Observational feedback:

What reanalysis tells us about the quality of observations

(and what observations tell us about other climate information)

Paul Poli

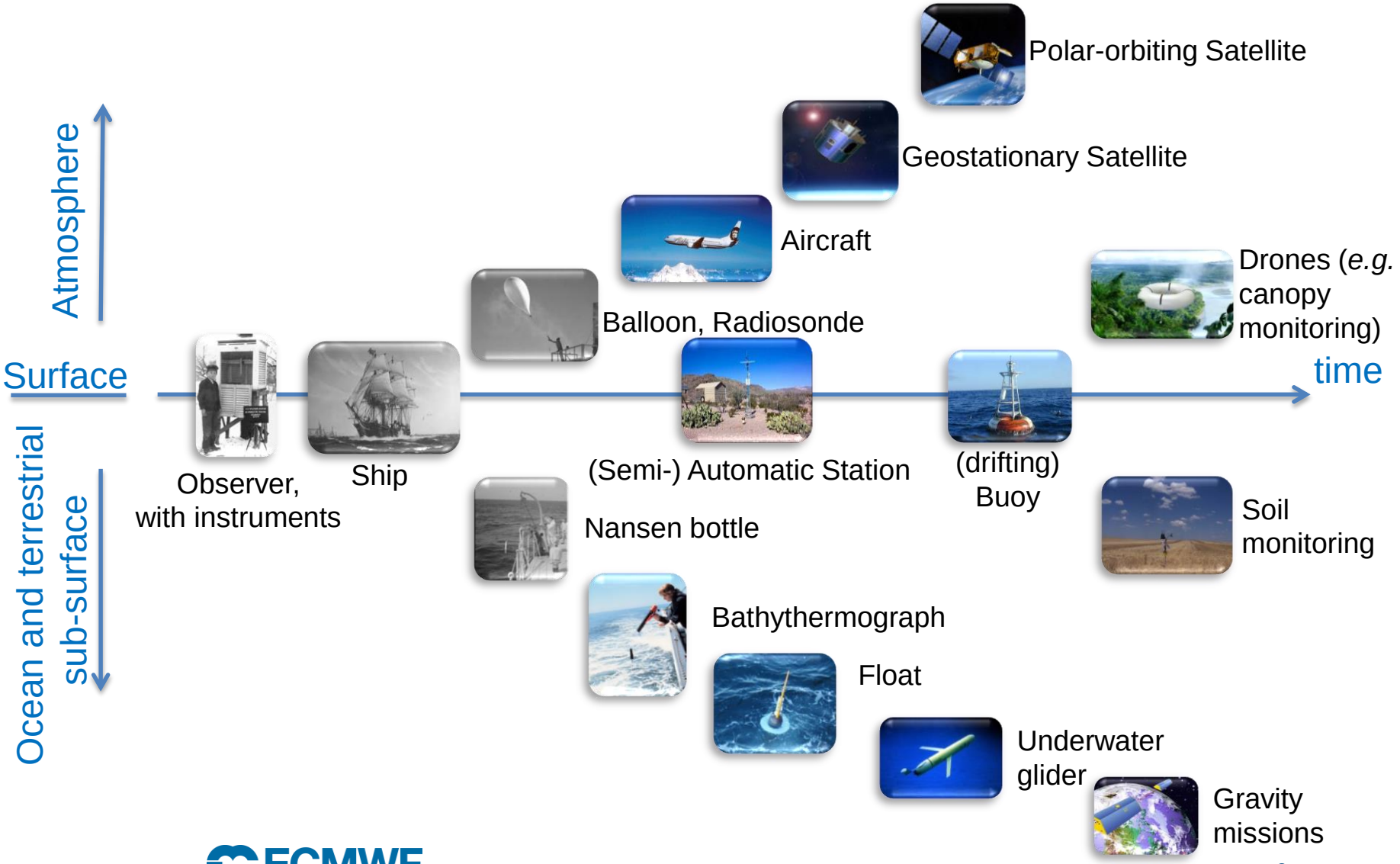
Acknowledgements: Dick Dee, Paul Berrisford, Hans Hersbach, David Tan, Shinya Kobayashi, and participants of the *Core-Climax coordination meeting towards exchanging observation feedback*

ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Outline

- What is reanalysis observational feedback?
- What does it tell us?
- Practicalities

Observations are sources of information



Climate information sources

Public

Under-standing & Applica-tions

Feedbacks

Reanalysis

Integration of an invariant, modern version of a data assimilation system and Earth system model, over a long time period, assimilating observations

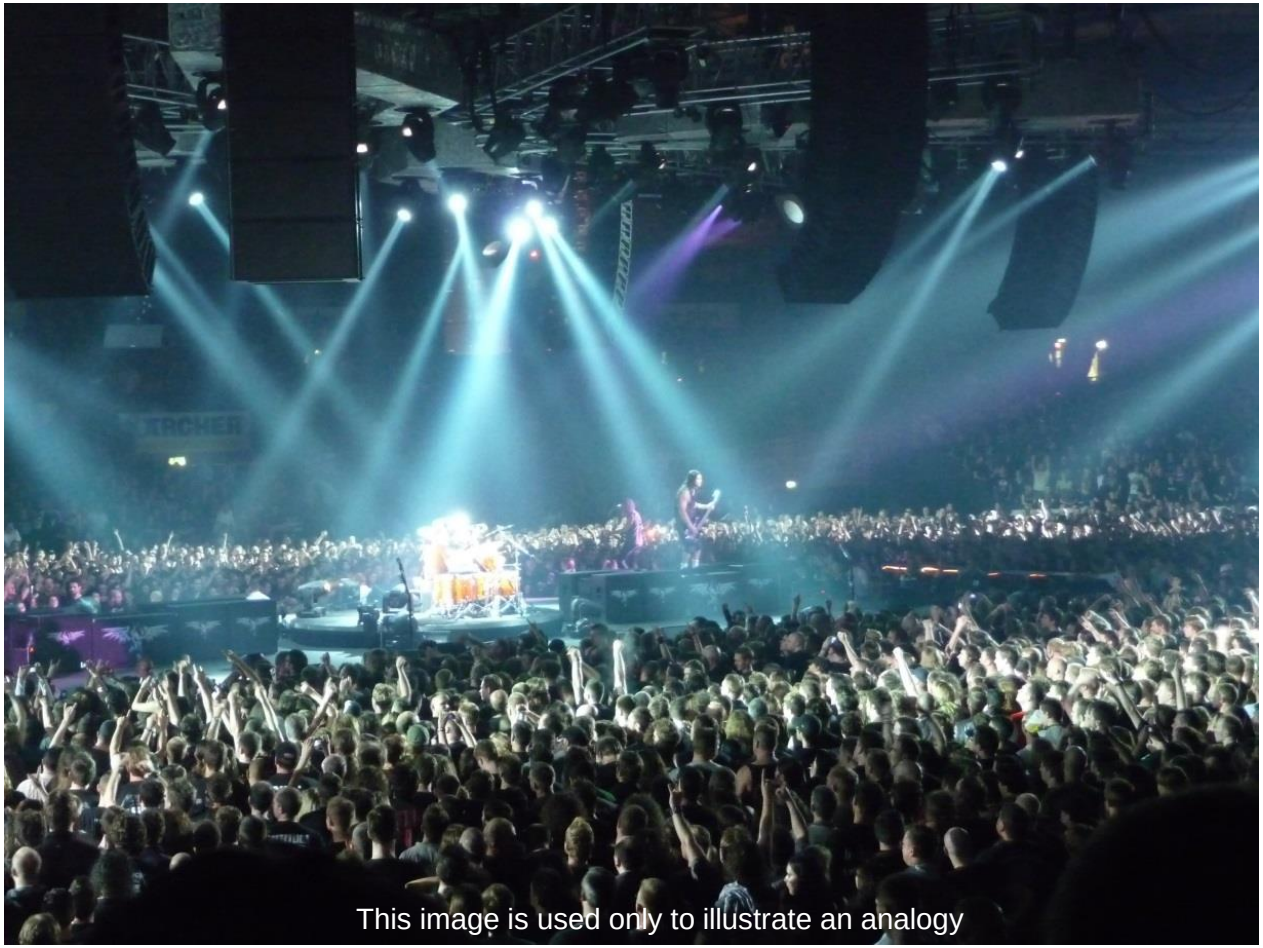
Obser-vations

Models

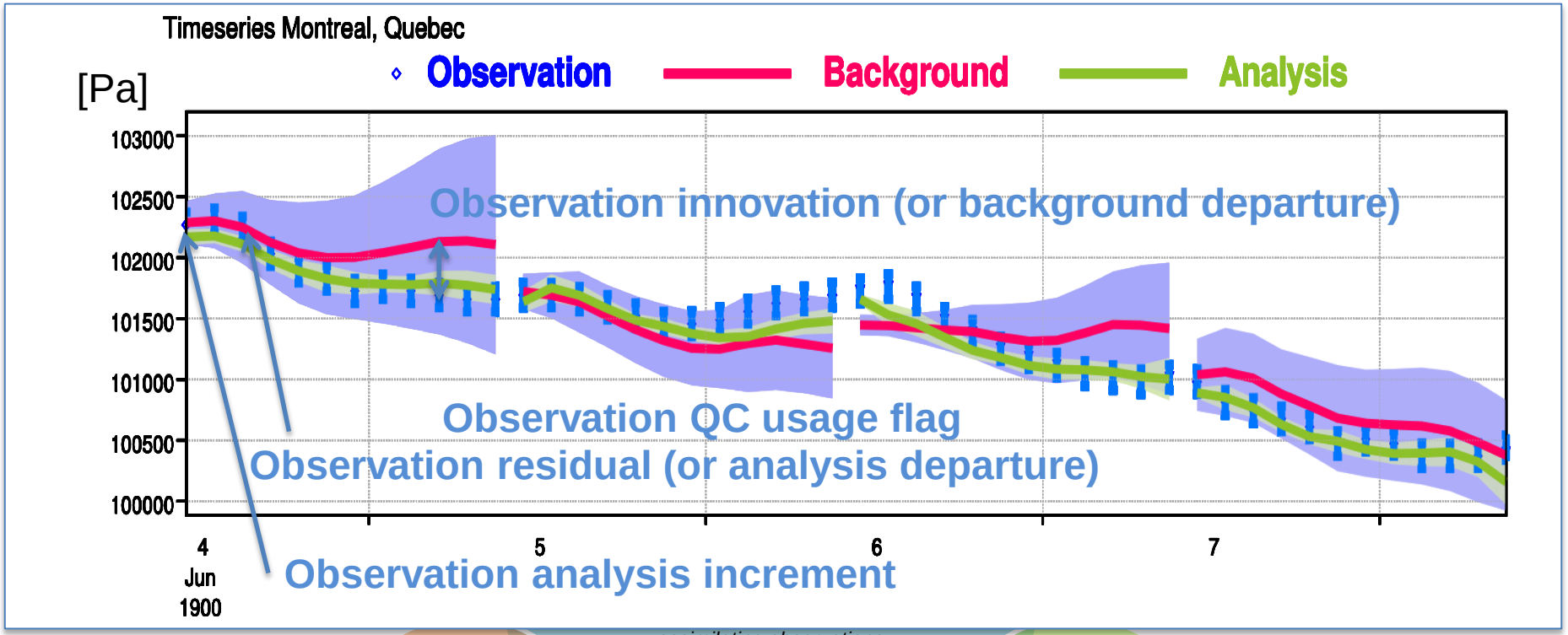
Instruments (orchestra, band …)

