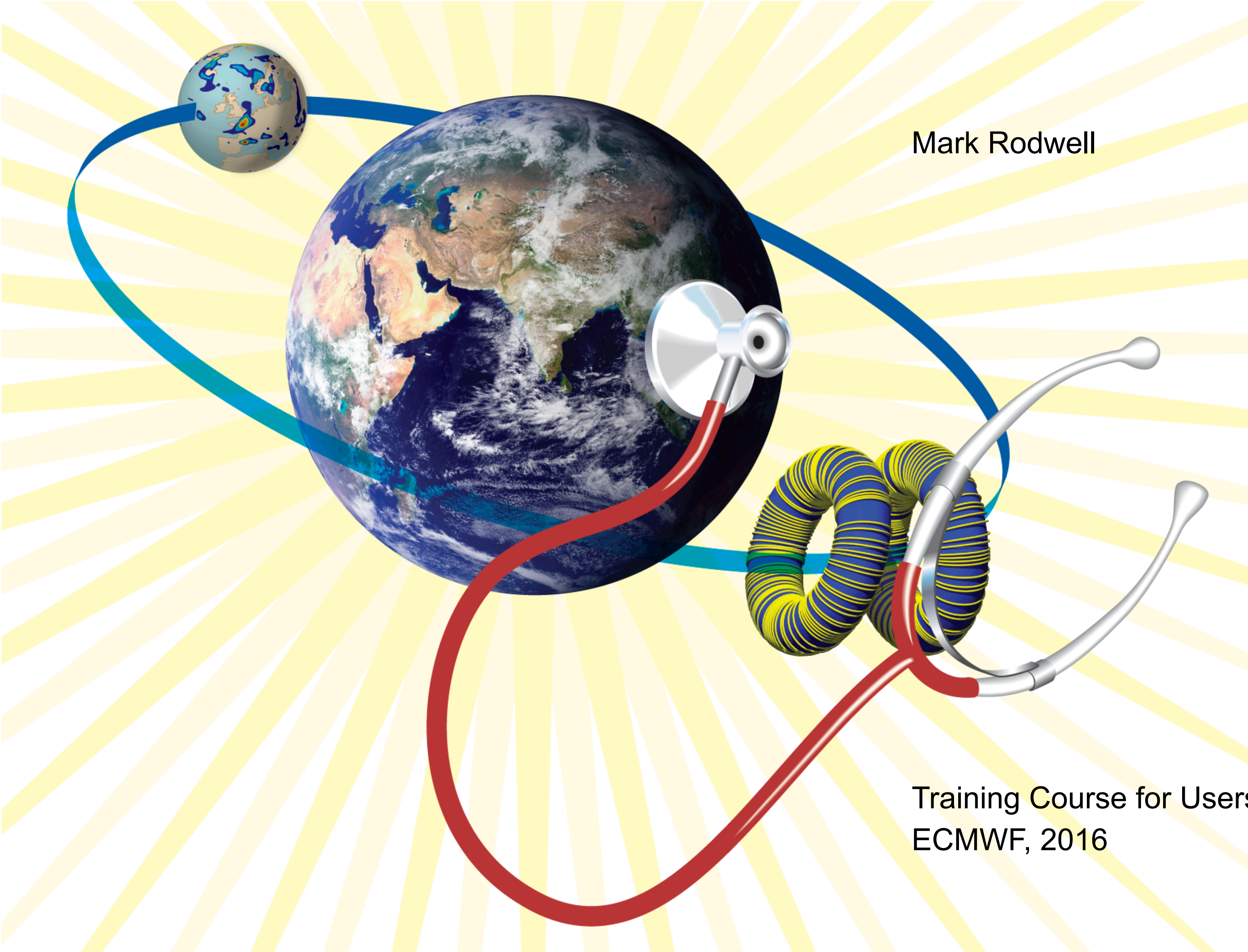
Model Errors & Diagnostic Tools



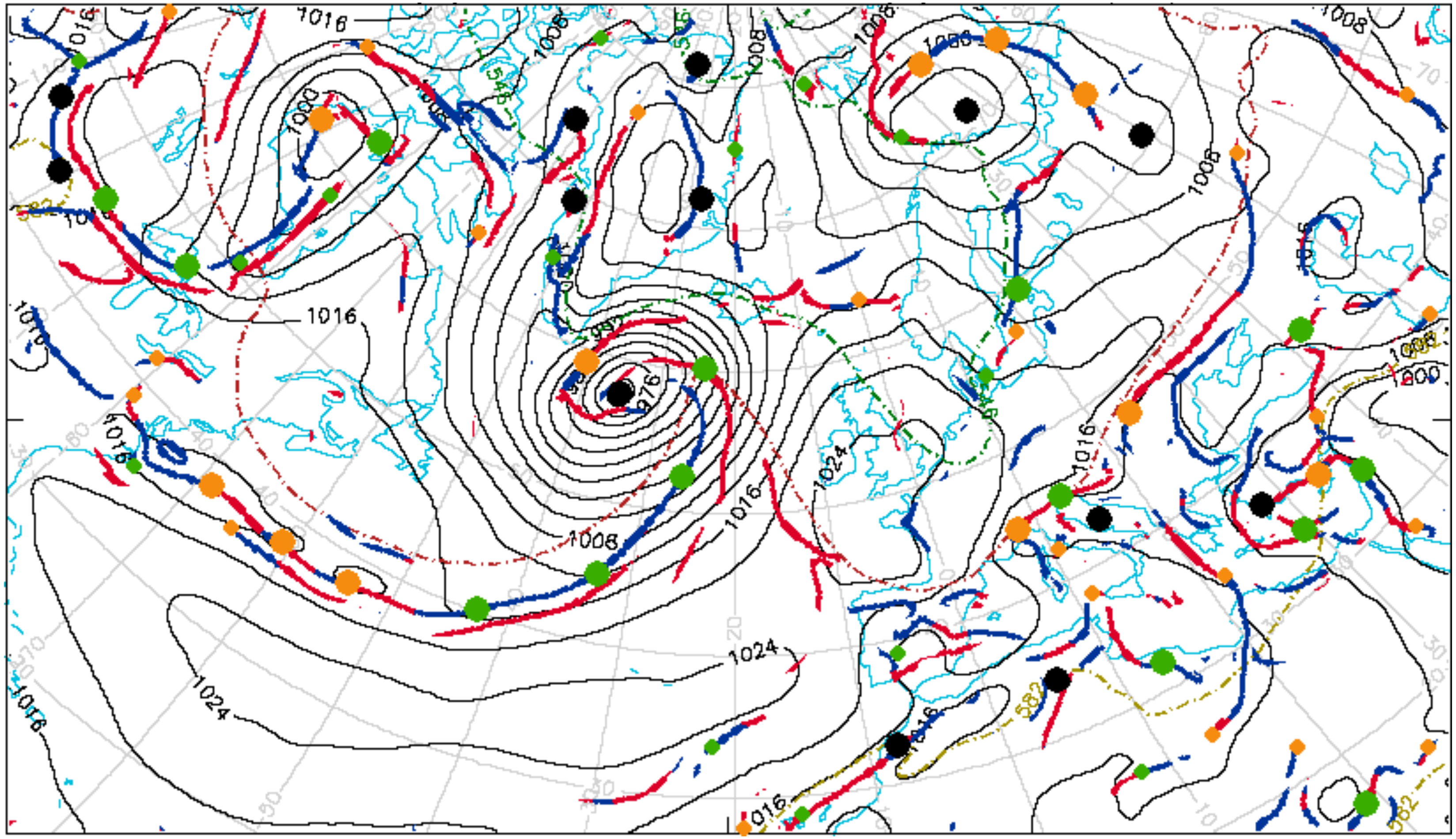
Mark Rodwell

Training Course for Users
ECMWF, 2016

Diagnostics

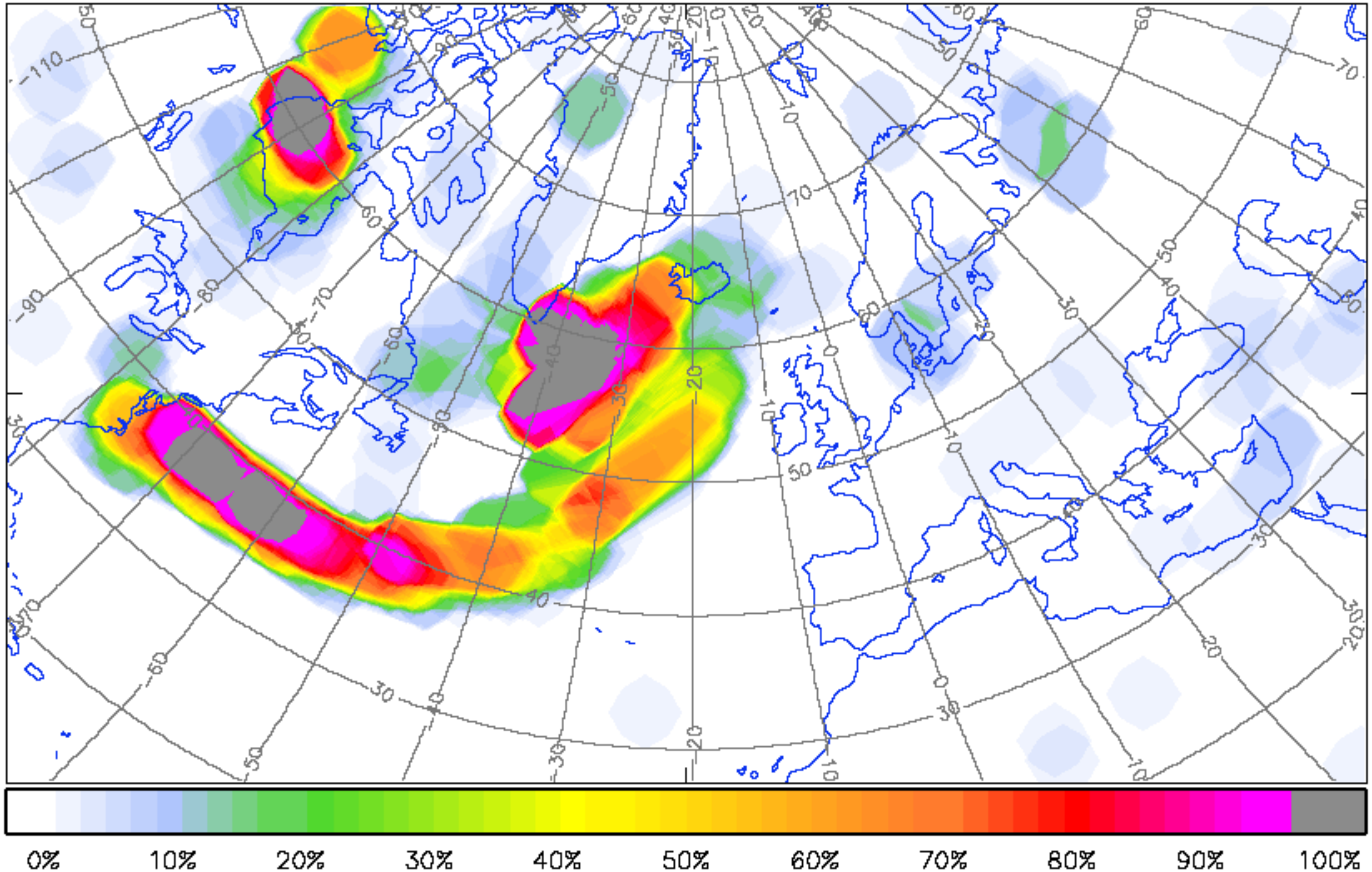
Strong extratropical cyclone on 2012/07/21

Surface analysis at 12UTC



Cold front   Warm front   Barotropic low   Frontal wave   Diminutive frontal wave

