Data Assimilation

Lars Isaksen

ECMWF

(European Centre for Medium-Range Weather Forecasts)

Special acknowledgements to Jean-Noël Thépaut and Massimo Bonavita

Further acknowledgements to Carla Cardinali, Patricia De Rosnay, Mike Fisher, Tim Hewson, Marta Janisková, Linus Magnusson, Tony McNally, Deborah Salmond and Mike Rennie

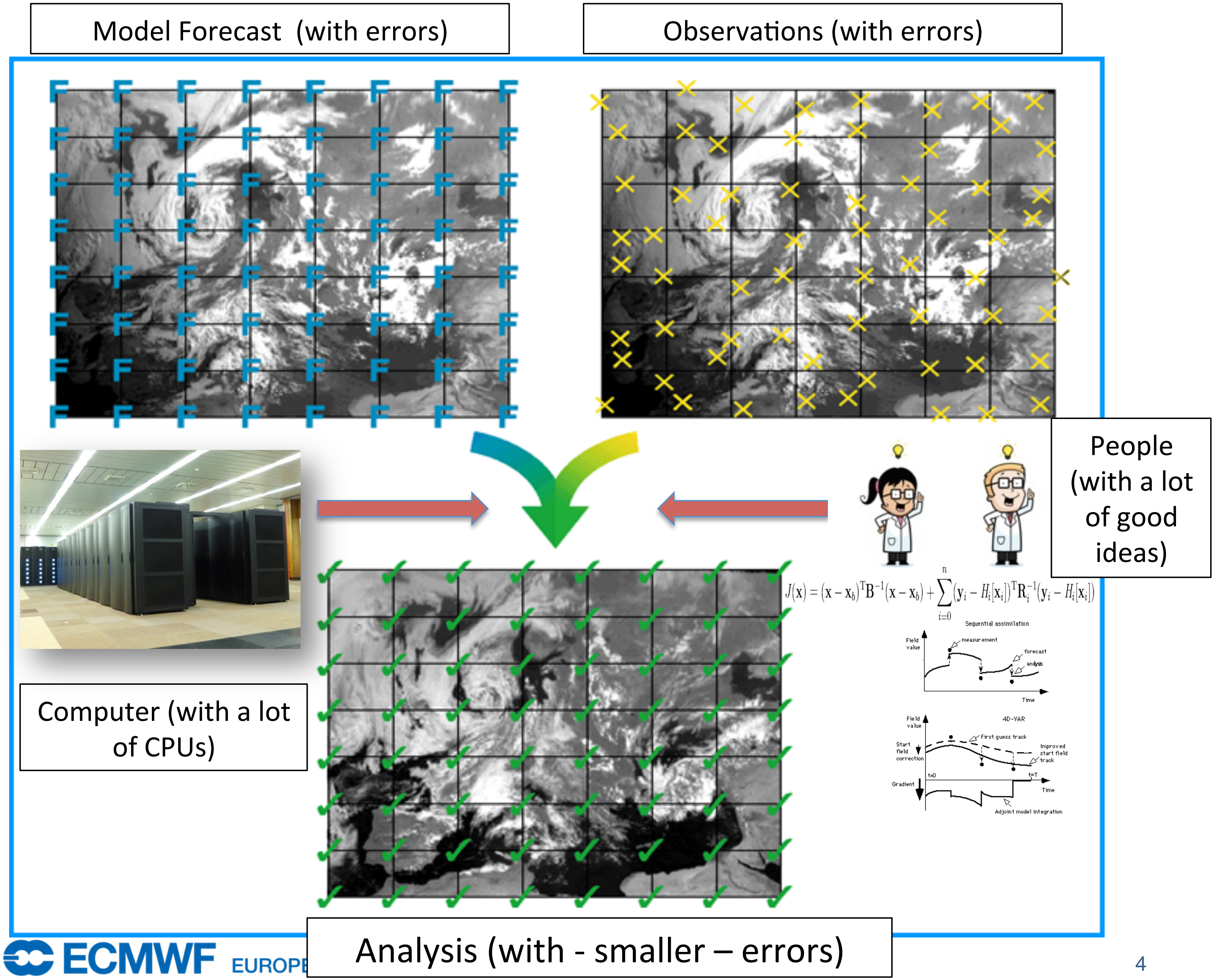
Overview of talk

- What is data assimilation?
- Observations used by the data assimilation system at ECMWF
- How to use satellite observations in data assimilation
- Observation errors and quality control
- Four dimensional variational data assimilation (4D-Var)
- Case studies
- Recent improvements of the data assimilation system
- Future challenges

Data Assimilation

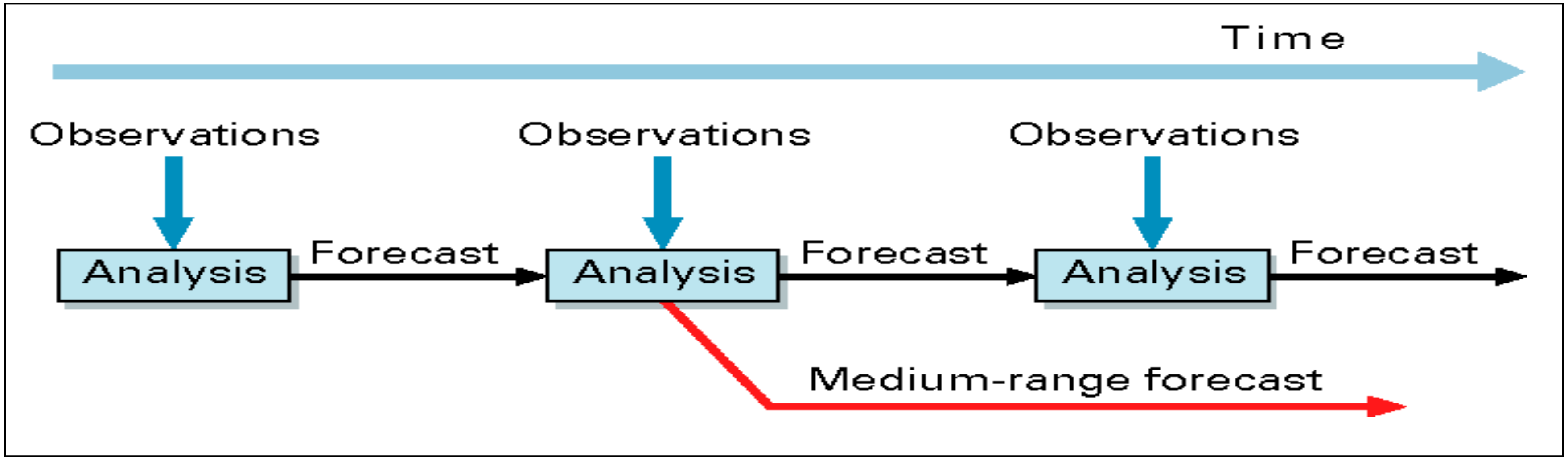
Process by which "optimal" initial conditions for numerical weather forecasts (NWP) are defined.

- The best NWP analysis is the one that leads to the best forecast
- Make the best estimate out of all information available
- Do it quickly – typically in less than 45 minutes on a large high performance computer



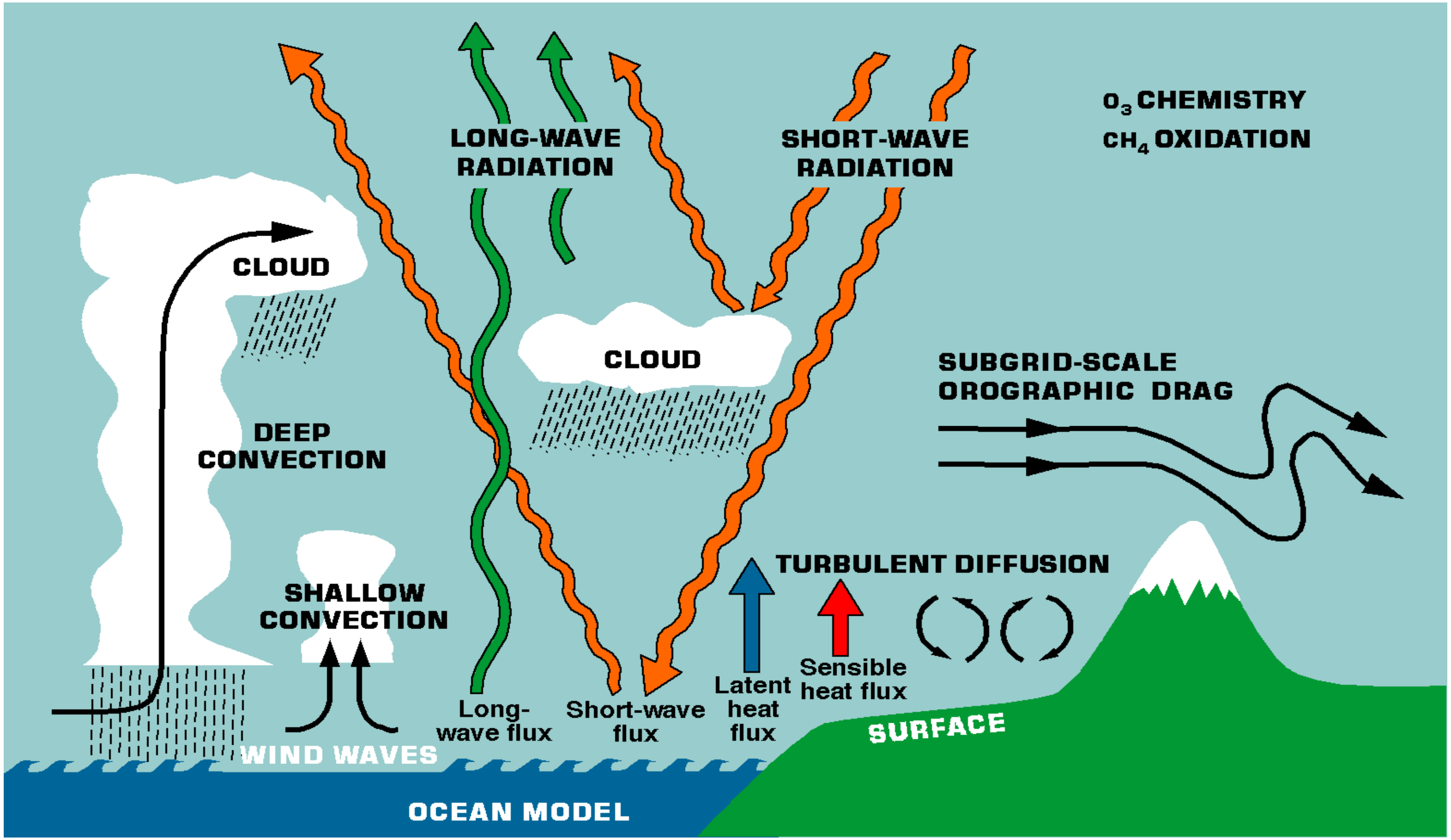
ECMWF EUROPE

Data assimilation



- The observations are used to correct errors in the short forecast from the previous analysis time.
- At ECMWF, twice a day about 20,000,000 observations are used to correct the 80,000,000 variables that define the model's virtual atmosphere.
- This is at ECMWF done by a 4-dimensional adjustment in space and time based on the available observations; this operation takes as much computer power as the 10-day forecast.

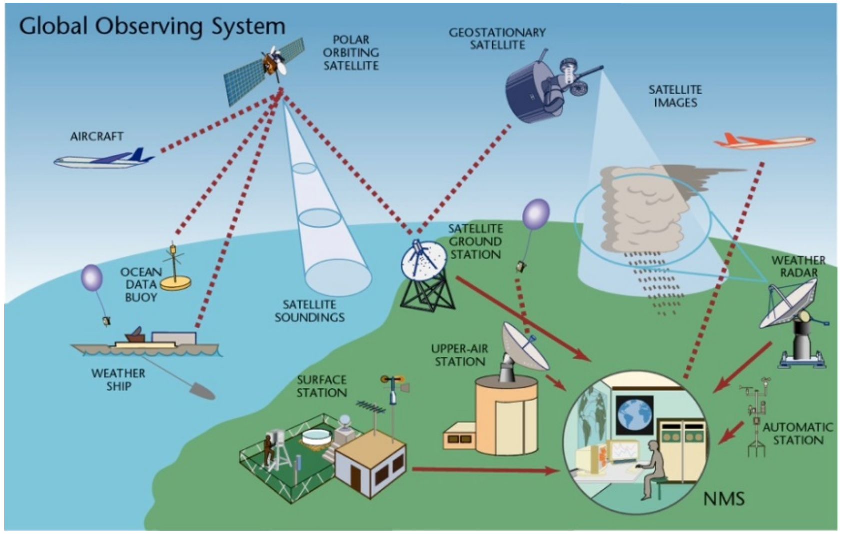
The forecast model is a very important part of the data assimilation system



Most important physical processes in the ECMWF model (see next lecture)

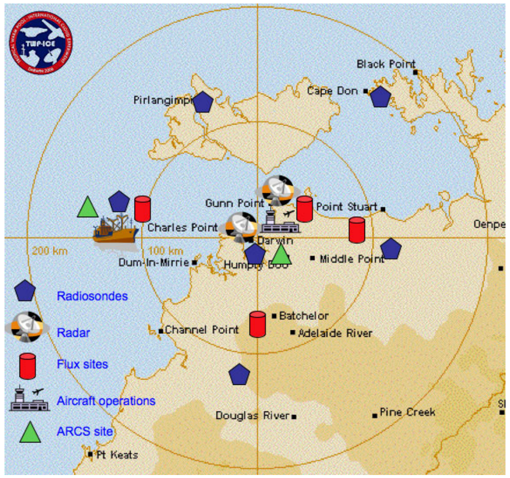
Observations used by the data assimilation system at ECMWF

WMO Integrated Global Observing System



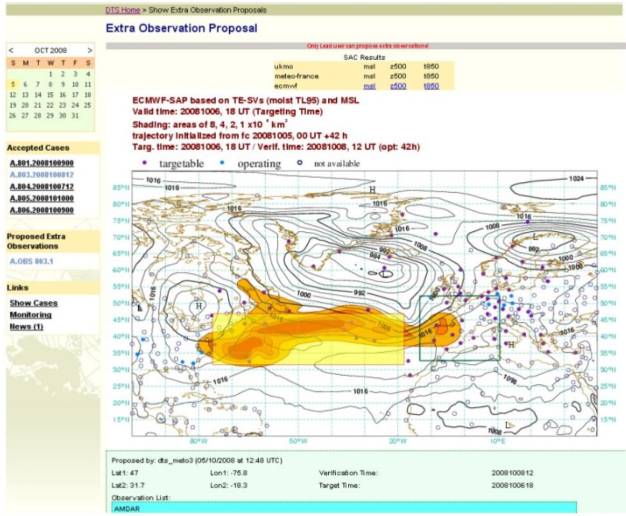
Courtesy: WMO

T
W
P
-
I
C
E

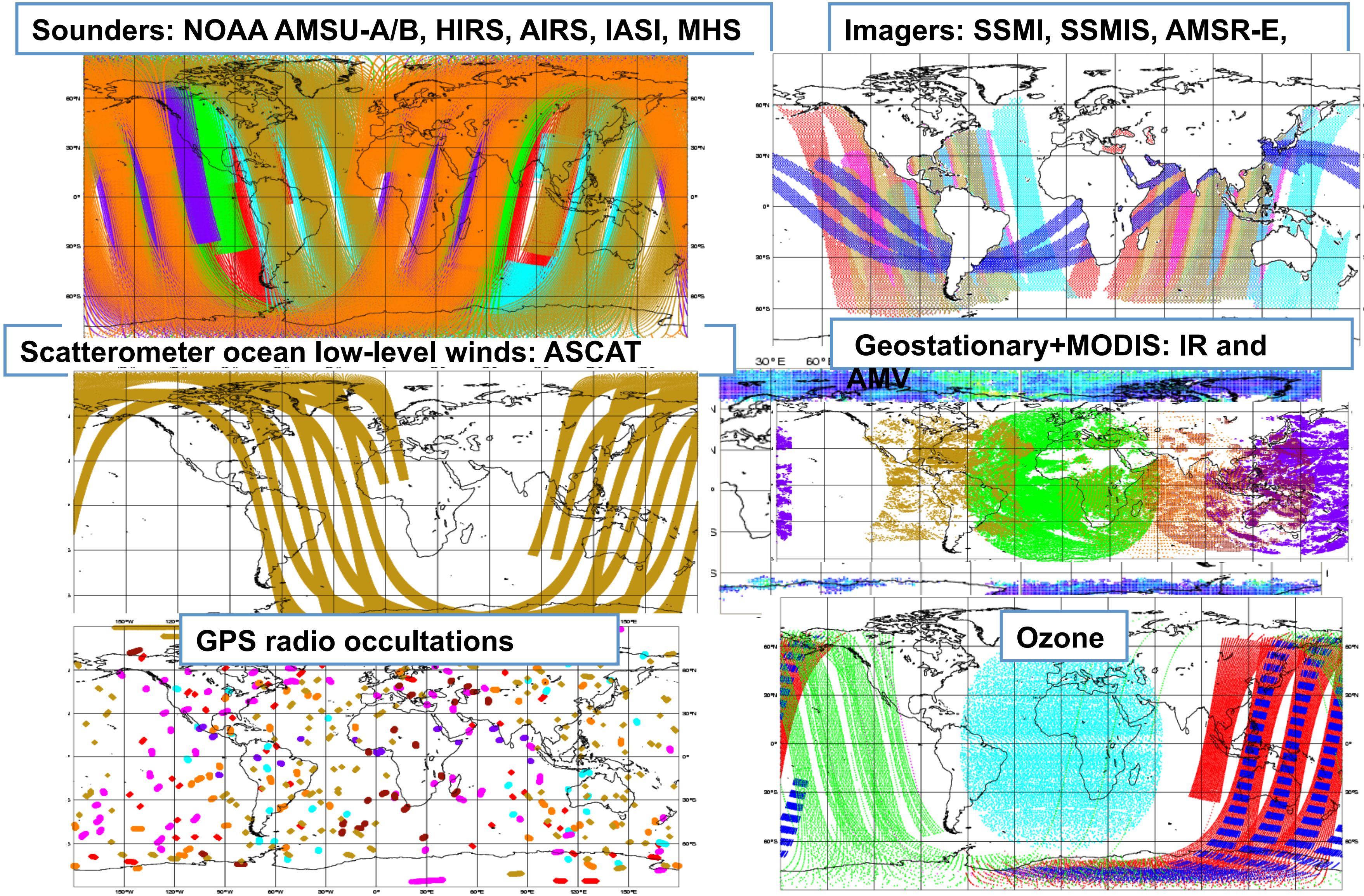


Supported by field campaign experiments,
Data targeting studies,
etc.

