Diagnostics at ECMWF

Mark Rodwell

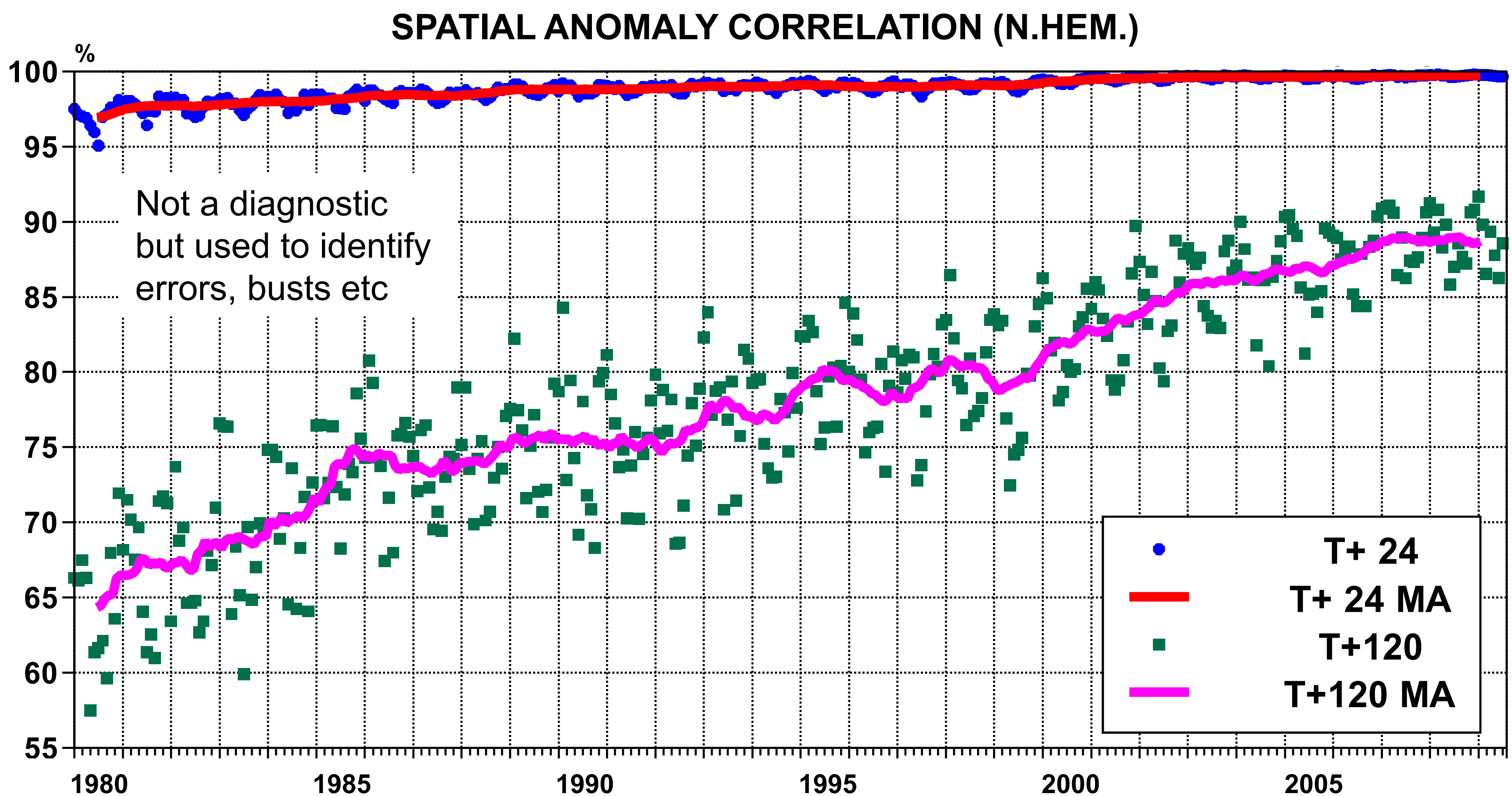
7 September 2009

Work with Thomas Jung

Outline

- Scores, Metrics and Diagnostics
- Diagnosis: The Changing Task
- The "Diagnostics Explorer"
- Using Analysis Increments & Initial Tendencies
- Scale-Dependent Verification
- Diagnostic Verification: Precipitation

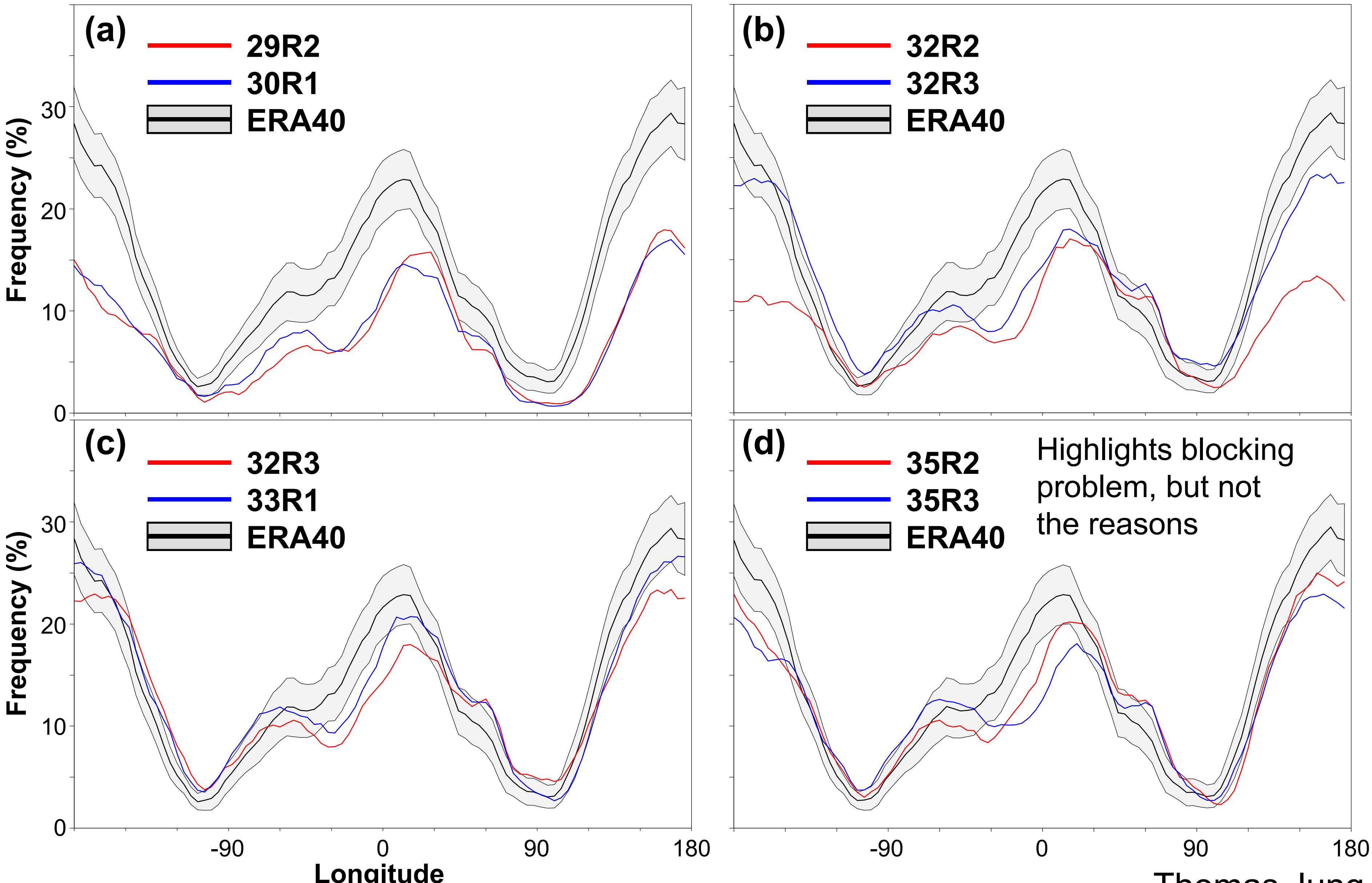
Scores: Verification of 500 hPa Geopotential



David Richardson

Based on ECMWF operational forecasts from 12 UTC analysis. MA=12 month moving average

Metrics: Blocking Frequency. DJF 1963-2006

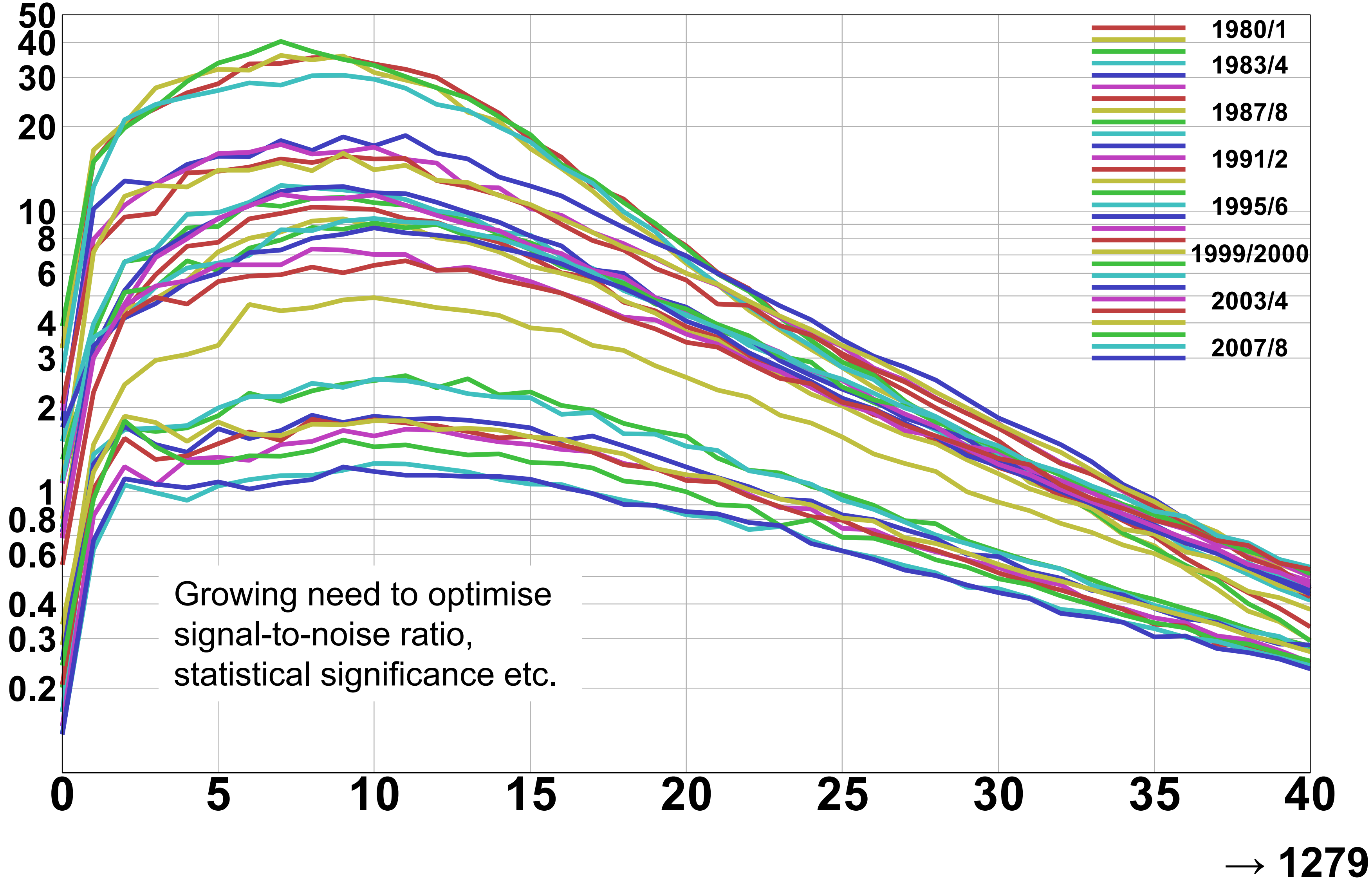


Highlights blocking problem, but not the reasons

Thomas Jung

Diagnosis: The Changing Task

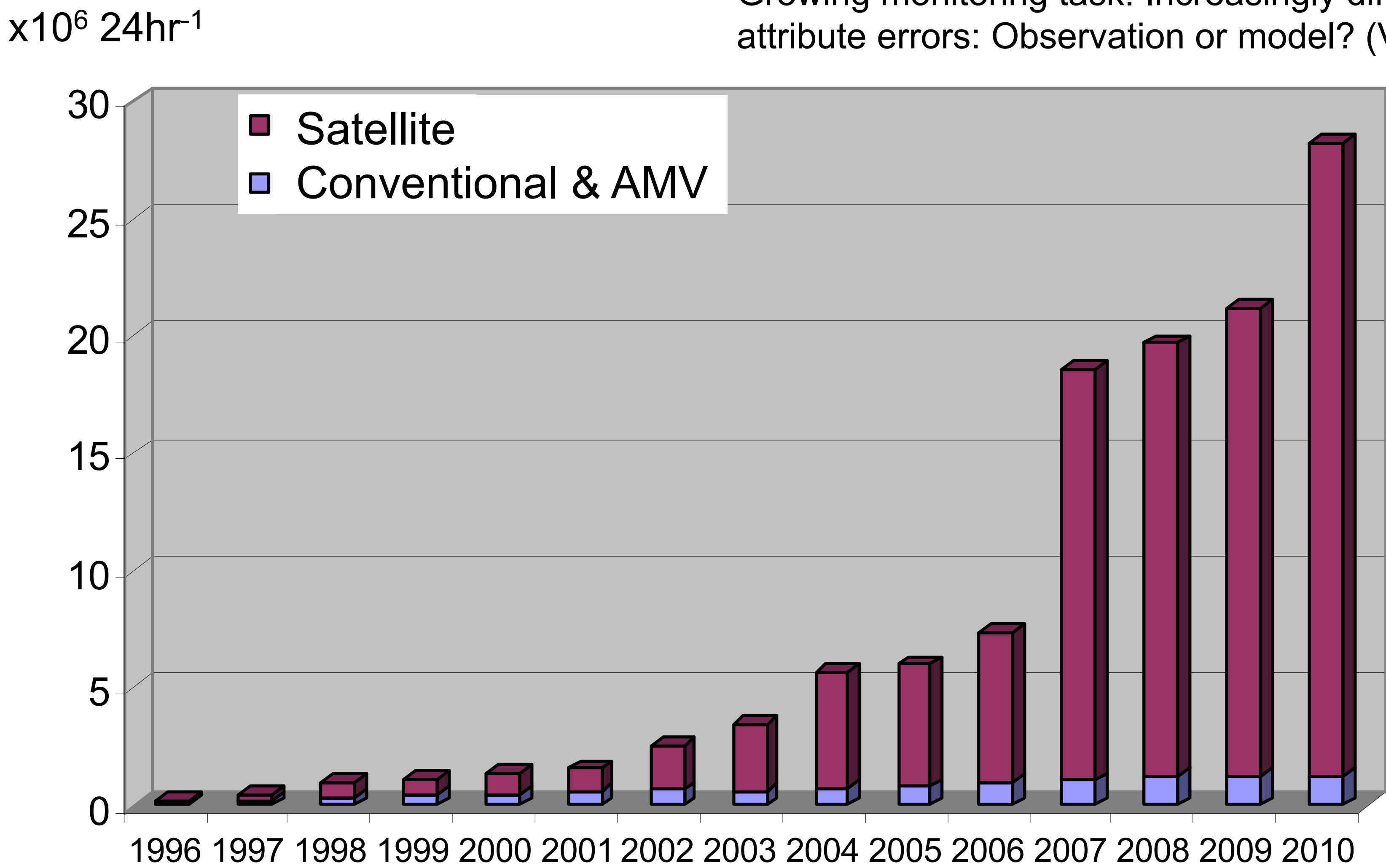
Temporal Variance of D+1 Error of Z500



Adrian Simmons (2006) - updated

Observational Data Volumes

Growing monitoring task. Increasingly difficult to attribute errors: Observation or model? (VAR-BC)



Peter Bauer and Jean-Noel Thepaut

Unit is millions of data values assimilated per 24 hour period

Diagnostics of the Diagnostics group

On-line Diagnostics: A 5D view of the IFS

IFS Component	Diagnostics
Data Assimilation	***Observation space – observation usage*** ● Many data sources including radiosonde and satellite ● Data count, first-guess departures (mean, rms), bias corrections
	Model space – analysis increments ● Prognostic and other parameters ● Mean, standard deviation, rms ● 21 pressure levels and zonal means
Weather forecast	***Forecast error*** ● Prognostic and other parameters ● Mean, standard deviation, rms ● 21 pressure levels and zonal means
	Scale-dependent error and activity ● Several parameters, levels and regions ● All spatial scales and selected spatial scales
Climate of atmospheric model and coupled model	***Seasonal-means of error*** ● Several diagnostics including geopotential height, winds, velocity potential, Hadley and Walker circulations, ocean waves
	Seasonal-means of variability ● Blocking ● ENSO teleconnections ● Empirical Orthogonal Functions ● Planetary and synoptic activity ● Power spectra ● Tropical waves (including Madden-Julian Oscillation)

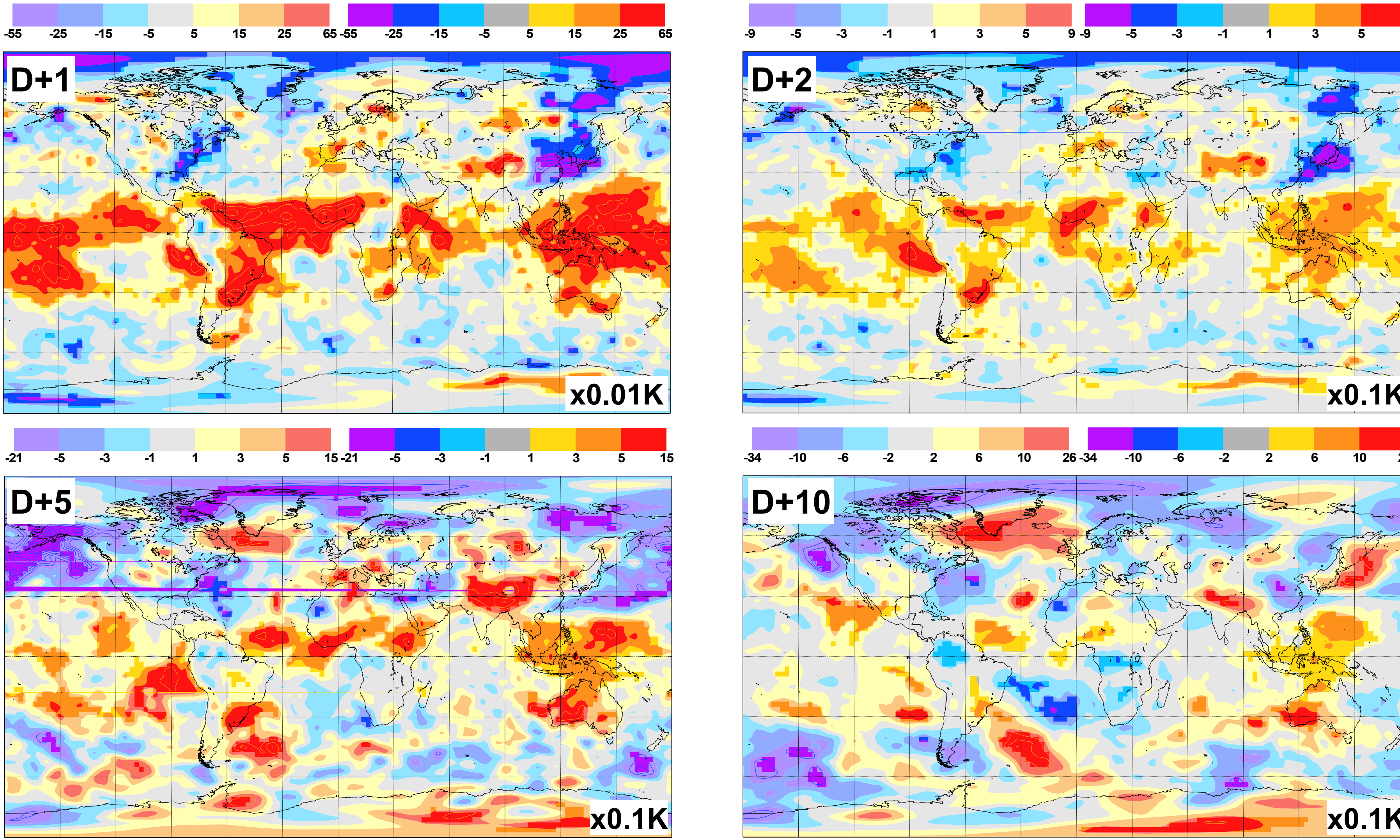
- All diagnostics are produced for operational forecasts (seasonal means) and "E-suites".
- Some diagnostics are produced for research experiments.
- "Initial Tendency" diagnostics will be added.
- **Aim:** Seamless and efficient diagnosis of entire forecasting and data assimilation system.
- Other sections produce more detailed diagnostics for their particular IFS component.

