Chapter 2. Re-Analysis of Historical Climate Data for Key Atmospheric Features

Convening Lead Author: Siegfried Schubert, NASA

Lead Authors: Phil Arkin, University of Maryland; James Carton, University of Maryland; Eugenia Kalnay, University of Maryland; Randal Koster, NASA

Contributing Authors: Randall Dole, NOAA**;** Roger Pulwarty, NOAA

KEY FINDINGS

- Reanalysis plays a crucial integrating role within a global climate observing system by producing comprehensive long-term, objective, and consistent records of climate system components, including the atmosphere, oceans, and land surface (Section 2.1).
- Reanalysis data play a fundamental and unique role in studies that address the nature, causes and impacts of global-scale and regional-scale climate phenomena (Section 2.3).
- Reanalysis data sets are of particular value in studies of the physical mechanisms that produce high-impact climate anomalies such as droughts and floods, as well as other key atmospheric features that affect the United States, including climate variations associated with El Niño-Southern Oscillation and other major modes of climate variability (Section 2.3).

- Observed global and regional surface temperature trends are captured to first order in reanalysis data sets, particularly since the late 1970s, although some regions continue to show major differences with observations (*e.g.*, Australia). Reanalysis precipitation trends are much less consistent with those calculated from observational datasets, probably due to deficiencies in current global reanalysis models (Section 2.4).
- While current reanalysis data have proven to be extremely valuable for a host of climate applications, it is important to understand that the overall quality of reanalysis products varies with latitude, height, time period, spatial and temporal scale, and quantity or variable of interest (Sections 2.1, 2.2, 2.3, and 2.4).
- Current global reanalysis data are most reliable in Northern Hemisphere middle latitudes, in the middle to upper troposphere, and on synoptic (weather) and larger spatial scales. They are least reliable near the surface, in the stratosphere, tropics, and polar regions (Sections 2.2, 2.3, and 2.4).
- Current global reanalysis data are most reliable on daily to interannual time scales. They are least reliable in the representation of the diurnal cycle and in the representation of decadal and longer time scales where they are most impacted by deficiencies in the coverage and quality of observational data and changes in observing systems over time (Sections 2.2, 2.3, 2.4).
- Current global reanalysis data are most reliable in quantities that are most strongly constrained by the observations (*e.g.*, temperature and winds), and least reliable for quantities that are highly model dependent, such as evaporation, precipitation, and cloud-related quantities (Sections 2.2, 2.3, 2.4).

- Substantial biases exist in various components of the atmospheric water cycle (*e.g.*, precipitation, evaporation and clouds), that limit the value of current reanalysis data for assessing the veracity of these quantities in climate models, as well as for practical applications. There are also significant biases in other surface and near-surface quantities related to deficiencies in representing interactions across the land-atmosphere and ocean-atmosphere interfaces (Sections 2.2, 2.3, 2.4).
- The comprehensive and multi-variate nature of reanalysis data provide value for understanding the causes of surface temperature and precipitation trends beyond what can be obtained from relatively incomplete observational datasets alone, even in the face of the noted biases in reanalysis-based trends (Section 2.4).
- Reanalysis data play a critical role in assessing the ability of climate models to simulate the statistics of climate – the means and variances (at various time scales) of basic variables such as the horizontal winds, temperature and pressure. In addition, the adjustments or analysis increments (*i.e.*, the "corrections" imposed on model states by the observations) produced during the course of a reanalysis provide a means to identify fundamental errors in the physical processes and/or missing physics that create climate model biases (Sections 2.2, 2.3).
- Reanalyses have had enormous benefits for climate research and prediction, as well as for a wide range of societal applications. Significant future improvements are possible by developing new methods to address observing system inhomogeneities, by developing estimates of the reanalysis uncertainties, by improving our observational database, and by developing integrated Earth system models and

analysis systems that incorporate key climate elements not included in atmospheric reanalyses to date (Section 2.5).

2.1. WHAT IS A CLIMATE REANALYSIS, AND WHAT ROLE DOES REANALYSIS PLAY WITHIN A COMPREHENSIVE CLIMATE OBSERVING SYSTEM?

2.1.1 Introduction

The world's weather and climate vary continuously on all time scales. The observation and prediction of these variations is vital to many aspects of human society. Extreme weather events can cause significant loss of life and damage to property. Seasonal to interannual changes associated with the El Niño-Southern Oscillation (ENSO) phenomenon and other modes of climate variability have substantial effects on the economy. Climate change, whether natural or anthropogenic, can profoundly influence social and natural environments throughout the world, with consequent impacts that can be large and far-reaching.

Determining the nature and predictability of climate variability and change is crucial to our future welfare. To address the threats and opportunities associated with weather phenomena, an extensive weather observing system has been put in place over the past century. Over the years, considerable resources have been invested in obtaining observations of the ocean, land, cryosphere, and atmosphere from satellite and surface-based systems, with plans to improve and expand these observations as a part of the Global Earth Observing System of Systems (GEOSS, 2005). Within this developing climate observing system, climate analysis plays an essential synthesizing role by integrating together data obtained from this diverse array of Earth system observations to enable improved descriptions and understanding of climate variations and change.

2.1.2 What is a Climate Analysis?

As discussed in Chapter 1, at its most fundamental level, an *analysis* is a detailed representation of the state of the atmosphere (and, more generally, other components of the Earth's climate system, such as oceans or land surface) that is based on observations. A number of techniques can be used to create an analysis from a given set of observations.

One common technique for creating an analysis is based on the expertise of human analysts, who apply their knowledge of phenomena and physical relationships to interpolate values of variables between observation locations. Such subjective analysis methods were almost universally employed before the advent of modern numerical weather prediction in the 1950s and are still used for many purposes today. While such techniques have certain advantages, including the relative simplicity by which they may be produced, they also suffer from key deficiencies that limit their value for numerical weather prediction and much climate research. An important practical deficiency, recognized in the earliest attempts at numerical weather prediction (Richardson, 1922; Charney, 1951), is that the process of creating a detailed analysis, for example, of the global winds, temperatures, and other variables through the depth of the atmosphere on a given day, is quite time consuming, often taking much longer to produce than the evolution of the weather itself. A second, more subtle deficiency is that physical imbalances between fields that are inevitably produced during a subjective analysis lead to forecast changes that are much larger than actually observed (Richardson, 1922). A third limitation of the subjective analysis method is that it is not reproducible. That is, the same analyst, given the same observational data, will generally not produce an identical analysis when given multiple opportunities.

Thus, by the early 1950s the need for an automatic, objective analysis of atmospheric conditions had become apparent. What made this goal feasible was the vital technological advance provided by the early computers of that day which, while quite primitive by today's standards, could still perform calculations far faster than previously possible.

The first objective analyses employed simple statistical techniques to interpolate data values from the locations where observations were made onto uniform spatial grids that were used for the model predictions. Such techniques are still widely employed today to produce many types of analyses, for example, global maps of surface temperatures and precipitation (Jones *et al*., 1999; Hansen *et al*., 2001; Doherty *et al*., 1999; Huffman *et al*., 1997; Xie and Arkin, 1997; Adler *et al*., 2003). However, purely statistical approaches, while of great value, also have limitations. In particular, they do not fully exploit known physical relationships among different variables of the climate system, for example, among fields of temperature, winds, and atmospheric pressure. These relationships place fundamental constraints on how weather and climate evolve in time. For this reason, statistical analysis techniques alone, while highly useful in representing fields of individual variables, are often less well-suited for applications that depend sensitively on relationships among variables, as in numerical weather prediction or in research to assess detailed mechanisms for climate variability and change.

An alternative objective analysis method, and the one that is the principal focus for this Report, is to estimate the state of the climate system (or of one of its components) by combining observations together within a numerical prediction model that represents mathematically the physical and dynamical processes operating within the system. This observations-model integration is achieved through a technique called data assimilation**.** One vital aspect of a comprehensive climate observing system achieved through data assimilation is the ability to integrate diverse surface, upper air, satellite and other observations together into a coherent, internally consistent depiction of the state of the global climate system. Figure 2.1 shows, for example, a snapshot of the coverage provided by the different atmospheric observing systems on 5 September 2003 that can be incorporated into such an analysis scheme.

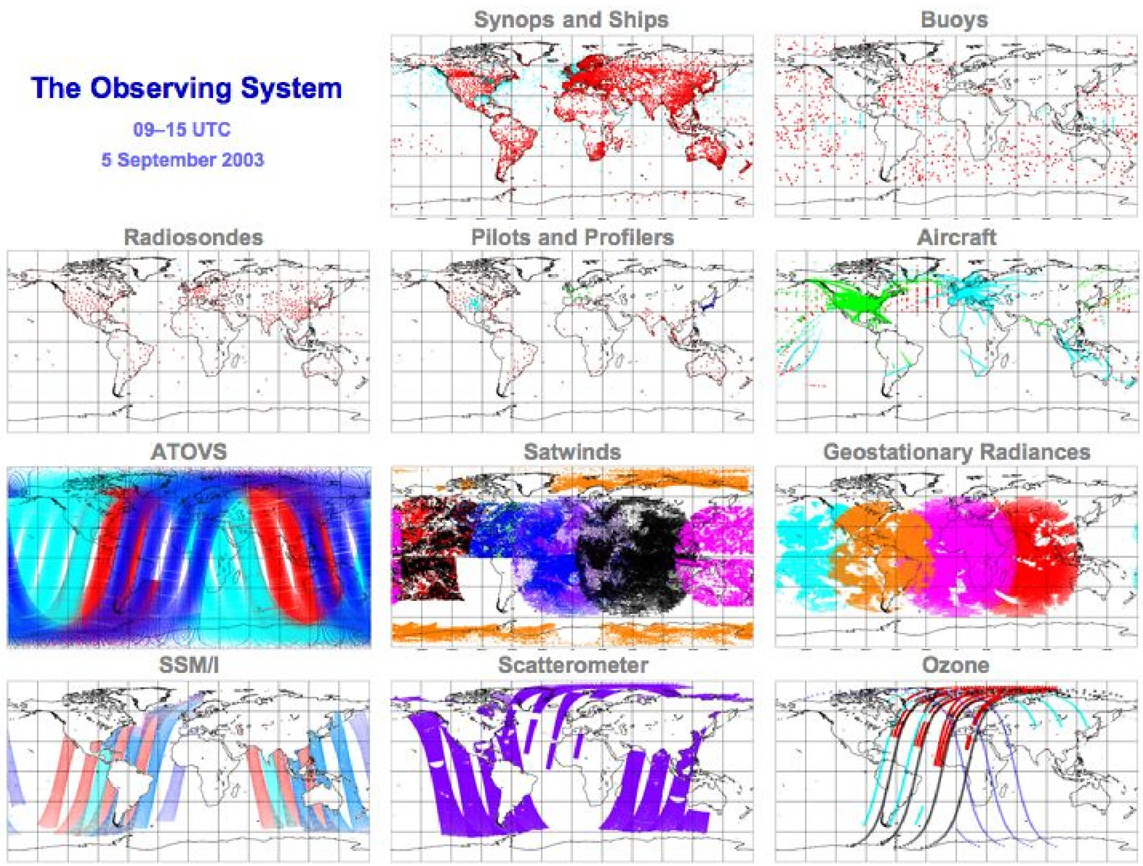

Figure 2.1 An example of the atmospheric data coverage provided by the modern observing systems (5 September 2003) for use in reanalysis. Taken from Simmons (2006).

How do we go about combining observations that have such different spatial coverage, sampling density and error characteristics? The basic method of data assimilation consists of mathematically combining a background field or "first guess" produced by a numerical prediction of the atmosphere (or oceans) with available observations in a way designed to minimize the overall errors in the analysis. Figure 2.2 shows schematically how data assimilation combines quality-controlled observations with a short-term model forecast (typically, a six-hour forecast) to produce an analysis that attempts to minimize errors in estimates of the atmospheric state that would be present from either the observations or model evaluated separately (for more details see Appendix 2.A).

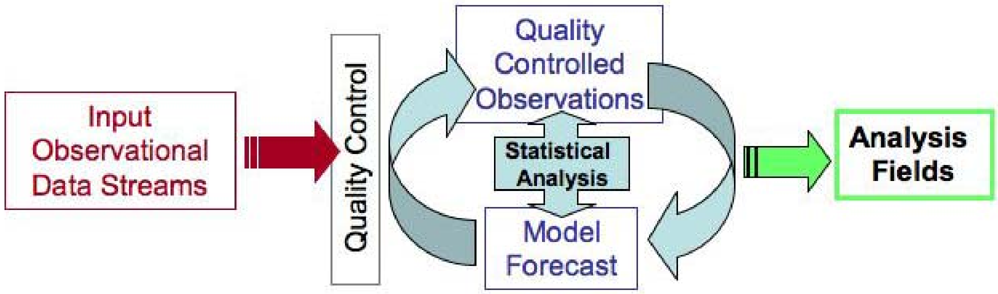

Figure 2.2 A schematic of data assimilation (adapted from a slide from Ricky Rood).

In practice, the quality of a global analysis is impacted by a multitude of practical decisions and compromises, involving the analysis methodology, quality control, the choice of observations and how they are used, and the model (see Appendix 2.A and discussion below). As one illustration of an analysis product, Figure 2.3 compares three different analyses produced from the observations available for 5 September 2003 (Figure 2.1) of the mid-troposphere pressure distribution (the geopotential height field) and total water vapor fields.

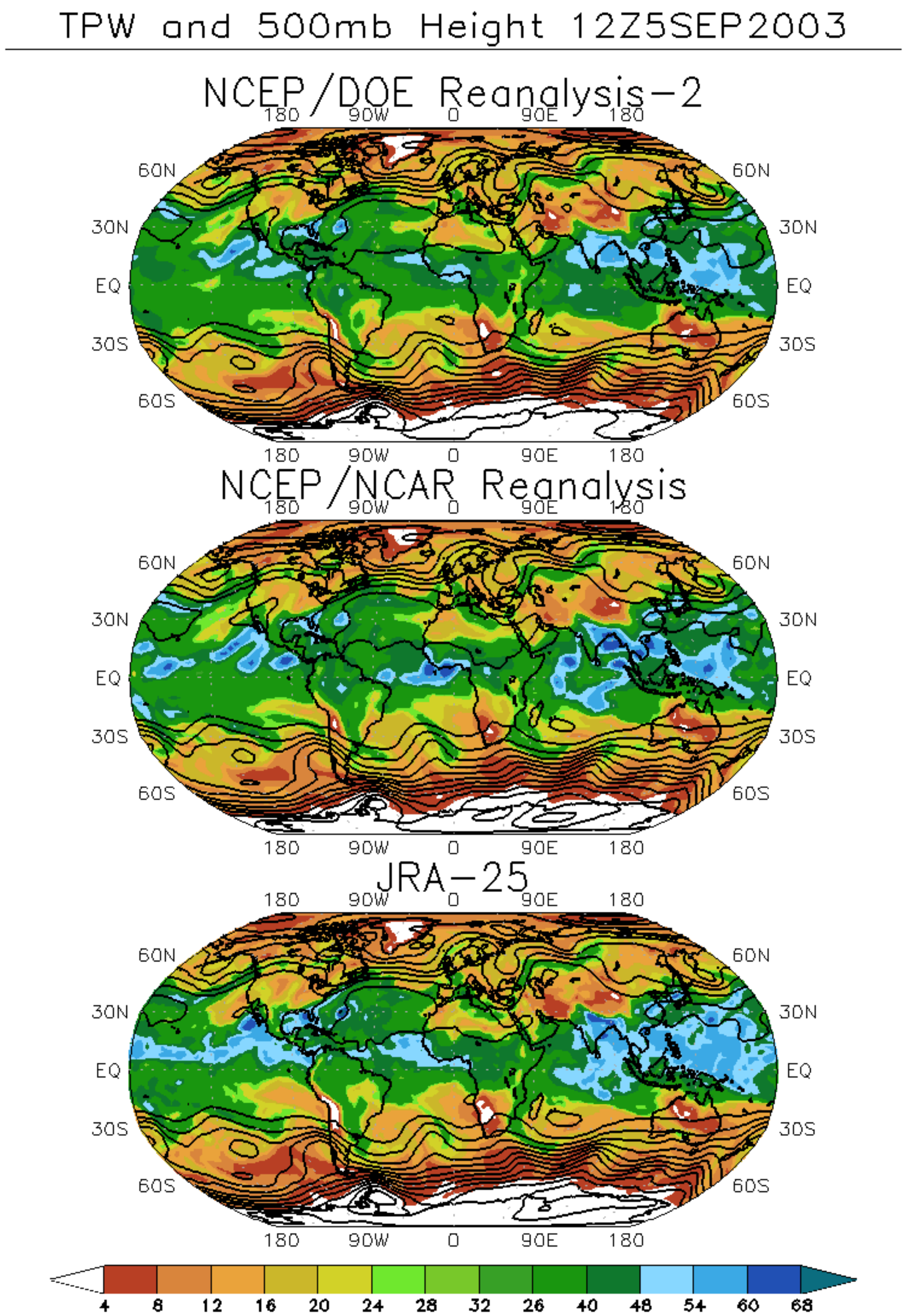

Figure 2.3 An example of the global distribution of the mid-tropospheric pressure field (contours are of the 500mb geopotential height field) and vertically integrated water vapor (shaded color - units are in mm) for 5 September 2003 from three different analyses.

We note that the two NCEP reanalyses were carried out with basically the same system (Table 2.1 – the NCEP/DOE reanalysis system corrected some of the known errors in the NCEP/NCAR system).

Table 2.1 Characteristics of existing atmospheric reanalyses.

Organization	Time Period	AGCM	Analysis scheme	Output	References
NASA DAO	1980-1994	2X2.5° Lat/lon- Δx~250 km, L20 (σ, top at 10mb), specified soil moisture	Optimal Interpolation (OI) with incremental analysis update	No longer available	Schubert *et al*. (1993)
NOAA NCEP and NCAR (R1)	1948-present	T62 - Δx~200km L28 (σ, top at about 3mb)	Spectral Statistical Interpolation (SSI)	http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis.html	Kalnay *et al* (1996)
NOAA NCEP and DOE (R2)	1979-present	T62 - Δx~200km L28 (σ, top at about 3mb)	Spectral Statistical Interpolation (SSI)	http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis2/	Kanamitsu *et al*. (2002) (Fixes errors found in R1 including fixes to PAOBS, snow, humidity, *etc*.)
ECMWF (ERA-15)	1979-1993	T106 - Δx~125km L31(σ-p, top at 10mb)	Optimal Interpolation (OI),1DVAR, nonlinear normal mode initialization	http://data.ecmwf.int/data/d/era15/	Gibson *et al* (1997)
ECMWF (ERA-40)	1957-2001	T159 - Δx~100km L60 (σ-p, top at 0.1mb)	3DVAR, radiance assimilation	http://data.ecmwf.int/data/d/era40_daily/	Uppala *et al*. (2005)
JMA and CRIEPI (JRA-25)	1979-2004	T106- Δx~125km L40 (σ-p, top at 0.4mb)	3D-Var, radiance assimilation	http://jra.kishou.go.jp/index_en.html	Onogi *et al*. (2005)
North American Regional Reanalysis (NARR)	1979-present	Δx= 32km L45	3D-Var, precipitation assimilation	http://nomads.ncdc.noaa.gov/#narr_datasets	Mesinger *et al*. (2006)

The three analyses show substantial agreement in mid-latitudes, especially for the pressure distribution. There is, however, substantial disagreement in the tropical moisture fields between the NCEP and JRA products. These differences indicate that there are insufficient observations and knowledge of physical processes (as reflected in the models) to tightly constrain the analyses and consequently, the uncertainties in the tropical moisture field are relatively large.

The numerical prediction model used for data assimilation plays a fundamental role in the analysis. It ensures an internal consistency of physical relationships among variables like temperatures, pressure, and wind fields, and provides a detailed, three-dimensional representation of the system state at any given time, including (for the atmosphere) winds, temperatures, pressures, humidity, and numerous other variables that are central for describing weather and climate (Appendix 2.A). Further, the physical relationships among atmospheric (or oceanic) variables that are represented in the mathematical model enable the model to propagate information from times or regions with more observations to other times or areas with sparse observations. At the same time, potential errors are introduced by the use of a model, as discussed in more detail later in this chapter.

Beginning in the 1970s, the sequence of initial atmospheric conditions or analyses needed for the emerging comprehensive global numerical weather prediction models were also used for climate analysis (Blackmon *et al*., 1977; Lau *et al*., 1978; Arkin, 1982). This unforeseen use of the analyses marked what could be considered a revolutionary step forward in climate science, enabling for the first time detailed quantitative analyses that were instrumental in advancing our ability to identify, describe, and understand many large scale climate variations, in particular, some of the major modes of climate variability described later in this chapter. However, the frequent changes in analysis systems needed to improve short-range numerical weather forecasts also introduced spurious shifts in the perceived climate that rendered these initial analyses unsuitable for problems such as detecting subtle climate trends. Recognition of this fundamental issue led to recommendations for the development of a comprehensive, consistent analysis of the climate system, effectively giving birth to the concept of a model-based climate reanalysis (Bengtsson and Shukla, 1988; Trenberth and Olson, 1988).

2.1.3 What is a Climate Reanalysis?

A climate reanalysis is an analysis performed with a fixed numerical prediction model and data assimilation method that assimilates quality-controlled observational data over an extended time period, typically several decades, to create a long-period climate record. This use of a fixed model and data assimilation scheme differs from analyses performed for daily weather prediction. Such analyses are conducted with models with numerical and/or physical formulations as well as data assimilation schemes that are updated frequently, sometimes several times a year, giving rise to “apparent” changes in climate that limit their value for climate applications. Climate analysis also differs fundamentally from weather analysis in that observations throughout the system evolution are available to be used, rather than simply those prior to the time when the forecast is initiated. While weather analysis has the goal of enabling the best short-term weather forecasts, climate analysis can be optimized to achieve other objectives, for example, to provide a consistent description of the atmosphere over an extended time period. However, current climate reanalyses evolved from methods developed for short-range weather prediction, and so have yet to realize their full potential for climate applications (see also Chapter 4).

Beginning in the late 1980s, several reanalysis projects were initiated to develop long time records of analyses better suited for climate purposes (Table 2.1). The products of these first reanalyses 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. They have produced detailed atmospheric climate records that have enabled successful climate monitoring and research to be conducted. They have provided a vitally needed test bed for improving prediction models on all time scales (see next section), especially for seasonal-to-interannual forecasts, as well as greatly improved basic observations and data sets prepared for their production. Reanalysis, when extended to the present as an ongoing climate analysis, provides decision makers with information about current climate events in relation to past events, and contributes directly to climate change assessments.

2.1.4. What Role Does Reanalysis Play within a Climate Observing System?

One of the key limitations of current and foreseeable observing systems is that they do not provide complete spatial coverage of all relevant components of the climate system. In fact, the observing system has evolved over the last half century mainly in response to numerical weather prediction needs, and hence is focused primarily on the atmosphere. This system today consists of a mixture of *in situ* and remotely sensed observations with differing spatial and temporal sampling and error characteristics (Figure 2.1). An example of the observations available for reanalysis during the modern satellite era is provided in Table 2.2.

Table 2.2 An example of the conventional and satellite radiance data available for reanalysis during the satellite era (late 1970s to present). These are the observations used in the new NASA MERRA reanalysis (Section 2.5.2).

