Unidata Community Equipment Request

The New Mexico State Climate Office and CARSAME Portal for Community Access to Meteorological, Satellite, and Model Archives

Submitted by

Dr. David DuBois
NM State Climatologist
Phone: 575-646-2974
Fax: 575-646-6041
Department of Plant and Environmental Sciences
MSC 3Q, PO Box 30003
New Mexico State University
Las Cruces, NM 88003
dwdubois@nmsu.edu

[redacted]

Signature of PI:

Neta Fernandez
Director, Office of Grants & Contracts
Phone: 575-646-1590
Fax: 575-646-2020
ogc@nmsu.edu

[redacted]

1. Project Summary

The Principal Investigator proposes the establishment of a community data portal at New Mexico State University using Thematic Real-time Environmental Distributed Data Services (THREDDS) and RAMADDA. The purpose of this portal will be to make available data sets to the public that have been archived at NMSU. These data sets have been archived for several years at the Center for Applied Remote Sensing in Agriculture, Meteorology and Environment (CARSAME) and New Mexico Climate Center. This project will increase that availability of near real-time satellite, numerical weather prediction model output, and surface weather station data to the Unidata community. Our portal will make use of a THREDDS Data Server to provide access to our data using visualization tools such as Unidata's Integrated Data Viewer (IDV). The RAMADDA application will be used to store existing and future case studies that can be viewed and used by researchers and students at NMSU and the community.

2. Project Description

Goals of the project include:

- To provide access to a archive of satellite, meteorological model, and observations that would be useful to the scientific community
- Increase awareness of local data sets available to researchers and students in New Mexico
- Provide a educational tool for storing analysis of events such as flooding, dust storms, and severe storms
- Contribute to the expansion of the Unidata community data and common data formats in our region and beyond
- Upgrade existing data storage hardware at the NM Climate Center and CARSAME

2.1. Motivation

The Center for Applied Remote Sensing in Agriculture, Meteorology and Environment or CARSAME has been the NMSU node on the Unidata Network since 2002. The initial reason for joining the network was to obtain the US Navy NOGAPS feed for operational, real-time modeling using MM5 in a hindcast mode for the previous 24hr period. The MM5 data are archived as are the NOGAPS data at CARSAME. In addition, after the NOGAPS server was not available, we began downloading the NAM 12-km (AWIP3D) data for operational, real-time modeling using WRF-ARW in a hindcast mode for the previous 30-hr period. These runs are centered over south central New Mexico as part of a study to better understand evaporation at the Elephant Butte reservoir. The runs consist of 3 domains with the outer two encompassing the whole of the Rio Grande Basin and the Chihuahuan Desert. Because of current data storage problems, we are only saving the NAM data and the inner 1-km domain. Eventually the lab would like to archive the two outer domains at 9-km and 3-km.

The CARSAME has been successful in archiving many datasets driven in part by requests by NMSU researchers. For example via LDM, the center is obtaining, in real time, NEXRAD data for New Mexico as well as MADIS data (Grid, LDAD, point) for the region and archiving these data. Via FTP, the center is obtaining RUC 13-km and 20-km data in real time.

From our satellite ground stations, we are downloading and processing AVHRR and GOES data to a variety of products that are georeferenced and in HDF file formats. CARSAME has successfully set up real time feeds for NASA MODIS data via a subscription from NASA/GSFC as well as direct broadcast data from Oregon State University. The Center is also receiving, in near real time, GSIP (GOES radiation data) data and processing to gridded, georeferenced arrays of SWD, SWU, LWD, and LWU.

The CARSAME laboratory is also receiving in near real time, USAF/AFWA LIS runs. The remote semsing archive started in 1997 with the acquisition of several satellite ground stations that were capable of receiving NOAA POES (Polar Operational Environmental Satellite). AVHRR data were received in real time and subsetted to our region of interest (20-40N, 95-115W) that encompasses the whole of the Rio Grande Basin as well as the Chihuahuan Desert. The data were calibrated, projected and saved as hdf files. We also received data from the DMSP satellites and SeaWiFS sensor on the SeaStar spacecraft which were similarly processed. Later, we added GOES capability. There was a gap in data while the ground stations were moved to a new location (approximately 2005-2007). We also receive on a subscription basis, MODIS data from Terra and Aqua in near real time from Oregon State University MODIS direct broadcast ground station and similar data from the Goddard Space Flight Center (GSFC). The GSFC data are somewhat unique in that we have archived collection 1-4 (various levels of data processing) as well as the more current collection 5 dataset.

