Towards higher spatial and temporal resolution data assimilation in ECMWF IFS

Ž. Zaplotnik and E. Orlandi
On behalf of colleagues from ESAS and DestinE

ECMWF

Outline

- Towards More Frequent Analysis Updates
  - Extending window
  - Continuously update analysis with latest observations available

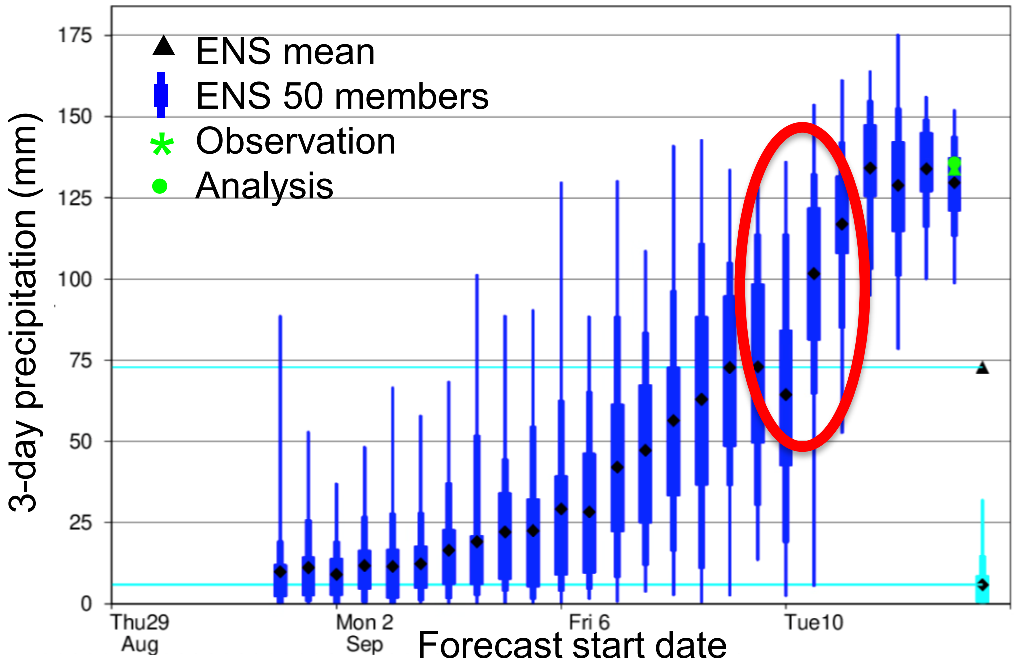
- Towards Higher Resolution Data Assimilation
  - Impact on initial conditions
  - Impact on predictability
  - Costs

Towards More Frequent Analysis Updates

- **Global analysis** and forecast are used to **initialize** and provide **boundary conditions** to high resolution Local Area Models.
- Evolution of LAM forecast is toward more frequent updates.
- Many downstream applications benefit from more frequent analysis updates
- Prediction of **extremes:**
  - Extreme cases with low predictability → timely analysis using latest observations.
  - Increase forecaster confidence in issuing alerts.

- **Hourly updates** in LAM require updated global boundary condition to extend their forecast validity.
- Most **WMO World Meteorological Centers** provide analysis updates 2 times a day.

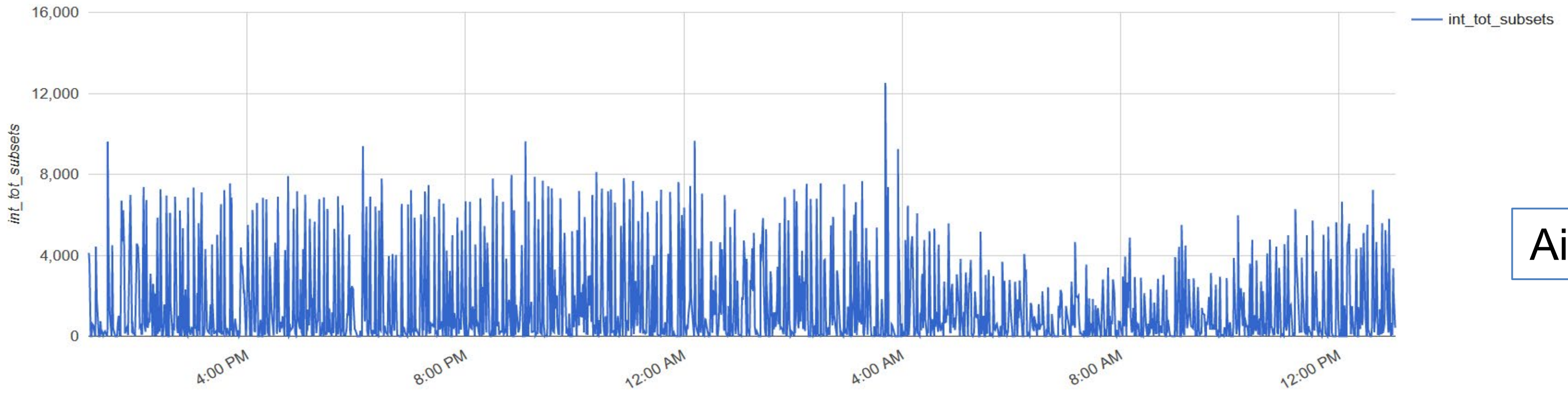
2024.09.13-16 Storm Boris



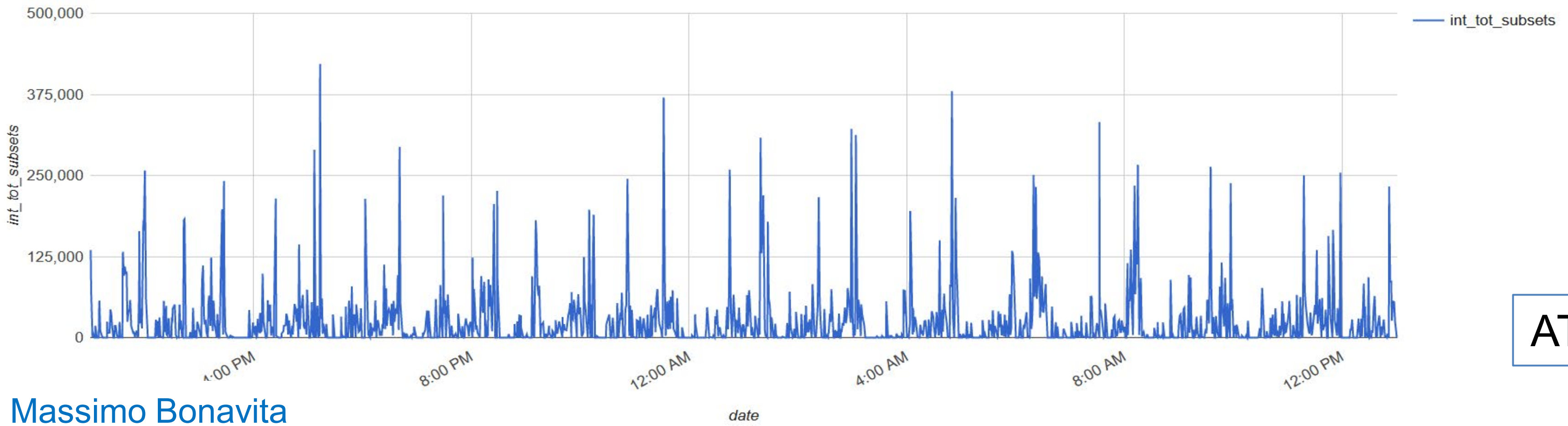
Linus Magnusson

Continuous inflow of observations

Timeline of number of obs received by the ECMWF acquisition and pre-proc. system (SAPP)



Aircraft obs.



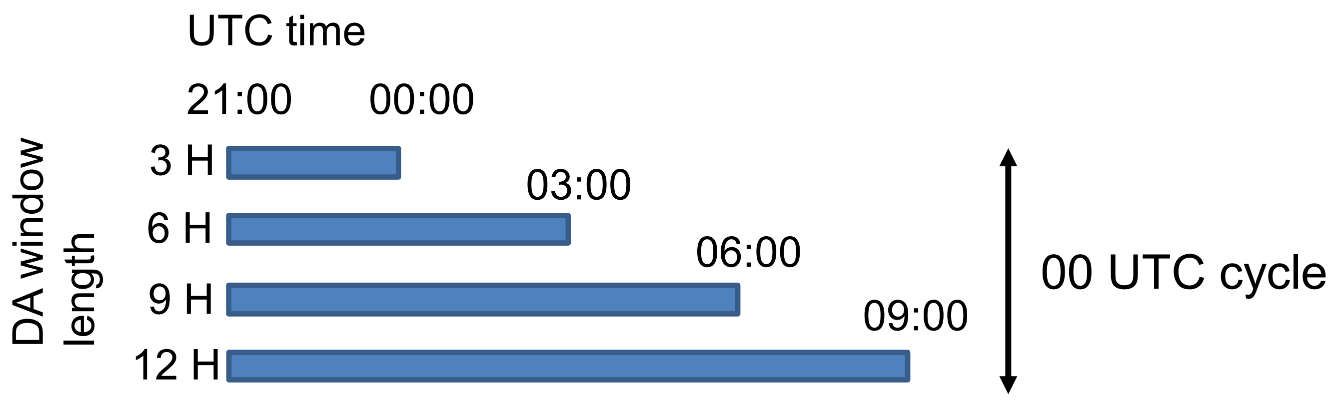
ATOVS rad.

Massimo Bonavita

Extending window DA

DA system with configurable window length

Within the same DA cycle, minimise the cost function on progressively longer windows



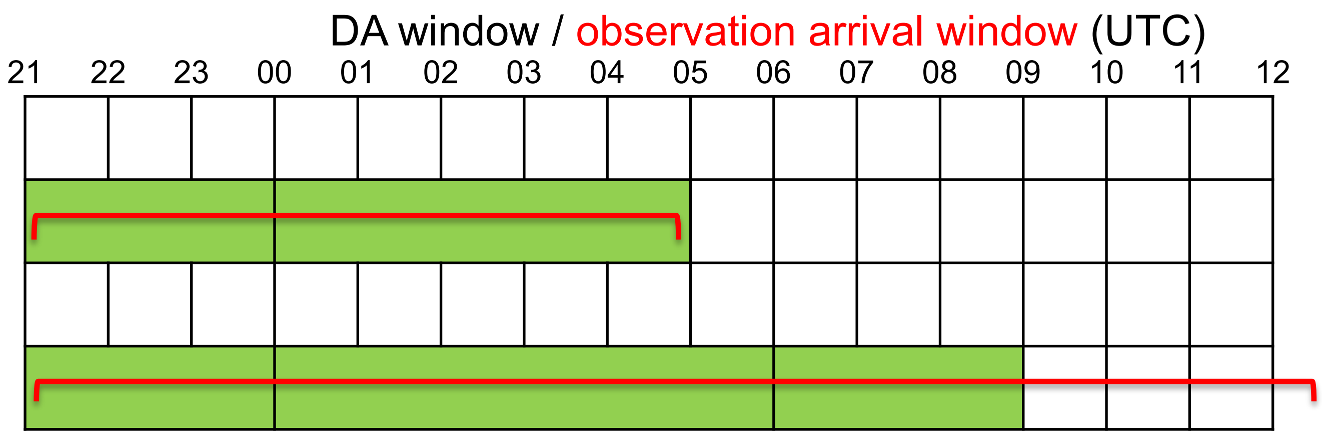
Each **window** produces an **updated analysis** using the **latest observations** available
The analysis produced in (*n-1*) window is used as **first guess** for next analysis update (*n*)
Background is the same for all analyses

Overarching objectives
Increase analysis update frequency
Improve forecast scores
Uniformly distribute use of computational resources

Experiment set-up

name	Trajectory resolution	DA window length (H)	Analyses / forecasts	Outer loop count	Outer loop 1	Outer loop 2	Outer loop 3	Outer loop 4
Control	TCo1279 (9 km)	8/12	2 / 1	8	TL255 (~80km)	TL319	TL399	TL511 (~40km)

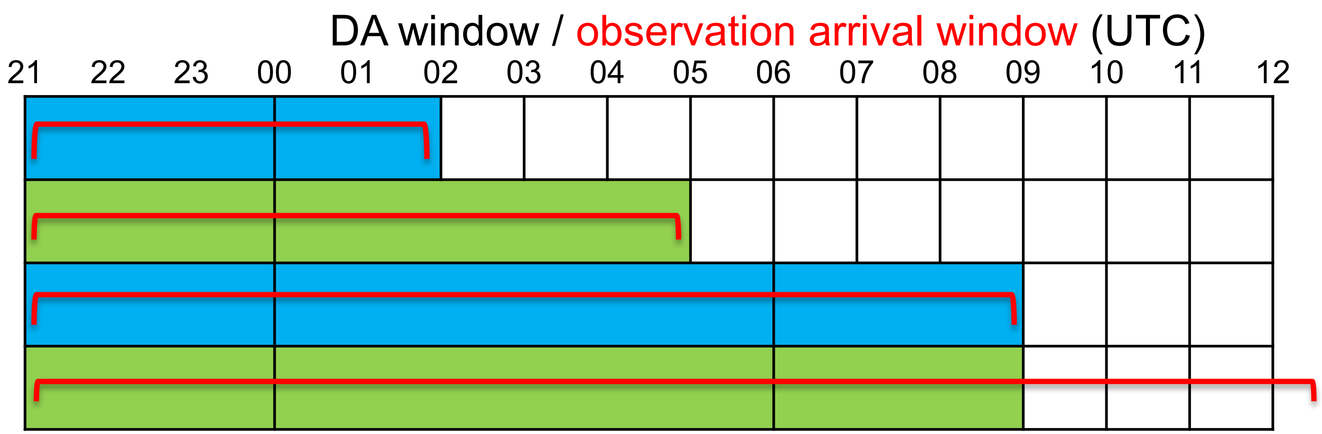
Ext Win	Contr	Window lengh (H)
	√	8
	√	12



Experiment set-up

name	Trajectory resolution	DA window length (H)	Analyses / forecasts	Outer loop count	Outer loop 1	Outer loop 2	Outer loop 3	Outer loop 4
Ext Win	TCo1279 (9 km)	5/8/12/12	4 / 3	16	TL255	TL319	TL399	TL511
Control	TCo1279 (9 km)	8/12	2 / 1	8	TL255 (~80km)	TL319	TL399	TL511 (~40km)

Ext Win	Contr	Window lengh (H)
✓	✗	5
✓	✓	8
✓	✗	12
✓	✓	12



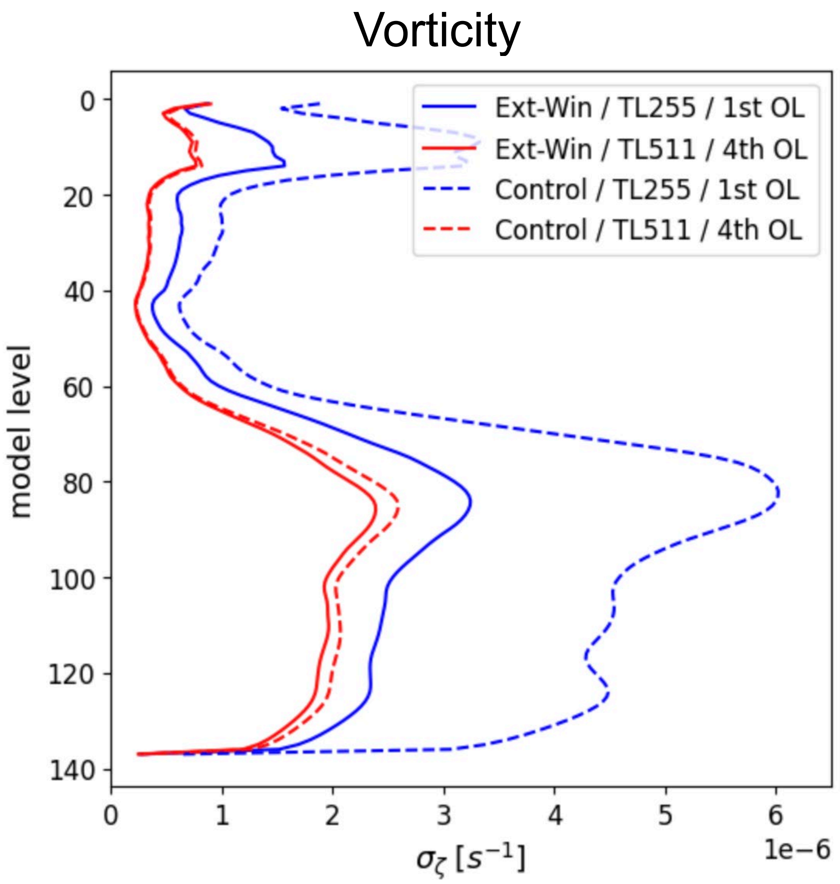
Global STDEV of analysis increments

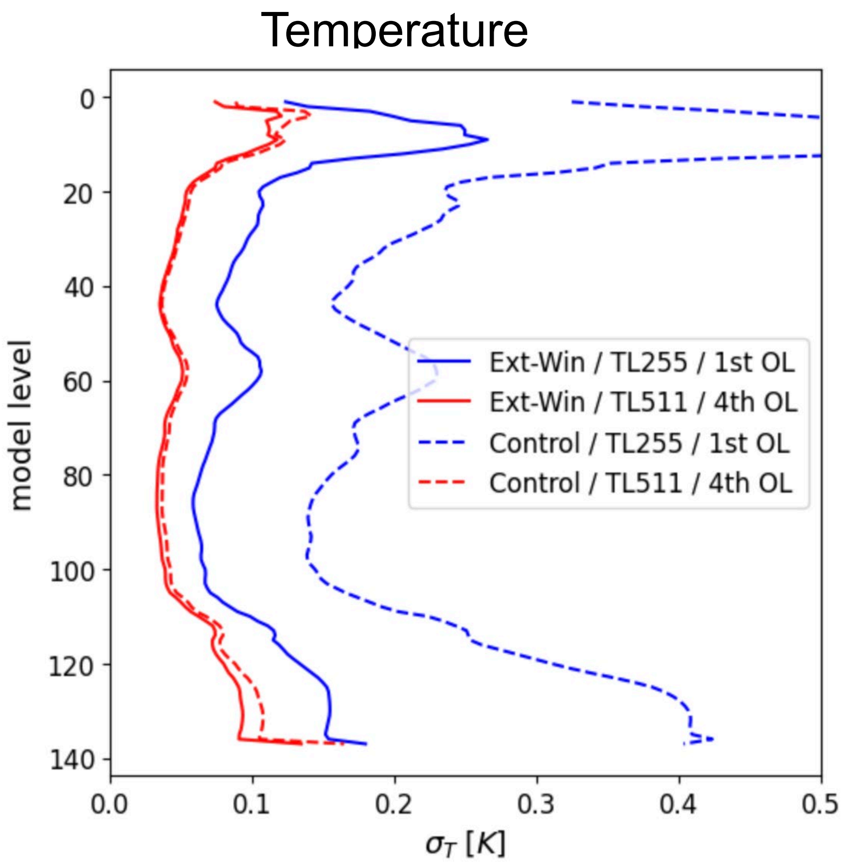
8H window analysis

- Measure the size of the increments using their **STDEV over the globe**
- **Ext-Win** uses **first guess** from the **5H window** analysis

