Introduction

Reanalysis products from NCEP, NASA and ECMWF have been widely used and continue to be used in a large number of scientific applications. The current ECMWF reanalysis ERA-40 has been designed taking into account the main areas of criticism concerning the ERA-15 analyses and also, before production, the expert opinions of an External Scientific Advisory Group. Extensive tests were carried out with the proposed system over limited periods before it was accepted for production.

This mid-term workshop, as anticipated in the project plan, was targeted towards reviewing different aspects of quality of early production analyses and seeking advice on solving possible problems. The areas covered by the presentations included the following:

- TOA and surface fluxes
- Evaluation of clouds
- Hydrological cycle and general circulation
- Use of conventional and satellite data
- Analysis quality over polar areas
- Stratospheric analysis
- Ocean wave analysis
- Analysis of ozone
- Detection of synoptic features

In addition to the Project Partners, as many as possible of the active ERA-40 users and experts on reanalysis were invited to describe their results and to contribute to discussions and recommendations in working groups concerning:

1. Observations for reanalysis
2. Assessment of the quality of the ERA-40 analyses
3. Trends in observations and re-analyses
4. Observing System Experiments for ERA-40
5. Distribution of ERA-40 data

The discussions and recommendations are summarized in the following five reports.

The workshop was both active and innovative, resulting directly in changes to the use of radiances in ERA-40. The effects of the changes benefit in particular the quality of analyses of the global hydrological cycle and of temperature over polar regions for the rest of the production and in a possible rerun. Among the discussions a strong view was expressed for a wider and easier access to the ERA-40 data for the worldwide research community, which in itself is a positive quality statement by the users.

These proceedings, a mixture of presentations and papers, provide ample information on ERA-40 in general and give a flavour of progress and concerns at the time of the workshop.

ECMWF wishes to thank all participants for contributing to a successful workshop, and the European Community (under contract EVK2-CT-1999-00027) and the World Climate Research Programme for financial support.

1. Working Group 1: Observations for Reanalysis

M. Fiorino, discussion leader/reporter

R. Jenne, J. Woollen, S. Saarinen, C. Genthon, N. Sokka, K. Onogi, L. Van de Berg, P. Viterbo, M. Dragosavac

1.1. Introduction

From the perspective of observations, reanalysis can be characterised as a data "quality control" process whereby a comprehensive and complete (at a given moment in time) set of observations are quality controlled (QC) against a 4-D, physically consistent representation of the atmosphere. The analysis fields and model diagnostics are the reanalysis. While these "gridded data byproducts" have received the greatest attention in the reanalysis user community, it is the data assimilation feedback (i.e., observation departures and QC flags) that gives special knowledge regarding observation quality.

The observations working group discussed the current state of the input observations to ERA-40 and emerging reprocessed and new sources, and then made recommendations regarding observation products for both scientific users of reanalysis grids as well as future reanalysis projects.

1.1.1. State of Conventional Observations for Reanalysis as of 200111

We reviewed input data sources and status, relative to the NCEP/NCAR and ECMWF archives, for two periods:

P1: 1957-1978 (pre FGGE)

P2: 1979-2000 (post FGGE)

P2 = NCEP R2 (NCAR V1 +) merged BUFR from
ECMWF MARS Archive
ECMWF component data sets (Antarctic, Canadian snow, JMA, ...)

P1 = NCEP BUFRised NCAR V3
ECMWF component data sets

There are two fundamental "phases" of observations:

1) original (archivist) in many forms;

2) adapted (filtered) for NWP into a single form (typically BUFR).

While both phases are necessary, the NWP centers may not be the most appropriate place for general maintenance of original data as they tend to adapt their observations to specific applications.

It was noted that the NCEP R2 archive is not complete in that some non-assimilation variables were not encoded. This is the "templating" issue imposed by the limited but common use of the possibilities offered by BUFR format, i.e., only fields in the template are encoded. However, the P1 translation of NCAR V3 to

NCEP BUFR followed a more general template to include quantities such as precipitation and significant weather.

1.1.2. Status of NCAR Archives

The NCEP archive and ERA-40 are missing some new data components (V3) received by NCAR as of 200002 for the period 1979-1994 (ERA-15).

