Enhancing Use of Surface Observations in IDV

A proposal to the Unidata Community Equipment Award Program

Unidata Program Center
ATTN: Community Equipment Awards
P.O. Box 3000
Boulder, CO 80307

Date: 3/22/07

Principal Investigator Name: John Horel
Title: Professor
Institution: University of Utah
Telephone number: (801) 581-7091
FAX number: (801) 581-4362
Street Address: 135 S 1460 E Rm 819, Salt Lake City, UT 84112

Email address: john.horel@utah.edu

Signature of PI:_____

Institution Official: Brent Brown
Title: Director, Office of Sponsored Projects
Telephone number: 801 581-4913
FAX number: 801 581-3007

Email address: brent.brown@osp.utah.edu

Signature of University Official: _____

1. Project Summary

Faculty, staff, and students in the Department of Meteorology at the University of Utah have actively participated in the Unidata program since its inception. An equipment award from the National Science Foundation in 1997 helped to develop the Meteorology Computation and Visualization Laboratory (CVL) in our department (see http://www.met.utah.edu/class/jimsteen/cvl/ . Since the 1997 award, University of Utah resources have been used to update and expand the lab, which now contains a total of 15 workstations and an instructor's console (Fig. 1). Server resources purchased with research funds are used as part of the Unidata Internet Data Delivery (IDD) system, including as a second level router of NEXRAD data to six downstream universities. These real-time data servers are 3-6 years old and in need of replacement.

Enrollment in several critical upper division undergraduate and graduate courses has increased in recent years such that the CVL is insufficient to meet demand. For example, our Spring 2007 mesoscale course has 25 students while the thermodynamics course has 20 students. There is insufficient space simply to move around the room, let alone allow students to work separately on the workstations. Since the CVL space is inadequate for classes with enrollment greater than 15, planning is underway to use a larger classroom housing more workstations to meet the needs for classes with up to 30 students.

Figure 1. Existing University of Utah Meteorology CVL.

Integrating automated surface observations from as many sources as possible with analysis and forecast model output, as well as satellite and radar imagery, is critical for students to visualize and understand the impacts of weather where it matters, the earth's surface. Access to surface observations from mesonets is particularly useful for assessing such weather impacts. In addition, surface analyses that incorporate mesonet observations are increasingly needed for educational, research, and forecasting applications (Horel and Colman 2005). The Real-Time Mesoscale Analysis (RTMA) developed by the National Centers for Environmental Prediction is now available for use in the classroom (Pondeca et al. 2007). RTMA grids of temperature, moisture, wind, pressure, precipitation, and cloud are available to the Unidata community through the Unidata CONDUIT distribution. The Integrated Data Viewer (IDV) supports displays of surface mesonet and RTMA surface analysis grids and provides an excellent framework within which to integrate all of the data streams. The PI participated in the Summer 2006 Unidata Workshop on *Expanding the Use of Models as Educational Tools in the Atmospheric & Related Sciences*, including presenting a tutorial to enhance students' understanding of data assimilation methods and ways to display mesonet data (see http://www.met.utah.edu/jhorel/unidata_2006.pdf).

MesoWest (http://www.met.utah.edu/mesowest/ , Horel et al. 2002) combined with the MADIS activities of ESRL (Miller et al. 2005) provide the infrastructure to access surface mesonet observations at over 15,000 locations throughout the country (Fig. 2). Although MesoWest and MADIS data streams are available via a number of protocols (e.g., web portals, ftp, LDM, MADIS via OPeNDAP), access to the MesoWest and MADIS data streams in IDV is awkward. There is no THREDDS/IDD archive in place to store and provide direct access to the data seamlessly to the Unidata community via IDV.

