P2.34 ECMWF 45-YEAR REANALYSIS DATA FROM NCAR

Joseph L. Comeaux II* and Steven J. Worley
National Center for Atmospheric Research, Boulder, Colorado

1. INTRODUCTION

Global atmospheric reanalysis projects have proven to be a major advancement for studies in meteorology, oceanography, climatology and other fields. Reanalyses use fixed data assimilation systems that eliminate inconsistencies found in archives of operational model output caused by changes to the operational models. These projects supply gridded data for early periods that were deficient (e.g. 1957-1969) or entirely lacking, and consistently apply advanced data assimilation systems over longer periods than was previously possible. Reanalysis projects have benefited from data recovery efforts undertaken at several institutions, including the National Center for Atmospheric Research (NCAR). These efforts have provided larger and higher quality input archives than were available for most operational models, particularly in the older periods. The resulting data products from reanalyses are a rich data resource, which have been extensively used by thousands of scientists (Kistler et al, 2001).

The European Centre for Medium-Range Forecasts (ECMWF) Re-Analysis (ERA-40) project has produced a comprehensive global analysis for the 45 year period covering September 1957 to August 2002. The project used the CY23r4 version of the ECMWF Integrated Forecast System (IFS) that was first employed in production runs on June 12, 2001. The ERA-40 version of the atmospheric model had 60 vertical levels, T159 spherical harmonic representation for the basic dynamical fields, and a reduced Gaussian grid with an approximate uniform spacing of 125km for surface and other fields. The atmospheric model was coupled to an ocean-wave model which resolved 25 wave frequencies and 12 wave directions at the nodes of a 1.5º grid. Multiple archives of *in situ* and satellite observations were assimilated by the model. These input data streams represented one of the largest and most complete collections of observations ever assembled. The ECMWF ERA-40 Project Report Series, available from http://www.ecmwf.int/publications/library/do/references/list/192, provides a more detailed description.

As a collaborating partner, the Data Support Section (DSS) of the Scientific Computing Division (SCD) at NCAR is the sole distributor for ERA-40 data products in the United States and some locations in Canada. All data are available from NCAR for non-commercial research only. Further information and access to the NCAR ERA-40 archive are available at http://dss.ucar.edu/pub/era40.

* Corresponding author address: Joseph L. Comeaux II, Nat'l Ctr. for Atmospheric Research, Boulder, CO 80307-3000; email: **joey@ucar.edu.**

2. ACCESS

The complete ERA-40 archive is over 33 Terabytes (TB) with most individual products ranging from 1-2 TB. The total primary analyzed data are roughly 5 TB. The primary goal of DSS is to provide efficient access to the entire ERA-40 archive. Thus, a delivery system has been designed that promotes access to both large and small amounts of data. It is straightforward to obtain complete individual products, data for short time periods, and subsets of limited variables and levels.

The most popular products can be obtained by directly downloading the archive files using the web or File Transfer Protocol (FTP). These monthly files contain all available fields and levels for the individual products. A web interface is provided for each online product that allows for single or multiple month data selections and the choice of web or FTP transfer. Web transfers are performed in real time and are best suited for small data requests. For FTP, users are first provided all needed download information through an e-mail message, after which the transfer is initiated by the user. The FTP transfers work well for both small and large data requests.

Access using subset extraction is also available for most online products. Through a web interface users can select single or multiple variables and levels, and time period. The output data can be sorted either synoptically or into time series by variable and level. These requests are processed in a delayed mode and generally take 1-3 days to complete, but may take up to 1-2 weeks, depending on the size and number of requests received. When a request is completed the user is notified by e-mail and provided all information needed to access the data, which must be transferred by FTP.

Some products are not available through the direct download or subset extraction methods. Typically, these products can be staged to a server for FTP transfer. In most cases, variable and level extraction is also possible. These individual requests are initiated and arranged through e-mail to the data specialist, who are identified throughout the online documentation, or by contacting Joey Comeaux, (joey@ucar.edu).

In cases where access to the ERA-40 archive through the Internet is not practical, the data may be copied to writeable media. Media options will be commensurate with the data request size and will be arranged by the data specialist.

Finally, the entire ERA-40 archive is also available on the NCAR Mass Storage System (MSS). Access to the MSS is restricted to users with NCAR computer accounts as provided by SCD. To apply for an account, please contact Database Services at dbs@ucar.edu or visit the web page at http://www.scd.ucar.edu/docs/access/accounts.html.

3. DATA PRODUCTS

The ERA-40 archive is divided into roughly 30 distinct data products. The tables and descriptions below group the data products into five categories and define the products by name, number of levels, number of variables, data volume for the 45-year period, dataset identification number, and available access options. The dataset identification number can be used in a standard web address template (http://dss.ucar.edu/datasets/dsnnn.n) to locate the data product at NCAR. The dataset web pages provide detailed descriptions of each product including metadata and supporting software and enable access to online data, if available.

