ATMOSPHERIC REANALYSES: A MAJOR RESOURCE FOR OCEAN PRODUCT DEVELOPMENT AND MODELING

Kevin E. Trenberth(1), Randall Dole(2), Yan Xue(3), Kazutoshi Onogi(4), Dick Dee(5), Magdalena Balmaseda(6), Michael Bosilovich(7), Siegfried Schubert(8), William Large(9)

(1) National Center for Atmospheric Research, P. O. Box, 3000, Boulder, CO 80307, USA, trenbert@ucar.edu

(2) NOAA Earth System Research Lab., PSD, 325 Broadway, Boulder, CO 80305, USA, Randall.M.Dole@noaa.gov

(3) NCEP/NOAA, 5200 Auth Rd, Camp Springs, MD 20746, USA, Yan.Xue@noaa.gov

(4) JMA, 1-3-4 Otemachi, Chiyodaku, Tokyo, 100-8122, Japan, Konogi@naps.kishou.go.jp

(5) ECMWF, Shinfield Park, Reading, RG2 9AX, UK, Dick.Dee@ecmwf.int

(6) ECMWF, Shinfield Park, Reading, RG2 9AX, UK, Magdalena.Balmaseda@ecmwf.int

(7) NASA/GSFC, Mail Code 610.1, Greenbelt, MD 20771, USA, Michael.Bosilovich@nasa.gov

(8) NASA/GSFC, Mail Code 610.1, Greenbelt, MD 20771, USA, Siegfried.d.schubert@nasa.gov

(9) National Center for Atmospheric Research, P. O. Box, 3000, Boulder, CO 80307, USA, wily@ucar.edu

Abstract: Atmospheric reanalyses have greatly improved our ability to analyse past climate variability. Further improvements to reanalyses, including expansion to encompass the ocean, land and sea-ice domains, hold promise for extending their use in climate change studies, research and applications. Other developments, such as the assimilation of observed cloud and coupled reanalysis, are needed before the considerable potential for providing reliable surface fluxes for the ocean is realized.

1. INTRODUCTION TO REANALYSIS

Atmospheric analyses provide a synthesis of the available observations in the context of a physical model. Global analyses have been routinely done since the late 1970s for purposes of numerical weather prediction (NWP). These atmospheric analyses were instrumental in shaping our understanding of climate variations on relatively short time scales, but the frequent changes in procedures used introduced many spurious variations in the perceived climate leading to a call to reanalyze the past observations using a constant state-of-the-art data assimilation system [1], [2].

The first generation of atmospheric reanalyses in the mid to late 1990s at NCEP [3] (called NCEP/NCAR), NASA/Goddard [4] and ECMWF (called ERA-15) [5] had substantial problems [6] that limit their use, particularly for global climate change and variability studies. A second limited version of the NCEP reanalysis (called NCEP/DOE) was run to address some problems [7] but is still a first generation reanalysis. The NCEP/NCAR and NCEP/DOE reanalyses, often referred to as Reanalysis 1 (R1) and Reanalysis 2 (R2), incorporated the atmospheric model operational in 1995. R1 and R2 are also extended into real time. The real time R1 has been used by the Climate Prediction Center of NCEP to produce global atmospheric monitoring and assessment products, while the real time R2 has been used to provide atmospheric initial conditions for the Climate Forecast System (CFS) used for operational seasonal forecasts [8]. The surface fluxes of R2 also provide the boundary forcings for the NCEP’s Global Ocean Data Assimilation System (GODAS), which in turn provides oceanic initial conditions for CFS [9].

In spite of their shortcomings, the reanalysis products have proven to be among the most valuable and widely used in the history of climate science, as indicated both by the number of scholarly publications that rely upon them and by their widespread use in current climate services. Besides being based on now outdated and low spatial resolution assimilation systems and containing significant systematic errors, another serious problem was effects of changes in the observing system that produced spurious changes in the perceived climate. As a result, trends and low frequencies are unreliable; this problem is exacerbated by model bias.

Two second generation global reanalyses, the ECMWF ERA-40 [10] and the Japanese Meteorological Agency 25-year reanalysis (JRA-25) [11] have addressed some of the shortcomings of the earlier reanalyses, but many of the problems tied to observing system changes and model deficiencies remain. Further reanalyses are underway and planned, see section 3. A summary of the current and planned reanalyses is given in Table 1.

