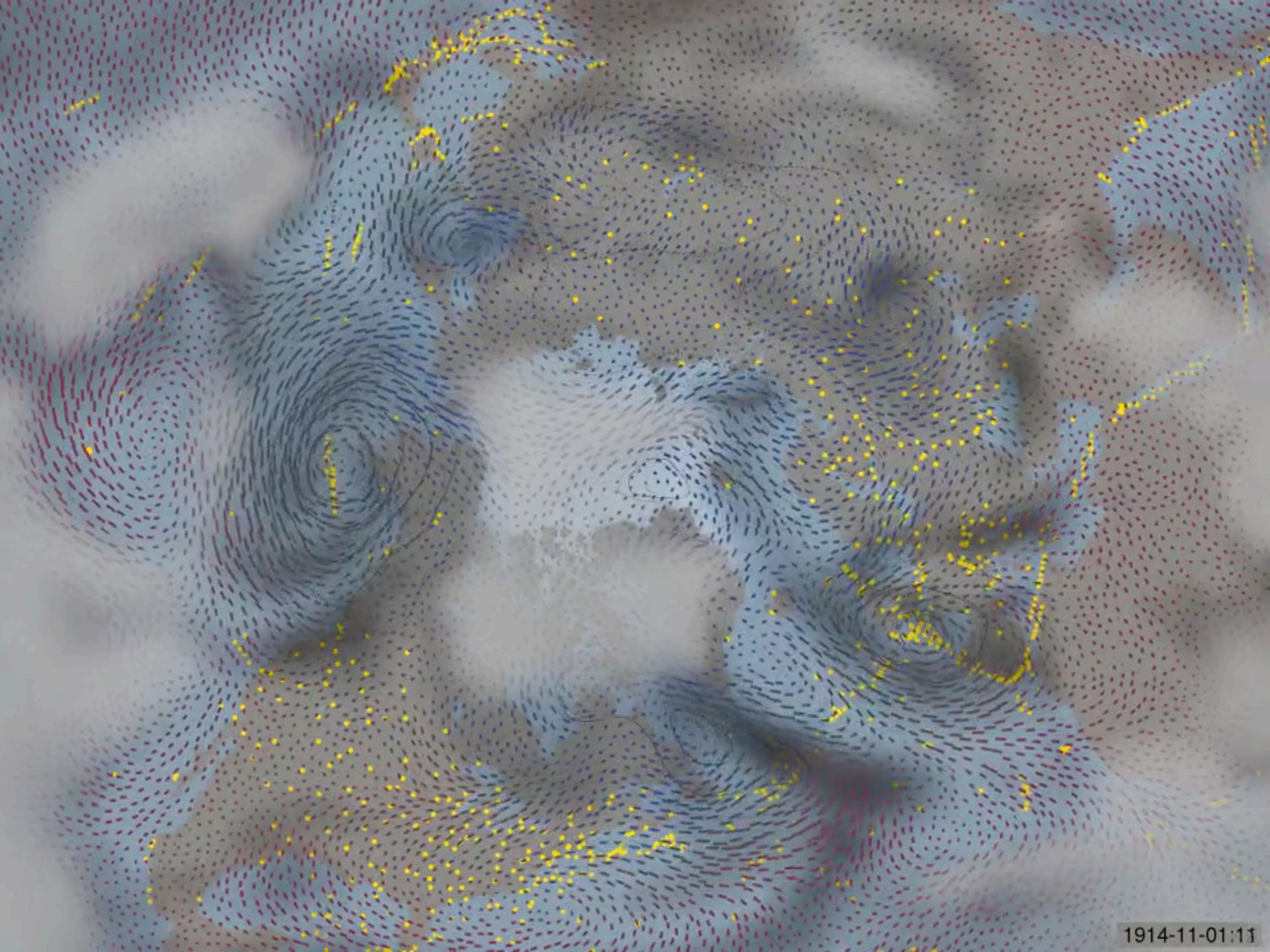

1914-11-01:11

Towards an improved 20th Century reanalysis version "2c" (1851-2012)

Gilbert P. Compo[1,2],
Jeffrey S. Whitaker[2],
Prashant D. Sardeshmukh[1,2],
Benjamin Giese[3], Philip Brohan[4]

[1]Univ. of Colorado/CIRES and
[2]NOAA Earth System Research Laboratory/PSD
[3]Texas A&M, Dept. of Oceanograhpy
[4]Hadley Centre UK Met Office

Quarterly Journal of the Royal Meteorological Society — Q. J. R. Meteorol. Soc. 137: 1–28, January 2011 Part A

RMetS
Royal Meteorological Society

Review Article
The Twentieth Century Reanalysis Project

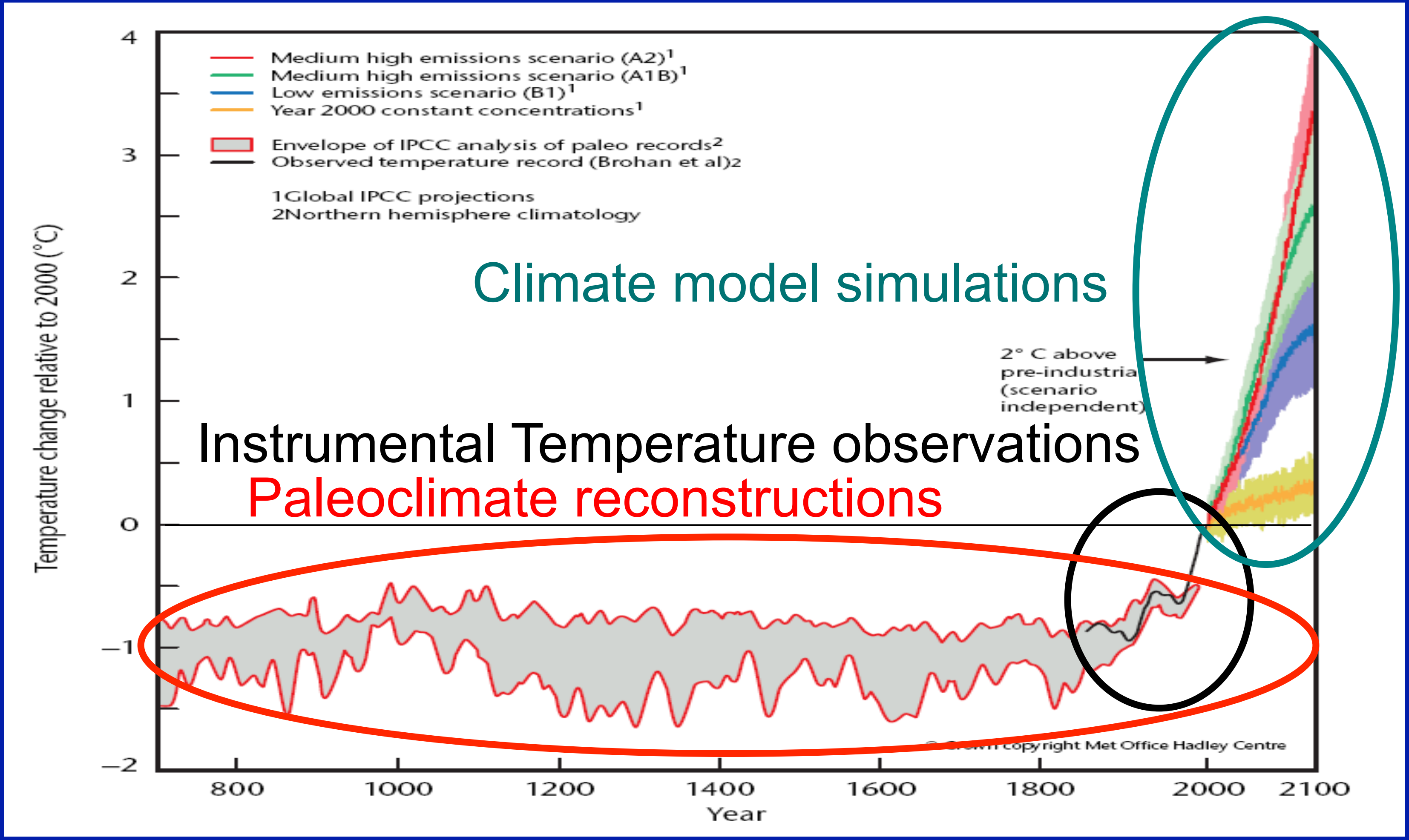
G. P. Compo,[a,b*] J. S. Whitaker,[b‡] P. D. Sardeshmukh,[a,b] N. Matsui,[a,b] R. J. Allan,[c†] X. Yin,[d] B. E. Gleason, Jr.,[e‡] R. S. Vose,[e‡] G. Rutledge,[e‡] P. Bessemoulin,[f] S. Brönnimann,[g,h] M. Brunet,[i,l] R. I. Crouthamel,[j] A. N. Grant,[g] P. Y. Groisman,[e,k] P. D. Jones,[l] M. C. Kruk,[d] 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]

[a]*University of Colorado, CIRES, Climate Diagnostics Center, Boulder, CO, USA*
[b]*NOAA Earth System Research Laboratory, Physical Sciences Division, Boulder, CO, USA*
[c]*ACRE Project, Hadley Centre, Met Office, Exeter, UK*
[d]*STG Inc., Asheville, NC, USA*
[e]*NOAA National Climatic Data Center, Asheville, NC, USA*
[f]*MétéoFrance, Toulouse, France*
[g]*ETH Zurich, Switzerland*
[h]*Oeschger Center for Climate Change Research, University of Bern, Switzerland*
[i]*Centre for Climate Change, Universitat Rovira i Virgili, Tarragona, Spain*
[j]*International Environmental Data Rescue Organization, Deale, MD, USA*
[k]*University Corporation for Atmospheric Research, Boulder, CO, USA*
[l]*Climatic Research Unit, University of East Anglia, Norwich, UK*
[m]*South African Weather Service, Pretoria, South Africa*
[n]*British Antarctic Survey, Cambridge, UK*
[o]*Dipartimento di Fisica, Università degli Studi di Milano, Milano, Italy*
[p]*Hong Kong Observatory, Hong Kong, China*
[q]*Norwegian Meteorological Institute, Oslo, Norway*
[r]*NOAA Climate Database Modernization Program, NCDC, Asheville, NC, USA*
[s]*Centro de Geofísica da Universidade de Lisboa, IDL, University of Lisbon, Portugal*
[t]*Environment Canada, Toronto, Ontario, Canada*
[u]*National 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

Illustration of Temperature Change relative to 2000



800 1200 1600 2000

Instrumental period is the short link between paleo reconstructions and climate model simulations

Courtesy UK Met Office Hadley Centre

A Major Goal of the 20th Century Reanalysis project and the international Atmospheric Circulation Reconstructions over the Earth initiative:

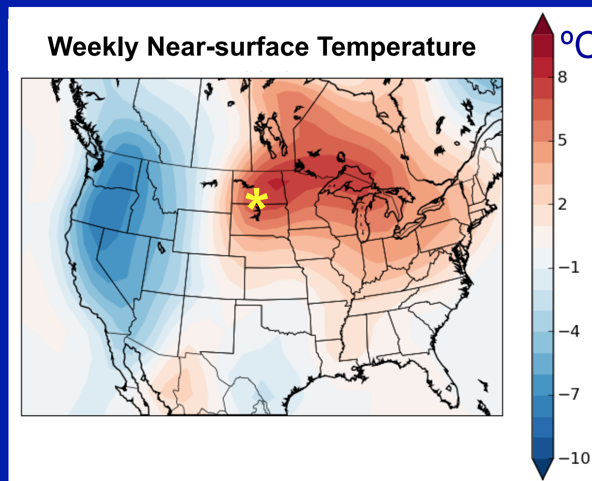
Use data assimilation to produce the longest possible *Instrument-based* estimates of global weather and climate for comparison to paleoclimate reconstructions and climate model simulations

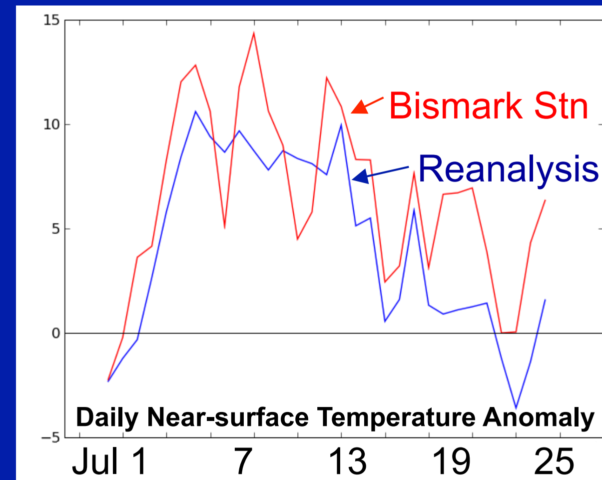
The 20th Century Reanalysis Project version 2c (1851-2011)

Summary: An international project led by CIRES and NOAA 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.*

Weekly-averaged anomaly during July 1936 North American Heat Wave (> 2,000 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 middle of the 19th century;

-Estimates of uncertainties in the basic reanalyses and derived quantities (e.g., storm tracks).

Examples of uses:

- 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

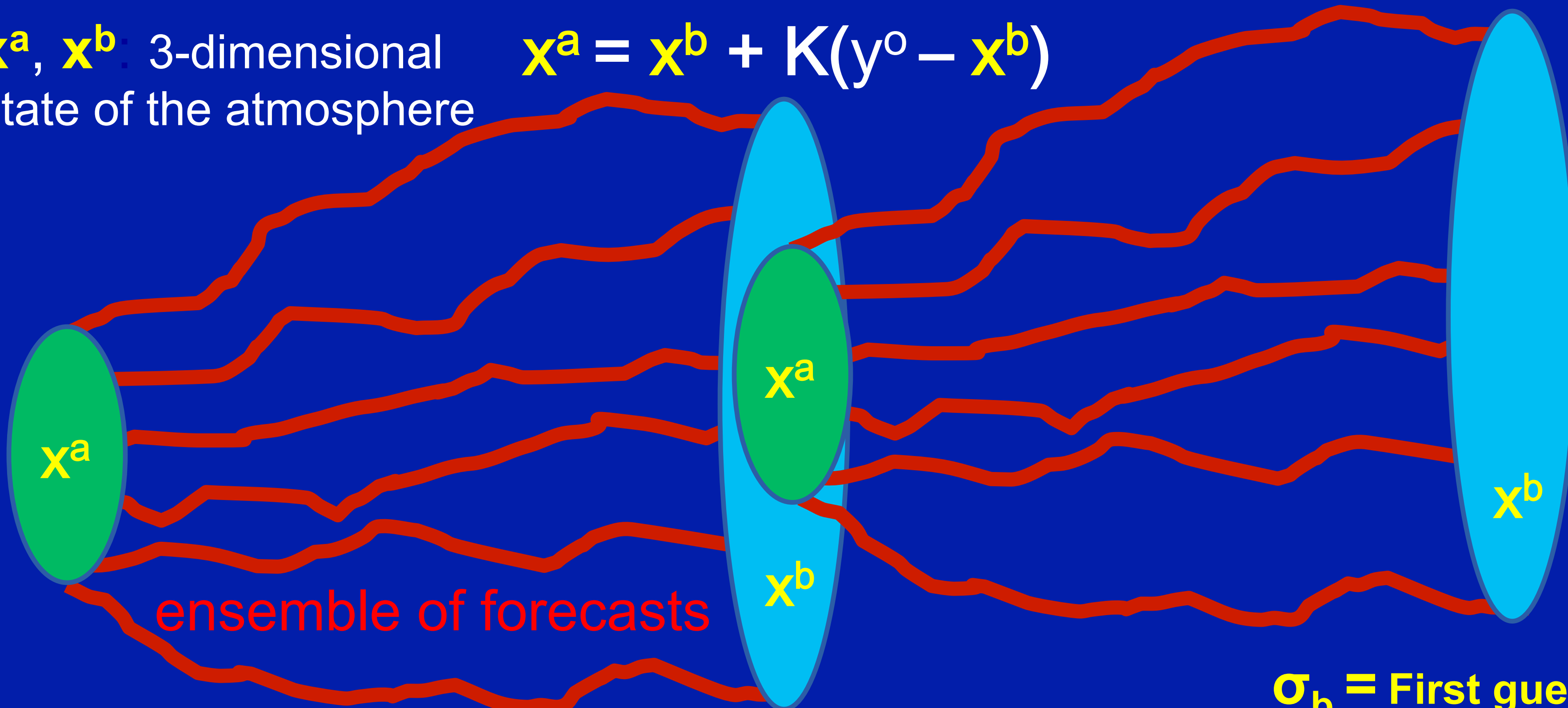
Compo et al. 2011

Ensemble Data Assimilation (Whitaker and Hamill, 2002)

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

x^a, x^b: 3-dimensional state of the atmosphere

$$x^a = x^b + K(y^o - x^b)$$



σ_a = analysis uncertainty

σ_b = First guess uncertainty

the weight K varies with the atmospheric flow and the observation network

analysis time (0Z) **analysis time (6Z)** **analysis time (12Z)**

Ensemble Filter Algorithm (Whitaker and Hamill, 2002)

Ensemble mean

$$\overline{\mathrm{x}}_j^{a} = \overline{\mathrm{x}}_j^{b} + \mathrm{K}\left(y^{o} - \overline{y}_j^{b}\right),$$

Ensemble deviations

$$\mathrm{x}_j'^{a} = \mathrm{x}_j'^{b} - \tilde{\mathrm{K}}\left(y_j'^{b}\right),$$

Sample Kalman Gain

$$\mathrm{K} = \mathrm{P}^{b}\mathrm{H}^{\mathrm{T}}(\mathrm{H}\mathrm{P}^{b}\mathrm{H}^{\mathrm{T}} + R)^{-1}$$
$$= \frac{1}{n-1}\sum_{j=1}^{n}\mathrm{x}_j'^{b}y_j'^{b}\left(\frac{1}{n-1}\sum_{j=1}^{n}y_j'^{b}y_j'^{b} + R\right)^{-1}$$

Sample Modified Kalman Gain

$$\tilde{\mathrm{K}} = \left(1 + \sqrt{\frac{R}{\mathrm{H}\mathrm{P}^{b}\mathrm{H}^{\mathrm{T}} + R}}\right)^{-1}\mathrm{K},$$

$\mathbf{x}_j = \mathbf{x} + \mathbf{x}'_j$ is pressure, air temperature, winds, humidity, etc. at all levels and gridpoints, every six hours.

y^{o} is only observations of hourly and synoptic **surface pressure**,

$\mathbf{y}^{b} = \mathbf{H}\mathbf{x}^{b}$ is guess surface pressure

20th Century Reanalysis v2c 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 atmospheric flow and the observation network every 6 hours

Using 56 member ensemble, new prescribed boundary conditions:

SODAsi.2c 18 member pentad SST and

COBE-SST2 monthly sea ice concentration

(corrects sea ice error in v2)

(*Giese et al. 2015*, *Hirahara et al. 2014*)

1851-2011:

T62 (~200km), 28 level NCEP GFS08ex atmosphere/land model

9 hour forecasts for 6 hour centered analysis window

- time-varying CO_2, solar and volcanic radiative forcing (Sato et al.)