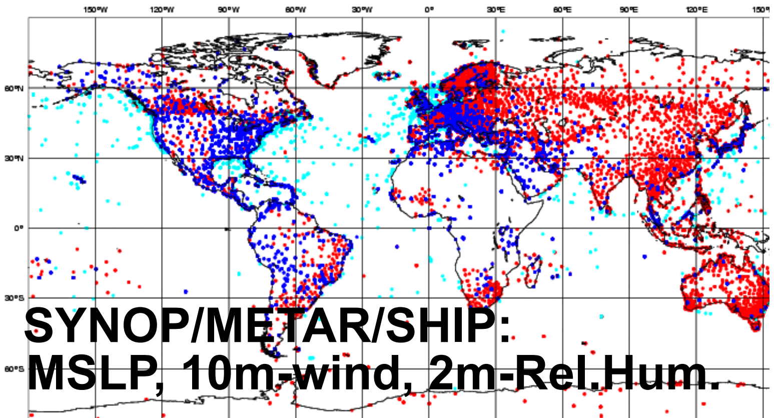
ECMWF Preview – DTS

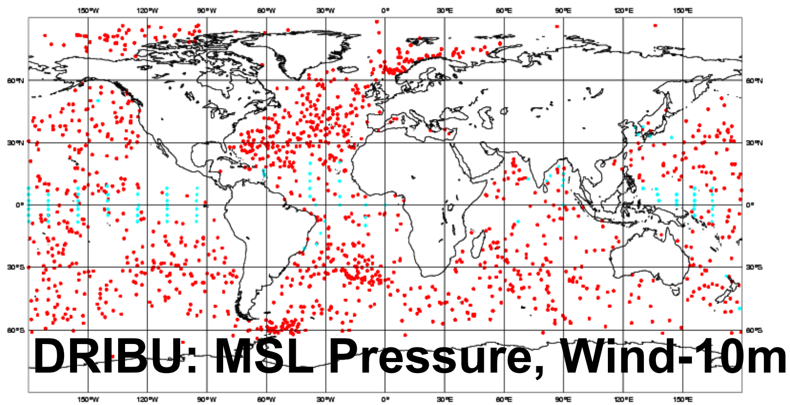


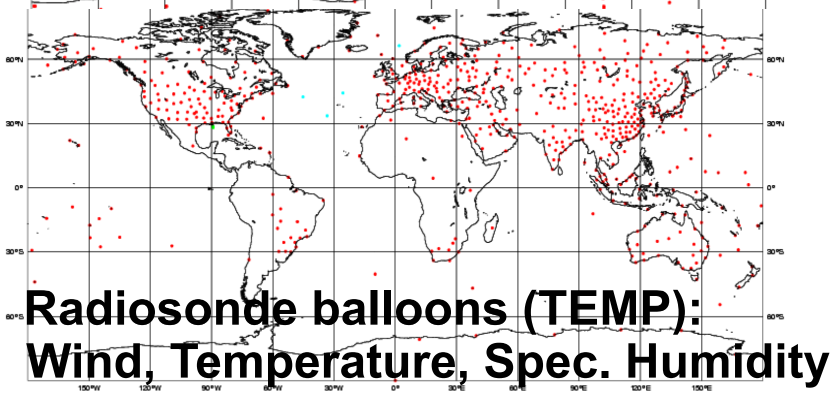
Satellite data sources used by ECMWF's analysis

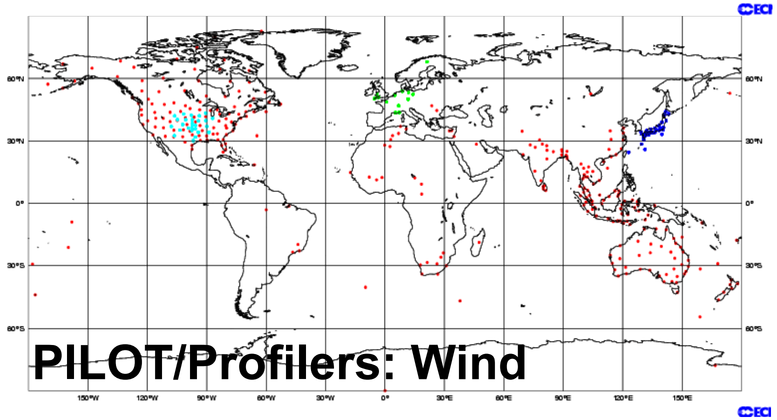


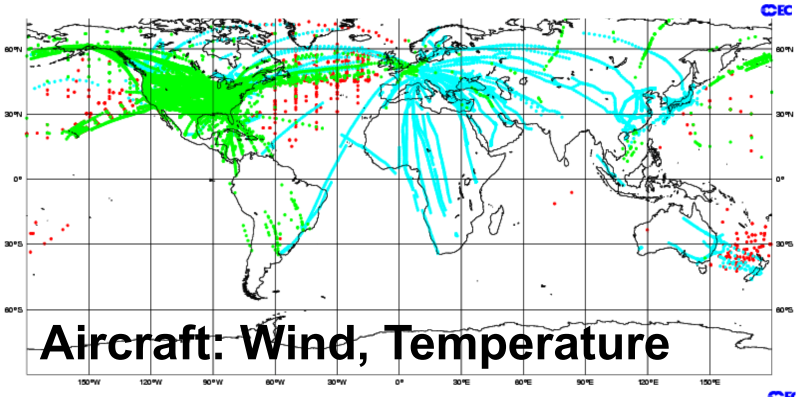
In situ observations used by ECMWF's analysis





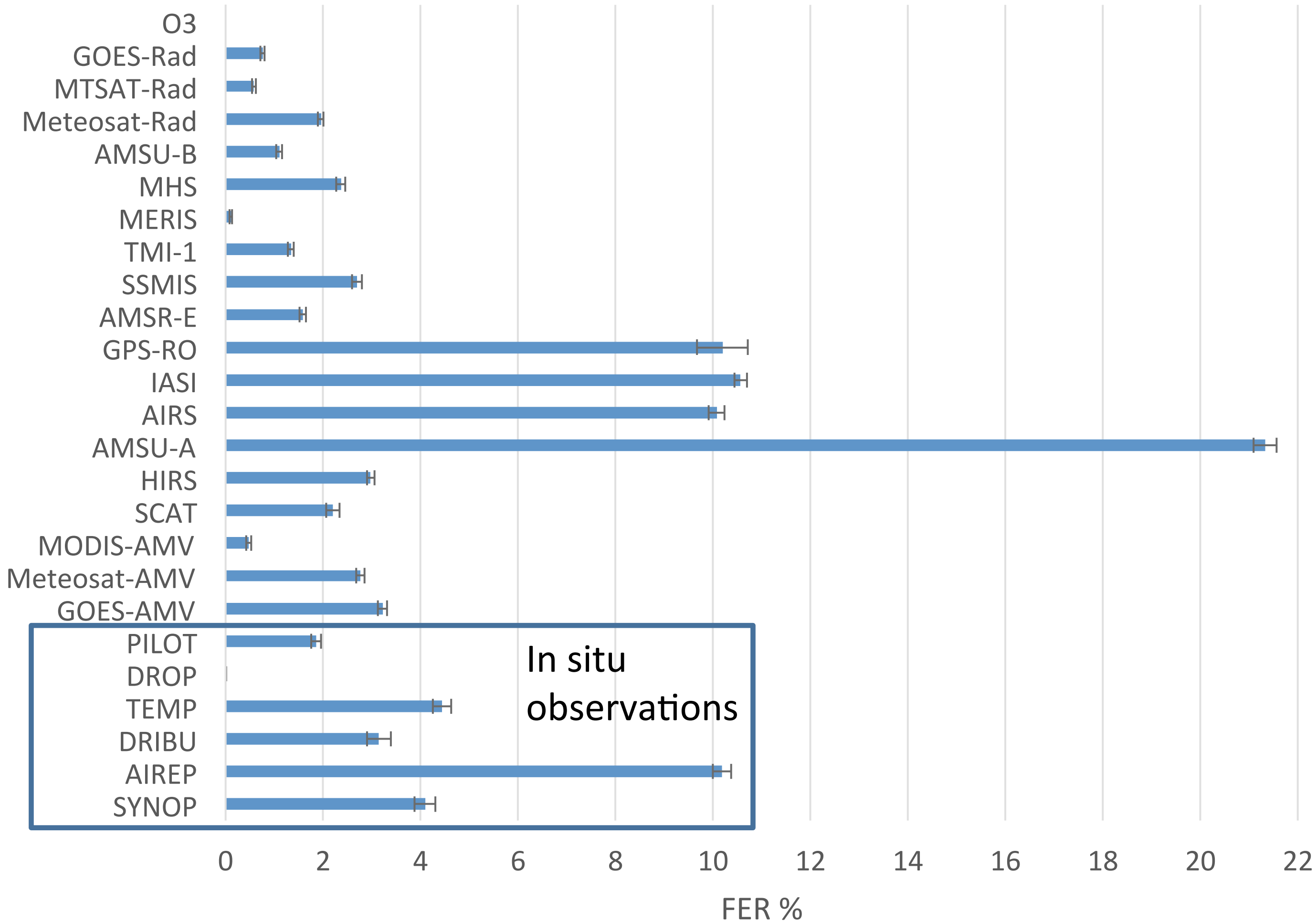






Note: We only use a limited number of the observed variables; especially over land.

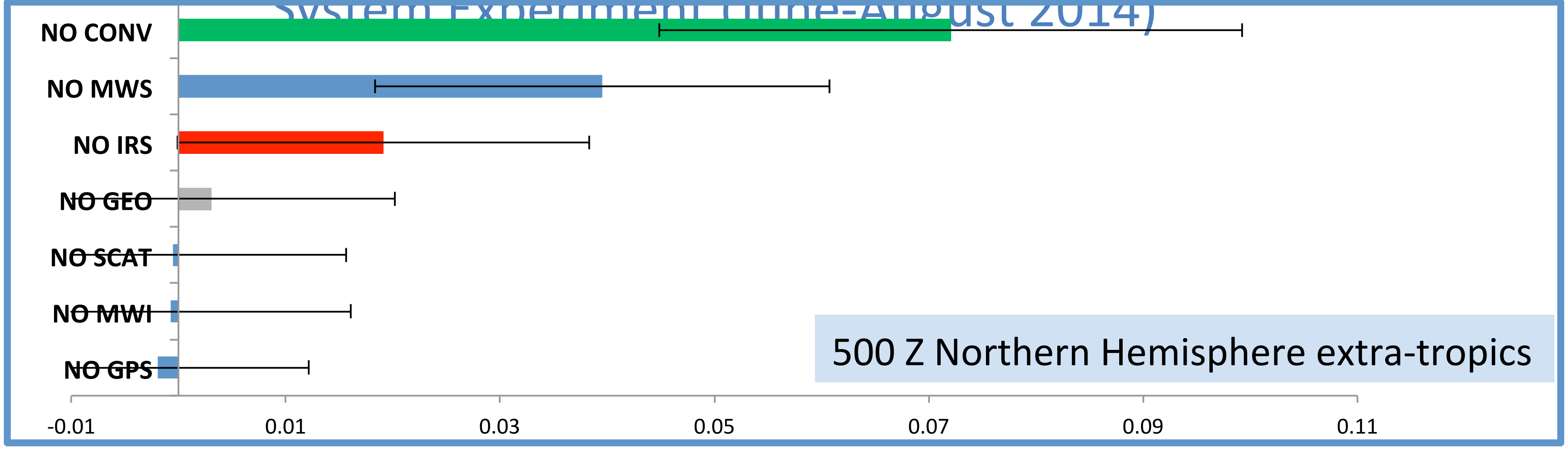
Which observations are most important for NWP?

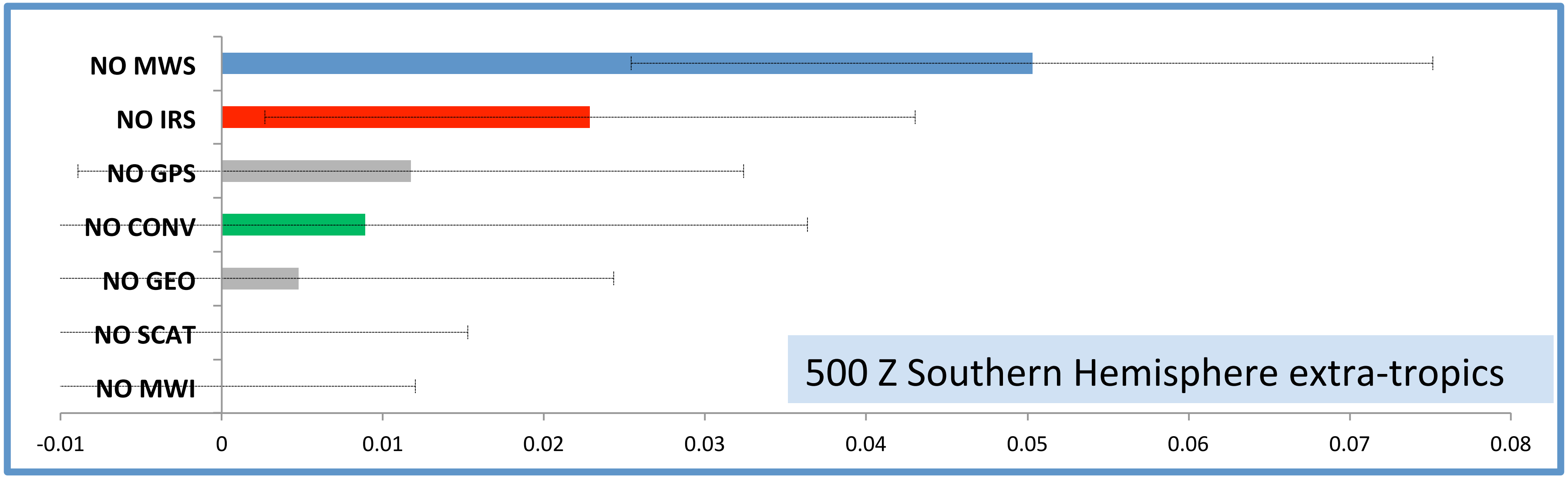


24h Forecast Error Reduction estimated using adjoint based diagnostics at ECMWF

Ranking observation Impact (on day 6 forecast)

