Report of the Workshop on "The Development of Improved Observational Data Sets for Reanalysis: Lessons learned and Future Directions"

28-29 September 2005
University of MD Conference Center, College Park, MD USA
Sponsored and support by: NASA, NOAA, NSF

Organizers:

Siegfried Schubert (NASA/GMAO)
Dick Dee (ECMWF)
Sakari Uppala (ECMWF)
Jack Woollen (NOAA/NCEP)
John Bates (NCDC)
Steven Worley (NCAR)

Key contributors:

Joey Comeaux (NCAR)
Russell Vose (NCDC)

1 May 2006

Table of Contents

Executive Summary

The workshop identified the required components and technical capabilities of a program for "The Development of Improved Observational Data Sets for Reanalysis," including international partnerships, institutional responsibilities, and priorities. This report summarizes the background information and recommendations for developing such a program including the coordination necessary for developing specific work packages (who does what) and their time dependencies. It was recognized that the timely development and shepherding of such a program requires strong leadership by individuals experienced with historical as well as real-time conventional and satellite data, and that the appointment of an individual or individuals to take on this responsibility be made as soon as possible at the international level. While determining specific resource requirements were beyond the scope of this workshop, the resources and funding of the program must allow for timely data refresh cycles commensurate with the needs of an on-going iterative process of reanalysis coordinated between the major centers. *As such, the key programmatic recommendation of the workshop is for the*

WCRP Observations and Assimilation Panel (WOAP) to appoint a working group of experts charged with developing a plan for "The On-going Development of Improved Observational Data Sets for Reanalysis", that describes the necessary resources, infrastructure, institutional commitments, and coordination on technical issues outlined in this report.

Action: Get commitments from member countries to carry out specific tasks as outlined below in the scientific and technical recommendations and action items.

The workshop also identified the need to raise the profile of activities concerning the pre, post, and bias processing of observations and the related science. It was recognized that the improved understanding of climate variability and change depends more than ever on the quality of the historical observations and their effective use in reanalyses.

The specific technical and scientific recommendations and action items are as follows:

Recommendation 1: All the main centers should prepare inventories of reanalysis observations (on the level of observation records)

Action 1a: Compare NCAR and NCDC inventories for surface and upper air.
Joey Comeaux will draft inventory record structure; will iterate with NCDC to reach mutual agreement. Russ Voss is the contact at NCDC

Action 1b: Draft inventory record structure for moving platforms.
Consult with ICOADS and WOB experts to assure compatibility.
Responsible individuals- Steven Worley and Joey Comeaux

Action 1c: Following Actions 1a and 1b, get inventories in established forms from JMA, ECMWF, NCEP, NASA, BoM for data not from NCDC, NCAR,

or in ICOADS and WOD. Contacts, JMA – Kasutoshi Onogi, ECMWF – Sakari Uppala, NCEP – Jack Woollen, NASA – Siegfried Schubert, BMRC – Peter Steinle, CRU East Anglia Univ – Phil Jones

Recommendation 2: A collaboration be formed that can sustain a data refresh cycle and create high quality merged datasets for reanalyses

Action 2a: Begin to collaboratively explore merging/combining activities with radiosonde data because of its vertical data structure and potentially large impact on reanalysis.

Action 2b: Add ancillary data (reanalysis feedback files – add specific fields; other expert calibration and inter-calibration information) to merged data set database.

Action 2c: Further inter-calibrate various measurements across successive generations of satellite sensors, with priority given to SSU data

Recommendation 3: Develop improved record tracking control for observations to further improve the use of feedback data from reanalyses targeted especially for data providers/developers:

Action 3a: Provide additional details on metadata efforts and standards for in situ and satellite data, including a pilot rich metadata standard for satellite data, by spring 2006 (John Bates)

Action 3b: Establish working group to determine standard set of record header information that satisfies recording track through the reanalysis system.

Action 3c: Provide a list of web sites relevant to data quality control for reanalysis and real-time NWP analysis. Post at the WMO WOAP web site. (Glenn White).

Recommendation 4: The observational, reanalysis, and climate communities should take a coordinated approach to further optimizing reanalysis for climate.

Action 4: Set up working group (Schubert, Dee, Whittaker, Thorne, etc.) to focus on the following reanalysis issues:

- All data vs. stable observing system including reduced resolution data
- What are the highest impact set of observations (priority of, e.g., SSU data)?
- Developing and testing adaptive bias-correction techniques for reanalysis
- Enhance community participation to interact and analyze reanalysis
- What can reanalysis do toward recommending sampling distribution and supporting the need for a "reference" upper air network?
- How can new observations best be tested/evaluated in reanalysis/OSEs?
- Establishment of ongoing relations with major data centers

1. Motivation

New observations can only tell us something about the climate if they can be put in the context of past observations. Are the observations indicating something unusual is happening? Are similar or related changes occurring in other parts of the globe, and/or in other quantities? Are they part of a trend? If so, what is forcing that trend, and will it continue? Are the observations during the course of a season indicating more or fewer extreme weather events compared with “normal”? The answers to such questions require climate data sets that are consistent over time and space, and that include estimates of both the climate state and its forcing.

Four-dimensional climate data assimilation, in which new observations are integrated with our historical record and with model fields provides the best approach to understanding such climate variability and ensuring that climate models are consistent with the full range of observed climate variations. Four dimensional data assimilation in fact provides information, not only about the quality of the models, but also about the quality of the observations. *As such, a systematic and iterative process of climate data integration based on four-dimensional data assimilation can facilitate both model development and efforts to improvement the long-term consistency and quality of the observations*.

The feasibility of employing data assimilation systems to reprocess past observations has been demonstrated with the production of several global “reanalysis” data sets in the U.S., Europe, and Japan. These datasets provide the most consistent representation of climate variability to date and are now widely used by the scientific community in a variety of applications including atmosphere-ocean interactions, seasonal prediction, climate monitoring, the hydrological cycle, and a host of regional and other diagnostic studies (see e.g., WCRP 1998; 2000, Arkin 2004, and Appendix).

The first generation of reanalyses, however, had problems that made them sub-optimal or even unusable for some applications. Perhaps the most serious problem for climate applications was that, while the assimilation system remained fixed, changes in the observing systems did produce spurious changes in the perceived climate (see e.g. Trenberth 1995, WCRP 1998; 2000). *These changes in the observing system have continued and, in fact, WCRP has identified the need to solve the problem of these spurious changes and trends as the most important problem in reanalysis.*

In light of these problems it is imperative that future reanalysis efforts include a long-term strategy for minimizing the impact of observing system changes (e.g., Schubert and White 2004). *A key component of that strategy must be a strong collaboration between the reanalysis and observational communities to improve our existing world-wide database of input observations.* This includes efforts aimed at the identification and correction of observational bias, correction of obvious errors, improved quality control, the merging of various data sets, rescue and organization to create more complete digital collections, improved handling of meta-data, as well as developing and testing adaptive bias-correction techniques that adjust to an ever-changing observing system.

