AGU100 ADVANCING EARTH AND SPACE SCIENCE

Journal of Advances in Modeling Earth Systems

COMMENTARY

10.1029/2018MS001434

Key Points:
- Historical reanalyses can bridge the gap between climate and weather by providing a century-long history of the weather
- The first coupled reanalysis of the twentieth century was recently released with promising results

Correspondence to:
L. C. Slivinski,
laura.slivinski@noaa.gov

Citation:
Slivinski, L. C. (2018). Historical Reanalysis: What, how, and why? *Journal of Advances in Modeling Earth Systems*, *10*, 1736–1739. https://doi.org/10.1029/2018MS001434

Received 10 JUL 2018
Accepted 20 JUL 2018
Accepted article online 2 AUG 2018
Published online 30 AUG 2018

Historical Reanalysis: What, How, and Why?

Laura C. Slivinski[1]

[1]Physical Sciences Division, University of Colorado Cooperative Institute for Research in Environmental Sciences, NOAA Earth System Research Laboratory, Boulder, CO, USA

Abstract Historical reanalyses combine observations of past weather with simulations from modern numerical weather prediction models to provide a consistent, global history of the weather. Recently, a new reanalysis was released that allows observations of the ocean to impact the atmosphere and vice versa. "CERA-20C: A Coupled Reanalysis of the Twentieth Century" by P. Laloyaux et al. (2018, https://doi.org/10.1029/2018MS001273) describes the first coupled centennial reanalysis, thereby providing more balanced estimates of the ocean-atmosphere system and allowing for a broader range of studies of the entire Earth system. Results suggest that similar methods could also be leveraged to improve modern weather forecasts.

A paper published earlier this year in JAMES by Patrick Laloyaux and colleagues describes a new view of the history of weather from 1901 to 2010—the CERA-20C *historical reanalysis* produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). CERA-20C includes several important technical advances, including the first use of a coupled ocean-atmosphere system in a centennial reanalysis. To appreciate why such an enormous effort is worthwhile, we first need to understand the significance of historical reanalyses in general.

1. Why Study the History of Weather?

By definition, weather is a short time scale phenomenon: One rainy day can ruin plans for a barbecue; clothing choices are generally decided based only on the temperature for the next 12 hr; and it only takes a 10-minute hailstorm to regret the decision to commute via bicycle. What is the importance of knowing the weather over the entire world for the past century?

First, knowledge of historical weather allows us to put important events in human history into a larger context. Consider the Dust Bowl of the 1930s that led John Steinbeck to write *The Grapes of Wrath*, or the relatively favorable conditions on D-Day that allowed the Allies to invade Normandy with some element of surprise. Without the context of the weather, these events cannot be fully understood.

Setting a context is also crucial to study relationships between weather and climate: Putting specific events in a climatic context is central to understanding the present and predicting the future. This applies especially to extreme weather events such as floods, droughts, and hurricanes. Observational studies require a large enough sample of events to be able to control for other factors, so a centennial history of weather is necessary for these high-impact, low-frequency events.

2. How to Get a History of Weather

Constructing a history of weather requires a thoughtful blending of observations and model simulations; neither alone is up to the task. To see why, imagine trying to understand the effects of climate change on rainfall in Australia by looking in-depth at the past century of precipitation. Observations alone provide a very narrow viewpoint, limited to observed variables at specific locations with weather stations that have lasted decades. These observations will also have varying accuracy further back in time. Unconstrained model simulations, on the other hand, might provide sufficient averages of precipitation trends but will not capture particular droughts or high rainfall events.

Reanalyses combine observations with model forecasts to get the best view of the entire picture: climatic trends, individual weather events, and estimates not only of historically observed variables like precipitation, temperature, and pressure but also of, say, soil moisture, upper-atmosphere winds, and middle-atmosphere vertical velocity. Observations and model forecasts are combined using the same data assimilation methods used by weather centers around the world to generate an *analysis*, or a best guess of the current state of the global weather system, in order to initialize their weather forecasts.