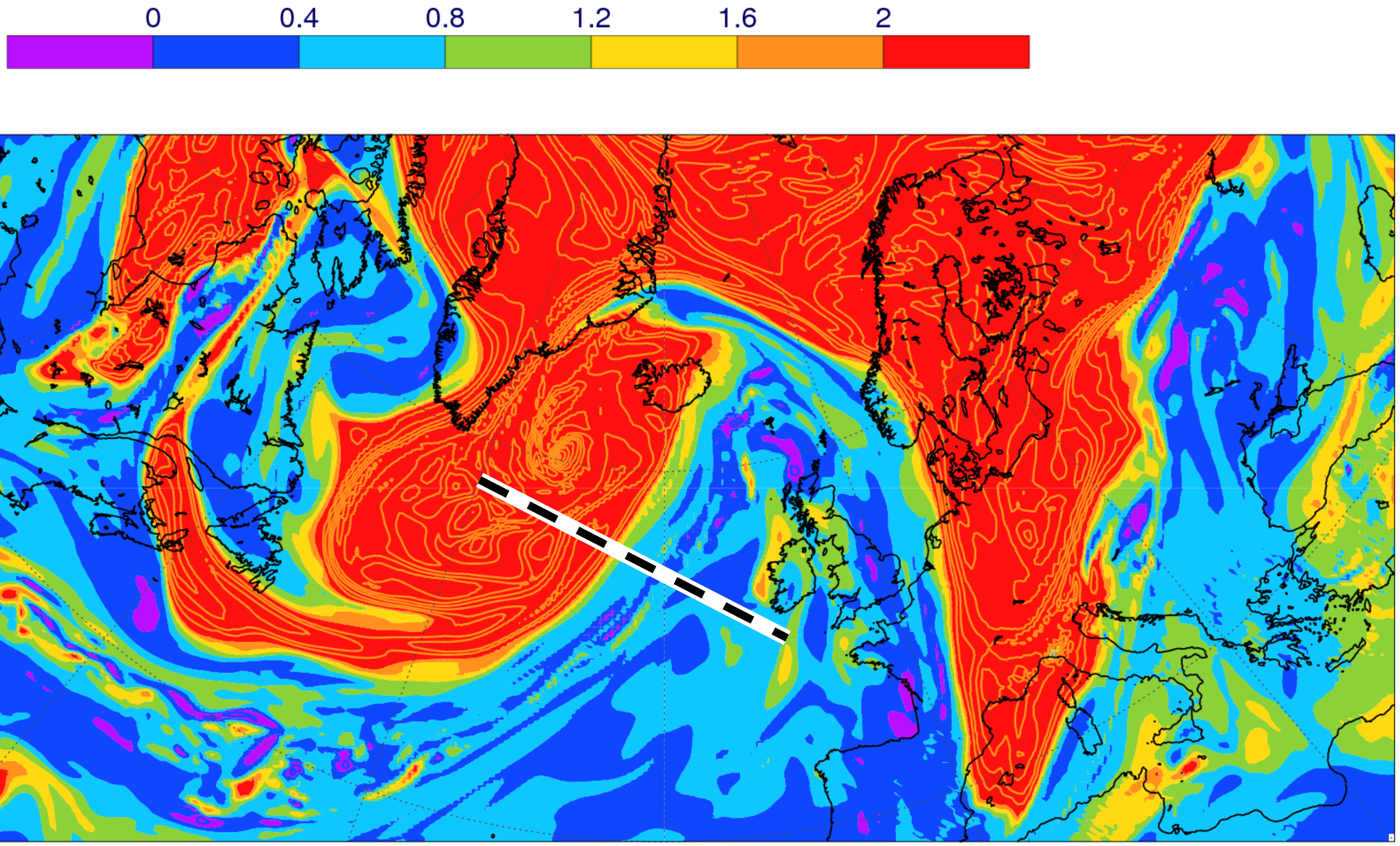
24-hr probability of gales on 2012/07/21



% members predicting gales (1km wind-speed > 34kts) within 300km in a 24-hour period.

Analysed Potential Vorticity on 330K at 12UTC

~400 hPa

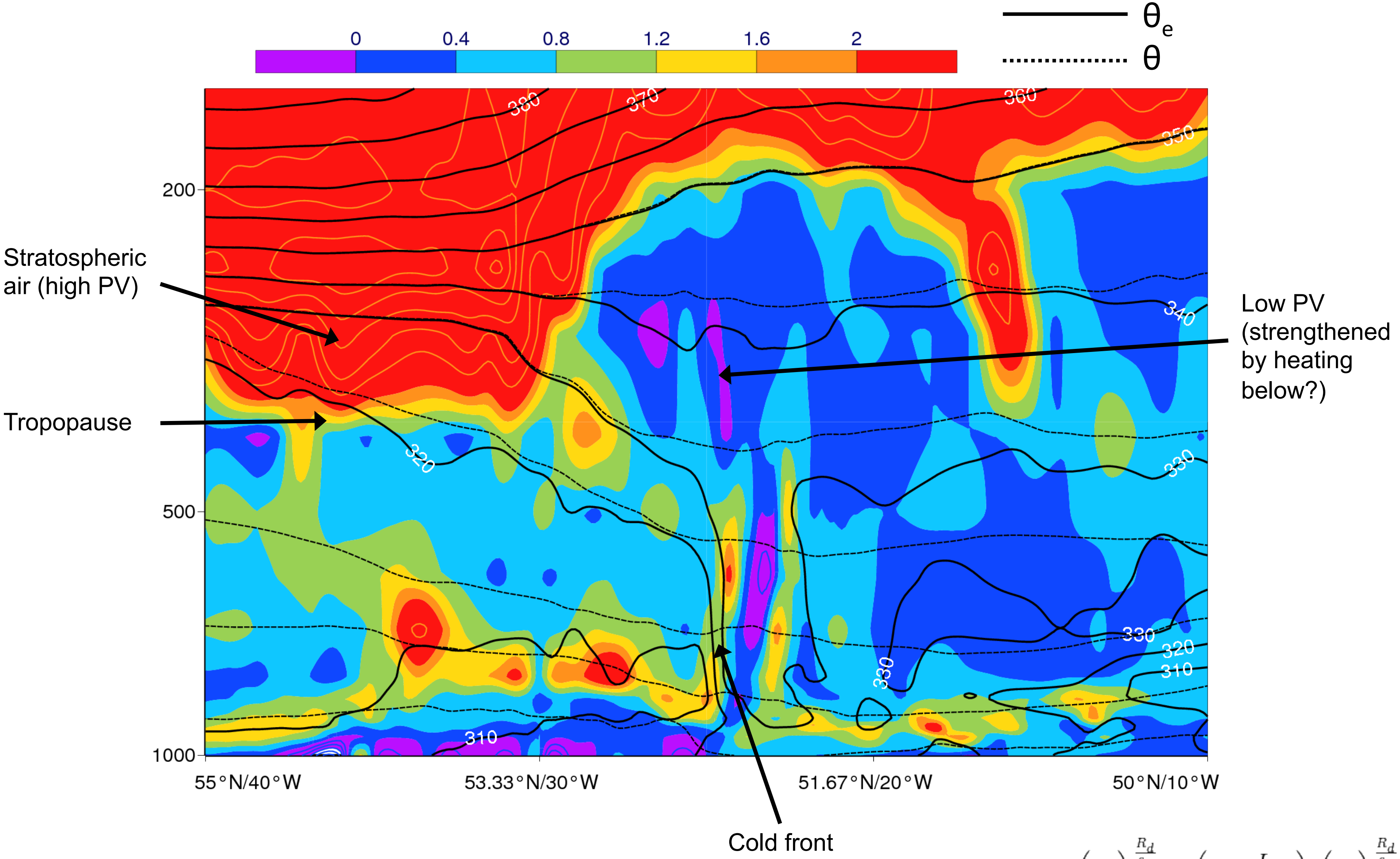


$$PV = -g\zeta_\theta \frac{\partial \theta}{\partial p}$$

PV tendencies are a nice way of looking at how the dynamics (advection) and physics (material tendency) contribute to the flow evolution

Contour interval: 0.4PVU below 2PVU. 2PVU above 2PVU

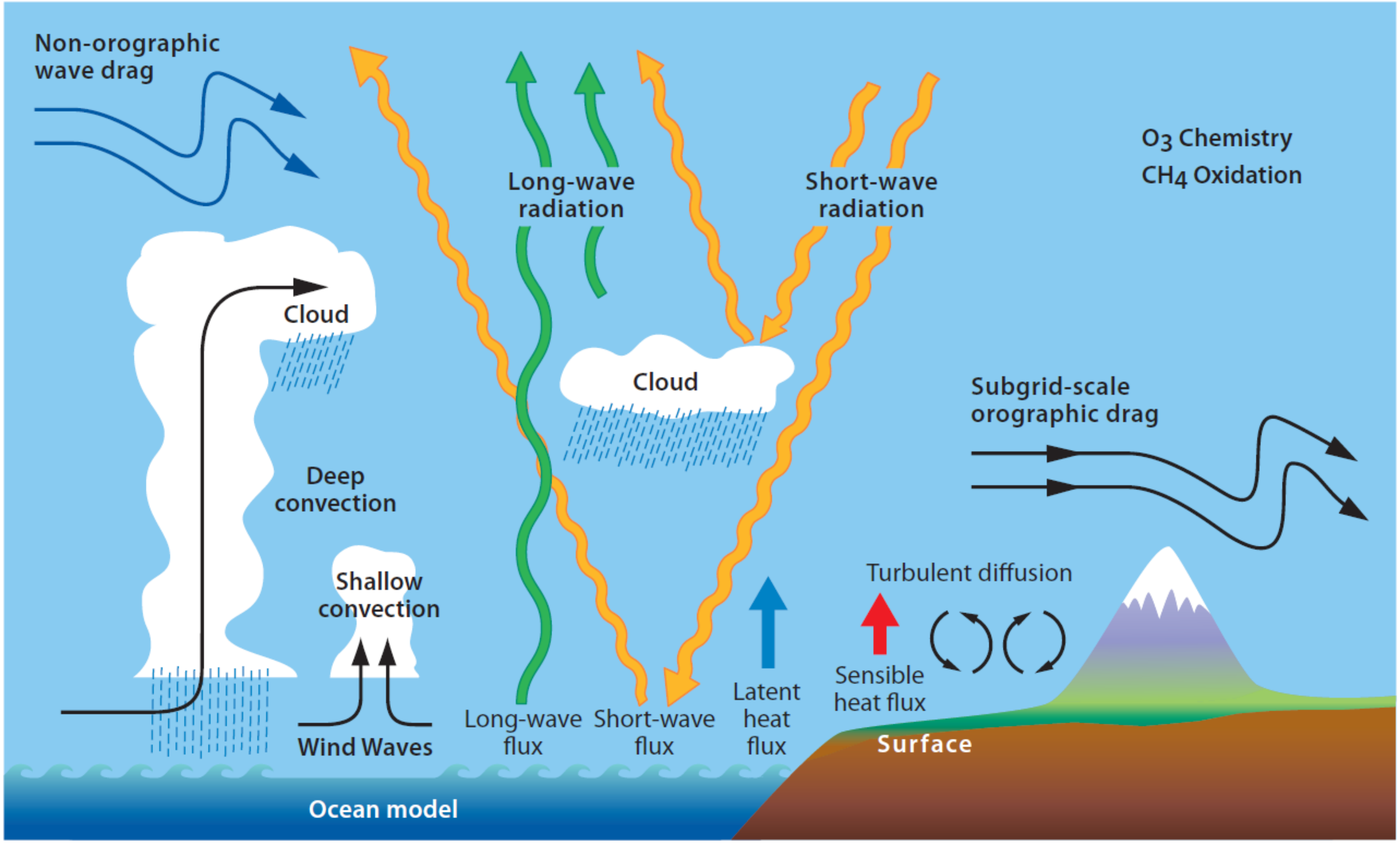
Analysed PV, θ, θ_e cross-section 18UTC



$$\theta_e = T_e\left(\frac{p_0}{p}\right)^{\frac{R_d}{c_p}} \approx \left(T + \frac{L_v}{c_p}r\right)\left(\frac{p_0}{p}\right)^{\frac{R_d}{c_p}}$$

Contour interval: 0.4PVU below 2PVU. 2PVU above 2PVU

The complexity of present-day model physics

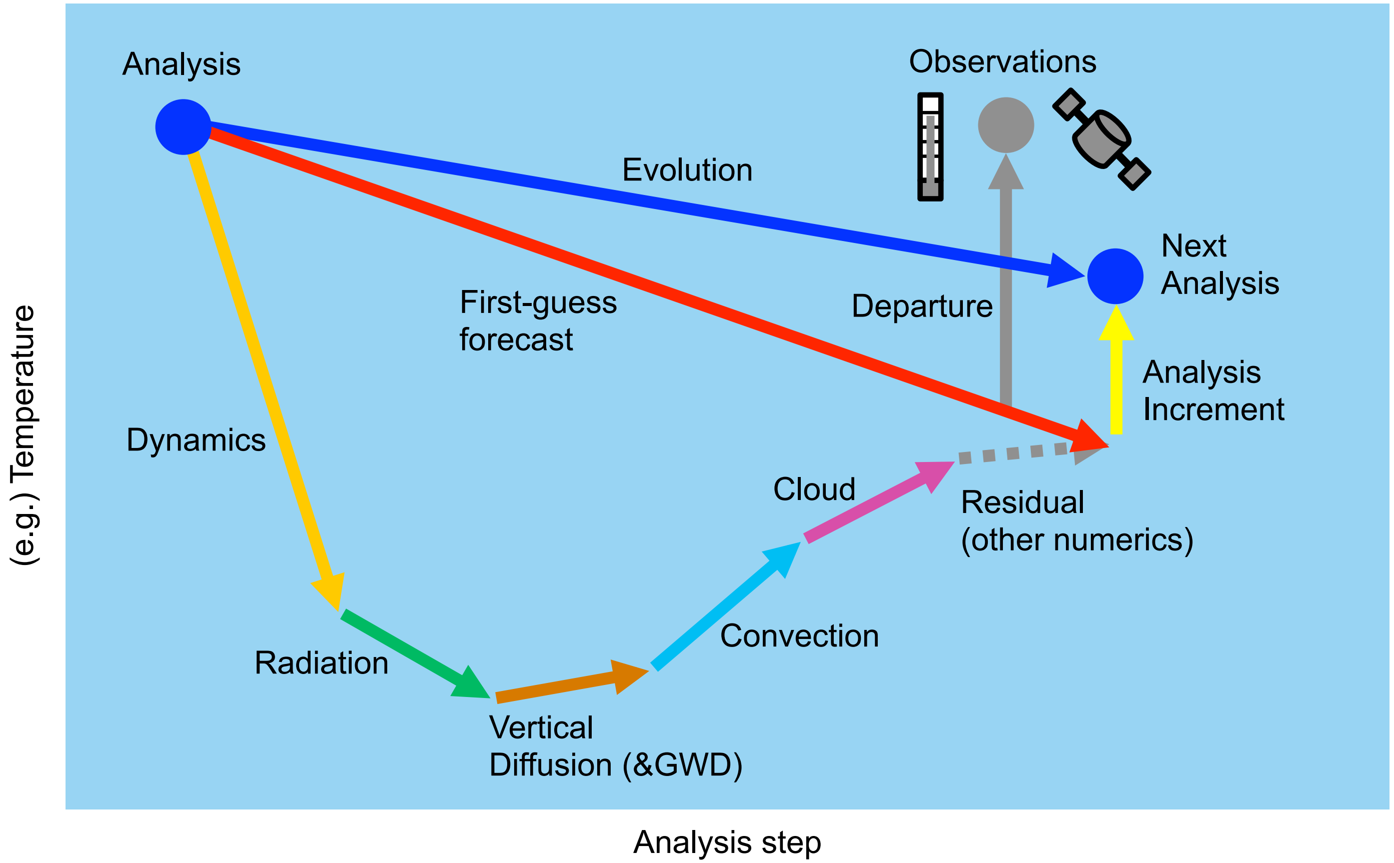


The complexity of today's models, with numerous interactions between physical processes and the resolved flow (including teleconnections), can make it very difficult to isolate the offending process(es). Single column and LES models can help, but these do not take into account the evolution of the resolved flow.