2. REANALYSES: EVALUATIONS, ADVANTAGES, PROBLEMS, AND SHORTCOMINGS

The comprehensive nature of the observing system necessary to document the behaviour of the global climate system is described in [12], while [13] outlines the needs for reanalysis to be an ongoing program, and [14] summarize the progress, issues and future promise.
Table 1. Summary of the main atmospheric reanalyses that are current or underway, with the horizontal resolution (latitude; T159 is equivalent to about 0.8°), the starting and ending dates, the approximate vintage of the model and analysis system, and current status.

<table>
<thead>
<tr>
<th>Reanalysis</th>
<th>Horiz. Res.</th>
<th>Dates</th>
<th>Vintage</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCEP/NCAR R1</td>
<td>T62</td>
<td>1948- present</td>
<td>1995</td>
<td>ongoing</td>
</tr>
<tr>
<td>NCEP-DOE R2</td>
<td>T62</td>
<td>1979- present</td>
<td>2001</td>
<td>ongoing</td>
</tr>
<tr>
<td>CFSRR (NCEP)</td>
<td>T382</td>
<td>1979- present</td>
<td>2009</td>
<td>in progress</td>
</tr>
<tr>
<td>C20r (NOAA)</td>
<td>2°</td>
<td>1891-2008</td>
<td>2009</td>
<td>in progress</td>
</tr>
<tr>
<td>ERA-40</td>
<td>T159</td>
<td>1957-2002</td>
<td>2004</td>
<td>done</td>
</tr>
<tr>
<td>ERA-Interim</td>
<td>T255</td>
<td>1989- present</td>
<td>2009</td>
<td>ongoing</td>
</tr>
<tr>
<td>JRA-25</td>
<td>T106</td>
<td>1979- present</td>
<td>2006</td>
<td>ongoing</td>
</tr>
<tr>
<td>JRA-55</td>
<td>T319</td>
<td>1958-2012</td>
<td>2009</td>
<td>underway</td>
</tr>
<tr>
<td>MERRA (NASA)</td>
<td>0.5°</td>
<td>1979- present</td>
<td>2009</td>
<td>in progress</td>
</tr>
</tbody>
</table>

In principle, through the reanalysis process, more complete and quality controlled observations are used with advantage of hindsight of problems in the original analysis, and with the benefit of a more up-to-date and constant state-of-the-art model. The full synoptic variability of the atmosphere and its evolution is captured. Comprehensive global gridded fields of variables and fluxes result from this process.

The several reanalyses that have been conducted have used a stable data assimilation system and have produced fairly reliable atmospheric climate records that have enabled (i) climatologies to be established; (ii) anomalies to be calculated; (iii) empirical and quantitative diagnostic studies to be conducted; (iv) exploration and improved understanding of climate system processes to be developed; and (v) model initialization and validation to be performed. The products provide the essential foundation for an accurate assessment of current climate ("climate nowcasts"), diagnostic studies of features such as weather systems, monsoons, El Niño-Southern Oscillation and other natural climate variations, seasonal prediction, and climate predictability. The reanalyses have provided a vitally needed test bed for model improvement on all time scales, especially for seasonal-to-interannual forecasts. Moreover, the basic assimilation and prediction systems are improved as deficiencies are identified and corrected by applying them both in reanalysis and routine weather and climate prediction.

Global reanalysis is also the foundation for regional reanalysis projects and downscaling where detailed climatologies can be prepared to support studies of local climate and climate impacts. There has been some progress in the use of reanalysis to investigate the difficult problem of the detection and attribution of long-term climate trends and variability. Reanalysis in the ocean and atmosphere has helped identify and correct deficiencies in the observational record, including the recovery of additional observations. Hence greatly improved basic observations and data bases are a side product of reanalyses.

The disparity in time scales of variability in the atmosphere and ocean are an issue to be properly addressed. It is not sufficient to use time average (e.g., weekly or even daily mean) values of the atmosphere for driving the ocean. At present the high frequency synoptic fluctuations are well reproduced by atmospheric reanalyses, but systematic biases preclude the use of the fields without an adjustment of some sort, and variability on longer time scales (especially decadal) is not well captured by current reanalyses (e.g., [16], [17]). The primary causes of this deficiency are the quality and homogeneity of the fundamental data sets that make up the climate record and the quality of the data assimilation systems used to produce reanalyses.

