In-situ and *some* **actively sensed** observations **plus** observation quality control

Sean Healy

Earth System Assimilation Section, ECMWF

Sean.Healy@ecmwf.int

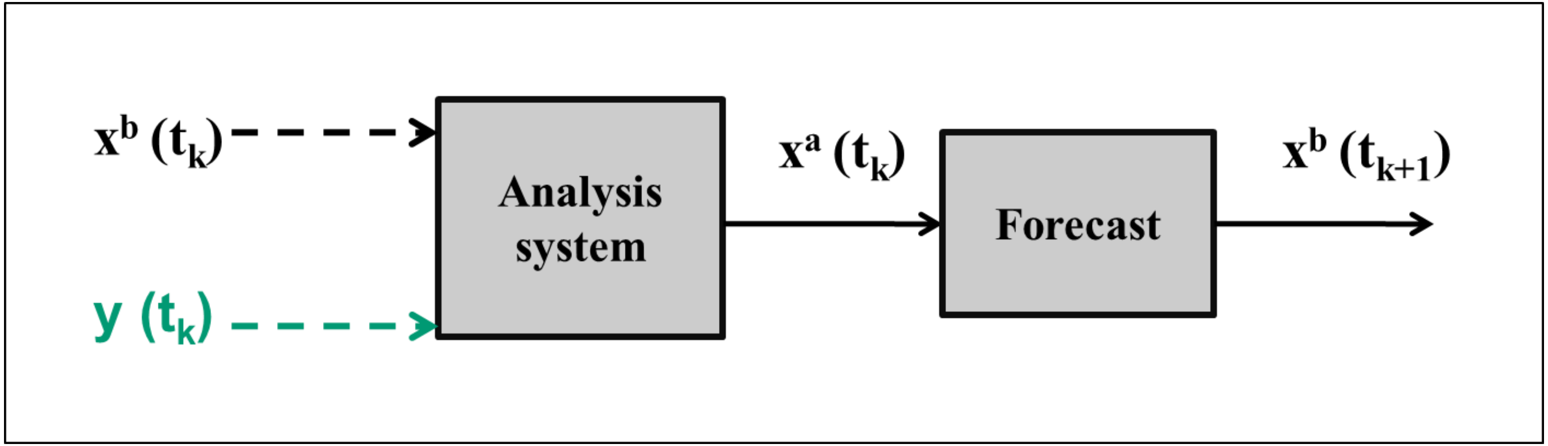
March 17, 2025

Bruce Ingleby, Giovanna De Chiara, Saleh Abdella, Lars Isaksen, Elias Holm. Mohamed Dahoui

Overview of lecture

- **Some jargon/definitions**
- **Review of <u>some</u> *in-situ* and *actively sensed* observations in global NWP**
  - How we assimilate the data, recent developments
- **Quality control (briefly!)**
  - What we (try to) do when the **actual observation errors are not what we expect or assume**, given the **assumed covariance matrices R**
- **Broad scope for lecture. Hopefully, it will "signpost" you to useful material.** **Plus, I can point you to the ECMWF experts**

You'll see versions of this many times in the next week



We combine a <u>diverse</u> set of observations with an NWP forecast in a *statistically optimal* (e.g., 3D-Var, 4D-Var, …) way to produce the "analysis"

This "analysis" is the initial conditions for the next forecast

We will be discussing a *subset* of the observations in $y(t_k)$ used here at ECMWF to produce the analysis

Useful data assimilation jargon

- The analysis is the *initial conditions* needed for the NWP forecast model run
- A ***previous*** forecast provides the background (or *a prior*) information to the analysis
- Observation operators, ***H***, enable observations and the model background to be **compared** in "**observation space**"
- In **observation space**, the differences we compute in the comparisons are called departures or innovations – "o-b"
  - They are central in providing observation information to the analysis
- These corrections, or increments, are added to the background to give the analysis (or *posterior estimate*)
- Observation operators also enable a comparison of observations and the analysis (analysis departures: "o-a")
- We'd expect abs(o-a)<abs(o-b) if the DA system is working correctly

Example: Statistics of departures

Background departures: $y - Hx_b$ (o-b)

Analysis departures: $y - Hx_a$ (o-a)

y = observations

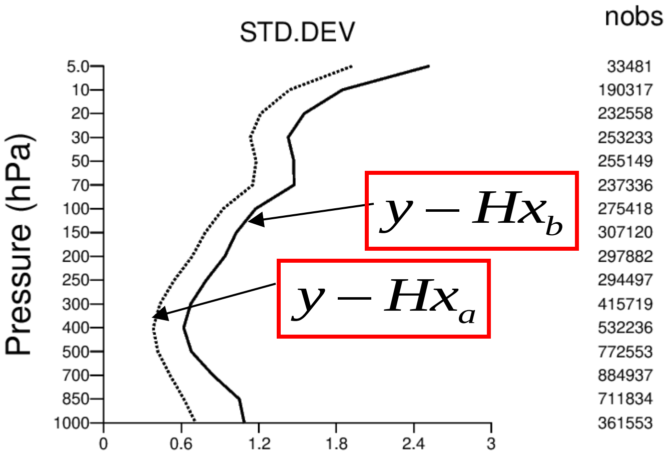
x_a = analysis state

x_b = background state

Number of observations

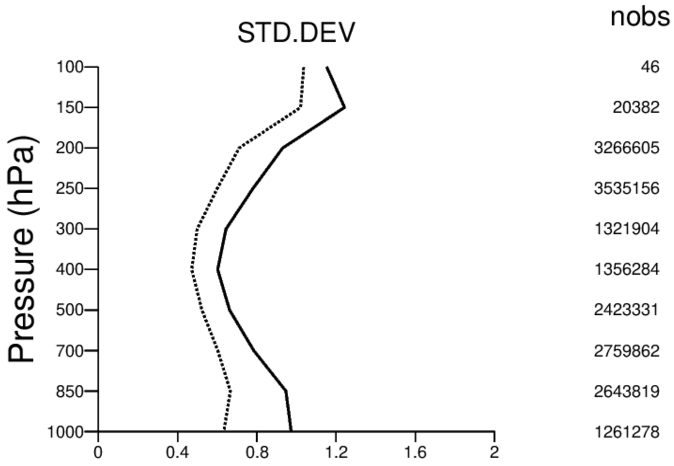
2020010200-2020020112(12)

Radiosonde temperature



2020010200-2020020112(12)

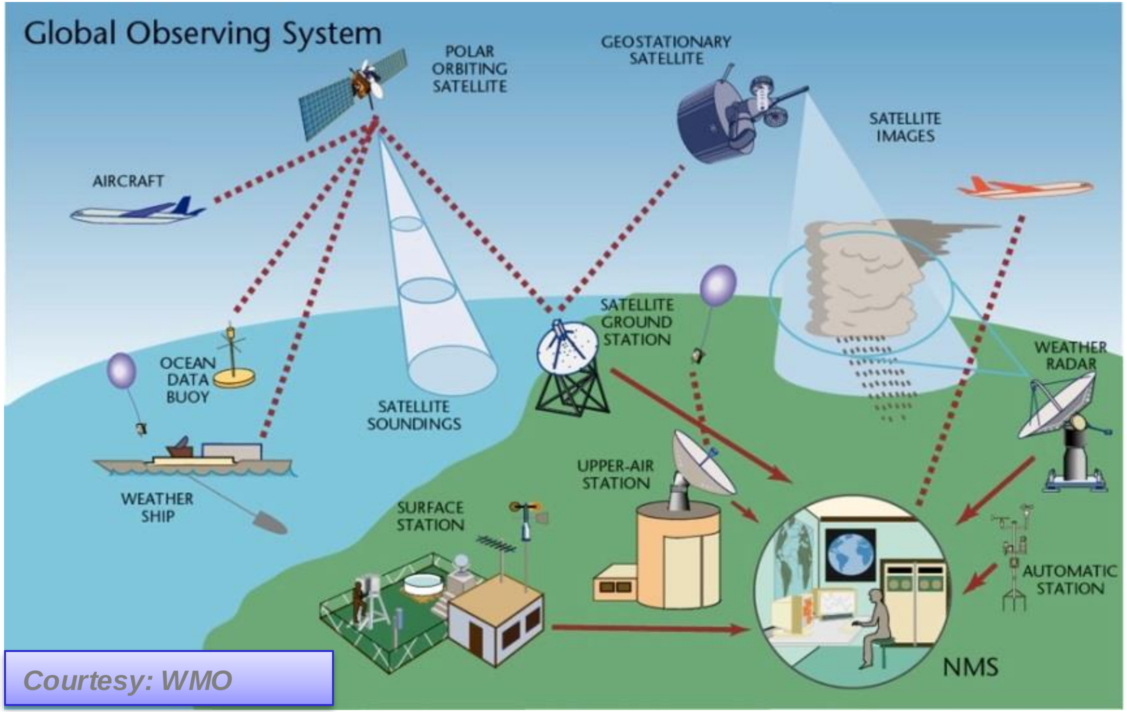
Aircraft temperature



- The standard deviation of background departures for both radiosondes and aircraft is around 0.7-1.0 K in the mid-troposphere.
- The standard deviation of the analysis departures is smaller – because the analysis has “drawn” to the observations.

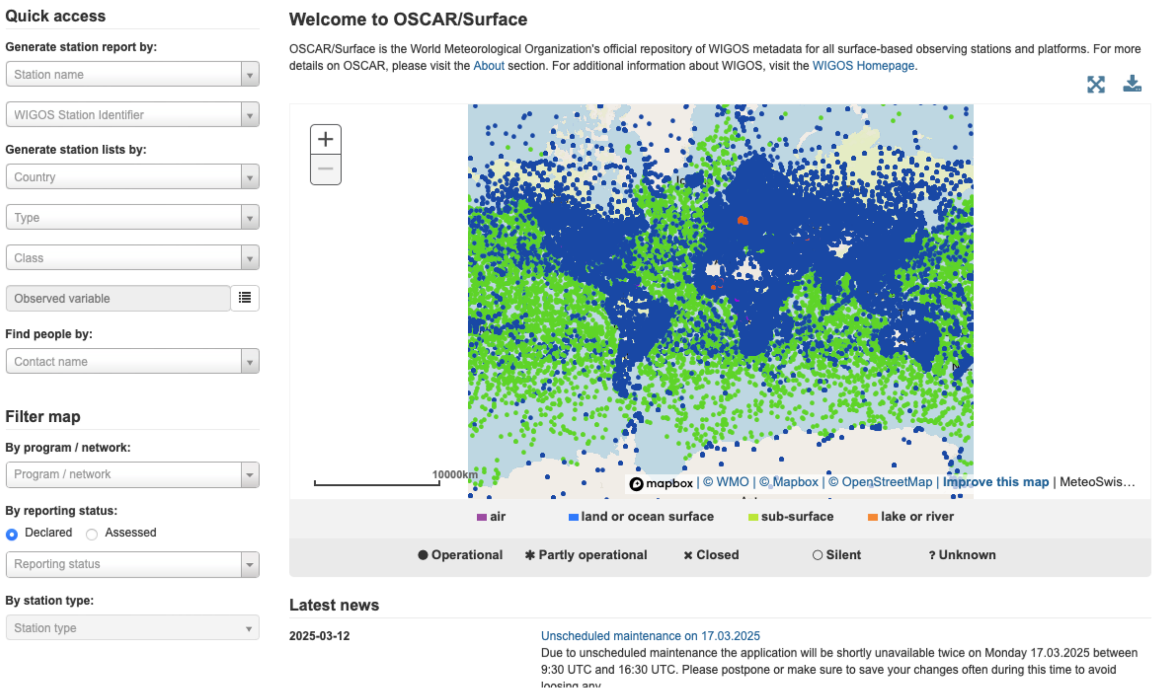
WMO Integrated Global Observing System

The WMO OSCAR database provides an excellent overview of the observations available



https://www.wmo-sat.info/oscar/
https://oscar.wmo.int/surface//index.html#/

WMO OSCAR (Observing Systems Capability Analysis and Review Tool)



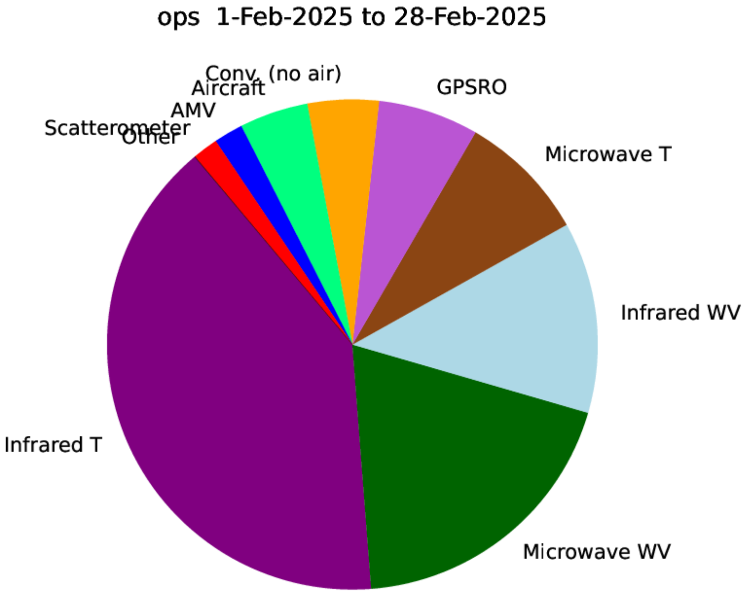
https://oscar.wmo.int/surface//index.html#/

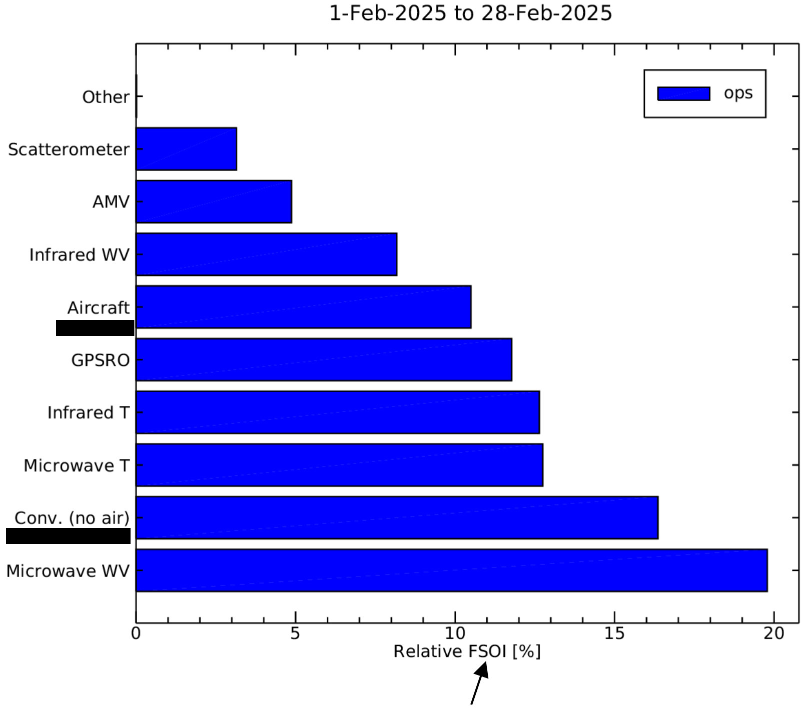
https://www.wmo-sat.info/oscar/

In-situ

- Sometimes called “**conventional**”
- Used since the very early days of NWP (1950’s). Now about 10 % of data we use
- Providing both surface and upper-air information. Most abundant in the NH
- Usually characterized by *relatively simple* forward operators, ***H***, because the measured quantities are geophysical (e.g., P, T, u, v, Q). **Simple, often “*messy*”, but they remain really a key component of global observing system!**
- Also useful for **forecast verification** and they help **constrain bias corrections** applied to satellite radiances
- **See really important review by**
  - Pauley P, Ingleby B (2022) **Assimilation of in-situ observations**. In: Park SK, Xu L (eds) Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. IV). Springer. Pages 293-371 in https://link.springer.com/book/10.1007/978-3-030-77722-7