Analysis Increments and Initial Tendencies

M.J. Rodwell

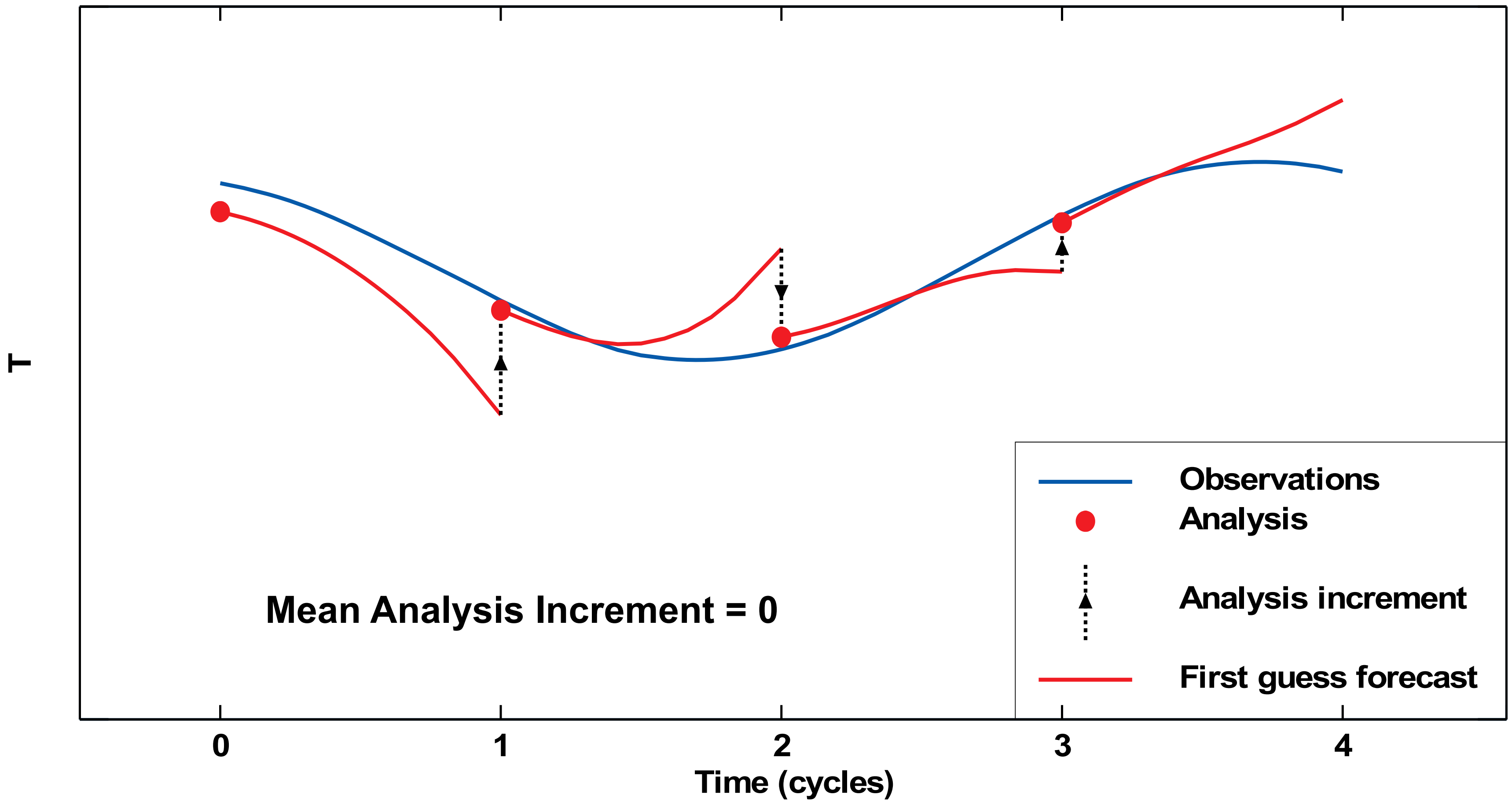
T500 Forecast Error as function of lead-time

Shorter lead-times localise error and increase signal-to-noise ratio



Based on DJF 2007/8 operational analyses and forecasts. Significant values (5% level) in deep colours.

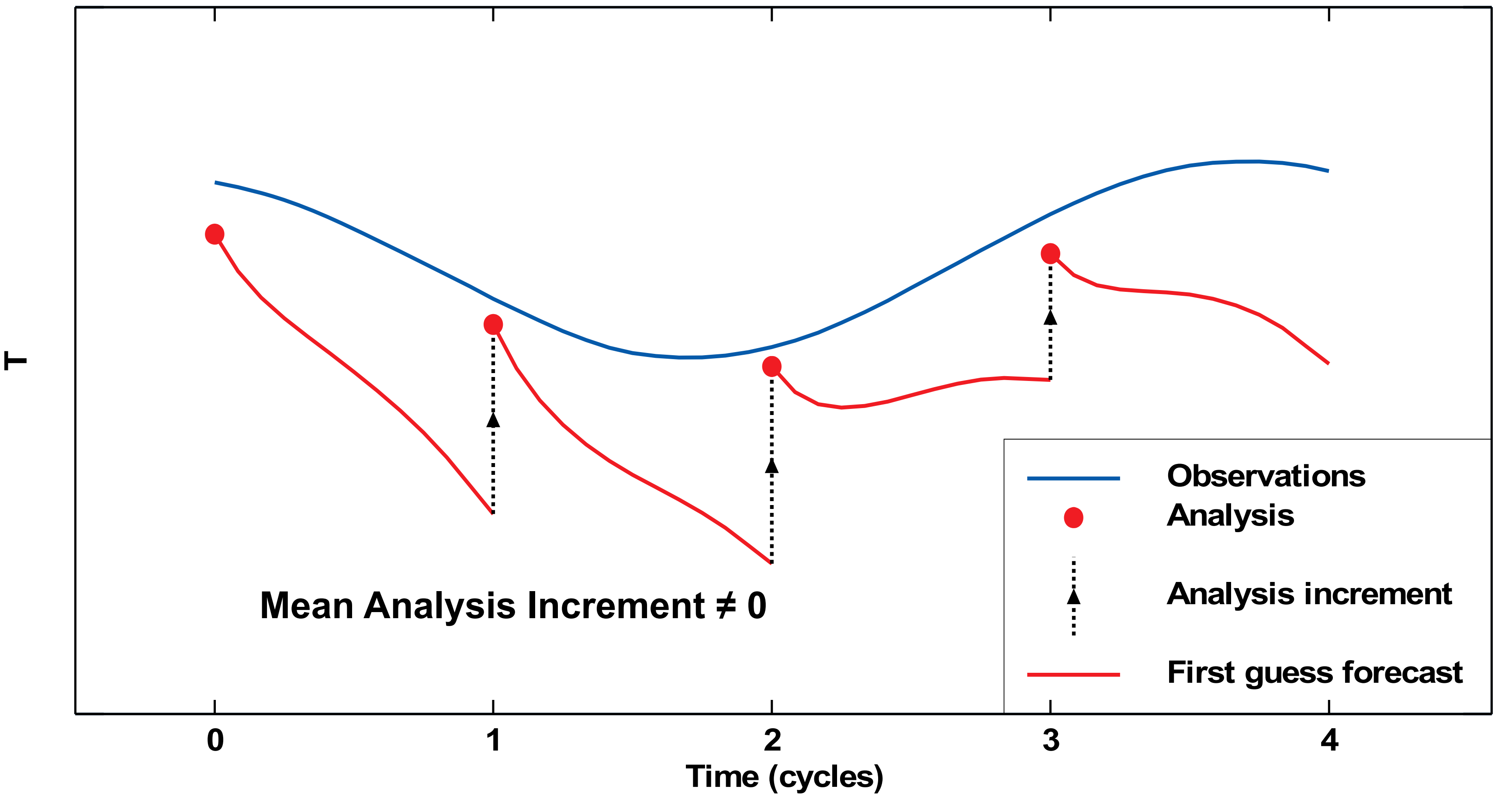
Data Assimilation Cycle: Perfect Model



For a perfect model, positive and negative increments tend to cancel

(Imperfect, unbiased observations)

Data Assimilation Cycle: Imperfect Model



−Mean Analysis Increment = Mean Net Initial Tendency (in appropriate units)

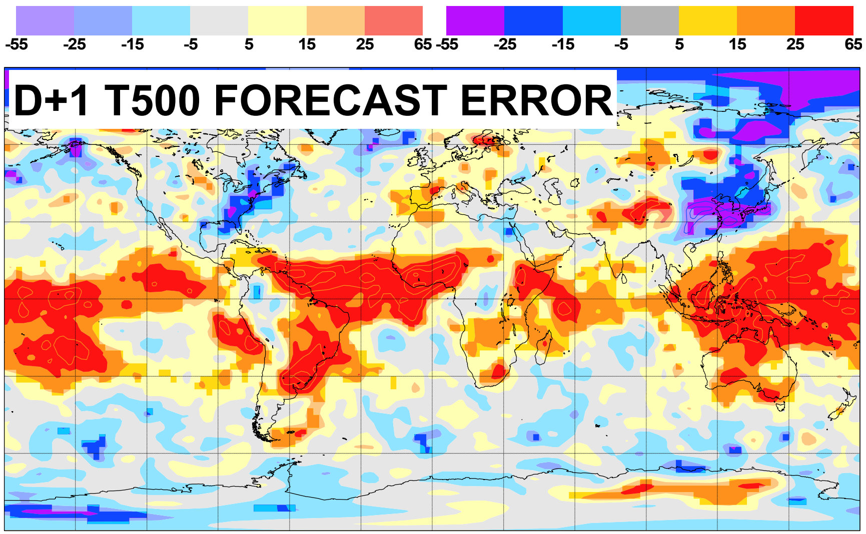
= Convective + Radiative + … + Dynamical mean tendencies

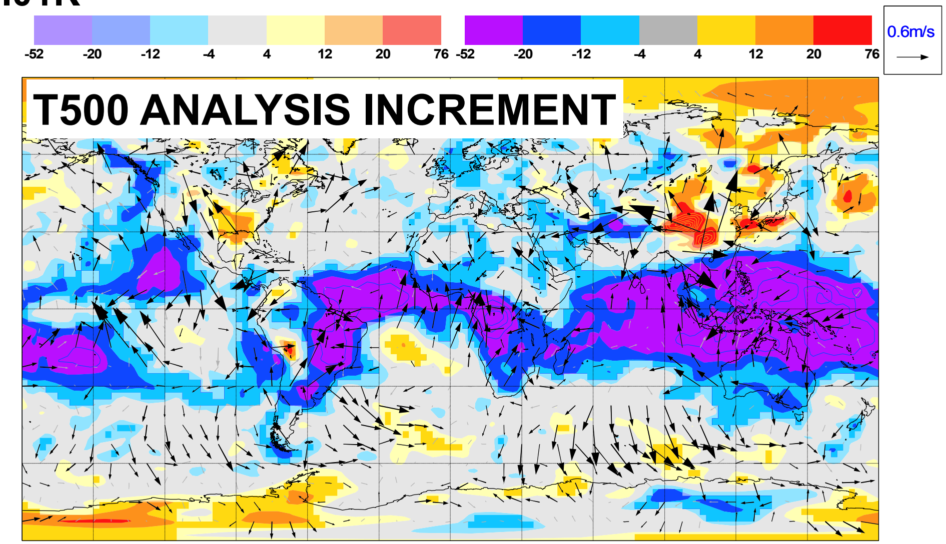
(summed over all processes in the model)

The use of "Initial Tendencies" was first proposed by Klinker and Sardeshmukh (1992)

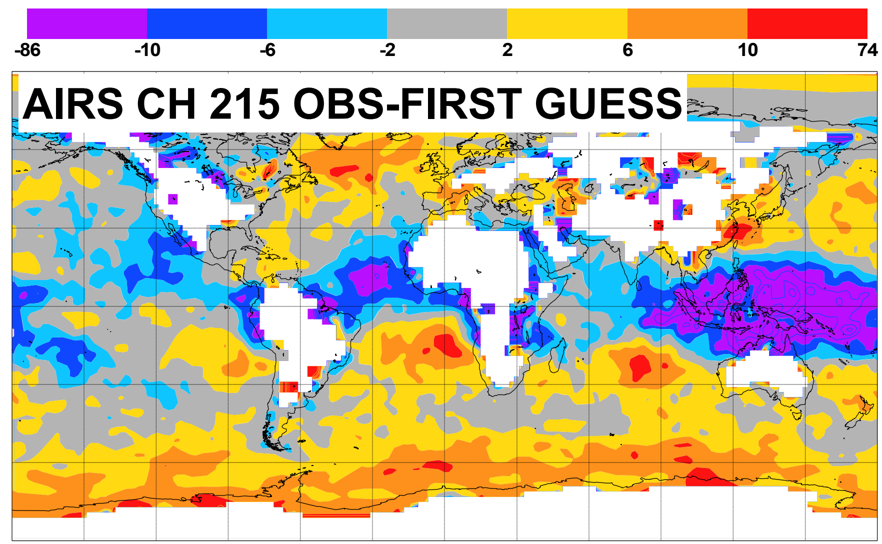
Confronting Models with Observations

UNIT=0.01K





- Every 1° square has data every cycle
  - ~6 Million data values
- Independent vertical modes of information:
  - IASI / AIRS: ~ 15
  - HIRS / AMSUA: ~ 5 (~ 2 IN TROP)
- Anchors (not bias corrected):
  - Radiosonde
  - AMSUA-14
  - Radio Occultation

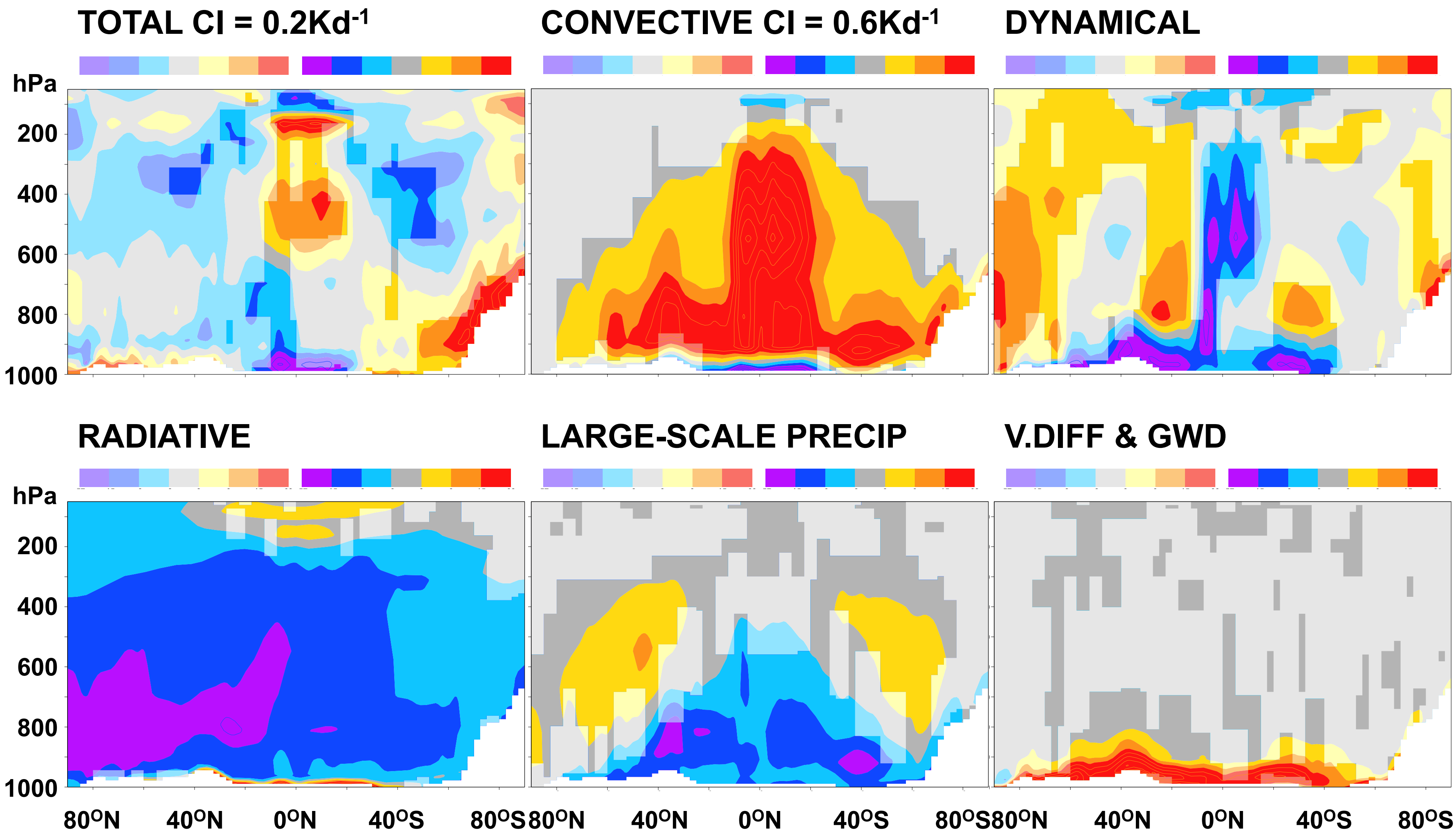


Based on DJF 2007/8 operational analyses and forecasts. Significant values (5% level) in deep colours.
AIRS CH 215 BRIGHTNESS TEMPERATURE ~T500

Zonal Mean Errors

Mean Temperature Tendencies

Total tendency should be zero for a perfect model. Why isn't it? Extratropics too!

