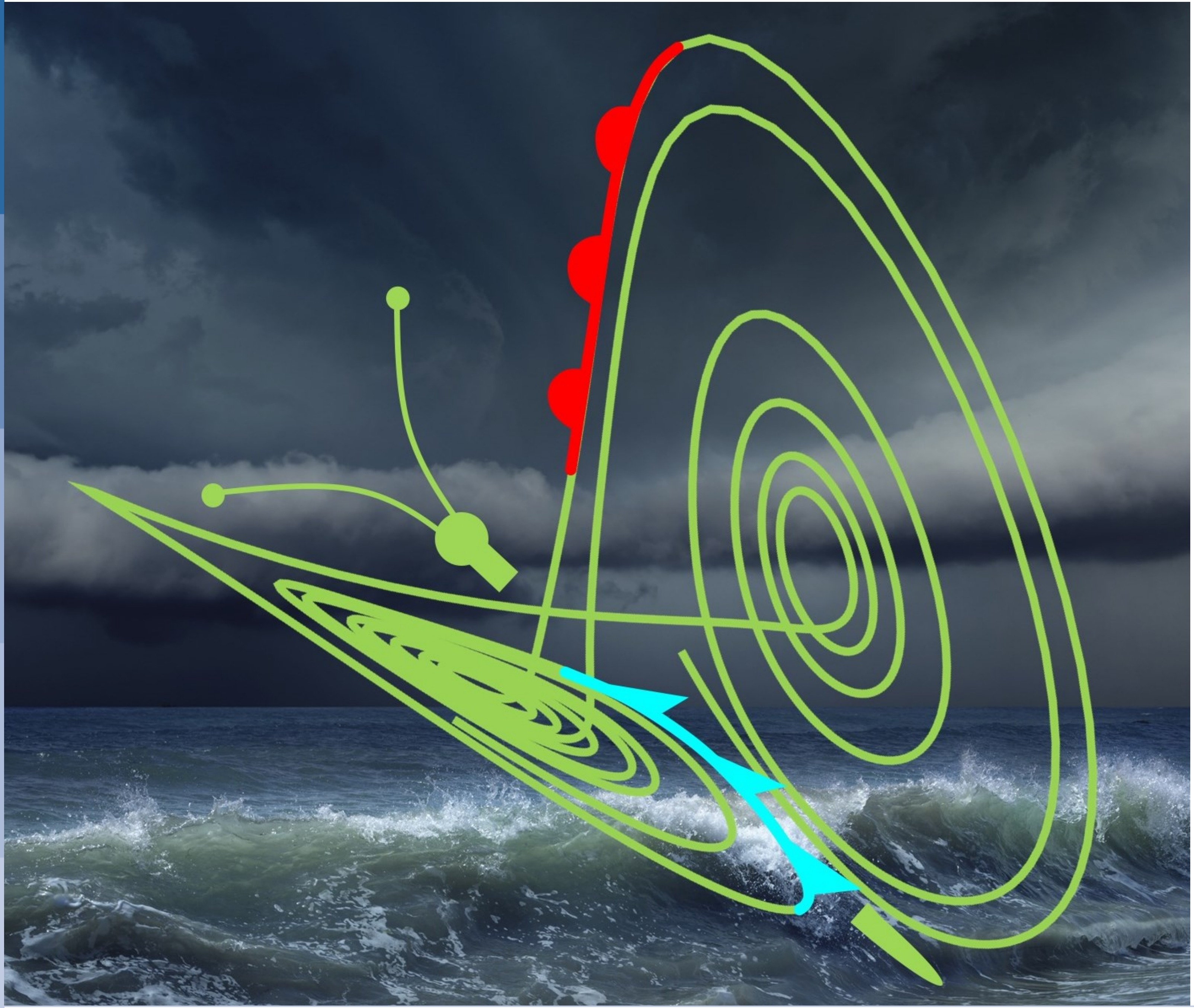


Predictability Diagnostics 2

Mark Rodwell
(with the support of ECMWF
and external collaborators)

ECMWF Training Course on
Predictability

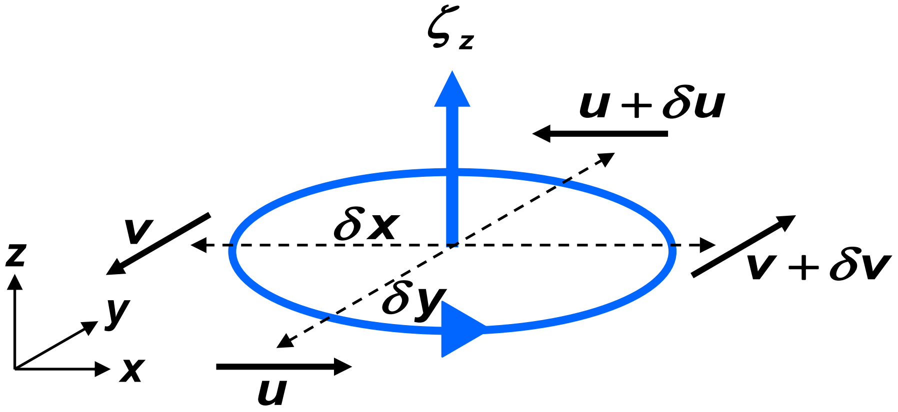
11 May 2017, ECMWF Reading

“All in a spin”

Motivation (2D flow) :

$$\zeta_z = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad \left(\equiv \hat{\underline{k}} \cdot \underline{\nabla}_z \times \underline{v} \right)$$

$\hat{\underline{k}}$ is the unit “vertical” vector and
$\underline{\nabla}_z \times$ is the horizontal curl operator

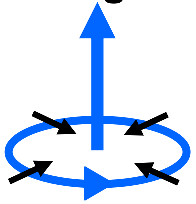


Curl of the 3D momentum equation in absolute frame of reference:

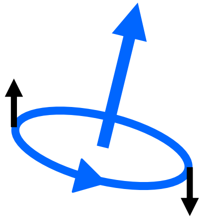
$$\frac{d\underline{\zeta}}{dt} = -\underline{\zeta}(\underline{\nabla} \cdot \underline{u}) + (\underline{\zeta} \cdot \underline{\nabla})\underline{u} + \frac{1}{\rho^2}\underline{\nabla}\rho \times \underline{\nabla}p + \underline{\nabla} \times \underline{F}_u$$

Lagrangian tendency in absolute vorticity

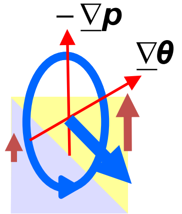
Divergence



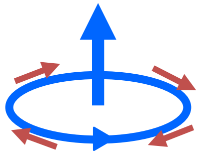
Tilting



Baroclinic



Friction



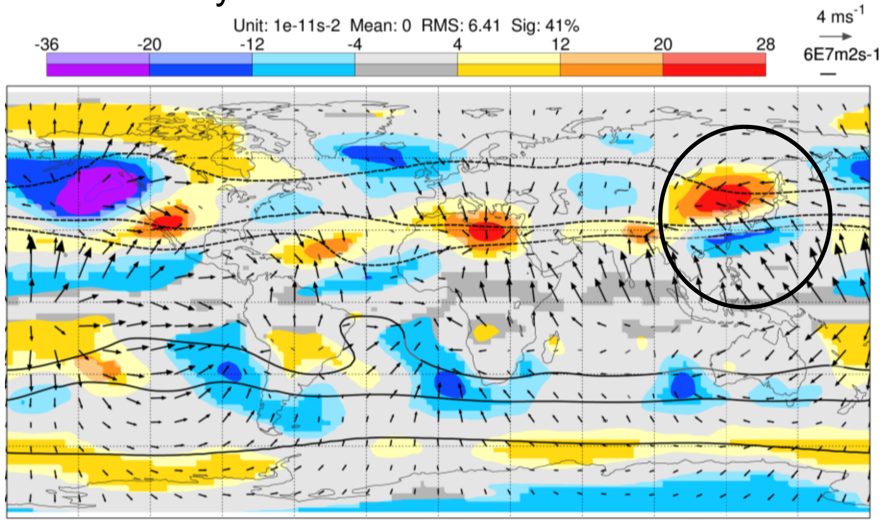
Shallow atmosphere approximation & assuming horizontal, barotropic, frictionless flow:

$$\frac{\partial \zeta}{\partial t} + \underline{v}_\psi \cdot \underline{\nabla}\zeta = -\underline{v}_\chi \cdot \underline{\nabla}\zeta - \zeta \underline{\nabla} \cdot \underline{v}_\chi$$

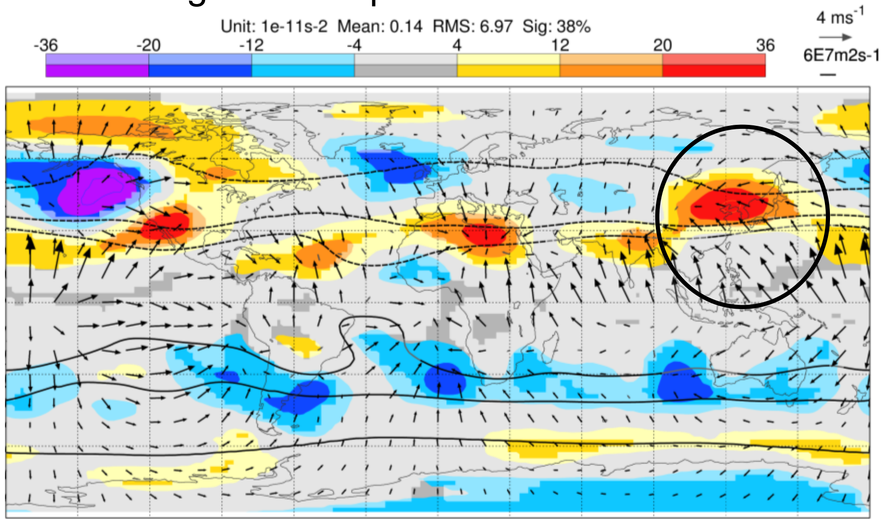
$$= -\underline{\nabla} \cdot (\underline{v}_\chi \zeta) \quad \text{"Rossby Wave Source"}$$

Terms in the Vorticity Equation

Mean Rossby Wave Source

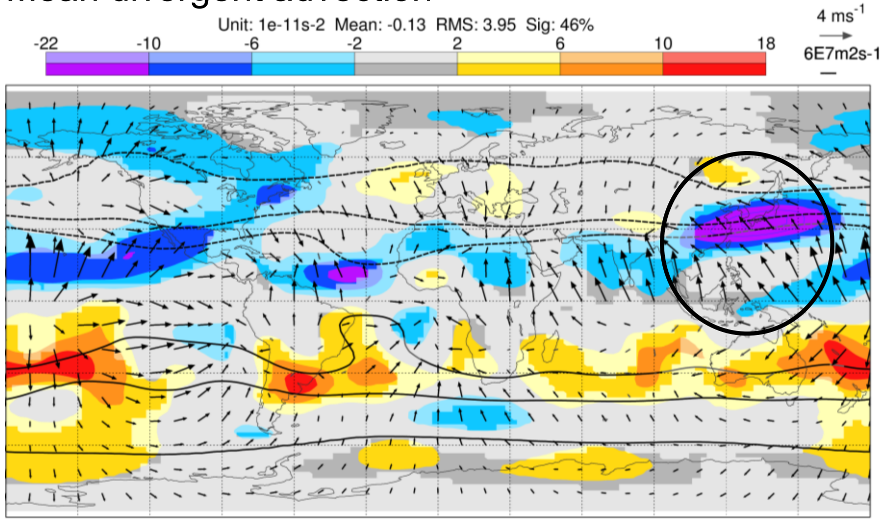


Mean divergence component

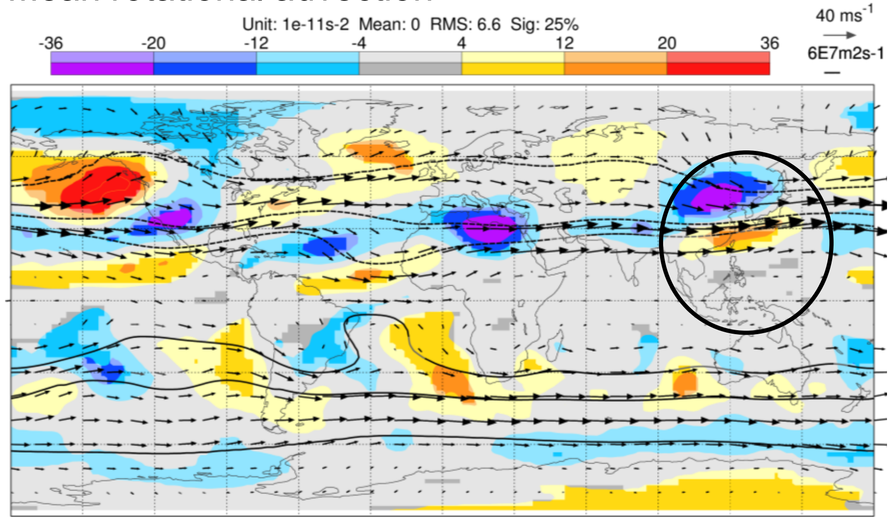


Convergence associated with "stretching"

