Developing 20th Century Reanalysis version 3 (1850-2014)

Gilbert P. Compo^{1,2}, Jeffrey S. Whitaker², Prashant D. Sardeshmukh^{1,2}, Benjamin Giese³, Philip Brohan⁴

¹Univ. of Colorado/CIRES and ²NOAA Earth System Research Laboratory/PSD ³Texas A&M, Dept. of Oceanograhpy ⁴Hadley Centre UK Met Office Quarterly Journal of the Royal Meteorological Society

Q. J. R. Meteorol. Soc. 137: 1-28, January 2011 Part A



Review Article The Twentieth Century Reanalysis Project

G. P. Compo,^{a,b*}J. S. Whitaker,^{b‡} P. D. Sardeshmukh,^{a,b} N. Matsui,^{a,b} R. J. Allan,^{c†} X. Yin,^d B. E. Gleason, Jr.,^{e‡} R. S. Vose,^{e‡} G. Rutledge,^{e‡} P. Bessemoulin,^f S. Brönnimann,^{g,h} M. Brunet,^{i,j} R. I. Crouthamel,^j A. N. Grant,^g P. Y. Groisman,^{e,k} P. D. Jones,¹ M. C. Kruk,^d A. C. Kruger,^m G. J. Marshall,ⁿ M. Maugeri,^o H. Y. Mok,^p Ø. Nordli,^q T. F. Ross,^{r‡} R. M. Trigo,^s X. L. Wang,^t S. D. Woodruff,^{b‡} and S. J. Worley^u

^aUniversity of Colorado, CIRES, Climate Diagnostics Center, Boulder, CO, USA ^bNOAA Earth System Research Laboratory, Physical Sciences Division, Boulder, CO, USA ^cACRE Project, Hadley Centre, Met Office, Exeter, UK ^dSTG Inc., Asheville, NC, USA *NOAA National Climatic Data Center, Asheville, NC, USA ^fMétéoFrance, Toulouse, France ⁸ETH Zurich, Switzerland ^hOeschger Center for Climate Change Research, University of Bern, Switzerland ¹Centre for Climate Change, Universitat Rovira i Virgili, Tarragona, Spain International Environmental Data Rescue Organization, Deale, MD, USA ^kUniversity Corporation for Atmospheric Research, Boulder, CO, USA ¹Climatic Research Unit, University of East Anglia, Norwich, UK ^mSouth African Weather Service, Pretoria, South Africa ⁿBritish Antarctic Survey, Cambridge, UK ^oDipartimento di Fisica, Università degli Studi di Milano, Milano, Italy PHong Kong Observatory, Hong Kong, China 9Norwegian Meteorological Institute, Oslo, Norway ¹NOAA Climate Database Modernization Program, NCDC, Asheville, NC, USA ⁵Centro de Geofísica da Universidade de Lisboa, IDL, University of Lisbon, Portugal ^tEnvironment Canada, Toronto, Ontario, Canada ^uNational Center for Atmospheric Research, Boulder, CO, USA *Correspondence to: Gilbert P. Compo, 325 Broadway R/PSD1, Boulder, CO USA 80305-3328. E-mail: compo@colorado.edu, gilbert.p.compo@noaa.gov

Compo et al. 2011 dx.doi.org/10.1002/qj.776

Special thanks to Chesley McColl, NCEP/EMC, NCDC, Hadley Centre, ACRE partners

The 20th Century Reanalysis Project (1871-2011)

Summary: An international project led by NOAA and CIRES to produce *4-dimensional* reanalysis datasets for climate applications extending back to the 19th century using an Ensemble Kalman Filter and *only surface pressure observations*.

Example: Weeklyaveraged anomaly during July 1936 United States Heat Wave (997 dead during 10-day span)

Daily variations compare well with in-situ data.



The reanalyses provide:

- -First-ever estimates of near-surface to tropopause 6-hourly fields extending back to the beginning of the 20th century;
- -Estimates of uncertainties in the basic reanalyses and derived quantities (e.g., storm tracks).
- Examples of uses (see <u>reanalyses.org</u> for more examples):
- •Validating climate models.
- •Determining storminess and storm track variations over the last 150 years.
- •Understanding historical climate variations (e.g., 1930s Dust Bowl, 1920-1940s Arctic warming).