Fractional increase in RMSE compared to control Observing System Experiment (June-August 2014)



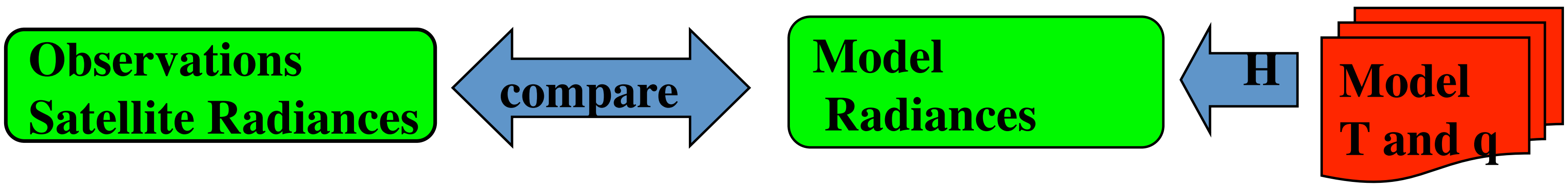


Tony McNally, ECMWF

How to use satellite observations in data assimilation

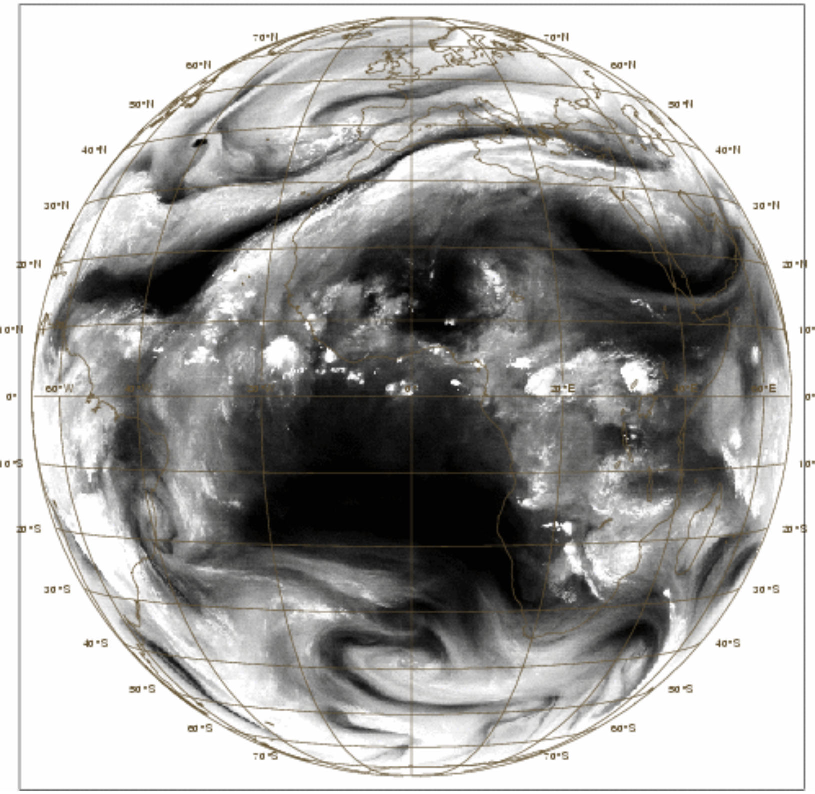
How we use satellite data in the analysis

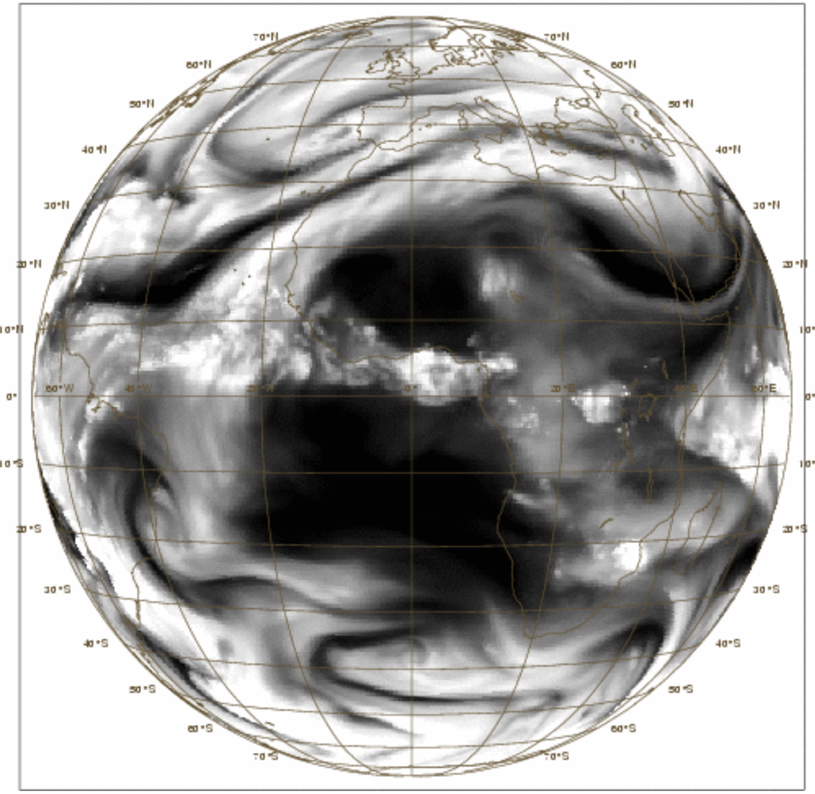
- Observations are not made at model grid points.
- Satellites measure radiances, NOT temperature and humidity .
- We calculate a model radiance estimate of the radiance measurement, using the so-called 'observation operator' (H).
- H performs a complex transformations of model variables (T,q) to radiances using a radiative transfer model .
- The model estimate is compared with the observed radiance.



Accurate radiative transfer models allows comparison of model and observed radiances

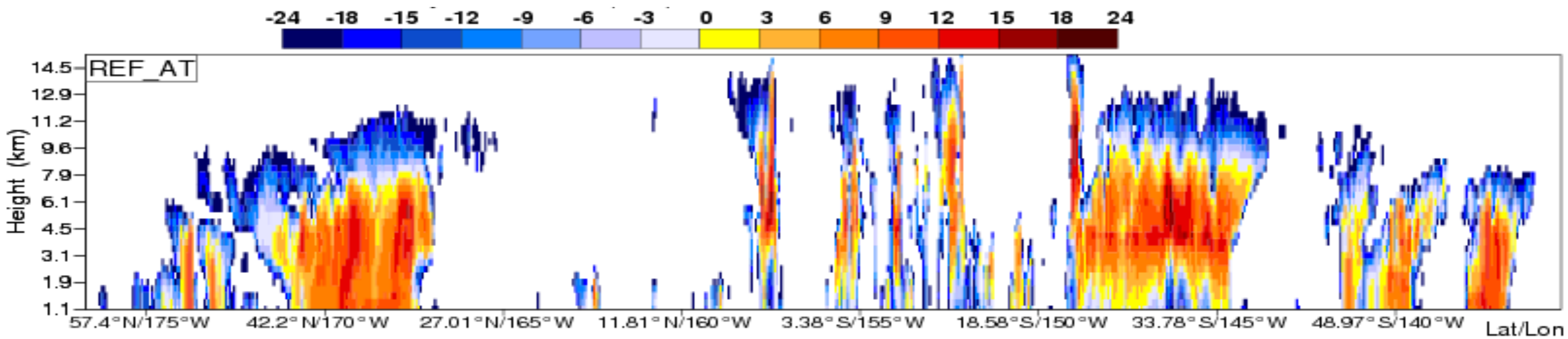
Observations Satellite Radiances ⟷ **compare** ⟷ **Model Radiances** ← **H** ← **Model T and q**



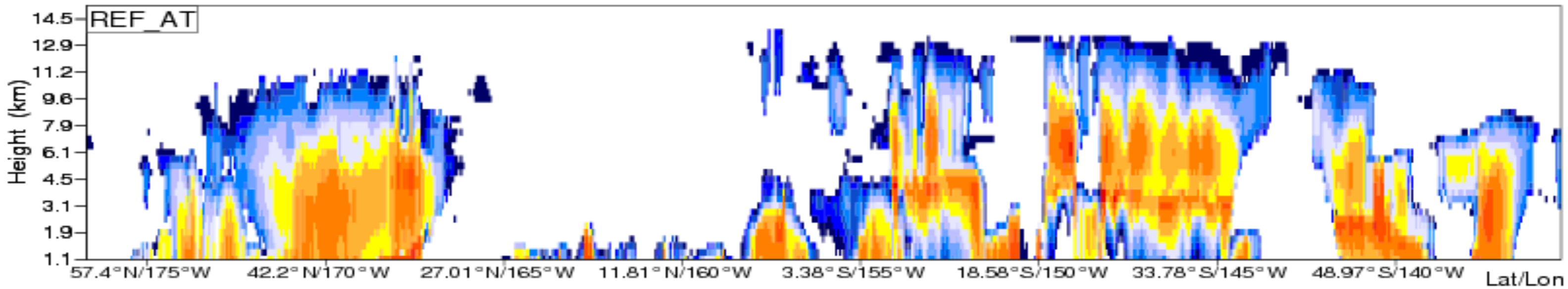


Models and observation operators have become much more realistic and accurate

CLOUDSAT is NASA satellite with a cloud profiling radar on-board



Cloud Radar Reflectivity - CLOUDSAT observations



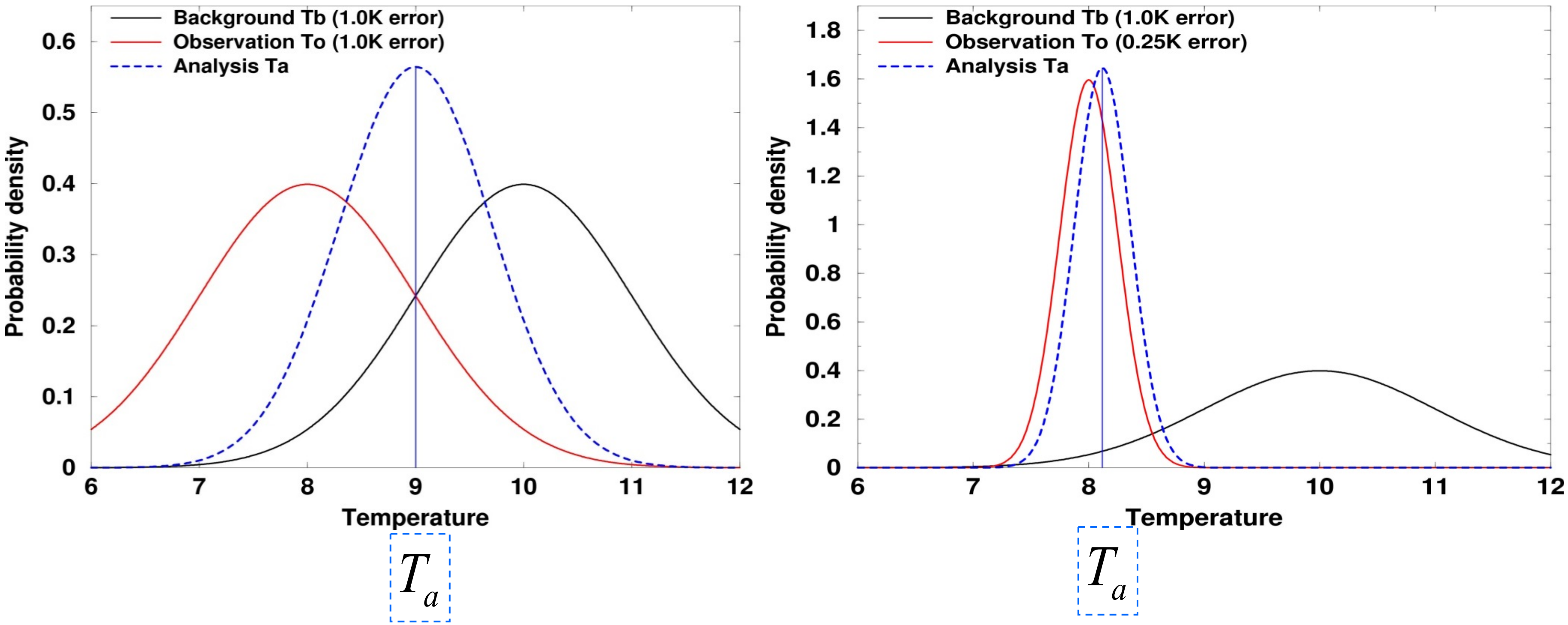
Cloud Radar Reflectivity - model equivalents

Observations and model background have errors
It is important to specify them accurately

Observed temperature: 8˚C

Short-range model background temperature: 10˚ C

Analysis: x˚C



Quality control of observations is very important

Data extraction

- Check out duplicate reports
- Ship tracks check
- Hydrostatic check

Thinning

- Some data is not used to avoid over-sampling and correlated errors
- Departures and flags are still calculated for further assessment

Blacklisting

- Data skipped due to systematic bad performance or due to different considerations (e.g. data being assessed in passive mode)
- Departures and flags available for all data for further assessment

Model/4D-Var dependent QC

- First guess based rejections
- VarQC rejections

Used data → Increments

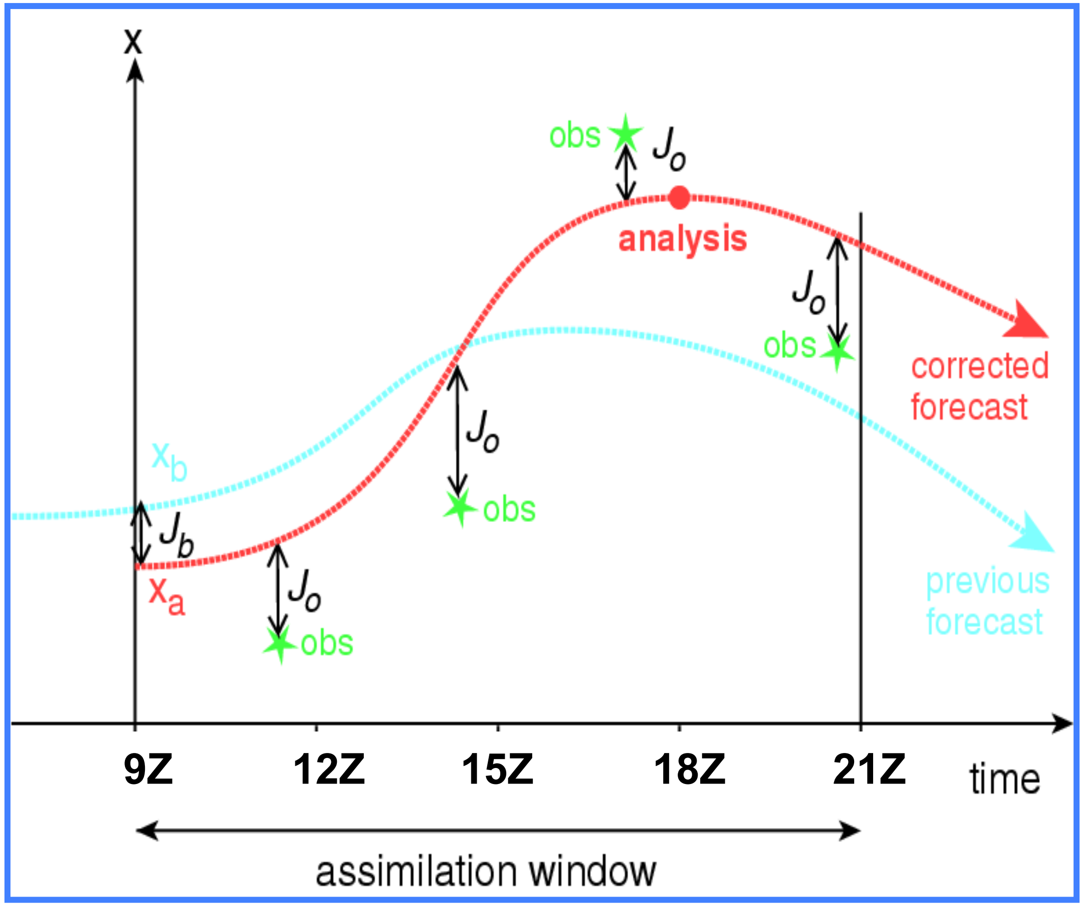
Analysis

Four dimensional variational data assimilation (4D-Var)