3.1 Model Resolution Analysis Fields

These data products (Table 1) are archived on the model native horizontal grids (either spherical harmonics or reduced Gaussian) at four analysis times per day (00, 06, 12, 18 UTC). Some fields that complement the analyses have been extracted from the 6-hourly forecast data and are included in these products. This promotes easier access to groups of related fields, and includes all of the chemical transport, net tendency, radiative tendency fields, some of the surface and single level fields, e.g. convective precipitation, surface sensible heat flux, and others. Analysis and first guess error fields will also be available from the MSS.

Product	No. of Levels	No. of Vars.	Vol. (GB)	Dataset ID	Access Options[1]
Surface & Single Level	1	108	509	ds117.0	OSM
Upper Air Pressure Level	23	9	796	ds117.1	OSM
Upper Air Model Level	60	11	1930	ds117.2	PSM
Isentropic	15	7	417	ds117.3	PSM
PV±2	1	7	31	ds117.4	PM
Chemical Transport	60	6	3400	ds117.5	M
Net Tendency	60	4	2280	ds117.6	M
Radiative Tendency	60	4	2280	ds117.7	M
Wave	1	23	375	ds123.0	PM

Table 1 ERA-40 Model Resolution Analysis Data Products

3.2 Model Resolution Forecast Fields

These data products include all the forecast fields archived on the native horizontal model grid. Collectively, these products are the largest entity in ERA-40 with roughly 20 TB. The forecasts extend to 36 hours beginning from the 00 and 12 UTC analyses, and out to six hours beginning at 06 and 18 UTC. Forecasts are available at 3 hour intervals out to 24 hours and 6 hour intervals thereafter.

Product	No. of Levels	No. of Vars.	Vol. (GB)	Dataset ID	Access Options[1]
Surface & Single Level	1	75	1971	ds121.0	M
Upper Air Pressure Level	23	9	3950	ds121.1	M
Upper Air Model Level	60	11	12700	ds121.2	M
Wave	1	23	349	ds123.1	M

Table 2 ERA-40 Model Resolution Forecast Data Products

3.3 Analysis Products on 2.5° Latitude-Longitude Grids

ECMWF has interpolated selected pressure level and surface analysis products onto equally spaced 2.5° latitude-longitude grids.

Product	No. of Levels	No. of Vars.	Vol. (GB)	Dataset ID	Access Options[1]
Surface & Single Level	1	55	73	ds118.0	OSM
Upper Air Pressure Level	23	11	353	ds118.1	OSM

Table 3 ERA-40 Analyses 2.5° Latitude/Longitude Data Products

3.4 Monthly Means, Variances, and Covariances

Monthly means are available for selected model resolution analysis fields, all model resolution forecast products, and the 2.5° data products. There are two monthly mean products, one computed for each synoptic hour (00, 06, 12 and 18 UTC) and another as a diurnal mean of the synoptic hours. Variances and covariances of these means are also available for some model level data as well as the surface pressure. Monthly means are also available for the 3 and 6 hour forecasts for the surface fields, and the 6 hour forecast for the pressure and model level products.

Product	No. of Levels	No. of Vars.	Vol. (GB)	Dataset ID	Access Options[1]
Model Resolution Analysis Monthly Means					
Surface & Single Level	1	108	21	ds119.0	OM
Upper Air Pressure Level	23	9	30	ds119.1	OM
Upper Air Model Level	60	11	81	ds119.2	OM
Isentropic	15	7	16	ds119.3	OM
PV±2	1	7	1	ds119.4	OM

[1] Access Option Key, O online for web and FTP download, S user subsetting available, P planned to be online soon, M available on the NCAR MSS.

Model Resolution Forecast Monthly Means					
Surface & Single Level	1	75	23	ds122.0	OM
Upper Air Pressure Level	23	9	30	ds122.1	OM
Upper Air Model Level	60	11	81	ds122.2	OM
2.5° Latitude-Longitude Grid Monthly Means					
Surface & Single Level	1	51	2	ds120.0	OM
Upper Air Pressure Level	23	11	11	ds120.1	OM
Model Resolution Analysis Variances, Covariances of Monthly Means					
Surface & Single Level	1	1	<1	ds119.5	OM
Upper Air Model Level	60	19	207	ds119.5	OM

Table 4 ERA-40 Monthly Means, Variance, and Covariance Data Products

3.5 ERA-40 Products Created at NCAR

The ERA-40 products furnished by ECMWF have been augmented with products created at NCAR. Surface pressure was computed as a supplement for the 2.5° resolution data products (as well as on the reduced Gaussian grid). The U and V wind components on the 2.5° latitude-longitude pressure level grids have been recomputed, because diagnostics indicated inconsistencies on the order of 0.5 m/s in the original data. Please see http://dss.ucar.edu/datasets/common/ecmwf/ERA40/docs/emos-diagnostics for more details. Monthly mean 2.5° U and V were also recomputed. In addition, fields are being synthesized on a 256x128 regular Gaussian grid at T85 spectral truncation for both model and pressure level data. This was designed and initiated to aid users of the NCAR Community Climate System Model directly, and others who desire ready access to fully synthesized fields that circumvents the need to deal with the spectral (spherical harmonic) form of these data products.