DATA SOURCE/TYPE	PERIOD	DATA SUPPLIER
Conventional Data		
Radiosondes	1970 - present	NOAA/NCEP
PIBAL winds	1970 - present	NOAA/NCEP
Wind profiles	1992/5/14 - present	UCAR CDAS
Conventional, ASDAR, and MDCRS aircraft reports	1970 - present	NOAA/NCEP
Dropsondes	1970 - present	NOAA/NCEP
PAOB	1978 - present	NCEP CDAS
GMS, METEOSAT, cloud drift IR and visible winds	1977 Š present	NOAA/NCEP
GOES cloud drift winds	1997 Š present	NOAA/NCEP
EOS/Terra/MODIS winds	2002/7/01 - present	NOAA/NCEP
EOS/Aqua/MODIS winds	2003/9/01 - present	NOAA/NCEP
Surface land observations	1970 - present	NOAA/NCEP
Surface ship and buoy observations	1977 - present	NOAA/NCEP
SSM/I rain rate	1987/7 - present	NASA/GSFC
SSM/I V6 wind speed	1987/7 - present	RSS
TMI rain rate	1997/12 - present	NASA/GSFC
QuikSCAT surface winds	1999/7 - present	JPL
ERS-1 surface winds	1991/8/5 Š 1996/5/21	CERSAT
ERS-2 surface winds	1996/3/19 Š 2001/1/17	CERSAT
Satellite Data		
TOVS (TIROS N, N-6, N-7, N-8)	1978/10/30 Š 1985/01/01	NCAR
(A)TOVS (N-9; N-10 ; N-11; N-12)	1985/01/01 - 1997/07/14	NOAA/NESDIS & NCAR
ATOVS (N-14; N-15; N-16; N-18; N-18)	1995/01/19 - present	NOAA/NESDIS
EOS/Aqua	2002/10 - present	NOAA/NESDIS
SSM/I V6 (F08, F10, F11, F13, F14, F15)	1987/7 - present	RSS
GOES sounder T_B	2001/01 - present	NOAA/NCEP
SBUV2 ozone (Version 8 retrievals)	1978/10 - present	NASA/GSFC/Code 613.3

A major strength of modern data assimilation methods lies in the use of a model to help fill in the gaps of our observing system. This can be considered as a very sophisticated interpolator that uses the complex equations governing the atmosphere's evolution together with all available observations to estimate the state of the atmosphere in regions with little or no observational coverage. Statistical schemes are employed that ensure that, in the absence of bias with respect to the true state of the atmosphere, the observations and model first guess are combined in an optimal way to jointly minimize observational and model errors, subject to certain simplifying assumptions such as normality of the statistics. This can be as simple as the model transporting warm air from a region that has good observational coverage (say over the United States) to a region that has little or no coverage (say over the adjacent ocean), or a more complicated "extrapolation", for example, where the model generates a realistic low-level jet in a region where such phenomena exist but observations are limited. The latter is an example of a phenomenon that is largely generated by the model, and only indirectly constrained by observations. This example highlights both the tremendous advantages and difficulties in using reanalysis for climate studies since it allows us, through a model (which is imperfect), to "observe" features that are indirectly or incompletely measured.

The use of a model also enables estimates of quantities and physical processes that are very difficult to observe directly, such as vertical motions, surface heat fluxes, latent heating, and many of the other physical processes that determine how the atmosphere evolves in time. Such quantities are in general highly model dependent and great care must be used in interpreting them. Any bias in the model fields or incorrect representation of physical processes (called parameterizations) will be reflected in the reanalysis to some extent. In fact, only recently have the models become good enough to be used with some confidence in individual physical processes. Until recently, most studies using assimilated data have taken an indirect approach to estimating physical processes by computing them as a residual of a budget that involves only variables that are well observed (see Section 3.2.3). Thus it is important to have a good understanding of which quantities are strongly constrained by the observations, and which are only indirectly constrained and depend critically on model parameterizations. In recognition of this problem, efforts have been made to document the quality of the individual products and categorize them according to how strongly they are observationally constrained (*e.g.*, Kalnay *et al*., 1996; Kistler *et al*., 2001).

Beyond their fundamental integrating role within a comprehensive climate observing system, climate analysis and reanalysis can also be used to identify redundancies and gaps in the climate observing system, thus enabling the entire system to be configured more cost effectively. By directly linking products to observations, a reanalysis can be applied in conjunction with other science methods to optimize the design and efficiency of future climate observing systems and to improve the products that the system produces.

Despite the usefulness of current reanalysis products, they also suffer from significant limitations. For example, they are affected by changes in the observing systems, such as the introduction of satellite data in 1979, and other newer remote sensing instruments (Figure 2.4). Such changes to the observing system strongly affect the variability that is inferred from reanalyses. In particular, inferred trends and low frequency variability are of limited reliability, a result exacerbated by model bias (*e.g.*, Figure 2.5 and discussion in Sections 2.3.2.2 and 2.4.2).

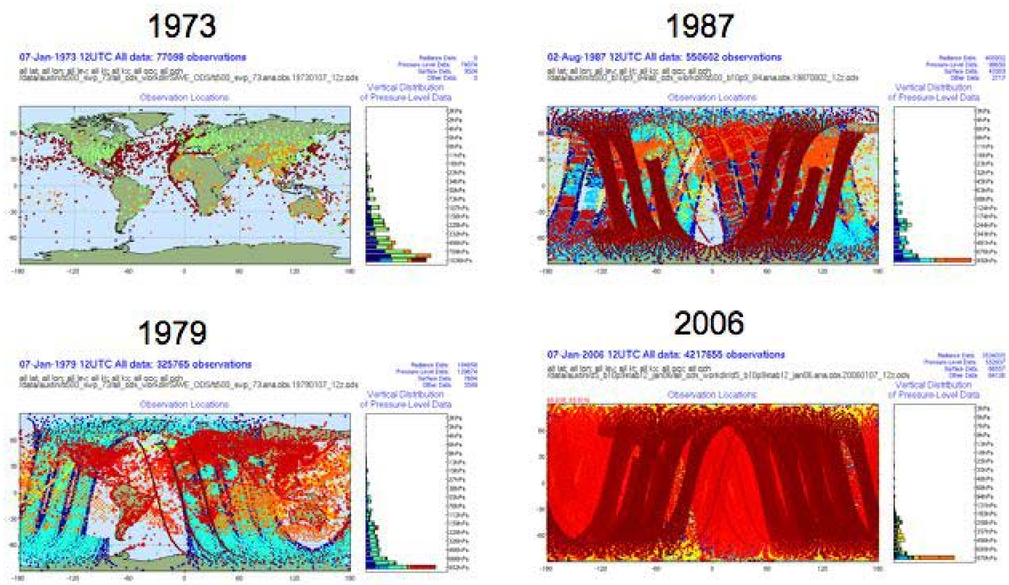

Figure 2.4 Changes in the distribution and number of observations available for NASA's MERRA reanalysis.

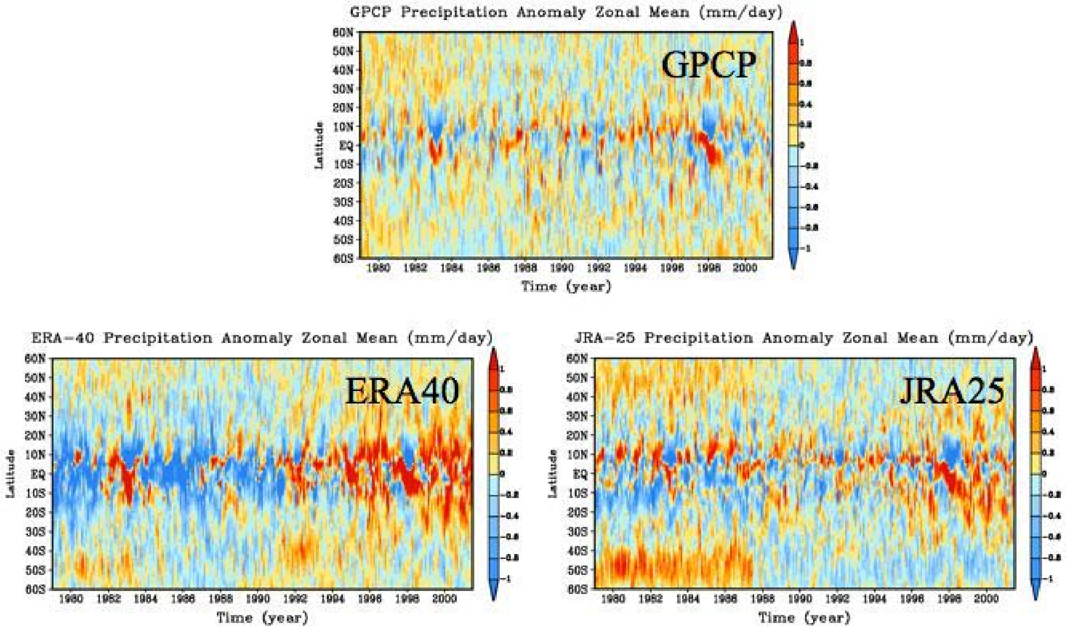

Figure 2.5 Trends and shifts in the reanalyses. The figures show the zonal mean precipitation from the GPCP observations (top panel), the ERA-40 reanalysis (bottom left panel), and the JRA-25 reanalysis (bottom right panel). Courtesy Junye Chen and Michael Bosilovich, NASA/GMAO.

The need to periodically update the climate record to provide improved reanalyses for climate research and applications has been strongly emphasized (*e.g.*, Trenberth *et al*., 2002b; Bengtsson *et al*., 2004a). Some reasons for updating reanalyses are: 1) to include critical or extensive additional observations missed in earlier analyses; 2) to correct erroneous observational data identified through subsequent quality-control efforts; and 3) to take advantage of scientific advances in models and data assimilation techniques, including bias correction techniques (Dee, 2005), and assimilating new types of observations, *e.g.*, satellite data not assimilated in earlier analyses. In the following sections, we discuss strengths and limitations of current reanalyses for addressing specific questions defined in the preface to this Report.

2.2. WHAT CAN REANALYSIS TELL US ABOUT CLIMATE FORCING AND THE VERACITY OF CLIMATE MODELS?

2.2.1 Introduction

Global atmospheric data assimilation combines various observations of the atmosphere (Figure 2.1) with a short-term model forecast to produce an improved estimate of the state of the atmosphere. The model used in the assimilation incorporates our understanding of how the atmosphere (and more generally the climate system) behaves and, ideally, can forecast or simulate all aspects of the atmosphere at all locations around the world.

As such, one can think about atmospheric data assimilation and reanalysis in particular, as a model simulation of past atmospheric behavior that is continually updated or adjusted by available observations. Such adjustments are necessary because the model would deviate from the "path" that nature took because the model is imperfect (our understanding about how the atmosphere behaves and our ability to represent that behavior in computer models is limited), and the information (observations) that we use to correct the model's "path" are incomplete and also contain errors. That is, we don't measure all aspects of the climate system perfectly – if we did, we wouldn't need to do data assimilation!

The above model-centric view of data assimilation is useful when trying to understand how reanalysis data can be applied to tell us about the veracity of climate models. It highlights the fact that reanalysis products are a mixture of observations and model forecasts, and their quality will therefore be impacted by the quality of the model. In large geographic regions with little observational coverage, a reanalysis will tend to reflect the climate of the model. Also, quantities that are poorly observed, such as surface evaporation, depend very much on the quality of the model's representation or parameterizations of the relevant physical processes (*e.g.*, in this case the model's land surface and cloud schemes). Given that models are an integral component of reanalysis systems, how then can we use reanalyses to help us understand errors in our climate models - in some cases the same model used to produce the reanalysis?

2.2.2 Assessing Systematic Errors

The most straightforward approach is simply to compare the basic reanalysis fields (*e.g.*, winds, temperature, moisture) with those that the model produces in free-running mode (a simulation that does not have the benefit of being corrected by the observations)[1]. The results of such comparisons, for example of monthly or seasonal mean values, can indicate whether the model has systematic errors such as being too cold or too wet in certain regions.

In general, such comparisons are only useful for regions and for quantities where the uncertainties in the reanalysis products are small compared to the model errors. For example, if the difference in the tropical moisture field between two reanalysis products (say NCEP/NCAR R1 and ERA-40) is as large as (or larger than) the differences between any one reanalysis product and the model results, then we could not reach any conclusion

[1] These are typically multi-year AGCM runs started from arbitrary initial conditions and forced by the observed record of sea surface temperatures (SST).

about the model quality based on that comparison. This points to the need for obtaining reliable uncertainty and bias estimates of all reanalysis quantities (*e.g.*, Dee and Todling, 2000) – something that has yet to be achieved in the current generation of reanalysis efforts. In the absence of such estimates, we can (as in the example above) get some guidance on uncertainties and model dependence by simply comparing the available reanalysis data sets. Such comparisons with reanalysis data are now routine and critical aspects of any model development and evaluation effort. Examples of such efforts span the climate modeling community and include the Atmospheric Model Intercomparison Project (AMIP) (Gates, 1992), the tropospheric-stratospheric GCM-Reality Intercomparison Project for SPARC (GRIPS) (Pawson *et al*., 2000), and coupled model evaluation conducted for the IPCC Fourth Assessment Report (IPCC, 2007).

Figure 2.6 illustrates a simple comparison between various atmospheric models and the first ECMWF reanalysis (ERA-15, see Table 2.1).

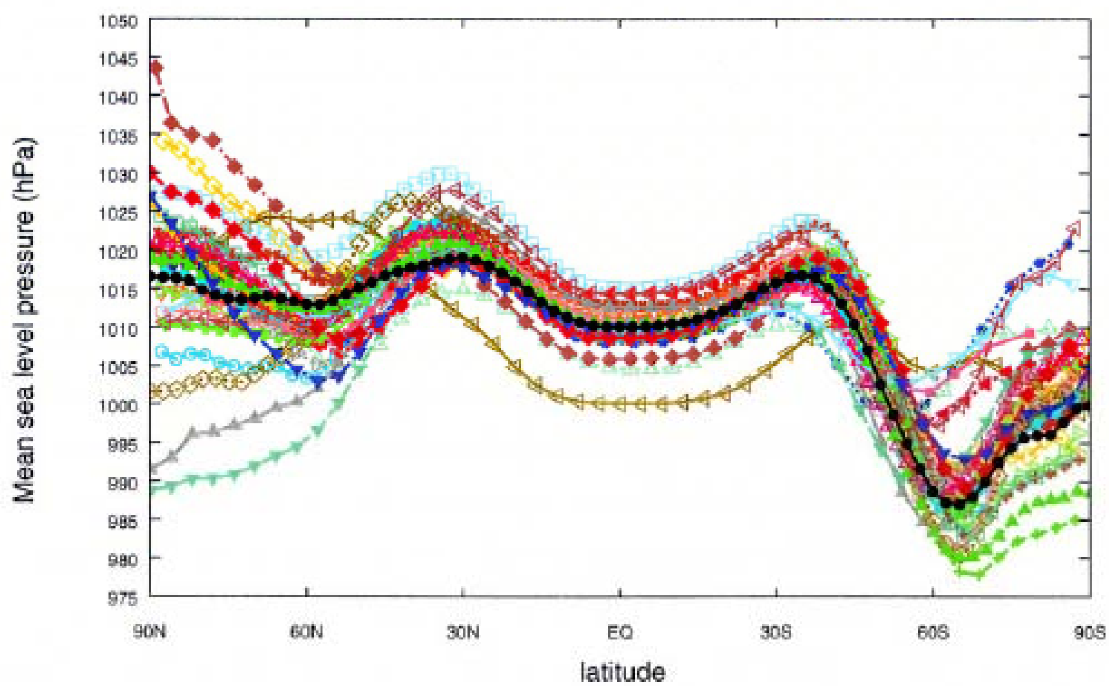

Figure 2.6 The zonal distribution of zonally-averaged sea level pressure simulated by the various AMIP models for DJF of 1979 to 1988 compared against the ECMWF (ERA-15) reanalysis (the black dots; Gibson *et al*. 1997). Taken from Gates *et al*. 1999.

The comparison shows considerable differences among the models in the zonal mean surface pressure, especially at high latitudes. It is interesting that the values scatter around the estimate provided by the reanalysis. Figure 2.7 shows an example of a more in-depth evaluation of the ability of AGCM simulations forced by observed sea surface temperatures to reproduce that part of the variability associated with ENSO.

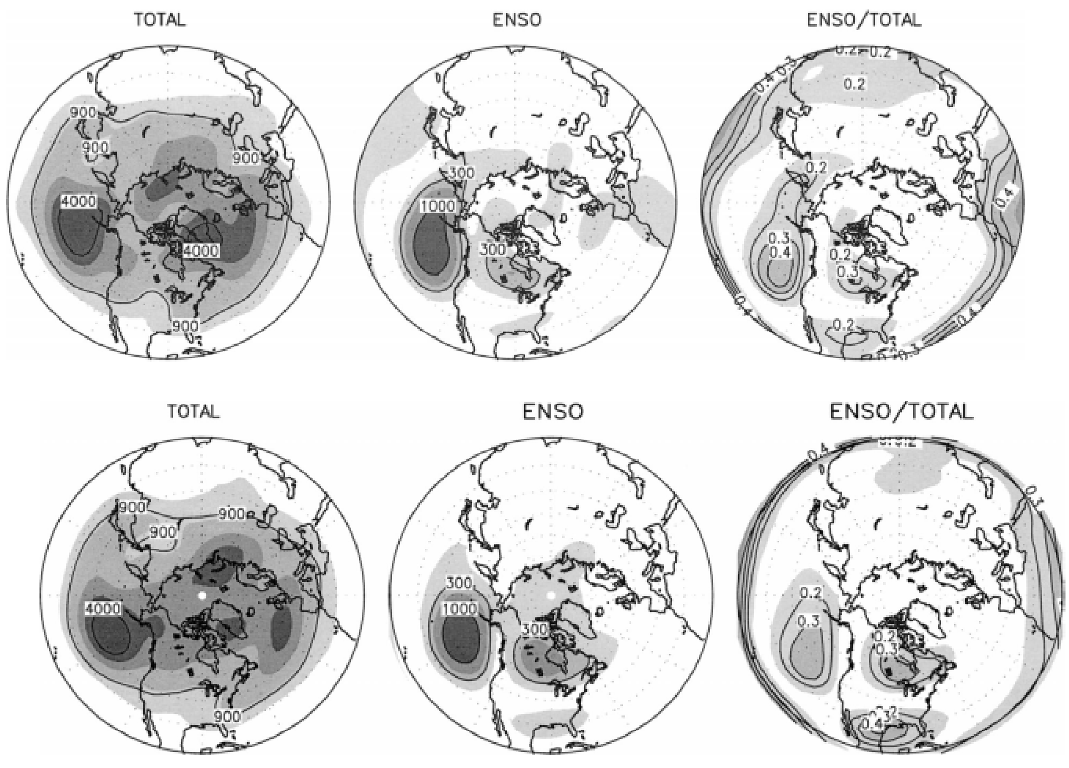

Figure 2.7 The left panels show the total variance of the time mean winter (December, January, February) 500mb height fields. The middle panels show that part of the total variance that is due to ENSO. The right panels show the ratio of the two variances (ENSO/Total). The top panels are from a reanalysis and the bottom panels are from GCM simulations forced with observed sea surface temperatures. The results are computed for the period 1950 to 1999, and plotted for the Northern Hemisphere polar cap to 20°N. The contour interval is 1000 (m^2) in the left and middle panels, and 0.1 in the right panels (taken from Hoerling and Kumar 2002).

In this case the comparison is made with the NCEP/NCAR R1 reanalysis for the winters (DJF) of 1950-1999. The comparison suggests that the models produce a very reasonable response to the ENSO-related sea surface temperature variations.

2.2.3 Inferences about Climate Forcing

While the above comparisons address errors in the description of the climate system, a more challenging problem is to address errors in the forcing or physical mechanisms (in particular the parameterizations) by which the model produces and maintains climate anomalies. This involves quantities that are generally only weakly or indirectly constrained by observations (*e.g.*, Kalnay *et al*., 1996; Kistler *et al*., 2001). Ruiz-Barradas and Nigam (2005) for example, are able to show that land/atmosphere interactions may be too efficient (make too large a contribution) in maintaining precipitation anomalies in the United States Great Plains in current climate models, despite rather substantial differences in the reanalyses. Nigam and Ruiz-Barradas (2006) highlight some of the difficulties that are encountered when trying to validate models in the presence of large differences between the reanalyses in the various components of the hydrological cycle (*e.g.*, precipitation and evaporation). This problem can be alleviated to some extent by taking an indirect approach to estimating the physical processes. In this case, a budget is computed in such a way that the reanalysis quantities that are highly model-dependent are determined indirectly as a residual of terms that are more strongly constrained by the observations (*e.g.*, Sardeshmukh, 1993). Nigam *et al*. (2000) show, for example, that the heating obtained from a residual approach appears to be of sufficient quality to diagnose errors in the ENSO-heating distribution in a climate model simulation.

Another approach to addressing errors in the forcing is to focus directly on the adjustments made to the model forecast during the assimilation (*e.g.*, Schubert and Chang, 1996; Jeuken *et al*., 1996; Rodwell and Palmer, 2007). These corrections can potentially provide a wealth of information about model deficiencies. Typically, the biases seen in, for example, the monthly mean temperature field, are the result of complex interactions among small errors in different components of the model that grow over time. The challenge to modelers is to disentangle the potential sources of error, and ultimately to correct the deficiencies at the process level to improve long-term model behavior.

An important aspect of the corrections made during data assimilation is that they are applied frequently (typically every six hours) so that the impact of the adjustments can be seen before they can interact with the full suite of model processes. In other words, the corrections made during the course of data assimilation give a potentially direct method for identifying errors in the physical processes that create model biases (*e.g.*, Klinker and Sardeshmukh, 1992; Schubert and Chang, 1996; Kaas *et al*., 1999, Danforth *et al*., 2007; Rodwell and Palmer, 2007). In fact, they can also give insights into missing model physics such as dust-forced heating in the lower atmosphere (Alpert *et al*., 1998), radiative heating in the stratosphere from volcanic eruptions (Andersen *et al*., 2001), and impacts of land use changes (Kalnay and Cai, 2003)– processes not represented in the models used in the first generation of reanalyses.

The development of a data assimilation system that provides unbiased estimates of the various physical processes inherent in the climate system (*e.g.*, precipitation, evaporation, cloud formation) is an important step in our efforts to explain, or attribute (see Chapter 3) the causes of climate anomalies. As such, reanalyses allow us to go beyond merely documenting what happened. We can, for example, examine the processes that maintain a large precipitation deficit in some region. Is the deficit maintained by local evaporative processes or changes in the storm tracks that bring moisture to that region, or some combination? As described in the next chapter, reanalysis data provide the first steps in a process of attribution that involves detection and description of the anomalies, and an assessment of the important physical processes that contribute to their development. Ultimately, we seek answers to questions about the causes that cannot be addressed by reanalysis data alone. Going back to the previous example, how can we disentangle the role of local evaporative changes and changes in the storm tracks? This requires model experimentation such as that described in the next chapter. It should be noted that even in that case, reanalyses play an important role in validating the model behavior.

2.2.4 Outlook

There are a number of steps that can be taken to increase the value of reanalyses for identifying model deficiencies, including: improving our estimates of uncertainties in all reanalysis products, balancing budgets of key quantities (*e.g.*, heat, water vapor, energy) (Kanamitsu and Saha, 1996; see also the next section), and reducing the spurious model response to the adjustments made to the background forecast by the insertion of observations (the so-called model spin-up or spin-down problem), especially when the adjustments involve water vapor and the various components of the hydrological cycle (Kanamitsu and Saha, 1996; Schubert and Chang, 1996; Jeuken *et al*., 1996). For example, Annan *et al*. (2005) proposed a method based on an ensemble of roughly 50 forecast integrations that estimates frictional and diffusive parameters. These and other approaches hold substantial promise of obtaining optimal estimates of uncertain model parameters from reanalyses, even for the very complex current climate models.

2.3. WHAT IS THE CAPACITY OF CURRENT REANALYSES TO HELP US IDENTIFY AND UNDERSTAND MAJOR SEASONAL-TO-DECADAL CLIMATE VARIATIONS, INCLUDING CHANGES IN THE FREQUENCY AND INTENSITY OF CLIMATE EXTREMES SUCH AS DROUGHTS?

In this section we examine the strengths and weaknesses of current reanalyses for identifying and understanding climate variability. This is an important step for addressing the more general issue of attribution (how well we understand the causes of climate variability) introduced in Chapter 1 and addressed more fully in Chapter 3. Understanding the connections between reanalysis, models and attribution is crucial for understanding the broader path towards attribution outlined in Chapter 1 (Box 2.1).