1.1.3. ERA-40 Data Flow

Most of the working group discussion concerned the form and flow of observations in the ERA-40 system and the new station height checking/flagging scheme. The basic classes of input files are:

- Conventional observations including satellite Cloud Drift winds
- SSM-I/ VTPR/ TOVS/ ATOVS 1c brightness temperatures
- ERS Scatterometer& Altimeter data
- SBUV/ TOMS/ GOME
- Synthetic Surface Pressure observations (PAOB)

All these data are presented to the analysis in the universal BUFR-format. Feedback versions of these input files are produced during the data assimilation, but are nearly 10X as large because of departure and flag information. The format of feedback information is also stored in a rather complex BUFR format.

1.1.4. Satellite Radiances

Satellite radiances go through a pre-analysis thinning process that cuts the volume by two orders of magnitude (~1/80 ratio). Even with thinning the data volumes are enormous because of feedback blowup factor.

In each 6 hour period the satellite data flow is:

satellite X ->
  prelcrad (thinning 1/81 ~ 10 Mb) ->
    preproc (~100 Mb) ->
      3dvar ->
        odb feedback (~300 Mb)

The increase in channels from about 50 in the TOVS satellites to 8000 in AIRS, etc. will make the data handling problem daunting for future reanalyses.

1.1.5. Conventional observations

Observations from the different conventional (including cloud drift winds) data sources have been inserted into the PREODB database. From these about 30 component data sets a merge is done for each 6 hour period to form the "consolidated" or "merged PREODB" data set presented to the ERA-40 3DVAR. The merged PREODB is considerably smaller than the radiance datasets and is of the order of 5 Mb and with feedback about 40 Mb.

1.2. Potential and reprocessed observations

EUMETSAT is reprocessing Meteosat winds for the period 1983-1988. extending the period if possible.

Similarly the US has full archives of GOES data from 1979-on, but there is a need to recalculated for 1979-88 due to the development of better wind algorithms

JMA is reprocessing GMS SATOB for 1987-1991, the original GMS for 1979 (FGGE) is held at U. of Wisconsin and JMA should be encouraged to do the reprocessing.

1.3. User communities and requirements

We expect that observations and feedback in BUFR and ODB forms will be of specialised application, especially for future reanalyses, but that monthly feedback statistics in both gridded and time series forms will likely satisfy general user requirements for metadata.

1.4. Future Reanalysis Programs

The calibration starting point for future renalysis should be the merged ERA-40 input observations since the ERA-40 merging represents a "best" filtering for 6-h 3DVAR-type data assimilations. Two upcoming reanalyses in the US are:

NASA DAO ReSTS - Reanalysis for Stratospheric Transport Studies

NCEP RR - NCEP Regional Reanalysis

These projects would have immediate benefit from merged PREODB.

1.5. Recommendations

We first encourage ECMWF to make observations, and observational metadata products that come out of ERA-40, available to future reanalysis projects and data assimilation research whenever possible and to allow for redistribution so as to limit impacts on data services.

1.5.1. Data Products

For data maintenance and archivist centers (e.g., NCEP, NCAR and ECMWF) the ECMWF MARS and other ECMWF components should be exchanged. It is also recommended that derived/ processed observations be made available to reanalysis programs; specifically the merged PREODB and the BUFR feedback.

Coordination between NCEP and ECMWF of the "templates" through which the source observations pass during the encoding process is needed. These templates often filter out fields because they are not of immediate application (e.g., temperature when the analysis only used heights).

1.5.2. Metadata Products

The most important observation product for general users is metadata. Basic documentation for ERA-40 is quite comprehensive, but updates as knowledge is gained, particularly as a function of the component data, is

encouraged. Further, a description of the merging process, from which a single stream of non-duplicate and "best" observations is presented to the data assimilation, is needed.

Other important metadata products include: 1) blacklists; 2) the TEMP sonde database and bias correction; and 3) feedback statistics in gridded form. The blacklists define the observations that the data assimilation cannot extract information from either because the observations have excessive errors or because the observation operators are not available or error knowledge is insufficient. The blacklist will certainly need to be reexamined for future reanalyses.

The correction of TEMP temperature biases, due to heating of the sensor by sunlight, is critical to the analysis in the upper troposphere and stratosphere. The databases of sonde types and the actual bias corrections should be made available. Monthly mean statistics by grid (e.g., the actual bias correction applied) should be sufficient.

The POSTODB feedback statistics process can generate monthly mean statistics on grids (e.g., 5x5 deg). Such gridded forms of the statistics would be extremely helpful to the general user. We specifically recommend that monthly means of the following statistics be generated:

Counts, rejects, flags, innovation, residual, increments

as a function of observation platform, e.g., TEMP, AMSU-B and SYNOP. A number of forms for such gridded data are available including XML.