Research into bias corrections and advanced reanalysis techniques is showing promise, and further reanalysis efforts are needed. A potentially significant advance would be the successful assimilation of observed cloud, such that both cloudy and clear areas persist into a forecast. There is little focussed effort in this area, but until progress is realized, radiation and precipitation...
products are likely to remain seriously deficient. A challenge is to improve estimates of uncertainty in the reanalysis products. Problems of biases in models and data are intricately connected with this challenge, because the ability to provide meaningful uncertainty estimates for reanalysis products ultimately depends on having information about the accuracy of the input data. Satellite instruments in particular can have substantial systematic errors that can dwarf the useful signal in the data. These errors are different for each instrument and can vary in space and time in a complex manner. In well-observed regions of the atmosphere it is possible to automatically correct such errors during the reanalysis procedure (e.g., see Fig. 1).

Some biases arise because the reanalyses are produced with specified sea surface temperatures (SSTs) that can not respond to large fluxes in one direction or the other, and thus there is an infinite source of heat and moisture at the surface that is not possible in a fully coupled system. Hence, even for a model that is perfectly in energy balance in the framework of the coupled system, as long as the model contains biases, the energy balance will be disrupted by specifying the SSTs. As typical errors in net surface energy flux are order 10 W m\(^{-2}\) on a global basis [18], and errors are much larger locally: they are non-trivial.

The products of special relevance to the ocean are the surface fluxes of radiation (shortwave solar radiation and longwave radiation), fresh water fluxes of precipitation and evaporation as well as runoff and river discharge from land, and fluxes of sensible heat and momentum. A serious impediment to the evaluation of these fluxes is the large range in observational estimates. However, useful constraints are that the net global ocean surface heat gain and water loss should be less than about 2 W m\(^{-2}\) and about 10 cm/yr (to balance runoff), respectively. Over the semi-enclosed Mediterranean Sea the climatological heat loss should be about 6.3 W m\(^{-2}\) and the water loss between -0.5 and -1.0 m/yr to balance the observed inflow near Gibraltar. Reanalyses also have the potential to provide fluxes of trace substances and gases, but again evaluation would be problematic.

### 3. FUTURE PROSPECTS

Successive generations of reanalyses have led to continual improvement of the basic dataset and improved results. New reanalysis efforts are underway in several institutions world-wide. The results to date suggest that it is important to advance the reanalyses by getting away from specified SSTs as a boundary condition and use at least a mixed layer ocean model (perhaps relaxed to observed SSTs), and surface flux diagnostics are revealing as to the model biases. A frontier for future research is the development of fully coupled data assimilation and reanalysis. Improved models with reduced biases are also needed, as are systems that more fully utilize existing observations (e.g., clouds).

#### 3.1. NCEP: CFSRR: Climate Forecasting System Reanalysis and Reforecast

A new reanalysis of the atmosphere, ocean, sea ice and land over 1979-2009 is being produced by NCEP under a project referred to as the CFS Reanalysis and Reforecast (CFSRR). It will be used to provide initial conditions for a reforecast of the new CFS over 1981-2009 for calibration and skill estimates of the CFS for operational seasonal prediction at NCEP. There are three main differences from R1 and R2: 1) much higher horizontal and vertical resolution (T382L64 (about 35 km) vs. T62L28) with sigma-pressure hybrid levels; 2) the guess forecast will be generated from a coupled atmosphere-ocean-ice-land system; 3) radiance measurements from the historical satellites will be assimilated. The oceanic reanalysis is similar to GODAS except the first guess will be provided by CFS and the Argo salinity will be included, and with 40 levels in the vertical, to a depth of 4737 km, and a high horizontal resolution of 0.25° in the tropics, tapering to a global resolution of 0.5° poleward of 10°N and 10°S.

The new CFS involves improvements to: the atmospheric data assimilation via the new NCEP Gridstered Statistical Interpolation Scheme (GSI); the physics and dynamics of operational NCEP Global Forecast System (GFS); the data assimilation of the ocean and ice with the NCEP GODAS, and a new GFDL MOM4 Ocean Model; the data assimilation of the land with the NCEP Global Land Data Assimilation System (GLDAS) and a new NCEP Noah Land Model.