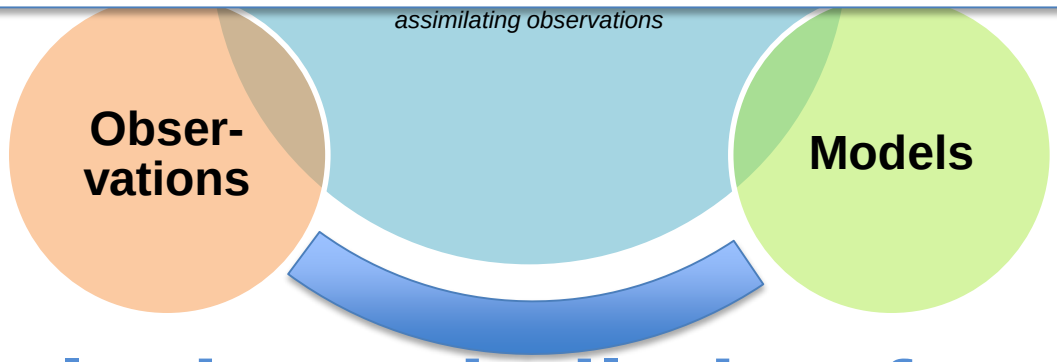
Conductor

Reanalysis is über-cool



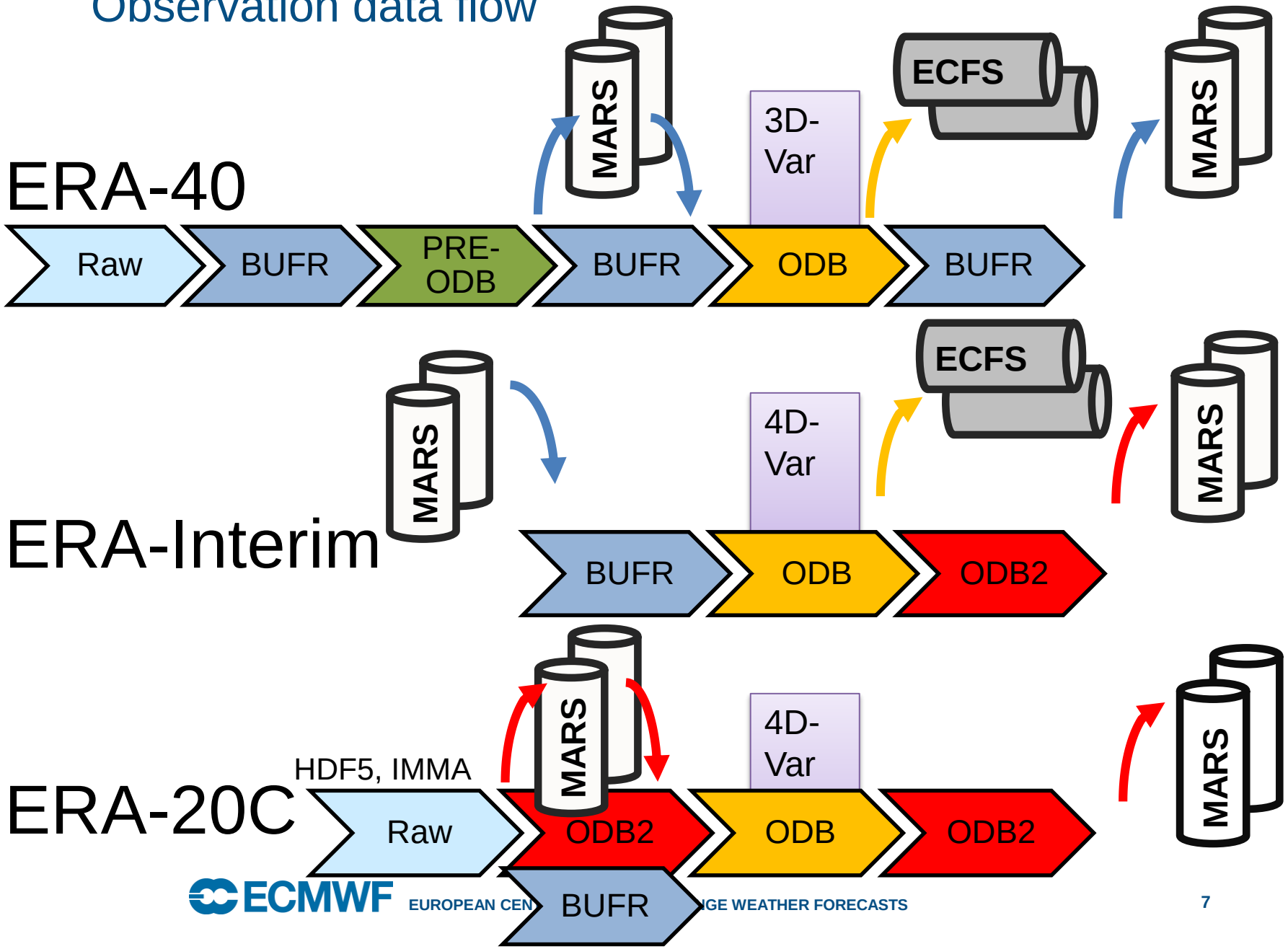
Climate information sources



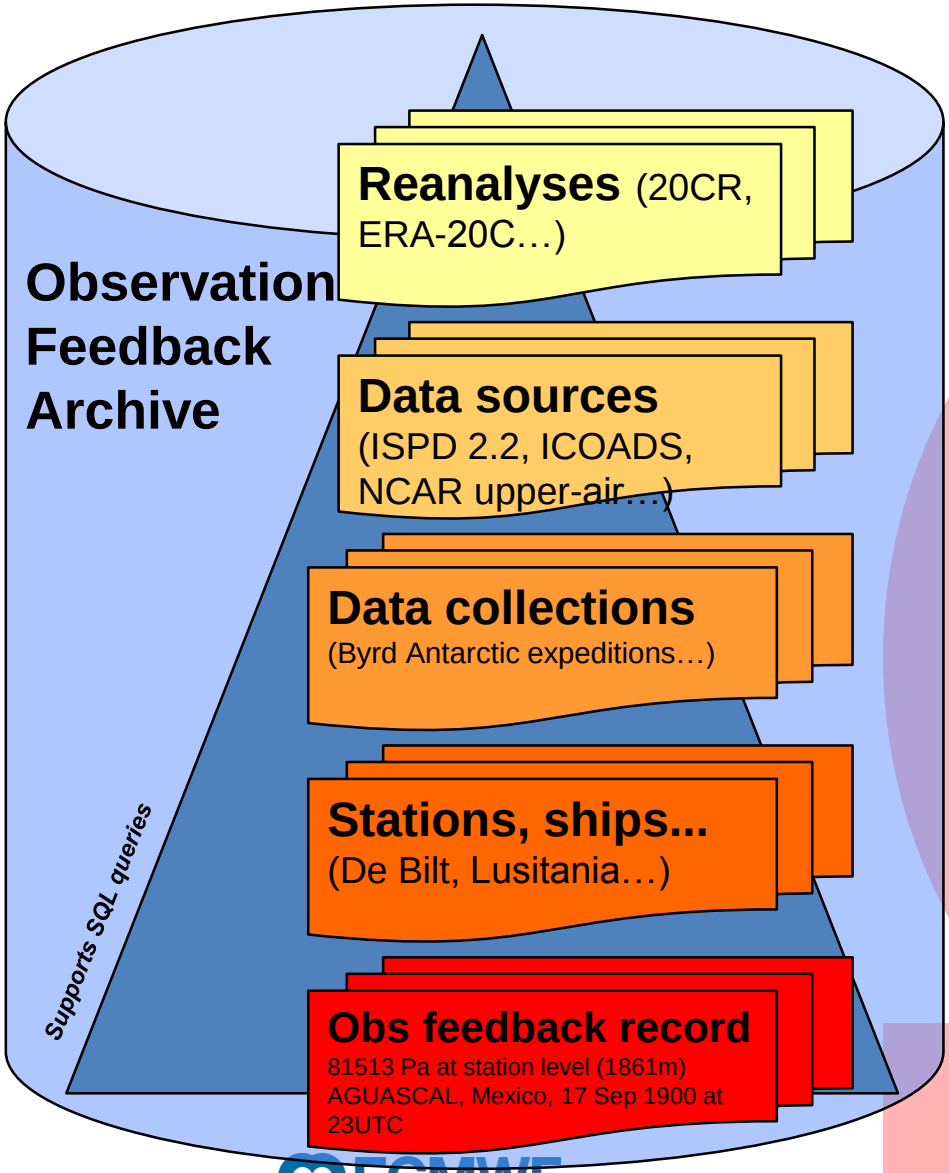


Reanalysis assimilation feedback

Observation data flow



ECMWF Observation Feedback Archive organization



Observation record attributes

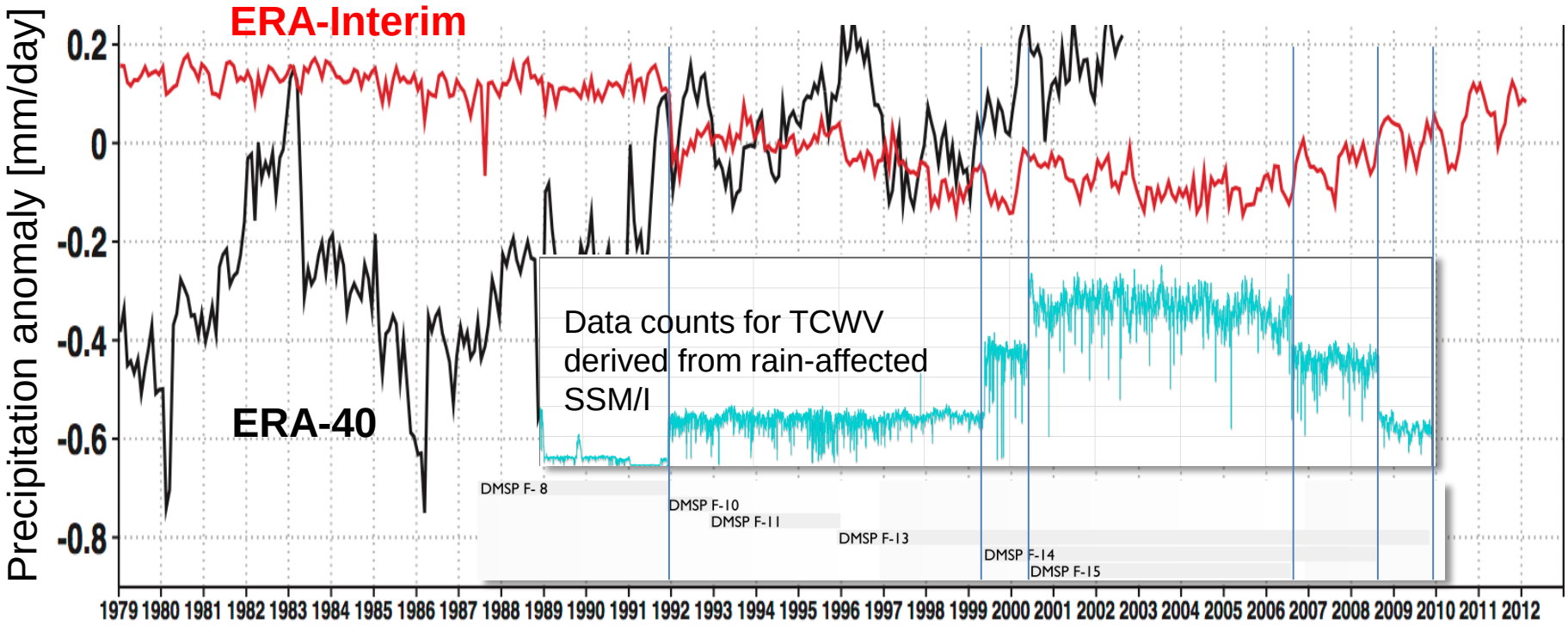
- Identification and metadata
  - Time and geolocation (lat, lon, alt, level)
  - Observation report type (buoy…)
  - Unique identifier
- Data
  - Geophysical variable (T,p,q…)
  - Reporting practice
  - Observation value
- Feedback-added attributes:
  - Ancillary (model) land-sea mask, orography, surface wind, …
  - Background & analysis departure
  - Assimilation QC and usage flags
  - Obs. systematic error estimate (bias or accuracy)
  - Obs. random error estimate (1-sigma precision)