Ext-Win

- Smaller adjustments to the estimated state are needed in successive updates
- Shadows reality closely and continuously
- Drop TL255 outer loop? → reduce computational footprint



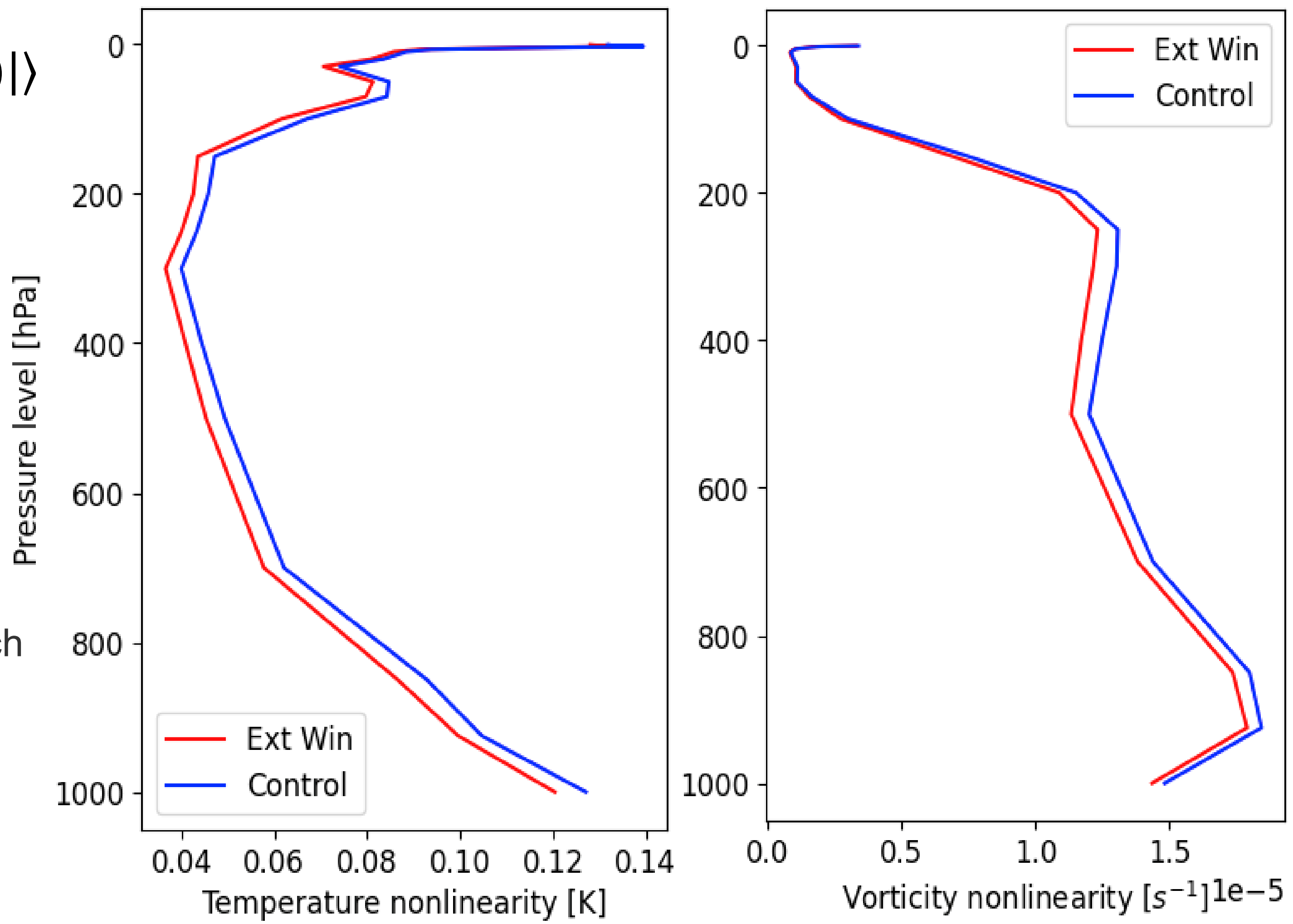


Trajectory non-linearity: 12H window analysis

Measure of nonlinearity:

$$\langle | M(\boldsymbol{x}^{n-1} + \delta\boldsymbol{x}^{n}) - (M(\boldsymbol{x}^{n-1}) + \mathbf{M}\,\delta\boldsymbol{x}^{n}) | \rangle$$

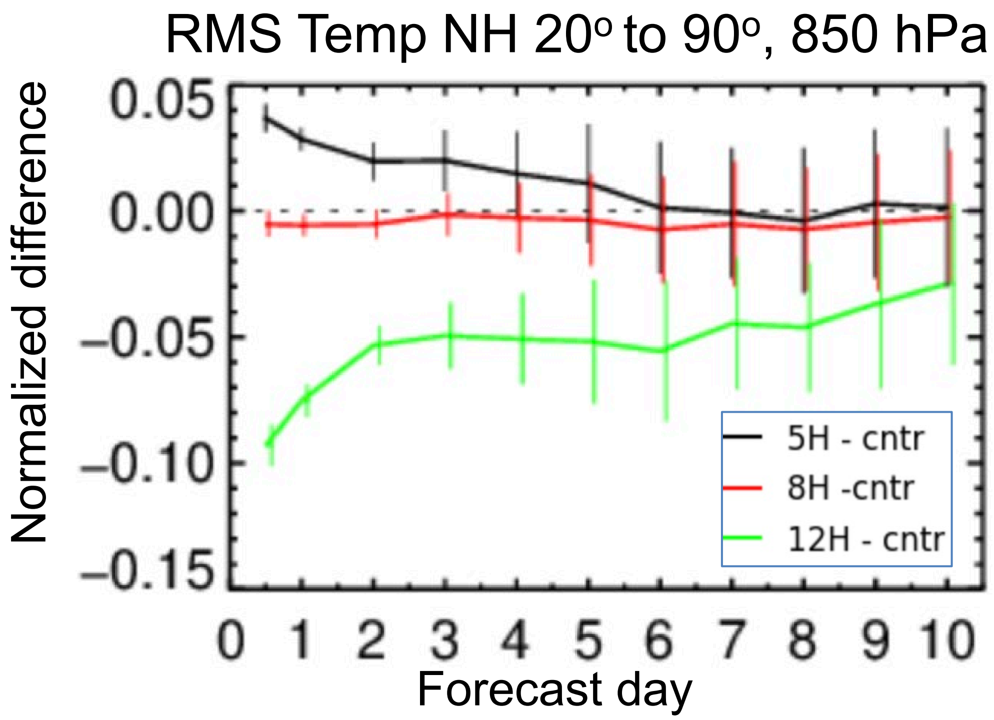
M = non-linear model
$\mathbf{M}$ = tangent linear model
$\delta\boldsymbol{x}$ = increment
$\boldsymbol{x}$ = state
n = minimization number

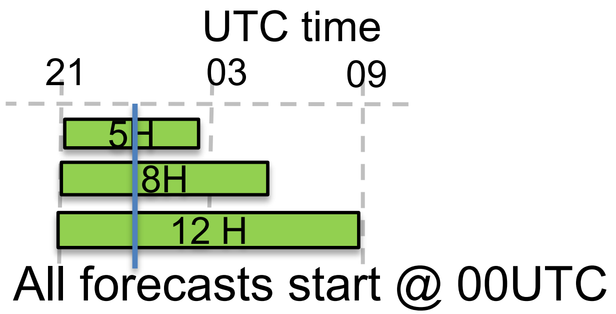


Ext-Win

- Additional iterations reduce the difference between non-linear and linear trajectory, which is the fundamental assumption made in incremental 4D-Var.
- Helpful toward high resolution modelling and assimilation of high res observations.

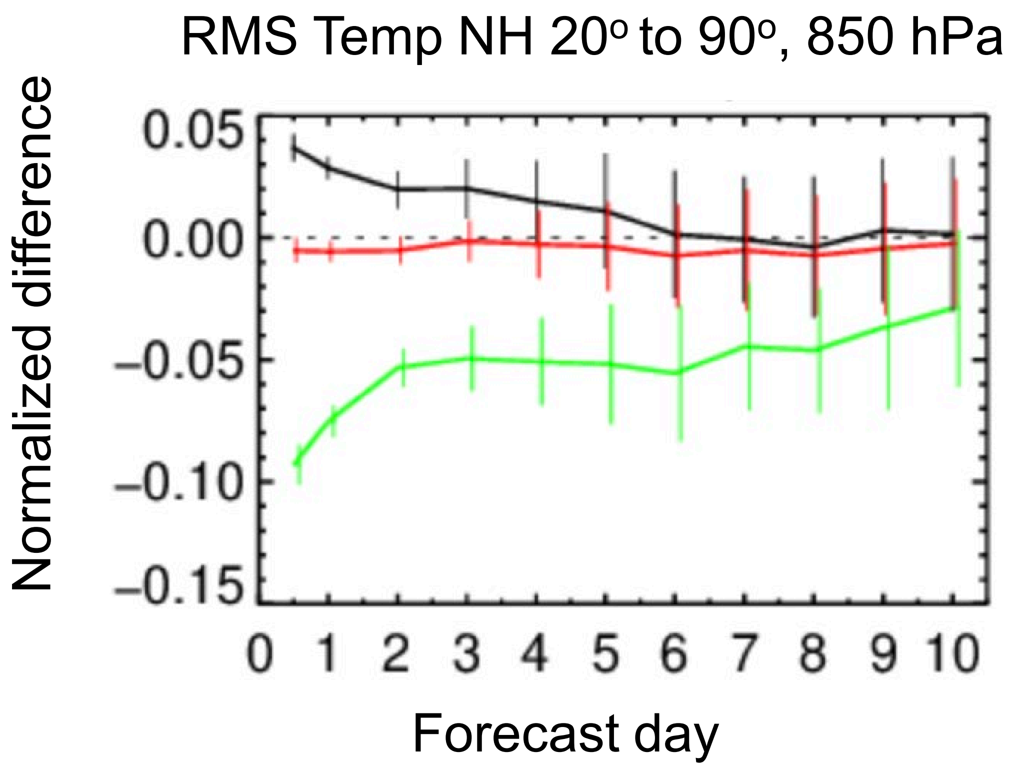
Effect on forecast scores of 3 hours of observations

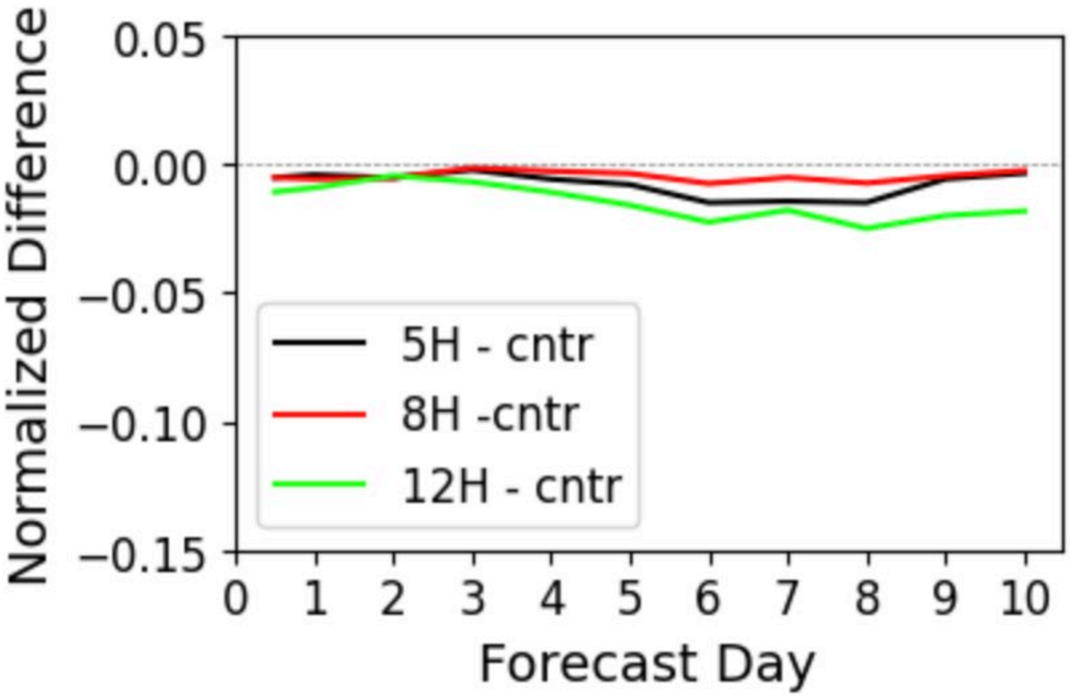




- **Unfair comparison**. All forecasts are initialized at the same time.
- **Black**/green curves show the effect of 3 hour less/more observations compared to operational forecast.
- Tradeoff between worse/better scores and time of analysis delivery.
- Forecast based on 5H DA window is delivered earlier and vice versa for 12H one

Forecast scores – fair



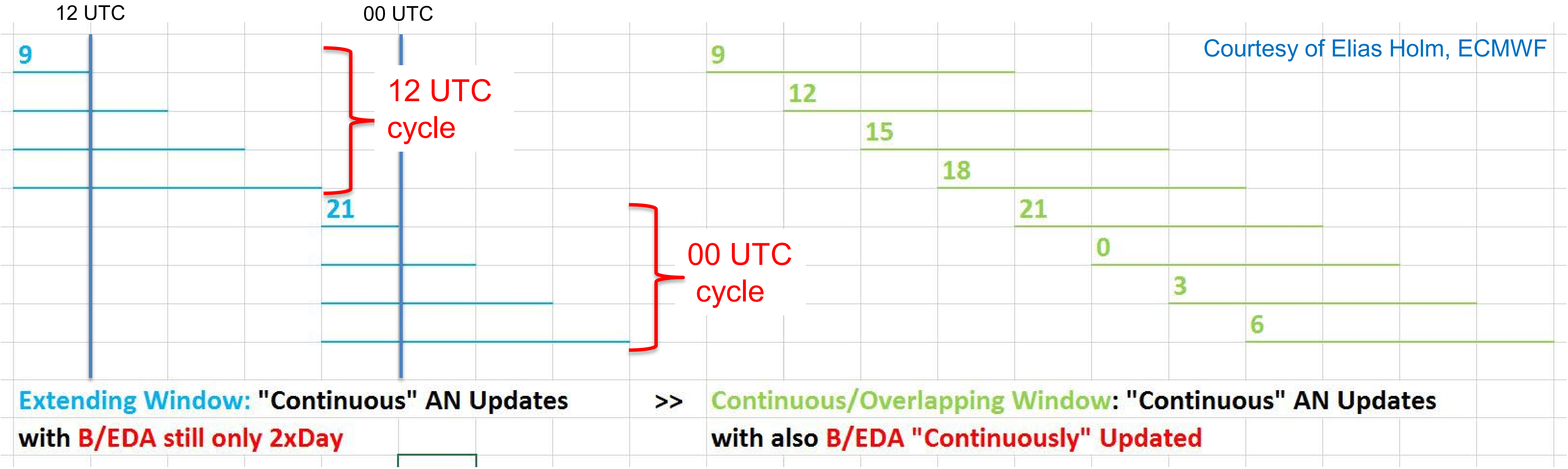


- Fair scores comparison (taking into account different DA window end) shows similar performance up to day 3.
- Not yet significant (45 days of exp) but incouraging scores from the longer window (12H) up to day 10.

Towards More Frequent Analysis *and EDA* Updates

There are 2 paths to pursue more frequent analysis and forecast updates:

- **Extending window →** EDA still only 2xDay
- **Continuous overlapping windows** → requires more frequent EDA updates as well
- More frequent EDA's require supercharged efficiency gains → **Emulate EDA members** from a small subset of full resolution EDA members (Wei Pan talk)



Conclusions and Outlook

- Evolution of forecast and analysis toward more frequent updates
- Extending window provides continuous analysis updates, making use of the latest observations availble
- Ext-win improves treating nonlinearities → helpful with high resolution model and observations
  - Drop low res (TL255) outer loop? → reduce computational footprint

Towards higher resolution data assimilation!

- ECMWF aims to continuously enhance the resolution of the IFS, thereby delivering forecasts with progressively finer details
  - Cycle 48r1: an improved resolution of ensemble forecasts (ENS): 18 km to 9 km
  - Cycle 49r1: an improved resolution of ensemble data assimilation (EDA): 18 km to 9 km

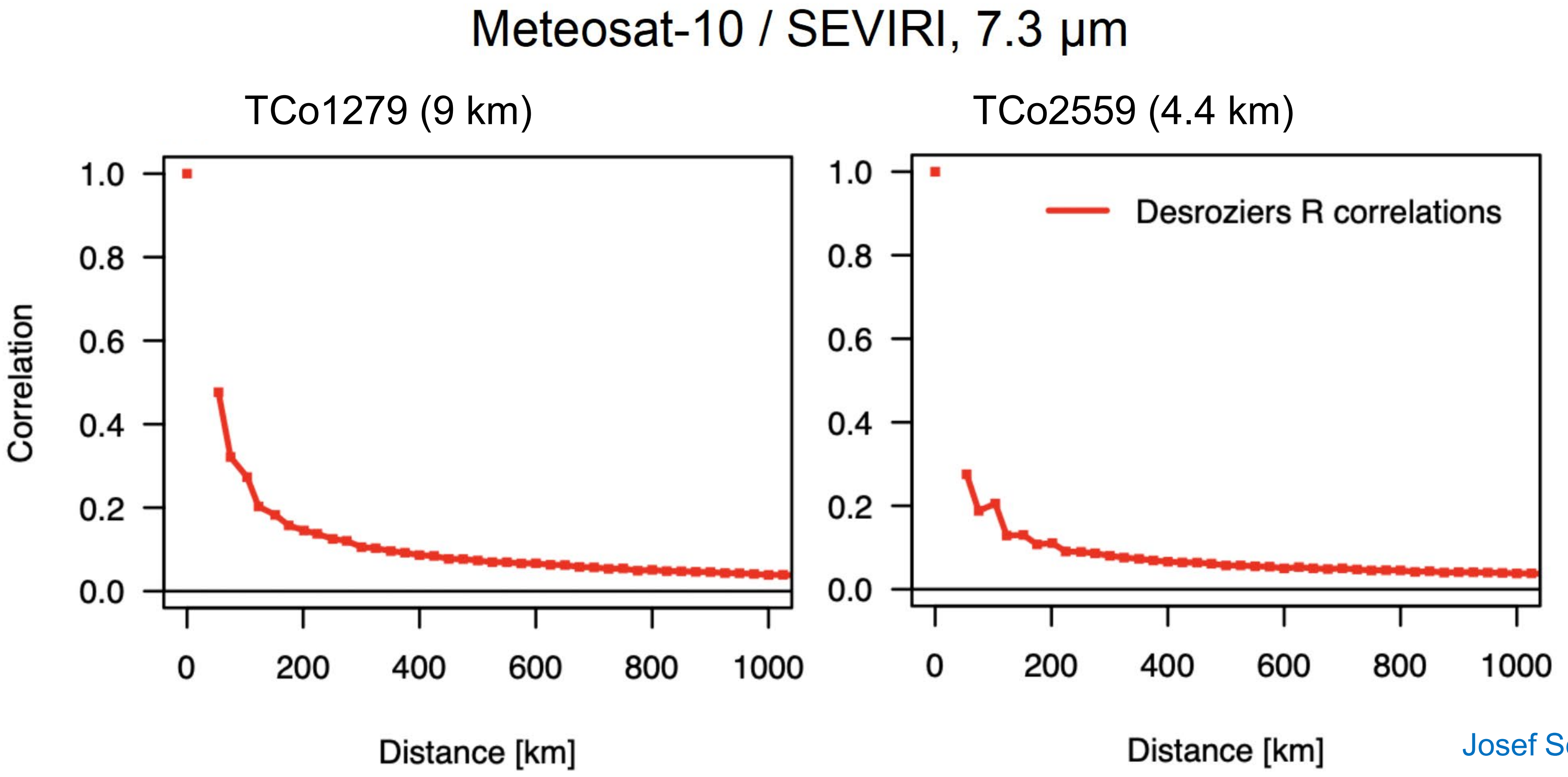
- At such resolutions, the IFS is able to resolve the mesoscale extreme-weather phenomena and quantify the probability of their occurrence

- The goal of Destination Earth Extremes digital twin is to provide prediction framework at kilometer-scale resolution, resolving the convective-scale phenomena
- Higher-resolution data assimilation supports these efforts!

