

Reply

—P. W. THORNE

Cooperative Institute for Climate and Satellites, and
NOAA/National Climatic Data Center,
Asheville, North Carolina

—R. S. VOSE

NOAA/National Climatic Data Center,
Asheville, North Carolina

We strongly welcome the input of Dee and colleagues (Dee et al. 2010) on our recent essay piece (Thorne and Vose 2010). The original essay was deliberately written as ideas for discussion. Dee et al. make many valid points. We feel in particular that Dee et al.'s laying out of plans for the data they plan to publish along with their next major reanalysis is a positive step. Too often such plans are not publicized in this way in advance of project inception. We restrict our discussion within the following sections to those points where we either need to clarify our initial intent or wish to comment on the points raised by Dee et al. Before that, there are two general issues that are not of scientific focus that we would like to clarify.

First, our essay was not intended to question the value of reanalyses. Nothing could be further from the truth; we would not have spent considerable time and effort to prepare the original piece if we did not believe reanalyses were a fundamental component of our climate monitoring capabilities. As both we and Dee et al. pointed out, reanalyses have unique capabilities that make them very important. Dee et al. are quite right to note that both reanalyses and more classical climate data records suffer from an incomplete, ever-changing observational basis. We did not make this clear enough in our original piece.

Second, Dee et al. (2010) contest our statement that reanalyses were never primarily designed to be long-term homogeneous. For those with an interest in the historical evolution of the atmospheric sciences we clarify this here. When our essay was given as seminars at two separate institutions while under preparation, an argument that reanalyses were designed to characterize trends was made by the author (PWT). In both cases a number of members of the audience who had been involved in or had intimate knowledge of the earliest and even more recent reanalysis efforts strongly claimed this was wrong. They suggested the characterization that made it into our essay—that reanalyses were designed to give the best estimate of the historical atmospheric state at each time step rather than to be long-term homogeneous.

REANALYSIS TECHNIQUE. In our simple error model we mixed data assimilation, parameterizations, and the physical model together, terming it model error. Dee et al. are entirely correct in noting that the physical model itself cannot logically be optimized—the physics are the same regardless of the observational constraint. However, the other two aspects could make a difference. We maintain that the operational model parameterizations and the assimilation scheme will have been tuned to produce a best fit to the modern-day observational mask. For this not to be the case would require numerical weather prediction models to be deliberately suboptimal, which would be perverse. The point we wanted to make, which we believe still stands, is that when you make a substantial change to the observations constraint—such as the introduction of Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) satellite data in late 1978—you may cause a change in the model field state, particularly where there are relatively few constraints (high atmosphere, poles, ocean regions). This could be minimized through systematic optimization to the extent possible/practical of sub-gridbox parameterizations and possibly the assimilation scheme to

minimize the potential impact. Dee et al. are correct in saying that much of this happened for the interim European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim); in our defense many of these papers were published after our submission of our piece. Other reanalysis centers may also have done so, but we are not aware of systematic publications in the same manner as now undertaken for ERA-Interim documenting such work. We would encourage even greater publication of such efforts in future.

CLIMATE SYSTEM METADATA. Dee et al. are quite correct to note that much use of metadata already occurs in present-day reanalyses, and they also make the important point that the issue is pervasive in the opposite direction. That is to say, reanalysis fields and metadata produced therein offer an appealing alternative (e.g., as neighbors or some other constraint) in creating more classical datasets, as proven by the efforts of Haimberger and colleagues for radiosondes (Haimberger et al. 2008). However, we maintain that the intended core of our initial point still stands; while reanalyses use some information on data quality and provenance in decision making, they do not, regrettably, currently have access to the full suite of information (e.g., a change in radiosonde at date X at station Y or n independently derived estimates of the bias at that time in either absolute or anomaly space). This would require the collation of all information on data quality from numerous climate data records, their conversion to a common format, and their ingest/rational application in the assimilation step. We would still maintain that, if undertaken comprehensively, this effort would serve as an additional valuable constraint upon the reanalysis fields. Otherwise the reanalysis efforts are deliberately being made blind to a wealth of understanding regarding the data, some of which has been built up over many decades of painstaking research.

ESTIMATING REANALYSIS UNCERTAINTIES. We agree with Dee et al. that a range of choices can be made with regard to input data and experimental design. We concur that there is certainly substantial value to the surface-only, century-scale reanalyses (Compo et al. 2006) highlighted by Dee et al. and other systemically distinct efforts and we would encourage these.