This workshop represents a first step in specifically addressing observational needs for future climate-oriented reanalysis efforts. The overall goal of the workshop was to develop a strategy for minimizing the impact of observing system changes in future reanalyses. This report summarizes the basic findings and recommendations of the workshop – these include recommendations for a world-wide inventory of data suitable for reanalysis, meta-data needs, research priorities, and collaborative projects that will accelerate progress on improving observational data sets for future climate reanalyses. It is acknowledged that reanalysis efforts are a cyclical process. Observations are collected and prepared to a reasonable optimum level at a point in time, an assimilation model is designed and tested, production runs analyze the data into a long-term climate record, results are studied, data problems and assimilation techniques are critiqued, new climate understanding is established, results foster activities to improve the data (fix discovered problems, add sources from throughout the period of record, refine bias corrections and homogenization methods), and advances in assimilation procedures and methods lead to another reanalysis that has strong promise to reveal more accurate climate insight.

The next section provides background on what has been learned from previous reanalysis efforts and workshops regarding observational needs. Section 3 presents emerging capabilities and provides a vision for the future based on the workshop presentations and discussions, and section 4 details our recommendations and associated initial action items.

2. Background

As mentioned above, several reanalysis data sets have already been produced. These "first and second generation" products include those produced by NCEP/NCAR (Kalnay et al. 1996), NCEP/DOE (Kanamitsu et al. 2002), NASA's Data Assimilation Office (Schubert et al. 1993), the ECMWF (ERA-15 – Gibson et al. 1994; ERA-40 – Simmons and Gibson 2000, Uppala et al. 2005), and the Japanese Meteorological Agency/Central Research Institute of Electric Power Industry (JRA-25). In addition, a new NASA/Global Modeling and Assimilation Office (GMAO) reanalysis (the Modern-Era Reanalysis for Research and Applications - MERRA) is scheduled to begin in 2006, and there are plans for follow-on reanalyses at both NCEP and ECMWF. Those efforts, by using fixed analysis systems, eliminate the artificial climate signals that occurred in analyses generated at the operational numerical weather prediction centers resulting from changes in the model and analysis systems.

Those reanalyses have provided vitally needed globally consistent data sets for weather and climate research (see e.g., Appendix) as well as an important test bed for model development and validation. Arguably, these reanalyses data products have been one of the most significant contributors to our understanding of climate variability in the past few decades. Nevertheless, extensive changes to the observing system strongly affect the variability that is inferred from reanalyses especially at longer time scales (e.g., Arkin et al. 2004). In particular, inferred trends and low frequency variability are of limited reliability, a result exacerbated by model bias. Budgets of momentum, heat and moisture calculated from reanalyses do not balance, reducing the confidence in diagnostic studies based on the products. The hydrological cycle suffers from sensitivity to approximations in the model physics, such as the handling of atmospheric convection on scales finer than the model grid. The diurnal cycle of cloudiness and precipitation over continents during warm seasons is poorly represented. The reanalyzed fluxes between the atmosphere and the surface exhibit unrealistic behavior, limiting both their utility for applications such as forcing models of ocean circulation and the ability to perform coupled assimilations of the atmosphere and ocean or land surface.

A number of issues must be addressed for achieving more accurate and less biased future climate analyses. The primary goal of reanalysis conducted so far has been to produce the best analysis, given available data. This inevitably makes the set of reanalyses inhomogeneous, reducing confidence in trends and long-term variability. Existing reanalyses have been created using four dimensional data assimilation, a process developed for numerical weather prediction, where the goal is to produce the best forecast, not the best analysis. Such constraints may sometimes limit the capability of analysis systems to utilize the full historical observation database. Furthermore, while true four-dimensional data assimilation capabilities have been developed, operation numerical weather prediction realities (the need to have an analysis and a forecast ready for use promptly) mean that actual analysis procedures do not use data after the time of the analysis in the same manner as data prior to that time. This constraint is not very relevant for climate analyses, and modified techniques may be needed.

An important, but perhaps under appreciated aspect of the first generation of reanalyses is that they fostered substantial improvements to the basic input observations and databases. New climate analyses and reanalyses can now take advantage of observations that have been rescued into digital form, collected from previously untapped sources, corrected for obvious errors, quality controlled, and quality checked as part of previous reanalysis efforts. While some of these data enhancements are an important by-product of reanalyses, there are as yet no agreed-upon standards on how or even what information about the observations should be archived in order to insure that each successive reanalysis can fully benefit from the observational knowledge base of prior reanalyses as well as contribute to an on-going, iterative improvement to our climate observations. This is largely a data stewardship problem that is only solved by close inspection of the data by experts and varied approaches that establish methods to improve the quality and/or accurately document the quality.

In fact such efforts to improve climate observations must go beyond the traditional reanalysis community to include reprocessing activities. Reprocessing involves applying a series of quality control and analysis procedures to observational data sets to develop climate data records. As reanalysis has developed and matured, it has become clear that both reanalysis and reprocessing efforts could benefit, mutually, from improved interaction and communication between the reanalysis and data steward communities. Reprocessing efforts have developed over time most often in response to specific scientific questions such as what has been the trend in surface temperature over the past 150 years, or does the upper tropospheric water vapor feedback act in the way that general circulation models simulate it? Answering such questions leads to a specific focus on one, or a limited set, of geophysical variables. The preparation of data sets, quality control, and reprocessing of data under this rubric has led to important advances in monitoring and understanding of the climate system. Over time, however, it has become clear that a more holistic approach is needed to monitor and predict the Earth system. To best make progress in this new framework, we must re-examine our old practices and develop a more integrated framework going forward.

For in situ data sets, the traditional reprocessing framework has resulted in the production of multiple data sets for a given geophysical variable. This is both a strength and weakness. For critical climate variables, an essential component of developing the best observational data set is to have multiple groups bring different perspectives to quality control and data set preparation, particularly in the early stages of the maturity of a product. As a particular product becomes more mature, however, it is increasingly important for product producers to exchange information amongst themselves and, where possible, come to consensus on particular aspects of the data processing. We are now at that point for both reprocessing and reanalysis efforts. This does not mean that just one group does data processing, it means that we recognize a progression and maturity of the data sets over time and can move on to newer or more difficult questions regarding the data. In addition, it is important to recognize that reprocessing and reanalysis are just different sides to the same coin and each must be supported, both independently and mutually, to make optimal progress.

Issues that must be addressed include, duplication of data sets, ad hoc approach to data preparation, difficulties in taking advantage of previous reanalysis, poor traceability, and bringing in work on observations that is independent of reanalysis. This workshop was held to identify and support processes for addressing these issues. It is clear that for an ongoing reanalysis of the climate system to be successful, different reanalysis groups and data centers must collaborate on common issues. For this reason, this workshop concentrated on more practical and technical issues associated with both reanalysis and reprocessing of climate data sets from in situ observations and from satellites.