We also recommend that monthly mean statistics (as above) be produced for TEMP and a "fixed" (TBD) number of SYNOP stations. NCEP has produced such a data set (in BUFR) for TEMP for the period 1948-present which has proven valuable for model verification.

1.5.3. Data Recovery and Improvement

Coverage and feedback statistics can be used to "target" data recovery in areas of poor coverage and or poor assimilation can be more clearly identified. Some known targets include:

- pre-1956 Canadian TEMP (needs digitisation)
- 1957- TEMP in a South African book (need digitisation)
- early 1900's remote surface observations in Norway and Sweden
- pre VTPR (1969-1972) SIRS-A/B satellite soundings

It is also suggested that modification of the the WMO station library might consider input from reanalysis (e.g., the station height correction scheme).

1.5.4. Data Archives

NCEP and ECMWF should maintain archivist (source) components as the merging process for reanalysis is specific to the ingest requirements of data assimilation.

2. Working Group 2: Assessment of the quality of the ERA-40 analyses.

Allan Betts, discussion leader/reporter

R. Allan, M. Serreze, D. Bromwich, G. Kelly, S Hagemann, T. Fuchs, P. Simon, P. Van Velthoven, F. Chevallier, Klaus Arpe, A. Beljaars, P. Kållberg

2.1. General

The assimilation scheme, which heavily constrains the re-analysis towards a realistic behaviour, is the major strength of ERA-40. This shows up in the acknowledged quality of the analysed fields (pressure, temperature, wind, etc...), despite some remaining problems in the handling of observations. In the areas where, and in the periods when, the observations are sparse or not available, ERA-40 relies on the physics of the forecast model. Similarly, those derived atmospheric quantities, calculated from the assimilated variables, like precipitation rates, are in part determined by the characteristics of the model physics. As a consequence, the weaknesses of the parameterisations, as their strengths, have a major influence on those ERA-40 data that are far from the assimilation scheme. Considering the pre-satellite period, it should be noted as well that the ECMWF physics was developed primarily for forecasting purposes rather than for long climate runs. During the pre-satellite years the model physics plays a more significant role and this should be realized.

2.2. Evaluation of clouds

The model clouds are overall well distributed by the dynamics, with realistic seasonal cycles. Individual tropical and extra-tropical cyclones are accurately detected. However, deficiencies have been identified in position and tropical cyclone forecast accuracy.

- The vertical distribution of the clouds is not correct. ERA-40 seems to have too much liquid water and too little ice. This is acknowledged as a long-standing problem in the ECMWF cloud parameterisation, which negatively impacts the top-of-the-atmosphere energy balance.
- On time-scales smaller than a month, the variations of the inter-tropical convergence zone are not well captured, in particular over land and during boreal winter in the Atlantic and the Eastern Pacific.
- The stratocumulus clouds off the west coast of the continents are optically much too thin.
- ERA-40 hardly captures those cyclonic structures where scales are not significantly larger than the ERA-40 numerical resolution.
- The tropical cyclones have large location error and are too spread.

2.3. Evaluation of the hydrological cycle

The ERA-40 hydrological cycle has several deficiencies, especially with regard to precipitation. The global water balance is not closed mainly due to the removal of net water out of the atmosphere (in the ITCZ) by the ERA-40 model, while the observations (probably HIRS) add water at each timestep. As a result P-E over the ocean is not negative. ERA-40 is generally wetter than ERA-15, and it definitely has too much precipitation over the tropical oceans.

The spin-up of the hydrological cycle within the first 24 hours of the forecast makes the global imbalance of the water budget and the large wet bias over the oceans even worse. The increase of precipitation in the extra-tropics is on the whole positive, since it reduces the winter dry bias of ERA-15 over Europe. There is also a large reduction of gridpoints with negative P-E values over land. Both ERA-15 and ERA-40 precipitation have a wrong diurnal cycle over land with maximum precipitation at noon.

A distinct erroneous feature is the major upwards shift in the global precipitation behaviour over the ocean in the middle of 1991 which may suggest that the Pinatubo eruption has influenced the satellite data assimilated by ERA-40 thereby amplifying an error which seems to be already in the ERA-40 system. It has to be investigated how the ERA-40 precipitation characteristics changed after the SSM/I bias correction error is eliminated from 1 January 1993 onwards, and after May 1993 when the influence of the erroneous data assimilation of ERS-1 ocean wave data has ended.