Motivation for higher-resolution DA

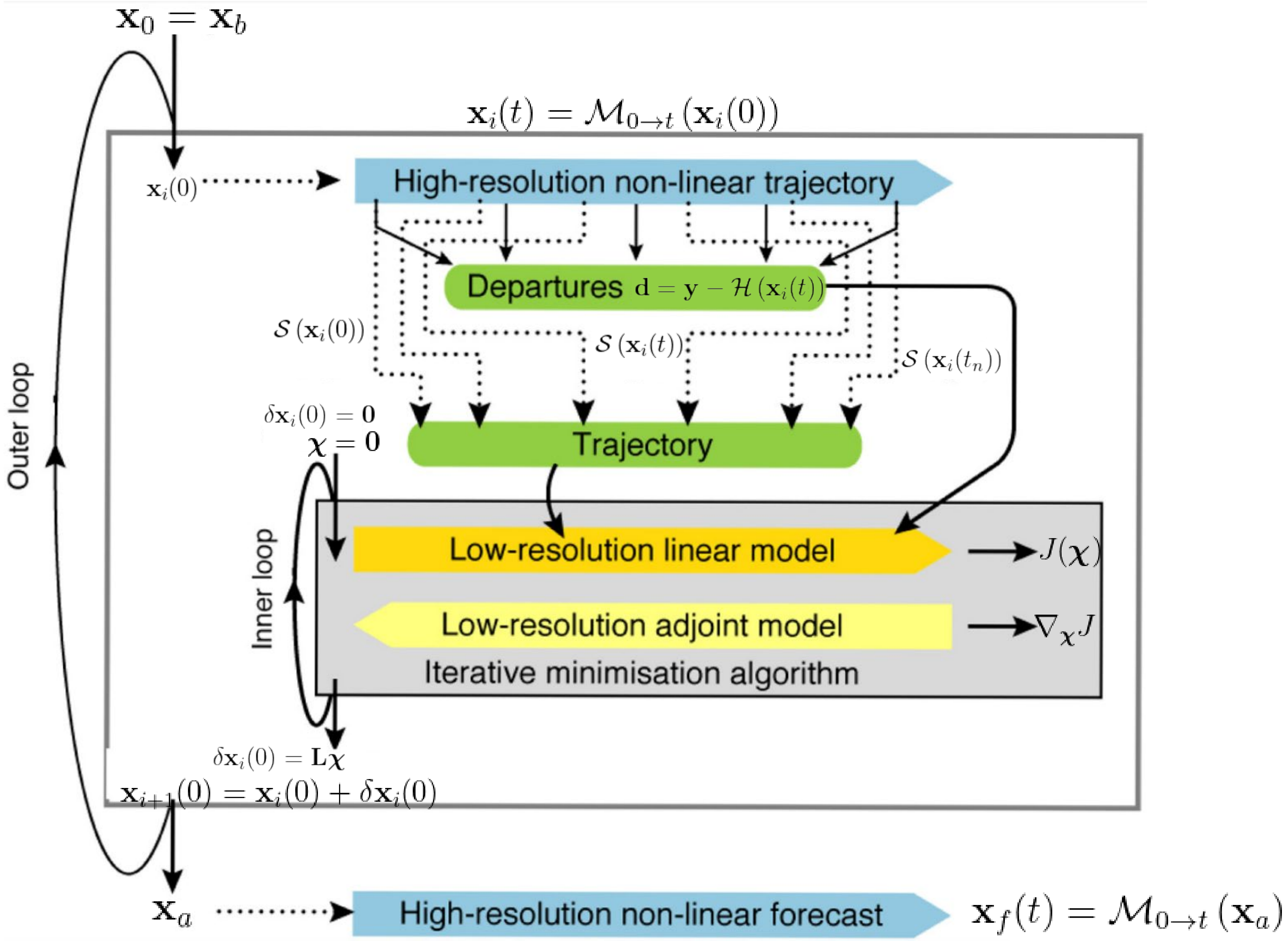
- Make use of greater portion of available observations
- Observation departures $\left(\mathbf{y} - H(\mathbf{x}_b)\right)$ are less correlated at short distances at higher resolutions,

→ less thinning is required to achieve uncorrelated errors



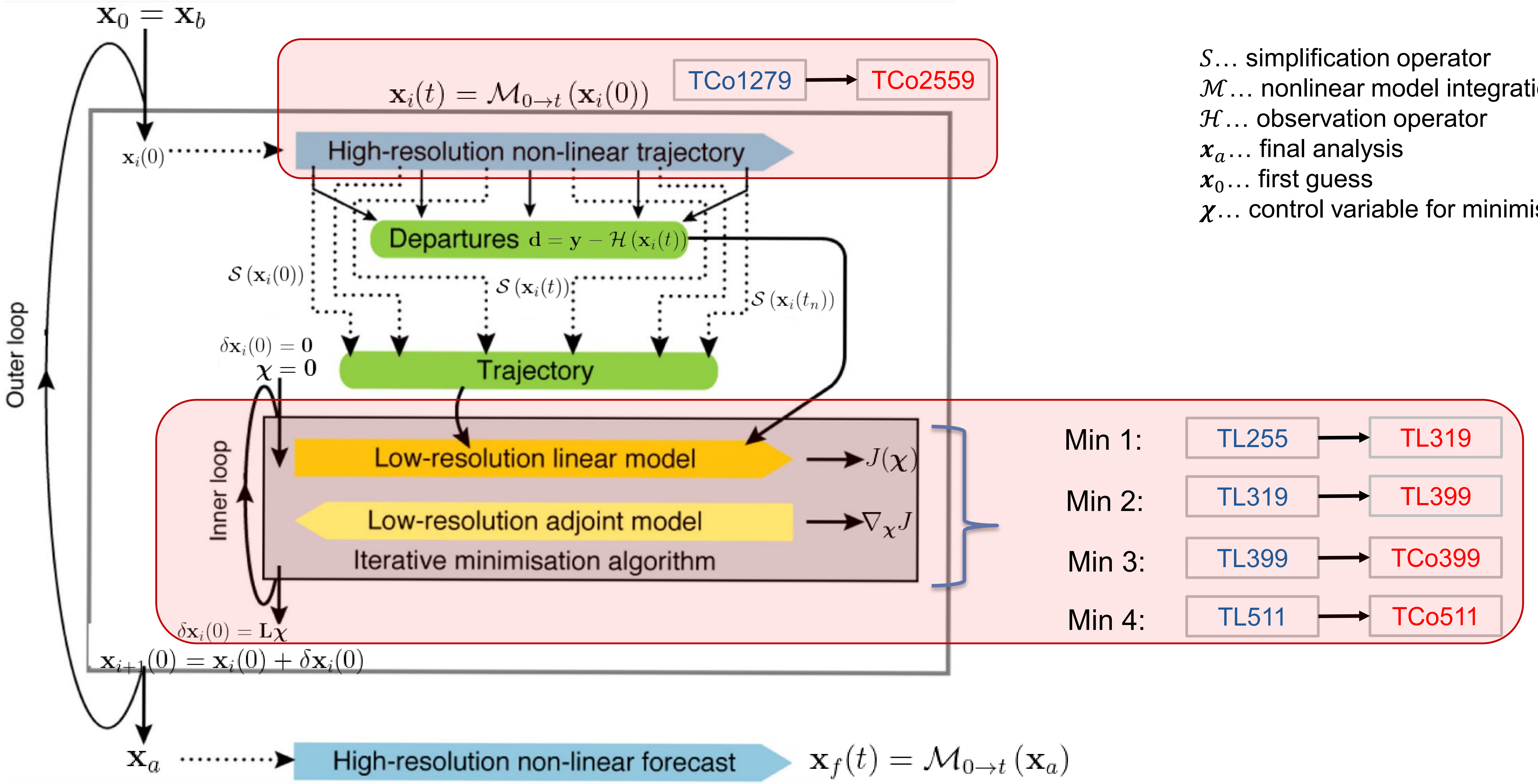
Josef Schröttle

4D-Var data assimilation in IFS



$\mathcal{S}$… simplification operator
$\mathcal{M}$… nonlinear model integration
$\mathcal{H}$… observation operator
$\boldsymbol{x}_a$… final analysis
$\boldsymbol{x}_0$… first guess
χ… control variable for minimisation

Towards higher spatial resolution
4D-Var data assimilation in IFS



$\mathcal{S}$… simplification operator
$\mathcal{M}$… nonlinear model integration
$\mathcal{H}$… observation operator
$\boldsymbol{x}_a$… final analysis
$\boldsymbol{x}_0$… first guess
$\boldsymbol{\chi}$… control variable for minimisation

Higher-resolution data assimilation

Setup

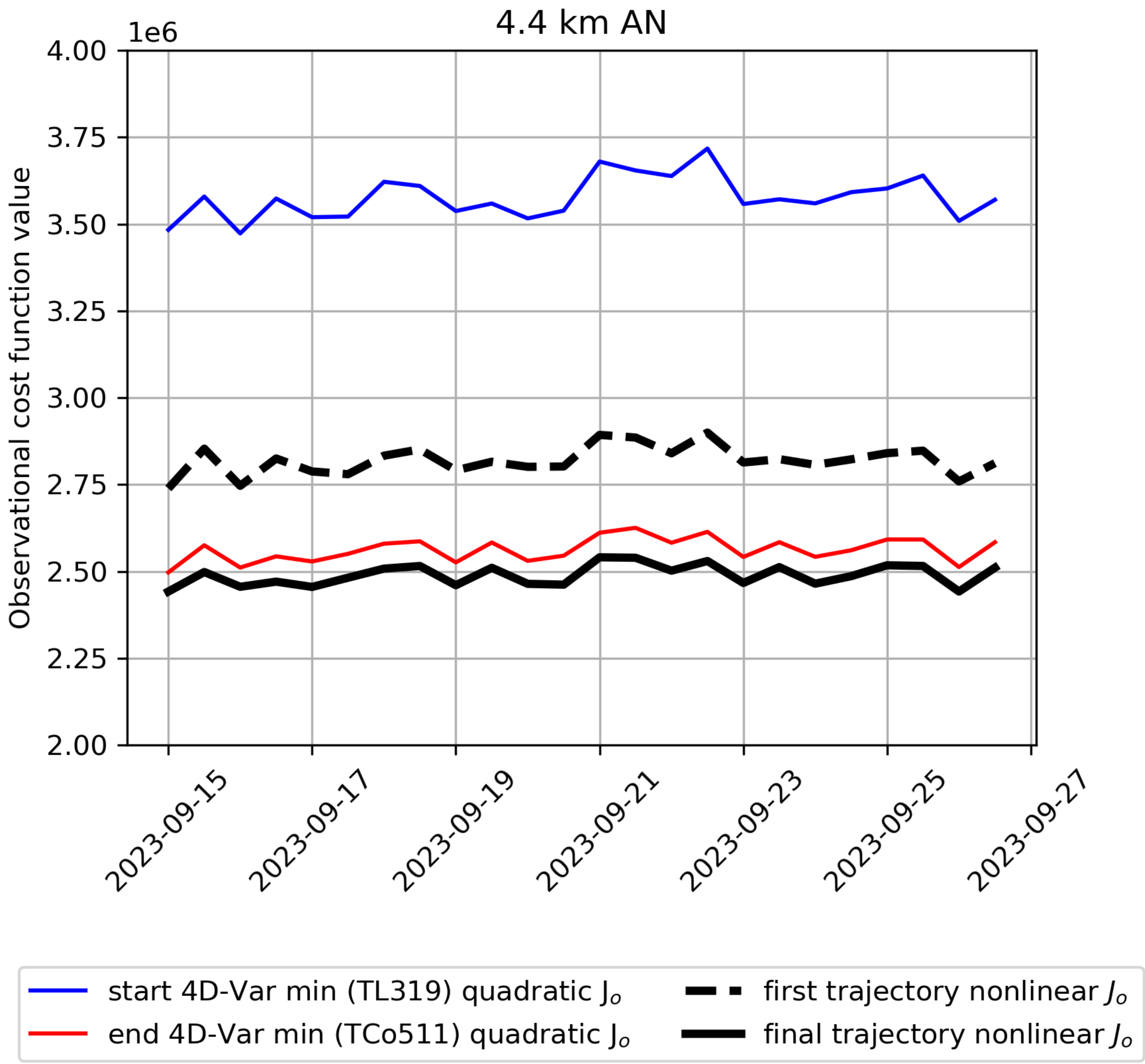
- TCo2559 (4.4 km) trajectory based on latest 49r1 DestinE forecast model with 200 s time step
- Improved resolution of minimisation (TL319/TL399/TCo399/TCo511) with relaxed cost function convergence criterion
- Observation time slots reduced from 1800s in operations to 400s
- High-resolution geostationary satellite data with reduced spatial thinning
- Long-window (12-hour) DA mode

Experiment data

September 1, 2023 – October 31, 2023 (now running further!)

Is the high-res 4D-Var setup optimal?

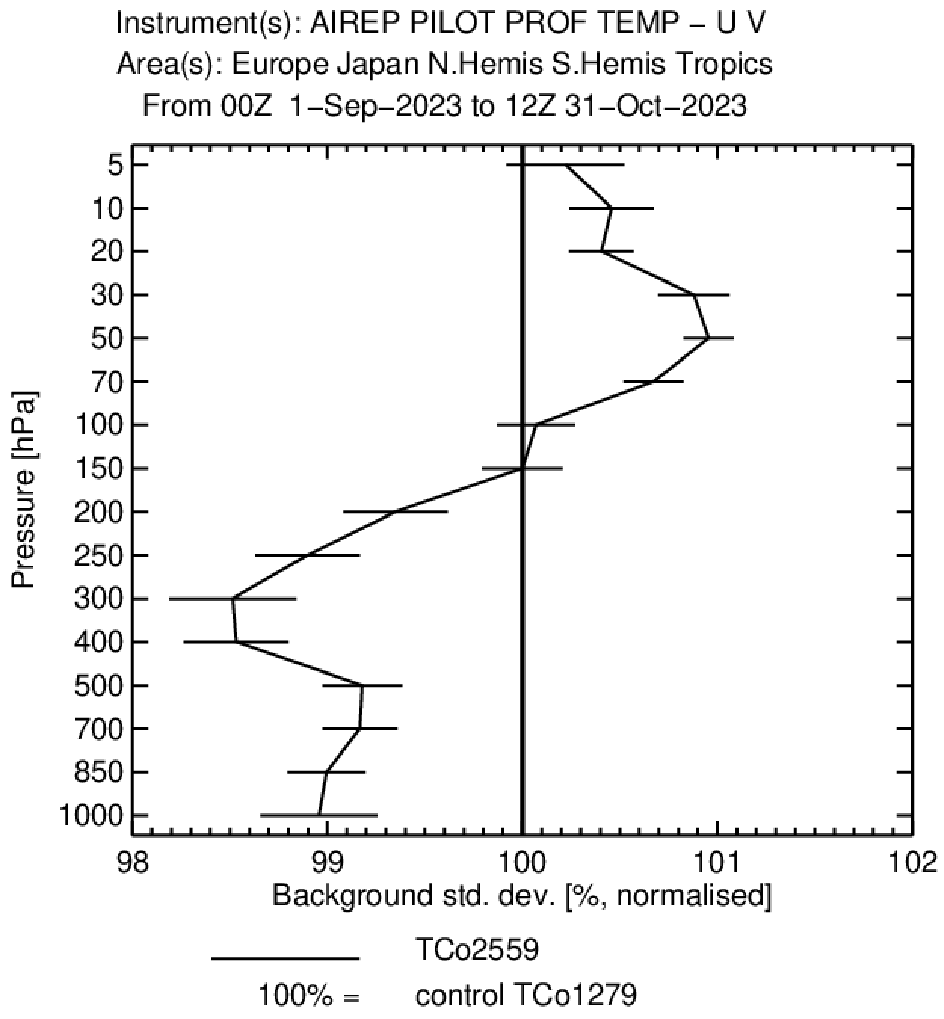
- Still not!!
- A mismatch between:
  - nonlinear cost function $\mathcal{J}_o$ at TCo2559 (4.4 km) and
  - quadratic cost function J_o at TL319 (55 km)
- Good match between:
  - nonlinear cost function $\mathcal{J}_o$ at TCo2559 (4.4 km) and
  - quadratic cost function J_o at TCo511 (20 km)
- Possible reasons for mismatch:
  - Resolution of the first inner-loop is too low?
  - Linearisation hypothesis less valid?



