Diagnosis of data assimilation systems by using normal modes

Nedjeljka Žagar

University of Ljubljana

Thanks to: J. Tribbia, J. Anderson, K. Raeder (NCAR), D. Kleist (NCEP), L. Isaksen and D. Tan (ECMWF)

ECMWF Workshop on Diagnostics of data assimilation system performance
15-17 June 2009

On normal modes

Dickinson and Williamson (1972, JAS): normal modes are useful

1. for addressing the problem of initialization of numerical weather prediction (NWP) models,
2. for the identification of model modes which have significance for climate simulations, and
3. for comparison of amplitudes of model modes with those observed in the real atmosphere.

Ad 1) extensively used (NLNMI), with struggle in the tropics (e.g. Heckley 1982; Cats and Wergen, 1983; Wergen 1988;)

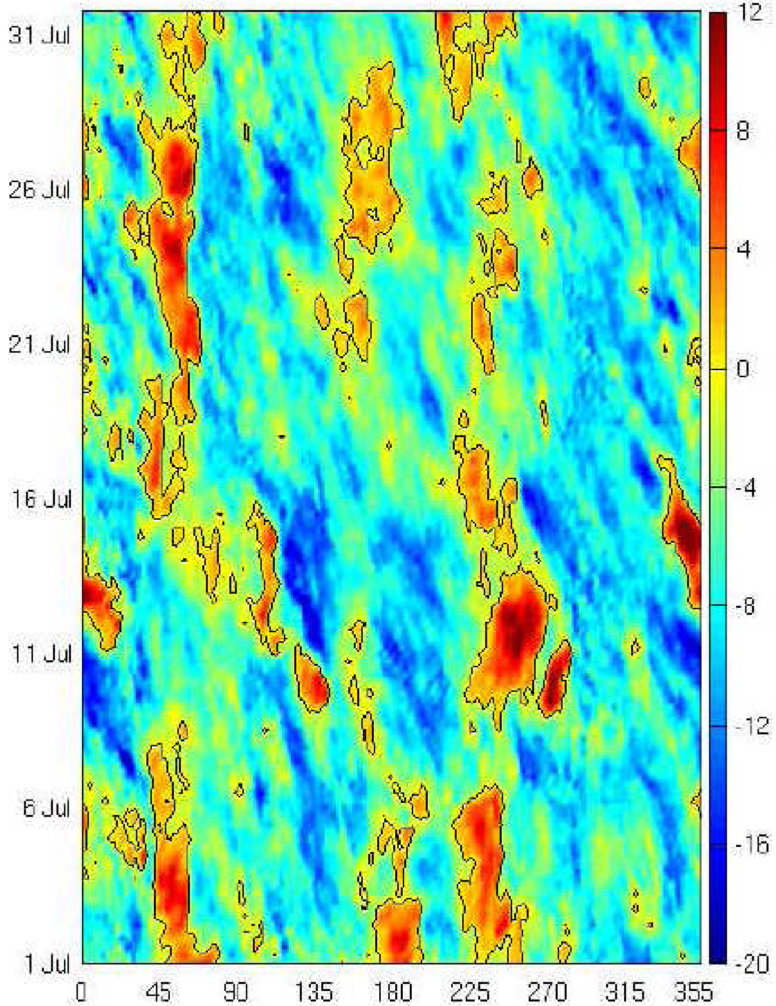
Ad 2-3) horizontal structures (Hough functions) used to identify large-scale Rossby waves (e.g. Hirooka, 2000; Madden, 2007); main application has been in the tropics (e.g. Wallace and Kousky, 1968); understanding large-scale tropical motions often relies on the normal modes on the equatorial β-plane.

Motivation

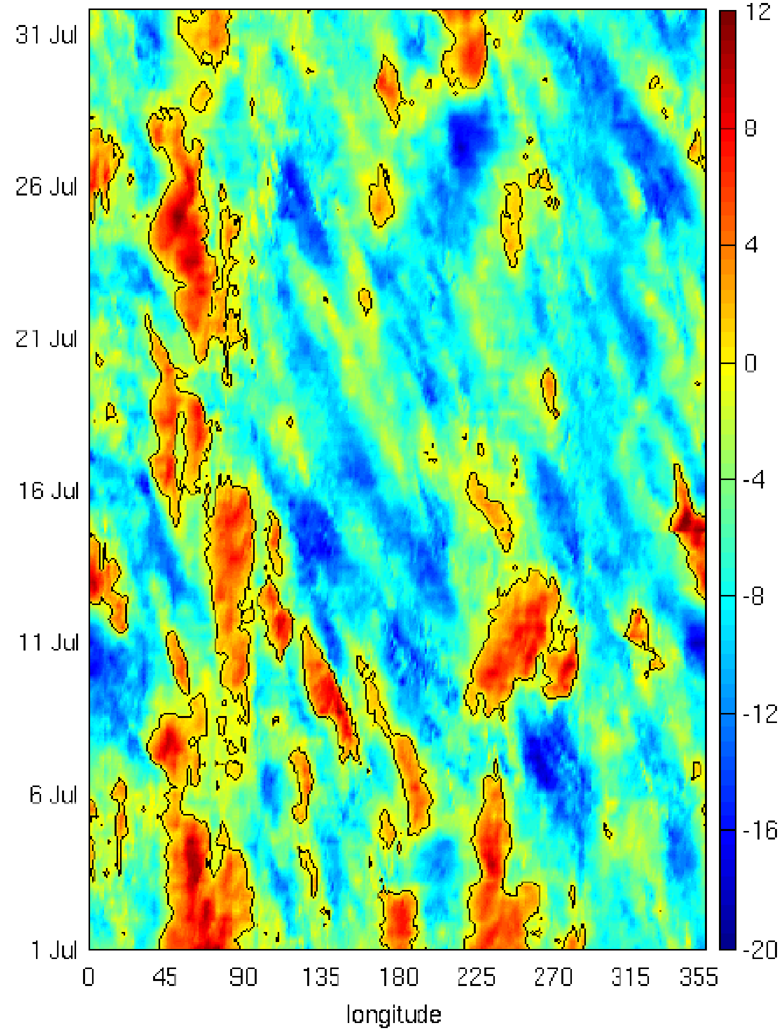
- The question of balance is fundamental to understanding how the atmosphere behaves
- Basic information source for understanding of global circulation and its variability have been the (re)analysis fields
- Divergent tropical circulations crucial, but unreliable from present (re)analysis. In mid-latitudes, inertio-gravity motions important for indirect circulation.
- State of the art NWP and climate models represent divergent motions much better. 4D-Var and abundant satellite observations provided lots of improvements of the tropical analyses and forecasts.
- Large-scale equatorial waves in recent years have been diagnosed from different mass-field observations. Quantification of their variance and dynamical relevance still not understood well.

