Comments on "Reanalyses Suitable for Characterizing Long-Term Trends"

—D. P. Dee, E. Källén, and A. J. Simmons,
European Centre for Medium-Range Weather Forecasts,
Shinfield Park, Reading, United Kingdom

L. Haimberger
Department of Meteorology and Geophysics,
University of Vienna,
Vienna, Austria

A recent article by Thorne and Vose (2010, hereafter TV) concerns the use of reanalysis data for characterizing long-term climate trends. The article raises legitimate questions about the ability to extract accurate climate information from time-varying observational datasets by means of model-based data assimilation techniques. TV predict that current approaches adopted by producers of reanalysis data are unlikely to result in climate-quality datasets. They make several recommendations for improving reanalysis methodology, and also propose a definition of "climate-quality" that is based on robust accuracy requirements.

The purpose of this note is to explain our views on these issues and to address some of TV's specific recommendations. To derive accurate and complete representations of climate variability and trends from observations is an ambitious goal. The difficulties involved, of course, are not specific to reanalysis. Fundamentally, the climate system is incompletely and inaccurately observed; data coverage, measurement techniques, and associated uncertainties are continually changing. Any such change can generate or modulate systematic errors in estimates of climate parameters. The idea behind reanalysis is to try to combine the observations by making optimal use of all available information. This includes metadata pertaining to data quality, as well as information about the physics of the climate system that can help us interpret and compare different pieces of data. In this way it becomes possible to expose the underlying uncertainties and to reduce their impact on the representation of climate parameters.

DOI: 10.1175/2010BAMS3070.1

REANALYSIS TECHNIQUE. As TV point out, reanalysis products are now the most widely used observational datasets in the atmospheric sciences. Reanalysis uses an atmospheric general circulation model to assimilate observational data from multiple sources into a dynamically coherent dataset. The role of the model in reanalysis is often misunderstood. Models contain errors like any other source of information; however, they undeniably express useful knowledge about relevant physical processes. Applying this knowledge potentially results in a more accurate and complete rendition of the system's past behavior than can be obtained from observations alone. A sufficiently realistic model can propagate information from locally observed parameters to unobserved parameters at nearby locations, and it can transmit information forward in time as well; this is the basis for numerical weather prediction. Reanalysis adds value to the instrumental record, for example, by using the model equations to extract information about winds from local pressure observations, and to improve rainfall estimates based on satellite measurements of temperature and humidity.

Given the crucial role of the assimilating model in reanalysis, it is naturally important to use the best model available. TV suggest that the effect of model errors on trend estimates should be minimized by learning from observing system experiments (OSEs). In fact, as part of the preparations for the latest European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim; see www.ecmwf.int/research/era) numerous data assimilation experiments were conducted, specifically to configure the variational bias correction scheme, to assess the impact of four-dimensional variational data assimilation (4D-Var), to determine the best model configuration to use, and to investigate the impact of different components of the observing system. Many of these experiments extended over periods of more than a year. Cases referred to in the literature include OSEs for stratospheric sounding data from the Stratospheric Sounding Unit (SSU) and Advanced Microwave Sounding Unit-A (AMSUA) (Kobayashi et al. 2009); for rain-affected radiance data from the Special Sensor Microwave Imager (SSM/I; Geer et al. 2008); for bias-corrected temperature data from radiosondes (Haimberger et al. 2008a); for reprocessed atmospheric motion vector wind vectors from Meteosat (Delsol et al. 2008); for ozone retrievals from the Global Ozone Monitoring Experiment (GOME) (Dragani 2010a, manuscript submitted to *Quart. J. Roy. Meteor. Soc.*); and for GPS radio occultation data from various instruments (Poli et al. 2010). Many other OSEs remain undocumented; however, the ERA-Interim reanalysis also derives considerable benefit from the OSEs and other data assimilation experiments routinely conducted to support the ongoing development of the forecasting system at ECMWF (e.g., Kelly and Thépaut 2007).

