

THE ECMWF RE-ANALYSIS (ERA) PROJECT - PLANS AND CURRENT STATUS

J. K. Gibson, P. Källberg, A. Nomura and S. Uppala
European Centre for Medium-Range Weather Forecasts (ECMWF)
Reading, U.K.

1. INTRODUCTION

The European Centre for Medium-range Weather Forecasts (ECMWF) has provided data from its archive for Meteorological Research for many years. The ECMWF Re-Analysis (ERA) project was devised in response to the wishes expressed by many users of these data. It is funded by the ECMWF Council, the European Community, the Universities of California and Maryland, the Japanese Meteorological Agency (JMA), and Cray Research Incorporated. In addition contributions in terms of data and assistance have been forthcoming from many quarters, including ECMWF Member States, NMC Washington, NCAR, and the Australian Bureau of Meteorology.

1.1 Objectives

The goal is to develop a global atmospheric data assimilation system to analyze in an internally consistent way the record of atmospheric data (from the earth's surface to 10 hPa) for the period 1979-1993, and then to produce a validated and documented 15-year data-set of assimilated data at high horizontal and vertical resolution. This data-set will provide global three dimensional descriptions of the velocity, temperature, geopotential and humidity fields for the atmosphere. In addition, analyses of surface parameters (temperature, pressure, soil moisture, etc.), surface fluxes of heat, moisture, radiation and other diagnosed quantities will be included.

2. ORGANISATION

2.1 The Data Assimilation System

The ERA data-sets will be derived from a special version of the ECMWF operational data assimilation system. In this system analyses of atmospheric variables are produced by optimal interpolation of irregularly-spaced observational data to produce increments to be added to global short-range forecasts. Since this system is shortly to be replaced by a three dimensional variational system, a strategy is being followed such that the new system can be used is sufficiently proven. The short-range forecasts will be produced with a global spectral model with a triangular truncation at wave number 106. The ERA production system, once established, will remain consistent, resulting in data unaffected by the periodic changes introduced over the years within a daily operational system.

The real-time ECMWF data collection (which has a 3-day cut-off) will be the basis for the observational data-set. This will be augmented by the following additional data:

- First Global Geophysical Experiment (FGGE) level II-B data
- Alpine Experiment (ALPEX) level II-B data
- NOAA Climate Analysis Centre and UK Meteorological Office delayed mode sea surface temperature (SST) analyses, which will serve as the oceanic lower boundary condition for the atmospheric data assimilation
- The COADS ship data-set
- additional radio-sonde data-sets available from National Meteorological Centres
- NOAA satellite radiance data for the 15-year period.

The development of a reliable and robust assimilation system suitable for the fast rate required to generate the assimilated data-sets within the project's time-scale is currently under way. This involves major software and data handling development, coupled with experimentation to select the appropriate version of the forecasting model and the spatial analysis procedure employed in the assimilation system.

2.2 Data-set production

Data-sets will be produced in three main phases:

Phase A1 - Pilot analyses for June 1985

Phase A2 - Analyses for 1979 - the FGGE year for which there exists a number of comparable data-sets

Phase A3 - Analyses for 1980-1993.

The pilot analyses and the FGGE year analyses will be used by the production team at ECMWF, and validation teams elsewhere to assess critically the quality of the analyses before embarking on the remaining period.

