J2.3

THE NOAA OPERATIONAL MODEL ARCHIVE AND DISTRIBUTION SYSTEM (NOMADS)

Glenn K. Rutledge*[1], Jordan Alpert[2], Ronald Stouffer[3], Benjamin Domenico[4], Lawrence Buja[5], Dean Williams[6], Brian Doty[7], Menas Kafatos[8], and Steve Hankin[9]

[1]National Climatic Data Center
Climate Data Division
Asheville, NC 28801

[2]National Centers for Environmental Prediction
Environmental Modeling Center
Camp Springs, MD 20736

[3]NOAA Office of Atmospheric Research
Geophysical Fluid Dynamics Laboratory
Princeton, NJ 08542

[4]University Corporation for Atmospheric Research
Unidata Program Center
Boulder, CO 80301

[5]University Corporation for Atmospheric Research
National Center for Atmospheric Research
Boulder CO. 80305

[6]Lawrence Livermore National Laboratory
Program for Climate Model Diagnosis and Intercomparison
Livermore, CA 94550

[7]Institute of Global Environment and Society, Inc.
Center for Ocean-Land-Atmosphere Studies
Calverton, MD 20705

[8]George Mason University
Center for Earth Observing and Space Research (CEOSR)
George Mason University
Fairfax, VA 22030

[9]NOAA Office of Atmospheric Research
Pacific Marine Environmental Laboratory
Seattle, WA 98115

1.0 INTRODUCTION

To address a growing need for retrospective Global Climate Model (GCM) and Numerical Weather Prediction (NWP) data, the National Climatic Data Center (NCDC) along with the National Center's for Environmental Prediction (NCEP) and the Geophysical Fluid Dynamics Laboratory (GFDL) have initiated a highly collaborative NOAA Operational Model Archive and Distribution System (NOMADS). NOMADS will provide retrospective access to model and observational data by a wide variety of users via the Internet and the Next Generation Internet (NGI). NOMADS is an inter-operable network architecture with fully integrated data access and manipulation tools using a distributed, near format independent client-server methodology.

NOMADS benefits from existing and emerging technologies to provide a *"Unified Climate and Weather Archive"*. To enable universal user access and system inter-operability, NOMADS has three primary data servers: 1) the Distributed Oceanographic Distribution System (DODS), 2) the GrADS-DODS Server (GDS), and 3) the Live Access Server (LAS). Traditional on-line data services through ftp are also available. Users with standard desktop data manipulation tools such as Ferret, MatLAB, AWIPS, IDL, GrADS and even typical Web browsers can access data in their preferred format. To provide for this format neutral data access, NOMADS uses these servers and the EXtensible Markup Language (XML).

* Corresponding author address: Glenn K. Rutledge, National Climatic Data Center, Asheville, NC 28801;
e-mail: Glenn.Rutledge@noaa.gov

2.0 BACKGROUND

Currently, NCEP NWP output and GFDL GCM simulations are available in real-time through a number of different channels. Historical data, the data needed for research and collaboration, are more difficult to obtain and the variety of products are much more limited. There is no single repository or effective distribution method for these data. Local piecemeal approaches at government laboratories, data centers, and universities fall short of filling the need for retrospective model data. Even in cases where the data are available, there exists no feedback mechanism to redeem practical value of research findings and facilitate these findings into operations. The primary U.S. National responsibility for the archive and service of weather and climate data rests with the NCDC. However, as the temporal and spatial resolution of models increases, the volume and varied formats of data presented for archive at NCDC using current communications technologies and data management techniques is inadequate. The scientific modeling community is a vast intellectual resource. This community is extremely interested in obtaining GFDL, NCAR, and NCEP, products for historical cases, and for operational and research purposes. The community is also interested in collaborating to improve weather and climate forecasts, but virtually no resources currently exist to support this collaboration.

3.0 BENEFITS

Under the NOMADS framework improvements to models will lead to longer lead times and improved accuracy for forecasts over all time scales. This results in fewer lives lost and less property damage. The NOMADS will allow researchers to more easily access and understand data thereby increasing scientific productivity. NOMADS provides a capability for scientists to assess the potential

impact of long term climate change on the Nation using increasingly detailed and complex models since various models will be more readily available for inter comparison.

The NOMADS proposed solution will reduce costs NOAA wide by integrating local efforts to archive and gain access to historical model data and for the first time create a unified archive of climate and weather data. The NOMADS project leverages off of many man years of software development within NOAA and elsewhere. Heavy dependence is made upon technology transfer and re-use of previously developed successful technologies such as the DODS, the Center for Ocean-Land-Atmosphere Studies (COLA) developed GDS, and the Pacific Marine Environmental Laboratory (PMEL) developed LAS. NOMADS allows an option for the integration of large data sets by performing some of the data manipulation, subsetting, and compression on the servicing host rather then transmitting large amounts of data across relatively slow speed lines. NOMADS available data sets can easily be expanded to include other data sets such as satellite and radar.