Mean divergent advection



Mean rotational advection

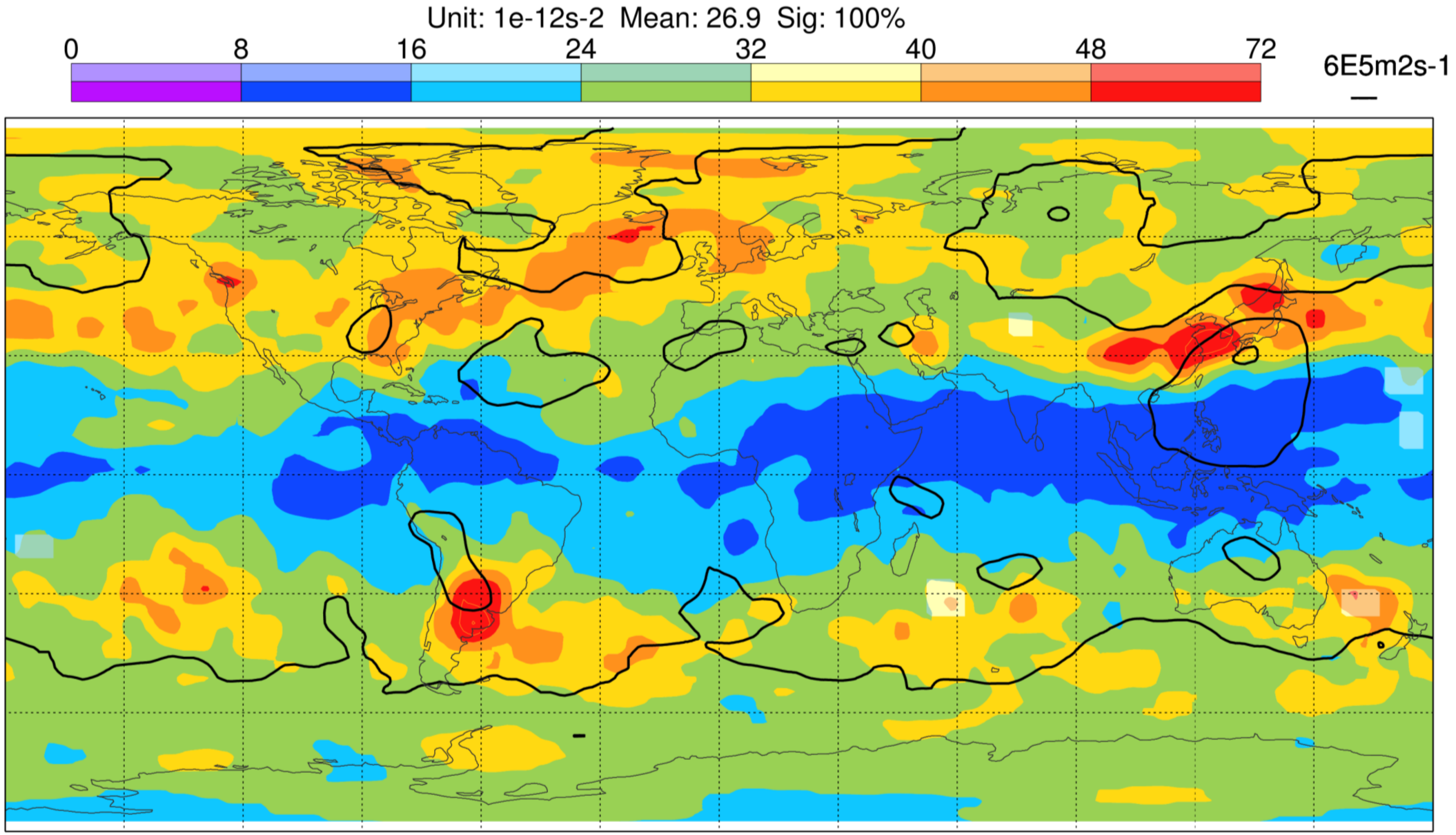


Advection of low planetary vorticity by divergent outflow

Adjustment of rotational flow balances RWS

Based on operational analyses for the period DJF 2015/16, with terms integrated between 100-300 hPa.

Day 1 RMS Error in Rossby Wave Source

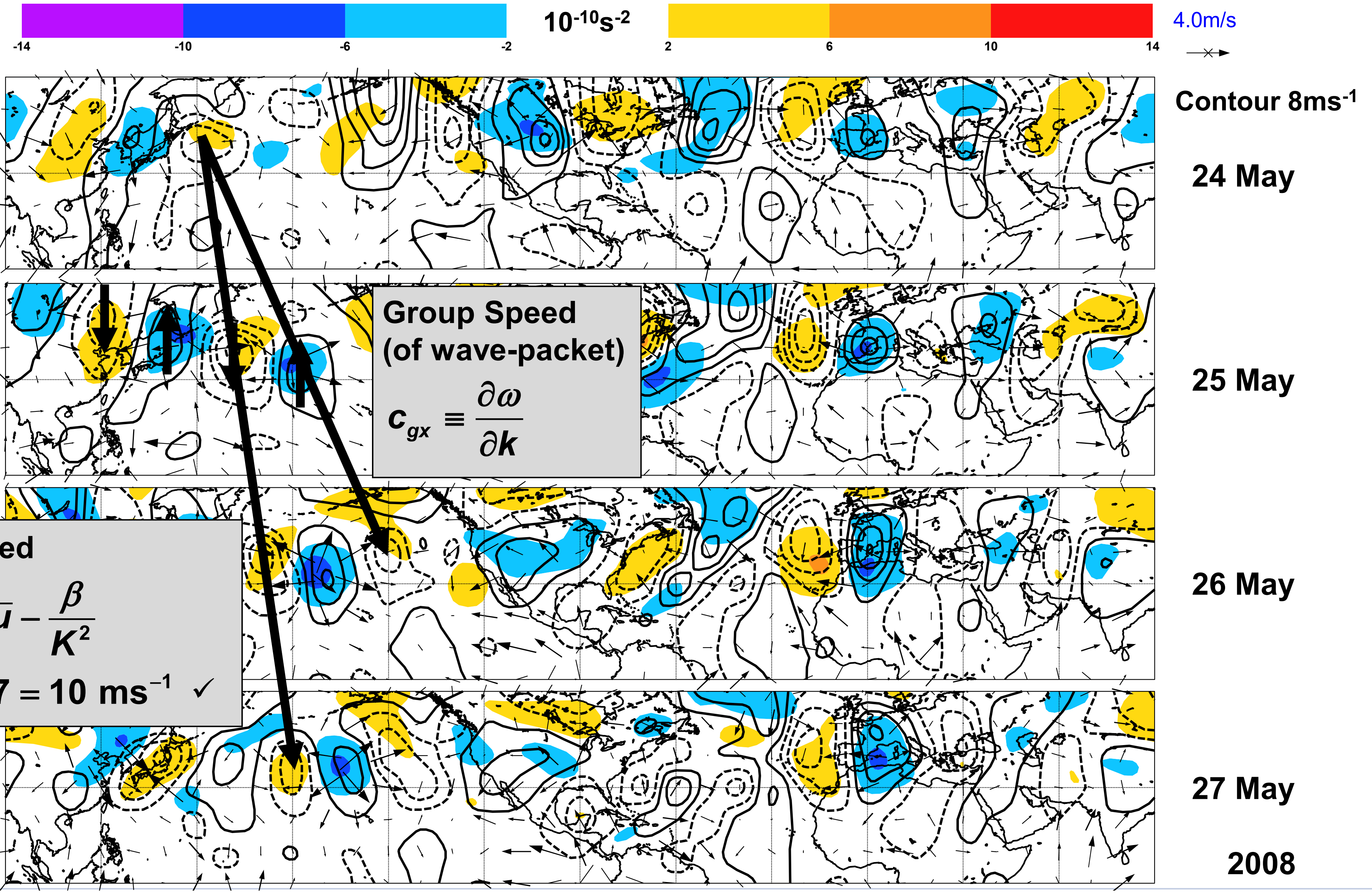


Based on operational analyses and forecasts for the period DJF 2015/16, with terms integrated between 100-300 hPa.

Extra-tropical Rossby waves

100–300 hPa

v_ψ, $\underline{v}_\chi$, RWS



Group Speed (of wave-packet)

$$c_{gx} \equiv \frac{\partial \omega}{\partial k}$$

Phase Speed

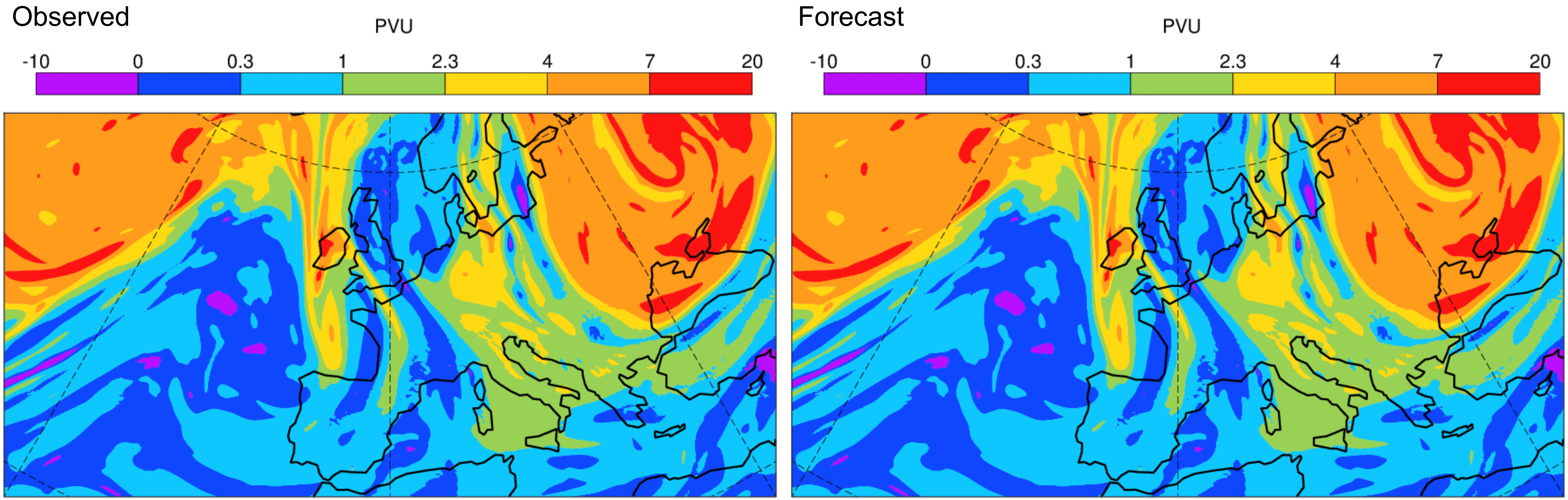
$$c_x \equiv \frac{\omega}{k} = \bar{u} - \frac{\beta}{K^2}$$

$$\approx 17 - 7 = 10 \text{ ms}^{-1} \checkmark$$

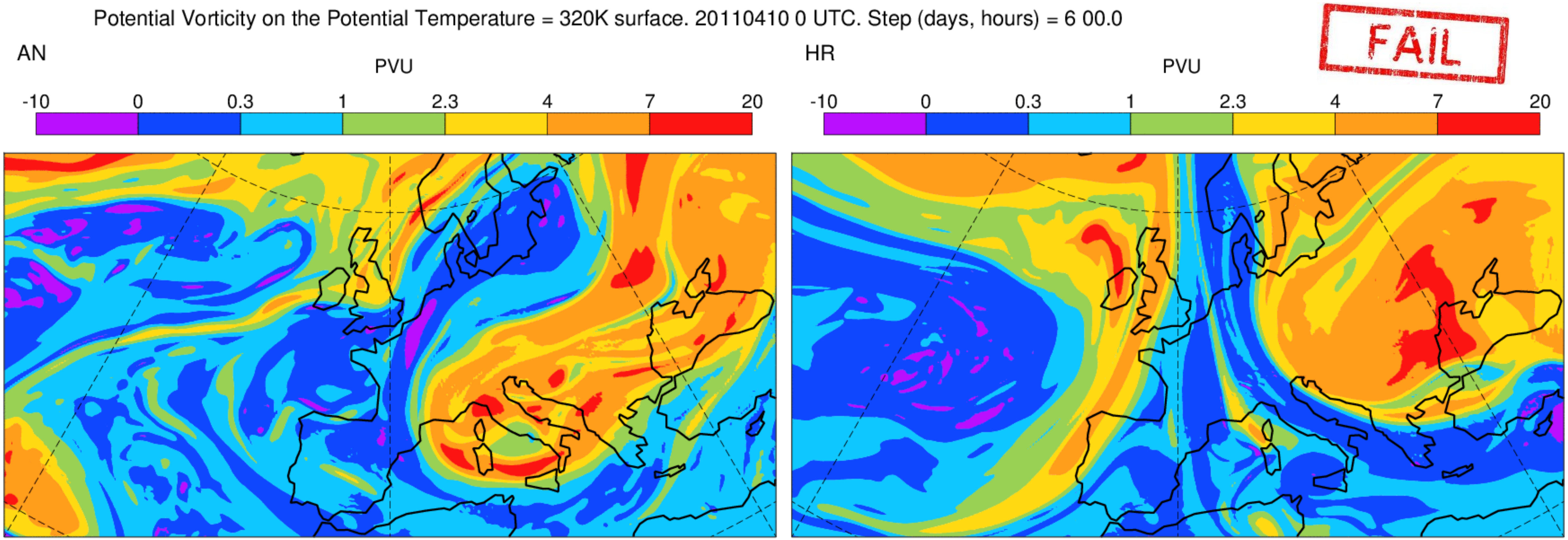
Flow dependent predictability and reliability