Box 2.1 The Complementary Roles of Reanalysis and Free-Running Model Simulations in the Attribution Problem

Section 2.3 demonstrates the value of reanalysis for identifying and understanding climate variability. By providing best estimates of the circulation patterns and other weather elements (moisture transport, evaporation, precipitation and cloudiness) present during observed extremes -- estimates that are temporally and spatially comprehensive and self-consistent -- reanalysis indeed offers a unique and profound contribution to the more general problem of attribution discussed in Chapter 3. Reanalysis is best positioned, for example, to provide a global picture of the prevailing anomalous circulation patterns associated with a given drought. By studying reanalysis data, investigators can hypothesize linkages between the drought and contemporaneous climate anomalies in other parts of the world (*e.g.*, anomalies in sea surface temperatures, or SSTs).

Reanalysis, however, is but one tool for addressing the problem. A drawback of reanalysis in this context is its inability to isolate causality -- to demonstrate unequivocally that one climate feature (*e.g.*, anomalous SSTs) causes another (*e.g.*, drought). Indeed, this drawback would extend to any imaginable set of direct observations of the atmosphere. To isolate causality, we need climate model simulations that are unconstrained by the assimilation of observational data. Such climate models can be forced in different ways to determine whether a certain forcing will cause the model to reproduce a climate anomaly of interest. For example, if an investigator suspects, perhaps based on an analysis of reanalysis data, that anomalous SSTs caused the severe drought in the southern Great Plains during the1950s, he or she could perform two simulations with a free-running climate model, one in which the 1950s SST anomalies are imposed, and one in which they are not. If only the first simulation reproduces the drought, the investigator has evidence to support the hypothesized role of the SSTs. An additional step would be to determine what caused the SST anomalies in the first place, and for that one would need further experiments with a fully coupled atmosphere/ocean/land model.

Such free-running modeling studies, of course, have their own basic deficiencies, most importantly the potential lack of realism in the climate processes simulated by an unconstrained (non-reanalysis) modeling system. This suggests an important additional role of reanalysis in the attribution problem. Not only can the reanalysis data help in the formulation of hypotheses to be tested with a free-running climate model, but the reanalysis data can (and should) be used to verify that the free-running model is behaving realistically, *i.e.*, that the variations in circulation and other climate processes in the free-running model are consistent (statistically and/or mechanistically) with what we have learned from reanalysis (see section 2.2). In effect, reanalysis and free-running model simulations are complementary tools for addressing the attribution problem, each with their own strengths and weaknesses. Only the unconstrained parts of a model can be used to address attribution (causality), implying the need for free-running models, but those unconstrained parts must be evaluated for realism, implying the need for reanalysis. Arguably, the best attack on the attribution problem is to use the reanalysis and free-running model approaches in tandem.

2.3.1. Climate Variability

The climate system varies on a wide range of time and space scales. The variability of the atmosphere in particular encompasses individual weather events that we experience every day, and longer-term changes that affect global weather patterns and can result in regional droughts or wet periods (pluvials) lasting many years. A primary goal of climate research is to understand the causes of these long-term climate variations and changes and to develop models that allow us to predict them.

On intra-seasonal to decadal time scales there are a number of key recurring global-scale patterns of climate variability that have pronounced impacts on the North American climate (Table 2.3). These include the Pacific North American pattern (PNA), the Madden-Julian Oscillation (MJO), the North Atlantic Oscillation (NAO) and the related Northern Annular Mode (NAM), the Quasi-Biennial Oscillation (QBO), El Nino-Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), and the Atlantic Multi-decadal Oscillation (AMO). These patterns, sometimes referred to as modes of climate variability or teleconnection patterns, can have pronounced effects on North American climate by shifting weather patterns and disrupting local climate features (*e.g.*, Gutzler *et al*., 1988; Hurrell, 1996).

Table 2.3 Characteristics of some of the leading modes of climate variability that are known to have a substantial impact on North American climate. The last column provides a subjective assessment of the quality of the atmospheric manifestations of these modes (and their impacts on regional climate) in current atmospheric reanalyses.

Phenomena	**Key references**	**Time scales**	**Link between atmosphere and ocean**	**Some impacts on North America**	**Consistency between atmospheric reanalyses**
Pacific/North American (PNA) pattern	Wallace and Gutzler (1981)	Subseasonal to Seasonal	Weak to moderate	West coast storms	good
Madden Julian Oscillation (MJO)	Madden and Julian (1994)	Approximately 30-60 days	Weak to moderate	Atlantic hurricanes	Fair to poor
North Atlantic Oscillation (NAO)	Hurrell *et al*. (2001)	Subseasonal to decadal	moderate on long time scales	East coast winters	good
Northern Annular Mode (NAM)	Thompson (2000);	Subseasonal to decadal	moderate on long time scales	East coast winters	Good to fair in stratosphere

	Wallace (2000)				
El Nino/ Southern Oscillation (ENSO)	Philander (1990)	Seasonal to interannual	strong	Winter in west coast and southern tier of United States, Mexico, warm season regional droughts	Good to fair on longer time scales
Pacific Decadal Oscillation (PDO)	Zhang *et al.* (1997)	decadal	strong	Drought or pluvials over North America	Fair to poor
Atlantic Multi-decadal Oscillation (AMO)	Folland *et al.* (1986)	decadal	strong	Drought or pluvials over North America, Atlantic hurricanes	Fair to poor

As we shall see in the following sections, the quality of the representation of these phenomena in reanalyses vary and depend on the time scales, locations, and physical mechanisms relevant to each of these modes of variability. The last column in Table 2.3 gives our expert assessment of the consistency of the atmospheric manifestations of these modes (and their impacts on regional climate) in current reanalyses based on such general considerations.

Figures 2.8 and 2.9 show examples of the connection between the PNA and NAO patterns and North American surface temperature and precipitation variations. The spatial correspondence between the reanalysis tropospheric circulation and the independently-derived surface fields show the potential value of the reanalysis data for interpreting the relationships between changes in the climate modes and regional changes in surface temperature and precipitation.

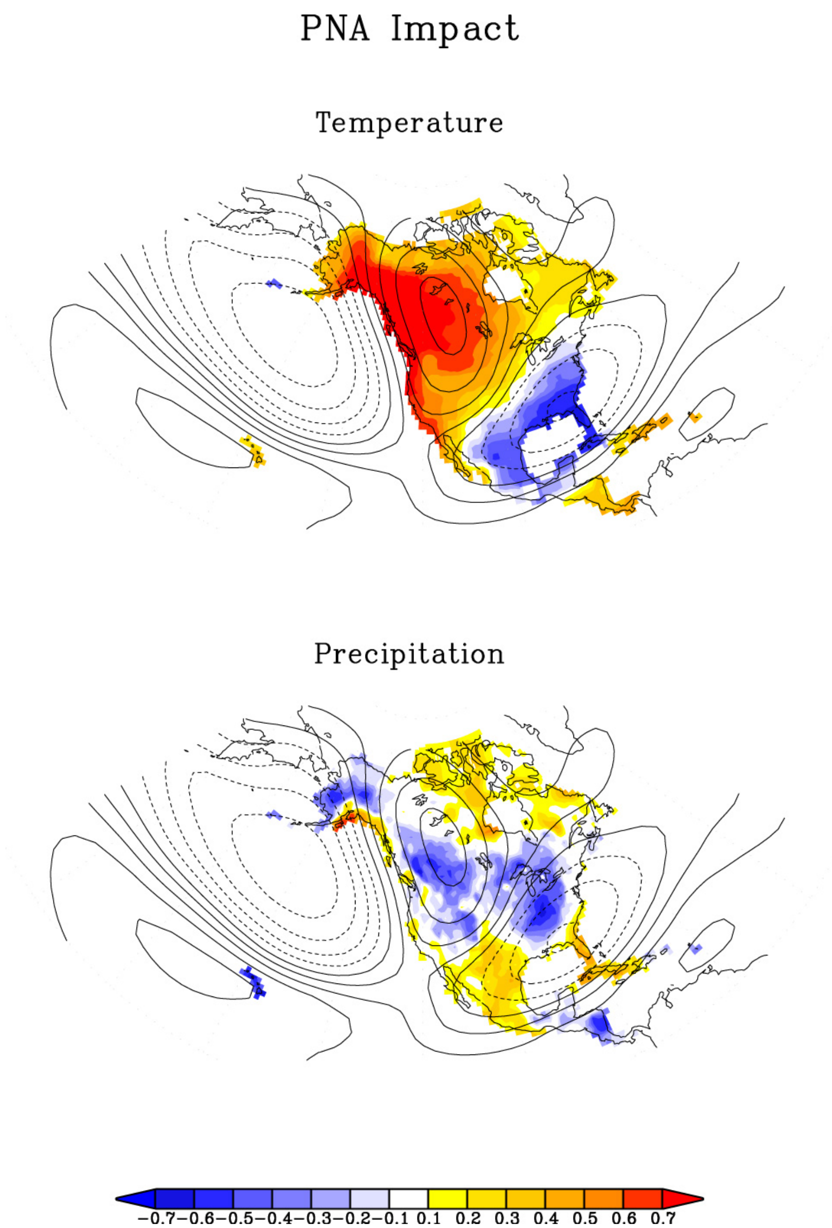

Figure 2.8 The correlation between the PNA index (Wallace and Gutzler 1981) and 500mb height field (contours). The shading indicates the correlations between PNA index and a) the surface temperature and b) the precipitation. The 500mb height is from the NCEP/NCAR R1 reanalysis. The surface temperature and precipitation are from independent observational data sets. The correlations are based on seasonal mean data for the period 1951 to 2006. The contours of correlation give an indication of the direction of the mid-tropospheric winds, and the positions of the troughs and ridges.

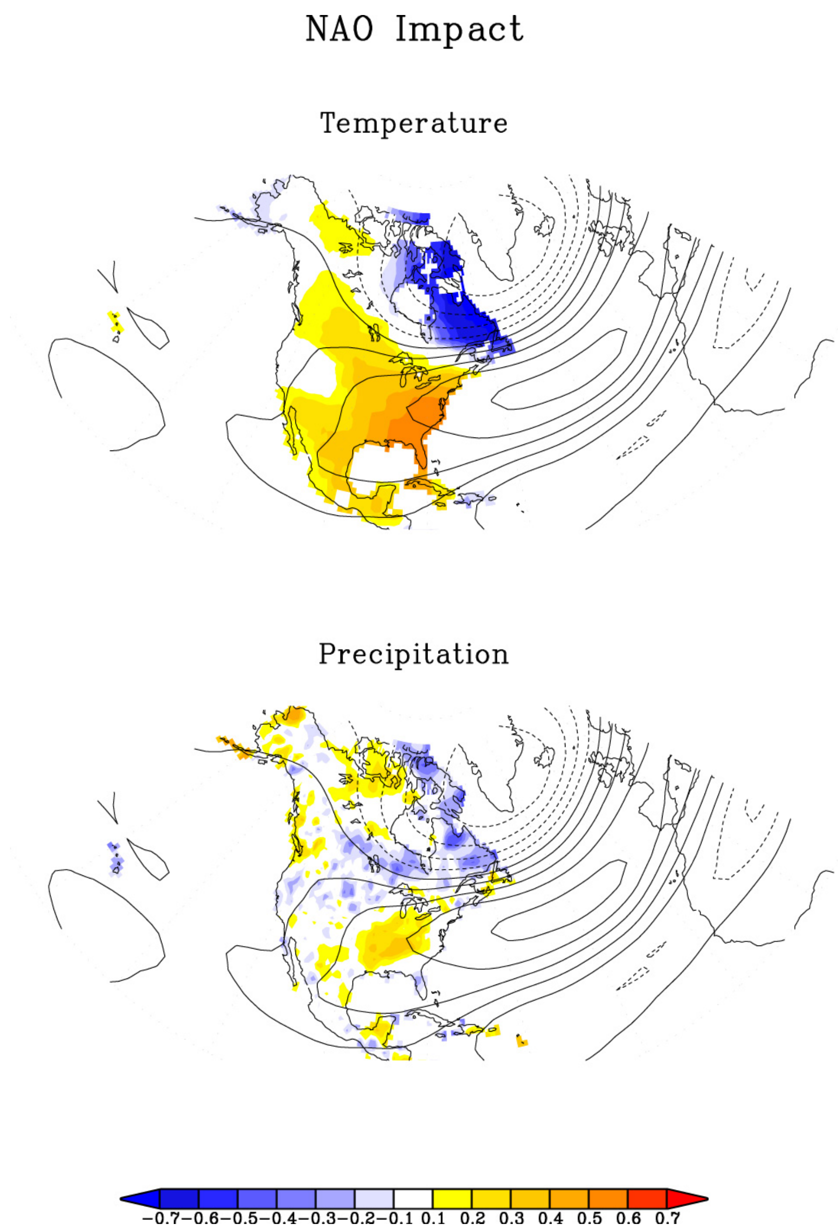

Figure 2.9 The correlation between the NAO index (Wallace and Gutzler 1981) and 500mb height field (contours). The shading indicates the correlations between NAO index and a) the surface temperature and b) the precipitation. The 500mb height is from the NCEP/NCAR R1 reanalysis. The surface temperature and precipitation are from independent observational data sets. The correlations are based on seasonal mean data for the period 1951 to 2006. The contours of correlation give an indication of the direction of the mid-tropospheric winds, and the positions of the troughs and ridges.

Specifically, during the positive phase of the PNA pattern, surface temperatures over western North America tend to be above average, and this can be related to an unusually strong high pressure ridge over the region as well as transport of warm Pacific air poleward along the west coast extending to Alaska. An upper-level trough centered over the southeast United States and the associated intensified north to south flow over the center of the continent facilitates the southward transport of Arctic air that produces a tendency toward below normal temperatures over the Gulf coast states. This same flow pattern is associated with transport of relatively dry polar air and a tendency to produce descending motions in the middle troposphere over the Missouri and Mississippi regions, both of which favor below normal precipitation, as observed. In contrast, the positive phase of the NAO pattern is accompanied by above average temperatures over the eastern United States and wetness in the Ohio Valley. The reanalysis data of tropospheric circulation help to interpret this relationship as resulting from a northward shifted westerly flow regime over the eastern United States and North Atlantic that inhibits cold air excursions while simultaneously facilitating increased moisture convergence into the region.

The above patterns arise mainly, but not exclusively, as manifestations of internal atmospheric variability (*e.g.*, Massacand and Davies, 2001; Cash and Lee, 2001; Feldstein, 2002, 2003; Straus and Shukla, 2002), and as discussed in Chapter 3, are also linked in varying degree to land surface and ocean variations. Understanding seasonal to decadal climate variability requires that we understand the physical mechanisms that produce these large-scale patterns, including how they interact with each other, and their coupling with the different climate system components (Chapter 3).

A key factor that limits our ability to fully understand such long-term variability has been the lack of long-term comprehensive and consistent observations of the climate system, including observations of the land and ocean, which are critical to understanding and predicting atmospheric variability on seasonal and longer time scales. Observations of each of these components of the climate system, while improving with the advent of the satellite era, are still far from satisfactory for addressing climate problems. In order to adequately address seasonal and longer variability, the observations need to cover many decades, span the globe, include all the key climate parameters, be consistent with our best physical understanding, and be continuous in time.

While these conditions are not fully met for any components of the climate system (see the following sections), the most advanced observational capabilities are of the atmospheric component. This system was developed primarily to support weather prediction, with major advances occurring with the advent of an upper air network of radiosondes in the 1950s, and with a near global observing system provided by the great increases in satellite measurements beginning in the late 1970s. While new efforts are underway to develop a true climate observing system spanning all climate system components and that provides continuity in time and space, the present climate observing system is inadequate for many applications (GEOSS, 2005).

2.3.2 Reanalysis and Climate Variability

One of the most important insights of the last few decades regarding our existing observational record was that we could leverage our investment in operational weather prediction by harnessing the prediction infrastructure (the global models and data assimilation methods for combining disparate observations) to develop a more consistent historical record of the atmosphere (Bengtsson and Shukla, 1988; Trenberth and Olson, 1988). This led to the development of several atmospheric climate reanalysis data sets (Schubert *et al*., 1993; Kalnay *et al*., 1996; Gibson *et al*., 1997). These data sets provided the first comprehensive depictions of the global atmosphere that, in the case of the NCEP/NCAR reanalysis (Kalnay *et al*., 1996) now span over 60 years. Studies using these and several follow-on reanalyses (Kanamitsu *et al*., 2002; Uppala *et al*., 2005; Onogi *et al*., 2005; Mesinger *et al*., 2006[2]) to examine seasonal to decadal variability of climate form the basis for this section (Table 2.1).

Over extended time periods, the reanalysis data provide the most comprehensive picture to date of the state of the atmosphere and its evolution. The reanalyses also provide estimates of the various physical processes such as precipitation, cloud formation, and radiative fluxes that are required to understand the mechanisms by which climate evolves. As we examine the utility of current reanalyses for identifying and understanding atmospheric variability, the critical roles of the model in determining the quality of the reanalysis must be recognized, and the impact of the spatial and temporal inhomogeneities of the observing system must also be appreciated. When assessing the utility of the reanalyses, we must also consider the nature of the problem that is being addressed. What is the time scale? What is the spatial scale? Does the problem involve the tropics or Southern Hemisphere, which tend to be less well observed, especially before the advent of satellite observations? To what extent are water vapor and clouds, or

[2] While not global, the North American Regional Reanalysis (NARR) has played an important role for studying regional climate variability. Two of its key strengths are the enhanced resolution, and the fact that precipitation observations were assimilated.

links to the land surface or the ocean important? These are important considerations, because assimilation systems used for the first generation of reanalyses evolved out of the needs of numerical weather prediction, which did not place a high priority on modeling details of the hydrological cycle or links to the land and ocean, which were deemed to be of secondary importance for producing weather forecasts from a day to a week in advance.

In the following subsections, we address the capacity of current reanalyses to describe and understand major seasonal-to-decadal climate variations by examining three key aspects of reanalyses: their spatial characteristics, their temporal characteristics, and their internal consistency and scope. We include in each subsection key examples of where reanalyses have contributed to our understanding of seasonal to decadal variability and where they fall short. We build on the results of two major international workshops on reanalysis (WCRP, 1997; WCRP, 1999) by emphasizing studies that have appeared in the published literature since the last workshop.

2.3.2.1 Spatial characteristics

The globally complete spatial coverage provided by reanalyses, along with estimates of the physical processes that drive the atmosphere, has greatly facilitated diagnostic studies that attempt to identify the causes of large-scale atmospheric variability that have substantial impacts on North American weather and climate (*e.g.*, the NAO and PNA). Our understanding of the nature of both the NAO and PNA has been substantially improved by studies using reanalysis products. Thompson and Wallace (2000), for example, provided a global perspective on the NAO, using reanalysis data to link it to the so-called Northern Hemisphere Annular Mode (NAM), and noting the similarities of that mode to another annular mode in the Southern Hemisphere. Reanalysis data have also been used to link the variability of the NAO to that in the stratosphere in the sense that anomalies developing in the stratosphere propagate into the troposphere, suggesting an intriguing source of potential predictability on intraseasonal time scales (*e.g.*, Baldwin and Dunkerton, 1999; 2001). Detailed studies made possible by reanalysis data have contributed to our understanding that both PNA and NAO modes of variability are fundamentally internal to the atmosphere, that is, they would exist naturally in the atmosphere without any anthropogenic or other "external" forcing (*e.g.*, Massacand and Davies, 2001; Cash and Lee, 2001; Feldstein, 2002; 2003; Straus and Shukla, 2002; see also next chapter on attribution). Straus and Shukla (2002), in particular, emphasized the differences between the PNA and a similar pattern of variability in the Pacific/North American region that is forced primarily as an atmospheric response to the tropical sea-surface temperature changes associated with ENSO.

In addition to improving our understanding of various global modes of atmospheric variability, reanalysis data allow in-depth evaluations of the physical mechanisms and global connections of high impact regional climate anomalies such as droughts or floods. For example, Mo *et al*. (1997), building on several earlier studies (*e.g.*, Trenberth and Branstator, 1992; Trenberth and Guillemot, 1996), capitalized on the long record of the NCEP/NCAR global reanalyses to provide a detailed analysis of the atmospheric processes linked to floods and droughts over the central United States, including precursor events tied to large-scale wave propagation and changes in the Great Plains low level jet (LLJ). Liu *et al*. (1998) use reanalysis data in conjunction with a linear model to deduce the role of various physical and dynamical processes in the maintenance of the circulation anomalies associated with the 1988 drought and 1993 flood over the United States.

Process studies focused on North America have benefited from the high resolution and improved precipitation fields of the North American Regional Reanalysis (NARR). They include studies of the nature and role of the LLJ (*e.g.*, Weaver and Nigam, 2008), land-atmosphere interactions (*e.g.*, Luo *et al*., 2007), and efforts to validate precipitation processes in global climate models (*e.g.*, Lee *et al*., 2007).

The above studies highlight the leading role of reanalysis data in the diagnostic evaluation of large-scale climate variability and of the physical mechanisms that produce high impact regional climate anomalies.

While reanalysis data have played a fundamental role in diagnostic studies of the leading modes of middle- and high- latitude variability and of regional climate anomalies, there are deficiencies that are particularly apparent in the stratosphere – a region of the atmosphere particularly poorly resolved in the first-generation reanalysis systems (*e.g.*, Pawson and Fiorino, 1998a; 1998b; 1999; Santer *et al*., 2003). Figure 2.10 shows an example of the substantial differences between the reanalyses that occur in the tropical stratosphere even in such a basic feature as the annual cycle of temperature.

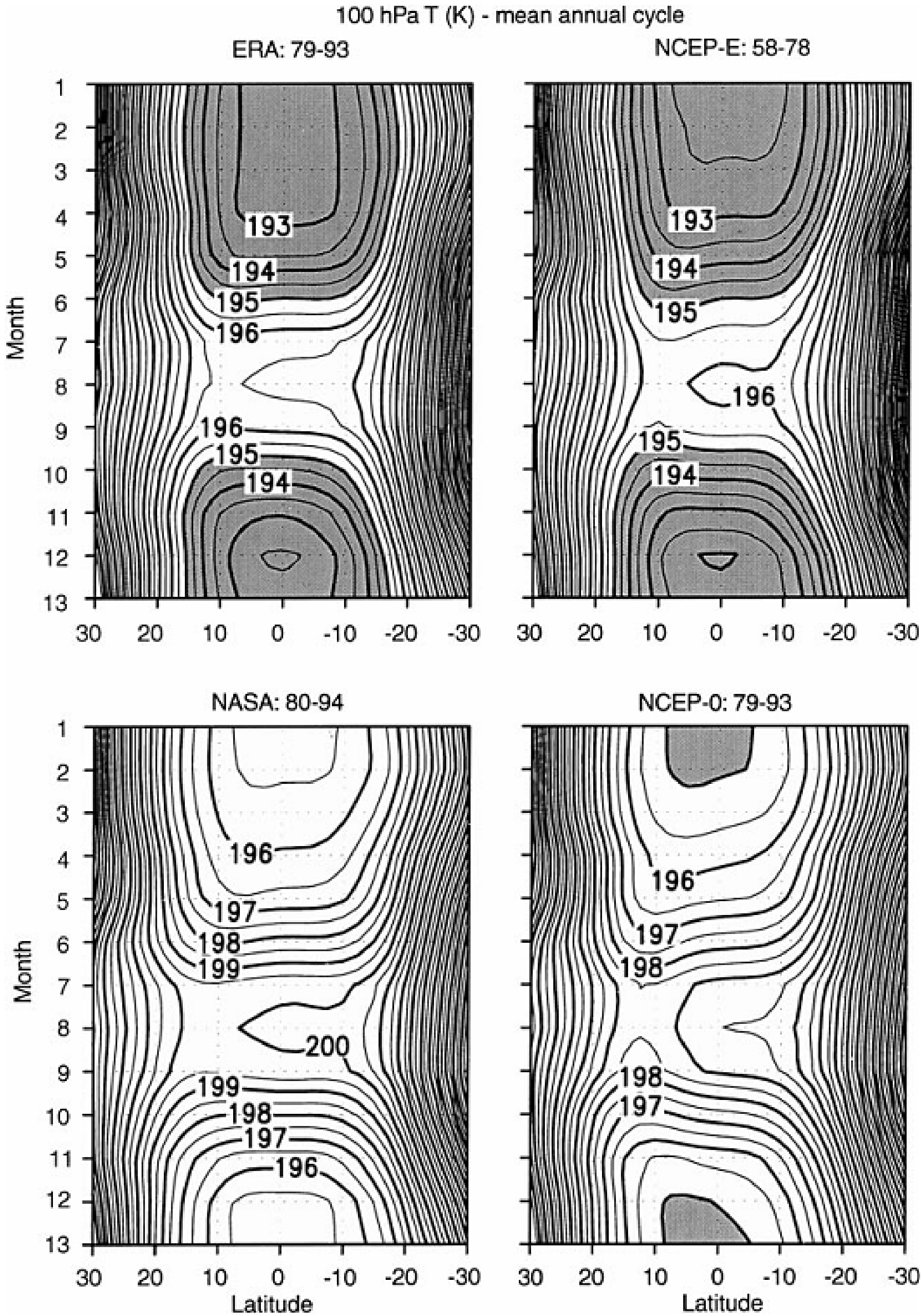

Figure 2.10 Latitudinal structure of the annual cycle in T(K) at 100 hPa for ERA (1979 to 1993, top left), NCEP-O (1958 to 1978, top right), NASA/DAO (1980 to 1994, bottom left), and NCEP-E(1979 to 1993, bottom right). The contour interval is 0.5 K. Temperatures lower than 195 K are shaded. Taken from Pawson and Fiorino (1999).