•Estimating risks of extreme events Compo et al. 2011

Support provided by US Dept of Energy Office of Science (BER), NOAA Climate Program

Ensemble Filter Algorithm (Whitaker and Hamill, 2002)

Ensemble mean Ensemble deviations $\mathbf{x}_{j}^{\prime a} = \mathbf{x}_{j}^{\prime b} - \tilde{\mathbf{K}}\left(y_{j}^{\prime b}\right),$ $\overline{\mathbf{x}}_{j}^{\mathrm{a}} = \overline{\mathbf{x}}_{j}^{\mathrm{b}} + \mathrm{K}\left(y^{\mathrm{o}} - \overline{y}_{j}^{\mathrm{b}}\right),$ $\mathbf{K} = \mathbf{P}^{\mathbf{b}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^{\mathbf{b}}\mathbf{H}^{\mathrm{T}} + R)^{-1}$ Sample $= \frac{1}{n-1} \sum_{j=1}^{n} x_{j'}{}^{b} y_{j'}{}^{b} \left(\frac{1}{n-1} \sum_{i=1}^{n} y_{j'}{}^{b} y_{j'}{}^{b} + R \right)^{-1}$ Kalman Gain Sample $\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{R}{\mathbf{H}\mathbf{P}^{\mathbf{b}}\mathbf{H}^{\mathrm{T}} + R}}\right) \quad \mathbf{K},$ Modified Kalman Gain

 $x_j = \overline{x} + x'_j$ is pressure, air temperature, winds, humidity, etc. at all levels and gridpoints, every six hours. y^o is only observations of hourly and synoptic <u>surface pressure</u>, $y^b = Hx^b$ is guess surface pressure

Ensemble Data Assimilation (Whitaker and Hamill, 2002)

20CR analysis is a weighted average of the first guess x^b and pressure observation y^o. Each observation is assimilated serially.

Xa

Хp

 x^{a} , x^{b} 3-dimensional $x^{a} = x^{b} + K(y^{o} - x^{b})$ state of the atmosphere

ensemble of forecasts

 σ_b = First guess uncertainty

Xp

 σ_a = analysis uncertainty

the weight **K** varies with the <u>atmospheric flow</u> and the <u>observation network</u>

analysis time

Xa

analysis time

analysis time

20th Century Reanalysis implementation of Ensemble Filter Algorithm

(Whitaker et al. 2004, Compo et al. 2006, Compo et al. 2011)

Algorithm uses an ensemble of GCM runs to produce the weight K that varies with the <u>atmospheric flow</u> and the <u>observation network</u> every 6 hours

Using 56 member ensemble, HadISST1.1 prescribed SST and sea ice monthly boundary conditions (*Rayner et al. 2003*)
<u>1871-2011</u>: T62, 28 level NCEP GFS08ex atmosphere/land model 9 hour forecasts for 6 hour centered analysis window - time-varying CO₂, solar and volcanic radiative forcing
Sampling and Model error parameterizations:

Covariance localization (4000 km, 4 scale heights) and
Latitude and time dependent multiplicative covariance inflation (alpha = 1.01 to 1.12) [*Anderson and Anderson, 1999; Houtekamer and Mitchell, 2001; Hamill et al. 2001; Whitaker et al., 2004*]

Every 5 years produced in parallel: 1871-1875,..., 1881-1885, ..., 1996-2000, except 1945-1951, 2001-2011 after 14 month spin-up

http://go.usa.gov/XTd Compo et al. 2011, doi:10.1002/qj.776

International Surface Pressure Databank version 2 (ISPD)

Synoptic and hourly observations assembled in partnership with



GCOS AOPC/OOPC Working Group on Surface Pressure

GCOS/WCRP Working Group on Observational Data Sets for Reanalysis

Atmospheric Circulation Reconstructions over the Earth (ACRE)

Land data Component: merged by NOAA NCDC, NOAA ESRL, and CU/CIRES

- 33 data sources
- 33,653 stations
- 1.7 billion obs
- **1768-2010**



<u>Tropical Cyclone Best Track data component</u>: IBTrACS merged by NOAA NCDC DATA ACCESS <u>rda.ucar.edu/datasets/ds132.0</u> (T. Cram, NCAR DSS; C. McColl CIRES) <u>reanalyses.org/observations/surface</u>