Animation of a very poor medium-range single forecast

Potential Vorticity on the Potential Temperature = 320K surface. 20110410 0 UTC. Step (days, hours) = 0 00.0



Animation of a very poor medium-range single forecast

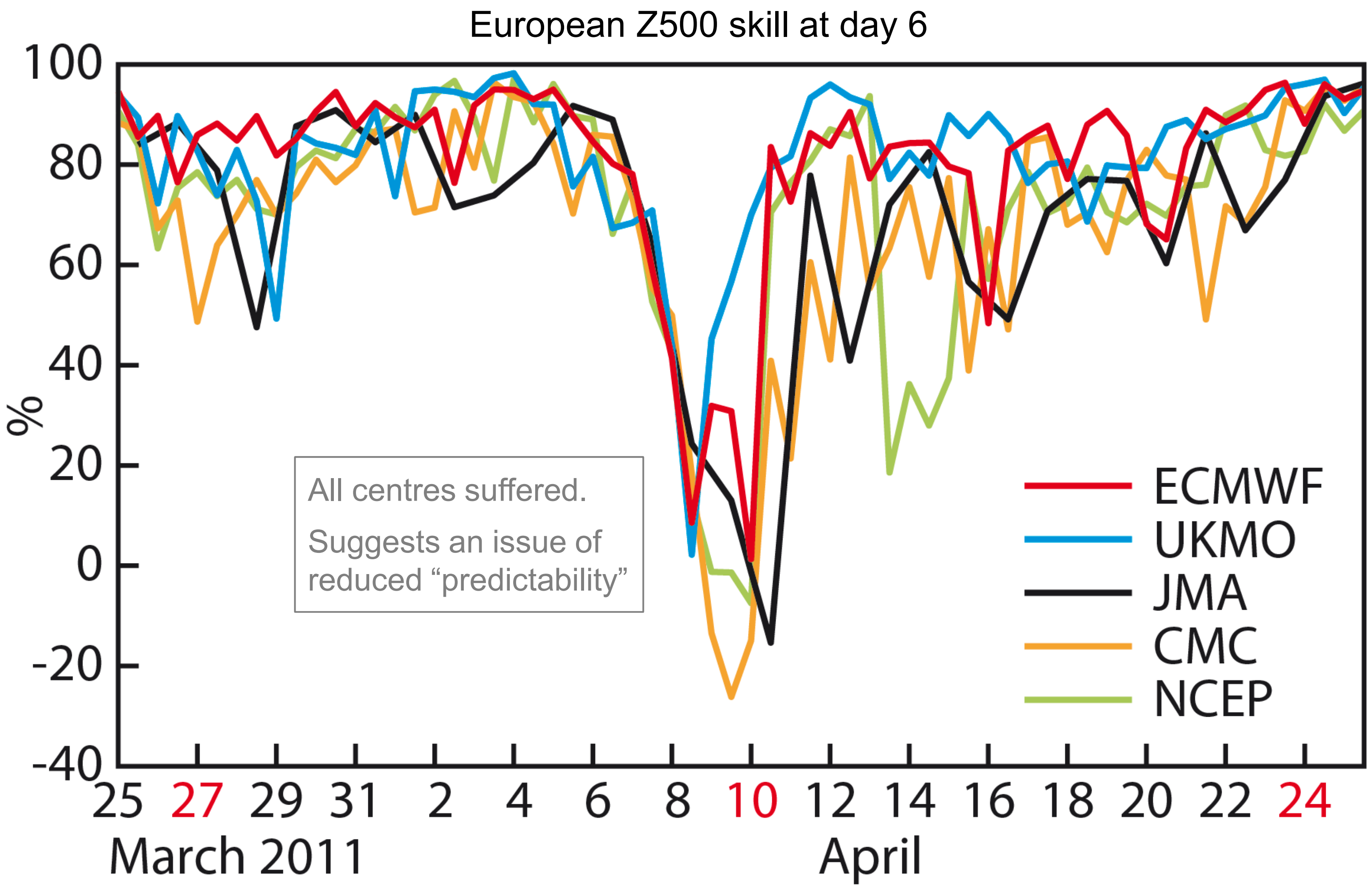


We see the mixing of air masses. The eventual block (high pressure) over Northern Europe is not well predicted

With a single forecast, it is easy to quantify the error (pointwise differences, pattern correlations etc.)

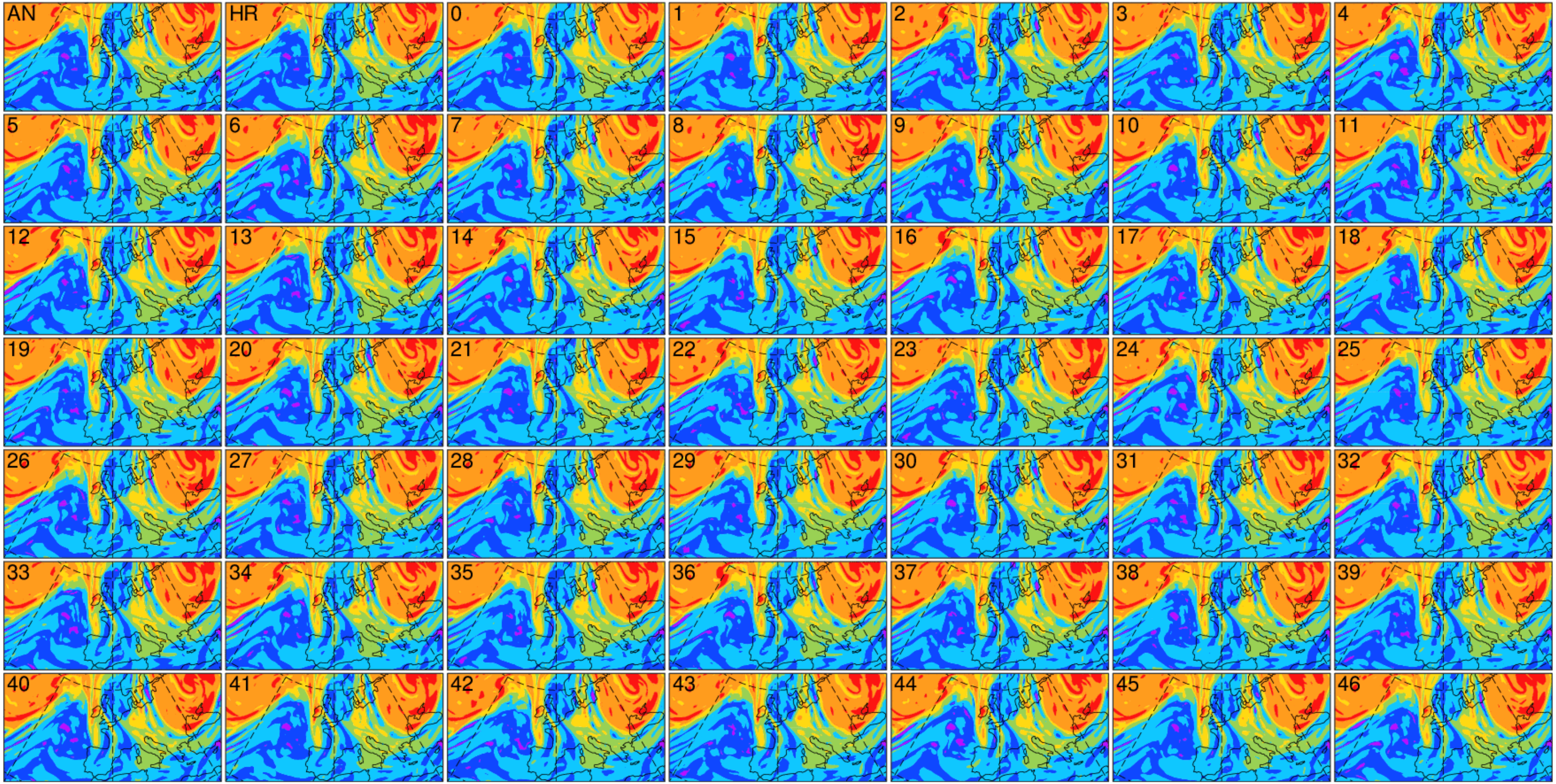
All forecast centres suffered

Rodwell et al, 2013, BAMS

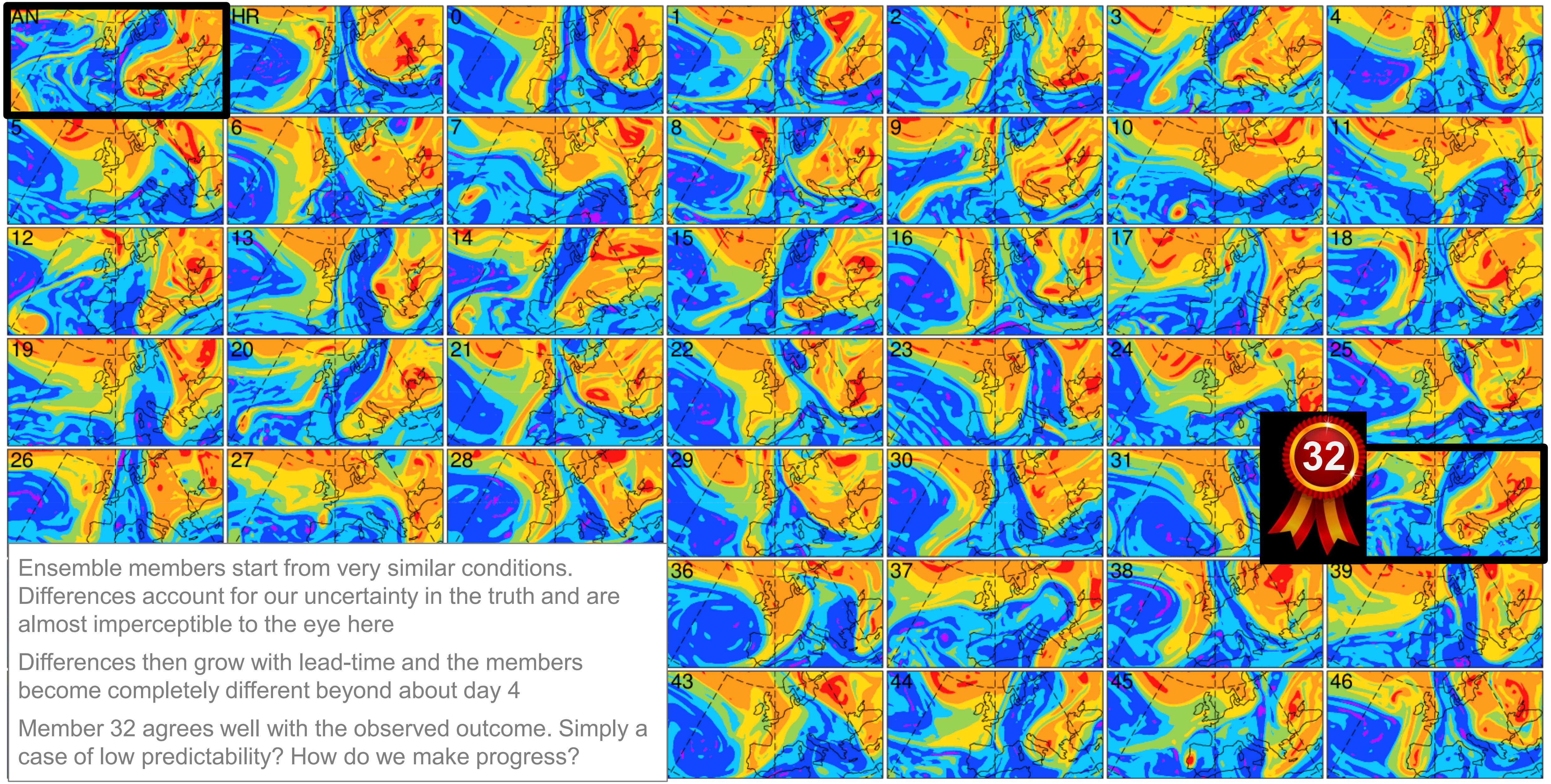


Spatial Anomaly Correlation Coefficient for 500 hPa geopotential height in [12.5°W –42.5°E, 35°N–75°N]. Date is forecast start

Potential Vorticity on the Potential Temperature = 320K surface. 20110410 0 UTC. Step (days, hours) = 0 00.0



Potential Vorticity on the Potential Temperature = 320K surface. 20110410 0 UTC. Step (days, hours) = 6 00.0