Another key problem area is in polar regions where the reanalysis models have deficiencies in both the numerical representation and in modeling of physical processes (*e.g.*, Walsh and Chapman, 1998; Cullather *et al*., 2000, Bromwich and Wang, 2005; Bromwich *et al*., 2007). Reanalyses to date are particularly deficient in the modeled polar cloud properties and associated radiative fluxes (*e.g.*, Serreze *et al*., 1998).

Variations in tropical sea surface temperatures, particularly those associated with ENSO, are a major contributor to climate variability over North America on interannual time scales (*e.g.*, Trenberth *et al*., 1998). Recent studies that use reanalysis data have contributed to important new insights on the linkages between tropical Pacific sea surface temperature variability and the extratropical circulation (*e.g.*, Sardeshmukh *et al*., 2000; Hoerling and Kumar, 2002; DeWeaver and Nigam, 2002), the global extent of the ENSO response (*e.g.*, Mo, 2000; Trenberth and Caron, 2000), and its impact on weather (*e.g.*, Compo *et al*., 2001; Gulev *et al*., 2001; Hodges *et al*., 2003; Raible, 2007; Schubert *et al*., 2008). An important aspect of many of the studies cited above is that they include companion model simulation experiments. In such studies the reanalyses are used to both characterize the atmospheric behavior and to validate the model results. This is an important advance in climate diagnosis resulting from increased confidence in climate models, and represents an important synergy between reanalysis and the attribution studies discussed in the next chapter.

While the reanalyses have proven themselves useful in many respects for addressing the problem of tropical/extratropical connections, they do have important deficiencies in representing tropical precipitation, clouds and other aspects of the hydrological cycle (*e.g.*, Newman *et al*., 2000). The Madden-Julian Oscillation or MJO is an example of a phenomenon where coupling between the circulation and tropical heating is fundamental to its structure and evolution (*e.g.*, Lin *et al*., 2004) – a coupling that is poorly represented in climate models. Current reanalysis products are inadequate for validating models, since those aspects of the MJO that appear to be critical for the proper simulation of the MJO (*e.g.*, the vertical distribution of heating) are poorly constrained by the observations and therefore are highly dependent on the models used in the assimilation systems (*e.g.*, Tian *et al*., 2006). Nevertheless, indirect (residual) approaches to estimating the tropical forcing from reanalyses have proven themselves useful, reflecting the greater confidence placed in the estimates of certain aspects of the large-scale tropical circulation (Newman *et al*., 2000; Nigam *et al*., 2000)

While the NAO, PNA and ENSO phenomena notably influence subseasonal to interannual climate variability, there is evidence that these modes also may vary on decadal or longer time scales. Understanding that behavior, as well as other possibly intrinsically decadal-scale modes of variability such as the PDO and the AMO require datasets that are consistent over many decades. We examine next the capacity of current reanalyses to address such longer time scale variability.

2.3.2.2 Temporal characteristics

A defining characteristic of the observing system of the last 100 years or so is that it varies greatly over time. Prior to the mid 20th century, the observing system was primarily surface-based and limited to land areas and ship reports, though some upper observations (*e.g.*, wind measurements from pilot balloons) were made routinely since the early 20th century (*e.g.*, Brönnimann *et al*., 2005). The 1950s marked the beginning of an upper air radiosonde network of observations, though these were primarily confined to land areas and especially Northern Hemisphere middle latitudes. The advent of satellite observations in the 1970s marked the beginning of a truly global observing system, with numerous changes subsequently to the observing system as new satellites were launched with updated and more capable sensors, and older systems were discontinued (Figure 2.2). This, together with sensor changes and the aging and degrading of existing sensors, makes the problem of combining all available observations into a temporally consistent long-term global climate record a tremendous challenge. Figure 2.11 provides an overview of the number of observations that were available to the NCEP/NCAR reanalysis (Kistler *et al*., 2001). These changes, especially the advent of satellite observations, have impacted the reanalysis fields, often making it difficult to separate true climate variations from artificial changes associated with the evolving observing system.

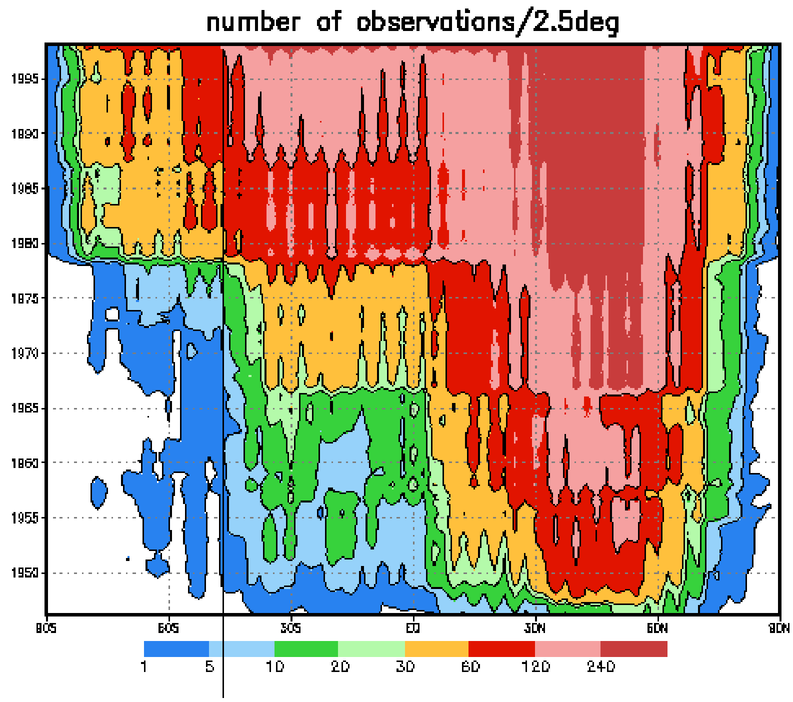

Figure 2.11 Zonal mean number of all types of observations available to the NCEP/NCAR reanalysis per 2.5° lat-long box per month from 1946 to 1998. A 12-month running mean has been applied. From Kistler *et al*. (2001)

The changes in the observing system have especially impacted our ability to study variability on interannual and longer time scales – the time scales at which changes to the observing system also tend to occur (*e.g.*, Basist and Chelliah, 1997; Chelliah and Ropelewski, 2000; Kistler *et al*., 2001; Trenberth *et al*., 2001; Kinter *et al*., 2004). The impact can be quite complicated, involving interactions and feedbacks with the assimilation schemes. For example, Trenberth *et al*. (2001) show how discontinuities in tropical temperature and moisture fields can be traced to the bias correction of satellite radiances in the ECMWF (ERA-15) reanalyses. Changes in the conventional radiosonde observations can also have impacts. For example the QBO, while clearly evident throughout the record of the NCEP/NCAR reanalysis, shows substantial secular changes in amplitude that are apparently the result of changes in the availability of tropical wind observations (Kistler *et al*., 2001). The major change in the observing system associated with the advent of satellite data in the 1970s represents a particularly difficult and important problem since it coincides with the time of a major climate shift associated with the PDO (*e.g.*, Pawson and Fiorino, 1999; Trenberth and Caron, 2000; Chelliah and Bell, 2004).

Despite these problems, reanalysis data can be very valuable in understanding long-term atmospheric variability, particularly if used in conjunction with other independent data. For example, Barlow *et al*. (2001) used NCEP/NCAR reanalyses of winds and stream function for the period 1958 to 1993, in conjunction with independent sea surface temperature, stream-flow, precipitation and other data to identify three leading modes of sea surface temperature variability affecting long-term drought over the United States.

A broad-brush assessment of the quality of the reanalyses is that the quality tends to be best at weather time scales and degrades as we go to both shorter and longer time scales. The changes in quality reflect both the changes in the observing system and the ability of the model to simulate the variability at the different time scales. At time scales of less than a day, deficiencies in model representation of the diurnal cycle, shocks associated with the insertion of observations, and an observing system that does not fully resolve the diurnal cycle combine to degrade analysis quality (*e.g.*, Higgins *et al*., 1996; Betts *et al*., 1998a). This problem contributes to errors in our estimates of seasonal and longer time averages as well. Unsurprisingly, the quality is best for weather time scales (*e.g.*, Beljaars *et al*., 2006) of one day to a week, given that the analysis systems and models used for atmospheric reanalyses so far were developed for numerical weather prediction. At interannual and longer time scales, the impact of the major atmospheric observing system changes, combined with the increasingly important connections with other components of the climate system, contribute to degrading reanalysis quality.

We emphasize here the important connections the atmosphere has to the land and ocean on seasonal and longer time scales. The assimilation systems for both these components are considerably less mature than for the atmosphere (discussed further in section 2.5). In fact, in the current generation of atmospheric reanalyses, the connection with the ocean is made by specifying sea surface temperatures from reconstructions of historical observations, and the land is represented in a very simplified form. We note that the simplified representation of the land can also contribute to deficiencies in representing the diurnal cycle, which is highly coupled to the land surface (*e.g.*, Betts *et al*., 1998b).

Model errors can have especially large impacts on quantities linked to the hydrological cycle such as atmospheric water vapor (*e.g.*, Trenberth *et al*., 2005) and major tropical circulations of relevance to understanding climate variations and change, such as the Hadley Cell (Mitas and Clement, 2006). Any bias in the model can, in fact, exacerbate spurious climate signals associated with a changing observing system. An example is a model that is consistently too dry in the lower atmosphere. Such a model may give a realistic tropical precipitation field when there are few moisture observations available to constrain the model, but that same model can produce very unrealistic rainfall when it is confronted with large amounts of water vapor information such as that coming from satellite instruments beginning in the late 1980s (Figure 2.5).

The impacts of the changing observing systems on current reanalysis products reflect the fact that little has been done to try to account for these changes. The philosophy to date has been to use all available observations in order to maximize the accuracy of the reanalysis products at any given time, while little consideration has been given to developing approaches that could ameliorate the temporal inhomogeneities over long time periods in the reanalysis products. This defect has been recognized, and efforts are now under way to carry out reanalyses with reduced observing systems that are fixed over time (*e.g.*, Compo *et al*., 2006), as well as other observing system sensitivity experiments that could help to understand if not ameliorate the impacts (*e.g.*, Bengtsson *et al*., 2004b,c; Dee, 2005; Kanamitsu and Hwang, 2006). Other efforts that can help include: model bias correction techniques (*e.g.*, Dee and da Silva, 1998; Chepurin *et al*., 2005; Danforth *et al*., 2007), improvements to our models (Grassl, 2000; Randall, 2000), and improvements to historical observations including data mining, improved quality control and further cross calibration and bias correction of observations (Schubert *et al*., 2006).

We next consider to what extent they are internally consistent. For example do they provide realistic surface fluxes that are consistent with the other components of the climate system (in particular the land and ocean), and moisture and energy budgets that are balanced?

2.3.2.3 Internal consistency and scope

One advantage of reanalysis products mentioned earlier involves the role of the model in providing internal consistency. By this we mean that the model enforces certain dynamical balances on the reanalysis fields that are known to exist in the atmosphere. An example is the tendency for the atmosphere to be in geostrophic balance (an approximate balance of the Coriolis and pressure gradient forces) in middle latitudes. One important implication is that the different state variables (the quantities that define the state of the atmosphere – *e.g.*, the winds, temperature and pressure) cannot take on arbitrary values but instead depend strongly on each other. That such constraints are satisfied in the reanalysis products is important for many studies that attempt to understand the physical processes or forcing mechanisms by which the atmosphere evolves (*e.g.*, the various patterns of variability mentioned above).

This, in fact, is at the heart of one fundamental advantage of model-based reanalysis products over univariate analyses of, say, temperature or water vapor observations. Reanalysis products provide us at any one time with a full multivariate, globally complete picture of the atmosphere together with the various forcing functions that determine how the atmosphere evolves in time. As such, in principle we are able to diagnose all aspects of how the climate system has evolved over the time period covered by the reanalyses. There is of course a key caveat: the results depend on the quality of the model as well as characteristics of model and observational errors used in the reanalysis. As mentioned earlier, the models used in the current generation of reanalyses were largely developed for middle-latitude numerical weather prediction, and have known deficiencies, especially in various components of the hydrological cycle (clouds, precipitation, evaporation) that are critical for understanding such important phenomena as the monsoons, droughts, and various tropical phenomena.

Given that models are imperfect, can model-based reanalysis products be used to validate model simulations (see also discussion in the previous section)? For example, by forcing models with the historical record of observed sea-surface temperatures, can we reproduce some of the major precipitation anomalies of the last hundred years or so (*e.g.*, Hoerling and Kumar, 2003; Schubert *et al*., 2004; Seager *et al*., 2005; see next chapter on attribution)? As we diagnose these simulations for clues about how the climate system operates, there is an increasing need to validate the physical processes that produce the regional climate anomalies (*e.g.*, drought in the Great Plains of the United States). There is a legitimate question over whether the reanalyses used in the validations are themselves compromised by model errors. However, evidence is growing that, at least in regions with relatively good data coverage, the reanalyses can be used to identify fundamental errors in the model forcing of hydrological climate anomalies (*e.g.*, Ruiz-Barradas and Nigam, 2005).

On global scales, the deficiencies in the assimilation models manifest themselves as biases in, for example, monthly mean budgets of heat and moisture, and therefore introduce uncertainties in the physical processes that contribute to such budgets (*e.g.*, Trenberth and Guillemot, 1998; Trenberth *et al*., 2001; Kistler *et al*., 2001). While there has been some success in looking at variability of the energy budgets associated with some of the major climate variations such as ENSO (*e.g.* Trenberth *et al*., 2002a), inconsistencies in certain budgets (especially the atmospheric energy transports) limit their usefulness for estimating net surface fluxes (Trenberth and Caron, 2001) - quantities that are a crucial for linking the atmosphere and the ocean, as well as the atmosphere and land surface. Deficiencies in the model-estimated clouds (and especially the short wave radiation) appear to be a primary source of the problems in the model fluxes both at the surface and the top of the atmosphere (*e.g.*, Shinoda *et al*., 1999). Figure 2.12 shows an example of estimates of the implied ocean heat transport provided by two different reanalyses and how they compare with the values obtained from a number of different coupled atmosphere-ocean model simulations.

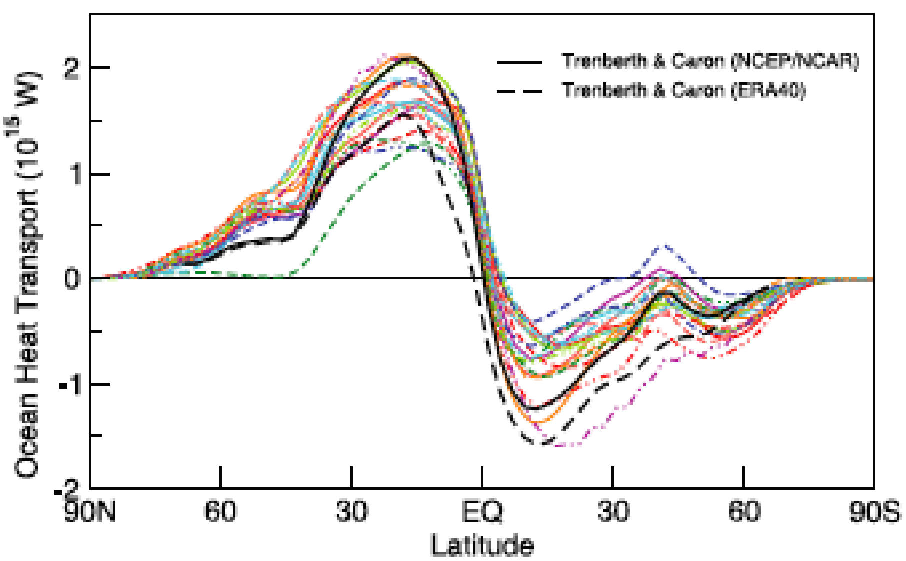

Figure 2.12 Annual mean, zonally averaged oceanic heat transport implied by net heat flux imbalances at the sea surface, under an assumption of negligible changes in oceanic heat content. The observationally based estimate, taken from Trenberth and Caron (2001) for the period February 1985 to April 1989, derives from reanalysis products from the National Centers for Environmental Prediction (NCEP)/NCAR (Kalnay *et al*., 1996) and European Centre for Medium Range Weather Forecasts 40-year reanalysis (ERA40;

Uppala *et al*., 2005). The model climatologies are derived from the years 1980 to 1999 in the 20th century simulations in the MMD at PCMDI. The legend identifying individual models appears in Figure 8.4 of the AR4 IPCC report (taken from chapter 8 of the IPCC AR4 report).

The internal consistency problem is compounded by the fact that current atmospheric reanalysis models do not satisfactorily represent interactions with other important components of the climate system (ocean, land surface, cryosphere). One result of this limitation is that the various surface fluxes (*e.g.*, precipitation, evaporation, radiation) at the interfaces between the land and atmosphere, and the ocean and atmosphere, are generally inconsistent with each other and therefore limit our ability to fully understand the forcings and interactions of the climate system (*e.g.*, Trenberth *et al*., 2001). While there are now important stand-alone land (*e.g.*, Reichle and Koster, 2005) and ocean (*e.g.*, Carton *et al*., 2000) reanalysis efforts in development or underway (see section 2.5), the long-term goal is a fully coupled climate reanalysis system (Tribbia *et al*., 2003).

2.4 TO WHAT EXTENT IS THERE AGREEMENT OR DISAGREEMENT BETWEEN CLIMATE TRENDS IN SURFACE TEMPERATURE AND PRECIPITATION DERIVED FROM REANALYSES AND THOSE DERIVED FROM INDEPENDENT DATA?

The climate of a region is defined by statistical properties of the climate system (*e.g.*, means, variances and other statistical measures) evaluated over an extended period of time, typically on the order of decades or longer. If these underlying statistical values do not change with time, the climate would be referred to as "stationary". For example, in a stationary climate a region's average monthly rainfall, say, during the 20th century would be the same as that in the 19th, 18th, or any other century (within statistical sampling errors). Climate, however, is fundamentally non-stationary; the underlying averages (and other statistical measures) do change over time. The climate system varies through ice ages and warmer periods with a timescale of about 100,000 years (Hays *et al*., 1976). The "Little Ice Age" in the 15th to 19th centuries (Bradley *et al*., 2003) is an example of a natural climate variation (an example of non-stationarity) with a much shorter timescale of a few centuries. Humans may be affecting climate even more quickly through their impact on atmospheric greenhouse gases (Hansen *et al*., 1981).

The search for trends in climatic data is, in essence, an attempt to quantify the non-stationarity of climate, as reflected in changes in long-term climate mean values. There are various methods for accomplishing this task (see CCSP SAP 1.1, Appendix 2.A for a more detailed discussion). Perhaps the most common approach to calculating a trend from a multi-decadal dataset is to plot the data value of interest (*e.g.*, rainfall) against the year of measurement. A line is fit through the points using standard regression techniques, and the resulting slope of the line is a measure of the climatic trend. A positive slope, for example, suggests that the "underlying climatic average" of rainfall is increasing with time over the period of interest. Such a trend calculation is limited by the overall noisiness of the data and by the length of the record considered.

Reanalysis datasets now span several decades, as do various ground-based and space-based measurement datasets. Trends can be computed from both. A natural question is: how well do the trends computed from the reanalysis data agree with those computed from independent datasets? This is one method for assessing the adequacy of reanalysis data for evaluating climate trends.

This question has been addressed in many independent studies. Here we focus on trends in two particular variables, surface temperature (or, more specifically, two meter height temperature, referred to here as T2M) and precipitation. Section 2.4.1 below describes the basic finding: reanalysis-based trends, though reasonable for T2M during certain periods, often do not agree with those derived from ground-based measurements. The reasons for the differences are many, as outlined in Section 2.4.2.

2.4.1. Trend Comparisons: Reanalyses Versus Independent Measurements

Simmons *et al*. (2004) provide the most comprehensive evaluation to date of reanalysis-based trends in surface temperature, T2M. Figure 2.13, reproduced from that work, shows their main result.

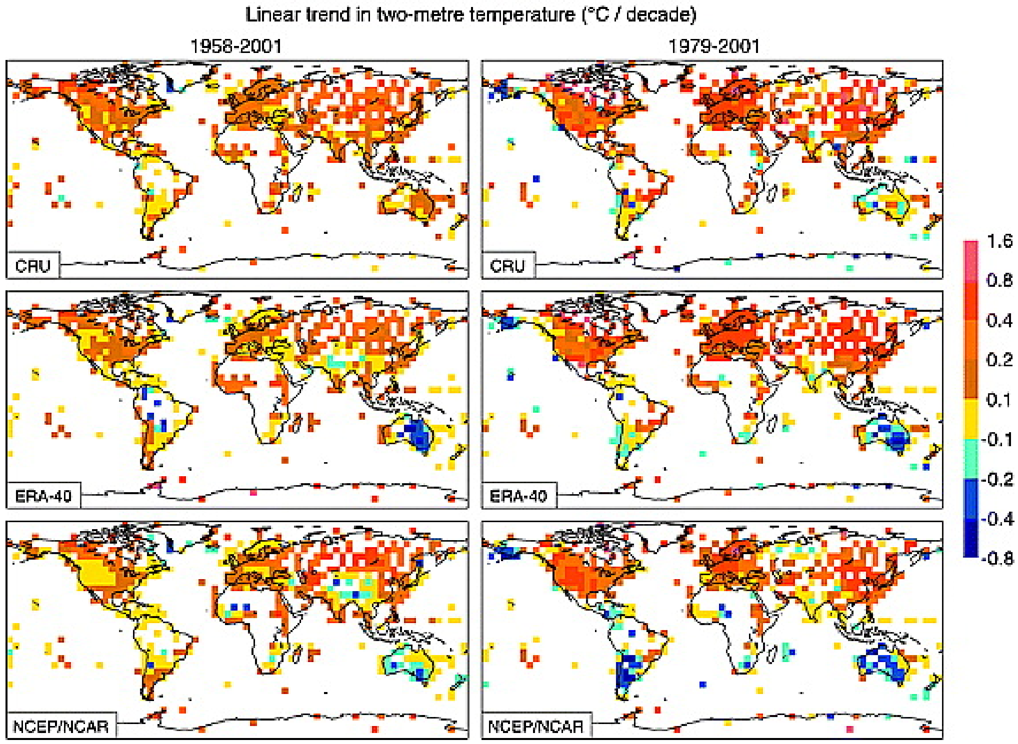

Figure 2.13. Trends in near-surface (2 meter) temperature inherent in an observational dataset (top), the ERA-40 reanalysis (middle), and the NCEP/NCAR reanalysis (bottom). Reproduced from Simmons *et al*., 2004).