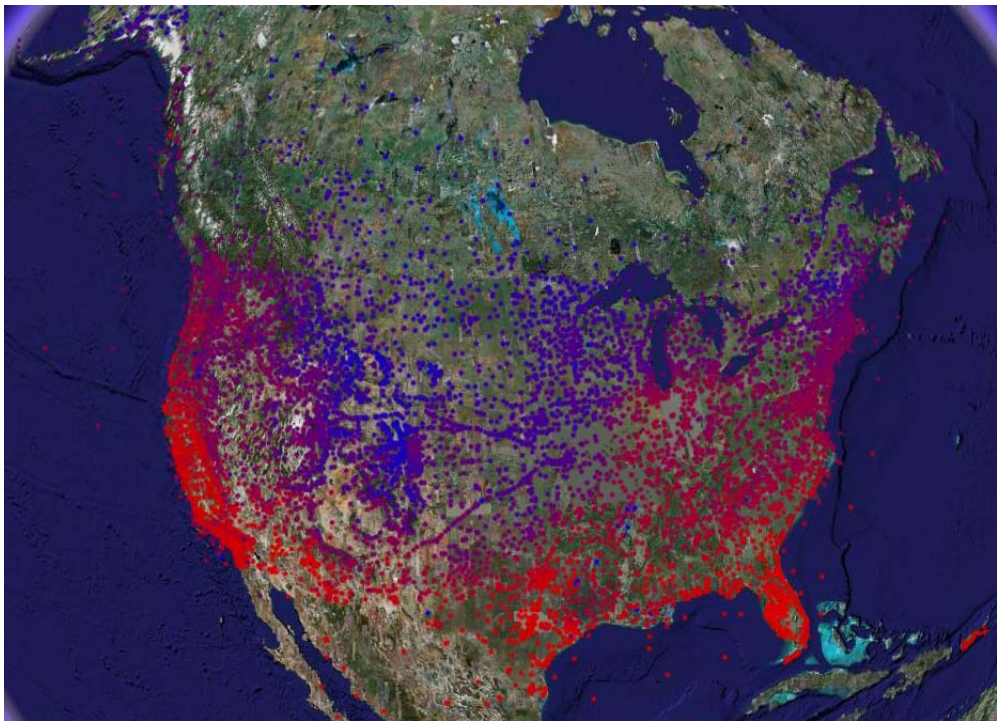

Figure 2. Stations available via MesoWest color coded by temperature (warm-red; cold-blue) and displayed using Google Earth software.

The proposed outcomes for this Unidata equipment award are:

- Replace our aging real-time data server to improve access at the University of Utah to resources available via IDD
- Add a THREDDS data server to provide real-time and retrospective access to surface mesonet observations for the entire Unidata community, as well as serve as a source for RTMA surface analyses during the most recent 2 months.
- Add 4 workstations in an expanded Meteorology CVL that will be available to students enrolled in laboratory classes at the University of Utah. Our students rely heavily on Unidata tools and resources in over a half dozen courses.

As will be detailed below, this project will help to: (1) make available surface mesonet and surface analyses that will be useful to the entire Unidata community; (2) enhance our ability at the University of Utah to participate in the IDD; (3) advance IDV capabilities; (4) improve local and community education and (5) lead to innovative research.

2. Project Description

a. Real-Time Data Server Upgrade

Our educational activities and research are hampered at the present time by the obsolete hardware available to retrieve, process, and archive the Unidata IDD data stream. While the hardware used by MesoWest and department web server has been upgraded recently, the two servers used for the IDD data stream have limited disk space, are prone to disk failure, and are slow. Since we are barely able to keep up with the data stream coming in, we have not advertised extensively our abilities to push MesoWest data via LDM or other means to the university community. *Hence, our obsolete real-time data servers limit our ability to fully participate in the Unidata IDD topology because of limited resources.*

While our IDD data server hardware is substandard, we are positioned well to be a more active IDD participant.due to our excellent overall computing infrastructure. The campus Center for High Performance Computing (http://www.chpc.utah.edu/) supports large bandwidth communications, including serving as the western node for the CRAFT Level-2 radar distribution. A computer room is available with adequate air conditioning and power and rack mounts for departmental servers.

