NOMADS:

A Climate and Weather Model Archive at the National Oceanic and Atmospheric Administration

BY GLENN K. RUTLEDGE, JORDAN ALPERT, AND WESLEY EBUISAKI

Never before has the United States maintained a digital archive for its operational weather models, and an innovative data access philosophy promotes interoperable access across the geosciences.

Historical and real-time observation-based weather data are available through a number of different channels [e.g., the National Climatic Data Center (NCDC), which is responsible for the long-term stewardship of such data]. In contrast, it is generally more difficult to obtain the input and output of high-volume weather and climate models and reanalysis products and data files that are useful in numerous applications (e.g., model improvement, intercomparison, validation and verification, extremes, climatology indices, and observational density spacing and process studies). Although various historical access solutions (Rutledge et al. 1998) exist at various

AFFILIATIONS: RUTLEDGE—National Climatic Data Center, Asheville, North Carolina; ALPERT AND EBUISAKI—NOAA/National Centers for Environmental Prediction, Camp Springs, Maryland
CORRESPONDING AUTHOR: Glenn K. Rutledge, National Climatic Data Center, Federal Building, 151 Patton Avenue, Asheville, NC 28801-5001
E-mail: Glenn.Rutledge@noaa.gov

The abstract for this article can be found in this issue, following the table of contents.

government laboratories and other institutions, a U.S. national archive of operational weather and climate models has simply not existed. To address the need for access to such model data, the NCDC, National Centers for Environmental Prediction (NCEP), and Geophysical Fluid Dynamics Laboratory (GFDL) have initiated the National Oceanic and Atmospheric Administration (NOAA) National Operational Model Archive and Distribution System (NOMADS). The primary goals of NOMADS are to act as a distributed front-end service to NCEP, and to improve access to a suite of popular model datasets and observations, ranging from numerical weather prediction (NWP) fields to Coupled Global Climate Model (CGCM) output to global and regional reanalysis. This paper describes the products currently available in NOMADS, its distributed and interoperable format-neutral approach to data management, and the NOMADS access tools that are currently available.

NOMADS servers exist at several locations in the United States, including the real-time (R/T) server at NCEP (i.e., the R/T NOMADS), the GFDL Data Portal providing NOMADS-compatible services for CGCM output and other data, and the real-time and historical NOMADS service at NCDC, which is the focus of this article.

FIG. 1. The ODC interface.

archive. This increases the possibility for a serially complete archive. Data quality control processes include checks on the raw Gridded Binary (GRIB; WMO 2001a) model data and its associated transmission and file headers to the actual product, thus ensuring the highest possible quality data are being archived and made available for distribution via NOMADS. NCDC works closely with NCEP if header or data errors are detected. Ingested models are also aggregated by model, grid, and forecast projection, which, along with locally generated "index" files (a file and variable identifier utility), greatly increases the access speed of requested files or subsets. NOMADS ingests approximately 250,000 individual grids a day, and with these index and file-level optimizations users can gain online access to any single model variable within seconds. The NCDC NOMADS security and access control is achieved through read-only Network File System (NFS) mounting of the NOMADS publicly accessible servers within NCDC's firewall to the on-line Redundant Array of Independent Disks (RAID) data servers behind the NCDC firewall. The online data currently account for approximately 20 terabytes (TB) of data with a doubling of this capability in the very near term. Offline access to NCDC's mainframe archive is by request only and at the user's request is staged to NOMADS data servers where users can then access the information using NOMADS-distributed technologies. Additional operational NCDC ingest, archive, and quality control process are also performed, including system loading, back-end disk clustering, load balancing, and other backup capabilities to form a 24/7 operation.

THE NOMADS ARCHITECTURE. A new paradigm for sharing data among climate and weather modelers is evolving. It takes advantage of the Internet and relatively inexpensive computer hardware. In this new framework, scientists put their data onto a computer on the Internet. Software running on the computer allows outside users to see not only their local data but also data on other computers running the same software and data format descriptions. NOMADS participants serve their datasets through a client–server relationship, that is, the datasets are Internet-ready and the display is done by the user (their client). The NOMADS approach to data distribution is applicable to a large set of data products, including both observational and numerical model–based files, and can serve users interested in time scales ranging from those associated with synoptic weather features to those related to decadal-to-centennial climate issues.

The operational NOMADS server at NCDC has three independent ingests of model data. Data feeds are compared against each other and repopulated if necessary on a daily basis before reaching the NCDC

ACCESS TOOLS. From the early stages of its initial development the NOMADS team (Rutledge 2001) sought to leverage off existing data access approaches and to advance the concept of distributed format-

neutral data access. In other words, the approach involved drawing upon existing agency and institutional solutions while developing partnerships and promoting the Open Source Project for a Network Data Access Protocol (OPeNDAP) transport protocol [formally called Data Object Design Studio (DODS); Gallagher and Milkowski 1995; Davis and Gallagher 1999]. As a result, NOMADS and other institutions provide access to distributed format-neutral data in several ways and not necessarily under the NOMADS collaboration. There are several user options for accessing NOMADS datasets, available for both advanced and novice users. A NOMADS Users' Guide is available to assist users in accessing NOMADS datasets (available online at http://nomads.ncdc.noaa.gov/guide/). The following access tools form the core of services while using NOMADS:

- The OPeNDAP Hyper-text Transfer Protocol (http) data transport protocol and the associated OPeNDAP Data Connector (ODC) program allows users to search for and retrieve datasets published by OPeNDAP data servers. OPeNDAP servers are located at major institutions around the world and serve a wide variety of data, including weather and climate, satellite imagery, ocean, and other datasets.

The ODC (Fig. 1) is downloaded to a users' desktop and provides a search-and-download capability to import the data into client applications like Grid Analysis and Display System (GrADS), Climate Data Analysis Tools (CDAT), Interactive Data Language (IDL) Ferret, Matlab, SPSS, Excel, and ArcView, or into databases such as Access and Oracle, and plot them with advanced graphics capabilities (many of these desktop tools are more fully described below). OPeNDAP and OPeDNAP-enabled clients are the core technologies to enable format-neutral access across disparate systems and data formats. More information on OPeNDAP and the ODC can be found at the OPeNDAP home page (online at http://opendap.org/).

