ECMWF IFS: from cycle 46r1 to 47r1 and beyond …

Michail Diamantakis

Many thanks to all ECMWF Research Department colleagues who provided input for this talk

41st EWGLAM and 26th SRNWP meeting

Sofia 2019

ECMWF strategy for global weather predictions

- Make skillful ensemble predictions of high impact weather up to two weeks ahead
  - Resolution increase is key in delivering this strategy - stronger ensemble signal with high resolution
- Today's numerical weather prediction system:
  - **Daily: 10 day, 9km horizontal resolution with 137 vertical levels deterministic forecast initialised by 4D-VAR analysis**
  - **Daily: 15 day forecast produced by 51 member 18km, 91 level ensemble with perturbed initial conditions to quantify uncertainty**
  - **Weekly: 46 day forecast (extended range) - 51 member 36km ensemble with 91 levels**
  - **Monthly: 7-months forecast (seasonal) – 51 member 36km ensemble with 91 levels**
  - **Copernicus Atmospheric Monitoring System:**
    - **Daily atmospheric composition forecasts (GHG, aerosols, dust, volcanic plumes etc)**
    - **Global re-analysis**
- 2025
  - Towards a seamless ENS prediction system with high horizontal (near 5km) and vertical resolution (137L)

IFS model

- Global hydrostatic primitive equation dynamical core with non-hydrostatic option available
- Spectral transform in the horizontal with vertical finite element discretization, semi-implicit, semi-Lagrangian timestepping:
  - stable at long timesteps (very efficient) and very accurate
- Initial conditions produced by 4D-VAR data assimilation system that relies on a tangent-linear and adjoint version of the non-linear model
- Continuous improvement of skill with release of a new operational cycle roughly once per year
  - CY45R1 June 2018
  - CY46R1 June 2019
  - CY47R1a Q2 2020 … ported on new super-computer in Bologna
  - CY48R1: single precision forecasts with increased vertical and horizontal ENS resolution subject to super-computer upgrade

Major 4D-VAR, ENS and model changes in CY46R1

- Continuous Data Assimilation
- Ensemble of Data Assimilations (EDA) increased from 25 members to 50 members
- New ENS initialization from 50-member EDA – members now exchangeable
- Weakly coupled data assimilation introduced for sea-surface temperature in the tropics
- Compute Simplified Extended Kalman Filter soil analysis Jacobians from EDA
- 1-hour radiation update frequency (timestep) in the ENS (same as in the 9km model)
- Convection scheme improvements (entrainment, CAPE closure, shallow convection).
- New version on wave physics based on Ardhuin et al 2010
- New 3D climatology replaced 2D CAMS aerosol climatology
- 2m diagnostic temperature improvements

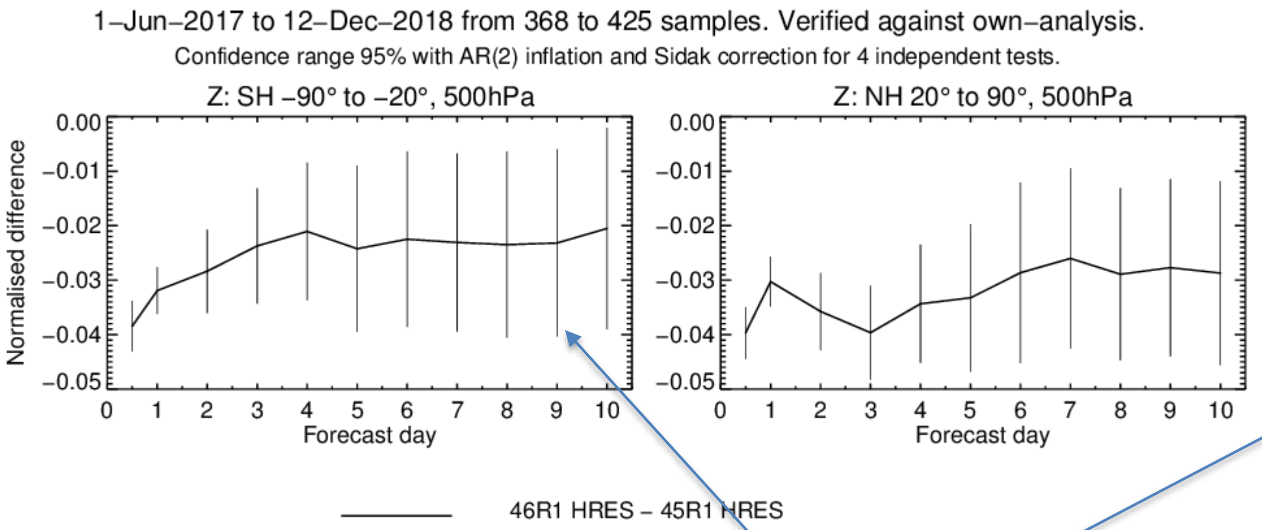
Furthermore, there have been additional important observational changes. Full details in:

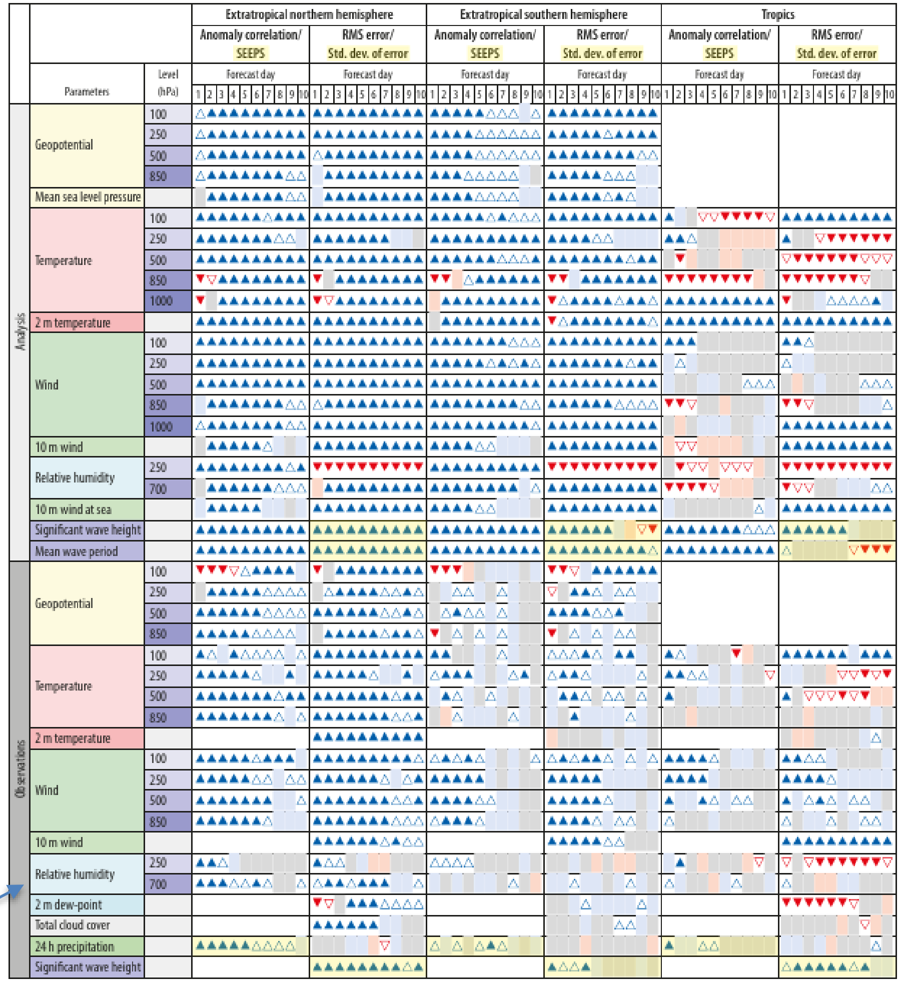
https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+cycle+46R1#ImplementationofIFScycle46R1-Meteorologicalcontentofthenewcycle

IFS 46r1: a cycle that clearly improves the forecast skill

Deterministic 9km (HRES) forecast:

- Reduction in RMSE and increase in ACC in geopotential height, MSL, winds, temperature, humidity, wave parameters, surface variables against own analysis and observations
- Improvements in stratosphere, upper air and surface





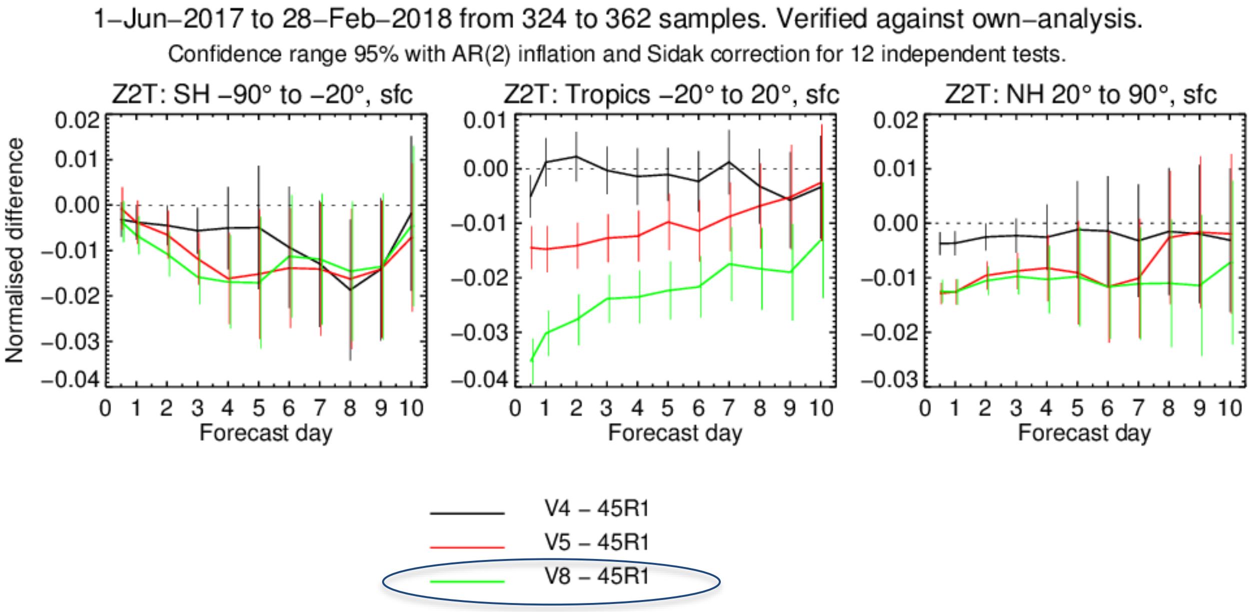
Negative or blue: better forecast

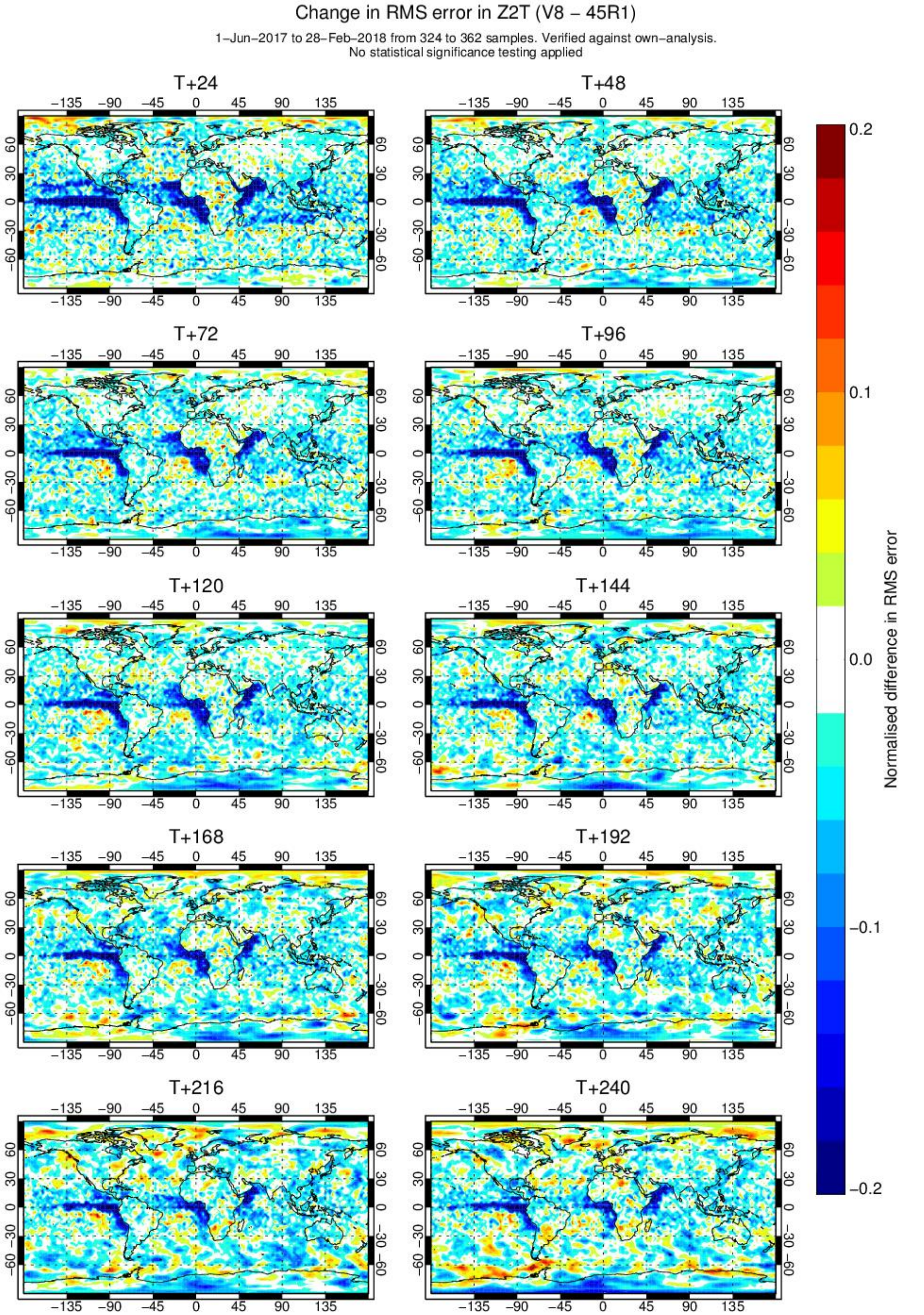
Symbol legend: for a given forecast step...