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

20th Century Reanalysis v2c implementation of Ensemble Filter Algorithm (same as v2)

(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 atmospheric flow and the observation network every 6 hours

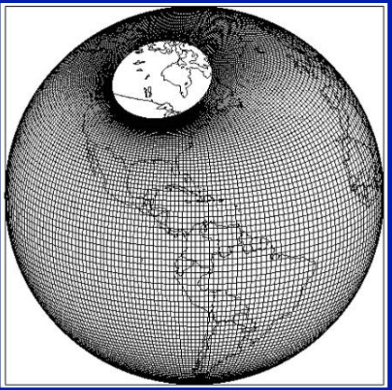
Sampling and Model error parameterizations:

-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: 1851-1855,…, 1881-1885, …, 1996-2000, .., 2006-2011 after 14 month spin-up

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

SODA sparse input v2
(1846-2011)



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

Giese et al. 2015

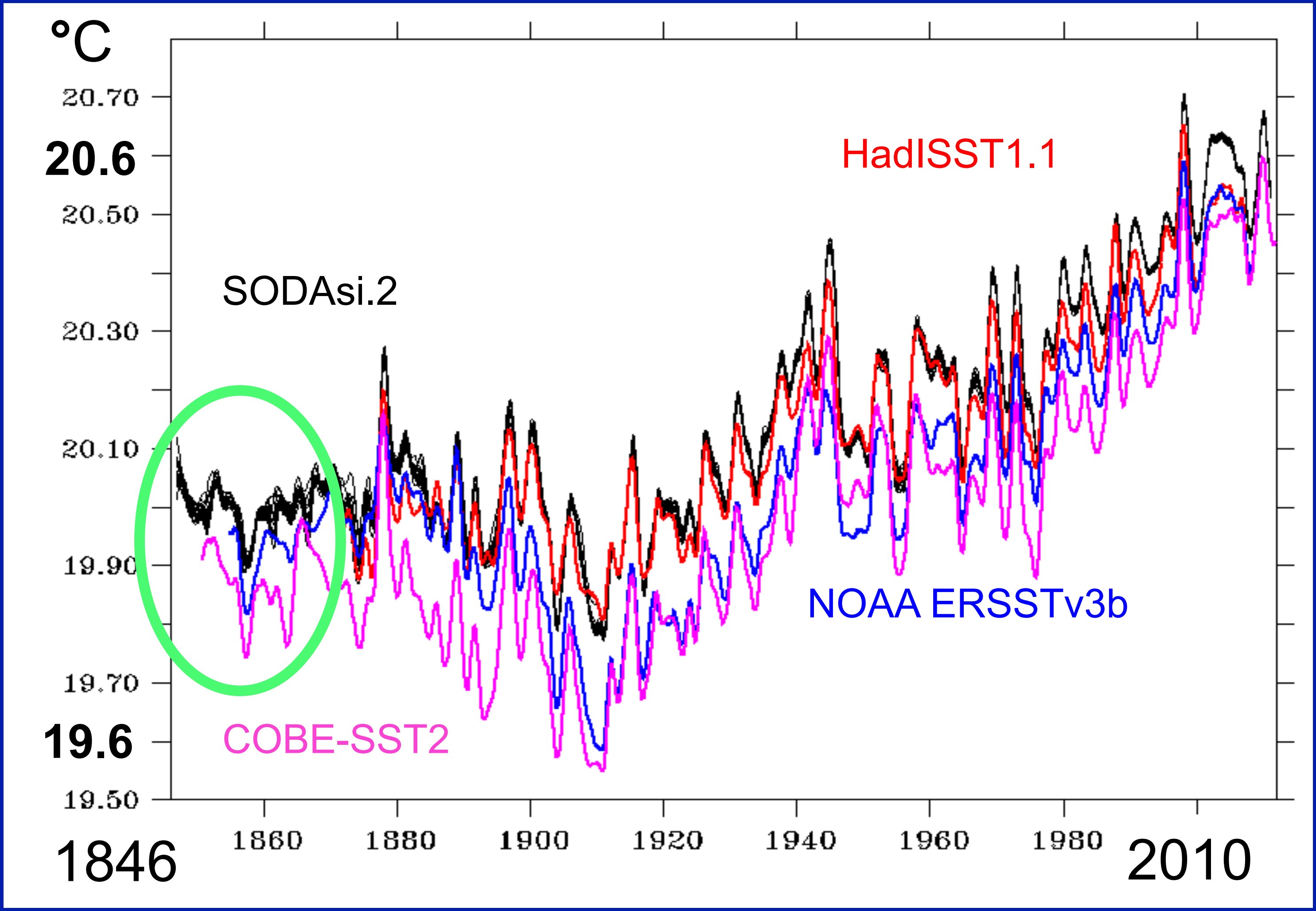
- **Winds**
  - 20CRv2 ensemble member daily stress (1949 – 2011)
  - 20CRv2 system with ISPDv3.2.4 and HadISST1.1 (1871-1948)
  - with ISPDv3.2.4 and climatological SST (1846-1870)

- **Heat and Salt fluxes**
  - Bulk formulae using 20CRv2 daily variables

- **SODAsi Observations**
  - Only ICOADS 2.5 SST data with Hadley Bucket Correction

Simple Ocean Data Assimilation sparse input v2
Global Ocean Annual Average (60N-60S)

Climo SST in 20CR 1846 to 1870 but variation in SODAsi



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

International Surface Pressure Databank version 3 (ISPD)

Subdaily observations assembled in partnership with

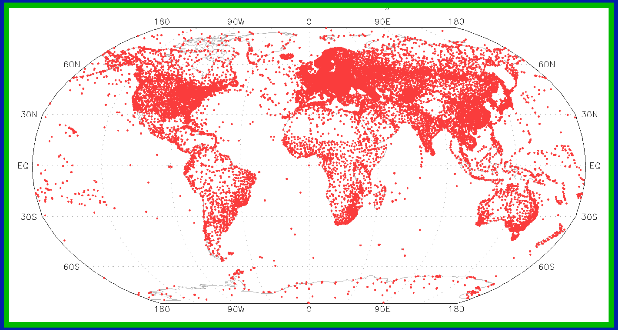
GCOS AOPC/OOPC Working Group on Surface Pressure

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

- **63 data sources**
- **33,653 stations**
- **1.7 billion obs**
- **1768-2011**



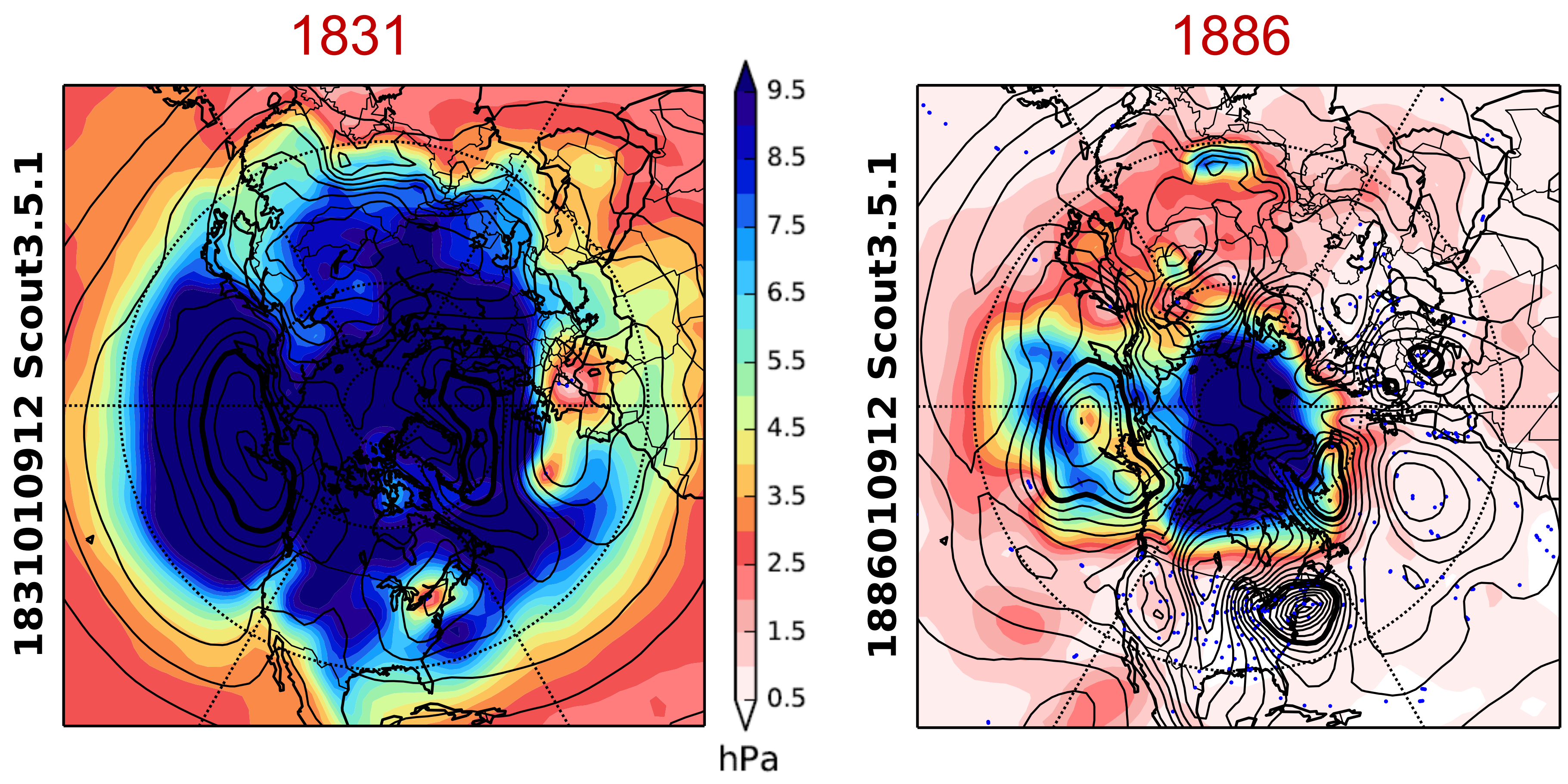
Marine data component: **ICOADSv2.5** merged by NOAA ESRL, NCDC, and NCAR;

Oldweather.org, ACRE marine data

Tropical Cyclone Best Track data component: **IBTrACS** merged by NOAA NCDC

DATA ACCESS rda.ucar.edu/datasets/ds132.0 (T. Cram, NCAR DSS; C. McColl CIRES)
Reanalyses.org/observations/surface , Cram et al. 2015

20CRv2c Analyses of Sea Level Pressure for selected dates in 1831 and 1886



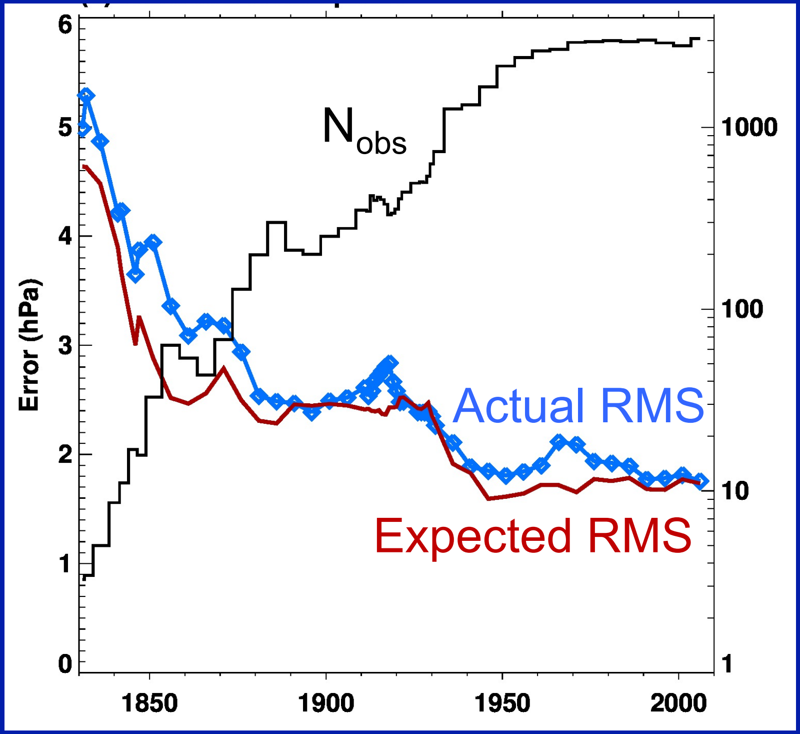
Contours-ensemble mean (ci: 4 hPa, 1000 hPa thickened)

Shading- blue: more uncertain, white: more certain

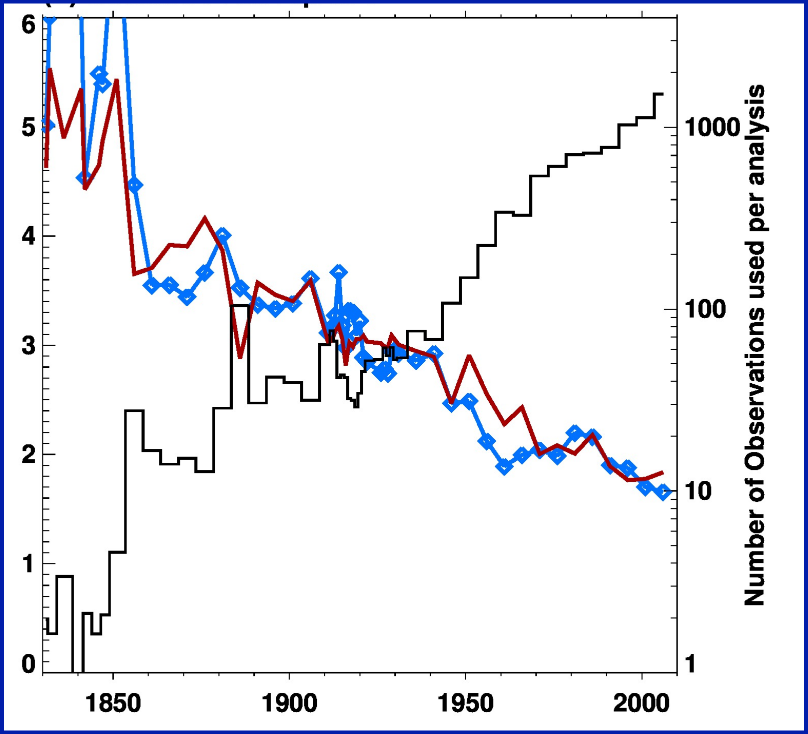
Analysis system responds to the observations and the flow, providing quantitative uncertainty for every variable at each analysis time.

v2c Surface Pressure uncertainty estimate poleward of 20(S,N)
blue actual RMS difference obs minus first guess
red expected RMS difference obs minus first guess

Northern Hemisphere



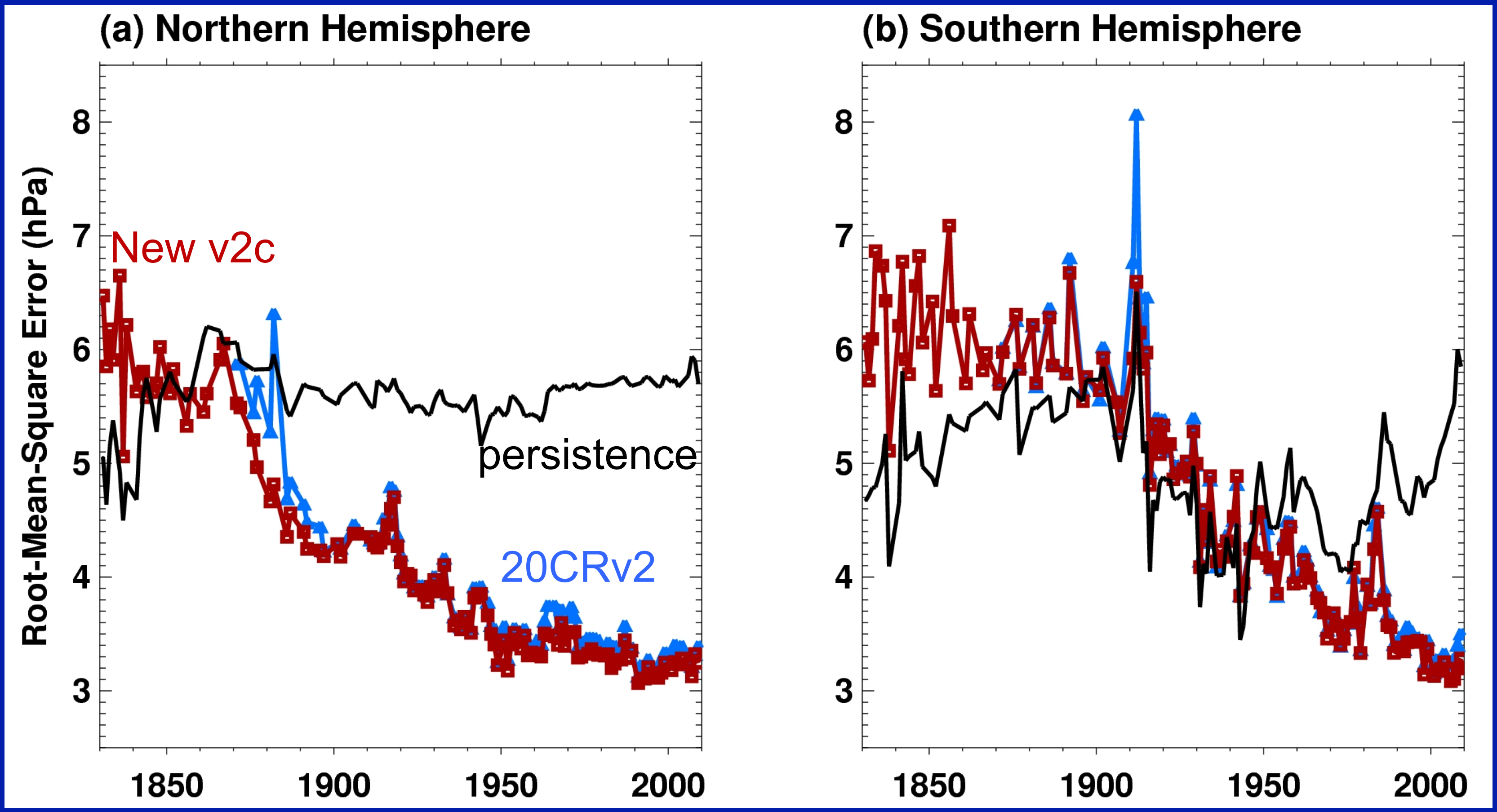
Southern Hemisphere



Uncertainty estimates are consistent with actual differences between first guess and pressure observations even as the network changes by three orders of magnitude over more than 150 years! (*This is not tuned*).

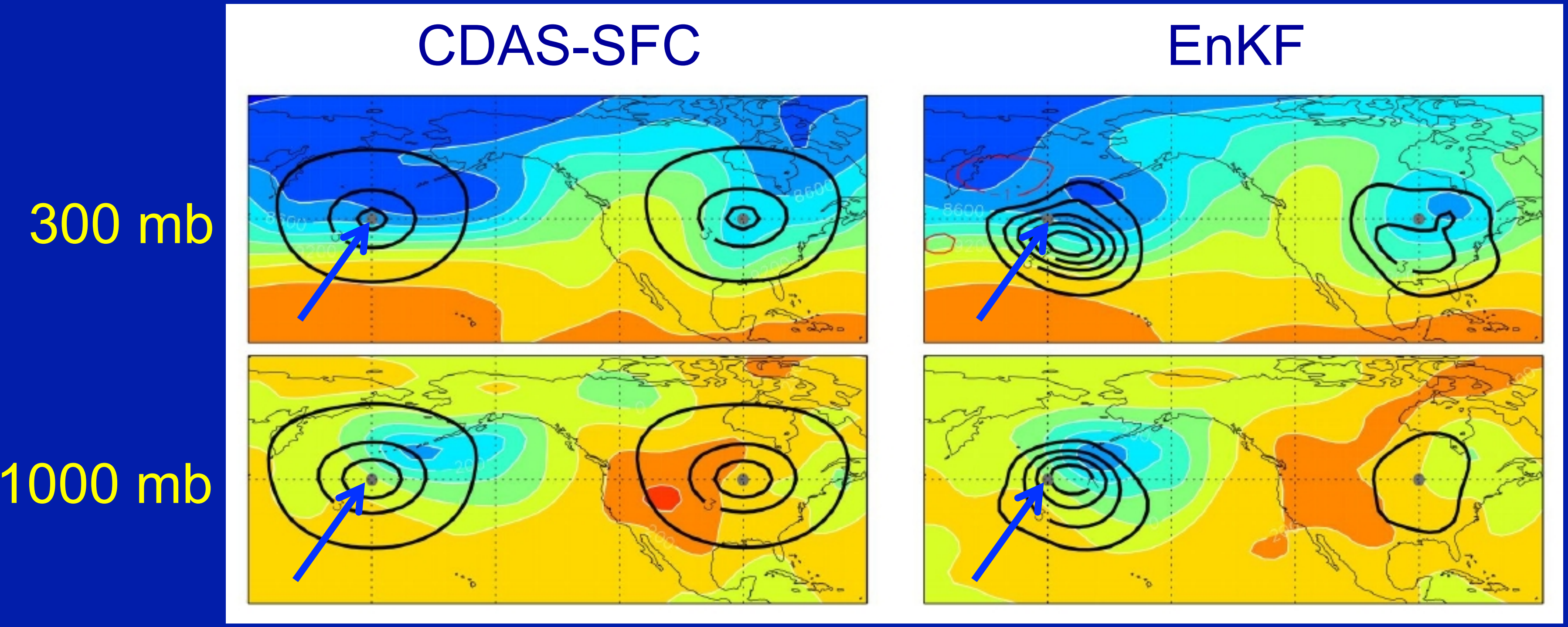
Adapted from Compo et al. 2011

Root Mean Square difference of Surface and Sea Level Pressure Observations and 24 hour Forecasts from 20CRv2 and v2c (*Jan-Dec*)



Northern Hemisphere 24 hr forecasts beat persistence even in 1850s. Southern Hemisphere has an analysis that produces forecasts comparable to persistence starting in 1900s. New v2c is an improvement.