3. Vision for the Future

a. The use of observations in future reanalyses

Current practice is for each reanalysis project to take responsibility for collecting historical observations as best as it can and merging the various observations into an input dataset for the analysis. This rather ad hoc approach to the collection and merging of data has worked well, but it often relied on the good will and personal devotion of a few key people in the responsible organizations.

As we begin work on the next generation of reanalysis products, it has become clear that it is not sufficient to simply use the observational input from a previous reanalysis, since many shortcomings in the data and their usage are usually identified in the evaluation of the earlier reanalysis products. Each successive reanalysis project can, in fact, extract a greater number of observations and more accurate information from them by taking advantage of new developments in data-assimilation systems, the handling of biases in observations, and in the metadata.

The common practice has been to create merged input datasets: one input stream for all the conventional observations, and separate input streams for each satellite system. This is a major task that requires continuous effort, and broad knowledge from data experts on the characteristics of the observing systems, including their history. New data continue to become available from both data archeology and the current suite of operational observing systems. The result is that the magnitude of this effort increases each year, since the massive amount of new operational data has to be merged into the input streams.

Reanalysis makes use of the input data and utilizes all the available metadata. At the same time reanalyses create the feedback metadata for each input stream. This metadata contains the information about how the reanalysis has used each datum with the departure information attached.

Future reanalysis projects are dependent on the activities of the data collection centers around the world. In our vision of reanalysis as an on-going iterative process, a more coordinated program is required to more efficiently update and enhance the input data streams. This includes taking advantage of the feedback from the previous reanalysis efforts in the creation of the next input dataset.

The creation of input datasets for reanalysis can be viewed as an important contribution to our stewardship of global climate observations, their maintenance for future use, and the extraction of reliable information concerning climate variability and change. It is a unified approach to the observing system and its history spanning both conventional and satellite data. It is a major effort that needs on-going institutional support. It also needs international support (e.g., WMO) to open data policies so that access to important data can be achieved worldwide.

Recent technological advances have been important not only in terms of improved data storage and data handling, but also in the availability of high performance data base systems. These systems allow the maintenance of different data sources separately using version control, so that reanalysis projects can extract or enquire new merged versions from the sources. Furthermore, the databases can be appended by new metadata and information from the feedback records of different reanalyses. This would facilitate studies based on observations alone, and should help to strengthen the ties between observationally based research and research based on reanalyses.

b. Emerging Capabilities

The last few years have seen the development of a number of promising new techniques for identifying and correcting observational bias and other errors. In particular, bias correction techniques have been developed for radiance data. The procedure, known as radiance bias correction, in fact has a long history in the quantitative use of satellite data in both retrievals and numerical weather prediction assimilation schemes. Radiance bias correction allows for correction of unresolved errors in a number of areas of the satellite sounding and inversion. These sources include forward radiative transfer errors, in situ validation instrument errors, errors in instrument end-to-end calibration, and errors in the numerical weather prediction models base climate state. Reduction of each of these sources of error is an ongoing effort by many research groups.

There are also extensive efforts aimed at the calibration of satellite instruments. The calibration and inter-calibration of radiance data for climate studies can be considered as a three-step process. The first step, referred to as the nominal calibration, involves using the best procedures to provide for the calibration of a single instrument on a single satellite. This step is also common to reanalysis for they also require the best absolute calibration of each instrument on each satellite. The second step is referred to as the normalized calibration. The normalization is usually accomplished by identifying one instrument as the baseline and then adjusting the systematic biases of the other instruments in the series to that baseline instrument. A final step involves adjusting the normalized calibration to some 'absolute' geophysical observation of the variable that is being retrieved, and is hence referred to as the absolute calibration.

There is also work being done to better take advantage of reanalysis feedback files. These files contain detailed information about the difference between the model first guess background field and each observational data set. Conceptually, it should be possible to use this information to improve the quality control of the observational data sets in an iterative manner. In reality, it is not that simple because the first guess background fields are not observations but a blend of model, and it's assumptions and simplifications of atmospheric processes, and observations. Nevertheless, there is a significant potential for using the reanalysis feedback files, for example, when observations and first guess background fields systematically differ over extended regions or time there is reason to investigate the observational data sources for possible errors. Gleaning and applying this additional quality control information is currently under utilized and is a new area of effort for the data stewards.

c. The Future

In order to optimize the use of observational data for both reprocessing and reanalysis, we are promoting the adoption of the concept of scientific data stewardship. Scientific data stewardship (SDS) is the new paradigm in data management consisting of an integrated suite of functions to preserve and exploit the full scientific value of environmental data. These functions are the careful monitoring of observing system performance for long-term applications, the generation of authoritative long-term records, for both reprocessing and reanalysis, from multiple observing platforms, the assessment of the state of the atmospheric, oceanic, land, cryospheric and space environments, and the proper archival of and timely access to data and metadata.

To promote full exploitation of the scientific value by current and future users, four functions, each with several constituent components, must be achieved. The first function is to provide real-time monitoring of the observing system performance for long-term applications. Such monitoring requires the establishment of tracking tools necessary for the detection of changes in the observing system as well as in the observation record. One example is the detection of small biases in the instrumental record. These biases can then be minimized or eliminated through efficient coordination with network operators.

The second function is generating authoritative, long-term records. This function will preserve and enhance the value of the irreplaceable historical data by conducting rigorous data analysis and research to validate and improve these authoritative records, and by reanalysis and reprocessing and enabling others to participate in these. For reanalysis, the primary techniques involve 3- and 4-dimensional variational analysis using a model to fuse together data from disparate observing systems like direct measurements from ground-based networks and indirect measurements from remote sensing instruments on satellites.

The third function uses the authoritative records to assess the current state of the environment and to put it in historical perspective. Long-term trends on local, regional or global scales can be determined and estimated for the future. In addition the authoritative records can be used to detect changes in environmental conditions between different time periods and different environmental regimes. The reanalysis framework provides a comprehensive framework for assessing the impact of different observing systems on the end product. At one end of the spectrum of reanalysis approaches all the available observations are used to obtain the best estimate of the earth system at all times. This type of reanalysis can be highly useful for studies of the complex interactions between, for example, the biological and geophysical processes. At the other end of the reanalysis spectrum, use of a consistent set of observations for different observing system epochs has led to a strategy of differing analysis time periods (see e.g., Arkin 2004).

The final function, insuring complete archival and access capabilities, requires that metadata, direct observations, and fundamental records from satellite and in situ platforms be comprehensive, complete and preserved, in perpetuity. Open, efficient access to the metadata, products, and data streams must be insured, and data made available in useful formats. The metadatabases become particularly important as the archives are cyclically refined. It is here that improved understanding of observations is captured, e.g. physics of the instrument design, basic and reanalysis-determined quality control, and lineage of use.