- The NCDC NOMADS has interactive Web browse, plot, and access tools (see the NOMADS Web pages at http://nomads.ncdc.noaa.gov):

1) "PLOT": The NCDC NOMADS "Web Plotter" originally developed by NCEP provides browse, limited online calculations, on- and offline access to data and variable subsets, long time series, Hovmoeller diagrams, and advanced plot and animation capabilities. "Plot" includes the GrADS display client (Fig. 2);
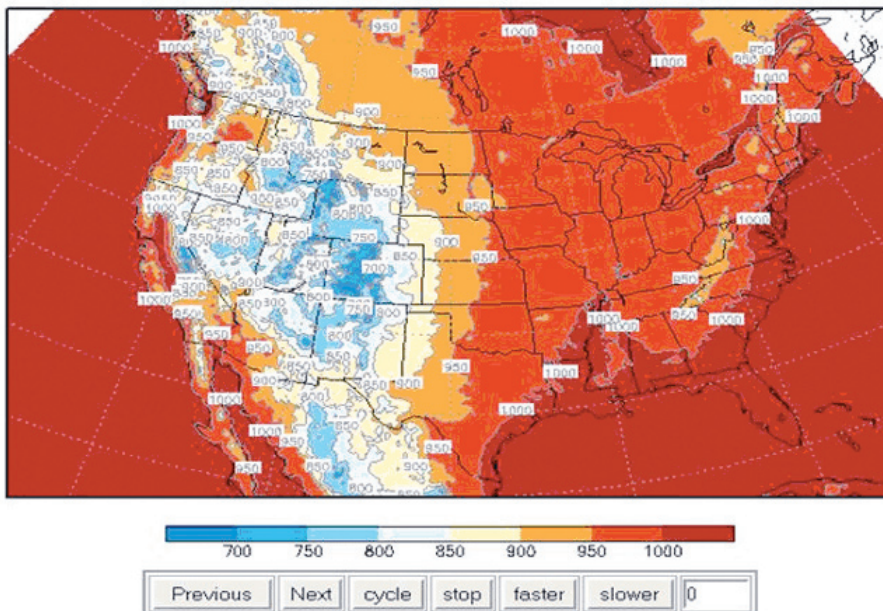

FIG. 2. The NOMADS Web Plotter: A GrADS animation example.

2) "FTP4U": The ftp4u capability (Ebisuzaki et al. 2004) provides access to binary GRIB data by subsetting through time, space, and variables using traditional but enhanced file transfer protocol (ftp) processes. Data are provided in raw GRIB format either online or from NCDC's offline archive. If the data are offline, the user fills out a Web request to move the data from the NCDC archive to the NOMADS online disks. This process can take up to several hours depending on the size of the request, at which time an e-mail notification procedure informs the user to return to NOMADS and access the requested data that are now online (Fig. 3).

3) "HTTP" AND "WGET" AND OPeNDAP CONSTRAINT EXPRESSIONS: The http link provides simple file-level access to download entire directories or files. This access method provides users with a well-known process to collect multiyear requests. The "wget" utility is a freely available network utility to retrieve files from the World Wide Web, using http and ftp, which are the two most widely used Internet protocols. Because wget is a noninteractive downloading utility, one can execute scripts containing wget commands. While http access is familiar to many users, it is highly recommended that users learn to subset their requests either by variable or time period for high-volume requests.

Fig. 3. Example of the "ftp4u" subsetting capability.

- Another very powerful tool for accessing high-volume NOMADS data is using the combination of http and an OPeDNAP "constraint" expression. Users can isolate individual model variables, dates, forecast time steps, and latitude and longitude using a Uniform Resource Locator (URL) to the data, with an OPeDNAP expression request appended to the URL. Then, when placed into a simple UNIX script, this command can find and access multiple forecast projections from many model runs. NCEP has demonstrated this capability to interrogate approximately 10 global ensemble members and 1,200 individual forecast probabilities in less than 1 min over the Internet (Alpert and Wang 2004). The NCEP example script at http://nomad5.ncep.noaa.gov/cgi-bin/var/ensprob2.pl demonstrates how OPeNDAP constraints can be made and then applied to custom situations. The script shows the user the actual composed OPeNDAP queries needed to access the ensemble forecasts and plots out a resultant event probability based on, in this example, the NCEP global ensembles.
- The GrADS Data Server (GDS; Adams et al. 2004) combines both the GrADS (Doty et al. 2001) desktop analysis tool [a freeware client from the Center for Ocean–Land–Atmosphere Studies (COLA)] and the OPeNDAP server to subset and exchange data in many formats with http. GDS data are fully OPeNDAP described to present the raw GRIB and Binary Universal Form for the Representation of Meteorological Data (BUFR; WMO 2001b) as OPeNDAP-compliant Network CDF (NetCDF) files. Users execute an open URL to data rather than keep the data locally. The NCDC and NCEP NOMADS systems use GDS as its primary indexing, inventory, and host-side data manipulation and subsetting tool for users (an example GrADS display is shown as part of the "plot" function in Fig. 2).
- The CDAT (Williams et al. 2002) is an open-source Python-based environment for scientific calculations and graphics with a focus on the needs of climate modelers. CDAT is primarily developed at the Lawrence Livermore National Laboratory (LLNL) in collaboration with the open-source community, which includes other institutions such as the British Atmospheric Data Centre in the United Kingdom and the Laboratoire des Sciences du Climat et de l'Environnement in France. (For more information on CDAT see Fig. 4 and http://cdat.sf.net.)
- The Live Access Server (LAS; Hankin et al. 2001) is a configurable Web server (Fig. 5) designed

to provide flexible access to georeferenced scientific data. It can present distributed datasets as a unified virtual database through the use of OPeNDAP networking. NWP GRIB and BUFR data are accessed by LAS though the GDS OPeNDAP capability. Ferret (Hankin et al. 1996) is the default desktop visualization application used by LAS, though other applications can also be used.

- The Interactive Data Viewer (IDV) from Unidata is a Java-based software framework for analyzing and visualizing geoscience data in two, three, and four (time) dimensions. The IDV brings together the ability to display and work with satellite imagery, gridded data, surface observations, balloon soundings, National Weather Service (NWS) Weather Surveillance Radar-1988 Doppler (WSR-88D) Level II and Level III radar data, and NOAA National Profiler Network data, all within a unified interface (Fig. 6). This desktop client can be downloaded from http://my.unidata.ucar.edu/content/software/IDV/index.html.