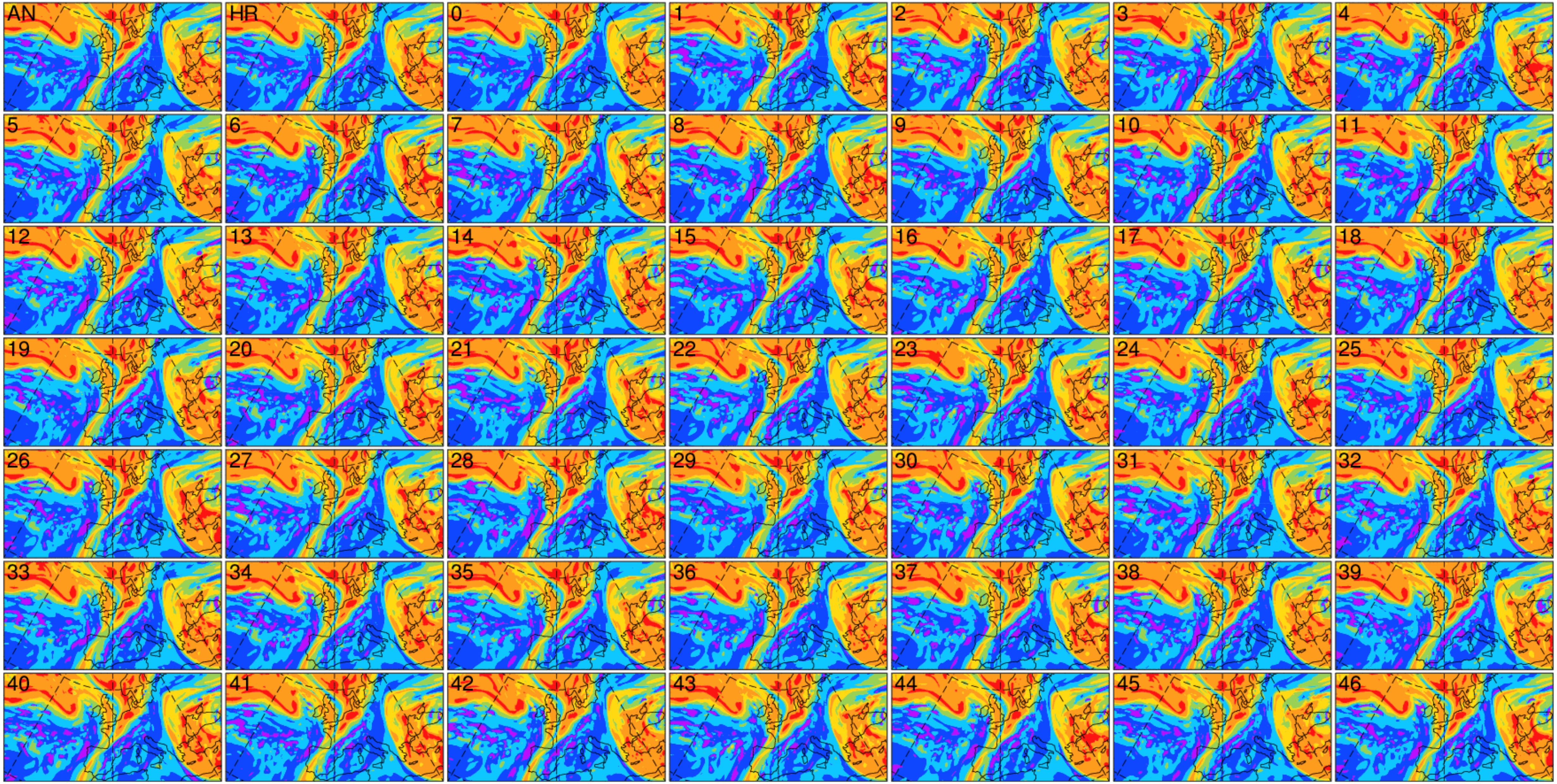


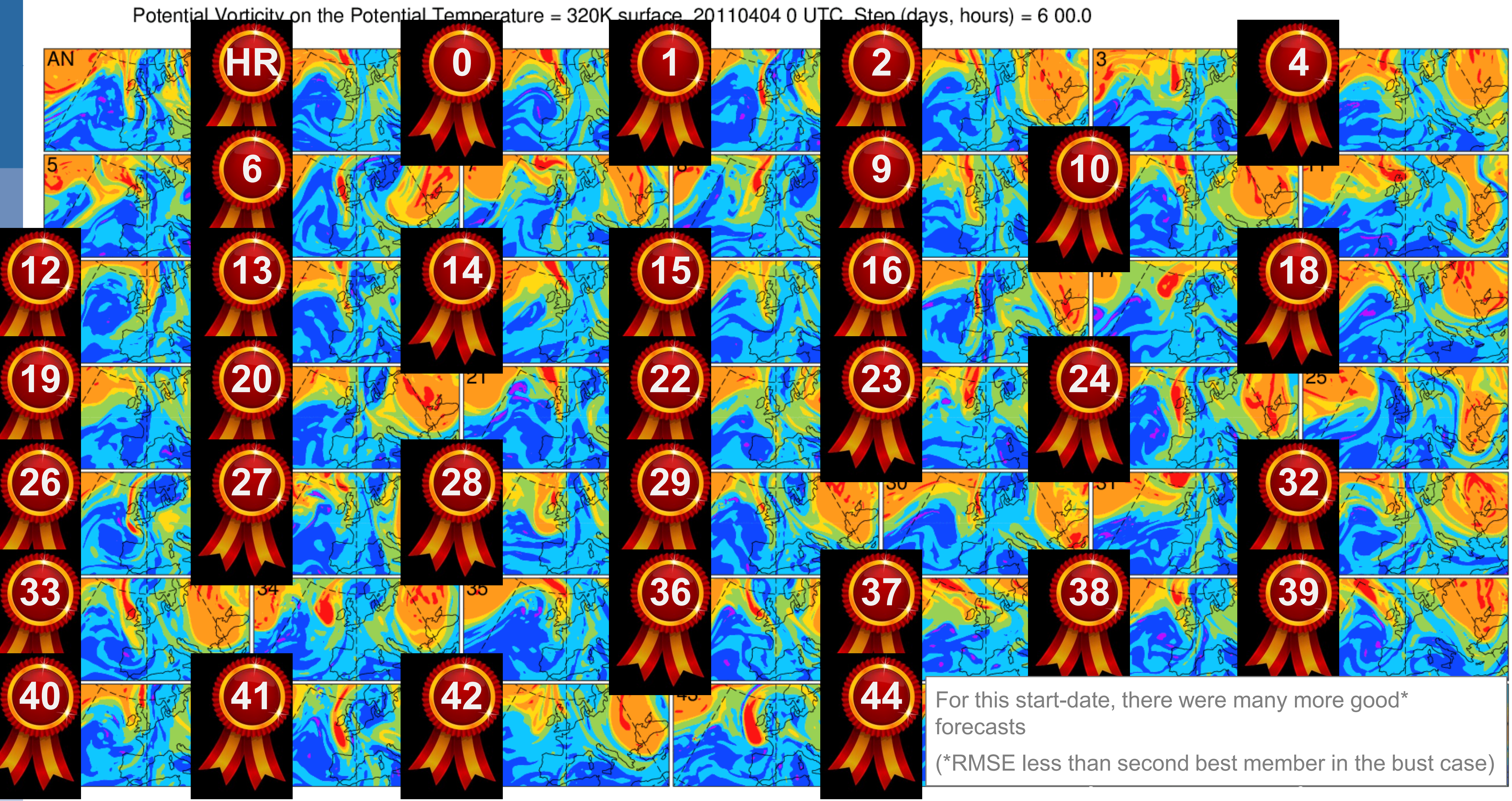
Ensemble members start from very similar conditions. Differences account for our uncertainty in the truth and are almost imperceptible to the eye here

Differences then grow with lead-time and the members become completely different beyond about day 4

Member 32 agrees well with the observed outcome. Simply a case of low predictability? How do we make progress?

Potential Vorticity on the Potential Temperature = 320K surface. 20110404 0 UTC. Step (days, hours) = 0 00.0

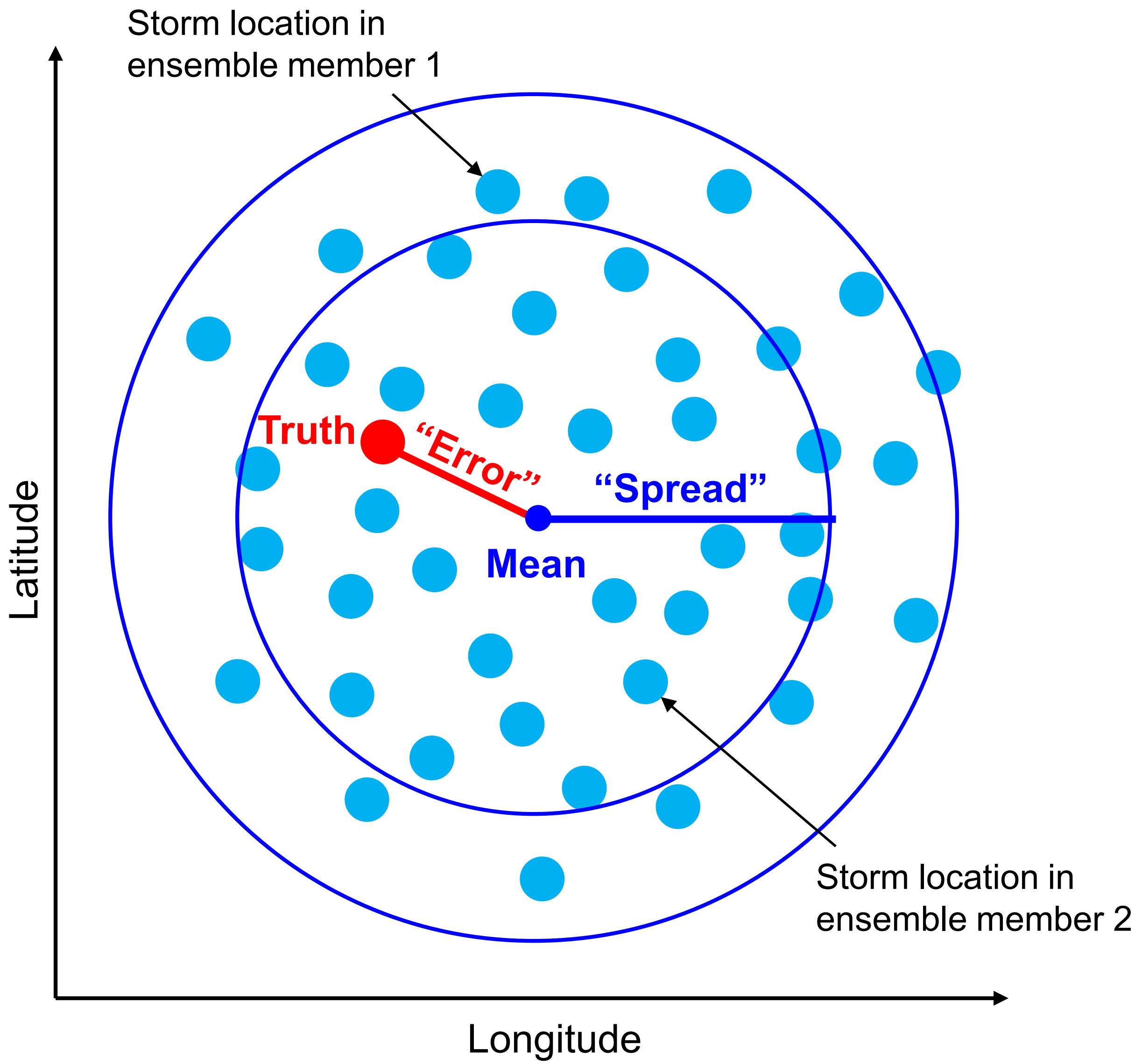




For this start-date, there were many more good* forecasts

(*RMSE less than second best member in the bust case)

Reliability and Sharpness – Example based on forecast of storm location



In a "**reliable**" forecast system, the truth can be considered as another ensemble member

Reliability is very useful: if we predict an event with probability 70%, it will happen with frequency 70%

A testable consequence of reliability is that:

average Error = average Spread

(averaged over many forecast start dates)

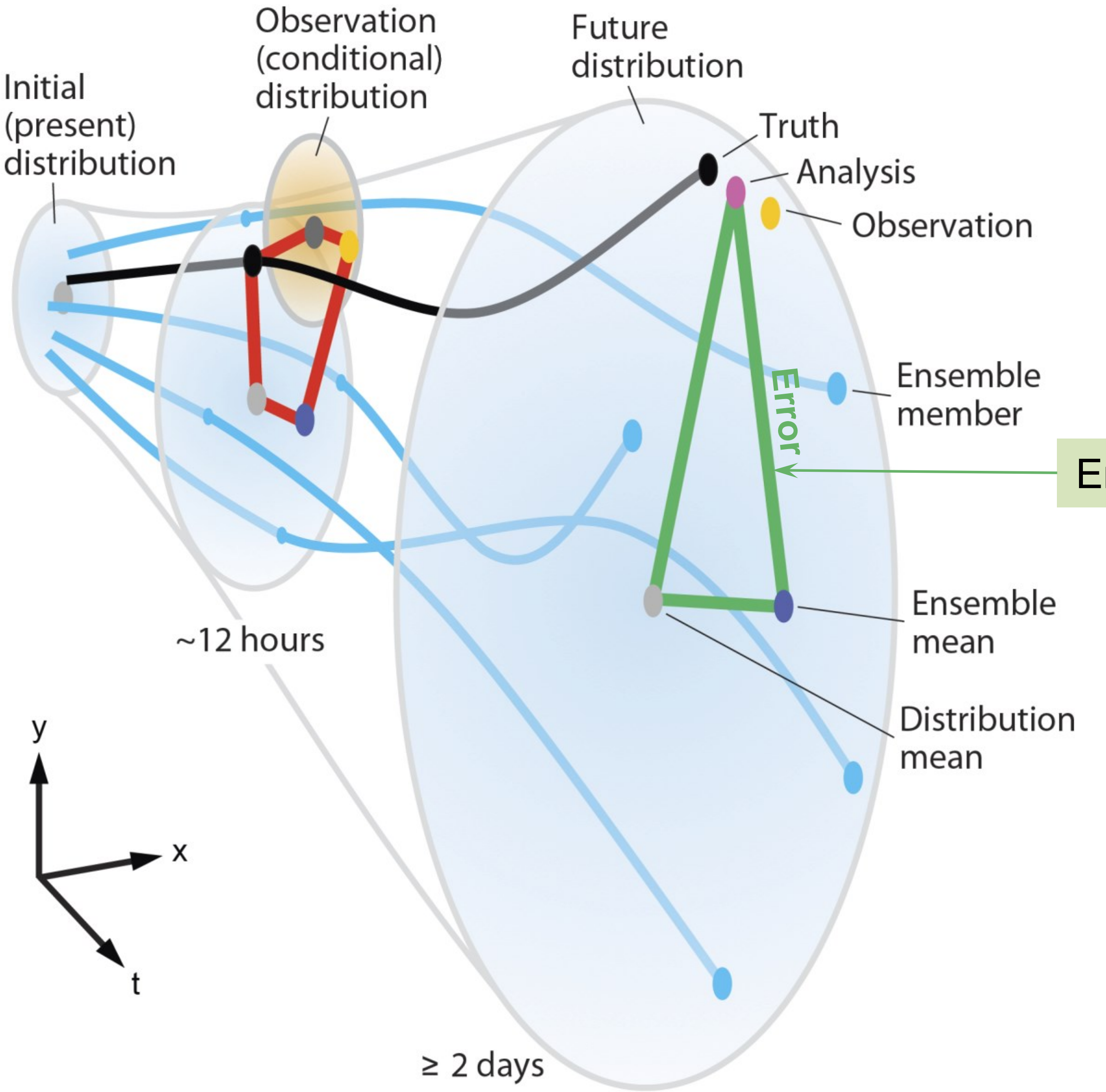
Given we had a reliable system, progress would be …