In-situ are roughly 10 % of the data we currently assimilate - **but they have a big impact despite their number**



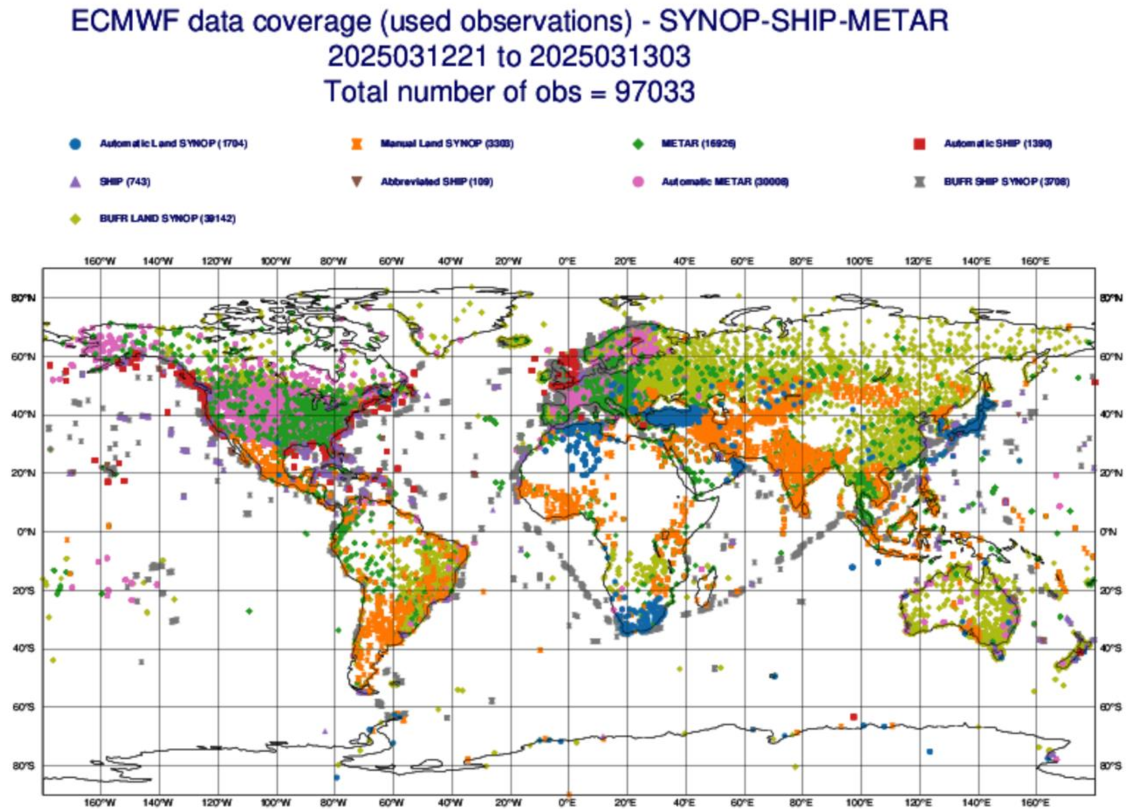


You will come across FSOI again later in the week
See diagnostics talk by Bruce Ingleby, March 19

In-situ data: *which parameters are assimilated in atmosphere analysis?*

Instrument	Parameters	Height
SYNOP **SHIP** **METAR**	pressure, dew-point **temperature,** **temperature,** pressure, wind vector	Station altitude, 2m Ships ~25m Station altitude
BUOYS	pressure, wind	MSL, 2-10m
TEMP **TEMPSHIP** **DROPSONDES**	temperature, humidity, wind vector	Profiles
PROFILERS	Wind vector	Profiles
Aircraft	temperature, wind vector, humidity	Profiles near airports + Flight level data

Geographical coverage



ECMWF

Geographical coverage

ECMWF | 50

Search site... Help Log in

Home About Forecasts Computing Research Learning Publications Anniversary

Number 178 - Winter 2024

Published in January 2024

View all Newsletters

Editorial

The pace is quickening

News

Capturing extreme rainfall events

AIFS: a new ECMWF forecasting system

ECMWF meets its representatives in 2023

ECMWF heating rates support PHILEAS aircraft campaign

How did ECMWF seasonal forecasts perform for the European summer of 2023?

A daily forecast with the prototype global Extremes Digital Twin of Destination Earth

EARTH SYSTEM SCIENCE

Improved two-metre temperature forecasts in the 2024 upgrade

Bruce Ingleby, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, Kenta Ochi (Japan Meteorological Agency), Ewan Pinnington, Patricia de Rosnay

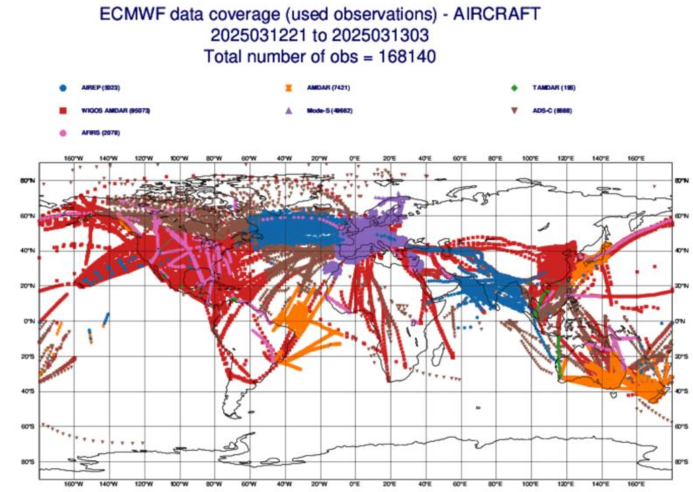
Two-metre temperature (T2m) is a key forecast variable. Here we describe how changes to ECMWF's Integrated Forecasting System (IFS), which are expected to become operational as part of IFS Cycle 49r1 in 2024, improve short-range forecasts of T2m. Physics changes in IFS Cycle 49r1 will include extensive improvements to surface vegetation fields and changes to make the interpolation of the temperature to 2 m more realistic. There have also been major upgrades to the data assimilation system, which establishes the initial conditions of forecasts (the 'analysis'). Currently, T2m ('screen temperature') and two-metre humidity ('screen humidity') from SYNOP weather station reports and METAR aviation weather reports are assimilated in a separate land surface analysis – primarily to update initial soil moisture and temperature conditions. However, only daytime screen humidity is assimilated in the atmospheric component of the data assimilation system, 4D-Var. In IFS Cycle 49r1, T2m will be assimilated in 4D-Var. This proved beneficial after extensive testing/tuning. Improvements in the separate snow data assimilation system as well as in T2m and soil moisture land data assimilation are also applied. Overall this results in better T2m forecasts, especially in northern hemisphere winter.

We recently improved our use of synop/metar 2m temperatures in operations

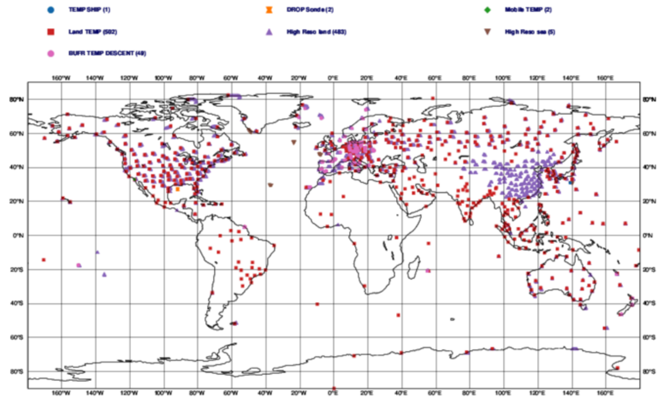
https://www.ecmwf.int/en/newsletter/178/earth-system-science/improved-two-metre-temperature-forecasts-2024-upgrade

Geographical coverage

Aircraft

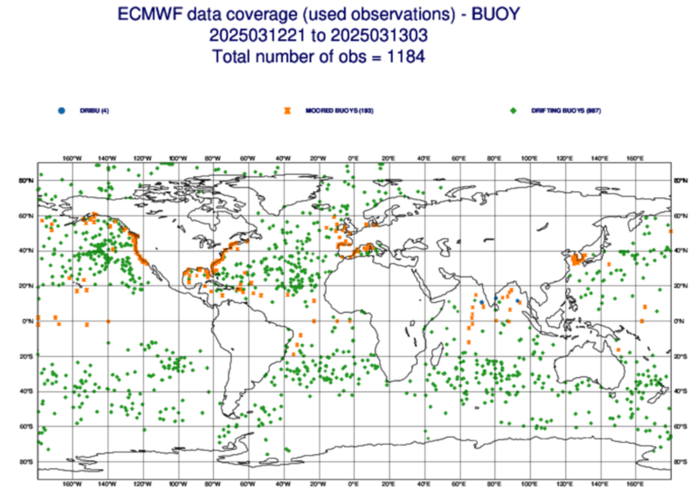


Radiosonde
Dropsonde



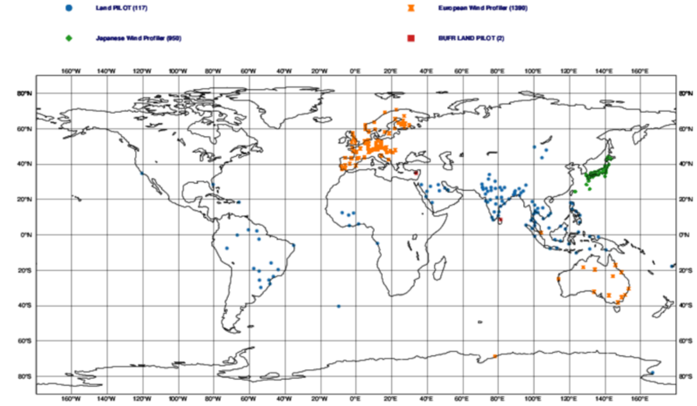
Geographical coverage

Buoy



F

Wind
profilers



Impact of various observing systems at ECMWF

Provided by Niels Bormann – 2021 annual seminar

https://events.ecmwf.int/event/217/contributions/2049/attachments/1397/2509/AS2021_Bormann.pdf

Observing system experiments – denying observation datasets

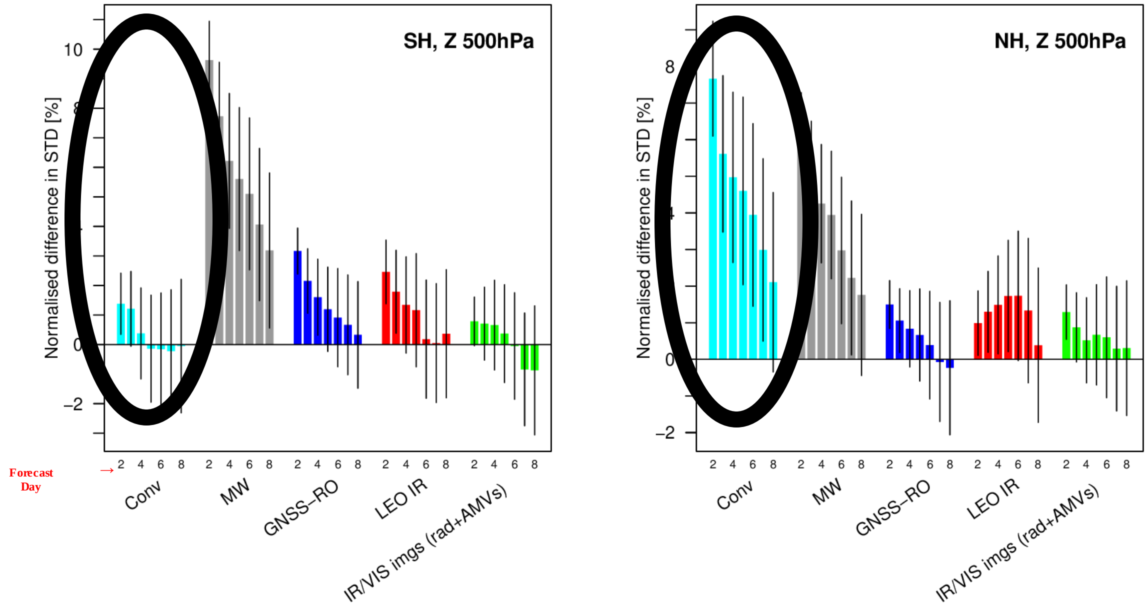
- **Periods, 6 months in total:**
  5 Sept – 2 Nov 2020
  1 Jan – 28 Feb 2021
  1 May – 30 June 2021
  (each + 4 days spin-up prior)

- **Denial** experiments compared to a full system for:
  - **Conventional in-situ observations**
  - MW radiances
  - IR sounders from LEO
  - IR/VIS imagers (AMVs + IR radiances)
  - GNSS-RO

- **Resolution: T_{CO} 399 (~25 km)**
- **Background error from operational system**

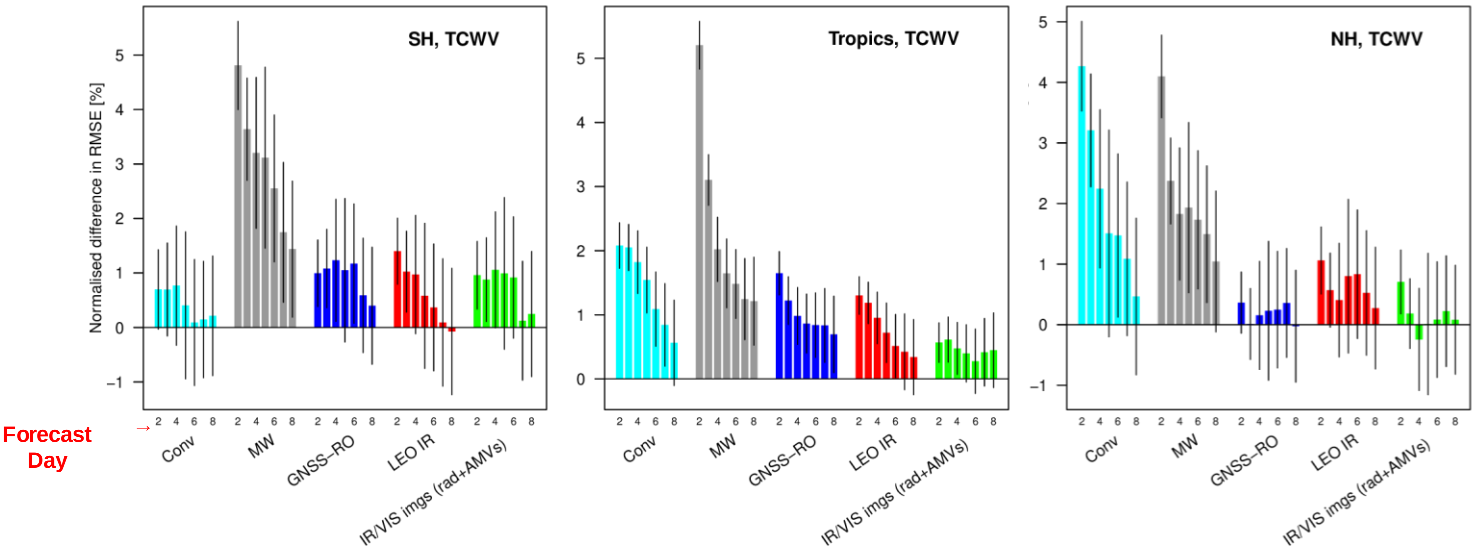
Forecast impact, day 2-8: 500 hPa geopotential