The surface fluxes in the NCEP operational GDAS have superior quality to those of R1 and R2 (see following comments). It is hoped that the CFSRR project will produce significant improvements in the NCEP surface fluxes, which will be used to improve ocean modelling, ocean analyses and ocean products in the near future.

To date the surface fluxes from R1 and R2 have been widely used in producing ocean simulations and ocean analyses, and have been evaluated using various observation-based products. Uncertainties in the latter must be accounted for in evaluations [17]. However, the mean zonal wind stresses in the tropical Pacific are too weak in R1 and R2, and its interannual variability is too low in R1 and too high in R2. Model errors are very large in the cold tongue region and more observations may be needed to constrain model solutions (Fig. 2). The latent heat fluxes in R1 agree with those in OAFlux [19] (not shown) -- although biases exist [20]. However, in R1 a very large transfer coefficient has been used to compensate for winds that are too weak and too much moist surface air. Net surface heat fluxes...
The mean precipitation (P) and evaporation (E) in R1 agree well with observation-based analyses, but both P and E in R2 have excessive amplitudes over most of the global ocean. The E-P in R2 is too low in the tropics, but is close to observations in the extra-tropics due to cancellation of biases in E and P [21]. Therefore, an important activity in CFSRR is to evaluate all flux components and continue to reduce errors.

3.2. NASA/Goddard: Modern Era Retrospective Analysis for Research and Applications (MERRA)

The NASA/Global Modeling and Assimilation Office (GMAO) atmospheric global reanalysis project is called the Modern Era Retrospective-Analysis for Research and Applications (MERRA). MERRA [22] is based on a major new version of the Goddard Earth Observing System Data Assimilation System (GEOS-5), that includes the Earth System Modeling Framework (ESMF)-based GEOS-5 AGCM and the new NCEP unified grid-point statistical interpolation (GSI) analysis scheme developed as a collaborative effort between NCEP and the GMAO. MERRA supports NASA Earth science by synthesizing the current suite of research satellite observations in a climate data context (covering the period 1979 to present), and by providing the science and applications communities with a broad range of weather and climate data.

Development and validation of the data assimilation system for MERRA emphasized improvements in hydrologic cycle estimates. MERRA products consist of a host of prognostic and diagnostic fields including comprehensive sets of cloud, radiation, hydrological cycle, ozone, and land surface diagnostics. One hourly 2D and surface fields (including lowest model level meteorology) will support the development of offline land and ocean surface model development. A special collection of data files are designed to facilitate off-line forcing of chemistry/aerosol transport models.

The model or native resolution of MERRA is 0.667° longitude by 0.5° latitude with 72 levels extending to a pressure of 0.01 hPa. Analysis states and two-dimensional diagnostics are available at the native resolution, while many of the three-dimensional diagnostics are available on a coarser 1.25° grid. MERRA production is complete for 1979 to 2005 as of August 2009.

MERRA products are available on-line at the Modeling and Assimilation Data and Information Service Center (MDISC) at http://disc.sci.gsfc.nasa.gov/MDISC/. Further information about MERRA and its status may be found at http://gmao.gsfc.nasa.gov/research/merra.

Preliminary evaluations of the MERRA energy budgets show that an imbalance at the surface of ~13 W m⁻² is largely related to an over estimate of incoming shortwave radiation over the ocean and to a lesser extent an underestimate of ocean evaporation [23]. MERRA was produced in 3 segments, enabling the temporal variability of the energy budget to be explored.
Figure 3. Annual means of ocean average surface flux anomalies from MERRA. For 1979 to 2005 the mean values are latent heat (LE) 87.7 W m\(^{-2}\); sensible heat (Hs) 11.7 W m\(^{-2}\); downwelling shortwave radiation (SWdn) 196.1 W m\(^{-2}\); and downwelling longwave radiation (LWdn) 342.5 Wm\(^{-2}\).