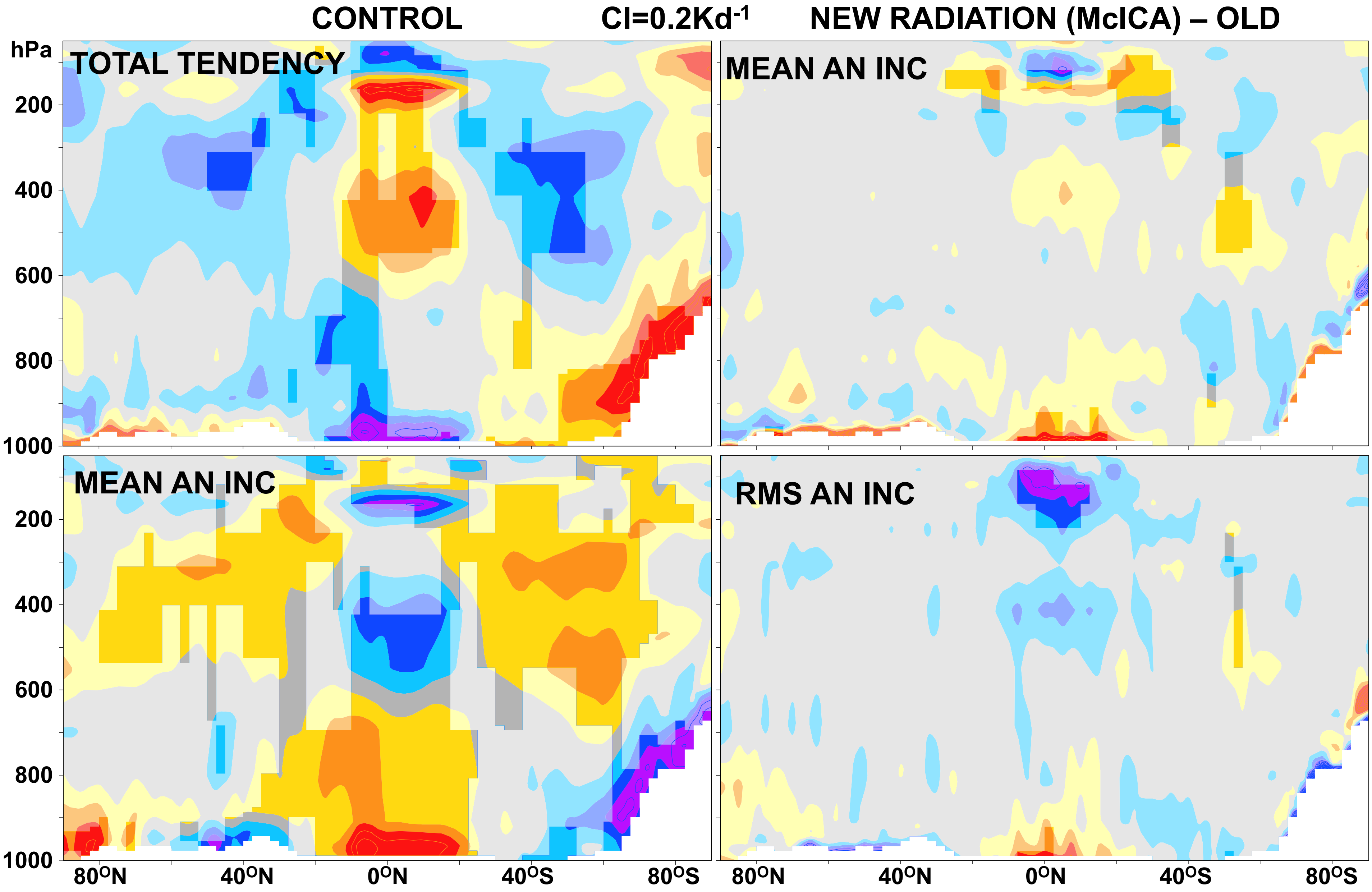


Mean tendencies are deduced on model levels. Y-axis shows approximate pressure value. Average is over December 2008, 4 forecasts per day, tendencies accumulated from T+1 to T+7. Model cycle 33R1, T_L159, L91

M.J. Rodwell

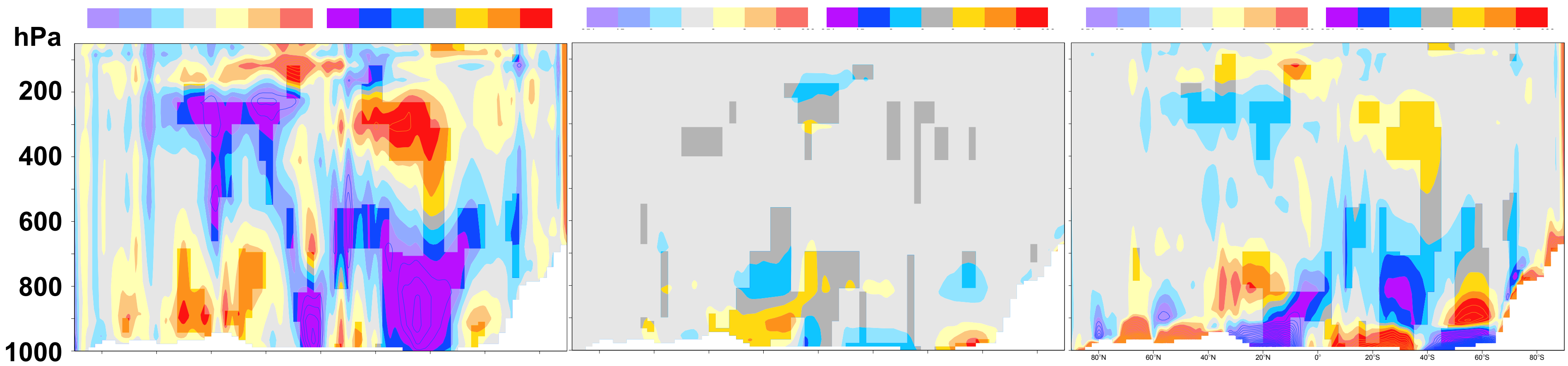
Impact of New Radiation Scheme

New radiation scheme reduced upper-tropospheric increments

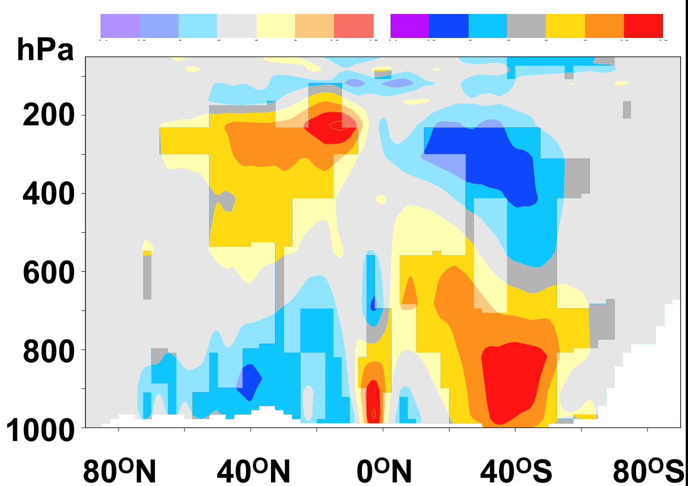


Mean Meridional Wind Tendencies

TOTAL CI=0.2ms^{-1}d^{-1} | CONVECTIVE CI=0.6ms^{-1}d^{-1} | DYNAMICAL



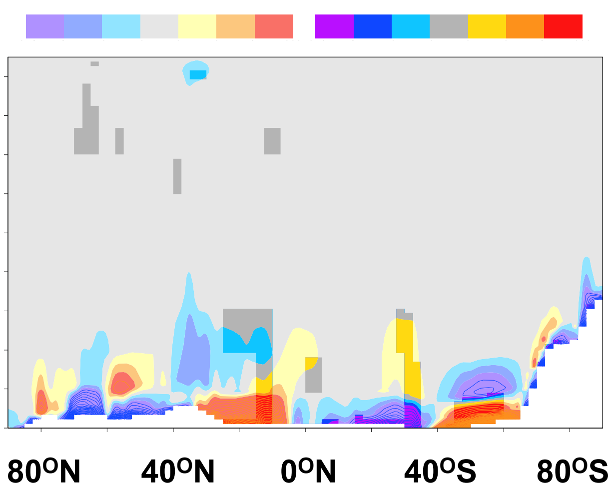
AN INC CI=0.4ms^{-1}d^{-1}



Either the dynamics in the upper troposphere are wrong, or there is a missing process.

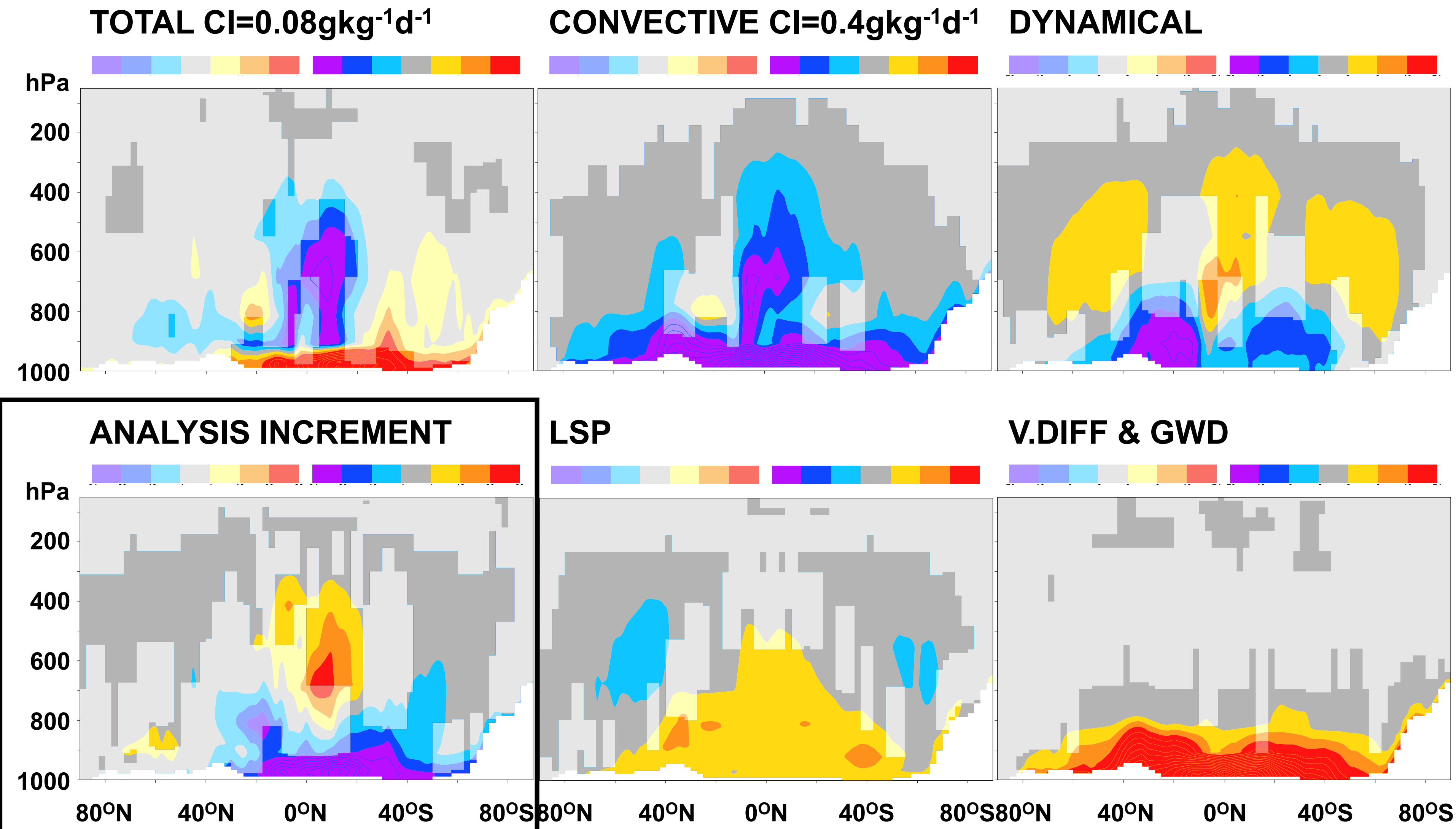
Hypothesis: need to increase tropical ascent and decrease radiative cooling and convective heating.

V.DIFF & GWD



Mean tendencies are deduced on model levels. Y-axis shows approximate pressure value. Average is over December 2008, 4 forecasts per day, tendencies accumulated from T+1 to T+7. Model cycle 33R1, T_L159, L91

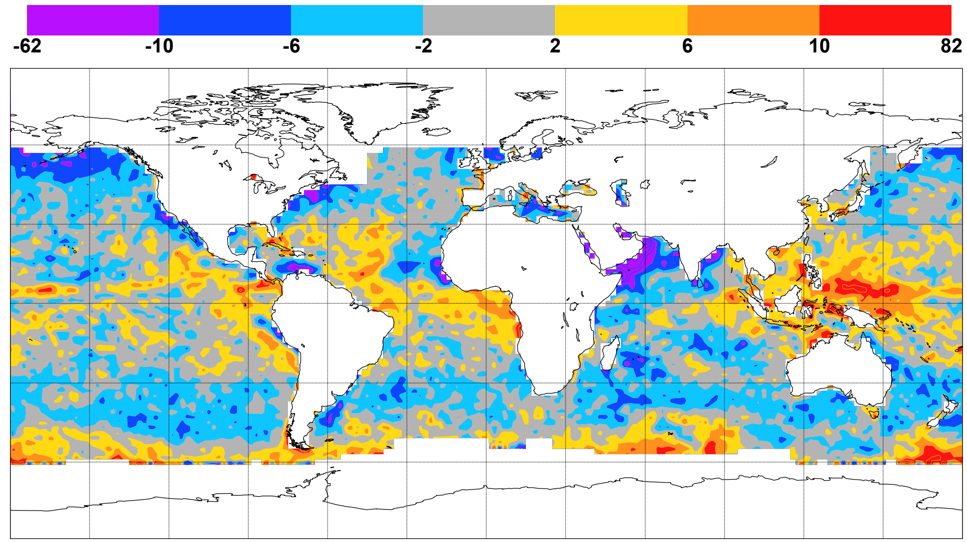
Mean Specific Humidity Tendencies



Mean tendencies are deduced on model levels. Y-axis shows approximate pressure value. Average is over December 2008, 4 forecasts per day, tendencies accumulated from T+1 to T+7. Model cycle 33R1, T_L159, L91

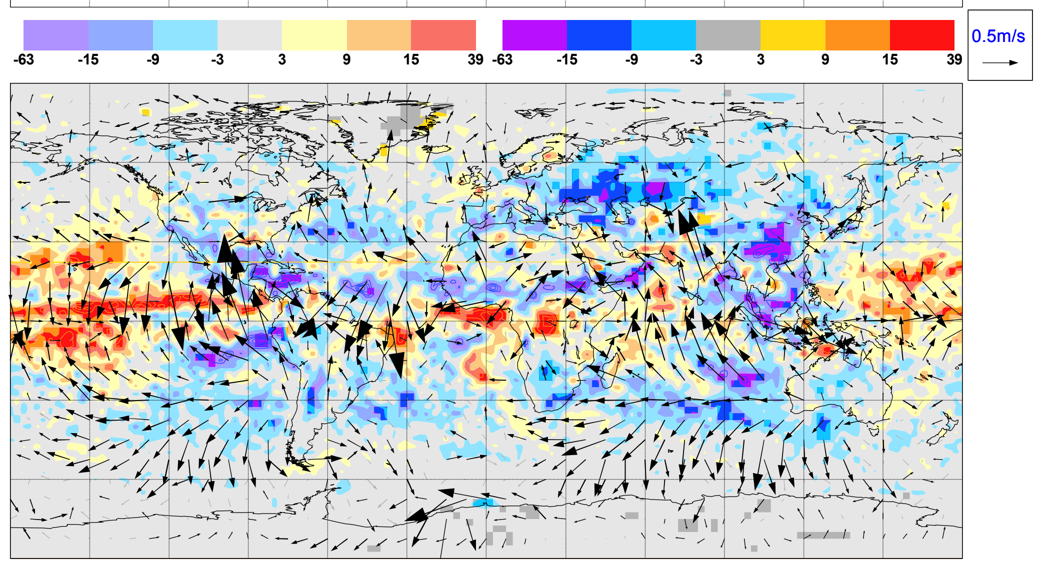
Specific Humidity at ~850 hPa

SSMI channel 3
Observation – First Guess
CI=0.4K



Drying increments in lower troposphere despite observations ‘wanting’ to moisten: rectification of noise at saturation point(?)