Verified against operational analyses, 3 periods combined



Forecast impact, day 2-8: Total column water vapour

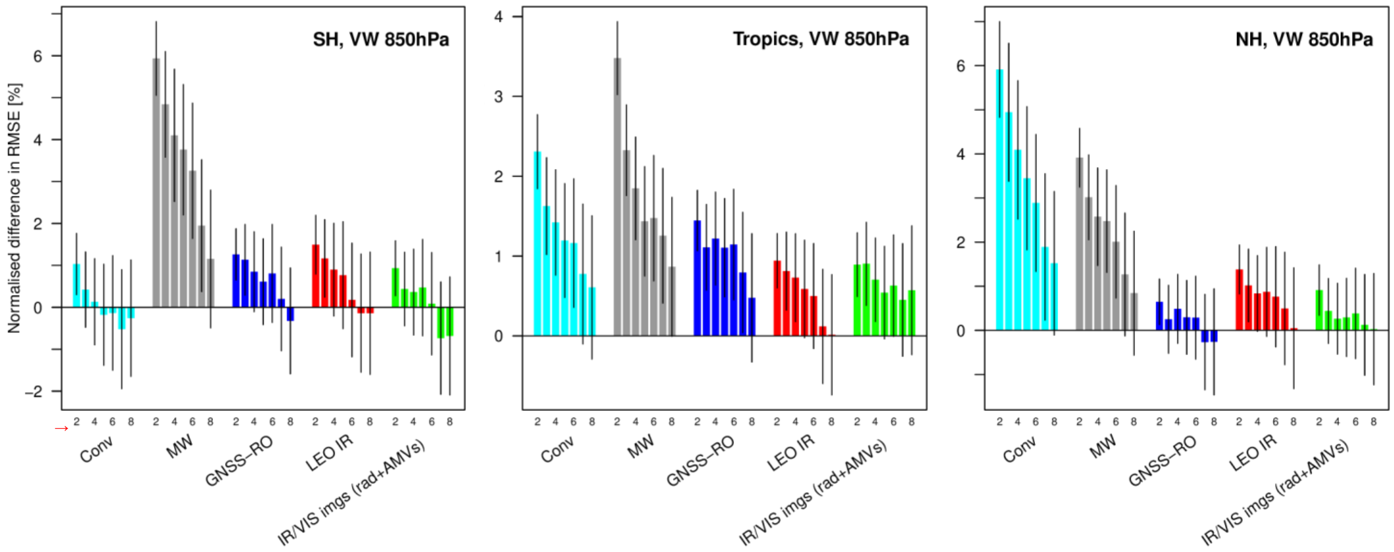
Verified against operational analyses, 3 periods combined



ECMWF

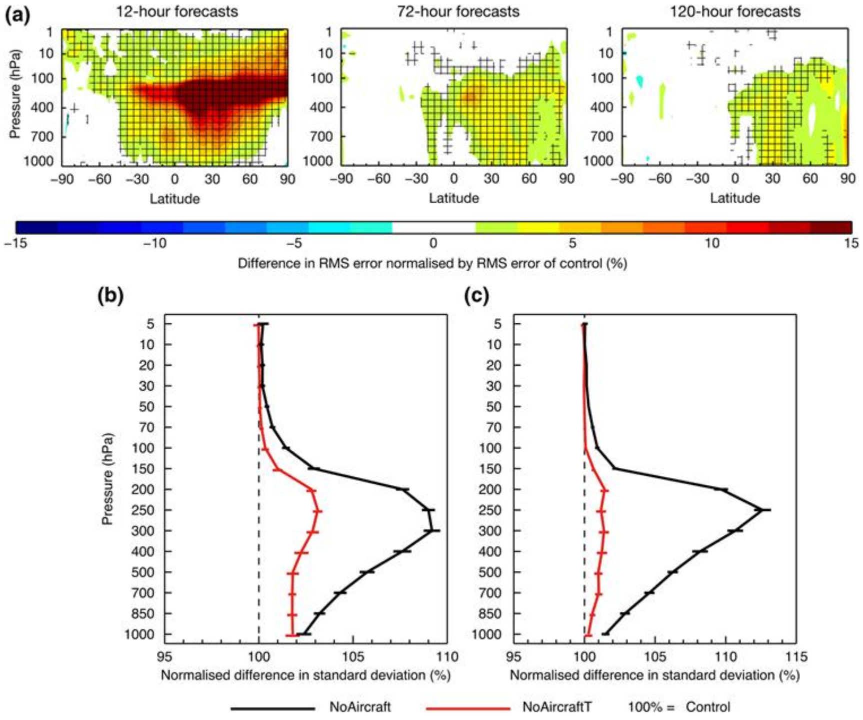
Forecast impact, day 2-8: Wind at 850 hPa

Verified against operational analyses, 3 periods combined



Forecast Day

Aircraft measurements of wind more important that temperature



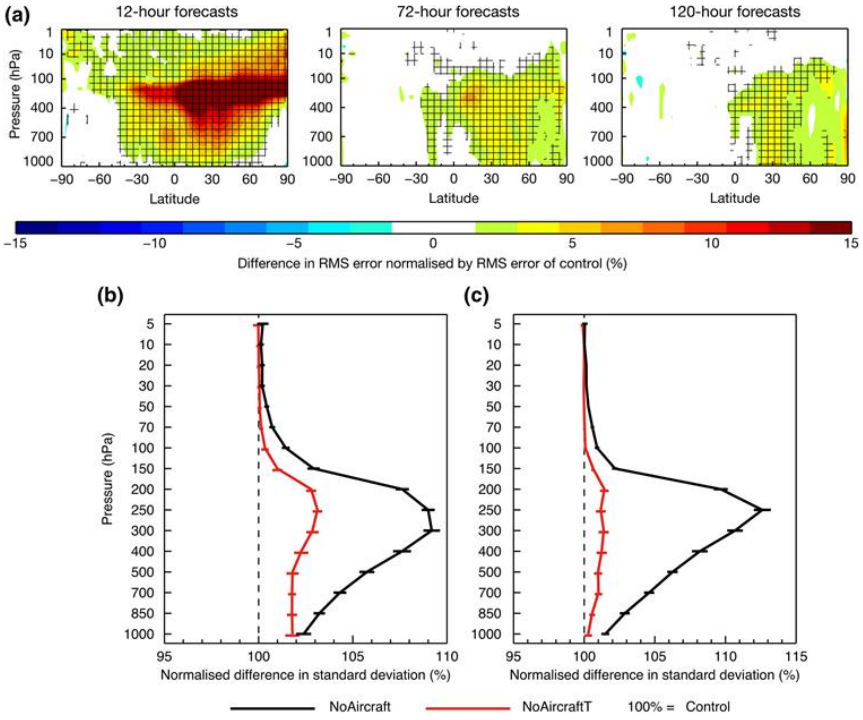
The short-range forecast fit to radiosondes degrades (>10% == HUGE)

b) TEMP c) wind vectors

The aircraft winds provide more information than the aircraft temperatures

Geophysical Research Letters, Volume: 48, Issue: 4, First published: 06 December 2020, DOI: (10.1029/2020GL090699)

Aircraft measurements of wind more important that temperature



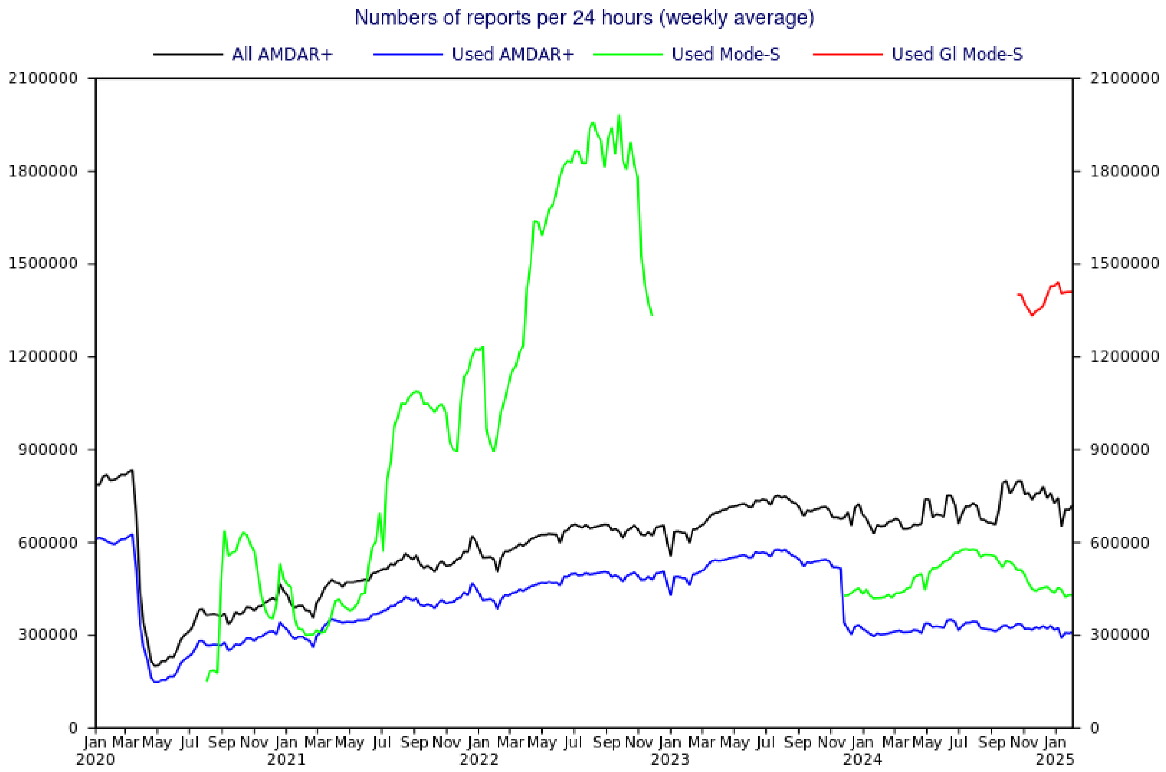
The short-range forecast fit to radiosondes degrades (>10% == HUGE)

b) TEMP c) wind vectors

The aircraft winds provide more information than the aircraft temperatures

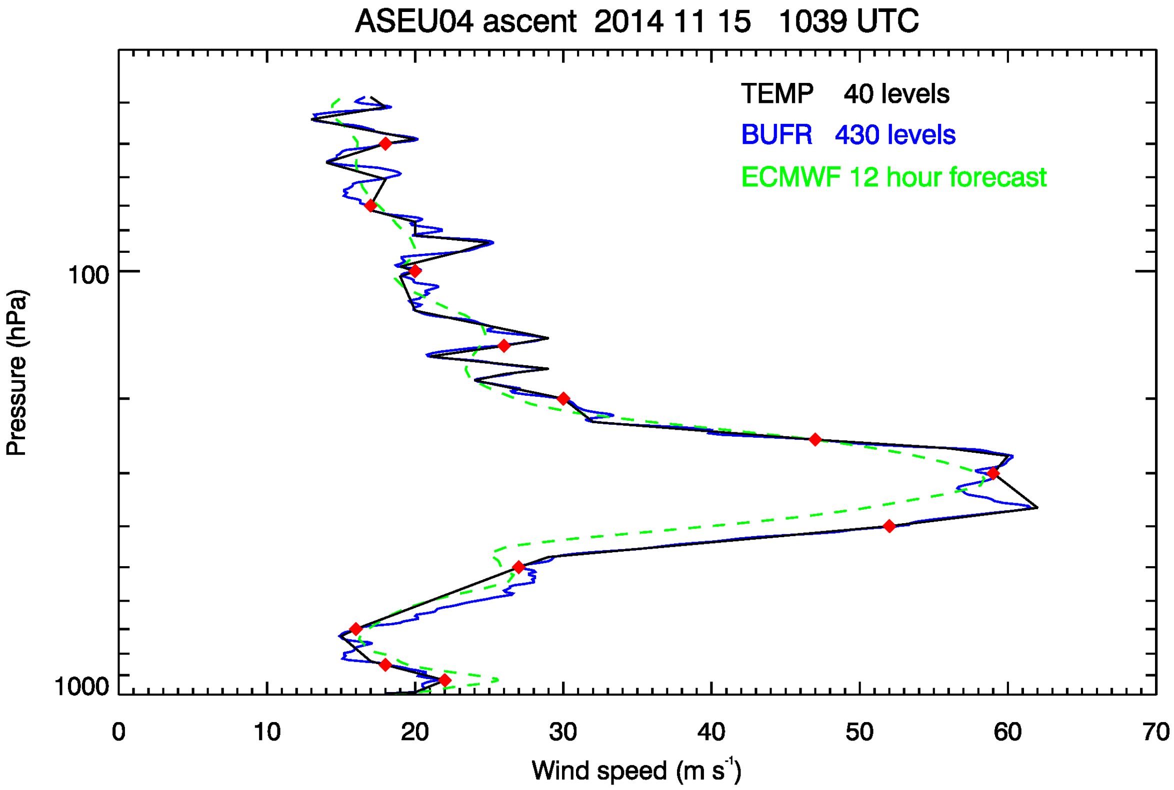
Geophysical Research Letters, Volume: 48, Issue: 4, First published: 06 December 2020, DOI: (10.1029/2020GL090699)

Number of aircraft measurements used at ECMWF



We can still improve the use of "established" observations, like radiosonde data:

BUFR radiosondes provide up to 8000 levels of measurements compared to less than 100 levels for TAC TEMP reports. A valuable improvement for data assimilation.



Bruce Ingleby, ECMWF

Accounting for radiosonde drift in data assimilation (we are improving the forward model H and reducing forward model error statistics, F)

- "*Old style*" radiosondes only provided the balloon **launch location**
- Native BUFR reports provides accurate location/time for each measurement
- The location/time information can be used to account for balloon drift in data assimilation

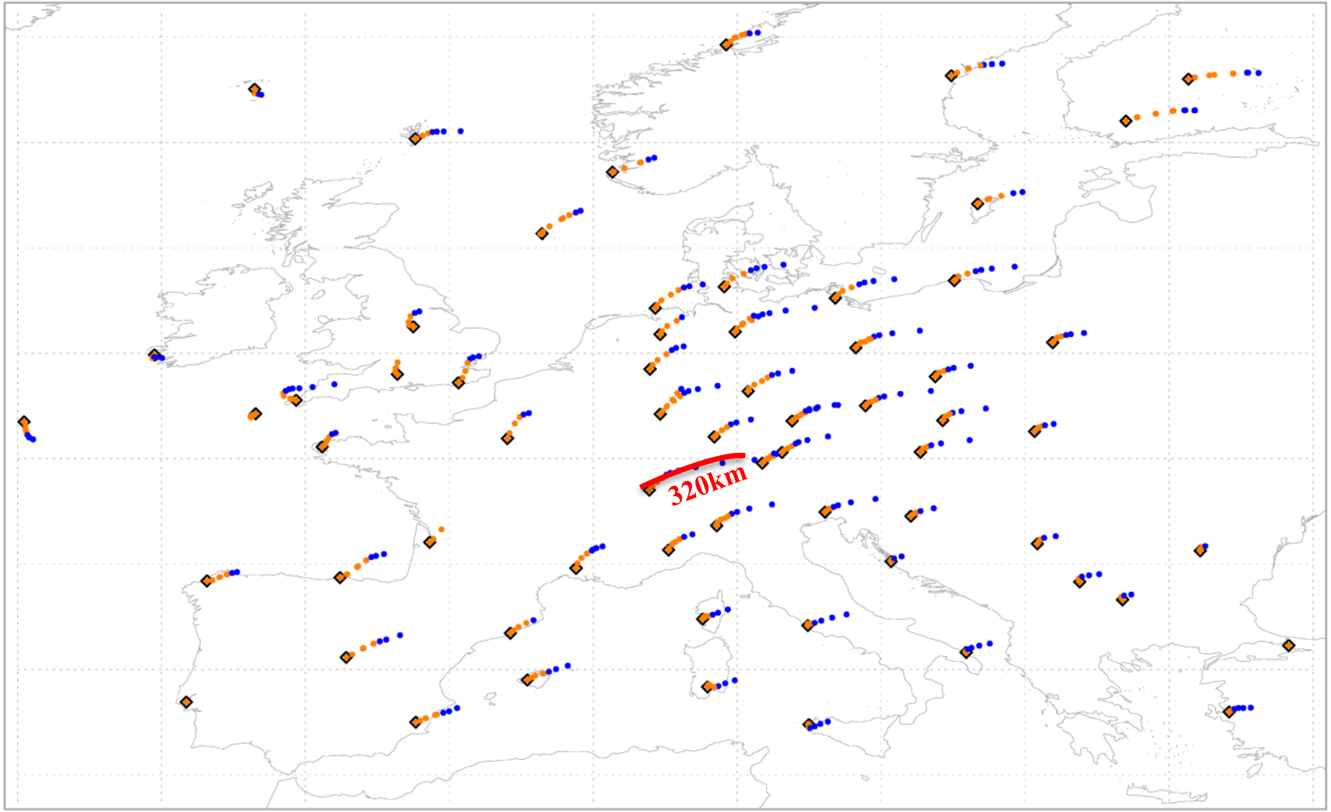
- We split the ascent into 15 minute chunks
- Was implemented at ECMWF in June 2018

- BUFR DROP (high-resolution dropsonde data was implemented at ECMWF in June 2019)
- In addition, descent data from BUFR radiosondes in Germany is now being used.