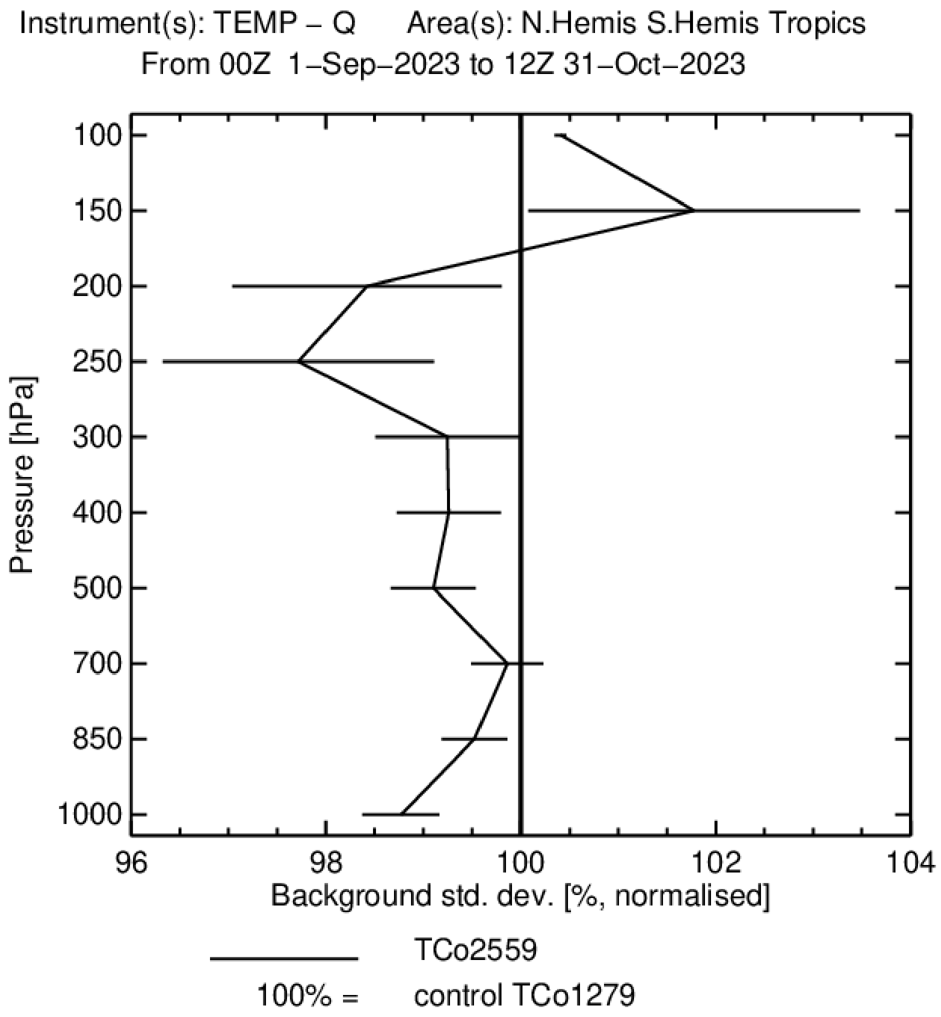
ECMWF

Improved initial conditions

WINDS - CONVENTIONAL

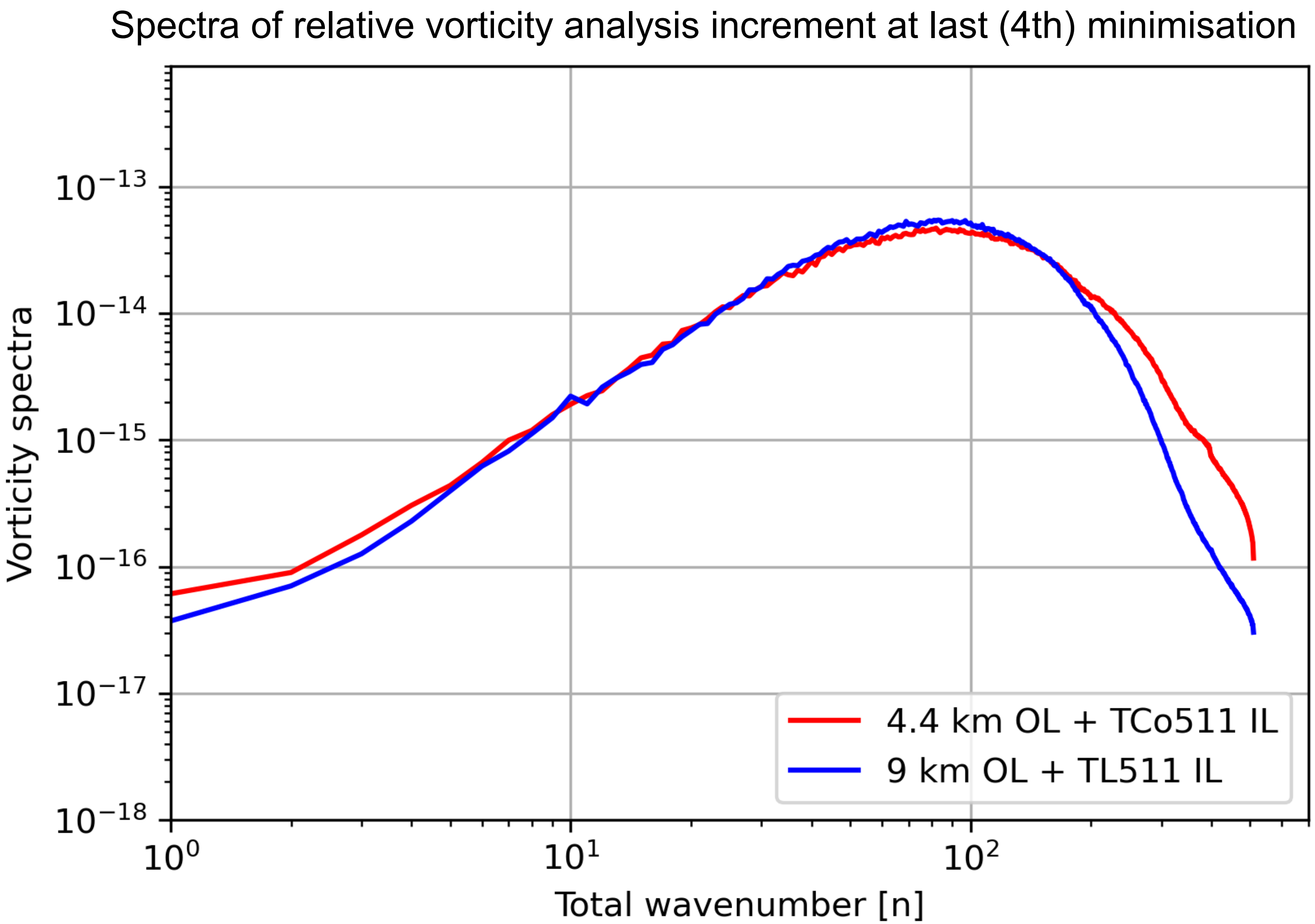


HUMIDITY - CONVENTIONAL



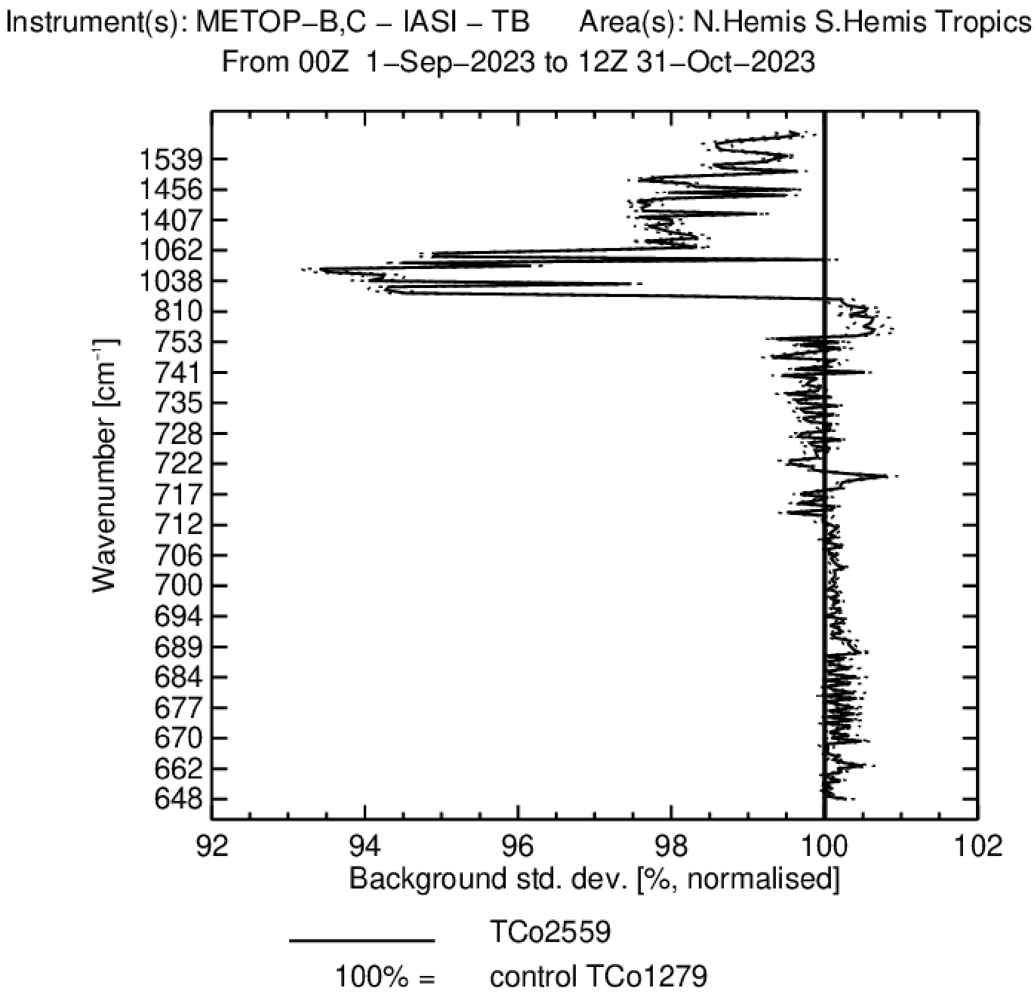
Improved initial conditions

- By employing higher-resolution minimisation (inner loop, i.e. TLM and ADM), a greater part of wind and humidity spectra is corrected



Improved initial conditions

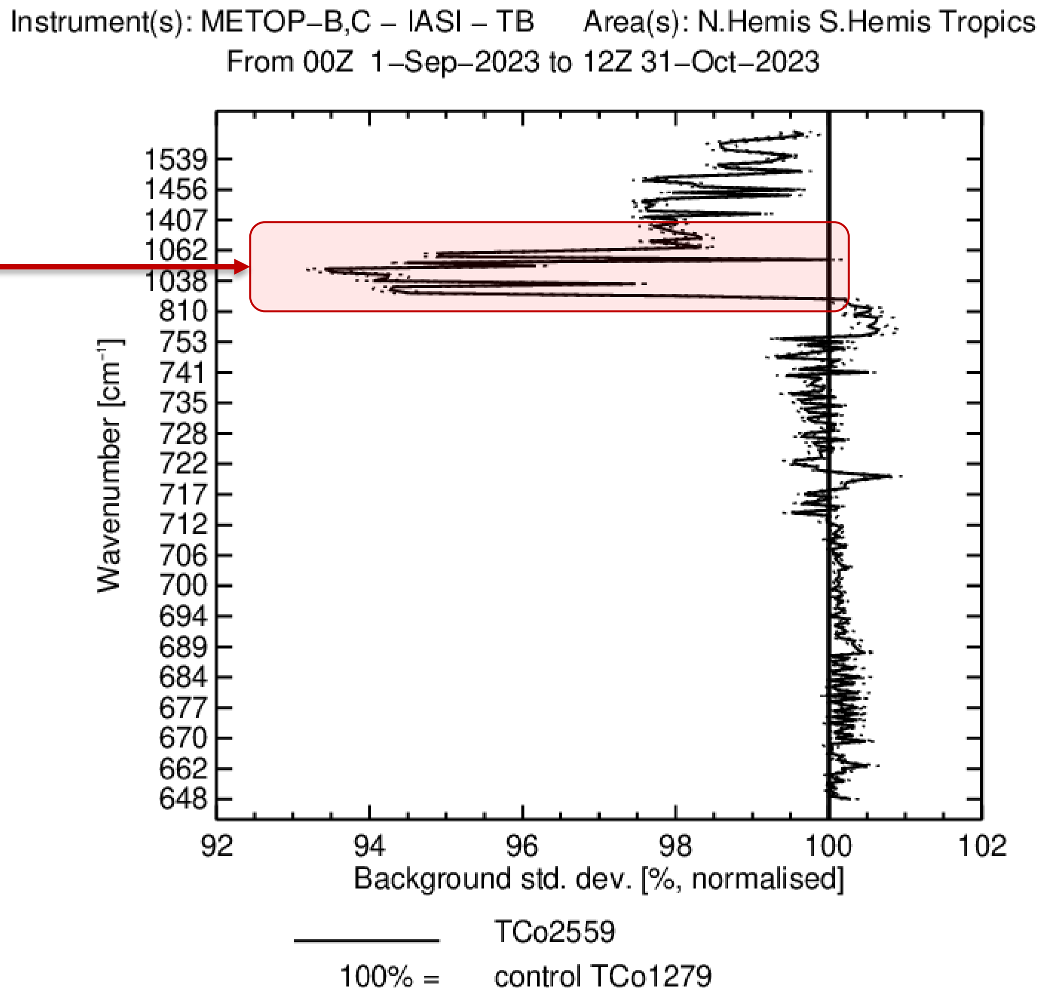
IASI hyperspectral sounder



Improved initial conditions

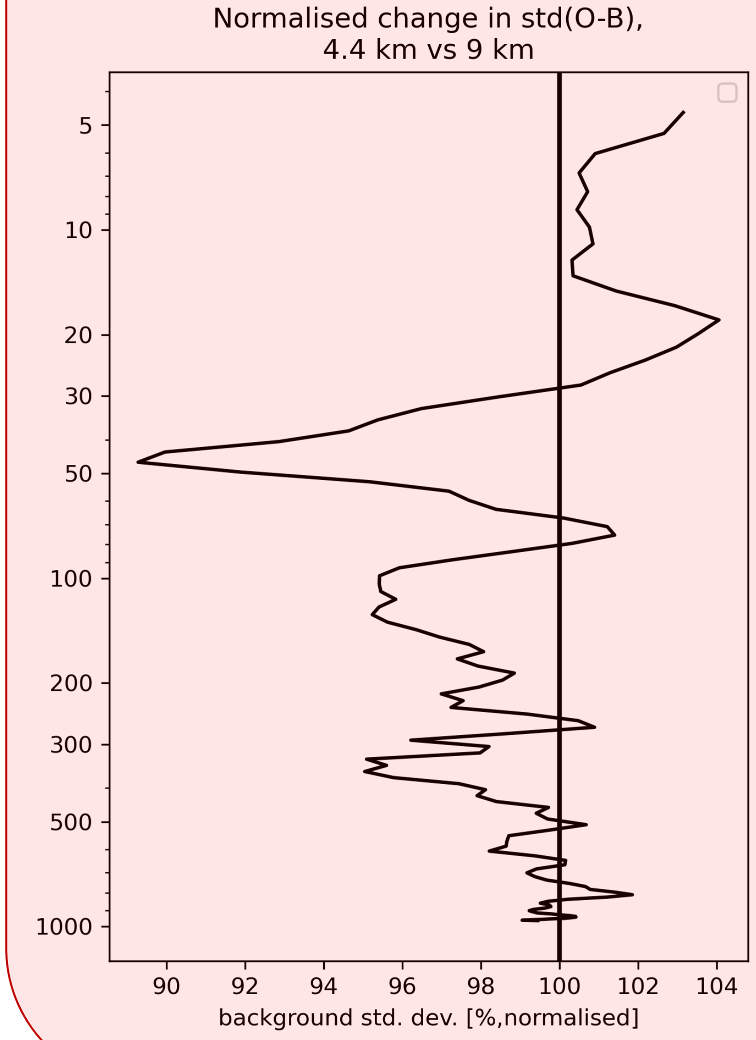
Better fit of short-range forecast to ozone-sensitive channels!

IASI hyperspectral sounder

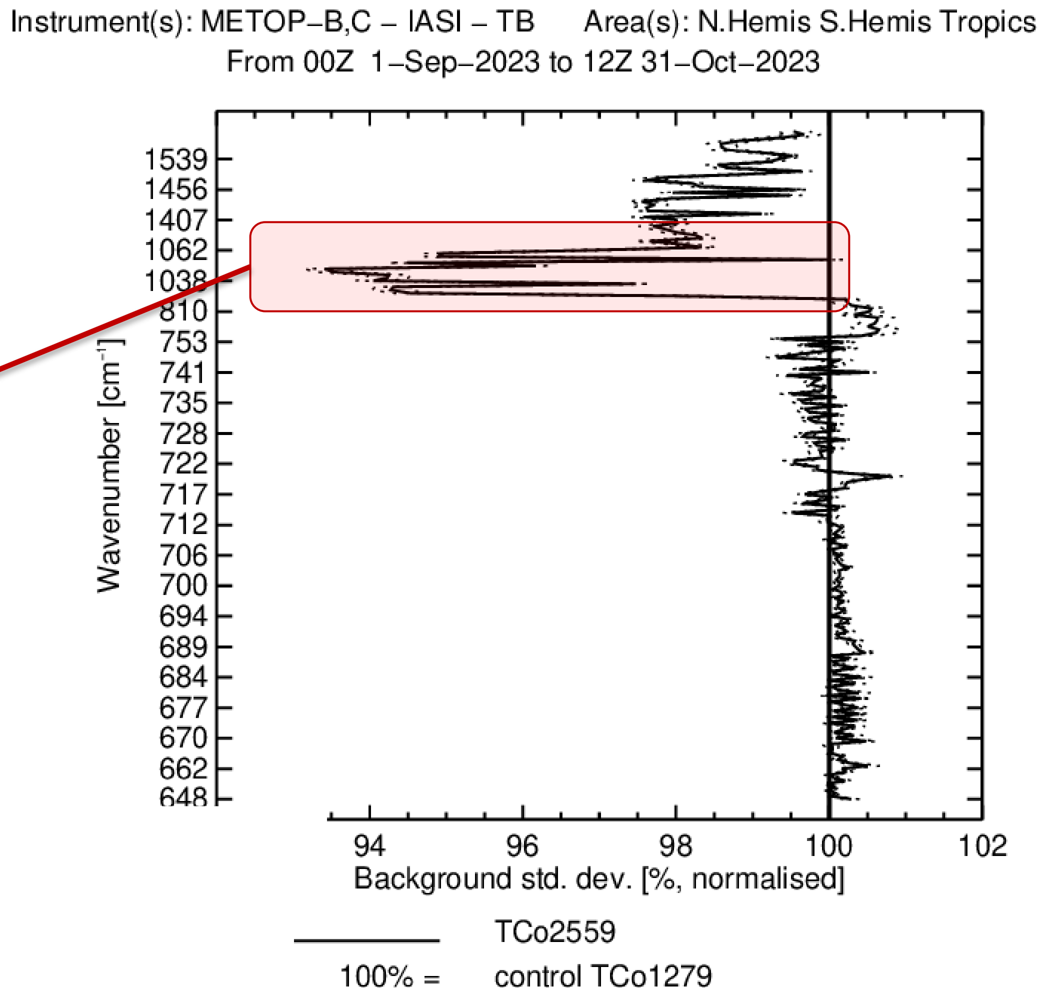


Improved initial conditions

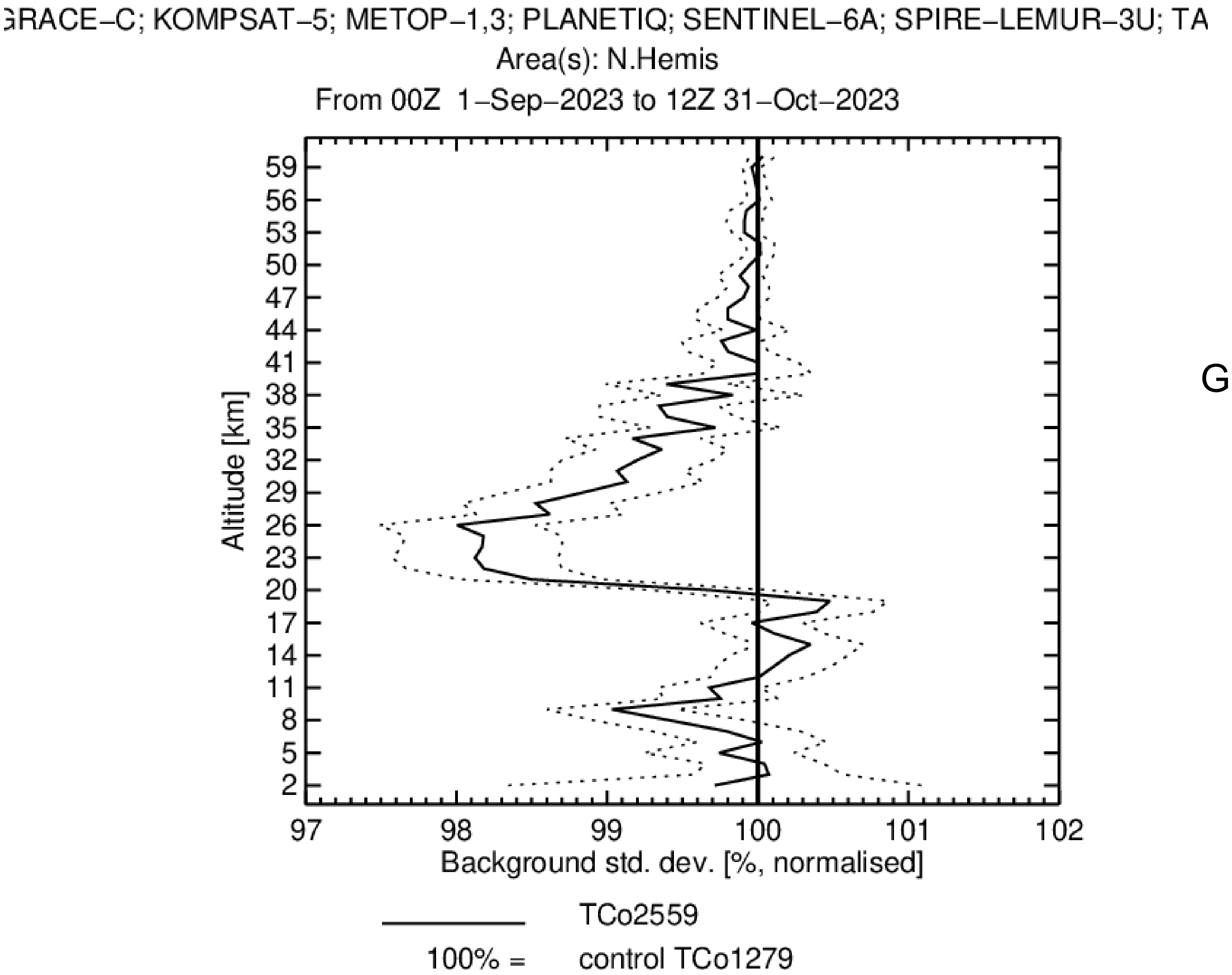
Verification of 0-12h FC against ozone sondes