Geopotential height first guess (colors) and analysis minus first guess (lines) for single pressure observations 1mb greater than first guess



Ensemble Kalman filter can extract spatially-varying structures relative to the flow and the previous observational density.

Compo et al. 2006

Example uses of reanalysis

1. Effectively doubling the reanalysis record length ☺
2. Validating climate models for large-scale synoptic anomalies during extreme periods, such as droughts (30's, 50's).
3. Better understanding events such as the 1920-1940's Arctic warming.
4. Determining storminess and storm track variations over last 100-150 years.
5. Developing and improving forecasts of low-frequency (e.g., Pacific-North America pattern, North Atlantic Oscillation) atmospheric variations and their interannual to decadal variability.
6. Understanding changing atmospheric background state associated with interdecadal hurricane activity.
7. Discovering previously undocumented hurricanes.
8. Homogenizing upper-air and other independent observations.
9. Estimating historical probability distributions for wind energy.
10. Calibrating paleoclimate reconstructions.
11. Determining Weather effects on heroic journeys, e.g., death of Mallory and Irvine climbing Mt. Everest.
12. Tracking icebergs in the vicinity of the *Titanic*

Example uses of reanalysis

1. Effectively doubling the reanalysis record length ☺
2. Validating climate models for large-scale synoptic anomalies during extreme periods, such as droughts (30's, 50's).
3. Better understanding events such as the 1920-1940's Arctic warming.
4. Determining storminess and storm track variations over last 100-150 years.
5. Developing and improving forecasts of low-frequency (e.g., Pacific-North America pattern, North Atlantic Oscillation) atmospheric variations and their interannual to decadal variability.
6. Understanding changing atmospheric background state associated with interdecadal hurricane activity.
7. Discovering previously undocumented hurricanes.
8. Homogenizing upper-air and other independent observations.
9. Estimating historical probability distributions for wind energy.
10. Calibrating paleoclimate reconstructions.
11. Determining Weather effects on heroic journeys, e.g., death of Mallory and Irvine climbing Mt. Everest.
12. Tracking icebergs in the vicinity of the *Titanic*.

Initial comparisons with ERA-20C

(*Poli et al. 2014*)

Assimilates

1) Pressure observations

    ISPDv3.2.6: International effort to recover 100s of new stations, ACRE & ERA-CLIM data rescue, over 33 new organizations contributing

2) ICOADS R2.5 marine near-surface winds

4D-Var algorithm

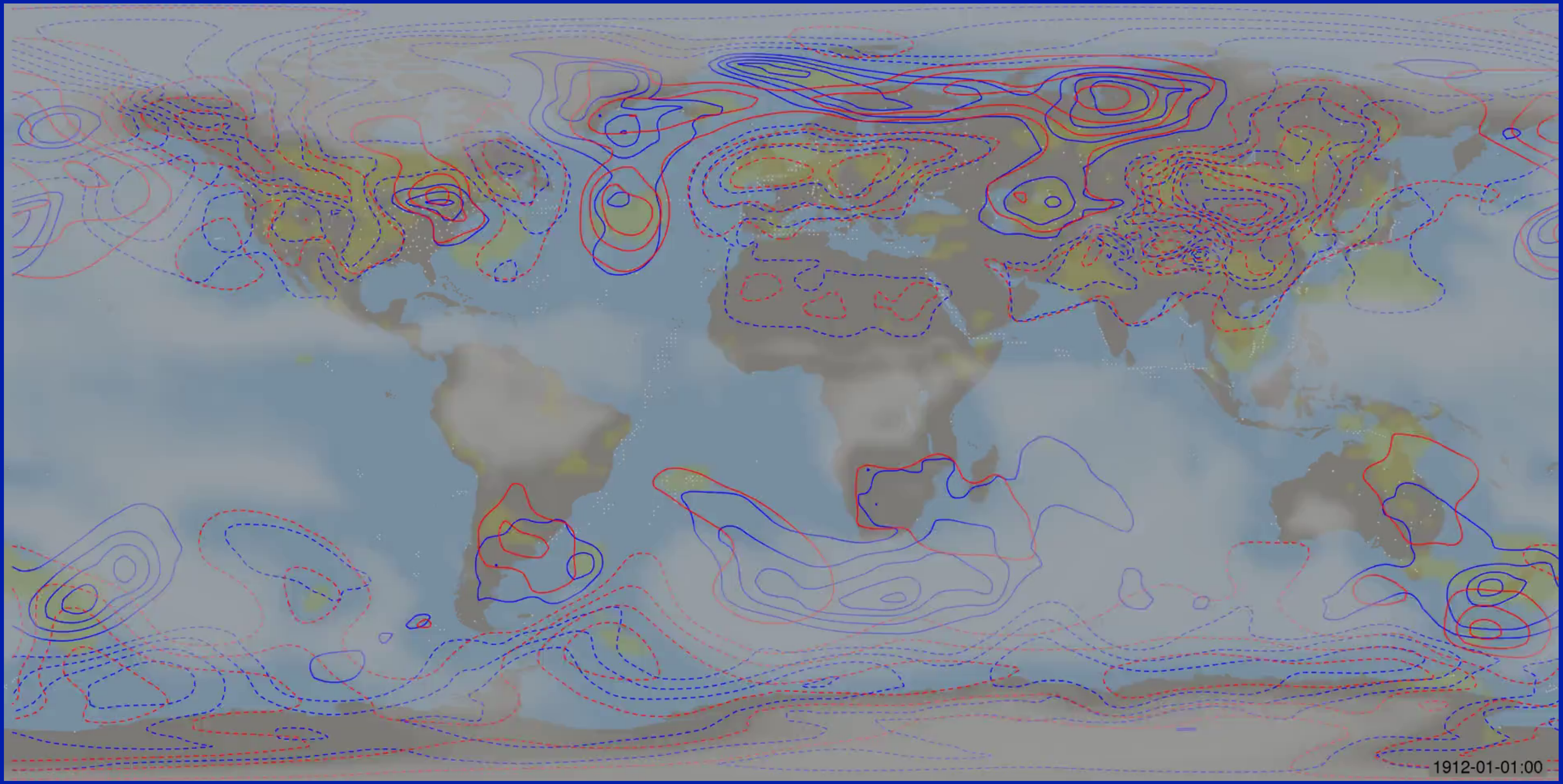
HadISST2.1.0.0 sea surface temperature and sea ice concentration

January 1912 Sea Level Pressure anomalies

20CRv2c ERA-20C

(gold=20CRv2c difference from ERA-20C > 3*spread)

Full Brohan movie at vimeo.com/109681668



Grey dots=observations in 20CRv2c. Grey Fog=where 20CRv2c very uncertain.

20CRv2c differences with ERA-20C large over land.

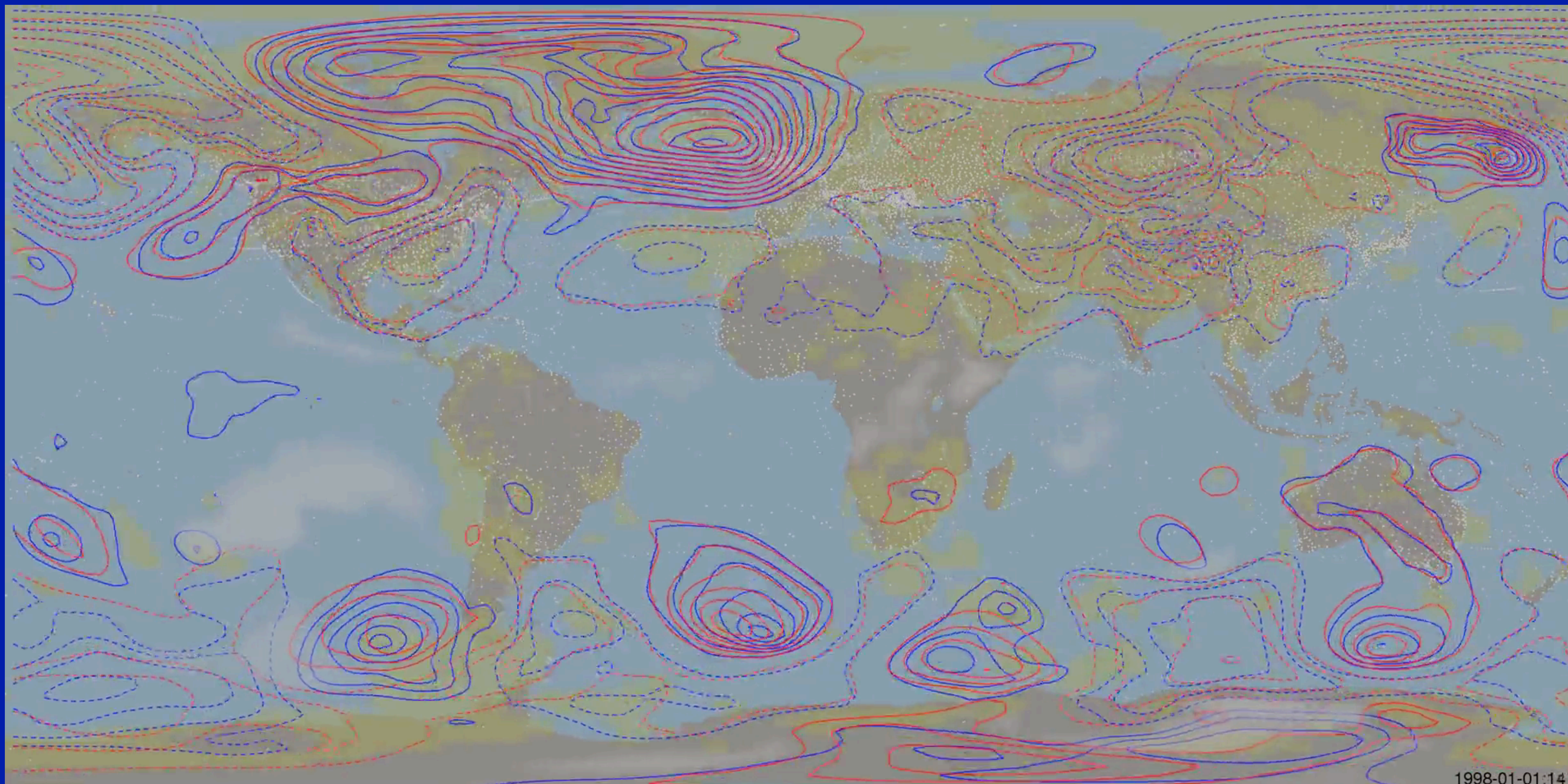
Suspect latitudinally-constant covariance inflation is too small there.

January 1998 Sea Level Pressure anomalies

20CRv2c ERA-20C

(gold=20CRv2c difference from ERA-20C > 3*spread)

Full Brohan movie at vimeo.com/109684347



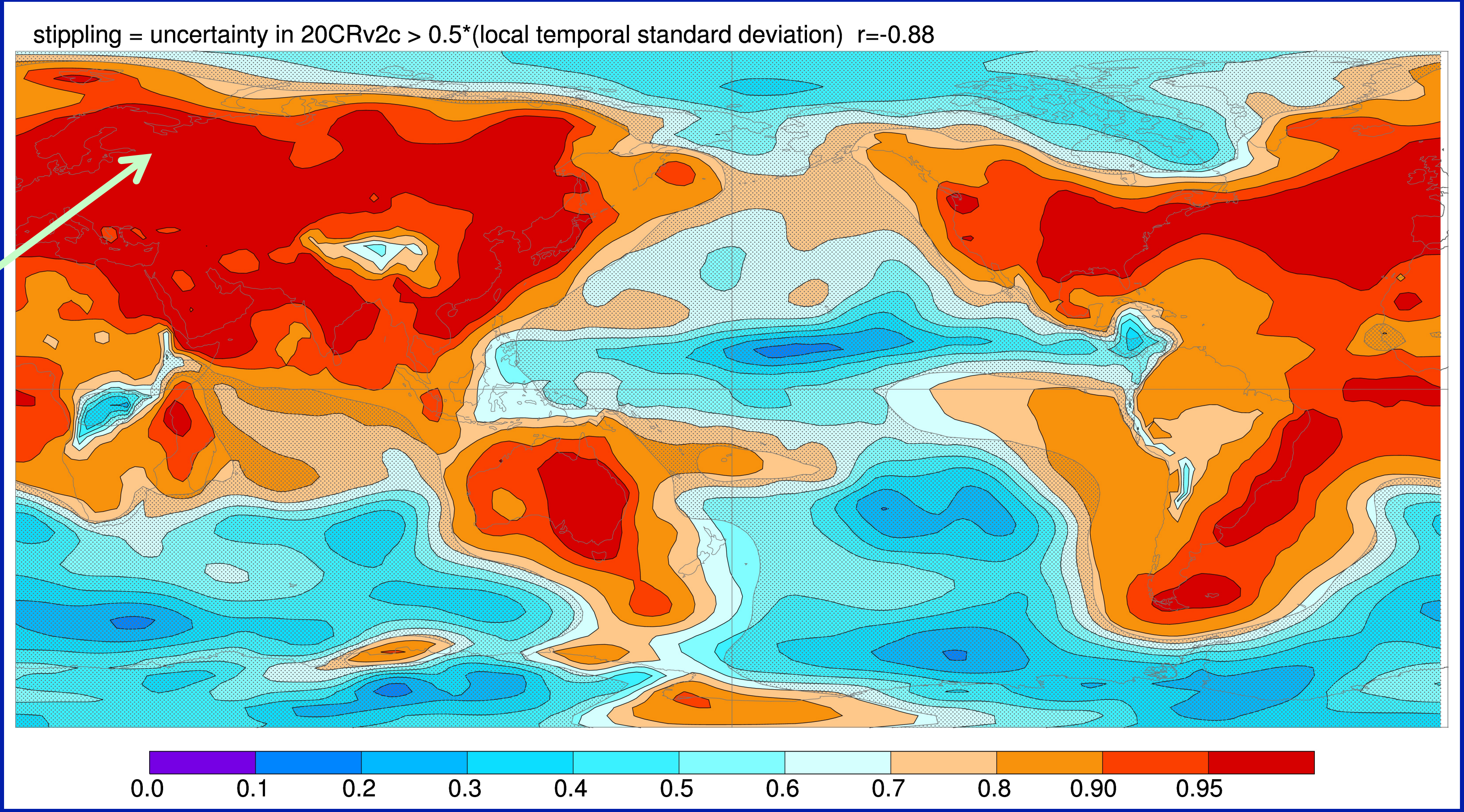
Grey dots=observations in 20CRv2c. Grey Fog=where 20CRv2c very uncertain.

20CRv2c differences with ERA-20C largest over land. Suspect latitudinally-constant covariance inflation is too small there.

Colors=Local Correlation between 20CRv2c and ERA20C daily Sea Level Pressure during 1912

Grey Stippling= 20CRv2c is uncertain

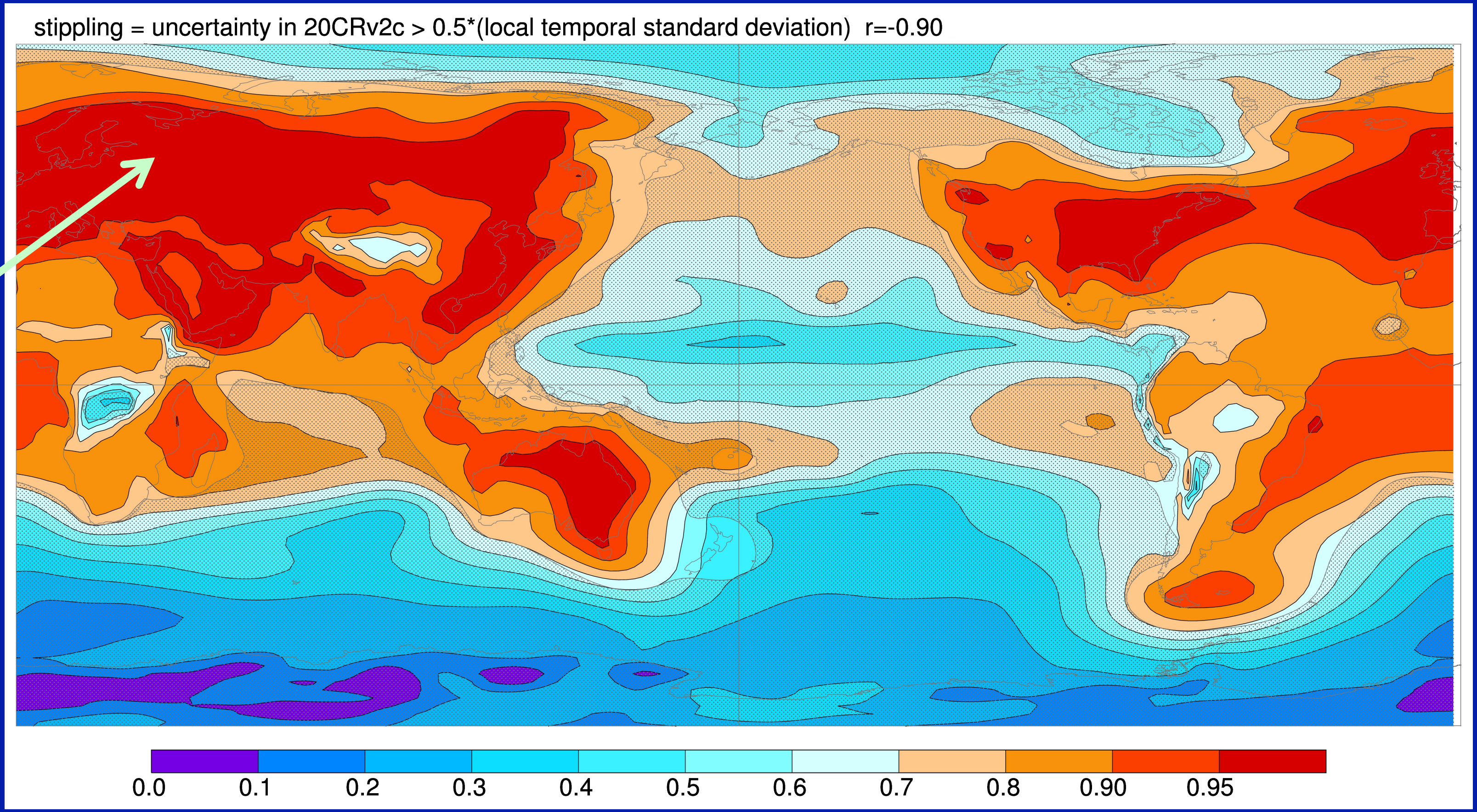
Local r >0.95



Uncertainty field from 20CRv2c ensemble standard deviation is a good proxy for the dataset agreement (pattern correlation = -0.88).

