Introduction to Re-Analysis: the ECMWF experience

Paul Poli, Franco Molteni

European Centre for Medium-Range Weather Forecasts, Reading, U.K.

Outline

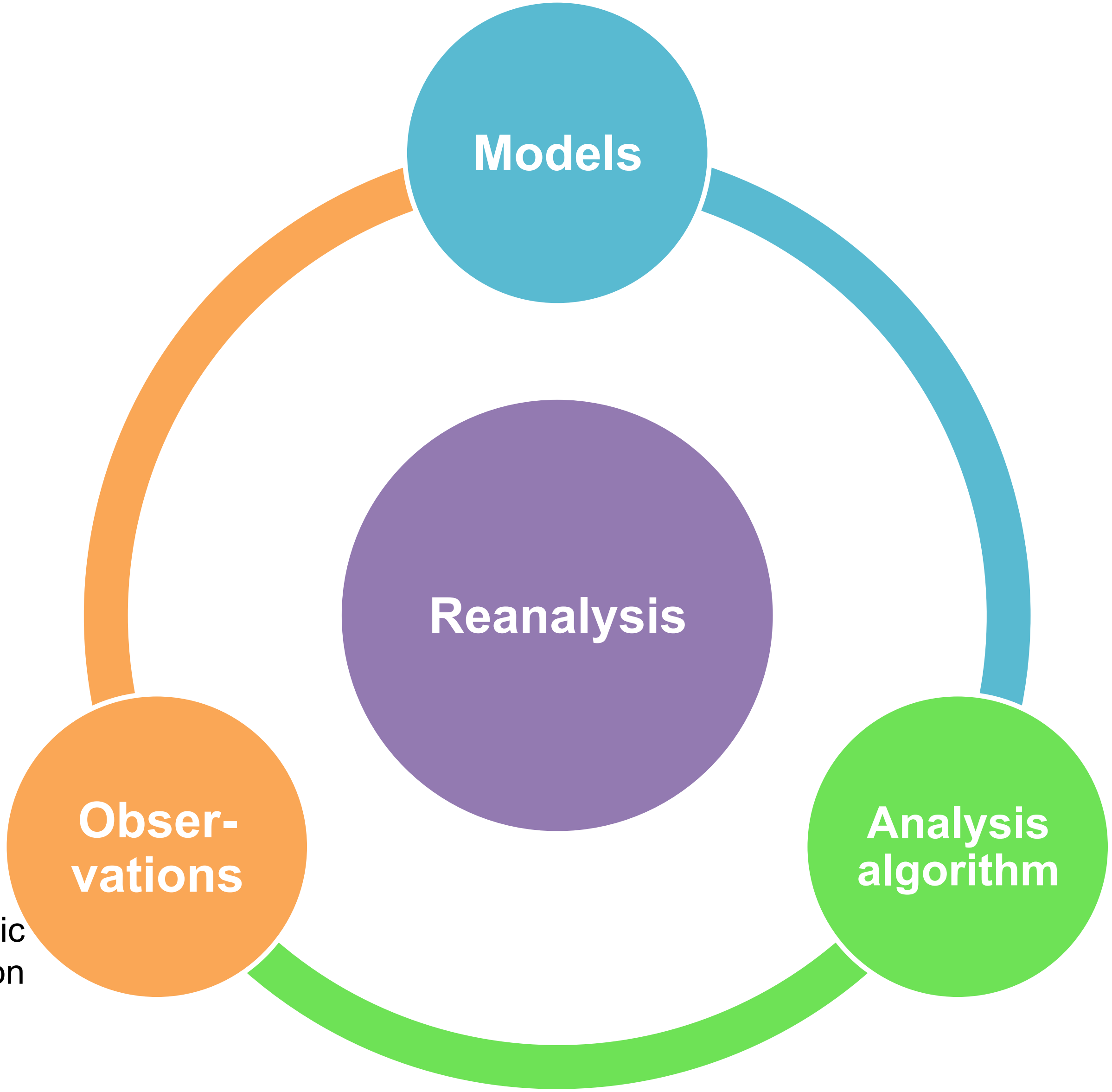
What is reanalysis ?
- Simple concepts
- Goals of reanalysis

How do we do reanalysis?
- Observations
- Model
- Data assimilation

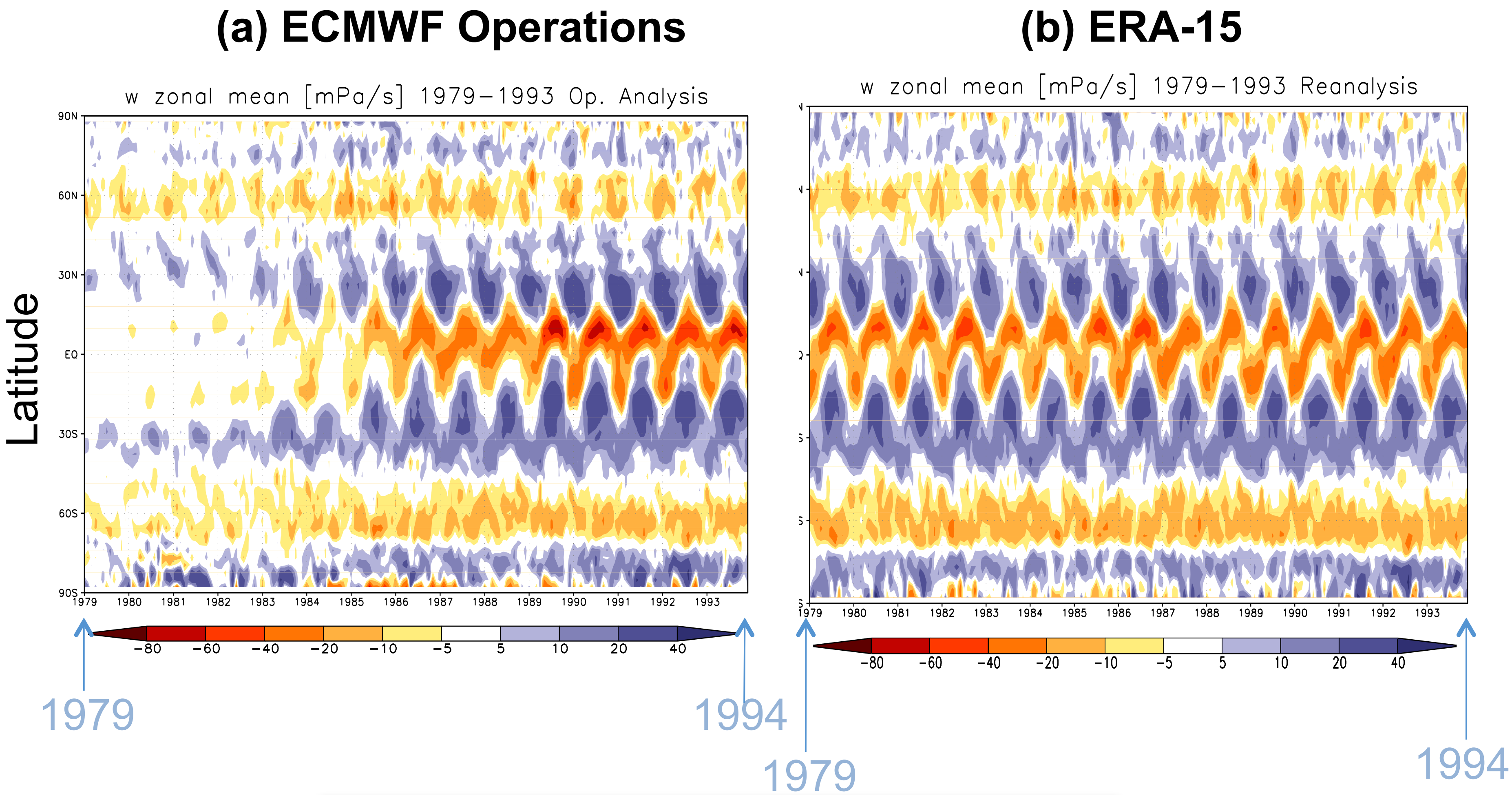
Reanalysis projects & results
- Projects
- Users
- Applications

Conclusions
- Summary
- Challenges ahead

Simple concepts

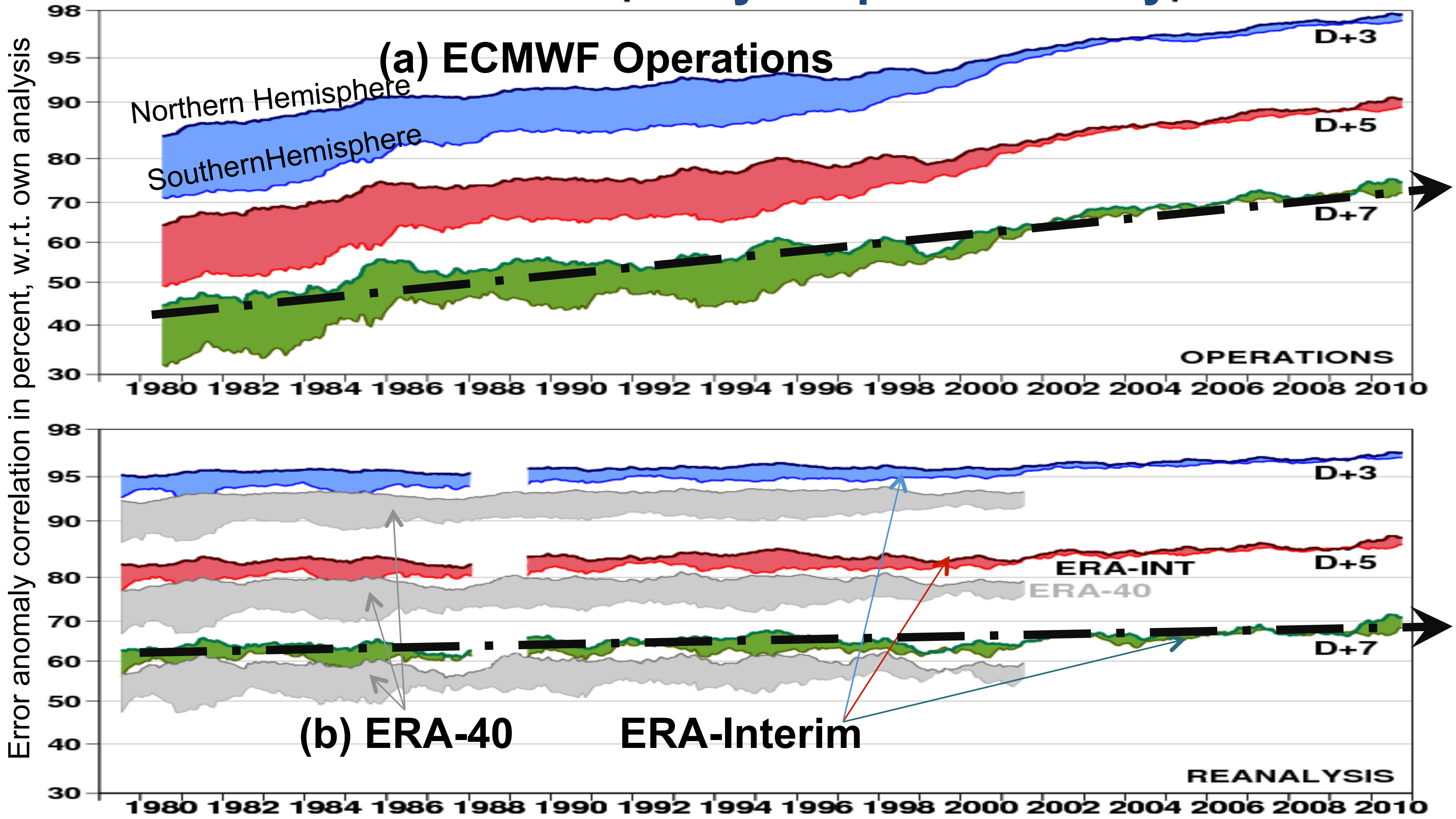


Why reanalysis?
1- Greater time-consistency of the products



Zonal mean vertical velocity at 500 hPa

Why reanalysis?
2- Greater time-consistency for predictability studies



Anomaly correlation of 500 hPa Z Forecasts

Why reanalysis?
3- Advantages against “observations-only” multi-decadal gridded datasets … for climate studies

1) How reanalysis deals with “missing data”

- Only assimilate observations when and where we have them!
- … instead of reverting to a crude, 2^{nd}-order, unphysical interpolation to “fill in the blanks”

2) Produced fields are space- and physically-consistent

- As specified by the NWP model

3) Use the widest variety of observations

- Not just temperatures, or winds, or humidities in isolation of each other…
- … but also pressures, satellite observations, … multi-variate approach