Linear regression was used, as described above, to determine trends from a purely observational T2M dataset (the CRUTEM2v dataset of Jones and Moberg, 2003), from the ERA-40 reanalysis, and from the NCEP/NCAR reanalysis. Two different time periods (1958 to 2001 on the left and 1979 to 2001 on the right) were considered. All three datasets show generally positive trends. The reanalyses-based trends, however, are generally smaller, particularly for the longer time period: the average trend for 1958 to 2001 in the Northern Hemisphere, in ^{o}C per decade, is 0.19 for the observations, 0.13 for ERA-40, and 0.14 for NCEP/NCAR. For the shorter and more recent period, the Northern Hemisphere averages are 0.30 for the observations, 0.27 for ERA-40, and 0.19 for NCEP/NCAR. Simmons *et al*. (2004) consider the latter result for ERA-40 to be particularly encouraging; they emphasize "the agreement to within ~10% in the rate of warming of the terrestrial Northern Hemisphere since the late 1970s." Stendel *et al*. (2000) note that for the ERA-15 reanalysis, which covers 1979 to 1993 using an earlier version of the modeling system, the trend in T2M over North America and Eurasia is too small by 0.14^{o} C per decade, relative to observations. Thus, in terms of temperature trends, the later ERA-40 reanalysis appears to improve significantly over the earlier ERA-15 reanalysis. Note from Figure 2.13 that the performance of ERA-40 and NCEP/NCAR varies spatially, with some very clear areas of large discrepancies that most likely represent reanalysis errors. Both reanalyses, for example, underestimate trends in India and grossly underestimate them in Australia. The NCEP/NCAR reanalysis does a particularly poor job in southern South America, a problem also noted by Rusticucci and Kousky (2002).

A similarly comprehensive evaluation of precipitation trends from reanalyses has not been published. Takahashi *et al*. (2006), however, do summarize the trends in total tropical (30^{o}S – 30^{o}N) precipitation over the relatively short period of 1979 to 2001 (Figure 2.14).

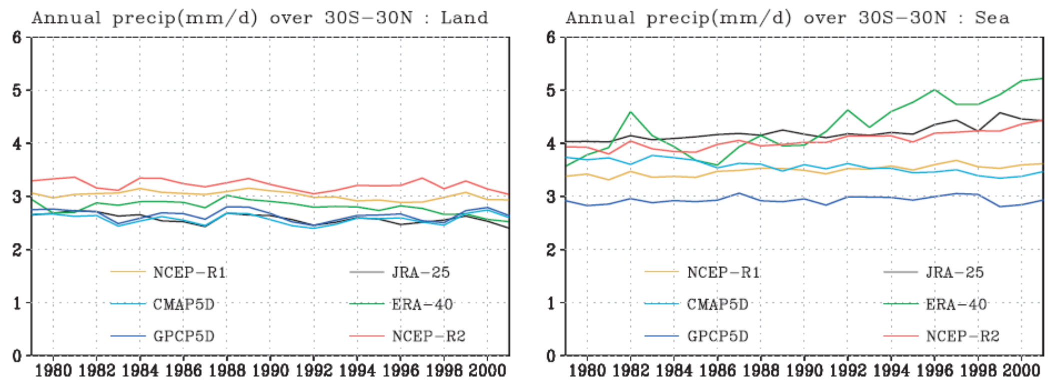

Figure 2.14. Annual tropical precipitation over land (left) and ocean (right) versus year from four reanalyses (NCEP-R1, NCEP-R2, JRA-25, and ERA-40) and from two observational datasets (CMAP5D and GCPC5D). Reprinted from Takahashi *et al*. (2006).

The biggest discrepancy between the observations and reanalyses is the large positive trend over ocean for ERA-40 and the smaller but still positive trends for the other reanalyses, trends that are not found in the observations. Similarly, Chen and Bosilovich (2007) show that the reanalyses produce a positive precipitation trend in the 1990s when global precipitation totals are considered, whereas observational datasets do not. By starting in 1979, the tropical analysis of Takahashi *et al*. (2006) misses a problem unearthed by Kinter *et al*. (2004), who demonstrate a spurious precipitation trend produced by the NCEP/NCAR reanalysis in equatorial Brazil. As shown in Figure 2.15, NCEP/NCAR produces a strong – and apparently unrealistic – increase in rainfall starting in about 1973, and thus an unrealistic wetting trend.

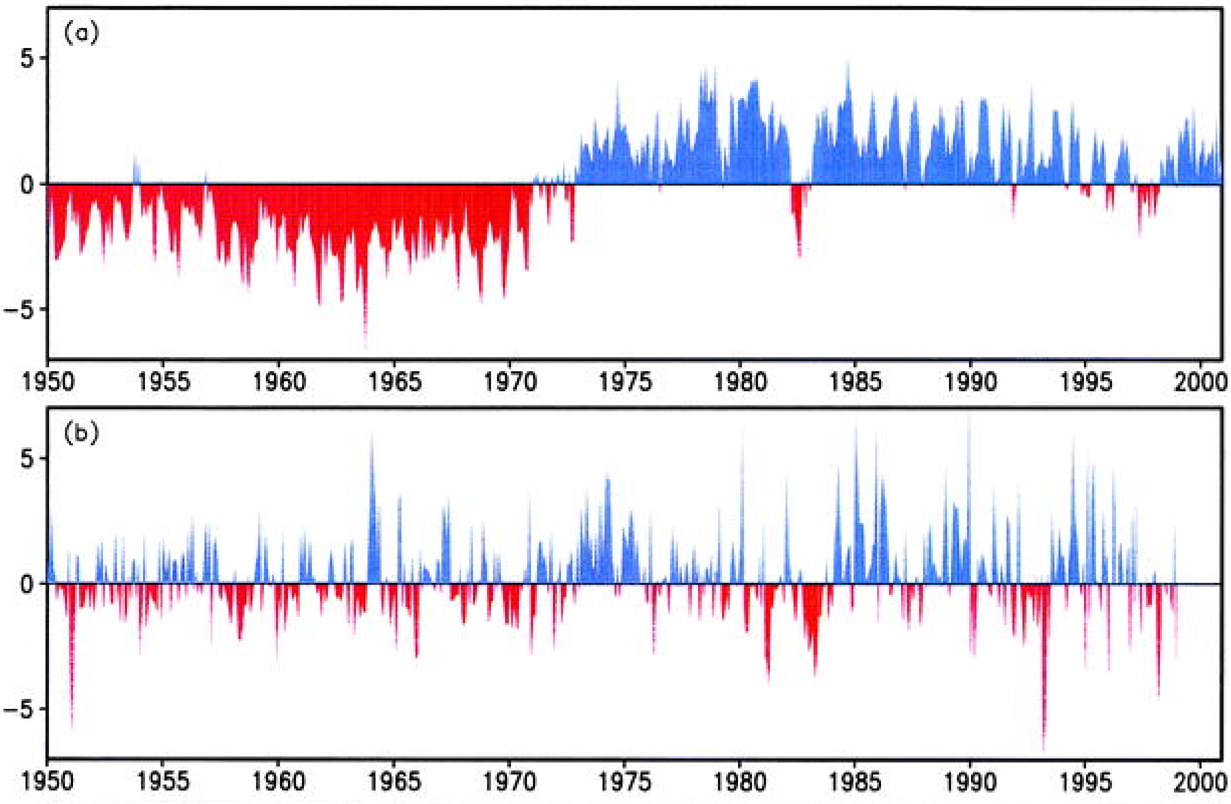

Figure 2.15. Time series of precipitation averaged over 10°S-equator, 55°-45°W, from (a) the NCAR/NCEP reanalysis, and (b) from an observational precipitation dataset. Reprinted from Kinter *et al*. (2004).

Similarly, Pohlmann and Greatbatch (2006) found that the NCEP/NCAR reanalysis greatly overestimates precipitation in northern Africa before the late 1960's but not subsequently, producing an unrealistic drying trend. Pavelsky and Smith (2006), in an analysis of river discharge to the Arctic Ocean, compared precipitation trends in the ERA-40 and NCEP/NCAR reanalyses with those from ground-based observations and found the reanalyses trends to be much too large, particularly for ERA-40. Figure 2.16 qualitatively summarizes these results.

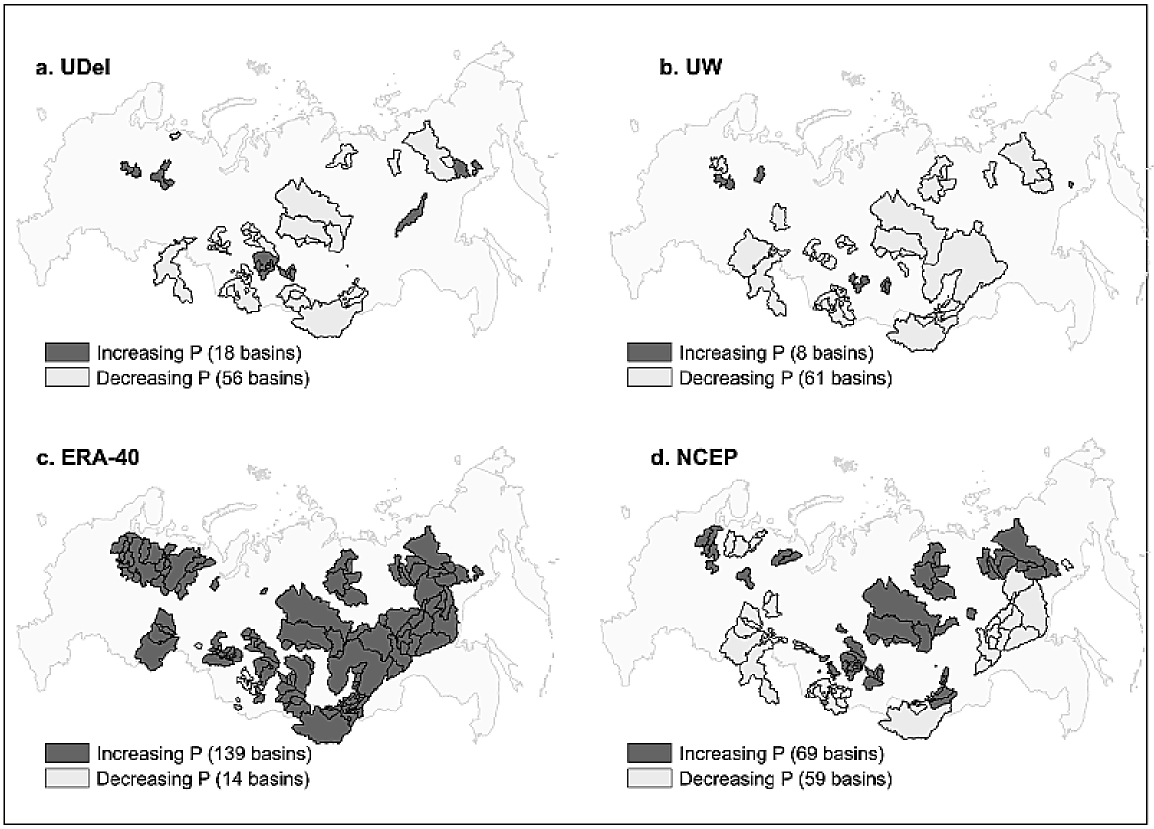

Figure 2.16. Identification of northern Asia river basins for which the computed precipitation trend is positive (a wetting trend) or negative (a drying trend), for four datasets: (top left) a dataset based on ground-based measurements of rainfall; (top right) a modified (improved) version of the first dataset; (bottom left) the ERA-40 reanalysis; and (bottom right) the NCEP/NCAR reanalysis. From Pavelsky and Smith (2006).

Identified for each dataset are the river basins with an increasing precipitation trend and those with a decreasing precipitation trend. For ERA-40, the vast majority of basins show an unrealistic (relative to ground observations) wetting trend.

2.4.2. Factors Complicating the Calculation of Trends

In summary, the previous studies indicate that observed temperature trends appear to be captured to first order by the reanalyses, particularly in the latter part of the record, though some problem areas (*e.g.*, Australia) show up clearly. Reanalysis-based precipitation trends appear to be much less consistent with those calculated from observational datasets. As described below, many studies have identified sources for errors with the reanalyses that can at least partly explain these deficiencies. It must be kept in mind, however, that trends produced from the observational datasets are themselves subject to errors for a number of reasons (see CCSP SAP 1.1, and also discussed below), so that the true deficiencies of the reanalyses-based trends cannot be wholly known.

First, and perhaps most important, a spurious trend in the reanalysis data may result from a change in the observations being assimilated. In particular, the late 1970s saw the advent of satellite data, an unprecedented increase of global-scale observations of highly variable quality. Consider now the example of a model that tends to "run cold" (has a negative temperature bias) when not constrained by data. Suppose this model is used to perform a reanalysis of the last 50 years but by necessity only ingests satellite data from the late 1970s onward. The first half of the reanalysis will be biased cold relative to the second half, leading to an artificial positive temperature trend (Figure 2.17).

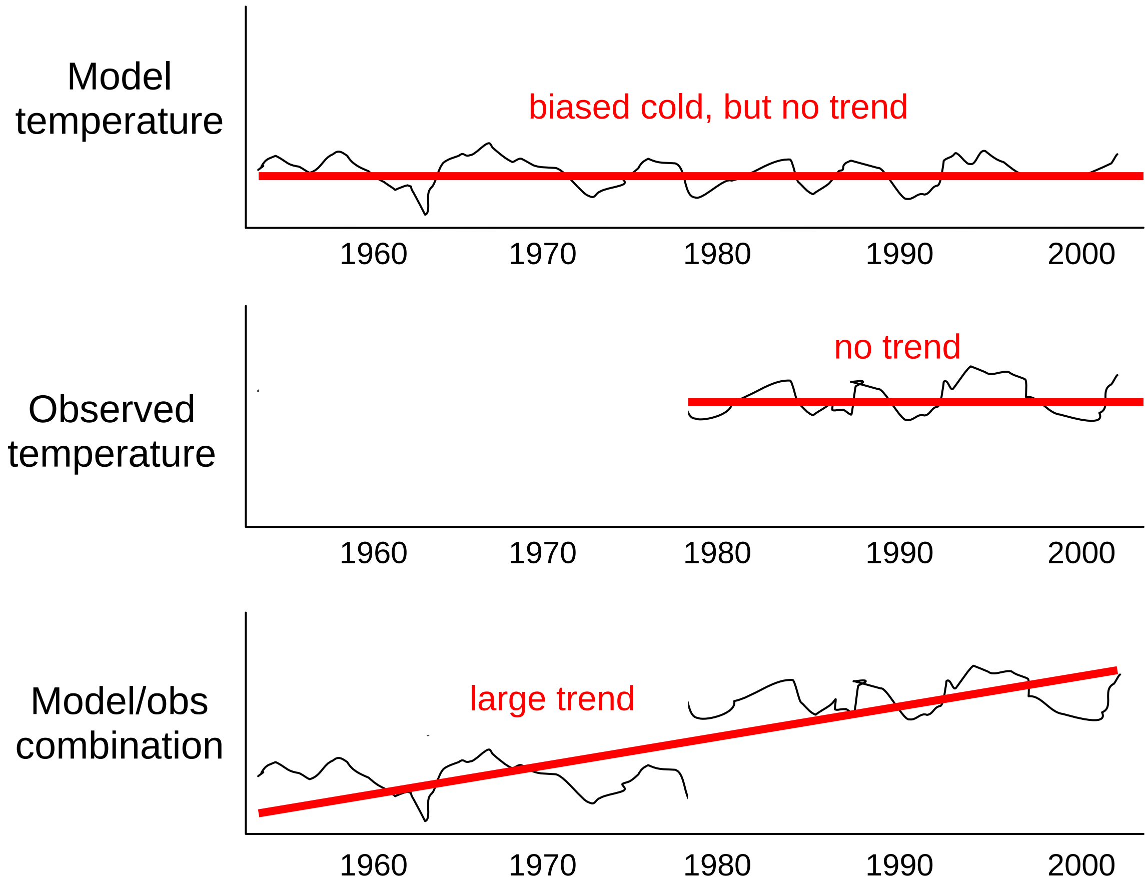

Figure 2.17 Idealized example showing how the correction of biased model data with observational data during only one part of a reanalysis period (in this case, from 1979 onward) can lead to a spurious temporal trend in the reanalysis product.

Bengtsson *et al*. (2004a) use this reasoning to explain an apparently spurious trend in lower troposphere temperature (not surface temperature) produced by the ERA40 reanalysis. Kalnay *et al*. (2006), when computing trends in surface air temperature from the NCEP/NCAR reanalysis, separate the 40-year reanalysis period into a pre-satellite and post-satellite period to avoid such issues. Note, however, that reanalyses are affected by other (non-satellite) measurement system changes as well. Betts *et al*. (2005) note in reference to the surface temperature bias over Brazil that "the Brazilian surface synoptic data are not included [in the ERA-40 reanalysis] before 1967, and with its introduction, there is a marked shift in ERA-40 from a warm to a cool bias in 2-m temperature."

Also, reanalyses that rely solely on the ingestion of atmospheric data may miss real trends in surface temperature that are associated with urbanization, cropland conversion, changing irrigation practices, and other land use changes (Pielke *et al*., 1999; Kalnay *et al*., 2006). The ERA-40 reanalysis, which does assimilate some station-based surface air temperature measurements, is less affected by this issue than the NCEP/NCAR reanalysis, which does not. This difference in station data assimilation may explain some (though not all) of ERA-40's better performance in Figure 2.13 (Simmons *et al*., 2004).

As mentioned above, calculating trends from observational datasets (the "truth" used for the evaluation of reanalysis-based trends) also involves errors, and introduces additional uncertainties when compared with reanalysis products for which values are provided on regular grids. For example, an important and challenging issue is estimating the appropriate grid-cell averaged temperature and precipitation values from point observations so that they can be directly compared with reanalysis products. Errors in representation may play a particularly important role. For example, the rain falling at one observation point may not be (and in fact, generally is not) representative of the rain falling over the corresponding model grid cell (which represents an area-average value). Rainfall measurements themselves are often sparse and distributed non-randomly, *e.g.*, in the mountainous western United States, much of the precipitation falls as snow at high mountain elevations, while most direct measurements are taken in cities and airports located at much lower elevations. Simmons *et al*. (2004) note that the gridded observational values along coastlines reflect mostly land-based measurements, whereas reanalysis values for coastal grid cells reflect a mixture of ocean and land conditions. Producing a gridded data value from multiple stations within the cell can lead to significant problems for trend estimation, since the contributing stations may have different record lengths and other spatial and temporal inhomogeneities (Hamlet and Lettenmaier, 2005). Jones *et al*. (1999) note that urbanization – urban development over time in the area of a sensor – can produce a positive temperature trend at the sensor that is quite real, but is also unrepresentative of the large grid cell that contains it.

Multi-decadal observational datasets are also strongly subject to changes in measurement systems. Takahashi *et al*. (2006) suggest that the use, starting in 1987, of a new satellite data product in an observational precipitation dataset led to a change that year in the character of the data. Kalnay *et al*. (2006) point to an artificial trend in observational temperature data induced by changes in measurement time-of-day, measurement location, and thermometer type. Jones *et al*. (1999) discuss the need, prior to computing trends, of adjusting or omitting station data as necessary to ensure a minimal impact of such changes.

Figure 2.18 gives a sense for the uncertainty inherent in trend computations from observational datasets.

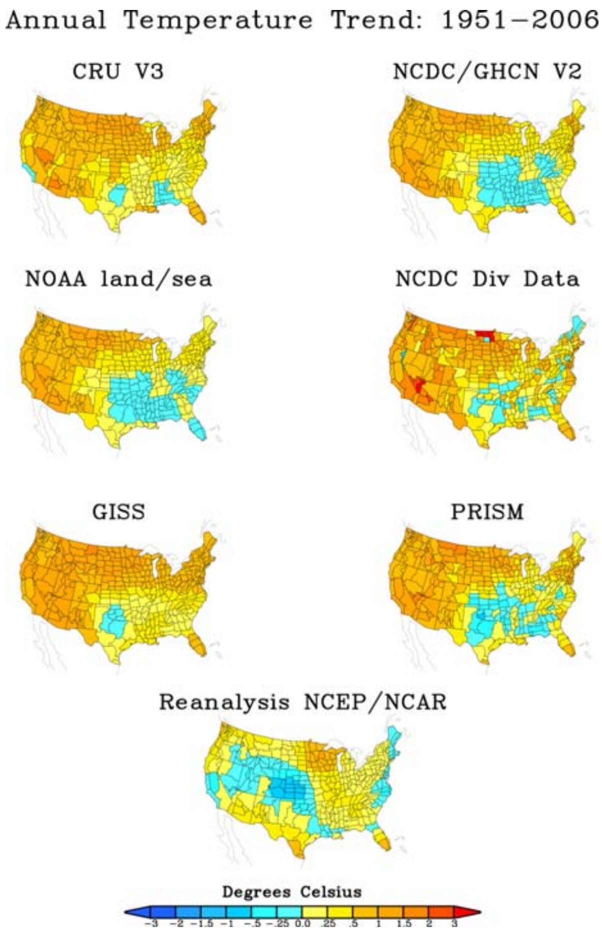

Figure 2.18. Annual temperature trends across the continental United States, as determined with six observational datasets and the NCEP/NCAR reanalysis (M. Hoerling, personal communication).

The top six maps show the spatial distributions (across the continental United States) of annual temperature trend as computed from six different datasets spanning 1951 to 2006, and the bottom map shows the trend computed from the NCEP/NCAR reanalysis. Of the seven maps, the reanalysis-derived map is clearly the outlier; the six observations-based maps all show a warming trend everywhere but in the South, whereas the reanalysis shows a general warming in the South and cooling toward the west. Even so, the six observations-based maps do not fully agree. The spatial extent of the cooling in the South is smaller in the GISS and CRU datasets than it is in the NCDC/GCHN dataset. The

NCDC climate division data show relatively high trend values in the west. Therefore it is important to recognize that we have no perfect "truth" against which to evaluate the reanalysis-based trends.

Other sources of uncertainty for both observations-based trends and reanalysis-based trends also merit mention. The mathematical algorithm used to compute the trends is important. Jones (1994a) uses the linear regression approach described above and the "robust trend method" of Hoaglin *et al*. (1983) and thereby computes two sets of trend values (similar, but not identical) from the same dataset. Also, part of the trend estimation problem is determining whether a computed trend is real, that is, the degree to which the trend is unlikely to be the result of statistical sampling variations. Groisman *et al*. (2004) describe a procedure they used to determine the statistical significance of computed trends. Even if all surface temperature data were perfect and the trend estimation technique was not an issue, the time period chosen for computing a trend can result in sampling variations, depending (for example) on the relationship to transient events such as ENSO or volcanoes (Jones, 1994b).

2.4.3. Outlook

While the above limitations hamper the accurate estimation of trends from either reanalyses or observational datasets, it is our assessment that it is likely that most of the trend differences shown in Figures 2.13 to 2.16 are related to limitations of the model-based reanalyses. Data sets that are derived directly from surface and/or satellite observations (such as those for surface air temperature, precipitation, atmospheric water vapor) will continue, at least for the near-term, to be the main tool for quantifying decadal and long-term climate changes. The observations-based trends are likely to be more trustworthy, partly because the relevant limitations in the observational data are better known and can, to a degree, be accounted for prior to trend estimation. This is less the case for existing reanalyses, which were not originally designed to be optimized for trend detection. Bengtsson *et al*. (2004a), examining various reanalysis products (though not surface temperature or precipitation), find that "there is a great deal of uncertainty in the calculation of trends from present reanalyses...". Note that reanalysis-based precipitation (for ERA-40 and NCAR/NCEP) and surface air temperature (for NCAR/NCEP) are derived solely from the models (*i.e.*, precipitation and surface temperature observations are not assimilated). Therefore, these fields are subject to inadequacies in model parameterization. The North American Regional Reanalysis is an important example of a reanalysis project that did employ the assimilation of observed precipitation data (Mesinger *et al*., 2006), producing, as a result, more realistic precipitation products.

It should be noted that reanalyses do have at least one advantage in analyzing trends. The complexity of describing and understanding trends is multi-faceted, and involves more than simply changes in mean quantities over time. Precipitation trends, for example, can be examined in the context of the "shape parameters" of precipitation probability distributions rather than total precipitation amount (Zolina *et al*., 2004). Observed precipitation trends in the United States reflect more than just an increase in the mean itself, being largely related to increases in extreme and heavy rainfall events (Karl and

Knight, 1998). Over tropical land, on the other hand, heavier rainfall events seem to be decreasing over the last 20 years, a trend that does, in fact, appear to be captured by reanalyses (Takahashi *et al*., 2006). Warming trends often reflect nighttime warming rather than warming throughout the full 24-hour day (Karl *et al*., 1991). Precipitation and temperature statistics are fundamentally tied together (Trenberth and Shea, 2005), so that precipitation and temperature trends should not be studied in isolation.