Figure from Peter Bechtold

Diagnosis within the data assimilation system

Schematic of the data assimilation process – a diagnostic perspective



Analysis increment corrects first-guess error, and draws next analysis closer to observations.
First-guess = sum of all processes.
Relationship between increment and individual process tendencies can help identify key errors.

"Initial Tendency" approach discussed by Klinker & Sardeshmukh (1992). Refined by Rodwell & Palmer (2007)

Single forecast process tendencies in T_{925} & $\underline{v}_{925}$

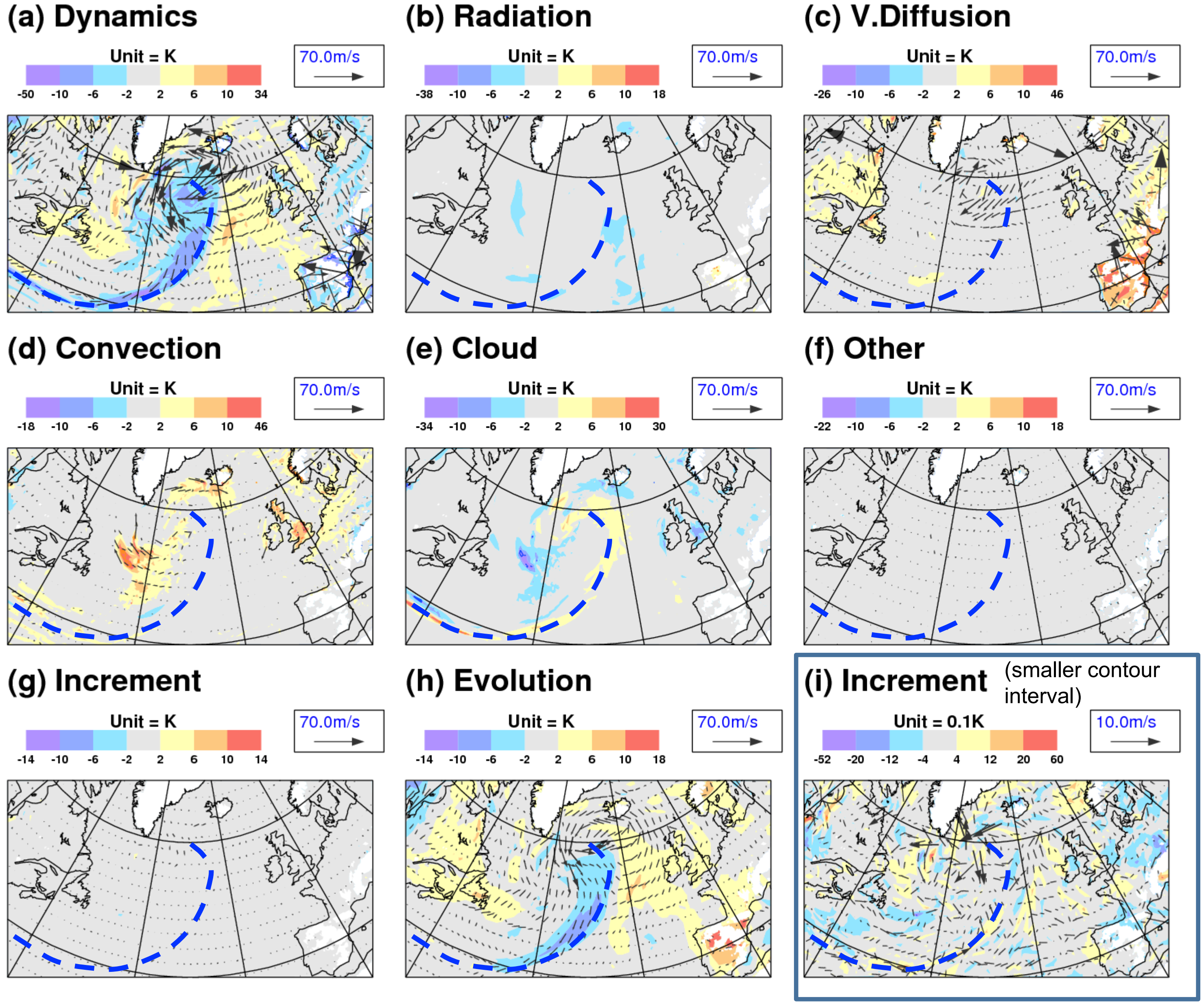
Surface cold front at 12UTC

Increment suggests too much frictional convergence(?) Vertical diffusion too strong?

Need to consider statistical significance. Use composites (see later)

Would be difficult to identify such an error by looking at longer lead-time forecasts or the model's climate

Tendencies are accumulated over first 12 hours of the forecast initialised at 6UTC on 20120721. Increment is at the end of tendency period (18UTC)



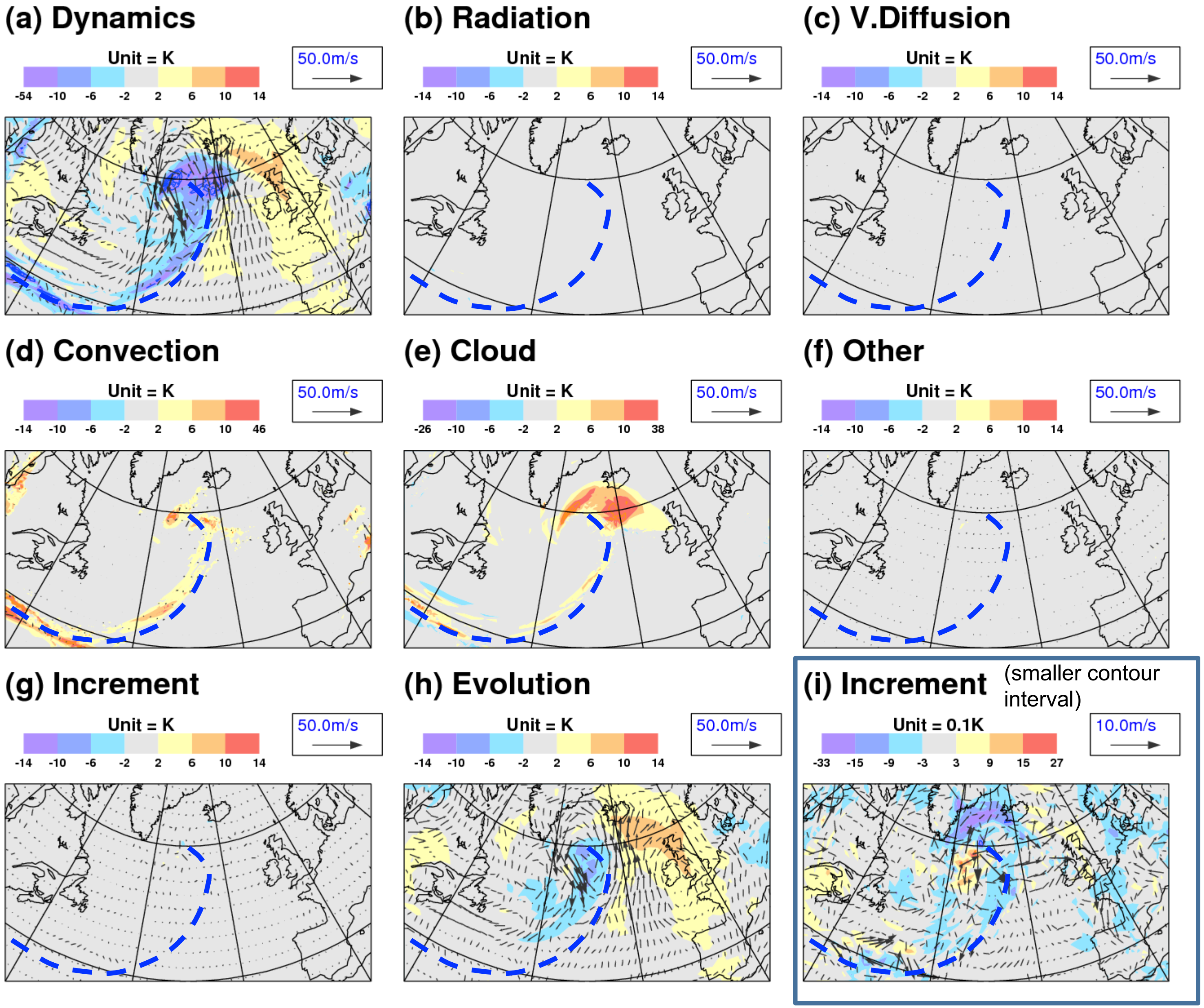
Diagnostics

Single forecast process tendencies in T_{500} & $\underline{v}_{500}$

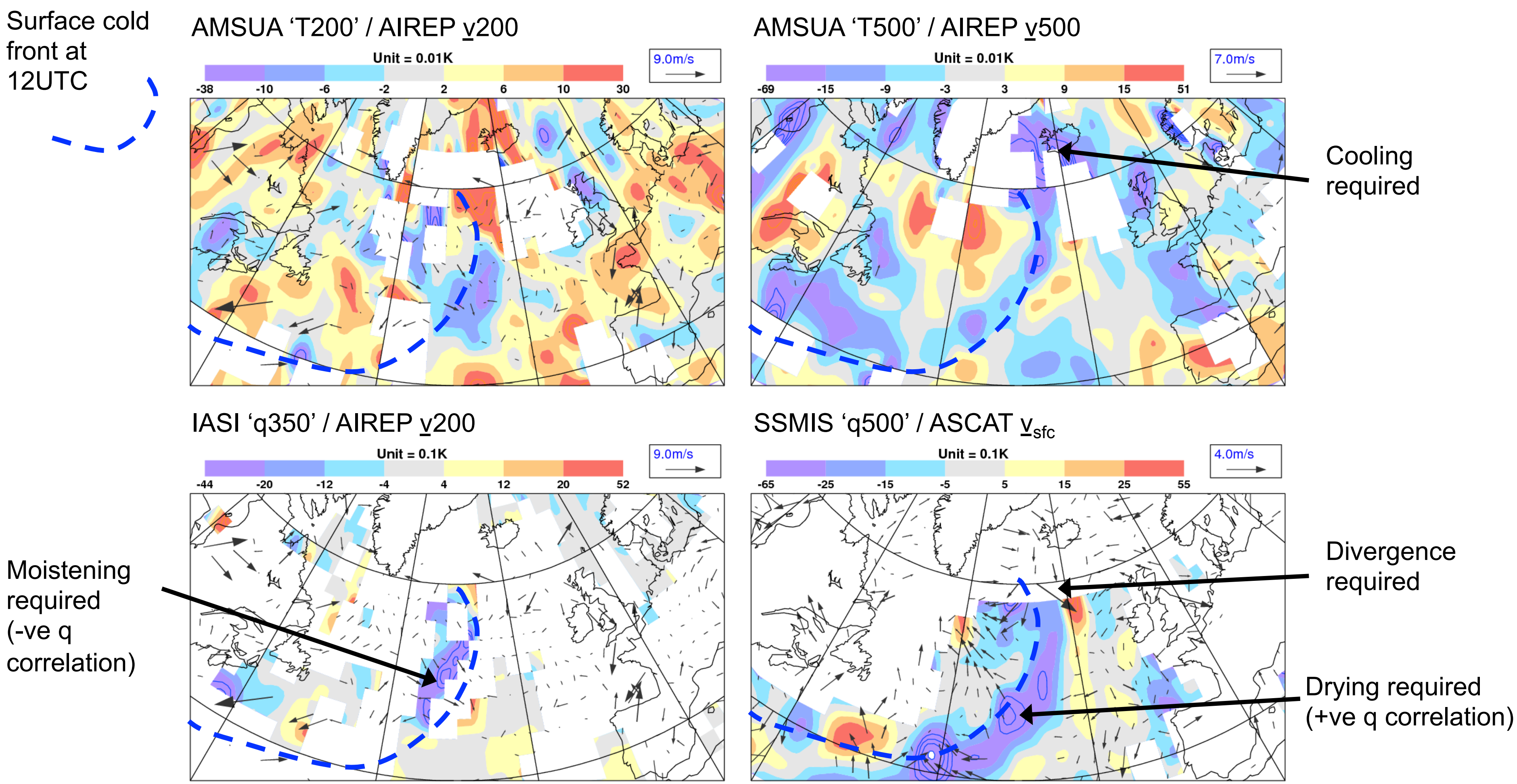
Surface cold front at 12UTC

Increment cools the edge of cloud-head and behind surface front. Cloud too strong?