Outline

- What is reanalysis observational feedback?
- What does it tell us?
- Practicalities

Assimilation feedback example 1

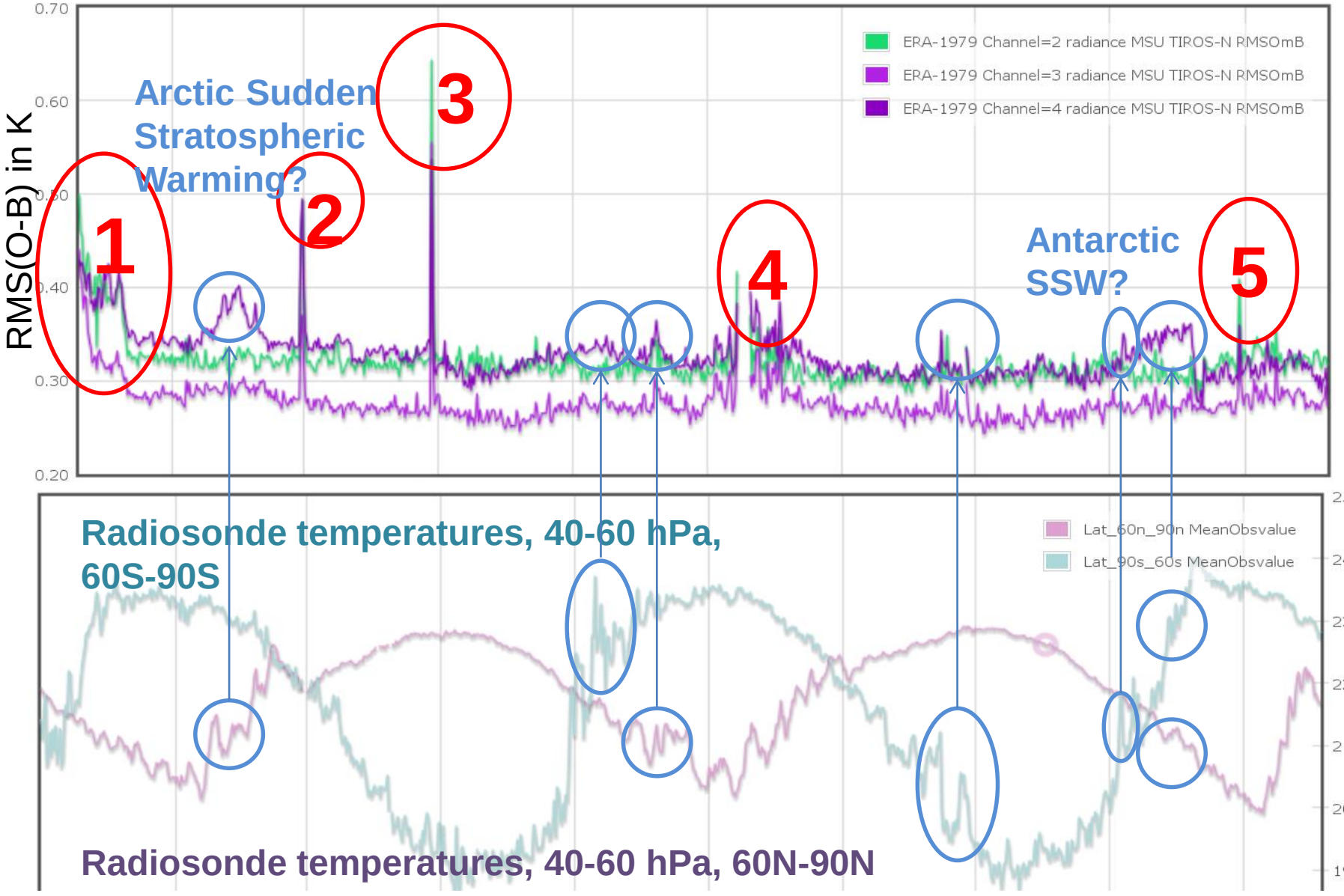
understanding the shifts in ERA-Interim water cycle



- Due to assimilation of rain-affected radiances from SSM/I, now fully understood *(Geer et al. 2008)*
- Effect scales (non-linearly) with the N. of assimilated SSM/I data

Assimilation feedback example 2

relationship between MSU and radiosonde records

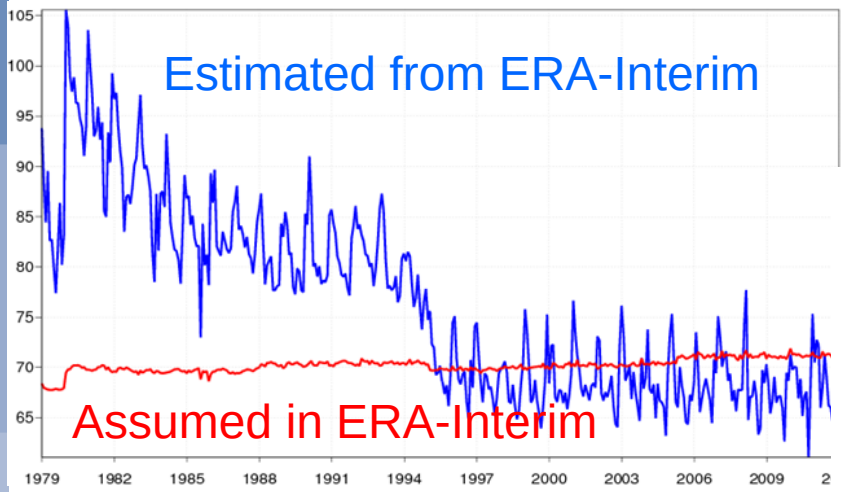


Assimilation feedback example 3

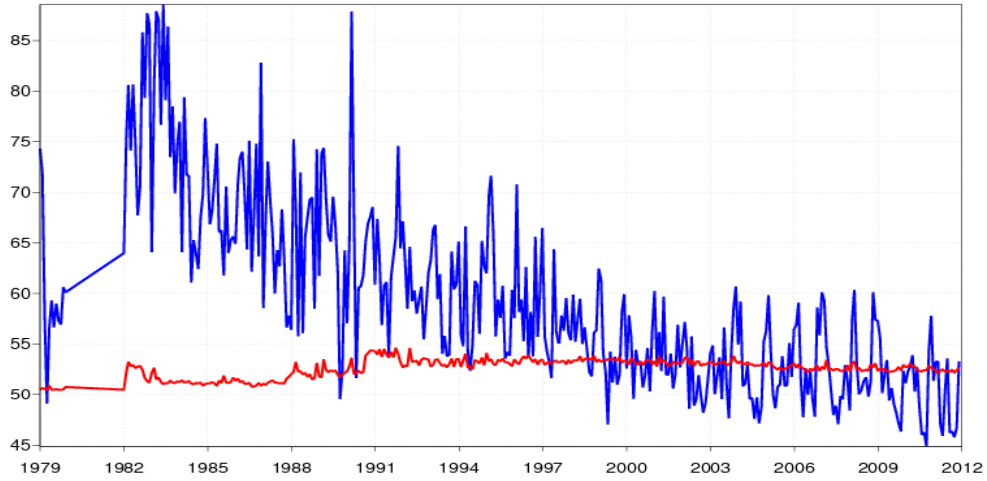
estimating surface pressure observation errors

[Pa]

Land SYNOP



Automatic land SYNOP

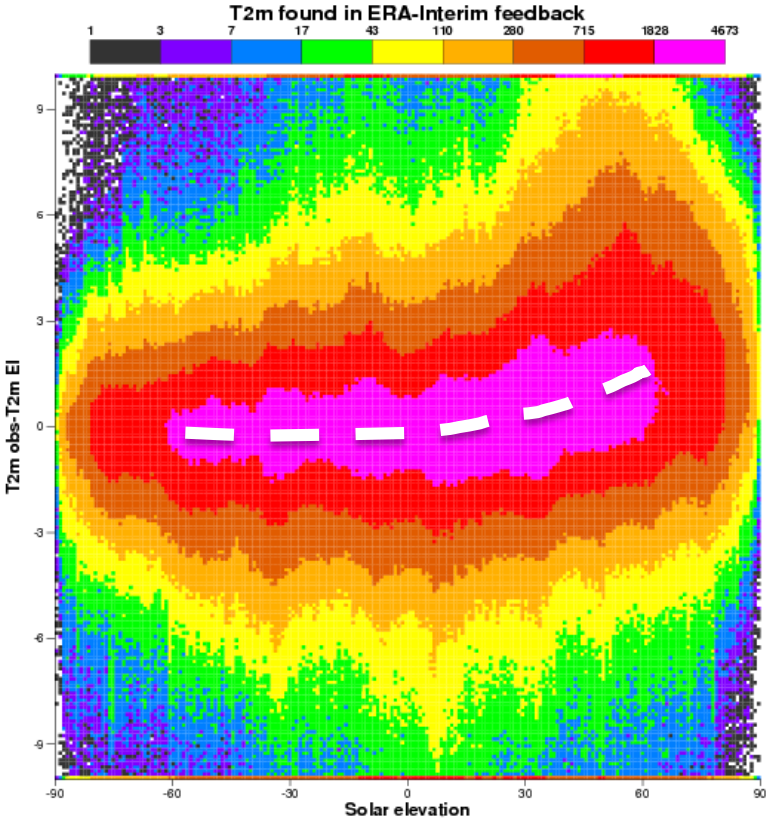


Using the method of Desroziers *et al.* (2005; Q.J.R. Meteorol. Soc. doi: 10.1256/qj.05.108)