2.4. High Latitude Validation

In comparison with ERA-15, forecast means of P-E from ERA-40 are roughly in balance with those computed from the atmospheric moisture budget. Over land areas, about half of the variance in observed monthly precipitation is typically captured, although performance varies by season and region. ERA-40 high latitude precipitation forecasts over land areas are not significally improved over ERA-15. Precipitation along the north Atlantic storm track is higher than in ERA-15 and seems to be closer to observations. However, summer precipitation is noticeably over-estimated over the central Arctic Ocean. The performance of ERA-15 is better in this area. Various aspects of the components of the surface mass balance of Antarctica have improved in ERA-40. However the interior of the icesheet is still too dry and affected by strong spin-up problems.

The large winter cold bias over land seen in ERA-15 2-m temperature has been addressed through adjustment of the surface and boundary layer parametrizations. However Antarctic surface temperatures are now too warm. This is manifested as higher sensible and latent heat fluxes in spring. However, compared to observations, ERA-40 appears to have a small warm bias in 2-m temperatures. Runoff is (mostly largely) overestimated over the two major ice sheets.

Of particular concern is a cold bias in lower tropospheric temperatures over extreme high latitudes of both hemispheres. As a result, 500hPa heights exhibit negative mean seasonal biases with respect to ERA-15 of approximately 30 m locally. This is most likely related to a problem in TOVS-HIRS retrievals. (An OSE to address the problem is underway)

Topographic representations of Greenland and Antarctica are improved in ERA-40 and the erroneous elevation of Vostok station has been corrected. However, incorrect station elevations still appear to be a problem.

The representation of sea ice concentration and extent in ERA-40 has been greatly improved for the satellite era. Improved parametrizations have also led to more realistic temporal variability in 2-m temperatures over sea ice.

2.5. Radiation budget

Clear-sky radiation budget and water vapour distribution for low latitudes are reasonable. Further analysis of the interannual variability is required to assess the quality of these products and whether they will be of use in the evaluation of climate models.

The all-sky top of atmosphere radiation budget is poor, mainly due to the unrealistic radiative effects of cloud. The cloud greenhouse effect is underestimated in the tropics possibly due to cloud ice, while tropical convective clouds are too reflective. Also the stratocumulus cloud influence on the radiation budget is unrealistically small. However, seasonal variability of fluxes seem reasonable.

While the cloud radiative effect in ERA-40 is poor, the cloud distribution and the atmospheric circulation appear consistent with observed radiation and cloud fields.

Subsampling of observations using ERA-40 dynamical parameters may provide a valuable technique for verifying the climate model developed within the Met Office. This requires verification of the atmospheric general circulation.

2.6. 6 Synoptic features

The detection rate of Tropical Cyclones (TC) in ERA-40 is quite good, ~90%, and somewhat better than in ERA-15. There are however quite large position errors in the analyses, and the subsequent forecasts are not good. This is likely to be due to the comparatively low resolution of ERA-40 for TC diagnostics and forecasting.

Tracking of synoptic features, both at the surface and e.g. at the PV=2 level, show a much better inherent consistency in ERA-40 than in ERA-15. This is believed to be due to the variational analysis in ERA-40. The optimum interpolation technique in ERA-15 with 'analysis boxes' produced more 'spotty' analyses with less consistency from analysis to analysis.

Synoptic 'maps' from ERA-40 and ERA-15 are generally very similar, although not negligible differences can be found over the southern oceans

2.7. Summary of identified strengths

- Skill of 10-day forecasts from ERA-40 is excellent – compared to ERA-15 and the early 1980's operations
- Much improved stratospheric analyses, temperature ('sudden warming'), winds ('QBO'), humidity ('extreme dryness in lower stratosphere') and ozone (when TOMS/SBUV available)
- Improved mid-latitude storm track precipitation
- ITCZ over Africa improved
- Soil water nudging over land has decreased
- High latitude frozen physics and water budget has improved
- Severe cold bias in high-latitude near-surface temperature has been replaced by a small warm bias
- Improved representation of sea ice and associated 2-m temperatures
- Model clouds generally well distributed by the dynamics.
- Clear sky radiative budget is reasonable.
- Good detection of tropical cyclones

- Better coherence and location of mid-latitude cyclones
- Improved representation of ocean 10-metre winds.
- Monthly mean 10-metre winds and waves are OK.
- ERA-40 waves are better than ERA-15, local and global improvements.
- Improved distribution and variability of total water column.