Colors=Local Correlation between 20CRv2c and ERA20C daily Sea Level Pressure averaged 1900-1919

Grey Stippling= 20CRv2c is uncertain



Local
r >0.95

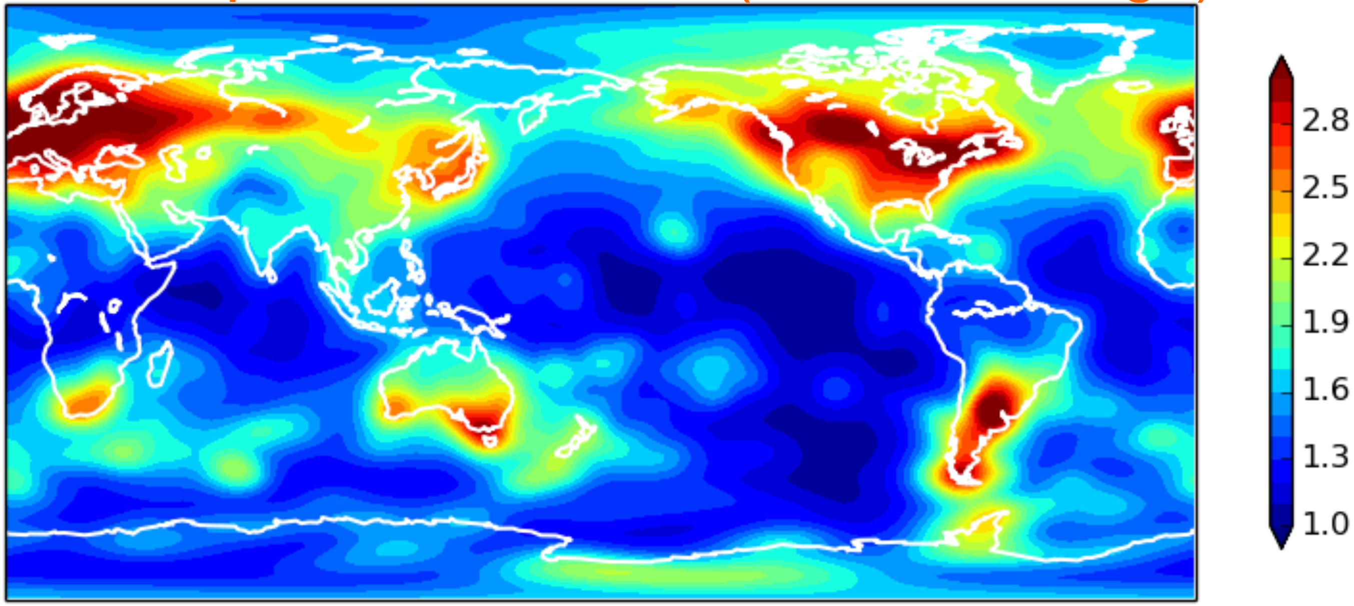
Uncertainty field from 20CRv2c ensemble standard deviation is a good proxy for the dataset agreement (pattern correlation = -0.9).

Developing 20CR version 3 Data Assimilation system

- Test Model: T126L64 GFS, circa 2013.
- 64 member EnKF
  - Varying covariance localization (based on estimated observation impact)
  - ‘Relaxation to prior spread’ inflation (larger where observations are dense).
  - New Quality Control that treats ob errors as non-Gaussian (based on Huber-norm VarQC algorithm developed at ECMWF).
  - No DFI - incremental analysis update (IAU) ending at analysis time instead.
- Sea-ice problem reduced.

New covariance inflation

Surface pressure inflation (annual average)



Relaxation to prior spread (RTPS) $\sigma^a \leftarrow (1-\alpha)\sigma^a + \alpha\sigma^b$

which implies $\mathbf{x}_i'^a \leftarrow \mathbf{x}_i'^a \sqrt{\alpha \frac{\sigma^b - \sigma^a}{\sigma^a} + 1}$

Here we use *a= 0.9*

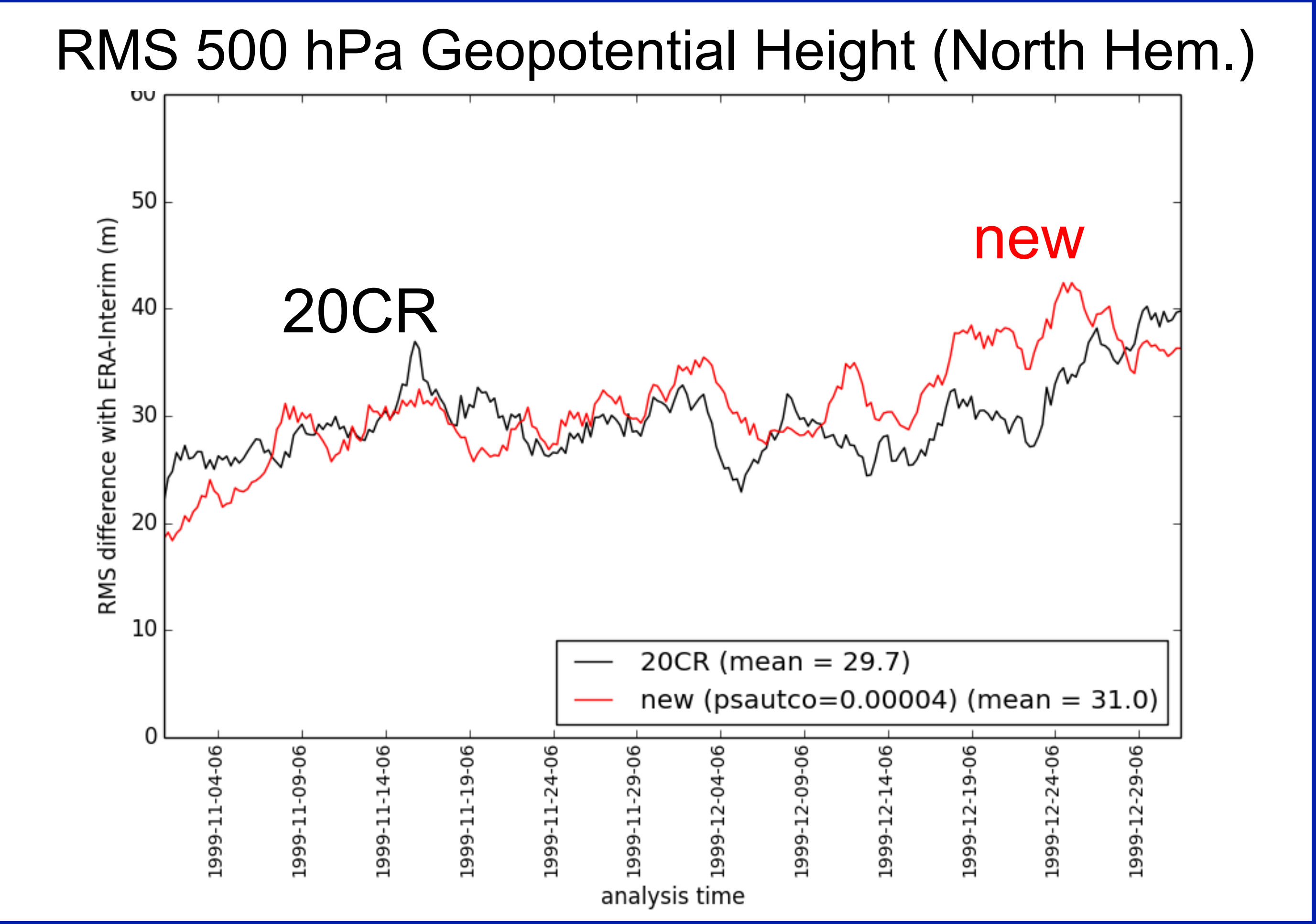
Previously, inflation was piecewise constant (NH,TR,SH)

Ref: *journals.ametsoc.org/doi/pdf/10.1175/MWR-D-11-00276.1*

Testing

- Start 1 September 1999 from 20CRv2 ensemble.
- Use same pressure obs and SSTs as 20CRv2, climatological sea ice.
- Run through 31 December 2000.
- Verify against ERA interim analyses.

First try using default GFS parameters

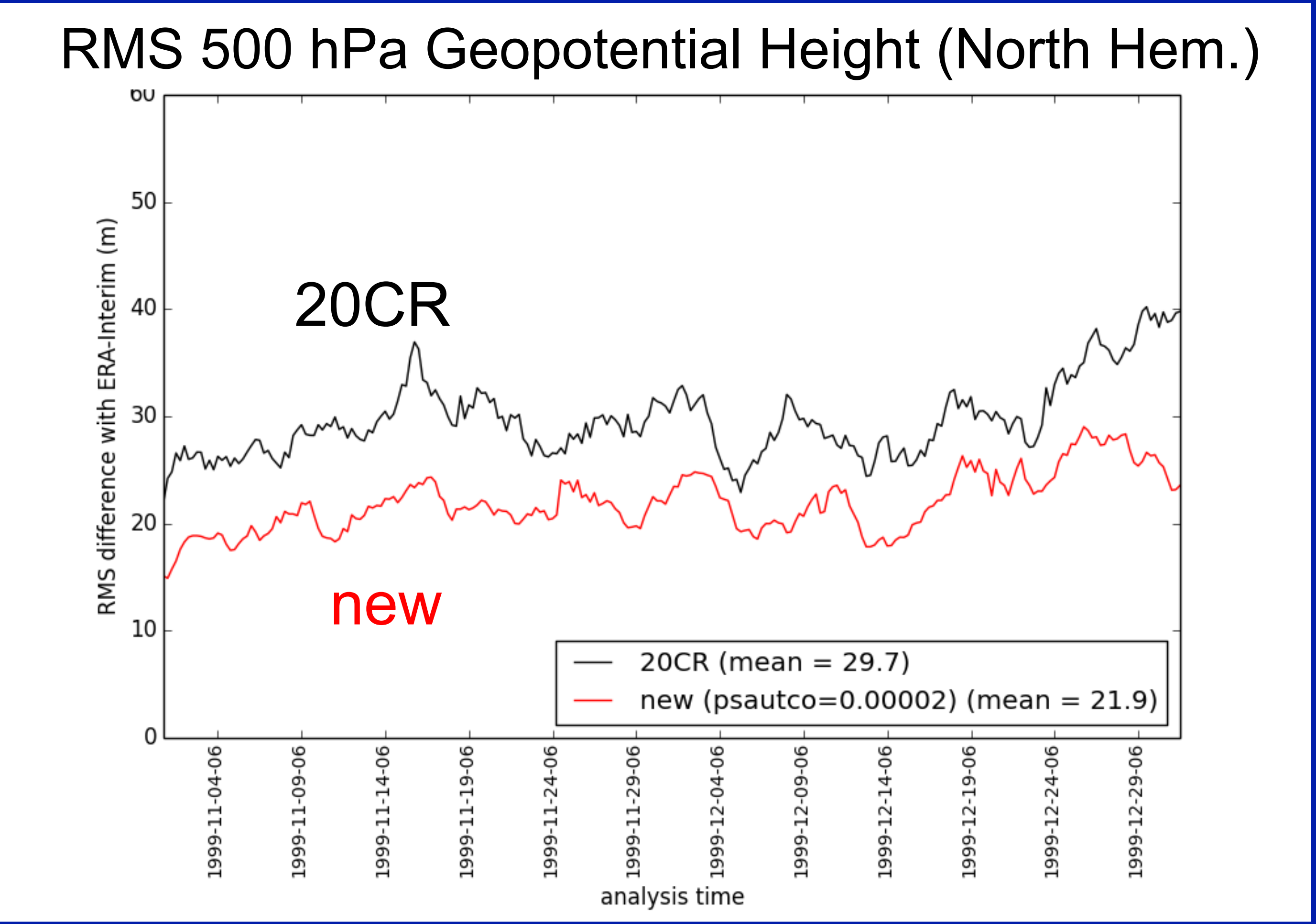


Worse than original 20CR. Large mid-upper tropospheric cold bias.

Cause: Not enough high cloud.

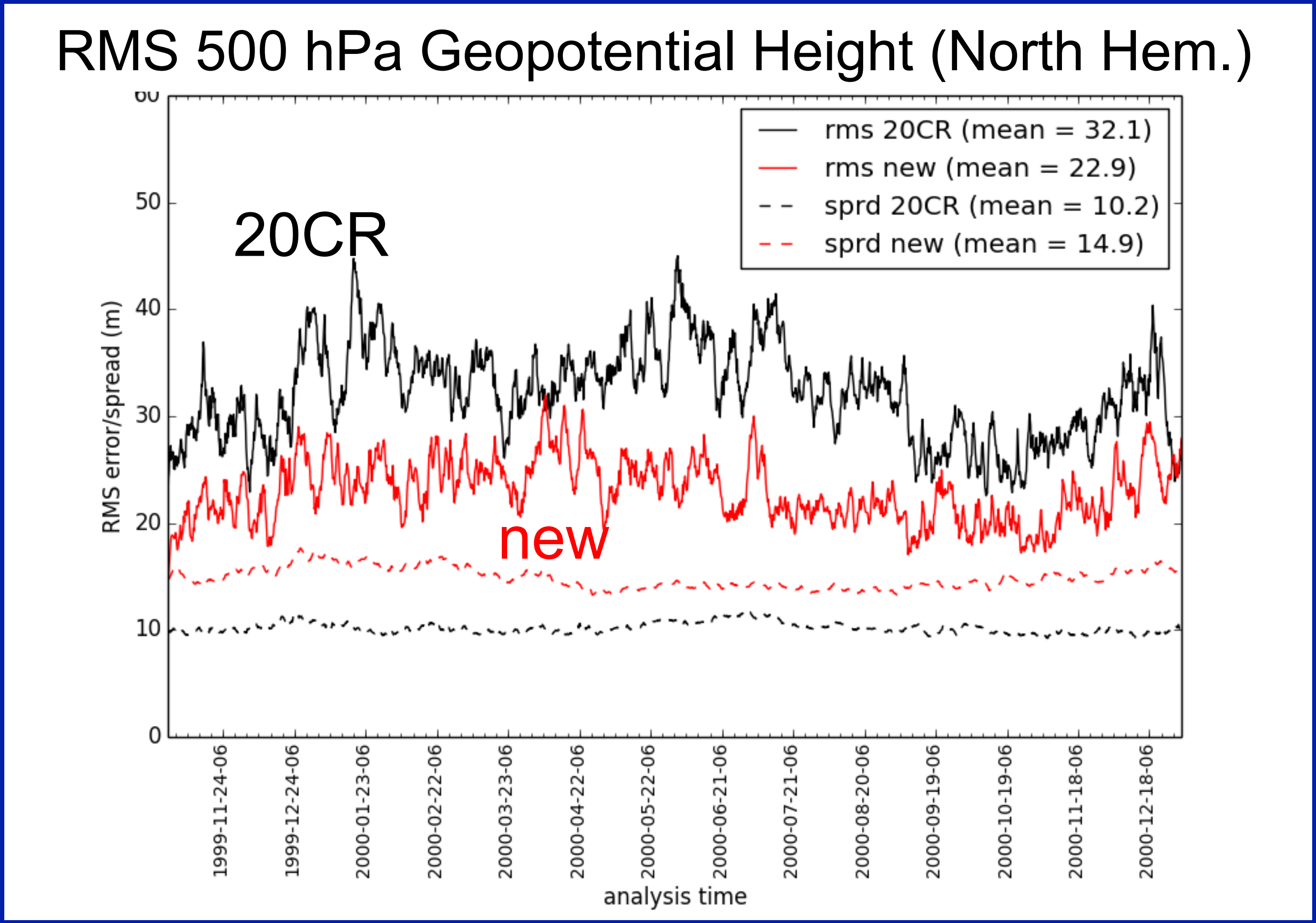
Solution: Reduce ice to snow auto-conversion rate in Zhao-Carr microphysics.

psautco halved



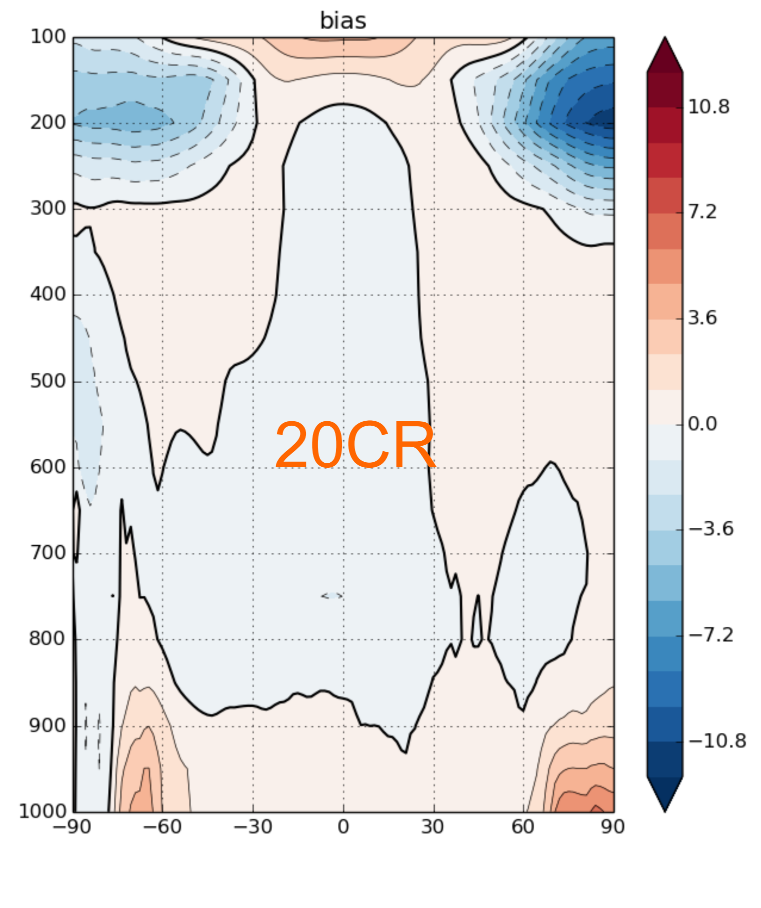
Now much better than 20CR in Northern Hemisphere (almost 25% reduction in error)!

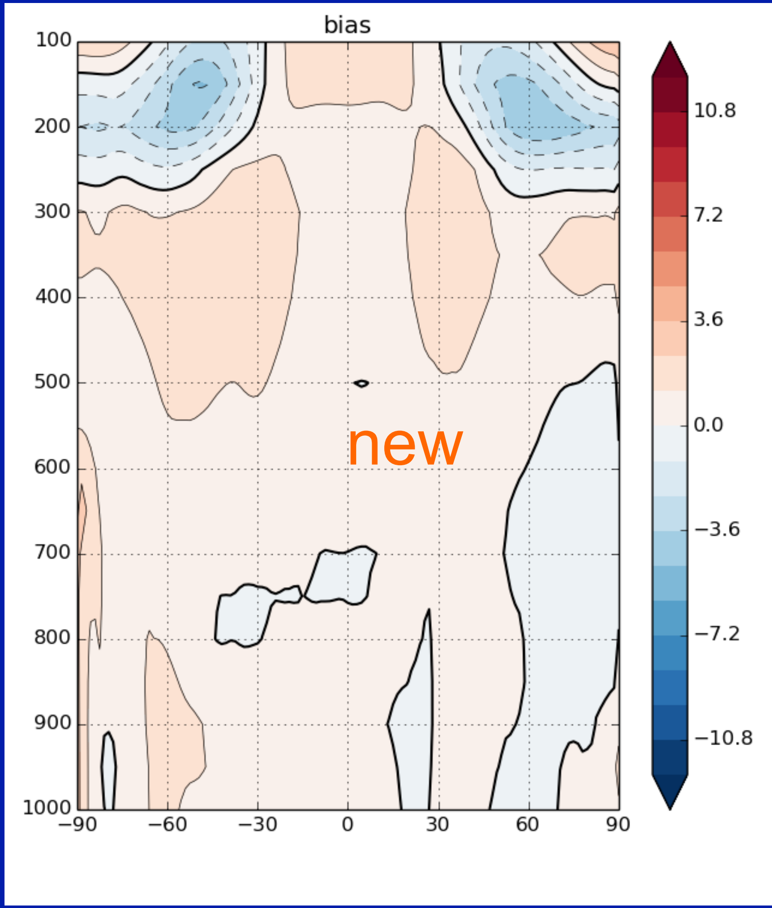
Results for full year



28% reduction in RMS,
spread increased by 45% (now much closer to error)

Zonal mean temperature bias





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

Historical Reanalysis Status and Plans

20th Century Reanalysis Project http://www.esrl.noaa.gov/psd/data/20thC_Rean

- **Fall 2014**: 1871-2012 (includes time-varying CO2, volcanic aerosols, GFS from NCEP). **Ensemble mean and spread and some individual member variables online now**.
  - http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.html (NOAA ESRL)
  - http://dss.ucar.edu/datasets/ds131.1 (NCAR)
  - http://portal.nersc.gov/20C_Reanalysis **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
  - British Atmospheric Data Center **Every member**

20CR v2c

- **Spring 2015**: 1851-2011

Very similar system to 20CRv2. 1851-2011

Corrects Sea ice issues using COBE-SST2 sea ice.

Uses ensemble of SODAsi.2 SST, more observations.