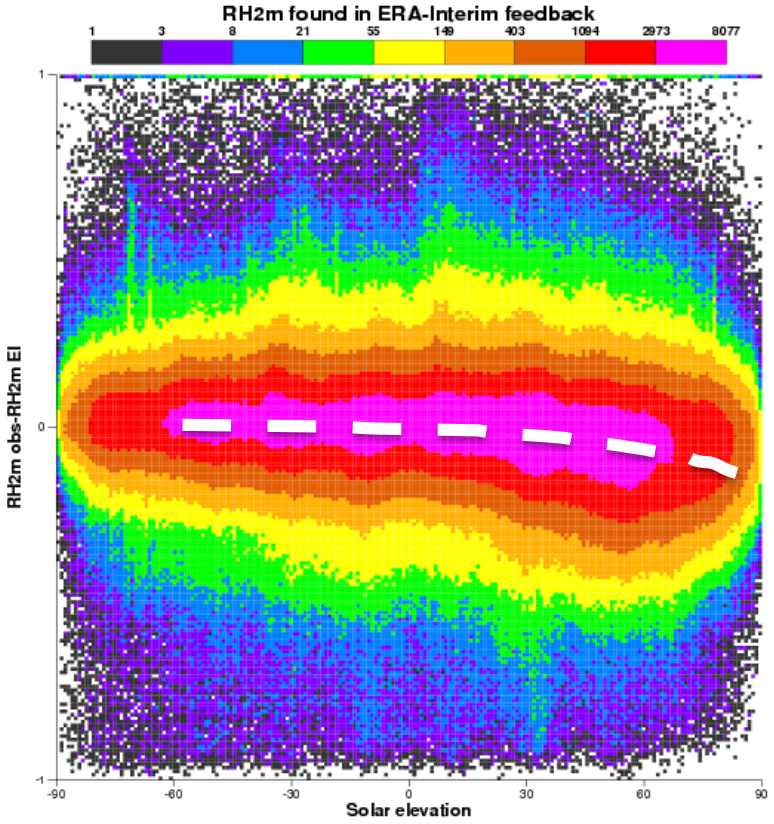
Assimilation feedback example 3

effect of solar radiation on temperature and relative humidity biases

19.9 million observations



18.9 million observations



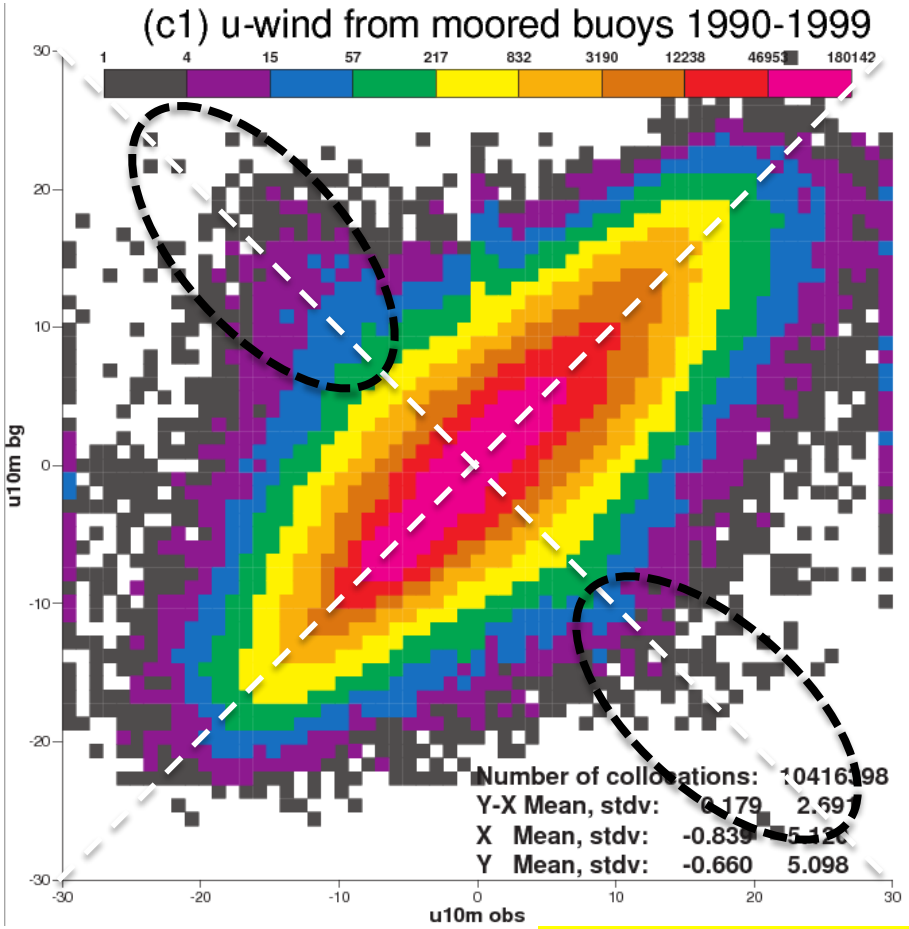
- T obs. mean departures from ERA-Interim analyses increase with solar angle: where is the bias, in obs. or ERA-Interim?

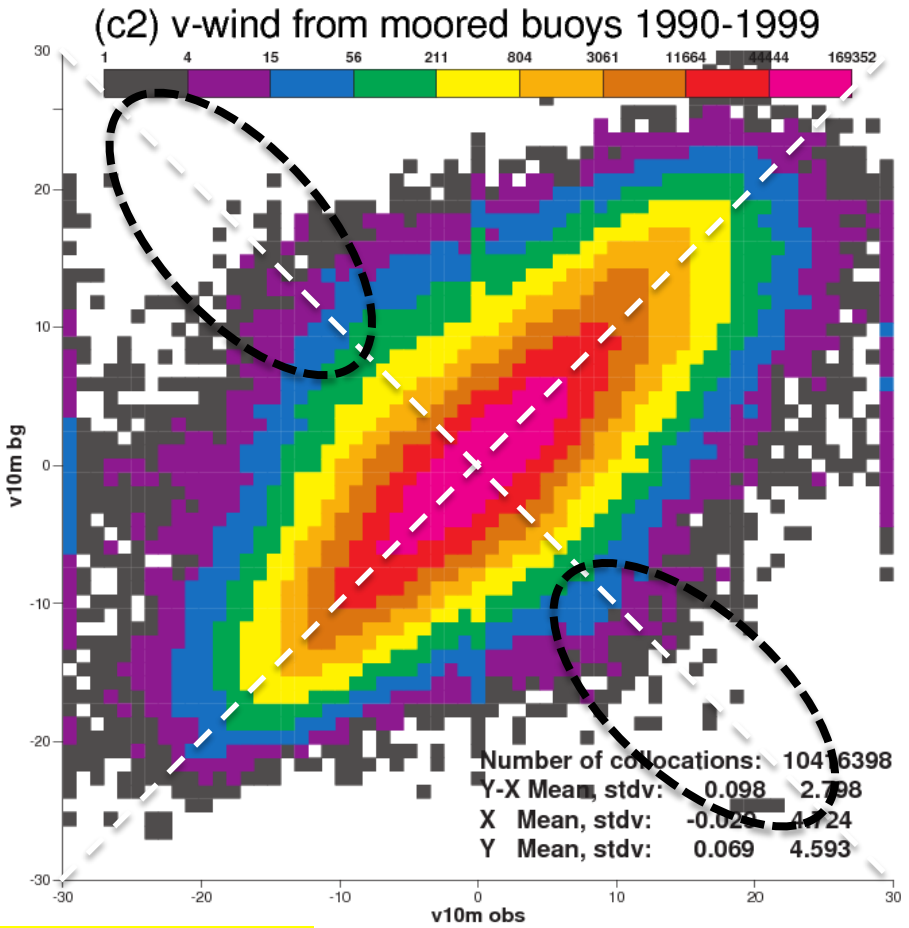
- RH obs. mean departures from ERA-Interim analyses decrease with solar angle: where is the bias, in obs. or ERA-Interim?

Based on land surface observations found in ERA-Interim obs. feedback, latitudes 20S-90S, Oct 1978-Jan 1989

Assimilation feedback to improve obs QC

Marine surface winds from buoys



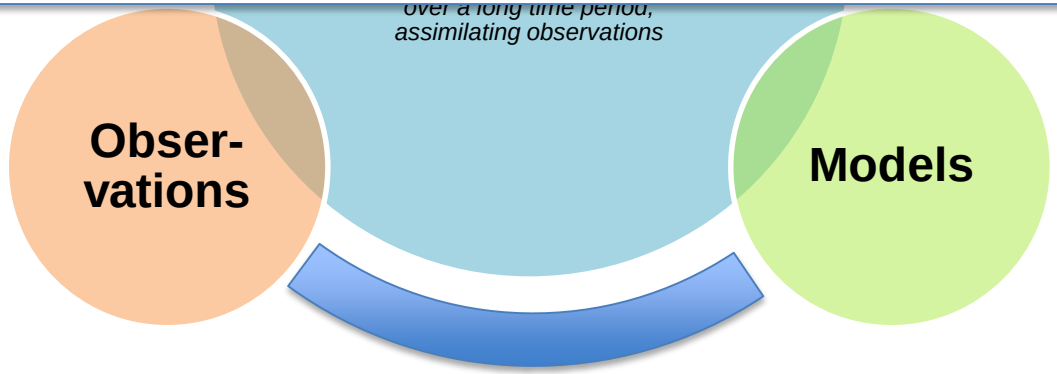


- Some winds have 180 degrees shift in direction! (u → -u and v → -v)

Climate information sources

81513 Pa at station level (1861m) AGUASCAL, Mexico, 17 Sep 2020 at 23UTC	81645 Pa analysed by ECMWF OPS, and obs. possibly biased by 50 Pa	81645 Pa analysed by ERA5, and obs. possibly biased by 45 Pa	82030 Pa estimated by ERA-20C which didn't use this obs. dataset	82030 Pa estimated by CMIP6 climate projection #1	82530 Pa estimated by CMIP6 climate projection #2

Allows to compare "observed vs. predicted"