4. Recommendations/Action Items

The following presents the main results of the workshop and outlines a set of recommendations and associated action items that were deemed most pressing to facilitate our vision for the future. It was recognized that the timely development and shepherding of such a program requires strong leadership by individuals experienced with historical as well as real-time conventional and satellite data, and that the appointment of an individual or individuals to take on this responsibility be made as soon as possible at the international level. While determining specific resource requirements were beyond the scope of this workshop, the resources and funding of the program must allow for timely data refresh cycles commensurate with the needs of an on-going iterative process of reanalysis coordinated between the major centers. *As such, the key programmatic recommendation of the workshop is for the*

WCRP Observations and Assimilation Panel (WOAP) to appoint a working group of experts charged with developing a plan for "The On-going Development of Improved Observational Data Sets for Reanalysis", that describes the necessary resources, infrastructure, institutional commitments, and coordination on technical issues outlined in this report.

Action: Get commitments from member countries to carry out specific tasks as outlined below in the scientific and technical recommendations and action items.

The workshop also identified the need to raise the profile of activities concerning the pre, post, and bias processing of observations and the related science. It was recognized that the improved understanding of climate variability and change depends more than ever on the quality of the historical observations and their effective use in reanalyses.

The meeting highlighted that recovery and updating of observational data is an ongoing effort and reanalyses themselves are part of this activity. A new reanalysis project needs to get the latest versions of data from various sources, and a single operational archive on its own does not provide enough information. Keeping track of the global observational dataset is crucial for the success of future reanalysis efforts. This task is increasing in complexity with many more satellite sources being available in recent years and through rescue and improvement of conventional data.

There was general agreement that one of the most pressing needs was for the various data centers to conduct a full international coordinated inventory of their conventional and satellite data holdings. The scope of the inventory would be global and include all observations relevant to reanalysis of the Earth System encompassing observations of the atmosphere, ocean, land, and cryosphere.

The specific technical and scientific recommendations, associated action items, and summaries of related discussion and information are as follows:

Recommendation 1: All centers should prepare inventories of reanalysis observations (on the level of observation records)

Action 1a: Compare NCAR and NCDC inventories for surface and upper air.
Joey Comeaux will draft inventory record structure, will iterate with NCDC to reach mutual agreement. Russ Voss is the contact at NCDC

Action 1b: Draft inventory record structure for moving platforms
Consult with ICOADS and WOB experts to assure compatibility
Responsible individuals - Steven Worley and Joey Comeaux

It was suggested to define data families that group similar qualities of observational data sets (e.g., land surface stations, ocean surface stations, upper air in situ observations, aircraft observations, satellite observations, etc.). For each observational data set family, compile an inventory of all data sets that will serve as a 'master list' of all data holdings available for that family. One needs to agree on a naming convention and catalog listing for each family. The reanalysis centers need the catalog list for all families, while the data centers may have individual experts for each family.

It was deemed important to identify data families and contacts for each family at the reanalysis centers and data centers. For each family conduct a comprehensive inventory compilation. Lead on each data family is (are): radiosondes (Thorne and Haimberger, with Durre and Comeaux); land surface (Comeaux and Voss); ocean surface and subsurface (Worley (ICOADS) and Boyer (WOD)); Aircraft (Comeaux and Voss); satellite (Bates). Maybe focus effort on the time periods that are most important, likely not the most modern period.

- Action 1c: Following Actions 1a and 1b, get inventories in established forms from:
  JMA, ECMWF, NCEP, NASA, BoM for data not from NCDC, NCAR, or in ICOADS and WOD. Contacts, JMA – Kasutoshi Onogi, ECMWF – Sakari Uppala, NCEP – Jack Woollen, NASA – Siegfried Schubert, BMRC – Peter Steinle, CRU East Anglia Univ – Phil Jones

The following provides information and issues regarding the main input data sets to be included in the inventory.

Conventional data:

Data types : surface, UA, aircraft, cloud drift winds from satellite, marine surface, and ocean profiles
Inventories by: dataset, station (carry location and ID), and monthly (for stationary observations).
UA (raobs and pibals – separately)
Moving platforms, a scheme is TBD.
Document past history, naming convention, QC/QA applied
May need a cross-organization glossary to match datasets
Separate inventory on undigitized and unavailable (at a major data provision center) data sources
Seek wider international participation through WMO. Contacts for, British Antarctic Survey, Met Office, Meteo France, DWD, LMD (France)
Specialized sondes exist (ozone, moisture, ground based microwave, etc), the centers holding these should be identified. These are good validation datasets. Ozone has been reanalyzed in ERA-40 using satellite data. Rocketsonde data has been used in the validation of stratospheric wind

analyses and should be made available to the reanalysis centers.
Other validation data - soil moisture datasets

Satellite data

What form should the inventory take?
Version,time range, gaps, granule size (orbit)?
Separate polar and geostationary satellites
Ideas and volunteers (John Bates in leading role)

NASA, NCDC, EUMETSAT, MSC (Japan), request contact from MSC for JAXA
Base first cut on what was used by ERA-40
Framework starting point may be orbit granularity
May not be easy to match inventory requirement because of very large extant archives
Logical start point is about 1978, but going back to cover the early 1970's (VTPR), or earlier would benefit research of the 1970's climate shift , NCDC could possibly make the VTPR data a higher priority
Above list of agencies is incomplete – need DMSP and many others
To get broad recognition could consider forming groups that could be designated as task groups under the WCRP/WOAP
Ozone satellite products (total column and layer profiles) are going through reprocessing cycles and inventories are needed
Need inventories of data jumps, problems, and known resolutions online and easily accessible
NOAA Polar Orbiter data guide is being expanded and put on line – this is part of the solution. Work at NCDC.
SIRS data should be examine for possible usage

Marine surface and subsurface

ICOADS and WOD are merged data source collections

ICOADS began 1983, first data edition 1987, now at Release 2.2 (October 2005) with global data covering 1794-2004 (Worley et. al., 2005).

WOD at NODC/OCL began about 1990, first edition WOD94, and WOD05 will be available approximately April 2006 with global data beginning in about 1900

Both collections have data from many sources and instrument types. Many data are yet to be rescued (digitized, saved from degrading magnetic media, transferred from remote collections), OCL and ICOADS are pursuing many options under very limited or decreasing budgets. More support would accelerate and sustain improvements. More contributions should be sought to fill data voids and extend the records. Overlaps, e.g. PMEL, coming from many sources can be a problem.