- ▲ 46r1 better than 45r1 statistically significant with 99.7% confidence
- △ 46r1 better than 45r1 statistically significant with 95% confidence
- 46r1 better than 45r1 statistically significant with 68% confidence
- no significant difference between 45r1 and 46r1
- 46r1 worse than 45r1 statistically significant with 68% confidence
- ▽ 46r1 worse than 45r1 statistically significant with 95% confidence
- ▼ 46r1 worse than 45r1 statistically significant with 99.7% confidence

FIGURE 2 HRES scorecard of IFS Cycle 46r1 versus IFS Cycle 45r1, verified by the respective analyses and observations at 00 and 12 UTC, based on 690 forecast runs in the period June 2017 to June 2019.

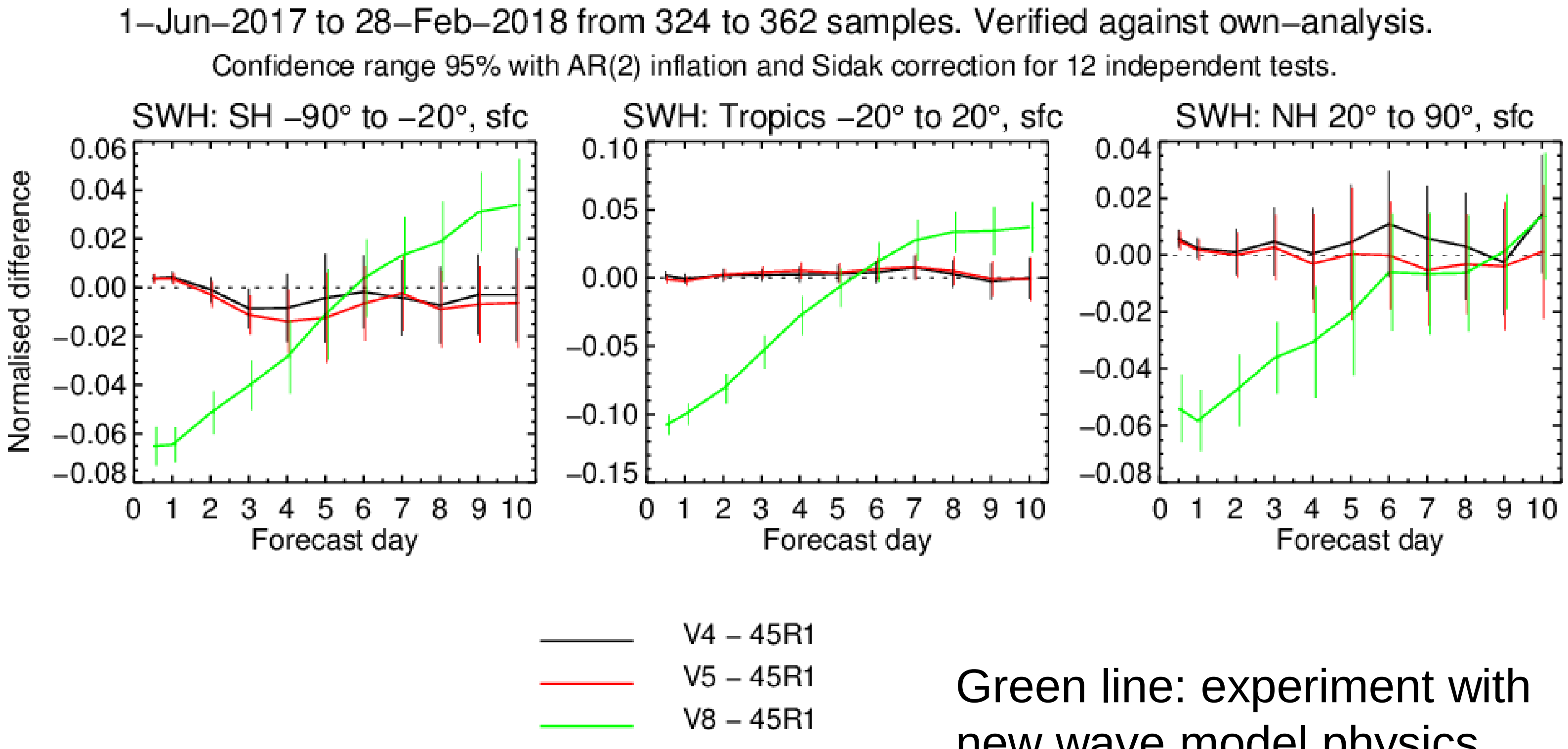
Improvements in 2m temperature





- Tco399 research experiment
- Improved 2m T: green line is the combined impact of:
  - *adjusting wet skin tile conductivity*
  - *tropical SST weakly coupled data assimilation (very positive in the tropics): replace OSTIA analysis by OCEAN5 analysis*

Wave model improvements: impact on "significant wave height" (SWH)



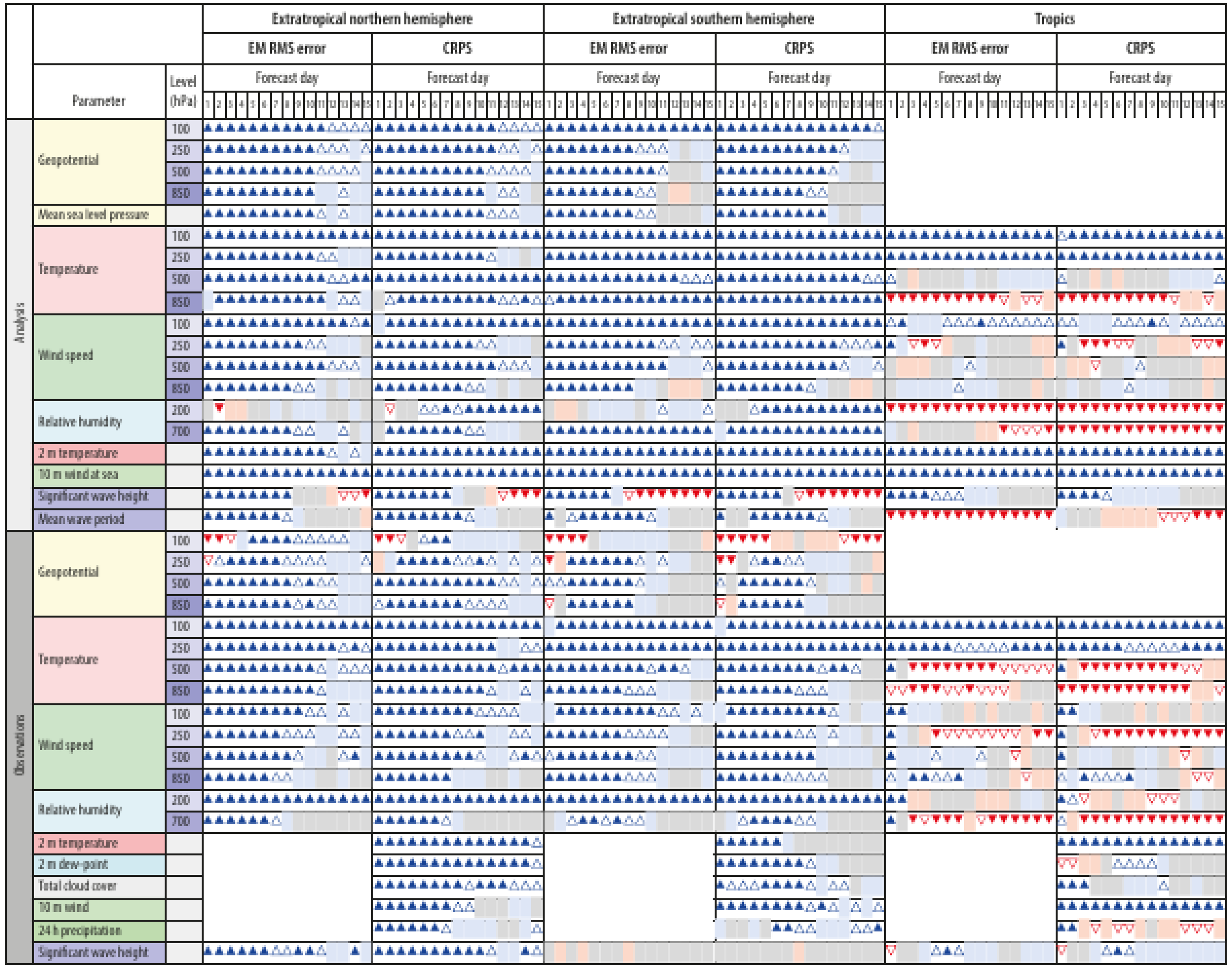
Green line: experiment with new wave model physics

- *Green line shows the positive impact of wave model changes in SWH (significant wave height)*
- *Increased RMSE error at +6d and beyond is due to increased model activity*
- *ACC remains positive through the forecast range (not shown here)*

Improvements in the Ensemble Prediction System forecast

EPS improvements due to combined effect of:

- Model improvements
- Increase frequency of radiation update to 1 hour from 3 hours
- Switch to 50-member EDA which abolishes the old symmetric +/- initialisation from 25 member and makes EDA members exchangeable



Symbol legend: for a given forecast step...

- ▲ 46r1 better than 45r1 statistically significant with 99.7% confidence
- △ 46r1 better than 45r1 statistically significant with 95% confidence
- 46r1 better than 45r1 statistically significant with 68% confidence
- no significant difference between 45r1 and 46r1
- 46r1 worse than 45r1 statistically significant with 68% confidence
- ▽ 46r1 worse than 45r1 statistically significant with 95% confidence
- ▼ 46r1 worse than 45r1 statistically significant with 99.7% confidence

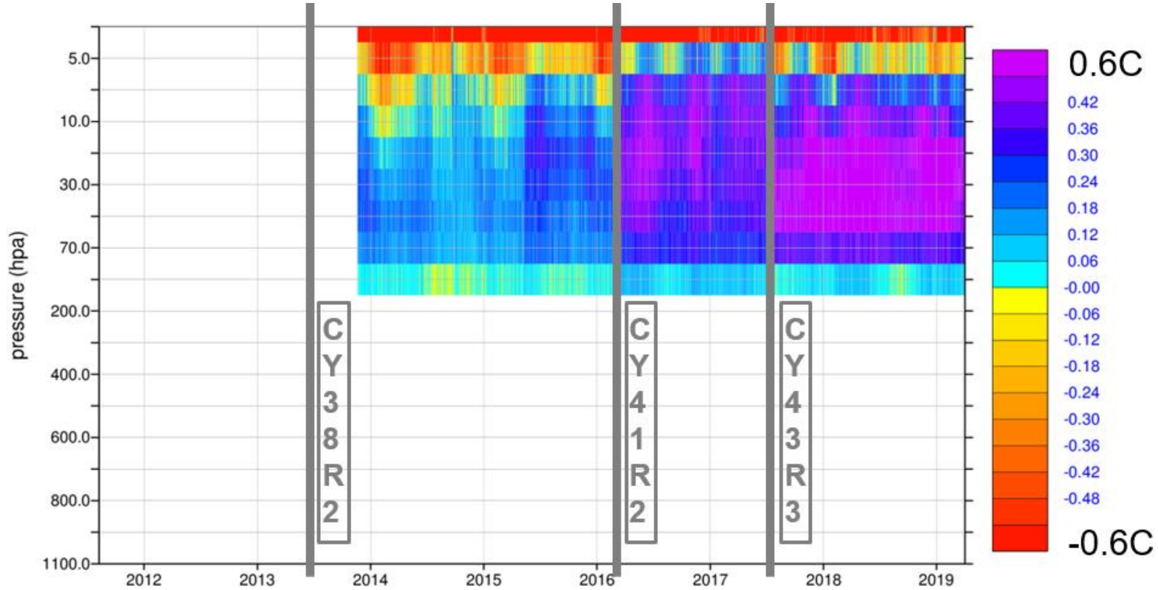
FIGURE 1 ENS scorecard of IFS Cycle 46r1 versus IFS Cycle 45r1 for medium-range forecasts up to forecast day 15, verified by the respective analyses and observations at 00 UTC based on 282 ENS forecast runs in the period June 2017 to June 2019.

CY47R1 : main 4D-VAR and model changes

- New weak-constraint 4D-Var for the stratosphere
- Continuous LWDA - use first guess from Early Delivery analysis
- Matching inner-outer loop timesteps of 450s in last minimization for compatible gravity wave speeds

- Radiation improvements (new MODIS albedo climatology depends on sun angle)
- Limiting Charnock parameter for high winds to improve Hurricane wind/pressure forecasts
- Quintic vertical interpolation in the semi-Lagrangian advection for temperature and humidity **improves temperature bias** in the tropics and reduces its resolution sensitivity (**see talk on numerical method developments on Wednesday**)

Temperature bias in the operational IFS model

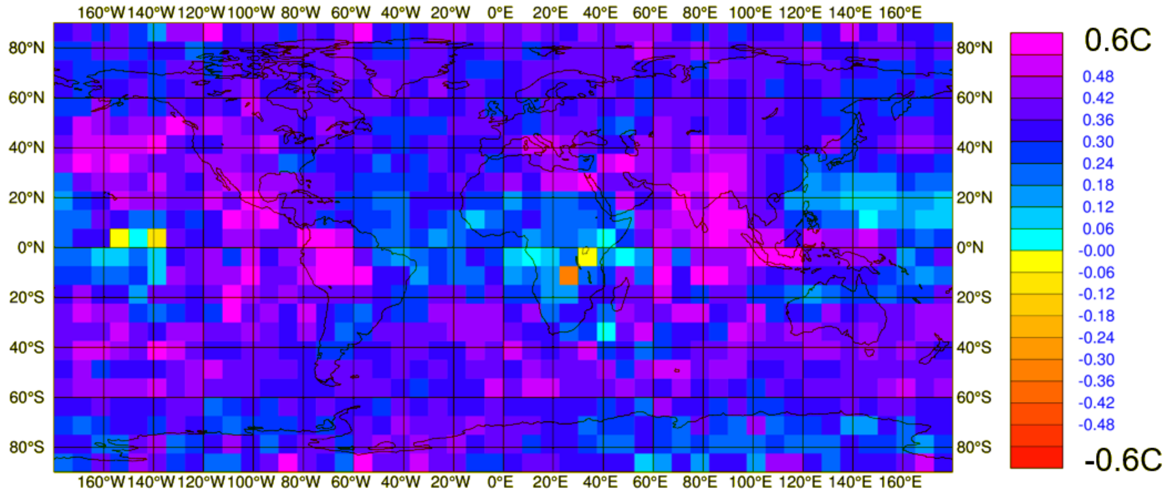


The short-term model bias is estimated by comparing the 12-hour first-guess trajectory with GPS-RO temperature retrievals