Zonal winds in July 2007 in operational analyses of ECMWF and NCEP

NCEP: u wind, ~370 hPa, along 5N

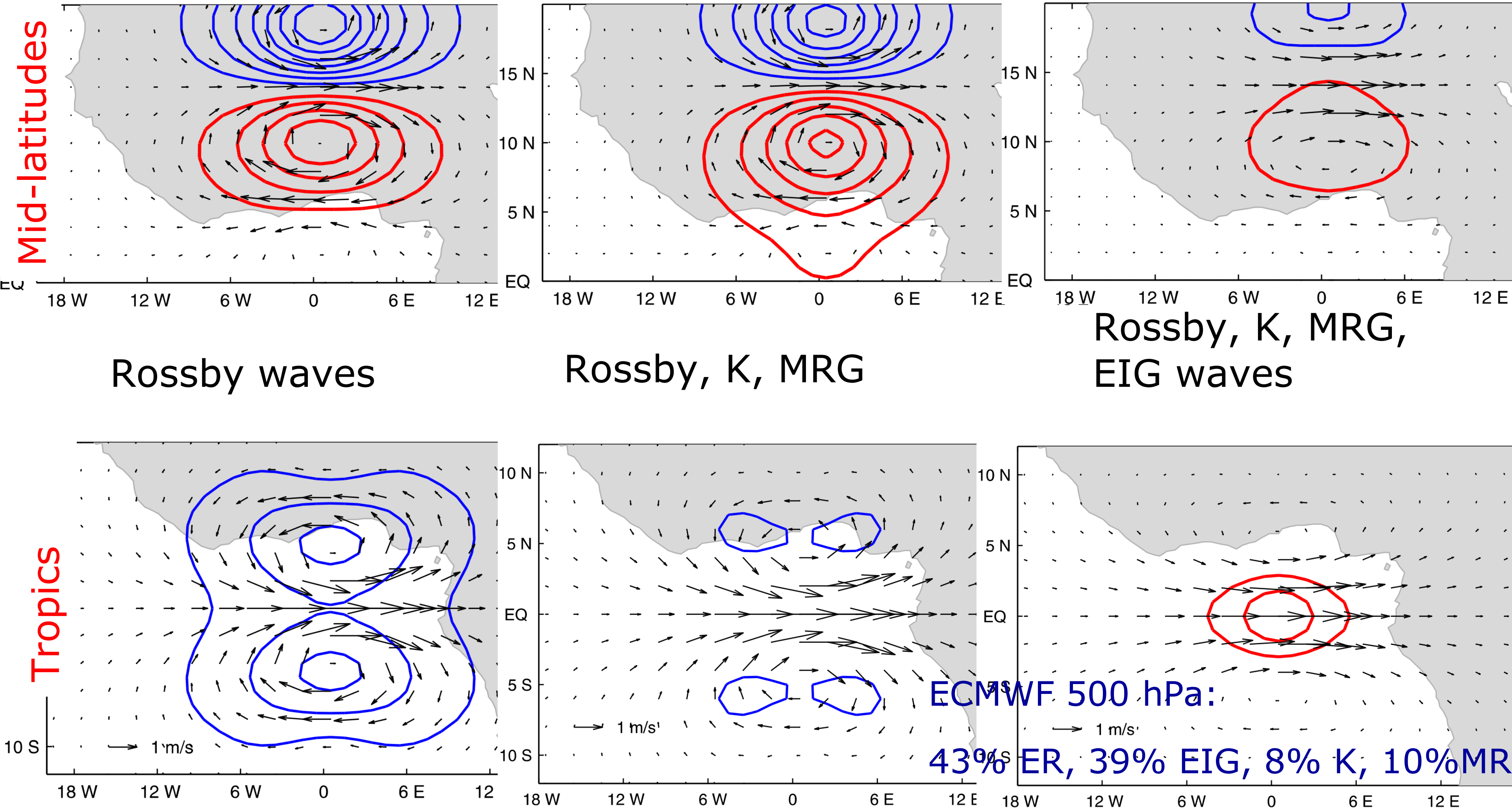


ECMWF: u wind, ~370 hPa, along 5N



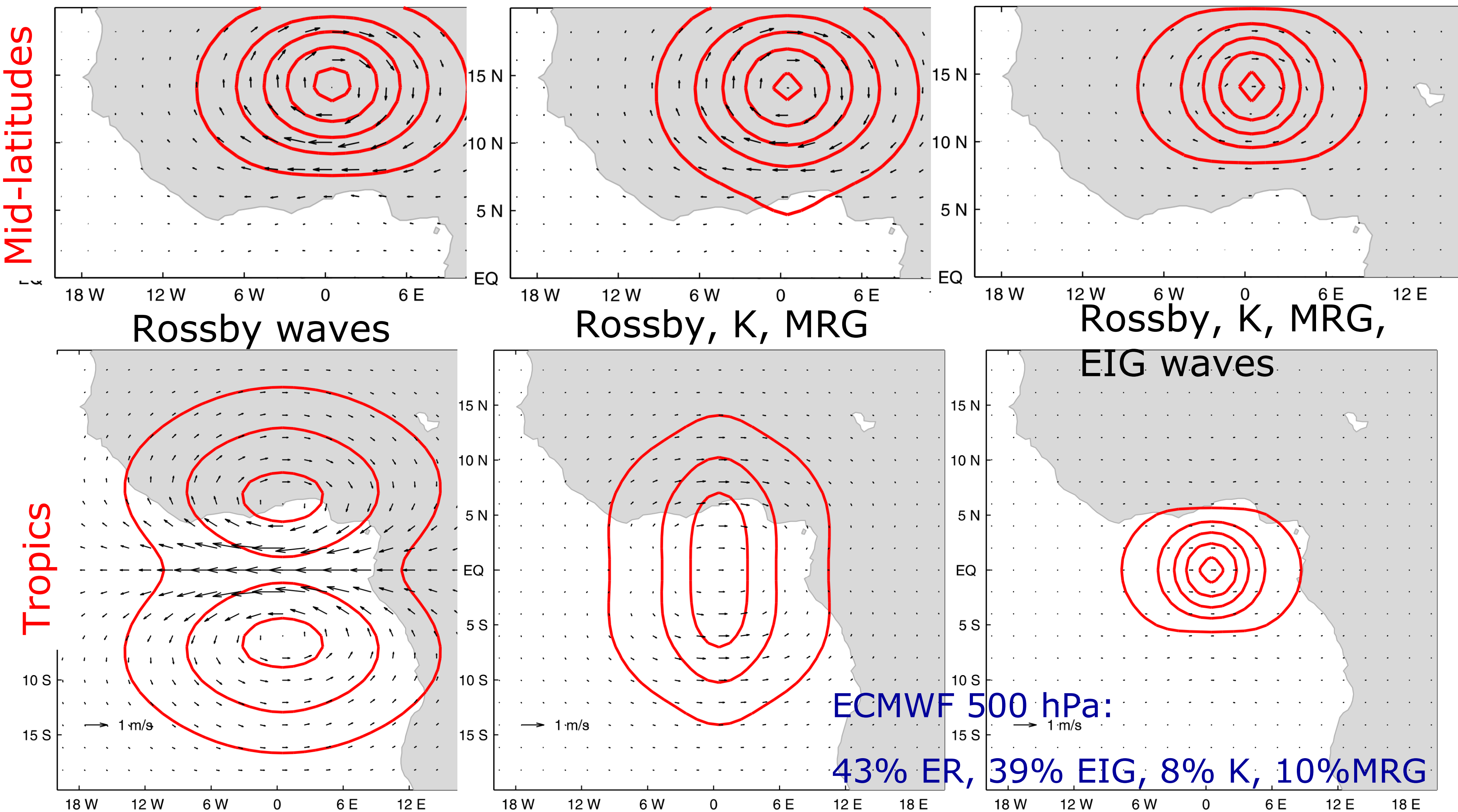
Multivariate 3D-Var with a simple model

Analysis increments due to a single zonal wind observation

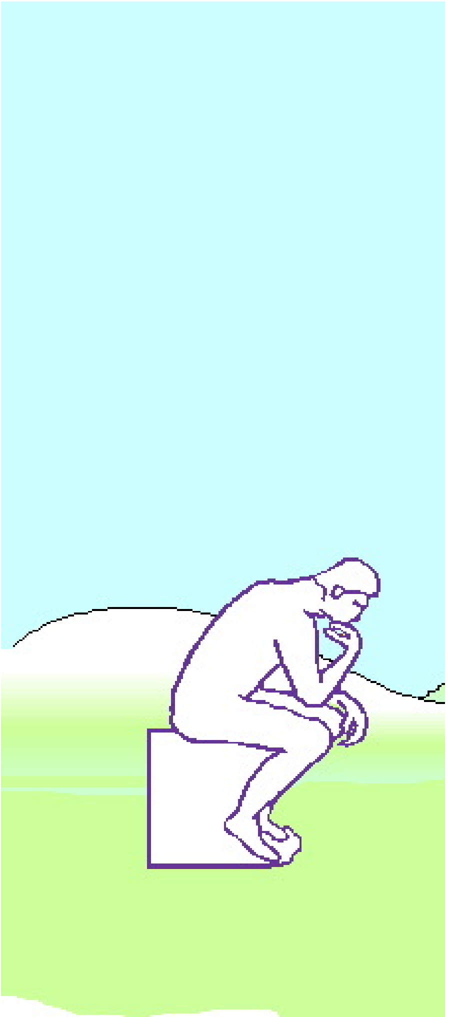


Multivariate 3D-Var with a simple model

Analysis increments due to a single zonal temperature observation

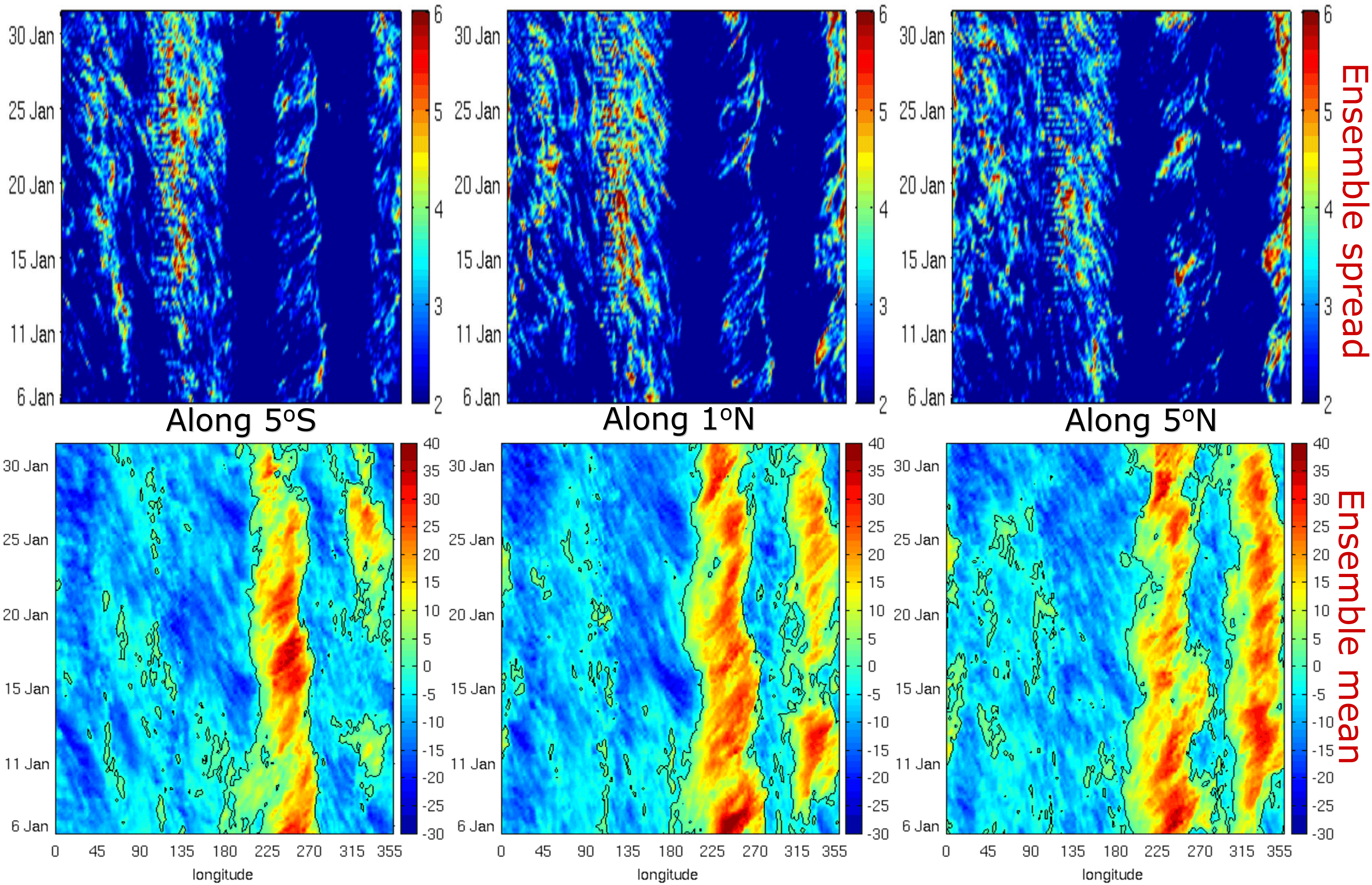


Questions

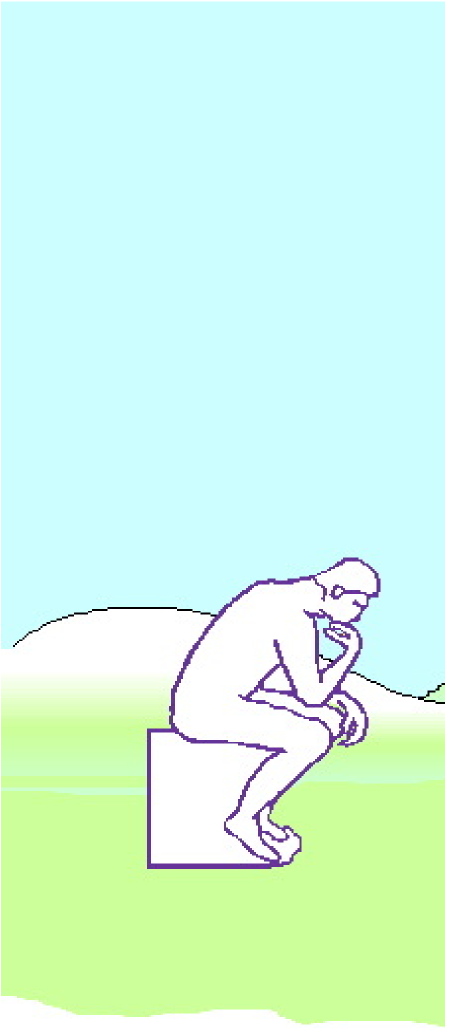


- How large part of the global short-range forecast errors pertains to the divergent motion i.e. inertio-gravity (IG) waves?
- How important are the large-scale tropical waves for the global data assimilation? How are the tropical forecast errors in the IG motion spread across the scales, time and motion types?
- How is the analysis uncertainty split between the balanced (ROT) and IG motion? How is it dependent on the modelling system and the assimilation methodology (4D-Var versus the ensemble Kalman filter (EnKF)?
- What is the potential of the EnKF due to flow-dependent background-error covariances in comparison to 4D-Var, especially in the tropics?

Flow dependency of the 6-hour zonal wind errors in the ensemble data assimilation system DART/CAM



Questions, cont.



- How large part of the global atmospheric flow pertains to the divergent motion i.e. inertio-gravity (IG) waves?
- How uncertain is this result i.e. how dependent is the distribution of the balanced/unbalanced energy on the analysis system and the assimilation methodology?
- What part of the flow variance on large scales belongs to the tropical waves? What is the relative importance of Kelvin wave, mixed Rossy-gravity waves, other equatorially trapped waves for the tropical flow variance?

Analysis systems

Four datasets for July 2007, every 6 hrs

ECMWF: operational analyses, 12-hour 4D-Var system, Cycle 32r2, T799 interpolated to N64 grid, 91 vertical level up to 0.01 hPa, DF as a weak constraint (little impact).

NCEP: operational analyses, grid-point 3D-Var system, T382 interpolated to N64 grid, 64 vertical levels up to 0.32 hPa, TL NNMI initialization.

DART/CAM: ensemble mean from the DART system, CAM version 3.1, T85 horizontal resolution, 26 vertical levels up to 3.5 hPa. Limited number of observations (conventional observations and AMVs), no initialization.

NCEP/NCAR reanalyses from the NCAR mass archive on N47 grid: 3D-Var system, T62 horizontal resolution, 28 vertical levels up to 2.7 hPa. Assimilattion of retrievals and no initialization.