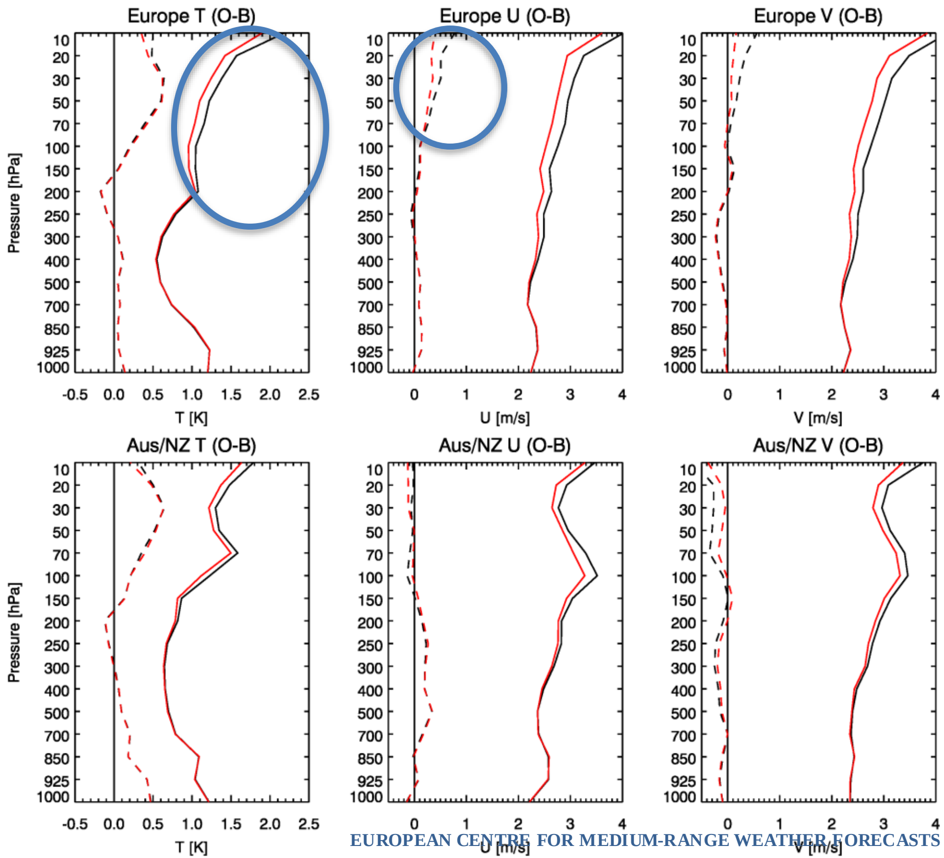
Example of large drift of radiosonde on a windy day

- Black diamonds – launch, levels to 100 hPa, levels above 100 hPa
- BUFR data not available for all countries at the time of this figure (Nov 2016)

2016-11-21 12 radiosonde drift (15 minute intervals)

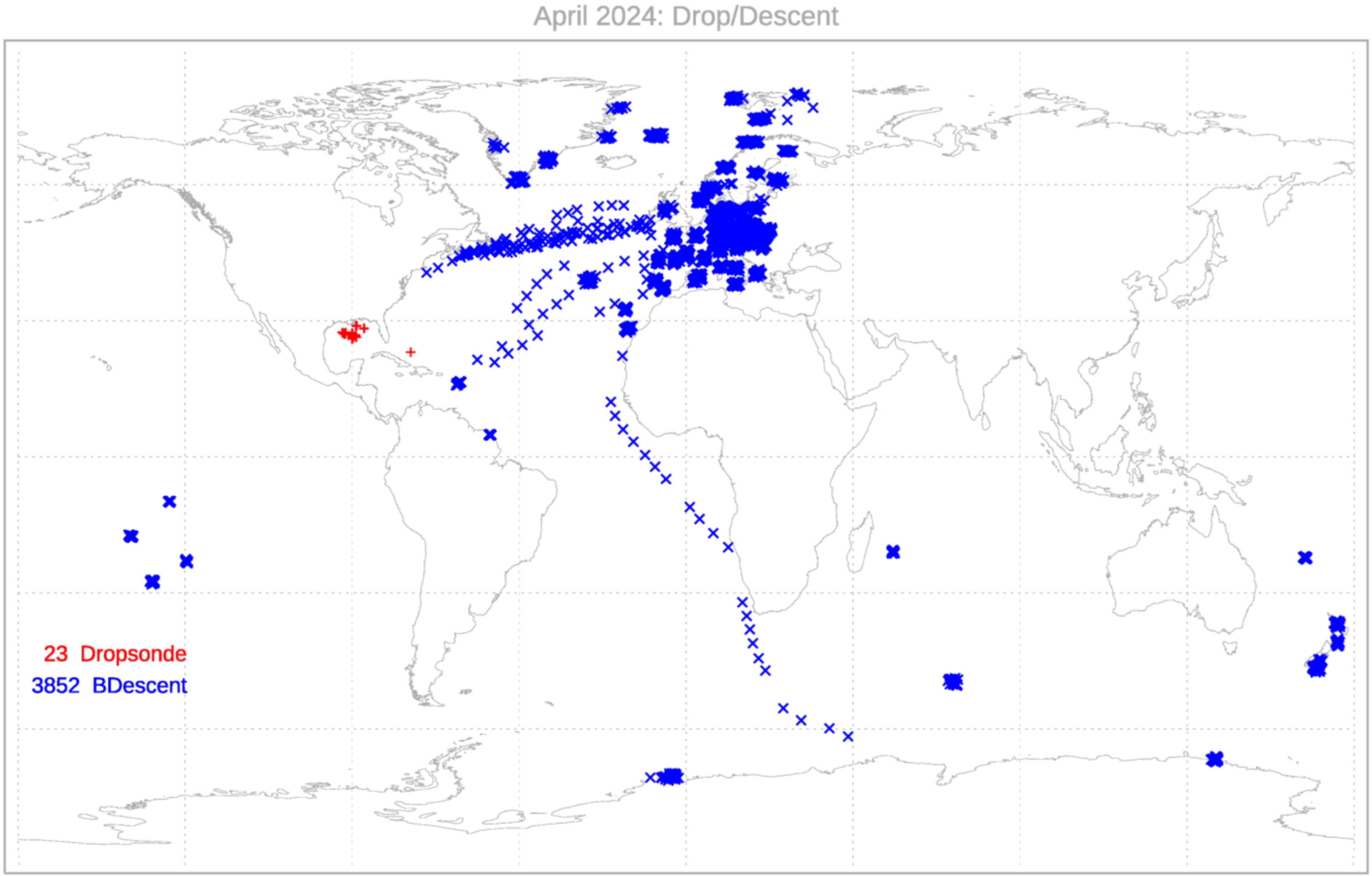


Impact of accounting for radiosonde drift in data assimilation
Mean and rms o-b statistics: Nov 2016



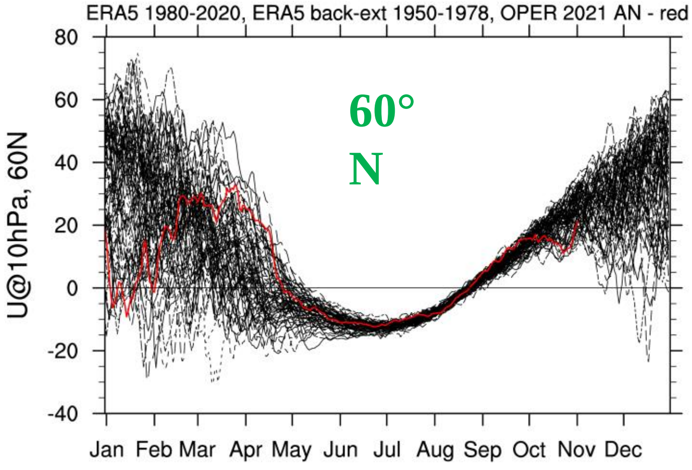
- Assimilated BUFR TEMP standard levels only (to get clean comparison)
- Good improvements at 200 hPa and above – including wind biases

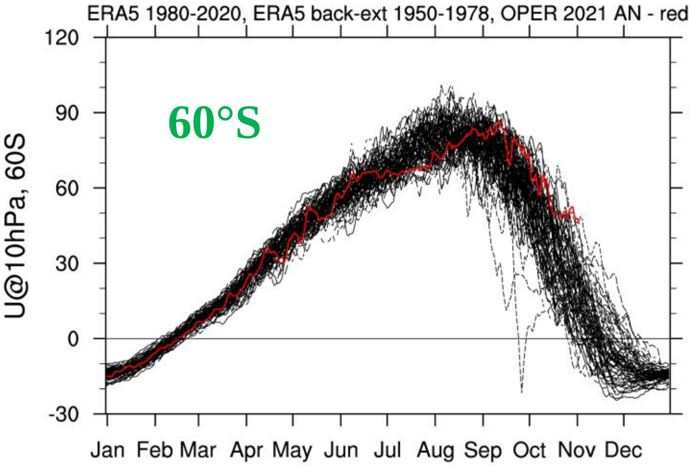
We now use sonde data from descents as well as ascent!



Stratospheric seasonality
- work with U10 hPa @ 60°N, 60°

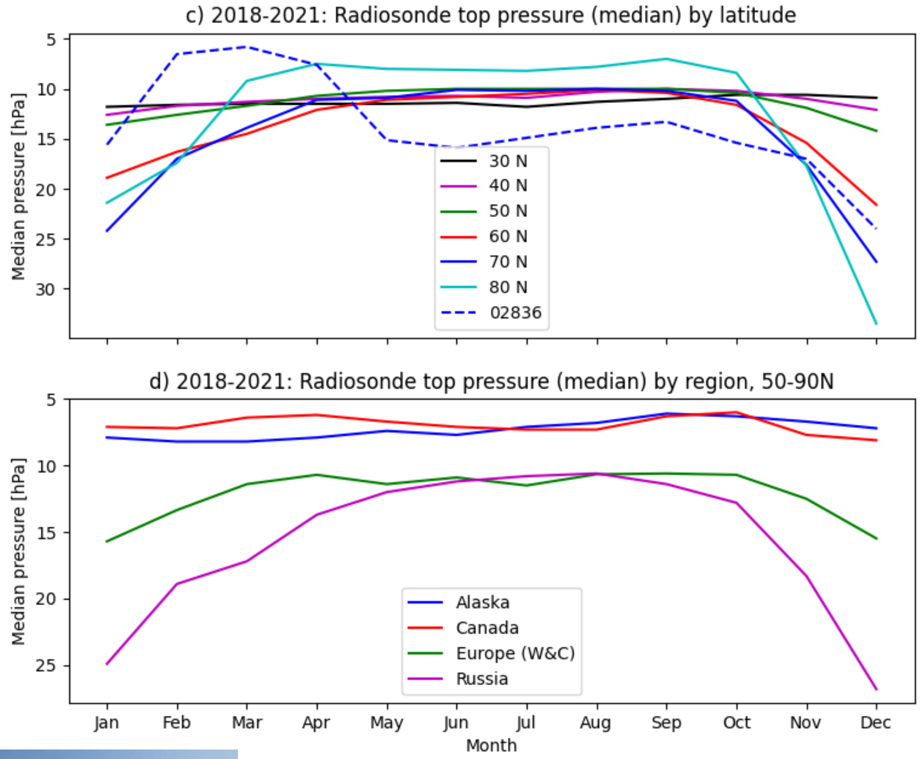
- Different years plotted 1950 to 2021
- Data from ERA5 (Hersbach et al, QJ, 2020)
- Winter: strong polar vortex
- SH vortex stable except when breaking down in Austral spring
- In NH winter planetary waves disturb the polar vortex
- Largest disturbances form stratospheric sudden warmings (SSWs)
- In summer there are about 4 months when nothing much happens – except a few gravity waves





Sondes more important in winter but they tend to burst at a lower height in winter!

- Seasonality in burst height is largest at hig latitudes
- Can see effect of larger balloon at Sodank (blue dashed line) – selected months
- Lower plot:
- Alaska: 600 g balloons
- Canada: 800 or 600 g
- Russia: 500 g
- Europe: varies, mainly 350 and 600 g
- Gas used will also affect height
- Question: Use bigger balloons in winter?



Some active satellite observation types

- **If you are on the NWP SAF training course, these observations are covered in much more detail!**

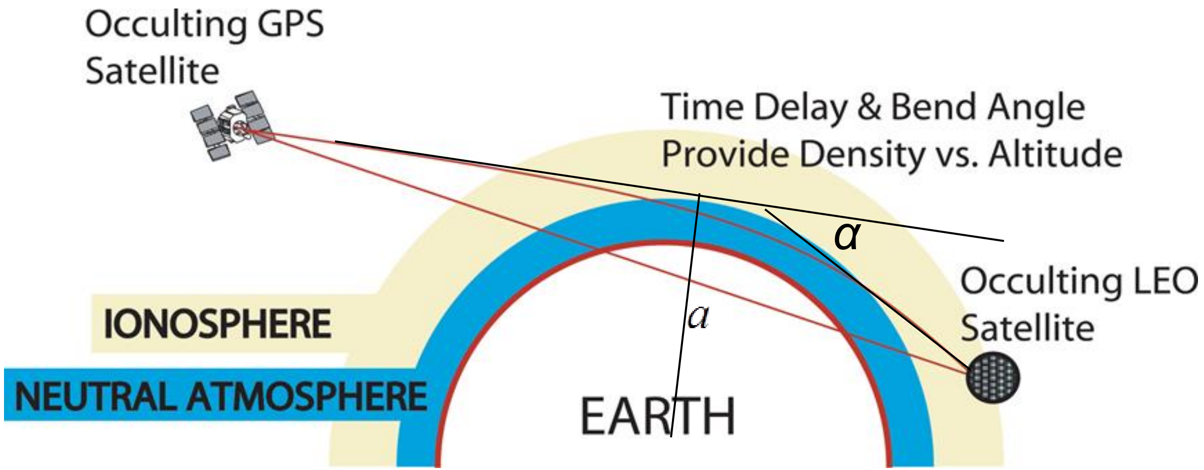
- More complicated forward operators, ***H***. Global datasets

  - GNSS Radio Occultation
    - Note that "ground-based GPS measurements" are **different**. They provides total column water information. Not covered here: EG, see, Bennitt, G. V., and A. Jupp, 2012: Operational Assimilation of GPS Zenith Total Delay Observations into the Met Office Numerical Weather Prediction Models. *Mon. Wea. Rev.*, **140**, 2706–2719, https://doi.org/10.1175/MWR-D-11-00156.1.

  - Scatterometer

  - Altimeter

Global Navigation Satellite System Radio Occultations
GNSS RO (GPS RO) geometry



As the LEO moves behind the Earth we obtain a profile of bending angles. The forward model $H(\mathbf{x})$ computes bending angle as a function of impact parameter (height), $\alpha(a)$.

The bending angle depends on temperature, humidity and pressure.