We propose to upgrade our real-time data servers to improve our reception of the IDD data stream, improve our ability to store large data sets (e.g., model grids and Level-2 radar fields), and potentially serve as a more active participant in the IDD topology. We propose to purchase a data server of the following class for $6000:

- 2 dual core Opteron processors
- 16GB of ram
- 3ware raid cards
- 8 150GB, 10,000 RPM sata drives

Specific hardware details may vary depending on memory and CPU prices at the time of purchase.

b. Improved Access to Surface Mesonet Observations and RTMA Analyses in IDV

MesoWest (http://www.met.utah.edu/mesowest) began over ten years ago to integrate the collection, archival, and distribution of weather observations available from hundreds of sources around the nation (Horel et al. 2002). MesoWest is used extensively for operational, research, and educational use with specific applications developed for fire weather operations (http://raws.wrh.noaa.gov). This effort requires extensive coordination with the Meteorological Assimilation Data Ingest System (MADIS) staff at ESRL to manage metadata, acquire new data sources, and deliver data to government and other users.

As shown in Fig. 3, considerable infrastructure as part of MesoWest is already in place to provide access to current and retrospective surface mesonet observations. This effort requires extensive coordination as well with the IT staff at NWS Western Region. Data from many

different mesonets are received and input into the master database located at the NWS Western Region. The master database at the University of Utah and that at Western Region are continually synchronized through a dedicated communication link. Users query the database through client databases located both at Western Region and the University of Utah.