A history of weather could in principle be strung together from these operational analyses, but since weather centers are continually updating their forecast models, assimilation methods, and observing systems, a long series of routine analyses contains artificial changes (Bengtsson & Shukla, 1988; Trenberth & Olson, 1988). To avoid some of these artifacts, reanalyses fix a forecast model and assimilation method and go back in time to *reanalyze* particular observation networks.

Fixing the observation network is critical as well: When reanalyses have assimilated all available observations at any given time, they have produced spurious changes when the observing system significantly changes (e.g., when satellite data are introduced.) Traditional reanalyses that assimilate upper-air observations and satellite data can therefore only extend as far back as these data are globally and consistently available: that is, 1957 for upper-air observations (Kalnay et al., 1996; Kobayashi et al., 2015) and 1979 for satellite data (Dee et al., 2011; Rienecker et al., 2011; Saha et al., 2010). To reconstruct the long records needed for climate studies and investigations of extreme events, most historical reanalyses only use conventional surface-level observations, such as surface pressure and marine winds, sacrificing a level of accuracy for consistency. These reanalyses include the NOAA-CIRES Twentieth Century Reanalysis (20CR; Compo et al., 2011) project and ECMWF's first historical reanalysis, ERA-20C (Poli et al., 2016), which each span a century or more. They have been used to investigate changes in frequency and intensity of droughts, floods, and tropical cyclones (Brooks, 2013; Emanuel, 2010; Kelley et al., 2015), to better understand El Niño/Southern Oscillation teleconnections (Deser et al., 2017), to study particular storms (Moore & Babij, 2017; Stucki et al., 2012), and even to determine the climate suitable for the TseTse fly and its relationship to African economic development (Alsan, 2015).

3. Rescuing Lost Observations

While modern observation collection is mainly automated, compiling the observations needed for a historical reanalysis still relies heavily on humans. Many early observations were recorded by hand in ship logbooks that now reside in museums, private collections, and government archives, among others. To use these observations in a reanalysis, they must first be *rescued*. Sometimes, this is done by individuals meticulously reading through cracked, yellowed journal pages in the archives of a museum basement, looking for weather observations that have not yet been included in historical reanalyses. These observations must then be digitized: In addition to international (Allan et al., 2011), university, and national meteorological service efforts, many are currently typed into databases by thousands of volunteers around the world. (See www.oldweather.org for an example.) The observations must be further processed to remove duplicate or unrealistic values and to account for small changes in location or surrounding landscape of the weather station. It can take years from the discovery of a source of historical weather observations until they are ready for assimilation into reanalysis; this substantial effort is motivated by the knowledge that many of these observations will significantly improve future reanalyses. In the early twentieth century, when the observation network is fairly sparse, a series of observations from even a single weather station or ship's journey can make an impact. As more observations are rescued, new historical reanalyses can benefit from denser observation networks than their predecessors.

4. The First Coupled Reanalysis of the Twentieth Century

Though CERA-20C benefits from the increasingly large set of historical observations, the most notable achievement is the use of a coupled ocean-atmosphere model along with a new, strongly coupled data assimilation algorithm (Laloyaux et al., 2016). This coupled system is an improvement over both atmosphere-only reanalyses and ocean-only reanalyses. Atmospheric reanalyses use prescribed sea surface temperatures and only assimilate atmospheric observations; ocean reanalyses often use a fixed near-surface atmospheric state as a boundary condition and only assimilate ocean observations. For example, when observations are assimilated into an ocean reanalysis, the atmosphere is not allowed to adapt, and inconsistencies can develop between the ocean and atmosphere. This can lead to spurious multidecadal trends in air-sea fluxes as the inconsistent ocean observations and atmospheric boundary conditions play tug-of-war on the ocean state. Conversely, a coupled system can assimilate both atmospheric and oceanic observations at once and allows each state to react and adapt to the other, leading to greater consistency between the ocean and atmosphere. CERA-20C has no spurious trend in air-sea fluxes because the two states are in balance at the ocean/atmosphere interface, instead of fighting against each other. Note, however, that removing these spurious trends does not imply removal of real climate change signals. Improved air-sea interactions also lead to more accurate estimates of variables in the boundary layer between the ocean and atmosphere, which can affect predictability of El Niño events (An, 2008; Ham & Kang, 2011; Laloyaux et al., 2018).