Analyses for selected dates in 1894 and 1914



Sea Level Pressure

500 hPa Geopotential Height,

Surface Pressure uncertainty estimate poleward of 20(S,N) blue actual RMS difference red expected RMS difference

Northern Hemisphere



Uncertainty estimates are consistent with actual differences between first guess and pressure observations even as the network changes over more than 100 years! *Compo et al. 2011*

Southern Hemisphere

8

Daily column ozone measurements and 20CR daily ozone at Arosa, Switzerland (46.8N, 9.7E)

20CR has prognostic ozone from partial CHEM2D-OPP: production and loss depends on ozone mixing ratio and includes seasonal and spatial variation. Only gas phase chemistry.

Anomaly comparison spanning 1924 to 1963



20CR ozone field has large scale fluctuations that reflect ozone Highs associated with weather extremes, e.g., cold air outbreaks

(Bronnimann and Compo 2012)

Climate application: Global warming over land

The observed increase in near-surface air temperature over land (2 m air temperature, hereafter TL_{2m}) is a core indicator of global warming (e.g., *IPCC AR4, Trenberth et al. 2007*).

Accuracy of datasets documenting the increase continues to be debated (e.g, *Pielke et al. 2007, Fall et al 2011, Montandon 2011, Jones and Wigley 2012*) -including before the US Congress (*Christy 2012*)

Why? The record of TL_{2m} consists of observations taken irregularly in space and time using a variety of instruments and measurement techniques (e.g., *Karl et al. 1986, Peterson et al. 1998, Pielke et al. 2007, Brohan et al. 2006, Jones and Wigley 2010, Hansen et al. 2010, Parker 2011, Christy 2012, Vose et al. 2012).*



Compo et al. 2013 dx.doi.org/10.1002/grl.50425

- Centers producing datasets of near-surface air temperature over land have addressed these issues
 [e.g., Jones and Wigley 2010, Karl et al. 1986, Karl and Williams 1987, Peterson et al. 1998, Brohan et al. 2006, Hansen et al. 2010, Vose et al. 2012, Jones et al. 2012]
- Plus, estimated the associated uncertainty. [e.g., CRUTEM3: Brohan et al. 2006, CRUTEM4: Jones et al. 2012, NOAA MLOST: Vose et al. 2012]
- But, debate continues [e.g., Pielke et al. 2007, Jones and Wigley 2010 Fall et al. 2011, Montandon et al. 2011, Christy 2012]
- Take completely different approach: ignore all TL_{2m} observations and look at 20CR, which assimilates only pressure observations over land!

Annual anomalies of TL_{2m} from 20CR, CRUTEM4, average of 5 other instrumental datasets (1901-2010)

Other TL_{2m} datasets

1901-2010: r=0.84 to 0.92

1952-2010: r=0.95 to 0.96

Shading: 95% confidence interval



 TL_{2m} from stations and 20CR has consistent largescale annual, decadal, and centennial TL_{2m} variations ¹²

Advances and Improvements towards *Sparse Input Reanalyses* spanning 19th-21st centuries over the next 2-10 years

- 1. More land and marine observations back to early 19th century, especially Southern Hemisphere and Arctic.
- 2. User requirements for, and applications of, reanalyses (take <u>survey</u>)
- 3. Higher resolution, improved Quality Control methods, possibly other surface variables (e.g., wind, T, Tropical Cyclone position)
- 4. Uncertainty in forcings (e.g, CO2, solar, SST)
- 5. Possibly Multi-model (e.g., NASA, NCAR, NCEP, GFDL, ESRL) Available 2015 – 20CRv3 (1850-2014)

Available 2018 – NOAA Climate Reanalyses, suite spanning (1850-2018) Requires international cooperation, e.g., Atmospheric Circulation Reconstruction over the Earth initiative

http://www.met-acre.org

Developing 20CRv3: <u>Higher resolution testing 1999</u>

20CRv2 system but with

GFS2013 T126L64: about 1 degree and better resolved stratosphere (top at 0.64 hPa, up from 5.6 hPa in 20CRv2)