2.2. Equipment Details

The following items are requested to build the New Mexico portal from its existing storage archive.

CD-ROM ripper: Kintronics, Inc. DG2. The first hardware request is for an automated CD reader since the majority of the archive exists on CD media and not on hard drives. The reason for using CDs is that they are portable and stable. CDs are cost effective being about $0.35 each but are not acceptable for our purpose in providing data to the community. Our archive of CDs is of the order of 50-60 thousand CDs. It is only recently that it has been possible to consider placing these data back to spinning disk as well as keeping currently acquired data on spinning disk.

Mass storage server: Dell MD1000. Data pulled from the CDs and the NM Climate Center server will be stored on a high capacity direct attached storage array such as a Dell PowerVault MD1000. This external storage array server will have a capacity of 30 TB.

THREDDS and RAMADDA Server: Data will be processed and accessed by the public using a server such as a Dell PowerEdge R815. This server will consist of 4 AMD Opteron™ 6164HE 1.7 GHz, 12-Core Processors with 32 GB of 1333 MHz UDIMMs Ram, and two 250 GB 3.5 inch SATA 7200 RPM RAID level 1 drives. This will be run using Novell SUSE Linux Enterprise Server 10 as the operating system.

2.3. Contribution to Research:

We anticipate that the THREDDS data portal will serve the research community not only at New Mexico State University but the region and across the border into Mexico. The New Mexico State THREDDS Data archive will be a key research component of the Office of the State Climatologist that will provide a wealth of satellite, model, and observations for the research community.

One of the research areas that will benefit from this archive are in the study of wind erosion and air quality in the southwestern US. Several faculty members at New Mexico State University and University of Texas El Paso have active research project in the study of the sources and transport of dust in the Chihuahuan Desert region. These projects have made extensive use of long wave and visible AVHRR and GOES imagery to determine dust plume boundaries. Access to the THREDDS archive and new imagery would make the imagery much more available and provide much more opportunities for their use by others outside of the two universities.

Another area of active research that we anticipate using this resource is in the area of water resources and agriculture. Many requests for data that come to the Office of the State Climatologist are directly related to agriculture in the form of scheduling irrigation, reducing evaporation, and determining new locations for crops. Having better access to high resolution model predictions will greatly increase the use of these models in applications such as forecasting evapotranspiration in crop lands.

2.4. Contribution to Education:

Currently there are many opportunities for students in the NMSU Plant and Environmental Sciences Department to use the data in this portal. Most of the student users at present are in studying at the graduate level with projects that deal with evapotranspiration, irrigation scheduling, precision agriculture, and viticulture. In this graduate program there are 28 PhD and 24 MS students enrolled in 2011 with several student thesis extensively using remote sensing and numerical weather prediction models. The NMSU Department of Geography teaches the undergraduate level courses in meteorology and climatology. We would encourage the use of this THREDDS portal for use in student projects and using RAMADDA for case study analysis. The PI of this project, a new faculty member at NMSU, is in the process of expanding the upper level climate and meteorology courses and would greatly benefit from the portal in the classroom. The use of IDV would be promoted to visualize the various data sets in these classes. Through the connection of the Office of the State Climatologist, we would promote the THREDDS Data archive with our colleagues at the University of Texas El Paso and the University of New Mexico in Albuquerque.

2.5. Contribution to UNIDATA Community

The following near-real time and archived satellite imagery will be made available to the Unidata community through the THREDDS Data Server.

- NOAA 15, 17, 18 and 19 AVHRR imagery. Recently the NOAA-17 has developed problems with a scan mirror and the resulting imagery is unusable. Of these images we have archives of level 1b files, bands 1-5, NDVI Normalized Difference Vegetation Index, LST Land Surface Temperature, and Albedo reflectivity ratio
- NOAA GOES imagery clipped to the Chihuahuan Desert region

Near real-time and archived model output files will be made available to the Unidata community through the proposed project. The THREDDS Data Server will include the following model based data sets.