Tendencies are accumulated over first 12 hours of the forecast initialised at 6UTC on 20120721. Increment is at the end of tendency period (18UTC)



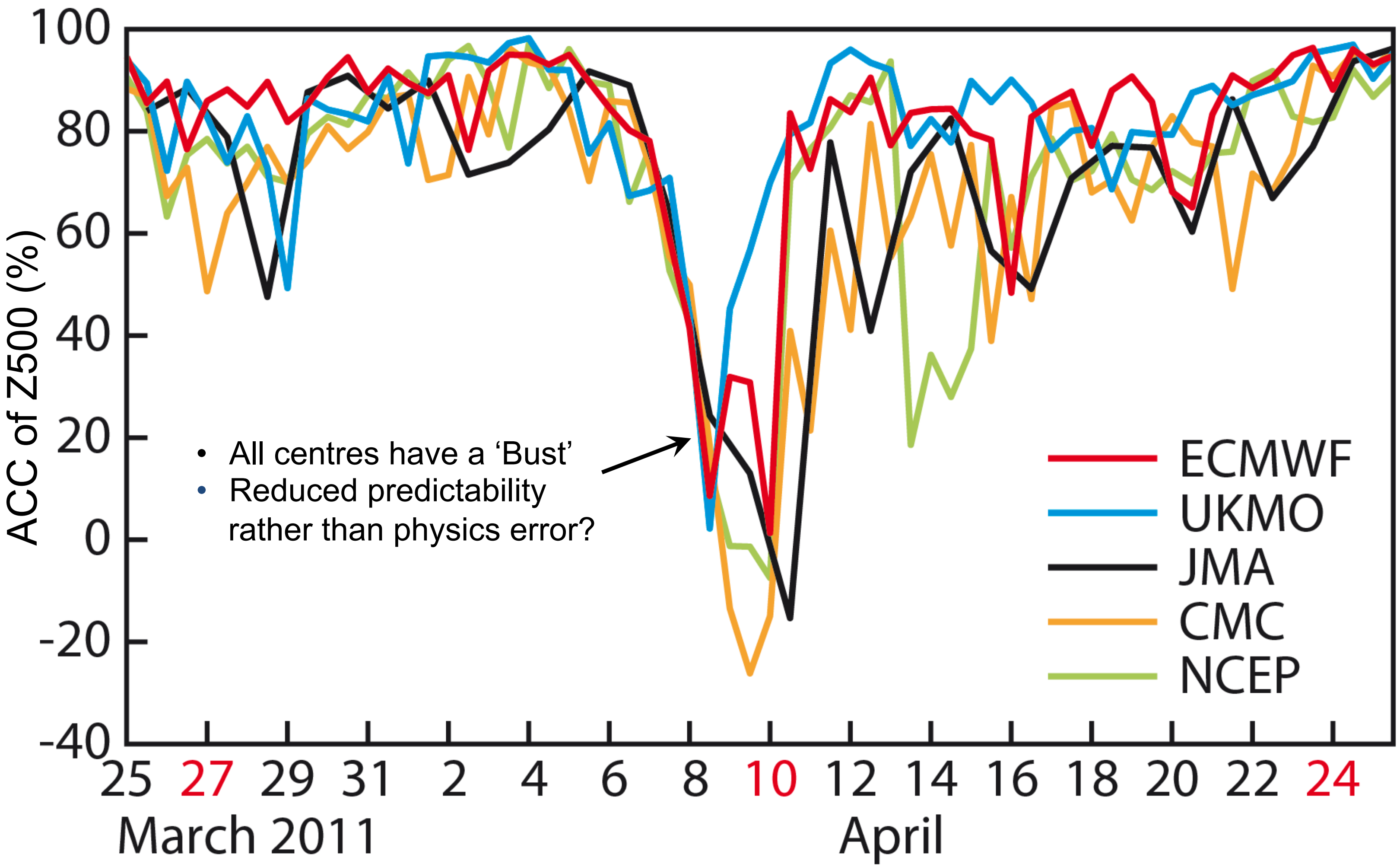
Observation – forecast (in period 9-21UTC)



Some satellite radiances can be equated to temperatures or humidities at a rough height.
For these, it is easy to relate the observation departures to the analysis increments

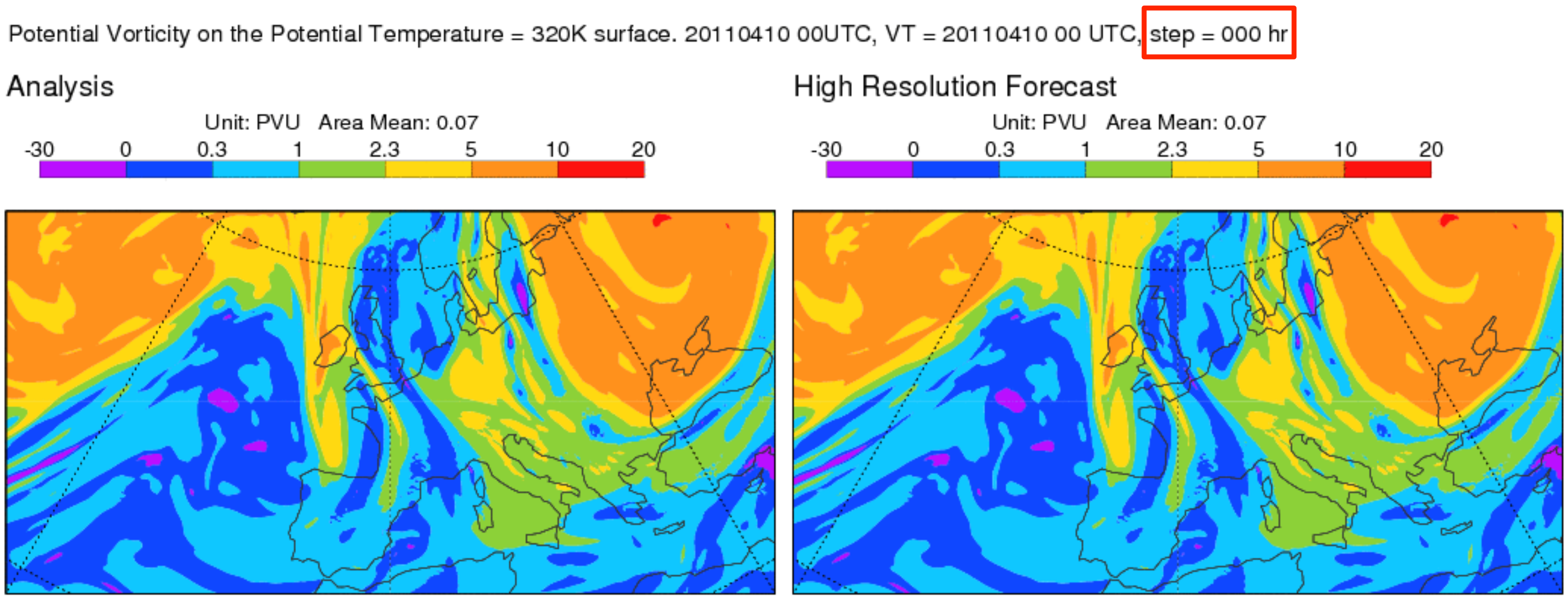
Observation minus first guess for selected observation types / channels over analysis window 9-21UTC on 2012/07/21

Skill of single forecasts (Europe, leadtime = 6 days)

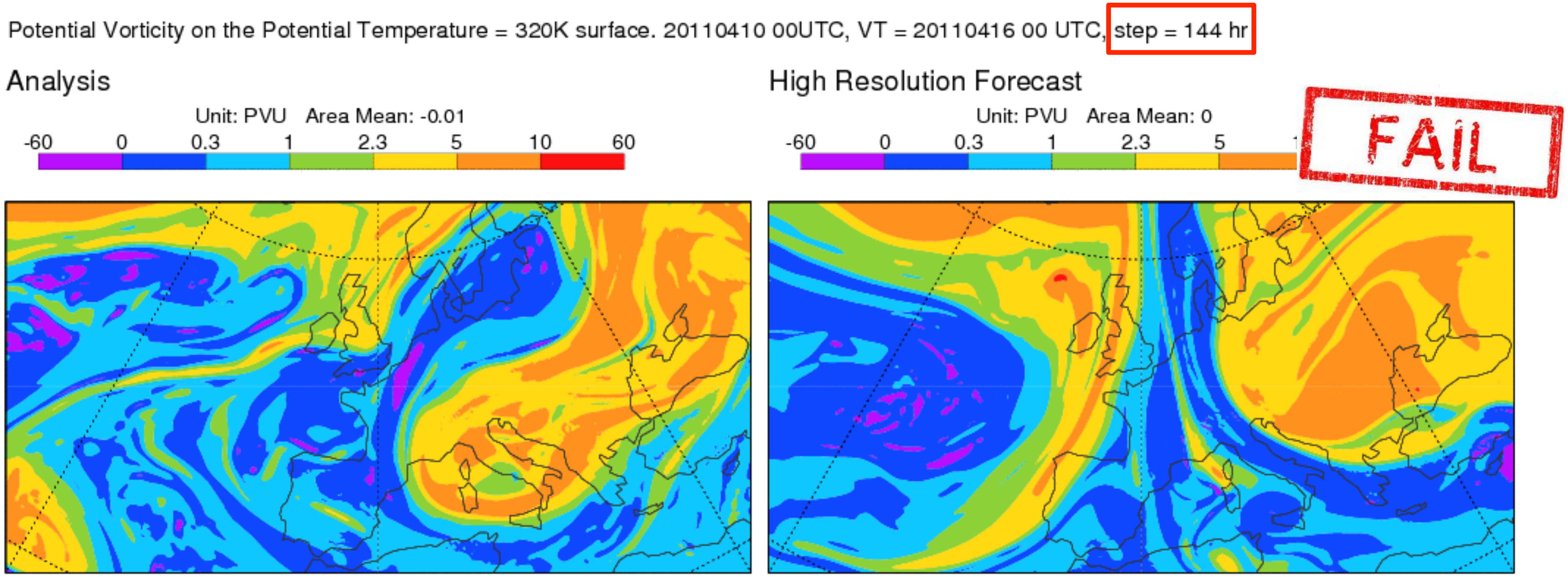


Score is the spatial Anomaly Correlation Coefficient (ACC)x100 for 500 hPa geopotential height (Z500) over Europe (12.5°W –42.5°E, 35°N–75°N). The date shown is the forecast start date

Deterministic forecasting (initial conditions)



Deterministic forecasting (flow evolution to day-6)