How a reanalysis center proceeds will quite rightly depend upon its scientific best judgment, and Dee and colleagues make a strong case for their chosen approach. We would maintain that there is value to undertaking a range of approaches. However, we would maintain that our intended core point that ensembles are important remains valid and applies to any class of reanalysis. Because the reanalysis system's behavior is unlikely to be linear, single-switch Observing System Experiment (OSE)-type experiments probably give limited information on what would happen if, for example, one simultaneously perturbed three model parameterizations, took out the Microwave Sounding Unit (MSU) and/or Advanced MSU (AMSU-A) data, and used a slightly perturbed sea surface temperature (SST) boundary. In other words, the most robust way to ascertain uncertainty across a range of space and time scales is likely to be through a reasonable range of random and systematic experimentation to produce an ensemble of realizations, as has proven informative in more classical dataset construction (Titchner et al. 2009). Precisely how to design this ensemble optimally is an area where we would defer to the reanalysis centers, but such an approach would be extremely valuable and intuitive to end users.

CLIMATE QUALITY DATASETS. We agree with Dee et al. that how you assign the quality of the dataset is vexatious and difficult. We also applaud their intended distribution model, which will undoubtedly be of great benefit to the user community. We would urge consideration of coordination with other reanalysis centers to ensure they output similar metadata and the adoption of common formats for this wealth of metadata to aid in its acceptance and usage as a valuable tool.

However, their approach makes the implicit assumption that the user base is "expert" and able to

TABLE I. Citation counts for Uppala et al. (2005) and Kalnay et al. (1996) taken from ISI. The Third WCRP International Conference on reanalysis that officially cautioned against the use of NCEP in preference for newer reanalysis took place in February 2008. Clearly very substantial time has been required for most users to switch and a large number still have not.

Year	Uppala et al. citations	Kalnay et al. citations
2006	67	679
2007	194	784
2008	350	898
2009	410	517
2010 (to 7/12)	209	157

make rational decisions using the wealth of information output by the reanalysis. Our experience with serving numerous datasets is that this assumption is not valid for the majority of users. In fact, many do not go to the primary source to obtain data and do not read the papers or the websites due to time constraints. So, we would still maintain that some objective assessment is required if these users are to make sensible choices. In this regard, we undertook a simple survey of Institute for Scientific Information (ISI) of Kalnay et al. [1996; National Centers for Environmental Prediction (NCEP)] against Uppala et al. [2005; 40-yr ECMWF Re-Analysis (ERA-40)] on citation rates. It is well recognized that ERA-40 as a second-generation reanalysis building upon both ERA-15 and NCEP experiences and lessons learnt is in many senses of better quality than NCEP. Indeed, the most recent Third World Climate Research Program (WCRP) International Conference on reanalysis was explicit in this regard and made a recommendation for users to switch to the newest generation of products. Yet it is only in this current year that ERA-40 citations have started to inch ahead (Table 1). Even allowing for latency in the publication process, this implies that a substantial number of users have been using a first-generation product when a newer and arguably better product had long been available. We would maintain that without a rational decision making basis users will continue to make suboptimal decisions. Regardless of how it is achieved, some means of benchmarking reanalysis performance and suitability of use for a variety of purposes is required if reanalyses are going to be used optimally. This benchmarking needs to be simple, consistent, and transparent.

This opens up an issue not covered in our initial piece. We believe there would be substantial value in putting reanalysis and the metadata in a common format through a common mirror portal akin to the Coupled Model Intercomparison Project (CMIP) archive. A common set of benchmarking could

then be performed enabling optimal nonexpert user choices to be made while also allowing the expert community full access to relevant data and metadata. We suspect that, if adopted, the reanalyses would become even more used and successful.

ACKNOWLEDGMENTS. We thank D. Dee and A. Simmons of ECMWF for taking the time to read our response and make sure that we had understood their piece fully.

REFERENCES

- 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., E. Källén, A. J. Simmons, and L. Haimberger, 2011: Comments on “Reanalyses suitable for characterizing long-term trends.” *Bull. Amer. Meteor. Soc.*, **72**, 65–70.
- Haimberger, L., C. Tavolato, and S. Sperka, 2008: Toward elimination of the warm bias in historical radiosonde temperature records: Some new results from a comprehensive intercomparison of upper-air data. *J. Climate*, **21**, 4587–4606.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- 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.
- Titchner, H. A., P. W. Thorne, M. P. McCarthy, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2009: Critically reassessing tropospheric temperature trends from radiosondes using realistic validation experiments. *J. Climate*, **22**, 465–485.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.