IASI hyperspectral sounder

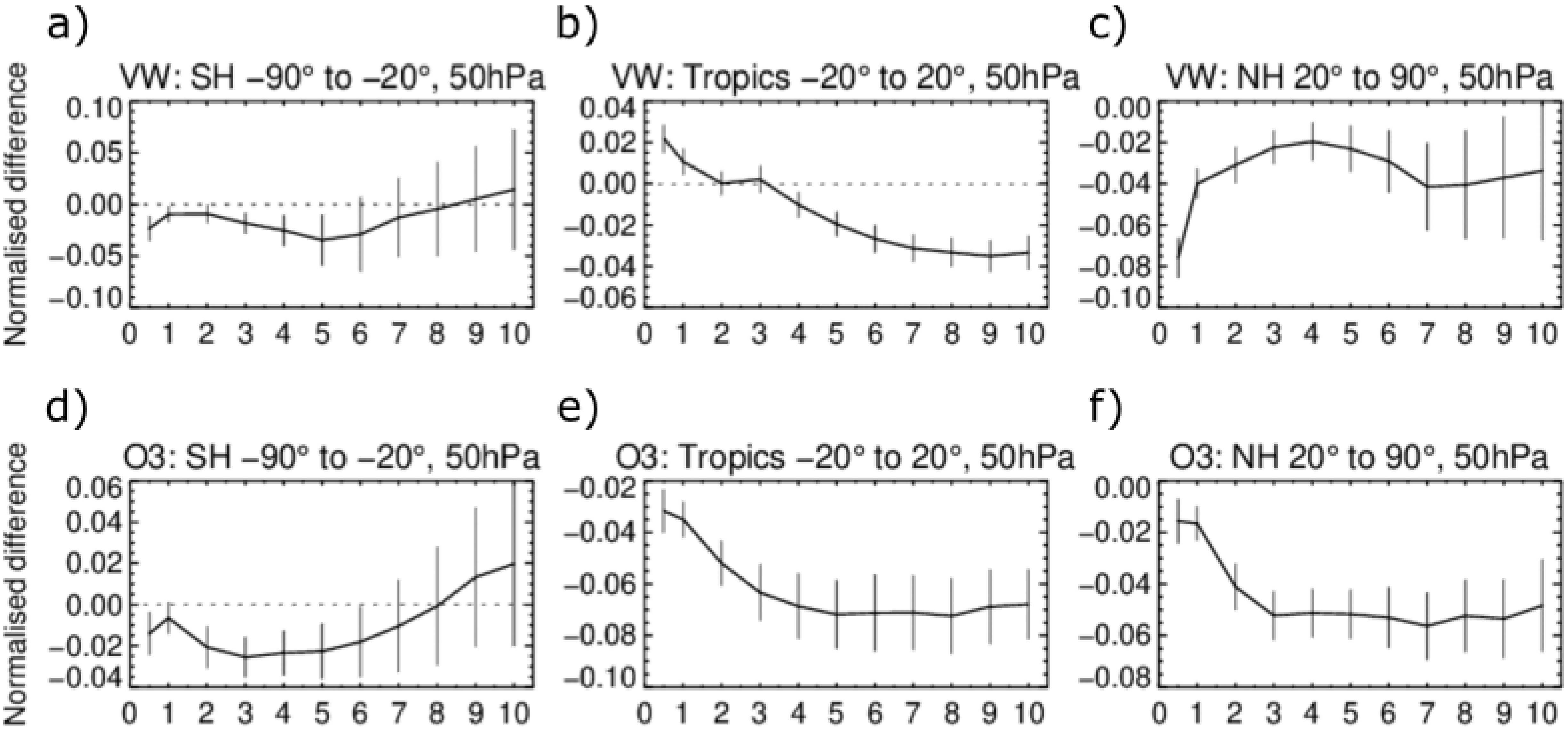


Improved initial conditions



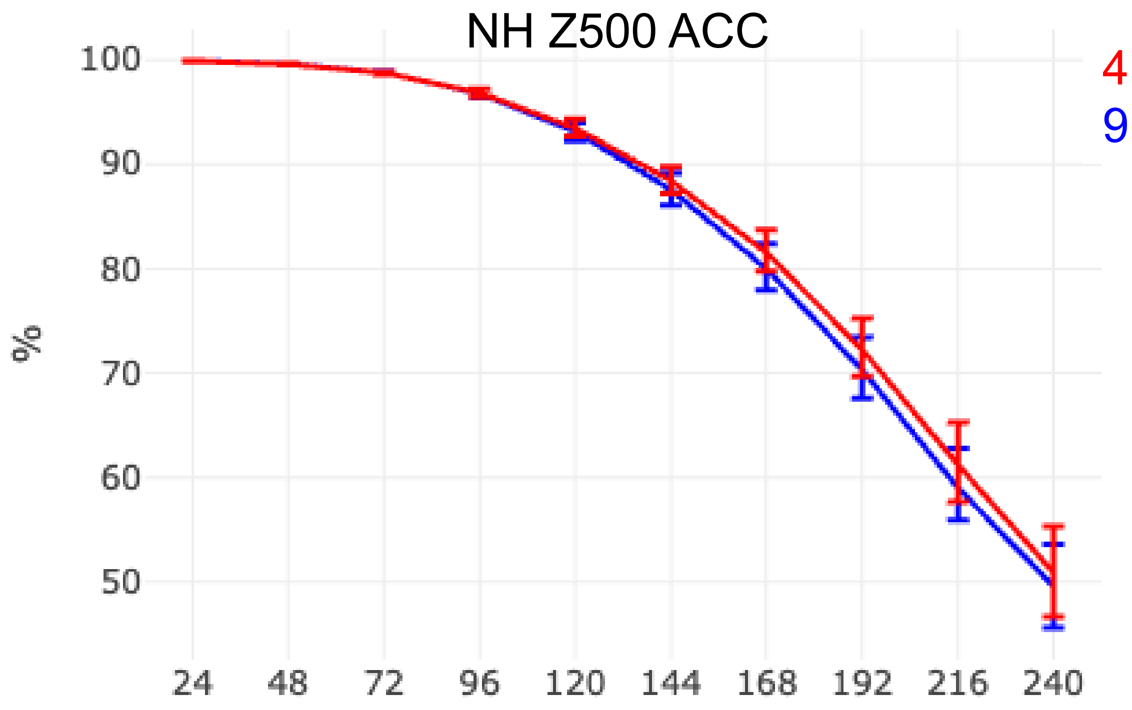
GNSSRO – Northern Hemisphere

Improved stratospheric predictability

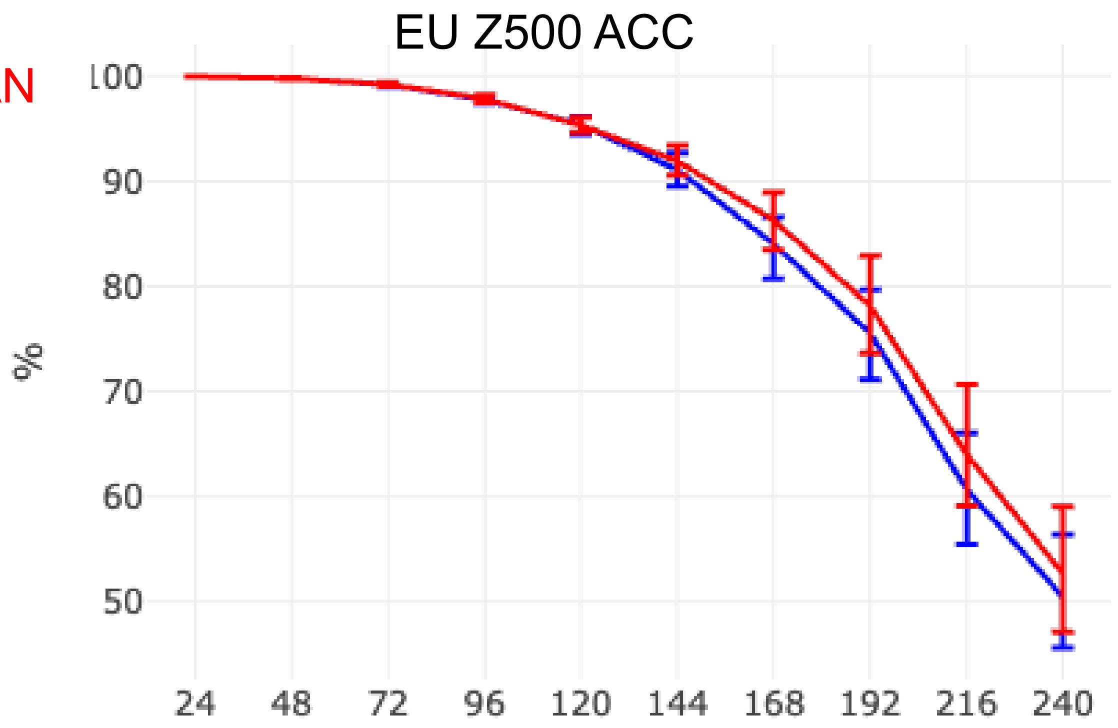


Improved IFS forecasts

- 6-12 hours of additional forecast skill, measured both
  - **against own analyis (below)** as well as
  - against observations