Snow

- NSIDC – station and satellite holdings, need to document what they have
- Early snow depth data from Japan?
- Review CLIC program for sources
- National archives in Scandinavia
- Snow depth and snow coverage are important
- Simultaneous 2m temperature useful
- Recommendation: Gain a better understanding of what snow coverage and snow depth are available.
- Sakari and Steven to ask NSIDC (Roger Barry, Ron Weaver) to provide a NSIDC inventory
- Design plan to find more sources available in international holdings

Land surface data

- Soil moisture diagnostic requirements
- Need good 2m temperature and near surface atmospheric moisture
- Soil moisture – direct analysis

–New satellite data are still in Beta phase

- Continued research on SMMR radiance as a proxy for soil moisture is recommended
- 9 year record from SMMR
- Current sensors AMSR-E and TRMM, future sensors SMOS, Aquarius, may add new information
- GRACE gravity-> terrestrial water storage estimates are very large temporal and spatial scale and maybe not appropriate for reanalysis - only a two year record
- River runoff – probably for validation only

–Availability USGS and GRDC (http://grdc.bafg.de)

SST

- ERA-40 used 2DVAR from NCEP for Nov. 1981 and after, and the Hadley Center HadISST for earlier
- JRA-25 uses the COBE Analysis (Ishii 2005)
- Cautionary note: Carry out sensitivity studies for various time periods to determine appropriate SST fields
- Increased resolution over time can affect the reanalysis model. How sensitive is the model to improving resolution? Sensitivity tests should be run.
- Bias corrected pre-1940 is important. There are bias throughout the record (e.g. more buoys than ships in modern time)
- Time resolution and space resolution are both considerations
- Ice limits are crucial at the SST boundaries, older data is very smooth and derived from low resolution estimates and modern era is well defined by satellites
- Prior to 1970 SH is based on climatology - can improvements be made here? Consider using new daily SST analysis from NCDC, 11/1981
- What is more important, uncertainty estimates on SST fields or higher resolution SST, if limited resources are available?

•Some leaning toward uncertainty, issue needs further discussion and has implications for resource deployment
•SST analyses are currently in a phase of rapid development (higher resolution via using multiple satellites). Reanalysis teams should contact the SST experts.
•Nick Rayner may convene a sea ice working group at the up coming MARCDAT-II meeting (Oct. 05, Exeter). It is suggested that the sea ice specification is currently the weakest element of SST analyses.

Sea ice

•Ice coverage for pre-satellite period, get latest opinions and work from John Walsh
•Need reconstructed ice limits in the SH – pre-satellite
•Reference ice development work in Rayner's paper, combined many sources to create long-term ice edge. Maybe, new information at the Oct. 2005 MARCDAT-II meeting
•COBE, Ishii 2005, - what sea-ice was used?
•GLAS satellite has potential to provide sea ice thickness

Data types not addressed

•Precipitation - Reanalysis projects need access to global daily precipitation data in order to validate the performance of hydrological cycle. Monthly precipitation analyses (e. GPCP) are valuable, but have errors of their own.
•Independent satellite based cloud analyses needed for validation

Other Issues:
–Could a very high-resolution land surface elevation dataset be used to verify, confirm, or supply better station elevation?
–A true reference dataset doesn't exist, the best we have are carefully adjusted monthly mean anomaly data
–Are there opportunities to develop reference sondes with standard procedures?
–GUAN (Global Upper Air Network) key stations with long-term records
RATPAC and UK products (monthly means) – useful to compare reanalysis statistics
–Sonde anomalies are largely related instrument changes
–Homogeneous time series are most available from US locations
–Need better instrument identification to go with the UA stations, allowing for basic bias adjustment
–Archives need documented QC procedures

Recommendation 2: A collaboration be formed that can sustain a data refresh cycle and create high quality merged datasets for reanalysis

•<u>Definitions and Objectives</u>
–Data joining = joining observation datasets from different archive sources
-Simplest way to merge data, basically just duplicate elimination
–Data combining = joining records in the same observations from different sources

-More complicated, involves combining observed components between duplicates
-Clearly results in maximum quantity of observed data produced

–Data merging = "data joining" + "data combining"

–Why should we design and develop ways to merge data?
-Global historical archives are fragmented in many different pieces
-Archives of the same observation record may have differences because:
-Difference in raw receipts creates different observations
-Decoder difference creates observation differences
-Different post decoder processing create observation differences
-Best ways to merge various observation types has not been definitively evaluated

–Communication between merging collaborators is key, including
-Access to source dataset components and resultant merged datasets
-Access to documentation, provenance, feedback, and other qualitative information
-Collaborative efforts to assess dataset utility and improve quality and quantity
-Collaboration of effort is consolidation of available resources

• Benefits

–Reanalysis projects will have access to established starting data
–High quality merged datasets will be created, tested, and improved with use
–Documentation of the process development is itself an important outcome
–Valuable byproduct is a resource for observation based climate studies

• Challenges

–A multitude of decisions are necessary during data translations and merging. How can these necessary decisions be made *non-destructively* with regard to the needs of various different re-analysis systems and procedures, now and in the future?
–When will collaborative efforts to merge observations by consensus actually result in real savings in observation preparation for reanalysis projects?
–The degree to which the data should be screened is not a static requirement. Differences in assimilation systems to handle noisy data vary now and will vary differently in the future.
–A complex system allowing for frequent improvements in source archives to be easily incorporated in the merged datasets must be designed and maintained.
–All steps in merging must be documented and reproducible.

Action 2a: Begin to collaboratively explore merging/combining activities with radiosonde data because of its vertical data structure and potentially large impact on reanalysis.

Action 2b: Add ancillary data (reanalysis feedback files – add specific fields; other expert calibration and inter-calibration information) to merged data set database.

–How should feed back data from previous reanalyses be carried forward into the merged datasets? What elements are important? How will the information be used? (see recommendation 3)

–Radiosondes –Thorne; Haimberger – by date – to contact Imke Durre regarding the inclusion of ancillary data into the IGRA database (e.g., some fields from ERA-40 feedback files, estimated bias corrections from other investigators). NCAR has the land surface and upper air ERA-40 feedback files online. The files are stored in an ASCII format and a sub-setting user request form is available. UA (http://dss.ucar.edu/datasets/ds366.0/), surface (http://dss.ucar.edu/datasets/ds476.0/
–Land (GCOS-TOPC; ISLSCP; ISH) and ocean surface (ICOADS) –GWEBS, vegitation index, ICOADS has included the WMO ship metadata (e.g. ship length, anemometer height) on each applicable record for 1973-2004.
–Aircraft – T, RH
–Satellite - ISCCP B1U data, reprocessed motion vectors, inventory/document NOAA cloud drift wind data
–What is available from the ocean reanalyses that might be useful?

Action 2c: Further inter-calibrate various measurements across successive generations of satellite sensors, with priority given to SSU data

Recommendation 3: Develop improved record tracking control for observations to further improve the use of feedback data from reanalyses targeted especially for data providers/developers:

Action 3a: Provide additional details on metadata efforts and standards for in situ and satellite data, including a pilot 'rich metadata standard for satellite data, by spring 2006 (Bates)

Action 3b: Establish working group to determine standard set of record header information that satisfies recording track through the reanalysis system.