Ensemble assimilation datasets

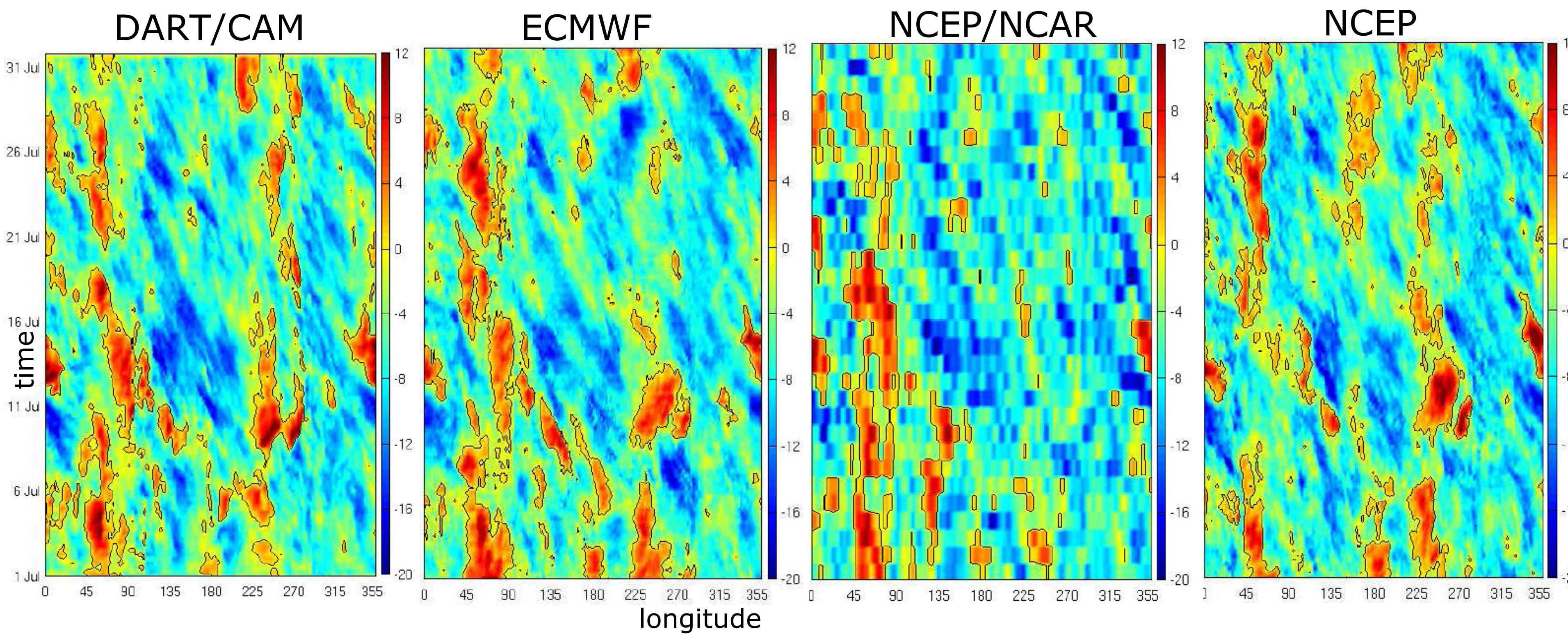
Two ensembles of global analyses and forecasts for July 2007:

- NCAR EnKF system DART/CAM (http://www.image.ucar.edu/DAReS/DART/): 80-member CAM ensemble on the horizontal resolution T85, 26 vertical levels up to 3.5 hPa. Limited number of observations (conventional observations and AMVs) in the troposphere. No moisture observation. The covariance localization and a time constant, spatially varying covariance inflation are applied.

- ECMWF 4D-Var ensemble: 21-member ensemble with 12 hour 4D-Var and model cycle 32r3. Operational 91 levels up to 0.01 hPa and new physical parameterizations which resulted in increased spread. Prepared by L. Isaksen and D. Tan of ECMWF.

Both datasets interpolated to N64 grid horizontally on all model vertical levels.

Tropical winds in 4 analysis datasets in July 2007 at 370 hPa along 5°N



3D normal mode expansion: discrete form

Following Kasahara and Puri (1981)

$$\mathbf{X} = (u, v, P/g)^T \qquad P = gz + RT_o(\sigma)q \qquad q = \ln(p_s)$$

$$\left(u_m, v_m, g^{-1}P_m\right)^T = \mathbf{S}_m^{-1}\sum_{\sigma=1}^{N_\sigma}\Pi_m(\sigma)\cdot\mathbf{X}(\lambda,\varphi,\sigma)$$

$$\chi_\upsilon = \frac{1}{2\pi}\int_0^{2\pi}\int_{-1}^{1}\left(u_m, v_m, g^{-1}P_m\right)^T\left(\mathbf{H}_\upsilon\right)^* e^{-ik\lambda}\,d\lambda\,d\mu$$

2-step forward projection

$\upsilon = (k, n, m, p)$ modal index

m - vertical mode index

k - zonal mode index

n - meridional mode index

p – motion type index

orthogonal 3D expansion basis:

$$\left\langle \Pi_m \mathrm{H}_{k'n'm'}, \Pi_{m'} \mathrm{H}_{k'n'm'} \right\rangle = \delta_{kk'}\delta_{nn'}\delta_{mm'}$$

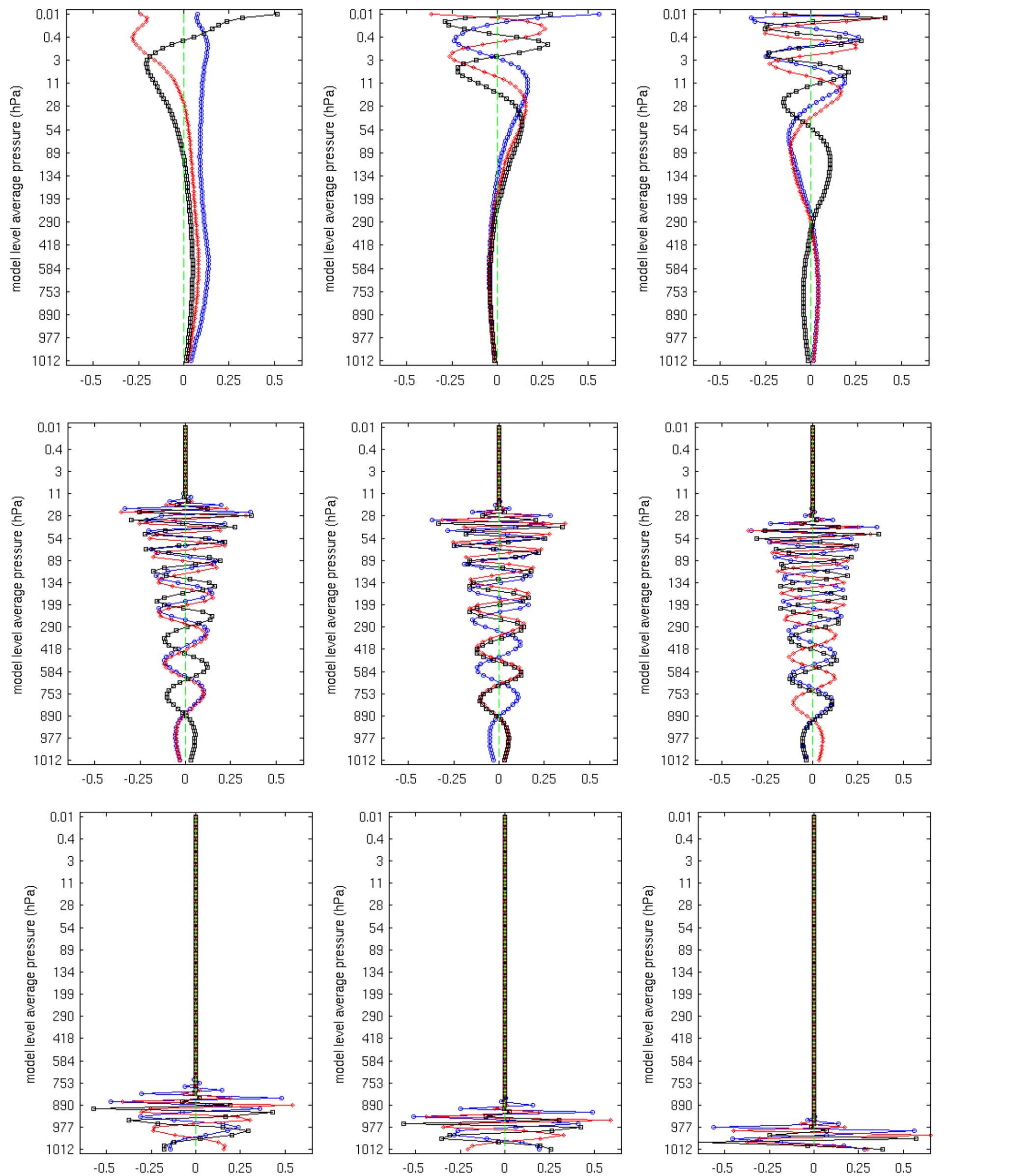
Energy partitioned into balanced (ROT) and inertio-gravity (IG) motions (eastward-EIG and westward-WIG) for each vertical mode

Vertical eigenfunctions for ECMWF system: finite-difference solution

Vertical structure equation:

$$\frac{d}{d\sigma}\left(\frac{\sigma g}{R\Gamma_o}\frac{d\Psi}{d\sigma}\right)+\frac{1}{H_{eq}}\Psi=0$$

H_{eq} - "equivalent depth"



Input information: vertical discretization, temperature profile, stability profile Γ_o

H_{eq} from 10 km to 8 mm
First 18 with H_{eq} > 100 m
Modes 19-38 between 100 m and 10 m, 39-66 between 10 and 1 m, and 66-91 below 1 m.

3D normal mode expansion: discrete form

Inverse projection:

common expansion coefficient

$$\mathbf{X}(\lambda,\theta,\sigma)=\sum_{m=1}^{N_m}\mathbf{S}_m\left(\sum_{p=1}^{3}\sum_{n=0}^{N_n}\sum_{k=-N_k}^{N_k}\chi_{knm}^{p}\mathbf{H}_{knm}^{p}(\sigma)e^{ik\lambda}\right)\Pi_m(\sigma)$$

normalization matrix

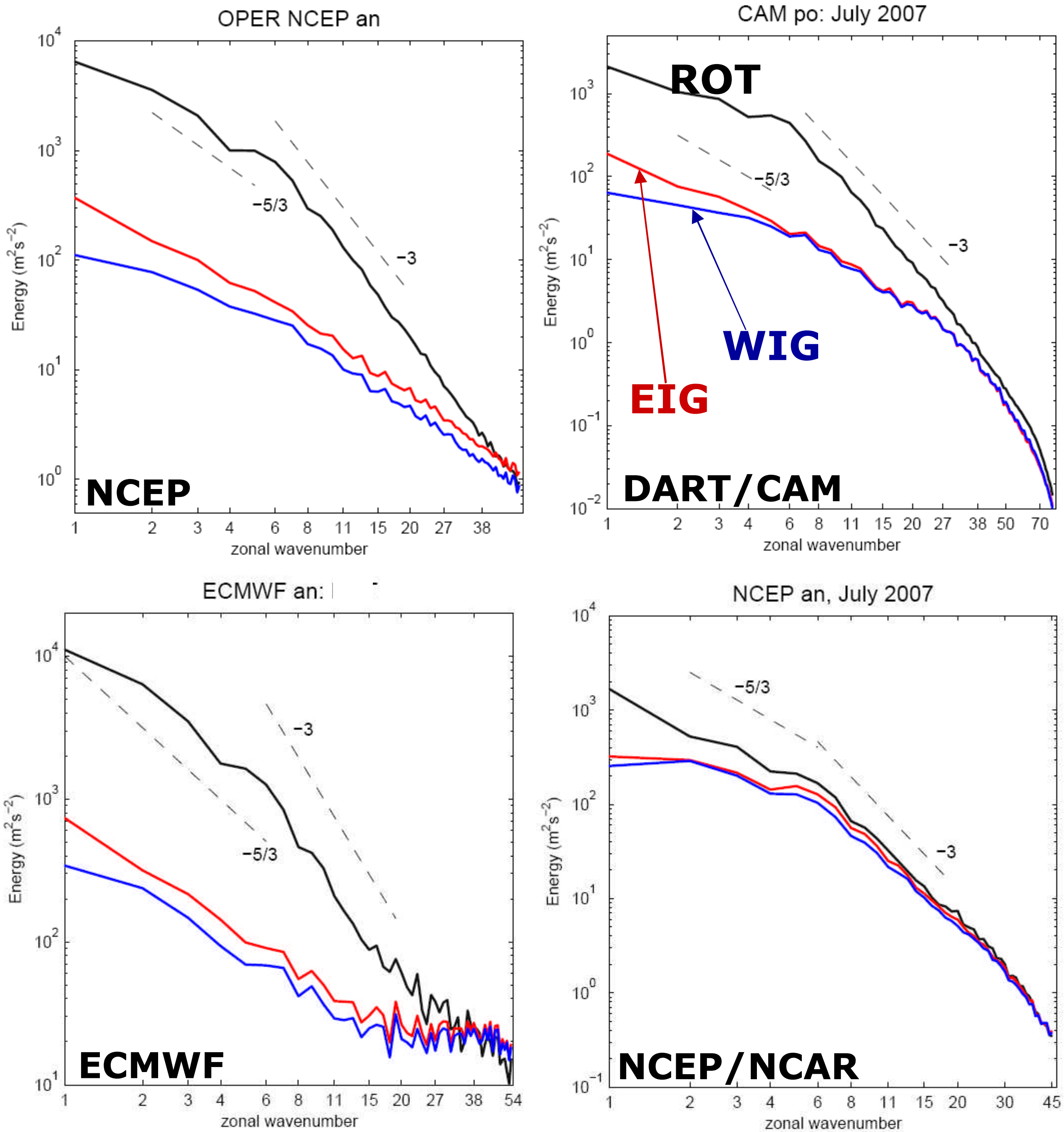
Hough functions

vertical eigenfunctions

Basic idea: select the expansion basis which provides the best fit (best correlation and variance fit to the input grid-point fields) ⇔ tuning of the truncation parameters N_k, N_n , N_m