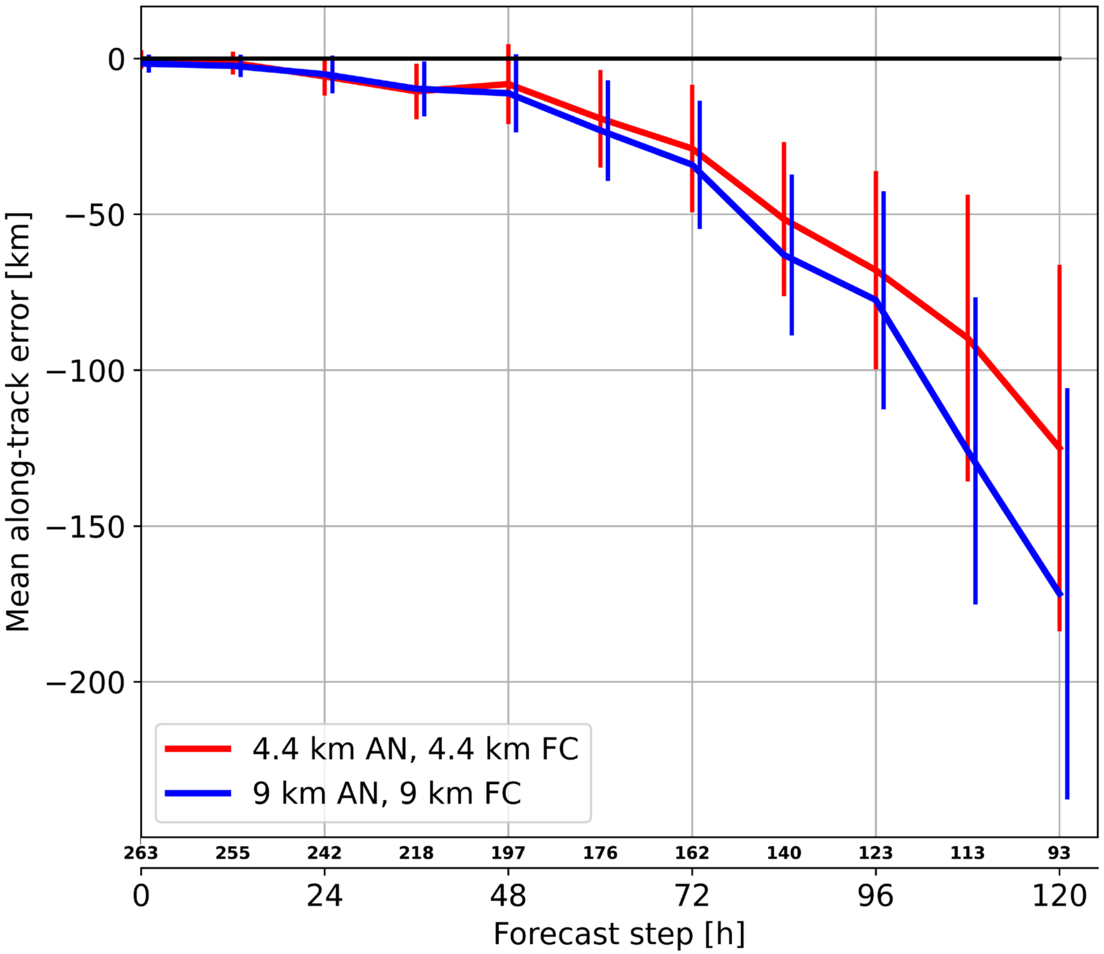


4.4 km FC + AN
9 km FC + AN

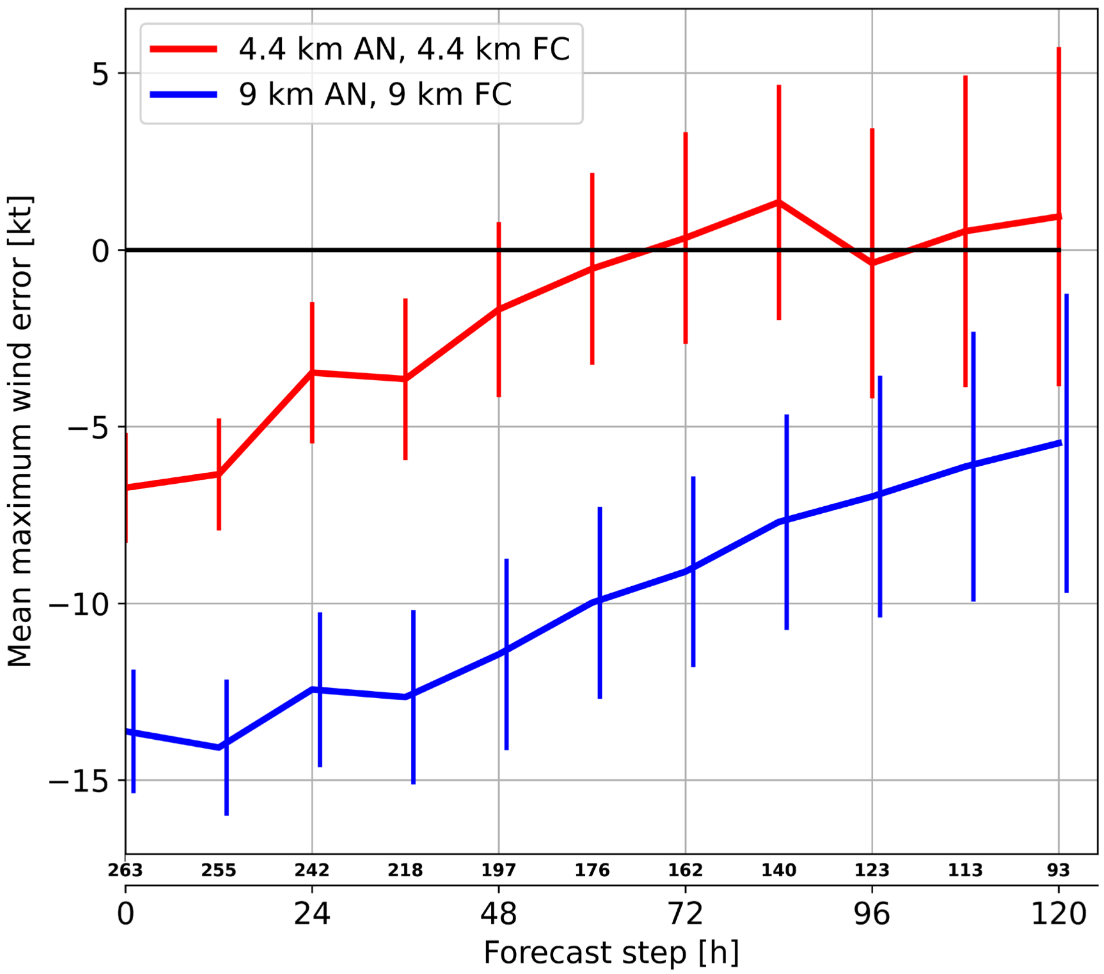


Improved TC forecasts

MEAN ALONG-TRACK ERROR

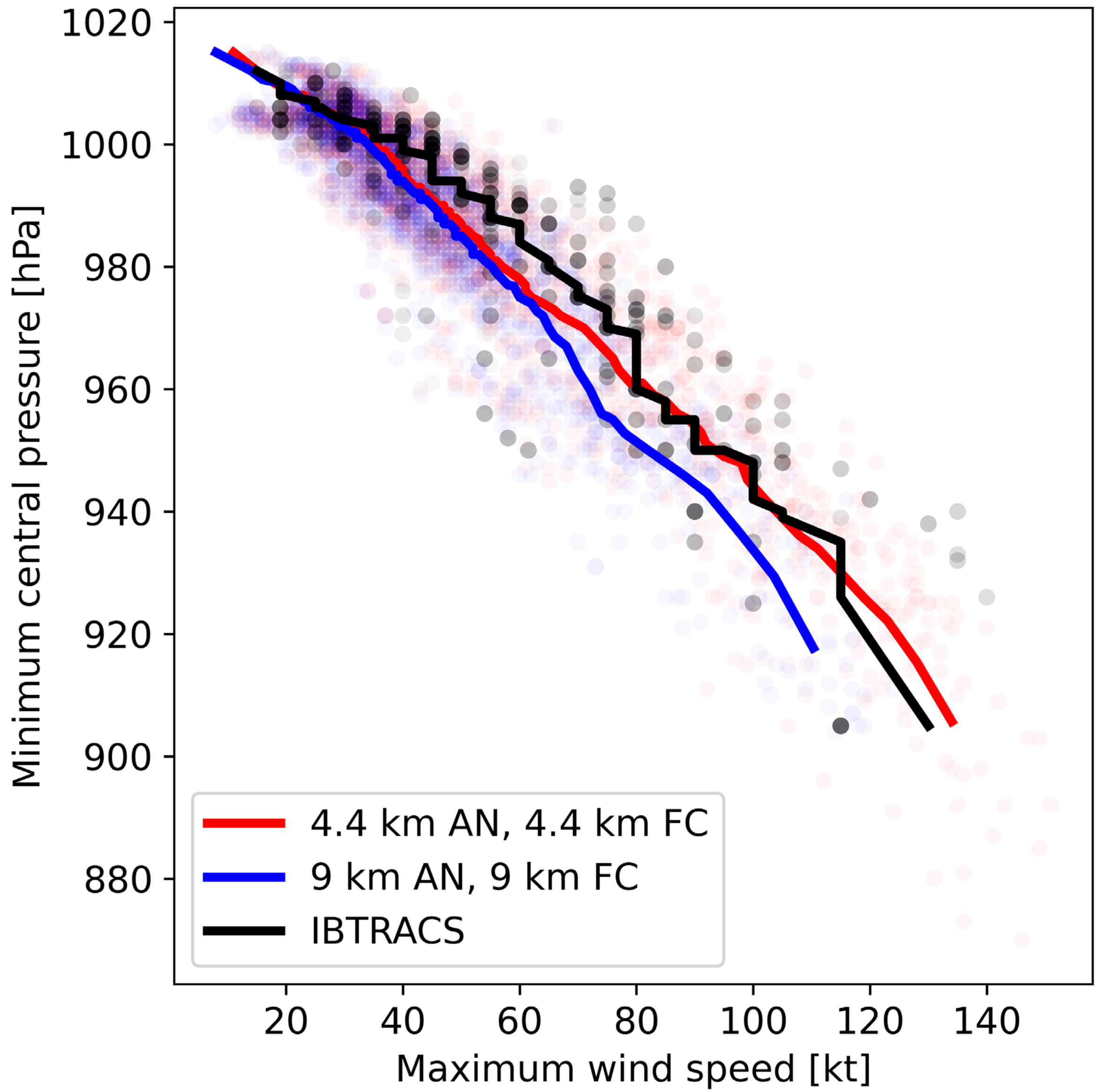


MEAN MAXIMUM WIND ERROR



Michael Maier-Gerber

Improved TC forecasts



ECMWF

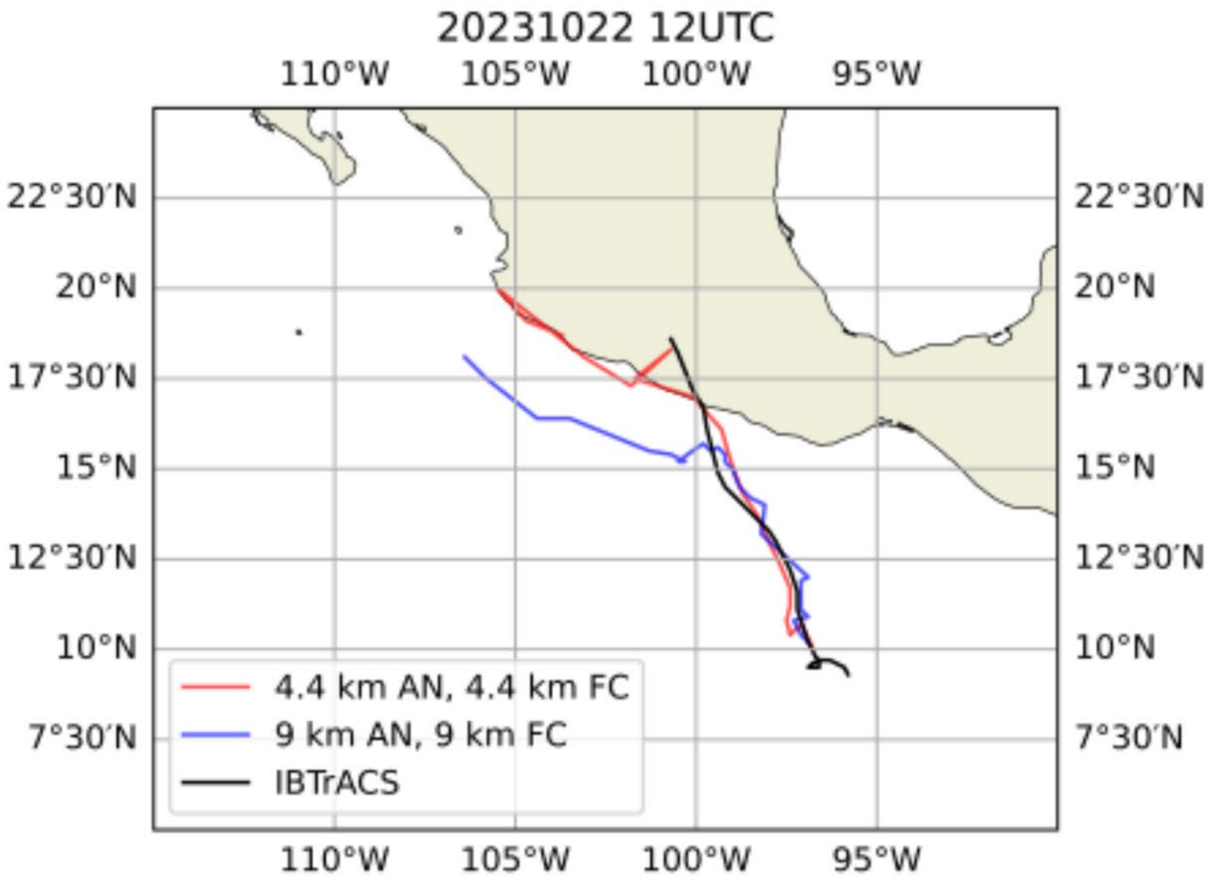
Michael Maier-Gerber

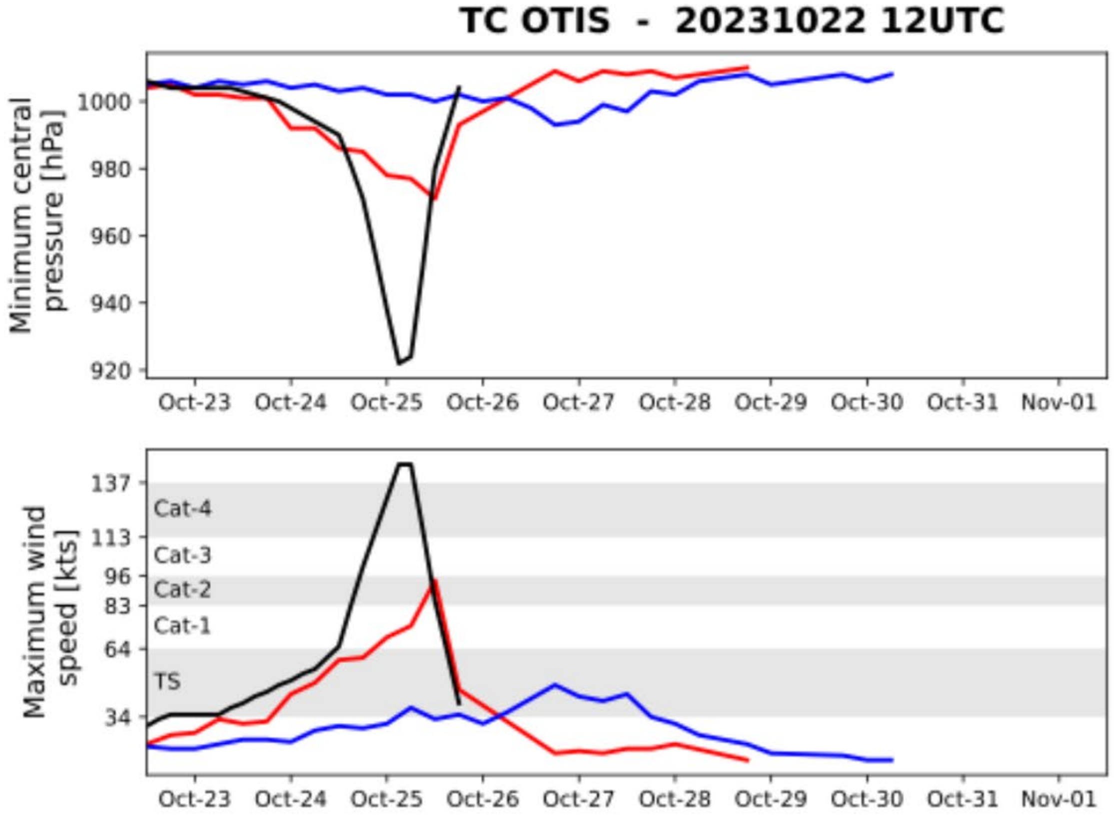
Extreme cases: hurricane Otis (October 23-25, 2023)

- East-Pacific hurricane which Hit Acapulco, Mexico on October 25th as a Category 5 storm
- p_{min} 923 hPa, maximum sustained winds 270 km/h
- Rapid intensification from a tropical storm to Cat-5 in less than 24 hours
- > 50 fatalities,$ 16 billion of damage
- None of the global or regional deterministic models or machine learning models were able to ”capture” the rapid intensification → a MAJOR forecast bust

Extreme cases: hurricane Otis (October 23-25, 2023)

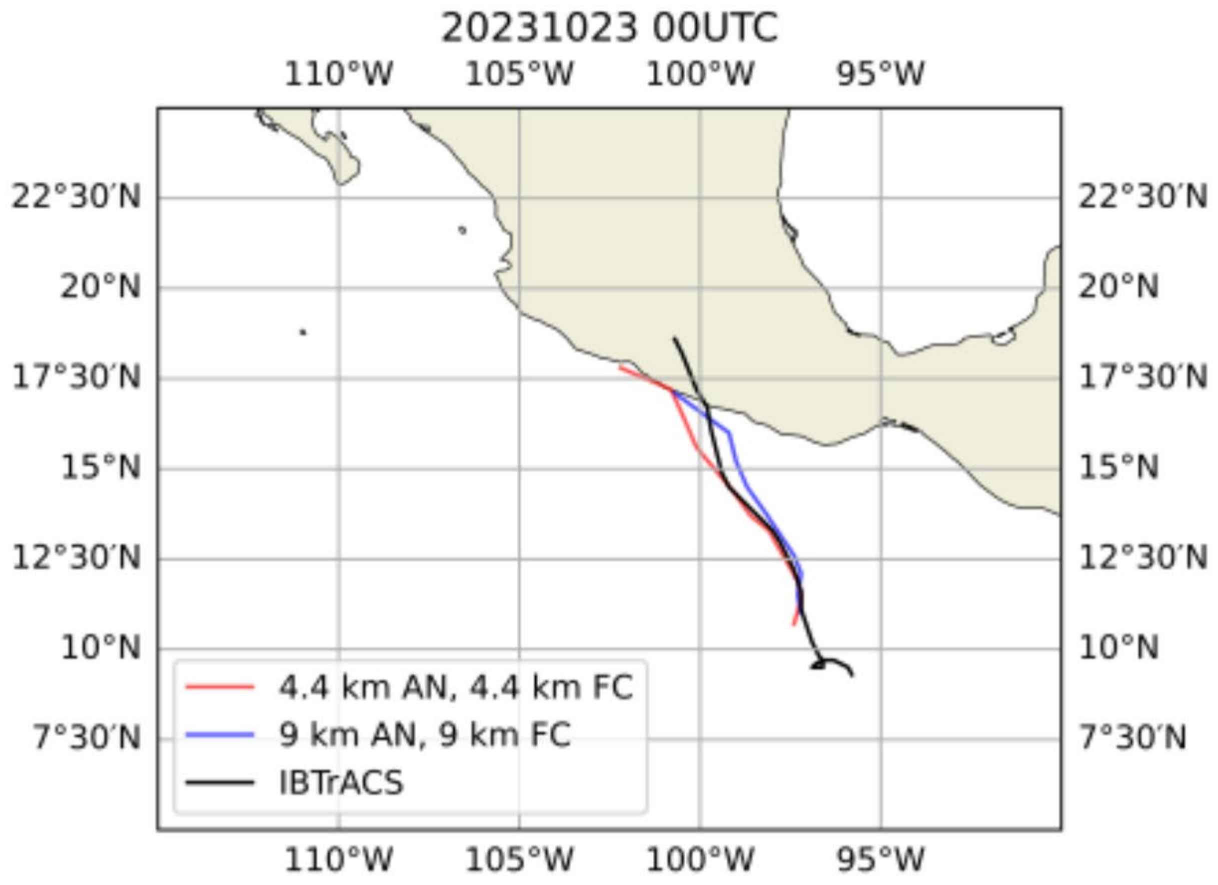
4.4 km AN, 4.4 km FC
9 km AN, 9 km FC

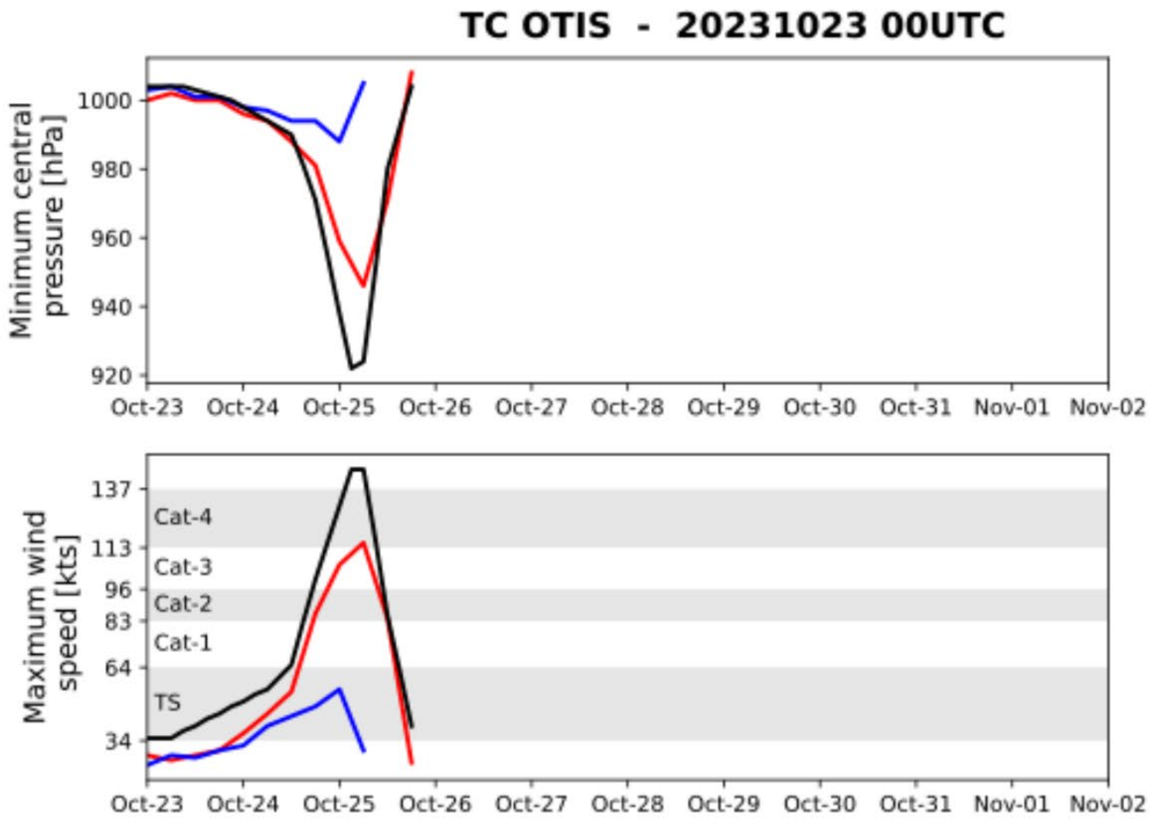




Extreme cases: hurricane Otis (October 23-25, 2023)