20CR v3

Spring 2017: 1851-2016

Higher resolution, improved algorithm and observational quality control

Development of Suite of NOAA Climate Reanalyses

Joint NOAA NCEP, ESRL, NCDC, Univ. of Colorado CIRES

- Tiered assimilation approach
  - *0. Boundary forced* (equivalent to AMIP); 1850-present
  - *1. Historical using only surface pressure*; 1850-present
  - *2. Modern using surface and conventional data*; 1946-present
  - *3. Satellite using conventional and satellite data*; 1973-present
- Assimilation System
  - Hybrid Global Statistical Interpolation [3D-Var]/Ensemble Kalman Filter (GSI/EnKF)
  - Possibly T254L64 (~50 km resolution) for 0-2
    - Higher resolution for 3. Satellite

20CR: version “1815”

20th Century Reanalysis "1815" implementation of Ensemble Filter Algorithm

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

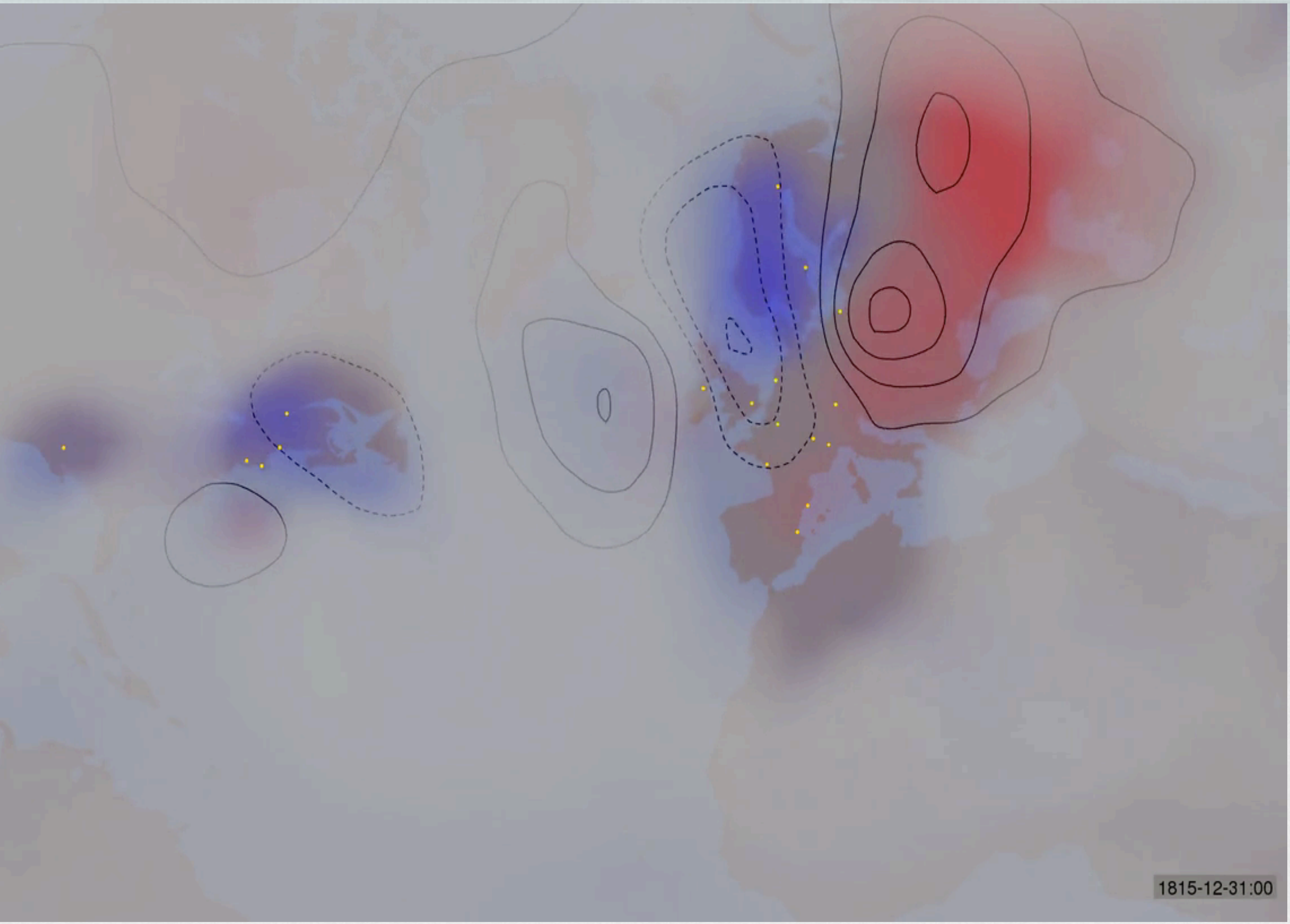
Algorithm uses an ensemble of GCM runs to produce the weight **K** that varies with the atmospheric flow and the observation network every 6 hours

Using 56 member ensemble,
prescribed 1861-1890 climatological boundary conditions:
COBE-SST2 monthly SST and sea ice concentration (*Hirahara et al. 2014*)

1815-1850: T62 (~200km), 28 level NCEP GFS08ex atmosphere/land model
9 hour forecasts for 6 hour centered analysis window
- time-varying CO_2 , 11 year repeating solar cycle, and

Specified monthly volcanic aerosol optical depth:
-No Aerosols, Gao et al (2008), Crowley (2008)

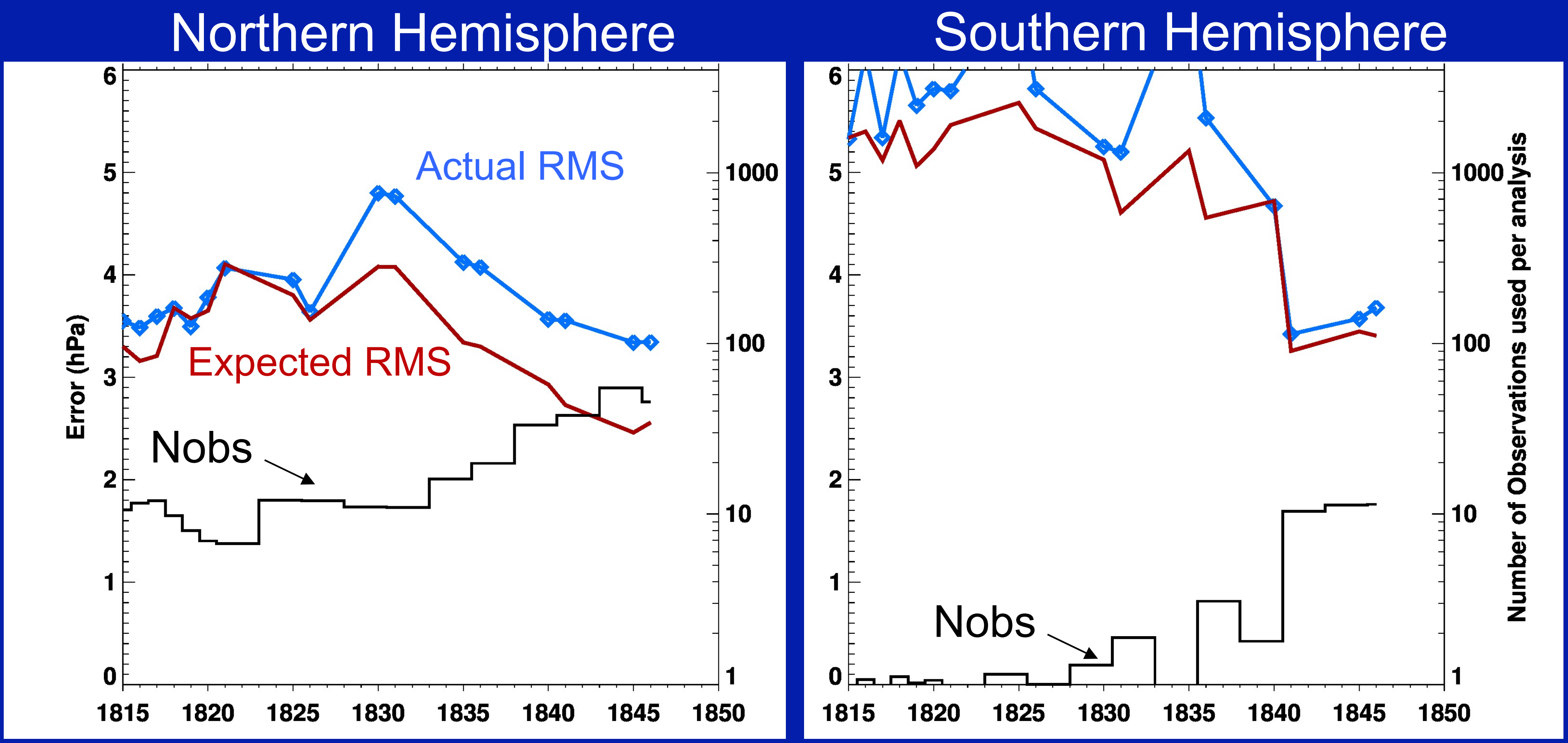
Every 5 years produced in parallel: 1816-1820,…, 1846-1850, after 14 month spin-up


1815-12-31:00

20CR:1815 Surface Pressure uncertainty estimate poleward of 20(S,N)

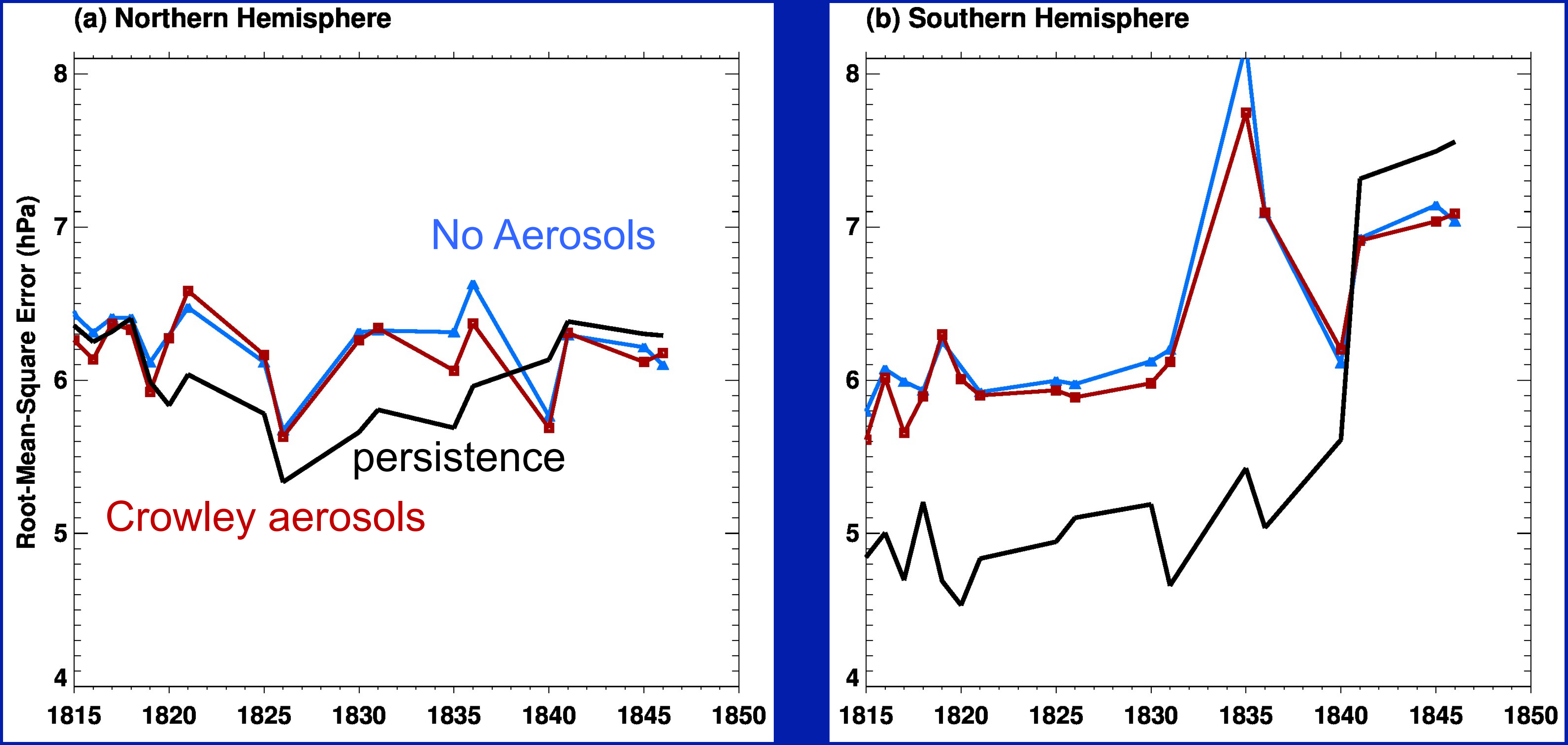
blue actual RMS difference

red expected RMS difference

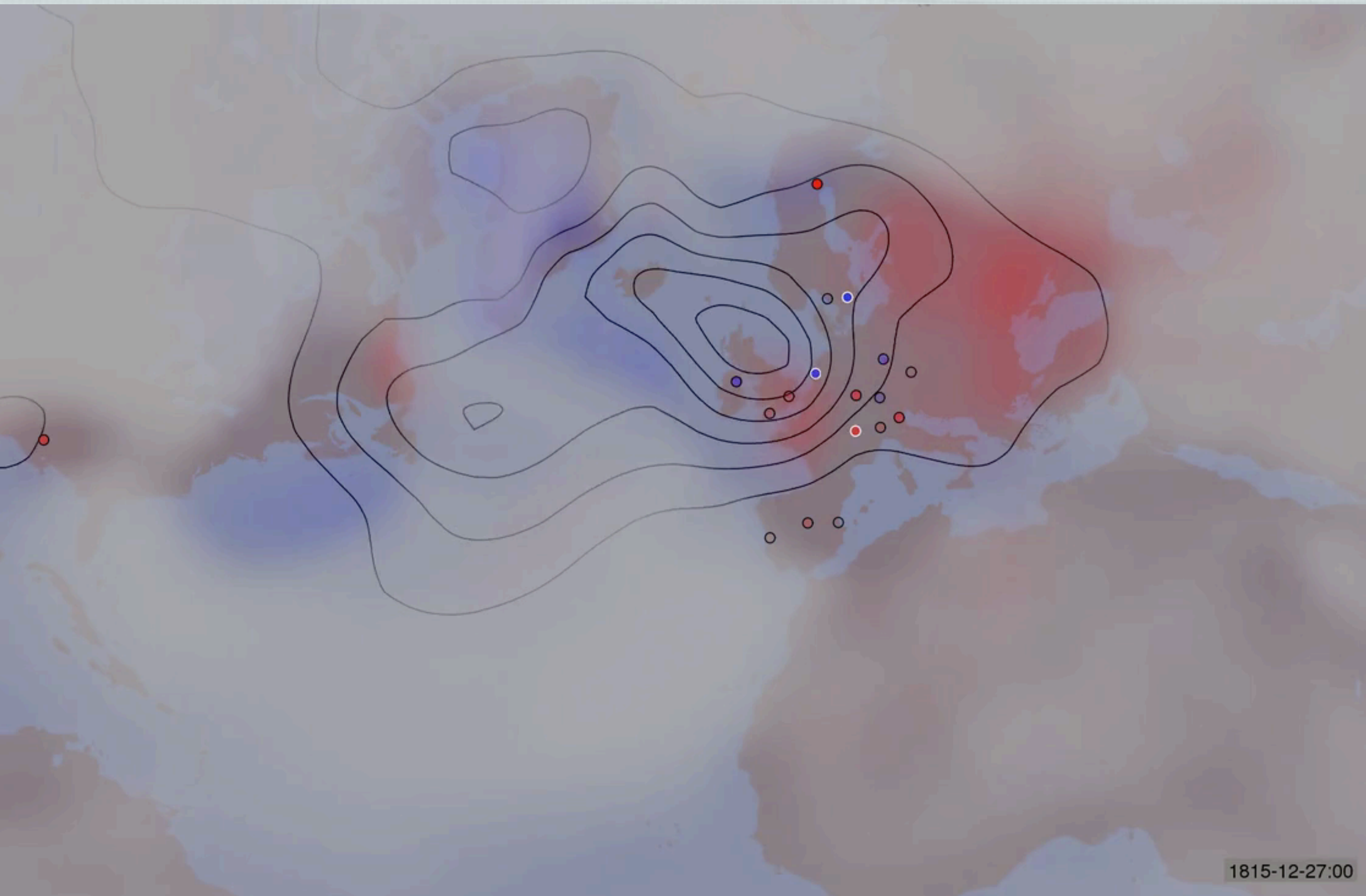


Uncertainty estimates are consistent with actual differences between first guess and pressure observations even in early 19th Century. Quantitative consistency degrades in NH after 1830s.

Root Mean Square difference of <u>Surface and Sea Level Pressure Observations</u> and 24 hour Forecasts from No Aerosols and Crowley Aerosols (*Jan-Dec*)



Northern Hemisphere 24 hr forecasts beat persistence even in 1815!
Southern Hemisphere <u>has an analysis</u> that produces forecasts comparable to persistence starting in 1840s with increased obs.

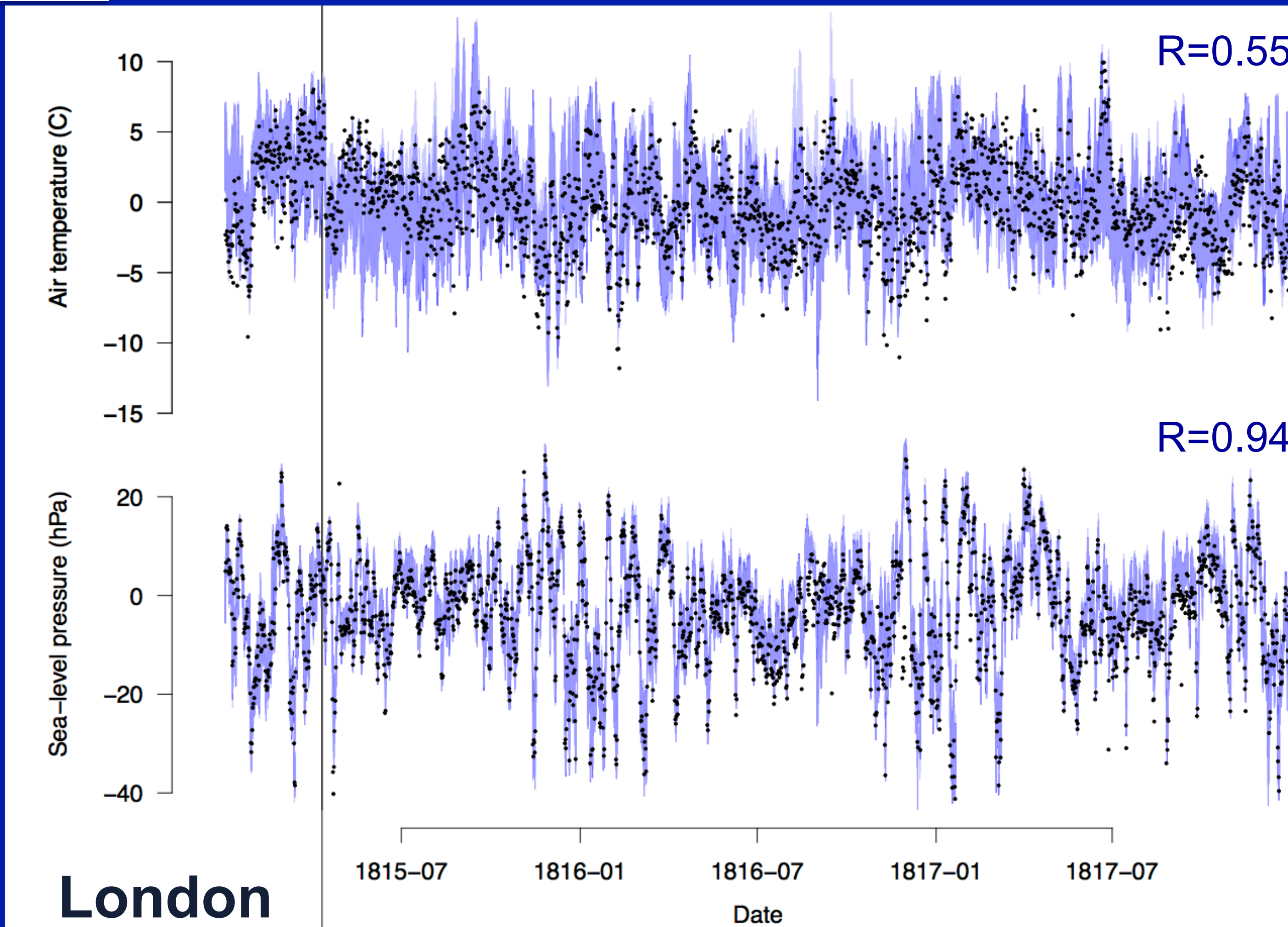

1815-12-27:00

Reconstructing the effects of Tambora 1815 and the Year Without a Summer of 1816

Comparison of anomalies from

Black dots: subdaily independent Air T and assimilated SLP from London

Purple swaths: 20CR-1815 ensemble range (1815 to 1817)



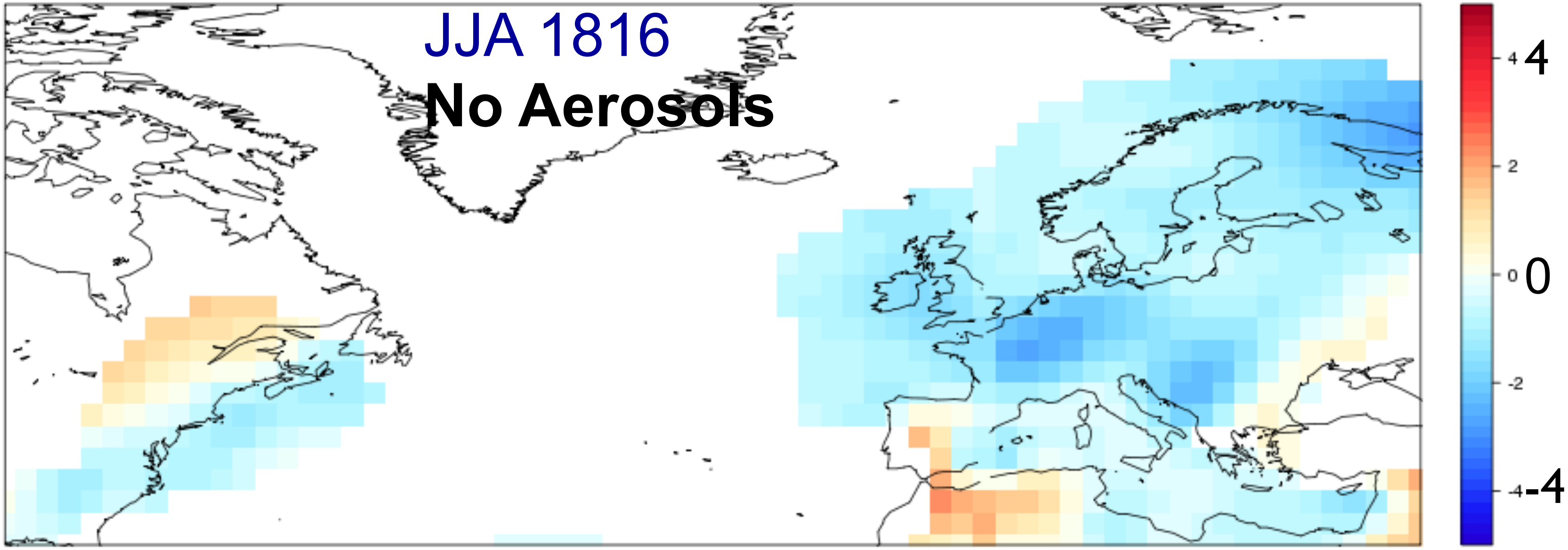
~4 other N. American and ~10 European Station, and ~10 Ship Obs assimilated each day.