Global Navigation Satellite System Radio Occultations
GNSS RO (GPS RO) geometry

As the LEO moves behind the Earth we obtain a profile of bending angles. The forward model $H(\mathbf{x})$ computes bending angle as a

Tangent point

α

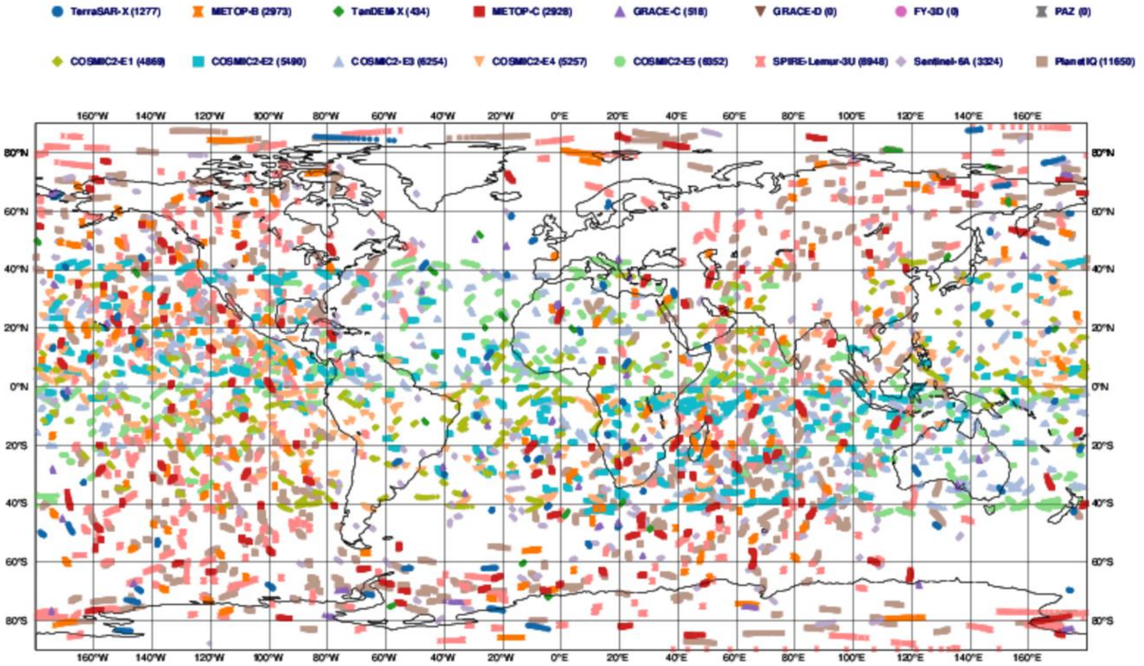
Key characteristics

- **Limb geometry means very good vertical resolution**
- **Can be assimilated without bias correction**

g

The bending angle depends on temperature, humidity and pressure.

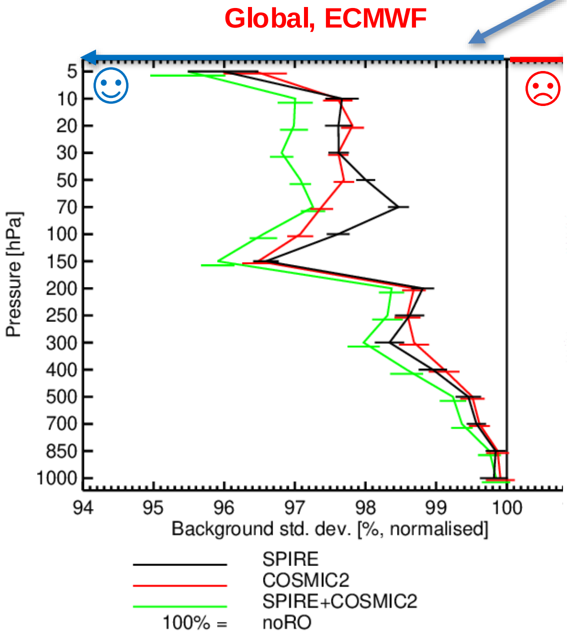
ECMWF data coverage (used observations) - GPSRO
2025031303 to 2025031309
Total number of obs = 60274

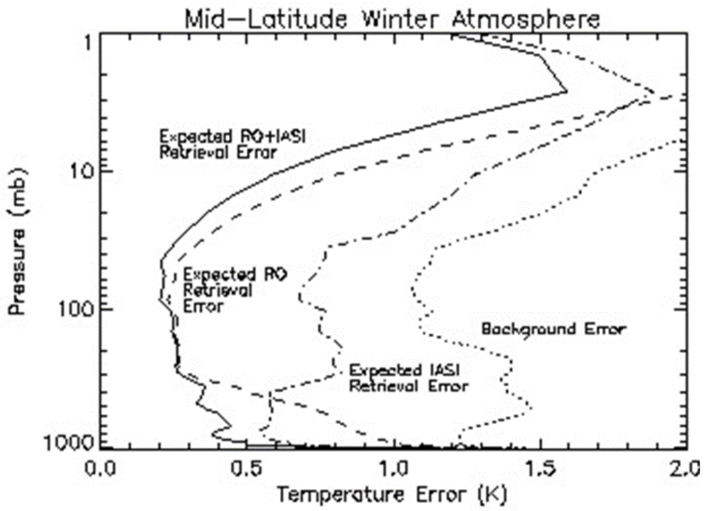


ECMWF

GNSS-RO has biggest impact in upper-troposphere/stratosphere
Fits to **radiosonde temperature** observations

Normalised standard deviation in (o-b) departure

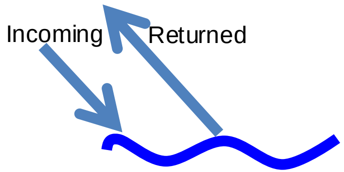




THEORY (2003)

Scatterometer

✓ A Scatterometer is an active microwave instrument (side-looking radar)
- Day and night acquisition
- Not affected by clouds



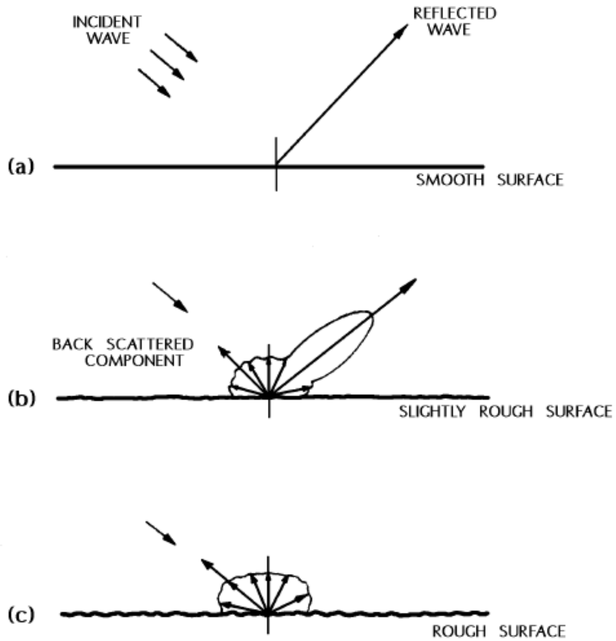
✓ The return signal, *backscatter* (σ_0 *sigma-nought),* is sensitive to:
- **Surface wind** (ocean)
- Soil moisture (land)
- Ice age (ice)

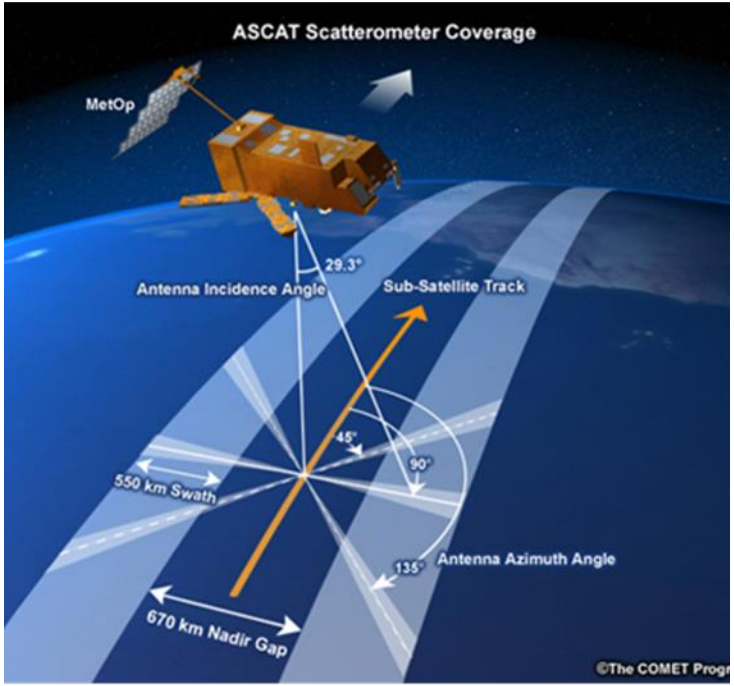
✓ Scatterometer was originally designed to measure ocean wind vectors:

- Measurements sensitive to the ocean-surface roughness due to capillary gravity waves generated by local wind conditions (surface stress)

- Observations from different look angles: wind direction

Dependency of the backscatter on... Wind speed (Bragg scattering)





EG, ASCAT

We measure be back scatter from three directions

- Fore/mid/aft

Triplet of backscatters used in a geophysical model function (GMF) to provide vector wind information.

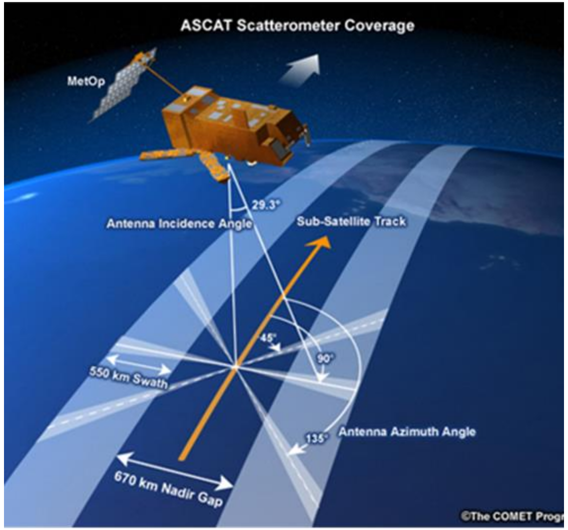
But the vector wind solutions are ambiguous!

How can we relate backscatter to wind speed and direction?

The relationship is determined empirically by developing a Geophysical Model Function (**GMF**)

- Ideally collocate with *surface stress* observations
- In practice with buoy and 10m model winds

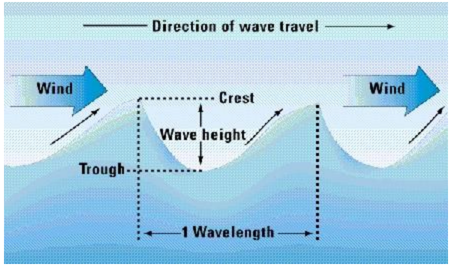
$$\sigma_0 = GMF(U_{10N}, \phi, \theta, p, \lambda)$$



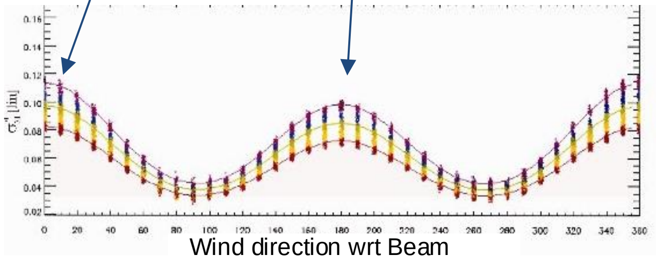
U_{10N}: *equivalent* neutral wind speed
ϕ : wind direction w.r.t. beam pointing
θ : incidence angle
p : radar beam polarization
λ : microwave wavelength

Dependency of the backscatter on... Wind direction

upwind
downwind

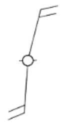


upwind
downwind

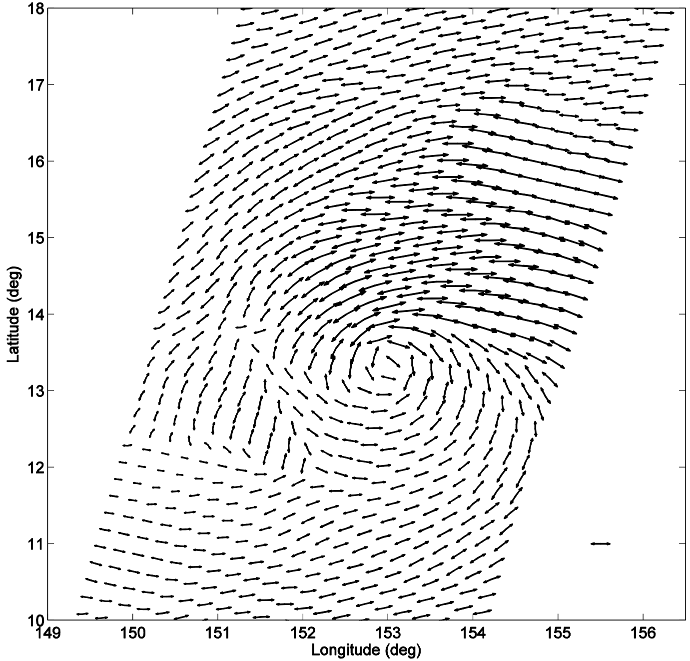


Wind Direction Ambiguity removal

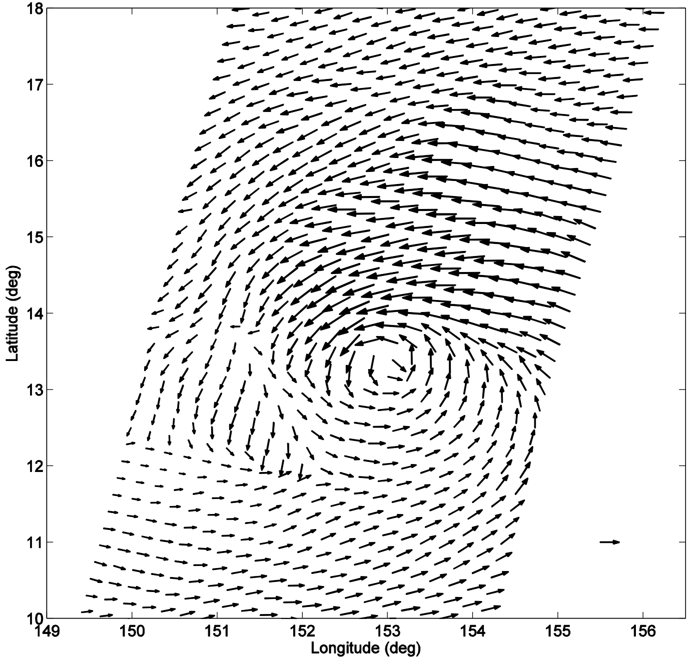
- **Each wind vector cell has usually two possible solutions for wind direction and speed**



- **The correct solution is determined during the 4D-Var**



Ambiguities provided



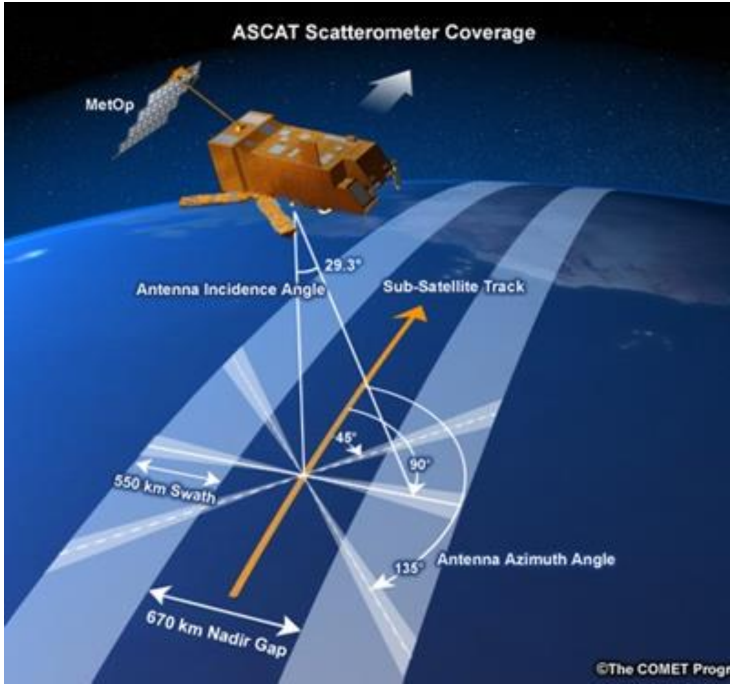
Ambiguities selected

Past, present and future scatterometers

Used on European platforms (1991 onwards):

✓ SCAT on ERS-1, ERS-2 by ESA
✓ ASCAT on Metop-B/C by EUMETSAT
✓ **SCAT on EPS-SG planned until 2040**

- Frequency ~5.3 GHz
- Wavelength ~5.7 cm
- Three antennae
    - Enables estimation of both wind speed and wind direction



Also, Chinese scatterometer data available now, including:

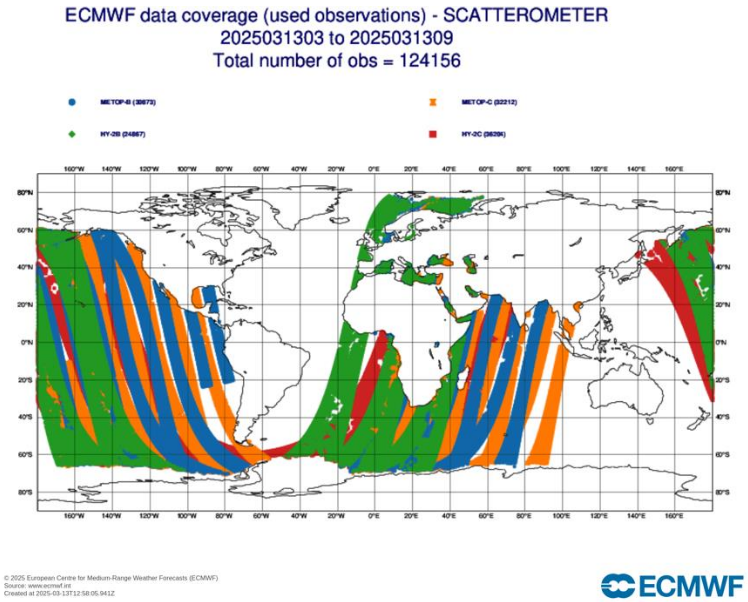
✓ HY-2B, HY-2C (HY-2D will be tested)

Why is Scatterometer important?

