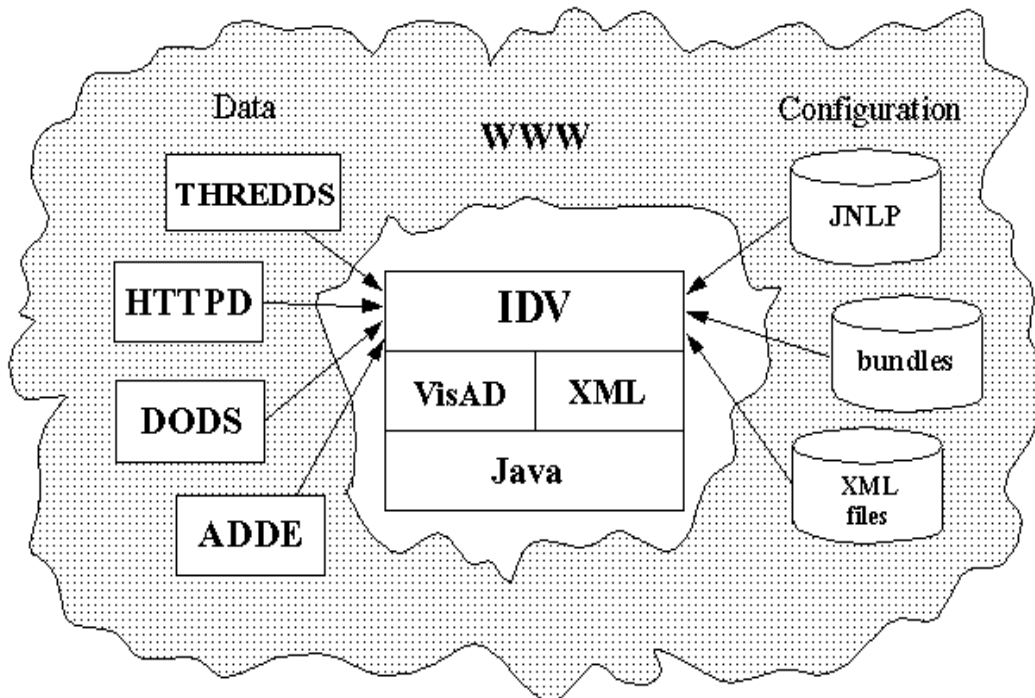

Figure 2. IDV can use the Internet for data access and configuration

Feature	Description
Integrated displays of a variety of data types	Grid, image, radar, sounding, surface observations
Support for a variety data access methods	Directly reads netCDF, Vis5D, McIDAS AREA and map (OUTL*) files, and ESRI Shapefiles. Provides access to other data formats (e.g. GINI, NEXRAD Level III, HDF) through ADDE and DODS
Multiple display types	Displays 3-D Isosurfaces, contours, color-filled contours, pseudo-color images, station plots, Skew-T, line plots
Interactive probes	Data readouts, vertical profiles from gridded data, time series, range and bearing
User defined formulas	Supports users creation of derived products from a library of mathematical and meteorological operations
Bundling of user preferences	Customization for specific uses through "bundles". Users can share saved data views, complete with annotations
Integrated HTML viewer	Allows user interface customization through HTML forms, creation of pedagogical materials for classroom exercises
Easy configuration	Uses XML files for configuration of UI and functionality.
Integrated documentation	Access to help and tutorial from the application

Table 1. Integrated Data Viewer features

application.   The state (loaded data sources and data depictions) can be saved to an XML "bundle" at any time.   These bundles can be used to pre-configure the application, or can be loaded interactively while the IDV is in use. Since they are text-based documents, the bundles can be distributed as e-mail attachments or accessed from a link on a Web server. This allows for the ready sharing of interesting collections of data/visualizations, case studies and pedagogic materials over the Internet.

For example, a researcher on a field project could load in a dataset that is accessible from a THREDDS catalog and provide a customized view of the data that would illustrate some scientific principle. Using the annotation capabilities in the IDV, she could provide an accompanying textual description of the particular view, and/or annotate the display directly. She can save that view in a bundle and e-mail the XML bundle to a colleague. Her colleague can start the IDV and load in the bundle to see the same view that the sender had, complete with annotations.

4. HTML VIEWER

Along with traditional sources of data, the IDV also integrates HTML documents and access to web servers and provides a unique HTML-based method for creating customized user interfaces.

Web pages can be viewed in the special HTML viewer component.   This viewer can interpret special embedded links to IDV displays, controls and IDV bundles within a standard HTML document. While these special tags are ignored in a typical Web browser like Netscape Navigator, the custom viewer interprets these tags and knows how to display the IDV components.   This allows an educator to create an entire class exercise in HTML, including powerful interactive, three-dimensional displays of remote data (Figure 3).

The HTML viewer can also be used to create web pages that control the main IDV, thereby creating a customized user interface. Since the viewer can display standard HTML forms, a simple user interface to load in a predefined data source and customized view can be created in such a form.