ECMWF use a 4D Variational (4D-Var) Data Assimilation method

Around 20,000,000 observations within a 12-hour period are used simultaneously in one global (iterative) estimation problem

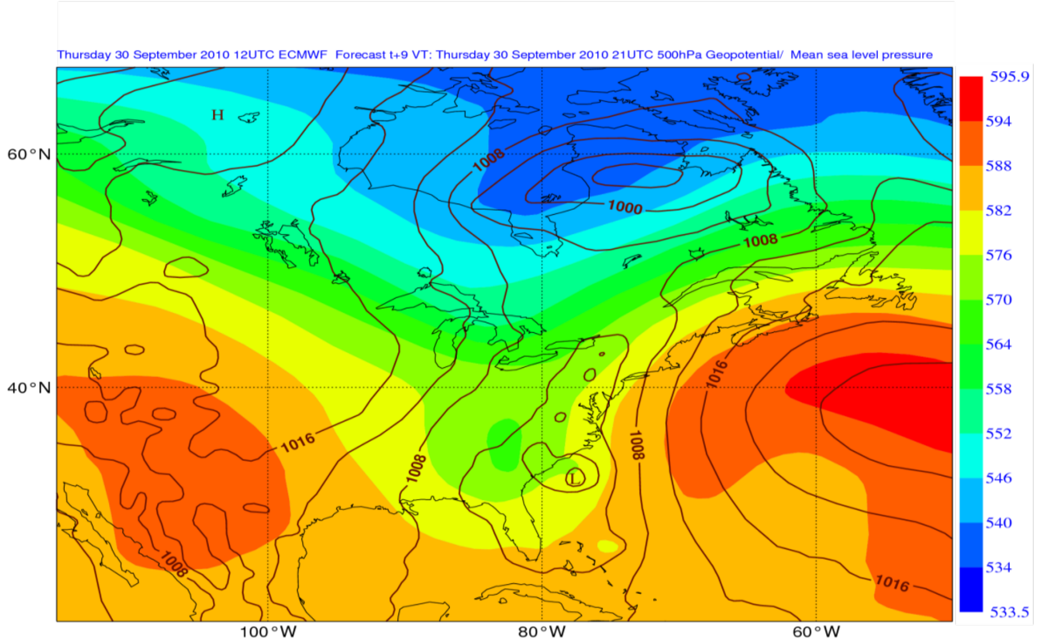
- "Observation – model" values are computed at the observation time at high resolution: 16 km
- 4D-Var finds the 12-hour forecast that take account of the observations in a dynamically consistent way
- Based on a tangent linear and adjoint forecast models, used in the minimization process at lower resolution
- 80,000,000 model variables (surface pressure, temperature, wind, specific humidity and ozone) are adjusted



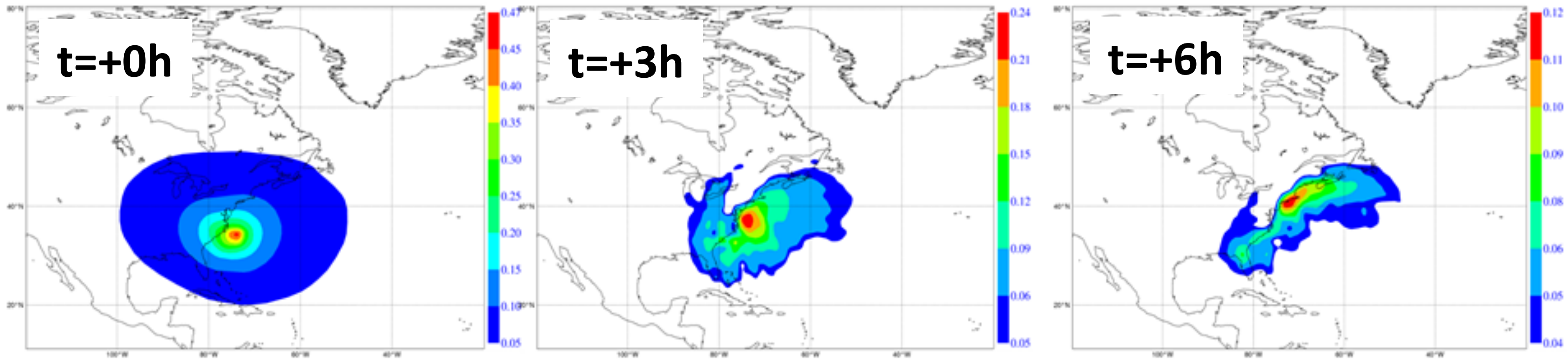
Incremental 4DVar

MSLP (contours) and 500 hPa geopotential height (shaded) background fields

Analysis change by adding an observation at the start of the assimilation window

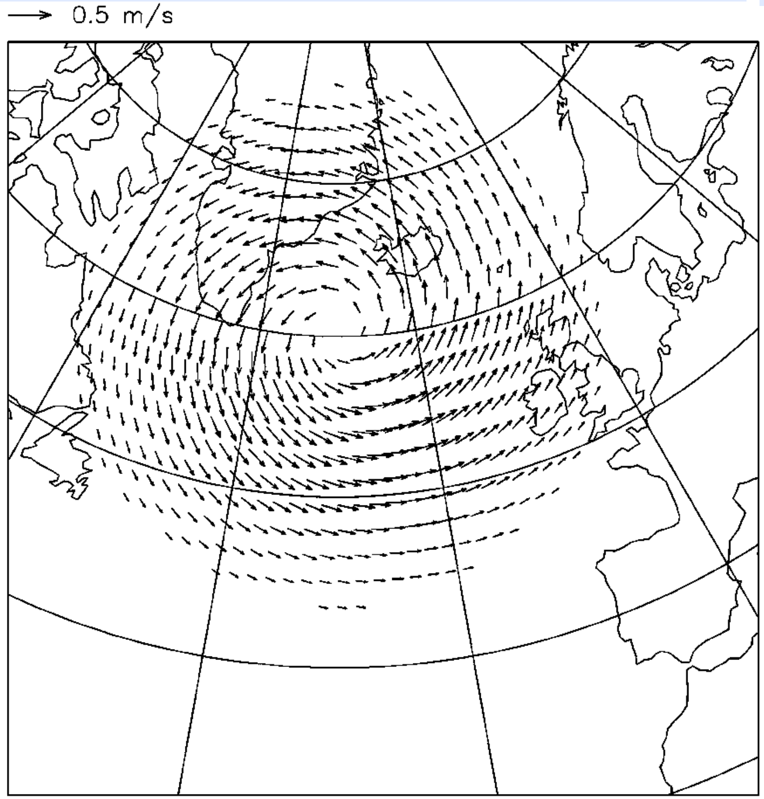


Temperature analysis increments for a single temperature observation at the start of the assimilation window: $x^a(t)-x^b(t) \approx$ $\mathbf{MBM^TH^T}(y-Hx)/(\sigma_b^2+\sigma_o^2)$

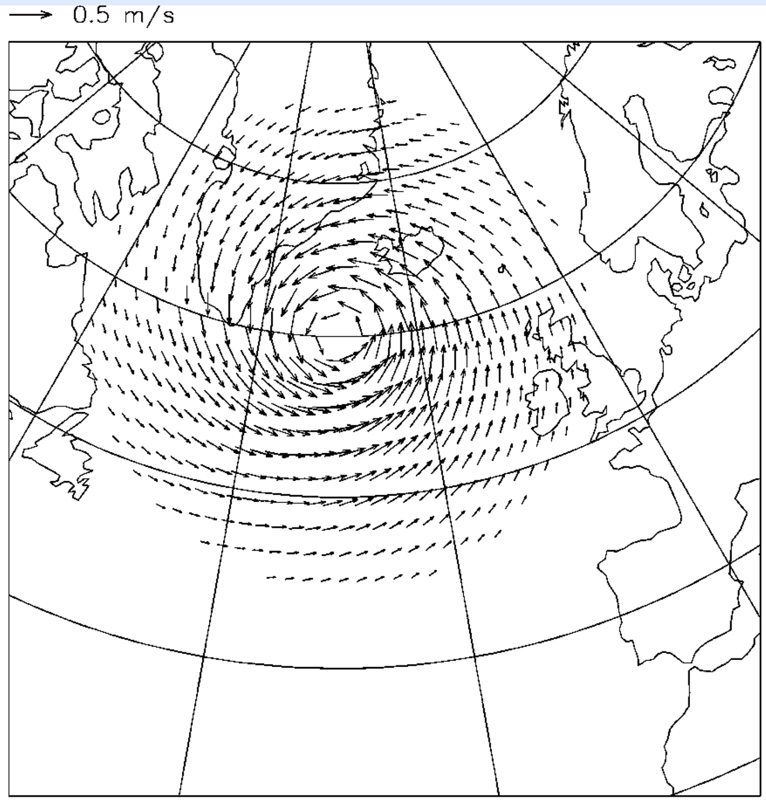


The Balance Operator ensures height/wind field approx. balance is retained in the extra-tropics

wind increments at 300hPa

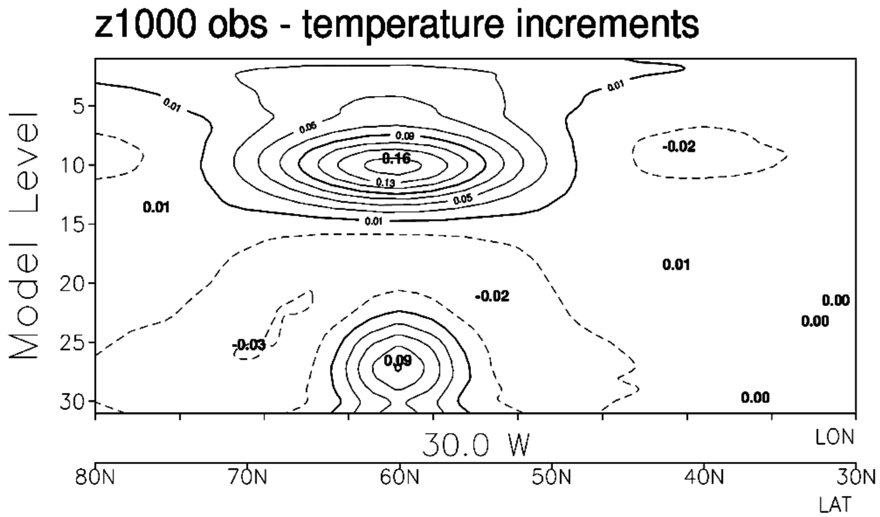


wind increments 150 metre above surface

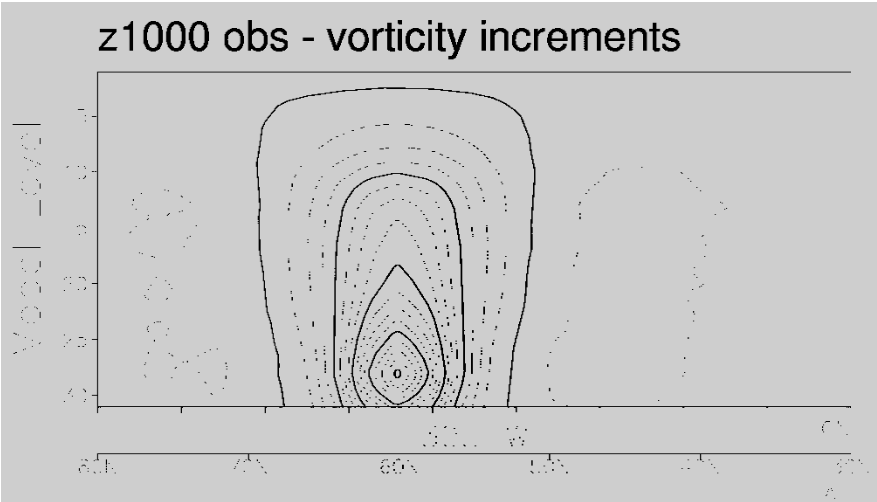


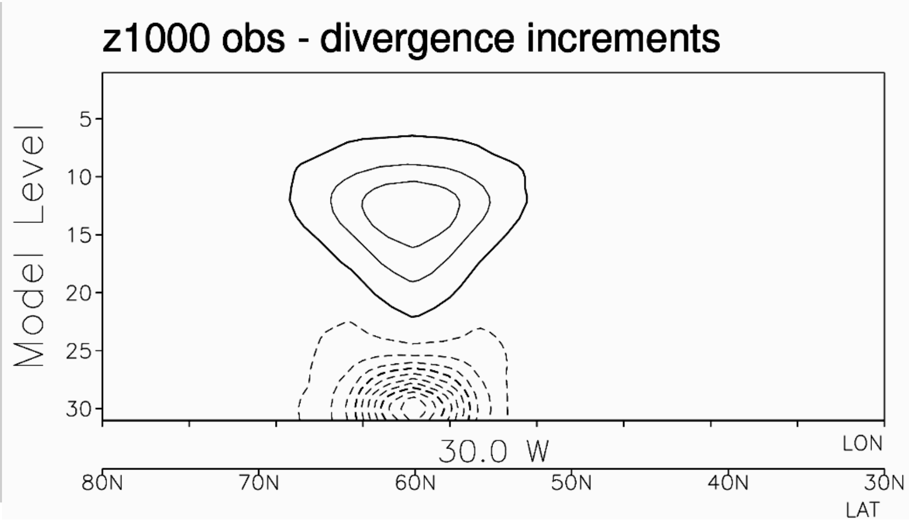
Wind increments obtained from a single surface pressure observation

J_b: Ensures that the model fields are adjusted meteorologically consistently in the region close to the observation location

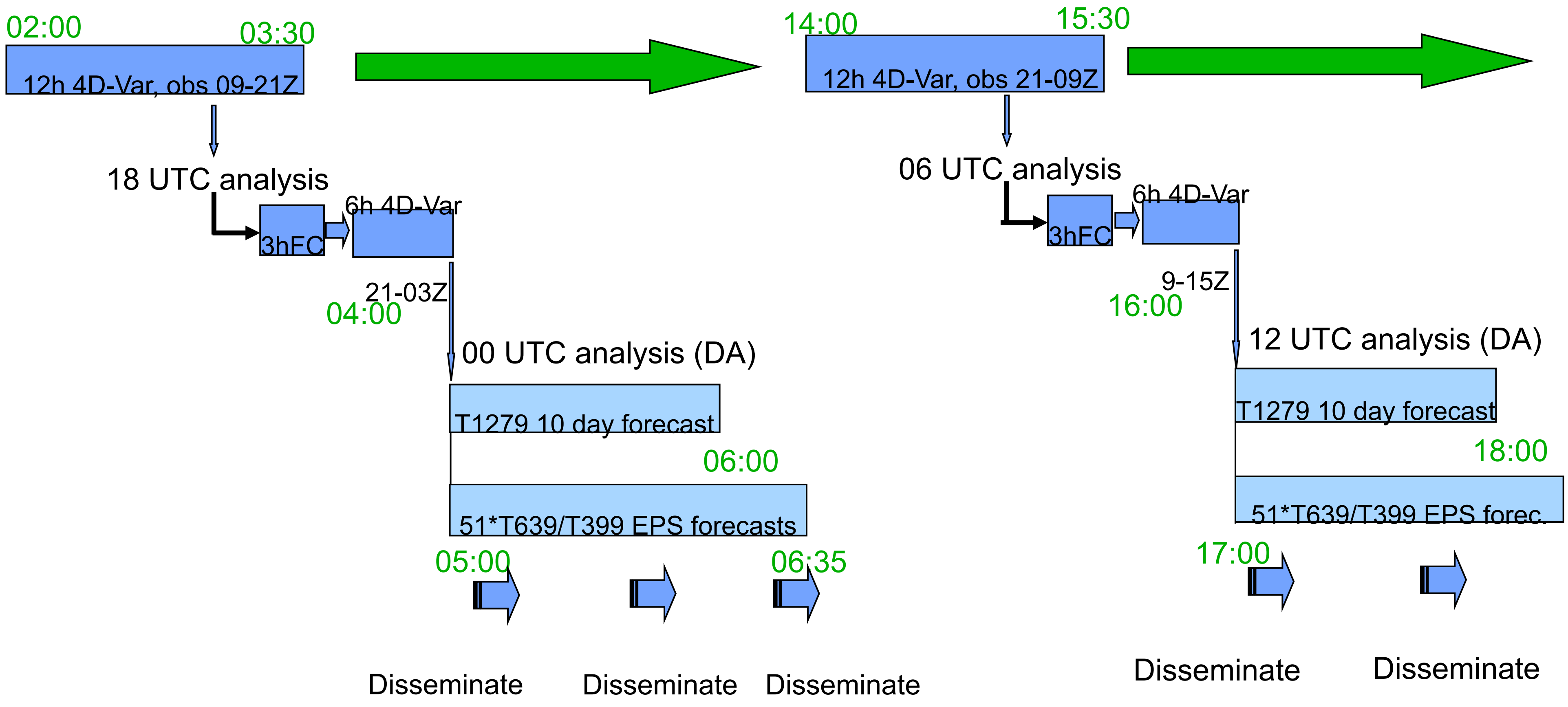


Increments due to a single observation of geopotential height at 1000hPa at 60N with value 10m below the background.