Since this is the first reanalysis to use strongly coupled data assimilation, though, there are still a few rough edges. Laloyaux et al. (2018) point out a drift in global ocean heat content as the ocean component of CERA-20C moves away from the initialization towards the model's preferred state; this emphasizes an inconsistency in the coupled and uncoupled models. While this is likely due to a model error, it also provides direction for focusing efforts to improve the models.

5. Estimating Confidence

Another important advance of CERA-20C over ECMWF's previous centennial reanalysis, ERA-20C, is the treatment of uncertainty. It is unfortunately impossible to perfectly recreate past weather in a reanalysis. Observations can be inexact due to measurement errors; weather forecast models have limitations due to finite resolution and inaccurate representation of true physics, for example; and there is inherent uncertainty in the dynamics of the weather system. Consider again the example of historical Australian rainfall, and suppose that the time series of regionally averaged precipitation over Australia from the reanalysis yielded a slightly negative trend over time. Does this mean that Australian precipitation is *significantly* decreasing? Without errors bars, the significance of this trend cannot be determined. Since observation networks are sparser and less accurate in, say, 1900 than 2010, the error bars in earlier time periods should generally be larger than those in recent time periods. To understand whether trends or other apparent signals are significant, it is important to be able to quantify the uncertainty of estimates from a reanalysis.

Unlike ERA-20C, CERA-20C provides an ensemble of 10 different possible realizations for each variable, at each time. Each ensemble member uses a different realization of stochastic noise within the assimilation system to represent observational and physical uncertainty. The ensemble can then be used to quantify uncertainty by examining the *spread* or how different the realizations are from each other. For example, the 10 realizations will likely be very similar to each other in 2010, when the user can be fairly confident in the ensemble average as an accurate estimate; on the other hand, the 10 realizations may differ significantly in 1900, implying less confidence. Conversely, the single estimate from ERA-20C cannot be used to measure confidence.

Laloyaux et al. (2018) provide some examples of how to interpret ensemble spread as confidence and how the uncertainty from CERA-20C changes in time, but they also discuss the shortcomings of quantifying uncertainty using only 10 realizations. Generally, larger ensembles provide more accurate uncertainty estimates. CERA-20C's relatively small ensemble is often too confident, which can lead to erroneous conclusions regarding significance.

6. The Feedback Loop of Progress

Along with CERA-20C, the forthcoming 20CR Version 3 generated by NOAA, CIRES, and the U.S. Department of Energy will add to the rather small collection of historical sparse-input reanalyses. Like CERA-20C, the existing 20CR data sets consist of ensembles of realizations, and Version 3 will provide an 80-member ensemble. The ability to compare different historical reanalyses provides a further opportunity for confidence estimation based on whether the reanalyses agree or disagree. An *ensemble of ensembles* from multiple reanalyses can be used to estimate *meta-confidence*, as well as provide interested users with opportunities for cross-comparison, for verification, and to determine the data set most relevant for them.

Coupled data assimilation is still in the early stages of the field, and CERA-20C provides an important step towards using this method in broader contexts, such as operational forecast systems that use satellite data. In an operational setup, the output from a data assimilation system is used to initialize forecasts. Consistent analyses from a coupled system may provide better initial conditions for, and thus improve, modern weather forecasts. In the context of historical reanalysis, a coupled system can also allow for more detailed studies: For instance, a coupled reanalysis should capture the cold sea surface temperature in the wake of a tropical cyclone, while an atmosphere-only reanalysis may capture the low pressure of the cyclone itself but not the cold wake.

Besides their enormous value in providing a historical view of weather, reanalyses are part of a continuous cycle of improvement: The reanalysis data can be used to improve models by illuminating errors across temporal and spatial scales, reanalyses can motivate further data rescue, and improved models and denser observation networks can then be used to produce the next generation of more accurate reanalyses. Historical reanalyses can also provide a testing ground for novel methods, such as strongly coupled data assimilation, prior to implementation in operational weather forecast centers. Recent advances in historical reanalysis provide clearer views of the past, while illuminating pathways for future research in the fields of data rescue, observation processing, data assimilation, and Earth system modeling.