Predicting "**sharper**" (tighter) distributions **while retaining reliability**

(A more predictable day should also have a sharper distribution)

Reliability in ensemble forecasting

Adapted from Rodwell et al. (2015) QJRMS



Error^2 = EnsVar + Residual

Error ≈ Spread

(when averaged over enough start dates - a consequence of statistical "reliability")

Progress: While maintaining this relationship, we wish to reduce both error and spread (i.e. "shrink diagram in x,y")

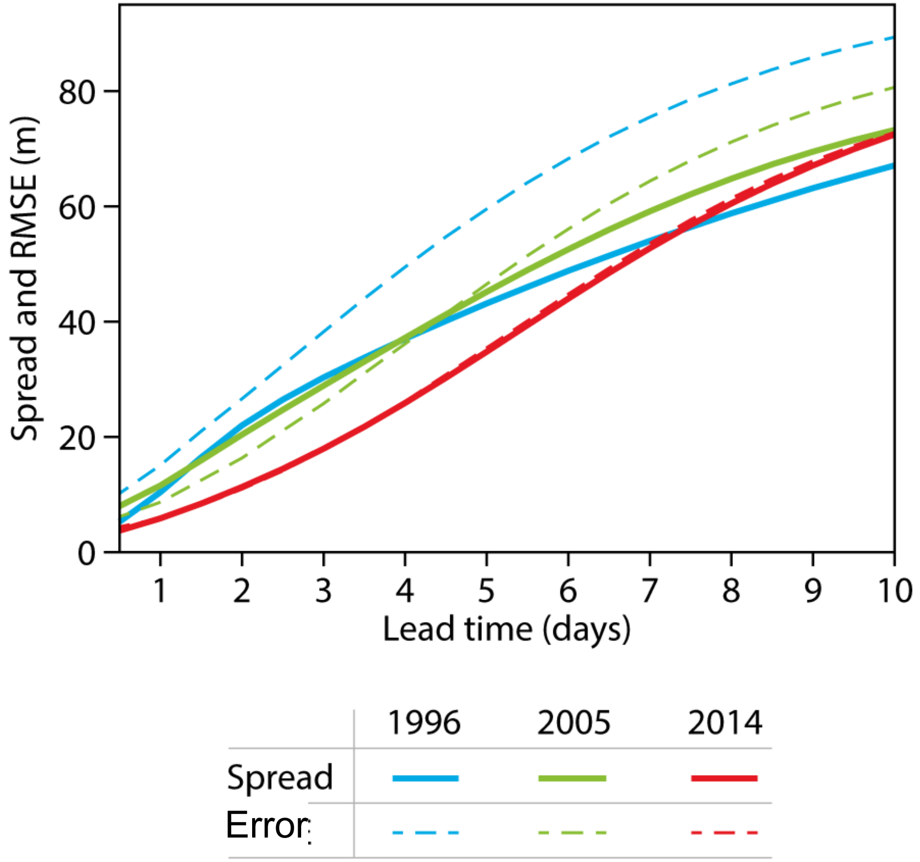
Proper scores (Brier, CRPS, Ignorance etc.) help ensure that development decisions lead us in this direction

(Cross-terms on squaring have zero expectation. EnsVar is scaled variance to account for finite ensemble-size)

Ensemble spread and error

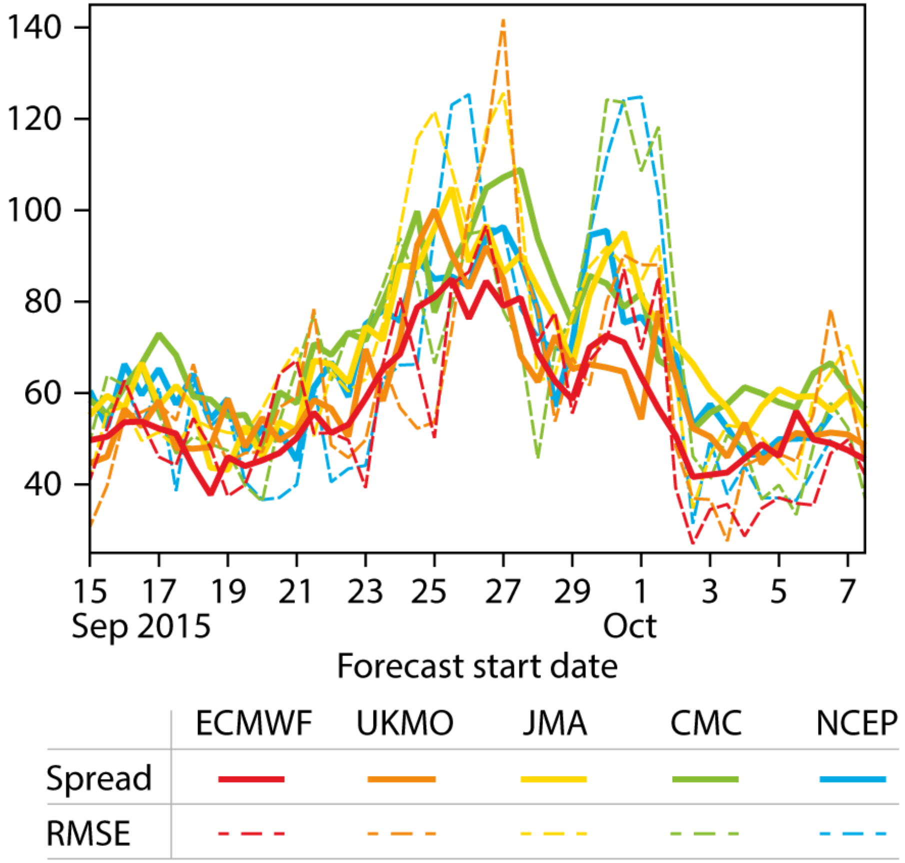
Z500

Northern Hemisphere, annual mean



Europe, day 6

Rodwell 2016, ECMWF Newsletter



Overall Error and Spread have reduced and come into alignment; due to better observations, initial conditions, forecast model and better representation of uncertainty

… but uncertainty varies from day-to-day. The real reason we make ensemble forecasts. What causes this, and how can we evaluate it in our forecasts?

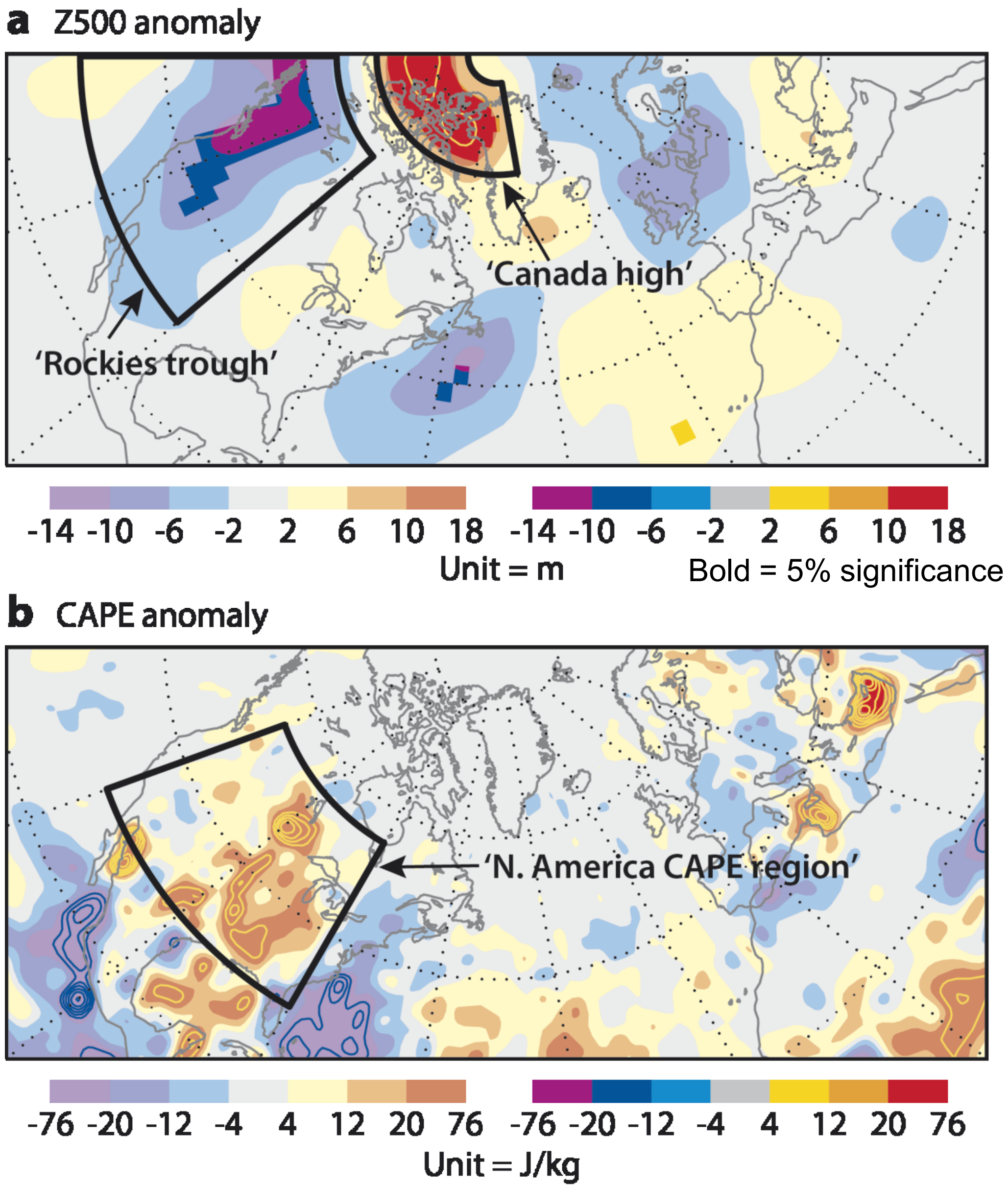
To make progress, we must avoid too much chaos, and look at the growth of uncertainty at very short lead-times

500 hPa geopotential height (Z500). "Error" is RMS of ensemble-mean error

Spread = ensemble standard deviation (scaled to take account of finite ensemble size)

Average initial conditions of 584 single forecast “busts” over Europe at day 6

Rodwell et al, 2013, BAMS



Trough over the Rocky mountains, with high convective potential ahead

Conducive to the formation of mesoscale convection

Can average over such cases to evaluate flow-dependent reliability and thus our model uncertainty

(Subsequent evaluation requires independent data to avoid misleading results)

‘CAPE’ = Convective Available Potential Energy

Mesoscale convection over Kansas

Systems grow to typically 500km in scale, with embedded convective cells and tornados

The Jetstream and mesoscale convection: “The piano string and hammer”