2.8. Summary of identified weaknesses

- Too large oceanic precipitation in the tropics.
- Sudden shift (increase) in tropical precipitation, total water column and cloud albedo in mid-1991
- P-E over oceans is erroneously positive
- Spin-up of the hydrological cycle
- Convective precipitation over land has maximum too early in the day (around noon)
- Overestimation of OLR, possibly due to underestimated cloud ice
- Tropical convective clouds too reflective, liquid water content overestimated.
- Stratocumulus badly represented
- High ocean-wave heights underestimated, low heights overestimated.
- Erroneous ERS-1 wave data used November 1992 to May 1993
- Position errors in detected tropical cyclones.
- Cold bias in lower troposphere in the central Arctic and Antarctic
- Too much precipitation in central Arctic in summer
- Before 1979 ozone is free-running may and not be useful.
- Trends before the satellite era may not be reliable, except in areas rich in conventional data.
-

2.9. Recommendations for further validation

- Surface radiation budget and surface albedo.
- Arctic stratus in summer and seasonal cycle of cloud radiative forcing
- Clouds in general – hard to do due to the paucity of global validation data sets.
- Consider how to validate streams 2 and 3 given the lack of independent data.
- AMIP-run(s) with the ERA-40 system should be carried out.
- Dry biases in RS humidity data in the early years.
- A General Circulation Atlas should be prepared
- Validation should produce specific recommendations for future users. (e.g. best period for precipitation)
- Encourage further validation against field experiments from the 1990s.
- Permafrost – there are independent datasets (ERS-1 and N.American soil temp data.)
- Stratospheric validation, using rocket sonde data.

2.10. Recommendations for future users

- Production system documentation
- Documentation of major issues.
- Documentation of all observational changes –
  - new systems,
  - error assumptions,

  - blacklistings,
  - bias corrections,
  - and other problems
- Publish observational and other changes on the website a.s.a.p.

3. Working Group 3: Trends in observations and re-analyses

K. Trenberth, discussion leader/reporter

K. Arpe, D. Bromwich, L. Heimberger, E. Holm, B. Hoskins, P. Simon, A. Simmons, P. Van Velthoven

3.1. What are the issues?

The ERA-40 re-analysis system is designed to synthesize the observational data available at any particular time to produce the best possible analysis for that time. It uses a fixed model, but assimilates data from an observing system that has changed considerably over the period of the re-analysis. ERA-40 may exhibit trends (such as cooling of the lower stratosphere) that are present in the assimilated observations, and may through the data assimilation produce trends in related variables for which no observations are assimilated. Re-analysis also provides comprehensive long-term descriptions of the atmosphere that may help further the understanding of observed trends.

However, there are several sources of problems relating to trends and low-frequency variability in re-analyses:

- Model bias, which causes the analyses to drift towards the model climate in the absence of observations;
- Incomplete representation of trends in atmospheric composition and boundary conditions of the model;
- Changes in the observing system (see below);
- Perturbations such as the Pinatubo eruption.

Consequently, there is a considerable potential for any trends and low frequency variability to be entirely spurious as the size of the expected signal is often less than the noise due to changes in the observing system.

3.2. Model bias and specification

It has been proposed for ERA-40 that an AMIP-style integration of the assimilating model be carried out for the period of the re-analysis using the same SST and other specified boundary conditions and composition changes as used in the data assimilation. **It is recommended that:**

- The experiment be carried out as an important aid to understanding some of the differences in analysis characteristics that will occur among periods with different observational coverage. An ensemble of integrations would be preferred to sample model variability.

ERA-40 includes changes in radiatively active gases and SST, but does not include changes in solar input, aerosol content or land use over the re-analysis period. **It is recommended that:**

- Sensitivity experiments be carried out for limited periods to document sensitivity to neglected changes;
- A more complete representation of trends in atmospheric forcings, including changes in solar irradiance, atmospheric composition and surface boundary conditions, be included in future re-analyses if sufficiently well known.

3.3. Observing system changes

There have been many changes to the global observing system, and within individual components of the system, over the period of ERA-40. The period has seen the establishment and enhancement of a satellite observing systems and of reports from commercial aircraft and drifting ocean buoys. The radiosonde observing system has undergone considerable change.

