

**P1.28 ACCESS TO REAL-TIME AND RETROSPECTIVE NUMERICAL WEATHER PREDICTION MODELS
AT THE NATIONAL CLIMATIC DATA CENTER
AND THE
NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION**

Glenn K. Rutledge *
National Climatic Data Center
Scientific Services Division
Asheville, NC 28801

Jordan Alpert
National Centers for Environmental Prediction
Environmental Modeling Center
Camp Springs, MD 20736

William Ted Smith
National Climatic Data Center
STG, Inc
Asheville, NC 28801

1. BACKGROUND

Numerical Weather Prediction (NWP) models are available in real-time through a number of different channels. However, historical model input and output which are needed for research, model improvement, publications, training, and collaboration, are more difficult to obtain. Local piecemeal approaches at government laboratories, data centers, and universities fall short of filling the need for access to retrospective models and associated data. Even with increases in available bandwidth and technological advances in distribution and access systems, access to retrospective models has been very limited. In fact, until now a U.S. National digital archive of operational weather models has simply not existed.

A new historical model archive at the National Climatic Data Center (NCDC) and a new real-time data access service at the National Centers for Environmental Prediction (NCEP) under a collaborative project called the NOAA Operational Model Archive and Distribution System (NOMADS) is now operational. Both systems are considered pilot projects with plans for future expansion.

The data that are available include both real-time and retrospective NWP initialization and restart files, NWP output, reanalysis, limited ocean data and climate General Circulation Models. Access to these data in a distributed open and format neutral manner is provided to promote improvements to operational models by providing a feedback mechanism from the research and Regional modeling communities to the operational communities, and promote model intercomparison efforts.

The NOMADS site at NCDC can be found at www.ncdc.noaa.gov/oa/model/model-resources.html or

* *Corresponding author address:* Glenn Rutledge, National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801; e-mail: glenn.rutledge@noaa.gov.

go to the NCDC Home Page, Climate, and then Model Resources.

2. PARTNERSHIPS FOR DISTRIBUTED DATA ACCESS

NOMADS adopted and actively supports the growing use and agreements necessary for format independent distributed access. Leveraging off of many distributed data access projects from many cooperating sources (see the NOMADS Home Page for list of collaborators), NOMADS uses the OPeNDAP (formerly DODS), binary data transport protocol, and an XML based protocol, to provide data over the internet. The servers at NCEP and NCDC include the GrADS-DODS Server (GDS), and traditional Web based services. The distributed data access features with is the core of the NOMADS data access philosophy for temporal and spatial sub-setting is provided by the GDS, and a local web based access interface was developed jointly by NCDC and NCEP. All of the access capabilities for NOMADS are detailed at the NCDC model resources site using the "Model Data Access Guide" at nomads.ncdc.noaa.gov/user-guide.html.

One of the major successes of NOMADS is the development of partnerships between many of the leading weather and climate science groups, both research and operational, into the ocean community. The OPeNDAP partnership is now stronger due to the NOMADS collaborators, and a global vision for open, interoperable approach to data access, is several steps closer toward a reality.

NOMADS is actively partnering with and supported by the University Corporation for Atmospheric Research (UCAR), Unidata Program Center's Thematic Real-time Environmental Distributed Data Services (THREDDS), as well as the Digital Library for Earth System Education (DLESE), and the National Science Digital Library (NSDL). Alignment with these activities, OPeNDAP, and other NOMADS partners such as new developments within PMEL's Live Access Server and the Center for

Ocean-Land-Atmosphere Studies' (COLA) GDS, ensures that NOAA can play a lead role in the national multi-organization efforts to provide users with coherent access to the vast array of environmental data stored in data collections around the country and the world.

Never before has the U.S. maintained a digital library for gridded model output. In the past NCEP maintained a short term collection for its own use, however the need for a publicly available historical model input and output archive was not justifiable. Within the last half decade or less however, the need for historical model data, and model initialization and boundary conditions for NWP has never been greater. With the increasing power of desktop computers and the availability of Regional models, a new paradigm for distributed and local model development is emerging. These regional modeling centers and other research centers are a vast NOAA resource and will continue to grow. A two-way feedback mechanism needs to be developed and fostered at all levels of the geo-sciences community. NOMADS allows these activities to grow since both computer and research scientists are actively engaged in solving these issues from a bottom-up approach based upon what users use and need most.