54 cases

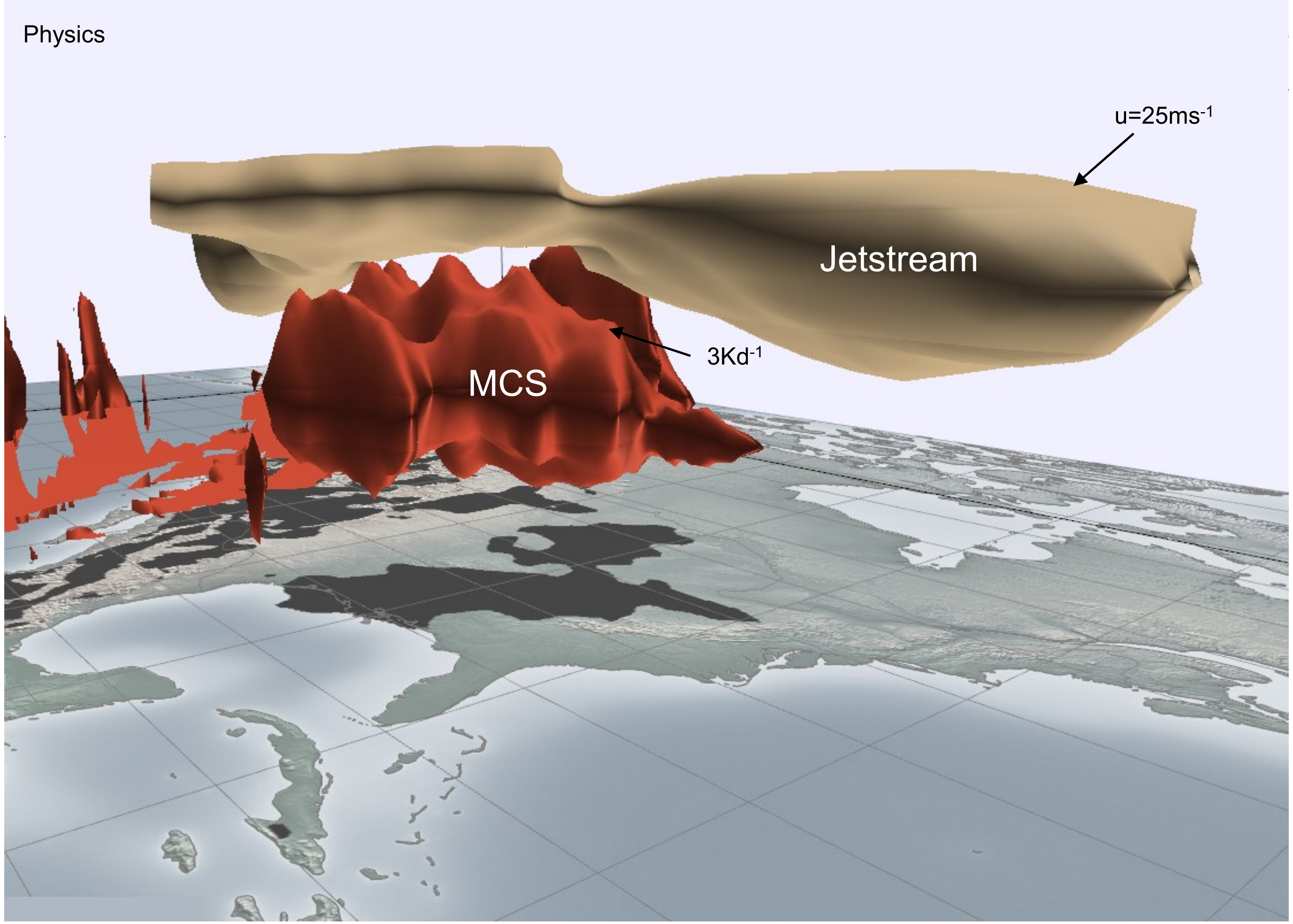
Met3D: Marc Rautenhaus

If we don’t hit the string hard enough, the wave in the string will be too weak

If we hit the string at the wrong time, the wave will arrive over Europe at the wrong time

We do not know when to press the key (mesoscale convection itself involves chaotic uncertainty)

What we want is that the ensemble members generate such convection with the “right” uncertainty



Reliability in ensemble data assimilation

Chaos makes it difficult to identify problems in the medium-range using the spread-error relationship

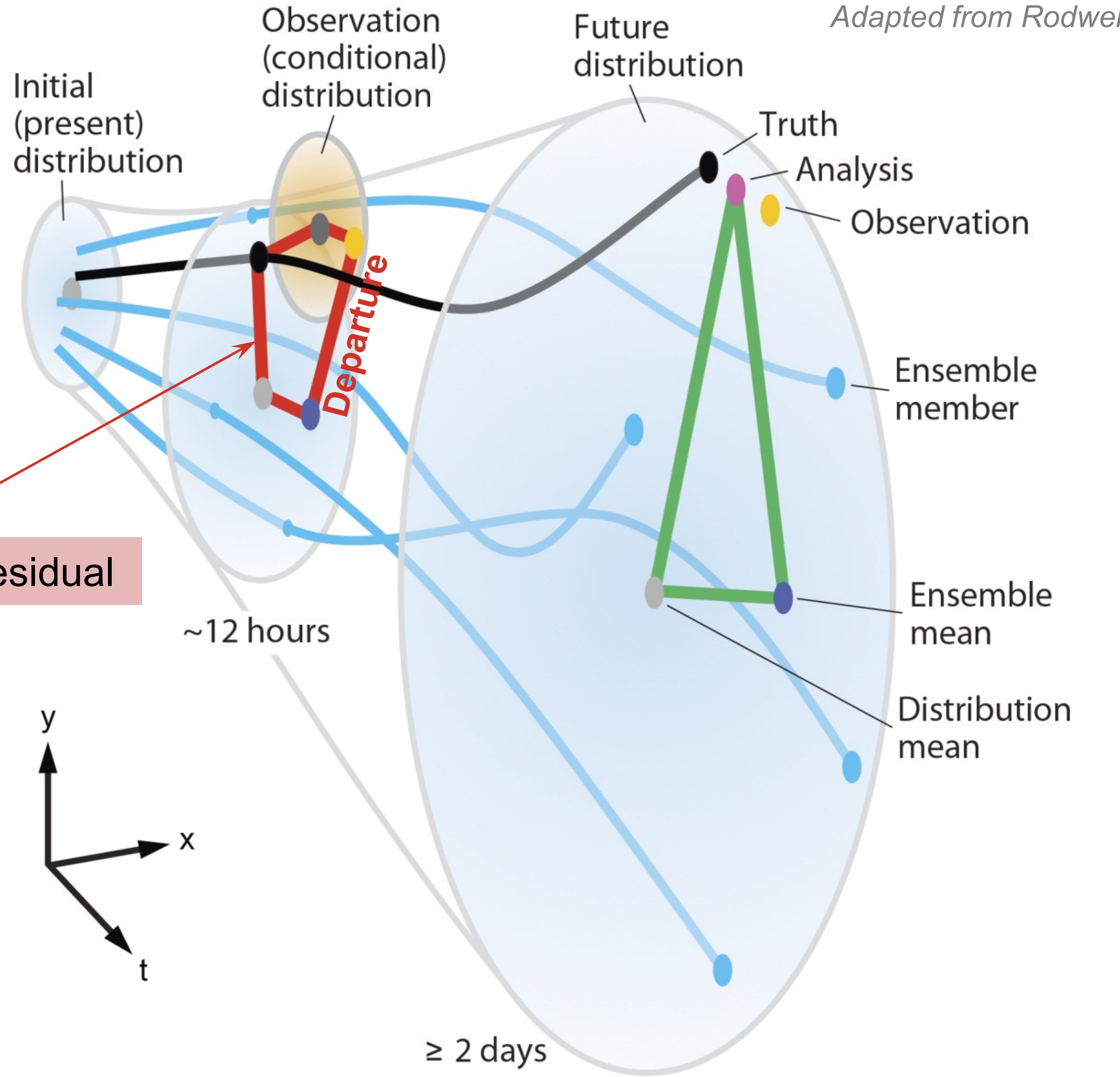
Go to much shorter lead-times – within ensemble data assimilation process

Need to take account for observation error

Obtain diagnostic equation to evaluate "instantaneous" growth of uncertainty

Depar^2 = Bias^2 + EnsVar + ObsUnc^2 + Residual

Adapted from Rodwell et al. (2015) QJRMS



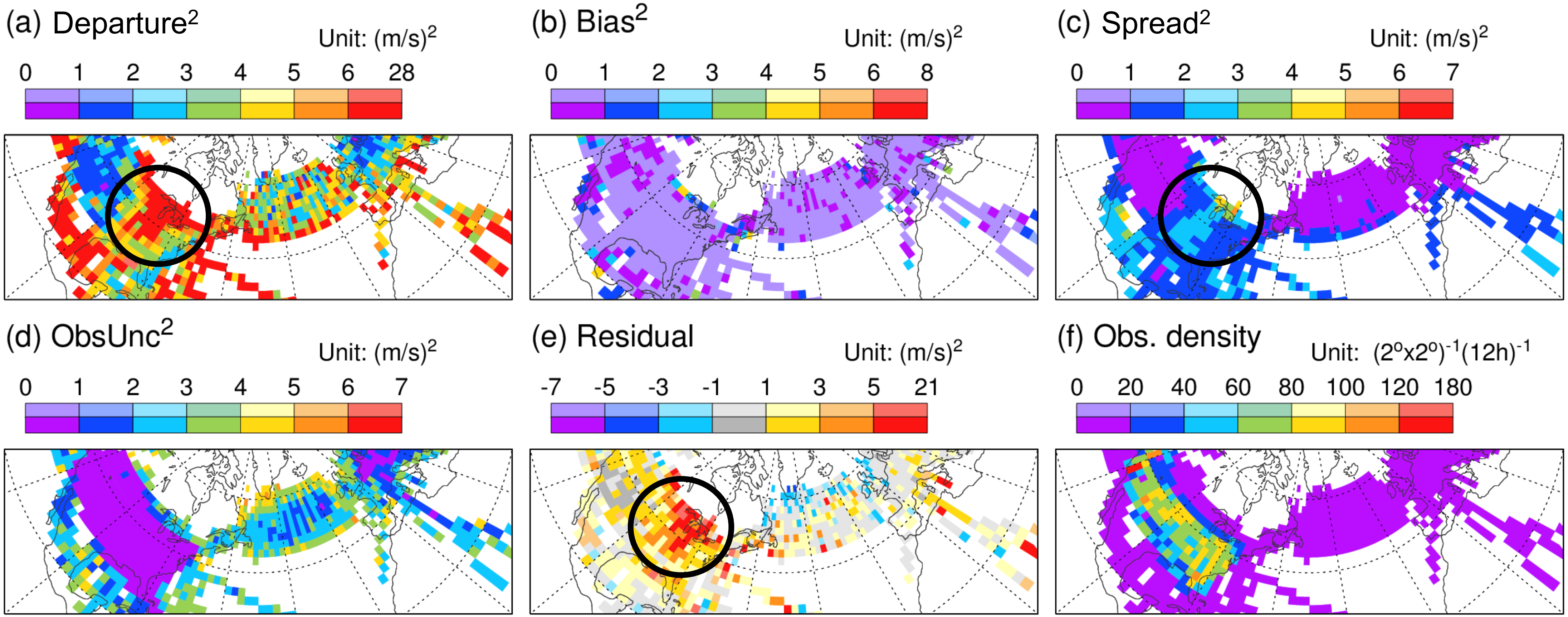
(Cross-terms on squaring have zero expectation. EnsVar is scaled variance to account for finite ensemble-size)

Evaluating model uncertainty in upper-level winds during “Rocky trough” situations

54 cases

Rodwell 2016, ECMWF Newsletter

Relative to aircraft west-east wind observations at 200hPa (±15)



Ensemble spread highlights enhanced uncertainty around the Great Lakes

Large errors ensue

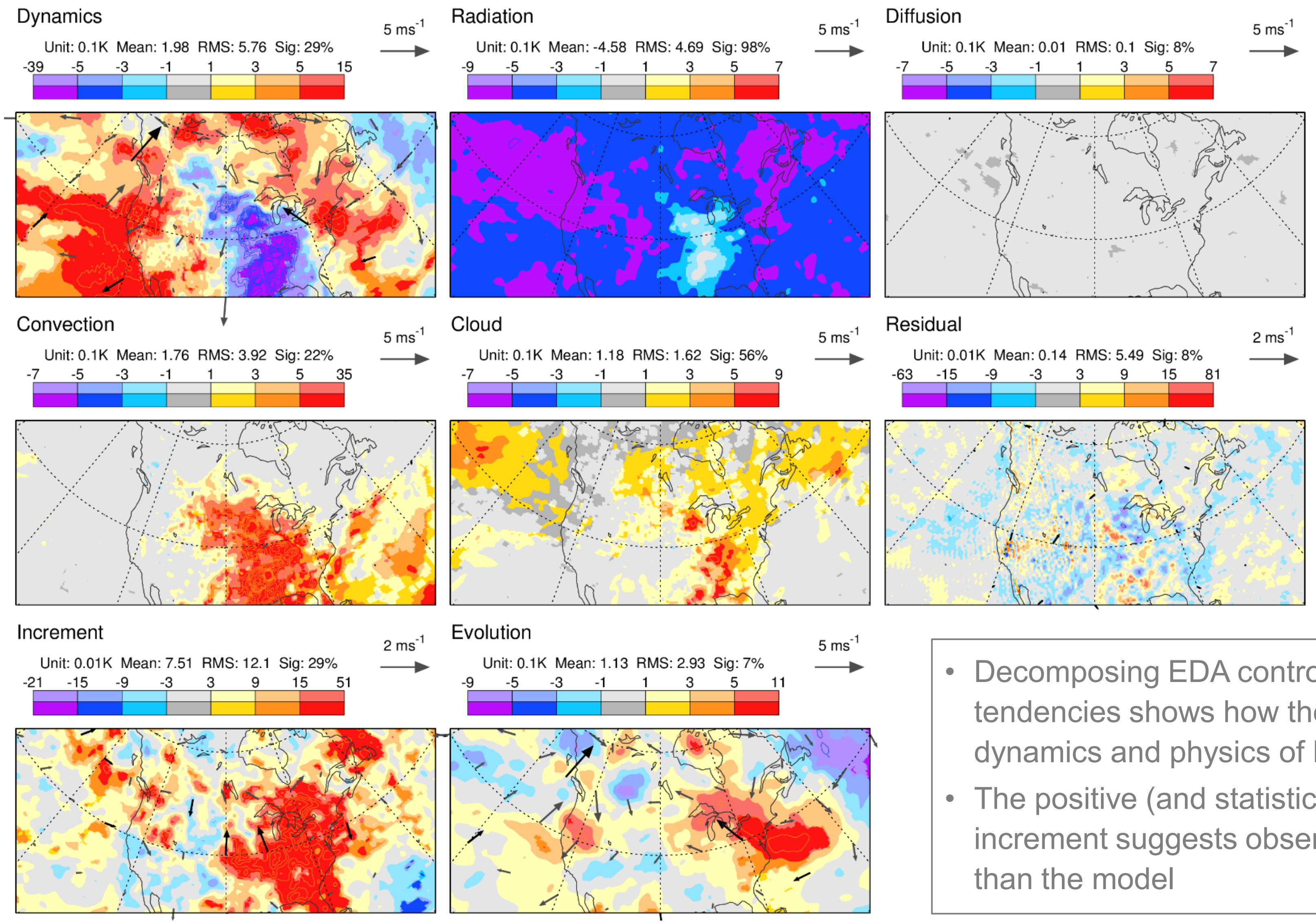
Errors are relative to observations that are also uncertain but, even if we take this into account, there appears to be too little spread (and model uncertainty) in this flow situation