The significant advantage that these OPeNDAP-enabled clients and servers have over traditional servers is that they provide direct access to data across multiple servers in the OPeNDAP format-neutral (NetCDF) form. For example, data served by the GDS meet OPeNDAP standards so that the raw GRIB data are seen by clients as a standard NetCDF file. Users execute an open URL to data rather then opening a local file on their disk. For high-volume data requests, data can be requested by variable, temporal, and spatial extents to reduce the overall volume of the request. Further, on many
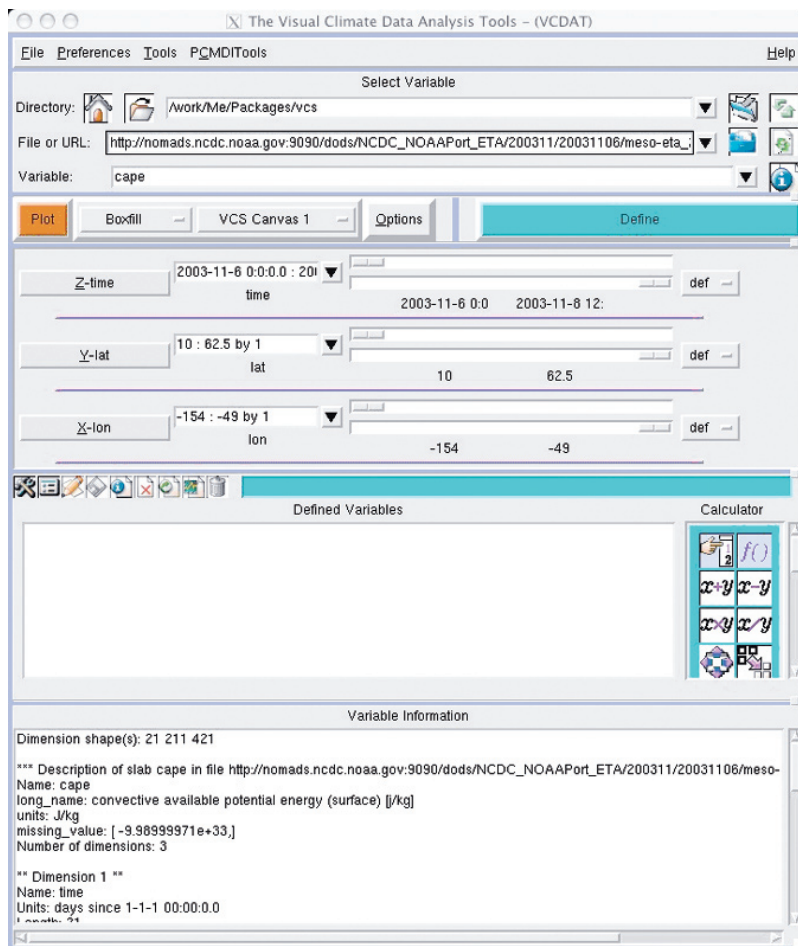

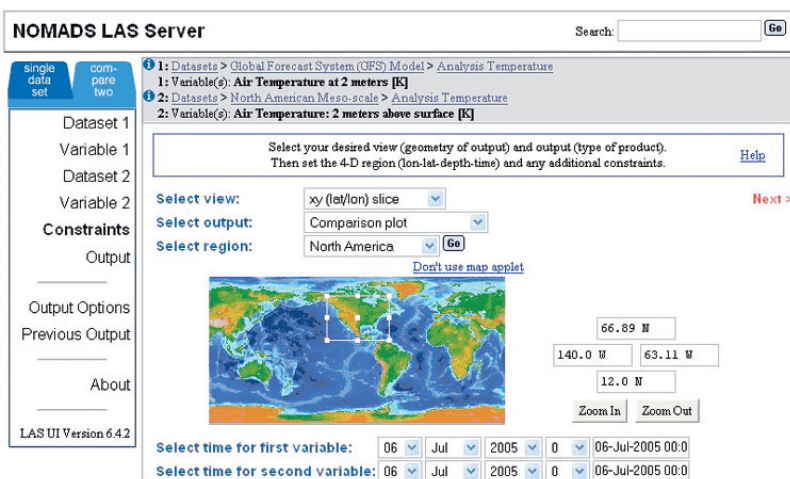FIG. 4. PCMDI's CDAT user interface.

FIG. 5. NOMADS live access server at NCDC: model intercomparison on the fly.

OPeNDAP-enabled servers, and the GDS in particular, users can request server-side computations on these data and even cache the results of a previously calculated value in a temporary directory for use at a later date. To use
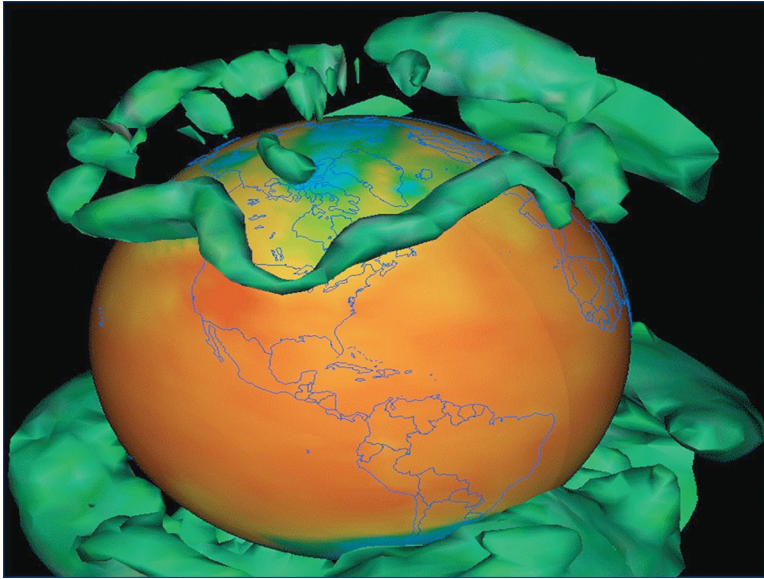
Fig. 6. Unidata's IDV showing upper-tropospheric wind data obtained remotely from a NOMADS server.

the NCDC NOMADS GDS, navigate to the data of interest under the "gds" Web heading using a common Web browser. Following the directory structure based on model and date, locate the OPeNDAP metadata under the "info" section of the OPeNDAP files. Then, copy the "DODS url" (Fig. 7) found there into your client as an appropriate open statement (e.g., in GrADS: "sdfopen"). The client now has all of the necessary information on that data file as to where it is, and how to decode, plot, and retrieve individual data elements contained within that file. Once this information is known, users can then develop scripts to automate the "distributed" access of the desired model data by date, time, model run, or variable.