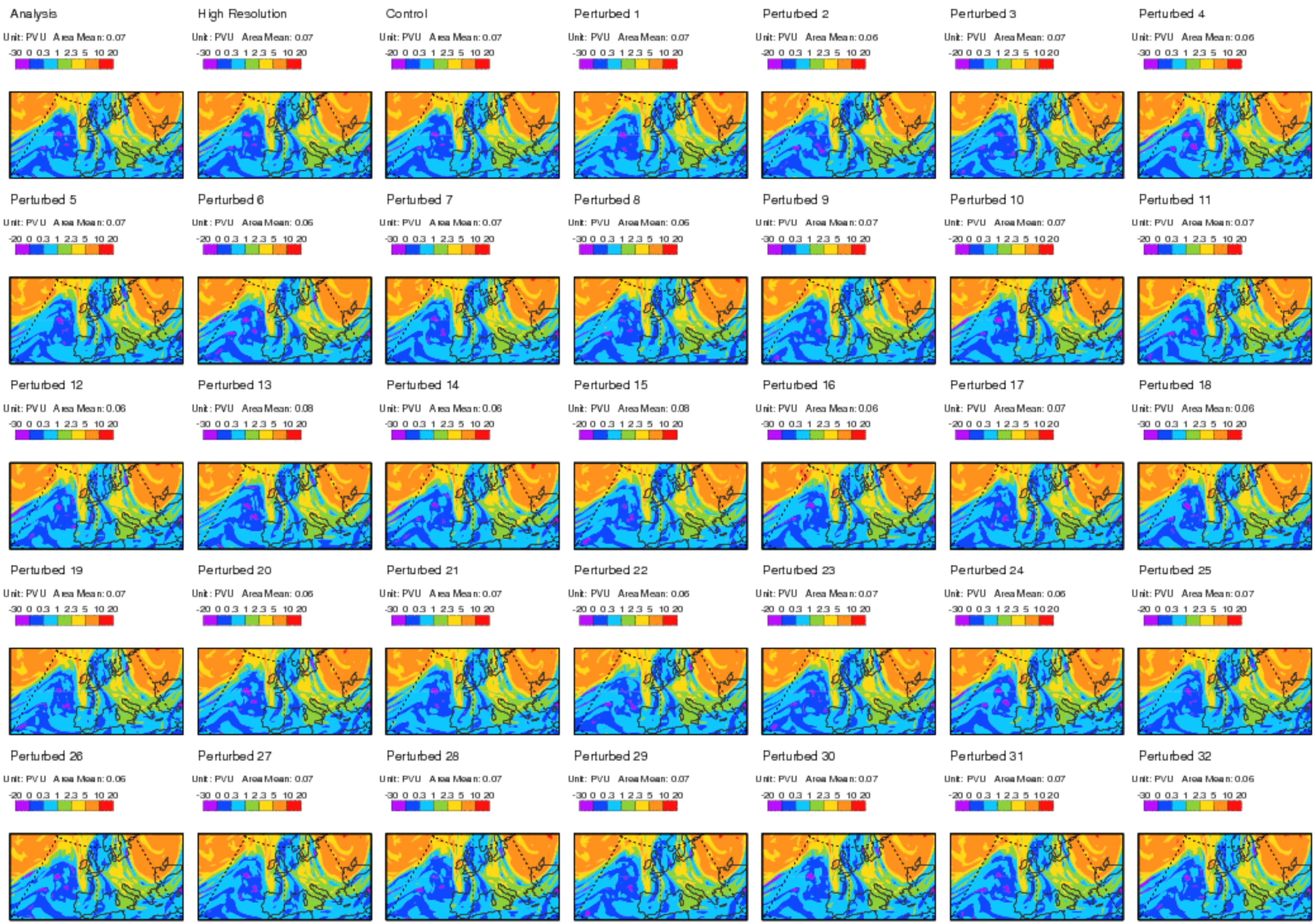


It is difficult, by day-6, to disentangle model error from the natural growth of initial condition uncertainty (chaos)

Diagnostics

Ensemble forecasting (initial conditions)

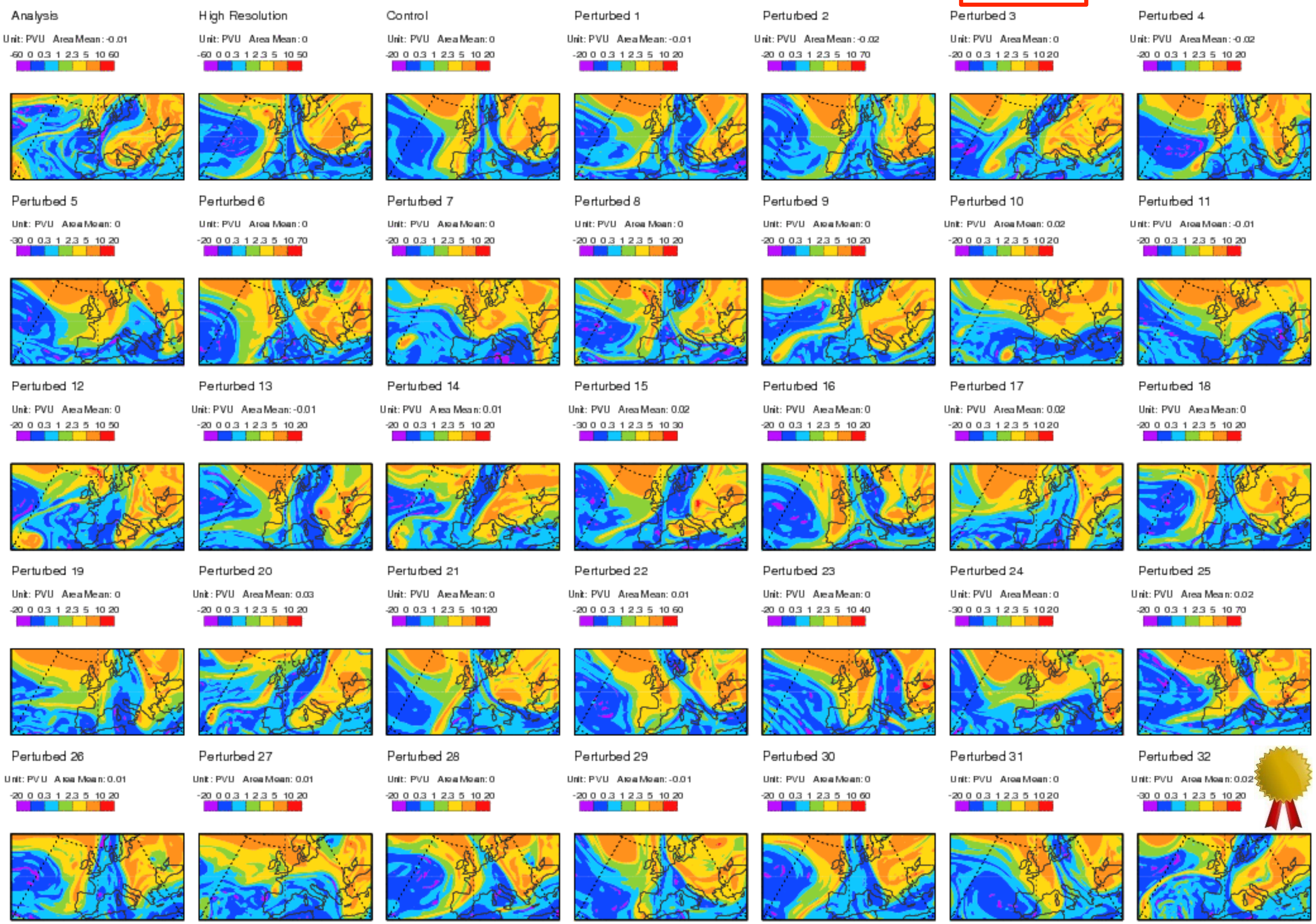
Potential Vorticity on the Potential Temperature = 320K surface. 20110410 00UTC, VT = 20110410 00 UTC, step = 000 hr



Diagnostics

Ensemble forecasting (flow evolution to day-6)

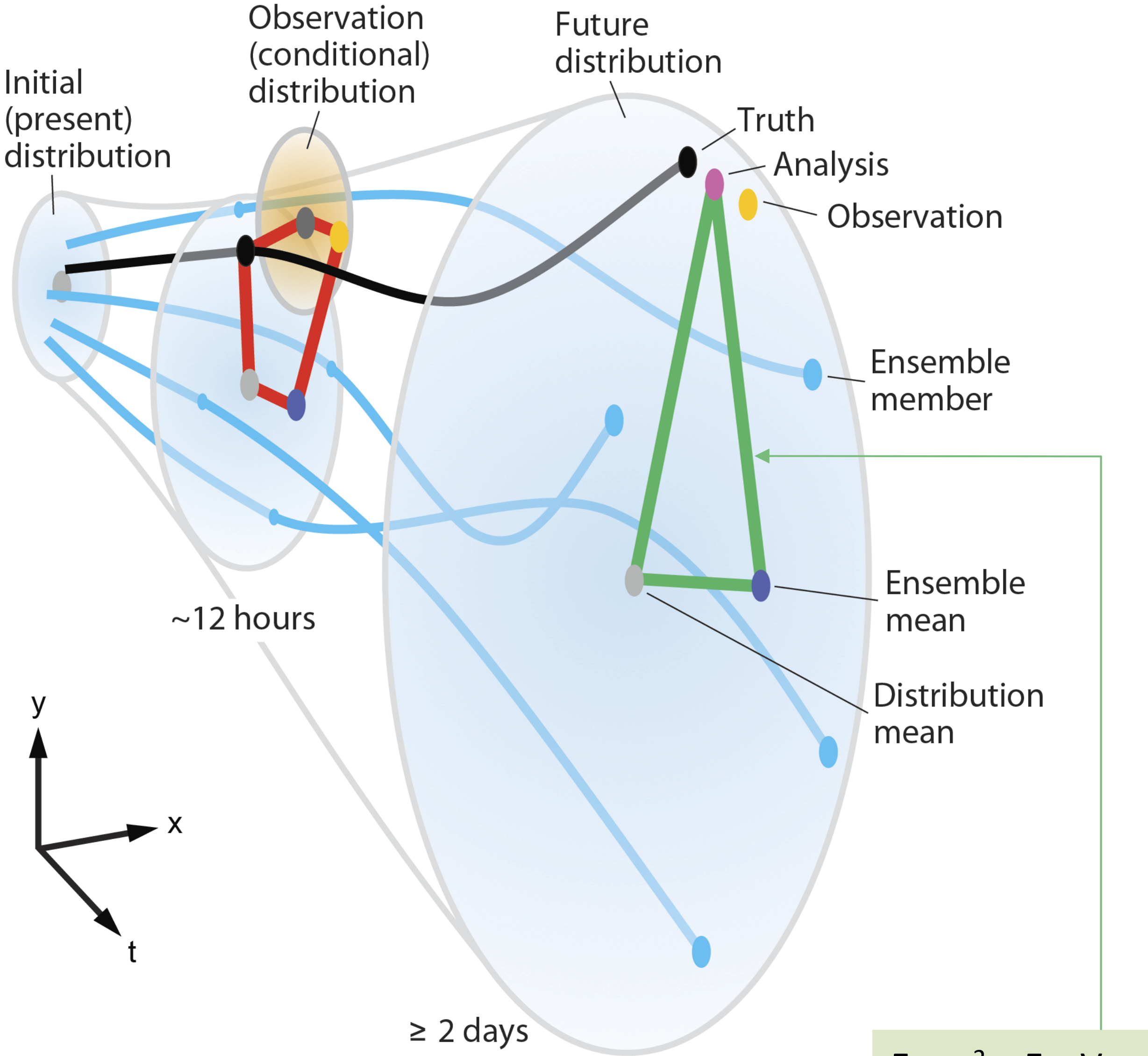
Potential Vorticity on the Potential Temperature = 320K surface. 20110410 00UTC, VT = 20110416 00 UTC, step = 144 hr



Reliability in ensemble forecasting

The importance of reliability is the motivation for using ‘proper’ scores (such as the Brier Score or CRPS).

Reliability (at all leadtimes) should reduce ‘jumpiness’ of ensemble forecasts



Error^2 = EnsVar + Residual

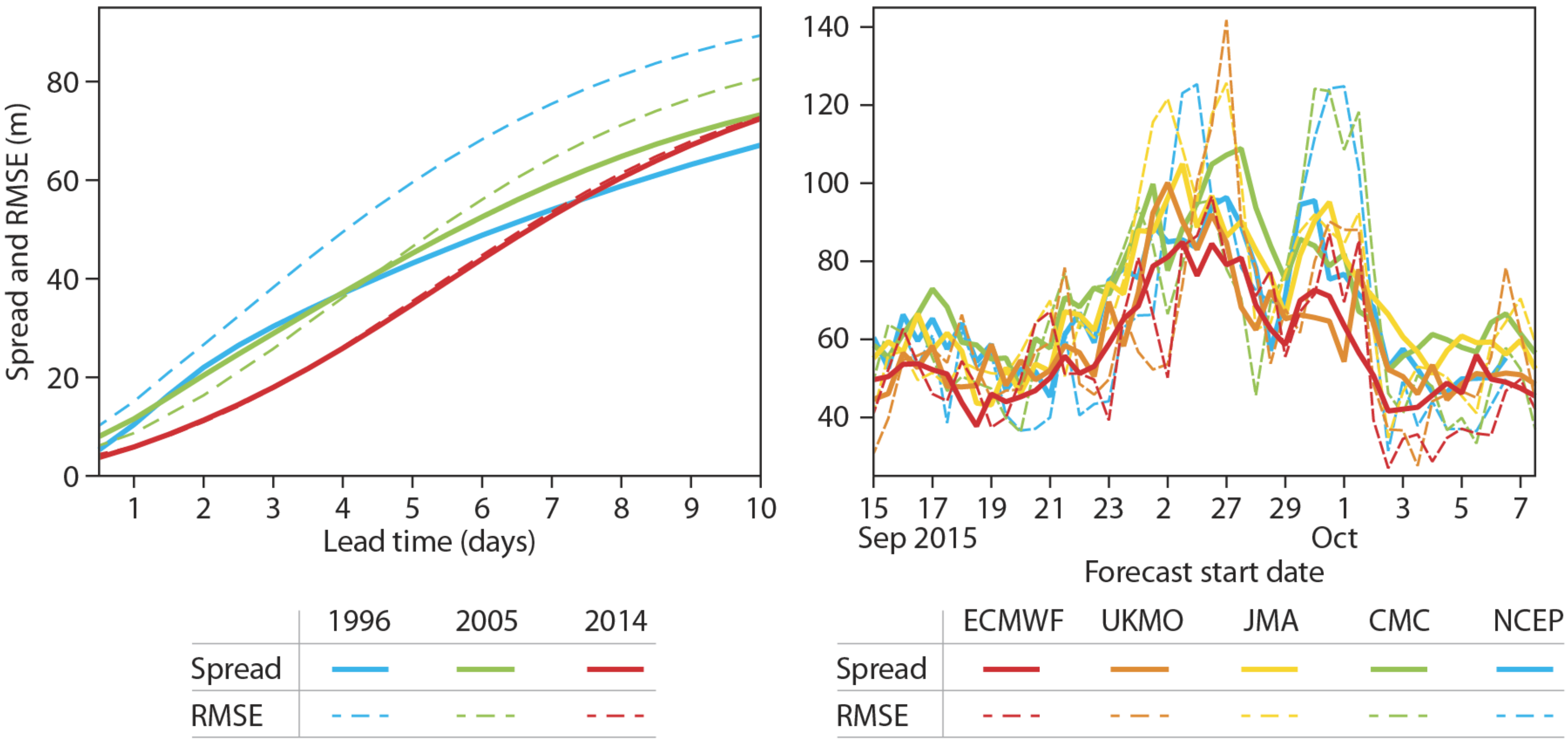
(Cross-terms on squaring have zero expectation. EnsVar is scaled variance to account for finite ensemble-size)