Depar^2 = Bias^2 + Spread^2 + ObsUnc^2 + Residual

Reliability $\Rightarrow$ $\mathbb{E}$[Residual]=0

Initial tendency budget from control forecast during “Rocky trough” situations

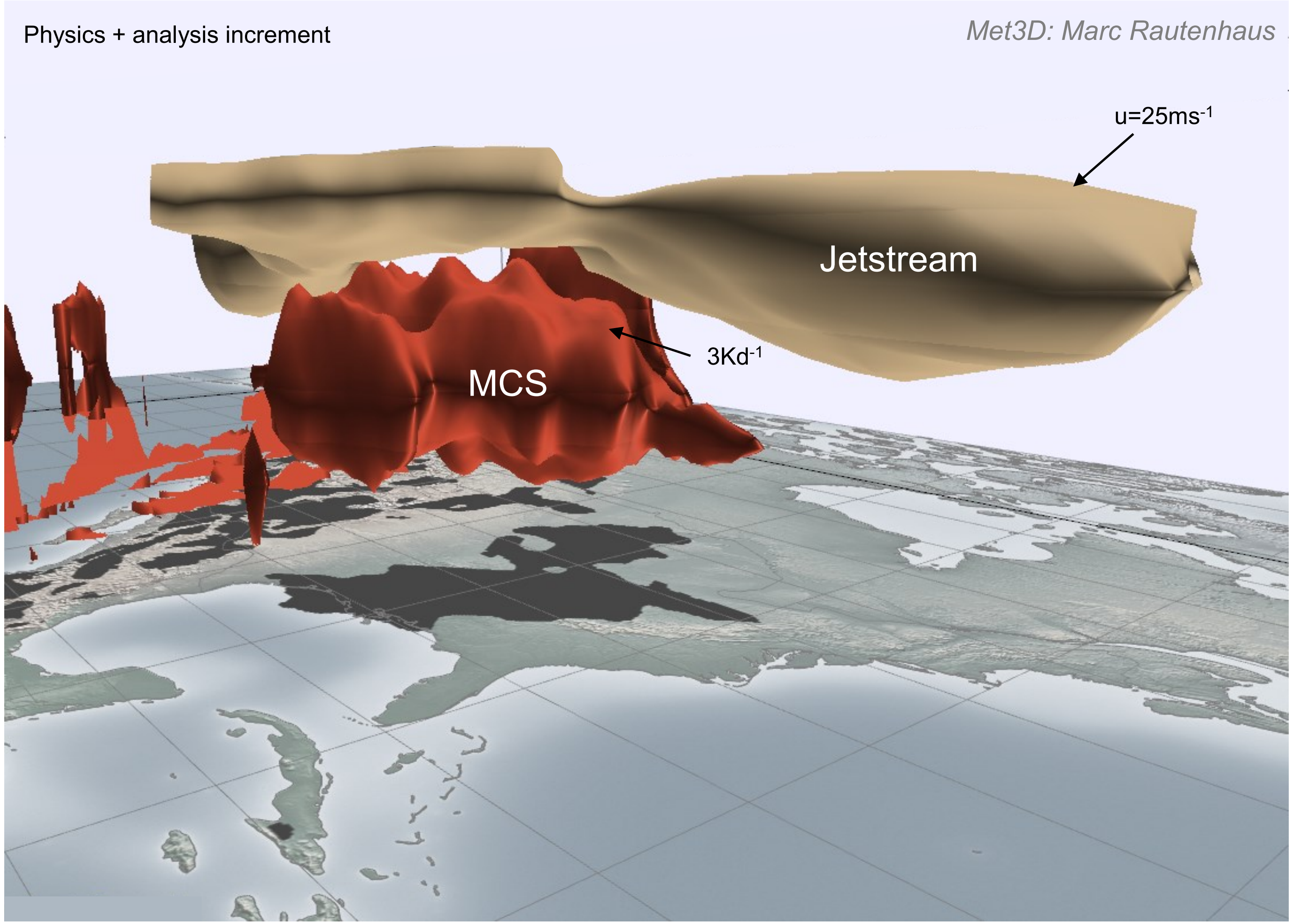
T300, 54 cases



- Decomposing EDA control forecast into process tendencies shows how the model represents dynamics and physics of MCS
- The positive (and statistically significant) increment suggests observations are warmer than the model

Process tendencies accumulated over 12hr background, the analysis increment, and evolution of the flow

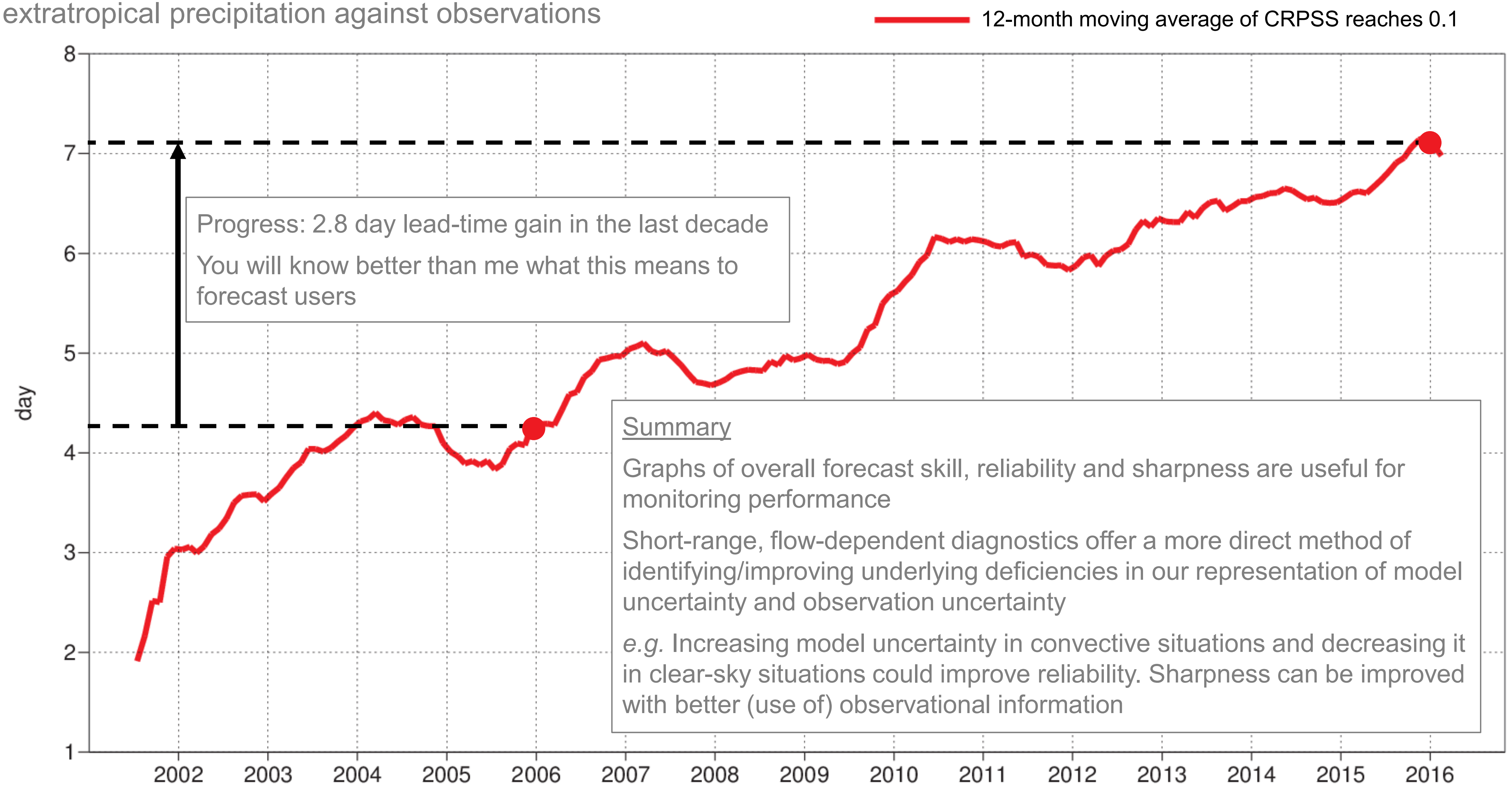
MCS – Jetstream interaction (composite)



- *Increments emphasize model systematic error: MCS does not interact enough with Jetstream*
- *Also need to strengthen stochastic physics to increase background variance?*

Trend in probabilistic forecast performance & Summary

CRPSS, extratropical precipitation against observations



Progress: 2.8 day lead-time gain in the last decade

You will know better than me what this means to forecast users

Summary

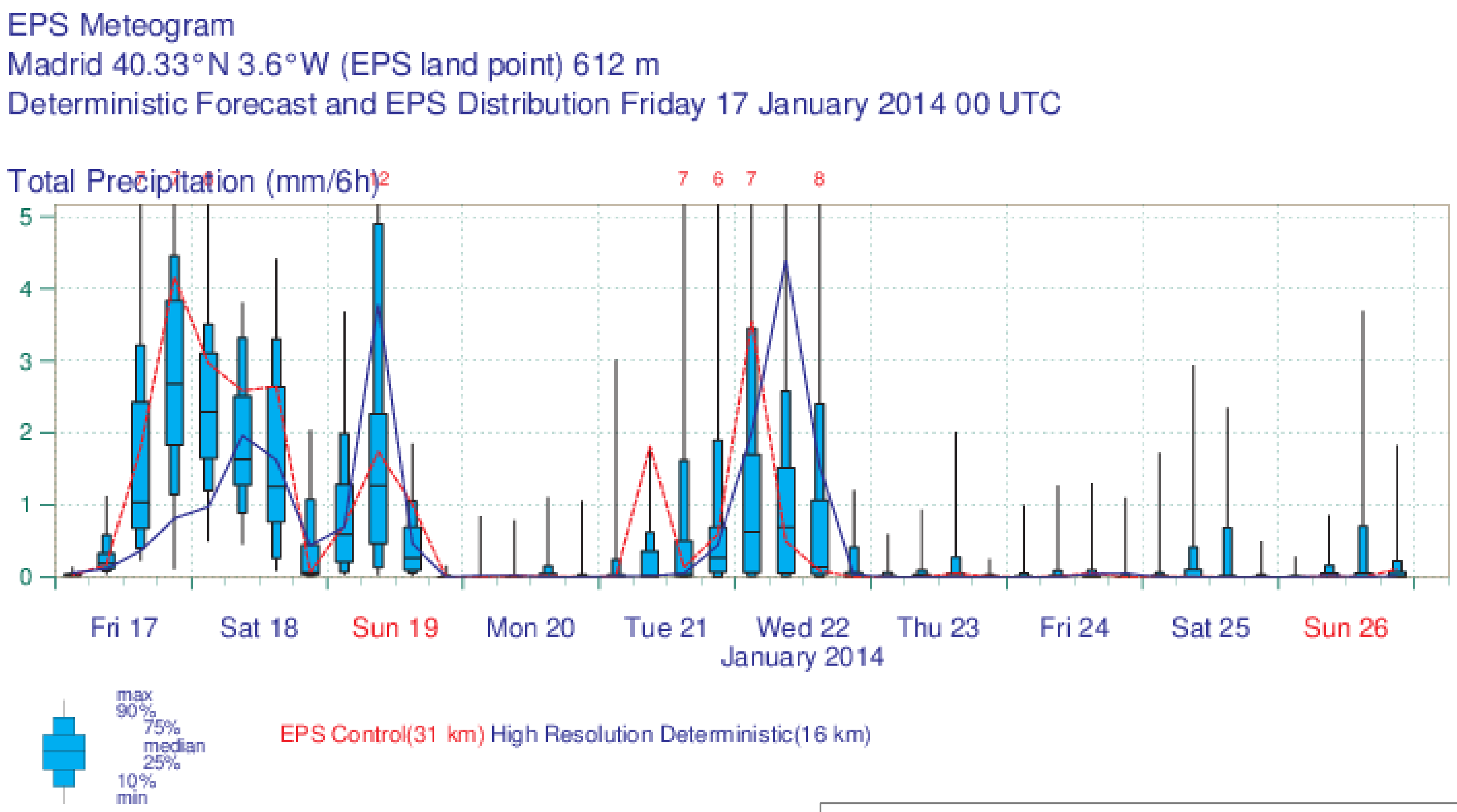
Graphs of overall forecast skill, reliability and sharpness are useful for monitoring performance