4.0 SYSTEM ARCHITECTURE

NOMADS provides three primary client-server data servers. The DODS, the GDS and the LAS. These servers provide access to collections of data. A DODS client requests data from the servers using URL's to describe the desired data. DODS servers translate data into format neutral network objects, allowing participating modelers to serve data in their native file formats. Clients once re-linked with DODS access these format neutral objects through standard Applications Program Interfaces (APIs). Client connectivity is built upon multiple API's, including standards such as netCDF and HDF, and GRIdded Binary (GRIB) format, enabling DODS/GDS to integrate with standard desktop tools - custom models, AWIPS, GrADS, Matlab, IDL, Ferret, etc. The LAS is also used in NOMADS and is a well-known Internet based data access and manipulation tool. The GDS allows DODS enabled transactions but also has the capability for enhanced server side data manipulation. In GDS, GrADS expressions can be provided in the body of the URL request to manipulate data on the server providing only the resultant back to the user. Since it appears that network connectivity continually lags behind desktop processing speeds, GDS allows Internet users access to information they would not normally consider possible.

Since most weather model data is stored in the GRIded Binary (GRIB) data format and most climate model data is stored in the NetCDF data format, NOMADS provides for the conversion between these two formats. Converting between the data formats has been a very difficult task in the past. The DODS software allows users to work in whatever data format is the most advantageous to them, allowing the data format to be hidden from them. The combination of the quality control and independent data format will for the first time provide users with a seamless interface to model data. Climate users will be able to analysis weather data and weather users will have access to climate data. Never before has this capability existed and NOAA has the opportunity to be at the forefront in this development of a unified climate and weather archive.

The idea of a distributed client-server architecture is not new, but is now catching on across multiple scientific disciplines. One concern is the potential lack of coordination between the major data centers. Or, different user desktop applications have changed and the data formats they use are no longer valid. NOMADS allows for growth of data sets and forever changing data formats. If the current list of applicable formats no longer applies or is no longer in use, the NOMADS framework is intended to allow for a new XML descriptor file to conform to the new data "standard." Although XML is supposed to be a generic and general language, in real life it requires the prior agreement on a vocabulary by all parties. Thus, full coordination is required for a successful inter-operability of systems for data sharing. This coordination already exists among the NOMADS collaborators. The sheer number of collaborating institutions contributing their expertise and collections under NOMADS is impressive. NOMADS has the potential to bridge the gaps of inter-operability between systems and thus provide users with a metadata of geosciences information, visualization and analysis tools, and research opportunities.

During this next project year, the NOMADS project will develop a metadata description template for the Binary Universal Form for the Representation of meteorological data (BUFR) format and provide these data under the GDS server. This format is used for the NCEP model input fields, and other data provided to NCDC in BUFR format. Increased on-line server capability will be added at GFDL, and an automated data and metadata quality control (QC) processes will be developed for the NOMADS architecture. These QC processes have been developed by NOMADS collaborator's at the Department of Energy's (DOE) Lawrence Livermore National Laboratory (LLNL) and will be implemented for NOMADS by GFDL.

5.0 DATA AVAILABILITY

When operational (within the next year), NOMADS will provide retrospective access to the entire NOAAPort AWIPS model output from NCDC's NOAAPort ingest capability (Rutledge, G.K., et. al., 2000). NCEP's run history information and the Global and Regional Data Assimilation files including spectral coefficients will be available. Additionally, GFDL GCM data that are now currently available are the R-15 and R-30 (150 and 300 year) climate model simulations. For further information on current available GCM data use: http://nomads.gfdl.noaa.gov/. For currently available NWP and other data see the PMEL NOMADS site at: http://ferret.wrc.noaa.gov/nomads and the NCDC site at: http://nomads2.ncdc.noaa.gov:9090

6.0 ADDITIONAL COLLABORATORS

In addition to GFDL, LLNL, NCDC, NCEP, PMEL, and COLA, NOMADS collaborators include the National

Center for Atmospheric Research (NCAR), NOAA's Climate Diagnostics Center (CDC), the Forecast Systems Laboratory (FSL), the George Mason University (GMU), the University Consortium for Atmospheric Research (UCAR) Unidata Program, the National Severe Storms Laboratory (collaborating with the University of Wisconsin, Space Science Engineer Center).