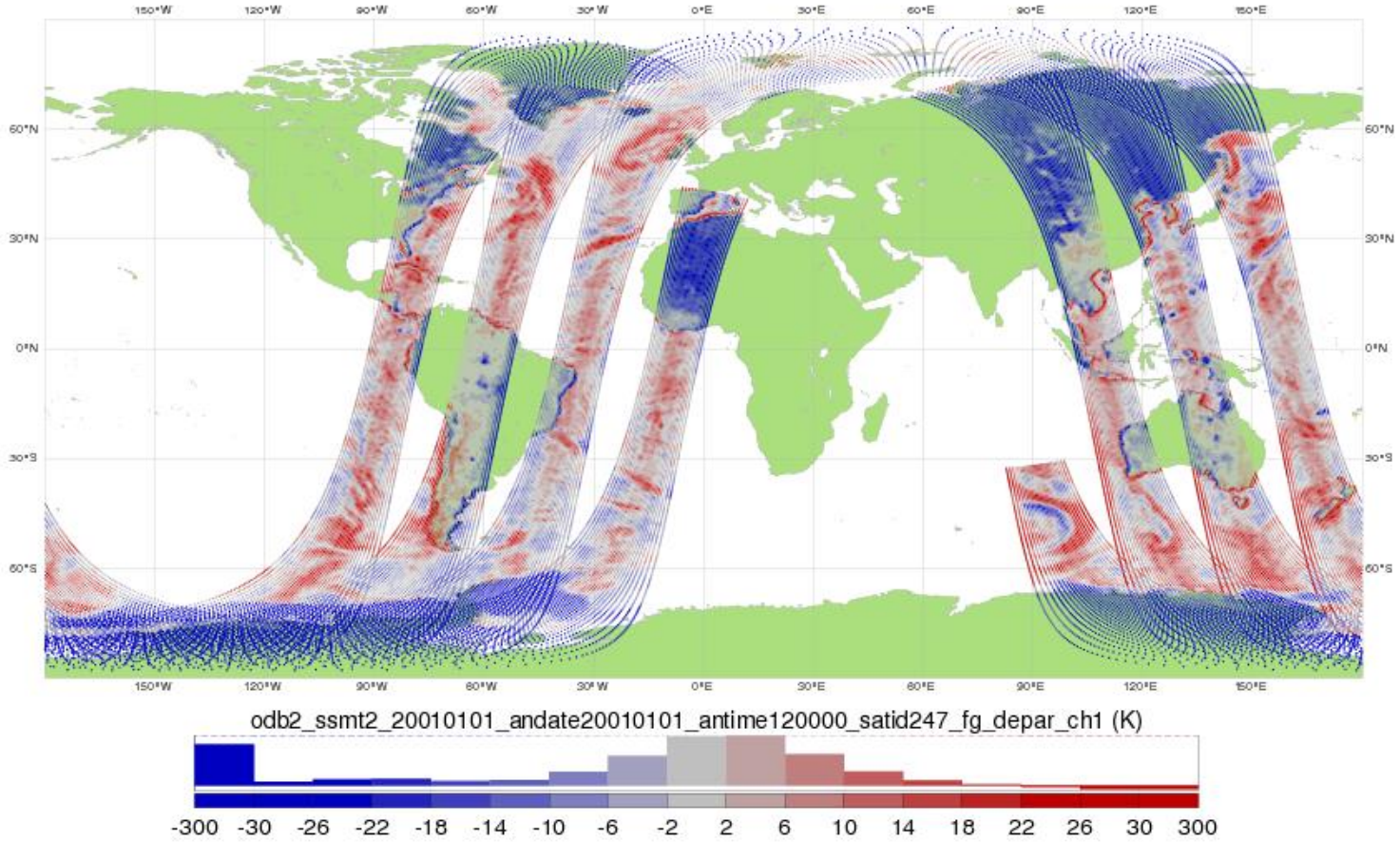


Independent feedback

Independent feedback example 1

Spotting inconsistencies in SSMT2 observations by comparing with an instrument simulator (RTTOV)

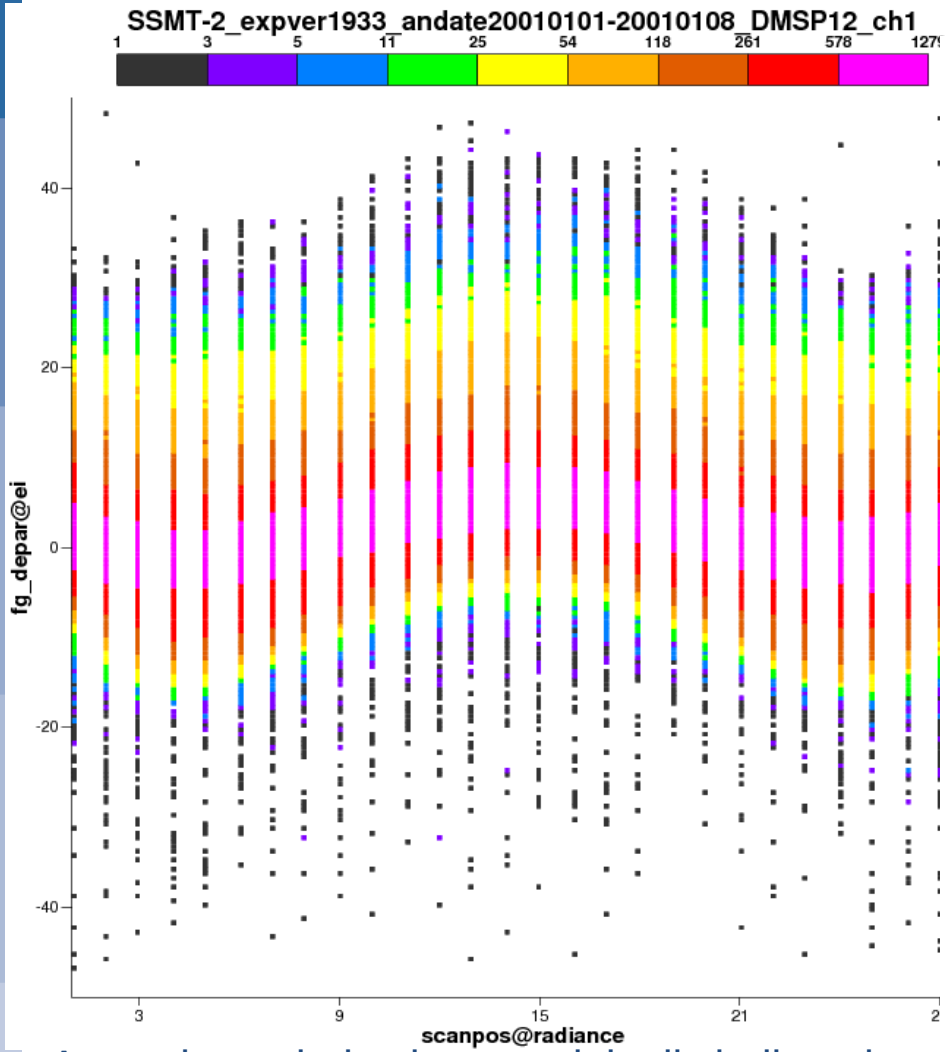
SSM/T-2, DMSP F14, 6-hour period around 20010101, 12UTC

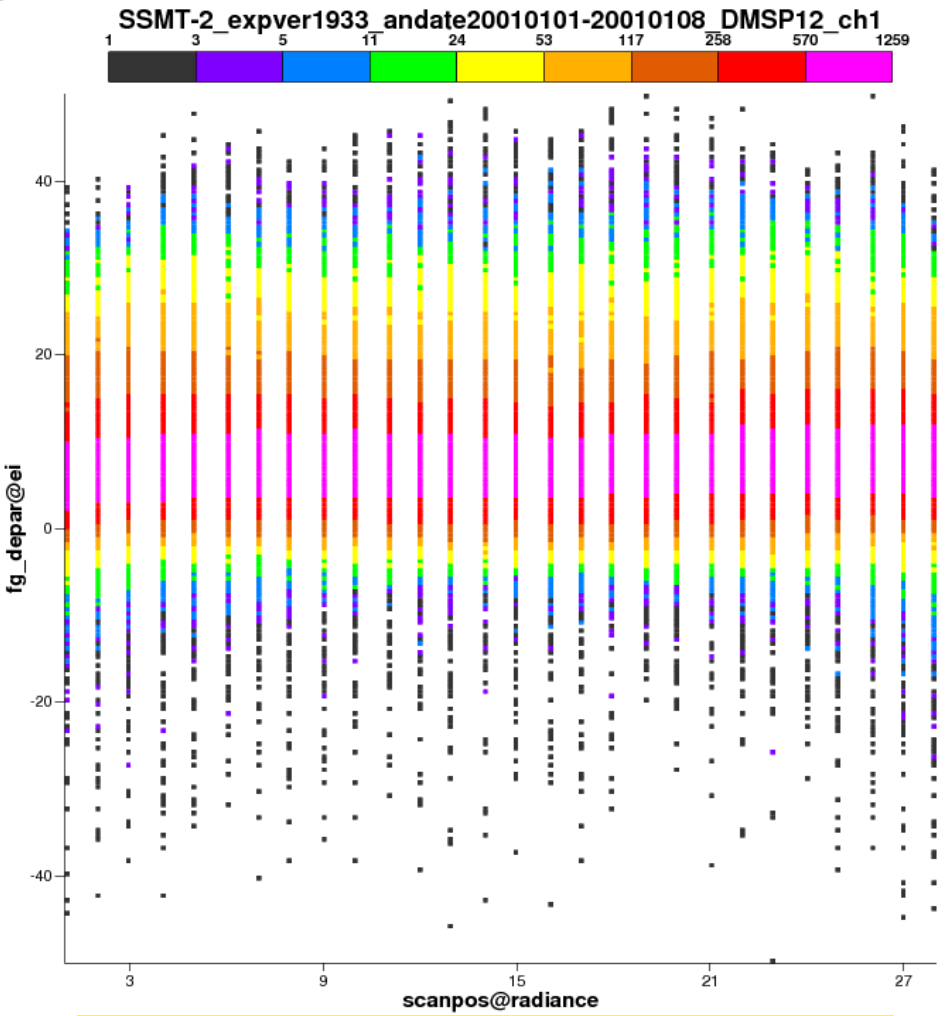


Inaccurate geolocation?