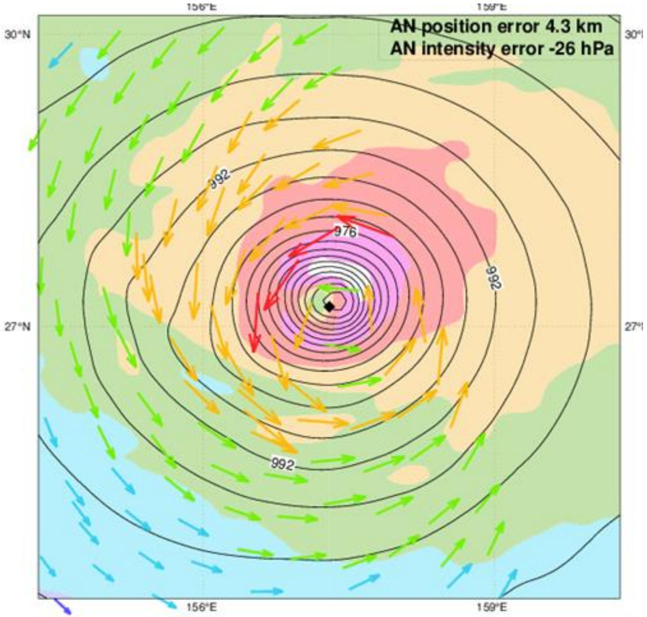
The scatterometer provides the ocean surface wind information (ocean wind vectors).

Ocean surface winds:

- affect the full range of ocean movement
- modulate air-sea exchanges of heat, momentum, gases, and particulates
- direct impact on human activities

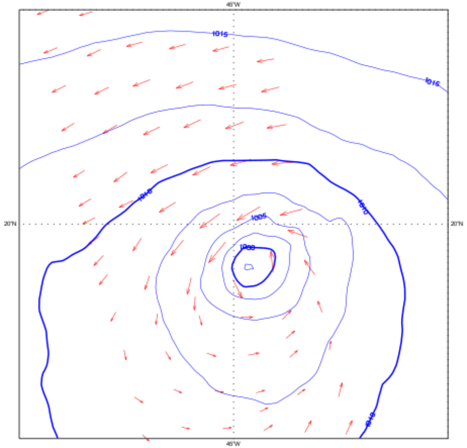


Important data source in tropical cyclones

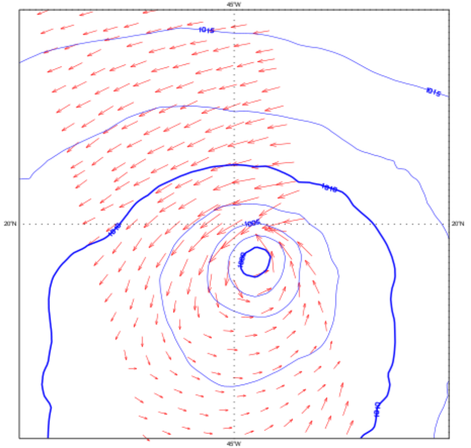


Some improvements in SCAT usage

- Increased the SCAT usage (reducing the thinning applied) in 2023



100 km thin



50 km (48R1, 2023)

- SCAT observation sensitive to the relative motion between the atmosphere and ocean
  - At the moment, we ignore the ocean current **but this will change in next operational cycle (50R1)**

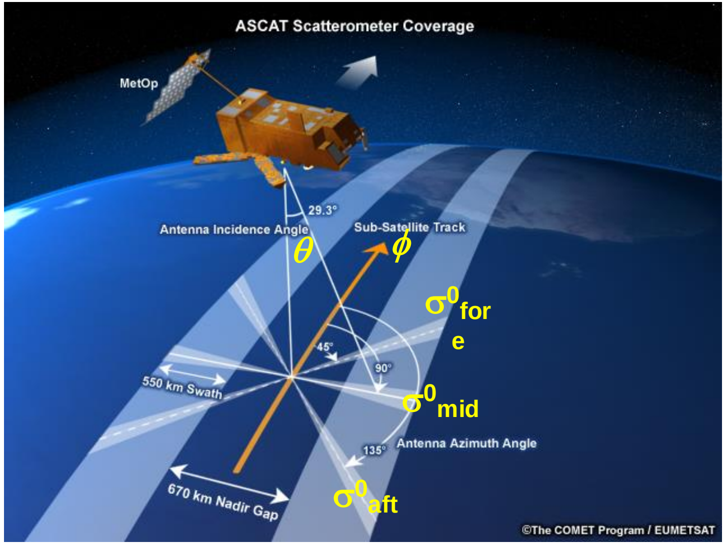
- Tested the direct assimilation of sigma0 – rather than assimilating ambiguous vector winds (more controversial)
  - we now handle non-linearity better in DA
  - **Revisit** the SCAT sigma0 problem and train a neural network to compute $\sigma_0 = GMF(U_{10N}, \phi, \theta, p, \lambda)$

SCATT Data Assimilation

Current approach

$\sigma^0_{fore}, \sigma^0_{mid}, \sigma^0_{aft}$
$\theta_{fore}, \theta_{mid}, \theta_{aft}, \phi$

→ U_{10N}, V_{10N} → IFS/4DVar



Plan

$\sigma^0_{fore}, \sigma^0_{mid}, \sigma^0_{aft}$
$\theta_{fore}, \theta_{mid}, \theta_{aft}, \phi$

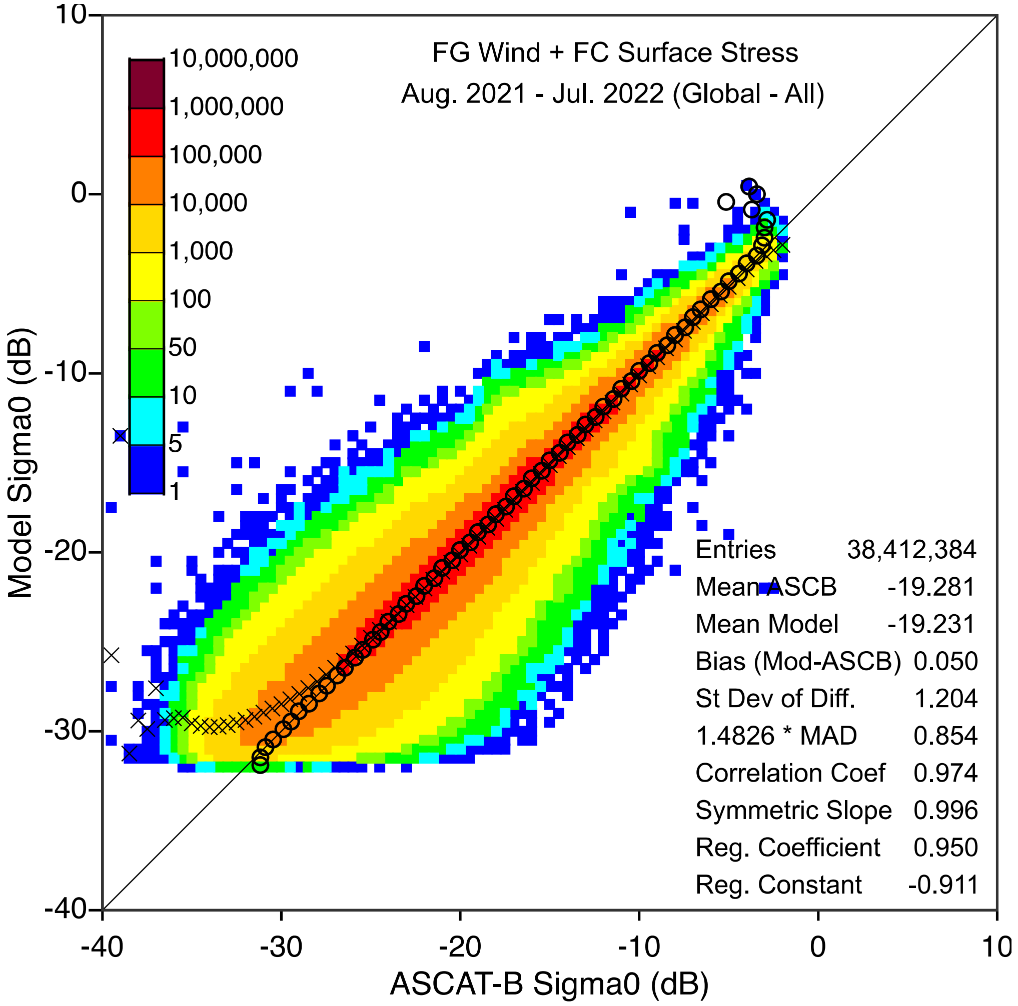
- sea-state,
- ocean currents,
- precipitation,
- … etc.

ML

→ IFS/4DVar

Training against model first guess (FG) wind

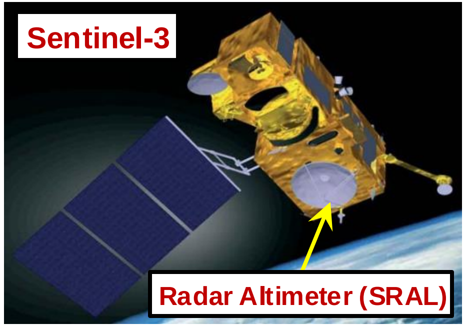
- **Can we assimilate sigma0 directly?**



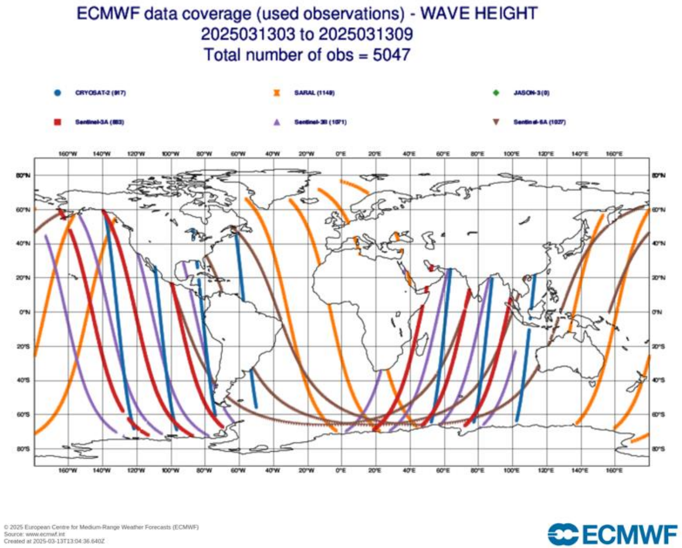
ECMWF

Radar Altimeters

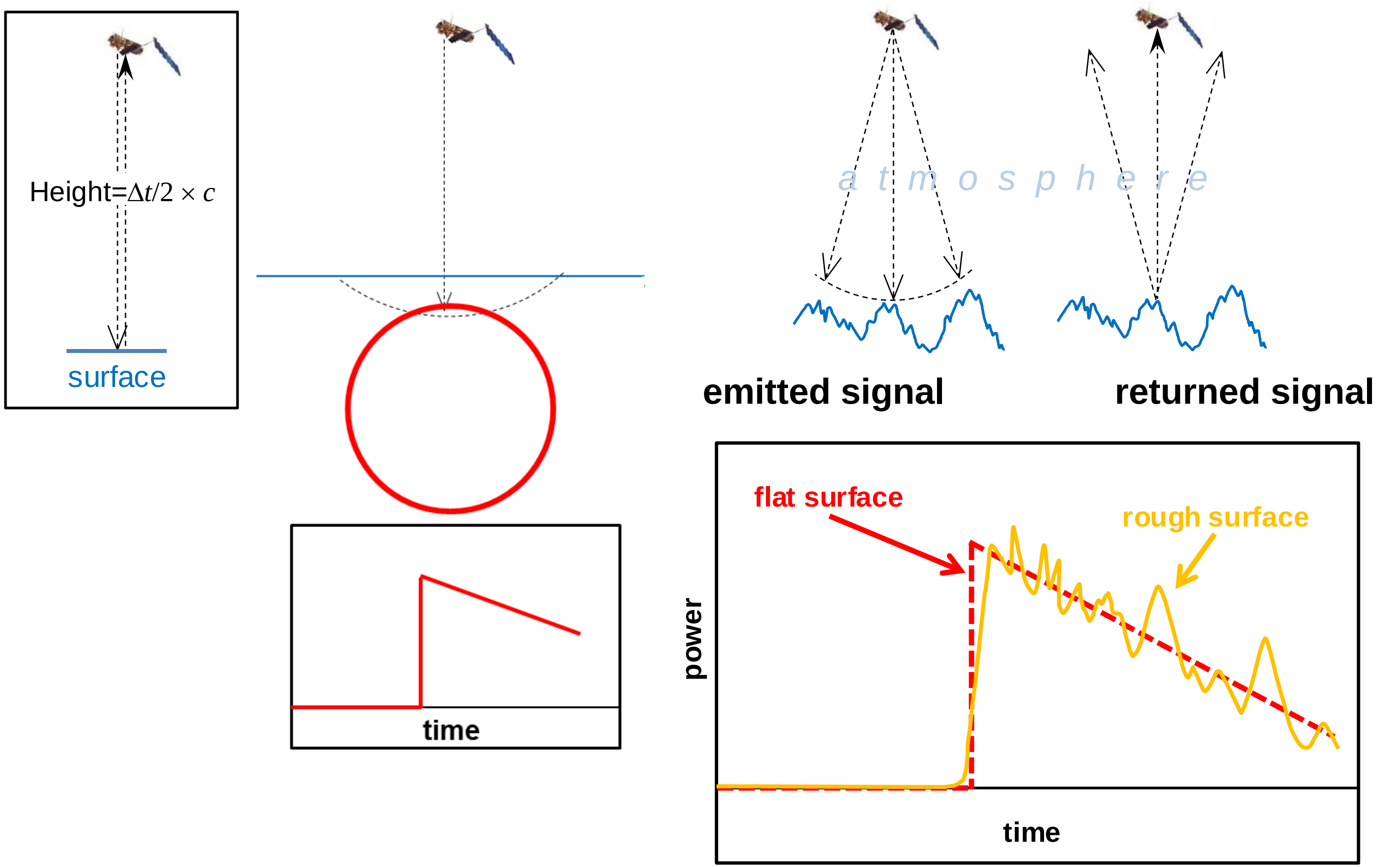
- ✓ Radar altimeter is a nadir looking instrument.
- ✓ Specular reflection.
- ✓ Electromagnetic wave bands used in altimeters:
  - Primary:
    - Ku-band (~ 2.5 cm) – Jason-3, Sentinel-3A/B/6
    - Ka-band (~ 0.8 cm) – SARAL/AltiKa (only example)
  - Secondary:
    - C-band (~ 5.5 cm) – Jason-3, Sentinel-3a,3b,6