–GMAO rep. tbd
–NCAR Joey Comeaux – draft meta dataset to be initiated at NCAR, circulate to the WG
–NCDC Russ Vose
–NCEP Jack Woollen
–JMA contact through Kasutoshi Onogi
–ECMWF Sakari Uppala

Once metadata are compiled, the next question is, given this metadata, how do we share with the community the quality control procedures that different groups have found useful over time? There are currently no comprehensive tools or procedures to do this. Web-based tools, such as web forum pages, may be a way to conduct community-based sharing of quality control methods. For a start, just compiling what is currently being done at data centers and numerical weather prediction/reanalysis centers would be of value.

Action 3c: Provide a list of web sites relevant to data quality control for reanalysis and real-time NWP analysis. Post at the WMO WOAP web site. (Glenn White, NCEP).

Meta data information

Types of metadata must be defined and hierarchies developed for what you do with those metadata. A useful framework for defining metadata and its functions within the context of digital archives may be found in the Open Archive Information System Reference Model (OAIS – RM adopted as ISO 14721 in 2003 available from http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html). Metadata, or described here as preservation description information, has standard information areas which then are defined more specifically for each information type. The OAIS RM defines the following information types for metadata:

- Reference Information: This information identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information. It also provides those identifiers that allow outside systems to refer, unambiguously, to this particular Content Information. Examples of these systems include taxonomic systems, reference systems and registration systems. In the OAIS Reference Model most if not all of this information is replicated in Package Descriptions, which enable Consumers to access Content Information of interest.

- Context Information: This information documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects existing elsewhere.

- Provenance Information: This information documents the history of the Content Information. This tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. This gives future users some assurance as to the likely reliability of the Content Information. Provenance can be viewed as a special type of context information.

- Fixity Information: This information provides the Data Integrity checks or Validation/Verification keys used to ensure that the particular Content Information object has not been altered in an undocumented manner. Fixity Information includes special encoding and error detection schemes that are specific to instances of Content Objects. Fixity Information does not include the integrity preserving mechanisms provided by the OAIS underlying services, error protection supplied by the media and device drivers used by Archival Storage. The Fixity Information may specify minimum quality of service requirements for these mechanisms.

Reference information metadata is also commonly referred to as catalog metadata (or also as collection metadata) in reference to the old library card catalog indexing of bibliographic information. To perform the inventory function discussed above, this catalog metadata must be agreed to and this is then what is cross checked at the reanalysis and data centers. Common standards for the collection and indexing of geospatial information is specified in the Federal Geographic Data Committee and in the new ISO metadata standard (ISO 19115 of 2003). These standards for catalog metadata also include discipline specific extensions, such as the remote sensing extensions for remote sensing within the FGDC. Whenever possible, data centers and reanalysis centers should use these standards.

Efforts to provide improved metadata have recently expanded. One such effort is the Metadata Integration and Improvement Initiative (MI3) Station History System, developed and implemented as a central repository of station information for NOAA and non-NOAA observing systems at NCDC. With a web-based user interface accessing a relational database, MI3 provides users immediate, flexible online access to rich station details for tens of thousands of current and historical observing stations. Similar efforts are in the pilot stage for satellite metadata.

Recommendation 4: The observational, reanalysis, and climate communities should take a coordinated approach to further optimizing reanalysis for climate.

Action 4a: Set up working group (Schubert, Dee, Whittaker, Thorne, etc.) to focus on the following reanalysis issues:

- All data vs. stable observing system including reduced resolution data
- What are the highest impact set of observations (priority)?
- Developing and testing adaptive bias-correction techniques for reanalysis
- Enhance community participation to interact and analyze reanalysis
- What can reanalysis do toward recommending sampling distribution and supporting the need for a “reference” upper air network?.
- How can new observations best be tested/evaluated in reanalysis/OSEs?
- Establishment of ongoing relations with major data center

5. Agenda

The presentations are available at:

http://www.infonetic.com/tis_conferences/dio/

Click on agenda on the left panel.

Wednesday, September 28

8:20 am: Siegfried Schubert – Introduction and Overview

8:30am: Kevin Trenberth – “WCRP perspective on reanalysis and reprocessing"

8:50am: Sakari Uppala - "ECMWF experience and the role of observations in future reanalyses"

9:10am: Kazutoshi Onogi - "JRA-25: progress and observations overview"

9:30pm Jack Woollen – “NCEP experience and data holdings for reanalysis”

9:50am break

10:20am Joey Comeaux - “Reanalysis Efforts at NCAR - Past, Present and Future”

10:40am: John Bates - "The status of historical satellite radiance archive" (including VTPR, TOVS, SSMI, ATOVS, SSU, AMSU, Geostationary) and methods for satellite to satellite calibration"

11:00am Fuzhong Weng "Community radiative transfer model developed for satellite data assimilation"

11:20am: Russ Vose - "The status of historical upper air data and conventional land surface observations"

11:40am: Tom Smith - "Prospects for improving SST/ ICE analyses”

12 :00pm: Dick Dee –"Adaptive bias correction techniques”

12:20 lunch

1:20 Siegfried Schubert - Charge to working groups, review/revise agenda if needed

Begin breakout sessions (divide into three working groups- see below)

Thursday, September 29 (8:30-5:30pm), break 10:00am, lunch noon, break 3:00pm

8:30am: begin meeting in plenary session

-talk by Tsengdar Lee (NASA hdqrtrs),

-discussion of any cross-cutting issues

e.g., creating merged super collections for reanalysis input

a. QA/QC, bias adjustment considerations
b. Duplicate elimination – choosing the best observation?
c. Creating a process in which data collection improvements can be easily be made available for the next reanalysis

-coordinate/synchronize/reshape groups as needed

- continue breakout sessions

Friday, September 30 (8:30-noon), break 10:00am

Working group chairs present reports in plenary session, propose collaborative projects