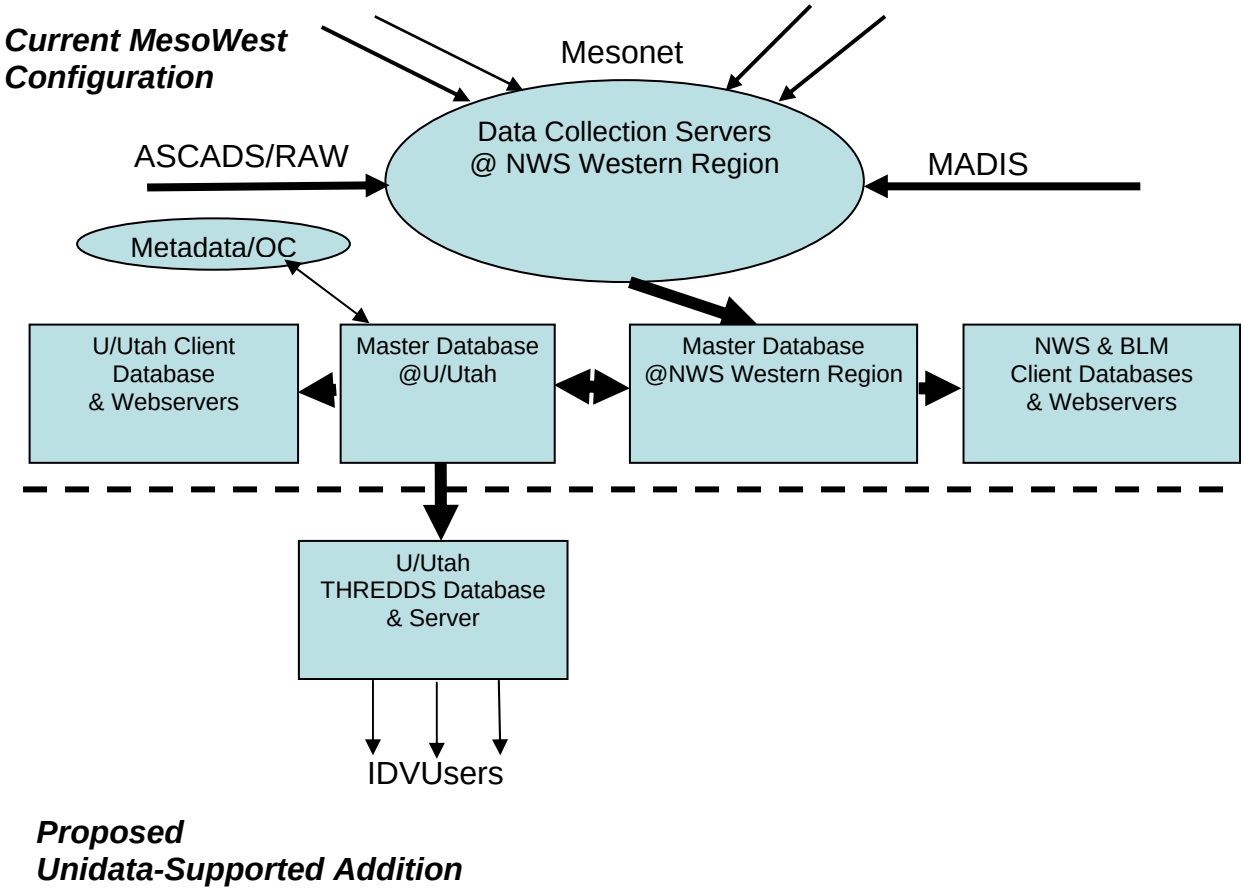

Figure 3. Considerable infrastructure is already in place (above the heavy dotted line) in order to support the collection, archival, and access to surface mesonet observations. The additional THREDDS database and server would allow IDV users to query the MesoWest database dynamically.

Usage of MesoWest by students at the University of Utah and elsewhere is extensive through the primary web portal (http://www.met.utah.edu/mesowest). Thousands of users access each day our web portal, which provides a convenient way to examine current weather, as well as retrospective case studies. We have added the ability to display graphical overlays of radar reflectivity and surface conditions (see Figs. 4 and 5). However, to fully integrate MesoWest data with other resources in IDV is not particularly straightforward at the present time. For example, NetCDF files of current conditions are available for use at the University of Utah but those files are not available to external users. NetCDF files can be obtained from MADIS as well via ftp, LDM or OPeNDAP. However, the time stamp convention used in these files results in the user managing hundreds of potential observation times during a day.