TV recognize that NWP centers invest considerable effort in improving their forecast models, but they point out that this does not necessarily result in a model configuration that minimizes long-term biases caused by changes in the observing system. It is not clear, however, how practically to optimize model parameterizations for reanalysis, or even how properly to define such an objective. TV do not offer any concrete proposals in this direction. We do know that data assimilation benefits from the predictive skill inherent in a good NWP model because the vast majority of available observations relate directly to atmospheric variability at weather scales. It is difficult to imagine how one could improve the representation of climate time scales in reanalysis unless the weather scales are accurate to begin with.

CLIMATE TRENDS FROM REANALYSES. The primary aim of reanalysis has always been to construct a long-term homogeneous record of the atmosphere, contrary to statements made by TV at the outset of their paper. The obvious first step is to use a fixed model and data assimilation scheme to remove the impact of forecast system changes on the representation of climate signals. This step is necessary but clearly not sufficient; it is well known by now (e.g., Dee 2005) that changes in the observations and their error characteristics can also result in artificial shifts and spurious trends in a reanalysis. Good progress has been made recently in data assimilation to address this issue (Dee and Uppala 2009). Nevertheless, given the frequency of major changes in the observing system, it is indeed very difficult to estimate climate variability on decadal time scales, and to derive accurate trends from these estimates. It must be recognized that this applies to any type of observational analysis, simply because the information available is uncertain, incomplete, and variable.

For example, there is currently no accepted best estimate of temperature trends since 1979 in the tropical troposphere. Estimates from radiosonde measurements vary depending on station selection criteria, quality control procedures, and bias corrections applied; and similarly for estimates based on radiance data from microwave sounders [e.g., see Fig. 6 in Santer et al. (2008)]. Given the scarcity of in situ upper-air observations in the tropics, it seems obvious that uncertainties in trend estimates there can only be reduced by combining multiple sources of information. A specific example of how thermal wind balance constraints can be used to infer tropical atmospheric temperature trends from wind data is given by Allen and Sherwood (2008). This study illuminates the potential of using dynamical balancing constraints to combine different atmospheric datasets. Reanalysis performed with a model-based data assimilation system can be viewed as an advanced implementation of this idea.

TV argue that a fundamentally different approach to reanalysis is needed in order to achieve climate-quality data products. In fact, the approach followed so far has led to a great deal of progress in the ability to represent climate signals accurately and consistently. For example, Simmons et al. (2004, 2010) show that surface temperature and humidity trends from the 45-yr ECMWF Re-Analysis (ERA-40; Uppala et al. 2005) and ERA-Interim are in excellent agreement with estimates obtained from climatological land station data from the Hadley Centre, at the locations where such data are available. The reanalyses additionally provide observationally constrained values over other land areas, including the poorly observed high latitudes and parts of the African and South American continents.

At higher levels in the atmosphere, the ERA-Interim reanalysis is consistent with state-of-the-art climate temperature time series data derived from observations only. Figure 1 shows estimates of globally averaged lower-stratospheric and upper-tropospheric temperature anomalies and decadal trends for the period 1989–2009, obtained from Microwave Sounding Unit (MSU) radiance data, from radiosonde temperature observations with and without bias correction, and from ERA-Interim, which uses

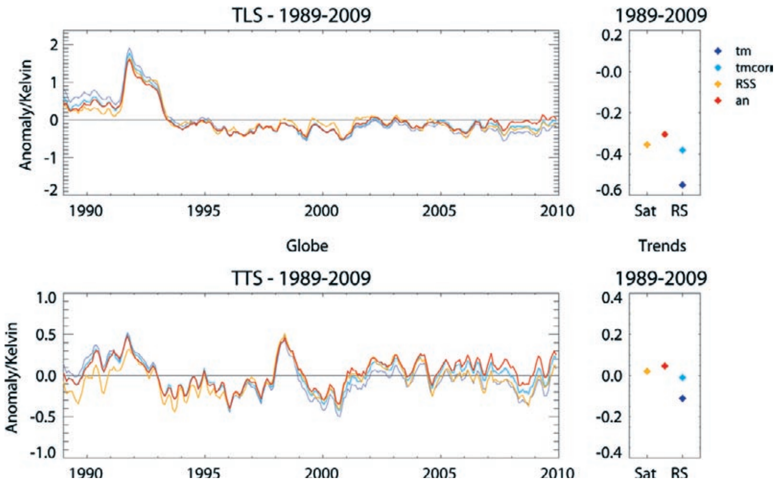

Fig. 1. MSU-equivalent global monthly temperature anomaly time series 1979–2009 for the lower stratosphere (TLS) and upper troposphere (TTS), and corresponding decadal trend estimates. Dark blue: unadjusted radiosonde data. Light blue: bias-corrected radiosonde data from RAOBCORE v1.4 (see Haimberger et al. 2008b; ERA-Interim background fields are used as reference). Orange: MSU anomaly time series from Remote Sensing Systems. Red: temperature data from ERA-Interim.