Improved covariance inflation and localization: accounts for spatial inhomogeneity of observations

Incremental analysis update: should reduce shocks from observations

Testing: New GCM Better use of observations

After 2 month spin-up, 12 months produced

First try using default GFS parameters

Northern Hemisphere 500 hPa Geopotential Height Root mean square difference with ERA-Interim



Worse than original 20CR. Mid-upper tropospheric cold bias. Cause: Not enough high cloud. Solution: Reduce ice to snow auto-conversion rate in Zhao-Carr microphysics. <u>1</u>

Halve the ice to snow autoconvergence rate

Northern Hemisphere 500 hPa Geopotential Height Root mean square difference with ERA-Interim



Changing default GFS ice to snow autoconvergenc rate significantly improves the New data assimilation compared to 20CR. Almost a 25% reduction in error.

Zonal mean temperature bias (2000)



Upper tropospheric warm bias in new system (too much high cloud?) Near surface temperature bias due to ice problem fixed. Stratospheric cold bias reduced, but not eliminated.

Developing 20CRv3: <u>Scout v3.3.8 testing 1831-1986</u>

20CRv2 system but with

Simple Ocean Data Assimilation with Sparse Input v2 (SODAsi.2): 18 member ensemble of daily SSTs (1846-2011)

ISPDv3.2.8: International effort to recover 100s of new stations, new marine observations from <u>Oldweather.org</u>, <u>ACRE</u> data rescue, over 33 new organizations contributing

Testing:

Effect of some accounting for uncertainty in SST Utility of increased density observations

After 14 month spin-up, 6 months produced for every 5th year of test years.

SODA Sparse Input 2 (1846-2011)

- 18 Ensemble Members
- Parallel Ocean Program v2.0.1
 - 0.4° longitude x 0.25 ° to 0.4 ° latitude with 40 levels



- Winds
 - 20CRv2 ensemble member daily stress (1949 2011)
 - 20CRv2 system with ISPDv3.2.4 and HadISST1.1
 - (1871-1948)
 - with ISPDv3.2.4 and climatological SST (1846-1870)
- Heat and Salt fluxes
 - Bulk formulae using 20CR daily variables
- Observations
 - Only ICOADS 2.5 SST data with Hadley Bucket Correction

Giese et al. 2014

Analyses of Sea Level Pressure for selected dates in 1831 and 1886

1831

1886



<u>Contours</u>-ensemble mean analysis (ci: 4 hPa, 1000 hPa thickened) <u>Shading</u>- blue: more uncertain, white: more certain Early analyses need more observations to advance.





Northern Hemisphere 24 hr forecasts beat persistence even in 1860s! Southern Hemisphere has an analysis that produces forecasts comparable to persistence starting in 1950s.

Sea Level Pressure obs from Oldweather.org HMS New Zealand





Shaded blue and red bars show $\pm 2 \sigma$ from reanalyses

Demonstrable improvement in data-sparse area of New Zealand!

1918 Fog of Ignorance (Scout pdf compared to climatology) and Glow of Discovery (Scout pdf improves upon 20CR)
 1 January 1918 Full Movie at http://vimeo.com/79549694



Grey dots=observations in 20CRv2 *P. Brohan, UK Met Office* Gold dots=New observations used in Scout3.3.8 from ACRE partners Black contours = Sea Level pressure

Historical Reanalysis Status and Plans

http://go.usa.gov/XTd

- Data Access: Analyses and ISPD (with feedback) freely available from NCAR, analyses from NOAA/ESRL, DOE NERSC, ESGF, and BADC.
- Fall 2013: 1871-2011 (includes time-varying CO2, volcanic aerosols, GFS from NCEP). Ensemble mean and spread, and selected individual member variables online now.
 - <u>http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.html</u> (NOAA ESRL)
 - <u>http://dss.ucar.edu/datasets/ds131.1</u> (NCAR)
 - <u>http://portal.nersc.gov/20C_Reanalysis</u> Every member (US Dept of Energy, NERSC)

NERSC High Performance Storage System direct-from-tape distribution

- Earth System Grid Federation ana4MIPS distribution and validation for IPCC AR5 and CMIP5
- British Atmospheric Data Center Every member
 - <u>http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_DE_6ae84cbc-177b-11e2-9c9c-00163e251233</u>