Given these (and other) examples of trend complexity, one advantage of a reanalysis dataset becomes clear: a proper analysis of the mechanisms of climate trends requires substantial data, and only a reanalysis provides self-consistent datasets that are complete in space and time over several decades. Clearly, given Figures 2.13 to 2.16, future reanalyses need to be improved to support robust trend estimation, particularly for precipitation. Climate researchers, however, may still find that for many purposes the comprehensive fields generated by reanalyses, together with their continuity (*i.e.*, none of the gaps in time that are a common feature in observational data) and spatial coverage provide value for understanding the causes of trends beyond what can be gained from observational data sets alone. For example, by providing estimates of trends in middle latitude circulation patterns and other weather elements (features that tend to have a robust signal in reanalyses – see section 2.4), reanalyses can provide insights into the nature of observed surface temperature and/or precipitation trends.

2.5 WHAT STEPS WOULD BE MOST USEFUL IN REDUCING SPURIOUS TRENDS AND OTHER MAJOR UNCERTAINTIES IN DESCRIBING THE PAST BEHAVIOR OF THE CLIMATE SYSTEM THROUGH REANALYSIS METHODS? SPECIFICALLY, WHAT CONTRIBUTIONS COULD BE MADE THROUGH IMPROVEMENTS IN DATA RECOVERY OR QUALITY CONTROL, MODELING, OR DATA ASSIMILATION TECHNIQUES?

As discussed previously, there are several reasons why our current approaches to assimilating observations for climate reanalysis can lead to spurious trends and patterns of climate variability. The instruments we use to observe the climate may contain systematic errors, and changes in the types of instruments over time may introduce false trends into the observations. Even if the instruments themselves are accurate, the spatial and temporal sampling of the instruments changes over time and thus may alias shorter time scale or smaller space scale features, or introduce spurious jumps into the climate record. In addition, the numerical models used to provide a background estimate of the system state contain systematic errors that can project onto the climate analysis. In the case of the ocean, changes in the quality of the surface meteorological forcing will be an additional source of false trends. Here we address issues of systematic instrument and data sampling errors as well as model and data assimilation errors as a backdrop for recommending improvements in the way future reanalyses are performed. Specific recommendations are given in Chapter 4.

2.5.1 Instrument and Sampling Issues

Prior to the middle of the 20th Century the atmosphere and ocean observing systems consisted mainly of surface observations of variables such as sea level pressure, winds, and surface temperature, though some upper air observations were already routinely made early in the 20th century (Brönnimann *et al*., 2005). Much of the marine surface data are already contained in the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) data set (Worley *et al*., 2005) but much also remains to be included. Considerable surface land data also exist, though these are currently scattered through several data archives, including those at the National Climatic Data Center (NCDC) and National Center for Atmospheric Research (NCAR). Many additional surface datasets remain to be digitized. The state of this surface land data should improve as various land data recovery efforts get under way (Compo *et al*., 2006). Any attempt to reconstruct climate in the first half of the 20th Century must rely on these surface observations almost exclusively and thus these data recovery efforts remain a high priority (Whitaker *et al*., 2004; Compo *et al*., 2006).

In 1936, the United States Weather Bureau began operational use of the balloon-deployed radiosonde instrument, thus providing routine soundings of atmospheric pressure, temperature, humidity, wind direction and speed for daily weather forecasts. By the International Geophysical Year of 1958 the radiosonde network expanded globally to include Antarctica and became recognized as a central component of the historical observation network that climate scientists could use to study climate. As a climate observation network, radiosondes suffer from two major types of problems. First, the instruments themselves contain systematic errors (Haimberger, 2007). For example, the widely used Vaisala radiosondes exhibit a dry bias that needs to be removed (Zipser and Johnson, 1998; Wang *et al*., 2002). Second, some radiosonde stations have moved to different locations, introducing inhomogeneities in the record (Gaffen, 1994).

Two additional observing systems were added in the 1970s. Aircraft observations increased in 1973, along with some early satellite-based temperature observations. In 1978 the number of observations increased dramatically in preparation for the First GARP Global Experiment, known as FGGE. The increase in observation coverage included three satellite-based vertical temperature sounder instruments (MSU/HIRS/SSU), cloud-tracked winds, and the expansion of aircraft observations and surface observations from ocean drifters. The impact of this increase in observations (particularly dramatic in the Southern Hemisphere) has been noted in the NCEP/NCAR and NCEP/DOE reanalyses (Kalnay *et al*., 1996; Kistler *et al*., 2001).

Currently the radiosonde network consists of about 900 stations. Most of these are still launched from continents in the Northern Hemisphere. Of these stations only ~600 launch radiosondes twice a day. Most of these launches produce profiles that extend only into the lowest levels of the stratosphere, at which height the balloons burst. A further troubling aspect of the radiosonde network is the recent closure of stations, particularly in poorly sampled Africa and the countries of the former Soviet Union.

As indicated above, the number of atmospheric observations increased dramatically in the 1970s with the introduction of remotely sensed temperature retrievals, along with a succession of ancillary measurements (*e.g.*, Figure 2.1). The temperature retrievals are made by observing the intensity of upwelling radiation in the microwave and infrared bands and then using physical models to relate these intensity measurements to a particular temperature profile. Interestingly, the problem of unknown systematic errors in the observations and the need for redundant observations has been highlighted in recent years by a false cooling trend detected in microwave tropospheric temperature retrievals. This false cooling trend has recently been corrected by properly accounting for the effects of orbital decay (Mears *et al*., 2003).

Like its atmospheric counterpart, the ocean observing system has also undergone a gradual expansion of *in situ* observations followed by a dramatic increase of satellite-based observations (Figures 2.19 and 2.20).

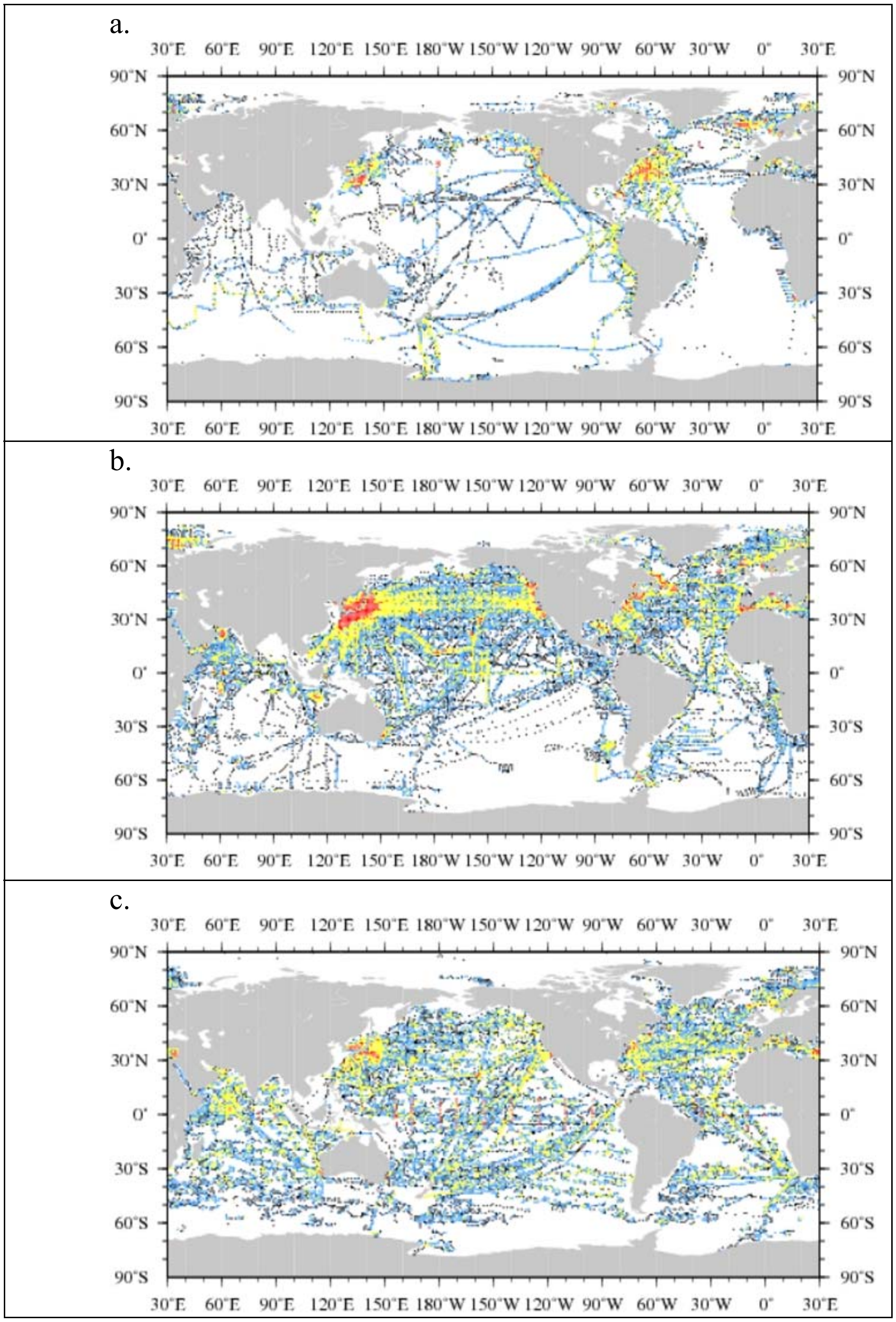

Figure 2.19 Distribution of temperature profile observations in the World Ocean Database showing 40,000 profiles extending to 150m depth for 1960 (panel a), 105,000 profiles for 1980 (panel b), and 106,000 profiles for 2004 (panel c) (http://www.nodc.noaa.gov/OC5/indprod.html).

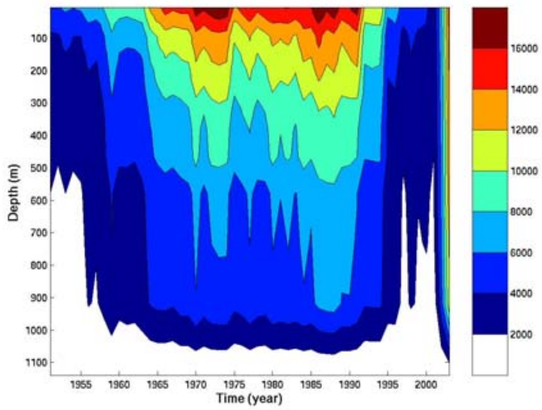

Figure 2.20 Distribution of salinity observations as a function of depth and time in the upper 1000m from the World Ocean Database 2001 (Carton and Giese, 2007). The decrease in salinity observations in 1974 resulted from the closure of the ocean weather stations, while the decrease in the mid 1990s resulted from the end of the World Ocean Circulation Experiment and the effects of the time delay in getting salinity observations into the data archives. The recent increase in salinity observations is due to the deployment of the Argo array.

Prior to 1970 the main instrument for measuring subsurface ocean temperature was the mechanical bathythermograph. This instrument was primarily deployed along trade shipping routes (Northern hemisphere) and recorded temperature only in the upper 280m, well above the oceanic thermocline at most locations. In the late 1960s the expendable bathythermograph (XBT) was introduced. In addition to being much easier to deploy, the XBT typically records temperature to a depth of 450m or 700m. Beginning in the late 1980s moored thermistor arrays have been deployed in the tropical oceans beginning with the TAO/Triton array of the tropical Pacific, but expanding into the Atlantic (PIRATA) in 1997 and most recently into the tropical Indian Ocean. These surface moorings typically measure temperature and less often salinity at fixed depths to 500m.

Two major problems have been discovered in the historical ocean temperature sampling record. The first is that much of the data were missing from the oceanographic centers.

The 1974 version of the World Ocean Atlas contained 1.5 million profiles. Thanks to great efforts by Global Oceanographic Data Archaeology and Rescue (GODAR) the latest release of the World Ocean Database (WOD2005) contains nearly 8 million profiles (Boyer *et al*., 2006). Such data archaeology and rescue work needs to be continued. A second problem arises from the fact that like its atmospheric counterpart the radiosonde, the XBT instrument was not designed for climate monitoring. XBT profiles are now known to underestimate the depth of the measurement by 1 to 2.5% of the actual depth (Hanawa *et al*., 1995). Unfortunately, the compensating drop-rate correction is different for different varieties of XBTs while less than half of the XBT observations identify the variety used. Some of the XBT observations collected since the late 1990s have already had a drop-rate correction applied without accompanying documentation, while there is evidence that the drop-rate error has changed over time, being more severe in the 1970s (AchutaRao *et al*., 2007).

For the last half of the 20th Century the main instrument for collecting deep profiles of ocean temperature as well as profiles of salinity was one or another version of the Salinity Temperature Depth or Conductivity Temperature Depth (we will refer to as the CTD) sensor. The CTD profiles are quite accurate, but are fewer in number than XBT profiles by a factor of five. As a result, diagnoses of historical changes in deep circulation must remain largely in the realm of speculation.

Since 2003 a new international observing program called Argo (Roemmich and Owens, 2000) has revolutionized ocean observation. Argo consists of a set of several thousand autonomous drifting platforms that spend most of their time at mid levels of the ocean, currently about 1000 m depth. At regular intervals, generally ten days, the Argo drifters sink and then rise to the surface, recording a profile of temperature and salinity, which is then transmitted via satellite to data archival centers. The introduction of Argo has greatly increased ocean coverage in the Southern Hemisphere and at mid-depths everywhere, and also greatly expanded the number of salinity observations. Argo is also gradually being expanded to measure variables such as Oxygen which are important for understanding the movement of greenhouse gases.

Further dramatic expansions of the ocean observing system have resulted from application of satellite remote sensing. This process began in the 1980s with the introduction of infrared and microwave sensing of sea surface temperature, followed in the early 1990s by the introduction of continuous radar observations of sea level, and then in the late 1990s with regular surface wind observations from scatterometers.

The availability of ocean data sets as well as general circulation models of the ocean has led to considerable interest in the development of ocean reanalyses (Table 2.3). The techniques being employed are rather analogous to those being employed for the atmosphere. One such example is the Simple Ocean Data Assimilation (SODA) ocean reanalysis of (Carton *et al*., 2000). Like its atmospheric counterpart, this reanalysis shows distinctly different climate variability when the massive satellite data is included.

We next turn to issues regarding the collection and interpretation of reanalysis-relevant land surface data. First, global scale *in situ* measurements of land states (soil moisture, snow, ground temperature) are essentially non-existent. Scattered measurements of soil moisture data are available in Asia (Robock *et al*., 2000), and snow measurement networks provide useful snow information in certain regions (*e.g.*, SNOTEL, <www.wcc.nrcs.usda.gov/snotel/>), but grid-scale *in situ* averages that span the globe are unavailable. Satellite data provide global coverage; however, they have their own limitations. Even the most advanced satellite-based observations can only measure soil moisture several centimeters into the soil, and not at all under dense vegetation (Entekhabi *et al*., 2004). Also, existing satellite-based estimates of surface soil moisture, as produced from different sensors and algorithms, are not consistent (Reichle *et al*., 2007), implying the need for bias correction. Time-dependent gravity measurements may provide soil moisture at deeper levels, but only at spatial scales much coarser than those needed for reanalysis (Rodell *et al*., 2007). Snow cover data from satellite are also readily available, but the estimation of total snow amount from satellite data is subject to significant uncertainty (Foster *et al*., 2005).

There are now a number of recommendations put forth by the community (*e.g.*, Schubert *et al*., 2006) to make progress on issues regarding data quality and the improvement of the world's inventories of atmospheric, ocean and land observations. These include the need for all the major data centers to prepare inventories of observations needed for reanalysis, to form collaborations that can sustain a data refresh cycle and create high quality datasets of all instruments useful for reanalyses, to develop improved record

Box 2.2 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 (Bosilovich *et al*. 2006) 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-present), and by providing the science and applications communities with of a broad range of weather and climate data with an emphasis on improved estimates of the hydrological cycle.

MERRA products consist of a host of prognostic and diagnostic fields including comprehensive sets of cloud, radiation, hydrological cycle, ozone, and land surface diagnostics. A special collection of data files are designed to facilitate off-line forcing of chemistry/aerosol models. The model or native resolution of MERRA is ⅔ degree longitude by 1/2 degree latitude with 72 levels extending to 0.01 hPa. Analysis states and 2-dimensional diagnostics will be made available at the native resolution, while many of the three-dimensional diagnostics will be made available on a coarser 1.25° latitude °—1.25° longitude grid. Further information about MERRA and its status may be found at <http://gmao.gsfc.nasa.gov/research/merra/>

tracking control for observations, and to further improve the use of feedback data from reanalyses targeted especially for data providers/developers. Furthermore, the observational, reanalysis, and climate communities should take a coordinated approach to further optimizing the usefulness of reanalysis for climate. In fact, these recommendations have now been taken up by the WCRP Observations and Assimilation Panel (WOAP).

2.5.2 Modeling and Data Assimilation Issues

Spurious trends may also be introduced into the reanalyses by systematic errors in the models used to provide background estimates for data assimilation and by incomplete modeling of those systematic errors in the data assimilation algorithm. Atmospheric models include numerical representations of the primitive equations of motion along with parameterizations of small-scale processes such as radiation, turbulent fluxes, precipitation, *etc*. Model integrations begin with some estimate of the initial state, along with boundary values of solar radiation and sea surface temperature, and are integrated forward in time. While the first generation of global reanalyses (Table 2.1) had resolutions on the order of 100 to 200 km , the latest reanalysis efforts (NASA's Modern Era Retrospective-Analysis for Research and Applications or MERRA – see Box 2.2, and NOAA's Reanalysis and Reforecasts of the NCEP Climate Forecast System or CFSRR- see Box 2.3) have horizontal resolutions of about 50 km or less. Regional models have much finer resolution, currently approaching one kilometer, and time steps of seconds. Such improvements in resolution have improved representation of physical processes such as the strength and position of the storm tracks and thus have improved simulation of climate variability and reduced model bias.

However, despite these increases in resolution, many important physical processes still cannot be explicitly resolved in current global models, such as convection, cloud formation, and precipitation of both water and ice. Thus these processes must be parameterized, or estimated from other, presumably more accurately simulated, model variables. Inaccuracies in these parameterizations are a major source of uncertainty in numerical simulation of the atmosphere and are a cause of false trends, or bias, in atmospheric models. Of course, even if the initial conditions and parameterizations were nearly perfect, the presence of atmospheric instabilities (*e.g.*, Farrell, 1989; Palmer, 1988) will inevitably lead to model forecast errors.

Ocean models also include representations of the primitive equations, with parameterizations for processes such as mixing and sea ice physics. Ocean models exchange thermodynamic, radiative and momentum fluxes with the atmosphere. Horizontal resolution of current global ocean models is approaching 10 km, in order to resolve the complex geometry of the ocean basins and the oceanic mesoscale. However, despite this fine resolution such models still exhibit systematic errors, suggesting that the small horizontal and vertical scales upon which key processes such as vertical mixing, convection, and sea ice formation are still not being resolved (Smith *et al*., 2000).

In most analyses exchanges between ocean and atmosphere are one-way in the sense that the ocean reanalysis is controlled partly by atmospheric fluxes, while the atmospheric reanalysis is controlled partly by specified sea surface temperature. Thus the fluxes in the reanalyses computed for the ocean and for the atmosphere, which should be the same are in reality inconsistent. The alternative procedure of carrying out both reanalyses in a fully coupled atmosphere/ocean model would ensure consistency. But a consequence of doing this combined analysis is that the surface exchanges are less strongly constrained and thus initial efforts at a combined analysis are found to contain considerable systematic errors in both fluids (Collins *et al*., 2006; Delworth *et al*., 2006). Correcting these systematic errors will present a major challenge for future efforts to develop consistent and accurate atmosphere/ocean reanalyses. NCEP is currently carrying out the first coupled ocean-atmosphere reanalysis, with encouraging results, but it is too early to know the extent to which the fluxes and trends are reliable (Box 2.3).

The land surface component of an atmospheric model also provides fluxes of heat, water, and radiation at the atmosphere's lower boundary. The key difficulty in producing realistic land fluxes is the tremendous amount of spatial variability (relative to that found in the atmosphere or ocean) in the properties that control these fluxes – variability, for example, in topography, vegetation character, soil type, and soil moisture content. Such variability is very difficult to deal with for two reasons. First, given the spatial resolutions used for global reanalyses (now and in the foreseeable future), we cannot properly resolve the physical processes that control the land surface fluxes, so the small-scale processes must be parameterized. Second, even if the processes could be resolved, we lack the high resolution global measurements required for many of the relevant land properties.

Despite these limitations, land models have been used in numerous Land Data Assimilation System (LDAS) projects. The current LDAS approach is to drive regional or global arrays of land surface models with observations-based meteorological forcing (precipitation, radiation, *etc.*) rather than with forcing from an atmospheric model. This allows the land models to evolve their soil moisture and temperature states to (presumably) realistic values and to produce surface moisture and heat fluxes for diagnostic studies (Figure 2.21).

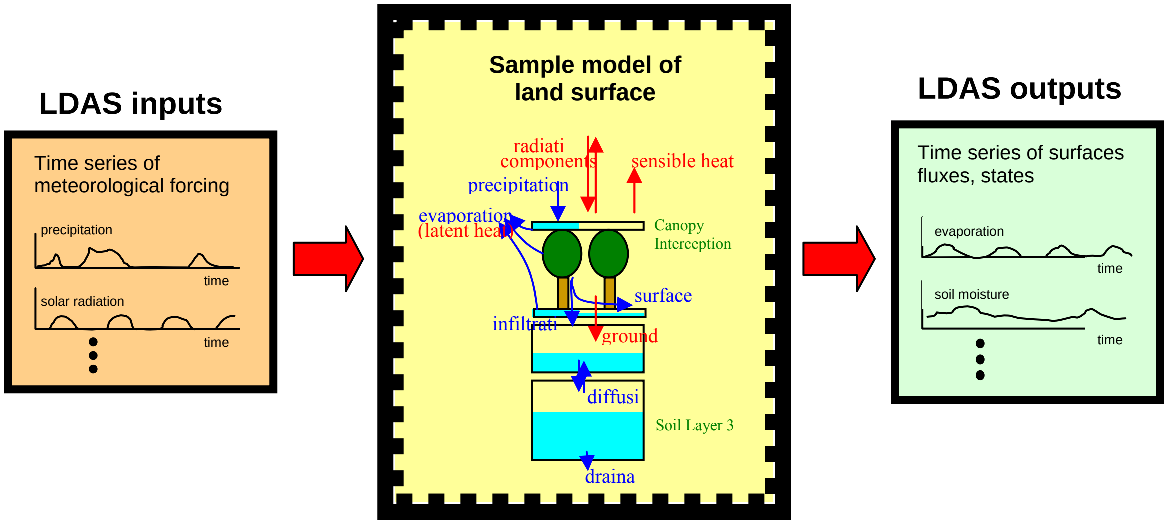

Figure 2.21. Schematic showing the inputs and outputs of a typical LDAS system.

A partial list of current LDAS projects is provided in Table 2.4. The LDAS framework is amenable to true assimilation, in which satellite- derived fields of soil moisture, snow, and temperature are incorporated into the gridded model integrations, using emerging techniques (*e.g.*, Reichle and Koster, 2005; Sun *et al*., 2004).

Table 2.4 A partial list of current Land Data Assimilation System (LDAS) projects.