6. Attendees

	Last	First	Company
1	Arkin	Phil	ESSIC, University of Maryland
2	Ballish	Bradley	NCEP / NCO / PMB
3	Barkstrom	Bruce	ASDC, NASA Langley Research Center
4	Bates	John	NOAA/NCDC
5	Bosilovich	Michael	NASA/GMAO
6	Boyer	Tim	US NODC
7	Bromwich	David	Byrd Polar Research Center,Ohio State University
8	Brubaker	Nicole	SAIC/GSFC
9	Cahalan	Robert	NASA/GSFC
10	Carton	James	University of Maryland
11	Chang	Yenui	NASA GSFC
12	Chen	Jinnye	NASA/GMAO
13	Comeaux	Joseph	NCAR
14	Dattore	Robert	NCAR
15	Dee	Dick	GMAO
16	Ebisuzaki	Wesley	CPC/NCEP
17	Fiorino	Mike	LLNL
18	Garcia	Hernan E.	NOAA / NODC, Ocean Climate Laboratory
19	Gelaro	Ron	NASA/GMAO
20	Gelman	Mel	NOAA / CPC
21	Haimberger	Leopold	University of Vienna
22	Herdies	Dirceu	NOAA/NCEP
23	Higgins	Wayne	NOAA / CPC
24	Ji	Ming	NOAA
25	Kalnay	Eugenia	UMD
26	Kaplan	Alexey	LDEO of Columbia University
27	Kim	Gi-Kong	NASA Goddard Space Flight Center
28	Kistler	Bob	NCEP/EMC
29	Kumar	Krishna	NCEP Central Operations
30	Le Marshall	John	JCSDA
31	Lee	Tsengdar	NASA Headquarters
32	Legler	David	U.S. CLIVAR Office
33	Liu	Emily	GMAO/GSFC
34	Lucchesi	Robert	NASA/GMAO
35	Miller	Christopher	NOAA
36	Mo	Kingtse	Climate Prediction Center
37	Onogi	Kazutoshi	Japan Meteorological Agency
38	Pawson	Steven	NASA GSFC
39	Pegion	Philip	NASA GSFC/SAIC
40	Redder	Christopher	NASA GSFC
41	Reichle	Rolf	NASA Goddard Space Flight Center, GMAO
42	Salstein	David	AER/NASA/UMBC
43	Schmid	Claudia	NOAA/AOML
44	Schubert	Siegfried	NASA Goddard Space Flight Center
45	Seidel	Dian	NOAA Air Resources Lab
46	Shi	Wei	NOAA/NWS/NCEP/Climate Prediction Center
47	Sienkiewicz	Meta	NASA Goddard Space Flight Center,SAIC
48	Smith	Thomas	NOAA/NCDC & CICS/ESSIC
49	Stokes	Diane	NOAA/NWS/NCEP/EMC
50	Suarez	Max	GMAO

51	Thorne	Peter	Hadley Centre, Met Office
52	Trenberth	Kevin	NCAR
53	Uppala	Sakari	Head of Reanalysis Section
54	Verter	Frances	NASA GSFC
55	Vose	Russ	NCDC
56	Weng	Fuzhong	NOAA/NESDIS/Office of Research&Applications
57	Whitaker	Jeff	NOAA Climate Diagnostics Center
58	White	Glenn	GCWMB/EMC/NCEP/NWS/NOAA
59	Woollen	Jack	NOAA/NCEP
60	Worley	Steven	NCAR
61	Wu	Man-Li	NASA Goddard Space Flight Center,GMAO
62	Yang	Runhua	GMAO
63	Zhang	Banglin	GSFC GMAO/SAIC
64	Zhou	Jiayu	NOAA/NWS/OST
65	Zhu	Yanqiu	NASA GSFC/SAIC

Appendix – Examples of popularity of reanalyses

A. Citations – NCEP/NCAR reanalyses

Dear Dr. Kalnay:

I am writing to inform you that your paper, "The NCEP/NACR 40-year reanalysis project,"
(Bull. Amer. Meteorol. Soc. 77[3]: 437-71, March 1996), has been selected as a highly cited paper in the field of Geosciences by ISI Essential Science Indicators (ESI).

ESI is a Web-based compilation of science indicators and trend data derived from ISI's databases, focusing on highly cited papers, authors, organizations, journals, and nations. It combines these data with editorial content to highlight important results. In our latest analysis, your paper was among the 10 most-cited articles identified in Geosciences over the past decade. Being highly cited generally reflects the high regard in which your work is held by your fellow scientists, and its value to the scientific community as a whole.

.

.

.

Sincerely,
Jennifer L. Minnick
Editorial Coordinator
ISI Essential Science Indicators

B. ERA-40 at NCAR: Data User Summary for 2005

NCAR/CISL/DSS continued to place emphasis on providing ERA-40 data to NCAR users and researchers throughout the United States in 2005. Improved access and new products resulted in a two-fold increase in the amount of data served in 2005 relative to 2004.

The new products are shown in Table 1. They are 6-hourly and monthly mean collections at T85 and T106 resolutions, and Feedback observations. The users ability to get the data they need has also been enhanced by creating web-based request interfaces and subsetting software for some products.

Table 1, New ERA-40 data products created at NCAR

Product	**UA levels**	**Surface and Single levels**	**GRIB**	**netCDF**
Monthly				
256x128 Gaussian (T85)	23.Plvl, 60.Mlvl	yes	yes	yes
320x160 Gaussian (T106)	23.Plvl		yes	
6-hourly				
256x128 Gaussian (T85)	23.Plvl, 60.Mlvl	yes	yes	yes
320x160 Gaussian (T106)	23.Plvl, 60.Mlvl	yes	yes	
Feedback observations				
In situ locations, UA and surface archives, in ASCII format				

At NCAR users access the ERA-40 archive through a network-connected server (Web and FTP) or directly from the Mass Storage System (MSS). Table 2 provides statistics for number of unique users, number of data files, and total amount of data delivered for these modes of access.

Table 2, ERA-40 User Access Metrics for 2005

	Web & FTP	**NCAR MSS**	**Total**
Number of Unique Users	70	80	147
Number of Data Files	32786	50324	83110
Data Amount (GB)	9610	25913	35523

ERA-40 is a significant research dataset and we are pleased to continue our cooperation with ECMWF on its curation and stewardship. Plans for 2006 include improved data discovery and access options (spatial subsetting) and increased real-time access through NCAR Community Data Portal.

References

Arkin, P., ed., 2004: *Ongoing Analysis of the Climate System: A Workshop Report*. University Corporation for Atmospheric Research, PO Box 3000, Boulder, CO 80307, 48 pp. Available at www.joss.ucar.edu/joss_psg/meetings/climatesystem/.

Gibson, J.K., P. Kallberg, A. Nomura, and S. Uppala, 1994: The ECMWF re-analysis (ERA) project- Plans and current status. Preprints, 10th Int. Conf. On Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, Nashville, TN, American Meteorological Society, 164-167.

Kalnay,E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetma, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NMC/NCAR 40-year Reanalysis Project. Bull.Amer.Meteor.Soc., 77,437-471.

Kanamitsu, M, W. Ebisuzaki, J. Woollen, S-K Yang, J.J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2):. Bull. Amer. Met. Soc., 83, 1631-1643.

Schubert, S.D., R. B. Rood and J. Pfaendtner, 1993: An Assimilated Data Set for Earth Science Applications *Bull. Amer. Meteor. Soc.*, 74, 2331-2342.

Schubert, S. D. and G. White 2004: Working Group Report on Implementation Plans for the CCSP Deliverable: "Re-analyses of historical climate data for key atmospheric features. Implications for attribution of causes of observed change. Available at www.joss.ucar.edu/joss_psg/meetings/climatesystem/.