- Demonstrated that surface-based reanalyses throughout the troposphere are feasible using advanced data assimilation and surface pressure observations.
- Effectively doubling the reanalysis record length from ~60 year to more than 150 years, allowing current atmospheric circulation patterns to be placed in a broader historical context. ③
- Increasing vertical levels not a panacea to resolve all stratosphere bias issues.
- Higher resolution and additional observations will further improve these reanalyses.
- For status updates, email
 - jeffrey.s.whitaker@noaa.gov,
 - compo@colorado.edu

<u>Reanalyses.org</u>: A collaborative website for research and questions about atmospheric, oceanic, and land reanalyses

eanalyses.org		↓ 佘	
Advancing Reanalysis	Home Contact Us Login/R	egister	
HOME ABOUT ATMOSPHERE OCEAN OBSERVATIONS I	MEETINGS REPORTS	Search	
<section-header> Welcome to the commentation Reanalyses.org Home Page Members will need to login information. Recent Updates Recent Updates Frequent to the period 1871-2009 in the period 187</section-header>	Reanalysis is a scientific method for developing a comprehensive record of how weather and climate are changing over time. In it, observations and a numerical mode simulates one or more aspects of the Earth system are combined objectively to generate a synthesized estimate of the state of the system. A reanalysis typically extend several decades or longer, and covers the entire globe from the Earth's surface to well above the stratosphere. Reanalysis products are used extensively in climate res and services, including for monitoring and comparing current climate conditions with those of the past, identifying the causes of climate variations and change, and prep climate predictions. Information derived from reanalyses is also being used increasingly in commercial and business applications in sectors such as energy, agriculture, resources, and insurance. Using a collaborative Wiki framework, the goal of reanalyses.org is to facilitate comparison between reanalysis and observational datasets. Evaluative content provided by reanalysis developers, observationalists, and users; and links to detailed data descriptions, data access methods, analysis and plotting tools, and dataset references are available. Discussions of the recovery of observations to improve reanalyses is also a focus. The wiki framework encourages scientific discussion between solves.	that Is over earch aring water ween	
Disposito Studies: Climate Variability - 01/08/2014 News (4 December 2013): Please take p 12/17/2013 Data Rescue - 12/17/2013 Data Rescue -	part in the Reanalysis User and Application Survey (closing 31 January 2014)		
12/17/2013 IOpics • Overview of current atmospheric reanalyses - 12/17/2013 • Overview of Current Atmospheric Reanalyses • Reanalyses.org Home Page - 12/04/2013 • Overview of Current Ocean Reanalyses • Upper-air - 12/02/2013 • Atmospheric Reanalyses Plotting and Data Manipulation Tools • 20CR: ensemble mean wind-speeds appear wrong - 11/20/2013 • jra-references -			
Please take part in the Reanalysis User and Application Survey			

Thank you to organizations contributing observations to ISPD:

All Russia Research Institute of Hydrometeorological	NIWA
Information WDC	NOAA Climate Database Modernization Program
Atmospheric Circulation	NOAA Earth System Research Laboratory
Reconstructions over the Earth (ACRE)	NOAA National Climatic Data Center
Australian Bureau of Meteorology	NOAA National Centers for Environmental Prediction
Australian Meteorological Association, Todd Project Team	NOAA Northeast Regional Climate Center at Cornell L
British Antarctic Survey	NOAA Midwest Regional Climate Center at UIUC
Cook Islands Met Service	NOAA Pacific Marine Environmental Laboratory
Danish Meteorological Institute	Norwegian Meteorological Institute
Deutscher Wetterdienst	Oldweather.org
EMULATE	Ohio State U. – Byrd Polar Research Center
Environment Canada	Portuguese Meteorological Institute (IM)
ETH-Zurich	Proudman Oceanographic Laboratory
European Reanalysis and Obs for Monitoring	SIGN - Signatures of environmental change in the
GCOS AOPC/OOPC WG on Surface Pressure	observations of the Geophysical Institutes
GCOS/WCRP WG on Obs Data Sets	South African Weather Service
Hong Kong Observatory	UK Met Office Hadley Centre
Icelandic Meteorological Office	U. of Bern, Switzerland
IBTrACS	U. of Colorado-CIRES/Climate Diagnostics Center
ICOADS	U. of East Anglia-Climatic Research Unit
IEDRO	U. of Giessen –Dept. of Geography
JAMSTEC	U. of Lisbon-Instituto Geofisico do Infante D. Luiz
Japan Meteorological Agency	U. of Lisbon-Instituto de Meteorologia
Jersey Met Dept.	U. of Mebourne
Lamont-Doherty Earth Observatory	U. of Milan-Dept. of Physics
KNMI	U. of Porto-Instituto Geofisca
MeteoFrance	U. Rovira i Virgili-Center for Climate Change
MeteoFrance – Division of Climate	U. of South Carolina
Meteorological and Hydrological Service, Croatia	U. of Toronto-Dept of Physics
National Center for Atmospheric Research	U. of Washington
Nicolaus Copernicus University	World Meteorological Organization - MEDARE
Niue Met Service	ZAMG (Austrian Weather Service)