Independent feedback example 2

"Guessing" instrument characteristics a posteriori



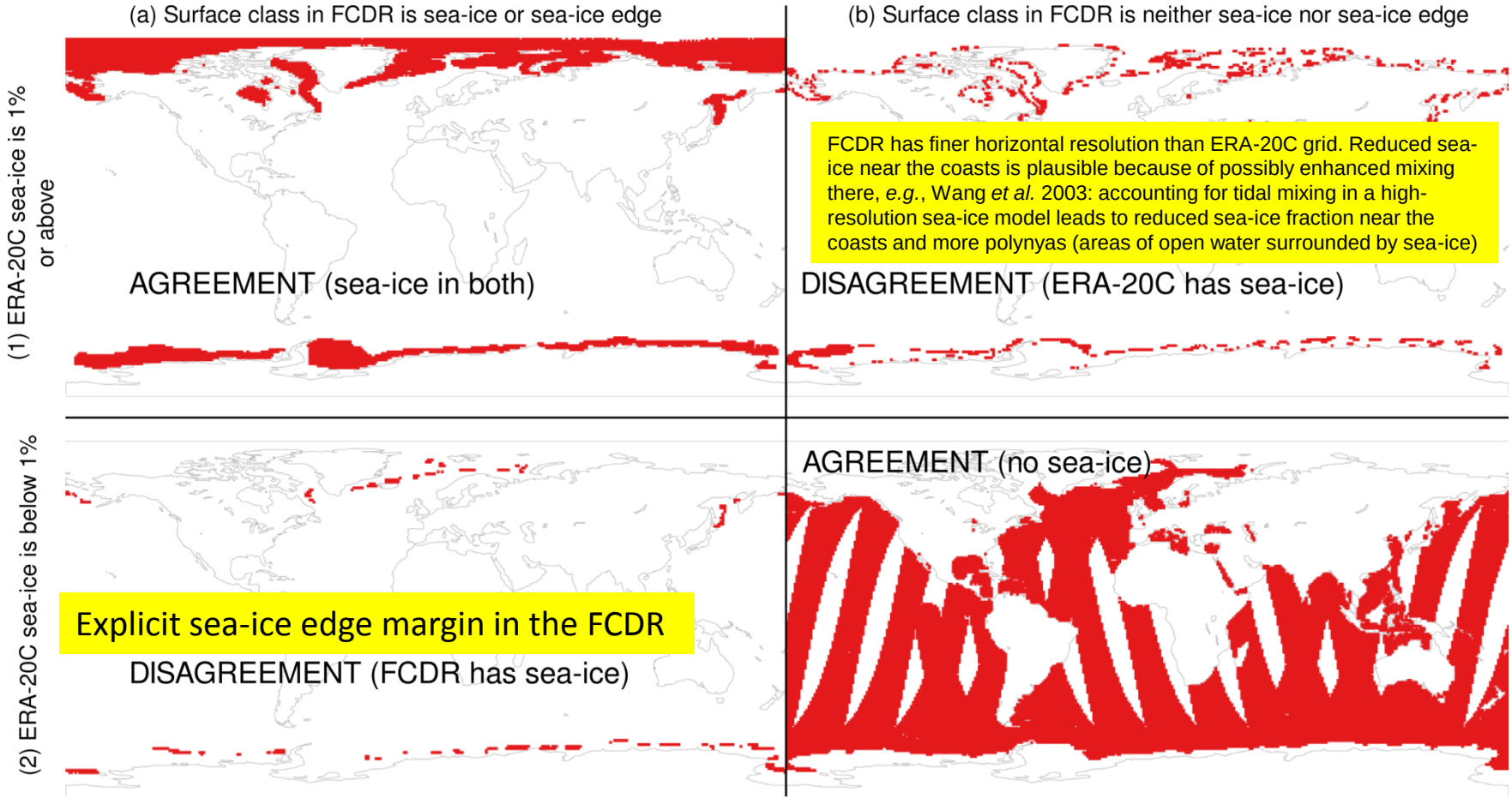


Assuming polarization as originally believed (vertical at nadir, rotating with viewing angle) 91 GHz chan.

Assuming polarization horizontal at nadir, rotating with viewing angle

Independent feedback example 3

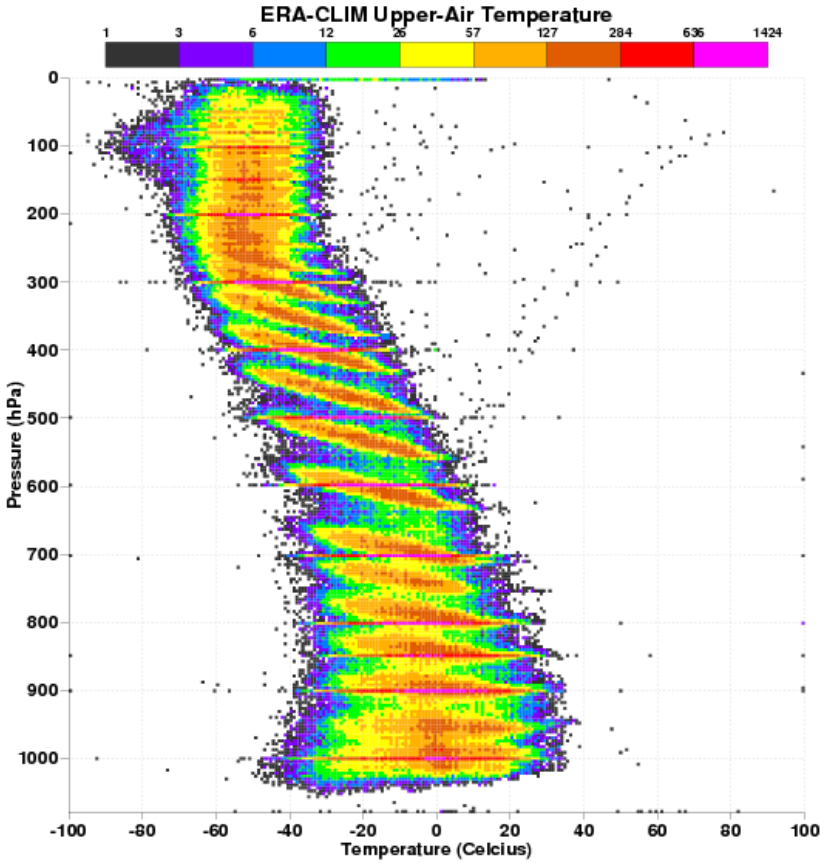
Assessing ancillary sea-ice data in SSM/I FCDR



Contingency table for the sea-ice reported in the FCDR (columns a, b) and matching sea-ice fraction from ERA-20C (HadISST2.1.0.0) at 1% or above (rows 1, 2), for 19970116, satellite F13

Independent feedback to improve obs QC

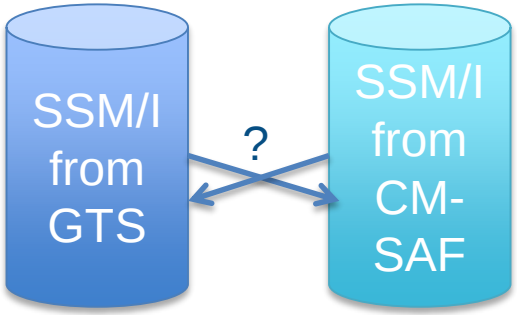
Data recovered after imaging and digitization efforts



CHUAN dataset available on PANGEA

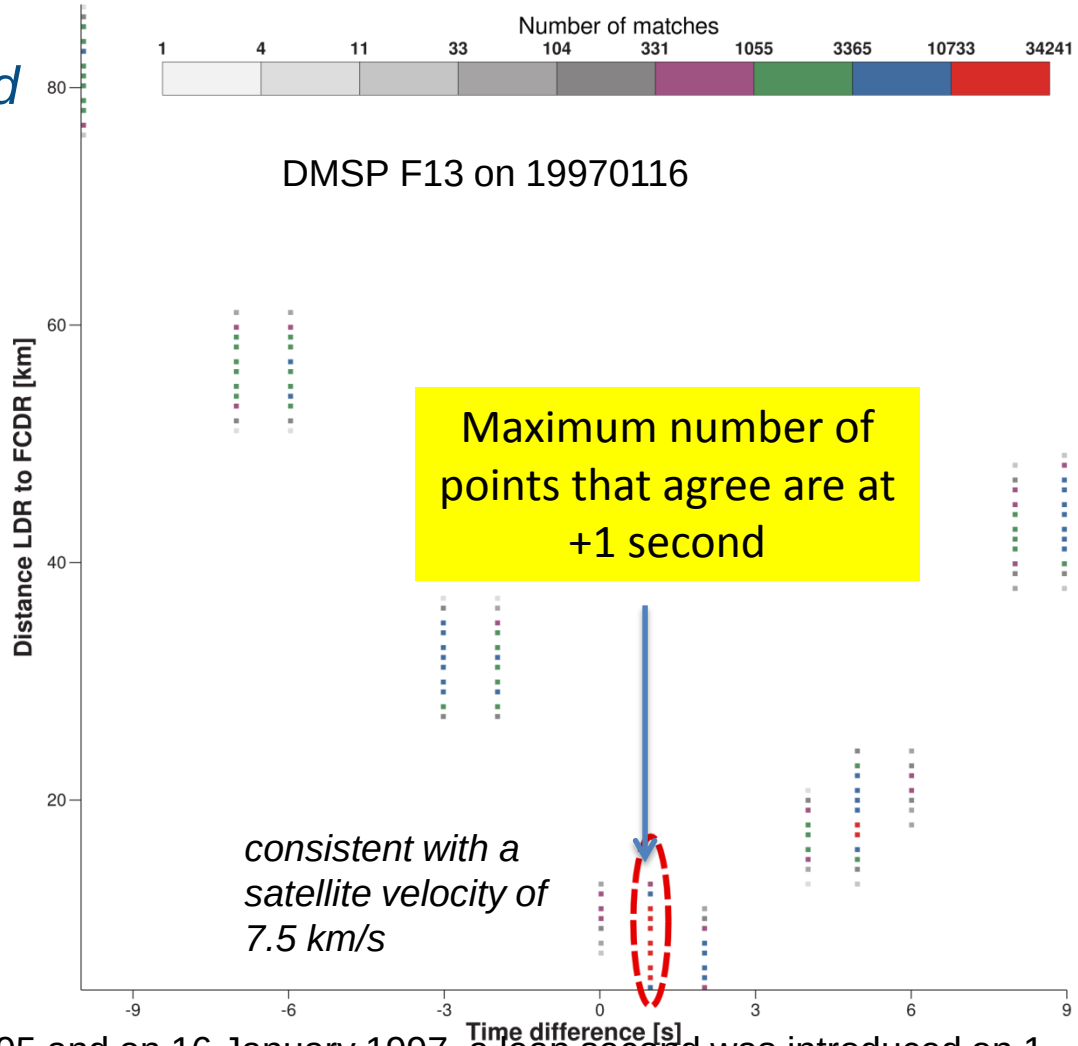
http://doi.pangaea.de/10.1594/PANGAEA.821222

Intra-observation feedback
Example 1:
finding a leap second



Look for best-match in terms of brightness temperature among data:

- Identical satellite
- Identical scan position numbers
- Identical channel numbers
- Time differences within 10 seconds
- Location differences within 100 km.