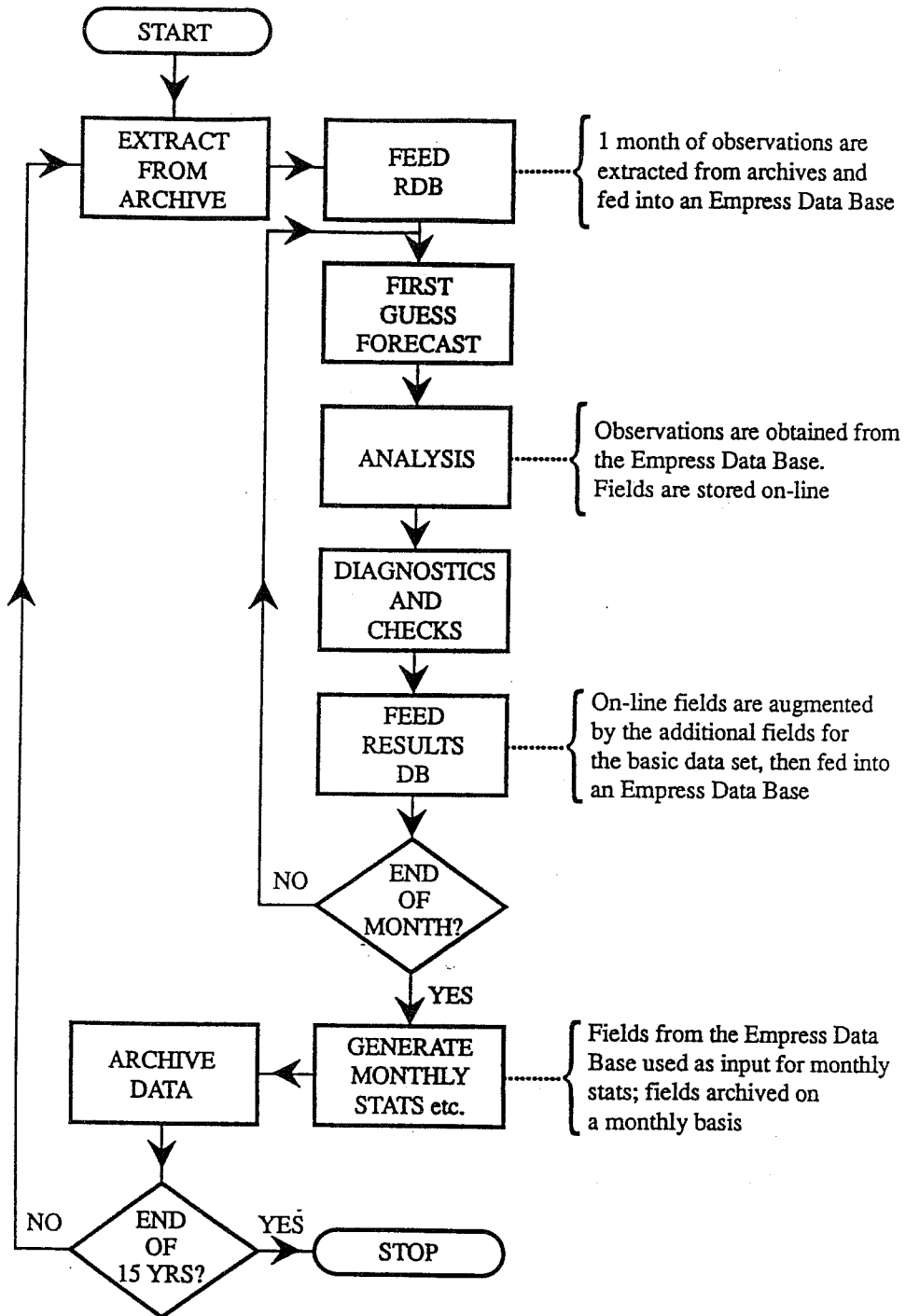


Figure 1. Production Methodology

Figure 1 illustrates the production methodology. Observational data will be pre-fetched from archives, one month at a time, organised in data base form, and checked before use. Immediate results from the 4 times daily first guess forecasts and analyses will be retained on-line to facilitate initial diagnosis and checks. Suitable results from each forecast/analysis cycle will be entered into a data base prior

to the generation of monthly statistics and addition to the archive. The analysis phase will require the ECMWF 16-processor Cray C90. Silicon Graphics file servers will be used for the on-line data bases and pre- and post-processing.

2.3 Validation

The validation exercise is designed to provide rapid, critical and comprehensive evaluation of the data-sets as they are produced, to detect defects as soon as they arise. This will be done through diagnostic studies of the analyses as they become available. Validation effort will be shared with scientists from the funding agencies, and with ECMWF's partners within the terms of the EC funded project.

The specific diagnostic tasks are:

- V1 Validation against global observations used and not used in the analyses; validation through regular forecast experiments. (ECMWF and Spain)
- V2 Validation by comparison with earlier analyses from ECMWF and re-analyses from NMC (Washington) - this will be particularly important at the very start of the project; validation of global energetics and the hydrological cycle; validation using ocean models by using the diagnosed surface fluxes to drive an ocean model, the resulting evolution of the SST will be compared with XBT data. (Max Plank Institute, Germany)
- V3 Comparison of clear sky outgoing radiances diagnosed from the analyses with measurements from the Earth Radiation Budget Experiment (ERBE) using simulations with the SAMSON system - this will help validate in particular the analyzed humidity field. (U.K.)
- V4 Validation by using surface fluxes to drive an Alpine snow-mantle model, the resulting evolution of the mantle will be compared with observations. (France)
- V5 Verification of the analyses over Greenland against observations. (Denmark)
- V6 Validation of the analyzed Planetary Boundary Layer structure using data at the Cabauw Site in The Netherlands. (Netherlands)
- V7 Diagnostics of atmospheric low-frequency variability deduced from the analyses will be compared with similar diagnostics from previous analyses produced by ECMWF and re-analyses produced by NMC Washington. (University of Bologna, Italy)
- V8 Validation of the analyses over the Mediterranean region. (Istituto per lo Studio delle Metodologie Geofisiche Ambientali, Modena, Italy)

Results from these studies will provide rapid assessment of the analyses produced during phases A1

and A2; the results of which may redefine details of the analysis system - and will also provide an essential component of the scientific documentation.

2.4 Archives and Data Services

The additional observational data, and a comprehensive set of re-analysis results will be added to the ECMWF archive. Appropriate data will be made available through ECMWF's Data Services.

Four main validated and documented data sets will be produced, forming the basis for data services:

ERA Basic Analysis Data Set

ERA Advanced Analysis Data Set

ERA Precipitation, Moisture, and Energy Flux Data Set

ERA Climate and Statistics Data Set

The Basic Data Set will provide selected variables in compact form at low resolution and will be particularly useful for users with limited data processing resources. It will contain geopotential, temperature, vertical velocity, two components of wind, and relative humidity on a 2.5 degree grid at 15 standard pressure levels, together with a number of surface and diagnostic parameters. One year of such data would require 10 standard 6250 bpi magnetic tapes.