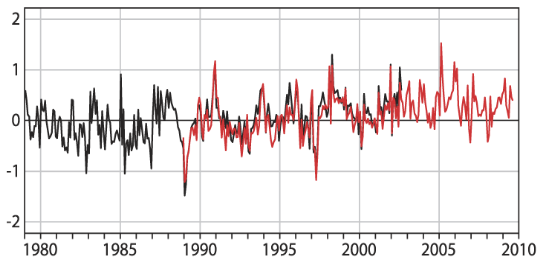
4) All observations are evaluated/used in a consistent way

- Accuracy and precision explicitly taken into account
- Seamless quality control (QC) procedures, across all observation types
- The background prediction provides a unique advantage for QC

Conclusion

- So, yes reanalysis combines lots of difficulties due to changes in observations input… but like with ANY OTHER observations-based dataset, the basic challenge is the same (change in observations’ quality and quantity over time). The difference is, we try to do things in a consistent manner, by applying the same methodology of ***data assimilation*** for all observations

Summary of the goals: reanalysis products should be consistent …

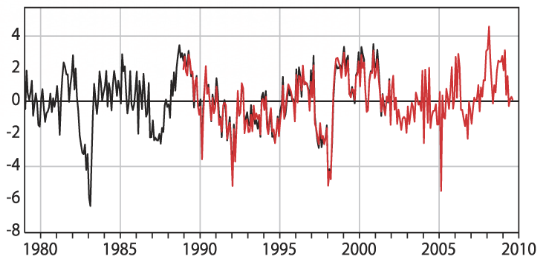


2-metre temperature anomaly (°C) over Africa

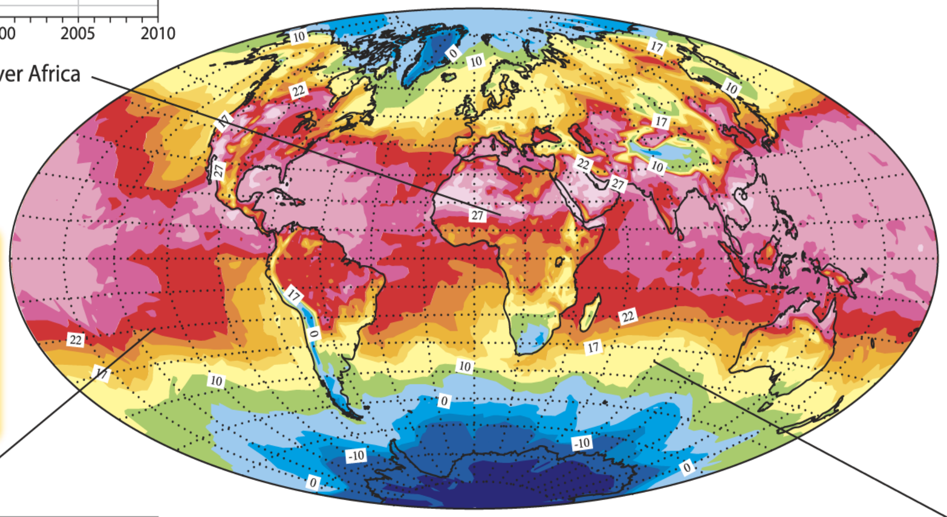
…in Time

…across Atmospheric Parameters

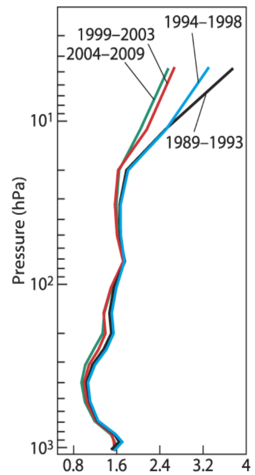
Southern Oscillation Index (hPa)



ERA-Interim 2-metre temperature (°C)
15 August 2003 03 UTC



…in the Horizontal



Standard deviation of differences between ERA-Interim and radiosondes temperature (°C) in the southern hemisphere

…in the Vertical

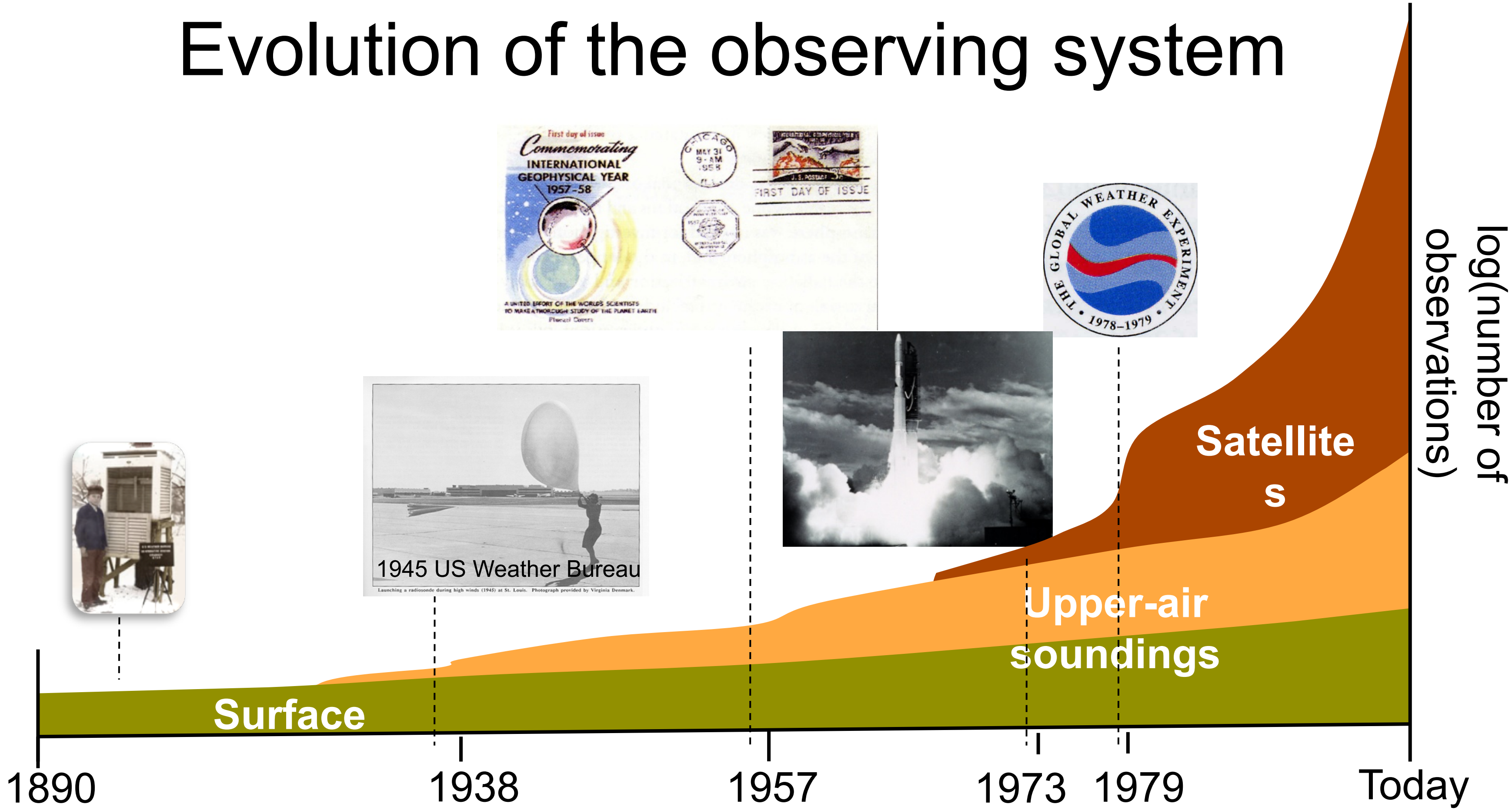
Outline

What is reanalysis ?	• Simple concepts • Goals of reanalysis
How do we do reanalysis?	• Observations • Model • Data assimilation
Reanalysis projects & results	• Projects • Users • Applications
Conclusions	• Summary • Challenges ahead

Reanalysis components
Part 1: Observations

Use as many observations as possible	• Goal being to produce the best estimate of the atmospheric state, at any given time and place
Use "good" observations	• Use corrected/reprocessed datasets when available • Focus efforts on long records
Keep track of what goes in/comes out	• Implement dedicated monitoring at key points of the assimilation (observation ingest, blacklisting, thinning, assimilation)
Keep that setup throughout the reanalysis	• Beware of large components of the observing system that suddenly disappear from the assimilation

Evolution of the observing system



1890: Manual stations, limited data exchange

1938: First radiosonde networks, systematic soundings

1957: International Geophysical Year: radiosonde network enhanced, especially in the Southern Hemisphere

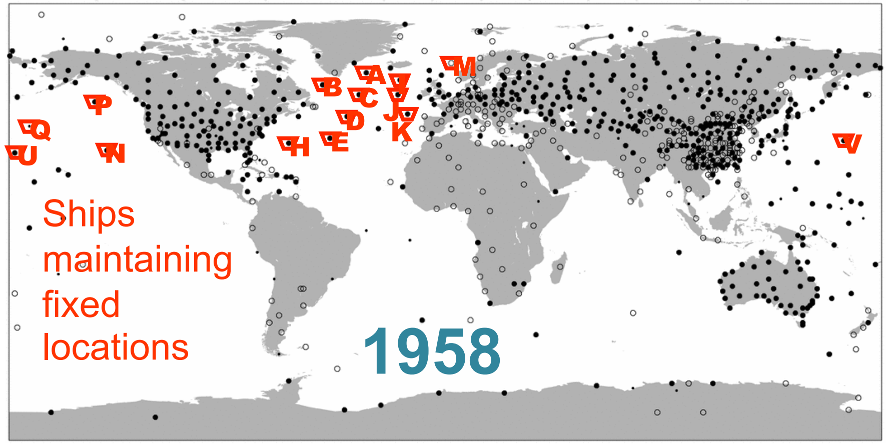
1973: First operational satellite soundings (NOAA-2)

1979: Improved sounding from polar orbiters; Winds from geostationary orbit; More data from commercial aircraft; Drifting buoys

Today: More satellites, aircraft, buoys, ocean drifters. Radiosondes: fewer, but probe higher. Better knowledge of instrument errors. More obs. per hour per sensor.

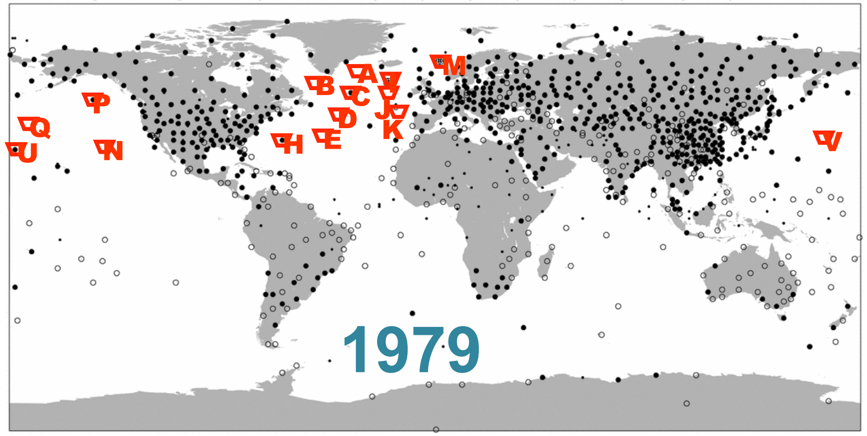
Evolution of radiosonde coverage

1°1° average in 3 categories: > 1 per week -> 0.5 per day: 0.5 per day -> 1.5 per day: > 1.5 per day



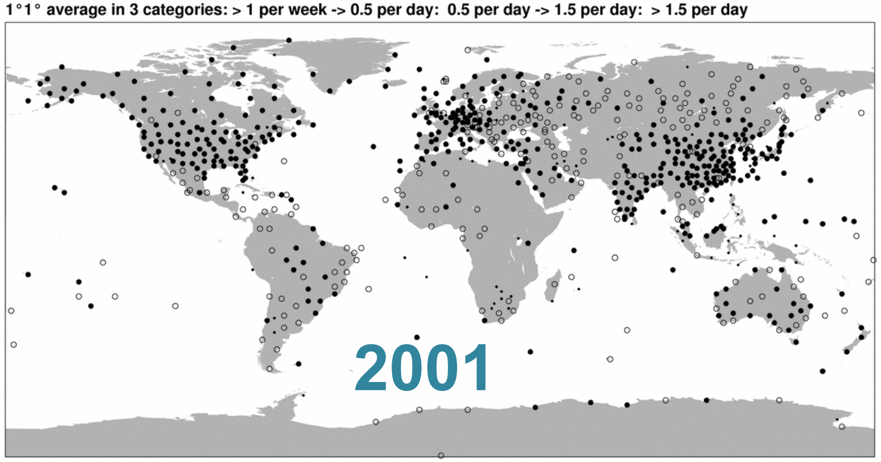
Average number of soundings per day: 1609

1°1° average in 3 categories: > 1 per week -> 0.5 per day: 0.5 per day -> 1.5 per day: > 1.5 per day