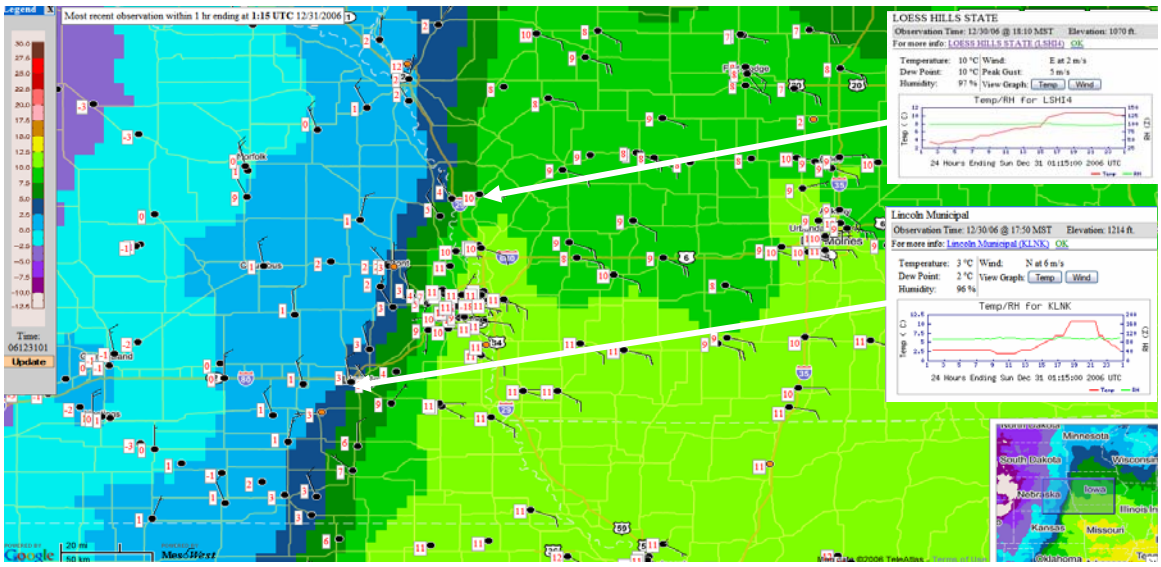

Figure 4. MesoWest temperature and wind observations across a strong frontal boundary in the central United States combined with the corresponding RTMA temperature analysis accessible via the MesoWest web portal. Similar capabilities are possible using IDV.

For efficient access to the mesonet data, we already support two types of queries. We store and support access to the most current observation from each station in a "Current" database table, which allows fast access to information when users are only interested in the current state. In addition, we support more complex queries (multiple times within a time interval, past times within the past 10 years, etc.) that span a variety of monthly (and yearly) database tables. These tables are subdivided into parameters to provide efficient access to the data (i.e., if someone is interested in precipitation it is not necessary to search through temperature and wind observations). Currently, the MesoWest database requires 160 gbytes of disk storage and is growing at a rate of roughly 30 gbytes per year (we have less complete data during the early part of the record).

Based on conversations with Unidata staff, we propose to use software under development by Unidata staff that will provide efficient access for the Unidata community to the entire 10-year MesoWest database. As shown schematically in Fig. 3, rather than storing static files in netCDF or some other format, we will allow users to query efficiently the MesoWest mySQL database dynamically using Java netCDF library tools. Software development is underway at Unidata to provide a short (3-day) archive of METAR observations in a database and a URL servlet interface that allows IDV users to query that database. The intention of the Unidata software developers is to provide a general station observation dataset adapter that would create the link from any data repository (such as MesoWest) using Java netcdf library routines. In order to pass the data to IDV, the station observation dataset adapter would get the metadata required to query the MesoWest database. We would adapt these software tools to query both the "current" and complete MesoWest data tables described in the previous paragraph depending on the users' requests.

