

Sample Data Interactive Publication:

Interactive access, analysis and display of remotely stored datasets from a convective rain event over Colorado

Ben Domenico

Jeff Weber

**Unidata Program Center
UCAR
Boulder, CO 80307-3000**

Abstract

This paper is an example of what we call data interactive publications. With a properly configured workstation, readers can click on "hotspots" in the document which bring up an interactive analysis tool called the Unidata Integrated Data Viewer (IDV). The IDV will enable the readers to access, analyze and display datasets on remote servers as well as documents describing them. Beyond the parameters and datasets initially configured into the paper, the analysis tool will have access to all the other dataset parameters as well as to a host of other datasets on remote servers.

These data interactive publications are built on top of several data delivery, access, discovery, and visualization tools developed by Unidata and its partner organizations. For purposes of illustrating this integrative technology, an atmospheric/hydrologic case study has been chosen: a July 16, 2004 convective rain event over Colorado, which illustrates how components of this process fit together. The Local Data Manager (LDM), Open-source Project for a Network Data Access Protocol (OPeNDAP) and Abstract Data Distribution Environment (ADDE) services, Thematic Realtime Environmental Distributed Data Service (THREDDS) cataloging services, and the IDV are highlighted in this example of a publication with embedded pointers for accessing and interacting with remote datasets.

An important objective of this paper is to illustrate how these integrated technologies foster the creation of documents that allow the reader to learn the scientific concepts by direct interaction with illustrative datasets, and help build a framework for integrated Earth System science.

Notes on the underlying technology

The technology used in this publication enables readers to interact directly with datasets and catalogs of datasets, as well as documents stored on remote servers. To utilize the

technology effectively, it's necessary to have a well-configured desktop workstation (at least 512 MB of memory) with the Java Webstart application installed.

The images and animations that show up on the screen via the Unidata Integrated Data Viewer (IDV) are not just pictures; they are interactive visualizations generated from the datasets, portions of which are accessed from remote servers or from the disks on your own desktop workstation. Rather than viewing images generated by clicking on a web page, the processes involved in this publication are more akin to those one would use to access research datasets from data providers, convert those datasets into a format that is compatible with a data analysis and display tool at your local site, then using that tool to work with the data interactively.

In this sample case, the datasets you start with are the output of forecast models:

- the Eta model (Grid #211, 80 Km resolution) which is run every six hours at NCEP (National Centers for Environmental Prediction)
- a high resolution (4 Km), local version of the model called the Workstation Eta run at the Unidata Program Center (UPC).

The NCEP model output is approximately 52 MB in size. The Workstation Eta covers less area and time but is at a much higher resolution which results in an output file that is over 800 MB—i.e., nearly a Gigabyte of data. These datasets are stored on a computer at the National Center for Atmospheric Research; the IDV interacts with them via established client/server protocols. A key element of the client/server technology used in these documents and visualizations is that only the needed parts of the huge files are transferred as they are requested. In spite of that, it takes a while to transfer even that subset to render the images on your workstation.

Since these are not pre-generated pictures, you have access to the complete IDV visualization system and to all the data on the server; thus you can interact directly with the data to create your own visualizations. In fact, you can access other datasets and include them in the analysis.

Just Do It

If you don't already have Webstart installed, the instructions for doing so can be found at [The Unidata WebStart Instruction Page](#).