Under the NOMADS framework collaborations of note include Unidata's THematic Real-time Environmental Distributed Data Services (THREDDS) system, the George Mason University NASA Seasonal-to-Interannual (SI) Earth Sciences Information Partner (ESIP), and LLNL's Program for Climate Model Diagnosis and Intercomparison (PCMDI) project, allowing for a potential NOMADS participation within the scope of the U.S. DOE's Accelerated Climate Prediction Initiative (ACPI) project.

7.0 ADVANCED REQUIREMENTS

Even with the emergence of new web based services, the comparison of transient GCM results with the observational climate record is still difficult for several reasons. One limitation is the global distributions of a number of basic climate quantities, such as precipitation, are not well known. Similarly, observational limitations exist with model re-analysis data. Both the NCEP/NCAR (Kistler, R., Kalnay, E., et. al.) and the ECMWF (Gibson et. al., 1997) re-analysis eliminate the problems of changing model analysis systems but observational data also contain time-dependant biases by changing observational networks, station moves, and the assimilation of various remotely sensed data (Rutledge, G., et. al., 1991) using differing sensor instruments, or calibrations. These changes in input data are blended with the natural variability making estimates of true variability uncertain. The need for data homogeneity is critical to study questions related to the ability to evaluate simulation of past climate. One approach to correct for time-dependant biases and data sparse regions is the development and use of high quality "reference" data sets (Karl, T. et. al., 2000).

Beyond the ingest and access capability soon to be implemented with NOMADS is the challenge of algorithm development for the inter-comparison of large-array data (e.g., satellite and radar) with surface, upper-air, and sub-surface ocean observational data. The implementation of NOMADS will foster the development of new quality control processes by taking advantage of distributed data access. One major challenge facing the scientific community is the development of methodologies for the intercomparison of large-array observational data sets with model simulations. In the near future, NOMADS would include the development of algorithms to blend remotely sensed data with in situ surface and upper-air data, and then to use these blended fields for verification and validation of both weather prediction and climate models in both time and space (Rutledge, G.K., et. al., 2001).

NOMADS is a technology for collaboration and model development as exemplified by the Weather Research and Forecast (WRF) model. The contribution of the scientific community to the WRF development effort requires the transfer of data sets of model retrospective results for testing and comparison. NOMADS enables both retrospective and real time access to the suite of digital products from reanalysis and operational results to give a range of users, from commercial to university scientist, model comparison access and supply operational grade initialization for forecast model study. See http://www.wrf-model.org/documentation_main.html for more information regarding the WRF model.

8.0 ACKNOWLEDGMENTS

The author and NOMADS Principal Investigator (PI) gratefully acknowledge the combined efforts of all the NOMADS Team members for the current success of NOMADS. Special thanks go to Brian Doty and Jennifer Adams at COLA, Dean Williams and Mike Firoino at LLNL, Ethan Davis at Unidata, Dave Brown at NCAR, Ruixin Yang at GMU, and Chris Kerr at GFDL. Prototype funding has been provided by ESDIM 01-420E.

9.0 BIBLIOGRAPHY

Gibson, J. K., P. Kallberg, S. Uppala, A. Noumura, A. Hernandez, and E. Serrano, 1997: ERA Description. ECMWF Re-Analysis Project Report Series, 1. ECMWF, Reading, UK, 77 pp.

Karl, T., D. Easterling, P. Groisman, et al."Observed Variability and Trends in Extreme Climate Events: A Brief Review." Bulletin of the American Meteorological Society, 81, March 2000, 417-425.

Kistler, R., W. Collins, E. Kalnay, R. Reynolds, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, J. Janowiak, M. Kanamitsu, K. Mo, C. Ropelewski, R. Jenne, D. Joseph and M. Fiorino, 2001: The NCEP/NCAR 50-year Reanalysis: Monthly-means CD-ROM and Documentation. Bull. Amer. Meteorological Soc. 92, 247-267.

Rutledge, G.K., E. Legg, and P. Menzel, "Operational Production of Winds from Cloud Motions," Palaeogeography, Palaeoclimaotology, Palaeoecology, Vol. 90 No. 1-3 (Global and Planetary Change Section), 1991, 141-150.

Rutledge, G.K., A. Stanley, E. Page, L. Spayd, and J. Brundage, 2000: "Implementation of the NOAAPort Data Archive and Retrieval System (NDARS) at the National Climatic Data Center," Preprints 16th International Conf. on Interactive Information and Processing (IIPS) for Meteorology, Oceanography and Hydrology, Long Beach, Amer. Meteor. Soc., 492-495.

Rutledge, G.K., T. Karl, D. Easterling, L. Buja, R. Stouffer, "The NOAA Operational Model Archive and Distribution System (NOMADS): Bridging the Model/Observations Gap," (Invited), American Geophysical Union, (AGU) Boston MA., May 2001.