NOMADS also provides a catalog-level discovery and access component using Unidata's Thematic Real-time Environmental Data Distributed Services (THREDDS) Catalog Service (TCS) system (Domenico et al. 2002), and therefore includes easy access for their OPeNDAP-enabled IDV desktop client. The THREDDS Data Server (TDS), by Unidata, combines a TCS with integrated data-serving capabilities, including OPeNDAP and Web Coverage Services (WCS), with automatic catalog generation. The data format capabilities of the TDS include NetCDF-3, OPeNDAP 2, and Hierarchical Data Format (HDF)5 data models, into what is called the "Common Data Model." NCDC NOMADS has implemented and is testing the latest release of the TDS and expect it to become part of the operational NOMADS for search, use, and access services in the near future. Another catalog service being prototyped at the

NCEP R/T component of NOMADS is the Distributed Metadata Server (DIMES) server for search, discovery, and access (Yang et al. 2001).

Other search and access tools available to the user include metadata entries and direct portal access from the Global Change Master Directory (GCMD), and metadata descriptions as required under the Federal Geospatial Data Committee (FGDC) metadata schema. Formal NCDC archive documentation is also available for NOMADS datasets in text format from NCDC.

DATA AVAILABILITY. The NOMADS data management vision is to make model and associated data and information services uniformly available, providing seamless provision of and access to model input and output data, independent of format and how and where they are collected and stored (Rutledge et al. 2002). The NOMADS at NCDC has been distributing products since 2003 and currently has over 16 TB of online data available for direct online access. Today the NCDC NOMADS provides up to approximately 5 TB of model and observational data representing roughly 1 million downloads per month. Both the NCDC and NCEP NOMADS servers provide access to the model output as soon as it is available from the NCEP computers. The R/T NOMADS (Alpert et al. 2002) servers at NCEP provide short-term real-time services while the NCDC NOMADS provides both historical and real-time data. Additionally, a server located at NCDC provides a backup service to the NCEP server in a "mirror" configuration. This configuration allows researchers to transparently browse, extract, and intercompare online model data from the R/T NOMADS server at NCEP, and if needed, access historical data from the NCDC server forming a 24/7 operation. For a complete listing of data availability on the NOMADS systems at GFDL, NCDC, and NCEP, see the NCDC NOMADS pages online at http://nomads.ncdc.noaa.gov. Table 1 provides the NWP data at NCDC while Table 2 provides their associated period of record. The products currently available on NOMADS include the following.

NCEP NWP models.
• Global Forecast System (GFS) (Kanamitsu 1989; Kanamitsu et al. 1991; Iredell et al. 2002),

- North American Mesoscale (NAM; formerly Eta Model) (Black 1994), includes NCEP variables for the Environmental Protection Agency (EPA) Community Multi-scale Air Quality (CMAQ) model,
- Rapid update cycle (RUC) (Benjamin et al. 2003),
- NCEP Spectral Statistical Interpolation (SSI) Global Data Assimilation System (GDAS) model input (sigma) and restart files (Du et al. 2004),
- NCEP North American Regional Reanalysis (NARR) (Mesinger et al. 2006),
- NCEP–National Center for Atmospheric Research (NCAR) R1 and R2 global reanalysis and Atmospheric Model Intercomparison Project (AMIP)-II (Kalnay et al. 1996; Kistler et al. 2001; Kanamitsu et al. 2002),
- Climate Data Assimilation System (CDAS)(Kalnay et al. 1996),
- NCEP Regional Special Model (RSM) (Juang et al. 1997),
- NCEP global ensembles (Toth et al. 2002, 2005; Zhu et al. 2002),
- NCEP Short Range Ensemble Forecasts (SREF) (Du et al. 2004),
- NCEP Climate Forecast System (CFS) coupled climate model (Saha et al. 2005, manuscript submitted to *J. Climate*),
- NCEP Ocean Wave (Tolman 1999),
- Paleoclimate Model Intercomparison Project (PMIP) data (Anderson et al. 1989),
- NWS National Digital Forecast Database (NDFD) (Glahn and Ruth 2003).

[The NCEP operational suite of models are updated at least yearly. For the latest information on model numerics and other documentation see the NCEP documentation pages at www.emc.ncep.noaa.gov/ *and* www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html; *and the NWS Technical Procedures Bulletin (TPB) updates at* www.emc.ncep.noaa.gov/gmb/STATS/html/tpblist.html.]

In situ observations.
- NCDC Global Historical Climate Network (GHCN) surface temperature and precipitation anomalies (Peterson and Vose 1997),

FIG. 7. GDS OPeNDAP "Info" and "url" directory information.

- NCDC Integrated Global Radiosonde Archive (IGRA) upper-air reference quality dataset [formerly the Comprehensive Aerological Data Set (CARDS) (Durre et al. 2005)],
- NCDC Smith–Reynolds Extended Reconstructed Sea Surface Temperatures (ERSST) and climatologies (Smith and Reynolds 2004).
- NOAA Optimum Interpolation Sea Surface Temperature Analysis (Reynolds et al. 2002).

Satellite and radar observations.
- National Ocean Data Center (NODC) Advanced Very High Resolution Radiometer (AVHRR), Pathfinder sea surface temperature (SST) analysis datasets are provided through a link to NODC (Kilpatrick 2001).
- Other satellite and radar data access using OPeNDAP are currently being developed at NCDC, including a real-time and historical radar data access capability; a limited Geostationary Operational Environmental Satellite (GOES) satellite data access pilot; and, currently under development, the National Aeronautics and Space Administration (NASA) Geodynamics Experimental Ocean Satellite (GEOS)-4 finite-volume general circulation model under the NOMADS collaboration (see information online at http://map05.gsfc.nasa.gov/).

TABLE 1. NCEP NWP real-time and historical availability.

Model availability on NCDC and NCEP NOMADS						
Model	**Grid****	**Domain**	**Resolution**	**Cycles**	**Forecasts**	**Variables and Levels**
NAM Meso-Eta Model	218	United States	12.19 km	hourly 6 h	3–60 h	141 total variables, with 39 vertical levels
GFS	003	Global	1° lat/lon	6 h	3–180 h	139 total variables, with 26 vertical levels
Ensembles (hi/lo res)	3	Global	2.5° lat/lon and 1° lat/lon	6 h	6–180 h	134 total variables, with 26 vertical levels
SREF	212	United States	45 km	12 h at 0900 and 2100 UTC	3–63 h	169 total variables, with 39 vertical levels
CFS*	2	Global	2.5° lat/lon	0000 UTC 1–3, 9–13, 19–23 Jan last day: 1981–2003	Monthly avg Jan–Sep	64 total variables with 5 vertical levels
Global 0.5*	4	Global	0.5° lat/lon	6 h	3–180 h	163 Total variables, with 42 vertical levels
RUC-20	252	United States	20/40/80 km	1 h	1–12 h	92 total variables, with 37 vertical levels
NDFD	221	United States	~5.0 km	1 h	1–21 h	Surface only, 6 variables (to date)
NCEP NARR available on NCDC NOMADS						
Model	**Grid**	**Domain**	**Resolution**	**Cycles**	**Forecasts**	**Variables and Levels**
NARR	221(a)	North America	32.46 km	3 h	None	29 vertical levels, with 187 total variables
NARR	221(b)	North America	32.46	3 h	3 h	29 vertical levels, with 187 total variables
NARR monthly means A	221(a)	North America	32.46	3 h	None fields	Monthly avg of all analysis
NARR monthly means B	221(b)	North America	32.46	3 h	3 h	Monthly avg of all forecast fields
NARR monthly 3-h means A	221(a)	North America	32.46	3 h	None	Monthly 3-hourly time step avg of all analysis fields (0000–0300, 0300–0600 UTC, etc.)
NARR monthly 3-h means B	221(b)	North America	32.46	3 h	3 h	Monthly 3-hourly time step avg of all forecast fields

