Observations for Reanalysis

Nick Rayner, Met Office Hadley Centre

Symposium on Climate Reanalysis and Services for Society, University of Bern, 14th December 2017

Material provided by Peter Thorne, Philip Brohan, Mark McCarthy, Alexander Sterin, Sylvie Jourdain, Maria Antonia Valente, Rob Allan, John Kennedy
Overview

- Diversity and evolution of the climate observing system
- Data assembly – the importance of clear, transparent, traceable access to data
- The importance of data rescue
- Continuation of observations and ongoing timely production of data sets
- Quantifying uncertainty – errors and biases in observations
- Evaluation of reanalysis using independent observations
  - Not discussed here, but very important
The diversity and evolution of the climate observing system
Global Climate Observing System

Artist’s impression

• Concept now of coordinated instrumentation, optimally distributed

• Actually has evolved from handful of locations to complicated web of intermingling systems, set up for different purposes

• Need to know how to use it
Late C19th coverage
Beginnings of a global network
Global Climate Observing System

Evolution

By 1880 Met Offices in UK, Netherlands, etc operated own observatories
No international standardisation of details

US network 150 volunteers in 1850 and 2000 by 1891

OldWeather

NOAA

NOAA

KNMI

www.metoffice.gov.uk

© Crown copyright
A wealth of satellite measurements

From weather to weather and climate, including:

• Meteosat-1 launched in 1977. The Meteosat series continues today providing ability to create long-term climate records from weather forecasting satellites.

• ERS-1 launched in 1991. Series of high-quality climate measurements continued until 2012. To be continued with the Sentinel-3s.

• Sentinel series launched in 2014. Wide variety of different types of data for European Copernicus programme.
What do we need for provision of climate services?

• Long records of a century or more in length to enable us to characterise extremes
• Daily or sub-daily observations
• No non-climatic discontinuities
• Have information pretty much everywhere
• Have information updating in an ongoing manner (quickly)
• Clear, transparent traceable access to data
The importance of clear, transparent, traceable access to data
Data assembly

- As observing system was not designed but evolved, measurements made were not kept in one central location
- Today there is no one place from which to access all the observations we need for producing a reanalysis or for providing climate services
- In many cases, the same measurements were kept in multiple places and are duplicated, or almost duplicated, in different archives
- Can also have portions of the records for a particular place in different archives, sometimes in different countries; this can particularly arise if one country is a former colony of another
Example taken from the International Surface Temperature Initiative merged databank, drawing on three sources: GHCN-D; Japan; Russsource.
As observing system was not designed but evolved, measurements made were not kept in one central location.

Today there is no one place from which to access all the observations we need for producing a reanalysis or for providing climate services.

In many cases, the same measurements were kept in multiple places and are duplicated, or almost duplicated, in different archives.

Can also have portions of the records for a particular place in different archives, sometimes in different countries; this can particularly arise if one country is a former colony of another.

What users need to have confidence in the information is traceability of the information to the original measurements, clarity and transparency in documentation and open access to the information.

National, regional and global archives are fundamental components of the whole system and every reanalysis and service is built upon them and relies upon them wholly.
The importance of data rescue
Extending our climate observations

• Need a baseline of current weather and climate risks, against which to assess how climate change will affect extreme weather events and the risks of climate variability and change

• Rescue historical data from archives and pull through into improved data sets and reanalyses

• Various past and current international efforts, including Climate Database Modernization Program, Climate ACRE, I-DARE, ERA-CLIM, ERA-CLIM2, EURO4M, OldWeather, Weather Detectives, Data Rescue @ home, C3S Data Rescue Service, etc
Outcome: nearly 18 million historical observations new to science this year to improve global historical data sets and dynamical reanalyses