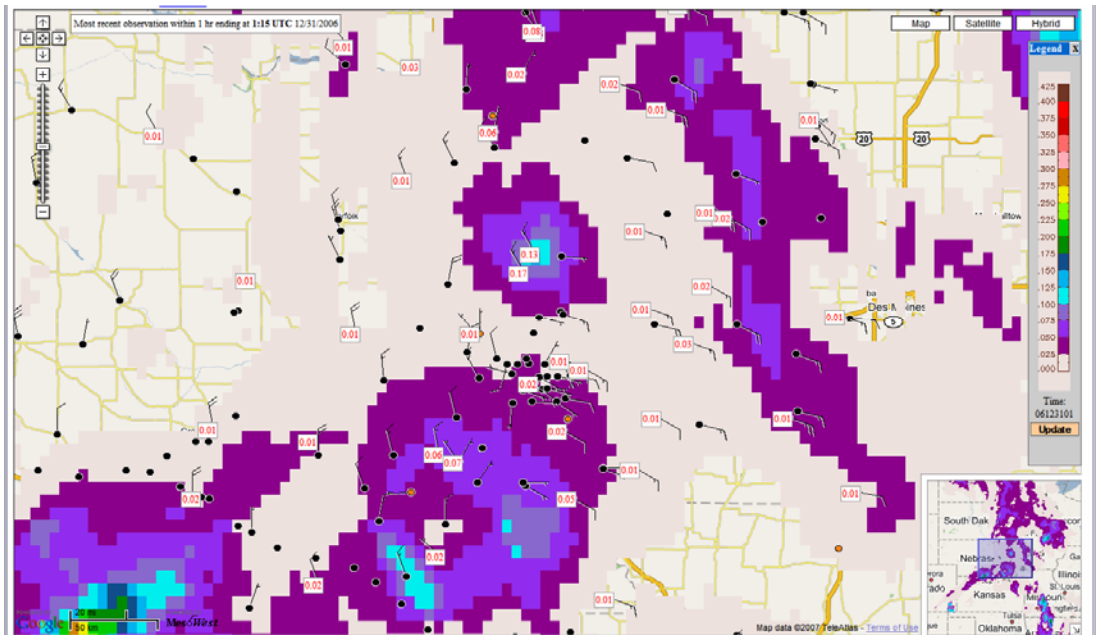

Figure 5. As in Fig. 4 except for MesoWest precipitation observations and the corresponding RTMA precipitation analysis.

RTMA surface analyses of temperature, moisture, wind, cloud, and precipitation at 5 km resolution over the CONUS are available now from NCEP. As shown in Figs. 4 and 5, these analyses help to visualize processes near the surface as well as provide context for surface observations. In addition, the techniques involved in data assimilation can be made less abstract by examining the analyses relative to the observations and the background fields from which the analyses are made. (The background field used by the RTMA is a downscaling of the 13 km RUC to 5 km using software developed by the Global Systems Division of ESRL.)

We propose to support a modest 60-day rotating THREDDS archive of RTMA grids (each hourly file in GRIB-2 format is roughly 16 mbytes), which will require storing roughly 30 gbytes of analyses continuously. The goal of this real-time archive is for educational purposes within the Unidata community rather than a long-term data repository. NCDC intends to host a complete archive of the RTMA grids as part of their NOMADS system.

Hardware requirements for the THREDDS database and server are identical to the proposed upgrade of our real-time data server. Essentially, this server will house the RTMA 60-day archive (order 30 gbytes) and a mirror of our MesoWest database (currently 160 gbytes and growing at roughly 30 gbytes per year). We propose to purchase a data server of the following class for $6000:

- 2 dual core opteron processors
- 16GB of ram
- 3ware raid cards
- 8 150GB, 10,000 RPM sata drives

Specific hardware details may vary depending on memory, disk, and CPU prices at the time of purchase.

c. Expansion of Meteorology CVL

As mentioned in the Project Summary, enrollment in several core undergraduate classes has increased in recent years such that the existing Computation and Visualization Laboratory (CVL) provides a poor learning environment due to overcrowding. Our department is currently housed in 2 separate buildings (INSCC and WBB), with the CVL currently located in the INSCC building. We will be requesting university support to remodel an existing larger computer lab in WBB to replace the CVL.

The proposed addition of 4 Linux workstations to the existing CVL will help to meet our immediate instructional needs as well as provide a foundation for the remodeled computer lab in the WBB building. Currently we have 15 workstations from 1-4 years old. We request funds from the University and typically obtain approval to upgrade 3-5 workstations every other year (we have already submitted an internal university proposal for such an upgrade for this year). We intend to request university funding for an additional 5 workstations next year in order to meet our anticipated critical needs for 25 workstation seats in core upper division undergraduate courses.

Besides direct class room instruction, the CVL is used extensively by students for a myriad of purposes and by faculty for community outreach, training and workshops. The existing CVL is the hub for our undergraduate experience accessible by students 24/7. The expanded computer lab will be coupled with an adjacent student lounge to continue to encourage students to study independently as well as in small group settings. For example, senior capstone projects tend to rely heavily on Unidata delivered data sets and software (IDV and GEMPAK).

Faculty in the Department of Meteorology support wide ranging research programs that rely in part on the IDD data stream for thesis research and journal publications. For example, the Mountain Meteorology Group has used the IDD data stream and Unidata visualization tools for at least 10 publications during the past couple of years (see http://www.met.utah.edu/research/mmg/publications/ for details).

Specifically, we propose to purchase 4 Linux workstations ($2000 each) with the following characteristics:

- 4GB memory
- AMD Dual core
- small form factor casings
- 20” wide screen flat panels

Hardware details will depend to some extent on memory prices at the time of purchase.