Operational schedule
Delayed Cut Off and Early Delivery suites

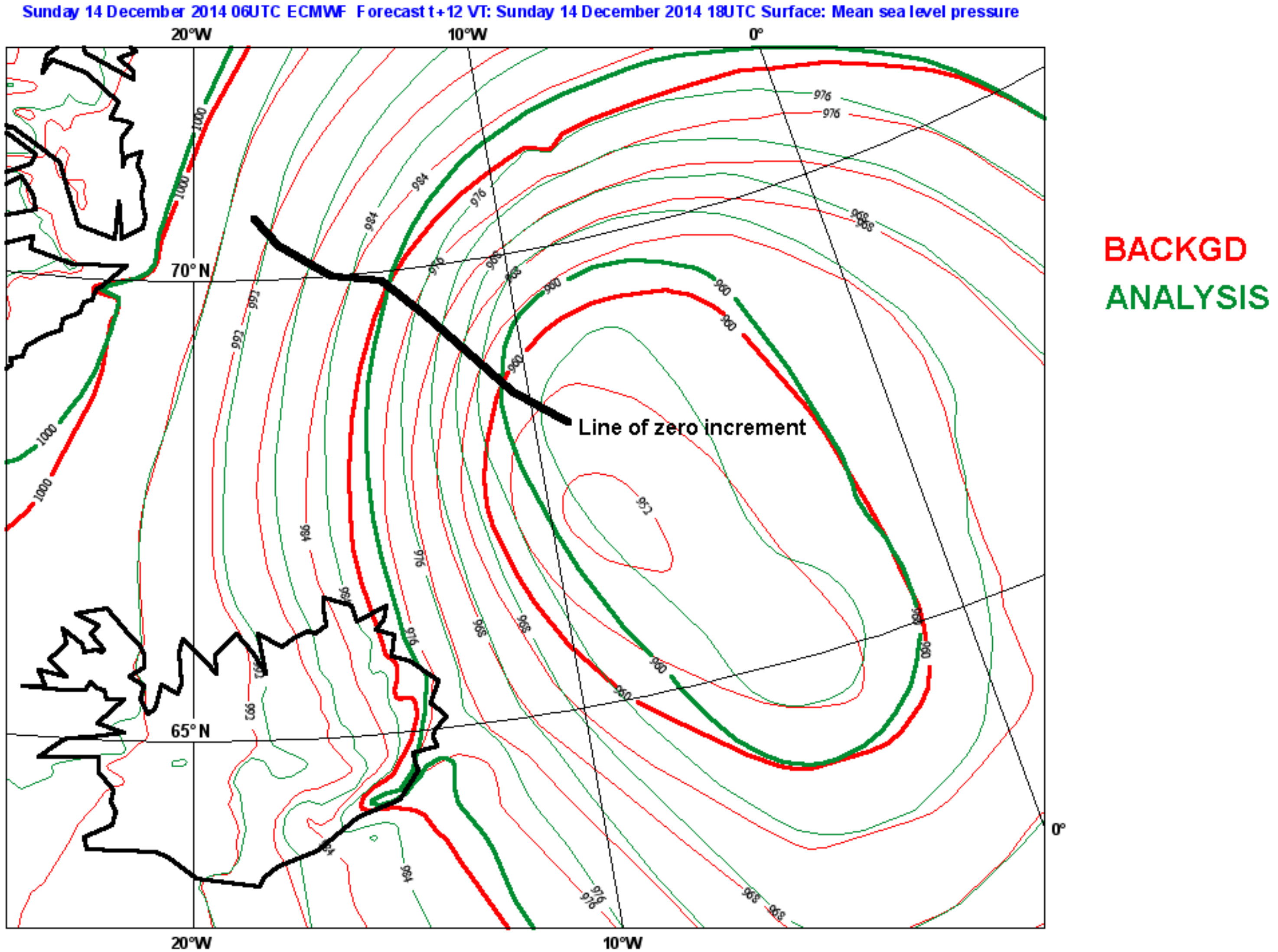


Case studies

HRES forecast skill –
a comment linked to jumpiness

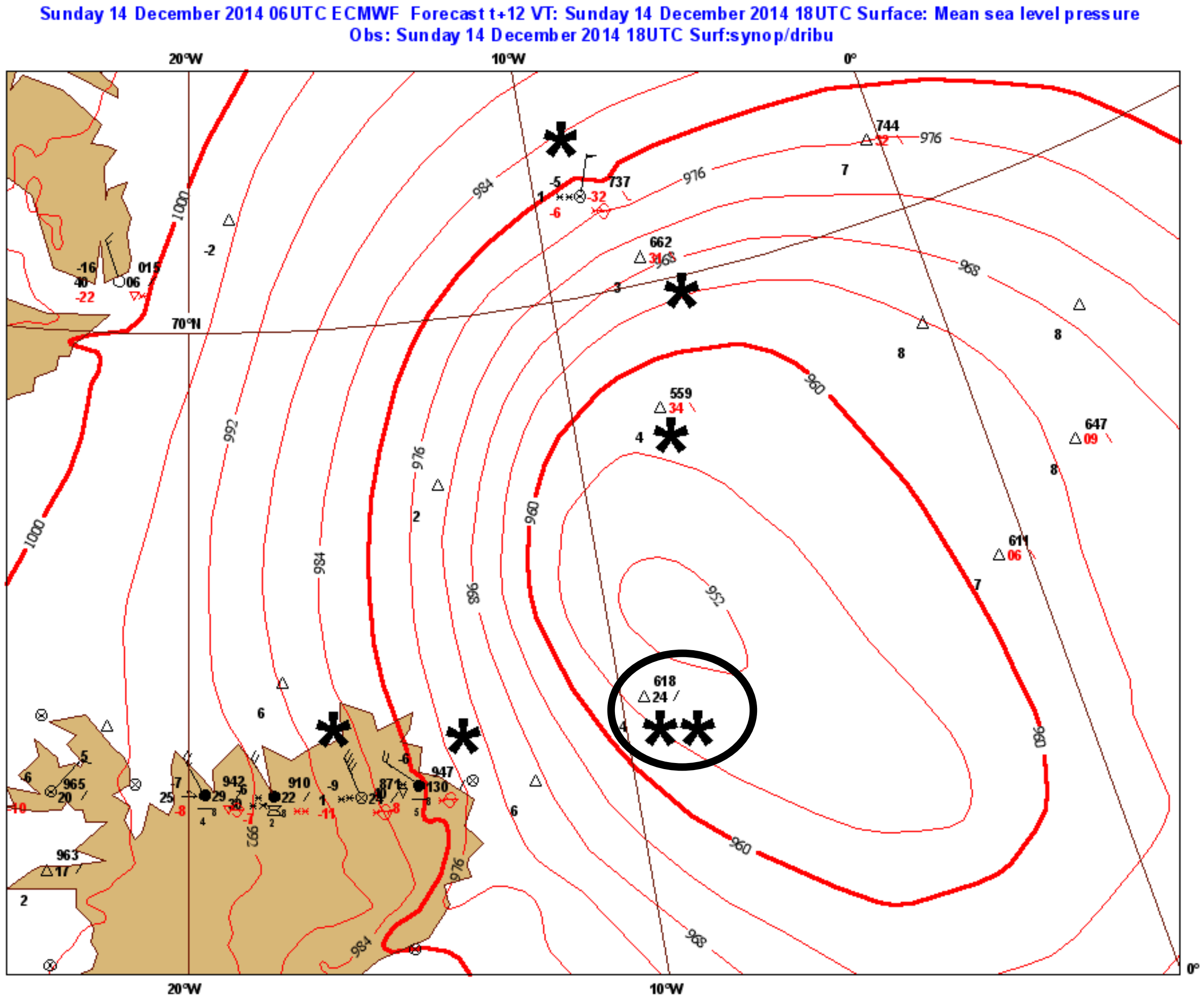
Case Study: Extra-tropical low near Iceland

14 December 2014



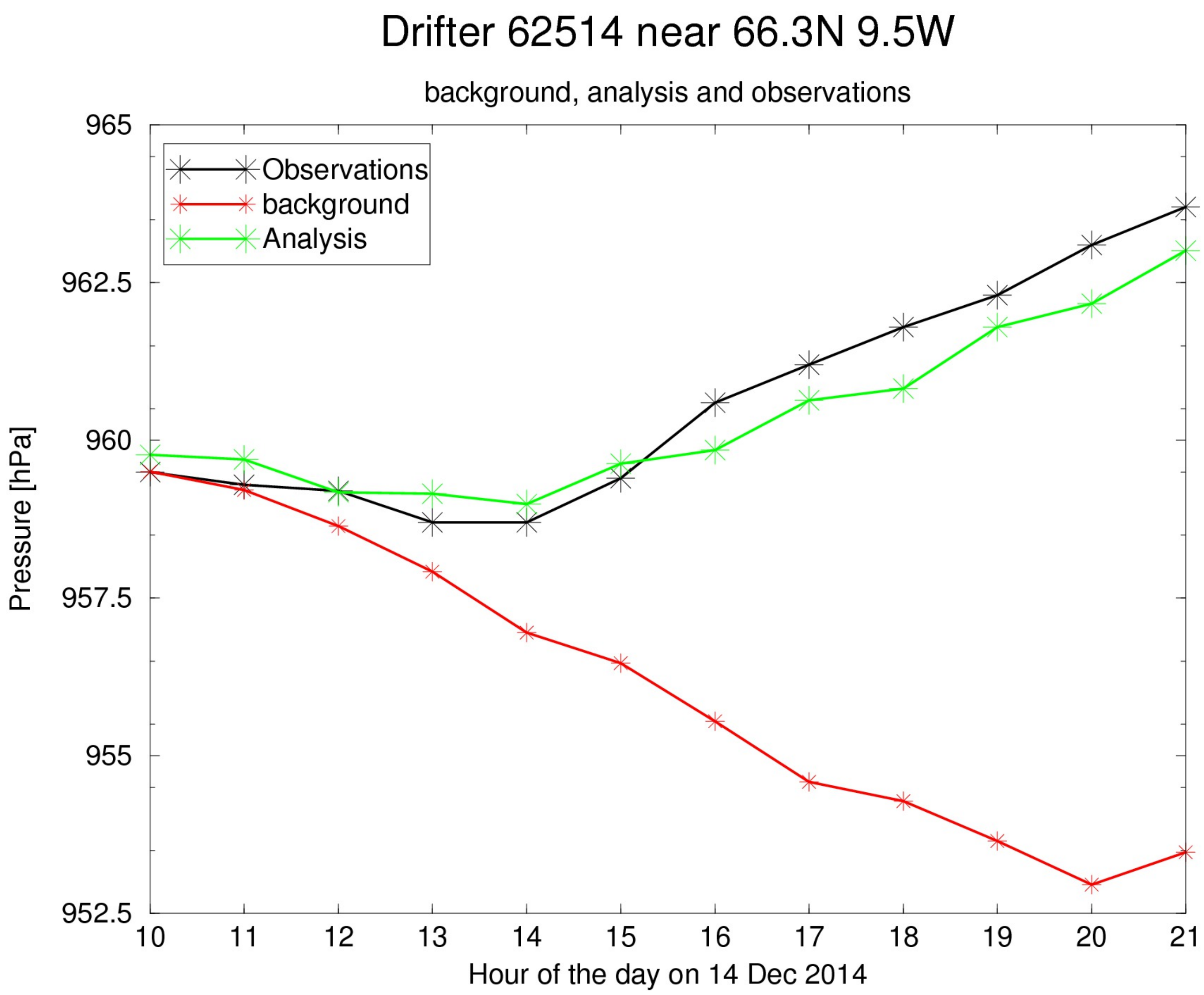
Case Study: Extra-tropical low near Iceland