Satellite data are first used in ERA-40 late in 1972. The VTPR temperature and humidity sounding data are replaced in 1979 by data from the TOVS (HIRS/MSU/SSU) system. Cloud-tracked winds from geostationary satellites covering all longitude zones were established for the global weather experiment in 1979. Ozone retrievals from TOMS and SBUV also become available around the same time. Marine surface winds and total column water vapour from SSM/I are used from 1987 onwards, and ERS altimeter and scatterometer data are available from 1992 onwards. The ATOVS (HIRS/AMSU) system begins to replace the TOVS system in 1998.

Quite apart from the introduction of new types of satellite observations, there are characteristics of specific observations that have to be taken into account in studies of trends and low frequency variability. The typical lifetime of individual satellites is a few years, and the number in orbit at any one time changes. There can be gradual changes in the orbital characteristics of satellites, drifts in measurements from particular instruments (such as due to gas-leakage from the SSU instrument), and changes in quality (and quality indicators) in derived quantities such as cloud-tracked winds and ozone retrievals. Measurements may also be perturbed by external events such as volcanic eruptions and cosmic storms, and by changes made by satellite operators, such as attitude changes which alter the solar heating of the platform.

There have been substantial changes in the general coverage of radiosonde observations. Changes have occurred in the location, time and frequency of launch, instrument accuracy and data processing for individual stations. The history of these changes is far from complete.

The planned observing system experiments in which assimilations are carried out both with and without new components of the observing system at the time of their introduction should help identify the reasons for some of the shifts expected to be seen in the time series of the completed ERA-40 analyses.

It is desirable to establish the consequences of more gradual observing-system changes through additional experiments. **It is recommended that:**

- a recent year be chosen and experiments be run using degraded sets of observations that simulate the data available from the observing systems of the mid 1980s, mid 1970s and late 1950s. Simulated weather-ship and VTPR observations could be constructed from the control ERA-40 (or operational 4D-Var) analyses. The experiments should cover at least a month from each of the four seasons.

3.4. Bias corrections

Bias corrections are typically based on observation-background differences, and carry a potential for corruption of re-analyses through biases in the background. Unintended perturbations can be perpetuated. This is a difficult area because of the absence of adequate "ground truth" or "baseline" measurements of uniform quality over the period of re-analysis.

Bias correction of satellite radiance measurements is essential for their use in re-analysis. The use of raw rather than pre-processed radiances in ERA-40 has enabled far fewer changes to the bias correction to be applied over the lifetime of an individual satellite. **It is recommended that:**

- ECMWF document the time series of satellite monitoring statistics and the bias corrections applied throughout the thirty years of ERA-40 during which radiance data are assimilated.

Although unsatisfactory, the closest to a baseline set of observations for detecting true trends in climate comes from the GUAN set of radiosonde stations, for which both the observational record and the metadata are nominally more complete. Tracking bias corrections for these subsets of stations may provide the only way to validate trends in tropospheric temperatures.

It is desirable to correct biases in the complete set of radiosonde measurements for re-analysis, although it is difficult to cope with undocumented changes of instrument type. Corrections should be applied with caution.

The feedback information for radiosonde data contains valuable information. Jumps provide a clear indication of instrument changes, or of changes in data pre-processing. Matching the biases and changes in ERA-40 feedback statistics with known characteristics and changes for the GUAN stations may identify unexpected changes in data for particular stations which would be important for GCOS. **It is recommended that:**

- the radiosonde feedback statistics for the whole of the ERA-40 period be extracted as a separate subset and made available to users in a simple, easy-to-use form;
- emphasis be placed on the GUAN network in the analysis of results.

3.5. Tracking performance of the data assimilation system

It is important that the performance of as many as possible aspects of the ERA-40 system be tracked over either the whole period or extended sub-periods for which suitable validation data are available. **It is recommended that this include:**

- Examination of time series of basic parameters such as the global mass of dry air and land temperature for selected regions;
- Examination of time series of the background and analysis fits to observations;
- Examination of time series of a variety of measures of forecast performance;
- Comparisons with time series of any direct observations that were not used by the data assimilation system, such as from Alpine summit stations, from surface-based Dobson spectrometers, from SAGE and from ocean-wave measurements;
- Comparisons with homogenized datasets such as the surface measurements from the European Climate Network and the U.S. Historical Climate Network;
- Comparisons with independently processed MSU-2 and MSU-4 data;
- Comparisons with documented trends from relevant GEWEX and SPARC datasets and reports;
- Comparisons with independent measurements such as river discharges.
-

It is furthermore important that the results of these studies, and the results arising from the studies of monitoring and feedback statistics be communicated effectively to the users of ERA-40. **It is recommended that this cover:**

- List of trends and jumps that are thought to be spurious, identifying all likely affected variables;
- List of other problems and unexplained features of the ERA-40 analyses noted during monitoring and validation, identifying the variables and time intervals concerned.