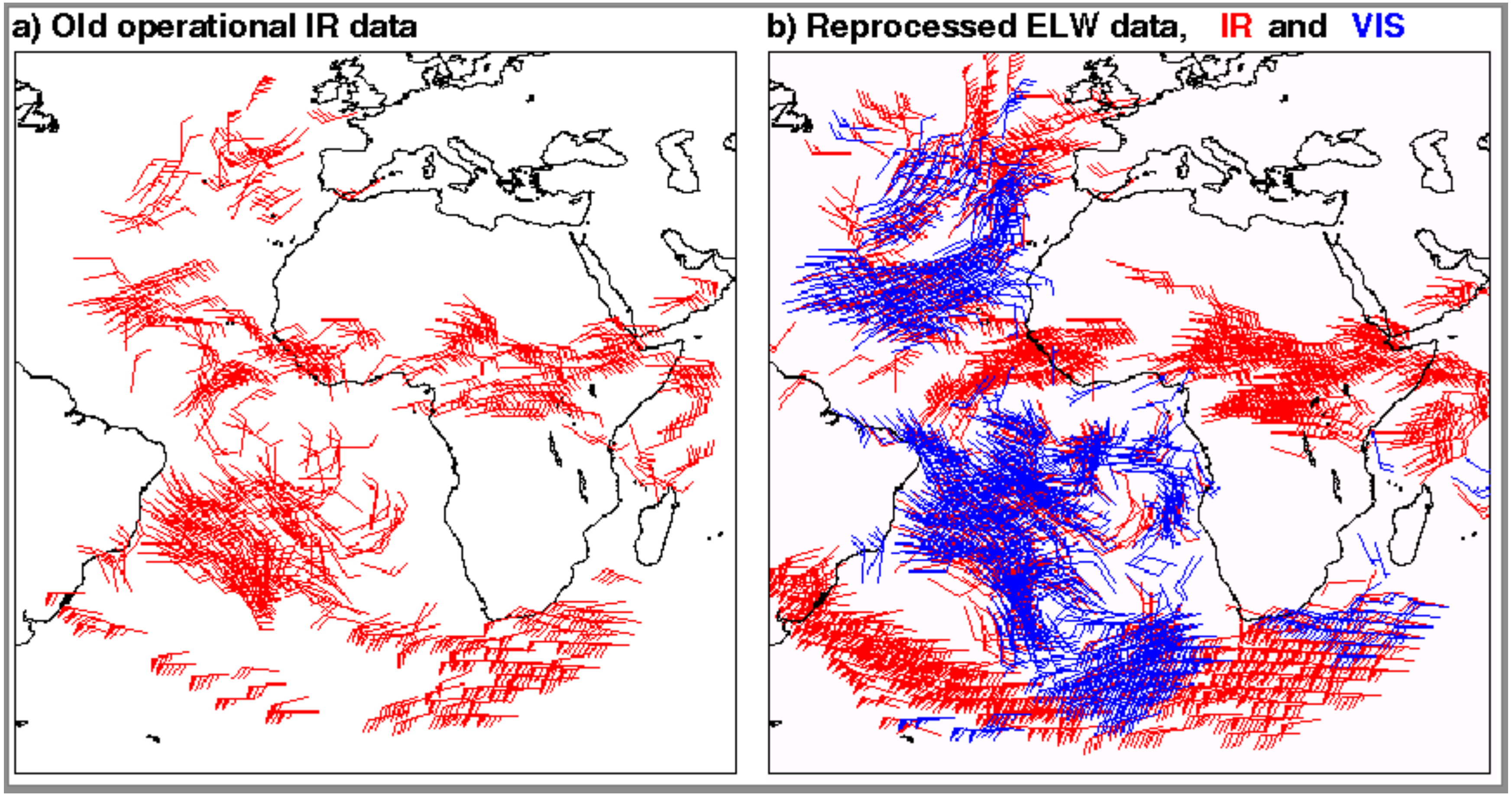
Average number of soundings per day: 1626

1°1° average in 3 categories: > 1 per week -> 0.5 per day: 0.5 per day -> 1.5 per day: > 1.5 per day



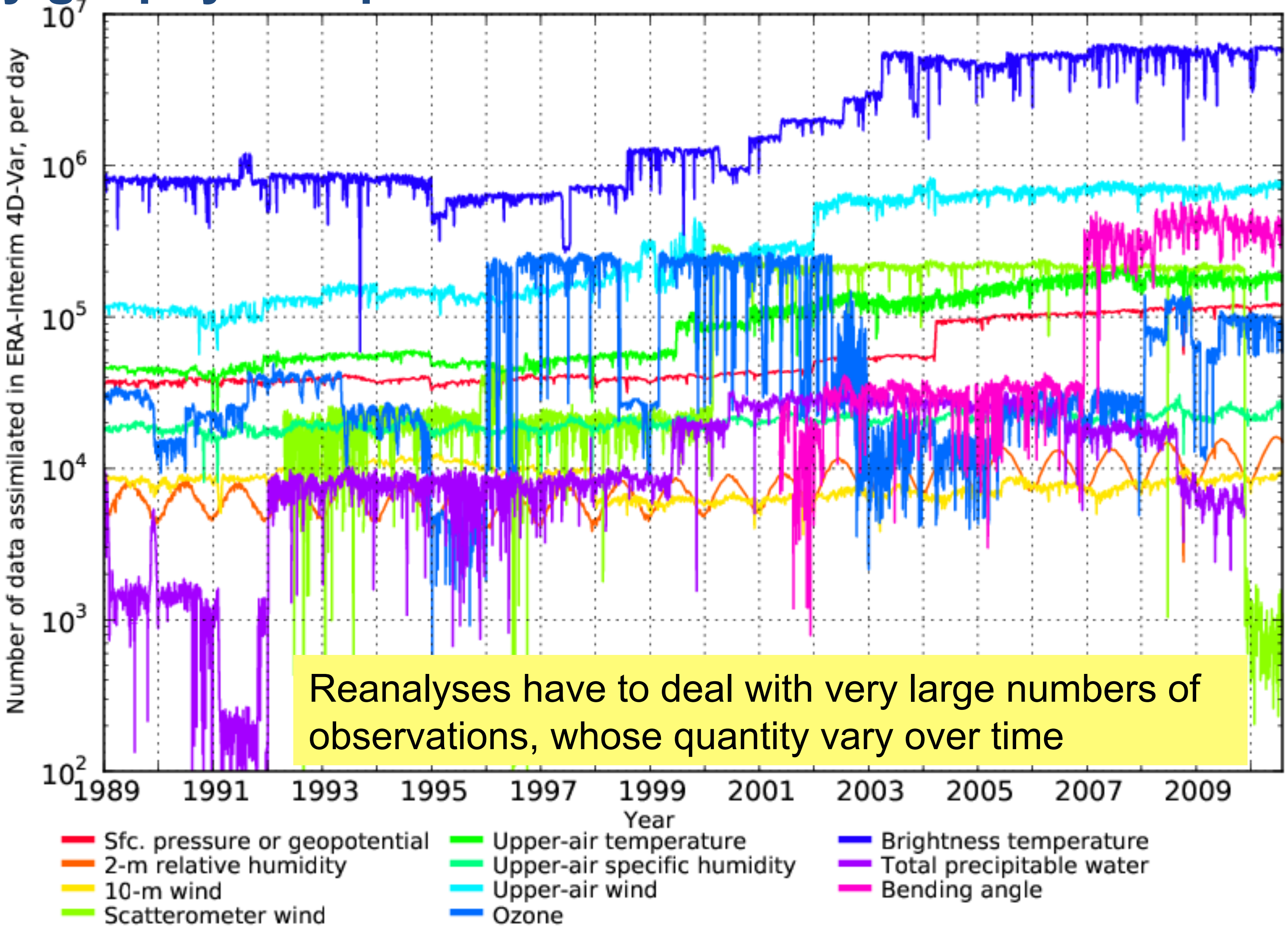
Average number of soundings per day: 1189

Example of improved data coverage, through reprocessing of Meteosat data into Atmospheric Motion Vectors



Early 1980s Expanded Low-resolution Winds

Number of observations assimilated in ERA-Interim, by geophysical parameter

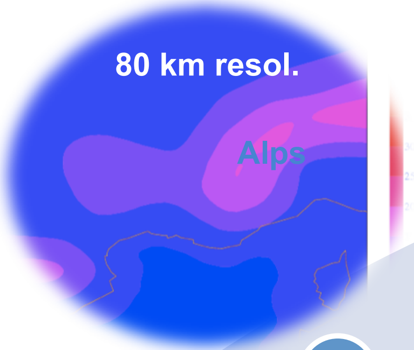


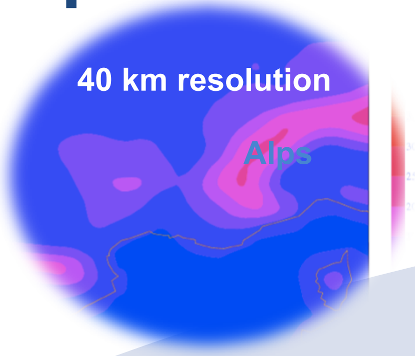
Reanalysis components
Part 2: forecast model

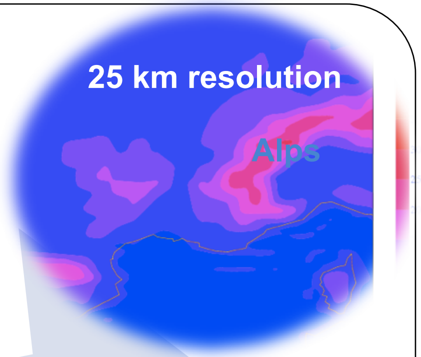
Use a fixed model version	•Dynamics, physics etc…
Use a fixed resolution	•Must be computationally affordable to cover several decades •For N decades to be run in 1 year: implies a factor N in run-time •10-year reanalysis in 1 year = run @ 10 days/day
Use the "best" model around	•Use the near-latest, stable, model version that was operational at some point •Not the time to start experimenting with new, untested model configurations
Shop around for good forcing data	•Ideally, one dataset per forcing type, to cover the whole time period
Keep that setup throughout the reanalysis	•Be extra careful with forcing data – any problem will map directly into the products! •Be extra careful when changing machine, compiler….

Illustration of resolution improvements

Orography of the Western Alps (500m contours)







Heritage

- **FGGE**
  1979
- **ERA-15:** 190 km resolution
  (1979-1993)
- **ERA-40:** 125 km resolution
  (1957-2002)

ERA-Interim
80 km resolution
(1979-present)

ERA-CLIM preparatory pilot reanalyses

- **ERA-20C**
  125 km resolution
  (1900-2000)
- **ERA-20C Land**
  25 km resolution
  (1900-2000)
- **ERA-SAT**
  40 km resolution
  (1979-onwards)

Reanalysis components
Part 3: Data assimilation

Use a fixed DA scheme

- Think ahead and put in it all that you need, observation-wise:
  - A blacklist that covers the entire reanalysis period
  - Observation operators, thinning, obs. errors etc…
- Use a B matrix that's fit for the job (ideally, flow-dependent…)

Test the DA scheme with various amounts of observations

- With few observations (1), and with all observations (2)
- To get a feeling for how the products will be affected when going from (1) to (2)

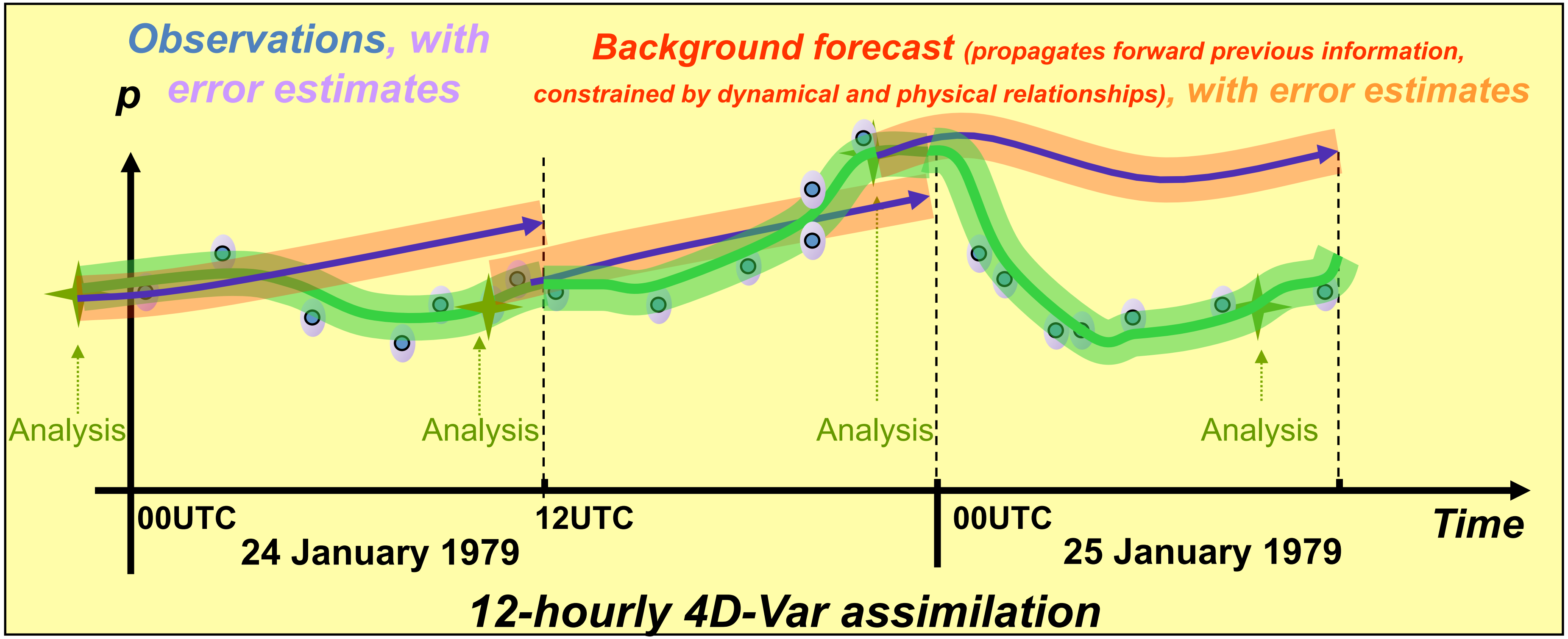
Do not ignore biases

- Implement bias correction algorithms, ideally with some underlying physical foundations:
  - Radiation corrections for radiosondes,
  - Radiative transfer model biases, etc…

Keep that setup throughout the reanalysis

- Be extra careful during run-time etc… so as not to lose the backward-compatibility (needed to rerun!)