Adapted from Rodwell et al. 2015

Ensemble spread and error

500 hPa geopotential height (Z500)

Northern Hemisphere, annual mean



Substantial improvements in sharpness and reliability. Due to:

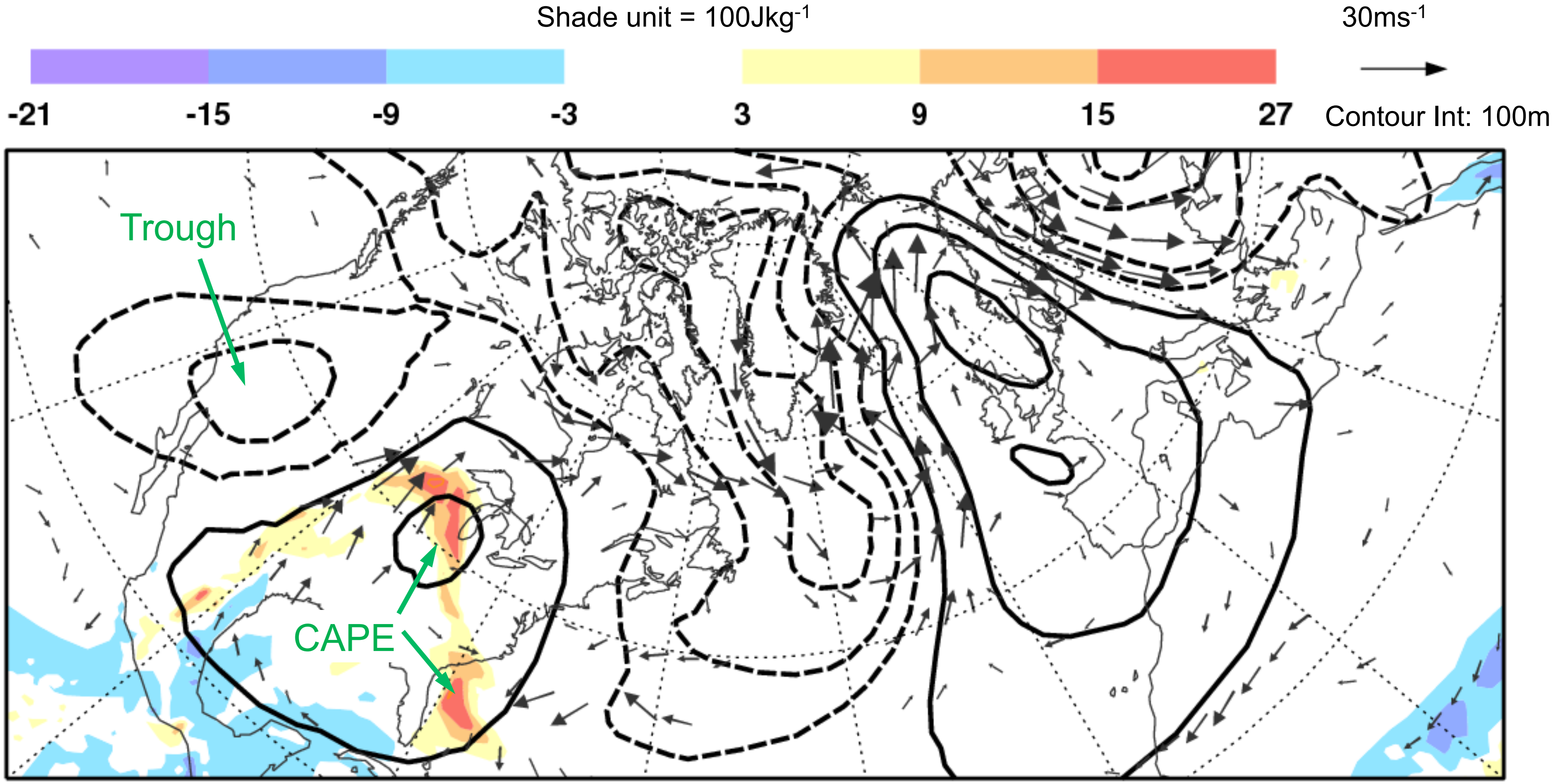
- Ensemble of data assimilations
- Stochastic physics
- Observations and modelling of observation error

- Spread agreement between centres indicates flow-dependent fluctuations in underlying predictability (reason for ensemble forecasting!)
- Need to assess flow-dependent reliability

RMSE is of ensemble-mean error. Spread = ensemble standard deviation (scaled to take account of finite ensemble size)

Initial conditions in 'bust' forecast. 10 April 2011

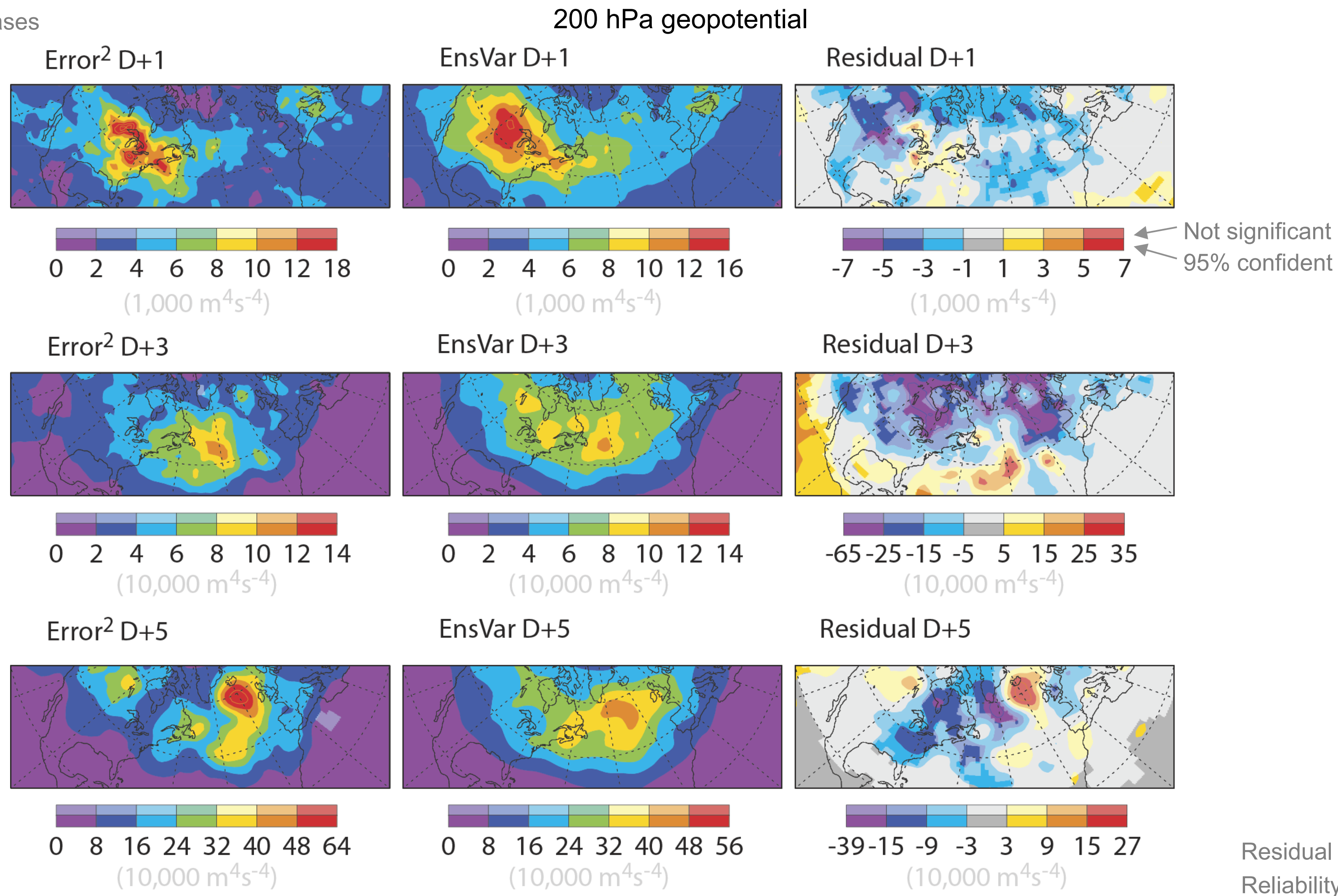
Z500, CAPE and v850 anomalies



The key features here are the trough over the Rockies, ahead of which is warm moist southerly flow and high Convective Available Potential Energy (CAPE), conducive to formation of mesoscale convective systems.

Spread-error for Trough/CAPE composite (⇒MCS)

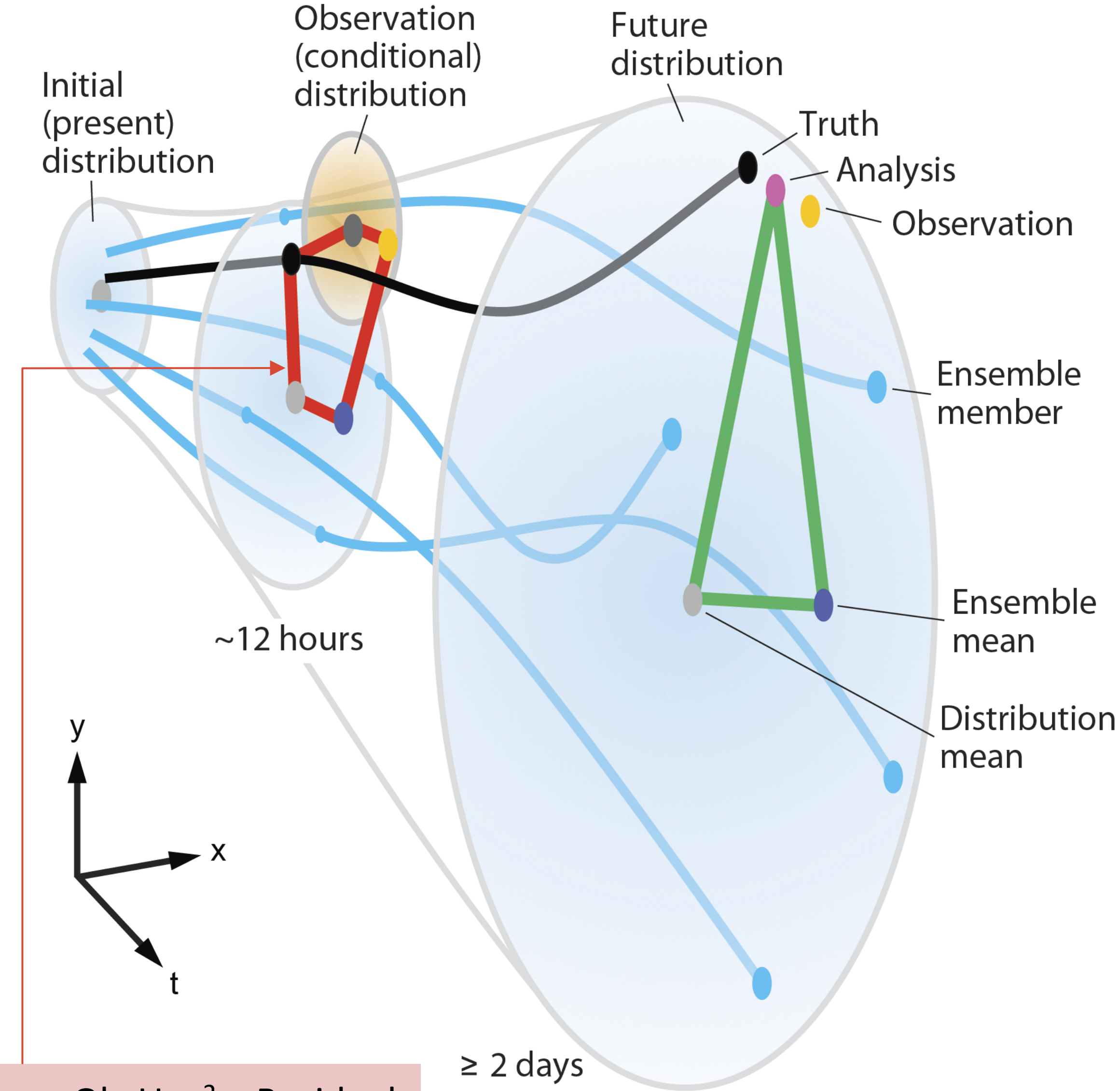
54 cases



Residual = Error^2 – EnsVar
Reliability ⇒ $\mathbb{E}$[Residual]=0

- Following conditions conducive to MCS development, enhanced errors and spread propagate east towards Europe → 'Busts' ✓
- Note: -ve residuals occur in non-trough/CAPE situation too.
- +ve residual at D+5 is not significant (Chaos? → Use bigger sample or shorter leadtime? Analysis uncertainty at D+1?) ✘