Examples of both kinds of HTML documents are   available   on   the   Unidata   Web   site
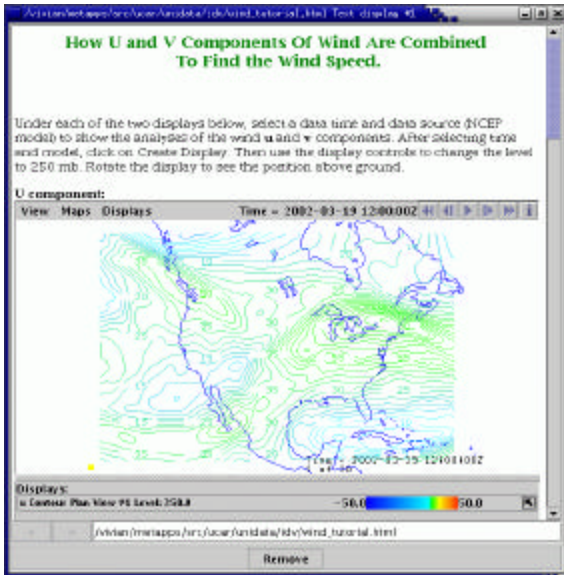
Figure 3. IDV display embedded in an HTML document

(http://www.unidata.ucar.edu/projects/metapps/idv_
html). As users develop new examples, these can
be linked from this site or made available through
the Digital Library for Earth Science Education
(DLESE).

5. WEB START DISTRIBUTION

The IDV takes advantage of Java's Web Start
(http://java.sun.com/products/javawebstart/)
technology to provide a seamless method for
internet-based software distribution. Using Web
Start, the necessary Java Archive (JAR) files can be
staged on a web server. The application can be
launched from a web link pointing to a Java Network
Launching Protocol (JNLP) file, which provides
configuration information for the application. When
a user clicks on the link, any JAR files or other
ancillary data that has changed since the last time
the application was launched are automatically
downloaded and installed on the user's system.
This capability provides a quick and easy approach
for the administration of applications and will be a
definite plus for general deployment of Unidata
software in the future.

Because the JNLP file can specify additional
configuration information besides the JAR files,
users can create custom JNLP files that will
automatically configure the application with custom
XML configuration files and preload data sources
and displays from bundles that are to be used in a
particular session. For example, an educator may
want to control the types of displays that are
available to the students as they progress through
an inquiry-based curriculum. In the early portion of
the curriculum, the educator may want to introduce
the concept of contour lines and have that be the
only type of display available to the student. As the
student progresses through the curriculum,
additional types of controls can be introduced to
provide displays of color-filled contours and 3D
isosurfaces. A customized JNLP file that points to
different configuration files can add each of these
features.

6. APPLICATION IN OTHER GEOSCIENCES

Although the IDV reference application was
designed upon the needs of the existing Unidata
atmospheric sciences community (Murray, 2001),
the framework tries to be agnostic in terms of the
data it can handle and it's context of use. Some
components are geared towards geo-referenced
data sets, but these need not be atmospheric in
nature. With the underlying general VisAD data
model, users in other geoscience disciplines (e.g.,
hydrology, oceanography) can also use the IDV for
analysis and visualization of their data sets.

The VisAD data model was designed to support
virtually any numerical data. Rather than providing a
variety of specific data structures like images, grids
and tables, the VisAD data model defines a set of
classes that can be used to build any hierarchical
numerical data structures. Metadata including
units, sampling topologies, error estimates and
coordinate transformations is incorporated into each
data object and is carried along through
mathematical operations. The IDV incorporates any
metadata available in the dataset or additional
metadata supplied by a THREDDS server.

For example, a researcher who wants to
calculate the air-sea temperature differences over a
portion of the Gulf Stream could use the IDV to find
a THREDDS data server that has an atmospheric
boundary layer temperature data set and another
server that has sea surface temperatures as part of
an oceanographic data set. He could use the
diagnostic capabilities of the IDV to subtract the
two datasets. During the process, the
temperatures from the second dataset are
resampled to the spatial and temporal locations of
the first before they are subtracted. If the two data
sets use different temperature units, these are
converted before values are subtracted.

As the Internet enables greater data sharing among scientists, it increases the problems associated with metadata and file format differences among scientists. Use of the VisAD common data model and THREDDS catalogs by the IDV mitigates this problem.

7. SUMMARY

Unidata's Integrated Data Viewer is a powerful visualization and analysis tool that is freely available to the geoscience community. It's web-enabled features help facilitate research and education by integrating diverse datasets, allowing customized user experiences and enabling collaborations.

8. REFERENCES

Baltuch, M. S., 1997:   Unidata's Internet Data Distribution (IDD) System: Two Years of Data Delivery. *Preprints*, *13th Intl Conf. on IIPS for Meteorology, Oceanography and Hydrology*, 478-480.

Domenico, B., J. Caron, E. Davis, R. Kambic and S. Nativi, 2002: Thematic Real-time Environmental Distributed Data Services (THREDDS): Incorporating Interactive Analysis Tools into NSDL, *J. Digital Information,* **2**(4).

Hibbard, W., 1998: VisAD: Connecting people to computations and people to people. *Computer Graphics,* **32**(3), 10-12.

Murray, D. and C. Murphy, 2001: Community Software Design – The Unidata Java-based MetApps Project. *Preprints, 17th Intl Conf. on IIPS for Meteorology, Oceanography and Hydrology*, 367-368.