Fig. 3 shows the annual global area average of the anomalies of some surface energy budget terms. With the decreases of incoming shortwave and increasing latent heat, both these quantities and the imbalance of energy at the ocean surface are improving as time moves toward more and better observations. This underscores the challenge as reanalyses move toward integrated system analyses, in which significant variations in surface forcing can occur with changes in the observing system.

In addition to MERRA, the GMAO is actively pursuing advanced capabilities in land surface, atmospheric constituent, and physical and biological ocean assimilation, with the goal of developing an Integrated Earth System Analysis (IESA) capability. As an interim step, the GMAO is developing a phased IESA capability that utilizes the MERRA product and an ocean data assimilation system to explore data assimilation in a fully coupled climate model.

### 3.3. ECMWF: ERA-interim and ERA-75

ECMWF is currently producing ERA-Interim, a global reanalysis of the data-rich period since 1989. Relative to the ERA-40 system, ERA-Interim incorporates many important model improvements such as resolution and physics changes, the use of four-dimensional variational (4D-Var) data assimilation, and various other changes in the analysis methodology.

The configuration of the ERA-Interim system and many aspects of its performance are described in ECMWF Newsletters 110 and 115 (see [24], [25] and http://www.ecmwf.int/publications/newsletters). Once it catches up with real-time in early 2009 ERA-Interim will be maintained as a Climate Data Assimilation System (CDAS). From the oceanic point of view, the ERA-Interim forcing fluxes represent an improvement with respect to ERA-40. The main differences are in the fresh water flux and in the solar radiation.

The differences in total solar heat flux (Fig. 4) have a large scale spatial structure, with more solar radiation in ERA-Interim mainly over the convective areas, and less solar radiation over the stratocumulus regions. There are also differences in the surface winds, which are generally stronger in ERA-Interim as a consequence of the increased horizontal resolution. Differences in the interannual variability of the wind fields from ERA-Interim and ERA-40 are large in the tropical Indian and Atlantic oceans, where it can exceed the 30% of the interannual variance.

Two ocean models have been used to evaluate the quality of ERA-Interim forcing fluxes, using ERA-40 forcing fluxes as a baseline. Such exercises are fraught with danger, because model and forcing error cannot be separated and there is always a distinct possibility of compensating error. Nevertheless, both models indicate a reduction of the mean SST error, especially over the warm pool area, which is likely a consequence of the improved solar radiation in ERA-Interim. Altimeter data have been used to evaluate the quality of the interannual variability of the model sea level, which is directly related to the interannual variability of the winds. Both models indicate that the interannual variability of the ERA-Interim winds is better than ERA-40 (which was already a big improvement over other wind products). Seasonal forecast experiments exhibit better skill scores when initialized with ERA-Interim fluxes versus those initialized with ERA-40.

As the name suggests, ERA-Interim represents a step towards ECMWF’s next generation reanalysis system. This reanalysis, tentatively called ERA-75, will span at least a 75-year period, extending back in time to the first half of the 20th century when substantial numbers of
upper-air meteorological observations began to be made available on a regular basis. Depending on available resources, the target is to begin producing ERA-75 in 2013.

3.4. JMA: The second Japanese reanalysis project JRA-55

Following the successful completion of the JRA-25 [11], the second Japanese atmospheric reanalysis project JRA-55 started in 2009. JRA-25 has many advantages and contributes substantially to climate monitoring and research, but as it covers only 26 years, it is not of sufficient length for climate monitoring. It has some deficiencies such as large temperature biases in the stratosphere (arising from an old radiation scheme) and it caused unnatural jumps in time series of mean temperature with changes of satellite data. The assimilation and forecast models used in the JRA-25 were taken from the operational NWP models as of April 2004.

JMA has updated many aspects of the operational NWP model both in data assimilation and forecast models since then and JMA NWP has significantly improved in recent 5 years (Fig. 5). In JRA-55 the NWP model to be used is much improved and the reanalysis period will be extended for 55 years from 1958 to 2012, which covers both cooler years before 1970s and warming years after 1980s. It includes the following major updates from JRA-25: 1) improved data assimilation (4D-var vs 3D-var); 2) increased model resolution (T319L60 with a reduced Gaussian grid system vs T106L40 in JRA-25); 3) variational bias correction for satellite radiance data to reduce the jumps; 4) new radiation scheme; 5) updated dynamical and physical processes; 6) increases in greenhouse gas (carbon dioxide) (vs constant in JRA-25); 7) updated 3-dimensional daily ozone profile data; 8) use of observation feedbacks from JRA-25 to detect observation change correctly.