→ bias increased with last horizontal resolution (Tco1279 in CY41R2)

→ bias increased with last radiative scheme (CY43R3)

→ GPS-RO reveals the large scale structure of the model error in the stratosphere



Bias between 70hPa and 100hPa (1 Oct. 2018 to 1 Feb. 2019)

New model error specification in weak-constraint 4D-Var

Unknown forcing is introduced in the model equations (additive, Gaussian, constant within the assimilation window, no cross-correlation with the background error)

$$\mathbf{x}_k = \mathcal{M}_{k,k-1}(\mathbf{x}_{k-1}) + \eta \qquad \text{for} \qquad k = 1, \dots, N.$$

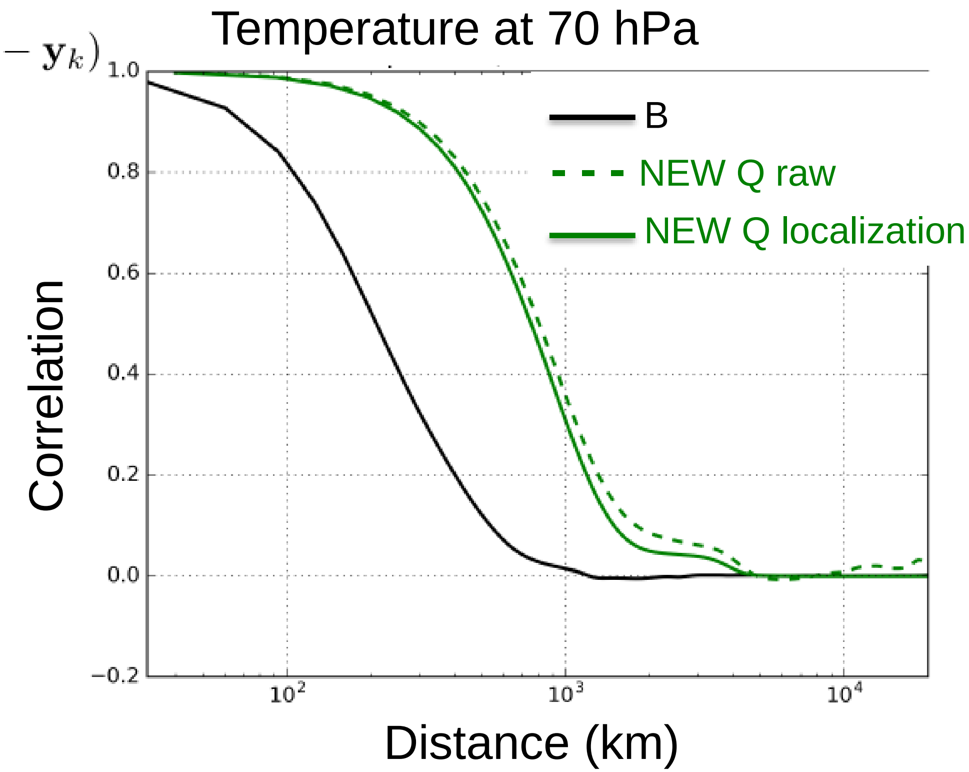
Observations can be fitted by changing the **initial state** or the **model forcing**

$$\begin{aligned} J(\mathbf{x}_0, \boldsymbol{\eta}) &= \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right)^{\mathrm{T}} \mathbf{B}^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^b\right) \\ &+ \frac{1}{2}\left(\boldsymbol{\eta} - \boldsymbol{\eta}^b\right)^{\mathrm{T}} \mathbf{Q}^{-1} \left(\boldsymbol{\eta} - \boldsymbol{\eta}^b\right) \\ &+ \frac{1}{2}\sum_{k=0}^{N} \left(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k\right) \end{aligned}$$

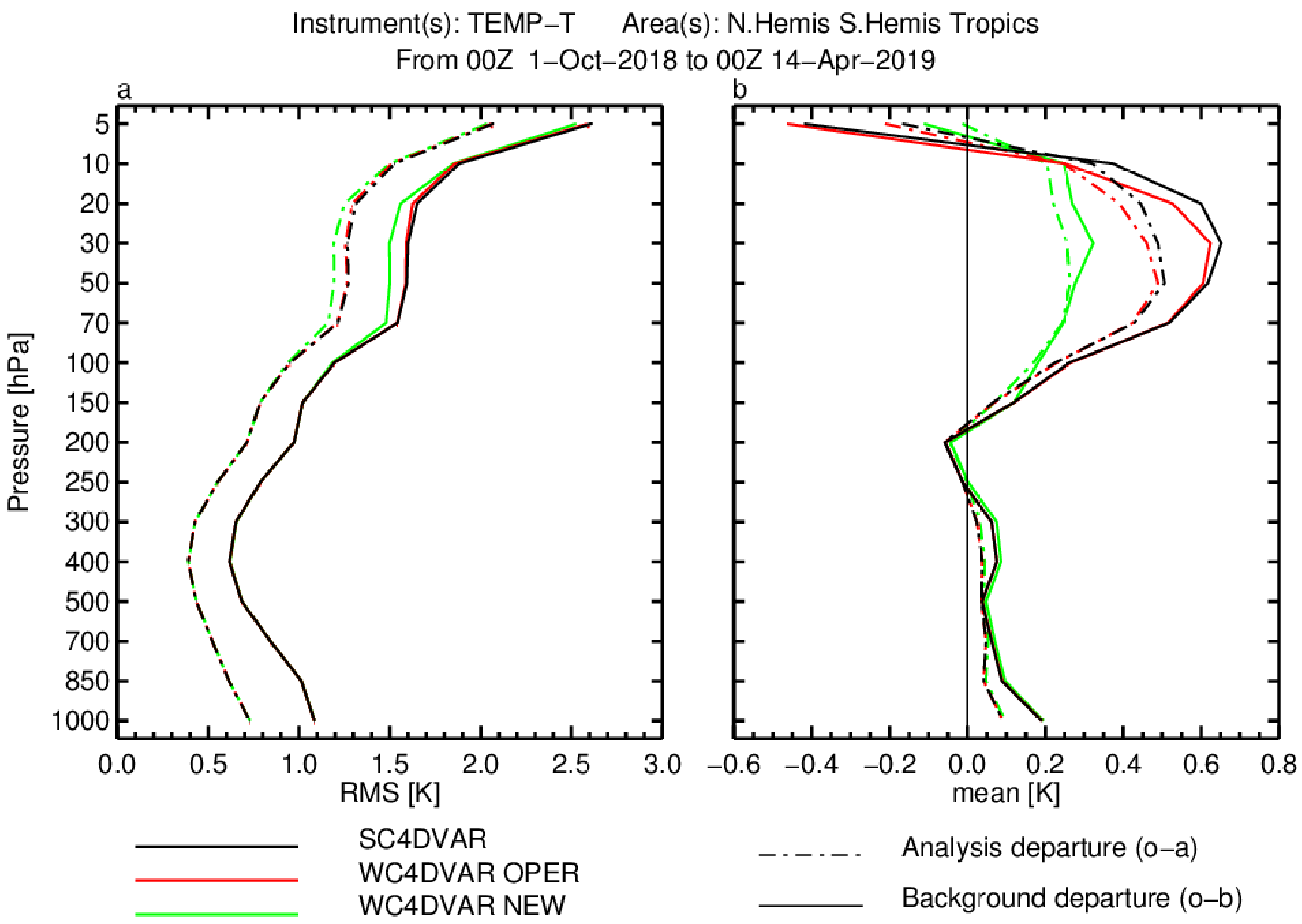
Horizontal correlation in the covariance matrices

→ **B** corrects the small scales

→ **Q** corrects the large scale



Weak-constraint 4D-Var with scale separation

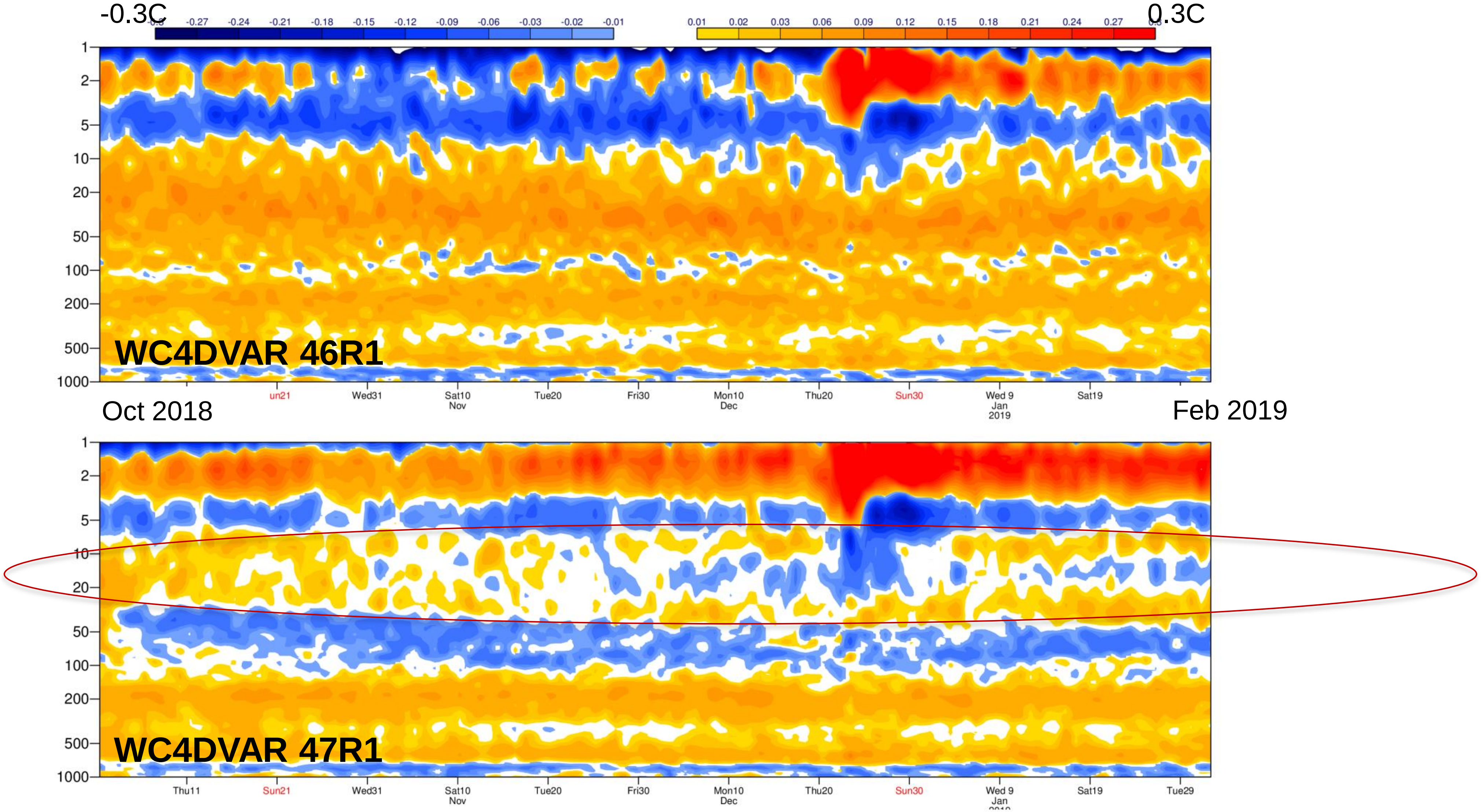


→ temperature bias is reduced up to 50% with respect to radiosondes and GPS-RO (RMSE is reduced up to 6%)

→ implemented in CY47R1 above 100hPa where the model bias is significant

Extra slide: Weak-constraint 4D-Var with scale separation

Timeseries of temperature increments δx



In 47R1, the stratospheric mean increment is closer to zero. Systematic error is corrected by the model forcing and random error is corrected in the initial condition

Matching timestep with non-linear model in the last minimization of 4D-VAR

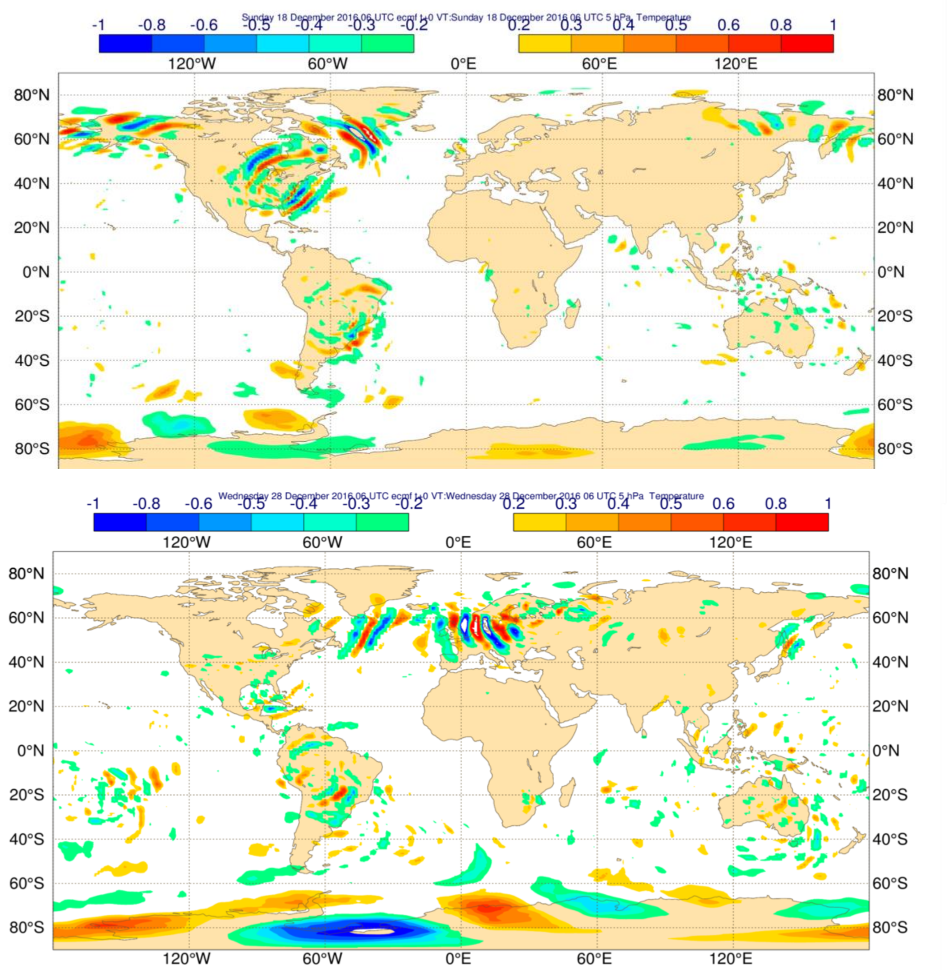
Currently:

- 4 outer loops in minimization all at lower resolution than forecast model
- Due to coarser resolution a longer timestep can be used for TL/AD models

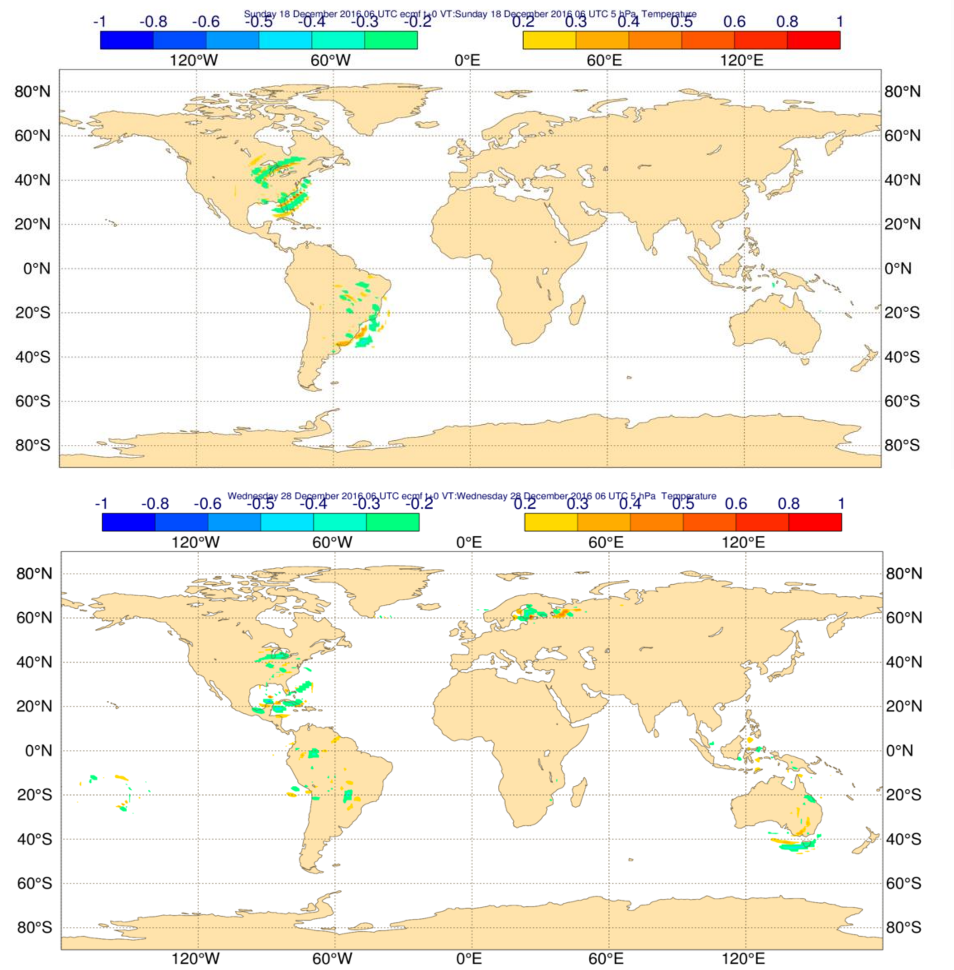
In CY47R1:

- Aiming to use the same timestep with the 9km forecast model (450s) at the last outer loop minimization
- Although more expensive it ensures consistency between the inner and outer loops which should make the speed of gravity waves more comparable.
- It has positive impact in scores increasing the overall minimization computational cost by 8%.

Difference between analysis increments computed by **nonlinear and linearised** models 9 hours in the assimilation window (**temperature ~5 hPa**)



Nonlinear Model tstep=450s; Linearised Model tstep=900s

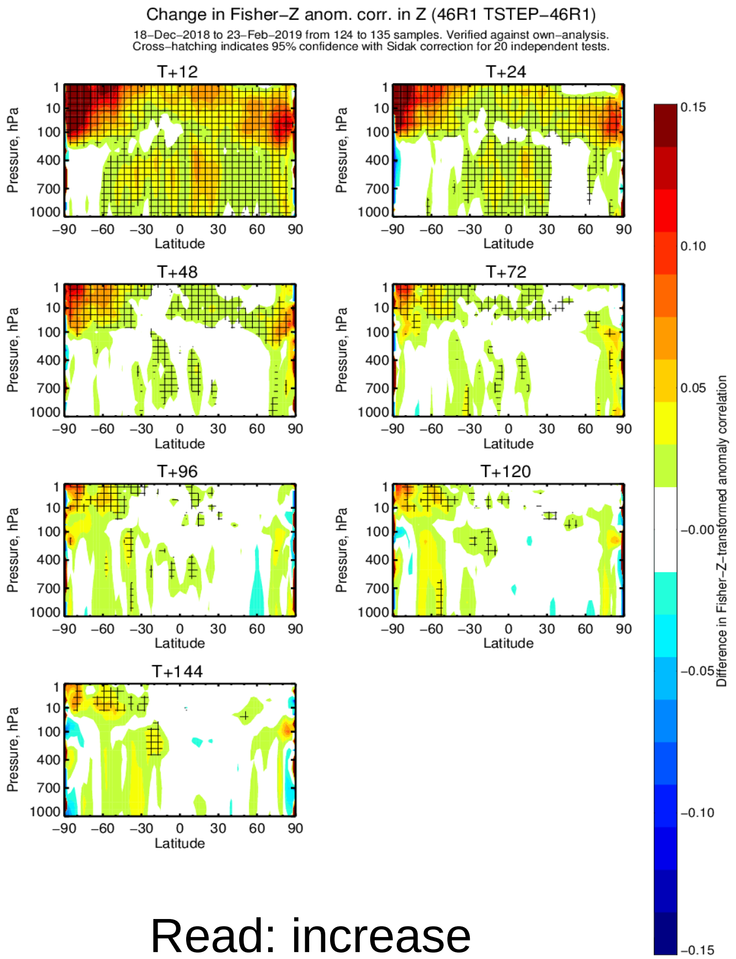


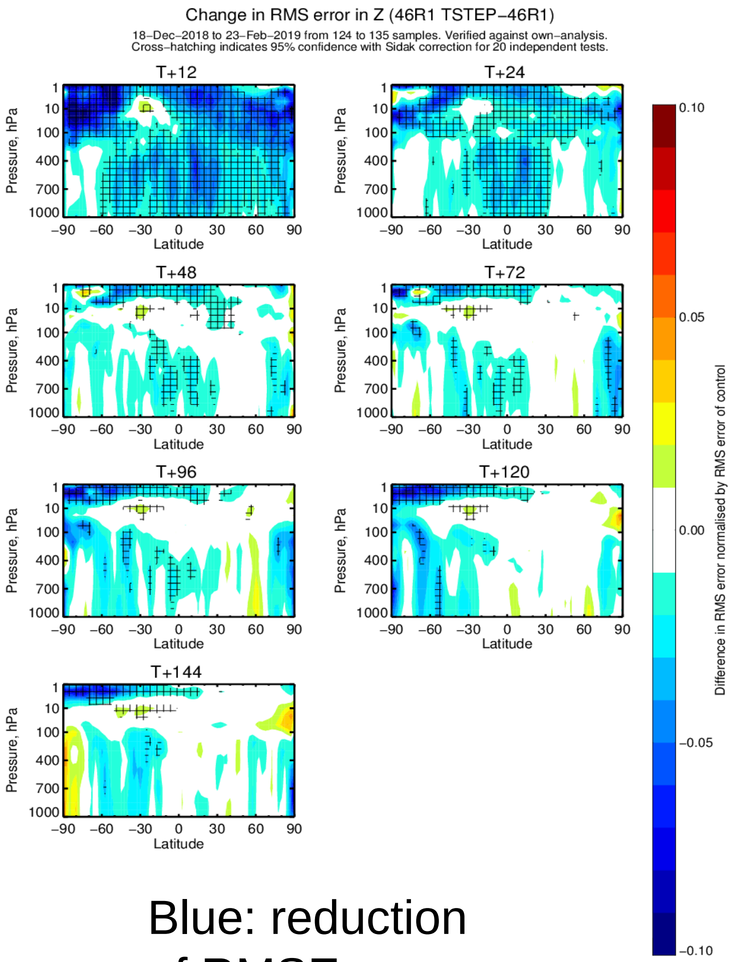
Nonlinear Model tstep=450s; Linearised Model tstep=450s

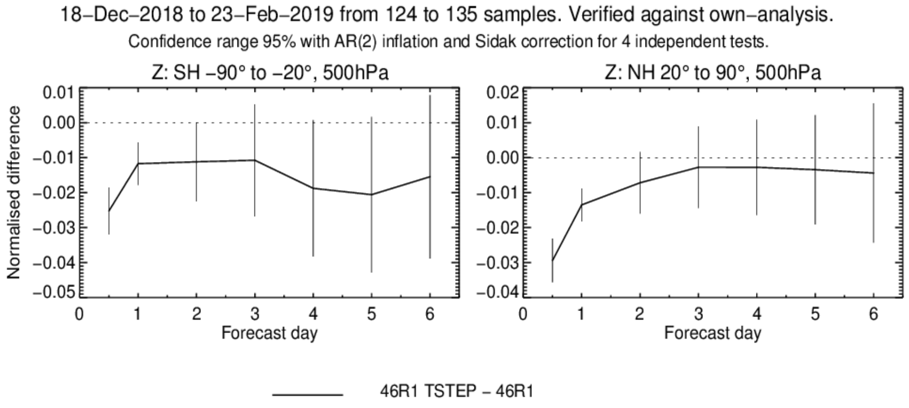
Slides courtesy of M. Bonavita

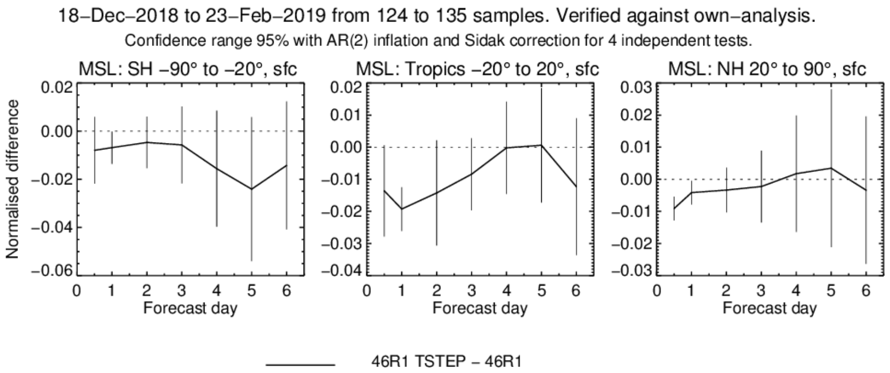
Inner/Outer loop time-step matching

- CY46R1 experimentation at HRES resolution (18/12/2018-23/02/2019)







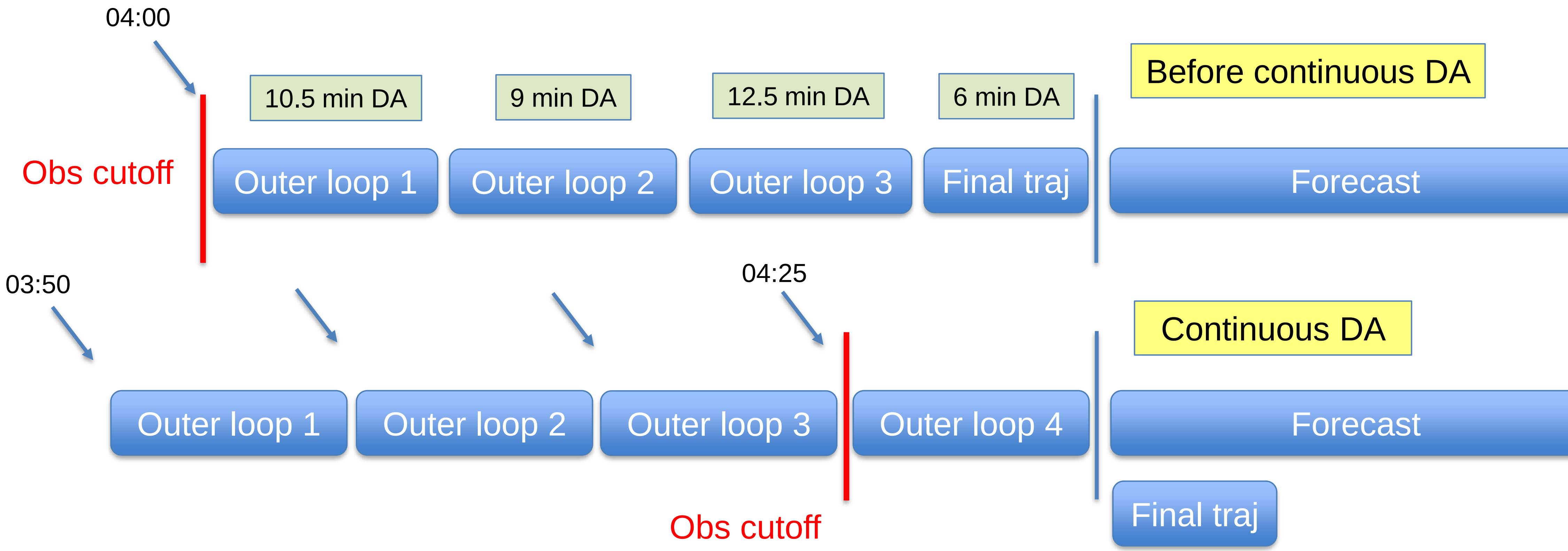


Read: increase in ACC against control

Blue: reduction of RMSE against control

Continuous Data Assimilation changes Early Delivery (DA) in CY46R1

- New observations added in each outer loop
- Early Delivery assimilation window from 6h to 8h, ensures **all** observations that have arrived can be assimilated (LWDA unchanged at 12h)
- Outer loops from 3 to 4
- Allows to assimilate more observations improving analysis quality