Data assimilation process – applied to reanalysis



This produces the "most probable" atmospheric state ***

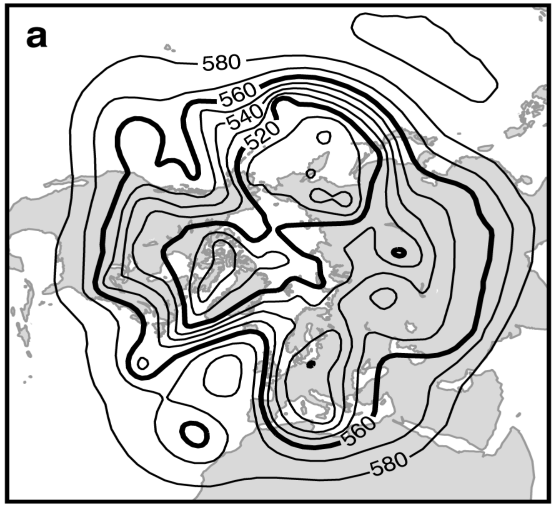
***In a maximum-likelihood sense, which is equivalent to the minimum variance, provided that **background and observation errors are Gaussian**, **unbiased**, **uncorrelated with each other**; all error covariances are correctly specified; model errors are negligible within the 12-h analysis window

For each analysis, construct a cost function and find its minimum:

$$J(\mathbf{x}) = \underbrace{(\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1} (\mathbf{x}_b - \mathbf{x})}_{\text{background constraint}} + \underbrace{[\mathbf{y} - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - \mathbf{h}(\mathbf{x})]}_{\text{observation constraint}}$$

$$\mathbf{h}(\mathbf{x}) = \mathbf{h}[M(\mathbf{x})]$$ simulates the observations

Illustration: why an advanced DA scheme is helpful in reanalysis

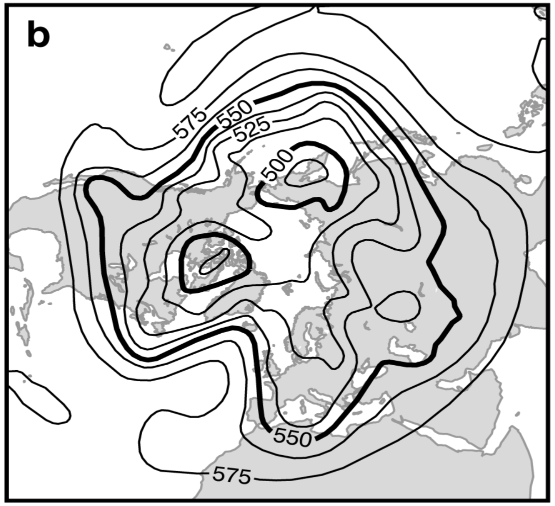


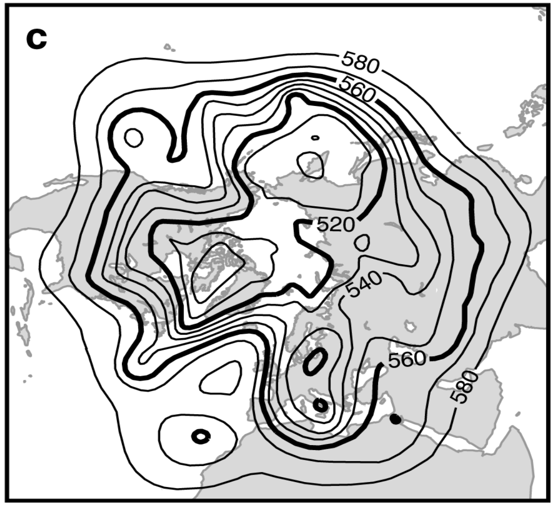
15 February 2005 00 UTC

4D-Var CONTROL

All observations (surface, radiosondes, satellite etc…)

3D-Var





4D-Var

Surface pressure observations only

Advances in data assimilation can help extract more information from historic data that could ever be thought possible at the time the observations were collected

Outline

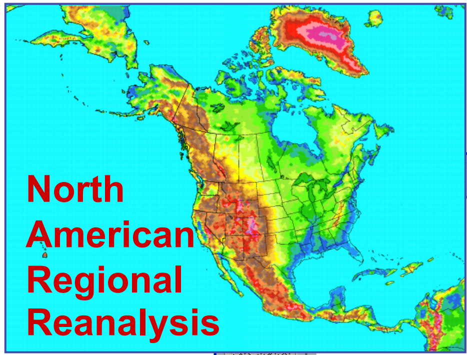
What is reanalysis ?	• Simple concepts • Goals of reanalysis
How do we do reanalysis?	• Observations • Model • Data assimilation
Reanalysis projects & results	• Projects • Users • Applications
Conclusions	• Summary • Challenges ahead

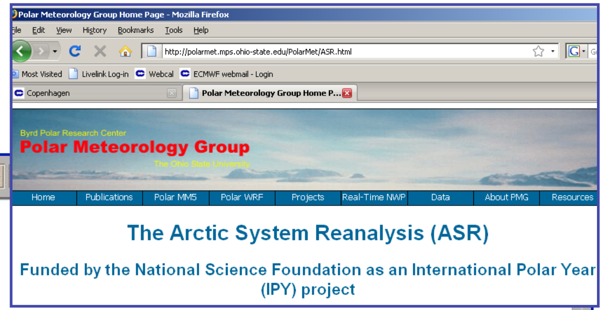
A short history of atmospheric reanalysis

- **1979**: **Observation datasets collected for the First GARP Global Atmospheric Research Program Experiment (FGGE)**: used *a posteriori* for several years, to initialize NWP models (= the first reanalyses!), to compare performances and progress
- **1983**: **Reanalysis concept first proposed** by Daley for monitoring the impact of forecasting system changes on the accuracy of forecasts
- **1988**: **Concept proposed again, but for climate-change studies**, in two separate papers: by Bengtsson and Shukla, and byTrenberth and Olson
- **1990s**: **First-generation comprehensive global reanalysis** products *(~OI-based)*
  - NASA/DAO (1980 - 1993) from USA
  - NCEP/NCAR (1948 - present) from USA
  - ERA-15 (1979 - 1993) from ECMWF – with significant funding from USA
- **Mid 2000s**: **Second-generation products *(~3DVAR)***
  - JRA-25 (1979 - 2004) from Japan
  - NCEP/DOE (1979 - present) from USA
  - ERA-40 (1958 - 2001) from ECMWF – with significant funding from EU FP5
- **Today: third generation of comprehensive global reanalyses *(~4DVAR or IAU)***
  - JRA-55 (1958 - 2012) from Japan
  - NASA/GMAO-MERRA (1979 - present) from USA
  - NCEP-CFSRR (1979 - 2008) from USA
  - ERA-Interim (1979 - present) from ECMWF

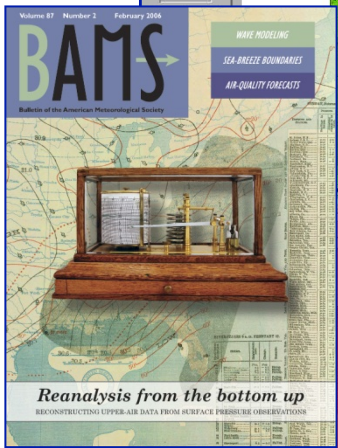
Atmospheric reanalysis: becoming more diverse

- Regional reanalysis and down-scaling from global reanalysis

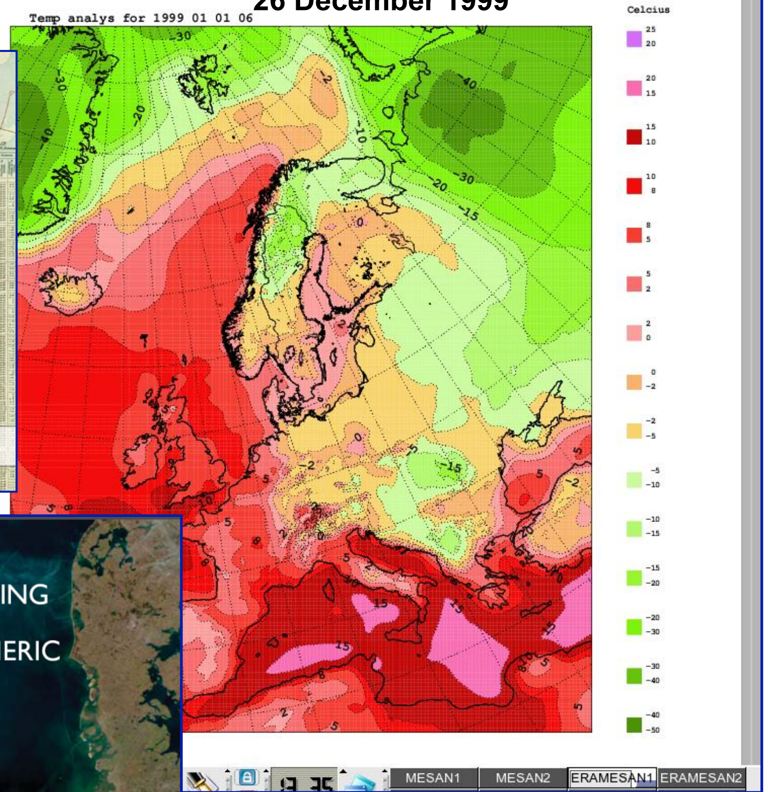




- Long-term reanalysis using only surface-pressure observations: 20^{th} Century Reanalysis (20CR)

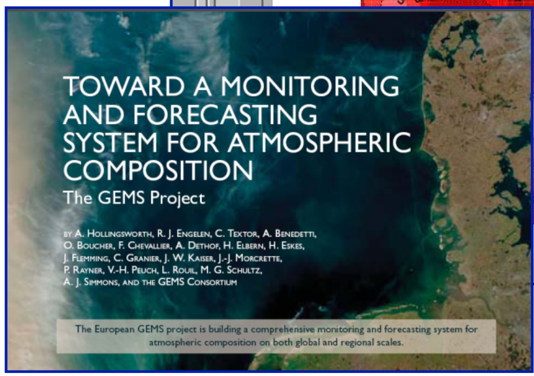


Maximum gusts
26 December 1999



2m temperature
6UTC, 1 January 1999

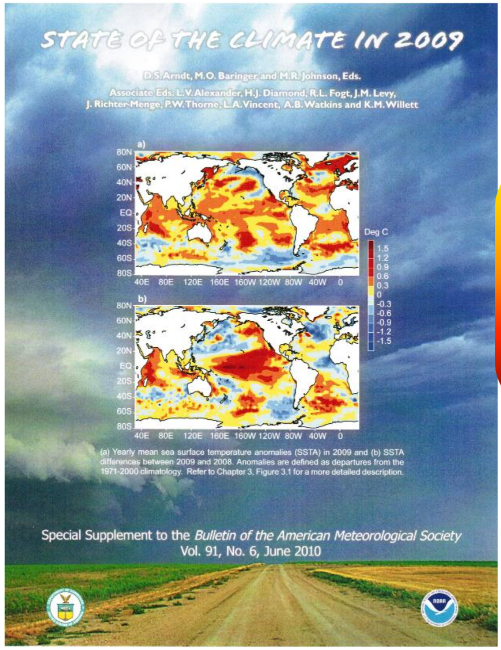
- Short-term reanalysis for chemistry& aerosols