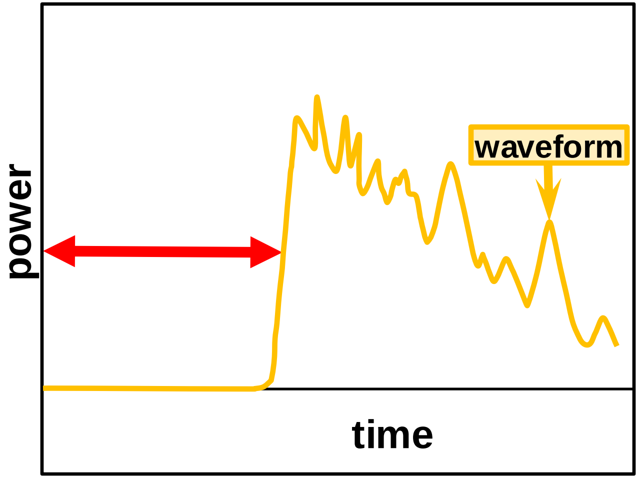
- ✓ Main parameters ***retrieved*** from an altimeter:
  - Sea surface height (*ocean model*)
  - Significant wave height (*wave model*) →
  - Wind speed retrievals (*used for verification*)



How Altimeter Works



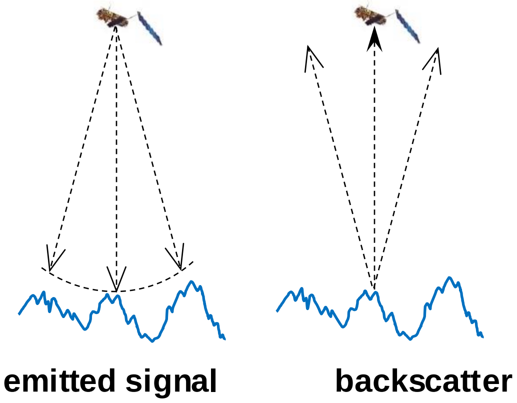
Sea Surface Height

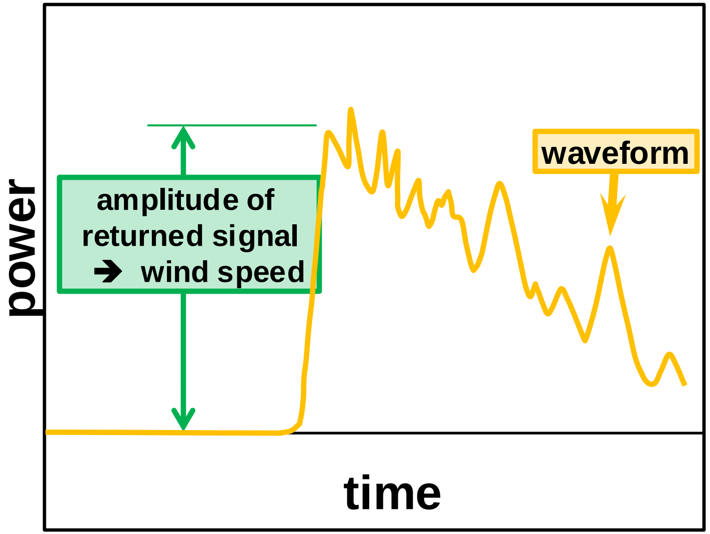


- ✓ Time delay → sea surface height
- ✓ Radar signal attenuation due to the atmosphere is caused by:
  - Water vapour impact: ~ 10's cm.
  - Dry air impact: ~ 2.0 m

  Correction made using radiometer and model data

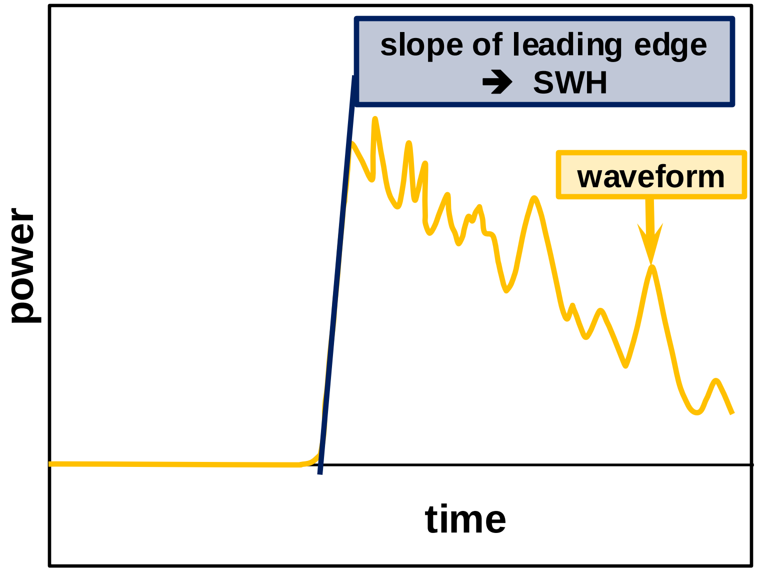
Surface wind speed





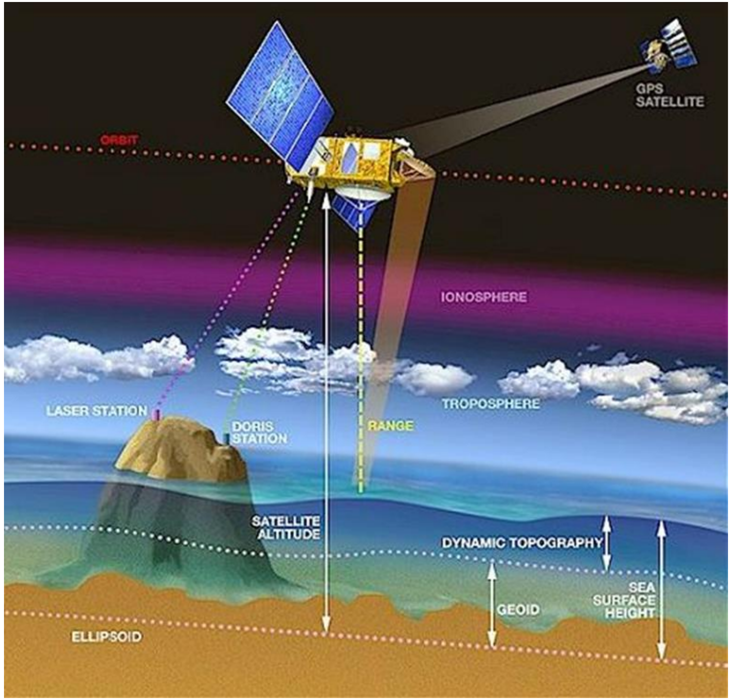
- ✓ Backscatter is related to water surface Mean Square Slope (MSS)
- ✓ MSS can be related to wind speed
- ✓ Stronger wind → higher MSS → smaller backscatter
- ✓ Errors are mainly due to algorithm assumptions, waveform retracking (algorithm), unaccounted-for attenuation & backscatter.

Significant Wave Height (SWH)



- ✓ SWH is the mean height of highest 1/3 of the surface ocean waves
- ✓ Higher SWH → smaller slope of waveform leading edge
- ✓ Errors are mainly due to waveform retracking (algorithm) and instrument characterisation.

Altimeter *corrections* applied to sea surface height

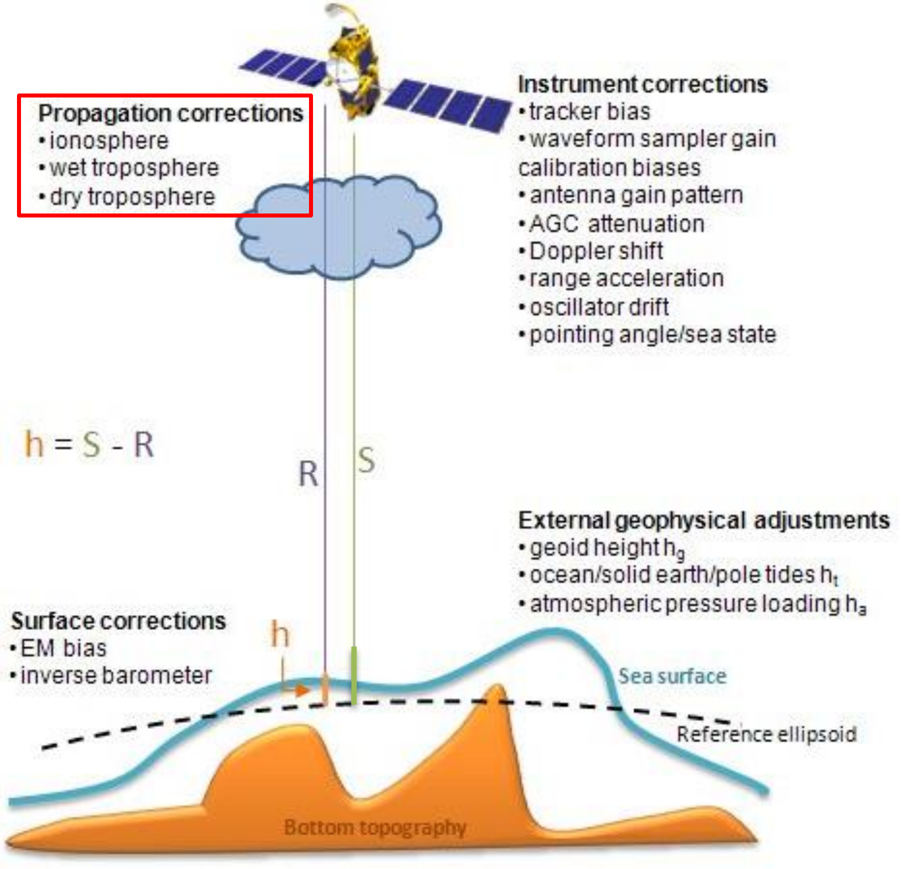


Sea Surface Height = Satellite altitude – Range - **Corrections**

ECMWF

Corrections to sea surface height measurements

- **Propagation corrections** – path delay of radar return signal due to:
  - **Ionosphere**: electron content of the atmosphere.
    - *Calculated by combining radar altimeter measurements acquired at two separate frequencies;*
    - *0 to 50 cm.*
  - **Wet troposphere**: cloud liquid water and water vapour in the atmosphere.
    - ***Retrieved** from **radiometer measurements** and/or estimated from meteorological models;*
    - *Correction ~ 0 to 50 cm.*
  - **Dry troposphere**: dry gases in the atmosphere.
    - *Calculated from meteorological models.*
    - *Related to surface pressure ~2.3 m.*

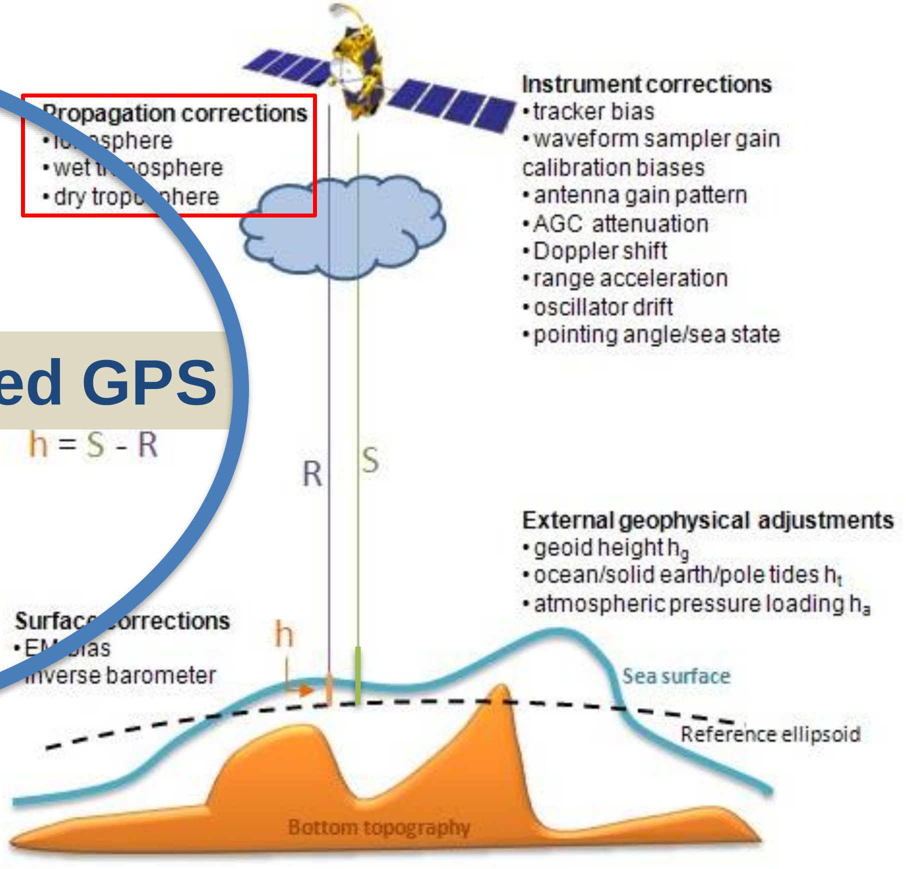


Corrections to sea surface height measurements

- **Propagation corrections** – path delay of radar return signal due to:
  - **Ionosphere**: electron content of the atmosphere.
    - *Calculated by combining radar altimeter measurements acquired at two separate frequencies;*
    - *0 to 50 cm.*
  - **Wet troposphere**: cloud liquid water and water vapour in the atmosphere.
    - ***Retriev*** ***radiometer*** *meteor* *models;*
    - *Correct* *o 50 cm.*
  - **Dry troposphere**: dry gases in the atmosphere.
    - *Calculated from meteorological models.*
    - *Related to surface pressure ~2.3 m.*

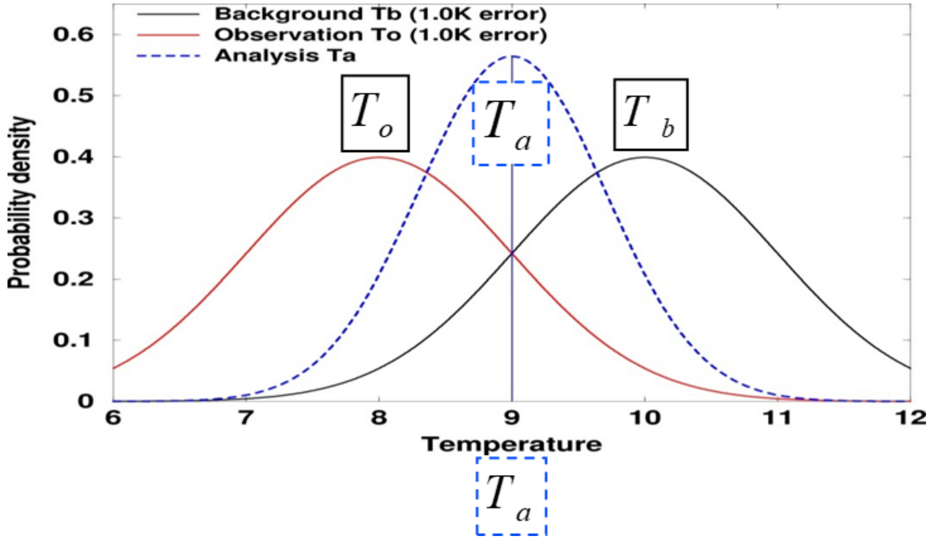
+

= ground-based GPS



Quality Control (QC)
Really important in DA methodology – but getting squeezed as training course grows

QC: The linear scalar temperature problem



- Assume the **standard deviation** of the background and observation errors are 1 K. The **assumed error statistics** determine the "gain matrix", **K**.