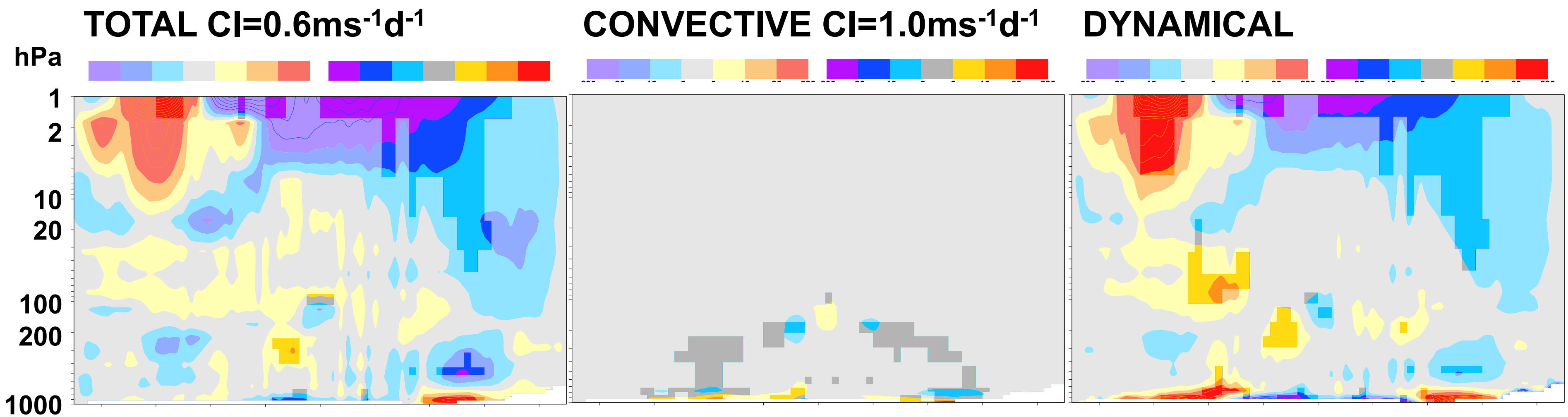
Analysis Increment
CI=0.06gkg^{-1}

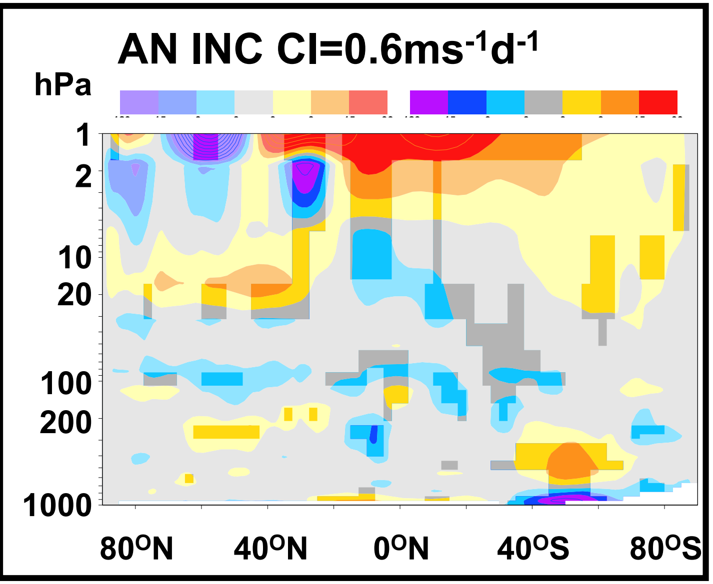


SSMI channel 3 all-sky microwave brightness temperature “first guess departures” and analysis increments are based on all 0 and 12 UTC data assimilation cycles 20090401—20090815 for IFS cycle 35R3 (E-suite), T_L799, L91. SSMI brightness temperature has a positive correlation with humidity.

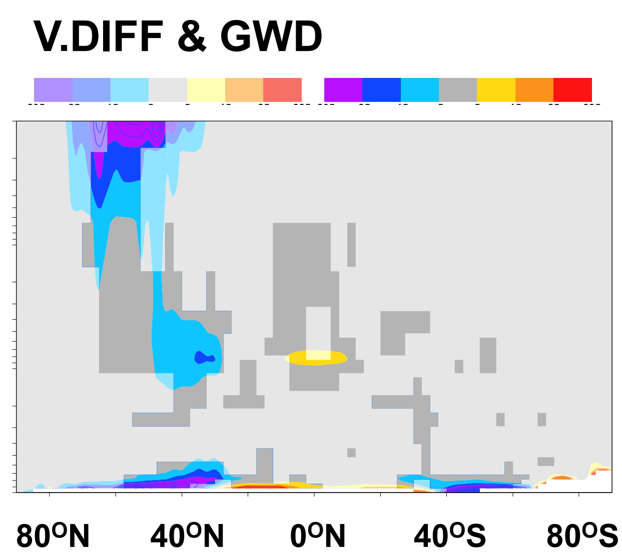
M.J. Rodwell

Mean Zonal Wind Tendencies (Trop & Strat)





Balance between dynamical and gravity-wave drag tendencies is not perfect

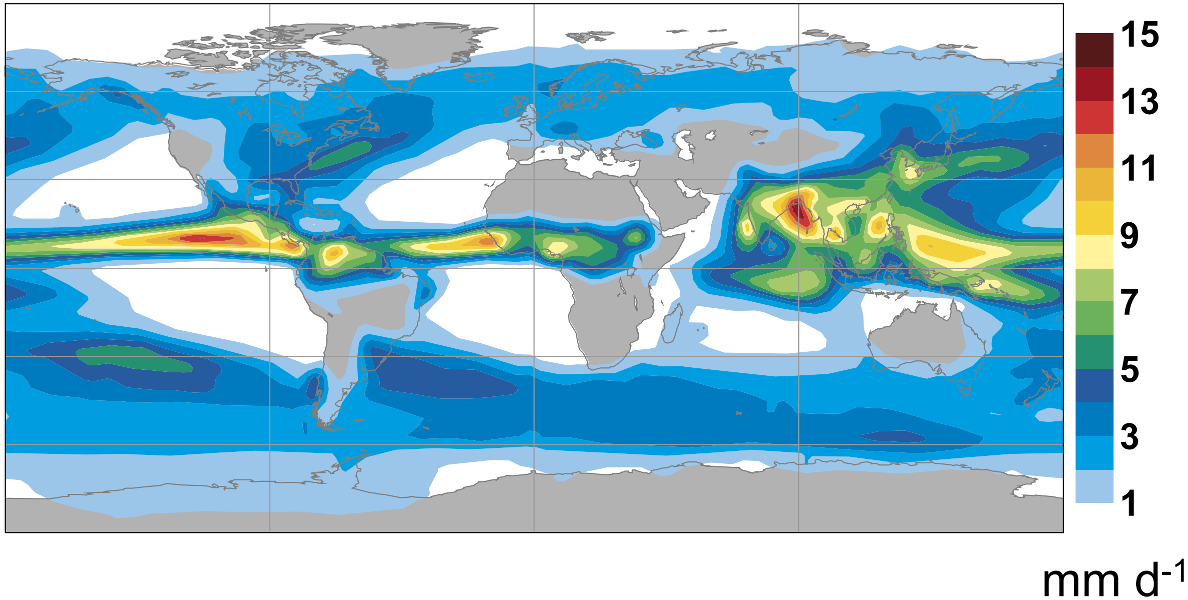


Mean tendencies are deduced on model levels. Y-axis shows approximate pressure value. Average is over December 2008, 4 forecasts per day, tendencies accumulated from T+1 to T+7. Model cycle 33R1, T_L159, L91

The Asian Monsoon

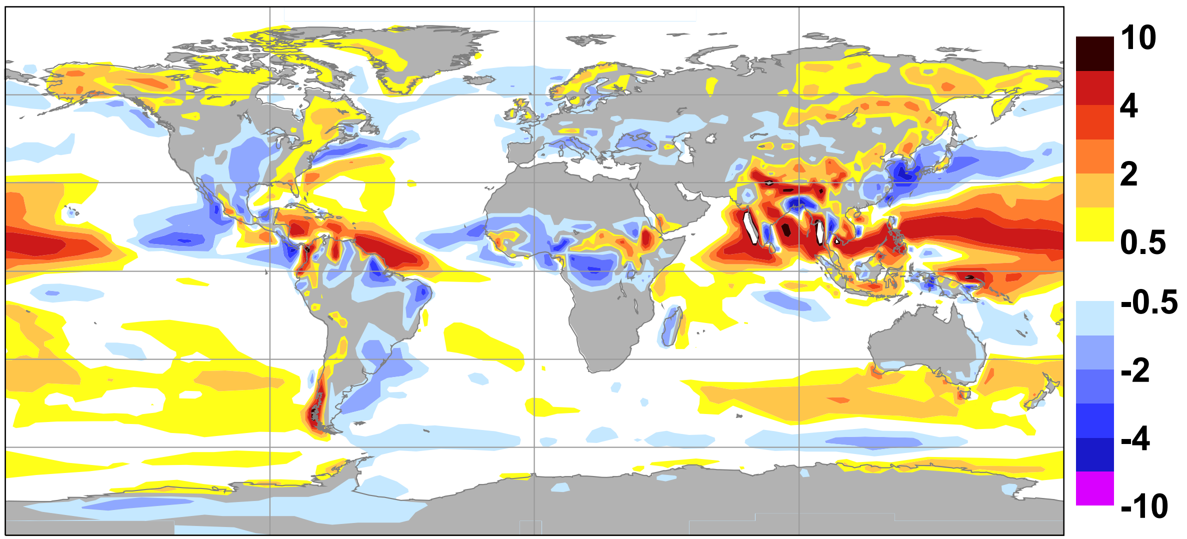
Precipitation JJA 1963-2006

OBSERVED PRECIPITATION GPCP



Too much rainfall in the Asian monsoon has been a long-standing problem

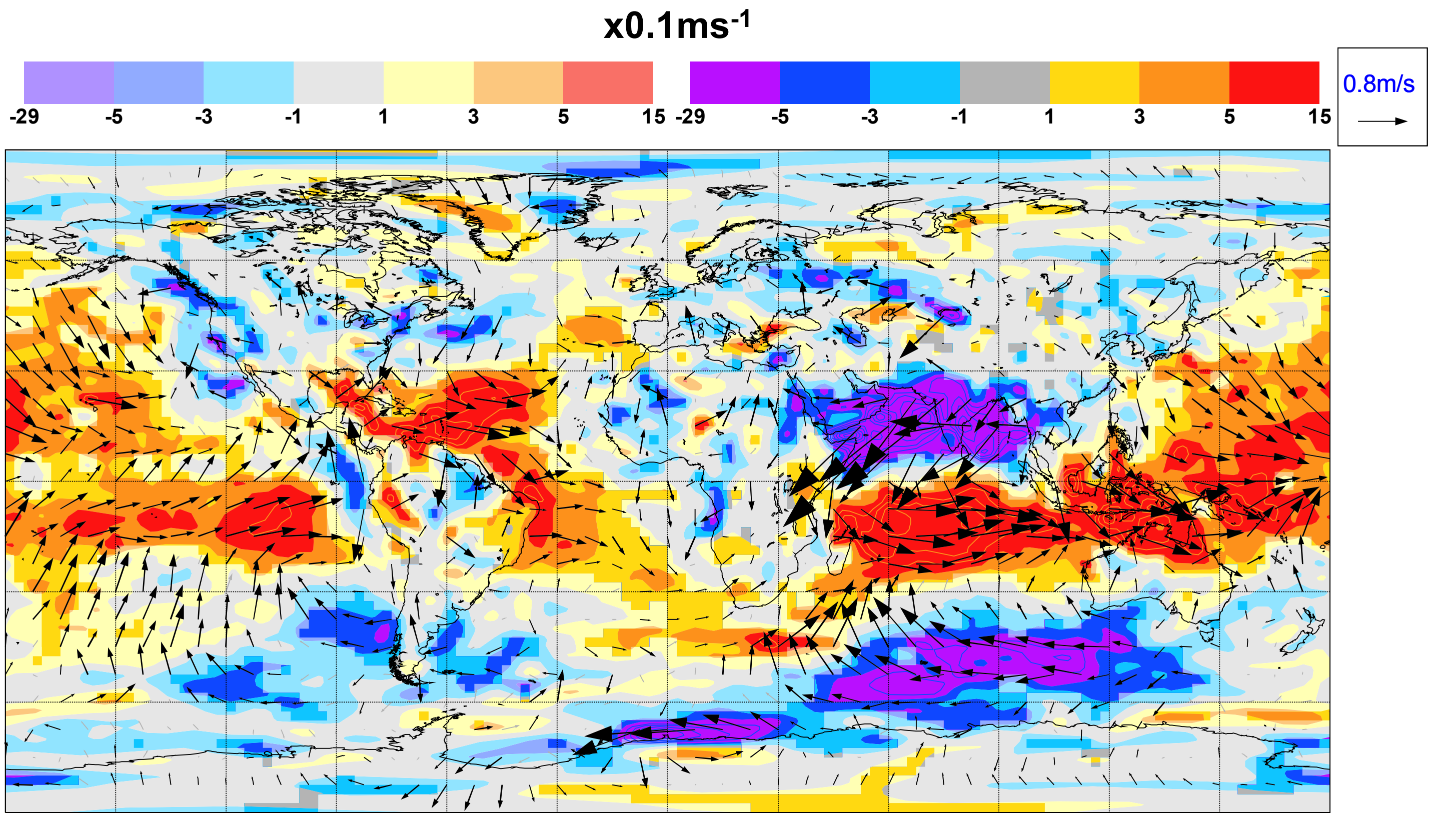
PRECIPITATION ERROR MODEL CLIMATE-GPCP



Model is cycle 33R1, T_L159 L91

M.J. Rodwell

JJA 2008 u and v 925hPa Analysis Increments

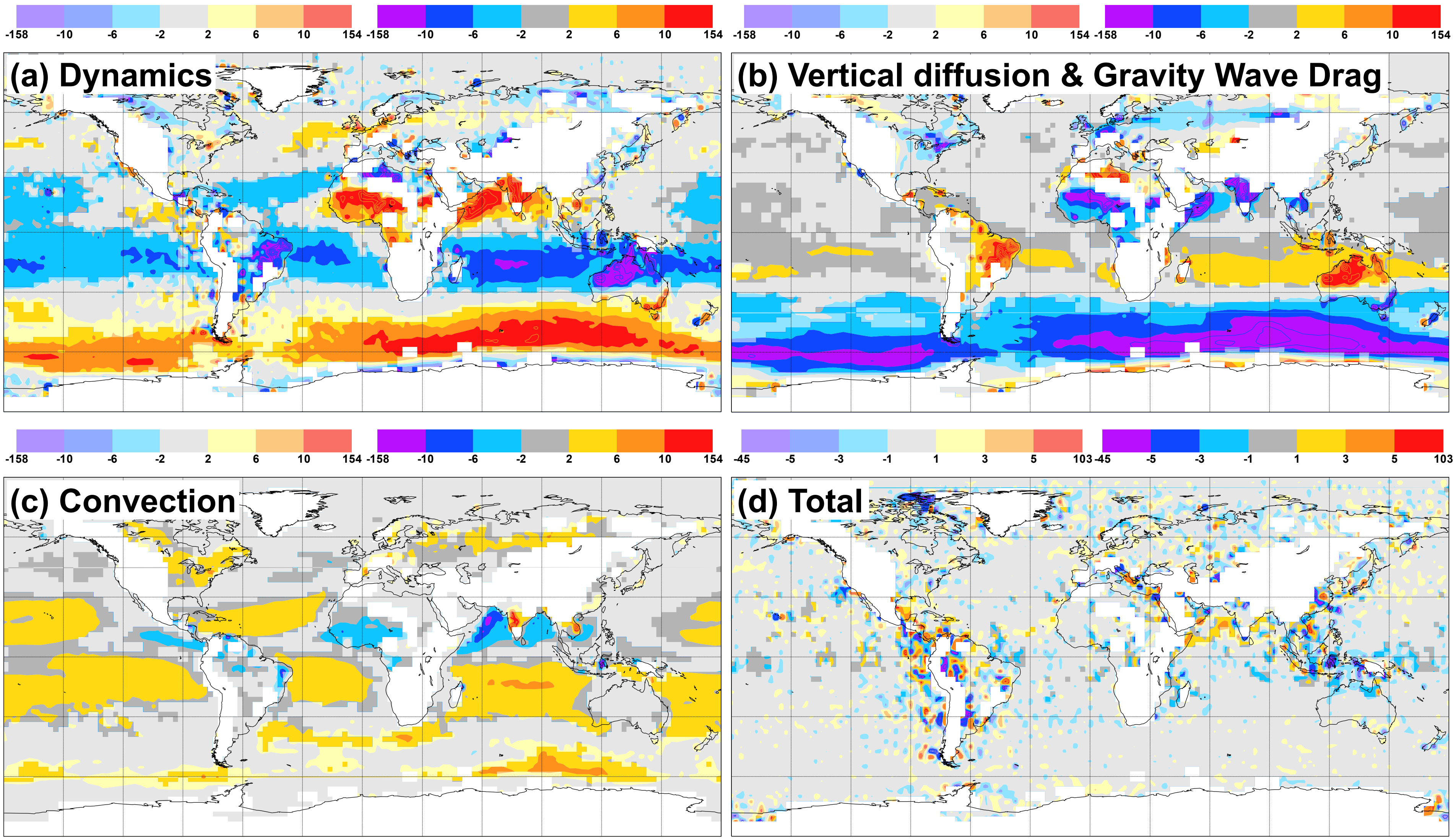


Analysis Increments indicate that the model wants to transport too much moisture into the monsoon: The root-cause of the monsoon error?