*The real-time data are available at NCEP and available for up to 2 weeks. These data, at the time of this writing, are being provided to NCDC for long-term preservation and NOMADS access.

**NCEP Office Note #388 (available at www.nco.ncep.noaa.gov/pmb/docs/.)

GFDL Coupled Global Climate Models (CGCM).
- GFDL Climate Model (CM) 2.0 experiments: the current generation model (Delworth et al. 2006),
- GFDL CM2.1 experiments (Delworth et al. 2006),
- GFDL_R30-c climate model experiments: previous generation model (Delworth et al. 2002),
- GFDL_R15_b climate model experiments: an older model (Dixon et al. 2002).

The GFDL climate models are available from GFDL online at http://nomads.gfdl.noaa.gov/.

As of 0000 UTC 25 May 2005, the NOAAPort ingest for the NOMADS NCEP NWP data archive was

Period of record for NOMADS datasets (as of 8 Aug 2005)			
NOAAPort (historical only)			
Model System	**Grid+**	**Online availability***	**Availability from deep archive**
GFS (Aviation Model)	All	1 Jun 2003–25 May 2005	1 May 2002–31 May 2003
GFS (Medium-Range Forecast)	201	1 Jun 2003–25 May 2005	1 May 2002–31 May 2003
	202		
	203	1 July 2003–25 May 2005	27 Jan 2003–31 May 2003
	205		
NAM early Eta Model	212	1 Jun 2003–25 May 2005	27 Jul 2002–31 May 2003
NAM Meso-Eta Model	211	1 Jun 2003–25 May 2005	27 Jul 2002–31 May 2003
	212		
	215		
	218	1 Jun 2003–25 May 2005	2 Oct 2002–31 May 2003
Rapid Update Cycle (RUC)	211/236	1 Jun 2003–25 May 2005	27 Jul 2002–31 May 2003
Real time and historical			
NAM Meso-Eta High resolution	003	1 Jan 2005–present	1 Mar 2004–31 Dec 2004
GFS high resolution	218	1 Jan 2005–present	1 Mar 2004–31 Dec 2004
North American Regional Reanalysis	221	1 Jan 1979–present (2005 in processing)	This dataset is completely online
Ensembles		Real time**	Currently being ingested
SREF		**	
CFS		**	
Global 0.5		**	
RUC-20	252	31 Oct 2005–present	
GDAS SSI model input		**	

*Reduced variable and resolution NOAAPort grids were discontinued on 5 May 2005 and replaced by higher-resolution NCEP NAM and GFS, grids 003 and 218.

**The real-time data are available at NCEP and available for up to 2 weeks. These data, at the time of this writing, are being provided to NCDC for long-term preservation and NOMADS access.

+NCEP Office Note #388

discontinued. These reduced resolution and variable grids were removed in favor of the higher-resolution and variable NCEP GFS and NAM grids.

Some of the most popular NOMADS datasets are more fully described below.

NCEP ENSEMBLES. The GFS ensemble forecasts are available from 0 to 7 days at 6-h intervals on a 1° × 1° grid; and 0–16-day forecasts at 6-h intervals are on a 2.5° × 2.5° grid. As noted earlier, as an example of an OPeNDAP service, the NOMADS team has developed a client application that accesses the NCEP global ensemble model forecast data to produce user-selected weather element event probabilities. Probability estimates can be defined simply as the percentage of ensemble forecasts of the total number that satisfy a specified event for some weather element. Ensembles are composed of many model realizations and access many large files for small amounts of data at specific locations, and for certain variables can be efficiently accomplished by the NOMADS GDS. In other words, the user does not have to download all of the high-volume files just to access the information contained in the one variable. The event probabilities are easily extended over ensemble model forecast times to show probability histograms defining a product for the forecast probability of user-

selected events. For access to the NOMADS ensemble event probability page see http://nomad5.ncep.noaa.gov/cgi-bin/var/ensprob2.pl.

THE NCEP CLIMATE FORECAST SYSTEM (CFS). The NCEP Climate Forecast System is a monthly mean time series of 53 commonly used variables from the retrospective forecasts of the NCEP Climate Forecast System (Saha et al. 2005, manuscript submitted to *J. Climate*) that have been archived to the NCEP NOMADS public servers in GRIB format. These include forecast monthly means of all 15 members initialized in each calendar month of the year, covering a period of 23 yr from 1981 to 2003. They also include an ensemble mean of the 15 members, and the verifying observations. The forecast lead time is from 0 to 8 months. The data can be found on the NCEP NOMADS server at http://nomad6.ncep.noaa.gov/.

NCEP NARR. The NCDC NOMADS is NOAA's primary access point for the new NARR. NARR is a reprocessing of the historical meteorological observations using NCEP's regional forecast model and associated 3D variational data assimilation (3DVAR) system. The products of NARR will be a new set of meteorological analyses covering the North American domain with a 32-km horizontal resolution, 3-h temporal resolution, and 50-hPa vertical resolution for October 1978 to the present. A "merged" dataset based on the analyses plus fluxes on the AWIPS grid is available on NOMADS at approximately 5-TB (60 MB every 3 h) volume. The NARR Monthly Means (NARRMON) dataset contains a monthly average (computed at NCEP) of all the fields in NARR. The NARRMON 3-h dataset is a monthly average of all fields, separated into eight 3-h time periods for each day throughout a month. See Table 1 for more information.