2-4 times more possible.

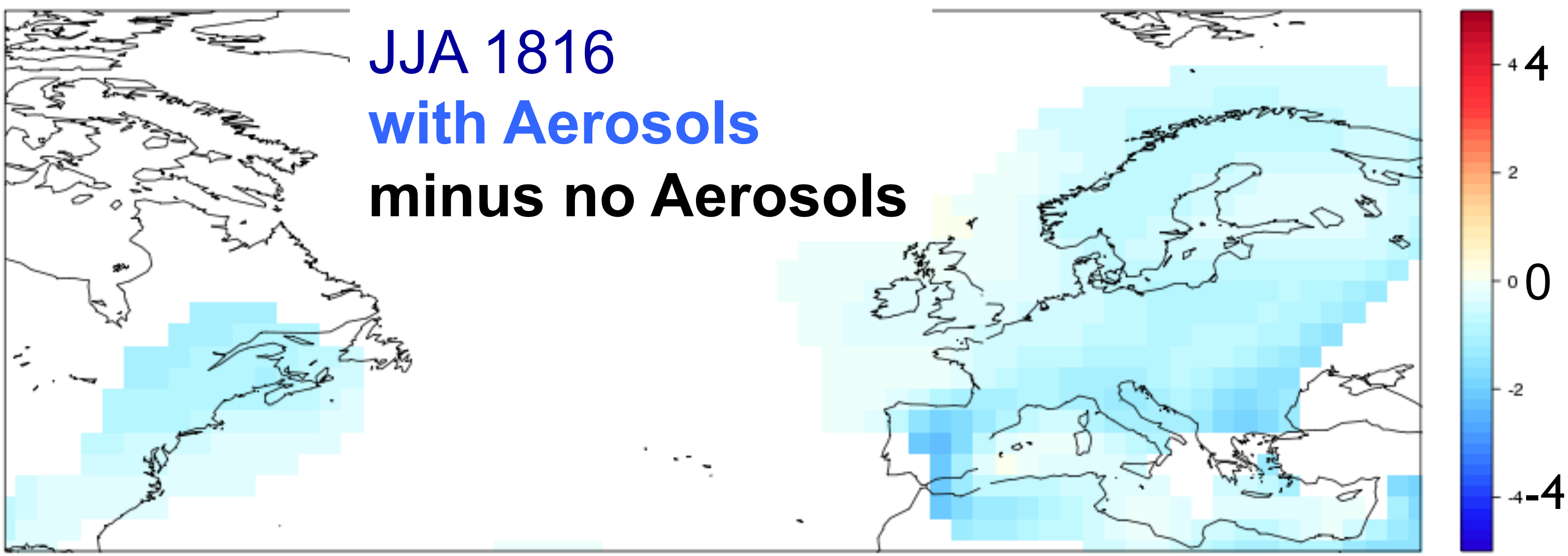
In regions such as Europe, 20CR-1815 compares well, showing skillful weather variability from the pressure observations. 1816 doesn't appear particularly anomalous in either dataset.

(Compo, Brohan, Whitaker, Broennimann, Brugnara, Allan, Sardeshmukh 2015)

(Compo, Brohan, Whitaker, Broennimann, Brugnara, Allan, Sardeshmukh 2015)

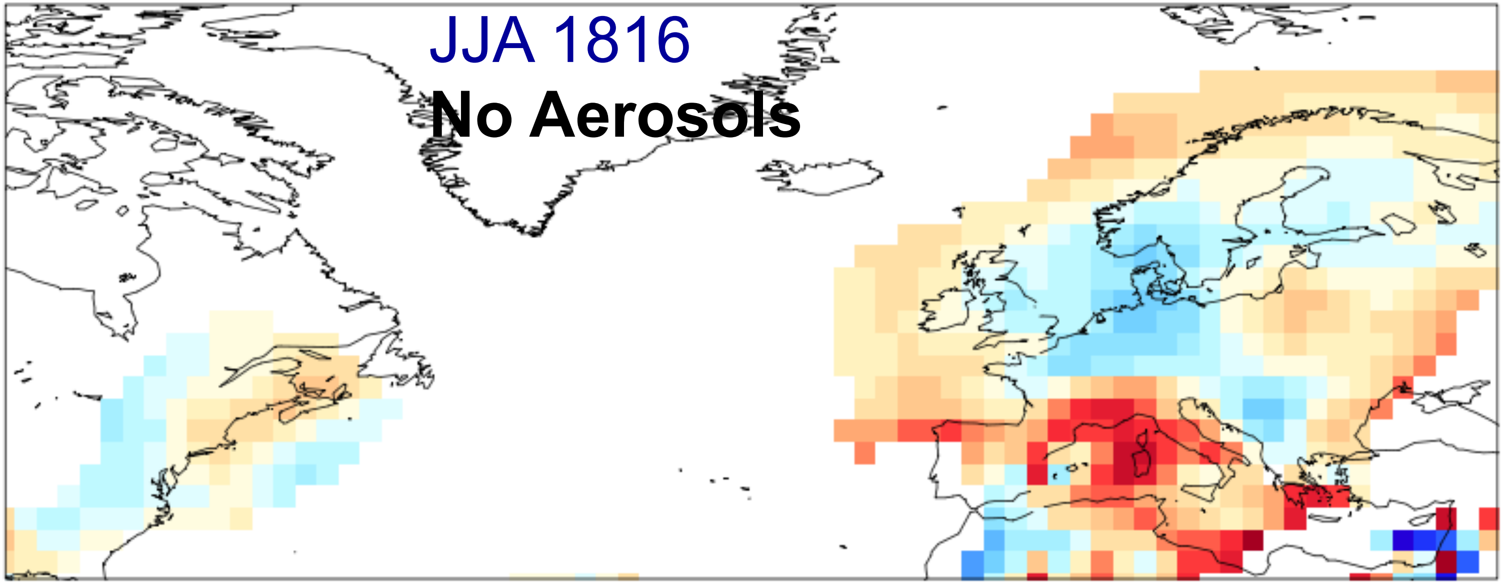


20CR JJA 1816 2m temperature anomaly (C) - **No Aerosols**

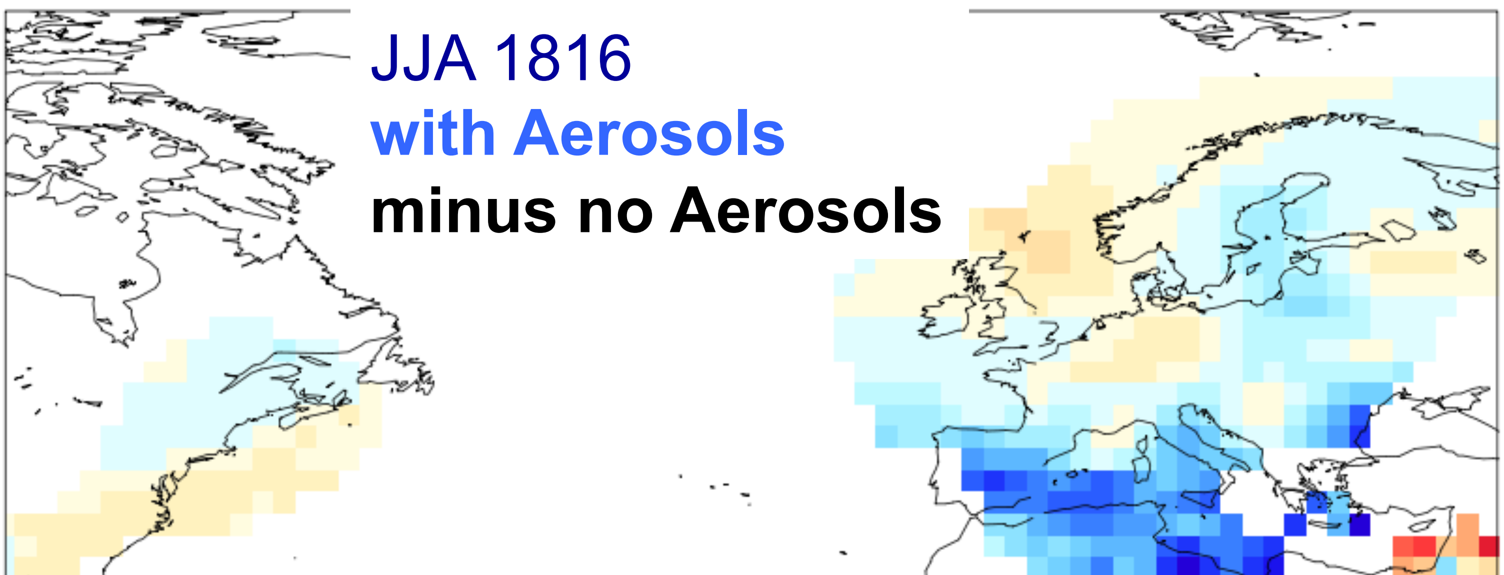


Additional cooling effect of adding Crowley aerosols is moderate but detectable.

(Compo, Brohan, Whitaker, Broennimann, Brugnara, Allan, Sardeshmukh 2015)



20CR JJA 1816 precipitation anomaly ratio – **No Aerosols**



Additional effect of adding Crowley aerosols

Conclusions

1. Demonstrated that surface-based reanalyses throughout the troposphere are feasible using advanced data assimilation and surface pressure observations.
2. Almost tripling the reanalysis record length from ~60 year to more than 160 years, allowing current atmospheric circulation patterns to be placed in a broader historical context. ☺
3. Used in many studies of climate, weather and water variability and extremes
4. Unexpected uses: e.g., iceberg risk during the Titanic (1912), coastal storm surge risk, and Tse fly variability.
5. Coordinating with NOAA's National Centers for Environmental Prediction and National Centers for Environmental Information on a consistent suite of NOAA Climate Reanalyses using upper-air and satellite observations.
6. Data assimilation algorithm improvements can be offset by a bad model. Need to improve both.
7. Future: improved algorithm, higher resolution model, longer span (200 years could be possible! Quality would depend on region).

Thank you to organizations contributing observations to ISPD:

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

The new 20CRv2c (1851-2013)

20CRv2 system but with

SODAsi.2: 18 member ensemble of daily SSTs (1846-2011)
COBE-SST2: monthly sea-ice concentration (1850-2011)

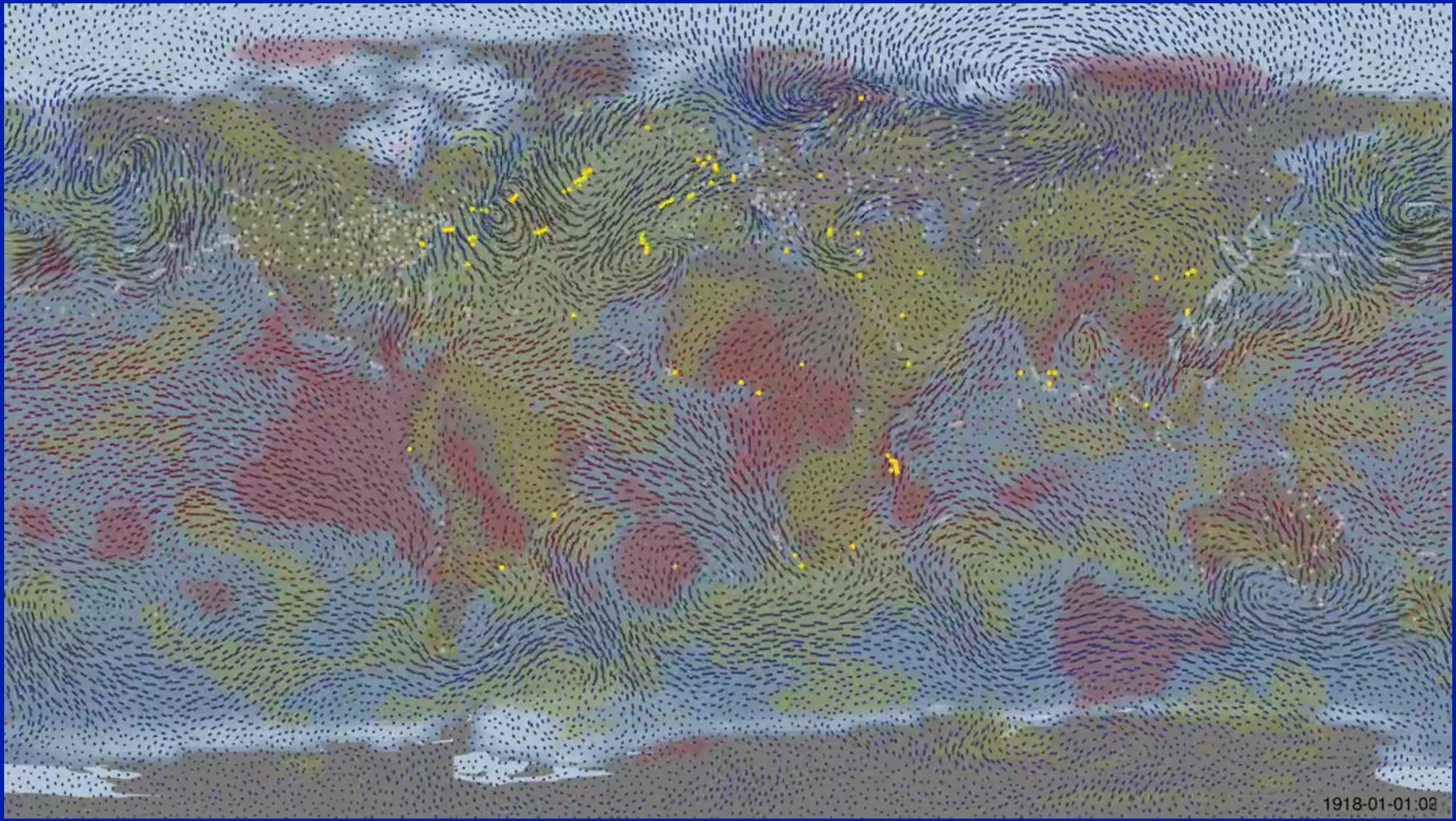
ISPDv3.2.9: International effort to recover 100s of new stations, new marine observations from Oldweather.org, ACRE data rescue,
over 33 new organizations contributing

Effect of some accounting for uncertainty in SST
Utility of new observations

After 14 month spin-up, 12 months produced for every 5th year.
1831, 1836,…,1851,1856,…,2007 complete

Compare to 1918 using same configuration but with
No Oldweather observations (spin up from July 1917)

January 1918
(20CRv2c worse than No Oldweather)
(20CRv2c improves upon No Oldweather)



Grey dots=observations in 20CRv2c and in No OldWx

Gold dots=New observations from Oldweather.org used in v2c

Vectors=low-level wind, red=warm, blue=cold

P. Brohan, UK Met Office

5 Quality Control steps are part of the 20CR data assimilation system

1. **Plausibility check**: reduce to SLP, is observation reasonable (between 880 and 1060 hPa)?
2. **Background check:** check difference of observation and first guess. Flag $|y_o - x_b| > 3(\sigma_o^2 + \sigma_b^2)^{0.5}$
3. **Buddy check**: can return observations that fail the Background check.
4. **Bias correction of stations**
5. **Thinning**: F-test σ_a^2/σ_b^2 before each observation is assimilated, only use observations that significantly reduce spread.

Buddy Check makes an analysis using only one observation at a time

1. Make an analysis using only the current *n*th observation.

   x_a = single observation analysis

2. Is the analysis error less than the first guess error when evaluated using the neighboring *k* observations within 1000 km

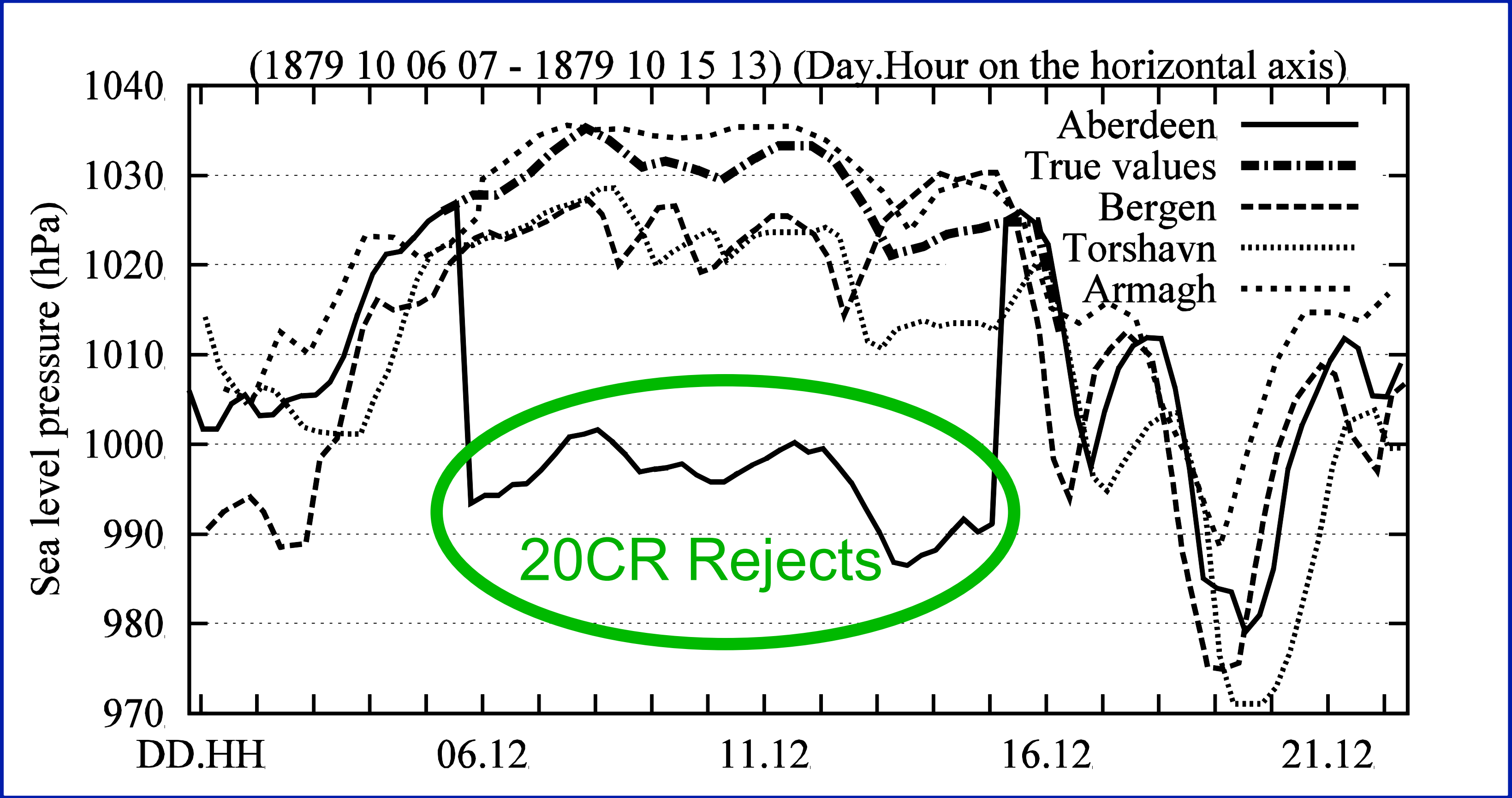
$$\Sigma_{k\neq n}|y_{ok}-x_a|^2 < \Sigma_{k\neq n}|y_{ok}-x_b|^2\ ?$$

Yes: Retain the observation even if it failed the Background check and allow it to be a buddy on next iteration.