Short-range, flow-dependent diagnostics offer a more direct method of identifying/improving underlying deficiencies in our representation of model uncertainty and observation uncertainty

e.g. Increasing model uncertainty in convective situations and decreasing it in clear-sky situations could improve reliability. Sharpness can be improved with better (use of) observational information

Optimising forecast configuration and usage

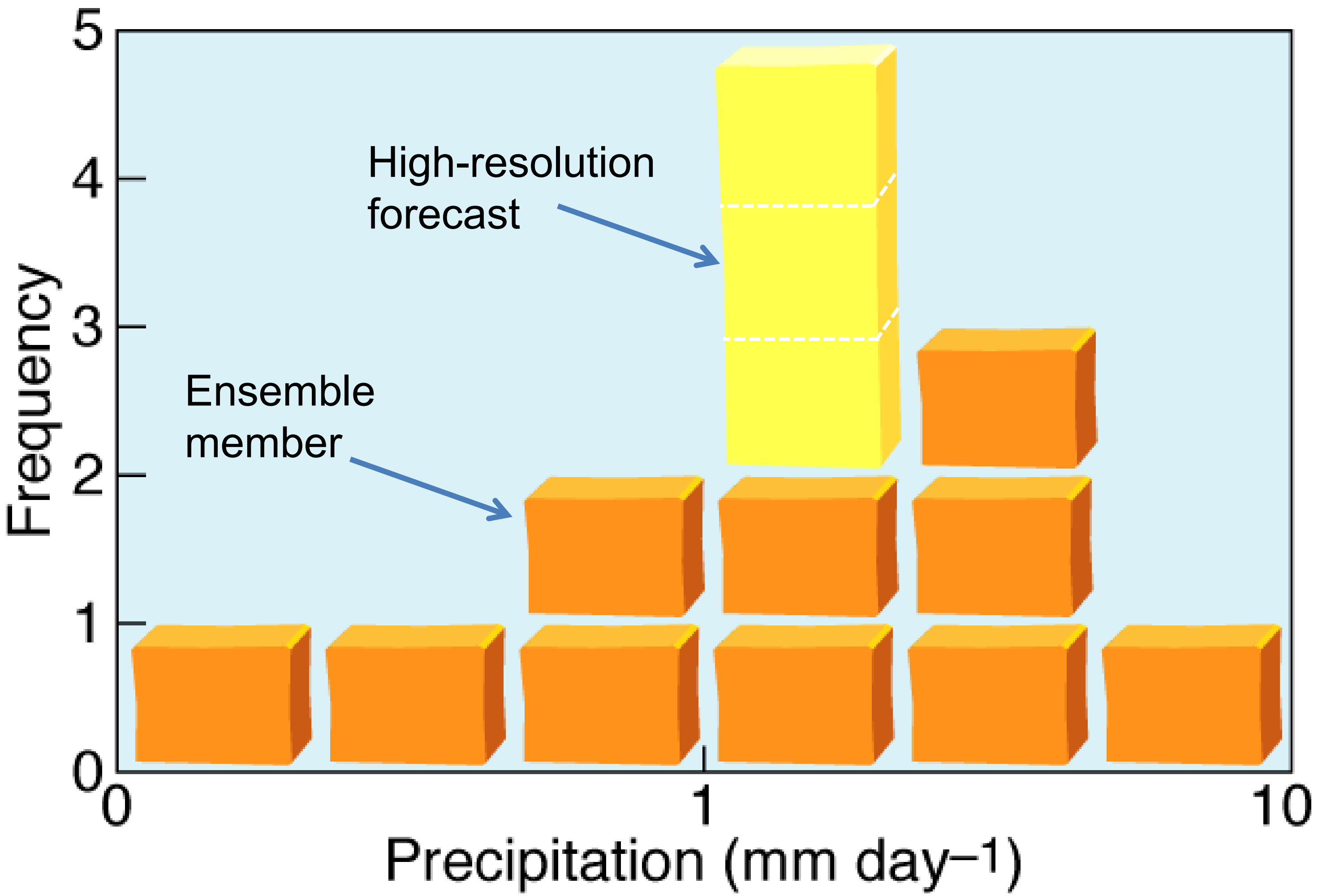
Ensemble and high-resolution information



- How do we optimally combine information from the ensemble and the high-resolution forecast?
- Is this dependent on lead-time?

Combined Prediction System: Methodology

Rodwell, ECMWF Newsletter 106 (2006)

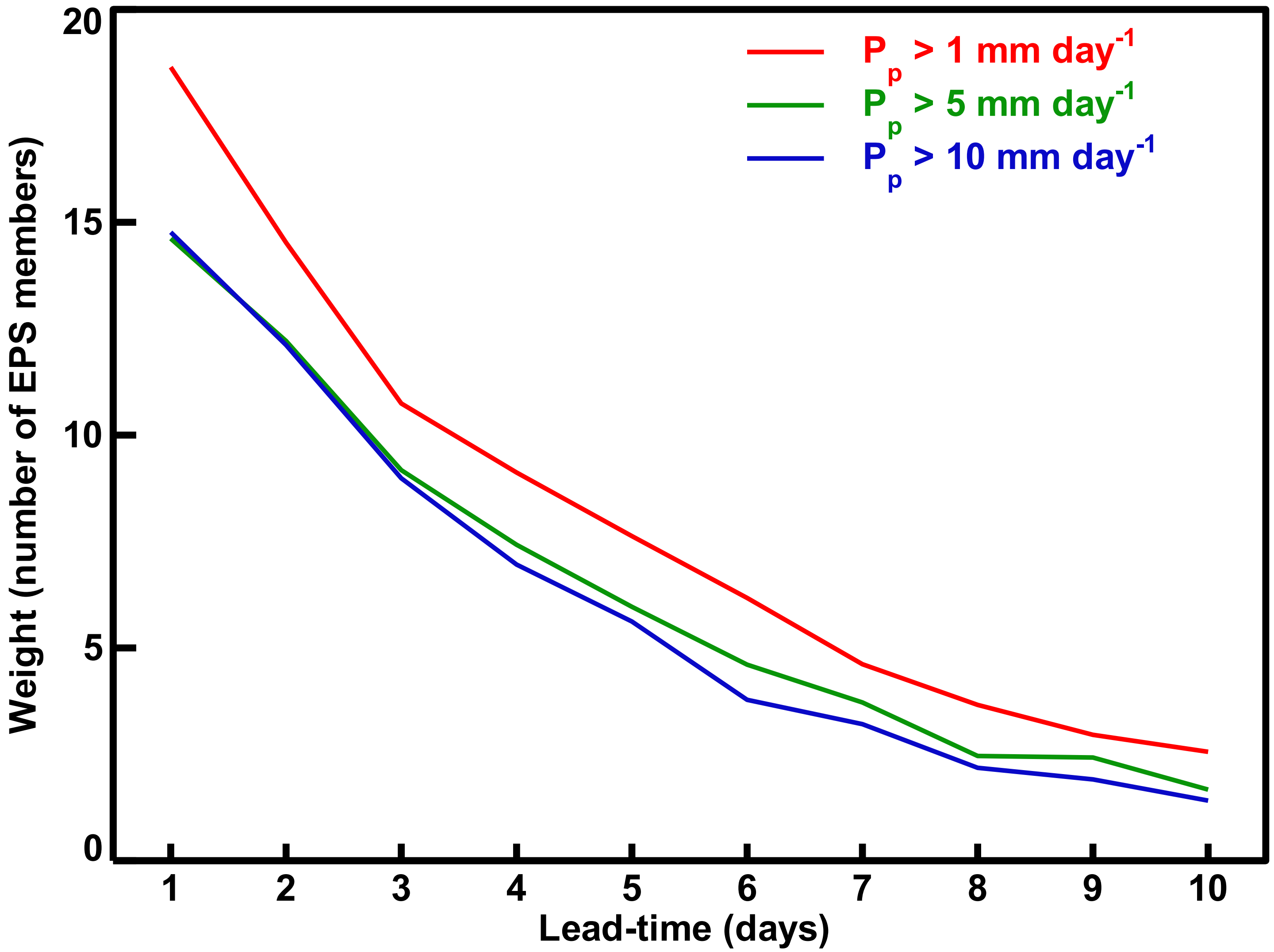


In the example, the weight for the high-resolution forecast is $w_{HRES} = 3$ and the probability of 1mm precipitation = 9/13

In the real case, find w_{HRES} that maximises (e.g.) Brier Skill Score or Ignorance score

Can do analytically by solving $dBSS/dw_{HRES} = 0$

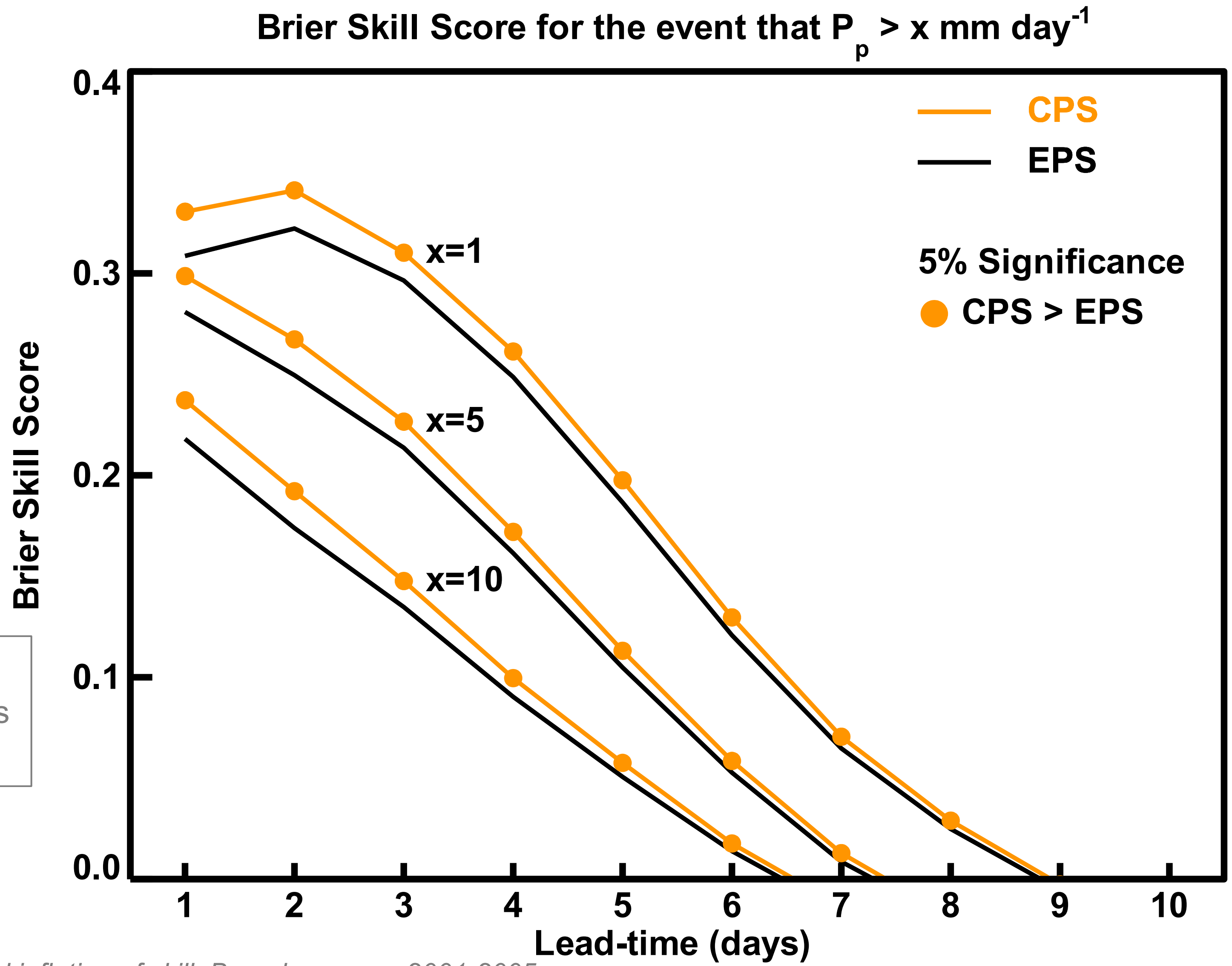
The weight to give the high-resolution system



At short lead-times, high-resolution system is very valuable. At longer lead-times weight →1. Based on years 2001-2005

Combined system is more skilful

Rodwell, ECMWF Newsletter 106 (2006)



The combined system is more skilful (on average) at all leadtimes and for each threshold

Results are cross-validated so no artificial inflation of skill. Based on years 2001-2005

Discussion

- Rossby waves and the “Rossby Wave Source” – simple models make useful diagnostics
- Flow-dependent reliability is key – EDA reliability budget seems useful
  - Effective, efficient, focuses on reliability - not sharpness
  - Development and diagnosis of ensemble data assimilation likely to be key to future NWP progress
- Instabilities as magnifiers of uncertainty
- Approaches to optimising system configuration and combining multiple sources of information

Previous talk

- Waves and spatio-temporal variability
- Initial tendencies approach