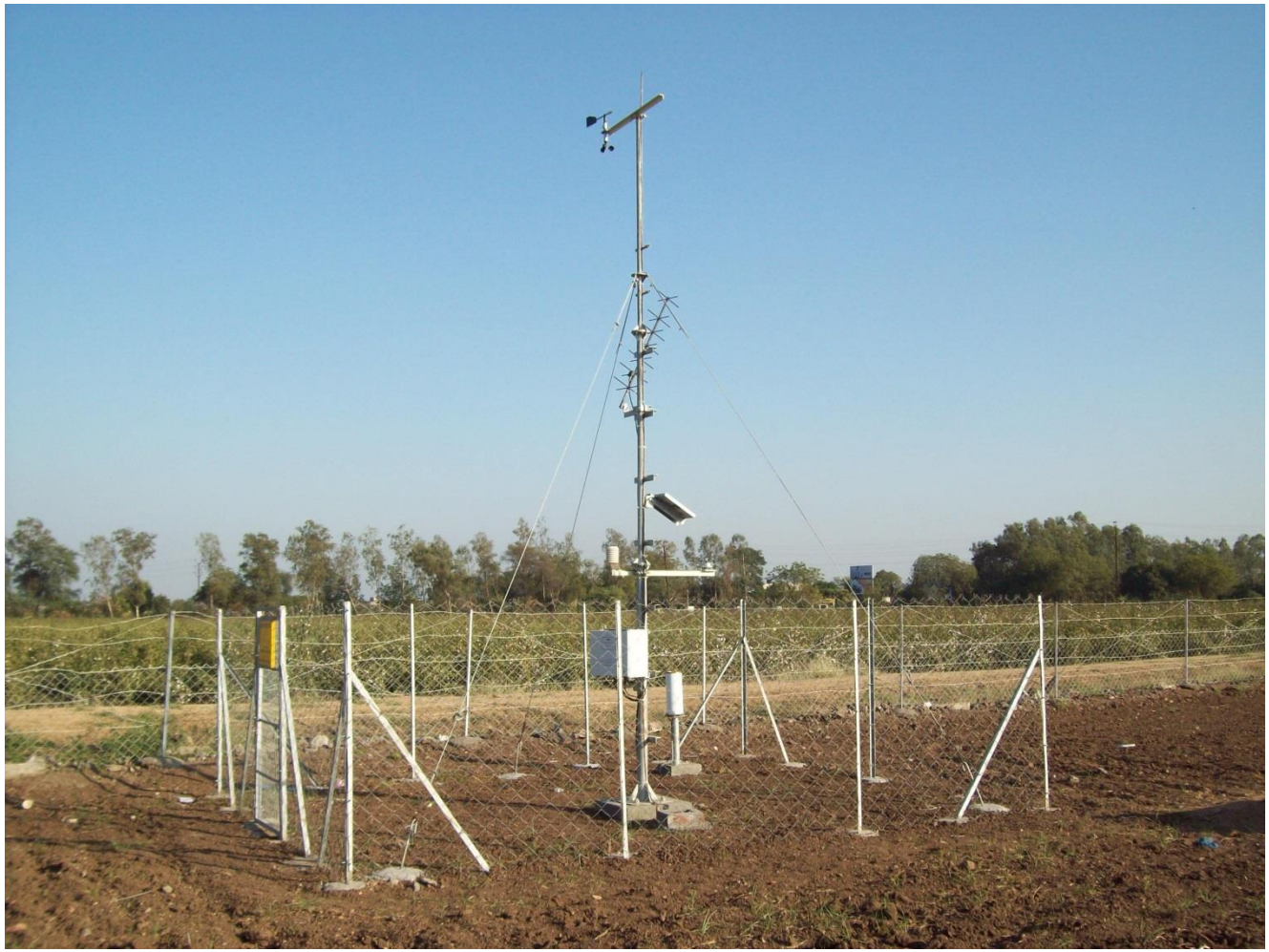
- Why? Between F13 launch in March 1995 and on 16 January 1997, a leap second was introduced on 1 January 1996 (Source: BIPM, http://www.bipm.org/en/bipm-services/timescales/time-ftp/publication.html#nohref), to keep the UTC (normally synced to atomic time) within 1s of mean solar time
- By the way, the last leap second was … yesterday! (30 June 2015 ended at 23:59:60)

Intra-observation feedback

Example 2: *variability of humidity with temperature*

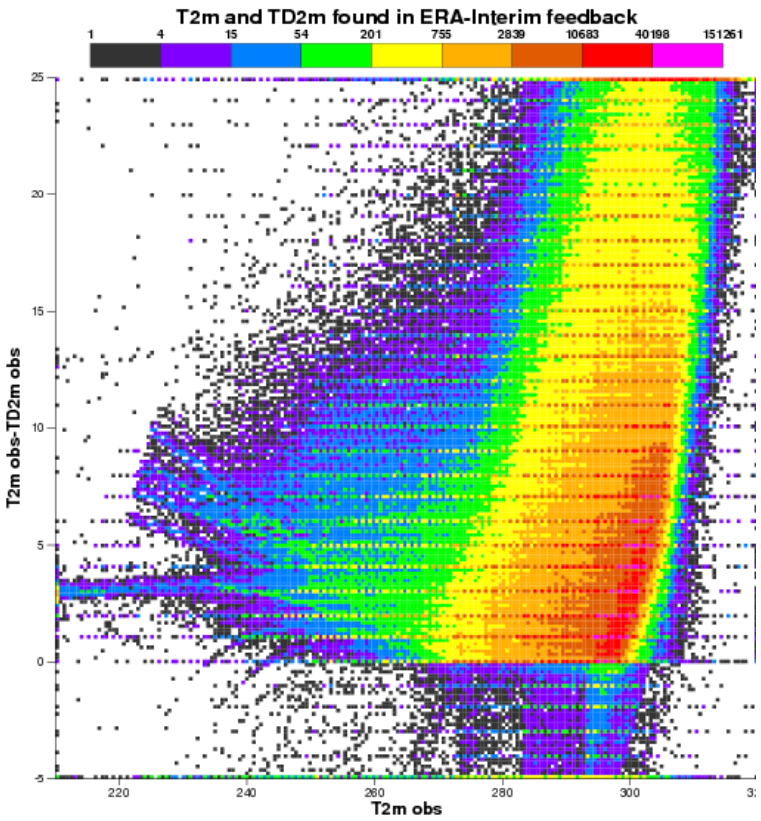
288.15 289.25 288.35 287.05 286.55 285.65 …

285.55 287.35 282.45 285.65 285.65 285.65 …

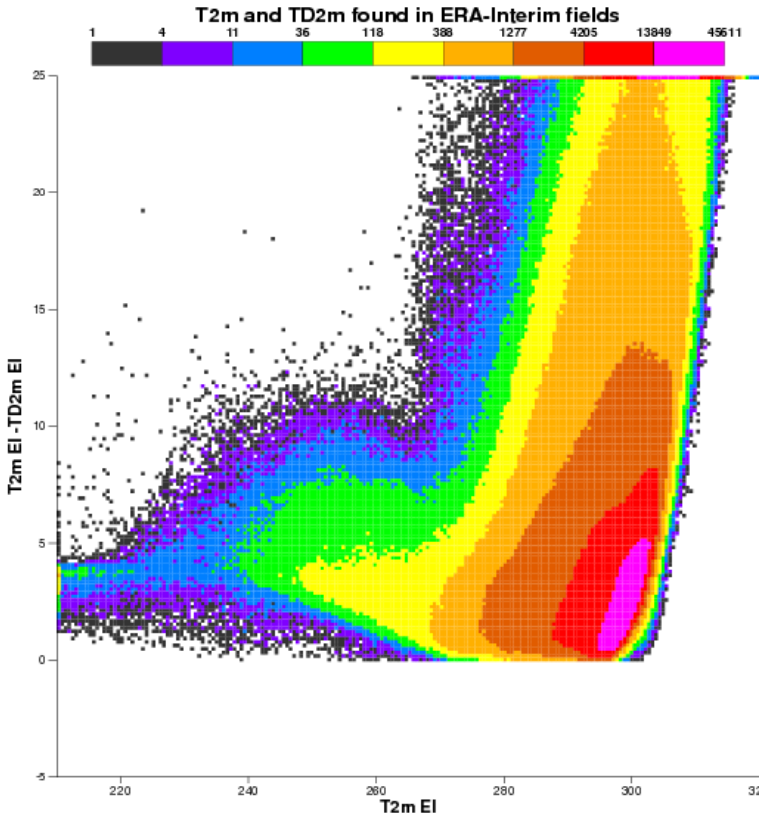


Intra-observation feedback
Example 2: *variability of humidity with temperature*

18.9 million observations



18.9 million observations



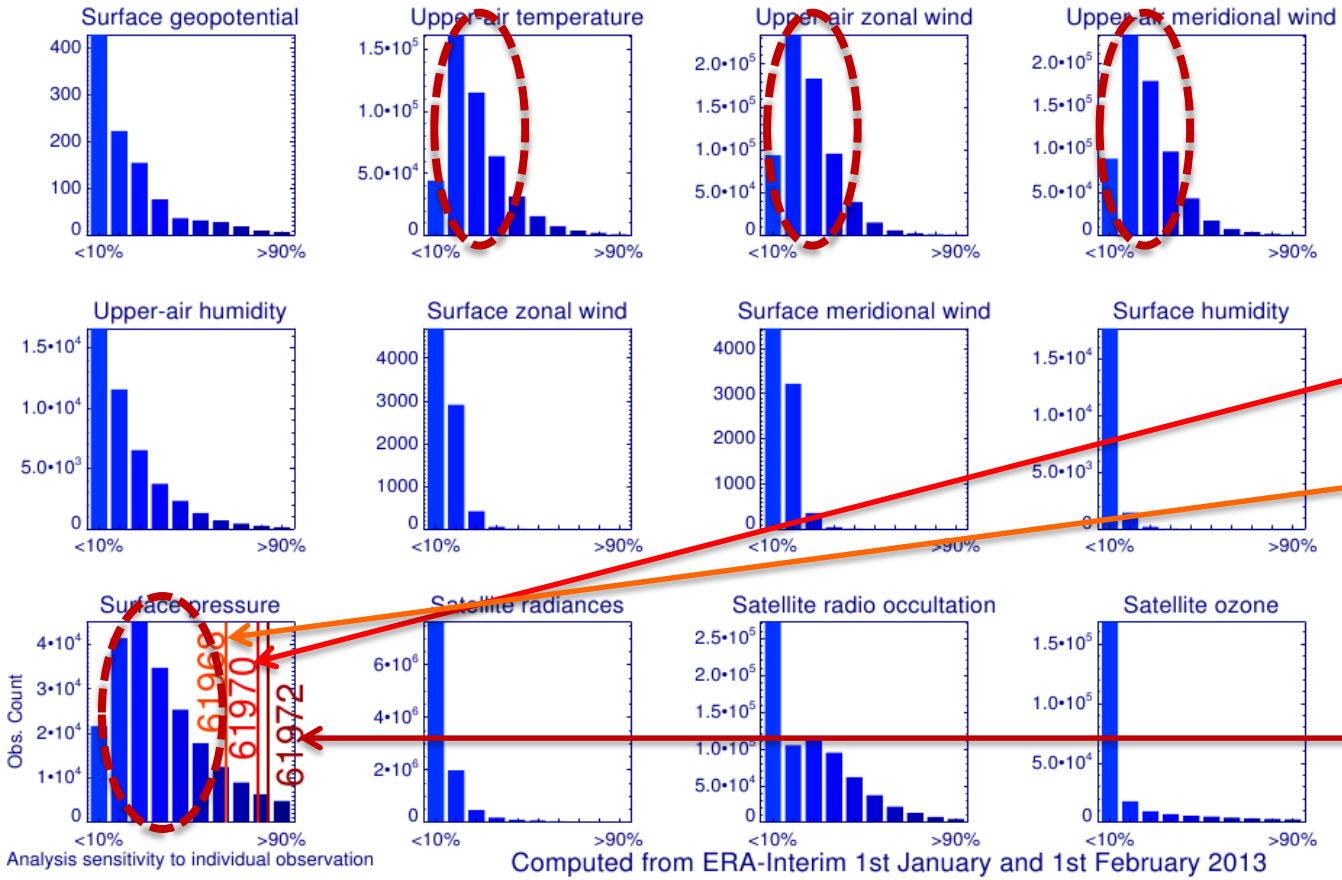
- Negative dew point temperature depressions!
- Depressions discretized with 1K accuracy!