the bias-corrected radiosonde data. All estimates represent vertical averages over deep atmospheric layers, equivalent to the information contained in measurements from MSU channels 3 [Temperature Troposphere/Stratosphere (TTS)] and 4 [Temperature Lower Stratosphere (TLS)]. Spurious trends due to time-varying biases in the uncorrected radiosonde time series are no longer present in the reanalysis.

CLIMATE SYSTEM METADATA. The suggestions made by TV concerning the use of observations in reanalysis add little to what is already done in practice. The traditional approach to reanalysis is to utilize all available information, from observations and models, in an attempt to optimize the quality of the ensuing data products. In practice this information includes metadata such as prior estimates of the error characteristics for all datasets, including bias adjustments; quality indicators from data providers obtained from their own data processing algorithms; and data selection rules (so-called blacklists) based on the collective experience of many users of the data, such as numerical weather prediction centers, other producers of reanalyses, and various research groups. For the purpose of reanalysis careful prior assessments are made to ascertain whether the available observations contain usable information, how accurate this information is likely to be, and how this information can best be extracted in a data assimilation system. TV underestimate the many assessments made before, during, and after reanalysis production in monitoring the observations. The data assimilation process generates invaluable information about the quality of the input observations by integrating and assessing all data in a single, self-consistent framework. This is a major benefit of reanalysis, as it permits, for example, correction of time-varying biases in satellite radiance measurements [see Fig. 4 in Dee and Uppala (2009)] and the detection and subsequent correction of nonphysical shifts in station data (Haimberger 2007).

ESTIMATING REANALYSIS UNCERTAINTIES. As indicated by TV, there is a tension between maximum use of observations and long-term homogeneity of the climate record. However, the two aims are not irreconcilable. Indeed, the vision of reanalysis is that both can be achieved simultaneously, by systematic assessment, intercalibration, and integration of the observational record in the context of physical models. It is difficult, but not impossible, to combine changing data records without introducing nonphysical changes in the mean signal and variability. Much has been learned from the first and second generations of reanalyses in this regard. Modern data assimilation schemes are designed to cope with changes in the observing system by reconciling observations with different biases using variational methods (Dee and Uppala 2009). Some of these biases can only be detected by incorporating multiple data sources in the reanalysis. In spite of the increasingly rapid and complex changes in the observing system, modern reanalyses of the past 20 years are the most accurate and homogeneous observation-constrained datasets available. These datasets can now be used, for example, to elucidate the role of accelerated sea-ice melting in Arctic warming (Screen and Simmonds 2010). Such a study, which requires data that are both accurate and consistent, would not have been possible had the reanalysis been restricted to the use of selected long in-situ records only.

For similar reasons we do not agree with TV's recommendation to use only satellite data from long-term operational missions in reanalysis. There are numerous examples of heterogeneous data records produced by successive research missions that have been successfully integrated into global reanalyses. Ozone data from a mix of research and operational platforms have been assimilated in ERA-40 and subsequently in ERA-Interim to create an increasingly accurate and coherent global ozone record (Dragani 2010a,b, manuscripts submitted to *Quart. J. Roy. Meteor. Soc.*). Similarly, the reanalysis of surface winds from scatterometer data provided by the European Remote Sensing Satellites (*ERS-1* and *ERS-2*) and Quick Scatterometer (QuikSCAT) platforms has produced an invaluable multidecadal record that can be extended using scatterometer data from current and future operational MetOp missions.