Reliability in ensemble data assimilation



$\text{Depar}^2 = \text{Bias}^2 + \text{EnsVar} + \text{ObsUnc}^2 + \text{Residual}$

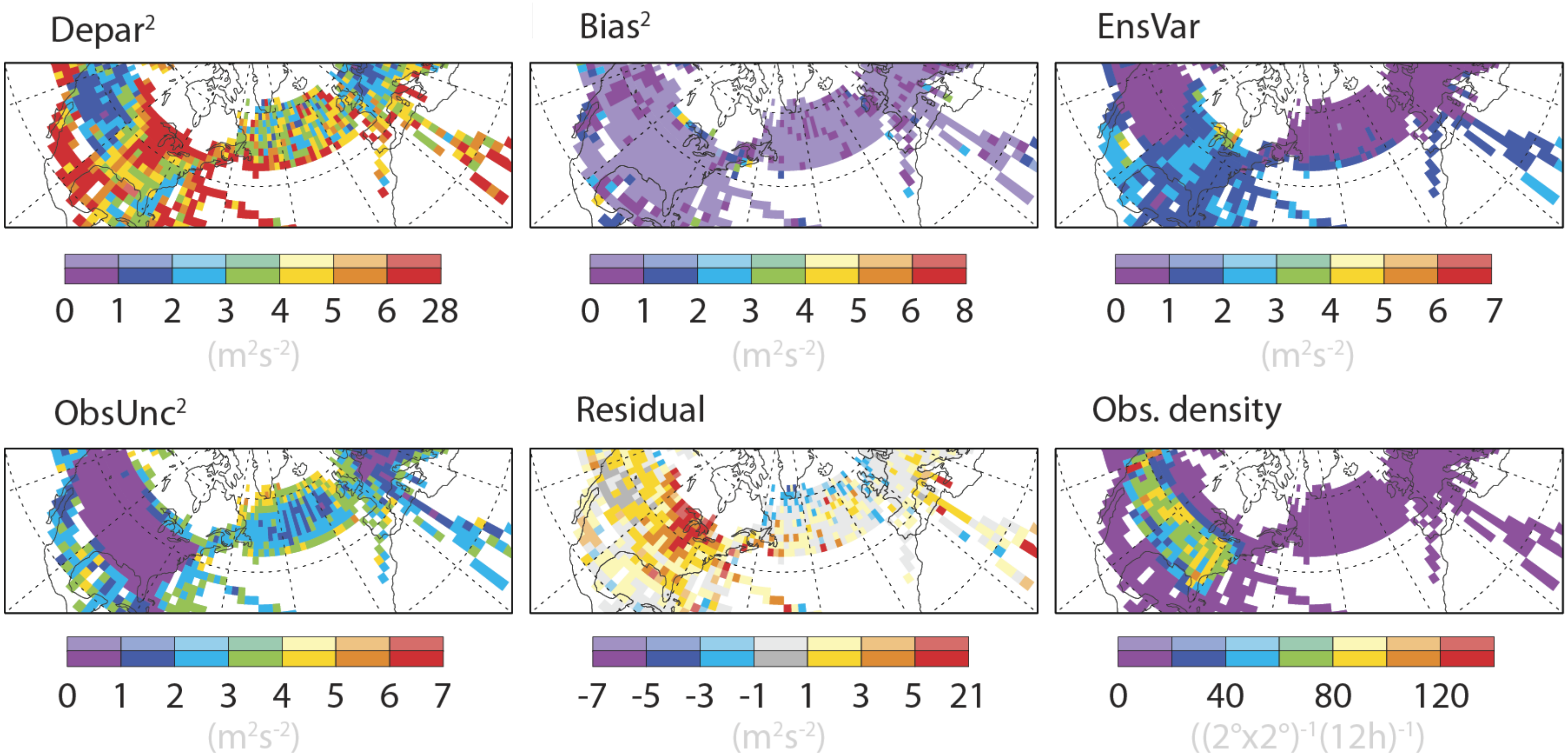
(Cross-terms on squaring have zero expectation. EnsVar is scaled variance to account for finite ensemble-size)

Adapted from Rodwell et al. 2015

EDA reliability budget: Trough/CAPE comp.

54 cases

Relative to aircraft observations of zonal wind 200hPa (±15)



- Residual highlights MCS, and suggests lack of background variance. (Obs uncertainty changes 2nd-order) ✓
- MCS uncertainty (existence, intensity, location) not well reflected in Jetstream uncertainty (with downstream consequences)
- Budget useful to diagnose biases, modelling of observation error and representation of model uncertainty (including stochastically-formulated process parametrizations)

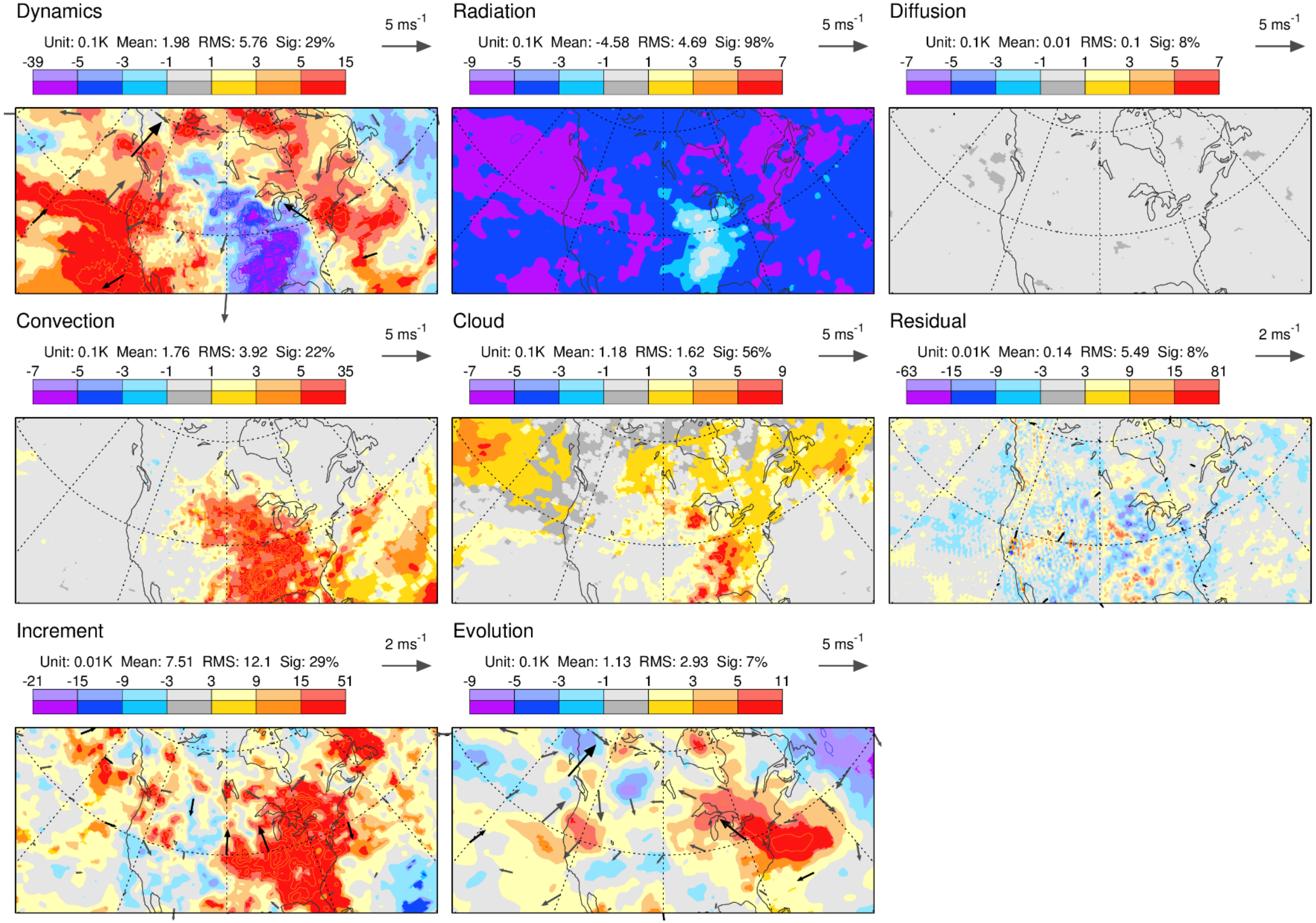
$Depar^2 = Bias^2 + EnsVar + ObsUnc^2 + Residual$

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

EDA = “Ensemble of Data Assimilations”

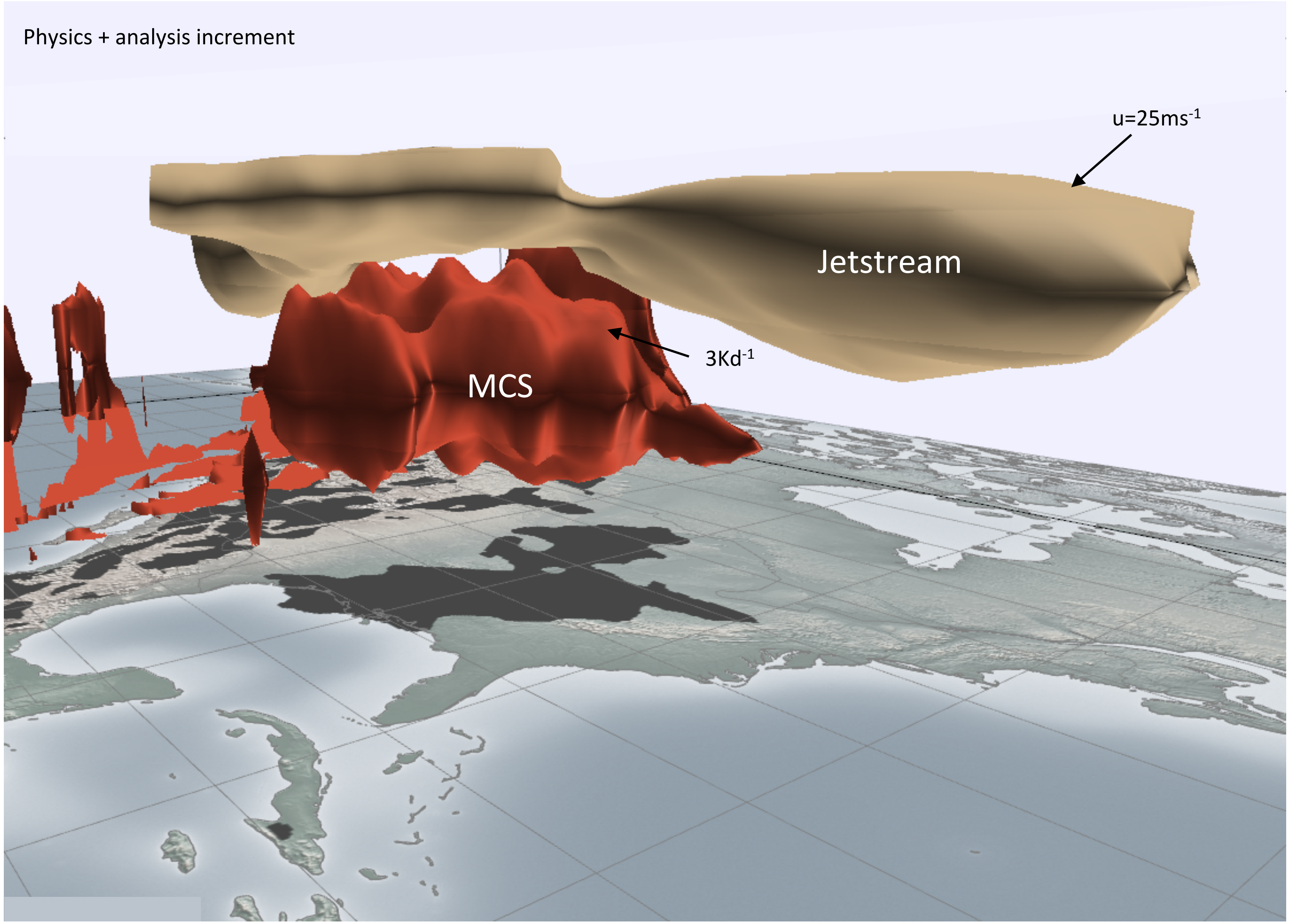
Single forecast tendencies: Trough/CAPE comp.