No: Reject the observation, even if it passed the Background check.

Example of correctly rejected data

SLP from multiple stations in the eastern North Atlantic
6-10 October 1878



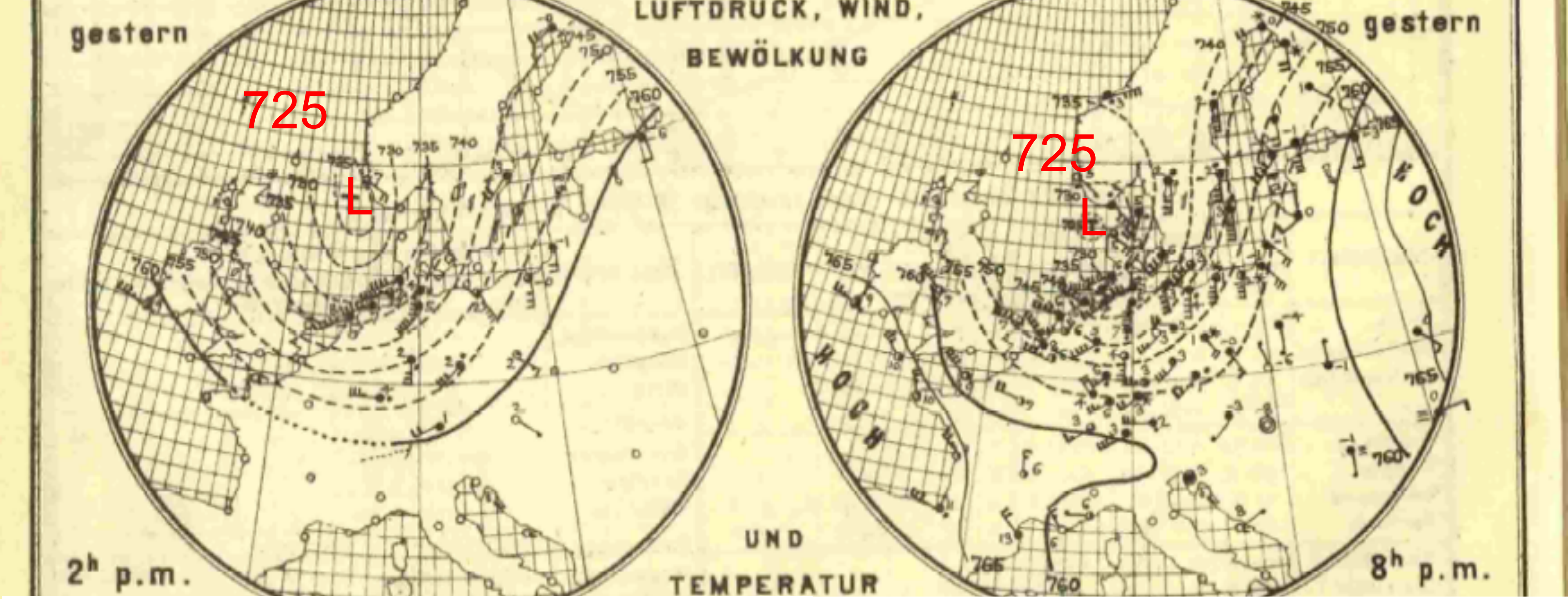
The extreme pressure gradient leads to large wind estimates in this decade if the data are used (Krueger et al. 2013).
20CR does not show such extremes.

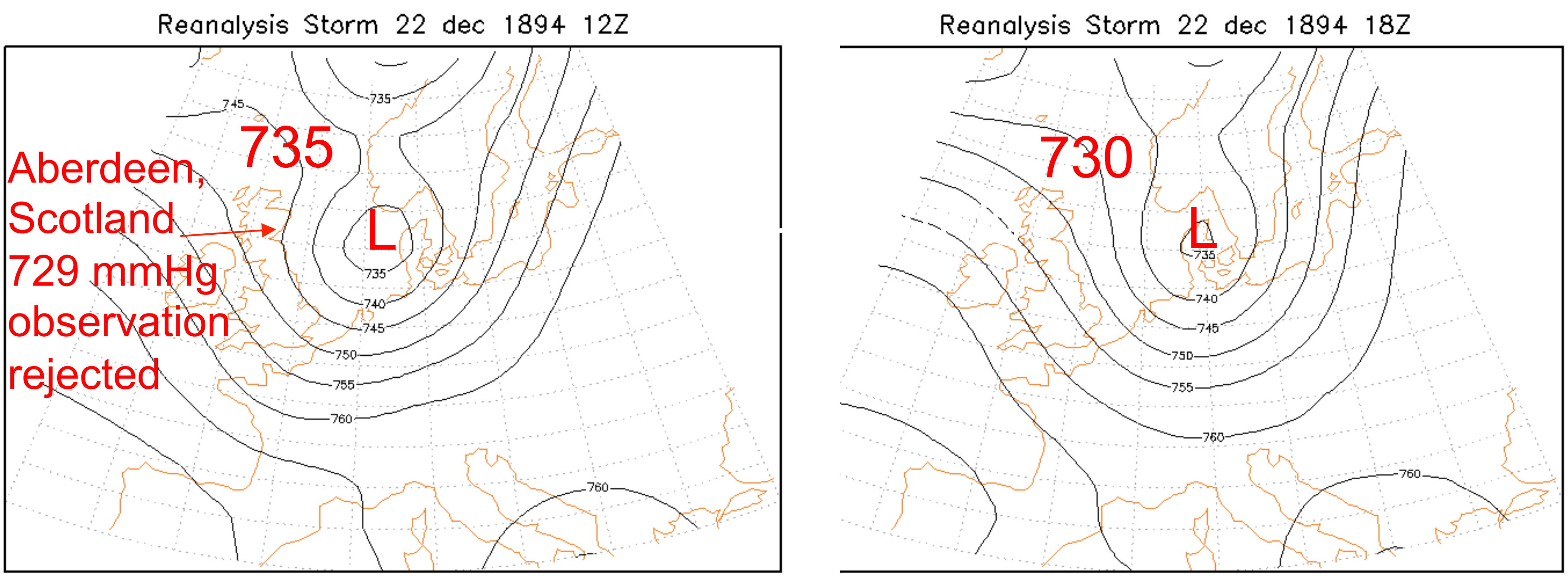
Wang et al. 2013

De Storm van 1894 (Zenit. 2010)

Henk de Bruin and Huug van den Dool

Frank Beyrich and **Britta Bolzmann** (DWD) provided 1894 weather maps of the *Seewarte Hamburg*

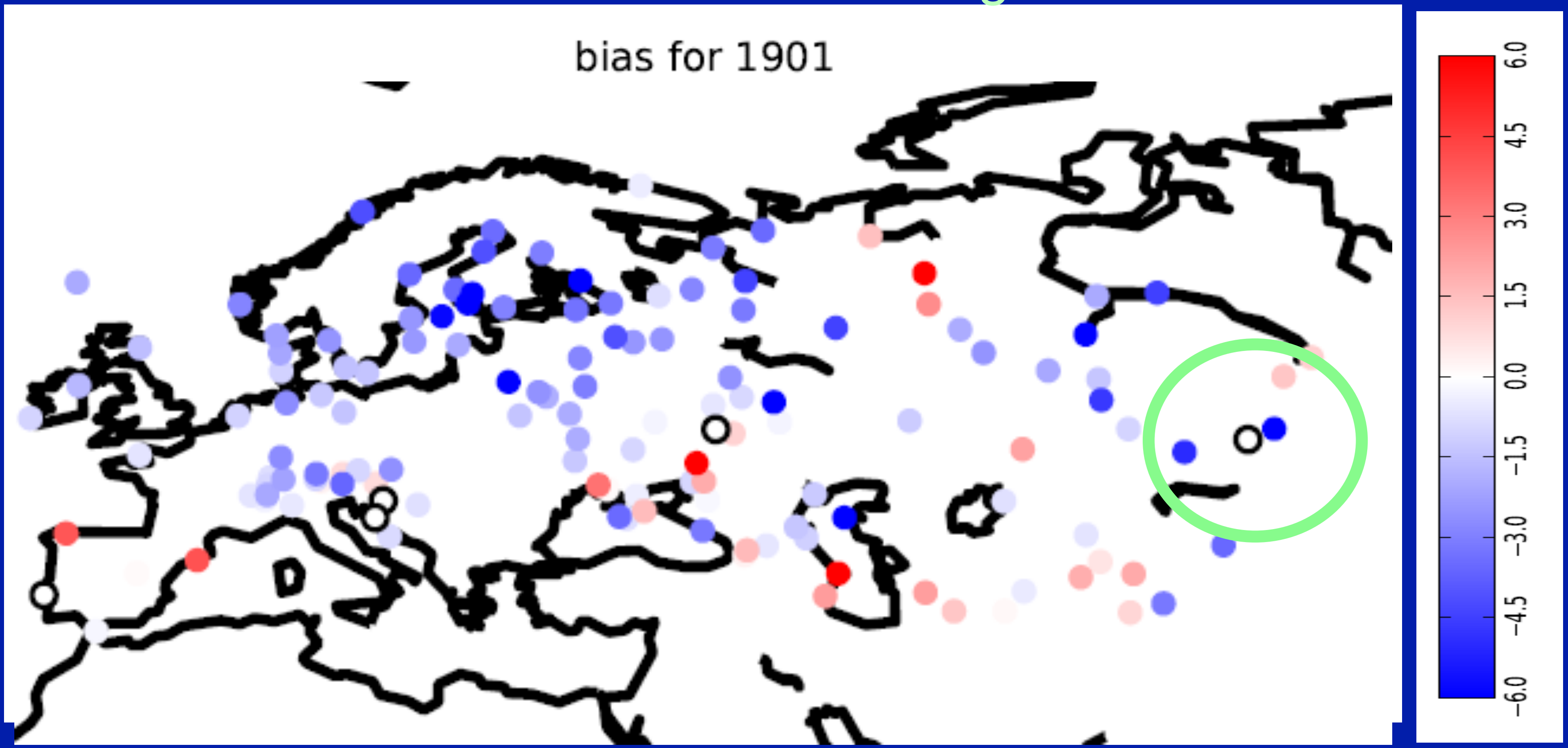




De Bruin and van den Dool (2010)

Station Bias correction algorithm corrects for Undocumented station moves, elevation errors

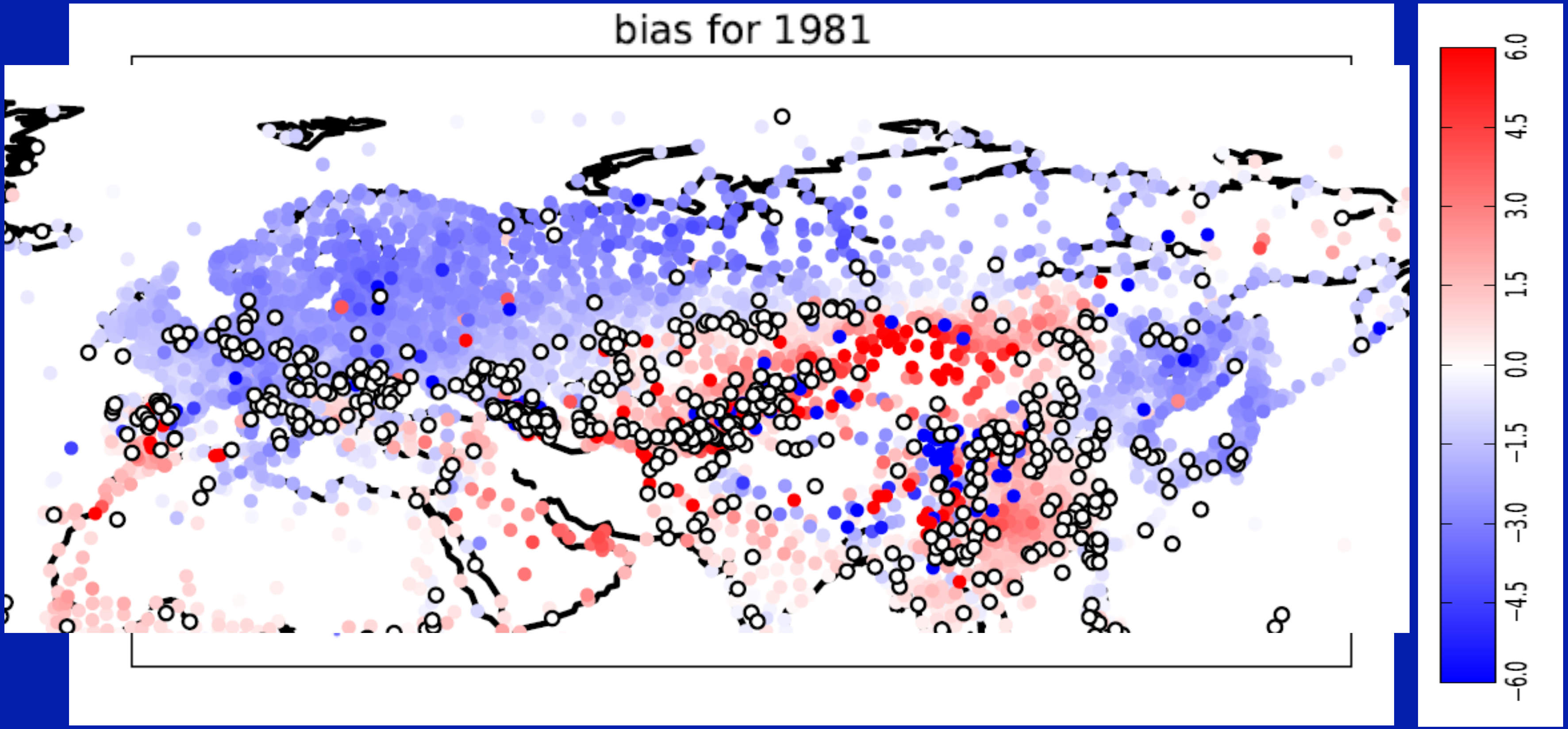
Remove statistically significant (paired t-test) differences between station observation and first guess.



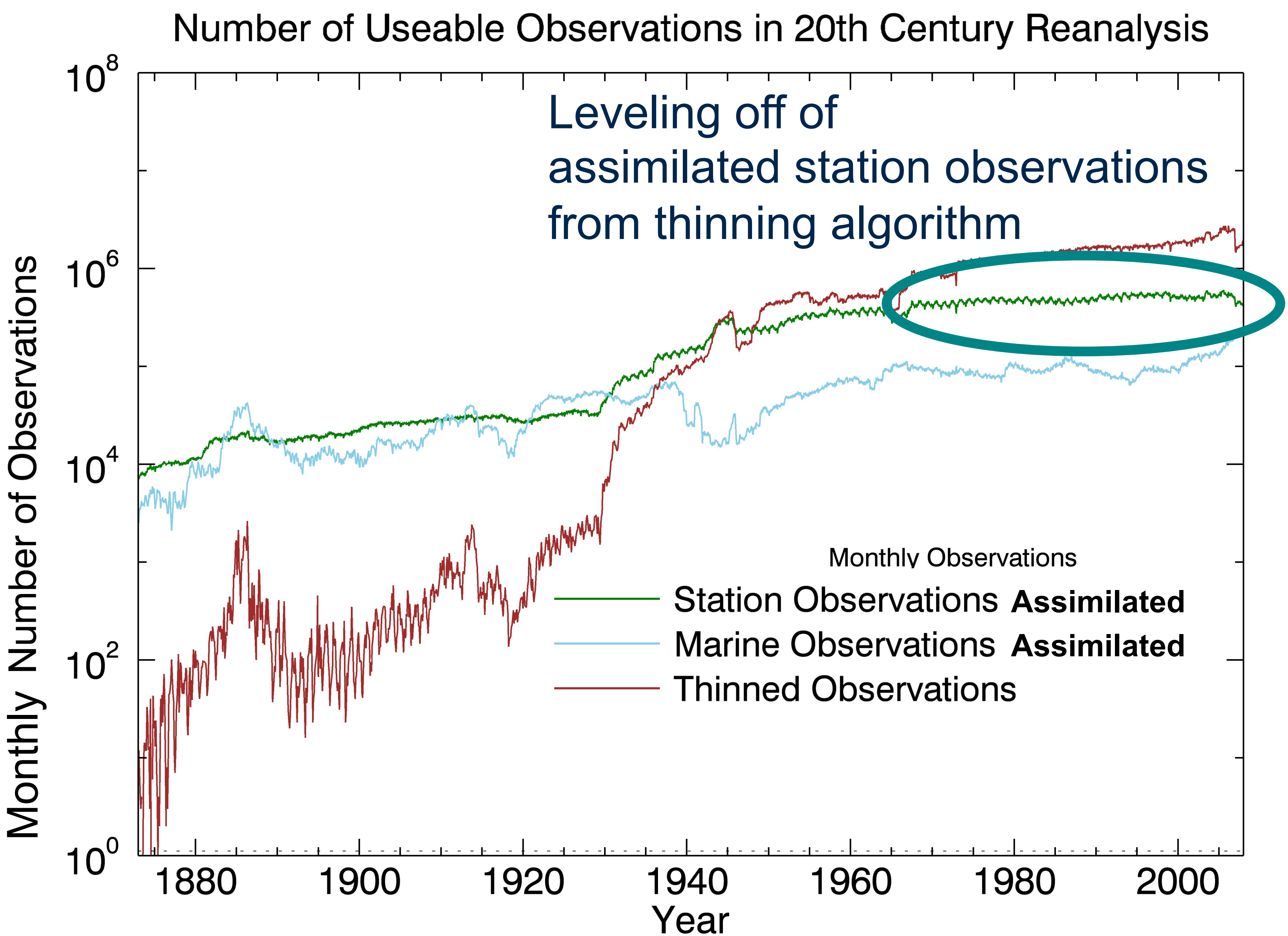
Large-scale coherence of the bias suggests large-scale model error may be attributed to stations. National boundaries suggest some issues may be national network wide.