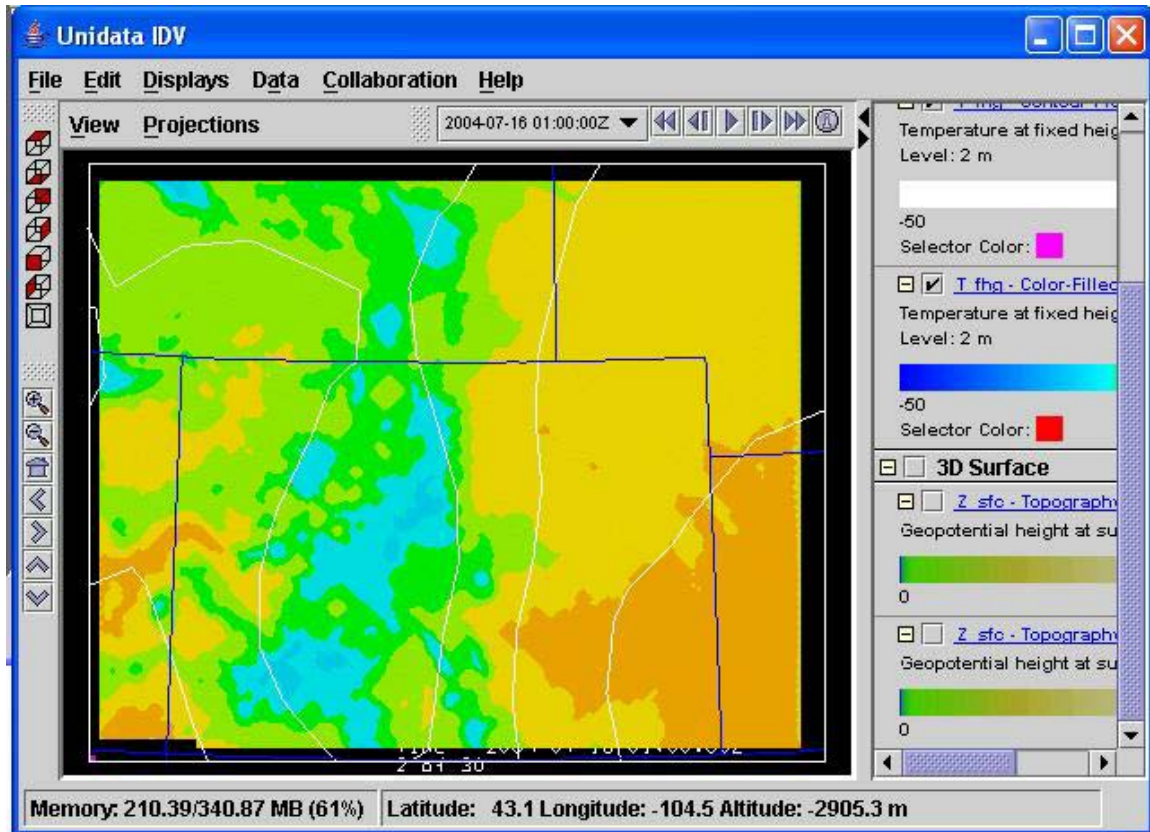
Once webstart is running on your workstation, the IDV configured with the appropriate datasets can be started by clicking on the following line:

[Temperature forecasts from NCEP national forecast and high-res. local Colorado Workstation ETA.](#)

What you see (initially): temperature forecast comparisons

These forecast models were run at nearly the same time. In fact the NCEP Eta model high resolution output (12 Km) was used to initialize the higher resolution Workstation Eta model which covers only the Colorado area in a nested grid.

The forecasts were for July 16, 2004, a day in which there was significant weather activity through Colorado. The visualization shown in this bundle compares the predicted temperatures at 2 meters above the surface. If you are unable to get the IDV running on your workstation, the screen dumps below give you a sense of the kind of visualization that would be displayed.