3.5. Historical 20th Century reanalysis project and plans

A promising new ensemble of reanalyses based solely on surface observations, and sea level pressure and sea surface temperature observations in particular, is underway with a goal to provide over 100 years of reanalyses along with uncertainty estimates [26]. The NOAA-CIRES led historical reanalysis project, the Twentieth Century Reanalysis Project (C20r), is using a state-of-the-art data assimilation system and surface pressure observations to generate a six-hourly, four-dimensional global atmospheric dataset spanning 1871-2008 to place current atmospheric circulation patterns into a historical perspective [27].

The first version has global coverage spanning 1908-1958, and two degree longitude-latitude horizontal resolution. Output includes the complete suite of surface momentum, heat, and moisture fluxes and winds necessary to force ocean and wave models. Output is available as of late March 2009 from NOAA ESRL, NOAA NCDC, and NCAR. More details on the project and data available can be found at http://www.cdc.noaa.gov/data/gridded/data.20thC_Rean.html.

This reanalysis will provide the first estimates of global tropospheric and stratospheric variability over such a long period. In the northern hemisphere, such reanalyses should provide useful descriptions of the atmospheric circulation and a physically consistent set of atmospheric variables dating from before 1900. The results provide ways to carry out consistency checks with data not used, such as early radiosondes. A preliminary experiment (Dr. B. Giese) suggests that SST variability is well-reproduced by an ocean simulation forced with C20r fields suggesting that oceanic data assimilation may be possible.

Also planned in 2011 is a 1° global reanalysis back to 1850 entitled Surface Input Reanalysis for Climate Applications (1850-2011), with data availability in 2012. This historical reanalysis will benefit from...
enhanced data recovery being facilitated by the international initiative Atmospheric Circulation Reconstructions over the Earth <http://www.met-international.org>, the NOAA Climate Database Modernization Program, and by working groups of GCOS and WCRP.

4. CONCLUDING REMARKS

Seasonal and longer term forecasts using a coupled climate model require initialization with both oceanic and atmospheric data, with the latter likely coming from atmospheric or coupled analyses. Methods and data for initializing coupled models for seasonal prediction have been tested [28] and results show that atmospheric reanalysis (atmospheric fluxes from ERA-40) are instrumental in the improvement of seasonal forecast skill. In most of the areas, the impact of atmospheric reanalysis is complementary to the assimilation of ocean observations, in that their contributions are additive, and all observations help.

Reanalyses to date have mainly focussed on producing the best set of analyses given the available observations. To reduce spurious trends, it is highly desirable to have one or more reanalyses performed with the objectives of producing the most consistent climate record, and which therefore confronts the observing system changes in new ways. This may require further observing system experiments and use of reduced observational datasets. Some plans are underway along these lines but have yet to take concrete form.

While the origins of reanalysis have been in atmospheric climate and weather, there have been significant studies of reanalysis (or synthesis) of ocean data [29], some of which use inverse modelling techniques. Because of the limited size of the historical ocean data sets, it has been necessary to develop novel techniques for increased homogeneity of ocean reanalysis. However, it is unclear if the observations and assimilation systems are sufficient to overcome model and forcing error. Nonetheless, adding ocean observations and their synthesis may provide further constraints on atmospheric reanalyses.

Other promising developments are occurring in sea ice, Arctic, and land surface reanalysis. There has also been initial development of coupled atmosphere-ocean data assimilation, which transforms the forcing problem and lays the foundation for future coupled reanalysis studies that may lead to more consistent representations of the energy and water cycles. Hence, with the ongoing development of analysis and reanalysis in the ocean, land and sea ice domains, there is huge potential for further progress and improved knowledge of the past climate record.

Reanalysis has proved to be as valuable for monitoring climate, climate research and applications as was believed when it was proposed twenty years ago. However, as the scope of global reanalysis grows, the research effort needed to optimise the benefits is so large that international cooperation will be essential. The potential for products useful for oceanography that are both bias free and which capture the variability on time scales from hours to decades is great.

5. REFERENCES