14 December 2014



Case Study: Extra-tropical low near Iceland

14 December 2014

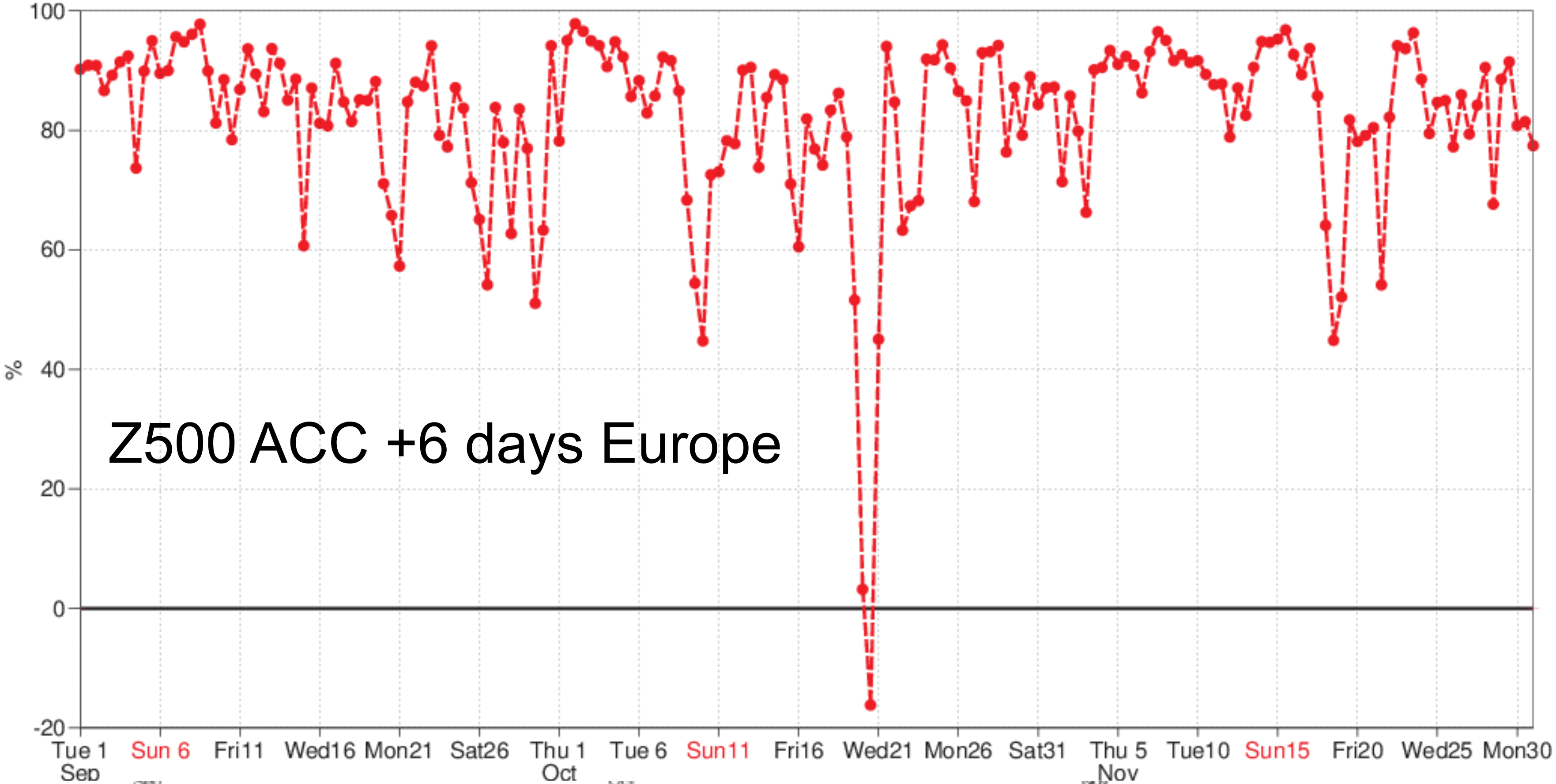


ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

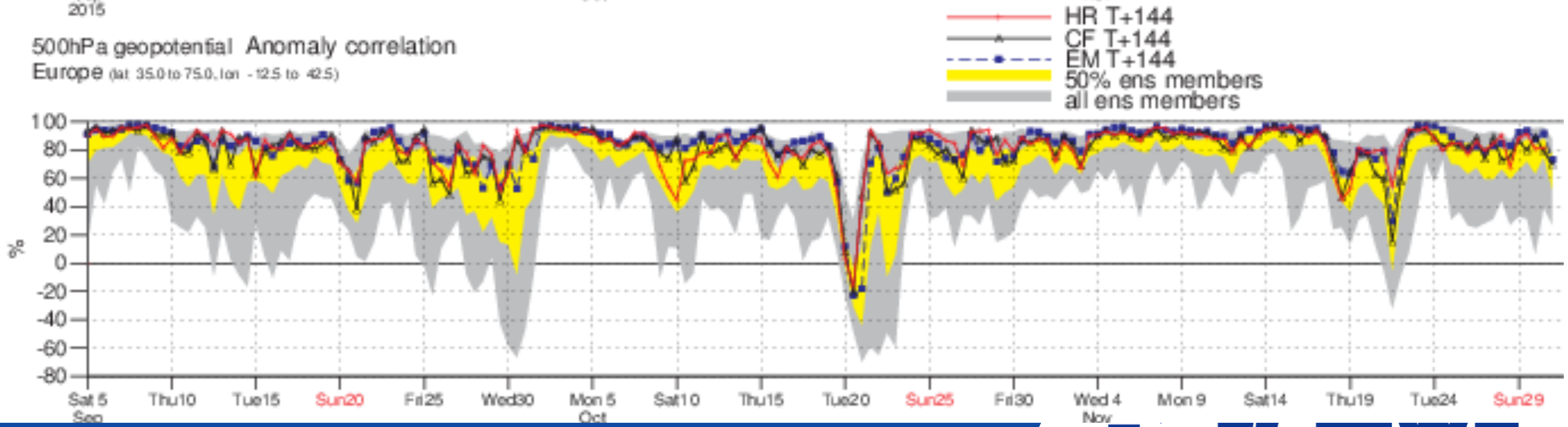
FD/RD SON 2015

Linus Magnusson

ECMWF

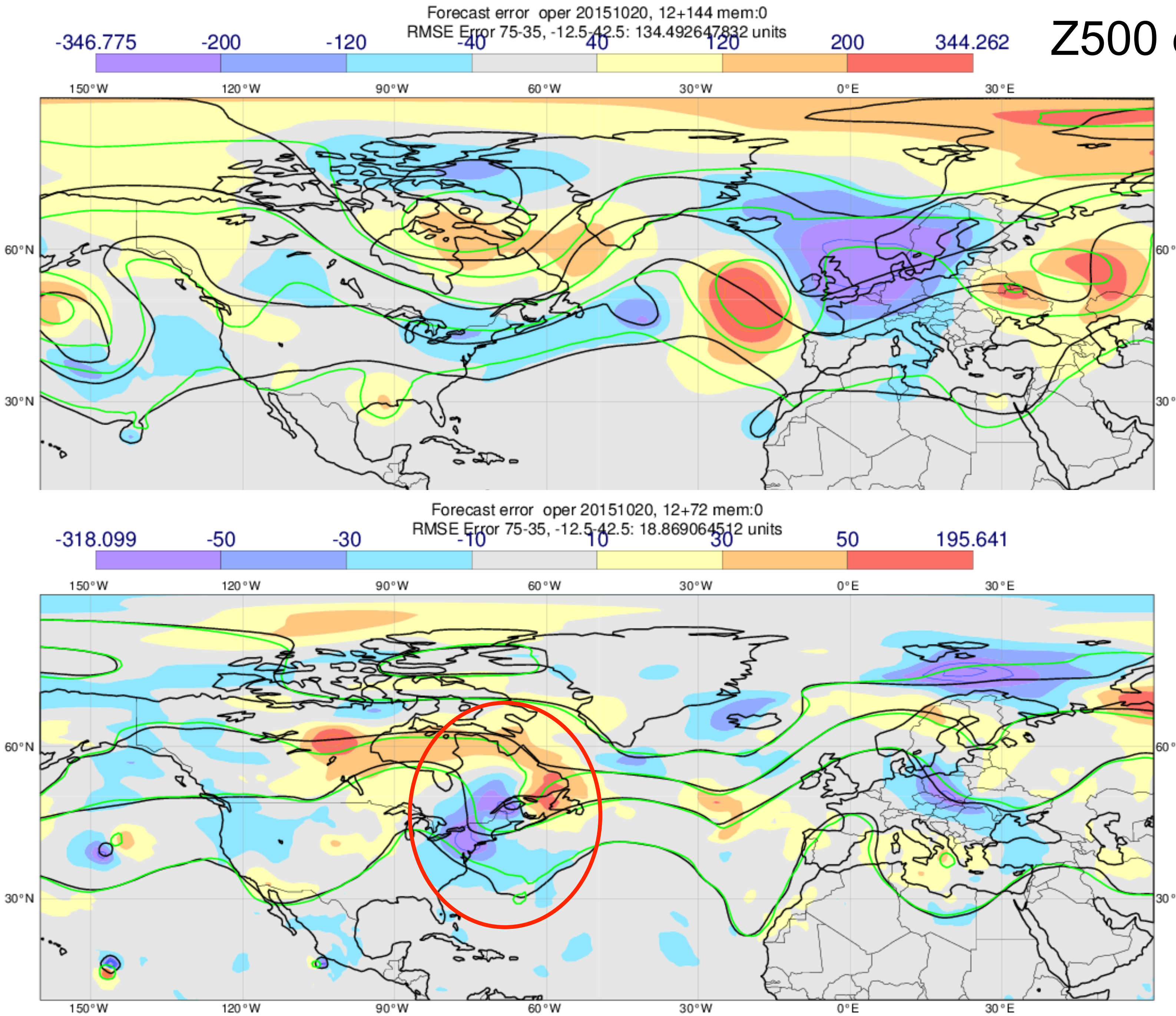


Z500 ACC +6 days Europe
%
100
80
60
40
20
0
-20
Tue 1 Sep
Sun 6
Fri11
Wed16
Mon21
Sat26
Thu 1 Oct
Tue 6
Sun11
Fri16
Wed21
Mon26
Sat31
Thu 5 Nov
Tue10
Sun15
Fri20
Wed25
Mon30



500hPa geopotential Anomaly correlation
Europe (lat 35.0 to 75.0, lon -12.5 to 42.5)
HR T+144
CF T+144
EM T+144
50% ens members
all ens members
%
100
80
60
40
20
0
-20
-40
-60
-80
Sat 5 Sep 2015
Thu10
Tue15
Sun20
Fri25
Wed30
Mon 5 Oct
Sat10
Thu15
Tue20
Sun25
Fri30
Wed 4 Nov
Mon 9
Sat14
Thu19
Tue24
Sun29

Z500 error

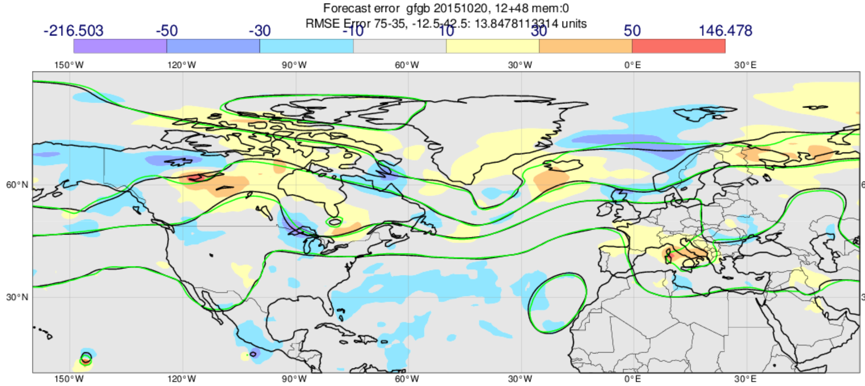


+6 days

+3 days

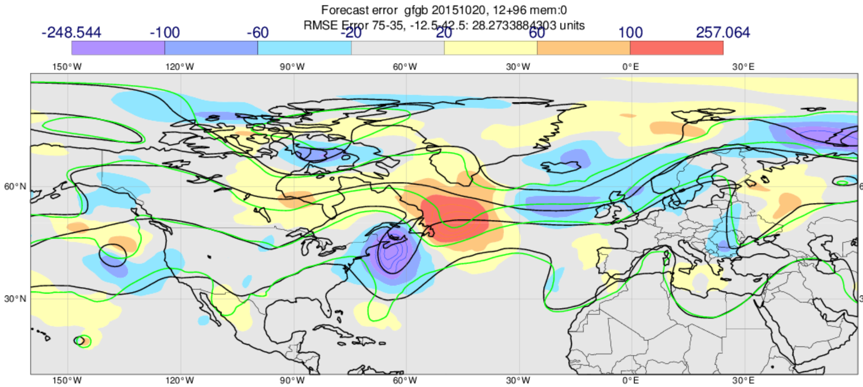
Control experiment 20 Oct. 12z

Z500 error +48h

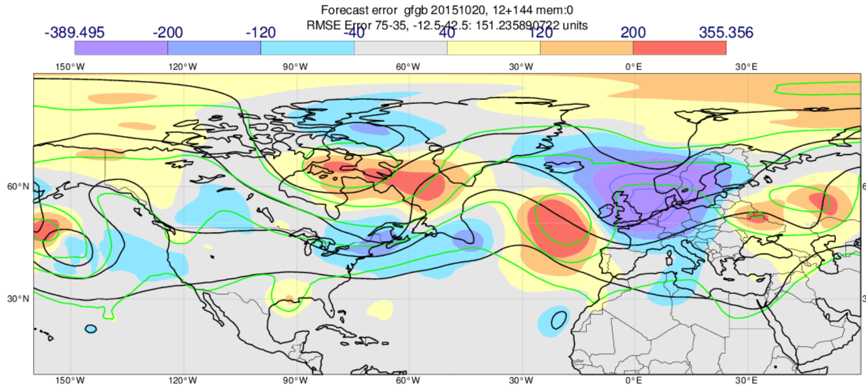


(forecast –black, analysis –green)

Z500 error +96h

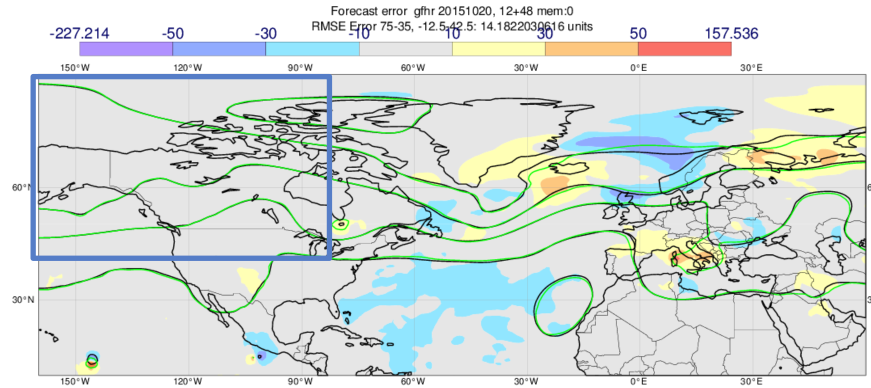


Z500 error +144h = 151 m



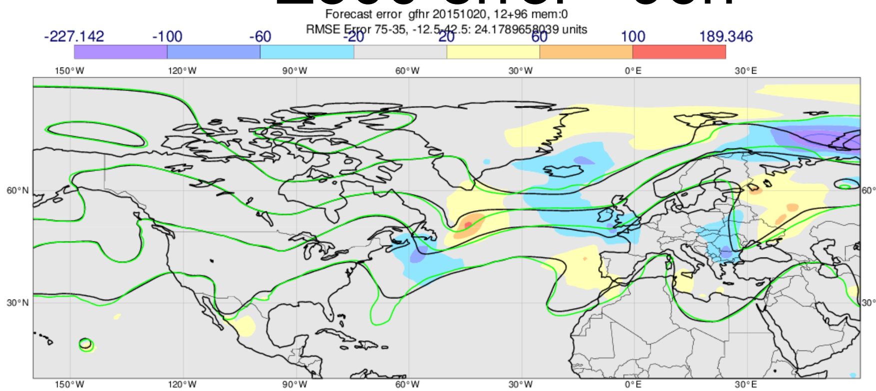
Relaxation in 40-90N, 180W-80W

Z500 error +48h

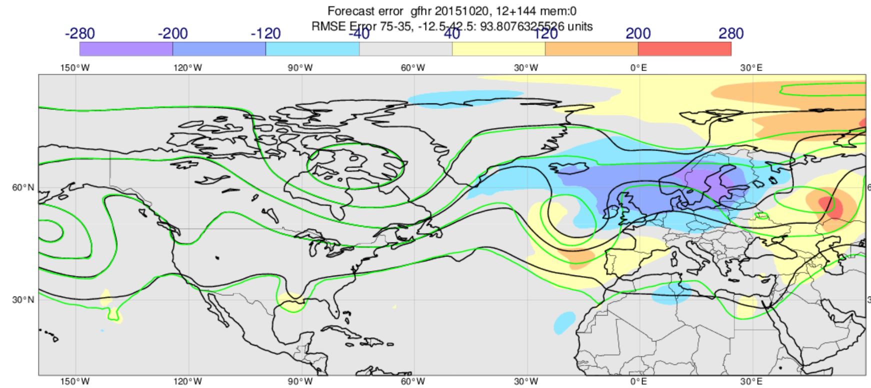


(forecast –black, analysis –green)