There is a good deal of work remaining however. Interoperable data access requires metadata management, and communications between collaborators and many of these pilot projects, including NOMADS require testing for operational demands. Conversion issues (e.g., grib to NetCDF), on some complex climate grids also remain but are slowly being addressed. Finally, the entire climate, weather and ocean communities need to develop an overarching plan, forge agreements, and a methodology to support distributed processing across organizational boundaries. Organizations and/or Programs such as CLIVAR, WCRP, USWRP, and the Climate Change Science Program could be used to provide this leadership.

3. NWP AVAILABILITY AND ACCESS

Traditional data access methods such as ftp, http, and wget are available under the NOMADS umbrella. The browse and plot capability on the NOMADS server is called "Web Plotter" developed by NCEP, and modified for use under NOMADS by NCDC. A data selection capability was added to the front end of Web Plotter for user selection of data for historical archive requests. These requests can span the on-line disks (currently up to 2 years), or reach into the NCDC archive for longer term access and manipulation. The Web plotter allows time series plots, lat/long vs. time (Hovmoller diagrams), and other host side data manipulation services.

Under NOMADS six primary categories of data are available at NCDC:

- 1) NWP Model Output (NCDC's NOAAPort ingest, and NCEP's Hi-Res fields)
 - 1a) FSL RUC model as provided by the NOAAPort Broadcast
- 2) NWP Model Input

(NCEP)

- 3) Reanalysis (at NCEP and NCDC): Global and Regional
- 4) Ocean models, weather and climate (Spring 2004)
- 5) AOGCM climate models at GFDL
- 6) Reference observational data sets (CARDS, ERSST's, GHCN, etc.)

Models and associated data will be stored at NCDC and retrieved by NOMADS from a 5 year rotating archive. Analysis products will be archived for the longer term however. Other data such as observational and reference data sets such as the Smith-Reynolds Extended Reconstructed Sea Surface Temperature's are currently available under the NOMADS framework at NCDC. However, the focus of this paper is model data availability.

3.1 Model Output Archive

The model output data currently available at NCDC for the period beginning May 2002 to the present include:

Eta Model

Data are available four times a day (00Z, 06Z, 12Z, 18Z), out to 48 hour forecasts at 3-hour intervals on various resolutions from 12km to 40Km on various grids and projections.

Global Forecast System

Data will be available four times a day (00Z, 06Z, 12Z, 18Z), out to 7 day forecasts at 3-hour intervals on a 1x1 degree global and regional grids from 60km to 200km (grids 201-213), and from 8 – 16 day forecasts on a 2.5x2.5 degree global grid in compressed GRIB grid #2.

Rapid Update Cycle (RUC) System

RUC fields as received on the NOAAPort system are available hourly out to 12hour forecasts in 3hrly forecast cycles. (FSL soon to provide a full suite of RUC model data on a NOMADS (DODS) Server at FSL.)

3.2 Model Input Archive

Under NOMADS and for the first time, the regional modeling community will have direct access to NCEP's real-time and historical model assimilation fields. As NCEP runs a series of computer analyses and forecasts operationally. One of the primary operational systems is the Global Data Assimilation System (GDAS) (Kanamitsu, 1989), which uses the spectral Medium Range Forecast model (MRF) for the forecast (Sela, 1980). In simple terms, for each run, unequally spaced conventional and remote sensed observations are assimilated with "first guess" data fields (forecasts from the previous model run), and dynamic imbalances in the data are reduced, resulting in "analyzed" data fields. Then the forecast is made. The analyzed data provides

an optimal representation of the real atmosphere on a grid or spectral coefficients of spherical harmonics. These can be compared to observations which have limitations due to measurement error or other instrument problems, and non-uniform spatial and temporal distributions of the observations.

Publicly available on the NOMADS servers at NCEP and NCDC is the GDAS dataset. GDAS consists of the minimum set necessary to re-generate NCEP analysis and forecast products (model re-start and initialization files). GDAS includes the Global Spectral Forecast Model (GSM), and the Spectral Statistical Interpolation (SSI) Cycling Analysis System (CAS) with triangular truncation (T) 170 and 28 sigma levels. To start the CAS, model spectral coefficients are provided on Gaussian grid in a sigma vertical coordinate system. These data represent the model's "ground truth", and the best estimate- in terms of analyzed fields- for scientific study. Data that are restricted may not be available.