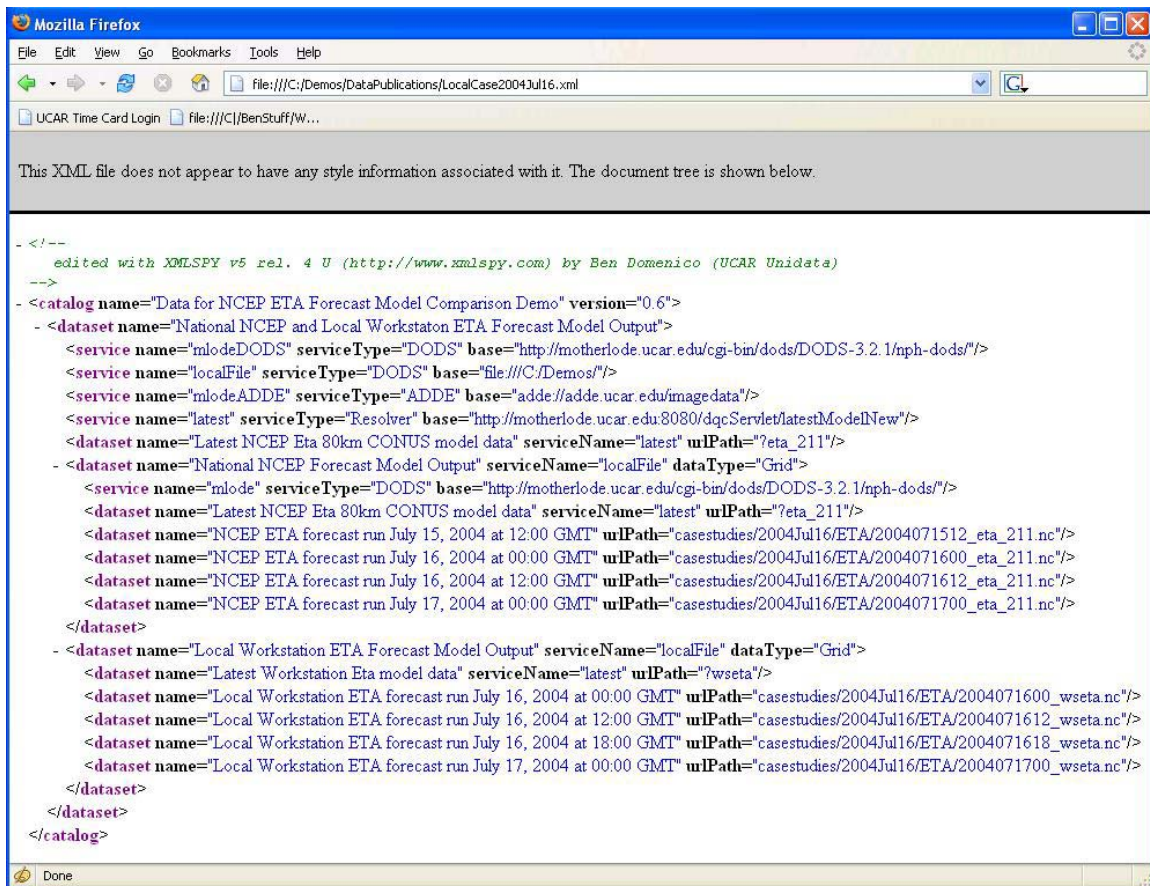


The Eta model output is shown as contour (plan view) white lines whereas the higher resolution Workstation Eta model output is shown as color filled contours. One can readily see the detail in the Workstation Eta model that is not visible in the NCEP model which covers most of North America. Using the controls in the upper right corner of the display window, the visualization can be animated through the entire forecast time. This animation reveals the higher temporal resolution of the Workstation Eta model since the color fill image changes for some frames where the white line contours of the national model stay constant because there is no national forecast for those intermediate times. The frames toward the end only show the NCEP model because the Workstation Eta forecast only goes out 24 hours in time.

Integrated text, tools, and data

As you will see, HTML files (such as this one) are also viewable within the IDV which accesses it from a Web server at the Unidata Program Center (UPC)--a different server from the one that provides the data. The catalogs that contain the URL pointers to the data files can be stored on still another network computer or on one's own workstation for that matter. In fact, in this example, the reader has two different views into the same HTML file. One is seen from within a Web browser and is used to invoke the IDV for interactions with the datasets; the other is actually seen from within the IDV itself. While this may be somewhat confusing at first, it illustrates the flexibility of this approach to integrating web-based documents with tools for interactive analysis and display of environmental datasets on remote servers.