See also similar in video form: https://vimeo.com/philipbrohan
Outcome: we can now map the UK’s wettest (Oct 1903) and driest (Feb 1891) months on record using 93,000 additional monthly rainfall observations for 1862-1909.
ERA-CLIM and ERA-CLIM2 have rescued >5.5 million station days of surface measurements and >1.1 million station days of upper air measurements.
C3S Data Rescue structure, Work Package interlinks & alignments with C3S Data Store, users & Lot 2

**WP0: Kick-start Workshop**
Workshop to bring together all partners in addition to DR community experts to solicit expert advice in fine tuning the project plans

**Copernicus Data Rescue Service**
Web based system to coordinate and facilitate global data rescue activities

**WP1 & 2**

- **Portal, Activity & Data Registry**
  - Explore and register past, current and future DR activities and metadata
  - User updatable global inventory of past, current and future activities and data series
  - Standardised activity and data series metadata and metadata QC
  - Searchable by location, variable, time period
  - Opportunity, gap and priority analysis tools

- **I-DARE**
  - Add past/present/future DR activity info
  - Add DR activity metadata upload facility
  - Add Copernicus DR Service portal
  - Interoperability
  - Promote use of I-DARE and Copernicus DR Service

**WP2**

- **Data Rescue Forum**
  - Promote community support and collaboration
  - Facility to support collaboration on DR activities

**WP3**

- **Data Rescue Tools**
  - Downloadable and online tools to undertake data rescue
  - Scanning and reformatting of images
  - Digitisation templates
  - Digitisation tools for common image types
  - Basic QC (daily values, inch to mm conversion etc.)
  - Metadata QC

- **Best Practice Guidelines**
  - User documents for DR services and activities
  - Archive work and imaging
  - Digitisation and metadata collection
  - Formatting and QC

**WP4: Capacity Building Workshops**
Engage with the existing DR community and broaden the community by promoting use of the Copernicus DR Service

**Data Upload Facility**
Upload digitised data to the Lot 2 Climate Data Store

- Submission of digitised data to international archives
- Link up with activity and metadata entries in the Copernicus DR Service portal
Continuation of observations and ongoing timely production of data sets
Need:
• Development of short-delay updates to monitoring data sets (particularly surface air temperature and precipitation), consistent with the long-term record
• Also development of short-delay updates to SST and sea ice data for boundary forcing of atmospheric models
Quantifying uncertainty – errors and biases in observations
Ways of achieving consistency

• Compare everything and develop empirical corrections, relative to a chosen reference
  • Risks picking the wrong reference and biasing the whole system
• Understand each data source physically and correct according to its own biases
  • Then compare to everything else and check consistency
  • But this requires good metadata, which is often lacking
  • However, this allows potential propagation of error structure
• Let the reanalysis handle it – still requires good understanding and metadata
Evolution of the SST observing system

A Call for New Approaches to Quantifying Biases in Observations of Sea Surface Temperature. Kent et al. (2017) BAMS
https://doi.org/10.1175/BAMS-D-15-00251.1
Not possible to represent uncertainties with one number

The mean or “best estimate” might not be a representative or physically realisable state of the system.

One solution is to represent uncertainties using ensembles:

• Multiple versions of the data with different choices made when constructing the dataset.
• Spread of the ensemble members represents underlying uncertainty.
• Very easy to use.
Structural uncertainty
E.g. bias adjustment method.

Parametric uncertainty
E.g. biases adjustment; number of EOFs; length scales.

Analysis uncertainty
E.g. EOF weights.
SST anomaly ensemble, January 1926

Observations

Member 1466

Member 69

Member 137

Member 396

Member 1346

Member 1194

Member 1059

Member 400
Compare to prototype HadISST2 realisations of same events

Different SST data sets

Ensemble of one data set
Summary

• Diversity and evolution of the climate observing system
• Data assembly – the importance of clear, transparent, traceable access to data
• The importance of data rescue
• Continuation of observations and ongoing timely production of data sets
• Quantifying uncertainty – errors and biases in observations
• Evaluation of reanalysis using independent observations
  • Not discussed here, but very important