The archived analysis data sets serve as model verification as well as the source for model reruns and retrospectives. Including the observations allows for cycling analysis systems to re-analyze the observations. Never before has this model input data and information been available to the public. A partial list of NOMADS planned available observations (with associated data format) include:

- Analysis Bias Corrected Information (ASCII)
- Ship / Buoy Observations (BUFR)
- Guess prep / guess output (BUFR)
- Observational Toss List (ASCII)
- Bogus Observations (BUFR)
- ACARS and Aircraft (BUFR)
- Wind Observations (BUFR)
- Analysis Ready Obs. (prepBUFR)
- Surface Analysis Restart Files (BUFR)
- Surface and Upper-Air observations (BUFR)
- Fixed Snow Field (GRIB)
- Previous 6 hour forecast (BUFR)
- "Post" Guess Output (spectral binary)
- Profiler (BUFR) / SST's (GRIB)
- MSU 14 and HIRS 14/15/16 (IEEE)
- SSMI Satellite obs (BUFR)
- NOAA (satellite) 15/16 AMSU - A/B
- TOVS 1B Radiances (IEEE)
- TOVS Satellite Obs (BUFR)
- GOES Satellite Obs (BUFR)
- O3 Sat Obs (binary) and ERS Sat obs
- SBUV: Satellite Wind Observations
- Radar VAD Winds (BUFR)

The NCEP GSM files transmitted to NCDC represent the first operational model restart and retrospective archive. The entire data set of model run history is too large to keep at this time. Thus, we have devised the minimum set necessary to regenerate, as closely as possible, an operational run with the NCEP system and allow for other test beds to utilize, run experiments, and other models to initialize from this

data set. The data set consists of conventional and remote sensed observations made ready to start NCEP cycling analysis system. Other NCEP operational models in addition to the GSM, such as the Eta and WRF regional models are coming soon. Additional data sets consisting of model run history in GRIB pressure vertical coordinate on a longitude/latitude grid will be present as determined by NOMADS panel.

"Post" is a FORTRAN program available from NCEP that will convert spectral coefficients to gaussian grid, sigma to pressure, and gaussian to latitude and longitude. Work continues to couple Post to the NOMADS user interface for source and executable downloads. The GDAS dataset under NOMADS will include the GSM and the SSI-CAS binary files and contains ~2.5Gb per day (4 cycles/day: 00Z, 06Z, 12Z, and 18Z). The binary files are raw data, which are acted on by NOMADS servers to produce useful grids. The analysis, initial condition and predicted fields on the model sigma levels are interpolated to the standard pressure levels in the POST program. The input file consists of the sigma level dependent variables and the output file consists of the pressure level variables in WMO standard GRIB. The file also contains several processed arrays (e.g. boundary layer parameters and tropopause parameters.)

In addition, using the Post program, the restart files are converted to GRIB data sets. The information contained in restart (in this case GSM sigma files) files represent the final analysis of conventional and remote sensed observations in the models vertical sigma coordinate and spectral coefficients. This file is needed to rerun NCEP models and analysis. The POST program changes this file to standard WMO GRIB containing fields in pressure vertical coordinate on a longitude/latitude grid. The GRIB data set can be used as a verification set. The conventional and remote sensed observations including quality control are part of the minimum set to restart the analysis cycling system or to start models directly from these initial conditions.

4. NORTH AMERICAN REGIONAL REANALYSIS

The new North American Regional Reanalysis (NARR) will become available at the NCDC NOMADS servers in December 2003 as NCDC will be archiving these data. The NARR is a 32km, 45 layer, EDAS derived system, from 1979 thru 2003, updated monthly. Features include a new land surface model, direct satellite radiance retrievals, and for the first time in any reanalysis, precipitation. It is expected that this data set when available in early 2004, will be widely used across the weather, climate and oceanographic communities.

5. SUMMARY

A new long term archive for climate, weather, and ocean research, and for operational real-time access has been established at NCDC and at NCEP. Access is available using both traditional (Web) and distributed data access methods.

The goals of these new services include input files

for the Regional modeling community; output files for the operational weather and climate community; and reanalysis and observational products for research in several geosciences. The distributed and sub-setting near format neutral philosophy being adopted by the climate, ocean and now weather communities is being supported and advanced by the NOMADS project. The NOMADS and Model Resources Pages at NCDC can be accessed at: www.ncdc.noaa.gov/oa/model/model-resources.html.

6. FURTHER INFORMATION

For further information on the NOMADS project see the 20th International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology (IIPS) paper 6.5: The NOAA Operational Model Archive and Distribution System (NOMADS): "A Status Report", at the 84th AMS Annual Meeting conference.

7. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the combined efforts of all the NOMADS Team members for the growing success of NOMADS. Thanks also to Don Hooper, Jennifer Adams, and Howard Diamond. Special thanks to the internal team at NCDC: Ted Smith, Alan Hall, Ken Schmidt, David Bowman, John Fauerbach and Danny Brinegar. Pilot funding has been provided by ESDIM 02-444E, and the NESDIS CIO's office; and finally direct support from the Director's of NCDC, NCEP, and GFDL.