- RUC 12-km and 20-km data to include hourly precipitation forecasts. The hourly precipitation forecast data consists of an analysis output as well as forecasts for 01hr, 02hr, 03hr, 06hr, 09hr, and 12hr from the cycle time. Four records for which precipitation are outputted for each forecast time are saved as are the original input data.
- WRF model predictions at 1-km resolution centered around Elephant Butte reservoir in central New Mexico as well as several years of MM5 runs over the CARSAME region with inner domain grid cell of 10km.

Point data from automated surface weather stations networks are being archived at the New Mexico Climate Center:

- NM Climate Center automated surface weather stations
- NM Viticulture
- RAWS
- Elephant Butte Irrigation District
- Navajo Agricultural Products Industry
- NRCS SNOTEL
- METAR sites

We will migrate our archived data sets using appropriate data formats for THREDDS such as NetCDF, and GeoTIFF. A schematic of the basic satellite, model, and point data sources are shown in Figure 1.

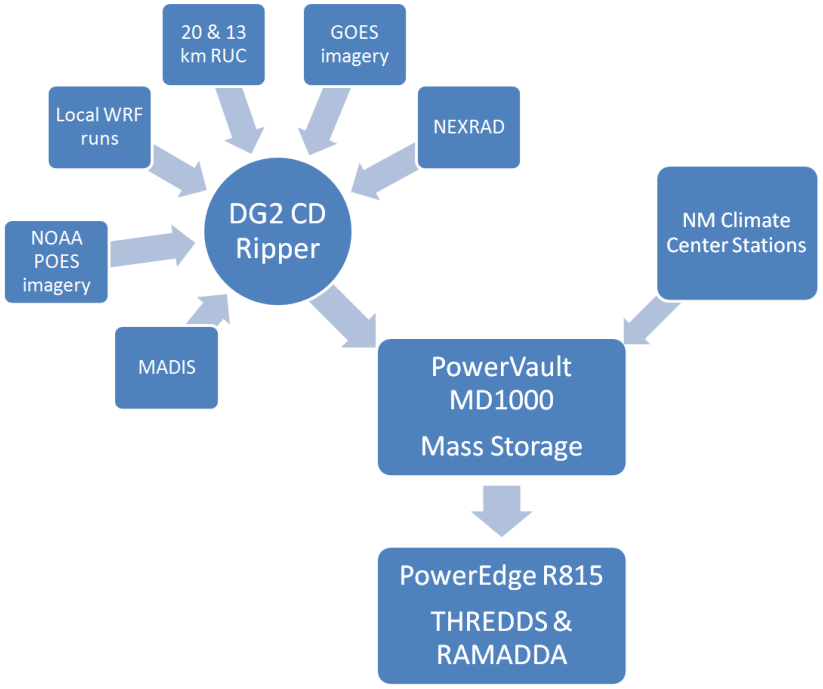


Figure 1. Sources of data in the Portal

2.6. Available computing resources at CARSAME and the NM Climate Center

Currently the New Mexico Climate Center and the CARSAME laboratory data are housed in separate servers and managed by different programs. This project would bring these two programs together in sharing data storage resources. Server and data storage resources currently in operation at the CARSAME laboratory include:

- Hardware available to act as the server for the public interface to THREDDS is a 4-core linux-based machine that is about 5 or 6 years old and multiple data collectors consisting of several low-end linux-machines as well as the windows boxes for archiving the data to CDs.
- Our satellite data are acquired by a 4-node 32-bit Sun HPC running Ubuntu GNU/Linux and MPICH2 and processed on a 16-node 64-bit Sun HPC running Ubuntu GNU/Linux and MPICH2

The NM Climate Center uses a single processor Pentium server running Ubuntu Linux to house ASCII data and a MySQL database. Data from the field is collected on a single processor Windows box running Campbell Scientific LoggerNet. One other single processor Linux machine is used for processing and data storage for various projects. Finally another single processor Linux box is used for the public webserver.