M.J. Rodwell

Initial Tendencies JJA 2008: u at 925 hPa

Unit = ms^{-1} over first 24h of forecast

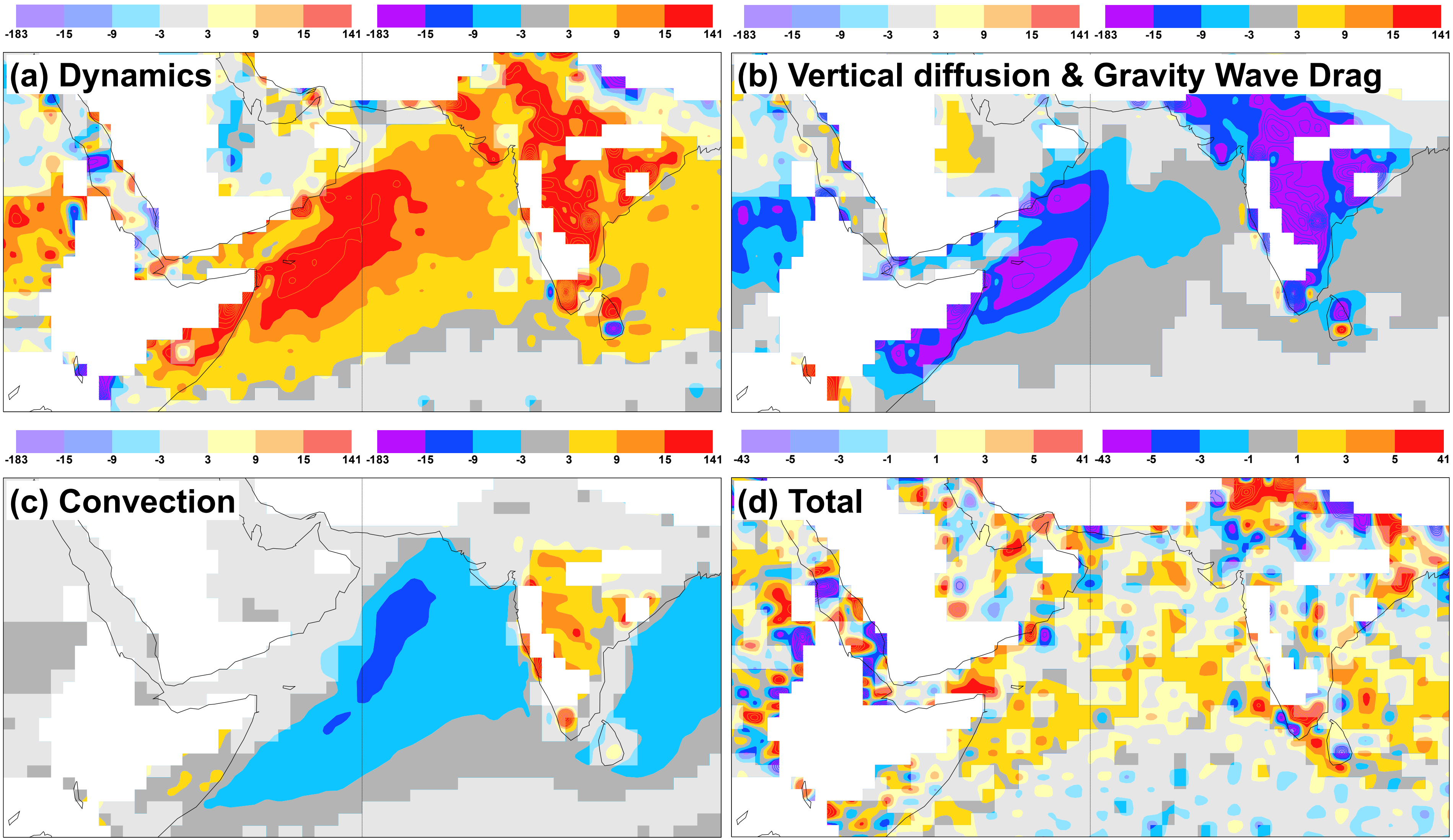


Arabian Sea region is unusual in having strong compensating tendencies

M.J. Rodwell

Initial Tendencies JJA 2008: u at 925 hPa

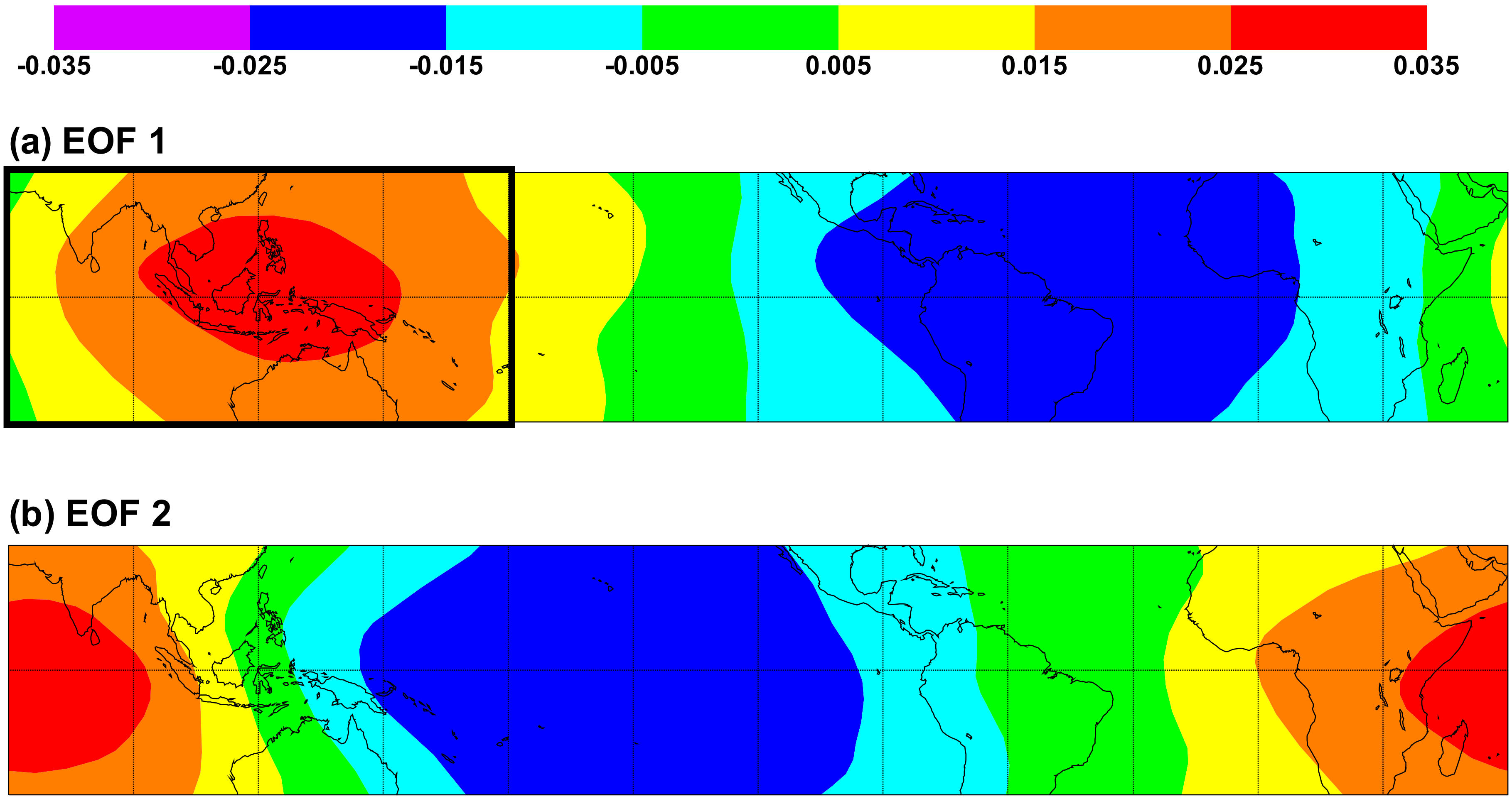
Unit = ms^{-1} over first 24h of forecast



Further work required to understand which tendency is at fault (other parameters, CRMs etc)

The Madden-Julian Oscillation (MJO)

EOFs of Vel.Pot. ERA-Interim Re-analysis

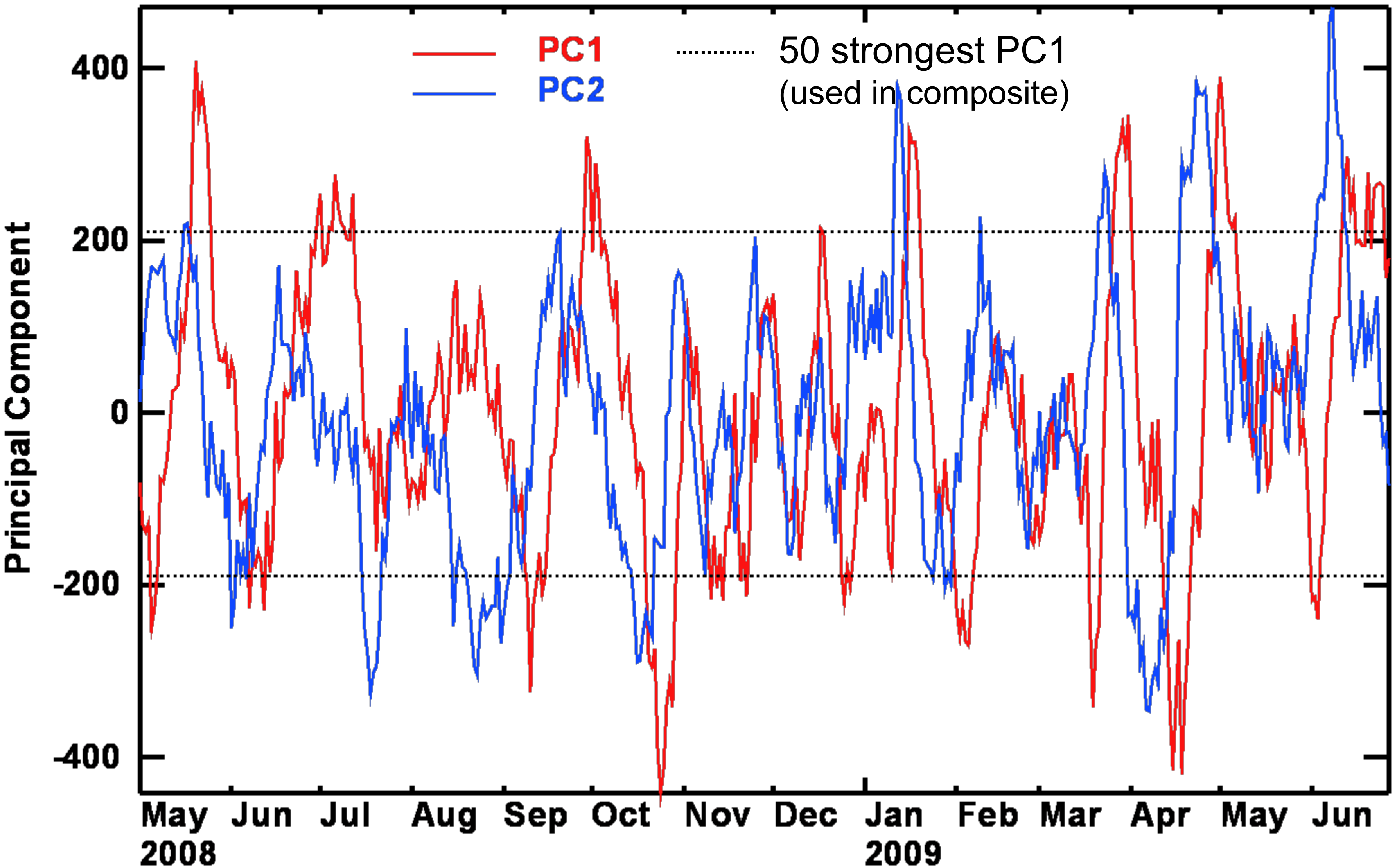


Will focus on the physics associated with EOF1 in the box indicated.

EOFs derived from ERA-Interim re-analyses for the period 19890101—20071231

M.J. Rodwell

PC1&2 of Vel.Pot. Operational Analyses



EOF2 leads EOF1 by a quarter period: indicating eastward propagation

MJO Phase Propagation In ECMWF Model

Convection over Maritime Continent

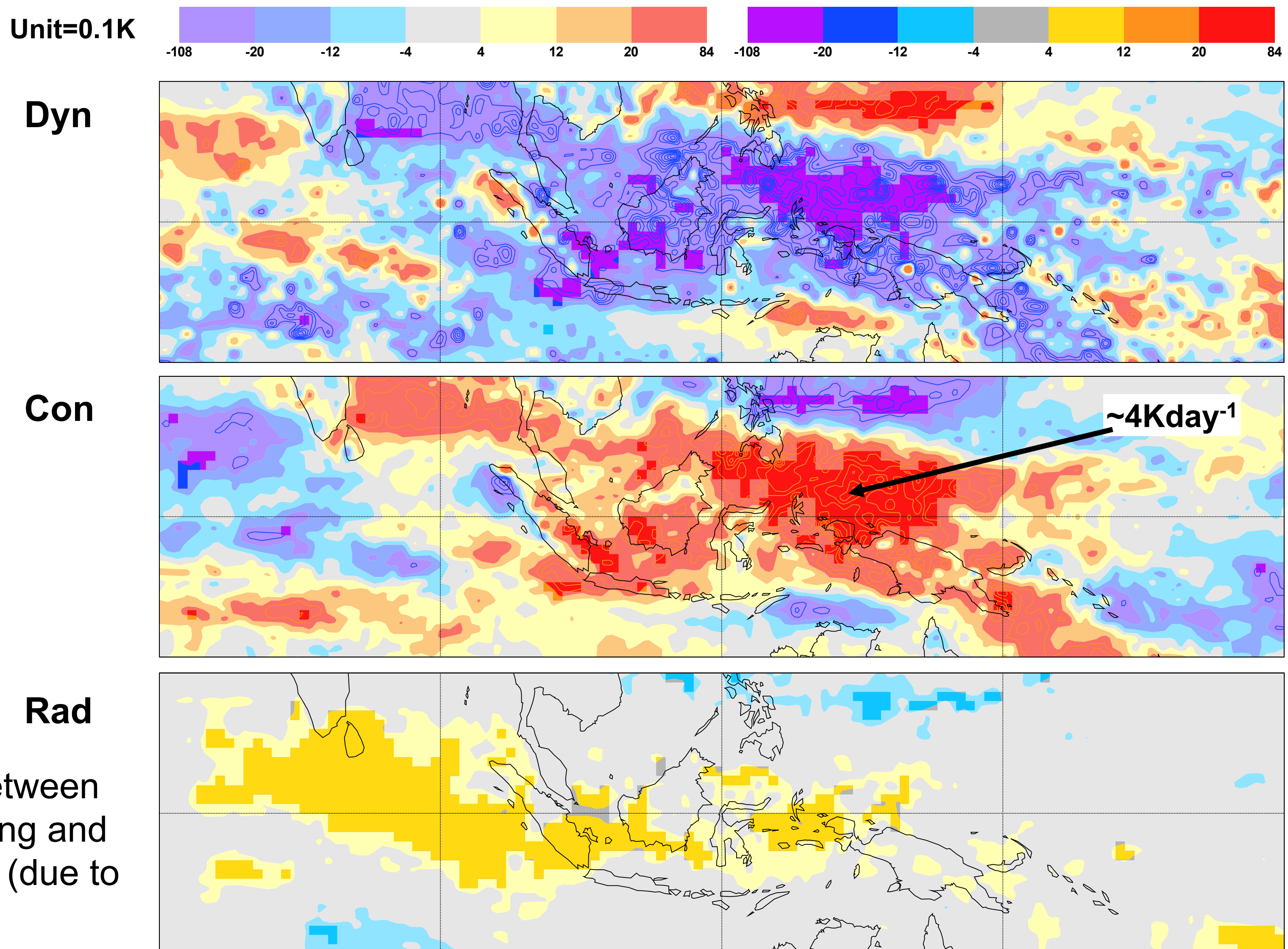
Phase Shift	1→2	2→3	3→4	4→5	5→6	6→7	7→8	8→1
OBS	71%	81%	81%	80%	86%	79%	68%	55%
Mod	71%	81%	80%	71%	72%	78%	65%	87%

Frederic Vitart

Model finds it difficult to propagate the MJO through the “Warm Pool”

Percentage of times that the MJO moves from one phase to the next within 10 days. Model cycle 32R3.

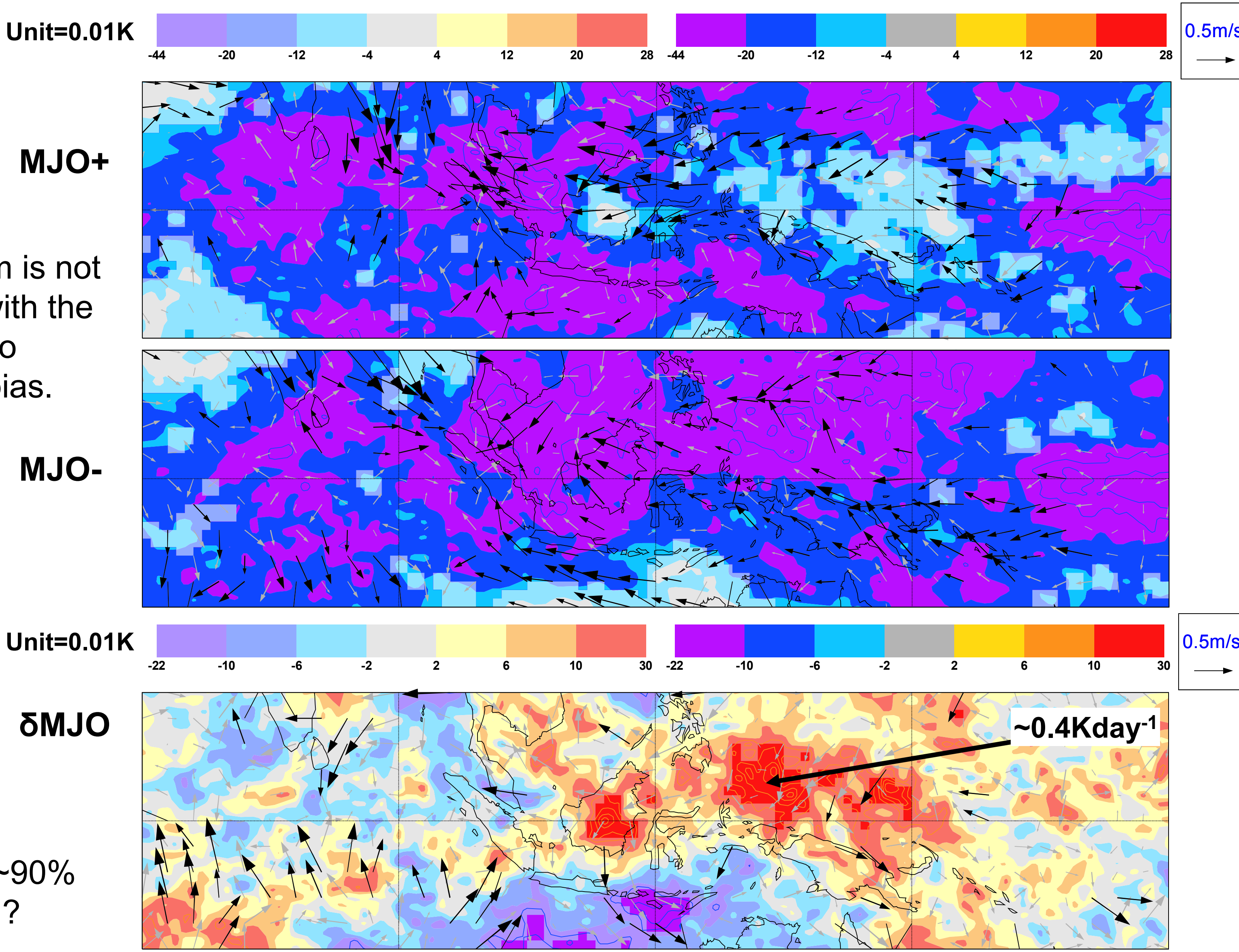
Initial Tendencies (First 24hr): T500, δMJO



Main balance between convective heating and dynamic cooling (due to ascent).
Radiation stabilises atmosphere behind MJO