Comparison of ISPD versions



New ISPDv3.2.8 has considerably more observations than v2

Simple Ocean Data Assimilation si.2 Global Ocean Annual Average (60N-60S)



SODAsi.2 trends and decadal variability are consistent with statistical reconstructions. SODAsi Generates interannual variations in late 1850s even when 20CR forcing had climatological SST. *Giese et al. 2014*



Linear Trends of TL_{2m} are consistent

CRU_TS3.10

20CR



Trends are similar, with differences in detail, over both 1901-2010 and more rapidly warming 1952-2010.

Comparing the spatial patterns and amplitudes of local trends of TL_{2m} from 20CR with (filled circles) eight observational datasets and from (open squares) SST-forced simulation AMIP20C

Cosine of angle is pattern correlation.

Radial distance is ratio of pattern amplitudes.



20CR compares better (r=0.69 to 0.81) to observed TL2m than to AMIP20C (squares), particularly in the recent period.

Annual anomalies of near-global TL_{2m} from 20CR, MLOST, UDEL (1901-2010)

TL2m

Difference



1901

2010 1901



Uncertainty estimates are largely consistent. Differences are not "urban warming".

Conclusions

- Ignore near-surface air temperature observations (TL_{2m}) and instead infer them from global observations of subdaily barometric pressure using a physically-based data assimilation system called the 20th Century Reanalysis (20CR).
- Independent estimate of TL_{2m} from 20CR demonstrates that, in spite of recently published and public concerns with the station temperature record, the temperature analyses used by the IPCC and many others are *reliable and robust estimates* of large-scale TL_{2m} variability and change.
- The 20CR reproduces both TL_{2m} annual variations and centennial trends, demonstrating the robustness of previous conclusions based on the temperature analyses regarding global warming (e.g., IPCC reports).

20CR data: http://go.usa.gov/XTd Contact: <u>compo@colorado.edu</u> Pattern correlation between 20CR and ERA40 and NCEP-NCAR Reanalyses of monthly anomalies of NH Tropospheric temperature (300 hPa geopotential height)



Reanalysis correlations are much higher than for SST-forced simulation. 20CR is more than an SST-forced simulation. Local Anomaly Correlation of 300 hPa geopotential height anomalies from 20th Century Reanalysis (20CRv2) and ERA40 (1979 to 2001)

Black curves show where NCEP-NCAR and ERA40 correlate > 0 975



Northern and Southern Hemisphere agreement are excellent between 20CRv2 and ERA40 when ERA40 has satellite observations.

Subdaily 500 hPa Geopotential Height anomalies from observations and 20th Century Reanalysis compare well.



Observations from CHUAN dataset (Stickler et al. 2010)

1905-2006 Measurements from kites, aircraft, registering balloon, and radiosondes at Lindenberg, Germany

Near Surface Annual Mean Temperature Anomalies for Land only (60N to 60S) 1979 to 2010



Correlations between 20CR and thermometer-based estimates are relatively high (0.94 to 0.95) [see Parker 2011]. Correlations with upper-air and satellite-based reanalyses also high (0.91 to 0.94). 95% error ranges are largely consistent. Compo, Sardeshmukh, Whitaker, Brohan, Jones, and McColl (2013)