Product	No. of Levels	No. of Vars.	Vol. (GB)	Dataset ID	Access Options[1]
2.5° Surface Pressure	1	1	1	ds118.0	OSM
2.5 ° U,V on Pressure Level	23	2	60	ds118.1	OSM
2.5° U, V Monthly Means	23	2	2	ds120.1	OM
T85 Pressure Level	23	11	1019	ds124.1	PSM
T85 Model Level	60	11	2657	ds124.2	PSM

Table 5 ERA-40 Data Products Created at NCAR

3.6 ERA-40 Observational Data

During the ERA-40 analysis, input data records were documented according to how they were processed by the data assimilation system. At each stage of the processing, pre-assimilation through final analysis, metadata fields were generated and appended to the corresponding original observational record. The appended metadata includes quality control information, error statistics, and departures from model values. This collection of data records, called the feedback data product, is archived in Binary Universal Form for the Representation of Meteorological Data (BUFR) format. These data are available on the NCAR MSS.

4. DATA FORMAT, SOFTWARE AND UTILITIES

All ERA-40 data have been archived in the GRIdded Binary (GRIB) format, Edition 1. ECMWF used its own local extensions for the parameter code tables. These include Table 2, Version 128 for most of the standard variables; Version 162 for the vertical integrals, net and radiative tendencies, chemical model transport, variances and covariances; and Version 140 for the wave model output. The code tables can be downloaded from http://www.ecmwf.int/publications/manuals/libraries/tables/tables_index.html. GRIB is a widely used format for gridded meteorological data and there are several utilities and software packages for decoding, regridding, and reformatting the data. Some of the more useful utilities and their web links are:

- WGRIB – created by Wesley Ebisuzaki from the National Centers for Environmental Prediction (NCEP), enables users to decode and reformat GRIB data (http://wesley.ncep.noaa.gov/wgrib.html).
- NCL - NCAR Command Language is designed specifically for data access, analysis, and visualization (http://ngwww.ucar.edu/ncl/). NCL can also regrid, and reformat the data into the network Common Data Format (NetCDF).
- GrADS – The Grid Analysis and Display System, created by Brian Doty from the Center for Ocean-Land-Atmosphere Studies (COLA), Institute of Global Environment and Society at the University of Maryland, is an interactive tool that is used for data access, analysis, and visualization (http://grads.iges.org/grads/).

ECMWF has an extensive FORTRAN software package for decoding and regridding GRIB data (http://www.ecmwf.int/products/data/software/interpolation.html). The DSS has produced a suite of programs, based on the EMOSLIB software library, for performing various regridding functions, including synthesizing the spherical harmonics data to Gaussian or regular latitude-longitude grids. These programs are available from the ERA-40 software web pages (http://dss.ucar.edu/datasets/ecmwf/ERA40/software). In addition, these pages have a detailed description of the GRIB format, examples for each of the utilities and software packages, documentation about the ERA-40 horizontal and vertical grids, and diagnostic evaluations performed at NCAR.

5. SUMMARY

The ERA-40 project has produced a valuable archive of analysis and forecast products covering a 45-year period, September 1957-August 2002. The archive includes model resolution analysis and forecast fields, an extensive collection of monthly means, and a 2.5º uniformly gridded low resolution product. The DSS of SCD at NCAR has also created several value-added products to assist researchers and will produce others in the future.

Access to the ERA-40 data products is offered in several ways:

- Many analyses are online and available for direct download of full product monthly files by web and FTP transfer. Subset extraction by variable and level is also possible by delayed mode processing for most online data.
- Products not available online may be staged to the server upon request.
- Data may be provided on writeable media, upon request.
- All ERA-40 data is available to NCAR users from the MSS.

The primary focus is to provide a high quality archive with an emphasis on data delivery. The NCAR ERA-40 archive has been designed to promote efficient access and data mining for a large variety of users and requests. It is actively supported by data specialists who provide consultation on the archive products, locally developed software, and individual requests that require non-routine data processing and possible data delivery on media. Many data services and access methods for the archive are now available while still more will be developed in the future. ERA-40 is undoubtedly a key reference research dataset and will be a valuable data resource for many years.

References

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247-268.