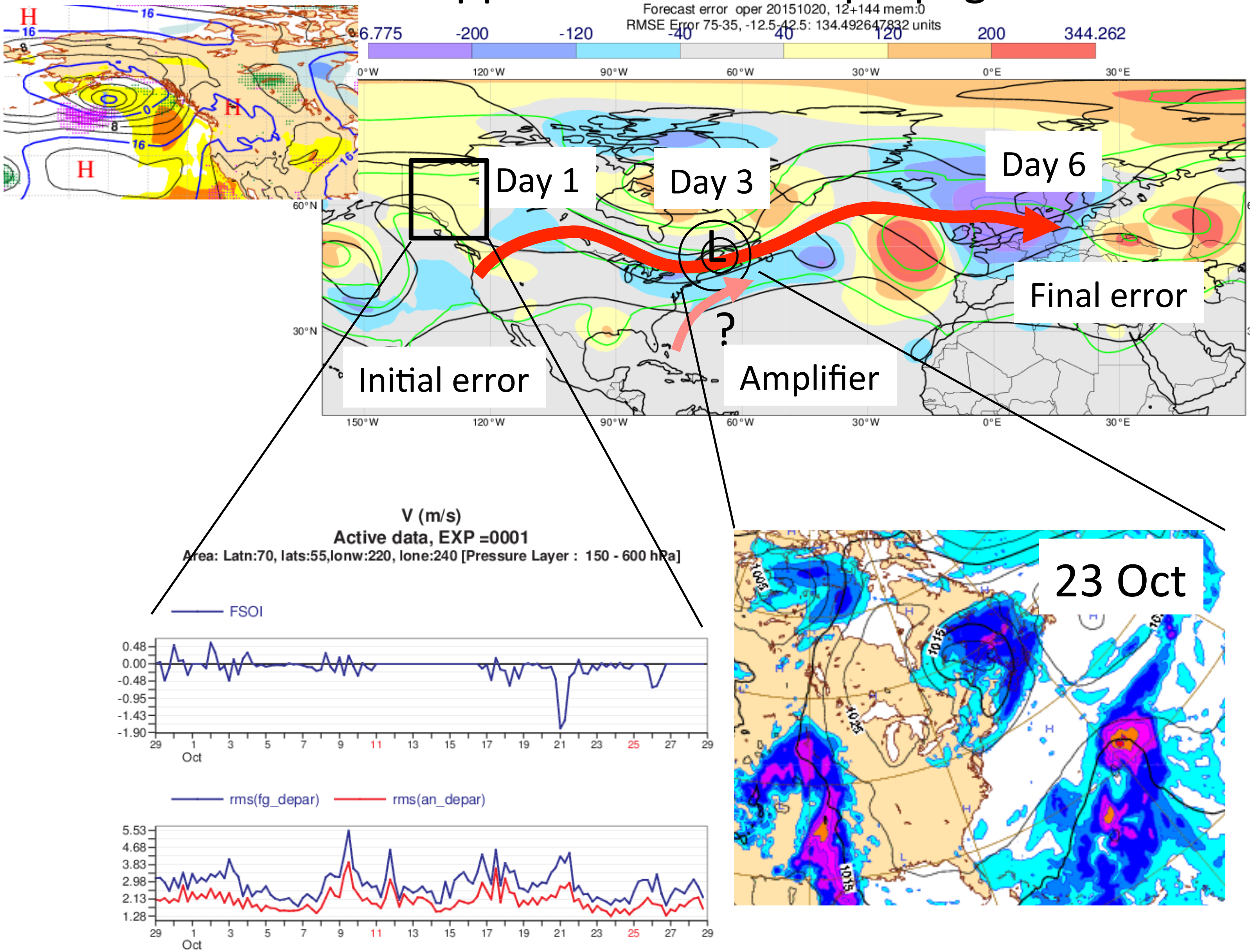
Z500 error +96h

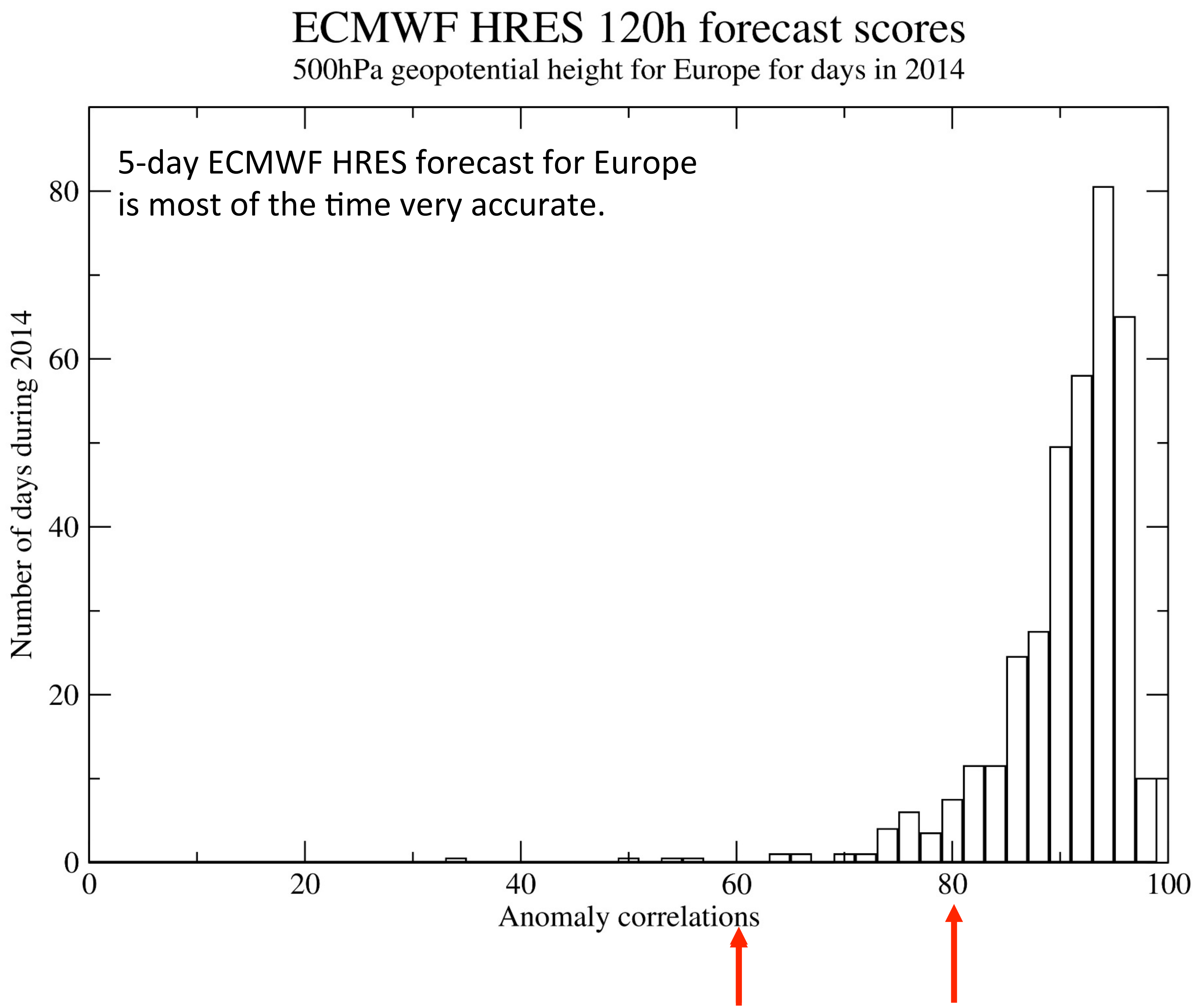


Z500 error +144h = 93 m



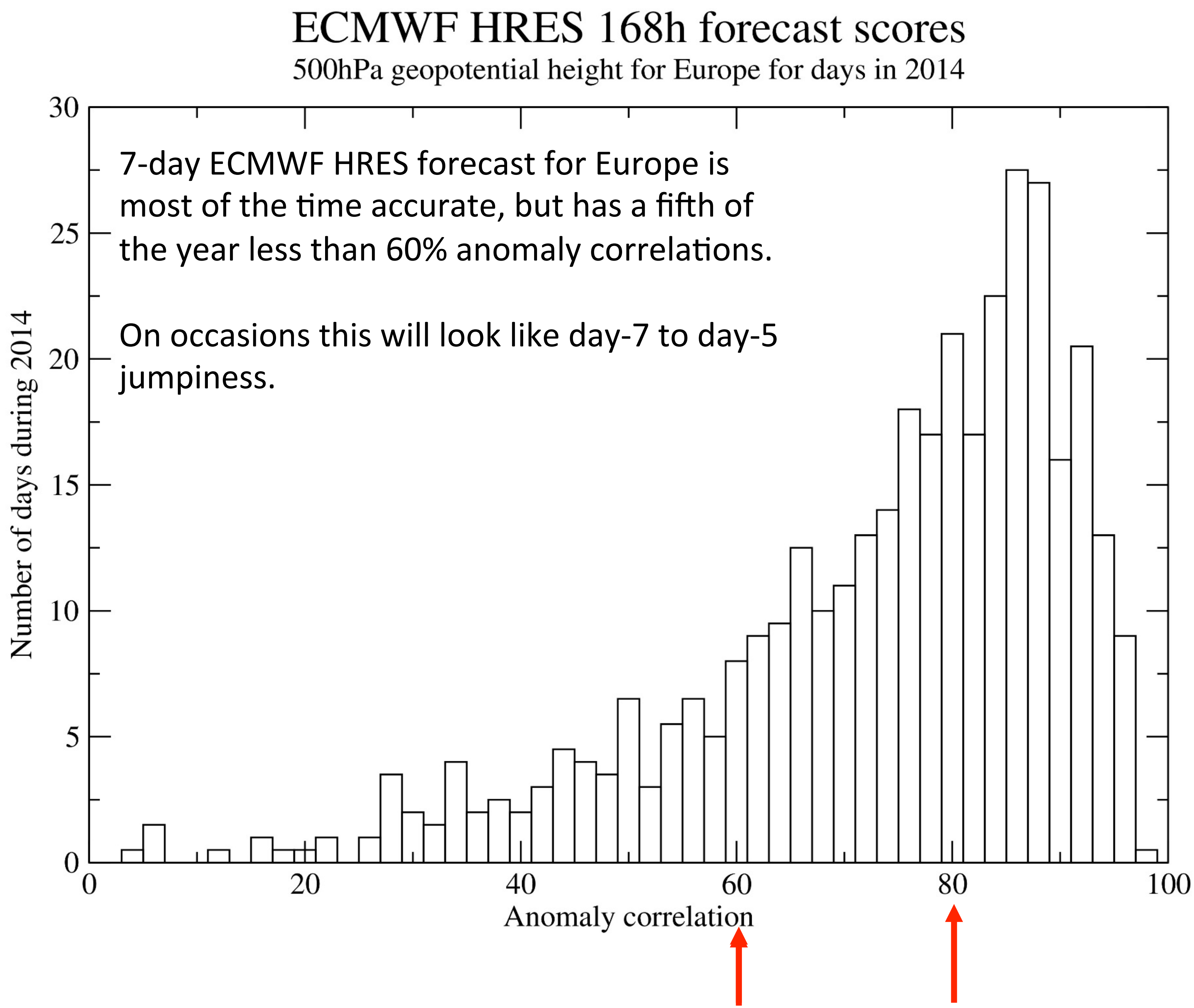
Approximated error propagation




ECMWF HRES 120h forecast scores
500hPa geopotential height for Europe for days in 2014
5-day ECMWF HRES forecast for Europe
is most of the time very accurate.
Number of days during 2014
0
20
40
60
80
0
20
40
60
80
100
Anomaly correlations

7-day ECMWF HRES forecast for Europe is most of the time accurate, but has a fifth of the year less than 60% anomaly correlations.

On occasions this will look like day-7 to day-5 jumpiness.



Recent improvements of the data assimilation system

Ensemble of Data Assimilations (EDA)

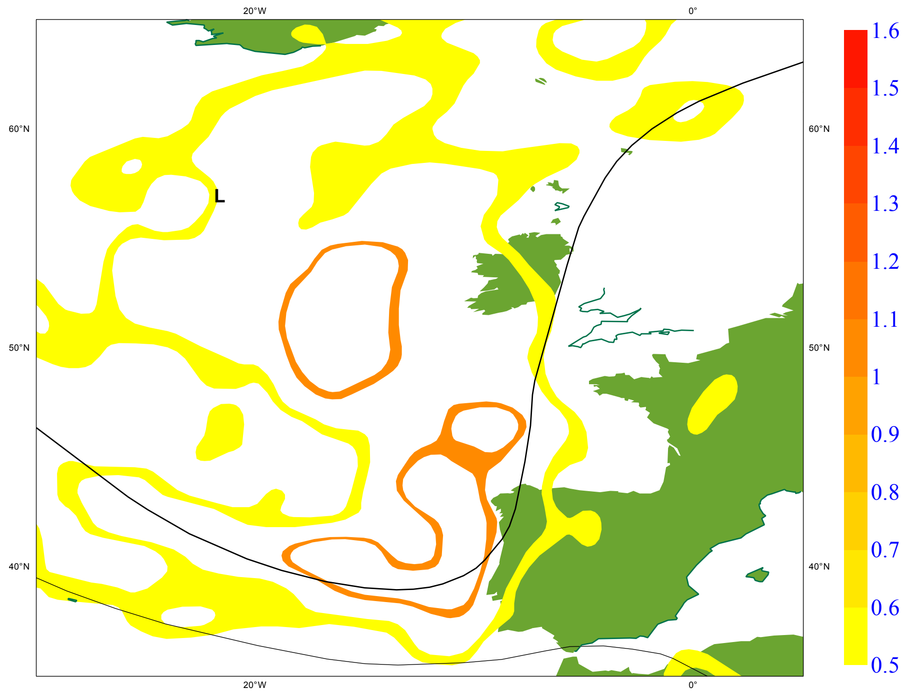
Run an ensemble of independent analyses with perturbed observations, model physics and Sea Surface Temperature fields.

25 EDA members plus a control at lower resolution.

Form differences between pairs of analyses (and short-range forecasts).

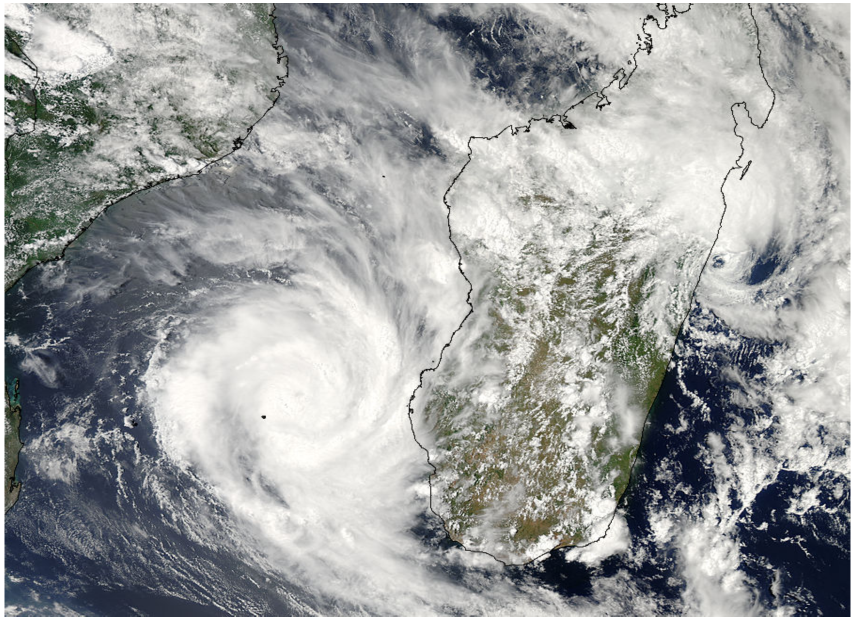
These differences estimates the statistical characteristics of analysis (and short-range forecast) error.

Yellow shading where the short-range forecast is uncertain →
give observations more weight in these regions.

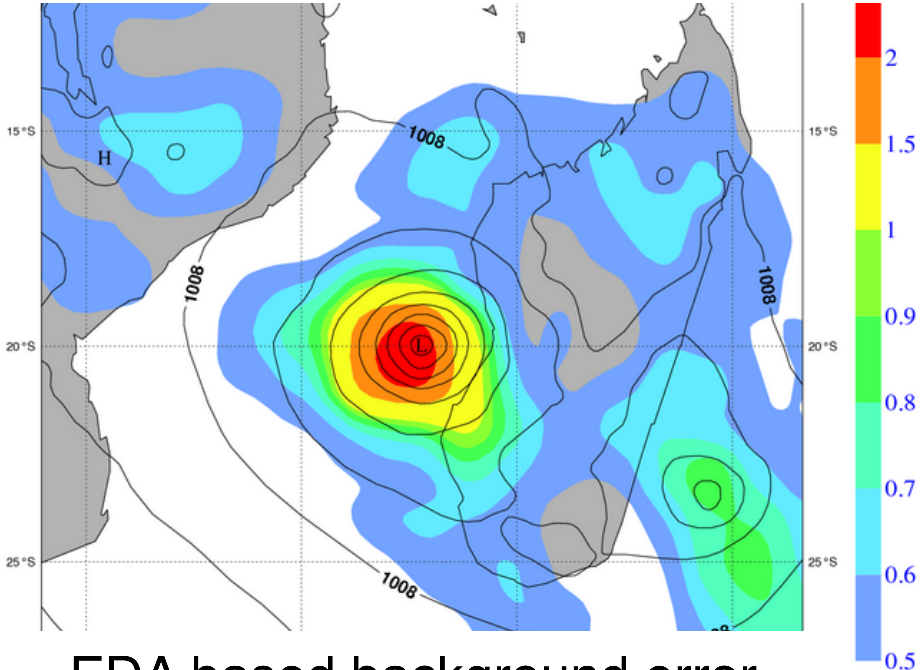


The EDA provides analysis and background uncertainty estimates

- To improve the initial perturbations in the Ensemble Prediction
- To calculate static and seasonal background error statistics
- To estimate flow-dependent background error covariances in 4D-Var
- To improve QC decisions and improve the use of observations in 4D-Var