Of note for NARR users: the data for each analysis time is split into two files so that the data will be compatible with the software program GrADS. For example, the "narr-b" file contains a 3-h forecast of the sensible heat flux whereas the "narr-a" file contains the average from the 0–3-h forecast. For example, suppose we set GrADS to have only one time variable to read analysis and forecast files:

- file A: analysis at 0000 UTC 1 January 2004,
- file B: 3-h forecast valid at 0000 UTC 1 January 2004.

GrADS allows you to combine datasets together so one can make a time series. If NOMADS combined all of the "A" files together, GrADS users would set the time to 1800 UTC 1 January 2004 and make a successful plot. On the other hand, if the A and B files were combined together, the GrADS client would not be able to distinguish between the analyses or the forecast valid time. To avoid this GrADS limitation, all of the data were put into the A files, except for the forecasts that would cause the timing confusion. Most users will only require the A files, while those users requiring a hydrological analyses increment, for example, will require the B files.

GFDL CGCM. The GFDL's CGCMs are available under the NOMADS framework, including model output and documentation of experiments performed using GFDL's current generation of global coupled climate models named GFDL CM2.0 and GFDL CM2.1. These two models are being applied to decadal-to-centennial (deccen) time-scale climate issues, including multicentury control experiments and climate change projections. To learn more about the GFDL CM2.0 and CM2.1 models and their output, one may visit http://nomads.gfdl.noaa.gov/CM2.X.

The GFDL CM2.0 and CM2.1 medium-resolution climate models consist of atmosphere, ocean, sea ice, and land surface model components coupled together. The two CM2.x models differ in some of their dynamics and physical parameterizations. While the atmosphere and land surface components use a different grid than do the ocean and sea ice components, both CM2.0 and CM2.1 share the same grid resolutions. The archived atmosphere and land surface model output is on a grid with approximately 2° grid spacing in the horizontal. While the atmospheric GCM utilizes 24 atmospheric levels, the vast majority of the atmospheric model output is stored on 17 standard pressure levels. The ocean and sea ice model grid resolution is approximately 1° in the horizontal with higher resolution in the Tropics, and "tripolar" elements in the Arctic. The ocean model component has 50 vertical levels.

Model output from 20 experiments (10 each for CM2.0 and CM2.1) are accessible from the GFDL Data Portal (see online at http://nomads.gfdl.noaa.gov/). These experiments were driven by forcing agents consistent with those requested by the Intergovernmental Panel on Climate Change (Houghton et al. 2001) for their fourth assessment report (AR4), and are applicable to research projects associated with the U.S. Climate Change Science Program (CCSP). The 10 experiments include the following:

- one preindustrial (circa 1860) control (300 yr of output available),
- three "Climate of the 20th Century" experiments (140 yr each; simulating 1861–2000),
- three Special Report on Emissions (SRES) climate change projections (A2, A1B, and B1 scenarios) in which forcing agents vary from 2001 to 2100 (the A1B and B1 experiments continue to year 2300 with forcing agents stabilized at 2100 levels),
- one "committed climate change" 100-yr-long model simulation, with forcing agents stabilized at year 2000 levels,
- two idealized forcing integrations in which atmospheric CO_2 increases 1% yr^{-1} to twice (the 2X run) and four times (the 4X run) its initial value, and is then held constant for a minimum of 150 additional years (220–300-yr experiment durations).

Some model output files from GFDL's older R30 and R15 climate model integrations (Delworth et al. 2002) are still available for those interested in previous generations of GFDL coupled climate models. See the GFDL Web site for more information on all these models integrations.

OTHER APPLICABLE TECHNOLOGIES AND PROJECTS. To respond to these changing requirements and mission goals a flexible user access infrastructure must become increasingly robust, responsive, efficient, secure, adaptable, and cost effective. NOMADS users have the capability to subset high-volume data; however, some researchers and others require access to entire suites of high-volume model data. One such effort that augments NOMADS for access to high-volume model data is the Earth System Grid (ESG) (Foster et al. 2001). The Department of Energy's (DOE's) Program for Climate Model Diagnostics and Intercomparison (PCMDI), NCAR, and other DOE sites initiated and continue to develop GRID-based intelligent filing systems and data management software to link storage devices located throughout the United States and the international climate research community. ESG uses the Globus Toolkit (Foster et al. 2004) grid technology among other services and data location and management techniques developed especially for climate and other high-data-volume users. The latest U.S. climate models from GFDL and NCAR for the Intergovernmental Panel of Climate Change (IPCC) Model Intercomparison Project (MIP) are being accessed through the ESG at PCMDI. GRID- and Web-based "services" have emerged as viable technologies that are increasingly utilized by government laboratories, corporate

institutions, and high-performance computing centers around the world. GRID and Web services are being developed and used in diverse applications such as high-energy physics, medical imaging, meteorology, and business applications. Recently completed efforts to define GRID software standards that align with standards from the World Wide Web, and Globus's adherence to these standards, have strengthened the interest and commitment of industry toward GRID computing (M. Govett, M. Doney, and P. Hyder 2004, unpublished manuscript).

It is the opinion of the authors that Web-based services will become an ingrained part of data access and computation in the near future as "Web services" using Extensible Markup Language (XML) (Bray et al. 1998) will become as transparent as http is today. One such exploratory effort—SciFlow (Yunck et al. 2004)—is a good example of distributed Web- and GRID-based service. Other data access efforts for high-volume model data access are being provided through Web and GRID-based portals, including the NCAR Community Data Portal (CDP) (online at https://cdp.ucar.edu/); NASA's Earth science community, especially through the Earth Systems Information Partners (ESIP) program; and the Earth System Modeling Framework (ESMF), an outstanding modular model coding project (online at www.esmf.ucar.edu/). Finally, international efforts, especially in Europe (Lawrence et al. 2004), are advancing interoperable data and system access and computation for both the GRID and Web.

NOMADS is monitoring the direction that the open-source and Web and GRID services communities are taking and leveraging these new tools as they develop. NOMADS is an active participant in many national and international exploratory and directed efforts, and has been selected as a candidate architecture to promote under the NOAA's Data Management Integration Team (DMIT) Global Earth Observations Integrated Data Environment (GEO-IDE) plan, and cited in plans developed by the Integrated Earth Observation System Architecture and Data Management Working Group (ADM), of the U.S. Group on Earth Observations (US-GEO). NOMADS is also involved in the Committee on Earth Observation Satellites (CEOS, online at www.ceos.org) CEOS GRID project; the Linked Environments for Atmospheric Discovery (LEAD) project (Droegemeier et al. 2004); and members and contributors to the Global Organization for Earth System Science Portal (GO-ESSP, online at http://essp.gfdl.noaa.gov). GO-ESSP is an international grassroots effort that emerged from the original NOMADS collaborator's to address distributed data

access to models and associated data, and actively works to resolve modelers' needs for interdisciplinary research and access. More information on GO-ESSP can be found online at http://go-essp.gfdl.noaa.gov/. Finally, NOMADS is currently working to include other model datasets, namely, from NASA (a new collaborator), as well as other model data, including the land (hydro) and ocean modeling communities.