4. Working Group 4: Observing System Experiments for ERA-40

L. Bengtsson, discussion leader/reporter

R. Rood, S. Caires, M. Kanamitsu, K. Hodges, P. Janssen, L. Isaksen, E. Andersson, S. Uppala

4.1. The need for Observing System Experiments within ERA-40

There is a general need to understand how reliable the ERA-40 analysis quality is during different periods from 1957 to 2001. We know from the operational and ERA-15 experience that the quality of observing systems in general is improving. The early period contains conventional observations only. The satellite observations become dominant towards the end of the period. The purpose of the Observing System Experiments is to be able to link features in the analysis to the characteristics in the observing systems. The following elements were identified to be dependent of the Observing System Characteristics:

A. ***Variability and trends due to changes in the observing systems***

- With emphasis on the stratosphere, tropics and Southern Hemisphere
- Systematic errors

B. ***Physical consistency***

- Water and energy cycles
- Integrated general circulation statistics (e.g. E-P fluxes)

C. ***Phenomenological studies***

- Storm tracks, Easterly waves
- Extreme events, Tropical cyclones

As a special and important element it was considered that there is a need for

D. ***Validation against independent data***

4.2. Objectives

Exploration of the information content of very degraded networks, such as surface data only, in preparation for future re-analysis activities

- Addition and Subtraction of Observation Types
- Treatment of data 3D-Var vs. 4D-Var
- Define a minimal observing system

A. Variability due to main observing system changes

For experimentation three main observing systems were defined:

- Present system
- Pre-FGGE = Minimal observing system - surface data, radiosondes, pilots
- FGGE

This roughly corresponds the different Streams in ERA-40 production and they can be re-constituted from the present system. Suggested period for experimentation is 1997 – 2000, which includes the latest El Nino and there is a change from TOVS to ATOVS in August 1998.

Specific experiments to be run for a season, or two

- Already promised as part of ERA-40
  - SSM/I, VTPR, TOVS, ERS-1 and –2, ATOVS
  - 3D-Var vs. 4D-Var
- General assessment w.r.t. different observing systems studies
  - Impact on hydrological and energy cycles, spin-up
- Automated + Space based system
- Altimeter wave heights

B. Physical consistency

Choose particular events where there are known or proxy changes in climate variables

- Pinatubo 199105-199109,
  - Investigate HIRS and SSMI impact on excessive marine rainfall
  - Explore the effect of volcanic aerosol in TOVS radiative transfer
- 1997/98 El Nino
  - Minimal vs. present systems
  - Impact and predictability
  - Validate with TRMM/ TMI

C. Phenomenological studies

As important features, the impact should be studied on:

- Storm tracks, Easterly waves
- Extreme events, Tropical cyclones

D. Validation against independent data

Special high-quality data sets not used in ERA-40 providing opportunities to validate at a higher level than normal

- EOLE, NSCAT, QuikSCAT, GPS, LIMS, UARS, TOPEX, Aircraft campaigns, METEOSAT radiances, TRMM/TMI/PR, CERES and field experiments

4.3. Supporting issues

- Ensemble AMIP run for the whole period, with same O3, greenhouse gases, SST, sea ice, model resolution, ...
  - To serve also as a baseline for the data impact studies
- Promote access to external users for ERA-40 related experimentation
  - Special research projects
  - Could ECMWF issue 'Announcements of Opportunity' to encourage member state active participation
- Promote cooperation with other re-analysis groups

5. Working Group 5: Distribution of ERA-40 Data

J. Slingo, discussion leader/reporter

M. Baldwin, P. Berisford, K. Fielding, Kumar Kolli, Bryan Lawrence, Roger Newson, Andreas Sterl, A. Hernandez

5.1. Background

We noted that:

- The ECMWF Scientific Advisory Committee has recommended that "Council should allow the widest and easiest possible access to ERA40 data by all research scientists.
- The ERA40 aims are to produce and promote use of a comprehensive set of global analyses description of the state of the atmosphere and land and ocean-wave conditions, and to foster European and international research by making the observations, the analyses, and study reports widely available

5.2. Existing Distribution Guidelines

The current implementation of the ECMWF guidelines for NWP data distribution now allows member of a research organisation (or well-defined community, e.g. SPARC, or an EU project) to obtain a copy of appropriate data and distribute it to members, provided that:

- The organisation has obtained permission to do so via a letter to the Director of the ECMWF, and
- All such users are registered, and access to the data is limited to those registered users, and
- The data use is restricted to a specific well-defined time-limited project.