54 cases T300 process tendencies accumulated over 12hr background, the analysis increment, and evolution of the flow



- 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

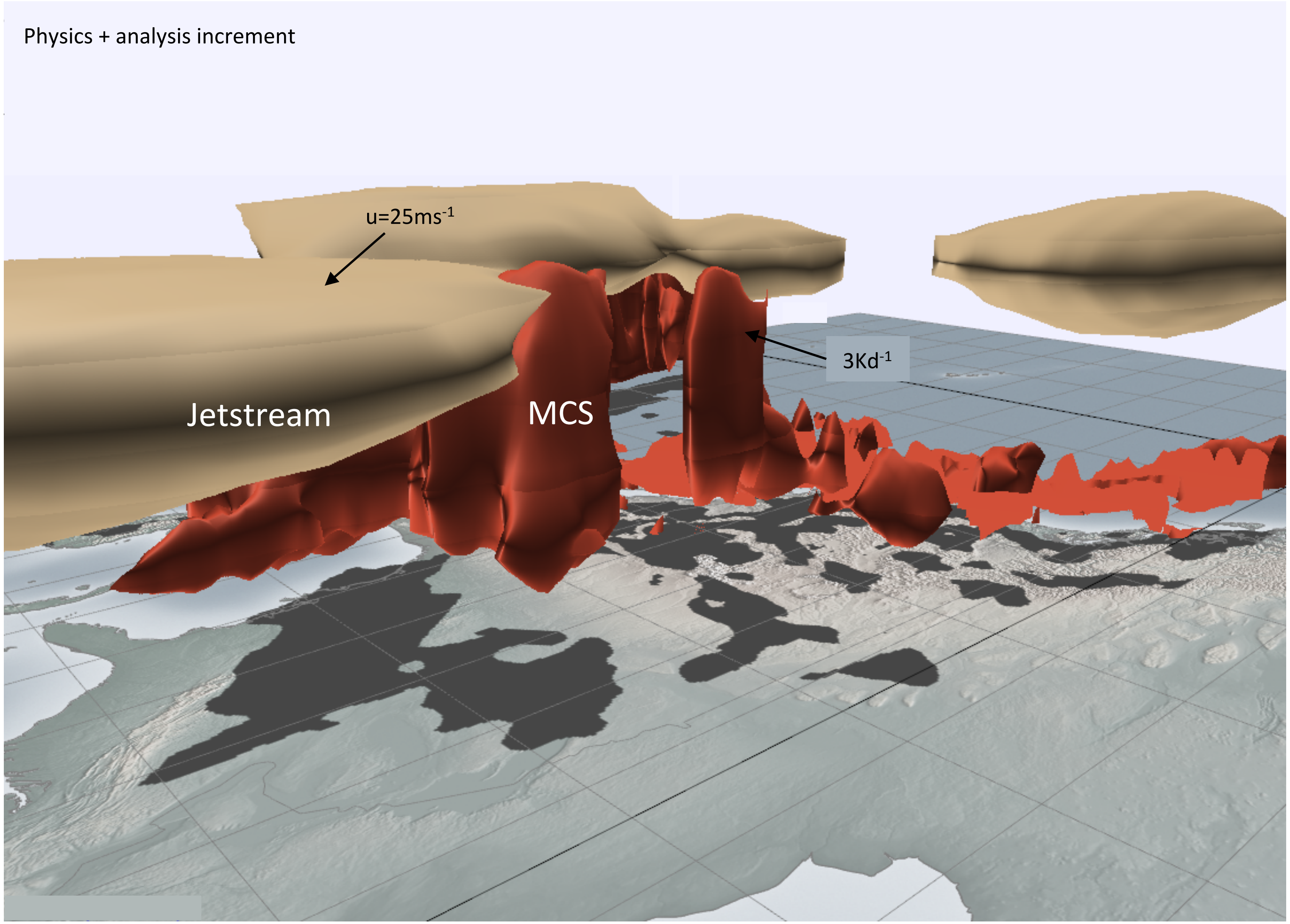
MCS – Jetstream interaction (composite)



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

Met3D: Marc Rautenhaus

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?

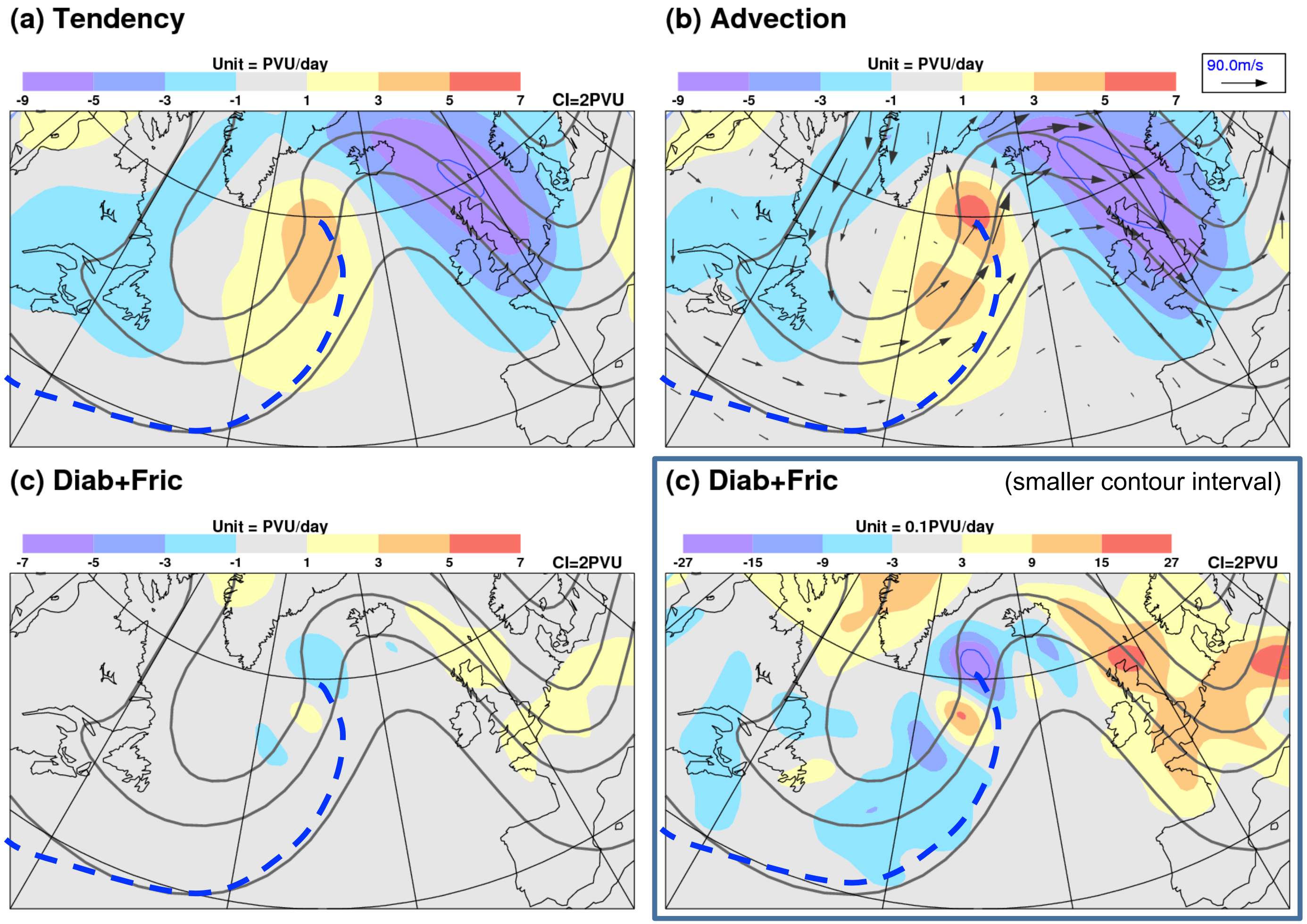
Met3D: Marc Rautenhaus

Summary

- Aim: flow-dependent diagnostics of mean-error & variance-error
- Short leadtimes essential to focus investigation on specific features & to permit smaller samples
- Improvements should be seen throughout the medium-range forecast

PV budget on θ=330K at 12UTC on 2012/07/21

~400 hPa

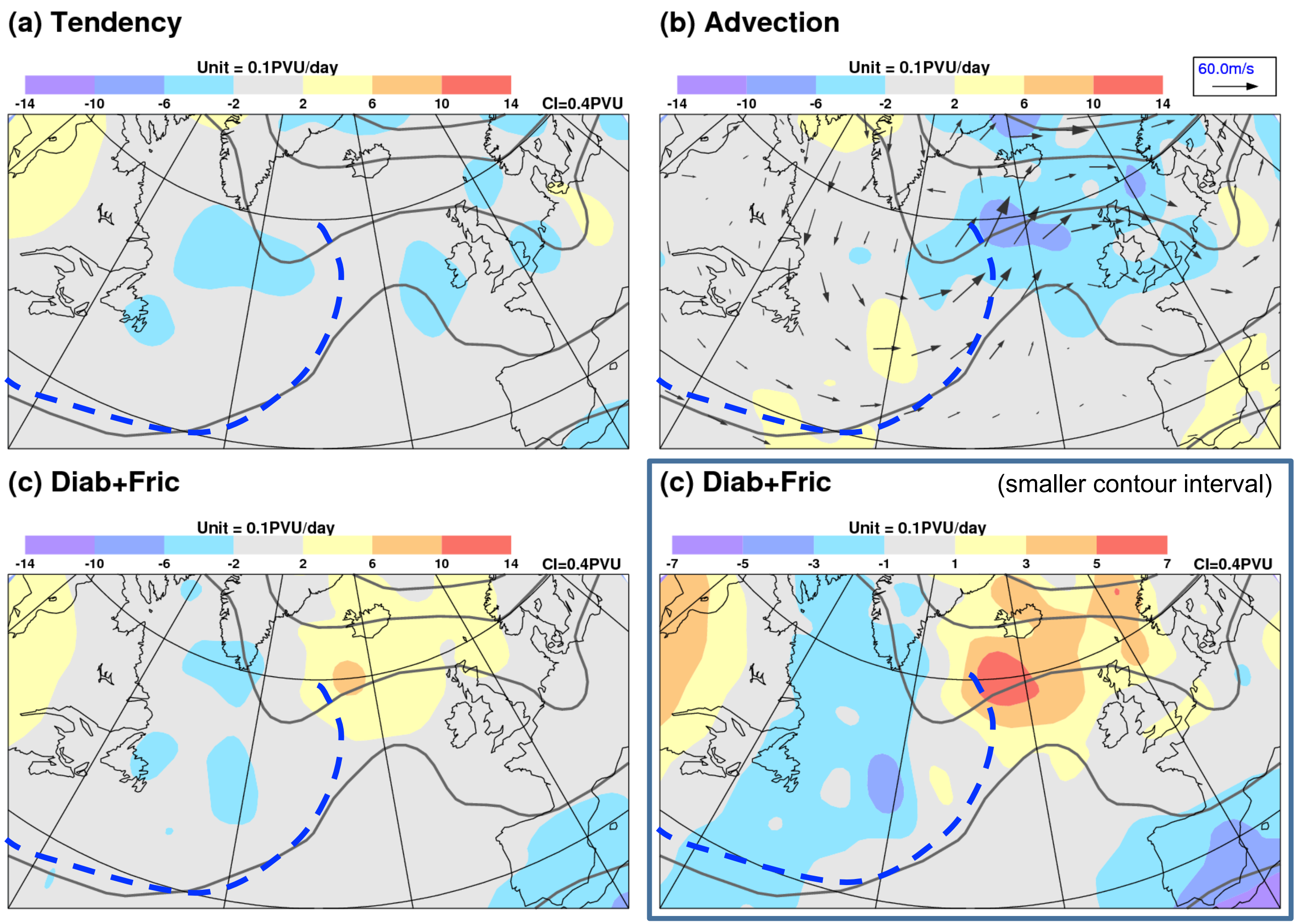


Advection moves +ve upper-level PV anomaly to the east.
Convection and cloud heating below act to destroy this along the front.

Interested in link to large-scale waves, and so terms are filtered to the scale of streamfunction anomalies

PV budget on θ=315K at 12UTC on 2012/07/21

~600 hPa



Diabatic effects strongly oppose PV advection at cloud head (at about 600hPa) (mainly due to changes to stratification below maximum in cloud heating?)

Interested in link to large-scale waves, and so terms are filtered to the scale of streamfunction anomalies