INTEROPERABLE DATA MANAGEMENT— THE NOMADS GOALS AND VISION.

Beyond improved and distributed access to models the fundamental issue that NOMADS seeks to address is how NOAA and its partners can organize its data files from its distributed climate and weather models and related observational data into a cohesive presence that facilitates real-time and retrospective climate and weather model analysis and intercomparisons. For the first time NOMADS users can access a historical long-term suite of model input, output, and observations to analyze and improve climate change and detection processes, to improve short-term NWP and seasonal forecasts, and to improve long-term global climate simulations under a distributed client–server framework. The goals of NOMADS are to

- improve access to NWP and GCM models and provide the observational and data assimilation products for regional model initialization and forecast verification,
- improve operational weather forecasts,
- develop linkages between the research and operational modeling communities and foster collaborations between the climate and weather modeling communities,
- promote product development and collaborations within the geoscience communities (ocean, weather, and climate) to study multiple earth systems using collections of distributed data under a sustainable system architecture, and
- act as a resource for the development of a long-term framework for systematic approaches to climate change detection efforts, climate and weather model evaluation, impacts studies, and other process studies.

To achieve this, NOMADS uses the OPeNDAP protocol and common data schema conventions, and suggests that new participants use and advance this service methodology and that data generators' and scientists work to provide their data in one of the many applicable data schema and formats.

The NOMADS has been developed to address model data access needs as outlined in the U.S. Weather Research Program (USWRP) "Implementation Plan for Research in Quantitative Precipitation Forecasting and Data Assimilation," which states a need to "redeem practical value of research findings and facilitate their transfer into operations." The NOMADS framework was also developed to facilitate model and observational data intercomparison issues as discussed in documents such as the Intergovernmental Panel on Climate Change (IPCC; Houghton et al. 2001), and the CCSP (online at www.climatescience. gov). Finally, NOMADS directly addresses goals as outlined in the National Research Council (NRC 2003) report "Fairweather: Effective partnerships in weather and climate services" recommendation number five that states

> The NWS should make its data and products available in Internet-accessible digital form. Information held in digital databases should be based on widely recognized standards, formats, and metadata descriptions to ensure that the data from different observing platforms, databases, and models can be integrated and used by all interested parties in the weather and climate enterprise.

The NOMADS data management vision is to make model data and information services uniformly available to all elements of NOAA, including the public and the operational and research communities, thus providing seamless provision of an interoperable access to model data independent of format, or how and where they are collected and stored. NOMADS has been developed as a unified climate and weather archive providing Web access to information so that users can make decisions about their specific research, operational, and education needs. This takes place on time scales from days (weather), to months (El Niño), decades, and centuries (climate change).

CONCLUSIONS.

A new paradigm for sharing data among climate and weather modelers is evolving. It takes advantage of the Internet and relatively inexpensive computer hardware. NOMADS participants serve their datasets through a client–server relationship in Internet-ready form. Both researchers and policy makers alike now expect our national data assets to be easily accessible and interoperable, regardless of their physical location. As a result, an effective interagency distributed data service requires the coordination of data infrastructure and management extending beyond traditional organizational boundaries.

NOMADS user access has seen a tremendous level of growth since its inception in 2001. The original

vision to improve the linkages between the research and operational modeling communities has been extended to foster new collaborations between and among the climate and weather communities. It is hoped that these collaborations will not only promote product development but new collaborations within the geoscience communities (ocean, weather, and climate) to study multiple earth systems using collections of distributed data under an integrated sustainable system architecture.

REFERENCES

Adams, J. M., J. Wielgosz, and B. E. Doty, 2004: GRADS version 1.9 and the GRADS-DODS Server. Preprints, *20th Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, 9.13.

Alpert, J. C., and J. Wang, 2004: The real time NOMADS project: Access to operational model data and value added products. *Proc. 20th Conf. on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, P1.25.

——, G. K. Rutledge, D. Williams, R. Stouffer, L. Buja, B. Doty, S. Hankin, B. Domenico, and M. Kafatos, 2002: The plan to access real-time NWP operational model data sets using NOMADS. *Proc. 18th Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Orlando, FL, Amer. Meteor. Soc., 73–75.

Anderson, D. W., W. L. Prell, and N. J. Barratt, 1989: Estimates of sea surface temperature in the Coral Sea at the last glacial maximum. *Paleoceanography,* **4,** 615–627.

Benjamin, S., G. Grell, J. Brown, and T. Smirnova, 2003: Mesoscale weather prediction with the RUC hybrid isentropic/terrain-following coordinate model. *Mon. Wea. Rev.,* **132,** 473–494.

Black, T., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting,* **9,** 265–278.

Bray, T., J. Paoli, and C. M. Sperberg-McQueen, 1998: Extensible Markup Language (XML) 1.0 specification, W3C REC-xml-19980210. [Available online at www.w3.org/XML/.]

Davis, E. R., and J. Gallagher, 1999: Using OPENDAP to access and deliver remote data. *Proc. 15th Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Dallas, TX, Amer. Meteor. Soc., 571–573.

Delworth, T. L., R. J. Stouffer, K. W. Dixon, M. J. Spelman, T. R. Knutson, A. J. Broccoli, P. J. Kushner, and R. T. Wetherald, 2002: Review of simulations of climate variability and change with the GFDL R30 coupled climate model. *Climate Dyn.,* **19,** 555–574.

——, and Coauthors 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate,* in press.

Dixon, K., T. Delworth, T. Knutson, M. Spelman, and R. Stouffer, 2002: A comparison of climate change simulations produced by two GFDL Coupled Climate Models. *Global Planet. Change,* **37,** 81–102.

Domenico, B., J. Caron, E. Davis, R. Kambic, and S. Nativi, 2002: THREDDS: Incorporating real-time environmental data and interactive analysis tools Into NSDL. *J. Digital Inf.,* **2,** xx–xx.

Doty, B. E., J. Wielgosz, J. Gallagher, and D. Holloway, 2001: GrADS and OPENDAP. *Proc. 17th Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Albuquerque, NM, Amer. Meteor. Soc., 385–387.