Total energy per mode (k,n,m) per unit mass: $$\sum_k\sum_n\sum_m gH_{eq}\left|\chi_{knm}\right|^2$$

Energy spectra for July 2007: analyses inter-comparison

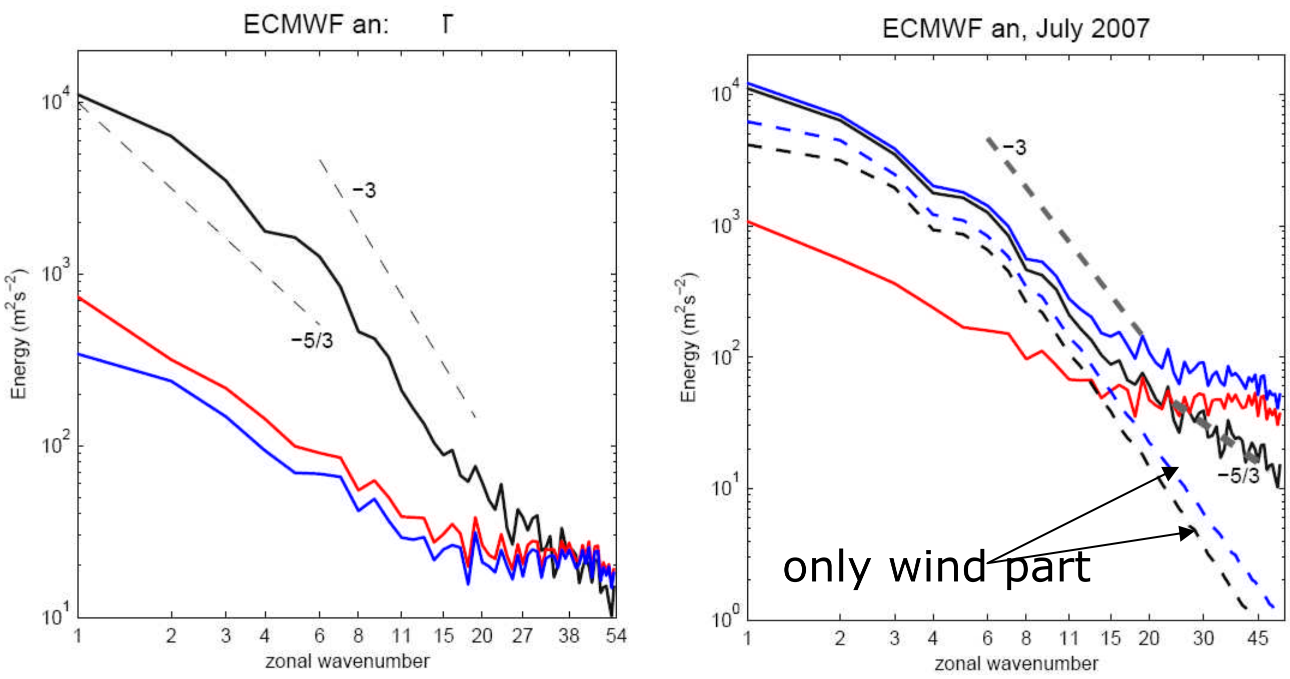


$$\sum_m \sum_n gH_{eq} \left|\chi_{knm}\right|^2$$

NCEP/NCAR

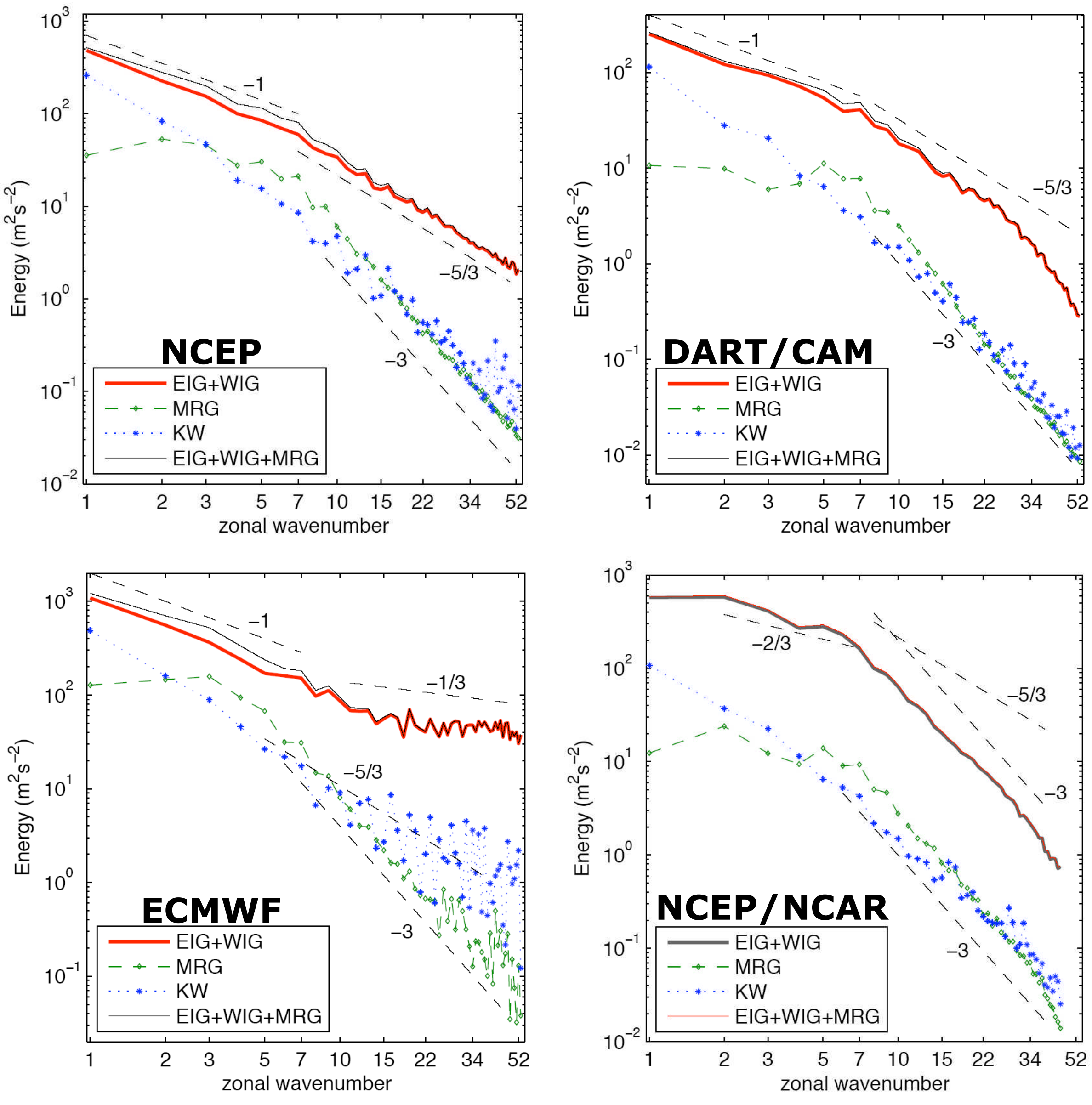
Significant IG motions also in SH mid-latitudes

Noisy spectra of ECMWF at smaller scales



Impact of the noisy surface pressure and orography fields when interpolated from T799 to the Gaussian grid N64 and vertical σ levels

Average energy spectra for Kelvin waves, mixed Rossby-gravity (MRG) and total inertio-gravity motions (EIG+WIG)



Energy percentages in various motions and analysis systems

Averaging over one month

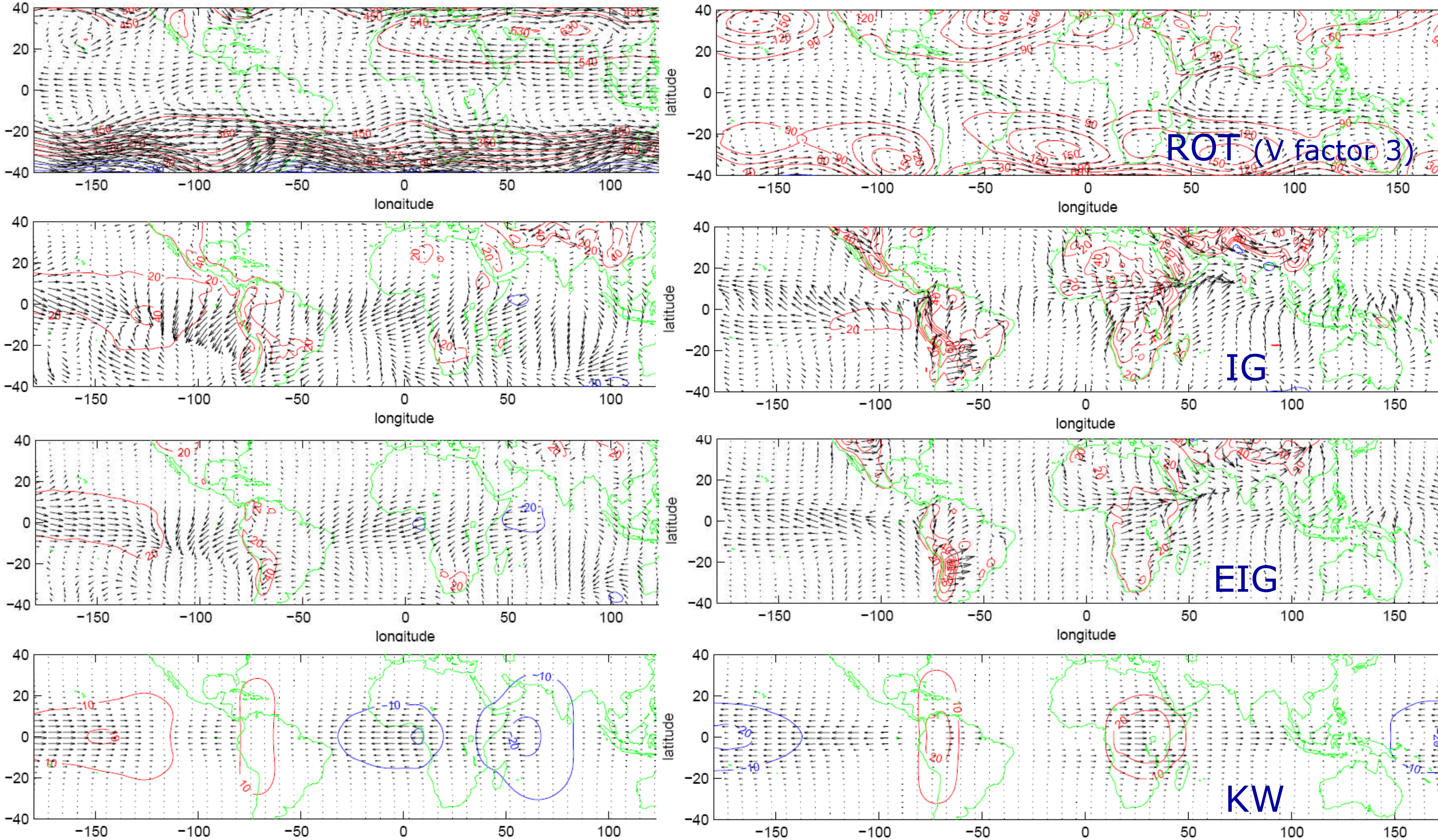
$$\sum_k \sum_n \sum_m gH_{eq} \left| \chi_{knm} \right|^2$$