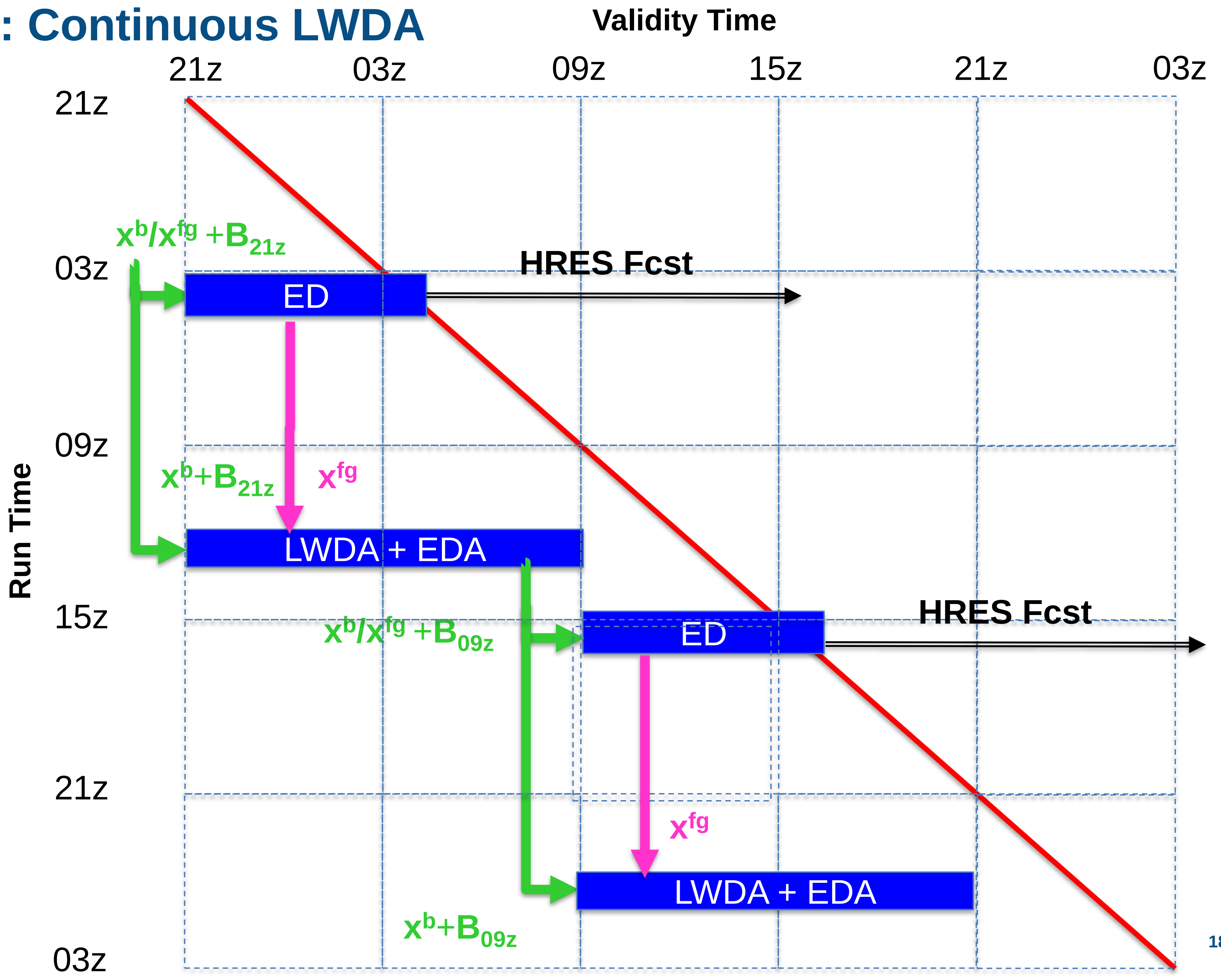
ECMWF

Slide courtesy of P. Lean

DA schedule 47R1: Continuous LWDA

Currently the "background" forecast provides "first guess" for LWDA

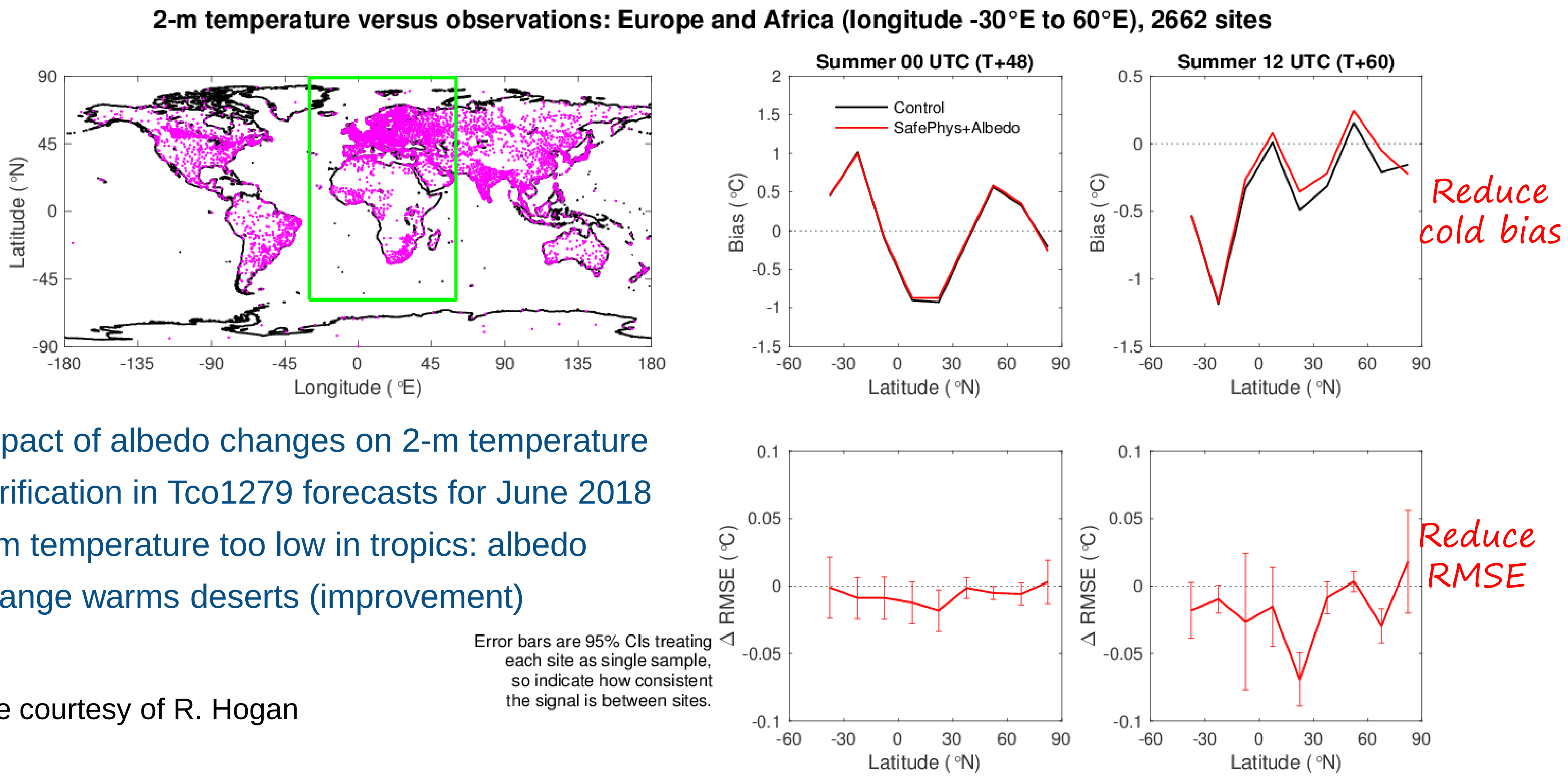
With continuous LWDA the ED analysis is used as "first guess" for the next LWDA: the LWDA analysis is warm-started from a better first guess.



Slide courtesy of E. Holm

ECMWF

Model improvements: Albedo changes for CY47R1



- Impact of albedo changes on 2-m temperature verification in Tco1279 forecasts for June 2018 2-m temperature too low in tropics: albedo change warms deserts (improvement)

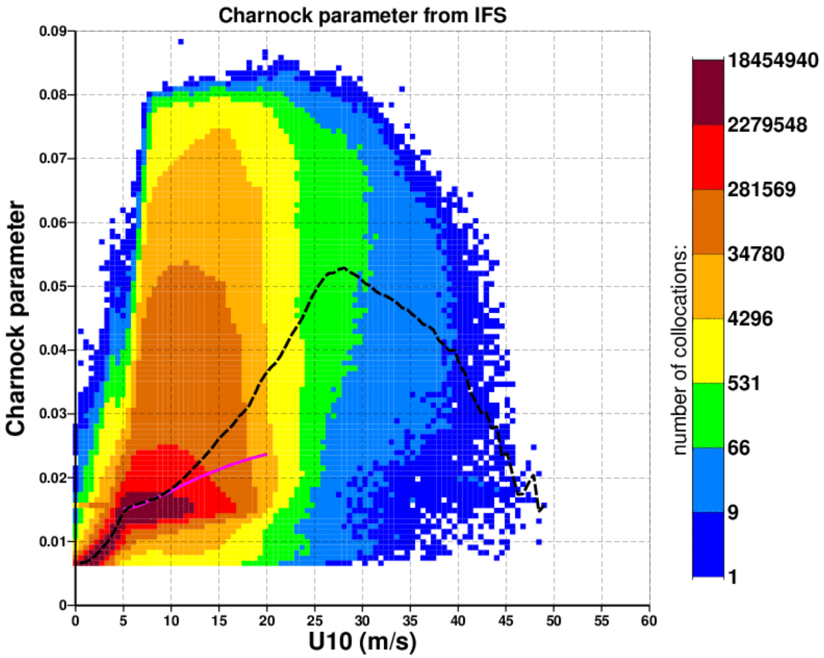
Error bars are 95% CIs treating each site as single sample, so indicate how consistent the signal is between sites.

Slide courtesy of R. Hogan

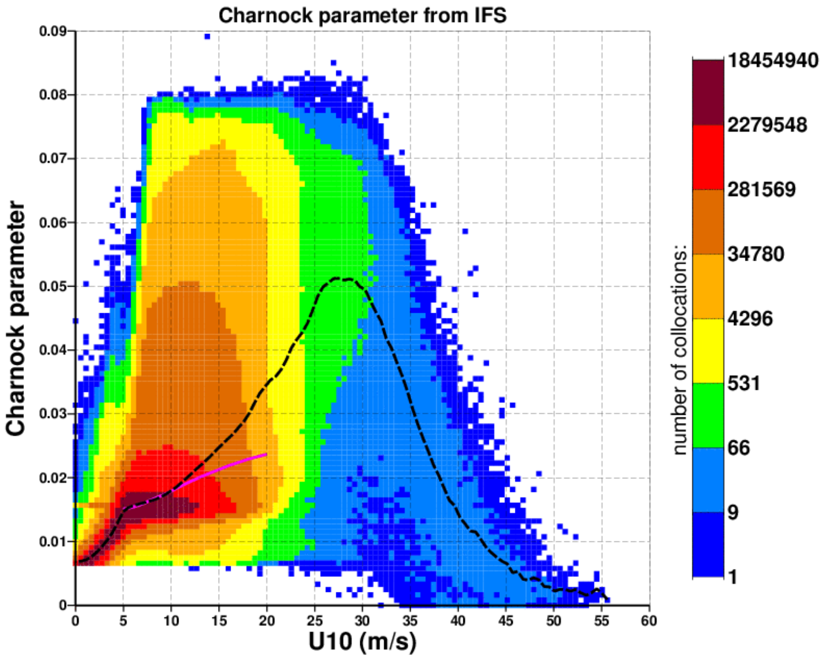
Charnock coefficient update for CY47R1

CY46R1

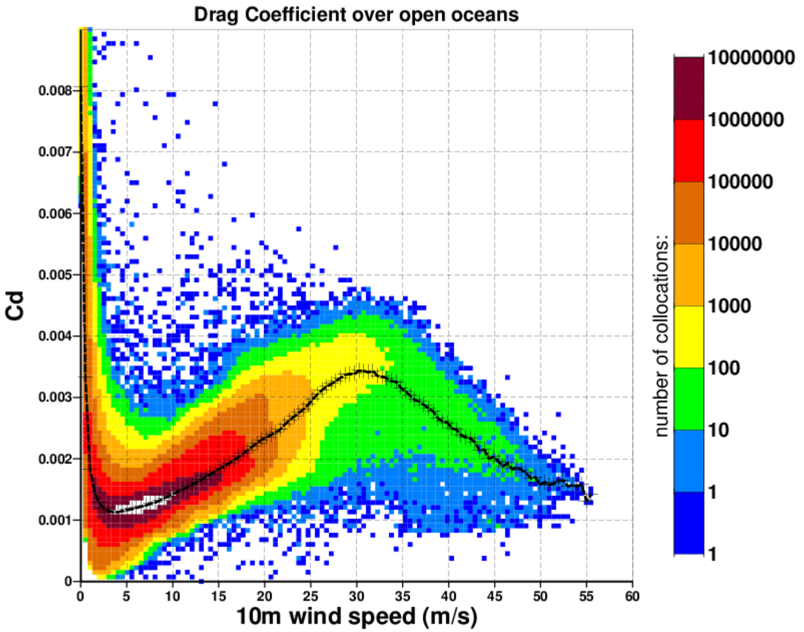
for CY47R1



Forecast data from stream da, class rd, expver h9f7, all Sea points with sea ice cover <= 0.3
from 20190902 12UTC, for steps from 1 to 240 by 1



Forecast data from stream da, class rd, expver h9ha, all Sea points with sea ice cover <= 0.3
from 20190902 12UTC, for steps from 1 to 240 by 1