The Advanced Data Set will provide data at higher resolution in time and space, giving access to the results of the re-analysis at the horizontal and temporal resolutions of the generating system. There will be options for users to obtain these data either on the grids used by the generating system (T106 spherical harmonics for upper air data, N80 Gaussian grid for surface and diagnostic parameters), or interpolated to a regular latitude/longitude grid of the user's choice. The parameters supported will be similar to those described for the basic data set above. In addition to the 15 standard pressure levels, data will be available on the 31 level hybrid vertical co-ordinate system used by the generating models. One year of 15 pressure level data would require 52 standard 6250 bpi magnetic tapes; approximately 88 tapes if the 31 model levels were specified.

A number of parameters of interest to the international scientific community are not computed within the analysis, but are potentially available from short range forecasts which provide first guess information for the analyses, provided such forecasts are extended in their range sufficient to meet such requirements. The Precipitation, Moisture and Energy Flux Data Set will be generated from such forecasts, and will provide parameters such as surface flux of sensible and latent heat, surface and top

of the atmosphere radiation fluxes, and rainfall. These will be retained at model resolution; interpolation to other resolutions will be supported. One year of such data at full resolution would require 10 standard 6250 bpi magnetic tapes.

The Climate and Statistics Data Set will contain monthly and seasonal means data together with a number of statistics which will be generated at the end of each month of analyses during production. These data will enable access to many of the climatological aspects of the results without recourse to the full daily data. Considerable discussion has taken place with eminent atmospheric scientists and potential users to determine the contents of this data set. One year of such data would require approximately 5 standard 6250 bpi magnetic tapes.

3. STATUS

3.1 Project Staff and Management

A team of 5 people, including an EC fellowship student, are currently occupied full time within the project. A further EC fellowship student will begin work on 10 January 1994.

An internal steering group together with the ECMWF Director reviews the project management. An external advisory group provides advice and facilitates liaison with the external scientific community and the potential users of the results.

3.2 Establishment of the Data Assimilation System

A programme of experimentation has been undertaken to determine to what extent the data assimilation system to be used should differ from that of the Centre's daily operations. This was necessary because the system, once established, should remain fixed for the full re-analysis production, and because the basic horizontal resolution to be used will be T106, not the operational resolution of T213. Results from these experiments to date indicate:

- a) the use of 31 vertical levels compared to 19 is beneficial in some aspects;
- b) the use of the "first guess at appropriate time" scheme provides insufficient benefit to the re-analysis to justify its additional cost in terms of computational resources;
- c) mean orography (as opposed to envelope orography) will be used;

Experimentation carried out within the ECMWF's Satellite Group was followed closely to establish how best to use the cloud cleared radiance data. Also considered were the optimal treatment of sea surface temperatures and sea ice coverage data, the inclusion of the Integrated Forecast System

and the Three Dimensional Variational Analysis schemes being developed at ECMWF.

3.3 Data acquisition

Work is at an advanced stage to convert FGGE and ALPEX data to the representation forms used within the ECMWF data assimilation system.

The sea surface temperature data produced by NMC Washington, the AMIP sea surface temperature data, and the sea surface temperatures produced by the UK Meteorological Office have been acquired. These have been subjected to comparisons, and examined in detail for various events including El Niño periods. Attention has been given to the representation of sea ice within the SST data, to enable decisions to be made as to whether additional sea ice information should be used.

NMC Washington have agreed to make available the COADS ship data-set as part of the co-operation established between the ECMWF and NMC re-analysis groups.

Various National Meteorological Centres, including a number of ECMWF Member States, have offered access to additional radio sonde and AIREP data. These offers will be taken up as manpower resources permit, beginning with those considered likely to have the most impact for the least effort.

NOAA satellite cloud cleared radiance data have been purchased jointly with NCAR. Delivery of these data have begun, starting with the 1979 data together with sufficient 1985 data to cover the experimental period, and will continue during the early part of the production phase. There are, unfortunately, gaps in these data (about 8 months are missing in total). To address the problem of these gaps, level I-B data is being obtained, and will be re-processed to obtain the II-B cloud cleared radiances.

3.4 Setting Up the Production System

The various components of the production system are being assembled and checked. To date the EMPRESS data base system to be used for handling observations has been established. The Supervisor-Monitor-Scheduler (SMS) is running on the ERA server; the environment for its use has been established and checked on both the Cray and SGI systems. Integration work is now in progress to generate more comprehensive suites necessary for full production.

4. THE SCHEDULE

The primary objective is to complete the re-analysis by the end of 1995. The anticipated schedule is set out in Figure 2.

1Q 1994	Phase A1 and A2, and associated validation complete
2/3Q 1994	Phase A3 complete to December 1982
4Q 1994	Phase A3 complete to December 1985
2Q 1995	Phase A3 complete to December 1989
4Q 1995	Phase A3 complete to December 1993

Figure 2. Schedule for the main tasks