	Motion type					
	$k > 0$			all k		
Dataset	ROT	EIG	WIG	ROT	EIG	WIG
DART/CAM	87.9	7.4	4.7	98.6	0.7	0.6
ECMWF	85.3	8.6	6.1	98.2	1.0	0.8
NCEP/NCAR	54.2	24.5	21.3	95.9	2.2	1.9
NCEP	91.3	5.8	2.9	98.7	0.8	0.5

	Dataset			
Energy ratio (%)	**DART/CAM**	**ECMWF**	**NCEP/NCAR**	**NCEP**
MRG : (IG+MRG)	8	12	4	15
KW : (IG+MRG)	20	17	7	25
MRG : (WIG+MRG)	18	25	7	34
KW : EIG	36	19	13	45

Mean tropical circulation at two levels

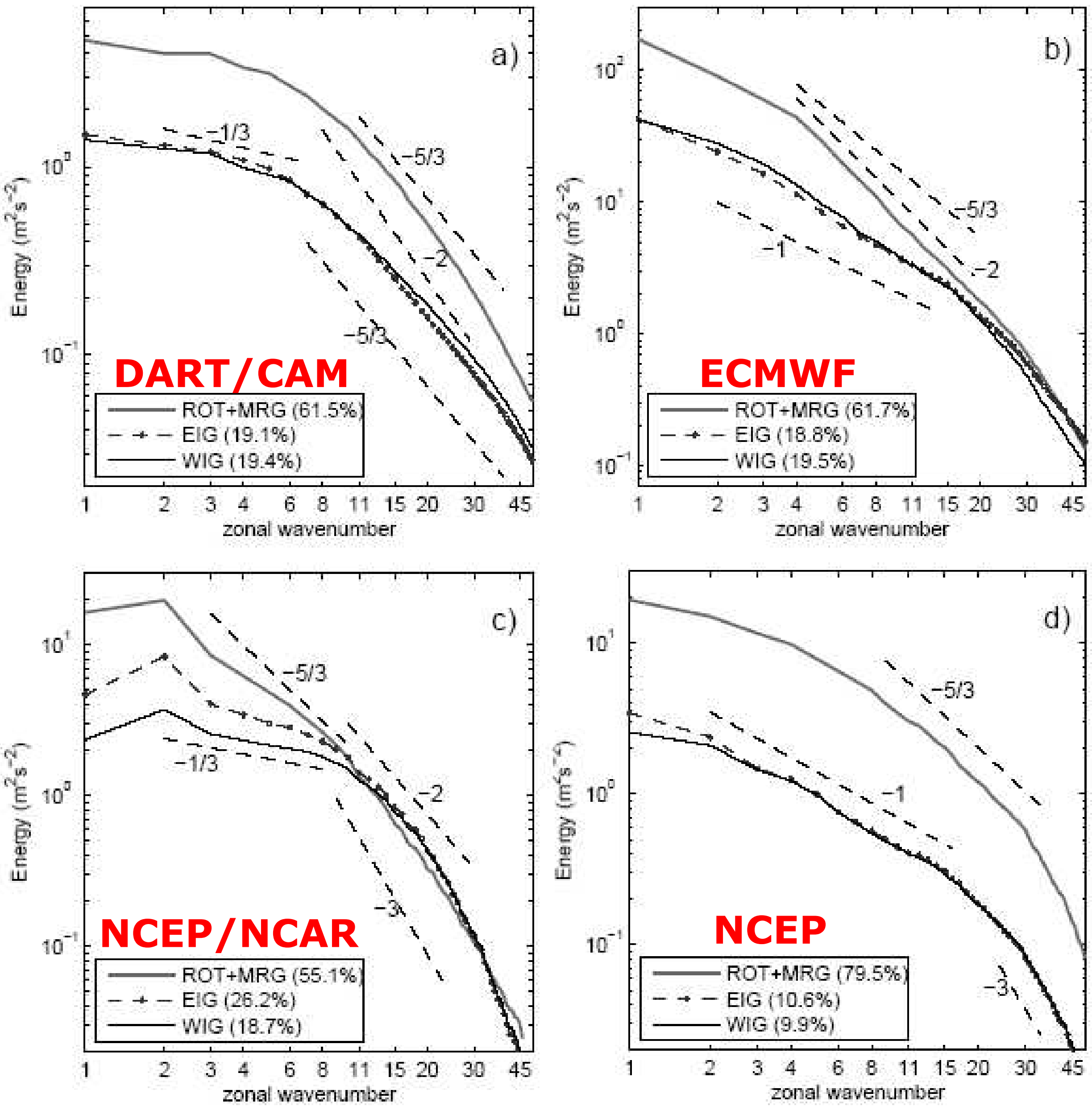
Level 15 (~269 hPa) | Level 23 (~868 hPa)



Tropics as envisaged by Gill (1980)

Analysis increments:

$$E_{\upsilon}^{in}(t) = gH_{eq,\upsilon}\left[\left(\chi_{\upsilon}^{an} - \chi_{\upsilon}^{bg}\right)\left(\chi_{\upsilon}^{an} - \chi_{\upsilon}^{bg}\right)^{*}\right]$$

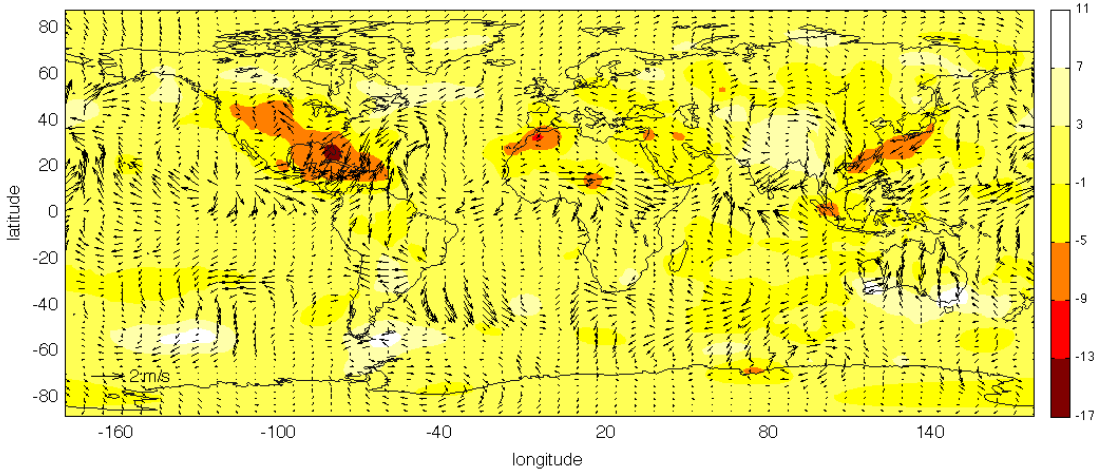


Average spectra of analysis increments

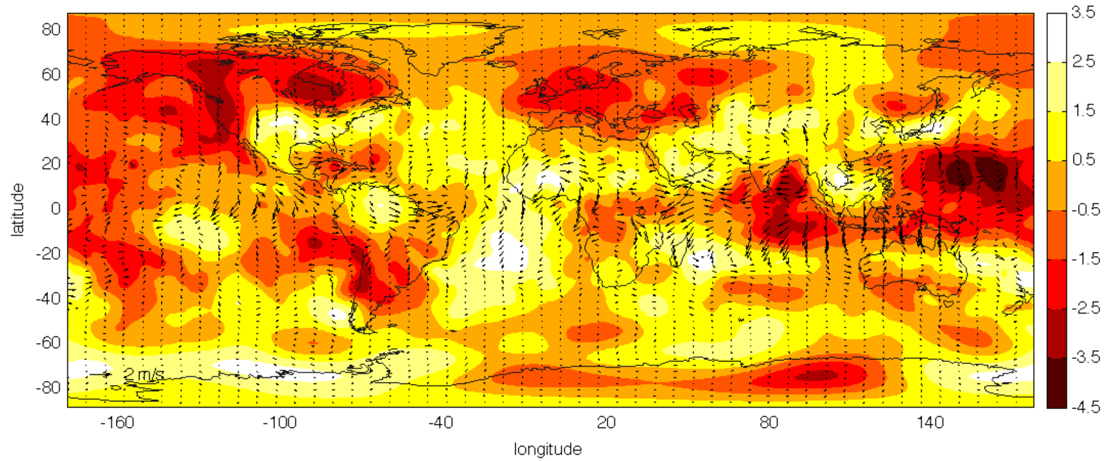
$$\sum_{m=1}^{N_m}\sum_{n=0}^{N_n}\overline{E_{\upsilon}^{in}(t)}$$

NCEP/NCAR:
% same as in the full wave fields, the only system where EIG part dominates over WIG also in increment fields

Time-averaged analysis increments (biases) for July 2007



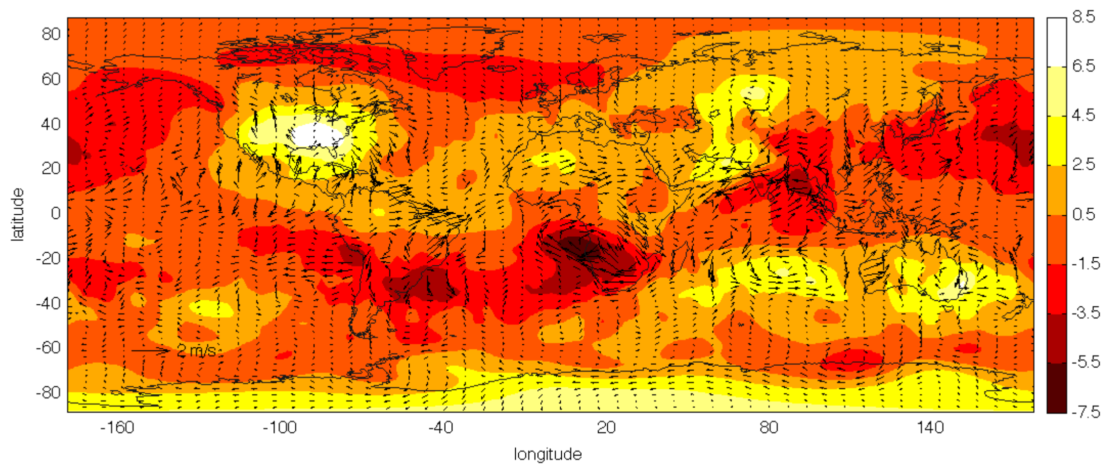
DART/CAM (level 13)



ECMWF (level 50)