M.J. Rodwell

Analysis Increments (12hr window): T500



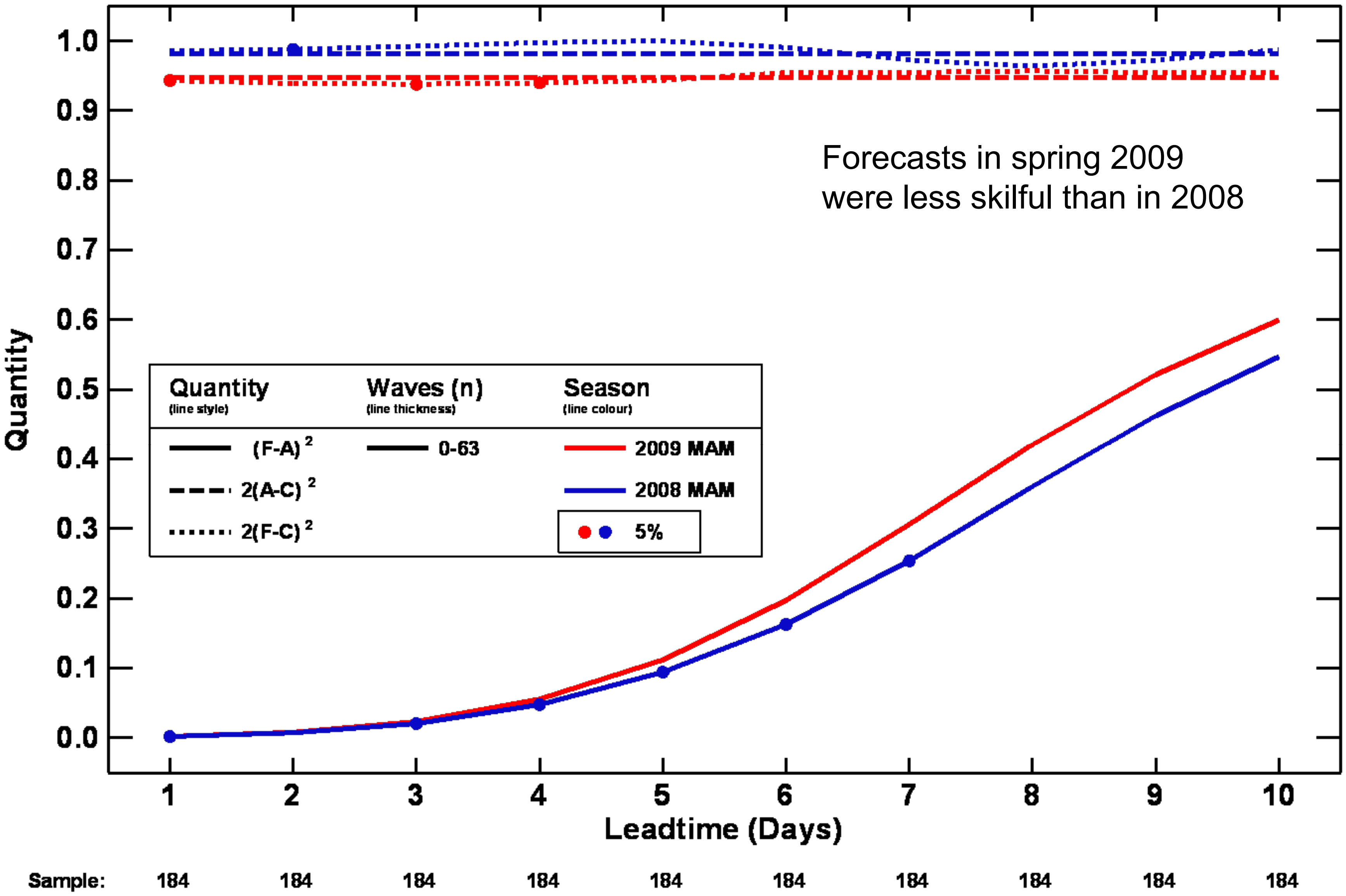
Main problem is not associated with the MJO. Need to reduce wet bias.

Model “MJO convection” ~90% of true signal?

Scale-Dependent Verification

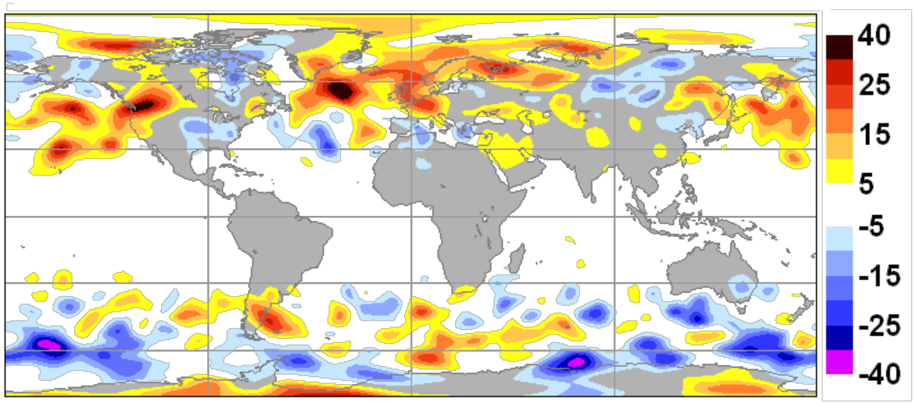
- Northern Mid-latitudes Spring 2009

Z500 Mean-Squared Error and Activity: Ops

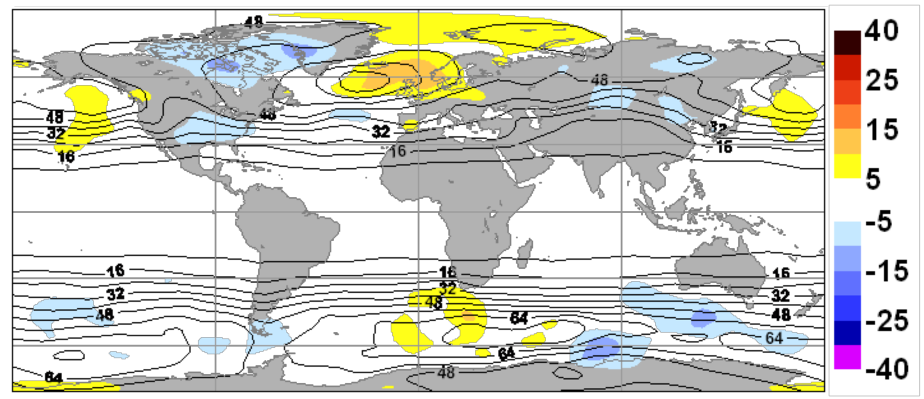


Z500 D+5 MAM: 2009-2008

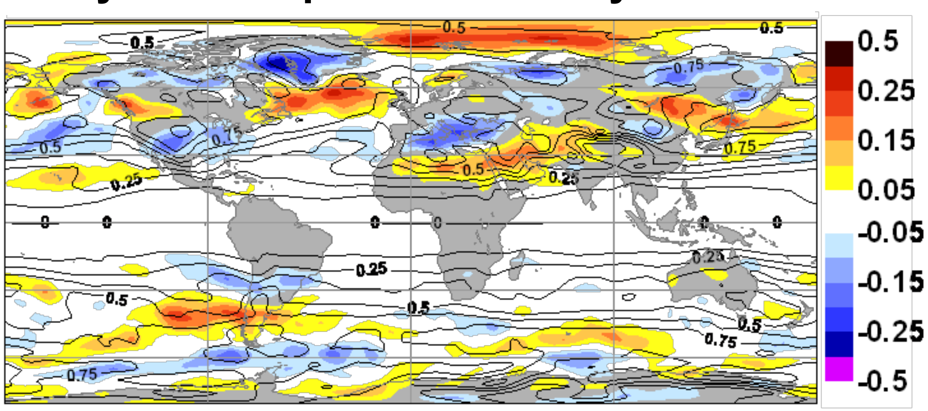
RMSE: Operational Forecast



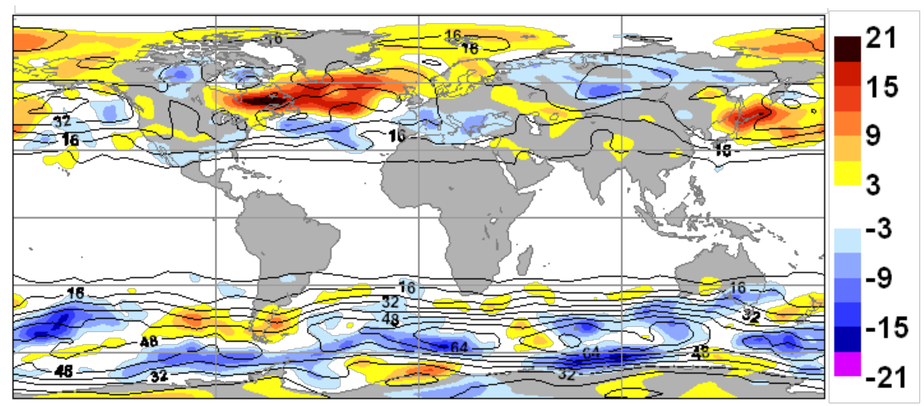
EPS Spread: Operational Forecast



Eady Index: Operational Analysis

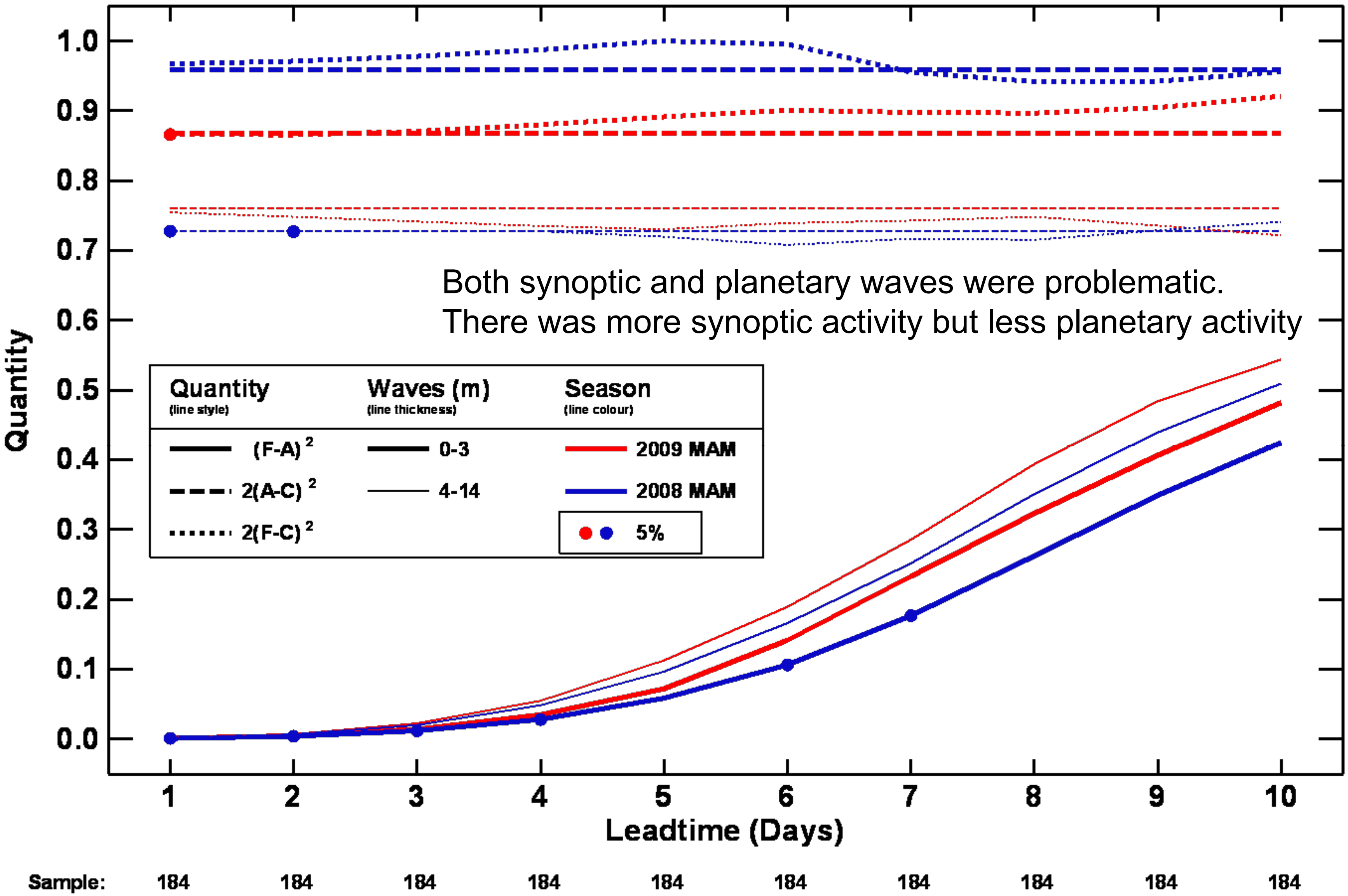


Synoptic Activity: Operational Analysis

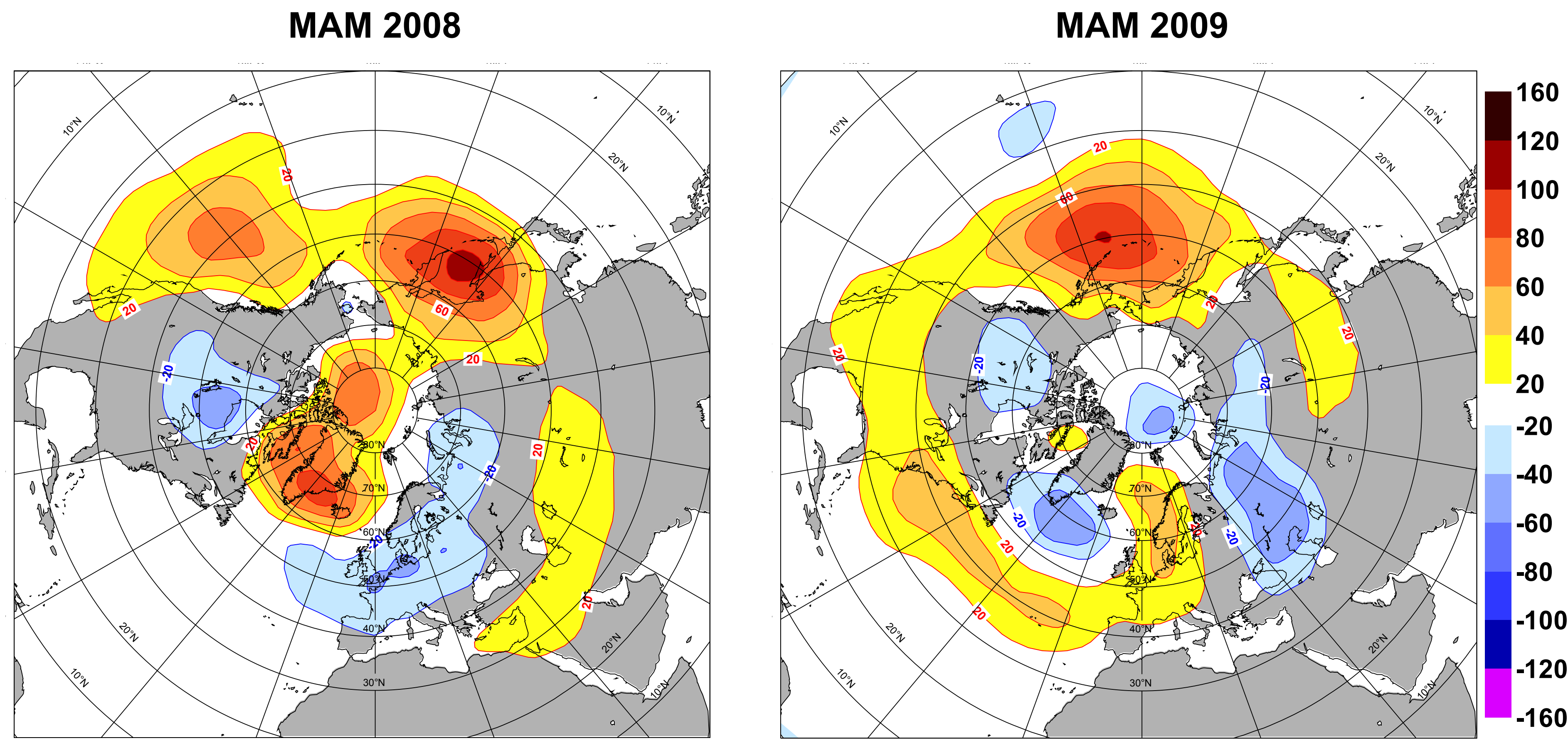


… More synoptic activity partly the reason?