Hurricane Fanele, 20 January 2009

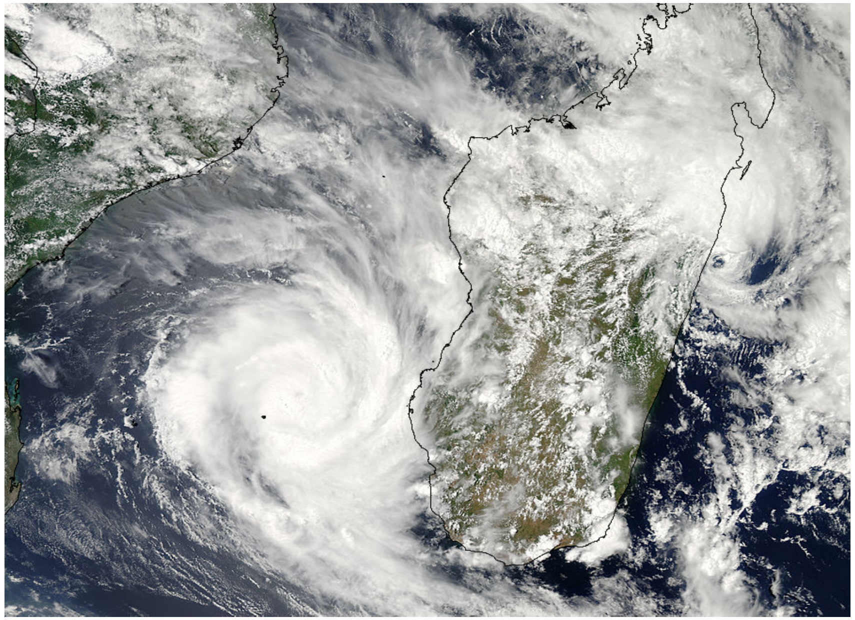


EDA based background error variance for Surface pressure

hPa

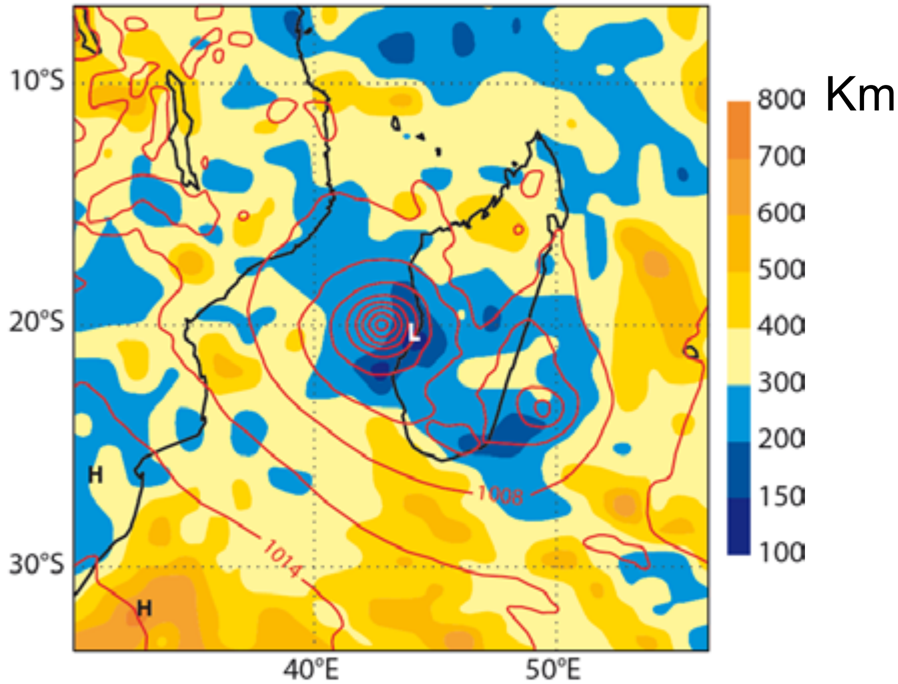
In November 2013 ECMWF will implement EDA based flow-dependent background error covariances in 4D-Var

The 25-member EDA has been used to estimate the background error covariance in 4D-Var.



Hurricane Fanele, 20 January 2009

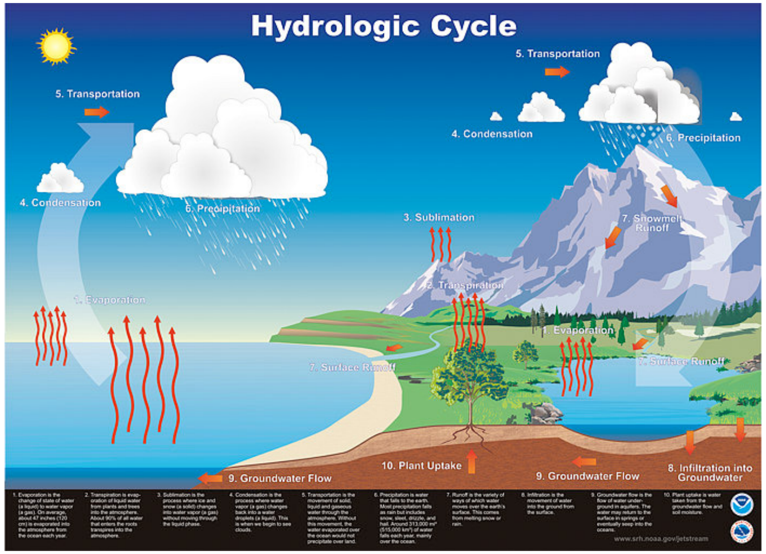
EDA based background error covariance length scale for Surface pressure



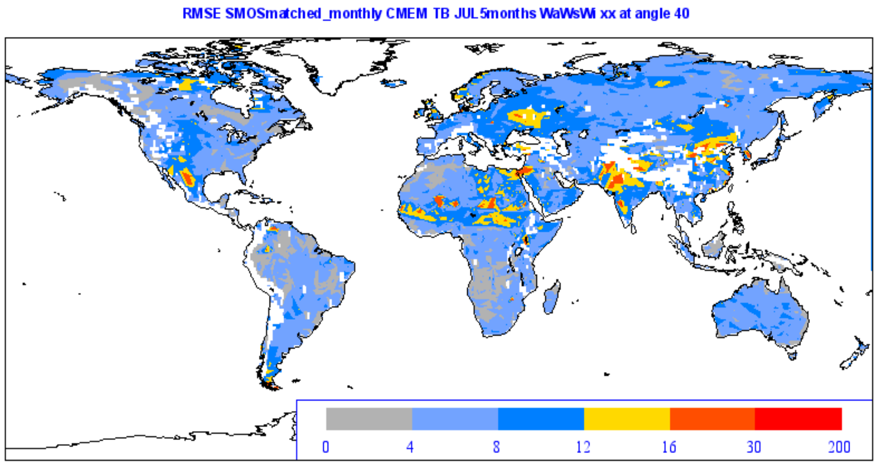
Land Data Assimilation

- Land surfaces: heterogeneities, range of spatial and time scales controlling the processes, reservoirs and fluxes.

- The Land Data Assimilation Systems (LDAS) make use of:
  - Processes and feedbacks represented with coupled land-atmosphere models (extension to carbon cycle available)
  - Data assimilation schemes, such as nudging, OI, EKF, EnKF, that update models states variables and/or surface parameters for NWP and climate applications
  - Routine Near Real Time observations with high information content about land surface variables (in-situ, SMOS, ASCAT, SMAP, etc.)

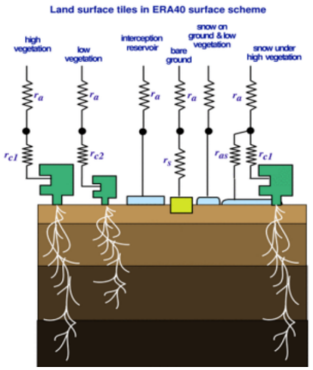


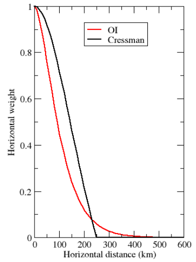
SMOS TB First Guess Departure (K) July 2012, RMSD=6.7K

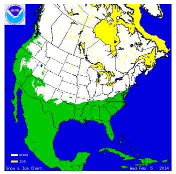


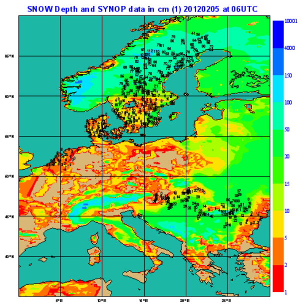
Snow in the ECMWF Data Assimilation System

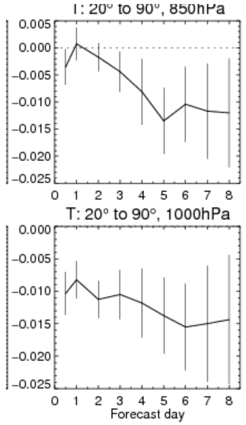
2009 2010 2011 2012 2013 2014 …











Snow Model
. Liquid Water
. Density
. Albedo
. Fraction

Snow Obs and DA
. Optimum Interpolation
. 4km IMS snow data
. Obs Quality Control
. IMS latency/acquisition

. Additional in situ obs
. WMO/SnowWatch action
. IMS data assimilation
. obs error revision

Snow Model & DA
. Multi-layer model
. Snow cover Fract
. BUFR SYNOP
. RT modelling
. Snow COST action

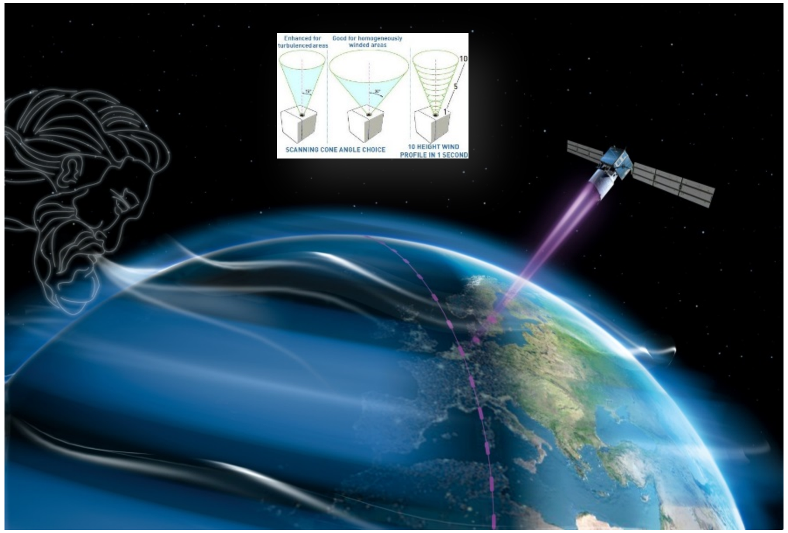
ECMWF Land Data Assimilation System:
https://software.ecmwf.int/wiki/display/LDAS/LDAS+Home

Future challenges

ADM-AEOLUS: An important wind profiling mission

An ESA Earth-explorer mission

Doppler wind lidar
Measures Doppler shift (due to wind) of backscattered UV laser light from the atmosphere



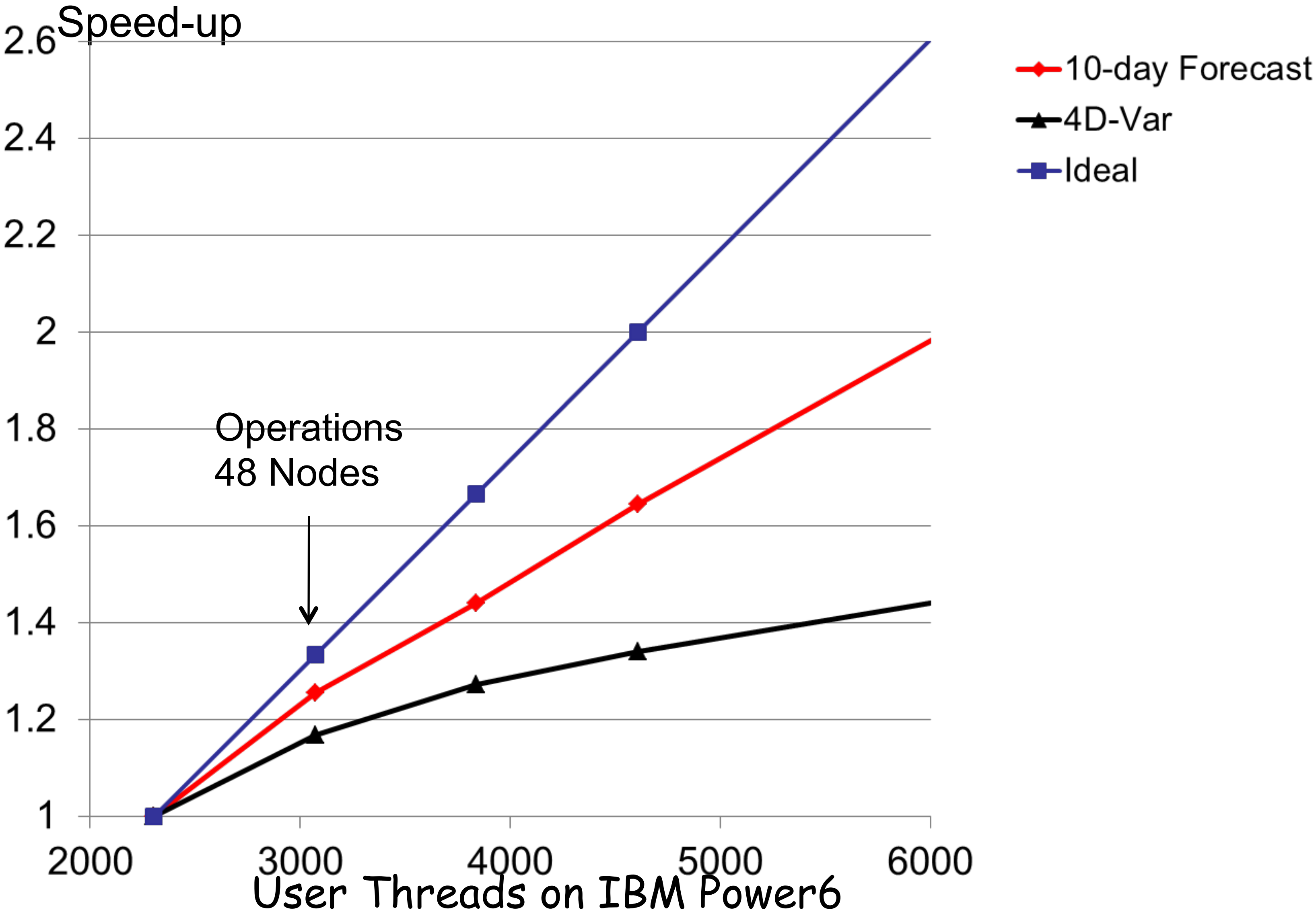
Courtesy: ESA

Main application is to improve global analyses and forecasts
Profiles of horizontal line-of-sight (HLOS) wind components

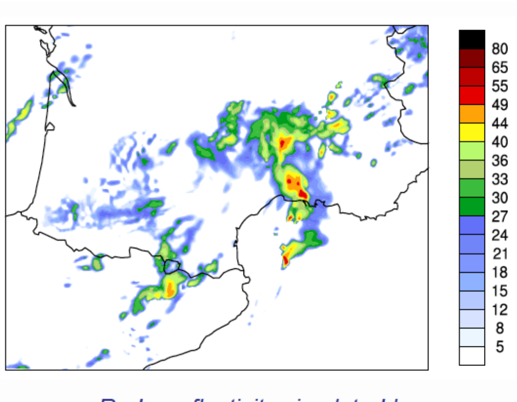
More wind profiles would greatly benefit the Global Observing System

Launch expected early 2017

Scalability of T1279 Forecast and 4D-Var

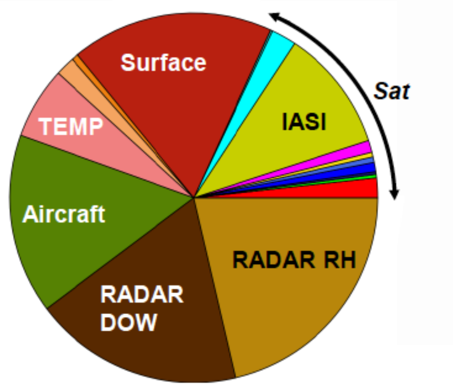


Challenges with Meso-scale Data Assimilation



Radar reflectivity simulated by AROME

Source: Thibaut Montmerle, Météo-France



Active obs in AROME for one rainy day

General

- Quick evolving processes
- Rapid updates requires (hourly or sub-hourly)
- Uncertainties and predictability

Remote sensing observations

- More timely use of information from GEO satellites
- Novel observations for convective scale DA
- Assimilate cloud-affected radiances
- Non-linear observation operators
- Accuracy and efficiency of radiative transfer in all-sky

Covariance modeling

- Traditional balance (e.g. geostrophic & hydrostatic) not applicable at high-resolution
- Impact on ensemble size
- Complex, non-linear, flow-dependent relationship between model variables
- Significant model error (in phase and amplitude)

Conclusions

- Data assimilation is combining information from a short-range forecast and observations to produce a more accurate estimate of the initial state
- An accurate estimate of the initial state is an essential requirement for producing accurate forecasts
- Observations are essential for data assimilation
- The potentials for improving the accuracy of the initial state are great (improved models and better use of observations)
- The best NWP data assimilation systems today are using hybrid variational/ensemble methods
- Efficient use of future HPCs will be a fundamental driver
- Specific challenges for meso-scale data assimilation

Thank You!