TCo1279 forecast
h9ha from 20190902 12 UTC, step 1 to 240 by 1

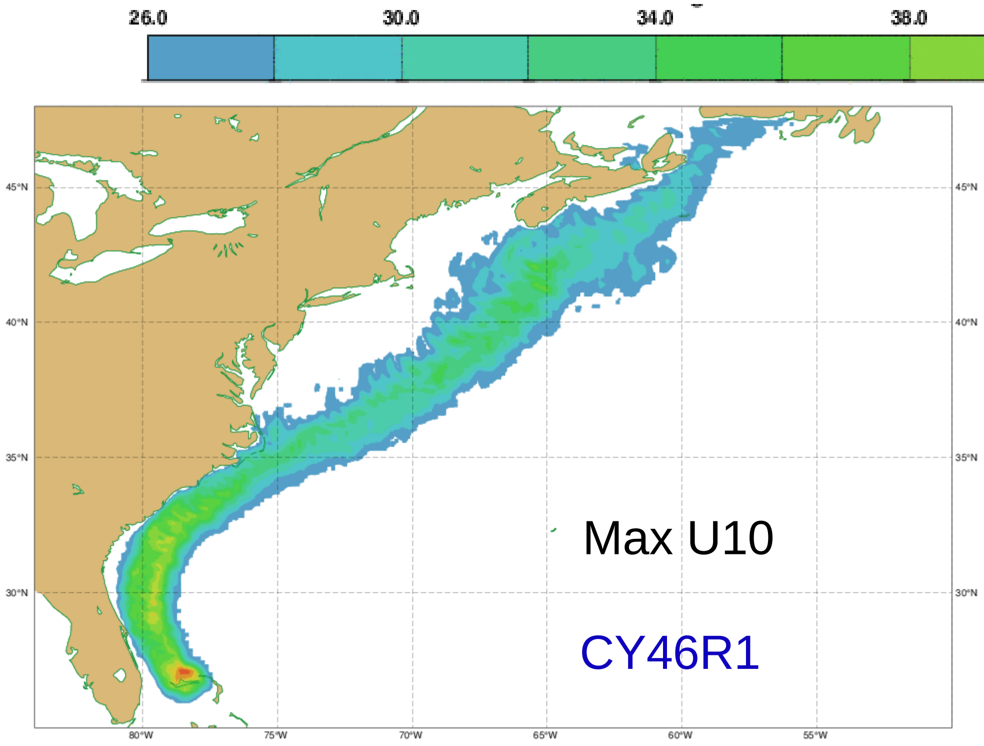
Tco1279 forecast from 2 September 2019, 12 UTC, hourly steps 1 to 240 hours, full coupling to NEMO, starting from DA initial conditions

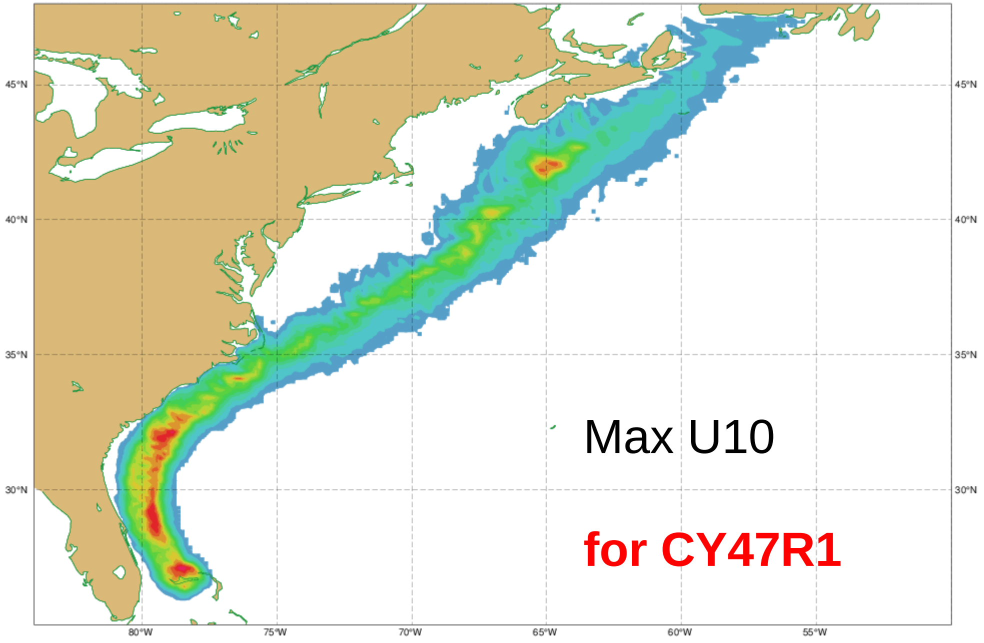
Work of J. Bidlot

Drag coefficient now fits better observations for strong winds over 30m/s compared with CY46R1 which was too big

ECMWF

Impact during Dorian, Tco1279 forecast from 2 September 2019, 12 UTC

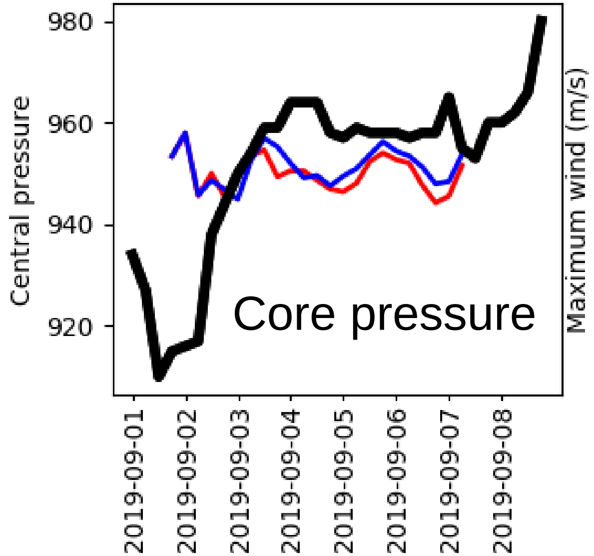


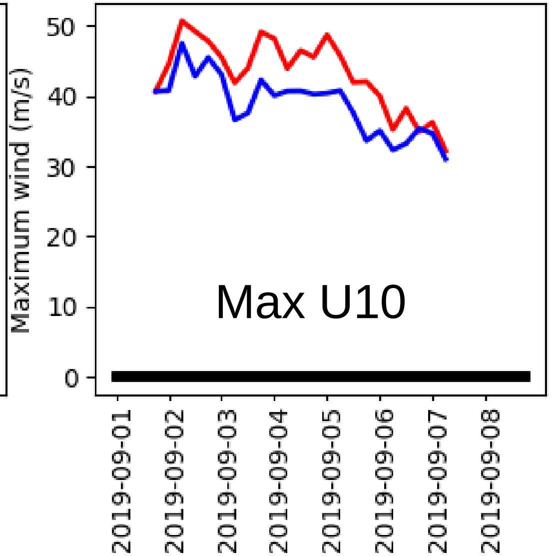


observations

CY46R1

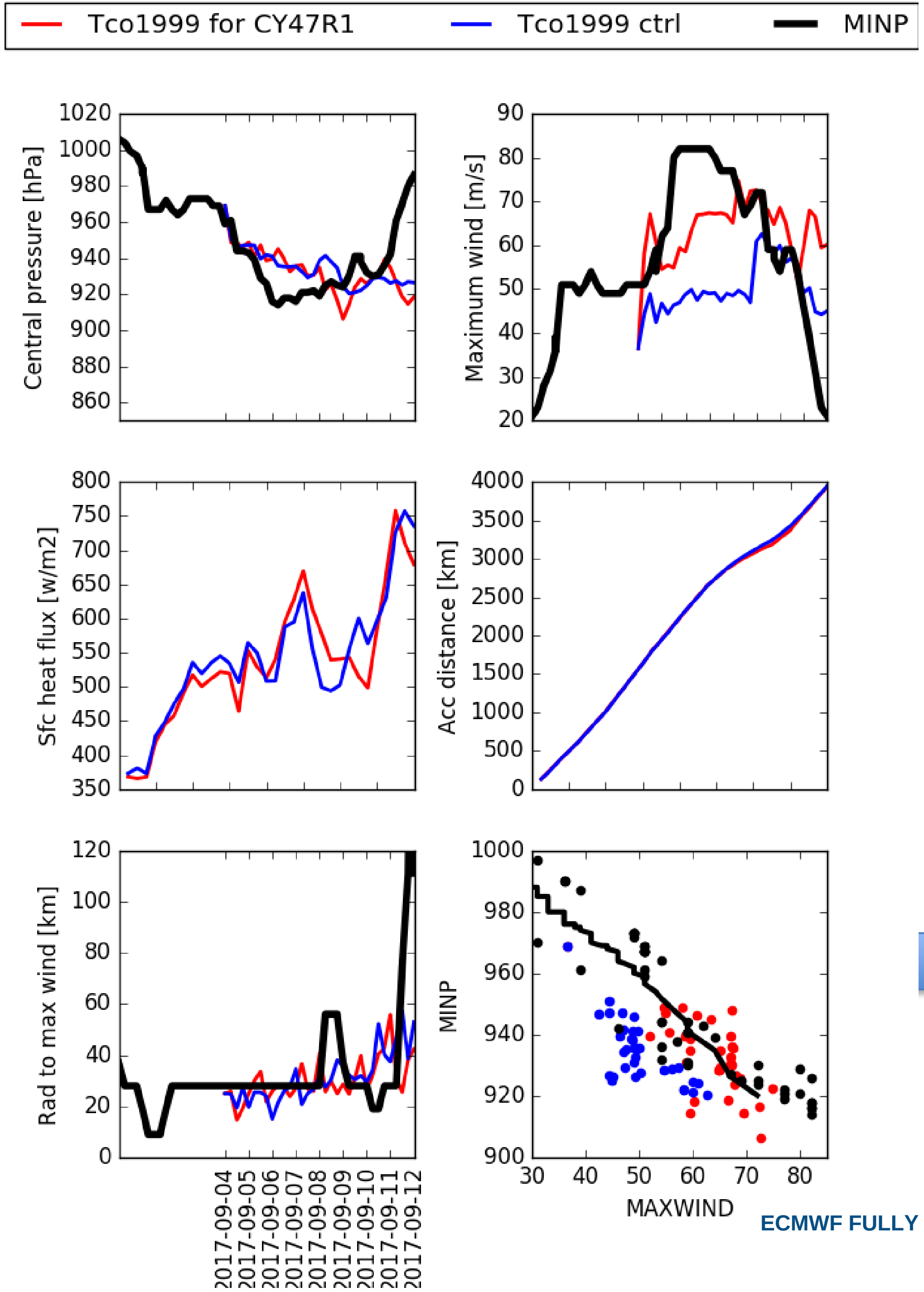
for CY47R1



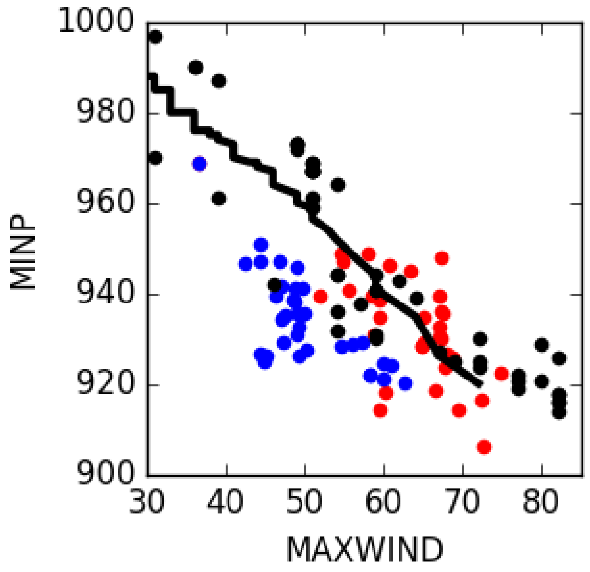


ECMWF

Irma, forecast from 20170904, 0 UTC:



CY46R1 Charnock
limitation on Charnock

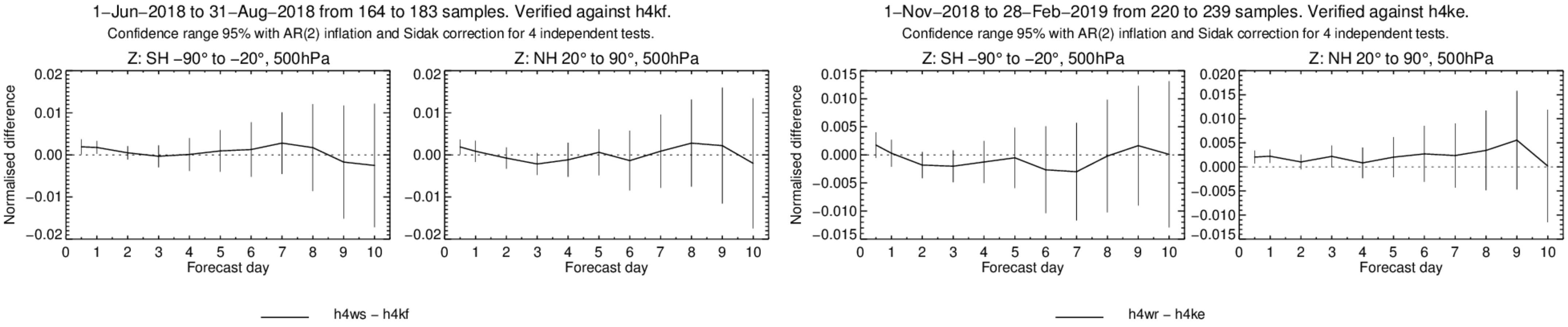


Output every 6 hours

ECMWF FULLY COUPLED SYSTEM

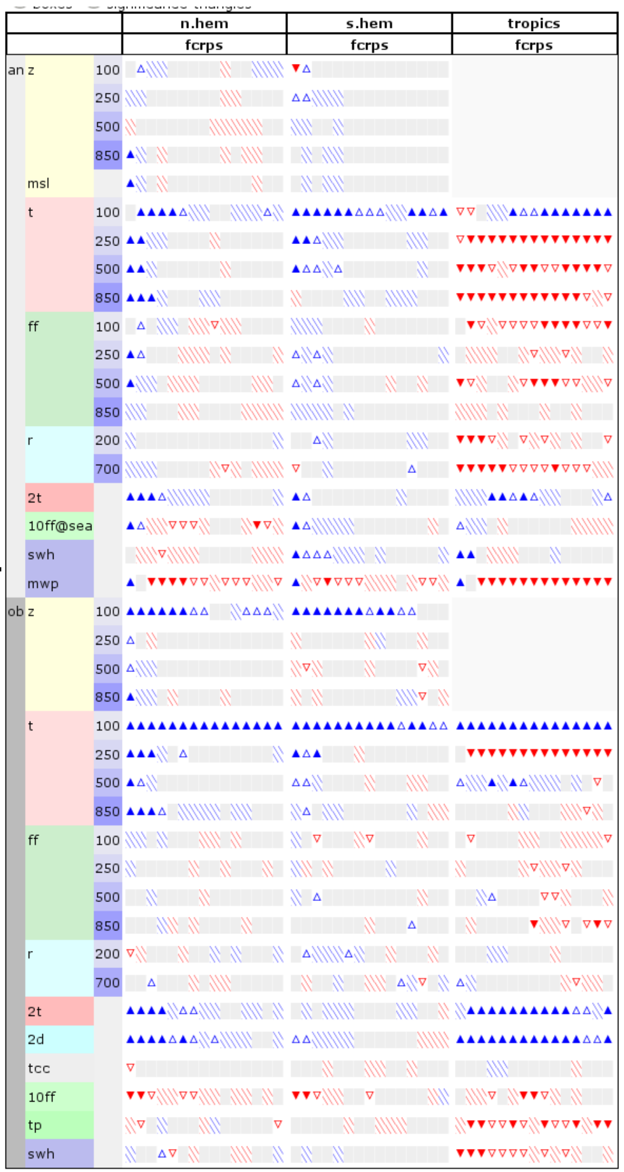
Beyond CY47R1: single precision forecasts for efficiency