One argument advanced by TV against the use of short-duration experimental data records is that these records often are poorly studied, and that limited resources should rather be directed at the improvement of the data record from operational missions. Yet, in discussing ways to robustly bracket the uncertainties in reanalysis data (in order to demonstrate climate quality), TV propose to use these supposedly questionable data as a validation benchmark. Clearly, if the data are of sufficient quality to be used for validation then they can be used to improve the reanalysis. Furthermore, records of short duration are useless for assessing locally the reliability of trend information from a reanalysis. It may well be reasonable in certain situations to withhold a limited number of selected observations for verification purposes, but these must be of the highest quality.

TV emphasize the necessity of using "out-of-sample" data to establish the accuracy of reanalysis output, but this is already done in reanalysis, as follows. During the reanalysis production it is customary to generate reforecasts initialized from the reanalyzed fields. Each reforecast is then verified against data that are valid at a later time; these are out-of-sample data that have not affected the forecast. The skill of the reforecasts depends on the accuracy, completeness, and physical coherence of the reanalyzed fields, in addition to the quality of the forecast model (see Fig. 2 in Dee and Uppala 2009). Similar methods are used in OSEs to assess the benefit of using specific data types in the reanalysis. For example, if a reanalysis at time t that includes Challenging Minisatellite Payload (CHAMP) data (an experimental GPS mission) improves the prediction of, say, radiosonde data at $t + 1$, then we have strong reason to include those GPS data in the reanalysis. Such an approach carries with it, of course, the possibility of a detrimental impact on unadjusted trend estimates if the newer data correct a bias present in the analyses for earlier times, but it helps ensure that the latest analyses are the best possible for quantifying the processes important for climate and for diagnosing recent climate anomalies. There is more to the climate quality of a dataset than an accurate direct depiction of trends, particularly when there is information from OSEs to quantify (and if desired adjust for) the systematic impact of changes in the observing system.

We agree with TV's suggestion that it can be useful to produce multiple reanalyses using different components of the observing system and involving different approaches to data assimilation. A good example is the ground-breaking Twentieth-Century Reanalysis Project (20CR; Compo et al. 2006), which uses only observations of sea-level pressure and sea-surface temperature data. Since satellite data are not used in this reanalysis, their introduction in the 1980s cannot affect the long-term consistency of time series of surface temperatures. However, large (but mostly unquantifiable) uncertainties remain, due to poor data coverage in large parts of the globe in early years and to significant changes in data quality throughout (e.g. Woodruff et al. 2008). Data quality control and data bias correction in this reanalysis system have been problematic given the absence of other types of observations, and away from the surface the climate trends are strongly influenced by model errors.

Nevertheless, specialized reanalysis projects such as 20CR can make a hugely important contribution to climate science because they stimulate data recovery efforts, provide a first globally integrated view of the data, and give insight into the quality and information content of those data. The use of additional observations in the reanalysis, assuming they are carefully prepared and properly assimilated, will further improve the accuracy of the reanalysis and decrease the uncertainties associated with the climate signals. TV instead suggest that a climate-quality reanalysis of the twentieth century should only make use of data from stations that have continuously observed throughout, but this would likely result in unusable trend estimates for all but a few regions of the globe.

CLIMATE-QUALITY DATASETS. The definition of "climate-quality data" proposed by TV is that uncertainties in the data must be demonstrably less than 10% of the expected multidecadal climate change signal, and this must be the case for a small set of key parameters that together encompass the climate system behavior. The range of indicators to be included in this requirement is left unspecified, but global surface temperature and large-scale precipitation are mentioned as examples. It is not entirely clear whether the 10% refers to the global average at any given time or the average over a fixed period. Relying on a definition that depends on an expected climate change signal could be problematic. In any case, our view is that such a rigid quantitative criterion is not very useful. We strongly doubt that any specific observational global data product extending over more than a few decades, whether obtained by reanalysis or otherwise, can be shown to meet it.