Project	Sponsor(s)	Spatial Domain	Unique Aspects	Reference	Project website
GSWP-2	GEWEX	Global, 1°	Separate datasets produced by at least 15 land models for the period 1986-1995	Dirmeyer *et al*. (2006)	http://www.iges.org/gswp2/
GLDAS	NASA, NOAA	Global, .25° to ~2°	Multiple land models; near-real-time data generation	Rodell *et al.* (2004)	http://ldas.gsfc.nasa.gov/
NLDAS	Multiple Institutions	Continental U.S., , 0.125°	Multiple land models; near-real-time data	Mitchell *et al*. (2004)	http://ldas.gsfc.nasa.gov/

			generation		
ELDAS and ECMWF follow-on	European Commission	Europe, 0.2°	True data assimilation of air temperature and humidity in some versions	Van den Hurk (2002); Van den Hurk *et al*. (2008)	http://www.knmi.nl/samenw/eldas/

Data assimilation offers a general way to correct a background estimate of the state of the atmosphere, ocean, and land surface consistent with available observations (Kalnay, 2003; Wunsch, 2006). However, most current data assimilation algorithms make several assumptions for reasons of efficiency or from lack of information that limit their effectiveness. These assumptions include: 1) that any systematic trends, or biases, in the observation measurement or sampling have been identified and corrected, 2) that the forecast model is unbiased, and 3) that the error statistics such as the model forecast error have linear, Gaussian characteristics.

However, several changes can be made to ameliorate these assumptions. Systematic errors introduced by expansions of the observing system can be reduced by the procedure of repeating the reanalysis with a reduced, but more homogeneous data set, excluding for example, the satellite observations. An extreme version of this approach is to use only surface observations (Compo *et al*., 2006). In that regard, atmospheric reanalysis schemes need to make better use of historical records of surface observations from land stations and marine platforms. This includes existing climate data sets (such as daily or monthly air temperature, pressure, humidity, precipitation, and cloudiness) that have already undergone extensive quality control for the purpose of climate variability and trend applications.

Box 2.3 Climate Forecast System Reanalysis and Reforecast Project (CFSRR)

The New Reanalysis and Reforecasts of the NCEP Climate Forecast System (CFSRR) is a major upgrade to the coupled atmosphere-ocean-land Climate Forecast System (CFS). This upgrade is being planned for Jan 2010 and involves changes to all components of the CFS including, the NCEP atmospheric Gridded Statistical Interpolation Scheme (GSI), the NCEP atmospheric Global Forecast System (GFS), the NCEP Global Ocean Data Assimilation System (GODAS) including the use of the new GFDL MOM4 Ocean Model, and the NCEP Global Land Data Assimilation System (GLDAS) including the use of a new NCEP Noah Land model.

There are two essential components to this upgrade: a new reanalysis of atmosphere, ocean, land and sea ice, and a complete reforecast of the new CFS. The new reanalysis will be conducted for the 31-year period (1979-2009). The reanalysis system includes an atmosphere with high horizontal (spectral T382, ~38 Km) and vertical (64 sigma-pressure hybrid levels) resolution, an ocean with 40 levels in the vertical to a depth of 4737 m and a horizontal resolution of 0.25 degree at the tropics, tapering to a global resolution of 0.5 degree northwards and southwards of 10N and 10S respectively, an interactive sea-ice model, and an interactive land model with 4 soil levels.

In addition to the higher horizontal and vertical resolution of the atmosphere, the key differences from the previous NCEP global reanalysis are that the guess forecast will be generated from a coupled atmosphere-ocean-land-sea ice system, and that radiance measurements from the historical satellites will be assimilated.

Nearly 1 Petabyte of data will be archived from the CFSRR, which will include hourly output at the highest resolution (0.5x0.5) for 37 atmospheric levels and 40 ocean levels. More information about CFSRR can be found at: <http://cfs.ncep.noaa.gov/cfsreanl/docs>

Systematic errors in the models may be explicitly accounted for and thus (potentially) corrected in the data assimilation algorithm, which then produces an analysis of both the model state and the model bias (*e.g.*, Dee and da Silva, 1998; Danforth *et al*., 2007). However, much additional work needs to be done to improve bias modeling. In addition to estimating and reducing bias, there is also a need to improve the representation of error covariances, and ultimately provide improved estimates of the uncertainties in all reanalysis products. New techniques such as the Ensemble Kalman Filter are being developed that are both economical and able to provide such estimates (*e.g.*, Tippett *et al*., 2003; Ott *et al*., 2004).

Looking ahead, a promising pathway for improved reanalyses is the development of coupled data assimilation systems along with methods to correct for the tendency of coupled models to develop bias. In this case the observed atmosphere, ocean, and land states are assimilated jointly into the atmosphere, ocean, and land components of a fully coupled climate system model. As already mentioned, the substantial bias in current coupled models makes this a significant challenge. Nevertheless, as we continue to improve our coupled models, this joint assimilation should ensure greater consistency of model states across the components because the states would be allowed to evolve together. For example, a satellite-based correction to a soil moisture value would be able to feed back on, and thereby potentially improve, overlying atmospheric moisture and temperature states. The overall result of coupled assimilation would presumably be a more reliable, and useful, reanalysis product. There are a number of efforts that are moving towards coupled data assimilation in the United States. These are focused primarily on developing more balanced initial conditions for the seasonal and longer forecast problem, and include the Climate Forecast System Reanalysis and Reforecast (CFSRR-see Box 2.3) Project at NCEP and an ensemble-based approach being developed at GFDL (Zhang *et al*., 2007). Also, the GMAO is utilizing the MERRA product (Box 2.2) and an ocean data assimilation system to explore data assimilation in a fully coupled climate model.

CHAPTER 2 REFERENCES

AchutaRao, K.M., M. Ishii, B.D. Santer, P.J. Gleckler, K.E. Taylor, T.P. Barnett, D.W. Pierce, R.J. Stouffer, and T.M.L. Wigley, 2007: Simulated and observed variability in ocean temperature and heat content. *Proceedings of the National Academy of Sciences*, **104(26)**, 10768-10773.

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U.

Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–Present). *Journal of Hydrometeorology,* **4(6)**, 1147-1167.

Alpert, P., Y.J. Kaufman, Y. Shay-El, D. Tanre, A. da Silva, S. Schubert, and Y.H. Joseph, 1998: Quantification of dust-forced heating of the lower troposphere. *Nature*, **395(6700)**, 367-370.

Andersen, U. J., E. Kaas, and P. Alpert, 2001: Using analysis increments to estimate atmospheric heating rates following volcanic eruptions. *Geophysical Research Letters*, **28(6)**, 991-994.

Annan, J.D., J.C. Hargreaves, N.R. Edwards, and R. Marsh, 2005: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Modelling*, **8(1-2)**, 135-154.

Arkin, P.A., 1982: The relationship between interannual variability in the 200 mb tropical wind field and the Southern Oscillation. *Monthly Weather Review,* **110(10)**, 1393-1404.

Baldwin, M.P., and T.J. Dunkerton, 1999: Downward propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research,* **104(D24)**, 30937-30946.

Baldwin, M.P. and T.J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294(5542)**, 581-584.

Barlow, M., S. Nigam, and E.H. Berbery, 2001: ENSO, Pacific decadal variability, and U.S. summertime precipitation, drought, and stream flow. *Journal of Climate,* **14(9)**, 2105-2128.

Basist, A.N. and M. Chelliah, 1997: Comparison of tropospheric temperatures derived from the NCEP/NCAR Reanalysis, NCEP Operational Analysis, and the microwave sounding unit. *Bulletin of the American Meteorological Society,* **78(7)**, 1431-1447.

Beljaars, A., U. Andrae, P. Kallberg, A. Simmons, S. Uppala, and P. Viterbo, 2006: The hydrological cycle in atmospheric reanalysis. In: *Encyclopedia of Hydrological Sciences.* Part 15. Global Hydrology. [Online], John Wiley & Sons. doi:10.1002/0470848944.hsa189. Article online posting date: April 15, 2006.

Bengtsson, L., S. Hagemann, and K. I. Hodges, 2004a: Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research,* **109**, D11111, doi:10.1029/2004JD004536.

Bengtsson, L., K.I. Hodges, and S. Hagemann, 2004b: Sensitivity of large-scale atmospheric analyses to humidity observations and its impact on the global water cycle and tropical and extratropical weather systems in ERA40. *Tellus A,* **56(3)**, 202-217.

Bengtsson, L., K.I. Hodges, and S. Hagemann, 2004c: Sensitivity of the ERA40 reanalysis to the observing system: determination of the global atmospheric circulation from reduced observations. *Tellus A*, **56(3)**, 456-471.

Bengtsson, L., and J. Shukla, 1988: Integration of space and *in situ* observation to study climate change. *Bulletin of the American Meteorological Society,* **69(10)**, 1130-1143.

Betts, A.K., J.H. Ball, P. Viterbo, A. Dai, and J. Marengo, 2005: Hydrometeorology of the Amazon in ERA-40. *Journal of Hydrometeorology,* **6(5)**, 764-774.

Betts, A.K., P. Viterbo, and E. Wood, 1998a: Surface energy and water balance for the Arkansas–Red River basin from the ECMWF reanalysis. *Journal of Climate,* **11(11)**, 2881-2897.

Betts, A.K., P. Viterbo, and A.C.M. Beljaars, 1998b: Comparison of the land-surface interaction in the ECMWF reanalysis model with the 1987 FIFE data. *Monthly Weather Review,* **126(1)**, 186-198.

Blackmon, M.L., J.M. Wallace, N.-C. Lau, and S.L. Mullen 1977: An observational study of the Northern Hemisphere wintertime circulation. *Journal of the Atmospheric Sciences,* **34(7)**, 1040-1053.

Bosilovich, M.G., S.D. Schubert, M. Rienecker, R. Todling, M. Suarez, J. Bacmeister, R. Gelaro, G.-K. Kim, I. Stajner, and J. Chen, 2006: NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA). *U.S. CLIVAR Variations,* **4(2)**, 5-8.

Boyer, T.P., J.I. Antonov, H. Garcia, D.R. Johnson, R.A. Locarnini, A.V. Mishonov, M.T. Pitcher, O.K. Baranova, and I. Smolyar, 2006: *World Ocean Database 2005* [Levitus, S. (ed.)]. NOAA Atlas NESDIS 60, National Environmental Satellite, Data, and Information Service, Silver Spring, MD, 182 pp.

Bradley, R.S., M.K. Hughes, and H.F. Diaz, 2003: Climate in Medieval time. *Science,* **302(5644)**, 404-405.

Bromwich, D.H., R.L. Fogt, K.I. Hodges, and J.E. Walsh, 2007: A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *Journal of Geophysical Research,* **112**, D10111, doi:10.1029/2006JD007859.

Bromwich, D.H., and S-H. Wang, 2005: Evaluation of the NCEP-NCAR and ECMWF 15- and 40-yr reanalyses using rawinsonde data from two independent Arctic field experiments. *Monthly Weather Review,* **133(12)**, 3562-3578.

Brönnimann, S., G.P. Compo, P.D. Sardeshmukh, R. Jenne, and S. Sterin, 2005: New approaches for extending the twentieth century climate record. *Eos, Transactions of the American Geophysical Union*, **86(1)**, 2,6.

Carton, J.A., G. Chepurin, X. Cao, and B.S. Giese, 2000: A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I: Methodology. *Journal of Physical Oceanography,* **30(2)**, 294–309.

Carton, J.A. and B.S. Giese, 2007: SODA: A reanalysis of ocean climate. *Monthly Weather Review,* Accepted.

Cash, B.A. and S. Lee, 2001: Observed nonmodal growth of the Pacific-North American teleconnection pattern. *Journal of Climate,* **14(6)**, 1017–1028.

Charney, J.G., 1951: Dynamic forecasting by numerical process.In: *Compendium of* Meteorology [Malone, T.F. (ed.)]. American Meteorological Society, Boston, pp. 470-482.

Chelliah, M., and C. Ropelewski, 2000: Reanalyses-Based Tropospheric Temperature Estimates: Uncertainties in the Context of Global Climate Change Detection. *J. Climate***. 13,** 3187–3205

Chelliah, M. and G.D. Bell, 2004: Tropical multidecadal and interannual climate variability in the NCEP–NCAR reanalysis. *Journal of Climate,* **17(9)**, 1777–1803.

Chen, J. and M.G. Bosilovich, 2007: Hydrological variability and trends in global reanalyses. American Meteorological Society 19th Conference on Climate Variability and Change, San Antonio, Texas, January 2007.

Chepurin, G.A., J.A. Carton, and D. Dee, 2005: Forecast model bias correction in ocean data assimilation. *Montlhy Weather Review,* **133(5)**, 1328-1342.

Collins, W.D., V. Ramaswamy, M.D. Schwarzkopf, Y. Sun, R.W. Portmann, Q. Fu, S.E.B. Casanova, J.-L. Dufresne, D.W. Fillmore, P.M.D. Forster, V.Y. Galin, L.K. Gohar, W.J. Ingram, D.P. Kratz, M.-P. Lefebvre, J. Li, P. Marquet, V. Oinas, Y. Tsushima, T. Uchiyama, and W.Y. Zhong, 2006: Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). *Journal of Geophysical Research,* **111**, D14317, doi:10.1029/2005JD006713.

Compo G.P., P.D. Sardeshmukh, and C. Penland, 2001: Changes of subseasonal variability associated with El Niño. *Journal of Climate,* **14(16)**, 3356-3374.

Compo, G.P., J.S. Whitaker, and P.D. Sardeshmukh, 2006: Feasibility of a 100 year reanalysis using only surface pressure data. *Bulletin of the American Meteorological Society*, **87(2)**, 175–190.

Cullather, R.I., D.H. Bromwich, and M.C. Serreze, 2000: The atmospheric hydrologic cycle over the Arctic basin from reanalyses. Part I: Comparison with observations and previous studies. *Journal of Climate*, **13(5)**, 923–937.

Danforth, C.M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model errors. *Monthly Weather Review*, **135(2)**, 281-299.

Dee, D.P. and A.M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quarterly Journal of the Royal Meteorological Society*, **124**, 269-295.

Dee, D.P. and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Monthly Weather Review*, **128(9)**, 3268–3282.

Dee, D.P., 2005: Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131(613)**, 3323-3343.

Delworth, T.L., A.J. Broccoli, A. Rosati, R.J. Stouffer, V. Balaji, J.A. Beesley, W.F. Cooke, K.W. Dixon, J. Dunne, K.A. Dunne, J.W. Durachta, K.L. Findell, P. Ginoux, A. Gnanadesikan, C.T. Gordon, S.M. Griffies, R. Gudgel, M.J. Harrison, I.M. Held, R.S. Hemler, L.W. Horowitz, S.A. Klein, T.R. Knutson, P.J. Kushner, A.R. Langenhorst, H.C. Lee, S.J. Lin, J. Lu, S.L. Malyshev, P.C.D. Milly, V. Ramaswamy, J. Russell, M.D. Schwarzkopf, E. Shevliakova, J.J. Sirutis, M.J. Spelman, W.F. Stern, M. Winton, A.T. Wittenberg, B. Wyman, F. Zeng, and R. Zhang, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *Journal of Climate*, **19(5)**, 643–674.

DeWeaver, E., and S. Nigam, 2002: Linearity in ENSO's atmospheric response. *Journal of Climate*, **15(17)**, 2446-2461 **.**

Dirmeyer, P.A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2, Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, **87(10)**, 1381-1397.

Doherty, R.M., M. Hulme, and C.G. Jones, 1999: A gridded reconstruction of land and ocean precipitation for the extended tropics from 1974 to 1994. *International Journal of Climatology,* **19(2)**, 119-142.

Entekhabi, D.; E.G. Njoku, P. Houser, M. Spencer, T. Doiron, Y. Kim, J. Smith, R. Girard, S. Belair, W. Crow, T.J. Jackson, Y.H. Kerr, J.S. Kimball, R. Koster, K.C. McDonald, P.E. O'Neill, T. Pultz, S.W. Running, J. Shi; E. Wood, J. van Zyl, 2004: The Hydrosphere State (Hydros) satellite mission: an Eearth system pathfinder for global mapping of soil moisture and land freeze/thaw. *TIEEE Transactions on Geoscience and Remote Sensing*, **42(10)**, 2184-2195.

Farrell, B.F., 1989: Optimal excitation of baroclinic waves. *Journal of the Atmospheric Sciences,* **46(9)**, 1193–1206.

Feldstein, S.B., 2002: Fundamental mechanisms of the growth and decay of the PNA teleconnection pattern. *Quarterly Journal of the Royal Meteorological Society,* **128(581),** 775-796.

Feldstein, S.B., 2003: The dynamics of NAO teleconnection pattern growth and decay. *Quarterly Journal of the Royal Meteorological Society*, **129(589),** 901-924.

Foster, J.L., C. Sun, J.P. Walker, R. Kelly, A. Chang, J. Dong, and H. Powell, 2005: Quantifying the uncertainty in passive microwave snow water equivalent observations. *Remote Sensing of Environment,* **94(2)**, 187-203.

Gaffen D.J., 1994: Temporal inhomogeneities in radiosonde temperature records. *Journal of Geophysical Research,* **99(D2)**, 3667-3676.

Gates, W.L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bulletin of the American Meteorological Society,* **73(12)**, 1962-1970.

Gates, W.L., J.S. Boyle, C. Covey, C.G. Dease, C.M. Doutriaux, R.S. Drach, M. Fiorino, P.J. Gleckler, J.J. Hnilo, S.M. Marlais, T.J. Phillips, G.L. Potter, B.D. Santer, K.R. Sperber, K.E. Taylor, and D.N. Williams, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the American Meteorological Society,* **80(1)**, 29-55.

GEOSS [Global Earth Observation System of Systems], 2005: *The Global Climate Observation System Implementation Plan*. <http://www.earthobservations.org/>

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ECMWF Reanalysis Project Report Series: 1. ERA description. 72 pp.

Grassl, H., 2000: Status and improvements of coupled general circulation models. *Science*, **288(5473)**, 1991-1997.

Groisman, P.Ya., R.W. Knight, T.R. Karl, D.R. Easterling, B. Sun, and J.H. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from *in situ* observations. *Journal of Hydrometeorology,* **5(1)**, 64-85.

Gulev, S.K., O. Zolina, and S. Grigoriev. 2001. Extratropical cyclone variability in the Northern Hemisphere winter from the NCEP/NCAR reanalysis data. *Climate Dynamics,* **17(10)**, 795-809.

Gutzler, D.S., R.D. Rosen, D.A. Salstein, and J.P. Peixoto, 1988: Patterns of interannual variability in the Northern Hemisphere wintertime 850 mb temperature field. *Journal of Climate,* **1(10)**, 949–964.

Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics, *Journal of Climate,* **20(7)**, 1377-1403.

Hamlet, A.F. and D.P. Lettenmaier, 2005: Production of temporally consistent gridded precipitation and temperature fields for the continental United States. *Journal of Hydrometeorology,* **6(3)**, 330-336.

Hanawa, K., P. Rual, R. Bailey, A. Sy, and M. Szabados, 1995: A new depth-time equation for Sippican or TSK T-7, T-6 and T-4 expendable bathythermographs (XBT). *Deep-Sea Research Part I: Oceaographic Research Papers*, **42(8)**, 1423-1452.

Hansen, J., D. Johnson, A. Lacis, S. Lebedeff, P. Lee, D. Rind, and G. Russell, 1981: Climate impact of increasing atmospheric carbon dioxide. *Science*, **213(4511)**, 957-966.

Hansen, J.E., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *Journal of Geophysical Research,* **106(D20)**, 23947-23963.

Hays, J.D., J. Imbrie, and N.J. Shackleton, 1976: Variations in the Earth's orbit: Pacemaker of the ice ages. *Science*, **194(4270)**, 1121-1132.

Higgins, R.W., K.C. Mo, and S.D. Schubert, 1996: The moisture budget of the central United States in spring as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. *Monthly Weather Review,* **124(5)**, 939-963.

Hoaglin, D.C., F. Mosteller, and J.W. Tukey, 1983: *Understanding Robust and Exploratory Data Analysis*. Wiley and Sons, New York, 447 pp.

Hodges, K.I., B.J. Hoskins, J. Boyle, and C. Thorncroft, 2003: A comparison of recent reanalysis datasets using objective feature tracking: Storm tracks and tropical easterly waves. *Monthly Weather Review*, **131(9)**, 2012-2036.

Hoerling, M.P., and A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *Journal of Climate,* **15(16)**, 2184–2203.

Hoerling, M.P. and A. Kumar, 2003: The perfect ocean for drought. *Science*, **299(5607)**, 691-699.

Huffman, G.J., R.F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The Global Precipitation

Climatology Project (GPCP) combined precipitation dataset. *Bulletin of the American Meteorological Society,* **78(1)**, 5-20.

Hurrell, J.W., 1996: Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophysical Research Letters*, **23(6)**, 665-668.

Hurrell, J.W., Y. Kushnir, and M. Visbeck, 2001: The North Atlantic Oscillation. *Science*, **291(5504)**, 603-605.

IPCC (Intergovernmental Panel on Climate Change) 2007: *Climate Change 2007: The Physical Science Basis.* Contribution of Working Group I to the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor, and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, UK, and New York, 987 pp. Available at <http://www.ipcc.ch>

Jeuken, A.B.M., P.C. Siegmund, L.C. Heijboer, J. Feichter, and L. Bengtsson, 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *Journal of Geophysical Research*, **101(D12)**, 16939-16950.

Jones, P.D., 1994a: Hemispheric surface air temperature variations: A reanalysis and an update to 1993. *Journal of Climate,* **7(11)**, 1794-1802.

Jones, P.D., 1994b: Recent warming in global temperature series. *Geophysical Research Letters*, **21(12)**, 1149-1152.

Jones, P.D., M. New, D.E. Parker, S. Martin, and I.G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, **37(2)**, 173-199.

Jones, P.D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *Journal of Climate,* **16(2)**, 206-223.

Kaas, E., A. Guldberg, W. May, and M. Déqué, 1999: Using tendency errors to tune the parameterization of unresolved dynamical scale interactions in atmospheric general circulation models. *Tellus A*, **51(5)**, 612-629.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society,* **77(3)**, 437–471.

Kalnay, E. and M. Cai, 2003: Impact of urbanization and land-use change on climate. *Nature*, **423(6939)**, 528-531.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. New York, Cambridge University Press, 341 pp.

Kalnay, E., M. Cai, H. Li, and J. Tobin, 2006: Estimation of the impact of land-surface forcings on temperature trends in eastern United States. *Journal of Geophysical Research,* **111**, D06106, doi:10.1029/2005JD006555.

Kanamitsu, M. and S. Saha, 1996: Systematic tendency error in budget calculations. *Monthly Weather Review,* **124(6)**, 1145-1160.

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). *Bulletin of the American Meteorological Society,* **83(11)**, 1631-1643.

Kanamitsu, M., and S.-O. Hwang, 2006: Role of sea surface temperature in reanalysis. *Monthly Weather Review,* **134(2)**, 532-552.

Karl, T.R., G. Kukla, V. Razuvayev, M. Changery, R.G. Quayle, R.R. Heim, D.R. Easterling, and C.B. Fu, 1991: Global warming: evidence for asymmetric diurnal temperature change. *Geophysical Research Letters*, **18(12)**, 2253-2256.

Karl, T.R., and R.W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bulletin of the American Meteorological Society*, **79(2)**, 231-241.

Kinter, J.L., M.J. Fennessy, V. Krishnamurthy, and L. Marx, 2004: An evaluation of the apparent interdecadal shift in the tropical divergent circulation in the NCEP-NCAR reanalysis. *Journal of Climate,* **17(2)**, 349-361.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP–NCAR 50–year reanalysis: Monthly means CD–ROM and documentation. *Bulletin of the American Meteorological Society*, **82(2)**, 247–267.

Klinker, E. and P.D. Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *Journal of the Atmospheric Sciences,* **49(7),** 608–626.

Lau, N.-C., H. Tennekes, and J.M. Wallace, 1978: Maintenance of the momentum flux by transient eddies in the upper troposphere. *Journal of the Atmospheric Sciences*, **35(1)**, 139–147.

Lee, M.-I., S.D. Schubert, M.J. Suarez, I.M. Held, N.-C. Lau, J.J. Ploshay, A. Kumar, H.-K. Kim, and J.-K.E. Schemm, 2007: An analysis of the warm-season diurnal cycle over the continental United States and northern Mexico in general circulation models. *Journal of Hydrometeorology*, **8(3)**, 344–366.

Lin, J., B. Mapes, M. Zhang, and M. Newman, 2004: Stratiform precipitation, vertical heating profiles, and the Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, **61(3)**, 296–309.

Liu, A.Z., M. Ting, and H. Wang, 1998: Maintenance of circulation anomalies during the 1988 drought and 1993 floods over the United States. *Journal of the Atmospheric Sciences,* **55(17)**, 2810–2832.

