

The preparation of global observations for reanalysis

Roy Jenne

NCAR

1. Introduction

Work at NCAR has resulted in the gathering and preparation of seven main types of data for reanalysis for 1948-2001. These data types are rawinsonde observations, winds only (pibals), aircraft data, satellite cloud winds, satellite soundings, surface 3-hour observations, ocean surface observations (COADS). The work to gather these started about 1967. NCEP and NCAR agreed in 1990 to carry out a joint reanalysis project. The interview period of work on the observations at NCAR was during 1991-2000.

2. Status of sending more observations from NCAR

Most of the observations that NCAR has been working on has been delivered to NCEP and ECMWF. About 3 sets will be completed during October-November 2001 (N. Pole raobs, Pacific profilers, Australia surface). ECMWF still has version 1 of UA station library locations for the data years 1979-2000. Version 3 was completed and inserted in data by 06/2000. Version 3.1 is next for data for 1994-2000 (completed by 07/2001). During Mar-Nov 2001 NCAR has prepared corrections for the worst errors in the locations of surface stations for the years 1975-2000.

3. The coverage and sources of data for ERA-40

The FGGE data set was available for most of 1979. ECMWF has its own archive of GTS data for 1980-2000, but these needed to be supplemented with other archives of satellite data, etc. Also, the NCEP and ECMWF GTS archives mostly duplicate each other, but there is probably a 5% gain in each by merging the two, which will be done. The observations for 1957-1978 for ERA-40 are nearly all from NCAR archives, which were built up over many years. They include data from former large US projects to gather data (especially by USAF), GTS from NCEP, USAF and Navy; NCDC data, and data from selected countries. We should all give thanks to the world's observers and to governments for these observations.

4. Adding more datasets of observations to reanalysis

In the future it could be useful to add more data to the present set of reanalysis observations. If this is done, checks should be made so that the new data are not adding in systematic errors. NCAR used a number of methods to detect systematic errors, including the massive comparison of reported data points between large component data sets that had some overlap. There are versions of some old datasets that do have systematic errors. If people use one of these from elsewhere and merge them into reanalyses, they will then introduce the errors again. Avoid this.

5. Documentation for the data archives

Many documents have been written at NCAR to describe the possible data sources for reanalyses, the processing of the data, data coverage, and the flow of data to NCEP and ECMWF. Many of these documents

are being gathered together to help describe the reanalysis observations. There will also be some summary documents to introduce this broad field without all of the complex details. There will be some information about the development and execution of reanalysis projects. The information will be on-line at NCAR and on a CD-Rom. About 9000 pages were scanned by Oct. 2001. These include some other topics such as ice ages, information about energy supplies, guide to sources of economic data etc. To find this information go to the NCAR Data Support home page and look for "Reanalysis: Guide to global observations and output" (<http://dss.ucar.edu>).

6. The future archives of reanalysis observations

NCAR is counting on a few centers to have copies of the datasets so that there are distributed data backups to prevent data loss caused by mechanical or human reasons. The component datasets must also be saved before all of the additional QC is done. In most cases additional data should be saved in both original form and processed form, especially if there was complexity in the transformations. Also, remember that some of the original datasets had information that was not carried along into the reanalysis. That is another reason for an archive of the original data. If the basic suggestions are carried out, it is much less likely that problems will be introduced in the future, and if some problems are found, it will probably be possible to fix them. Another reason to have several copies of these data in the world is that it will help to prevent the creation of data monopolies that are harmful to data security, data access, and data cost to users.

7. The additional value of observations for reanalysis

These observations can now define about 54 years of the world's weather history. They make it possible to do reanalysis projects. They are key components of data needed for programs like GCOS and CLIVAR. The observations, and the reanalyses output can also be used for much other research in meteorology, oceanography, and climate.