- Single precision already used for RD experimentation in CY46R1
- Results are overall neutral compared with double precision:
  - Also tested on extreme weather such as TCs: tracks and intensity unaffected
  - A small degradation near the poles at high resolution more noticeable at winter is currently investigated
- Reduces cost of atmospheric component of forecast by 40%
- We are aiming to have single-precision IFS operational form CY48R1

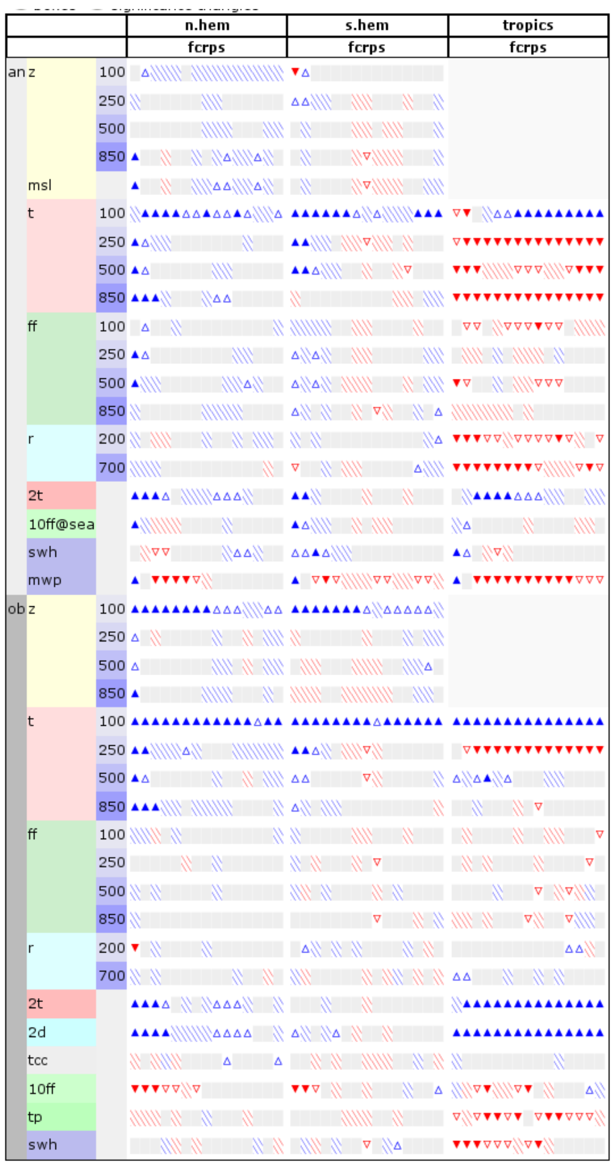


47r1.v6 vs 46r1: TCo399, 8-member ENS

double precision



single precision



- Consistent signal seen with single precision ensembles and double precision ensembles

Each error plot is converted into a sequence of symbols (e.g. ▼▽╲╲ ╲╲△▲) where each symbol indicates for given time step whether the experiment is significantly better or worse than the control on different confidence levels. Different colours are used for symbols comparing control and experiment activity or spread: purple ▼ indicates the experiment is more active or has larger spread than the control and green ▲ the opposite.

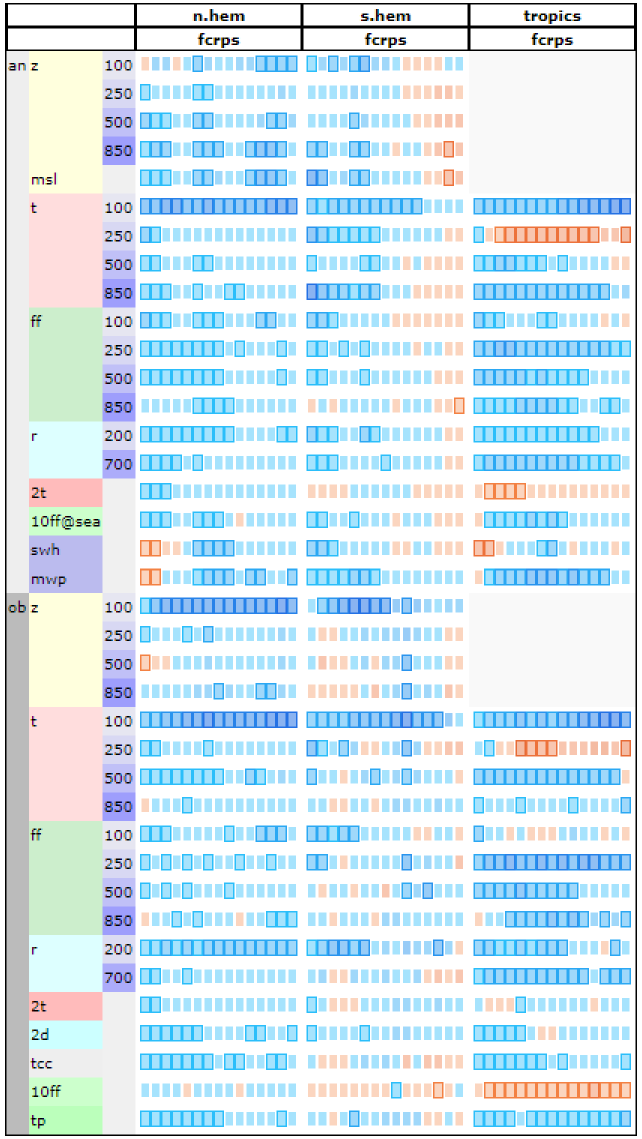
Symbol legend: for a given forecast step…

- ▲ experiment **better** than control statistically **significant with 99.7% confidence**
- △ experiment **better** than control statistically **significant with 95% confidence**
- ╲╲ experiment better than control statistically **significant with 68% confidence**
- not really any difference between control and experiment
- ╲╲ experiment worse than control statistically **significant with 68% confidence**
- ▿ experiment **worse** than control statistically **significant with 95% confidence**
- ▼ experiment **worse** than control statistically **significant with 99.7% confidence**

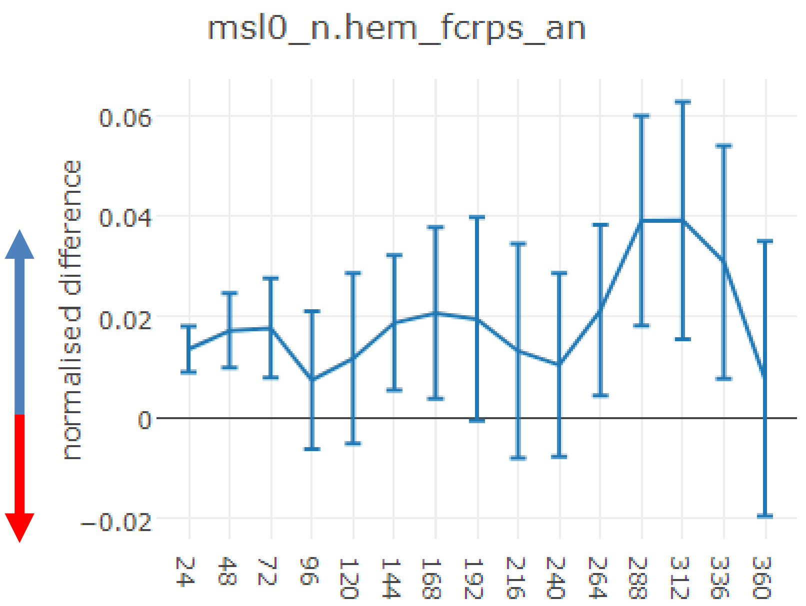
Slide courtesy of M. Leutbecher

Beyond CY47R1: Increasing vertical resolution and its impact

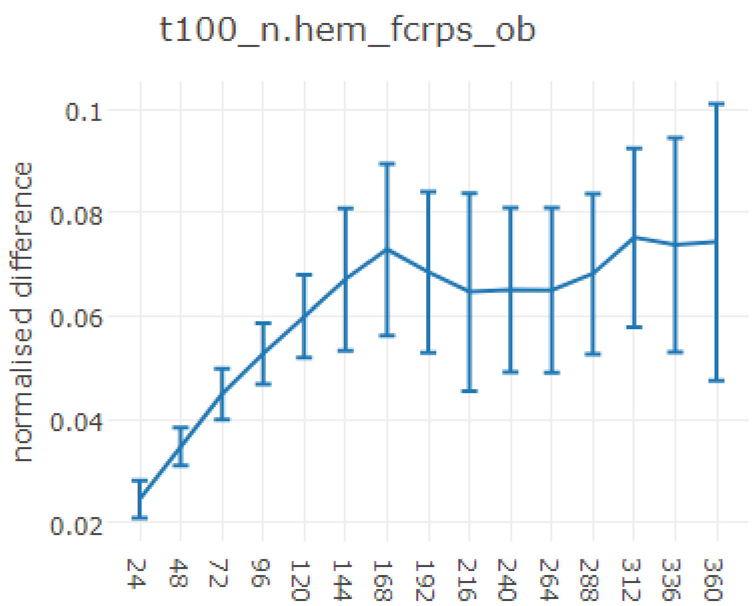
TCo639L137 vs TCo639L91, Summer, 3 Month, initialized every 2 day, 8 member ensemble, CY46r1



mslp vs analysis nhem



t100hPa vs obs nhem

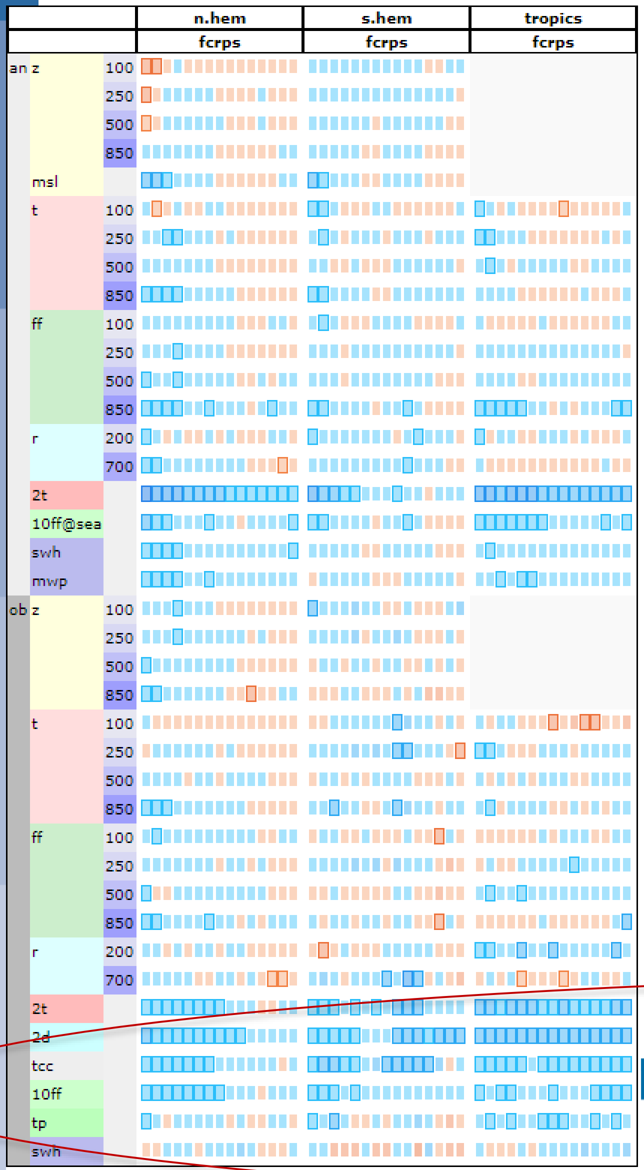


Slide courtesy of S. Lang

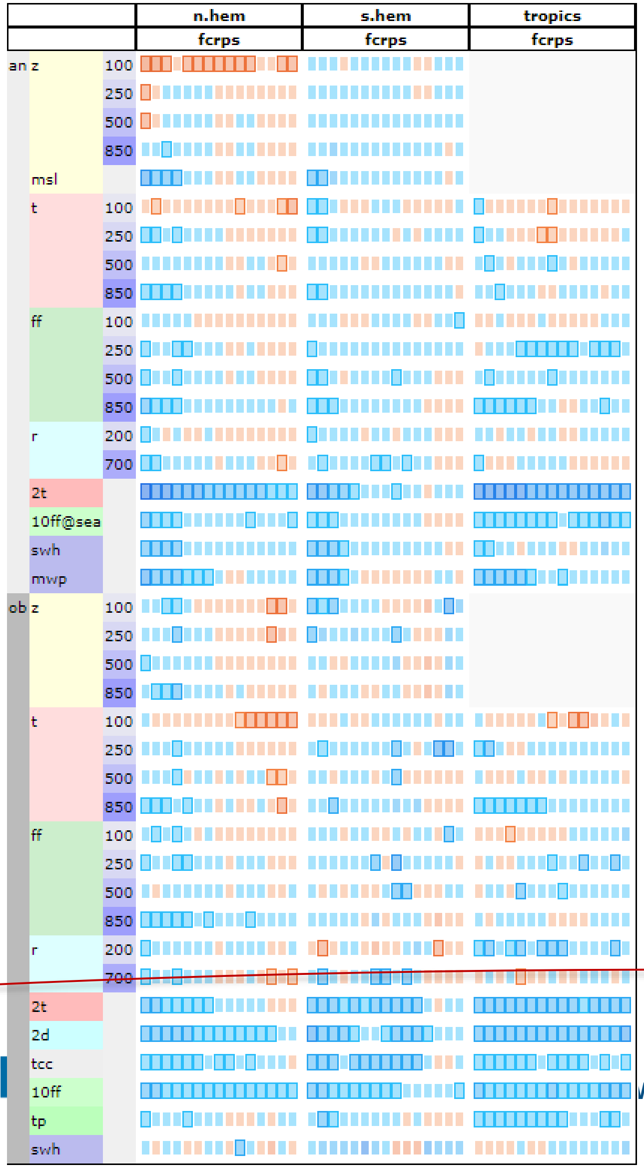
Increasing horizontal resolution and comparing against high vertical res reference