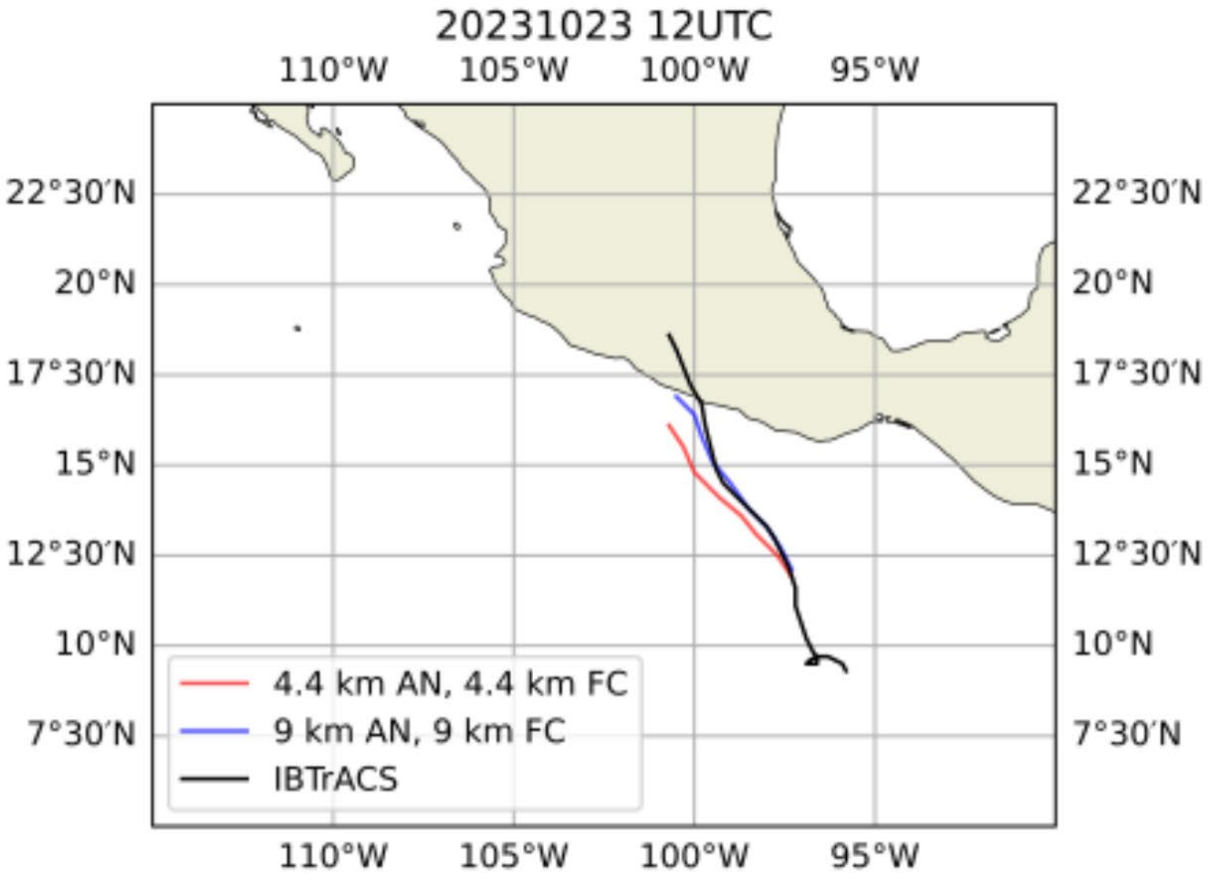
4.4 km AN, 4.4 km FC
9 km AN, 9 km FC

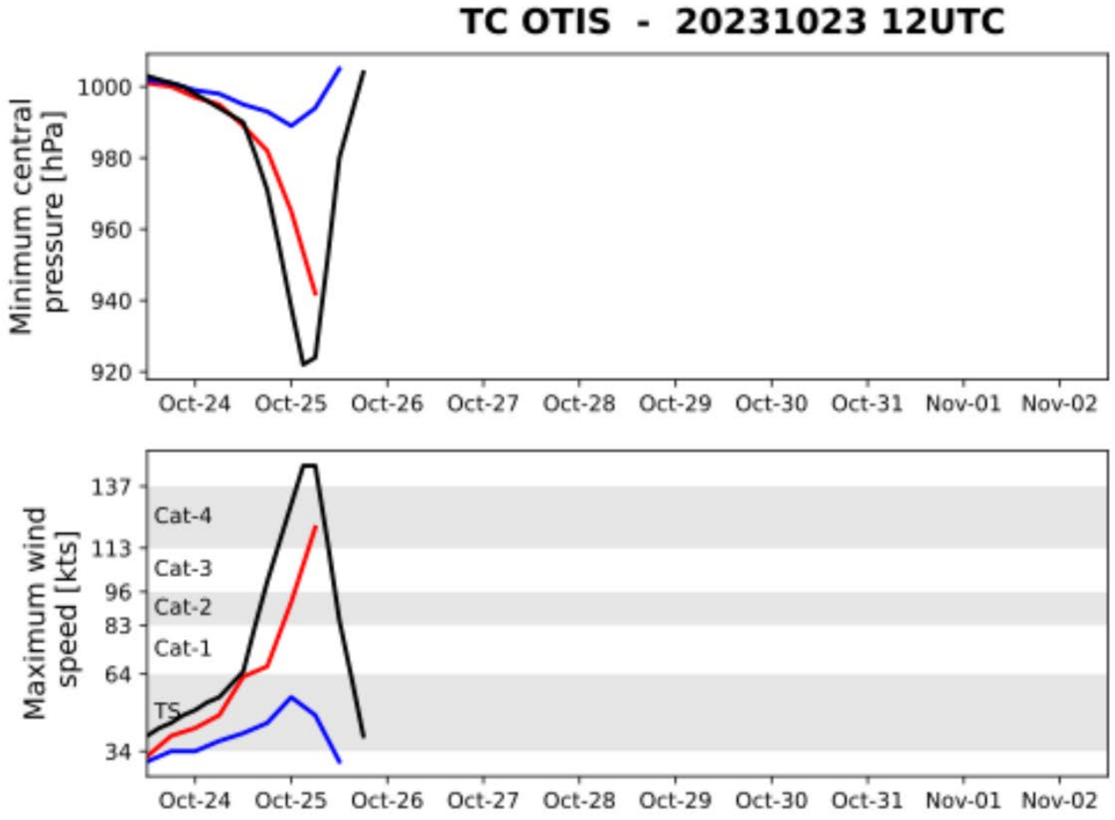




Extreme cases: hurricane Otis (October 23-25, 2023)

4.4 km AN, 4.4 km FC
9 km AN, 9 km FC





Costs

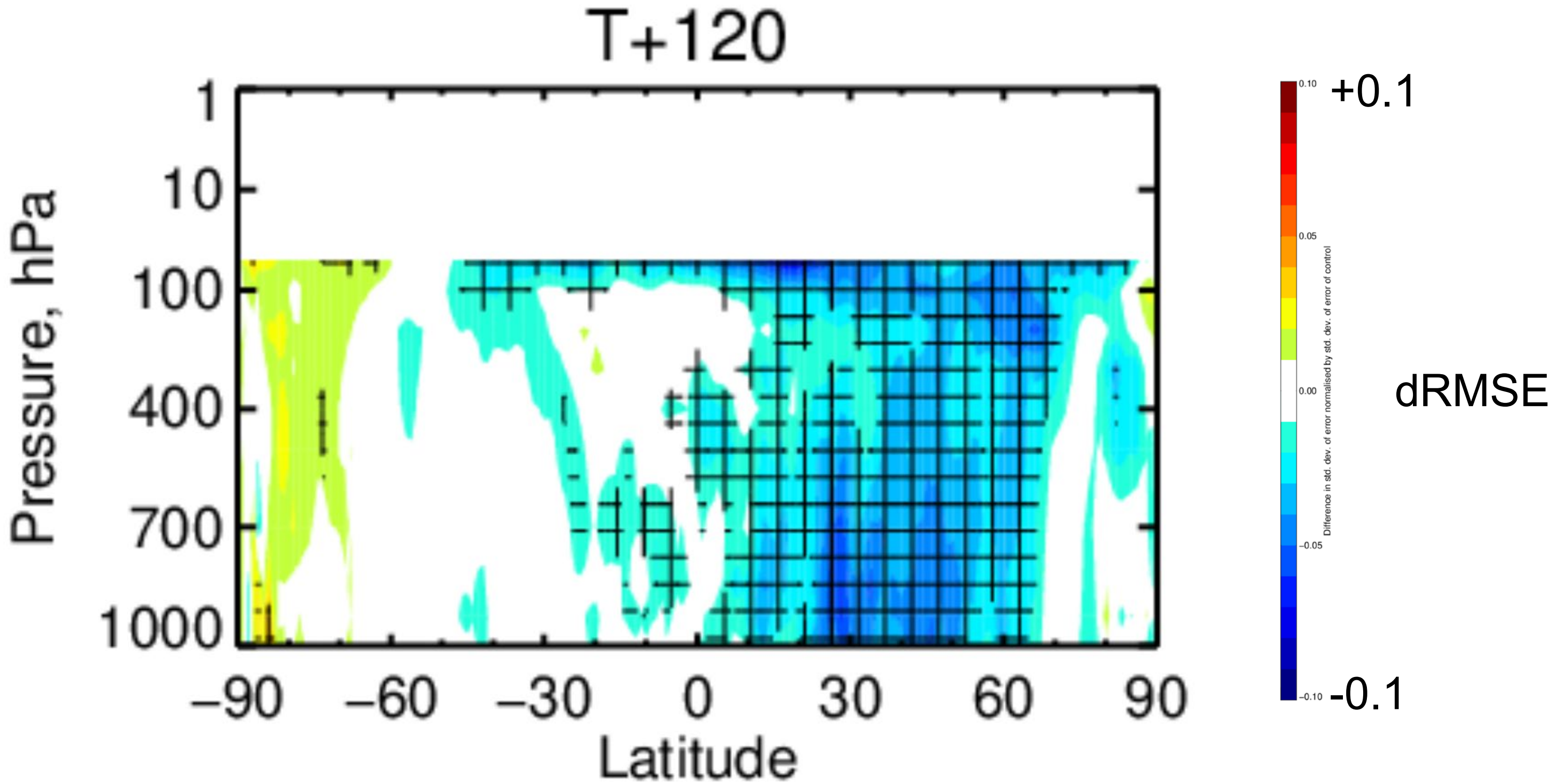
- High-resolution analysis (4.4 km) significantly more expensive than operational-resolution analysis (9 km) → ~3.5x
- Scaling it to 256 nodes on ATOS results in timings ~1.8x longer than operations on 128 nodes → 4.4 km DA likely achievable on next HPC

Task	Relative Cost Factor 4.4 km AN VS 9 km AN
Outer-loop iter 1	3.2
Outer-loop iter 2	3.3
Outer-loop iter 3	3.2
Outer-loop iter 4	4.1

Improved AIFS forecasts from 4.4 km analyses

- AIFS forecasts were initialised 4-times daily
(00, 06, 12, 18 UTC) between Sep 1 - Oct 15, 2023

- [BELOW] Verification of AIFS forecasts of vector wind against
initialising ("own") analysis for 5-day forecast



Conclusions and Outlook

Higher-resolution DA leads to:

- Improved initial conditions (→ better training data for AIFS ?)
- Significant improvement in the forecast skill compared to operational-like setup (6-12 hours of additional forecast skill in the medium-range)
- Enhanced AIFS forecast skill
- Still room for high-res DA setup upgrades!

Outlook:

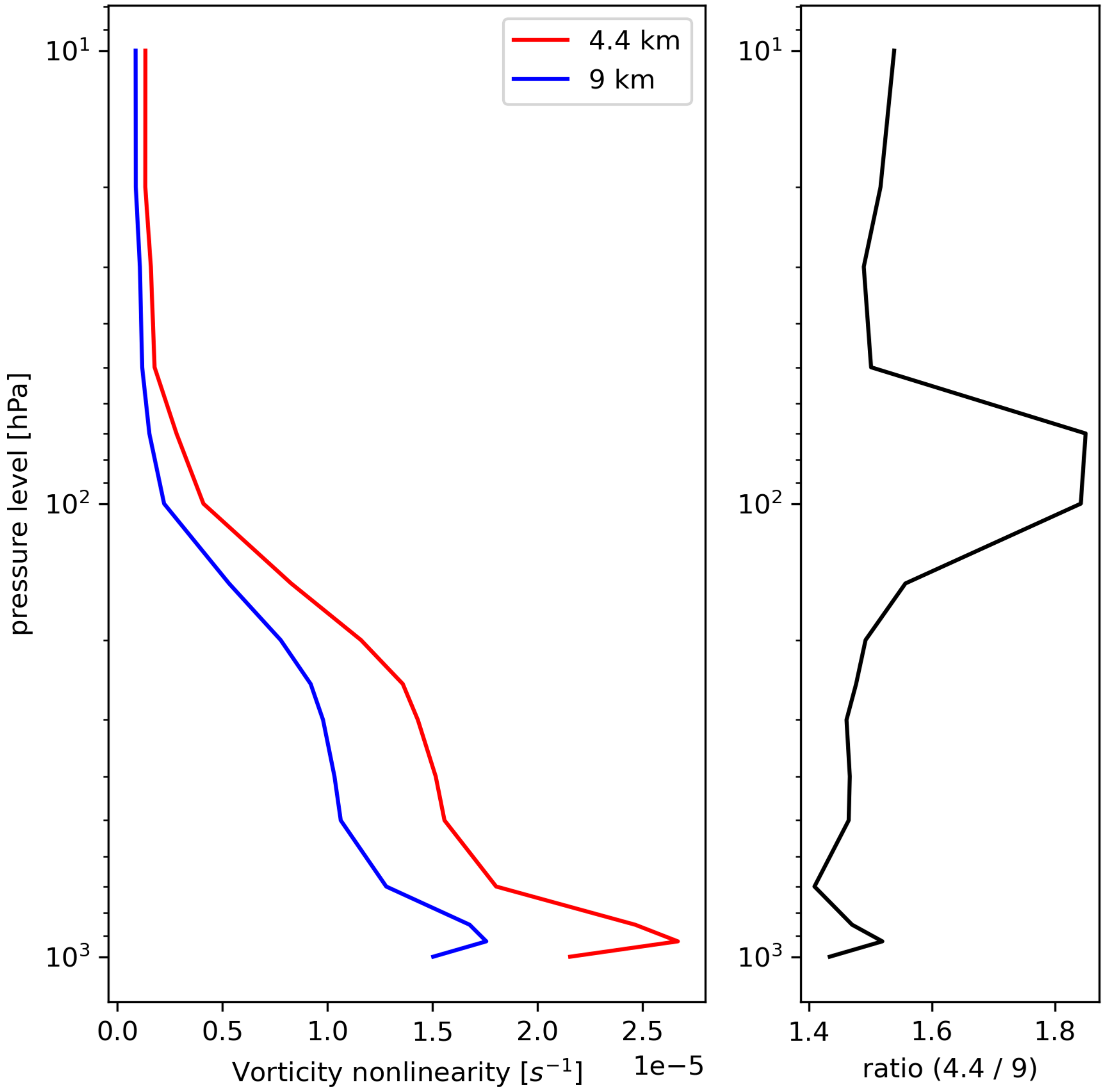
- address the issues in the initial condition in the tropical UTLS
- enhance the resolution of first inner loop!
- further enhance outer loop resolution to TCo3999 (2.8 km) and final inner loop resolution to TCo639 (18 km)

Backup slides

Is the high-res 4D-Var setup optimal?

- Still not!!
- A mismatch between:
  - nonlinear cost function $\mathcal{J}_o$ at TCo2559 (4.4 km) and
  - quadratic cost function J_o at TL319 (55 km)
- Good match between:
  - nonlinear cost function $\mathcal{J}_o$ at TCo2559 (4.4 km) and
  - quadratic cost function J_o at TCo511 (20 km)

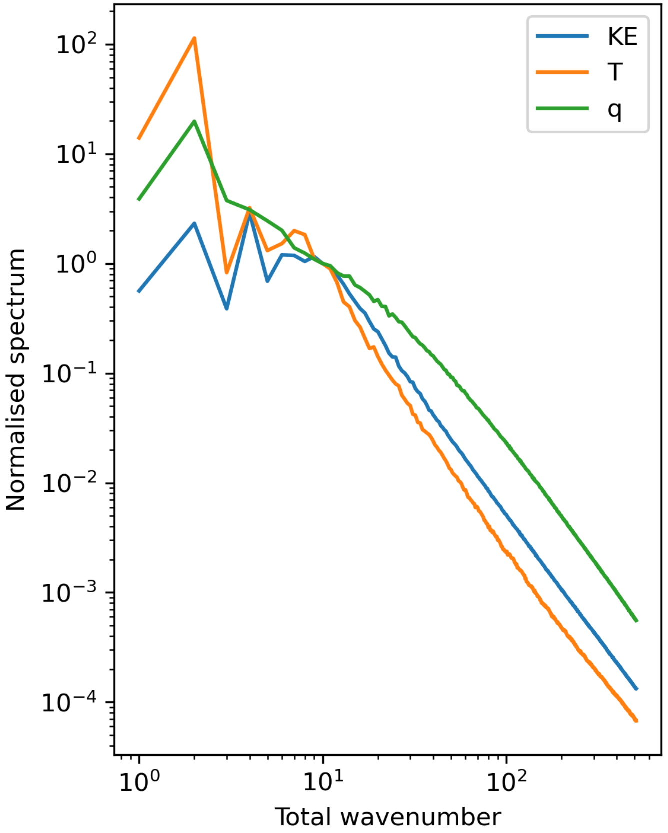
- Possible reasons:
  - Resolution of the first inner-loop is too low?
  - Linearisation hypothesis less valid? → Certainly!



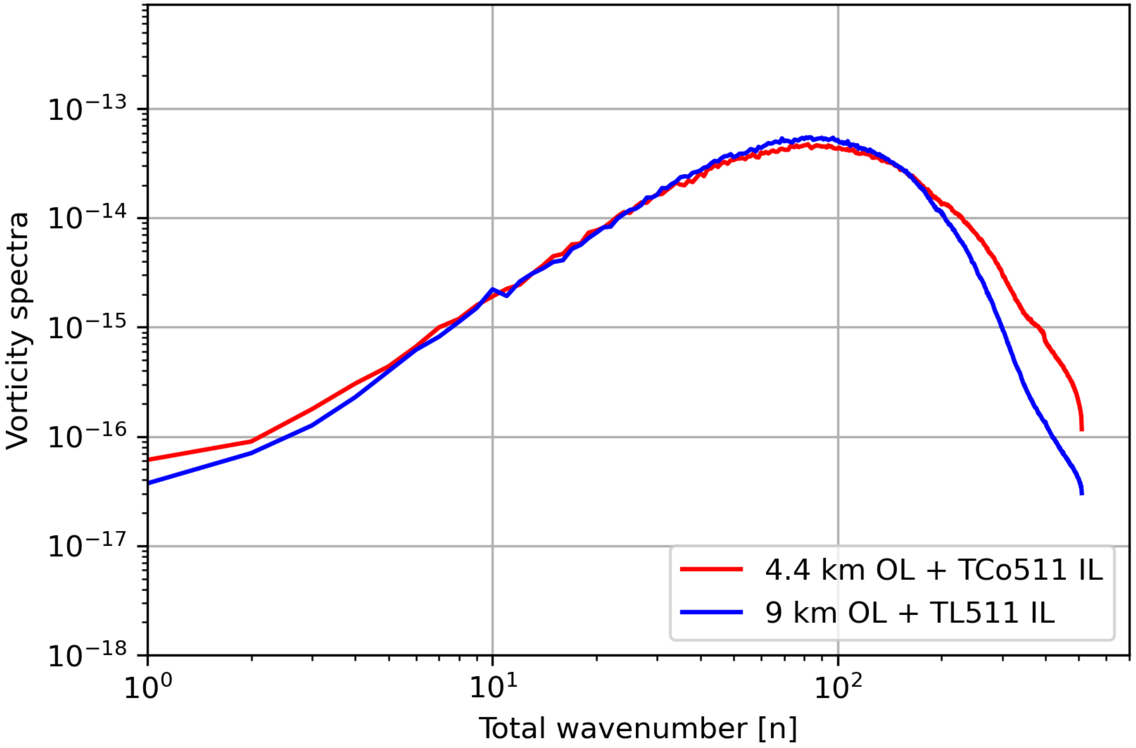
Improved initial conditions

- By employing higher-resolution minimisation (inner loop, i.e. TLM and ADM), we correct greater part of wind and humidity spectra