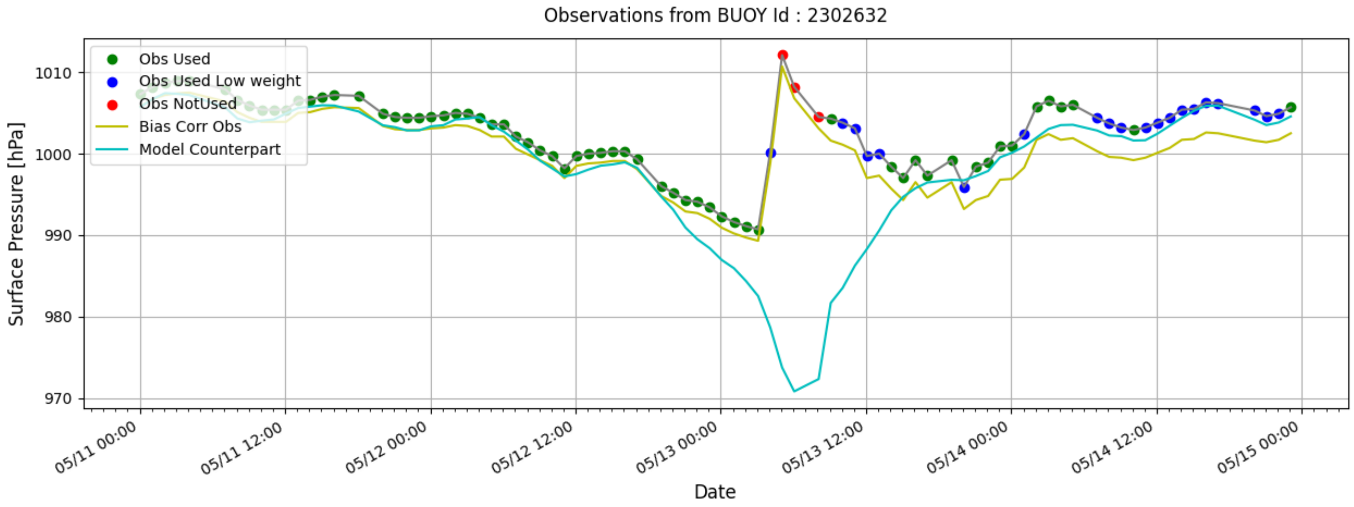
- If these errors are uncorrelated, the st. dev. of $(T_o - T_b)$ differences should be **about** $\sqrt{2}$K.

- All observations have errors – we accept that (**R** matrix). But what should we make of a difference of, ***say***, $(T_o - T_b) > 20$ K? The **actual errors** in this case are probably not consistent with the error statistics we've assumed in the **K** matrix.

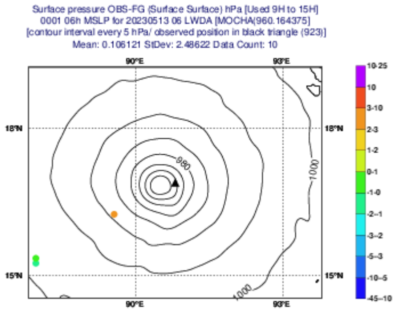
Large departures can be caused by …

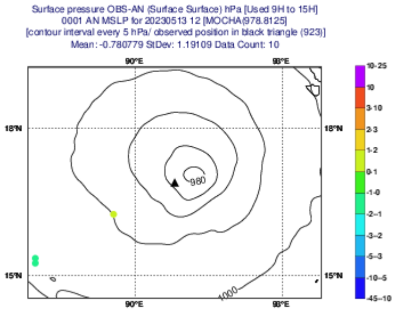
- **Either** the observation errors are large **or** the background (forecast) errors are large

- **A real example that caused problems at ECMWF: TC Mocha May 13 2023**

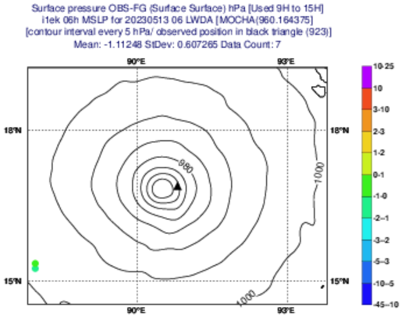


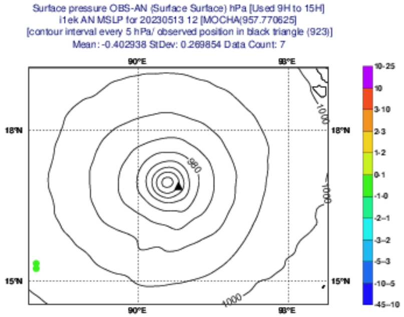
TC Mocha





Operations





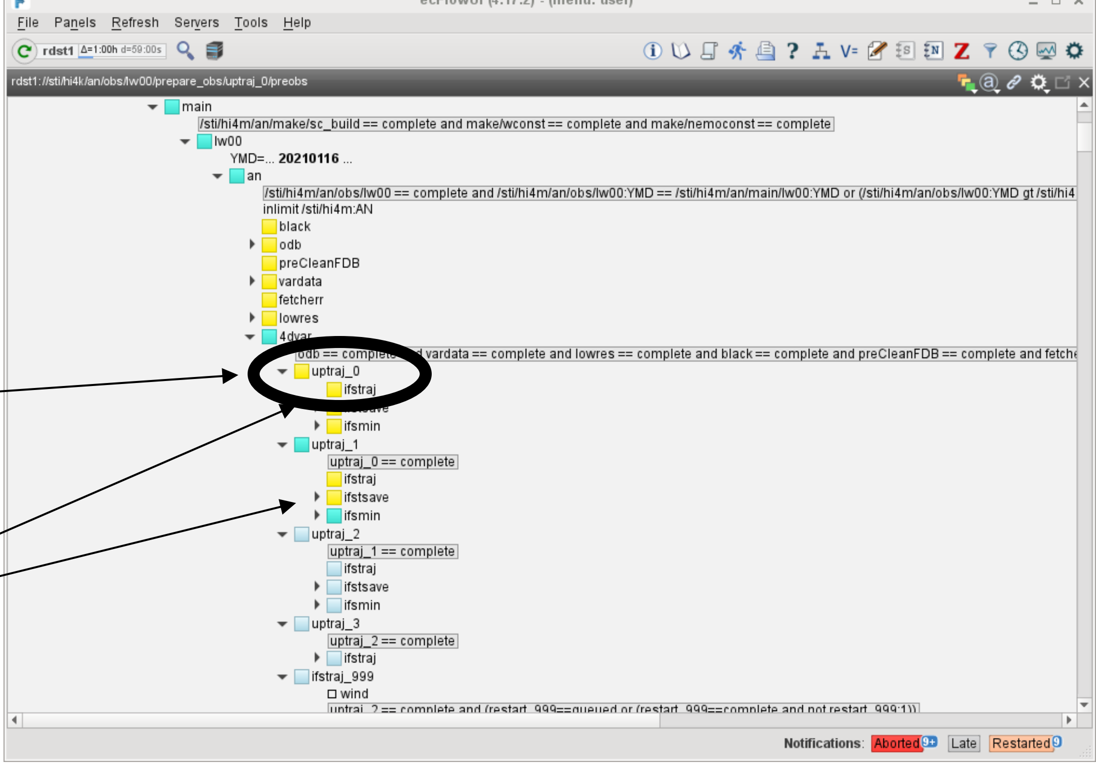
Remove ob

QC steps

- The “**first guess check**” should remove **really bad** data in our *1st trajectory*

- Then also rely on ***Variational QC*** and the **Huber norm** additional QC from the 1st trajectory to “down weight” the data if necessary

- Data rejected by first-guess check has **gone – it can’t come back!** But with VarQC/Huber, data can get more weight later if supported by other observations



Q. J. R. Meteorol. Soc. (1999), **125**, pp. 697–722

Variational quality control

By ERIK ANDERSSON* and HEIKKI JÄRVINEN
European Centre for Medium-Range Weather Forecasts, UK

What is the probability of an (o-b) of this size given R and B?
Normal departures and gross errors have different distributions

$$p^{QC} = (1 - A)N + Ap^{G}$$

The *a priori* probability of gross error

Assumed distributions

- The gross errors have a flat distribution

$$p^G = \frac{1}{2d}$$

- The ordinary departures a normally distribruted

$$N = \frac{1}{\sigma_o\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y - Hx}{\sigma_o}\right)^2\right]$$

Take –ln(P^{QC})=J_O^{QC}

$$J_o^{\mathrm{QC}} = -\ln\left[\frac{\gamma + \exp(-J_o^{\mathrm{N}})}{\gamma + 1}\right]$$

$$\nabla J_o^{\mathrm{QC}} = \nabla J_o^{\mathrm{N}}\left[1 - \frac{\gamma}{\gamma + \exp(-J_o^{\mathrm{N}})}\right]$$

$$\text{with } \gamma \text{ defined as}: \gamma = \frac{A\sqrt{2\pi}}{(1-A)2d}$$

Take $-\ln(P^{QC}) = J_O^{QC}$

$$J_o^{\mathrm{QC}} = -\ln\left[\frac{\gamma + \exp(-J_o^{\mathrm{N}})}{\gamma + 1}\right]$$

$$\nabla J_o^{\mathrm{QC}} = \nabla J_o^{\mathrm{N}}\left[1 - \frac{\gamma}{\gamma + \exp(-J_o^{\mathrm{N}})}\right]$$ **=1-PGE**

$$\text{with } \gamma \text{ defined as}: \gamma = \frac{A\sqrt{2\pi}}{(1-A)2d}$$

So, we <u>weight</u> the (o-b) departures by 1 minus the Probability of Gross Error (PGE). The *a priori* PGE, *A*, is updated based on the size of the (o-b) departure using *Bayes Theorem*!

The large (o-b) of 20 K in our scalar example would be multiplied by (1-PGE)

$$\nabla J_o^{QC} = \nabla J_o^{N}\left[1 - \frac{\gamma}{\gamma + \exp(-J_o^{N})}\right] \quad \text{=1-PGE}$$

$$\text{with } \gamma \text{ defined as} : \gamma = \frac{A\sqrt{2\pi}}{(1-A)2d}$$

In recent years we have also used the Huber norm

Quarterly Journal of the Royal Meteorological Society *Q. J. R. Meteorol. Soc.* 141: 1514–1527, July 2015 A DOI:10.1002/qj.2440

Royal Meteorological Society

On the use of a Huber norm for observation quality control in the ECMWF 4D-Var

Christina Tavolato[a,b] and Lars Isaksen[a*]

[a]*European Centre for Medium-Range Weather Forecasts, Reading, UK*

[b]*Department of Meteorology and Geophysics, University of Vienna, Austria*

*Correspondence to: L. Isaksen, ECMWF, Shinfield Park, Reading RG2 9AX, UK.
E-mail: lars.isaksen@ecmwf.int

The Huber norm is less conservative than VarQC

$$f(x) = \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left\{-\frac{\rho(x)}{2}\right\} \qquad (1)$$

with

$$\rho(x) = \begin{cases} \dfrac{x^2}{\sigma_o^2} & \text{for} \quad |x| \leq c\,, \\ \dfrac{2\,c|x| - c^2}{\sigma_o^2} & \text{for} \quad |x| > c\,, \end{cases} \qquad (2)$$

$x = y - H(\mathbf{x})$, the (o-b) in our terminology/notation!

The Huber norm is less conservative than VarQC

$$f(x) = \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left\{-\frac{\rho(x)}{2}\right\}$$

with

$$\rho(x) = \begin{cases} \dfrac{x^2}{\sigma_o^2} & \text{for} \quad |x| \leq c\,, \\ \dfrac{2\,c|x| - c^2}{\sigma_o^2} & \text{for} \quad |x| > c\,, \end{cases}$$

Derived from departure statistics
Can be asymmetric either side of peak.

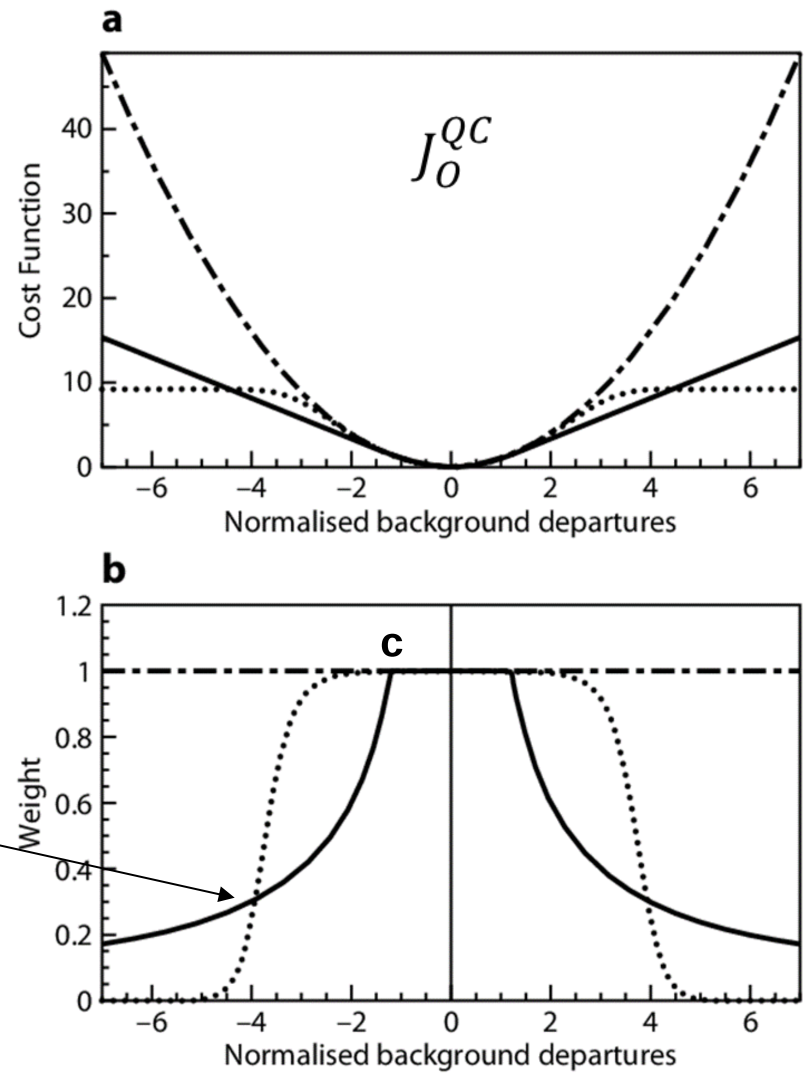
$x = y - H(\mathbf{x})$, the (o-b) in our terminology/notation!

COST function + weight

No QC: Gaussian
Solid line: Huber norm
Dotted line: “VarQC”

Huber norm gives more weight than VarQC in the “wings”

Should we be more conservative and revert to VarQC?



Summary

- Aim of data assimilation is to retrieve as much information from observations as possible and provide good initial conditions for the forecast model. We need
  - observation operator, $H(\mathbf{x})$
  - estimate of observation error statistics to provide the weighting, $\mathbf{R}$
- Impact of in-situ and actively sensed observations in global NWP
  - Impact of the data types, how we assimilate the data
  - We continue to develop and improve our use of in-situ data
- Quality control – a vital part of DA methodology
  - introduced the VarQC and Huber norm approach used at ECMWF
  - We need to screen out cases when their errors are not consistent with the $\mathbf{R}$ we assume