Z500 Scale-Dependent MSE and Activity: Ops

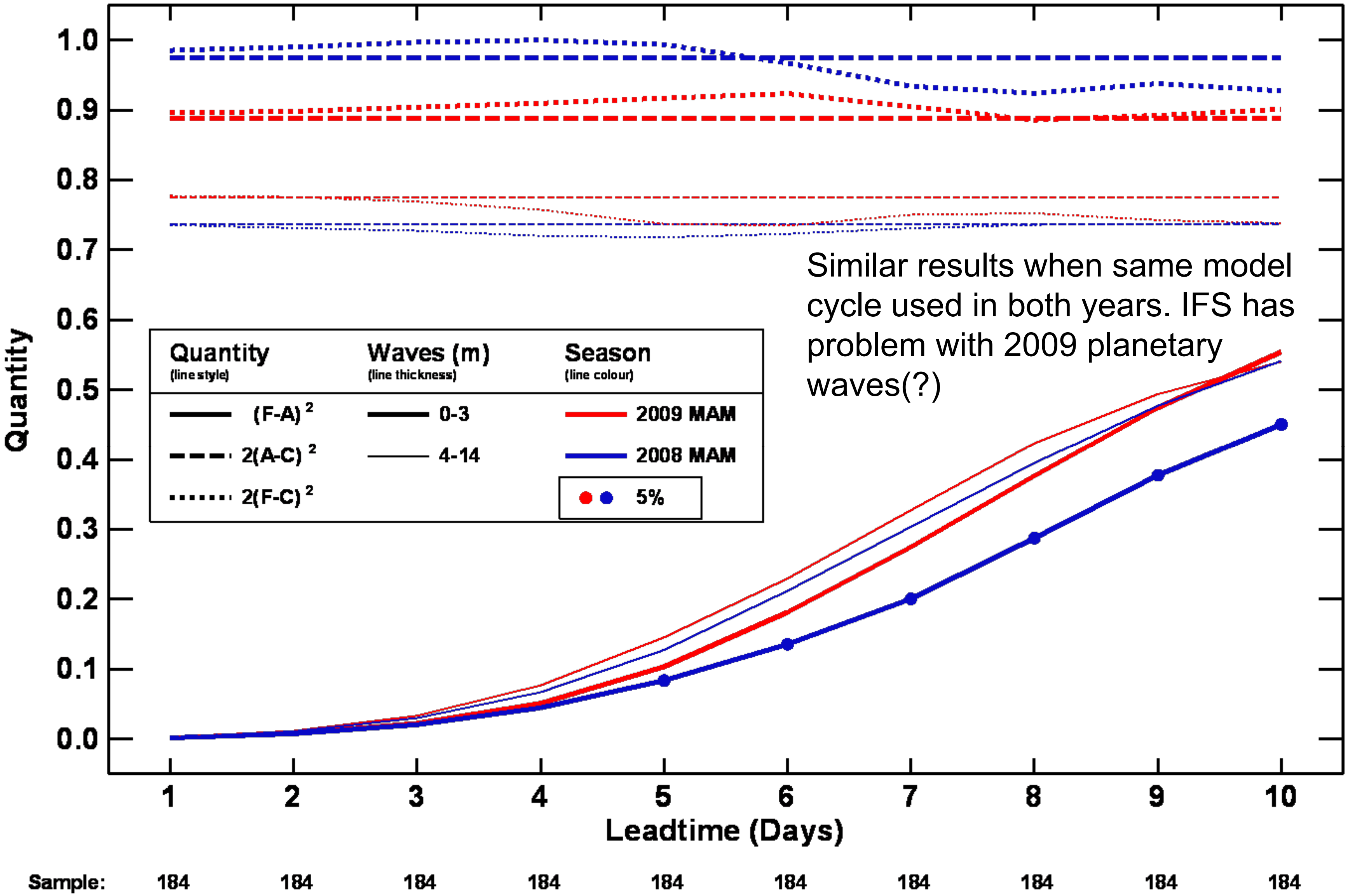


Mean Z500 Anomalies

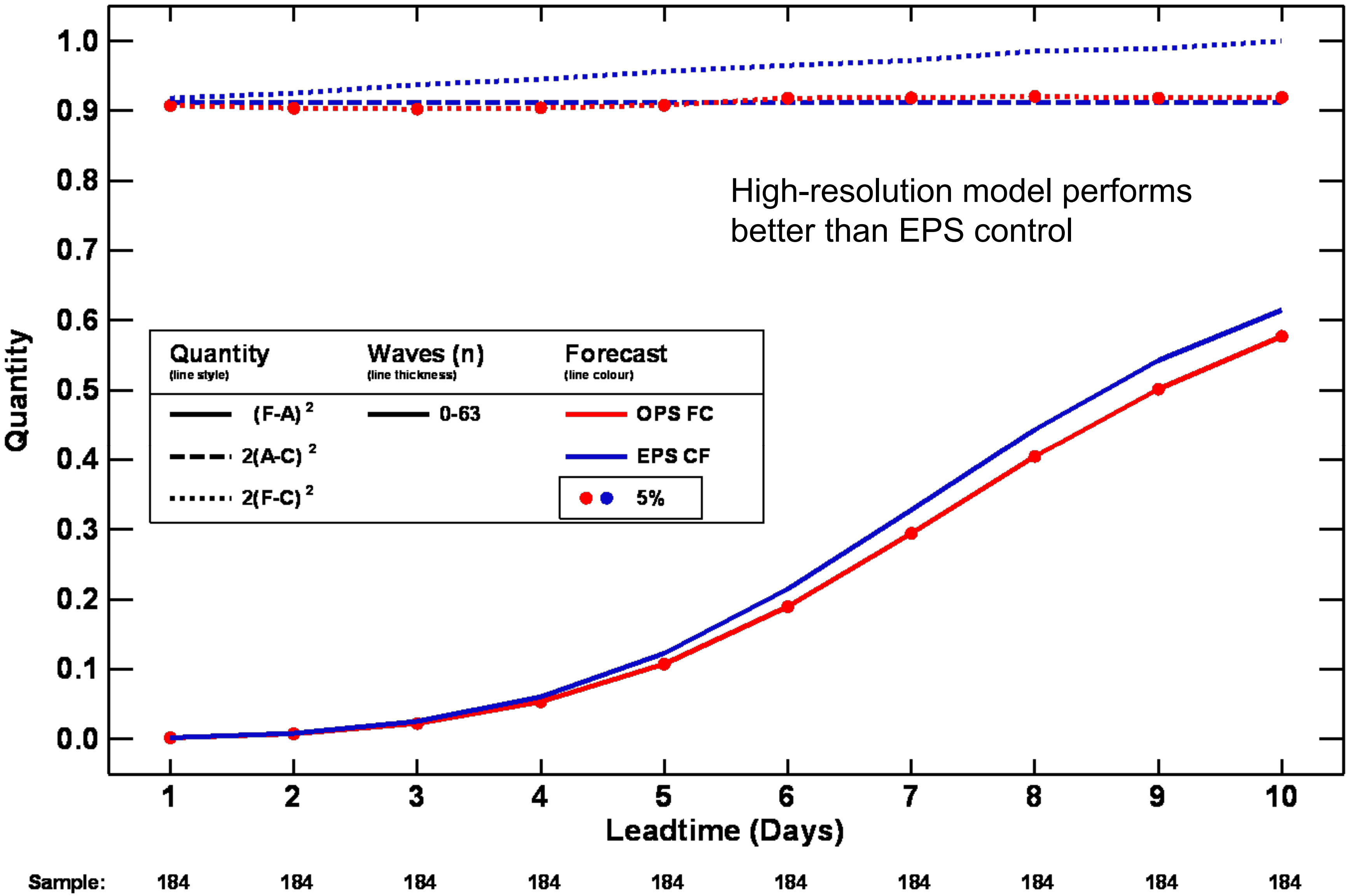


Planetary waves were very different in 2009 compared to 2008

Z500 Scale-Dept. MSE and Activity: ERA-Interim

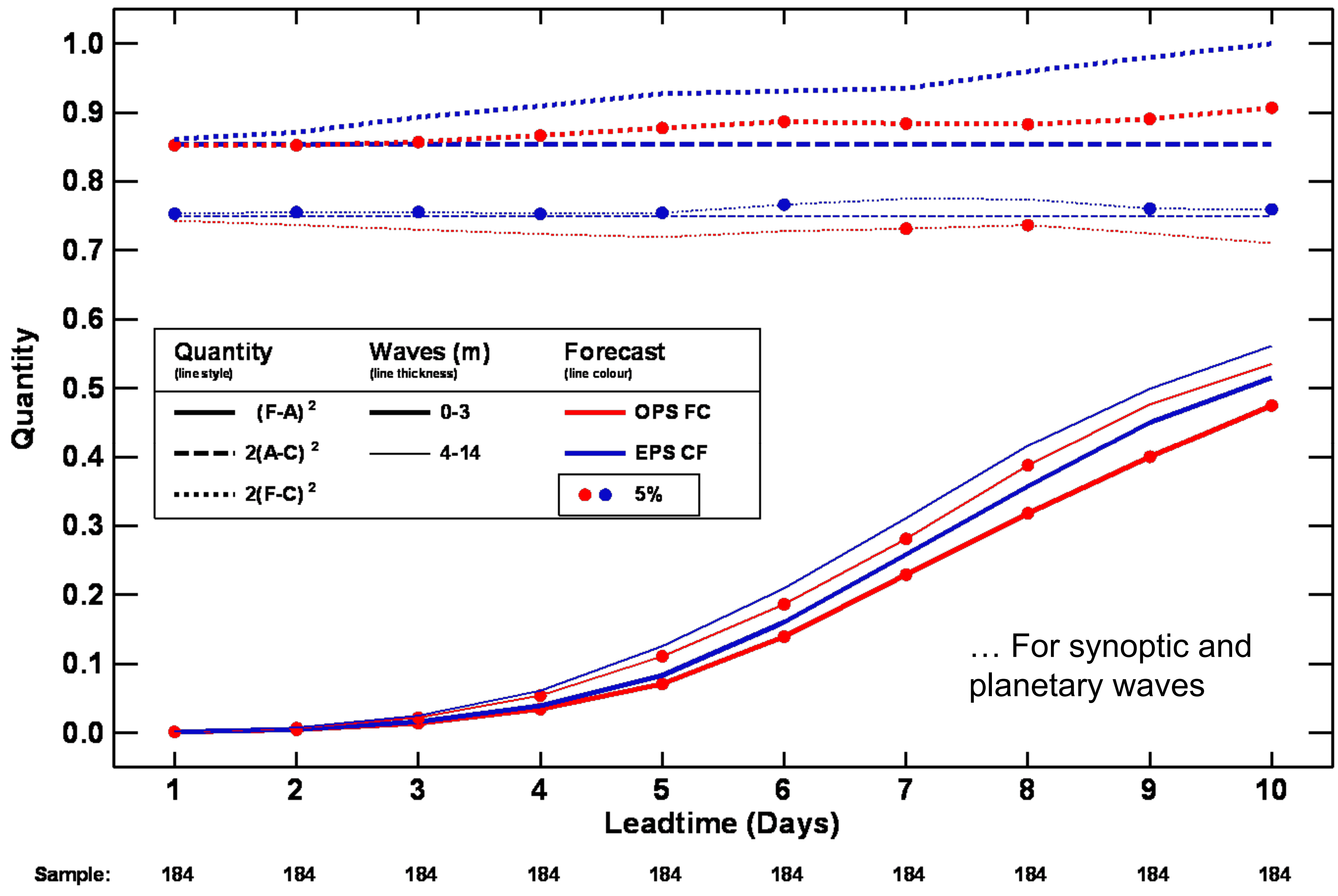


Z500 MSE and Activity: FC versus CF



Diagnostics
R&J 40

Z500 Scale-Dependent MSE and Activity: FC,CF



Diagnostic Verification

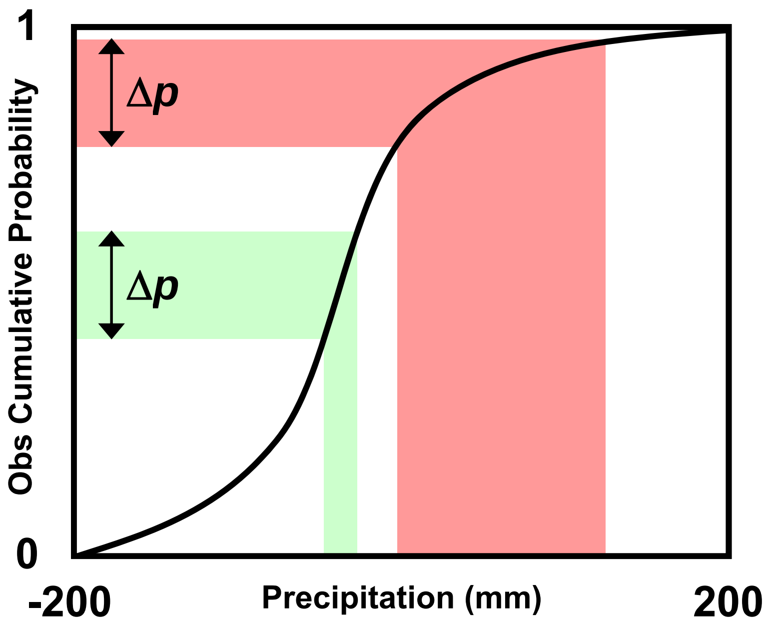
- Verification measures that aid diagnosis of error
- Verification of smaller-scale quantities: Precipitation

Diagnostics
R&J 42

Linear Error in Probability Space & Equitable Skill

- "LEPS" – Ward and Folland (1991) in terms of seasonal rainfall totals
- Equitable scores – Gandin and Murphy (1992)

Linear Error in Probability Space



Aim: to combine these two concepts into a score for daily precipitation forecasts

Equitable Categorical Skill Score

	Observation				
	Probability	p_0	p_1	...	p_n
Forecast	Category	0	1	...	n
	0	s_{00}	s_{10}	...	s_{n0}
	1	s_{01}	s_{11}	...	s_{n1}
	:	:	:	.	:
	n	s_{0n}	s_{1n}	...	s_{nn}

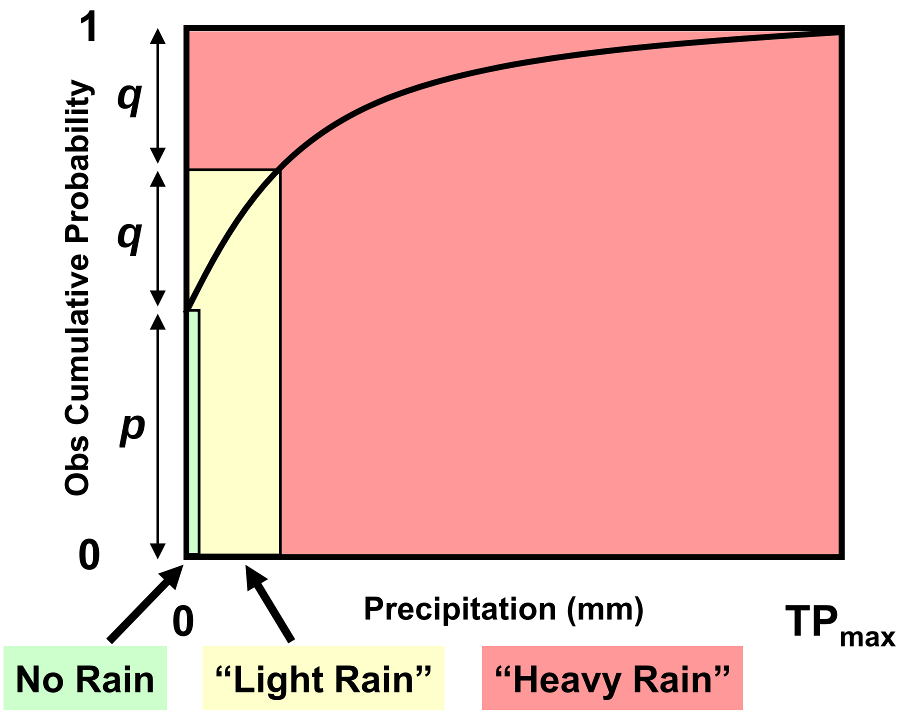
Perfect FC $\sum_i p_i s_{ii} = 1$

Constant FC $\sum_i p_i s_{ij} = 0 \quad \forall j$

Random FC $\sum_{i,j} q_j p_i s_{ij} = 0 \quad \sum_j q_j = 1$

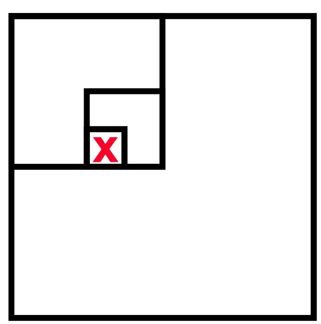
A New Score for Precipitation ("SLEEPS")

Use of Cumulative Distribution



Nearest Grid-box to Point-Observation

- Monitor scores as resolution improves
- 24hr accumulation improves representativeness



Semi-Linear error in probability space

		Observation		
	Probability	p	q	q
Forecast	Category	No Rain	Light Rain	Heavy Rain
	No Rain	0	c	c+a
	Light Rain	d	0	a
	Heavy Rain	d+b	b	0

- Clear link between forecast score and model error (compared to, e.g., correlation)

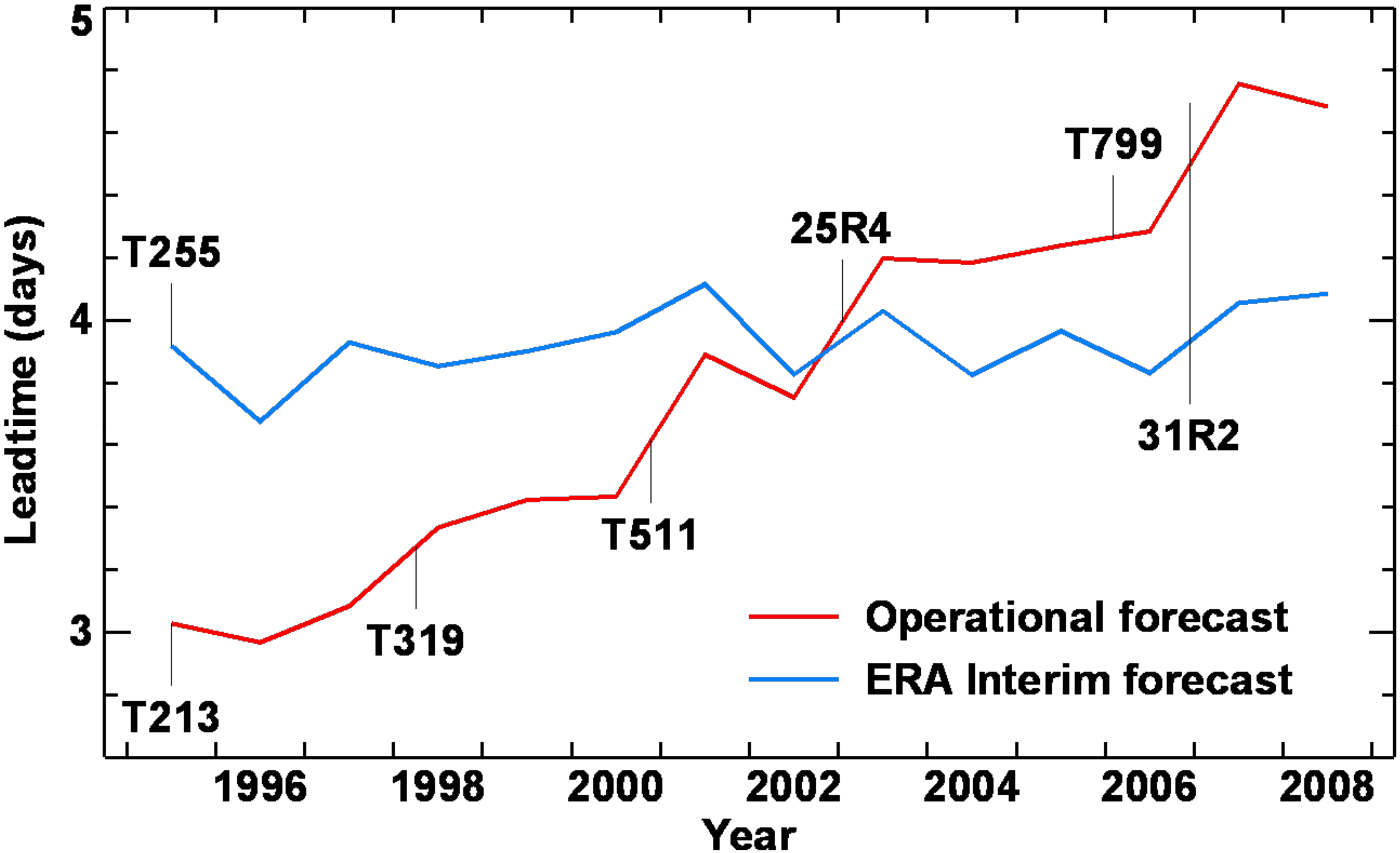
Enhanced Equitability Constraints

$$q(a+2c)=1 \qquad pd+qa=1$$

$$(p+q)b+pd=1 \qquad pd=qa$$

- Better score $\Rightarrow$ Better system
- Less sensitive to sampling uncertainty
- Aggregation of errors from different climatic regions