3. Budget

As detailed in Section 2, we are proposing to purchase 2 Linux servers and 4 Linux workstations to meet the project goals to improve community access to surface mesonet observations, participate to a greater extent in the Unidata IDD system, and improve student access to Unidata software and resources at the University of Utah.

Specifically, we are requesting a total of $20000 to purchase two data servers ($6000 each) (2 dual core opteron processors, 16GB of ram, 3ware raid cards, 8 150GB, 10,000 RPM sata drives) and 4 Linux workstations ($2000 each) (4GB memory, AMD Dual core, small form factor casings, 20" wide screen flat panels). Specific hardware details may vary depending on memory, disk, and CPU prices at the time of purchase.

This upgrade of our hardware capabilities will take advantage of considerable infrastructure already in place. The Center for High Performance Computing (CHPC) at the University of Utah is responsible for providing high-end computer services to advanced programs in computational sciences and simulations. CHPC has core competencies in operating large scale computing resources, advanced networking, scientific computing and simulations. The University of Utah is a member of the Internet*2* consortium and is connected to the ABILENE backbone via a OC12 link. We are discussing with CHPC staff further integration of our IDD data reception and delivery in the context of their broader large data storage systems. CHPC already has a robust infrastructure for data storage that utilizes multiple servers with multi-path access for disk space. The system has been purchased from Crosswalk, and is called I-Grid.

The Department of Meteorology supports a powerful computing environment with a network of more than 150 workstations and more than 20 terabytes of local disk storage. Department scientists use computer clusters with more than 1000 processors maintained by the CHPC to run various state-of-the-art regional and global atmospheric models.

MesoWest software is supported by the equivalent of 2 full-time staff members (3 staff and 1 undergraduate student provide this support) and partial support for a network administrator. All software development required to integrate MesoWest observations and RTMA analyses into the THREDDS, IDD, and IDV framework will be done using University of Utah staff resources. The Principal Investigator will oversee at no charge the development of the requisite software.

4. Project Milestones

Because of the pressing needs both to improve the classroom environment and provide access to mesonet observations, all equipment will be purchased by Fall 2007. Staff time will be devoted to utilize the general station observation dataset adapter under development by Unidata staff at the present time.

Summer 2007. Coordinate with Unidata staff to develop software to access MesoWest mySQL database dynamically using Java netCDF library tools

Fall 2007. Linux workstations installed in existing CVL to provide improved instructional environment for students

Fall 2007. Installation of 2 Linux servers completed and integrated into existing CHPC and Department of Meteorology networks

Winter 2008. After development and testing in cooperation with Unidata IDV programmers, provide direct, dynamic queries to surface mesonet observations in MesoWest database for IDV users throughout the Unidata community

5. References

Horel, J., M. Splitt, L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, D. Zaff, J. Burks, 2002: MesoWest: Cooperative Mesonets in the Western United States. *Bull. Amer. Meteor. Soc*., 83, 211-226.

Horel, J., B. Colman, 2005: Meeting Summary: A Community Meeting on Real-Time and Retrospective Mesoscale Objective Analysis: An Analysis of Record Summit. *Bull. Amer. Meteor. Soc.*, 86, 1477–1480.

Miller, P., M. F. Barth and L. A. Benjamin 2005: An update on MADIS observation ingest, integration, quality control, and distribution capabilities. *21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*.

Pondeca, M., G. Manikin, S. Park, D. Parrish, W. Wu, G. DiMego, J. Derber, S. Benjamin, J. Horel, S. Lazarus, L. Anderson, B. Colman, G. Mann, and G. Mandt, 2007: The development of the real time mesoscale analysis at NCEP. *23rd IIPS Conference*, San Antonio, TX.

See http://www.met.utah.edu/research/mmg/publications/ for additional related publications that relied on Unidata resources (IDD data stream and visualization software).