Wind, temperature, humidity spectra

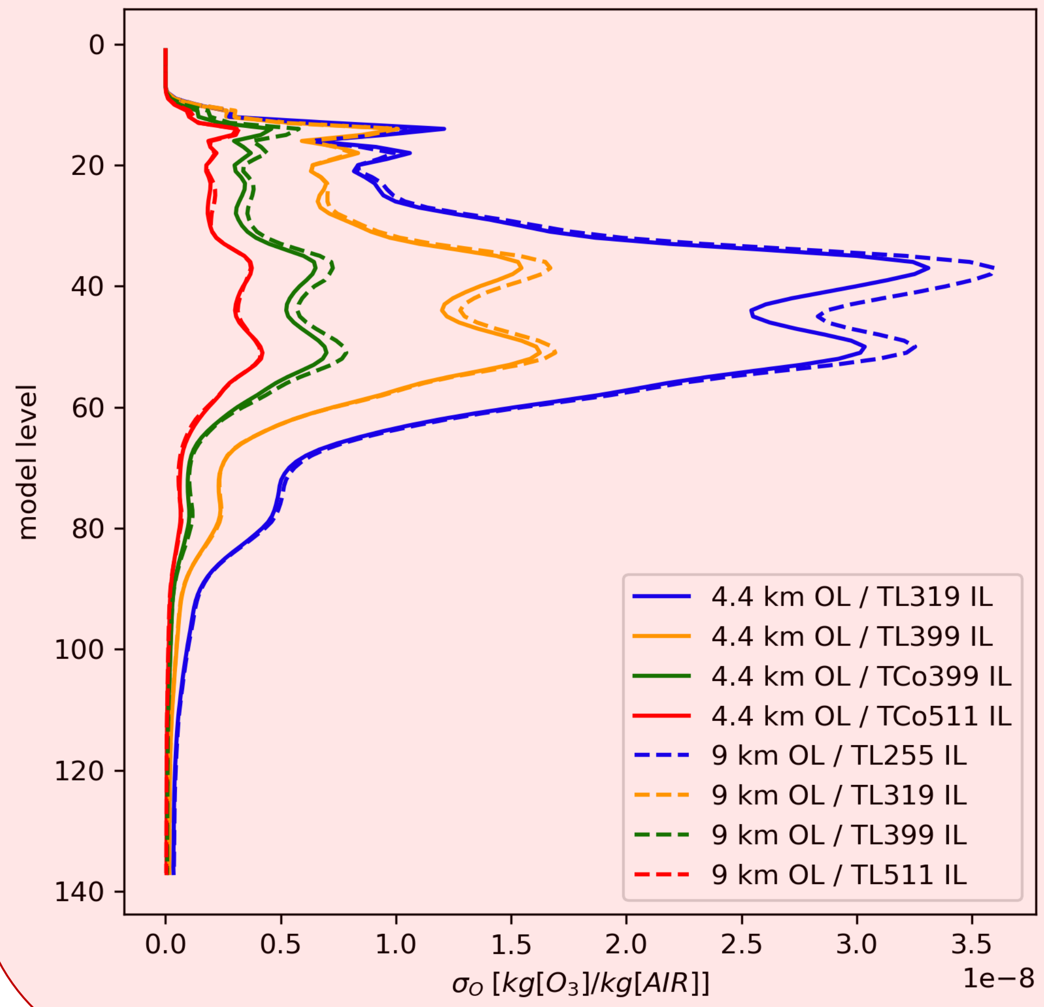


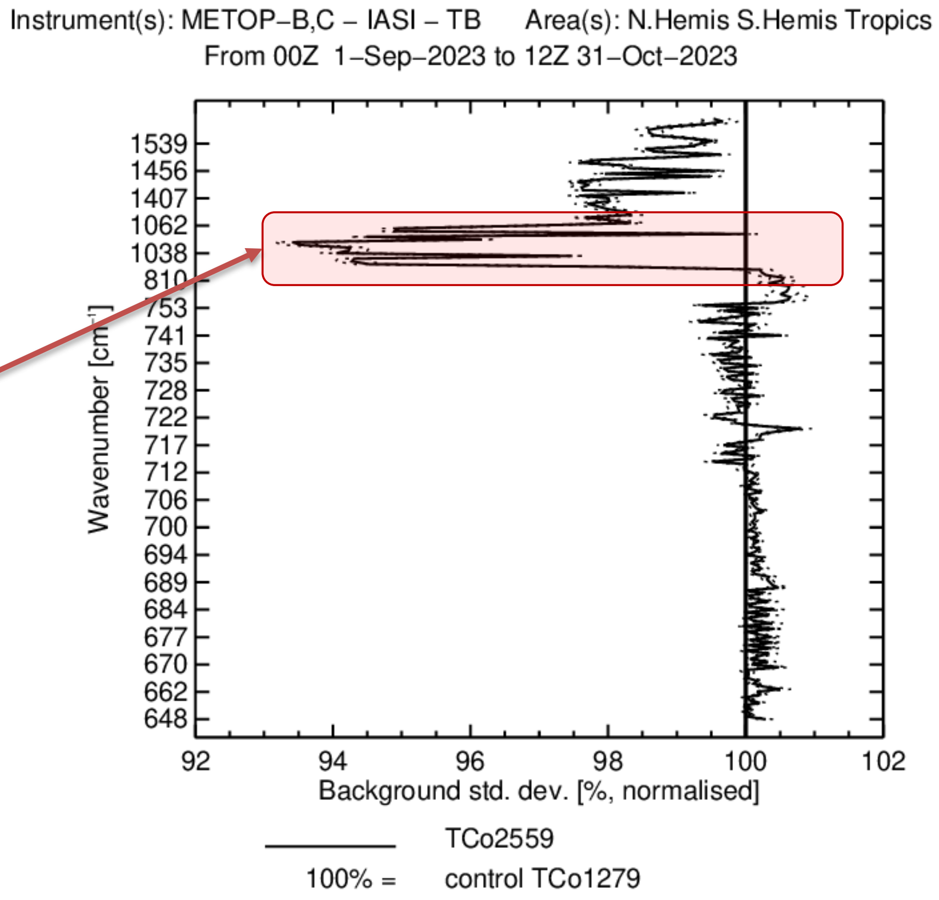
Spectra of relative vorticity analysis increment at last (4th) minimisation



Improved initial conditions

Less correction is applied to ozone fields in 4.4 km DA than in 9 km DA





Improved IFS forecasts

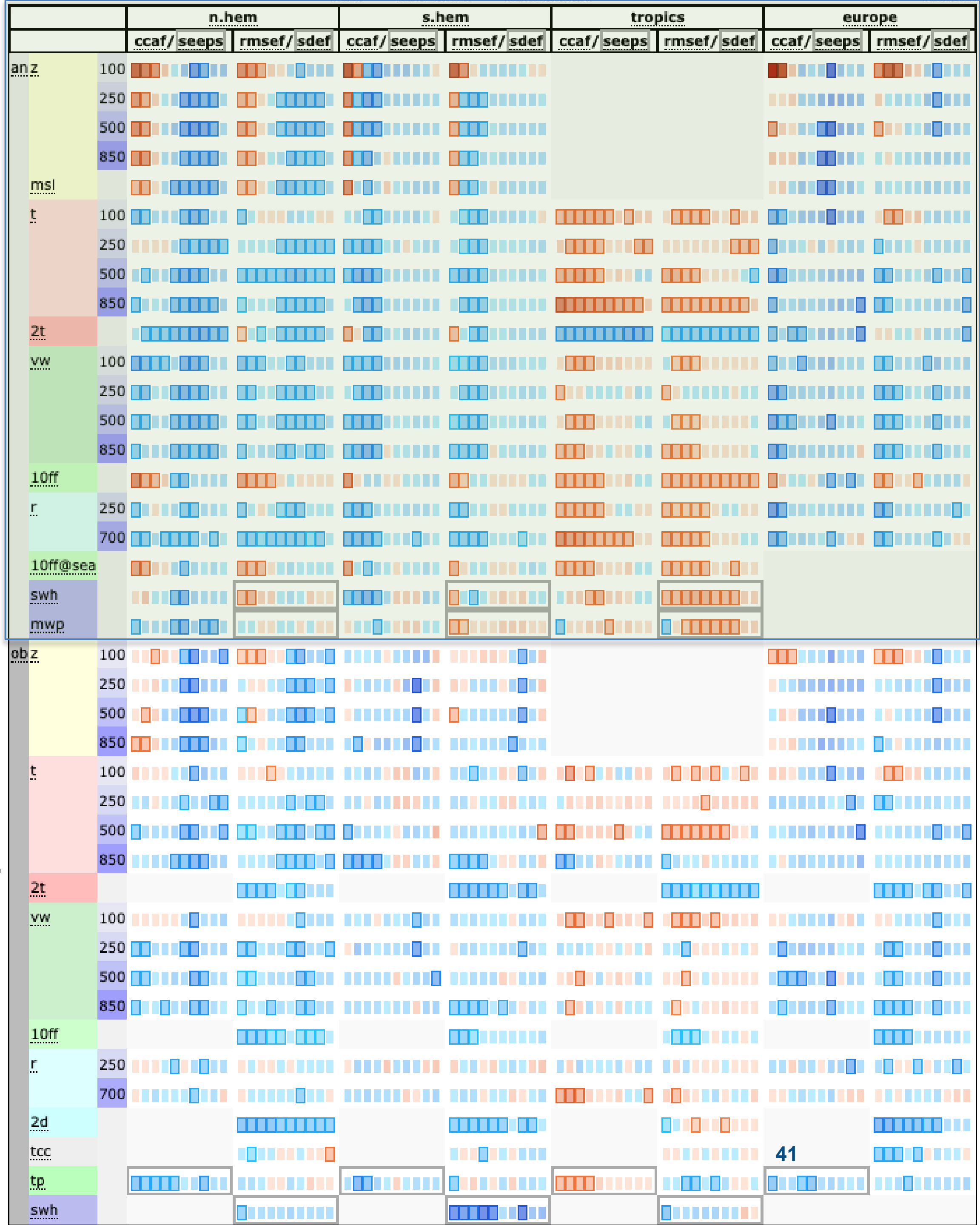
Scorecard:
4.4 km FC from 4.4 km AN vs
9 km FC from 9 km AN

September 1, 2023 – October 31, 2023

Blue means 4.4 km is better.
Red means 9 km is better.

Own analysis verification

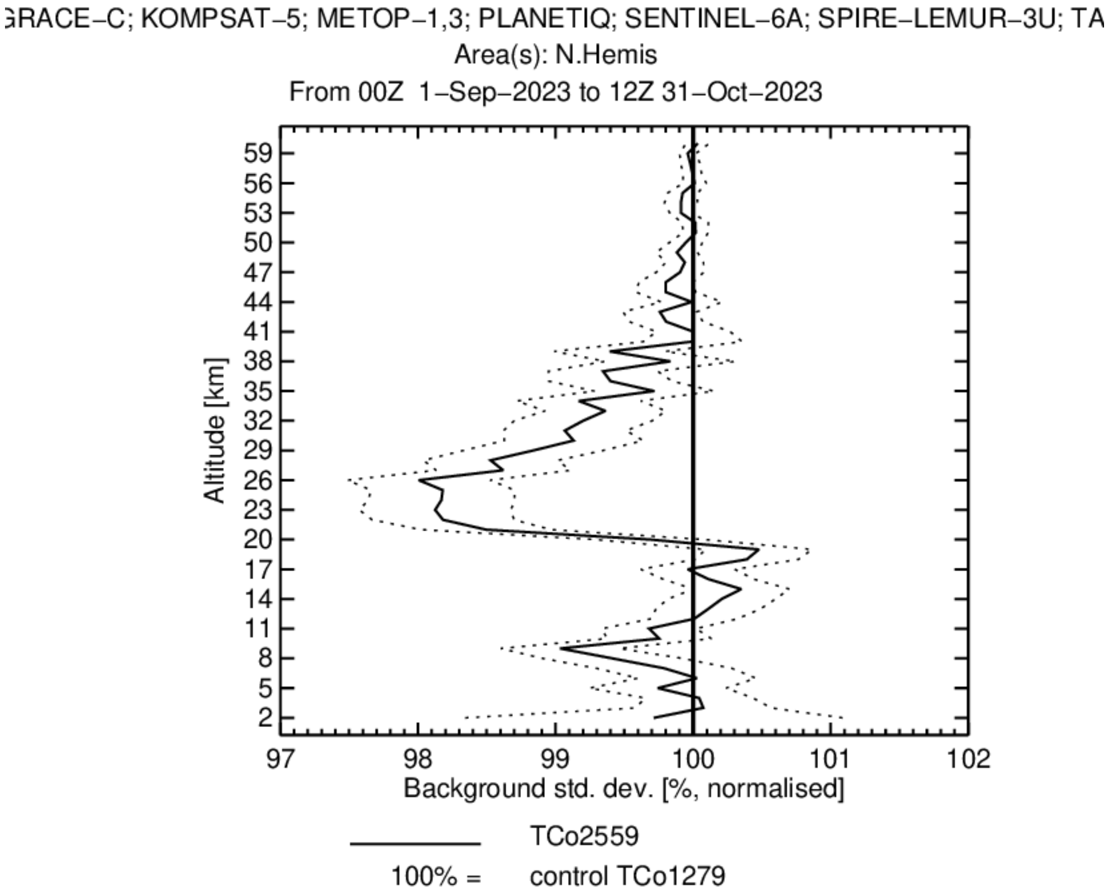
Verification vs. observations



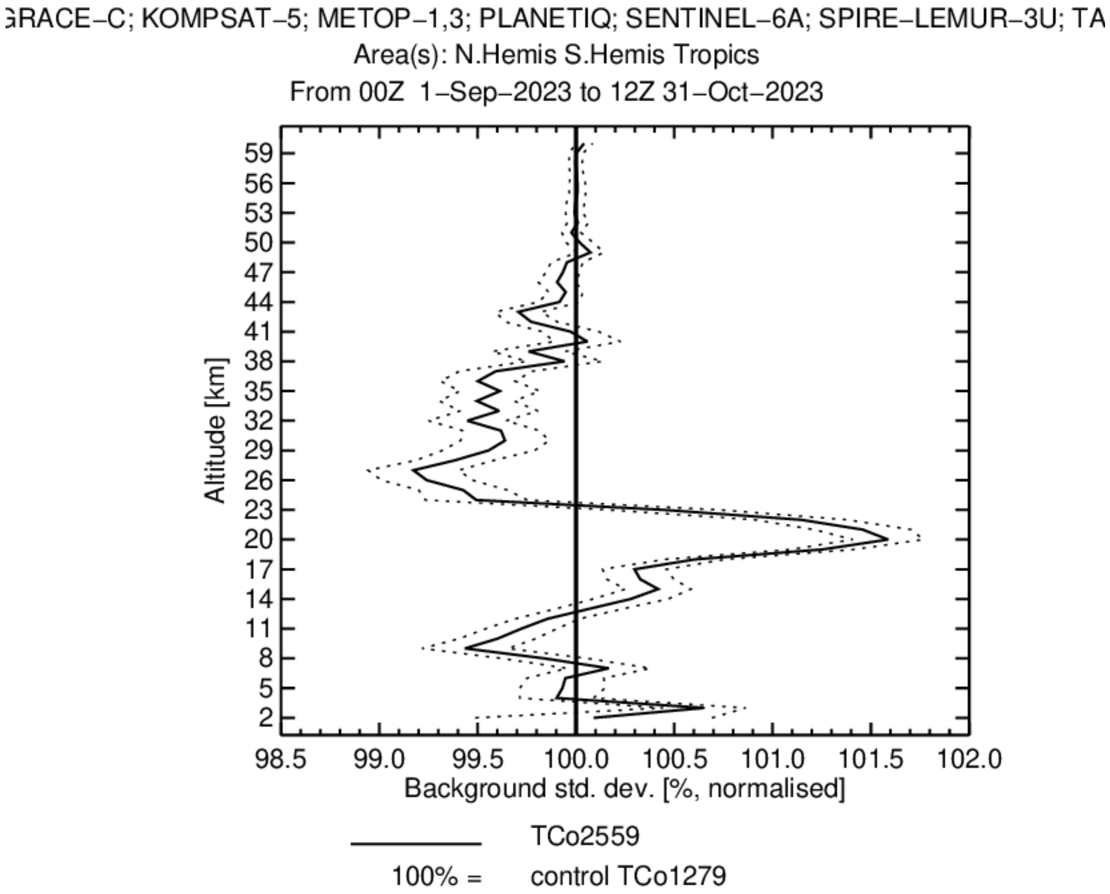
ECMWF

Improved initial conditions

GNSSRO – Northern Hemisphere



GNSSRO – Tropics

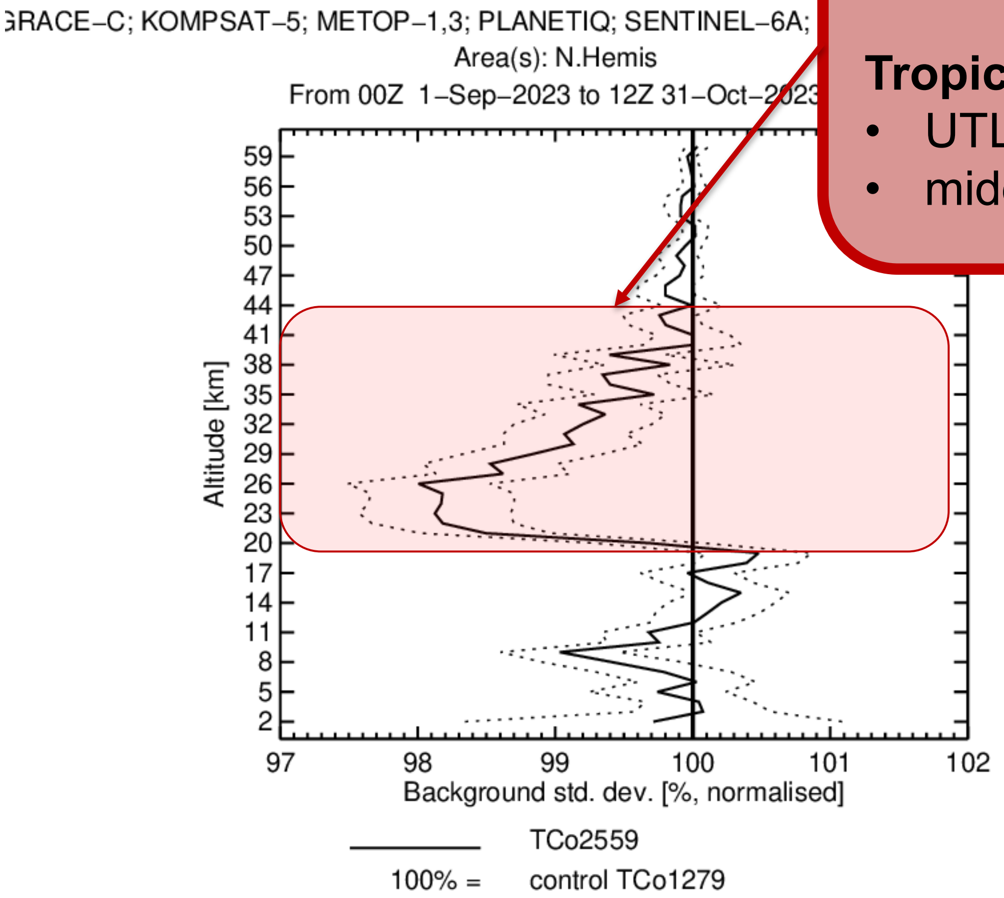


Improved initial conditions

Midlatitudes:
- stratosphere improved

Tropics:
- UTLS degraded,
- middle stratosphere improved

GNSSRO – Northern Hemisphere



GNSSRO – Tropics