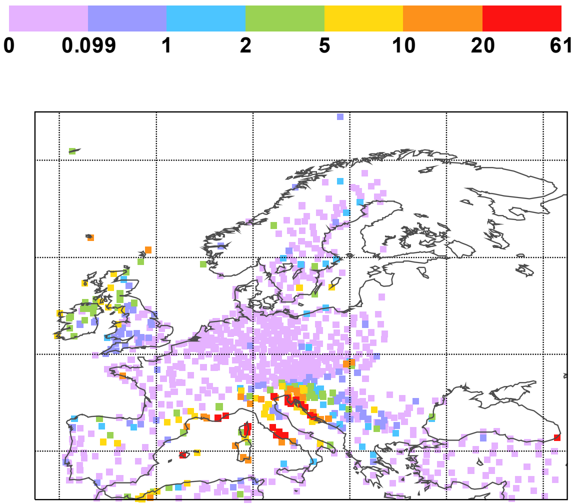
Extra-Tropical 'SLEEPS' Lead-time to Error of 0.6



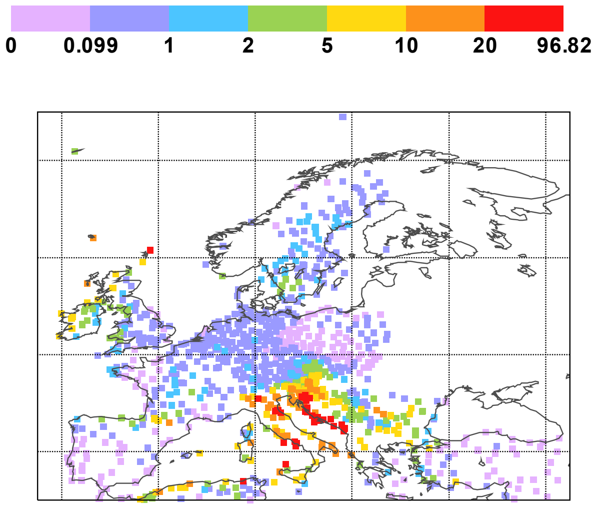
For the Northern Hemisphere, improvements must be due to model formulation, resolution or data assimilation methodology (not increasing number of observations)

Ten Worst D+4 Precip. Forecasts of 2008: #8

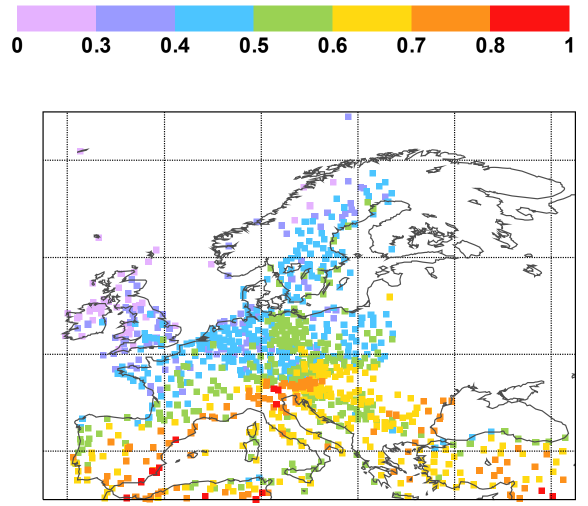
(a) Observation



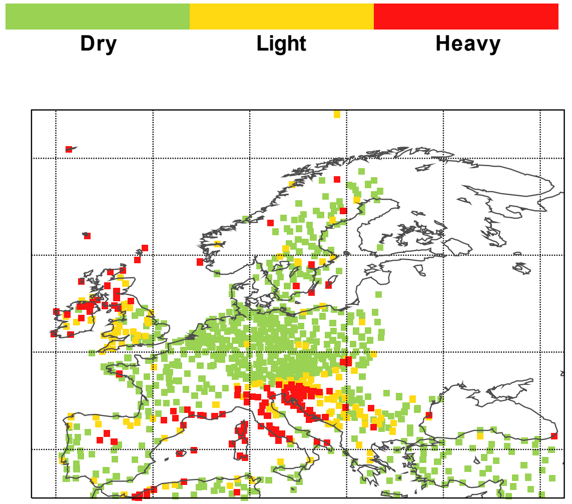
(b) Forecast



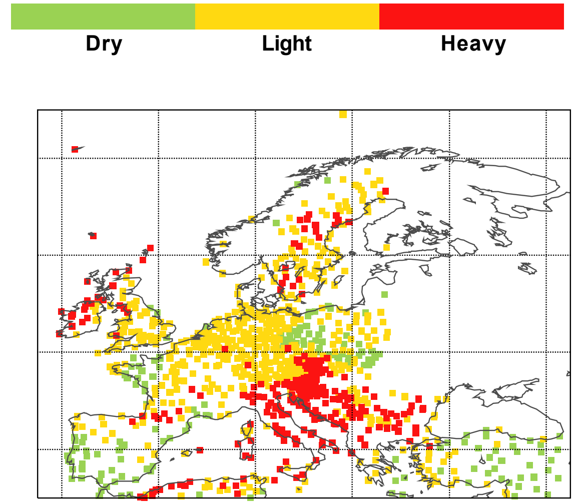
(c) Probability No Precip.



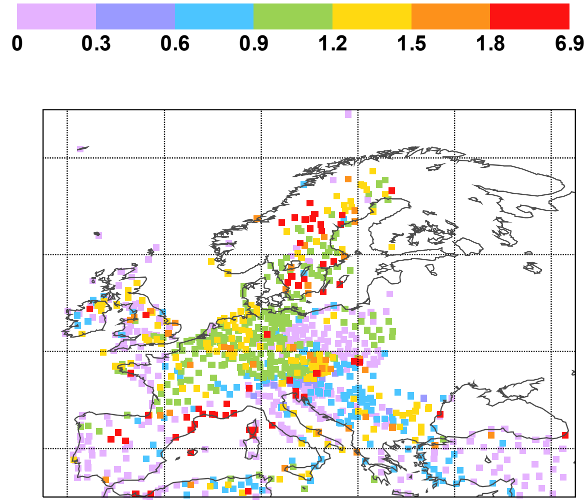
(d) Observed Category



(e) Forecast Category



(f) SLEEPS

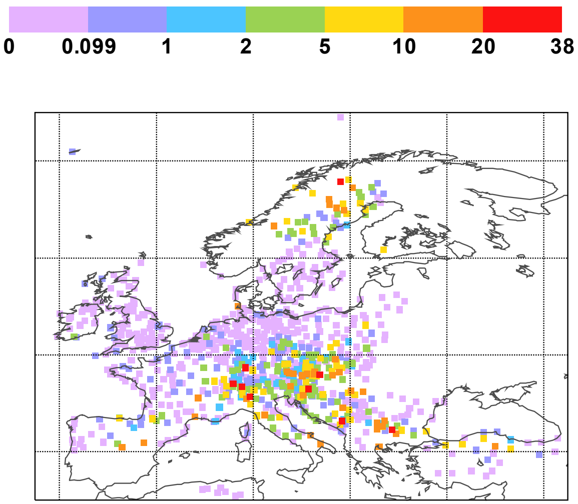


Forecast is penalised for not predicting dry weather in (generally wet) Northern Europe in winter

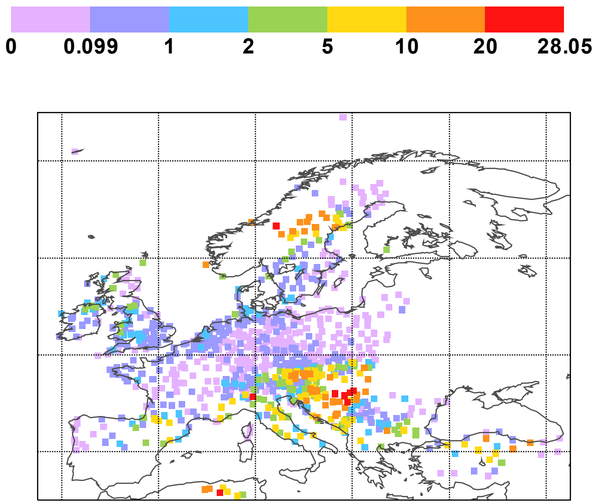
Diagnostics
R&J 46

Ten Worst D+4 Precip. Forecasts of 2008: #6

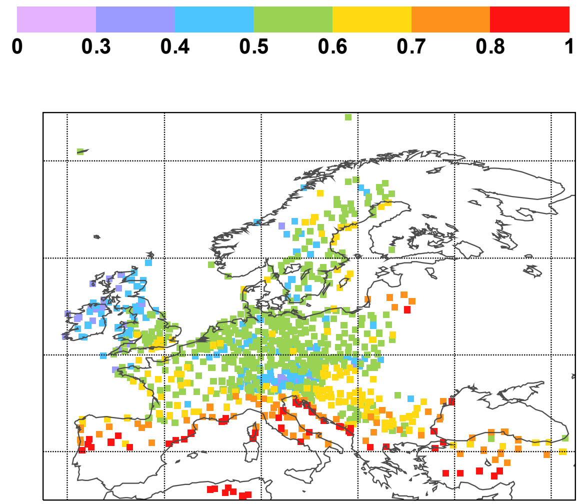
(a) Observation



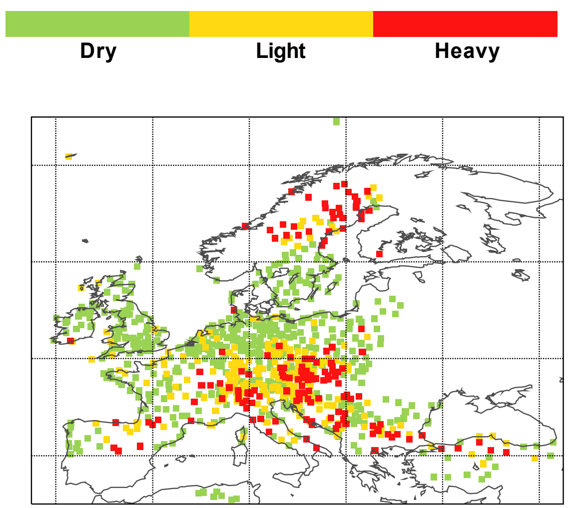
(b) Forecast



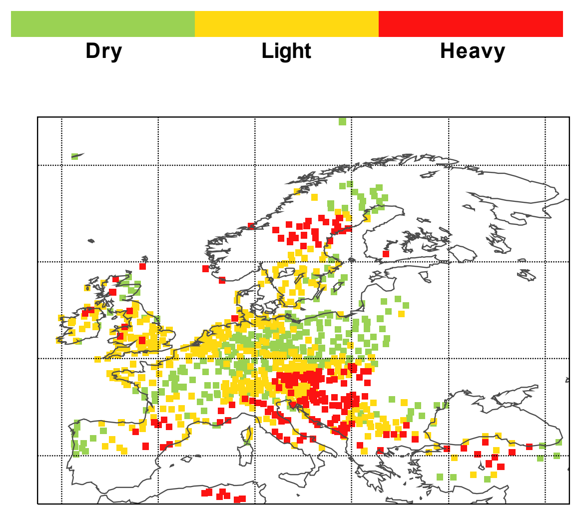
(c) Probability No Precip.



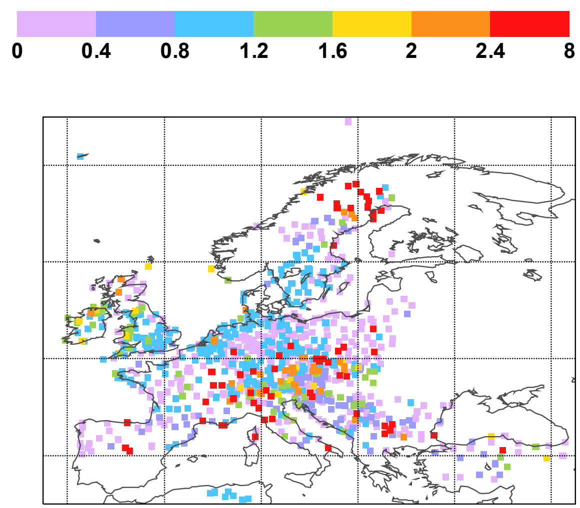
(d) Observed Category



(e) Forecast Category



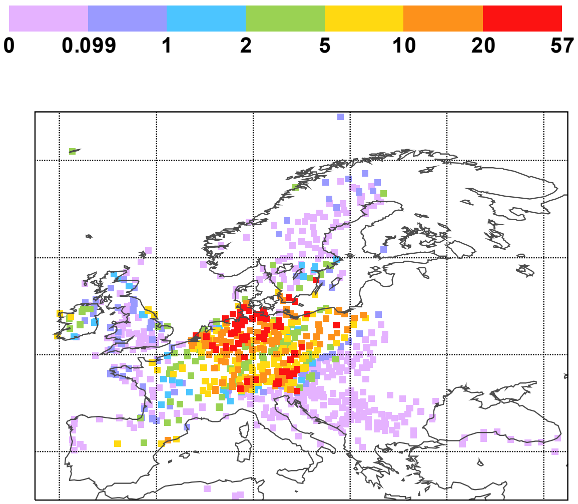
(f) SLEEPS



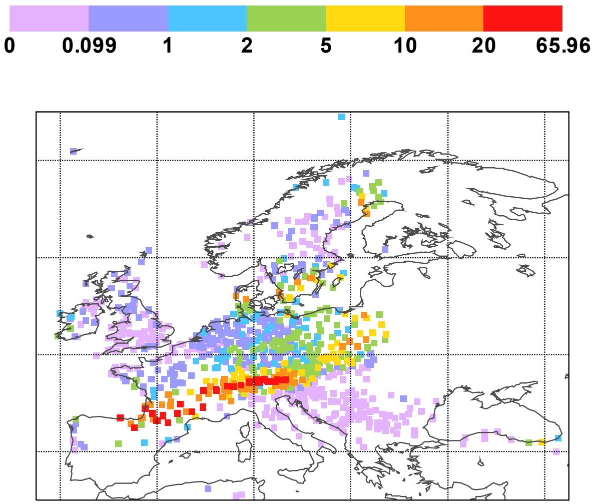
Forecast penalised for not predicting convective rain in (generally dry) Southern Europe in summer

Ten Worst D+4 Precip. Forecasts of 2008: #4

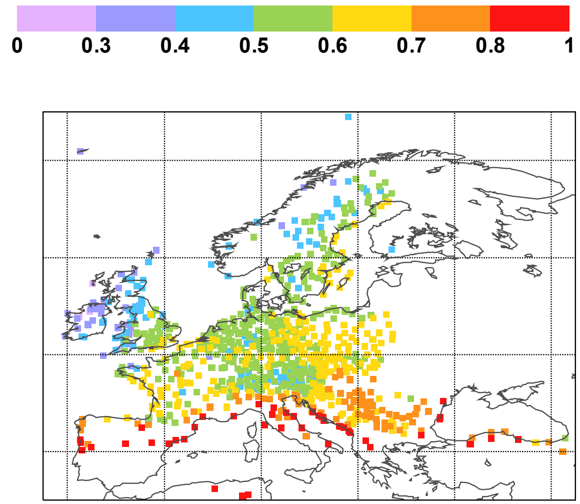
(a) Observation



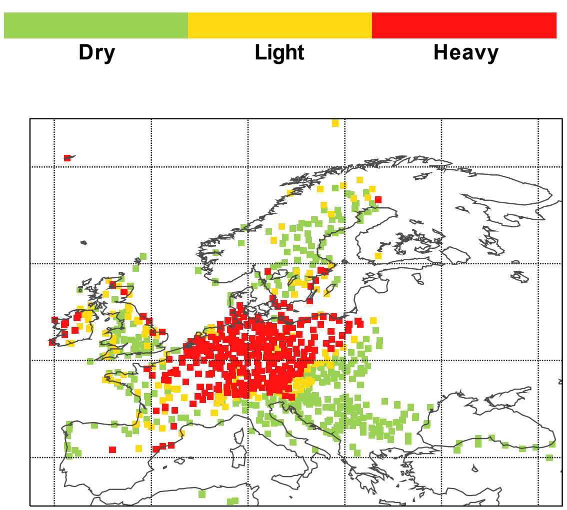
(b) Forecast



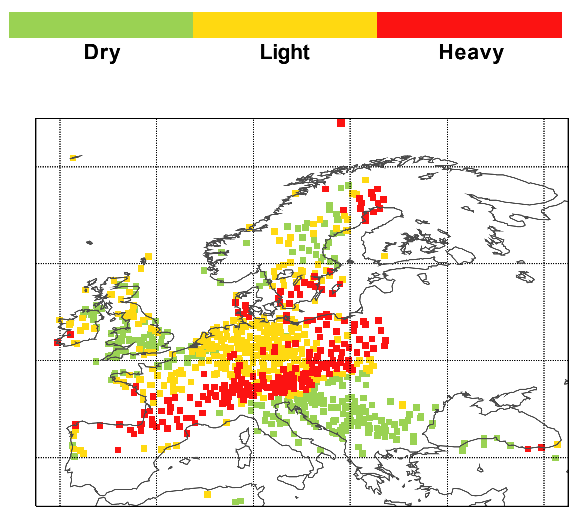
(c) Probability No Precip.



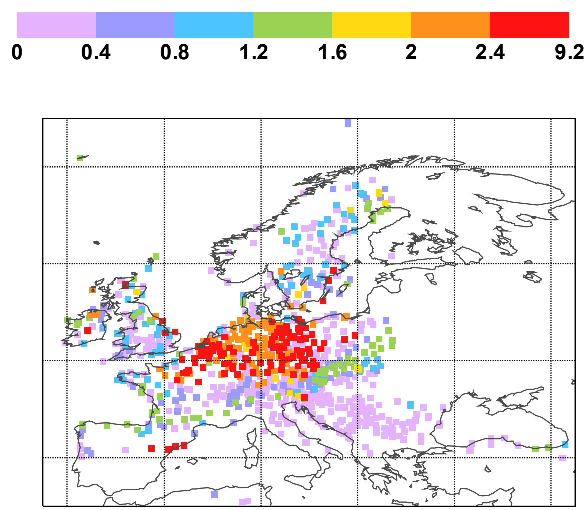
(d) Observed Category



(e) Forecast Category



(f) SLEEPS



Forecast penalised for not predicting heavy rain over (generally dry) northern Europe in summer

Summary

- **Diagnostics must become ever more powerful and precise**
  - Higher resolution
  - More observations
  - Better forecast models
  - Smaller signal-to-noise ratio
- **Initial Tendency diagnosis of model error**
  - Highlights Local causes – Hints at solutions
  - Higher statistical significance
  - Zonal-mean errors
  - Asian monsoon errors
  - MJO errors
- **Scale-dependent verification**
  - Separation of planetary waves (tropical origin?) and synoptic
- **Diagnostic Verification**
  - Careful design of scores can aid error detection: Precipitation