How users exploit reanalysis data

- Monitor the observing system
  - Feedback on observational quality, bias corrections and a basis for homogenization studies of long data records that were not assimilated
- Develop climate models
  - Use reanalysis products for verification, diagnosis, calibrating output,, …
- Drive users' models/applications
  - Use reanalysis as large-scale initial or boundary conditions for smaller-scale models (global→regional; regional→local), in various fields: ocean circulation, chemical transport, nuclear dispersion, crop yield, health warnings, …
- Use climatologies derived from reanalysis for direct applications
  - Ocean waves, resources for wind and solar power generation, insurance, …
- Study short-term atmospheric processes and influences
  - Process of drying of air entering stratosphere, bird migration, …
- Study of longer-term climate variability/trends
  - Requires caution due to changes in observations input
  - Lead to major findings in recent years in understanding variability

Growing recognition of reanalysis for climate application: BAMS “State of the Climate in 2009”



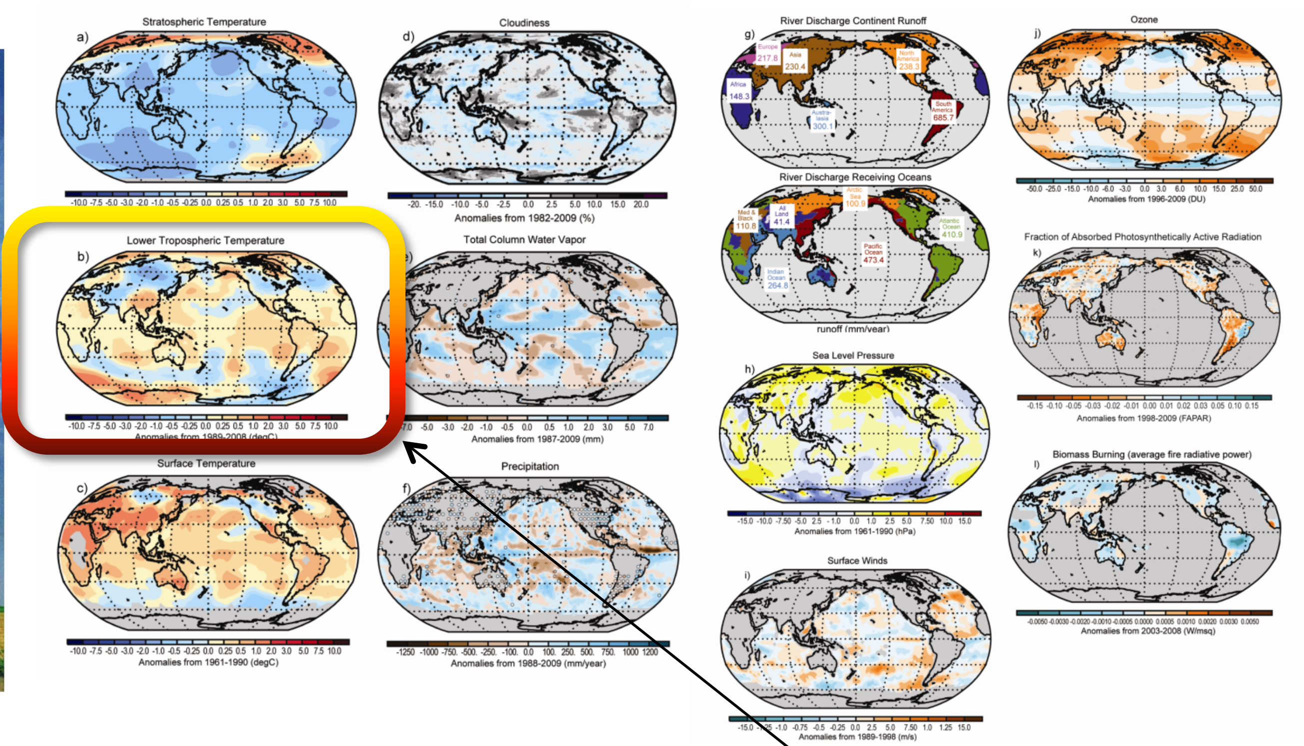
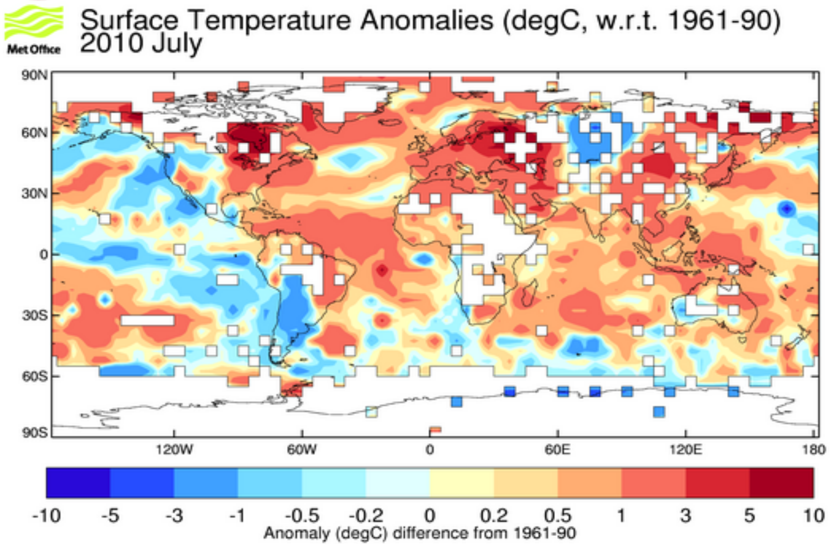


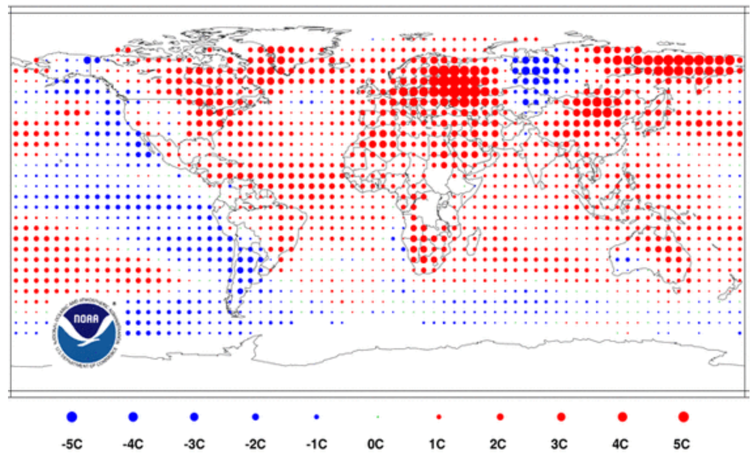
Plate 2.1. Global annual anomaly maps for those variables for which it is possible to create a meaningful 2009 anomaly estimate. Climatologies differ among variables, but spatial patterns should largely dominate over choices of climatology period. Dataset sources/names are as follows: lower stratospheric temperature (RSS MSU); lower tropospheric temperature (ERA-interim); surface temperature (NOAA NCDC); cloudiness (PATMOS-x); total column water vapor (SSM/I over ocean, ground based GPS over land); precipitation (RSS over ocean, GHCN (gridded) over land); river discharge (authors); mean sea level pressure (HadSLP2r); wind speed (AMSR-E); ozone (GOME2); FAPAR (SeaWIFS); Biomass Burning (GEMS/MACC). See relevant section text and figures for more details.