Station Bias correction algorithm for such issues as undocumented station moves, elevation errors

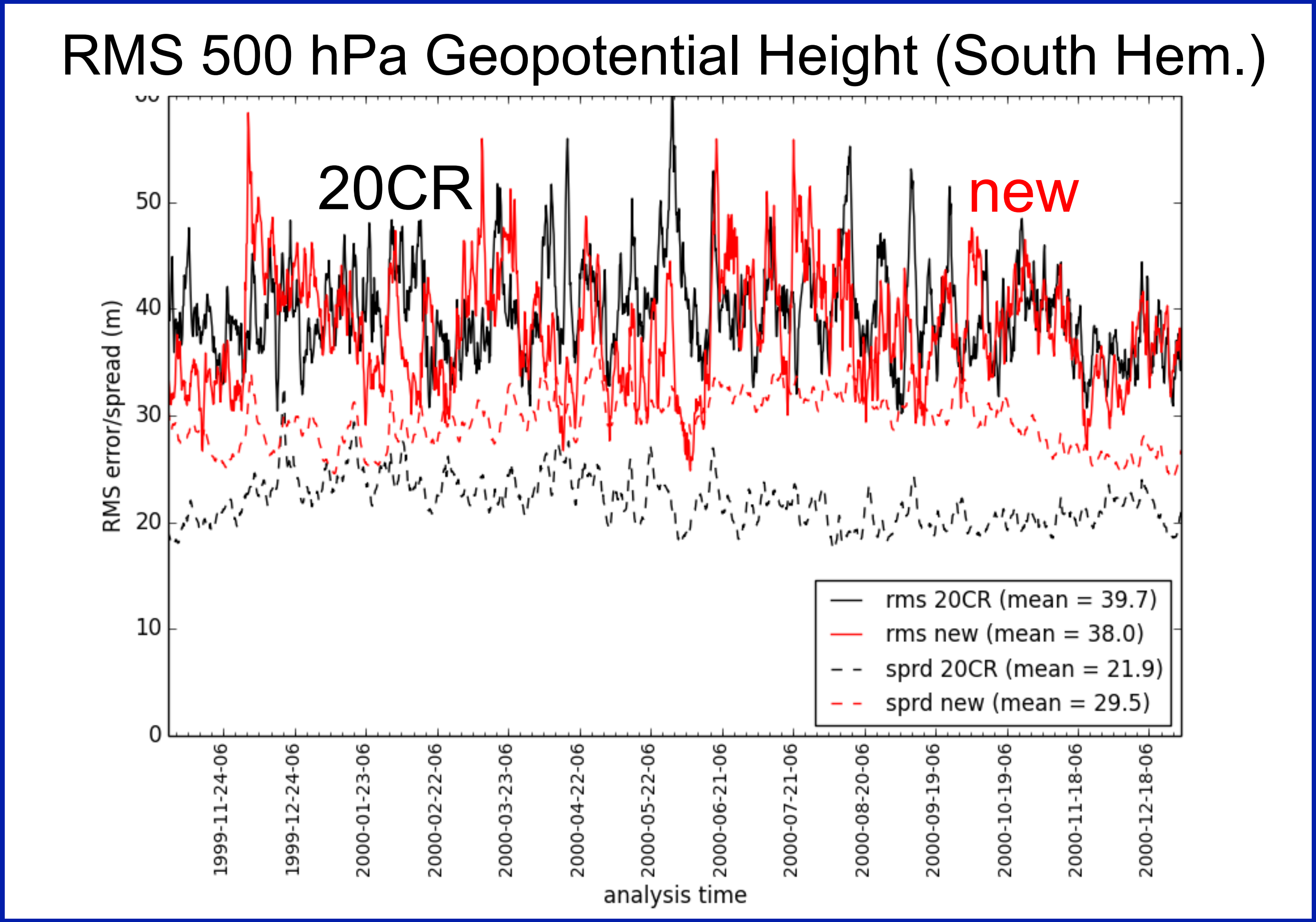
Remove statistically significant (paired t-test) differences between station observation and first guess.



Large-scale coherence of the bias suggests large-scale model error may be attributed to stations. National boundaries suggest some issues may be national network wide.


Number of Useable Observations in 20th Century Reanalysis
Leveling off of
assimilated station observations
from thinning algorithm
Monthly Number of Observations
10^8
10^6
10^4
10^2
10^0
1880
1900
1920
1940
1960
1980
2000
Year
Monthly Observations
Station Observations **Assimilated**
Marine Observations **Assimilated**
Thinned Observations

Results of v3 test for full year



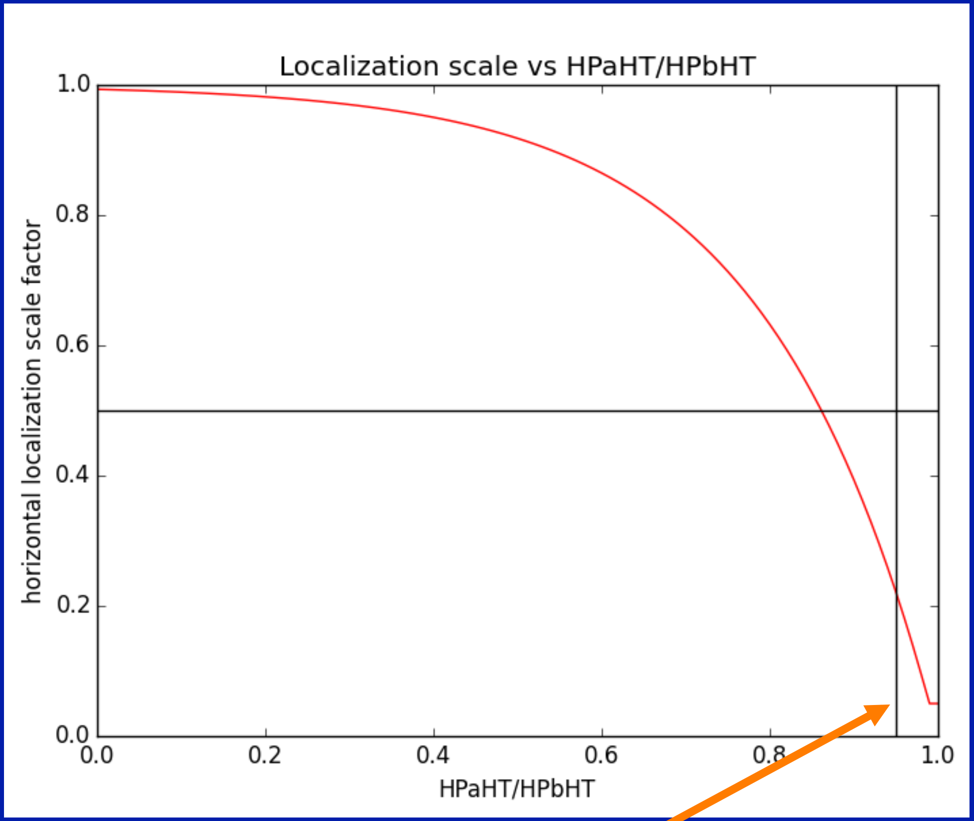
Not much reduction in RMS for Southern Hemisphere (but spread to RMS consistency improved)

Adaptive Covariance Localization

Length to zero influence is 4000 Km times ratio of

Analysis spread $HPaH^T$ from assimilating one observation to
First Guess spread $HPbH^T$.

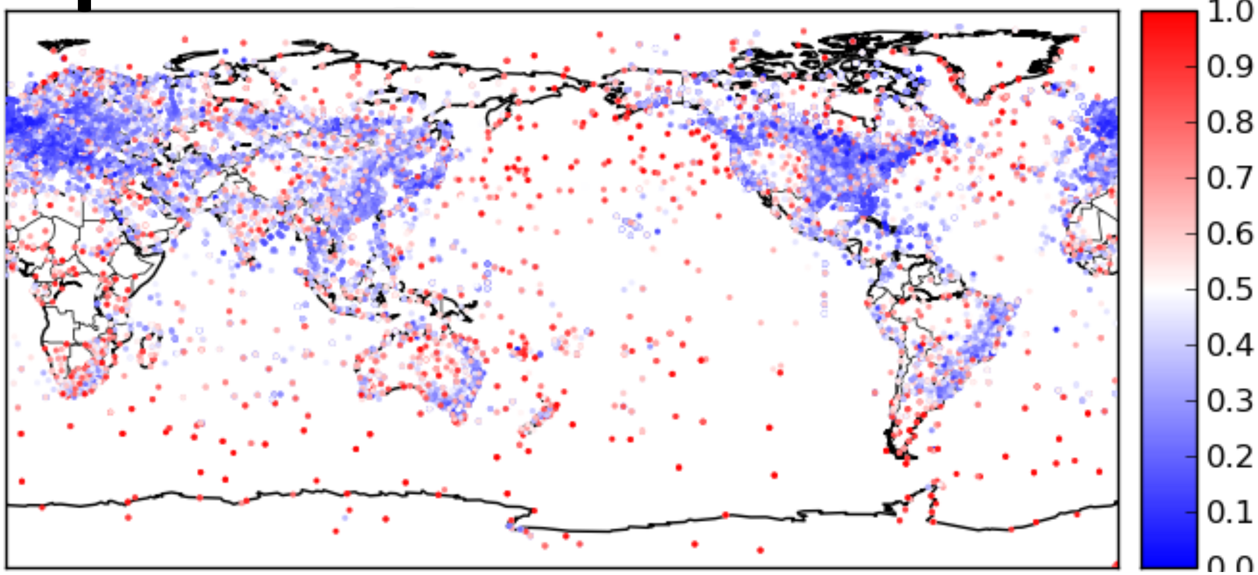
Analysis becomes First Guess for next observation.



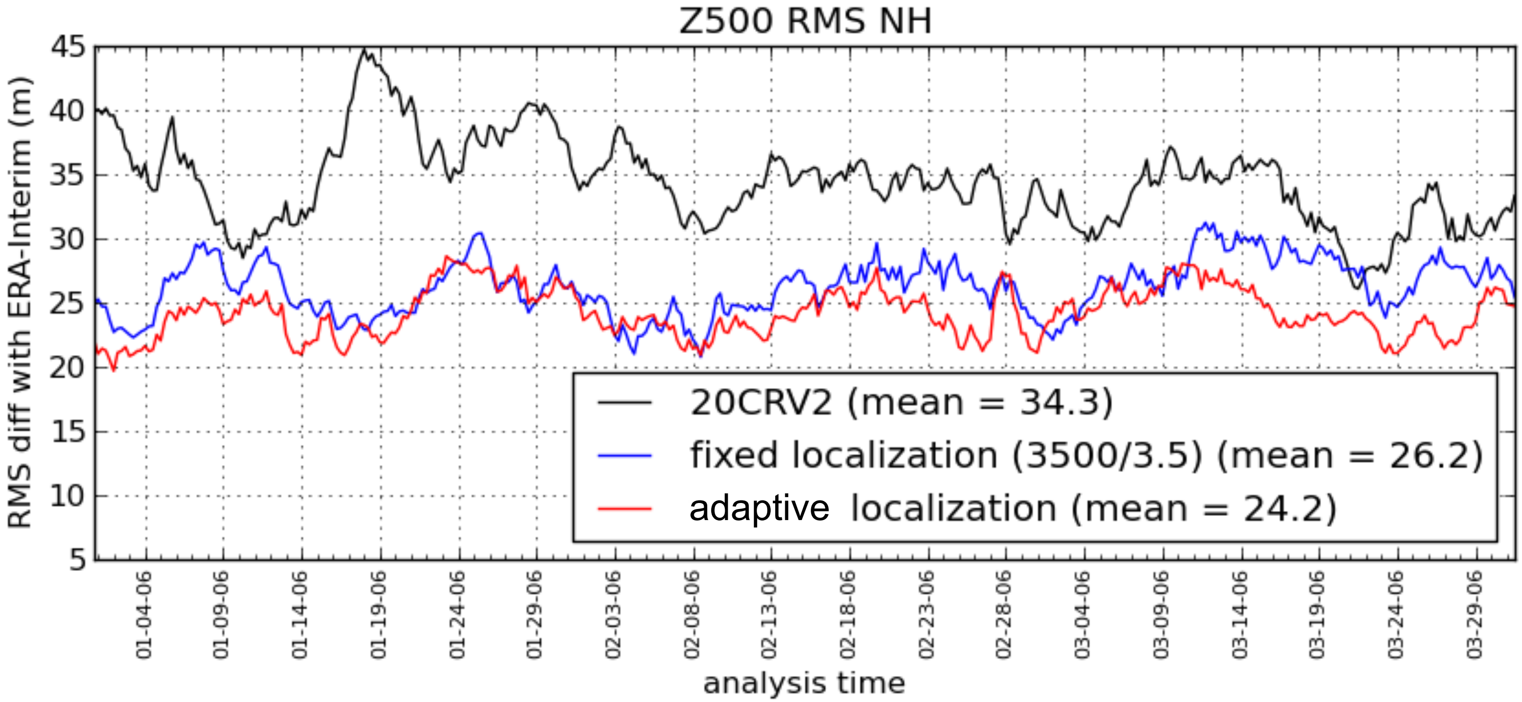
20CRv2 (and 2c) uses box car function.

Adaptive localization

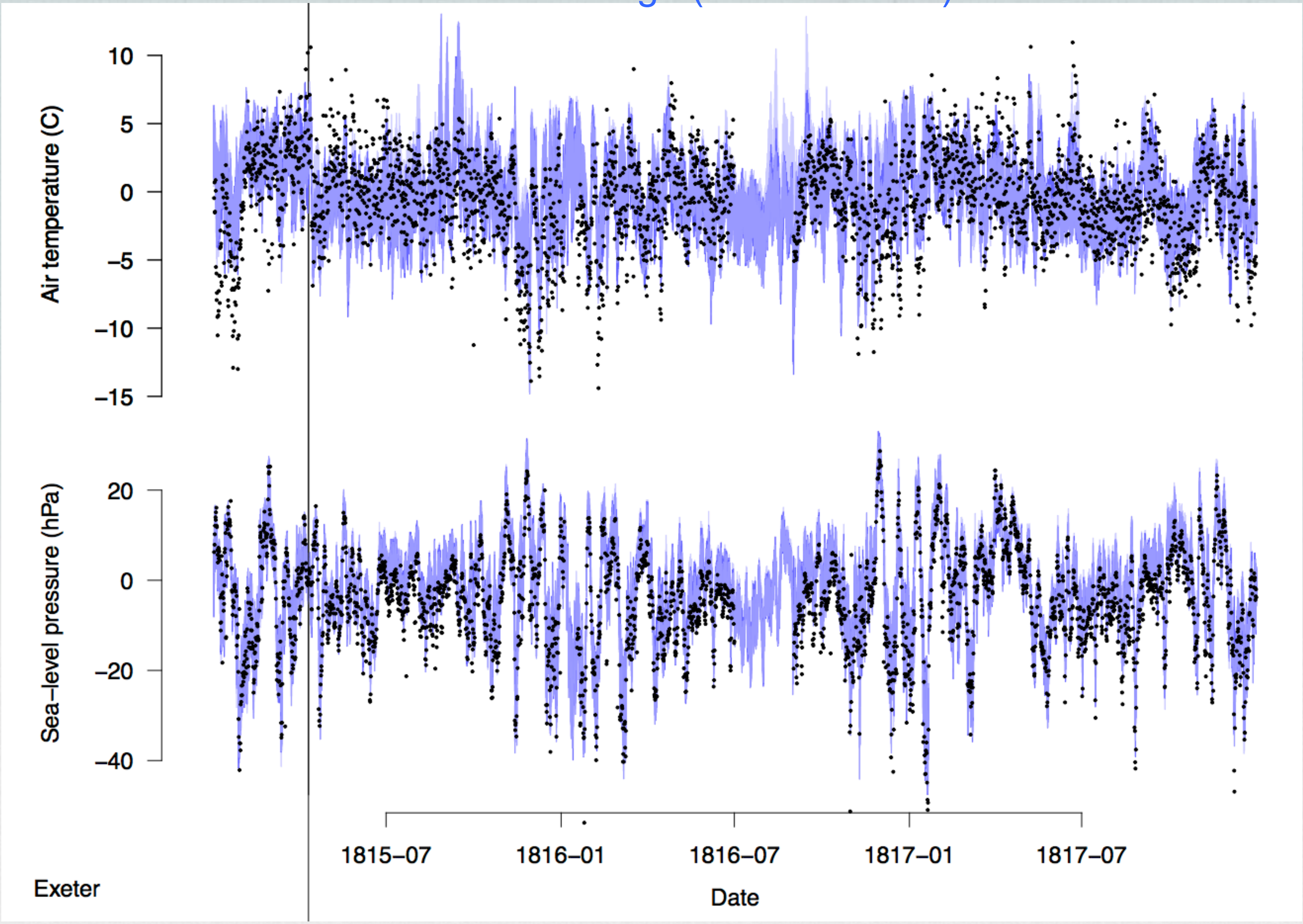
Dots are ps ob locations.
Blue → short localization
Red → broad localization



Adaptive better about 85% of the time

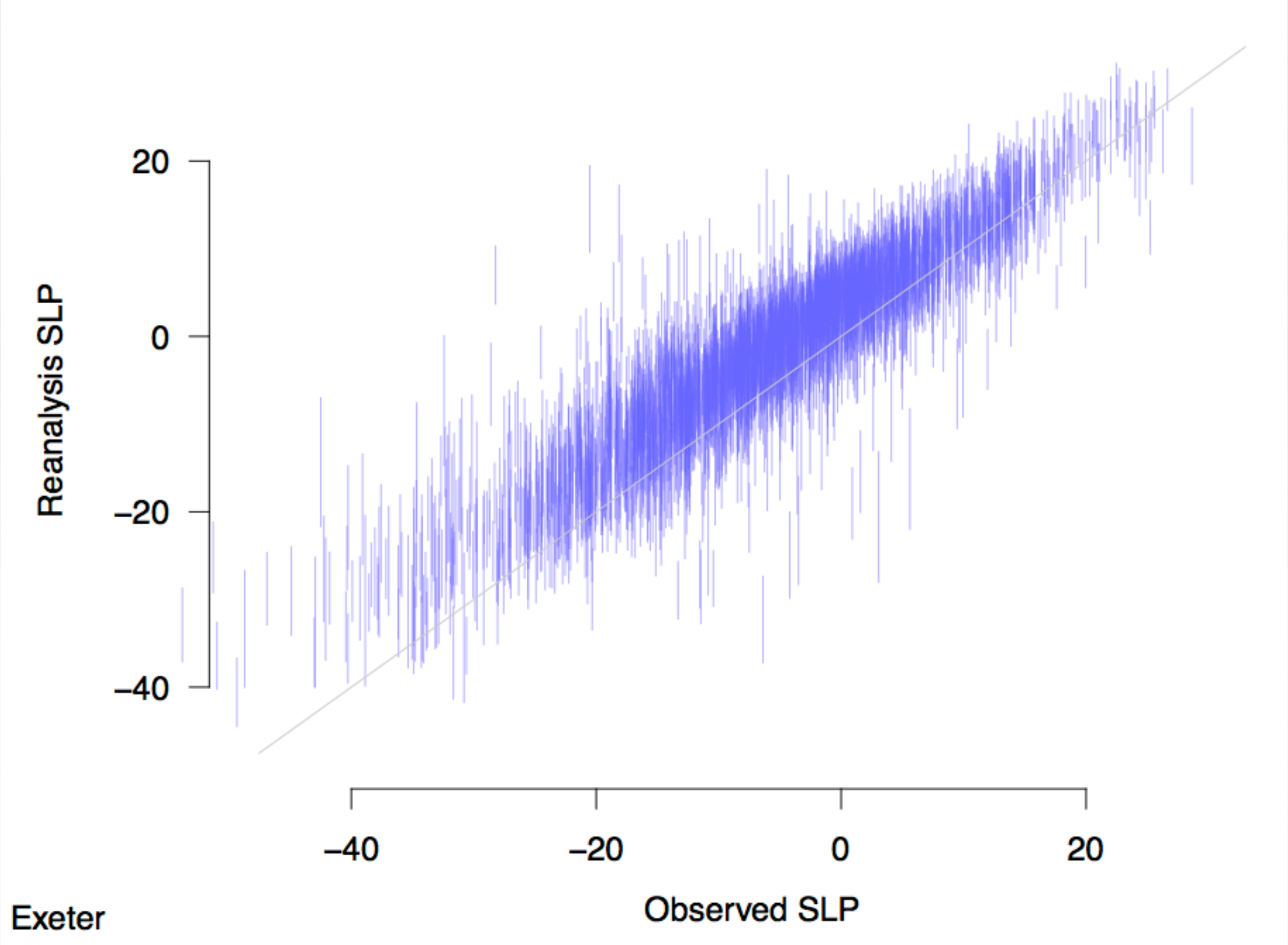


Comparison of anomalies of
Black dots: subdaily independent Air T and independent SLP from Exeter
Purple swaths: 20CR-1815 ensemble range (1815 to 1817)



Both variables compare well, though pressure is more precise. Extreme temperature anomalies are muted

Comparison of anomalies from
Subdaily independent SLP observations from Exeter vs.
Reanalysis (20CR-1815) 2*ensemble spread (1815 to 1817)



Low pressure extremes are muted, but otherwise pressure successfully predicted at the independent station.