Reference TCo639L137, summer, 3 month, 8 member ensemble, 45r1

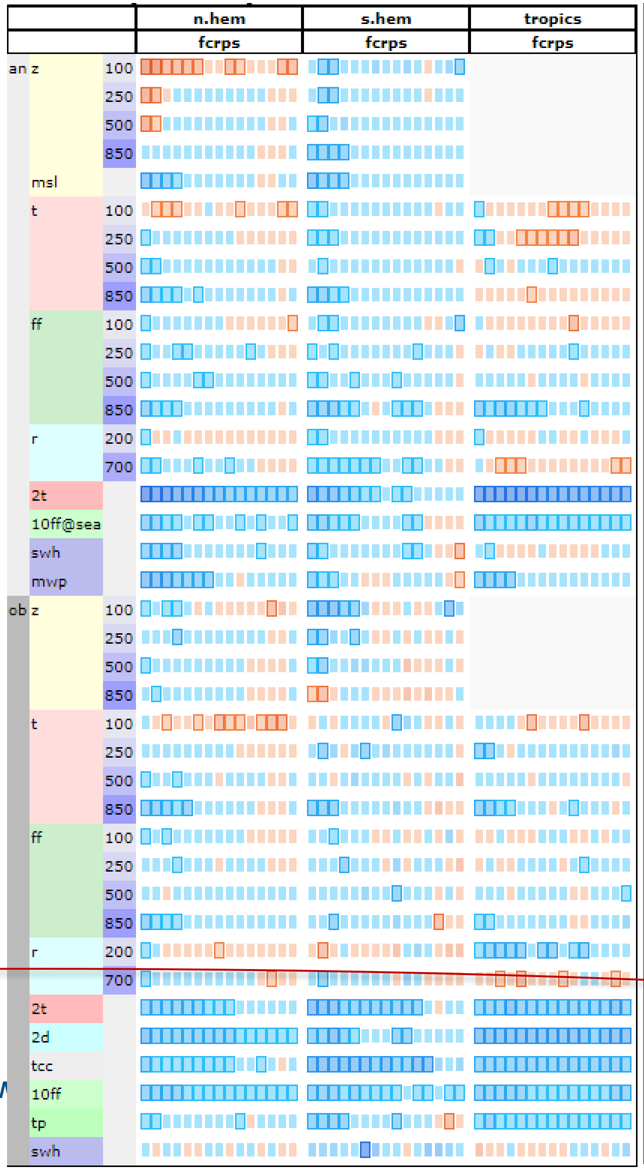
TCo799L137



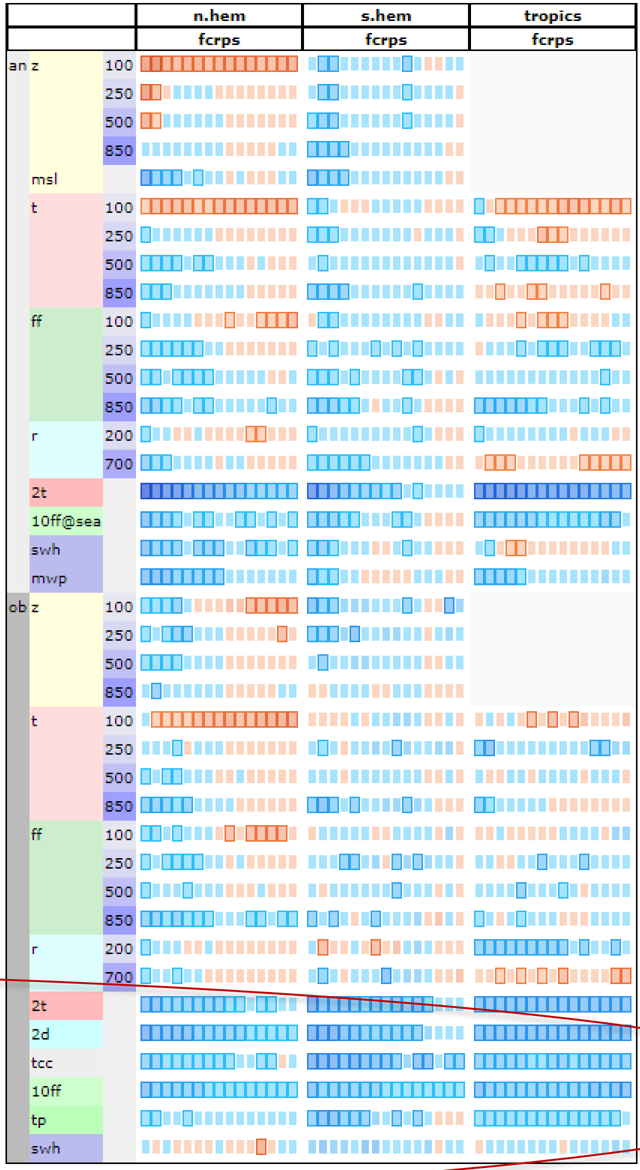
TCo911L137



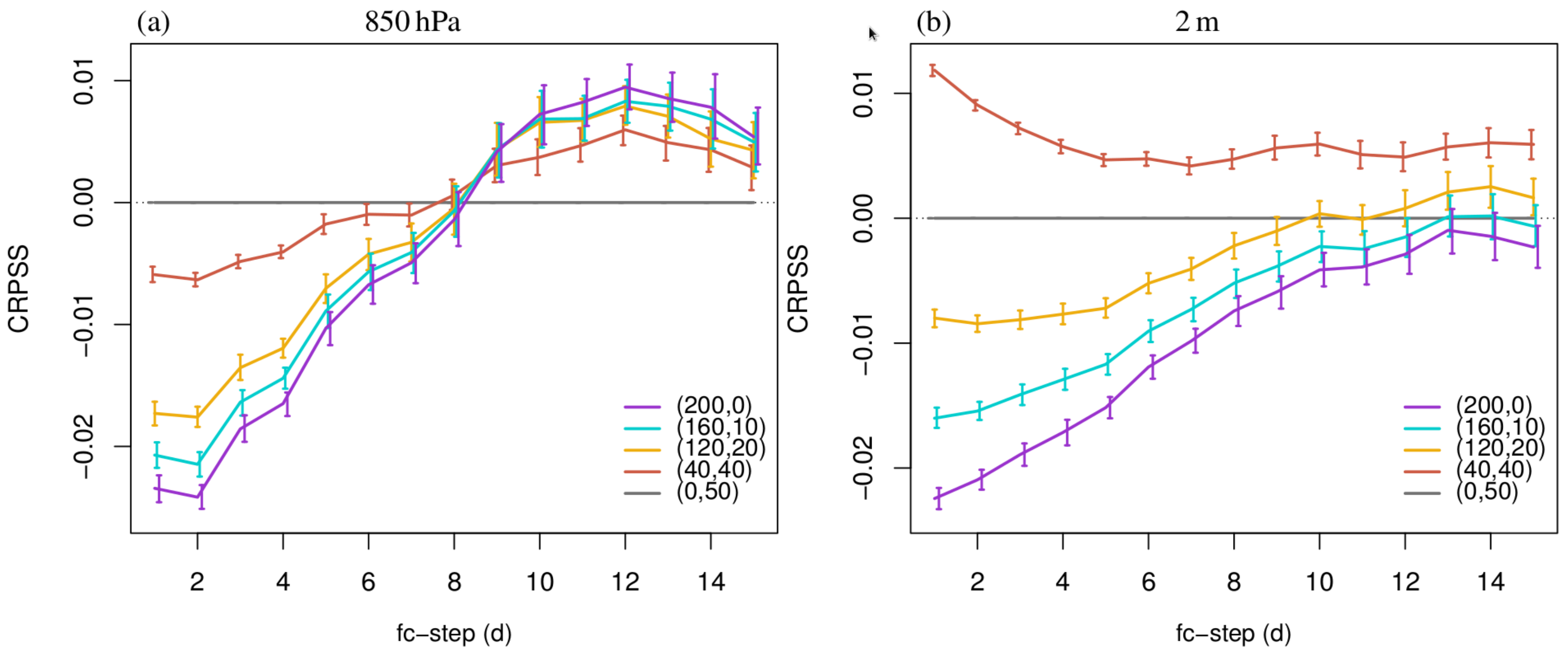
TCo1023L137



TCo1279L137



Extra slide: Dual-resolution ensembles? TCo399 and TCo639 example



- 850 hPa temperature and 2-metre temperature show different behaviour
- Dual-resolution improves 2-metre temperature
- For T850, single resolutions are preferred: higher-res up to Day 8 and lower-res after Day 8

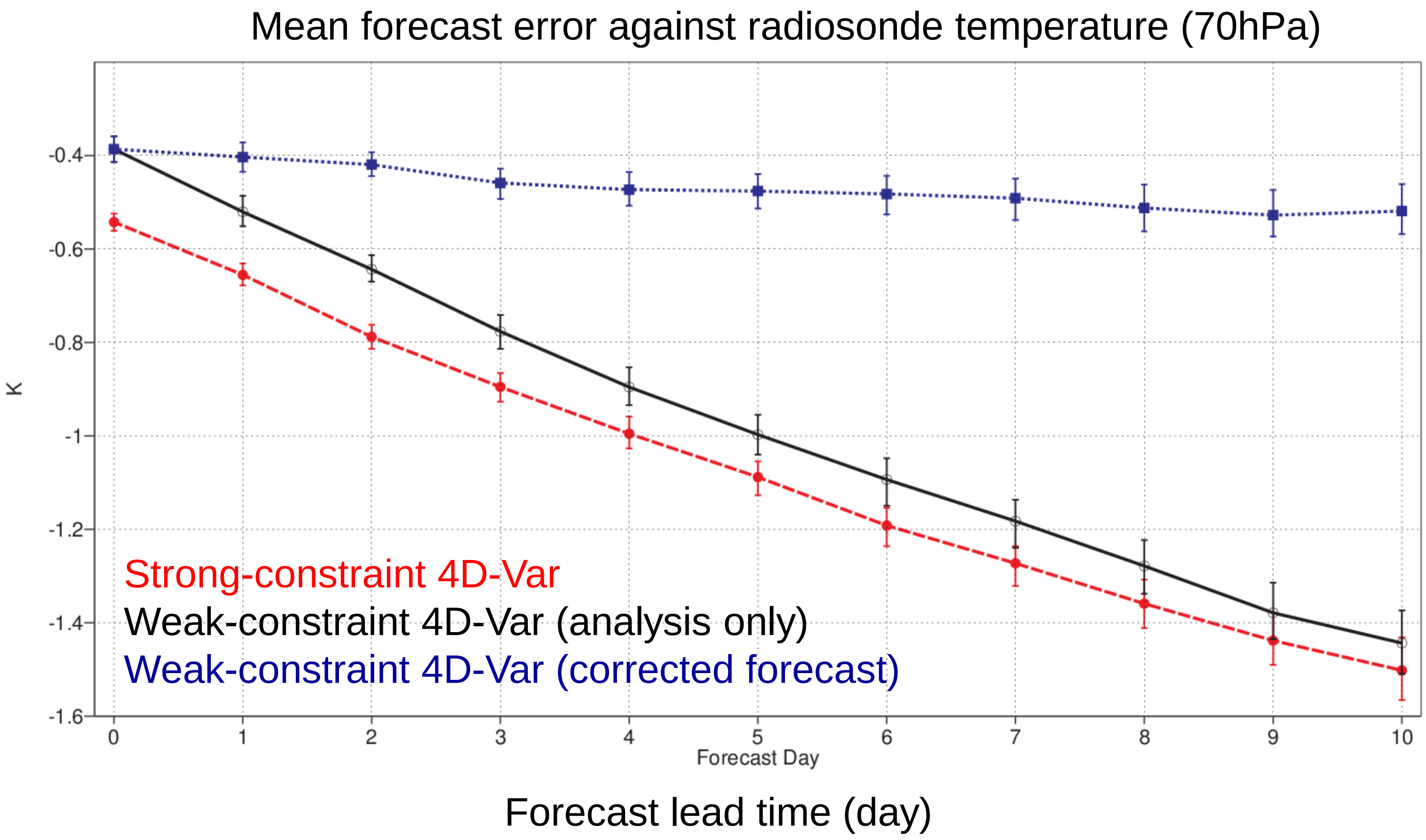
Analytic expression for expected CRPS of ensemble with k lower-res and m higher-resolution members based on extension of Weigel and Bowler (2009) toy model to account for error correlation φ between resolutions:

$$\frac{k}{k+m}A_1+\frac{m}{k+m}A_2-\frac{k(k-1)}{2(k+m)^2}A_3-\frac{m(m-1)}{2(k+m)^2}A_4-\frac{km}{(k+m)^2}\sqrt{A_5'-2\phi\beta_L\beta_H}$$

Large φ close to 1 favours single resolution and lower φ favours dual-resolution. Estimates of φ from verification and analytic expression explain behaviour (Leutbecher and Ben Bouallègue, 2019 in review)

Slide courtesy of M. Leutbecher

Impact of weak-constraint 4D-Var forcing on forecast skills



In the corrected forecast

→ the model error estimation η is applied as a constant model forcing over 10 days

→ the forecast model is not biased anymore and mean error does not increase