2.7. System Operation and Maintenance

The installation, operation, and maintenance will be performed by a systems administrator who currently works in the CARSAME laboratory. The administrator will also be overseeing the migration of

the archived data from the CD-ROMs to the mass storage server. Day to day operations will be carried out with undergraduate student research assistants. The system administrator and student salaries will be covered by departmental accounts and not from support from Unidata. Salaries for the PI and Director of CARSAME are covered by university budgets and external funding. Collaboration between the CARSAME laboratory and the NM Climate Center will be strengthened by the common research goal of the State Climatologist and the Director of CARSAME in providing relevant data to the community to support education, research, outreach, and extension.

3. Budget

The following table provides a listing of the equipment that is requested in this proposal and cost estimates based on prices at the time of writing. The CD Ripper and PowerEdge R815 servers do not incur indirect costs since they are over $5,000. Indirect costs are computed at a rate of 42.4%. NMSU will provide funds for shipping of the servers and any cost over-runs due to price increases over the quotes in the table.

Item	Quantity	Unit Cost	Total
CD Ripper: Kintronics, Inc. Big Ripper DG2 2-drive CD/DVD data ripper with 300 disc capacity	1	$5,495	$5,495
Direct Attached Storage Array: Dell PowerVault MD1000	1	$3,799	$3,799
Hard Disks for MD1000 30 TB Hard Drives (15 x 2 TB)	15	$82	$1,230
THREDDS & RAMADDA Data Server: Dell PowerEdge R815 server 4 AMD Opteron™ 6164HE 1.7 GHz 12-Core Processors 32 GB 1333MHz UDIMM Ram (16 x 2 GB) 500 GB Hard Drive (2 x 250 GB)	1	$7,284	$7,284
University Indirect Cost (42.4%)			$2,132
Shipping (NMSU to provide funds)	0	$500	$0
Total			$19,940

4. Project Milestones

Deployment of the THREDDS and RAMADDA Data portal will follow the schedule in the table below.

In months after date of award								
Task	1	3	6	9	12	15	18	24
Purchase hardware								
Develop THREDDS Data Server								
Migration of CD archive to mass storage array and data format conversion								
Testing of THREDDS and development of test cases for RAMADDA								
Roll-out of THREDDS and RAMADDA portal to community								

Appendix: BIOGRAPHICAL SKETCH of PI

David W. DuBois, Ph.D.
New Mexico State Climatologist
Department of Plant and Environmental Science
New Mexico State University, Las Cruces, New Mexico 88003



Education

University of Nevada, Reno	Atmospheric Sciences	Ph.D./May 2003
New Mexico State University	Physics	M.S./ May 1991
Rutgers University	Physics	B.A./ May 1988

Appointments

State Climatologist	Dept. of Plant & Env. Sci., NMSU	02/10 - present
Associate Research Scientist	Div. Atmospheric Sciences, DRI	07/09 – 02/10
Assistant Research Scientist	Div. Atmospheric Sciences, DRI	03/03 – 06/09
Manager, Modeling Section	New Mexico Environment Dept.	03/02 – 03/03
Environmental Engineer	New Mexico Environment Dept.	11/99 – 03/02
Research Assistant	Div. Atmospheric Sciences, DRI	08/95 – 11/99
Associate Scientist	SciTec, Inc., subsid. of TRW	06/91 – 08/95
Research Assistant	Dept. of Physics, NMSU	01/89 – 03/91

Selected Publications

Koracin, D., R. Vellore, D.H. Lowenthal, J.G. Watson, J. Koracin, T. McCord, D.W. DuBois, L.-W. Antony Chen, N. Kumar, E.M. Knipping, N.J.M. Wheeler, K. Craig, and S. Reid (2010). Regional source identification using Lagrangian stochastic particle dispersion and HYSPLIT backward-trajectory models. Accepted for publication in the Journal of the Air & Waste Management Association

Chen, L.-W.A., J.G. Watson, J.C. Chow, D.W. DuBois, L. Herschberger (2010) Chemical mass balance source apportionment for combined PM2.5 measurements from U.S. non-urban and urban long-term networks. Accepted for publication in Atmospheric Environment