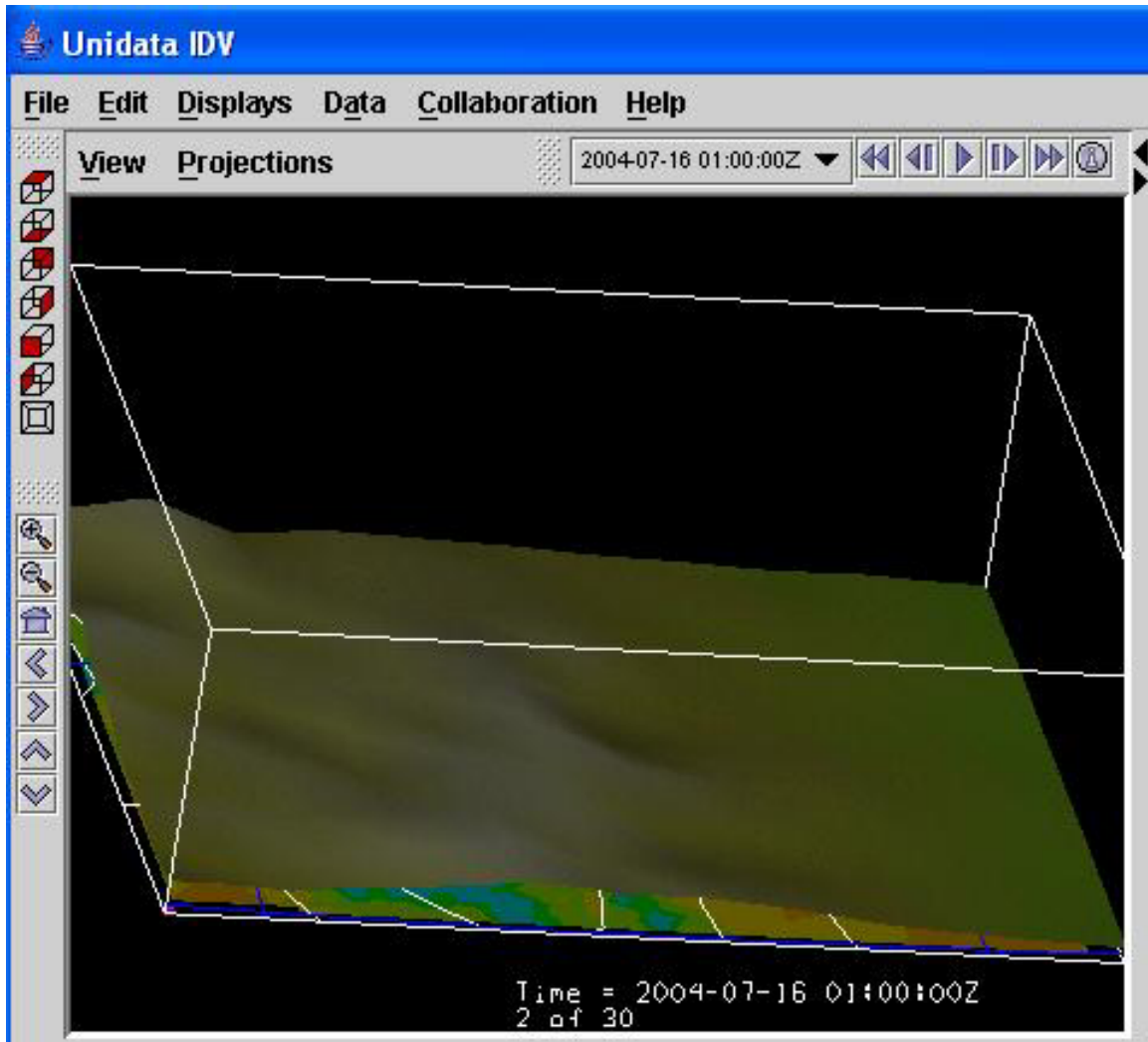
The catalogs pointing to the datasets are in the XML language. For those who understand XML or are curious about it, an [example THREDDSData catalog](http://my.unidata.ucar.edu/content/projects/THREDDSDataPublications/Case2004Jul16.xml) can be viewed at: <http://my.unidata.ucar.edu/content/projects/THREDDSDataPublications/Case2004Jul16.xml>