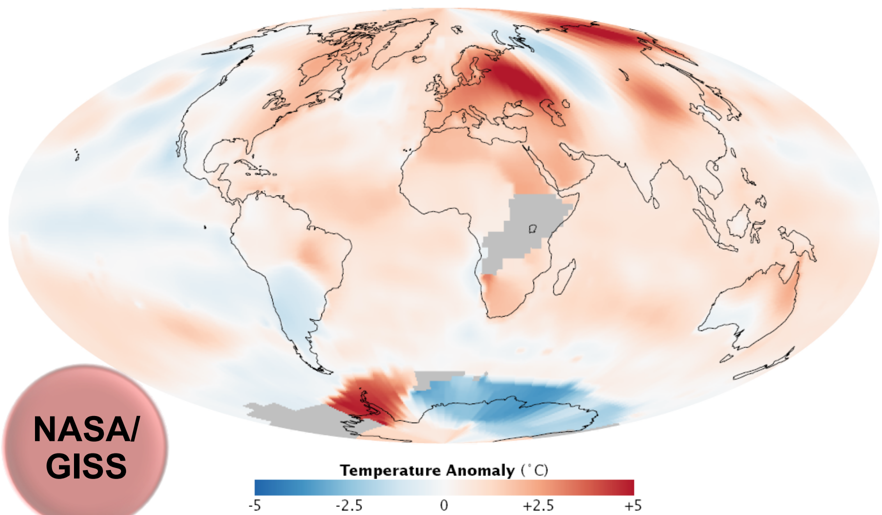
Temperature: Global anomalies July 2010

Hadley Centre

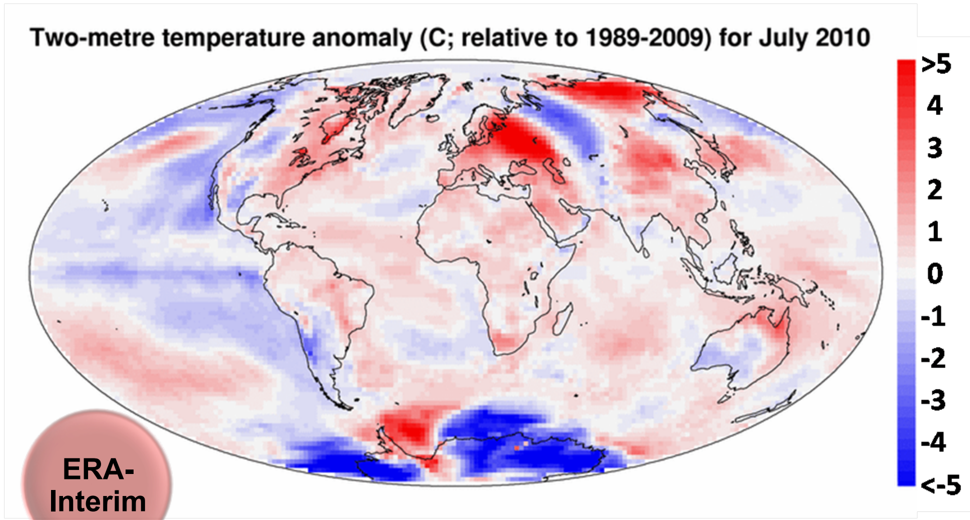


NOAA/ NCDC





NASA/ GISS

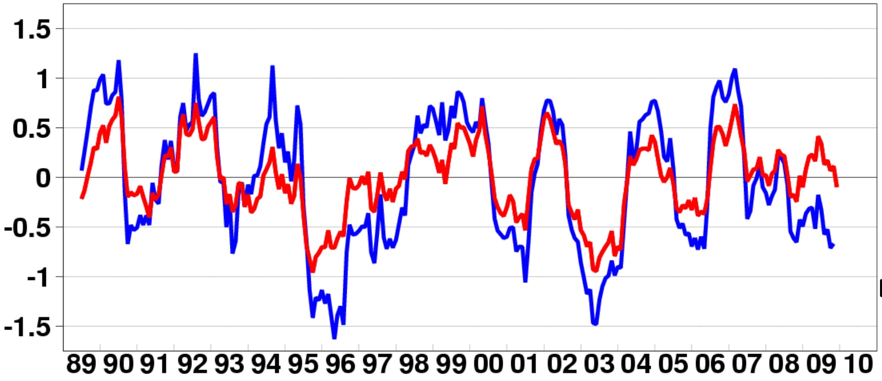


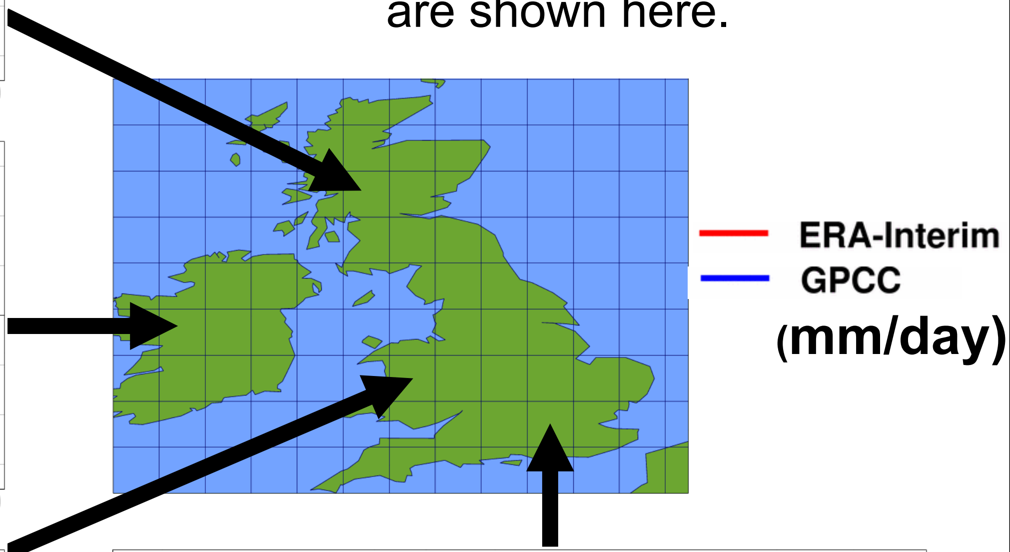
ERA-Interim

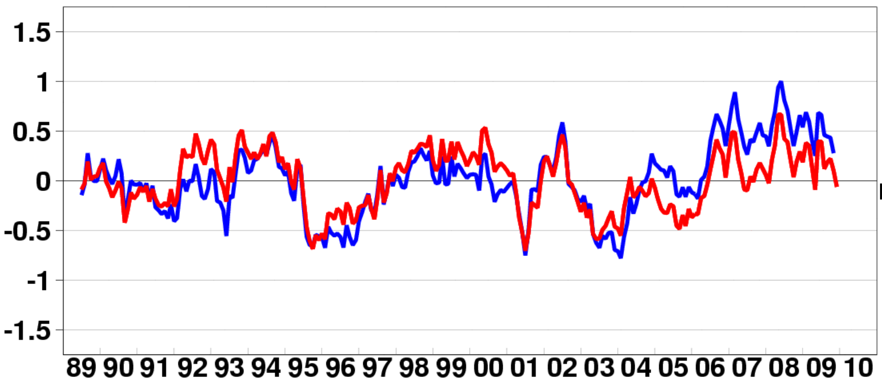
Rainfall: Regional anomalies for 1x1 degree boxes

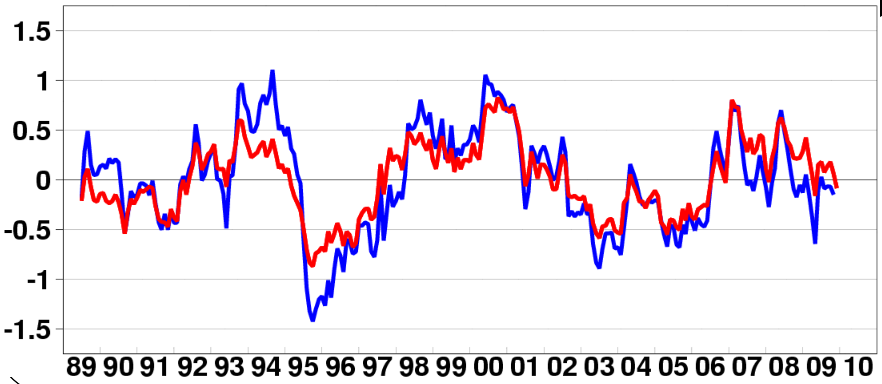
Anomalies are computed with respect to (1989-2009) means for each month from ERA and GPCC respectively.

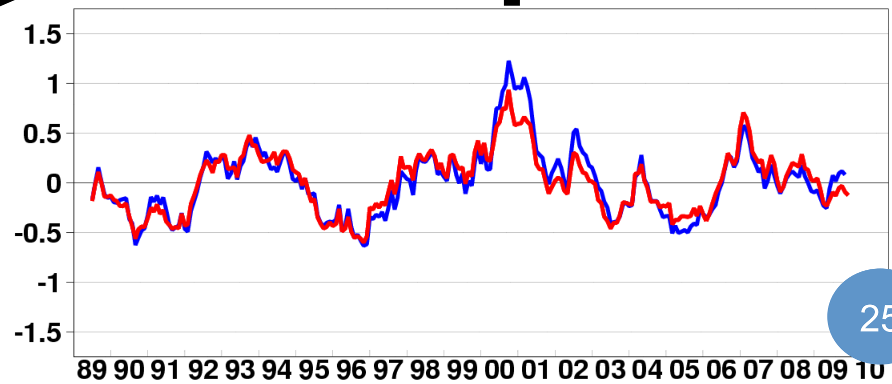
Time series of 12-month running means are shown here.