Luo, Y., E.H. Berbery, K.E. Mitchell and A.K. Betts, 2007: Soil moisture memory and land-atmosphere interactions in the North American Regional Reanalysis. Submitted to *Journal of Hydrometeorology*.

Madden, R.A., and P.R. Julian, 1994: Observations of the 40-50 day tropical oscillation: a review. *Monthly Weather Review,* **122(5)**, 814-837.

Massacand, A.C. and H.C. Davies, 2001: Interannual variability of the extratropical northern hemisphere and the potential vorticity wave guide. *Atmospheric Science Letters*, **2(1-4)**, 61-71.

Mears, C.A., M.C. Schabel, and F.J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. *Journal of Climate,* **16(22)**, 3650-3664.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American regional reanalysis. *Bulletin of the American Meteorological Society,* **87(3)**, 343–360.

Mitas, C.M. and A. Clement, 2006: Recent behavior of the Hadley cell and tropical thermodynamics in climate models and reanalyses. *Geophysical Research Letters*, **33**, L01810, doi:10.1029/2005GL024406.

Mitchell, K.E., D. Lohmann, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, B.A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R.W. Higgins, R.T. Pinker, J.D. Tarpley, D.P. Lettenmaier, C.H. Marshall, J.K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B.H. Ramsay, A.A. Bailey, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research,* **109**, D07S90, doi:10.1029/2003JD003823.

Mo, K.C., J.N. Paegle, and R.W. Higgins, 1997: Atmospheric processes associated with summer floods and droughts in the central United States. *Journal of Climate,* **10(12)**, 3028-3046.

Mo, K.C., 2000: Relationships between low-frequency variability in the Southern Hemisphere and sea surface temperature anomalies. *Journal of Climate,* **13(20)**, 3599-3610.

Newman, M., P.D. Sardeshmukh, and J.W. Bergman, 2000: An assessment of the NCEP, NASA, and ECMWF reanalyses over the tropical west Pacific warm pool. *Bulletin of the American Meteorological Society*, **81(1)**, 41–48.

Nigam S., C. Chung, and E. DeWeaver, 2000: ENSO diabatic heating in ECMWF and NCEP–NCAR reanalyses, and NCAR CCM3 simulation. *Journal of Climate,* **13(17)**, 3152–3171.

Nigam, S. and A. Ruiz-Barradas, 2006: Seasonal hydroclimate variability over North America in global and regional reanalyses and AMIP simulations: varied representation. *Journal of Climate,* **19(5)**, 815-837.

Onogi, K., H. Koide, M. Sakamoto, S. Kobayashi, J. Tsutsui, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, K. Kato, R. Oyama, T. Ose, S. Kadokura, K. Wada, 2005: JRA-25: Japanese 25-year re-analysis–progress and status. *Quarterly Journal of the Royal Meteorological Society*, **131(613)**, 3259-3268.

Ott, E. B.R Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil, and J.A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A*, **56(5)**, 415-428.

Palmer, T.N., 1988: Medium and extended range predictability, and the stability of the PNA mode. *Quarterly Journal of the Royal Meteorological Society,* **114(481)**, 691-713.

Pavelsky, T.M. and L.C. Smith, 2006: Intercomparison of four global precipitation data sets and their correlation with increased Eurasian river discharge to the Arctic Ocean. *Journal of Geophysical Research,* **111**, D21112, doi:10.1029/2006JD007230.

Pawson, S. and M. Fiorino, 1998a: A comparison of reanalyses in the tropical stratosphere. Part 1: thermal structure and the annual cycle. *Climate Dynamics,* **14(9)**, 631-644.

Pawson, S. and M. Fiorino, 1998b: A comparison of reanalyses in the tropical stratosphere. Part 2: the quasi-biennial oscillation. *Climate Dynamics*, **14(9)**, 645-658.

Pawson, S. and M. Fiorino, 1999: A comparison of reanalyses in the tropical stratosphere. Part 3: inclusion of the pre-satellite data era. *Climate Dynamics,* **15(4)**, 241-250.

Pawson, S., K. Kodera, K. Hamilton, T.G. Shepherd, S.R. Beagley, B.A. Boville, J.D. Farrara, T.D.A. Fairlie, A. Kitoh, W.A. Lahoz, U. Langematz, E. Manzini, D.H. Rind, A.A. Scaife, K. Shibata, P. Simon, R. Swinbank, L. Takacs, R.J. Wilson, J.A. Al-Saadi, M. Amodei, M. Chiba, L. Coy, J. de Grandpré, R.S. Eckman, M. Fiorino, W.L. Grose, H. Koide, J.N. Koshyk, D. Li, J. Lerner, J.D. Mahlman, N.A. McFarlane, C.R. Mechoso, A. Molod, A. O'Neill, R.B. Pierce, W.J. Randel, R.B. Rood, and F. Wu, 2000: The GCM–Reality Intercomparison Project for SPARC (GRIPS): Scientific issues and initial results. *Bulletin of the American Meteorological Society*, **81(4)**, 781–796.

Philander, S.G., 1990: *El Niño, La Niña, and the Southern Oscillation.* Academic Press, San Diego, CA, 293 pp.

Pielke, R.A., R.L. Walko, L.T. Steyaert, P.L. Vidale, G.E. Liston, W.A. Lyons, and T.N. Chase, 1999: The influence of anthropogenic landscape changes on weather in South Florida. *Monthly Weather Review*, **127(7)**, 1663-1673.

Pohlmann, H., and R.J. Greatbatch, 2006: Discontinuities in the late 1960's in different atmospheric data products. *Geophysical Research Letters*, **33**, L22803, doi:10.1029/2006GL027644.

Raible, C.C., 2007: On the relation between extremes of midlatitude cyclones and the atmospheric circulation using ERA40. *Geophysical Research Letters*, **34**, L07703, doi:10.1029/2006GL029084.

Randall, D., 2000: *General Circulation Model Development: Past, Present and Future,* D. Randall editor, Academic Press, 416 pp.

Reichle, R.H. and R.D. Koster, 2005: Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters*, **32**, L02404, doi:10.1029/2004GL021700.

Reichle, R.H., R.D. Koster, P. Liu, S.P.P. Mahanama, E.G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research,* **112**, D09108, doi:10.1029/2006JD008033.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, Reprinted by Dover Publications, New York, 1965, p. 46.

Robock, A., K.Y. Vinnikov, G. Srinivasan, J.K. Entin, S.E. Hollinger, N.A. Speranskaya, S. Liu, and A. Namkhai, 2000: The global soil moisture data bank. *Bulletin of the American Meteorological Society*, **81(6)**, 1281-1299.

Rodell, M., P.R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J.K. Entin, J.P. Walker, D. Lohmann, and D. Toll, 2004: The global land data assimilation system. *Bulletin of the American Meteorological Society,* **85(3)**, 381-394.

Rodell, M., J.L. Chen, H. Kato, J.S. Famiglietti, J. Nigro, and C.R. Wilson, 2007: Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, **15(1)**, 159-166.

Rodwell, M.J. and T.N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society*, **133(622)**, 129-146.

Roemmich, D. and W.B.Owens, 2000: The Argo Project: global ocean observations for understanding and prediction of climate variability. *Oceanography*, **13(2)**, 45-50.

Ruiz-Barradas, A. and S. Nigam, 2005: Warm-season rainfall variability over the U.S. Great Plains in observations, NCEP and ERA-40 reanalyses, and NCAR and NASA atmospheric model simulations. *Journal of Climate,* **18(11)**, 1808-1829.

Rusticucci M. and V.E. Kousky, 2002: A comparative study of maximum and minimum temperatures over Argentina: NCEP/NCAR reanalysis versus station data. *Journal of Climate,* **15(15)**, 2089–2101.

Santer, B.D., R. Sausen, T.M.L. Wigley, J.S. Boyle, K. AchutaRao, C. Doutriaux, J.E. Hansen, G.A. Meehl, E. Roeckner, R. Ruedy, G. Schmidt, and K.E. Taylor, 2003: Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. *Journal of Geophysical Research,* **108(D1)**, 4002, doi:10.1029/2002JD002258.

Sardeshmukh, P.D., 1993: The baroclinic χ [chi] problem and its application to the diagnosis of atmospheric heating rates. *Journal of the Atmospheric Sciences,* **50(8)**, 1099-1112.

Sardeshmukh, P.D., G.P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *Journal of Climate,* **13(24)**, 4268-4286.

Schubert, S.D., R. Rood, and J. Pfaendtner, 1993: An assimilated dataset for earth science applications. *Bulletin of the American Meteorological Society,* **74(12)**, 2331–2342.

Schubert, S. and Y. Chang, 1996: An objective method for inferring sources of model error. *Monthly Weather Review,* **124(2)**, 325-340.

Schubert S.D., M.J. Suarez, P.J. Pegion, R.D. Koster, and J.T. Bacmeister. 2004: On the cause of the 1930s dust bowl. *Science*, **303(5665)**, 1855-1859.

Schubert, S., D. Dee, S. Uppala, J. Woollen, J. Bates, and S. Worley, 2006: *Report of the Workshop on "The Development of Improved Observational Data Sets for Reanalysis: Lessons Learned and Future Directions"*. University of Maryland Conference Center, College Park, MD, 28-29 September 2005. 31 pp, 236 kb.. Available at <http://gmao.gsfc.nasa.gov/pubs/docs/Schubert273.pdf.>

Schubert, S.D., Y. Chang, M.J. Suarez, and P.J. Pegion, 2008: ENSO and wintertime extreme precipitation events over the contiguous United States. *Journal of Climate,* **21(1)**, 22–39.

Seager, R., Y. Kushnir, C. Herweijer, N. Naik, and J. Velez: 2005: Modeling of tropical forcing of persistent droughts and pluvials over western North America: 1856-2000. *Journal of Climate,* **18(19)**, 4065-4088.

Serreze, M.C., J.R. Key, J.E. Box, J.A. Maslanik, and K. Steffen, 1998: A new monthly climatology of global radiation for the Arctic and comparisons with NCEP–NCAR reanalysis and ISCCP-C2 fields. *Journal of Climate,* **11(2)**, 121–136.

Shinoda, T., H.H. Hendon, and J. Glick, 1999: Intraseasonal surface fluxes in the tropical western Pacific and Indian oceans from NCEP reanalyses. *Monthly Weather Review,* **127(5)**, 678–693.

Simmons, A.J., P.D. Jones, V. da Costa Bechtold, A.C.M. Beljaars, P.W. Kållberg, S. Saarinen, S.M. Uppala, P. Viterbo, and N. Wedi, 2004: Comparison of trends and low frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *Journal of Geophysical Research*, **109**, D24115, doi:10.1029/2004JD005306.

Simmons, A.J., 2006: Observations, assimilation and the improvement of global weather prediction - some results from operational forecasting and ERA-40. In:

Predictability of Weather and Climate [Palmer, T. and R. Hagedorn, (eds.)], Cambridge University Press, Cambridge (UK), New York, pp. 428-458.

Smith R.D., M.E. Maltrud, F.O. Bryan, M.W. Hecht, 2000: Numerical simulation of the North Atlantic Ocean at 1/10°. *Journal of Physical Oceanography,* **30(7)**, 1532-1561.

Stendel, M., J.R. Christy, and L. Bengtsson, 2000: Assessing levels of uncertainty in recent temperature time series. *Climate Dynamics*, **16(8)**, 587-601.

Straus, D.M. and J. Shukla, 2002: Does ENSO force the PNA? *Journal of Climate,* **15(17)**, 2340–2358.

Sun, C., J.P. Walker, and P.R. Houser, 2004: A methodology for snow data assimilation in a land surface model. *Journal of Geophysical Research,* **109**, D08108, doi:10.1029/2003JD003765.

Takahashi, K., N. Yamazaki, and H. Kamahori, 2006: Trends of heavy precipitation events in global observation and reanalysis datasets. *SOLA (Scientific Online Letters on the Atmosphere)*, **2**, 96-99, doi:10.2151/sola.2006-025.

Thompson, D.W. and J.M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate,* **13(5)**, 1000–1016.

Tian, B., D.E. Waliser, E.J. Fetzer, B.H. Lambrigtsen, Y. Yung, and B. Wang, 2006: Vertical moist thermodynamic structure and spatial-temporal evolution of the MJO in AIRS observations. *Journal of the Atmospheric Sciences,* **63(10)**, 2462-2485.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill, and J.S. Whitaker, 2003: Ensemble square root filters. *Monthly Weather Review,* **131(7)**, 1485-1490.

Trenberth, K.E. and J.G. Olson, 1988: An evaluation and intercomparison of global analyses from National Meteorological Center and European Centre for Medium

Range Weather Forecasting. *Bulletin of the American Meteorological Society,* **69(9)**, 1047-1057.

Trenberth, K.E. and G.W. Branstator, 1992: Issues in establishing causes of the 1988 drought over North America. *Journal of Climate,* **5(2),** 159–172.

Trenberth, K.E. and C.J. Guillemot, 1996: Physical processes involved in the 1988 drought and 1993 floods in North America. *Journal of Climate,* **9(6)**, 1288-1298.

Trenberth, K.E., G.W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperature. *Journal of Geophysical Research*, **103(C7)**, 14291-12324.

Trenberth K.E. and C.J. Guillemot, 1998: Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. *Climate Dynamics*, **14(3)**, 213–231.

Trenberth, K.E. and J.M. Caron, 2000: The Southern Oscillation revisited: Sea level pressures, surface temperatures, and precipitation. *Journal of Climate,* **13(24)**, 4358–4365.

Trenberth, K.E. and J.M. Caron, 2001: Estimates of meridional atmosphere and ocean heat transports. *Journal of Climate,* **14(16)**, 3433–3443.

Trenberth, K.E., J.M. Caron, and D.P. Stepaniak, 2001: The atmospheric energy budget and implications for surface fluxes and ocean heat transports. *Climate Dynamics*, **17(4)**, 259-276.

Trenberth, K.E., D.P. Stepaniak, and J.M. Caron, 2002a: Interannual variations in the atmospheric heat budget. *Journal of Geophysical Research,* **107(D8)**, 4066, doi:10.1029/2000JD000297.

Trenberth, K.E., T.R. Karl, and T.W. Spence, 2002b: The need for a systems approach to climate observations. *Bulletin of the American Meteorological Society,* **83(11)**, 1593–1602.

Trenberth, K.E., J. Fasullo, and L. Smith, 2005: Trends and variability in column-integrated atmospheric water vapor. *Climate Dynamics* **24(7-8),** 741-758.

Trenberth, K.E. and D.J. Shea, 2005: Relationships between precipitation and surface temperature. *Geophysical Research Letters*, **32**, L14703, doi:10.1029/2005GL022760.

Tribbia, J., M. Rienecker, E. Harrison, T. Rosati, M. Ji, 2003: *Report of the "Coupled Data Assimilation Workshop", April 21-23, 2003, Portland, Oregon*. 23 pp. Available at: <www.usclivar.org/Meeting_Files/SSC-11/CoupledDA_rept_final.pdf>

Uppala, S.M., P.W. Kållberg, A.J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J.K. Gibson, J. Haseler, A. Hernandez, G.A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R.P. Allan, E. Andersson, K. Arpe, M.A. Balmaseda, A.C.M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, J. Woollen, 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131(612),** 2961-3012.

Van den Hurk, B., 2002: European LDAS established. *GEWEX Newsletter*, **12(2)**, 9.

Van den Hurk, B., J. Ettema, and P. Viterbo, 2008: Analysis of soil moisture changes in Europe during a single growing season in a new ECMWF soil moisture assimilation system. *Journal of Hydrometeorology*, **9(1)**, 116-131.

Wallace J.M. and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review,* **109(4)**, 784–812.

Wallace J.M., 2000: North Atlantic Oscillation/Annular Mode: Two paradigms—One phenomenon. *Quarterly Journal of the Royal Meteorological Society,* **126(563)**, 791–805.

Walsh, J.E., and W.L. Chapman, 1998: Arctic cloud–radiation–temperature associations in observational data and atmospheric reanalyses. *Journal of Climate,* **11(11)**, 3030–3045.

Wang, J., H.L. Cole, D.J. Carlson, E.R. Miller, K. Beierle, A. Paukkunen, and T.K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 radiosonde – Application to TOGA COARE data. *Journal of Atmospheric and Oceanic Technology*, **19(7)**, 981–1002.

WCRP [World Climate Research Programme], 1997: *Proceedings of the first WCRP International Conference on Reanalyses. Silver Spring, Maryland, USA, 27-31 October 1997.* WCRP-104, WMO/TD-No. 876. World Meteorological Organization, Geneva (Switzerland), 1998, 461 pp.

WCRP [World Climate Research Programme], 1999: *Proceedings of the Second WCRP International Conference on Reanalyses. Wokefield Park, UK, 23-27 August 1999.* WCRP-109, WMO/TD-No. 985. World Meteorological Organization, Geneva (Switzerland), 2000, 452 pp.

Weaver, S.J. and S. Nigam, 2008: Variability of the Great Plains low-level jet: Large-scale circulation context and hydroclimate impacts. *Journal of Climate,* in press.

Whitaker, J.S., G.P. Compo, X. Wei, and T.M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Monthly Weather Review,* **132(5)**, 1190–1200.

Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, and N. Lott, 2005: ICOADS Release 2.1 data and products. *International Journal of Climatology*, **25(7)**, 823–842.

Wunsch, C., 2006: *Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications*. Cambridge University Press, Cambridge (UK), 371 pp.

Xie, P. and P.A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bulletin of the American Meteorological Society*, **78(11)**, 2539-2558.

Zhang, Y., J.M. Wallace, and D.S. Battisti 1997: ENSO-like interdecadal variability: 1900-93. *Journal of Climate*, **10(5)**, 1004-1020.

Zhang, S., M.J. Harrison, A. Rosati, and A. Wittenberg 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Monthly Weather Review*, **135(10)**, 3541-3564.

Zipser, E.J., and R.H. Johnson, 1998: Systematic errors in radiosonde humidities a global problem? In: *Proceedings of the Tenth Symposium on Meteorological Observations and Instrumentation, January 11-16, 1998, Phoenix, Arizona*. American Meteorological Society, Boston, pp. 72-73.

Zolina, O., A. Kapala, C. Simmer, and S.K. Gulev, 2004: Analysis of extreme precipitation over Europe from different reanalyses: A comparative assessment. *Global and Planetary Change*, **44(1-4)**, 129-161.

Appendix 2.A Data Assimilation

Data assimilation is an exercise in the calculation of conditional probabilities in which short model forecasts are combined with observations to best estimate the state of, for example, the atmosphere. Because of limitations in model resolution and errors associated with parameterization of unresolved physical processes, and because of the chaotic behavior of the atmosphere, the accuracy of a forecast is described by a probability distribution. Similarly, the accuracy of observations is also described by a probability distribution. In data assimilation these probability distributions are combined to form conditional probabilities, which are simplified by assuming these distributions are Gaussian. The conditional probabilities are used to create a more accurate *analysis* than can be obtained solely from either the forecasts or the observations. The same approach can be applied to the ocean, land surface, or cryosphere.

Atmospheric data assimilation proceeds through a succession of *analysis cycles* of (typically) 6 hours. At the beginning of each cycle, a 6 hour model forecast is carried out starting from initial conditions of atmospheric pressure, temperature, humidity, and winds provided by the previous analysis cycle, with observed boundary conditions such as sea surface temperature and snow cover. At the end of each cycle all available current observations are quality controlled, and the differences between the observations and the model forecast of the same variables are computed (these differences are known as observational increments or innovations). The observations may include the same variables observed with different systems (*e.g.*, winds measured from airplanes or by following the movement of clouds). They may also include observations of variables that do not directly enter the forecast such as satellite radiances, observations of which contain information about both temperature and moisture.

If the evolving probability distributions of the model forecasts and observations were known then it is possible to construct an analysis that is optimal in the sense of minimizing the expected variance of the error (difference between the analysis of a variable and its true value). In practice we do not know the probability distributions. Also, we cannot solve the computational problem of minimizing the error variance for realistically complex systems. In order to address these twin problems a number of simplifying assumptions are needed. The observational increments are generally assumed to be Gaussian. With this assumption a cost function can be constructed whose minimization, which provides us with the optimal analysis, leads to the Kalman Filter equations. A more severe assumption that the probability distribution of the forecast errors is time-independent gives rise to the widely used and simpler three dimensional variational type of data assimilation (3DVAR). Four dimensional variational data assimilation (4DVAR) is a generalization of the cost function approach that allows the forecast initial conditions (or other control variables such as diffusive parameters) to be modified based on observations within a time window.

Despite the use of simplifying assumptions, the Kalman Filter and 4DVAR approaches still lead to vastly challenging computational problems. Efforts to reduce the magnitude of the computational problems and exploit physical understanding of the physical system have led to the development of Monte Carlo approaches known as Ensemble Kalman Filter (EnKF). EnKF methods, like 4DVAR, can be posed in such a way that the analysis at a given time can be influenced by future observations as well as present and past observations. This property of time symmetry is especially desirable in reanalyses since it allows the analysis at past times to benefit to some extent from future enhancements of the observing system.

Table 2.5 Characteristics of some existing global ocean model-based reanalyses of ocean climate (extracted from: http://www.clivar.org/data/synthesis/directory.php)

CNES, Météo France, CERFACS	OPA8.2**,** 2°x2°x31Lev (~0.5°x2° tropics) ERA40 forcing	Multivariate 3D-Var (OPAVAR) for T & S profiles	1962-2001	cerfacs.fr/globc/overview.html
ECMWF	HOPE, 1°x1°x29Lev (1/3°x1° tropics)	OI	1959-2006	ecmwf.int/products/forecasts/d/charts/ocean/reanalysis/
ECCO-GODAE	MITgcm 1°x1°	4DVAR	1992-2004	www.ecco-group.org
ECCO-JPL	MITgcm and MOM4 1°x1°x50 lev	Kalman filter and RTS smoother	1993-present	ecco.jpl.nasa.gov/external/
ECCO-SIO	1°x1°	4DVAR	1992-2002	ecco.ucsd.edu
ECCO2	MITgcm, 18kmx 18kmx50Lev	Green's functions	1992-present	
ENACT consortium			1962-2006	www.ecmwf.int/research/EU_projects/ENACT/
FNMOC/GODAE				www.usgodae.org
GECCO			1950-2000	www.ecco-group.org
GFDL			1960-2006	www.gfdl.noaa.gov/
UK Met Office GloSea	GloSea OGCM 1.25°x1.25°x40Lev (0.3°x1.25°tropics) daily ERA40 fluxes with corrected precipitation	OI	1962-1998	www.metoffice.gov.uk/research/seasonal/glosea.html
NASA Goddard GMAO	Poseidon, 1/3°x5/8°	MVOI, Ensemble KF	1993-pres	gmao.gsfc.nasa.gov
INGV	OPA8.2 2°x2°x31 lev (0.5°x2° tropics) ERA40 and operational ECMWF fluxes	Reduced Order MVOI with bivariate T and S EOFs	1962-pres	
MEXT K-7	MOMv3 1°x1°x36lev NCEP2 reanalysis, ISCCP data.	4D-VAR	1990-2000	www.jamstec.go.jp/frcgc/k7-dbase2/eng/
MERCATOR-3	OPA8.2 2°x2°x31lev (~0.5° meridional at the tropics)	Singular Evolutive Extended Kalman (SEEK) filter	1993-2001	www.mercator-ocean.fr/html/systemes_ops/psy3/index_en.html
JMA MOVE/MRI.COM			1949-2005	www.mri-jma.go.jp/Dep/oc/oc.html
NOAA/NCEP GODAS	MOMv3 1°x1°x40Lev (1/3°x1° tropics) NCEP Reanalysis2	3DVAR	1980-pres	www.cpc.ncep.noaa.gov/products/GODAS/
BoM, CSIRO, POAMA	ACOM2 (based on MOM2)**,** 2°x2°x27Lev (0.5°x2° at high latitudes) ERA40	MVOI, ensemble KF	1980-2006	www.bom.gov.au/bmrc/ocean/JAFOOS/POAMA/
SODA	POP1.4, POP2.01, global ave 0.25°x0.25°x40Lev, ERA40, QuikSCAT	MVOI with evolving error covariances	1958-2005	www.atmos.umd.edu/~ocean/