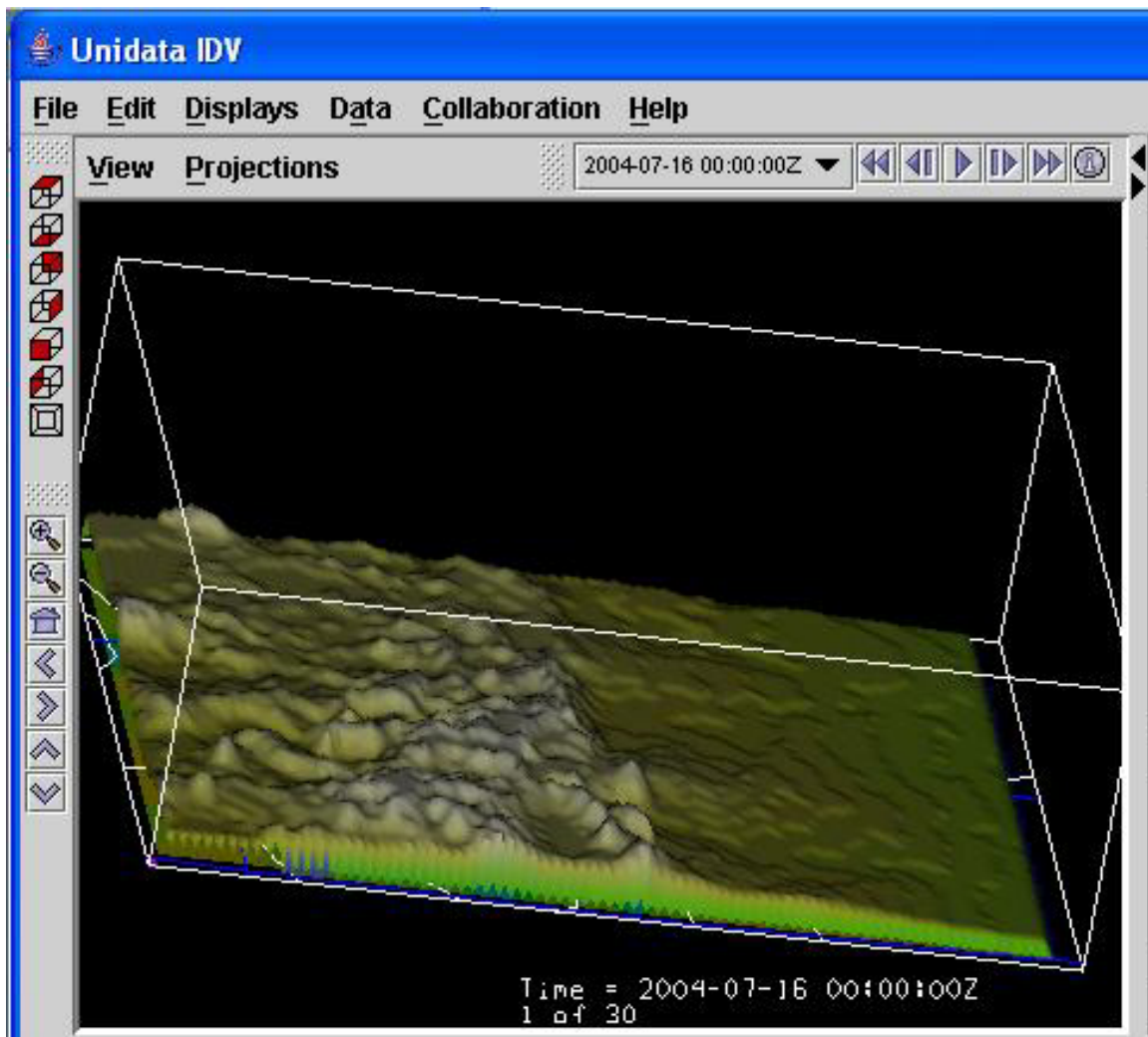
Topology from geopotential height at surface

Two additional displays are available but inactive in the IDV as it comes up. These are images of a parameter called geopotential height at the surface which serves as a "proxy" for the topography underlying the model. By clicking on the control lines at the bottom of

the image window, either one or both can be made visible. One of these is for the NCEP national model; the other is for the local workstation model. Comparing these images by turning them on and off provides some sense of the difference in resolution between the NCEP Eta and the Workstation Eta.



The image above provides a sense of the resolution of the topography in the NCEP national Eta model whereas the one below shows the increased detail of the Rocky Mountains in the local Workstation Eta model.