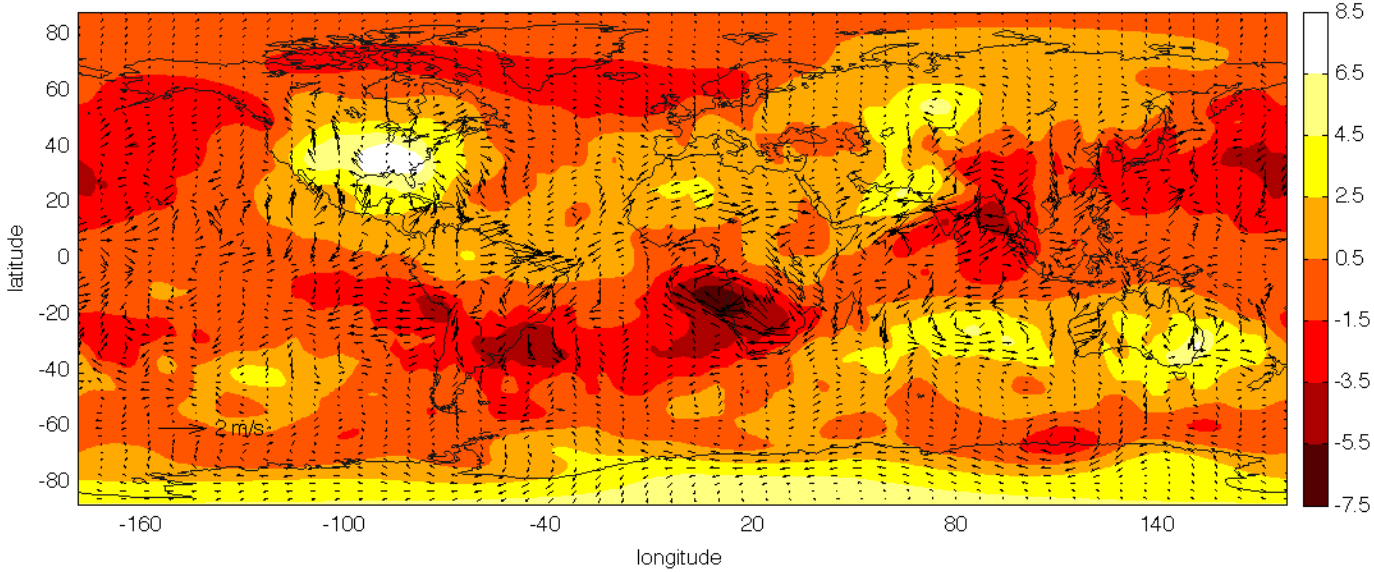
Model level close to 200 hPa

$$\overline{\Delta \chi} = \frac{1}{N}\sum_{t=1}^{N}\left[\chi_{\upsilon}^{an}(t) - \chi_{\upsilon}^{bg}(t)\right]$$



NCEP (level 30)

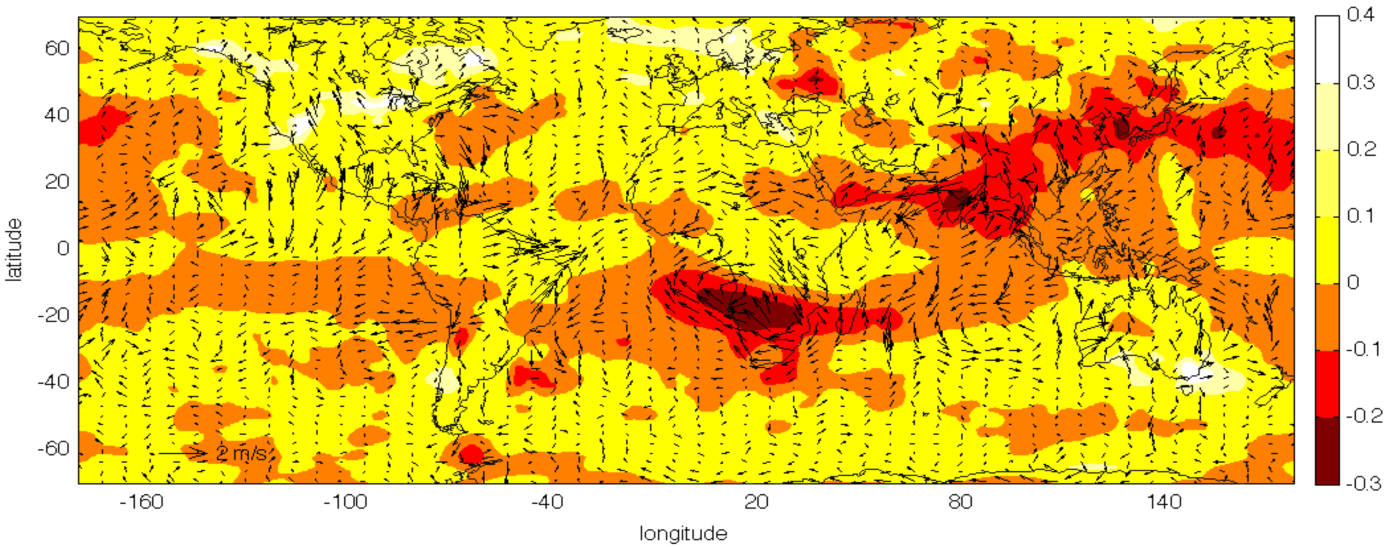
Verification: Level 30 NCEP



Modal space diagnosis:

(winds and *P*/g variable)

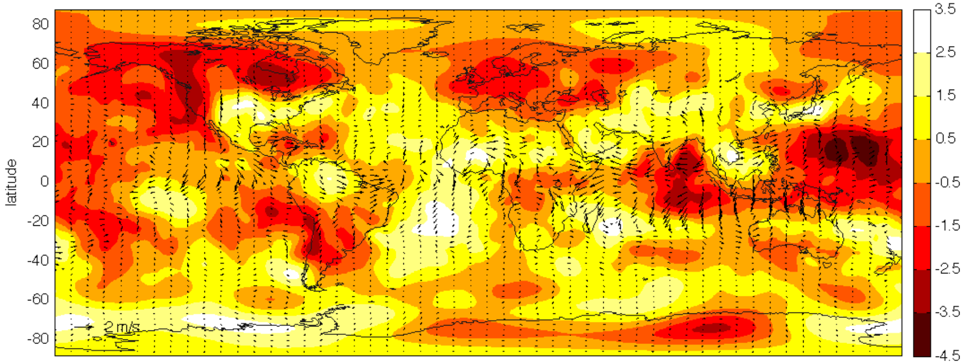
$$\overline{\Delta\chi} = \frac{1}{N}\sum_{t=1}^{N}\left[\chi_{\upsilon}^{an}(t) - \chi_{\upsilon}^{bg}(t)\right]$$



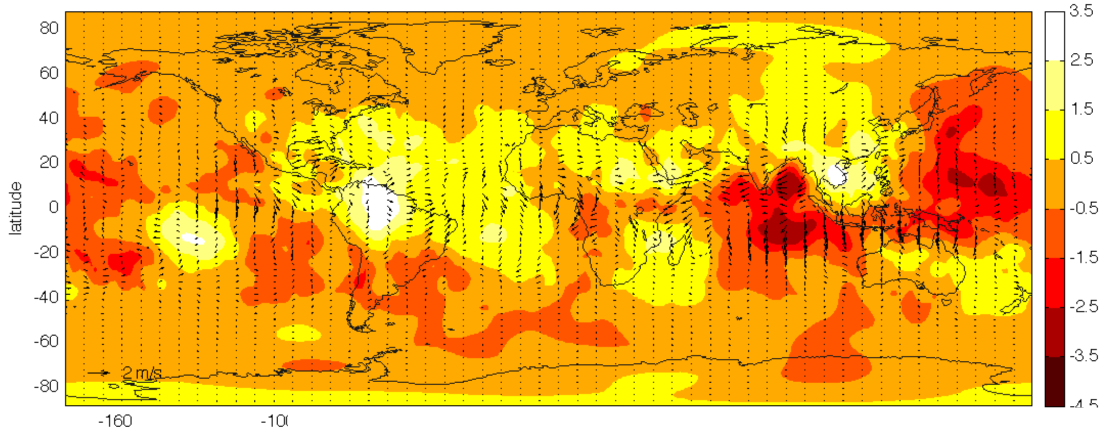
Physical space diagnosis:

(winds and T variable)

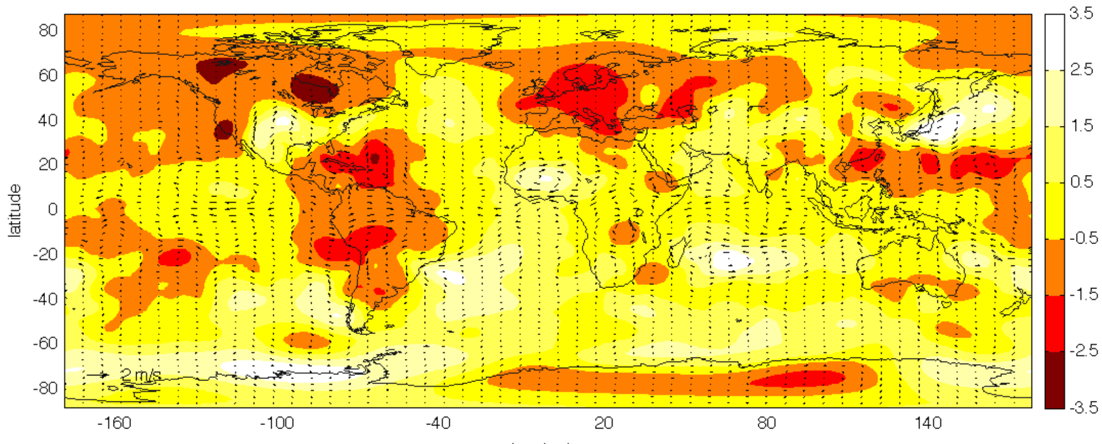
Biases split in ROT and IG parts: Level 50 ECMWF