Acknowledgments
This work was partially supported by the NOAA Climate Program Office and the Physical Sciences Division of the NOAA Earth System Research Laboratory. The author would like to thank Robert Pincus, Gilbert P. Compo, and Paul Dirmeyer for discussions that improved this article. ECMWF provides access to CERA-20C and other public data sets via a dedicated data portal (http://apps.ecmwf.int/datasets). Any views or opinions expressed herein are those of the author and do not necessarily reflect the views of NOAA or the Department of Commerce.

References

Allan, R., Brohan, P., Compo, G. P., Stone, R., Luterbacher, J., & Brönnimann, S. (2011). The international atmospheric circulation reconstructions over the Earth (ACRE) initiative. *Bulletin of the American Meteorological Society*, *92*(11), 1421–1425. https://doi.org/10.1175/2011BAMS3218.1

Alsan, M. (2015). The effect of the TseTse fly on African development. *American Economic Review*, *105*(1), 382–410. https://doi.org/10.1257/aer.20130604

An, S.-I. (2008). Interannual variations of the tropical ocean instability wave and ENSO. *Journal of Climate*, *21*(15), 3680–3686. https://doi.org/10.1175/2008JCLI1701.1

Bengtsson, L., & Shukla, J. (1988). Integration of space and in situ observations to study global climate change. *Bulletin of the American Meteorological Society*, *69*(10), 1130–1143. https://doi.org/10.1175/1520-0477(1988)069<1130:IOSAIS>2.0.CO;2

Brooks, H. E. (2013). Severe thunderstorms and climate change. *Atmospheric Research*, *123*, 129–138. https://doi.org/10.1016/j.atmosres.2012.04.002

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Deser, C., Simpson, I. R., McKinnon, K. A., & Phillips, A. S. (2017). The Northern Hemisphere extratropical atmospheric circulation response to ENSO: How well do we know it and how do we evaluate models accordingly? *Journal of Climate*, *30*(13), 5059–5082. https://doi.org/10.1175/JCLI-D-16-0844.1

Emanuel, K. (2010). Tropical cyclone activity downscaled from NOAA-CIRES reanalysis, 1908–1958. *Journal of Advances in Modeling Earth Systems*, *2*, 1. https://doi.org/10.3894/JAMES.2010.2.1

Ham, Y.-G., & Kang, I.-S. (2011). Improvement of seasonal forecasts with inclusion of tropical instability waves on initial conditions. *Climate Dynamics*, *36*, 1277–1290. https://doi.org/10.1007/s00382-010-0743-0

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R., & Kushnir, Y. (2015). Climate change in the Fertile Crescent and implications of the recent Syrian drought. *Proceedings of the National Academy of Sciences*, *112*(11), 3241–3246. https://doi.org/10.1073/pnas.1421533112

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan Series II*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., & Janssen, P. (2016). A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 65–78. https://doi.org/10.1002/qj.2629

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J. R., Broennimann, S., Buizza, R., et al. (2018). CERA-20C: A coupled reanalysis of the twentieth century. *Journal of Advances in Modeling Earth Systems*, *10*, 1172–1195. https://doi.org/10.1029/2018MS001273

Moore, G. W. K., & Babij, M. (2017). Iceland's Great Frost Winter of 1917/1918 and its representation in reanalyses of the twentieth century. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 508–520. https://doi.org/10.1002/qj.2939

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1058. https://doi.org/10.1175/2010BAMS3001.1

Stucki, P., Rickli, R., Brönnimann, S., Martius, O., Wanner, H., Grebner, D., & Luterbacher, J. (2012). Weather patterns and hydro-climatological precursors of extreme floods in Switzerland since 1868. *Meteorologische Zeitschrift*, *21*(6), 531–550. https://doi.org/10.1127/0941-2948/2012/368

Trenberth, K. E., & Olson, J. G. (1988). An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts. *Bulletin of the American Meteorological Society*, *69*(9), 1047–1057. https://doi.org/10.1175/1520-0477(1988)069<1047:AEAIOG>2.0.CO;2