More information about forecast models

For more general background information, a concise description of the contents of various forecast models is in the NCAR Scientific Computing Division [Documentation for Online Model Output Datasets](http://dss.ucar.edu/datasets/ds335.0/docs/): <<http://dss.ucar.edu/datasets/ds335.0/docs/>>

A more all-encompassing description of datasets used in the atmospheric and ocean sciences is available in an NCAR tech note in [An Introduction to Atmospheric and Oceanographic Datasets](http://www.cgd.ucar.edu/cas/tn404/): <<http://www.cgd.ucar.edu/cas/tn404/>>.

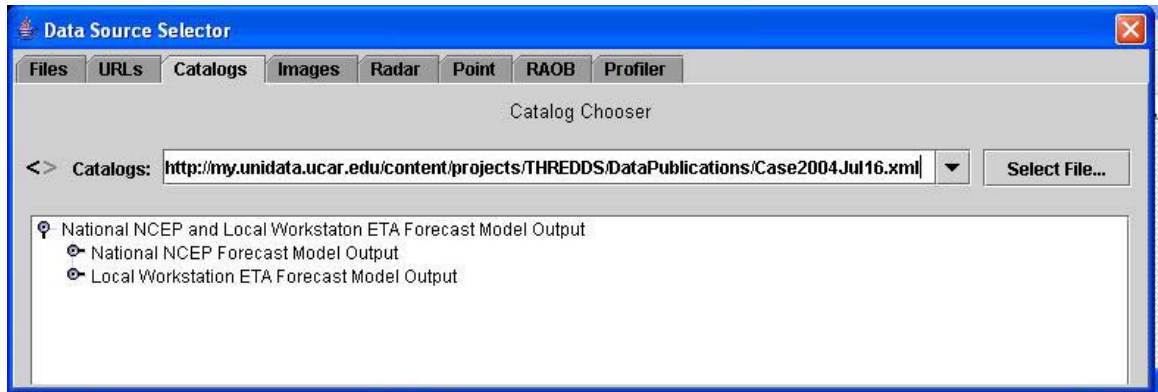
A somewhat outdated overview of the Eta model is given in [A Description of the Meso ETA Model](http://nimbo.wrh.noaa.gov/wrhq/96TAs/TA9606/ta96-06.html): <http://nimbo.wrh.noaa.gov/wrhq/96TAs/TA9606/ta96-06.html>.

The Workstation Eta is documented at the [Workstation Eta Homepage](http://wwwt.emc.ncep.noaa.gov/mmb/wrkstn_eta):
http://wwwt.emc.ncep.noaa.gov/mmb/wrkstn_eta.