Droegemeier, K., and Coauthors, 2004: Linked environments for atmospheric discovery (LEAD): A cyber-infrastructure for mesoscale meteorology research and education. *Proc. 20th Conf. on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, S6.1.

Du, J., and Coauthors, 2004: The NOAA/NWS/NCEP short-range ensemble forecast (SREF) system: Evaluation of an initial condition vs multi-model physics ensemble approach. Preprints, *16th Conf. on Numerical Weather Prediction,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, 21.3.

Durre, I., R. S. Vose, and D. B. Wuertz, 2005: Overview of the integrated global radiosonde archive. *J. Climate,* **19,** 53–68.

Ebisuzaki, W., J. Alpert, J. Wang, D. Jovic, and P. Shafran, 2004: North American Regional Reanalysis: End user access to large data sets. *Proc. 20th Conf. on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, 8.2.

Foster, I., and C. Kesselman, 1997: Globus: A metacomputing infrastructure toolkit. *Int. J. Supercomput. Appl.,* **11** (2), 115–128.

——, C. Kesselman, and S. Tuecke, 2001: The anatomy of the Grid: Enabling scalable virtual organizations. *Int. J. Supercomput. Appl.,* **15,** 200–222.

——, J. Nick, and S. Tuecke, 2002: Grid services for distributed system integration. *IEEE Comput.,* **35** (6), 37–46.

Gallagher, J., and G. Milkowski, 1995: Data transport within the distributed oceanographic data system. *Fourth Int. World Wide Web Conf.,* MIT Laboratory for Computer Science. [Available online at www.w3.org/Conferences/WWW4/.]

Glahn, H. R., and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.,* **84,** 195–201.

Hankin, S., D. E. Harrison, J. Osborne, J. Davison, and K. O'Brien, 1996: A strategy and a tool, FERRET, for closely integrated visualization and analysis. *J. Visualization Comput. Animation,* **7,** 149–157.

——, J. Callahan, and J. Sirott, 2001: The live access server and OPeNDAP: Web visualization and data fusion for distributed holdings. *Proc. 17th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* Albuquerque, NM, Amer. Meteor. Soc., 380–383.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis.* Cambridge University Press, 892 pp.

Iredell, M., H. Pan, and P. Caplan, 2002: Changes to the 2002 NCEP Operational MRF/AVN Global Analysis/Forecast System. National Centers for Environmental Prediction, Global Modeling Branch. [Available online at www.emc.ncep.noaa.gov/gmb/moorthi/gam.html.]

Juang, H.-M. H., S.-Y. Hong, and M. Kanamitsu, 1997: The NCEP regional spectral model: An update. *Bull. Amer. Meteor. Soc.,* **78,** 2125–2143.

Kalnay E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 335–342.

——, and Coauthors, 1991: Recent changes implemented into the Global Forecast System at NMC. *Wea. Forecasting,* **6,** 425–435.

——, W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643.

Kilpatrick, K. A., G. P. Podesta, and R. Evans, 2001: Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.-Oceans,* **106** (C5), 9179–9197.

Kistler, R., and Coauthors, 2001: The NCEP/NCAR 50-year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **92,** 247–267.

Lawrence, B., and Coauthors, 2004: "Googling" Secure Data. *Proceedings of the U.K. e-science All Hands Meeting,* S. J. Cox, Ed.

Mesinger, F., and Coauthors, 2005: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* in press.

NRC, 2003: Committee on Partnerships in Weather and Climate Services, Committee on Geophysical and Environmental Data. *Fair Weather: Effective Partnerships in Weather and Climate Services,* National Academies Press, 5.

Peterson, T. C., and R. S. Vose, 1997: An overview of the Global Historical Climatology Network temperature data base. *Bull. Amer. Meteor. Soc.,* **78,** 2837–2849.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate,* **15,** 1609–1625.

Rutledge, G. K., 2001: NOMADS. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology,* W. Zwieflhofer and N. Kreitz, Eds., World Scientific, 269–276.

——, V. Baer, G. DiMego, and J. Hayes, 1998: A user-driven data requirements and assessment approach for AWIPS/NOAAPORT. *Proc. of the 14th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology,* Phoenix, AZ, Amer. Meteor. Soc., 392–395.

——, and Coauthors, 2002: The NOAA Operational Model Archive and Distribution System (NOMADS). *Proc. 13th Symp. on Global Change and Climate Variations,* Orlando, FL, Amer. Meteor. Soc., J76–J78.

Smith, T. M., and R. W. Reynolds, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate,* **17,** 2466–2477.

Tolman, H. L., 1999: User manual and system documentation of WAVEWATCH-III version 1.18. NOAA/NWS/NCEP/OMB Tech. Note 166, 110 pp.

Toth, Z., O. Talagrand, G. Candille, and Y. Zhu, 2002: Probability and ensemble forecasts. *Environmental Forecast Verification: A Practitioner's Guide in Atmospheric Science,* I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 137–164.

——, ——, and Y. Zhu, 2005: The attributes of forecast systems: A framework for the evaluation and calibration of weather forecasts. *Predictability of Weather and Climate,* T. Palmer and R. Hagedorn, Ed., Cambridge University Press, in press.

Williams, D. N., R. S. Drach, P. F. Dubois, C. Doutriaux, C. J. O'Connor, K. M. AchutaRao, and M. Fiorino, 2002: Climate data analysis tool: An open software system approach. *13th Symp. on Global Change and Climate Variations,* Orlando, FL, Amer. Meteor. Soc., CD-ROM, J71.

WMO, 2001a: FM 92-VIII Ext. GRIB (Gridded Binary). Manual on Codes, WMO-306, Vol. 1.2, Part B.

——, 2001b: FM 94–VIII Ext. BUFR (General Regularly-Distributed Information) in Binary Form. Manual on Codes, WMO-306, Vol. 1.2, Part B.

Yang, R., X. Deng, M. Kafatos, C. Wang, and X. S. Wang, 2001: An XML-based Distributed Metadata Server (DIMES) supporting earth science metadata. *Proc. of the 13th Int. Conf. on Scientific and Statistical Database Management,* Fairfax, VA, IEEE, 251–256.

Yunck, T., B. Wilson, A. Braverman, E. Dobinson, and E. Fetzer, 2004: GENESIS: The General Earth Science Investigation Suite. *NASA Earth Science Technology Conf. 2004,* Palo Alto, CA, NASA. [Available online at http://esto.nasa.gov/conferences/estc2004/papers/alpl.pdf.]

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: On the economic value of ensemble based weather forecasts. *Bull. Amer. Meteor. Soc.,* **83,** 73–83.