All

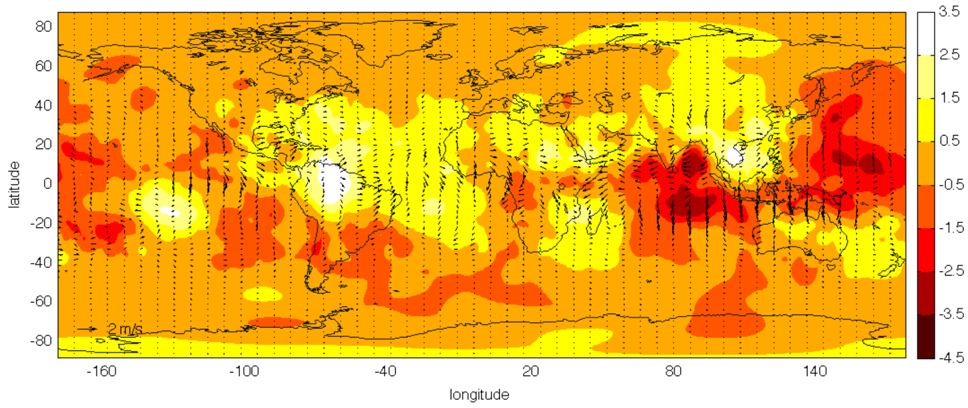


Inertio-gravity

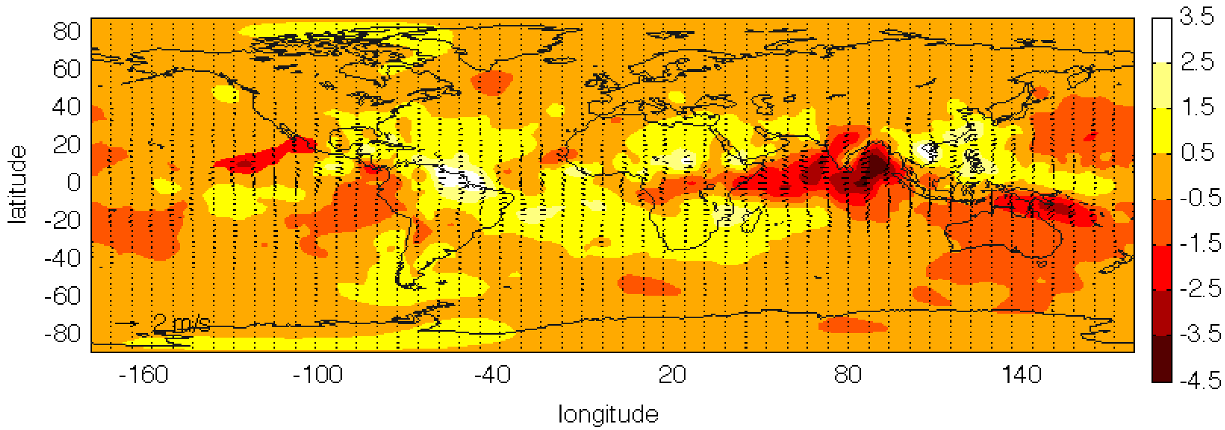


Balanced

IG split in EIG, KW parts: Level 50 ECMWF



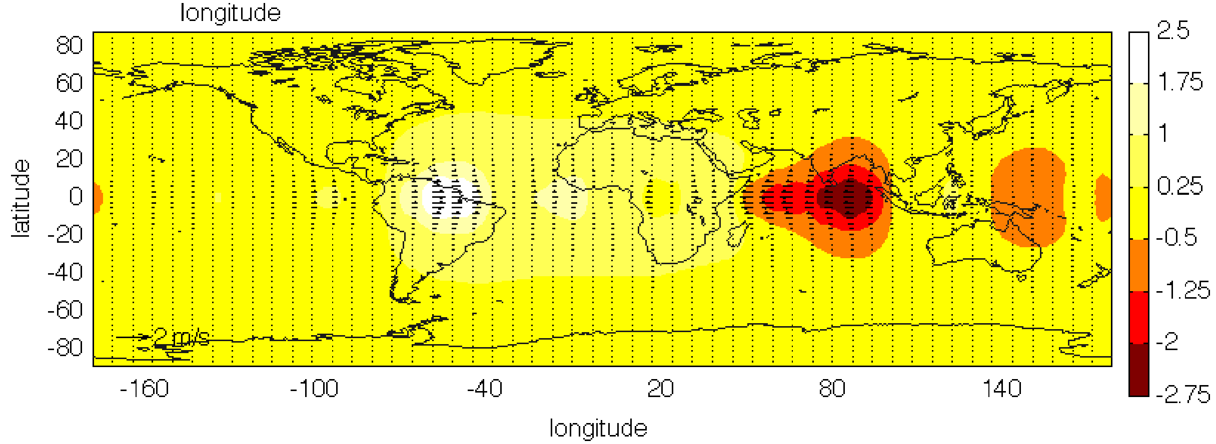
Inertio-gravity



Eastward Inertio-gravity

Kelvin wave

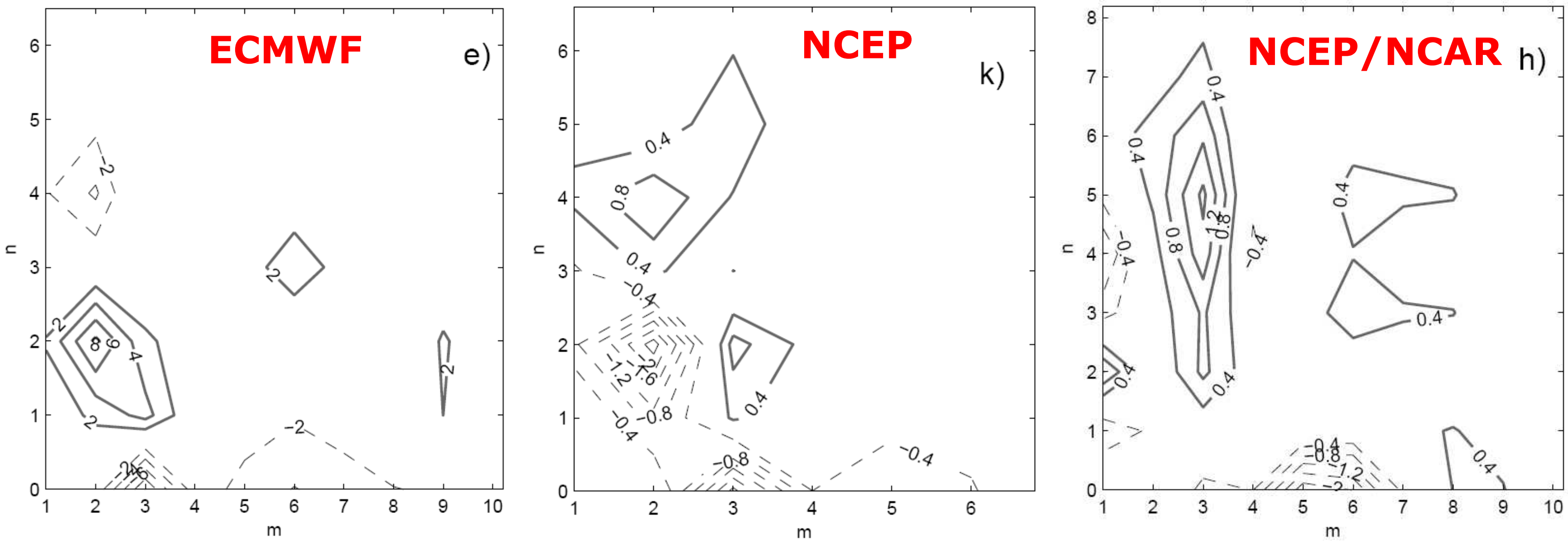
(Level 53)



Time-averaged energy difference

$$\overline{\Delta E_{\upsilon}} = \sum_{x} gH_{eq,\upsilon} \overline{\left[\chi_{\upsilon}^{an} \left(\chi_{\upsilon}^{an} \right)^{*} - \chi_{\upsilon}^{bg} \left(\chi_{\upsilon}^{bg} \right)^{*} \right]}$$

any of (*k*,*m*,*n*)



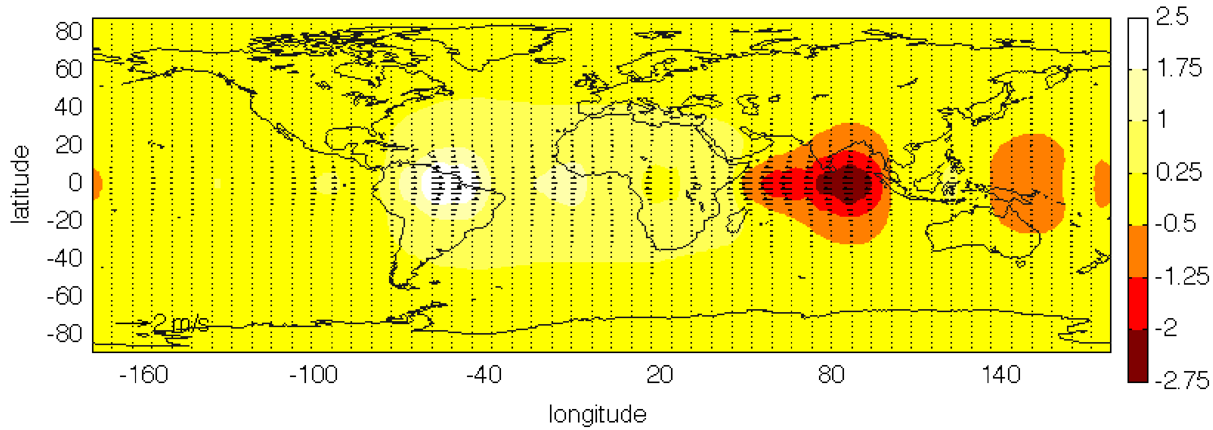
Biases in 3 systems similar for KW

Not in DART/CAM as few observations in the troposphere and no obs assimilated in the stratosphere

ECMWF system: Kelvin wave signal in July 2007

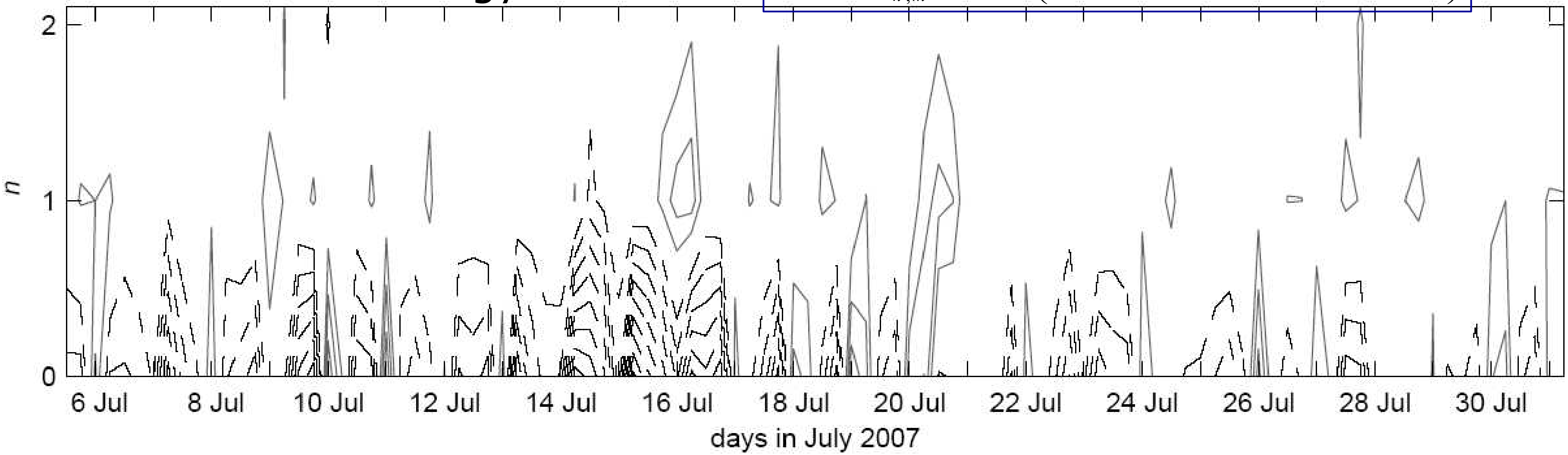
Time-averaged KW increment

(Level 53, ~240 hPa)