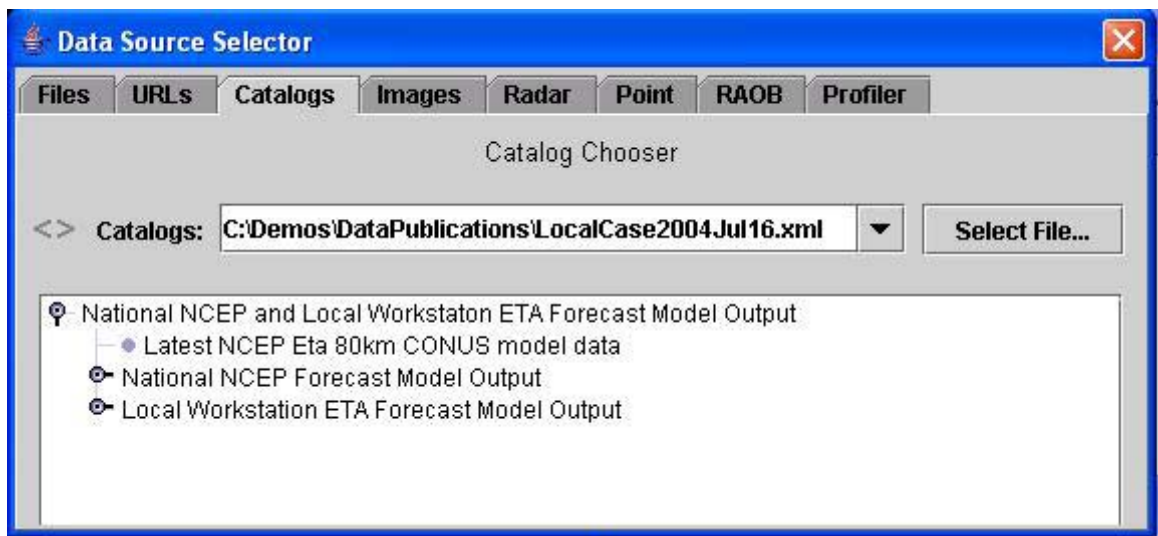
Exercises

One of the main goals of this mini case study is to provide an entry point for learning more about the capabilities of the client/server technology that makes it possible to interact directly with the datasets to learn a bit about the science behind the models in the process. The following activities should help in that process:

1. Working with the animation controls in the upper right of the image window provides a sense of how the different forecasts evolve in time.
2. Think about what sort of additional data would be useful to overlay to verify the accuracy of the predicted conditions.
3. To learn more about the science, the reader may wish to figure out why and how "geopotential height at the surface" serves as a proxy for the topography used in the forecast models.
4. To gain a better sense of the topologies, click on the #1 or #2 lines at the bottom of the image to turn on the topography displays; then hold down the right mouse button in the image window and move the mouse around to gain a 3D perspective on the image. If you rotate the image enough, you can actually view the temperature data from underneath the topography images. The next release of the IDV will allow you to "drape" the data over the topography.
5. To examine other runs of the forecast models and to visualize other parameters, a special THREDDS catalog of the datasets has been prepared in a [sample catalog](http://my.unidata.ucar.edu/content/projects/THREDDS/DataPublications/Case2004Jul16.xml) at:
<http://my.unidata.ucar.edu/content/projects/THREDDS/DataPublications/Case2004Jul16.xml>
6. You can view the XML source for that catalog by pasting the URL into your browser window; but more importantly, the catalog can be accessed from within the IDV by going through the Data -> New Data Source menus and selecting the "Cataloged Data" tab. Then paste the URL above as the "Catalogs" entry. Next, click on Update to gain access to all the national and local Colorado model run datasets saved in this case study. There you will find pointers to the latest run of both the national and local ETA model runs, so you can compare the current conditions to those in the case study collection.



It is worth noting that catalogs can be accessed from remote servers such as the entry in the above example from the my.unidata.ucar.edu web server and the one below that in a file on the local disk.



7. You can click on the icons at the beginning of each line in the catalog to open and close the lists within. Then you can highlight the model run of interest and click on "Add source" to make it available in the data selector where you can choose the Fields and Display you want to include in the visualization. Try a few different Fields and displays to gain a sense of how the IDV brings in subsets of data from remote servers and integrates it into the local interactive display.
8. A serious issue of concern is determining how much of the additional detail shown in the local Colorado model reflects reality so that it makes it possible to predict weather conditions more accurately for each locale in the state. Answering that question is an additional exercise for the reader.

References:

The Unidata WebStart Instruction Page

<http://my.unidata.ucar.edu/projects/metapps/webstart/IDV/index.html>

Temperature forecasts from NCEP national forecast and high-res. local Colorado Workstation Eta

<http://my.unidata.ucar.edu/content/projects/THREDDS/DataPublications/EtaWsEtaTemps.jnlp>

Documentation for Online Model Output Datasets

<http://dss.ucar.edu/datasets/ds335.0/docs/>

An Introduction to Atmospheric and Oceanographic Datasets:

<http://www.cgd.ucar.edu/cas/tn404/>

A Description of the Meso Eta Model

<http://nimbo.wrh.noaa.gov/wrhq/96TAs/TA9606/ta96-06.html>

Workstation Eta Homepage

http://wwwt.emc.ncep.noaa.gov/mmb/wrkstn_eta/

Example THREDDS catalog

<http://my.unidata.ucar.edu/content/projects/THREDDS/DataPublications/Case2004Jul16.xml>

Acknowledgment

This material is based in large part upon work supported by the National Science Foundation.