Chen, L.-W.A., J.G. Watson, J.C. Chow, D. DuBois, and L. Herschberger (2010). PM2.5 Source Apportionment in Minnesota, USA: Reconciling Receptor Models for Urban and Rural Monitoring Networks. Accepted for publication in the Journal of the Air & Waste Management Association

Chen, L.-W.A., J.G. Watson, J.C. Chow, D. DuBois, and L. Herschberger (2010). PM2.5 Source Apportionment in Minnesota, USA: Application of the Chemical Mass Balance Method to Urban and Rural Monitoring Networks. Accepted for publication in Environmental Science & Technology

Lowenthal, D.H., J.G. Watson, D. Koracin, L.-W.A. Chen, D. DuBois, R. Vellore, N. Kumar, E.M. Knipping, N. Wheeler, K. Craig, and S. Reid (2010). Evaluation of Regional Scale Receptor Modeling. Journal of the Air & Waste Management Association, 60(1):26-42.

Kavouras, I.G., V. Etyemezian, and D.W. DuBois. (2009). A geospatial screening tool to identify the sources of windblown dust Environmental Modeling and Software, 24(8): 1003-1011

Kavouras, I.G., Etyemezian, V., Xu, J., DuBois, D.W., Green, M., Pitchford, M. (2007) *Assessment of the local windblown component of dust in the western United States*. Journal of Geophysical Research 112, doi:10.1029/2006JD007832.

Xu, J., DuBois, D., Pitchford, M., Green, M., Etyemezian, V. (2006) *Attribution of sulfate aerosols in federal class I areas of the western United States based on trajectory regression analysis.* Atmospheric Environment 40, 3433-3447.

Kuhns, H., Gillies, J., Etyemezian, V., DuBois, D., Ahonen, S., Nikolic, D., Durham, C. (2005) *Spatial variability of unpaved road dust PM10 emission factors near El Paso, Texas.* Journal of the Air & Waste Management Association 55, 3-12.

DuBois, D, Green, M., Xu, J., Freeman, D., Etyemezian, V., Pitchford, M. (2004) *The Causes of Haze Assessment: An overview*. AIR AND WASTE MANAGEMENT ASSOCIATION'S VISBILITY SPECIALTY CONFERENCE, ASHEVILLE, NC.

Husar, R.B., Tratt, D., Schichtel, B.A., Falke, S.R., Li, F., Jaffe, D., Gasso, S., Gill, T., Laulinen, N.S., Lu, F., Reheis, M., Chun, Y., Westphal, D., Holben, B.N., Geymard, C., McKendry, I., Feldman, G.C., McClain, C., Frouin, R.J., Merrill, J., DuBois, D., Vignola, F., Murayama, T., Nickovic, S., Wilson, W.E., Sassen, K., Sugimoto, N. (2001) *The Asian dust events of April 1998*. Journal of Geophysical Research 106, D16, 18317.

Principal Professional Activities

1. Director of the New Mexico Climate Center. The Center's primary purpose is to maintain and collect meteorological data across the state of New Mexico and to operate an 15 site automated surface weather stations. 2010-present
2. Principal investigator for the New Mexico Department of Health, Border Air Quality Study. This is a four year study to assess air quality across the New Mexico/Mexican border and to correlate human health effects with adverse air quality conditions. 2010-present
3. Principal investigator for the NMSU/Desert Research Institute boundary layer profiling atmospheric research group. The group has used a 915-MHz radar wind profiler, SODARs, a tethered balloon system for the measurement of winds aloft, and a laser ceilometer to understand the three dimensional nature of particles in the atmosphere. 2003-present
4. Investigations in understanding regional haze and long-range transport of aerosols. Co-Principal investigator of a comprehensive regional haze study to determine the causes of haze at national parks and wilderness areas across the western U.S., Alaska and Hawaii. 1998-present

Student Support and Advising

Graduate thesis advising: 1 MS, 1 PhD at NMSU, 1 PhD at Univ. of Nevada Reno
Undergraduate sponsorship and advising: 1 in NMSU, 6 Univ. of Nevada Las Vegas