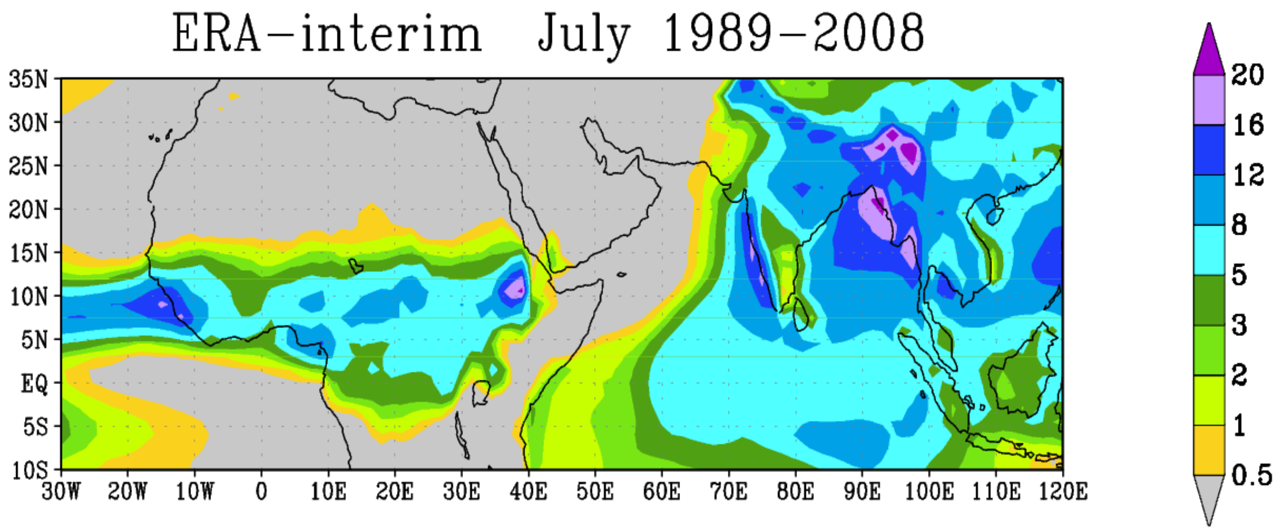


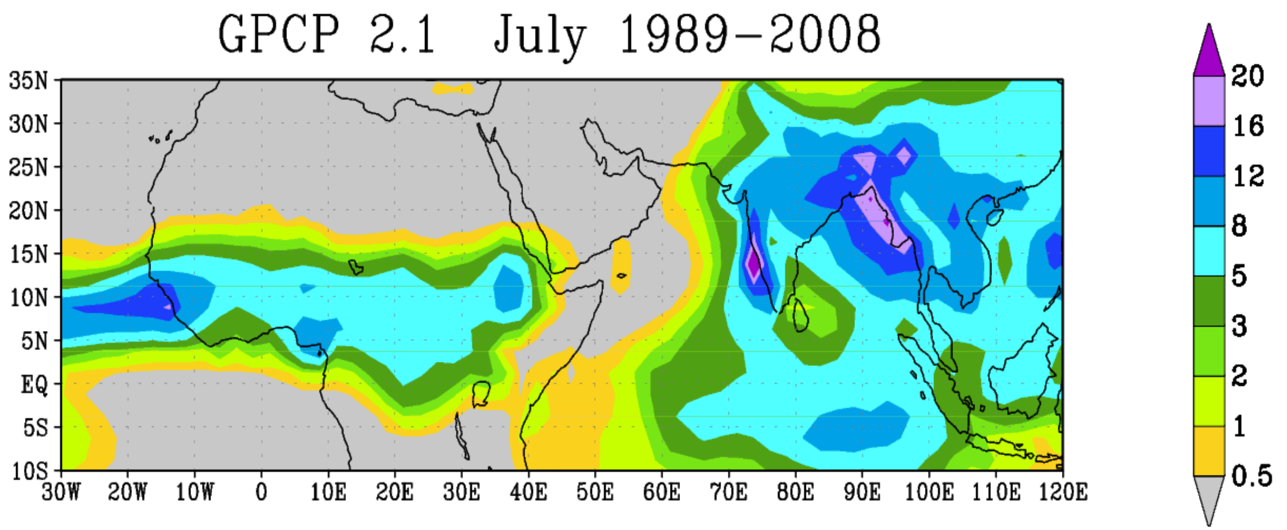






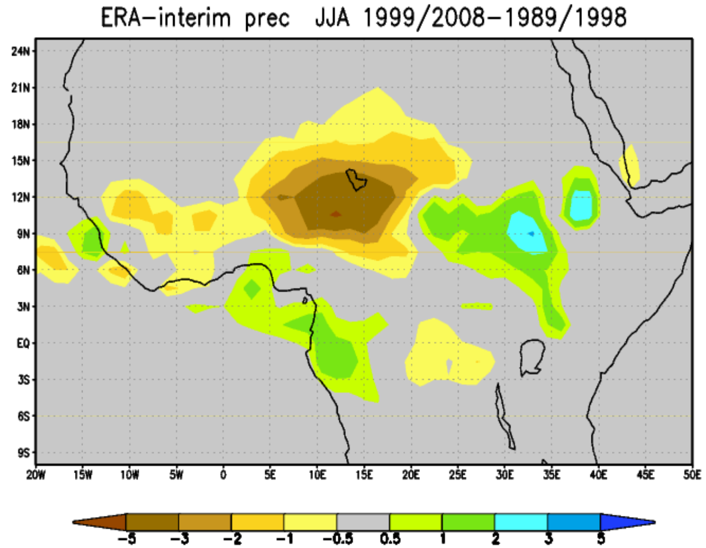
Rainfall climatology in July



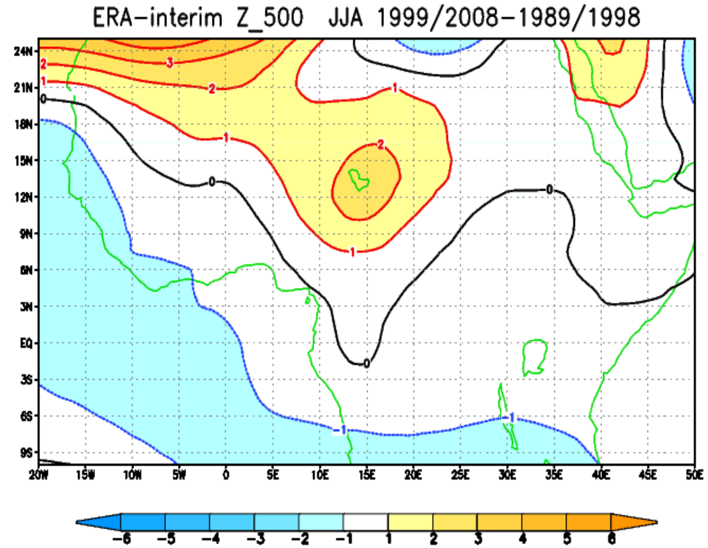


ERA-interim: trends over Africa in JJA

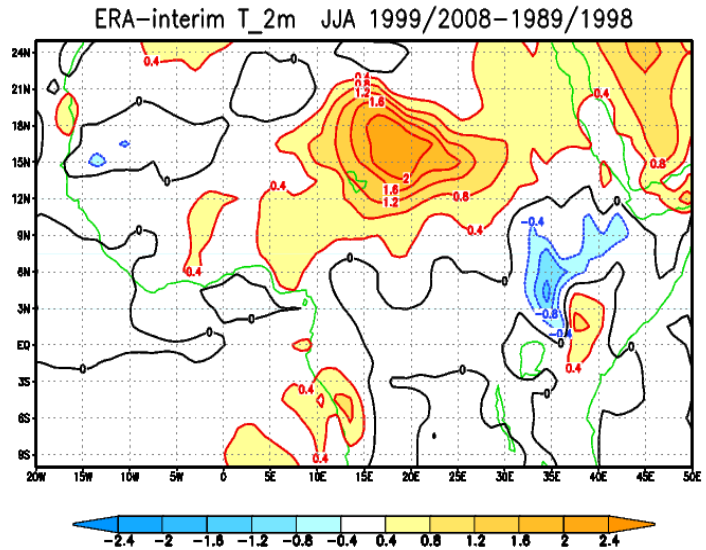
Prec



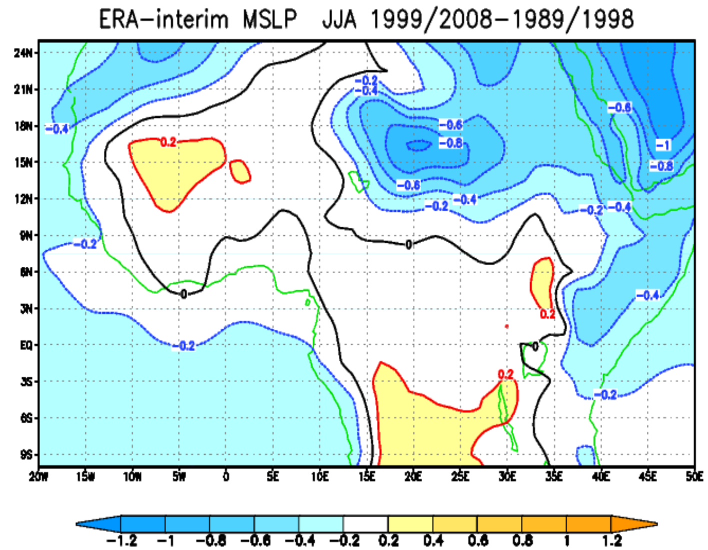
Z 500



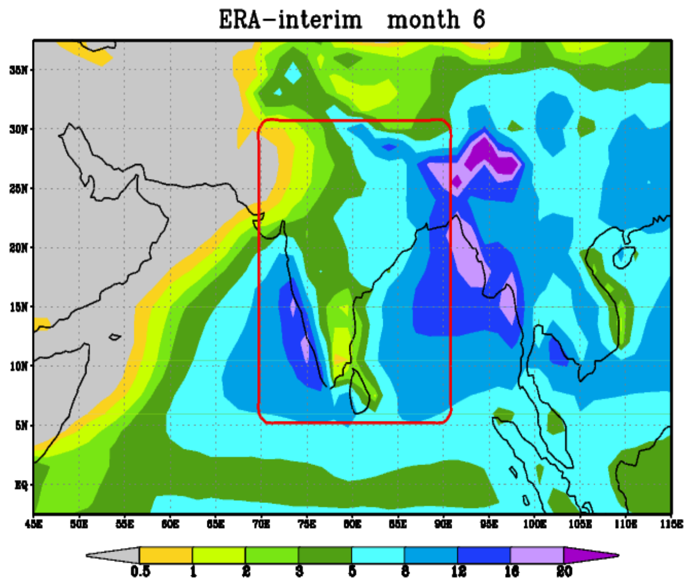
T_2m

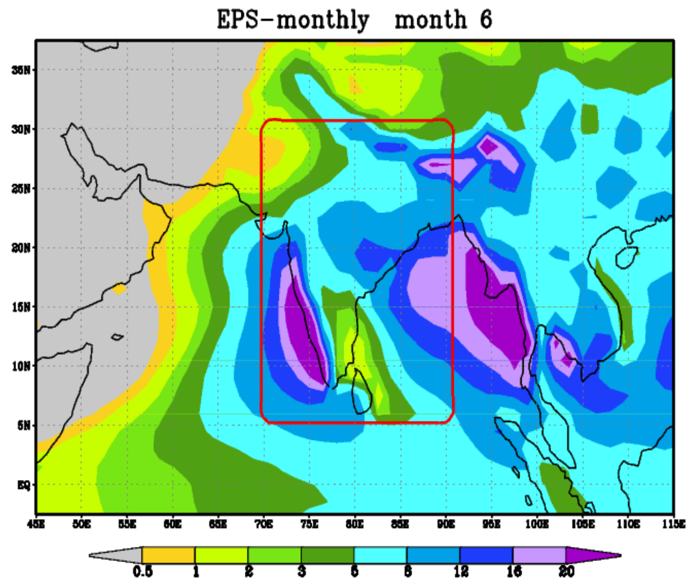


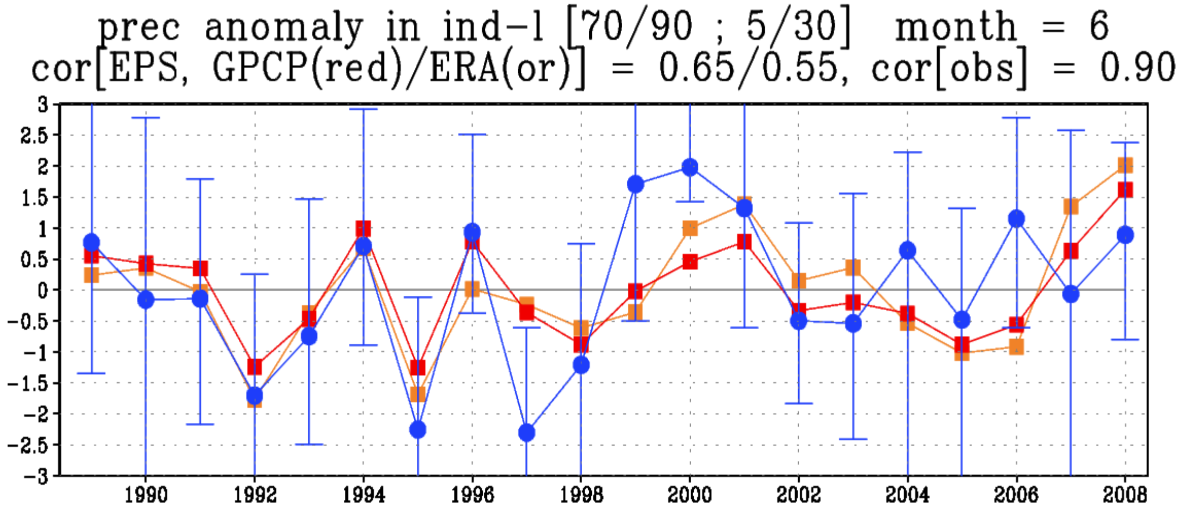
SLP



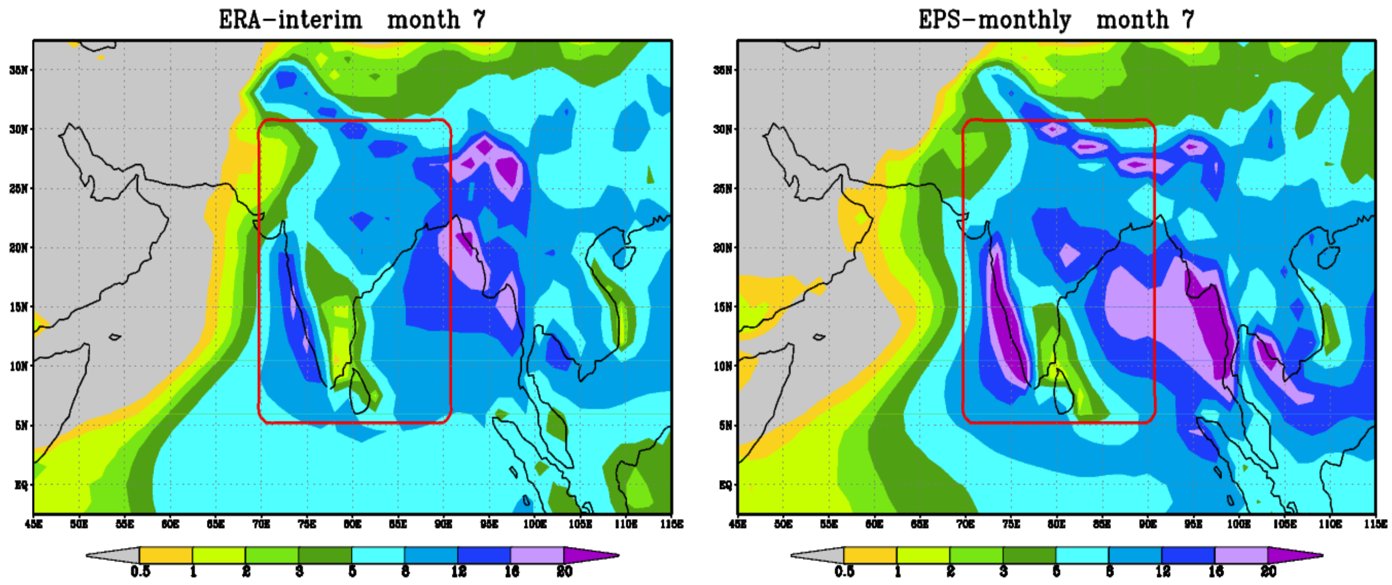
46-day EPS re-forecasts: India, June

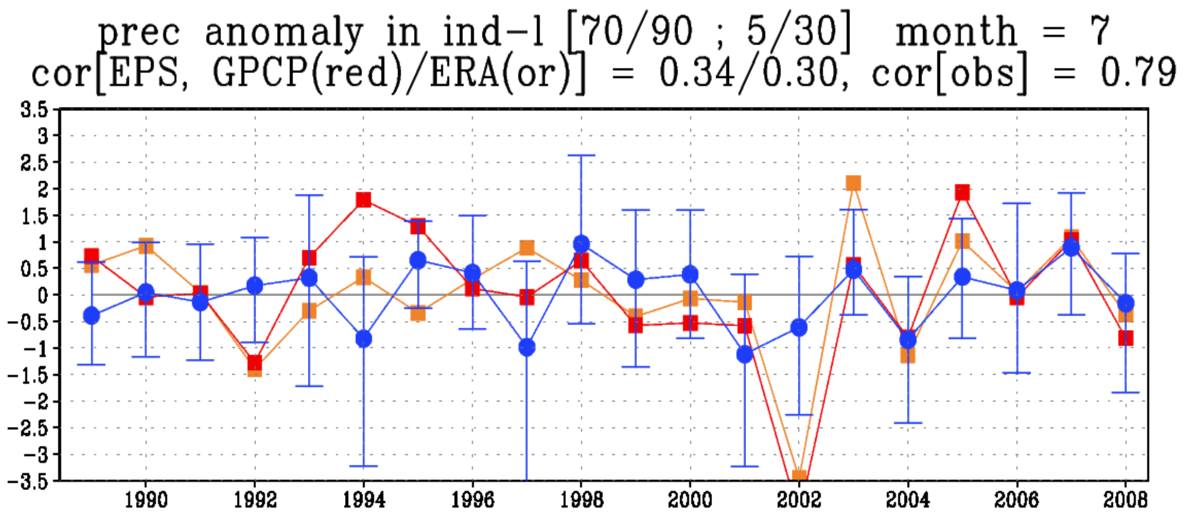




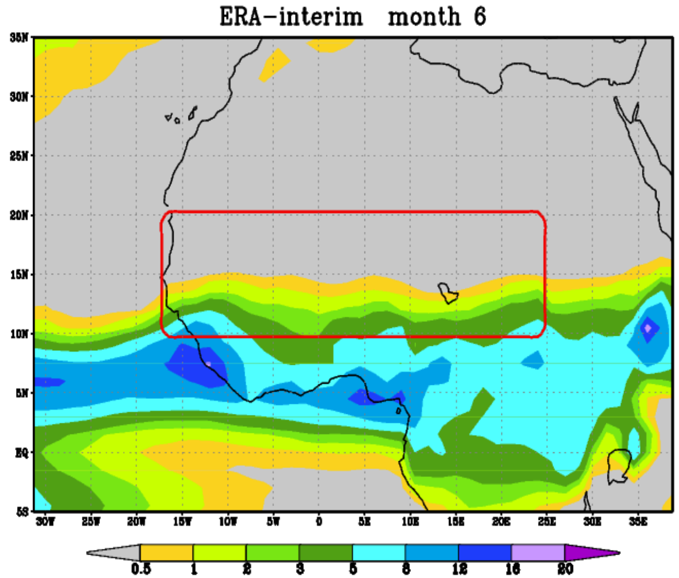


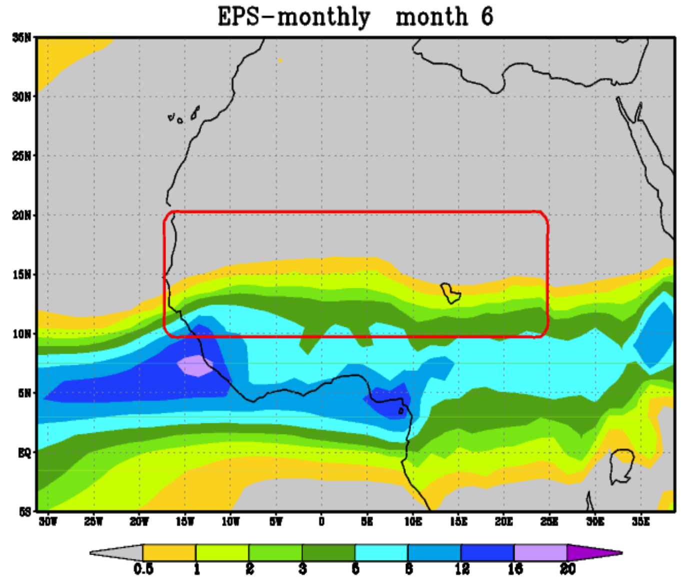
46-day EPS re-forecasts: India, July

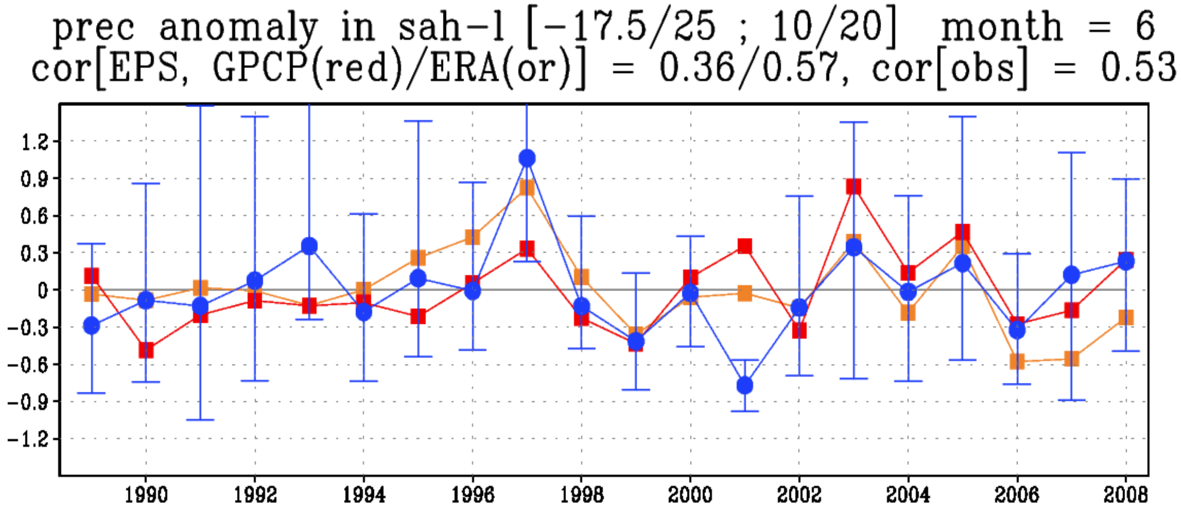




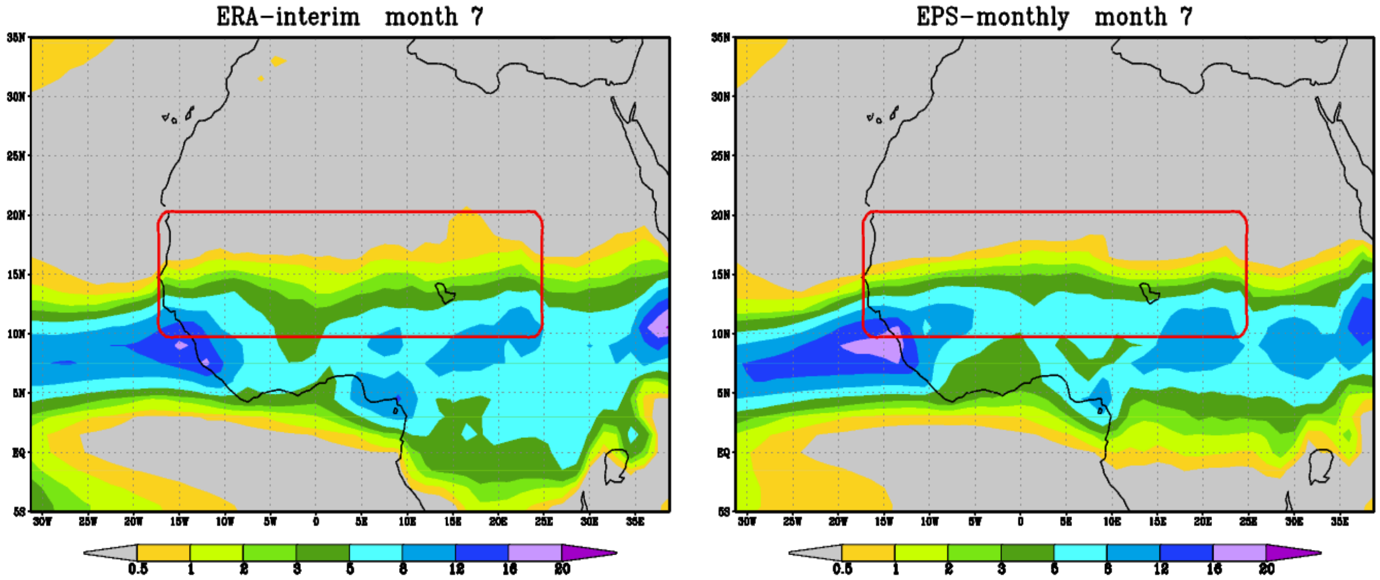
46-day EPS re-forecasts: Sahel, June

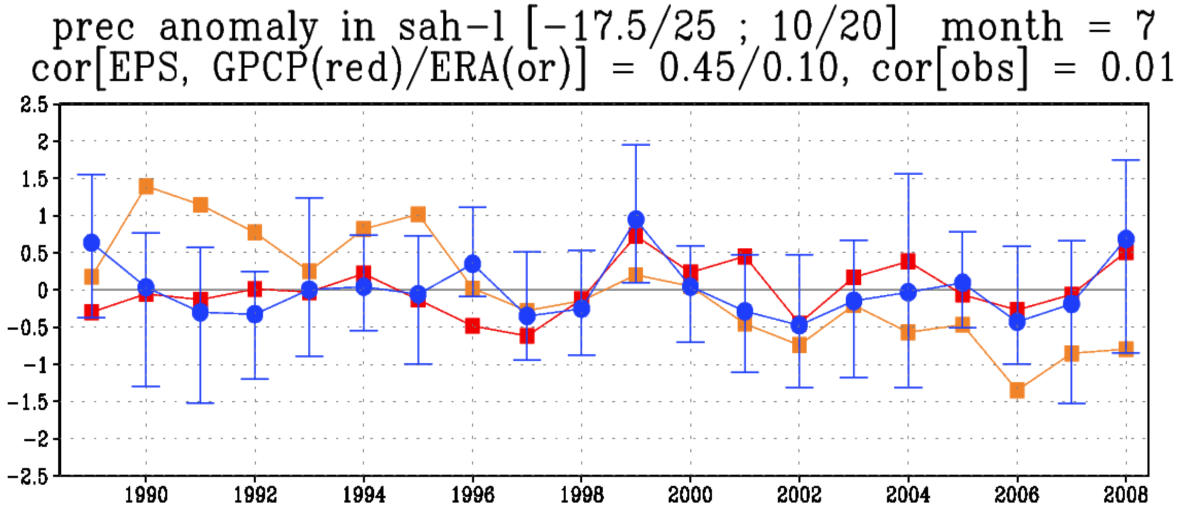






46-day EPS re-forecasts: Sahel, July





Outline

What is reanalysis ?	• Simple concepts • Goals of reanalysis
How do we do reanalysis?	• Observations • Model • Data assimilation
Reanalysis projects & results	• Projects • Users • Applications
Conclusions	• Summary • Challenges ahead

Summary of important concepts

- **Reanalysis does <u>not</u> produce “gridded observations”**
  - But it enables to extract information from observations in one, unique, theoretically consistent framework

- **Reanalysis sits at the end of the (long) meteorological research and development chain that encompasses**
  - observation and measurement collection,
  - observation processing and data exchange,
  - numerical weather prediction modelling and data assimilation

- **Unlike NWP, a very important concern in reanalysis is the <u>consistency in time, over several years</u>**

- **Reanalysis is bridging slowly, but surely, the gap between the “weather datasets” and the “climate datasets”**
  - Resolution gets finer
  - Reanalyses cover longer time periods, without gap
  - Helps different communities work together
  - Reanalysis has developed into a powerful tool for many users and applications

Current status of global reanalysis & Future outlook

- **It is worth repeating as all ingredients continue to evolve**:
  - Models are getting better
  - Data assimilation methods are getting better
  - Observation processing is improving
  - Old observations (paper records) are being rescued