Simmons, A.J. and J.K. Gibson, 2000: The ERA-40 project plan. The ERA-40 Project Report Series No. 1, European Centre for Medium-Range Weather Forecasts.

Trenberth, K.E., 1995: Atmospheric circulation climate changes. Climate Change, 31, 427-453.

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J. 2005: The ERA-40 Reanalysis. Quart. J. Roy. Meteorol. Soc., in press.

WCRP, 1998: Proceedings of the First WCRP International Conference on Reanalyses. Silver Spring, Maryland, 27-31 October 1997, WCRP-104, WMO/TD-No. 876.

WCRP, 2000: Proceedings of the Second WCRP International Conference on Reanalyses. Wokefield Park, Reading, UK, 23-27 August 1999. WCRP-109, WMO/TD-No. 985.

Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, and N. Lott, 2005: ICOADS Release 2.1 data and products. *Int. J. Climatol. (CLIMAR-II Special Issue)*, **25**, 823-842 (DOI: 10.1002/joc.1166).

Acronyms

2DVAR	Two-Dimensional Variational Analysis
AER	Atmospheric and Environmental Research, Inc.
AMIP-II	Atmospheric Model Intercomparison Project- II
AMSR-E	Advanced Microwave Scanning Radiometer - EOS
AMSU	Advanced Microwave Sounding Unit
AOML	Atlantic Oceanographic and Meteorological Laboratory
ASDC	Atmospheric Science Data Center at NASA Langley
ATOVS	Advanced Television and Infrared Observation Satellite Operational Vertical Sounder
B1U	B1 Uniform Data
BMRC	Bureau of Meteorology Research Center, Australia
BoM	Bureau of Meteorology, Australia
CCSP	Climate Change Science Program
CICS	Cooperative Institute for Climate Studies
CISL	Computational & Information Systems Laboratory
CLIC	Climate and Cryosphere Project
COBE	Centennial comprehensive marine dataset produced by JMA
CPC	Climate Prediction Center
CRU	Climate Research Unit (Brit met office)
DAO	former Data Assimilation Office (NASA)
DMSP	Defense Meteorological Satellite Program
DOE	Department of Energy
DSS	CISL SCD (Scientific Computing Division) Data Support Section
DWD	Deutscher Wetterdienst **(**Germany’s National Meteorological Service)
ECMWF	European Center for Medium-Range Weather Forecasts
EMC	Environmental Modeling Center NOAA
ENSO	El Niño–Southern Oscillation
ERA-15	first generation 15-year European Reanalysis
ERA-40	40-year European Reanalysis
ESI	Essential Science Indicators
ESSIC	Earth System Science Interdisciplinary Center
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FGDC	Federal Geographic Data Committee
FTP	File Transfer Protocol
GCOS	Global Climate Observing System
GCOS-TOPC	GCOS Terrestrial Observation Panel for Climate

GCWMB	Global Climate and Weather Modeling Branch (NWS)
GLAS	Geoscience Laser Altimeter System
GPCP	Global Precipitation Climatology Project
GMAO	Global Modeling and Assimilation Office (NASA)
GRACE	Gravity Recovery and Climate Experiment
GRDC	Global Runoff Data Centre (http://grdc.bafg.de)
GRIB	GRIdded Binary (bit-oriented data exchange format)
GSFC	Goddard Space Flight Center
GUAN	GCOS Upper-Air Network
GWEBS	Global Water and Energy Budget Study
HadISST	Hadley Centre Global Sea Ice and Sea Surface Temperature dataset
ICOADS	International Comprehensive Ocean Atmosphere Data Set
IGRA	Integrated Global Radiosonde Archive
ISLSCP	International Satellite Land-Surface Climatology Project
ISCCP B1U	International Satellite Cloud Climatology Project B1 Uniform Data
ISH	Integrated Surface Hourly (Land Data)
ISI	Institute for Scientific Information
ISO	International Organization for Standardization
JAXA	Japan Aerospace Exploration Agency
JCSDA	Joint Center for Satellite Data Assimilation
JMA	Japan Meteorological Agency
JRA-25	Japanese 25-year Reanalysis Project
LDEO	Lamont-Doherty Earth Observatory of Columbia University
LLNL	Lawrence Livermore National Laboratory
LMD	Laboratoire de Météorologie Dynamique (France)
MARCDAT-II	2nd International Workshop on Advances in the Use of Historical Marine Climate Data
MERRA	NASA’s Modern-Era Reanalysis for Research and Applications
MI3	The Metadata Integration and Improvement Initiative Station Information Management System
MSC	Meteorological Satellite Center (Japan)
MSS	Mass Storage System
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center, NOAA
NCEP	National Centers for Environmental Prediction
NCO	NCEP Central Operations
NESDIS	National Environmental Satellite, Data, and Information Service (NOAA)
netCDF	network Common Data Form
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration (USA)
NODC	National Oceanographic Data Center
NODC/OCL	NODC/Ocean Climate Laboratory
NSF	National Science Foundation
NSIDC	The National Snow and Ice Data Center
NWP	Numerical Weather Prediction

NWS	National Weather Service (NOAA)
PMB	Production Management Branch
PMEL	Pacific Marine Environmental Laboratory
OAIS	Open Archival Information System (Standard)
OAIS-RM	Open Archival Information System - Reference Model
OCL	Ocean Climate Laboratory
OSE	Observing System Experiment
OST	Office of Science and Technology (NWS)
QA	Quality Assurance
QC	Quality Control
RATPAC	Radiosonde Atmospheric Temperature Products for Assessing Climate
SAIC	Science Applications International Corporation
SDS	Scientific Data Stewardship
SH	Specific Humidity
SIRS	Satellite Infrared Spectrometer
SMMR	Scanning Multichannel Microwave Radiometer
SMOS	Soil Moisture and Ocean Salinity
SSMI	Special Sensor Microwave Imager
SST	Sea Surface Temperature
SST/ICE	Sea Surface Temperature and Ice Information
SSU	Stratospheric Sounding Unit (On NOAA polar orbiting satellite)
T/RH	Temperature and Relative Humidity
TIROS	Television and Infrared Observation Satellite program
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
UA Station	Upper Air Station
UK	United Kingdom
UMBC	University of Maryland, Baltimore County
UMD	University of Maryland
US	Upper Stratosphere
USA	United States of America
USGS	United States Geological Survey
VTPR	Vertical Temperature Profile Radiometer
WCRP	World Climate Research Programme
WG	Working Group
WMO	World Meteorological Organization
WOAP	WCRP Observation and Analysis Panel
WOB	Weather Research and Forecasting (WRF) Model Oversight Board
WOD	World Ocean Database
WOD05	World Ocean Database 2005
WOD94	World Ocean Database 1994

Acknowledgements

This workshop was supported by NASA’s Science Mission Directorate, NOAA’s Climate Prediction Center, and NSF’s Division of Atmospheric Sciences. Logistics support was provided by Kathy Regul of INFONETIC.