Instead, we would like to present an alternative vision of climate quality, which we intend to use as a guideline for future reanalysis projects. Our vision recognizes that the observational record has inherent uncertainties that are not always quantifiable. The key requirement for observational data products (such as reanalyses) to be useful for climate studies is that all available information pertaining to these uncertainties must be exposed and made accessible to the scientific community. Reanalysis products, in addition to the gridded climate parameters they usually provide, should include all observations and other datasets used as input and all metadata related to uncertainty (as discussed above), as well as qualitative guidance on the reliability of specific products. Users need to be able to assess the observational information content of specific reanalyzed parameters as a function of space and time, depending on whether those parameters have been directly observed or indirectly constrained by observations of other parameters. Making this information available to users, who ultimately set the requirements for their specific applications, allows them to draw meaningful inferences about the uncertainties in their end products (such as trend estimates). Error bars should only be provided where they can be rigorously justified. We believe that climate science is best served by providing open and complete access to all information used in producing the reanalysis data, including the observations themselves.

REFERENCES

Allen, R. J., and S. C. Sherwood, 2008: Warming maximum in the tropical upper troposphere deduced from thermal winds. *Nat. Geosci.,* **1,** 399–403.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850. *J. Geophys. Res.,* **111,** D12106, doi:10.1029/2005JD006548.

Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh, 2006: Feasibility of a 100 year reanalysis using only surface pressure data. *Bull. Amer. Meteor. Soc.,* **87,** 175–190.

Dee, D. P., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.,* **131,** 3323–3343.

——, and S. Uppala, 2009: Variational bias correction in ERA-Interim. *Quart. J. Roy. Meteor. Soc.,* **135,** 3323–3343.

Delsol, C., D. Dee, S. Uppala, N. Bormann, J.-N. Thepaut, and P. Bauer, 2008: Use of reprocessed AMVs in the ECMWF interim reanalysis. *Proc. Ninth Int. Winds Workshop*, Annapolis, MD, EUMETSAT, P.51. [Available online at www.eumetsat.int/Home/Main/AboutEUMETSAT/Publications/ConferenceandWorkshopProceedings/2008/SP_1232700640055?l=en.]

Geer, A. J., P. Bauer, and P. Lopez, 2008: Lessons learnt from the operational 1D+4D-Var assimilation of rain- and cloud-affected SSM/I observations at ECMWF. *Quart. J. Roy. Meteor. Soc.,* **134,** 1513–1525.

Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate,* **20,** 1377–1403.

——, C. Gruber, S. Sperka, and C. Tavolato, 2008a: Homogenization of the global radiosonde temperature and wind dataset using innovation statistics from reanalyses. *Proc. Third WCRP Int. Conf. on Reanalysis,* Tokyo, Japan, WCRP, G4.411. [Available online at http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/G4-411_ea.pdf.]

——, C. Tavolato, and S. Sperka, 2008b: Toward elimination of the warm bias in historic radiosonde temperature records—Some new results from a comprehensive intercomparison of upper-air data. *J. Climate,* **21,** 4587–4606.

Kelly, G. A., and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the global observation system through observing system experiments. *ECMWF Newsletter,* No. 113, ECMWF, Reading, United Kingdom, 16–28.

Poli, P., S. B. Healy, and D. P. Dee, 2010: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.,* **136,** 1972–1990.

Santer, B. D., and Coauthors, 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *Int. J. Climatol.,* **28,** 1703–1722.

Screen, J. A., and I. Simmonds, 2010: The central role of diminishing sea ice in recent arctic temperature amplification. *Nature,* **464,** 1334–1337.

Simmons, A. J., and Coauthors 2004: Comparison of trends and low-frequency variability in CRU, ERA-40 and NCEP/NCAR analyses of surface air temperature. *J. Geophys. Res.,* **109,** D24115, doi:10.1029/2004JD005306.

——, K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature and precipitation:

Inferences from reanalyses and monthly gridded observational datasets. *J. Geophys. Res.,* **115,** D01110, doi:10.1029/2009JD012442.
Thorne, P. W., and R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends: Are they really achievable? *Bull. Amer. Meteor. Soc.,* **91,** 353–361.
Uppala, S. M., and Coauthors, 2005: The ERA-40 Reanalysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.
Woodruff, S. D., H. F. Diaz, E. C. Kent, R. W. Reynolds, and S. J. Worley, 2008: The evolving SST record form ICOADS. S. Brönnimann et al., Eds., *Climate Variability and Extremes during the Past 100 Years,* Springer, 65–83.