Also it was noted that according to the contract with EU "products should be made available at an affordable price".

We recommend that

- "Analysis of ERA40" be accepted as a specific project (in effect, the last point be waived for ERA40 data).
- The possibility that some WMO training centres or other suitable organisations could act as nodes for redistribution online in developing countries should be considered.

5.3. Major Recommendation

- There should be an online (web) interface to as much of ERA40 as is practicable along the lines of the CDC NCEP/NCAR Reanalysis site.
- Provision of data from this site will require a funded distribution concept, not a self-funded distribution mechanism. Council should be asked to consider funding ERA40 data distribution and/or ECMWF should seek external funding e.g. EU, ESA, Eumetsat etc.
- ECMWF could consider the possibility of collaboration with external data centres to make this possible.
- All dissemination of ERA40 including WEB should be accompanied by scientific metadata (including technical and scientific documentation) in an easily understandable and accessible form..

5.4. Overall comments on proposed CD-Rom for ERA-40

- The proposed contents of the CDs do not justice to the advances of ERA40 over previous efforts and do not reflect user expectations of resolution and variety of contents.
- We recommend that WCRP/WGNE should be asked to provide guidance on the CD contents, perhaps based on standard diagnostic sets proposed (e.g AMIPII).
- We agree that the format of the data should be GRIB, but recommend that every CD should include a GRIB to NetCDF decoder for some common architectures.It was noted that for many researchers GRIB still is a difficult data format to decode.
- We recommend that a software CD should be available with suitable visualisation software, decoders, and graphics tools (e.g. GRADS if licensing can be dealt with).

5.5. Comments on proposed Yearly CD-Rom

- We found it difficult to identify who would be the users of this CD dataset. Rather than produce one yearly dataset, we recommend that:
  - Specific user communities be asked to describe their requirements for yearly data, e.g. stratosphere, boundary-layer and surface processes, general circulation etc.

If multiple yearly datasets are not practicable, then *we recommend an expansion of the yearly CDs to at least two per year, covering all 23 pressure levels for Z,T,U,V and a more comprehensive surface dataset, plus a range of moisture variables in the troposphere.*

5.6. Comments on proposed Monthly CD-Rom

This is a fundamental deliverable, and should be seen as the most important CD product. It needs to be as comprehensive as possible. We recommend that:

- a boxed set covering a wider range of variables and including metadata be made available. The variables should include at least those from the WGNE standard diagnostics for mean climate and variability plus stratospheric material.
- It should be offered at a price consistent of the stated aim of the widest and easiest possible access.

5.7. The Basic Dataset

- We understand that this will be freely available.
- We think the provision of this totally free dataset is an excellent concept which will be of great educational value.
  - We recommend adding surface temperature and surface winds.

5.8. Specialized Datasets

- We encourage joint projects between ECMWF and specific user communities which could produce their own datasets for distribution (either electronically on physical media).
- ECMWF should publicise this option.

5.9. Time Series Dataset

- As we understand it the proposed data set will be extracted at 6 hourly intervals into a special orthogonal MARS database. This will be a considerable advance on ERA15, but will be restricted to specific variables decided a priori.
- We recommend that the entire archive should be available as time-series data at a suitable rate.
  - We would note that the ECMWF could look at utilising external archives to help provide this service.

5.10. Observations

- We recommend that all observational datasets should be available for external onward distribution where licensing constraints are not an issue
- .ECMWF should consider how the quality control and analysis feedback on observations can be made available (again, would ESA or Eumetsat provide funding to help such an initiative?)

5.11. Metadata

We recommend that metadata concerning

- Quality of ERA-40 products including the "bad features",
- Documentation of major observing system changes
- Radiosonde station information
- Description of recommended use of parameters (which precipitation, which flux,...

should be made available in an easily accessible form