  - The technical infrastructure for running & monitoring improves constantly
  - **With each new reanalysis we improve our understanding of systematic errors in the various components of the observing system**

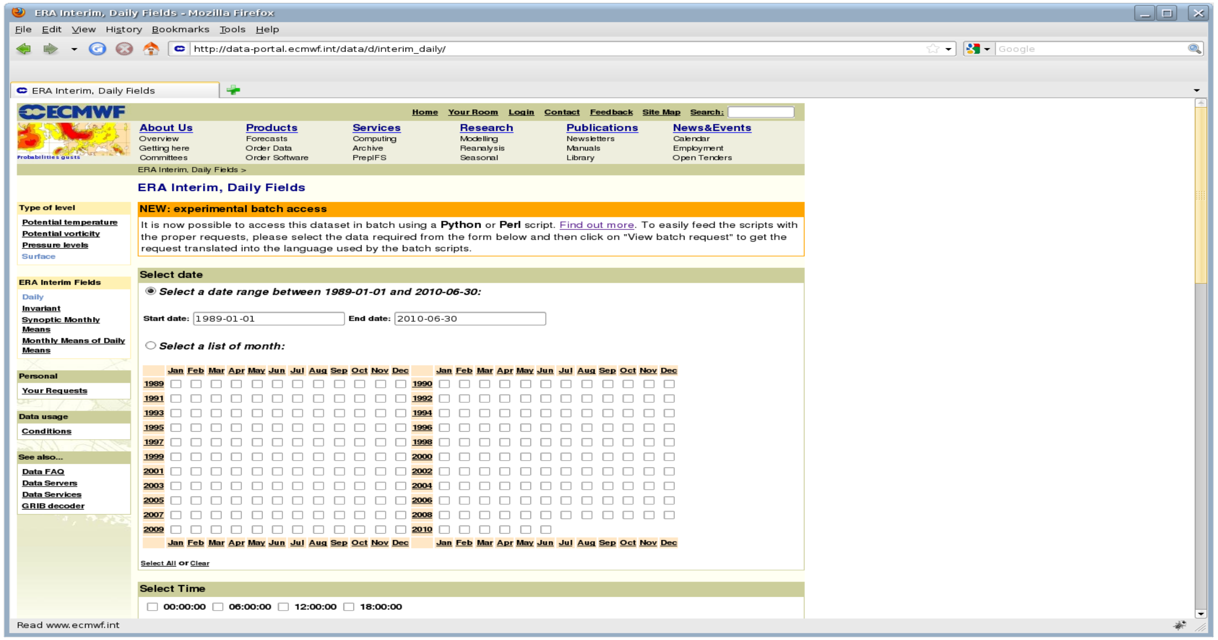
- **Major challenges for a future comprehensive reanalysis project**:
  - Bringing in additional observations (not dealt with in ERA-Interim)
  - *Dealing with model bias (ultimately responsible for problems with trends)*
  - Coupling with ocean and land surface
  - Making observations used in reanalysis more accessible to users
  - Providing meaningful uncertainty estimates for the reanalysis products

Further reading and on-line material

- Kalnay et al. (1996), "The NCEP/NCAR 40-Year Reanalysis Project", *Bull. Am. Meteorol. Soc.* **77** (3), 437-471
- Uppala et al. (2005), "The ERA-40 reanalysis", *Q. J. R. Meteorol. Soc.* **131** (612), 2961-3012, doi: 10.1256/qj.04.176
- Bengtsson et al. (2007), "The need for a dynamical climate reanalysis", *Bull. Am. Meteor. Soc.* **88** (4), 495-501
- SciDAC Review (2008), "Bridging the gap between weather and climate", on the web at http://www.scidacreview.org/0801/pdf/climate.pdf with contributions from G. P. Compo and J. S. Whitaker
- European reanalysis (ERA): http://www.ecmwf.int/research/era
- NCEP/NCAR reanalysis: http://www.cdc.noaa.gov/data/reanalysis/reanalysis.shtml
- NCEP CFSR: http://cfs.ncep.noaa.gov/cfsr/
- Japanese 25-year reanalysis (JRA-25): http://jra.kishou.go.jp
- NASA GMAO Modern Era Retrospective-analysis for Research and Applications (MERRA) http://gmao.gsfc.nasa.gov/research/merra/
- Dee et al. (2011), "The ERA-Interim reanalysis: configuration and performance of the data assimilation system ", *Q. J. R. Meteorol. Soc.,* **137** (656), 553-597
- http://www.oldweather.org
- http://www.data-rescue-at-home.org

ERA-Interim data availability and access

- Jan 1979 until Feb 2011, with monthly updates
- Resolution: T255L60, 6-hourly upper-air fields (3-hourly for surface)
- Analysis + forecast products; monthly averages
- Access to products:
  - Member state users: MARS: full access
  - All users: ECMWF Public Data Server: http://data-portal.ecmwf.int/data/d/interim_daily/
    Currently ***reduced resolution*** 1.5 deg x 1.5 deg and 37 pressure levels
    Will contain full resolution data… any time soon
    Years 1979-1988 available around mid-summer 2011



ERA-CLIM: data recovery efforts