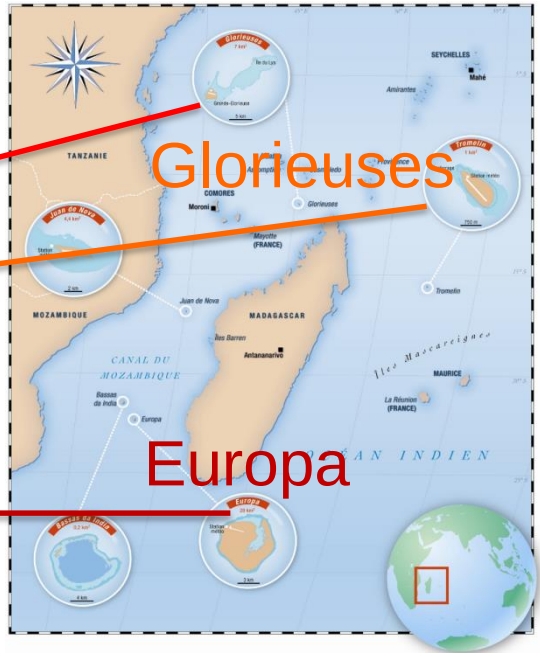
- Variability in ERA-Interim analyses for <u>exactly</u> the same sample of event

Based on land surface observations found in ERA-Interim obs. feedback, latitudes 20S-90S, Oct 1978-Jan 1989

Value or impact of observations in reanalysis

- Analysis sensitivity to observations (Cardinali *et al.* 2004, Q.J.R. Meteorol. Soc., doi: 10.1256/qj.03.205) computed routinely in ERA-Interim





Outline

- What is reanalysis observational feedback?
- What does it tell us?
- Practicalities

This is all very nice, but …

Has anybody here seen reanalysis

observational feedback before?

Reanalysis observational feedbacks available in 2015

- MERRA gridded innovation observation (GIO)

http://disc.sci.gsfc.nasa.gov/mdisc/data-holdings/merra-innov

NetCDF

- ECMWF ERA-40 and NCEP/NCAR

BUFR

http://apps.ecmwf.int/services/mars/catalogue/?type=af&class=e4&stream=oper&expver=1

http://rda.ucar.edu/datasets/ds090.0

- NOAA 20CR feedback is part of ISPD

HDF5

http://rda.ucar.edu/datasets/ds132.0/

- ICOADS value-added database

IMMA (ASCII)

http://icoads.noaa.gov/ivad/ (under development)

- ECMWF ERA-20C feedback

ODB ASCII

http://apps.ecmwf.int/datasets/data/era20c-ofa/

- But I concede that these are not the most user-friendly data to use, with the exception of GIO perhaps.
  - These are still disjoint efforts. No uniform language or harmonization, and no systematic effort across reanalyses. A meeting was organized to review the current situation, identify gaps, and propose ways forward for improvement

Meeting on reanalysis observation feedback, sponsored by

- Core-Climax is an EU FP7 project that ended yesterday (30 June 2015)
- Meeting was held at ECMWF, 11-13 November, 2014
- **Conclusions**
  - **reanalysis observation feedback contains highly valuable information to enhance understanding of Earth System Models and corresponding observations**, by confronting each other in the same variable, location and time.
  - Although **historical information exists about satellite observation data quality**, it is distributed between the several reanalysis centers' "blacklists" and the data providers' notes and logs. Much would be gained by exchanging all this information.
  - (…)
- Complete report and meeting presentations are available from http://www.coreclimax.eu/?q=Feedback

Meeting on reanalysis observation feedback, sponsored by

- **Reviewing current practices:**
  - **The satellite climate data record producers represented at the meeting (EUMETSAT and NOAA/STAR) express the need for quantitative, datum-level information, from the observation feedback produced by reanalyses.** This is to be used for detailed investigation of the impact of their products, and to improve their understanding of the data quality by comparison with the reanalysis, its quality control, the assimilation innovations (departures), and the bias corrections.
  - **Reanalysis producers do not generally provide convenient access to their observation feedbacks, which are all found to use complex and center-dependent data models, formats, attributes** (…)
- **Discussing ways forward to improve upon the current situation:**
  - **The group proposes a canonical definition of observation feedback,** whereby the smallest element is a record that identifies uniquely an observation datum, its position in space, time, the vertical or channel, and a series of basic, agreed, attributes, for which little confusion is possible.
  - (…)

One issue that came up as problematic

- Was Governance! It's o.k. to define some terms, but before investing in software to implement this, how to maintain definitions and evolve them?
- There is well-established governance for big-ticket contributors to the observing system:
  - WMO common tables section C for all (current) instruments and satellites
    - Though many defunct satellite sensors are not listed there
  - List of vessels and buoys ➔ In large collections such as ICOADS
  - Lists of upper-air sounding sites ➔ In large collections such as IGRA and CHUAN
- Governance less well-established for land surface stations ➔ badly need an ICOADS for land!
- No governance to define names for assimilation feedback

Before going further

- We need to speak the same language about the typology of observations
- We need to formalize the interactions, and how we name them
- We need to categorize the various levels of information contained in observations and models

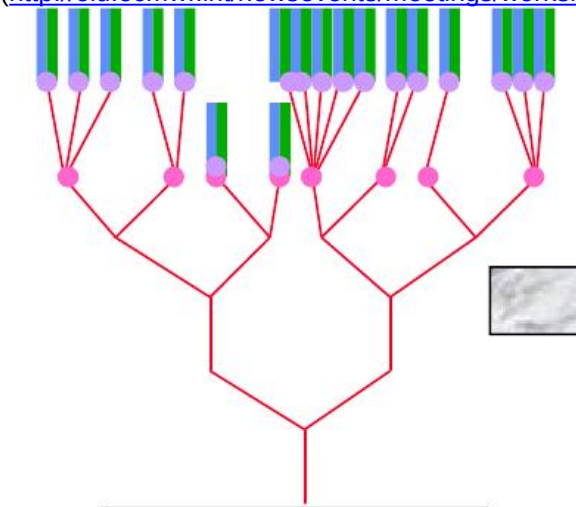
Concepts on how to exploit this information

- Reanalysis users enjoy the ‘regular’ hyper-cube that reanalysis offer: all spatio-temporal dimensions are covered, there is no gap, etc…
- Observations on the other hand present a highly irregular problem if considered as a hyper-cube:
  - Satellite sounders can have from 1 to several thousand channels
  - Radiosondes can have from ~100 to several thousand levels
  - The geophysical variables reported from surface stations typically depend on the route by which they were received (BUFR encoding for GTS)
- However, it is still possible to bring this problem back to a serial one, by considering the observing system as a tree of multiple branches

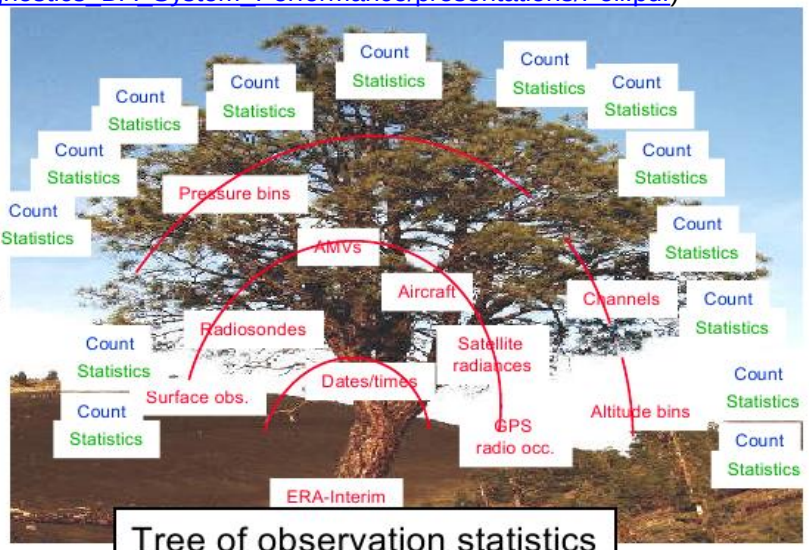
Concepts for a simple exploitation of observation feedback

- "Monitoring long data assimilation time series: a reanalysis perspective with Era-Interim" (2009)

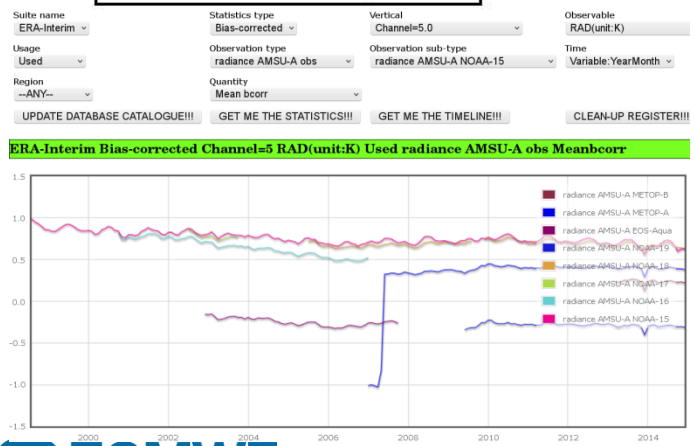
(http://old.ecmwf.int/newsevents/meetings/workshops/2009/Diagnostics_DA_System_Performance/presentations/Poli.pdf)



Tree of SQL requests



Tree of observation statistics



Browsable selection of dimensions, on-demand queries aggregating statistics from DBase

Snapshot of interactive tool built with: python, ODB, PostgreSQL, TURBOGEARS, MetPy, CherryPy, json.org, Google code, http://code.google.com/p/flot/

Conclusions

- Reanalysis observation feedback is a mine of information
  - Also allows to assess earlier projections (“obs vs. projected”)
- Does the reanalysis community keep its feedback to itself?
  - Not quite, but there is no standard for exchange and making available
  - Even though there are a the few datasets available
  - Some initial agreements have been reached between reanalysis producers, but turning these good intentions into practical, standardised datasets requires governance
- Opening up the access will attract users
  - To learn about observations and improve the record
  - Also as educational material for ‘big data mining’
- However, serving the ‘raw’ data will not be sufficient
  - To ease user uptake, one needs computing/aggregation, visualization facilities to allows ‘discovery’ of the observing system: movie example