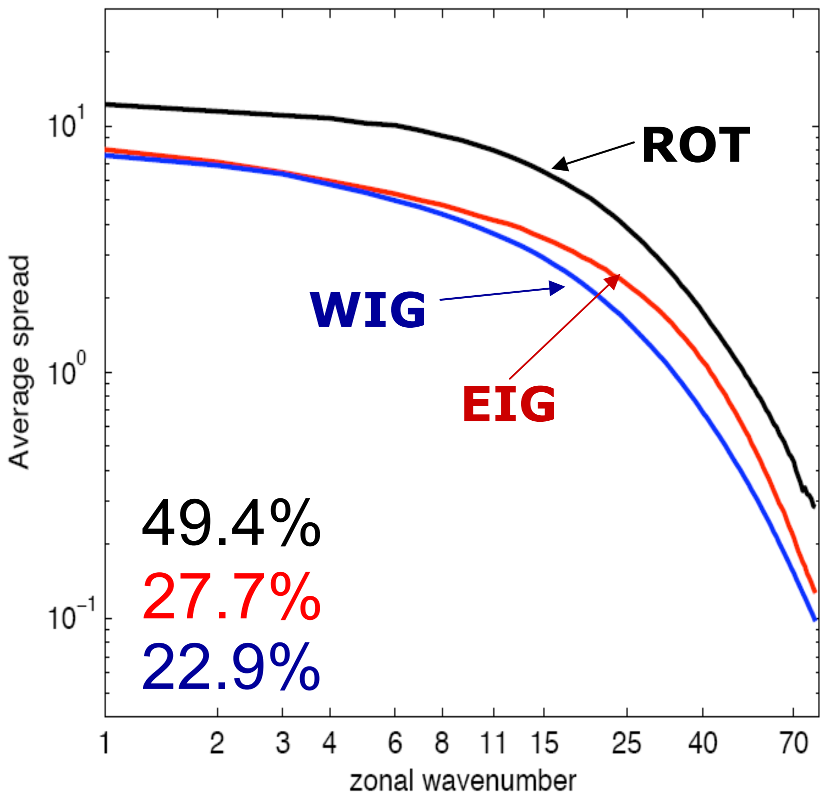


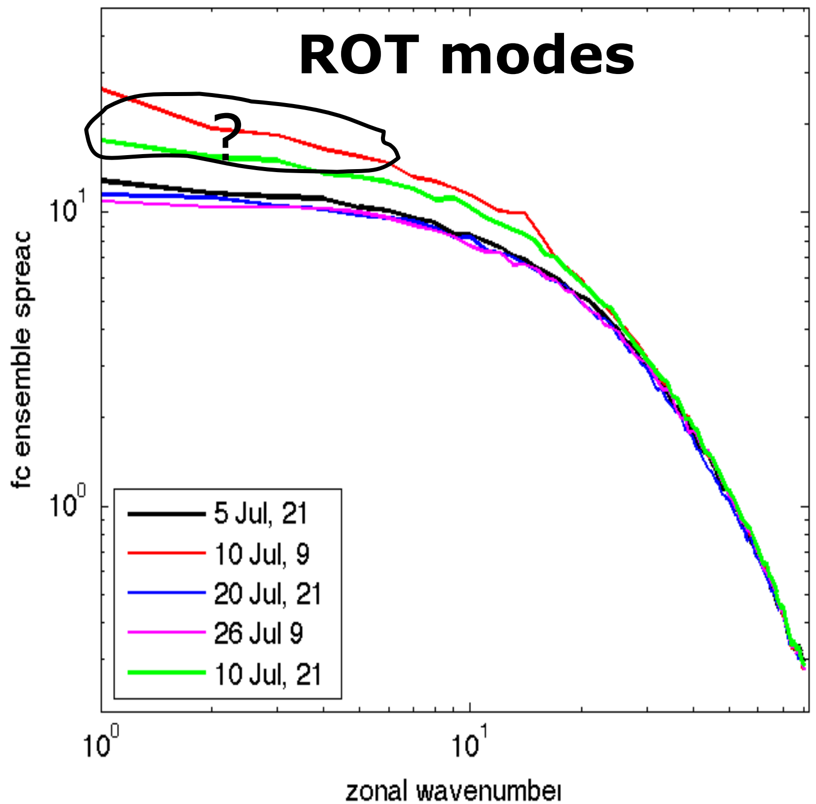
EIG increment energy evolution:

$$\Delta E_{\upsilon} = \sum_{k,m} gH_{eq,\upsilon} \left(\chi_{\upsilon}^{an} \left(\chi_{\upsilon}^{an} \right)^{*} - \chi_{\upsilon}^{bg} \left(\chi_{\upsilon}^{bg} \right)^{*} \right)$$



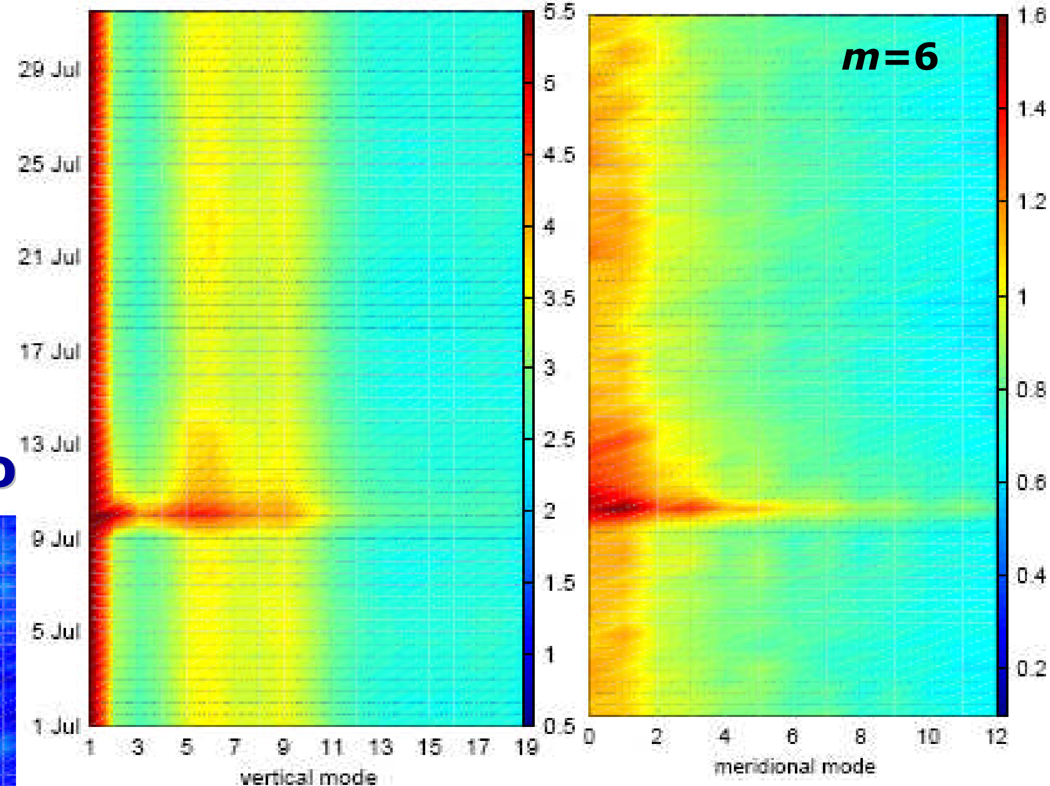
ECMWF 3-h fc ensemble: average spread vs. its variability



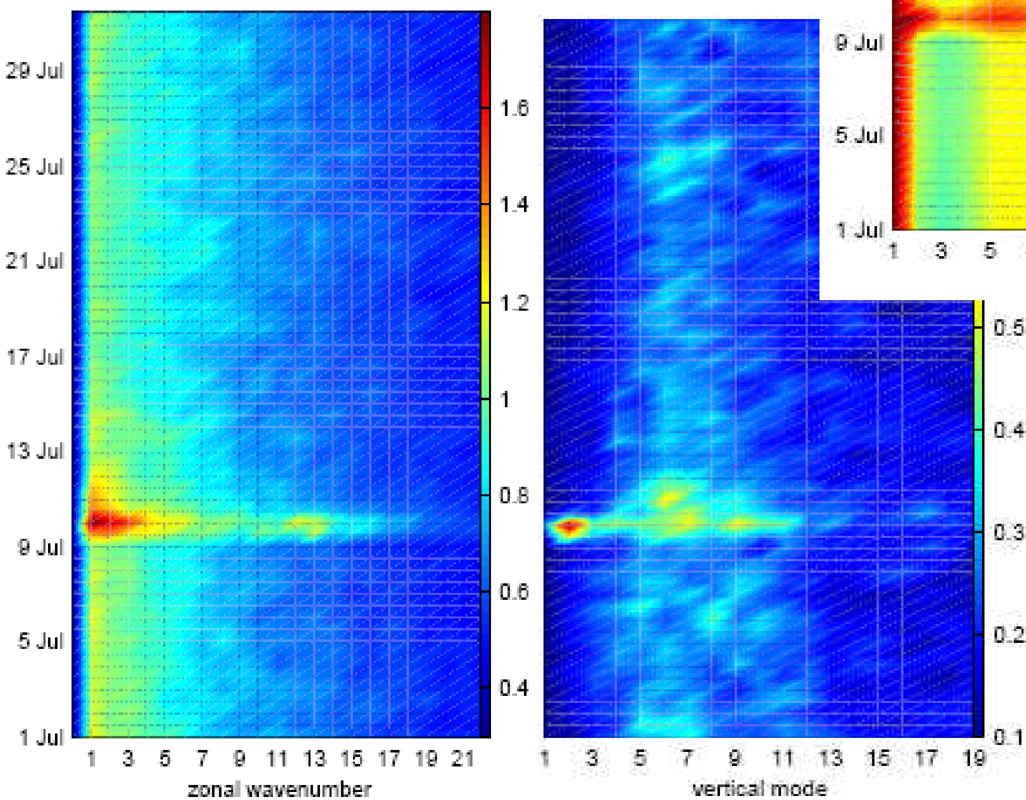


Something happened to the spread on 10 July, about 4 days before something happened to the increment

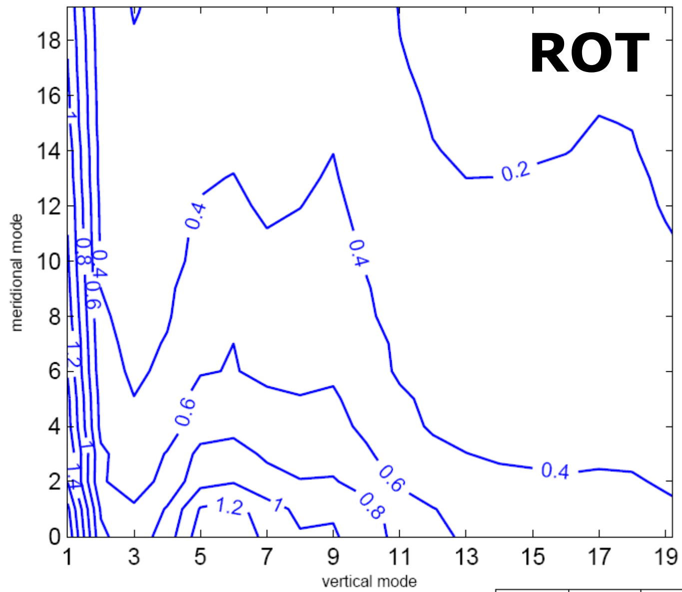
ROT modes fc-error evolution

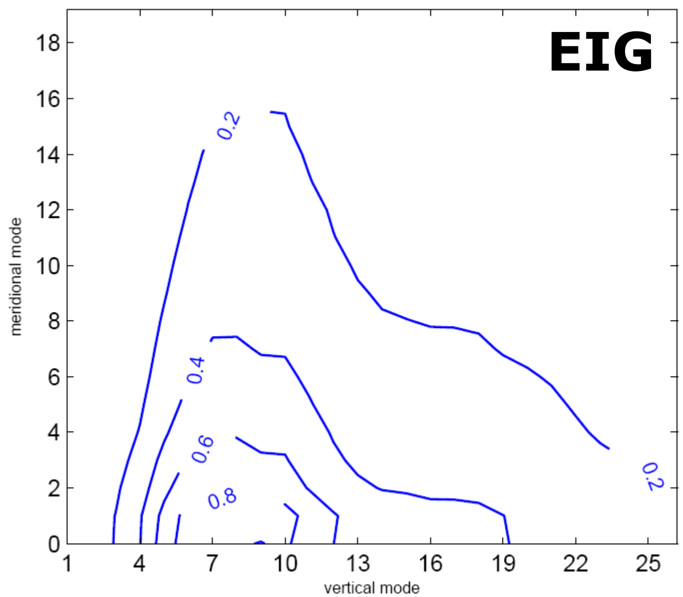


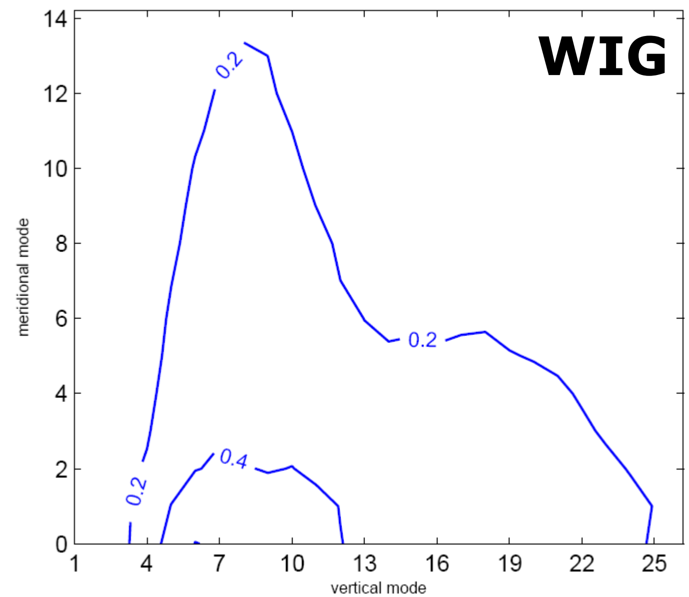
Kelvin wave fc-error evo



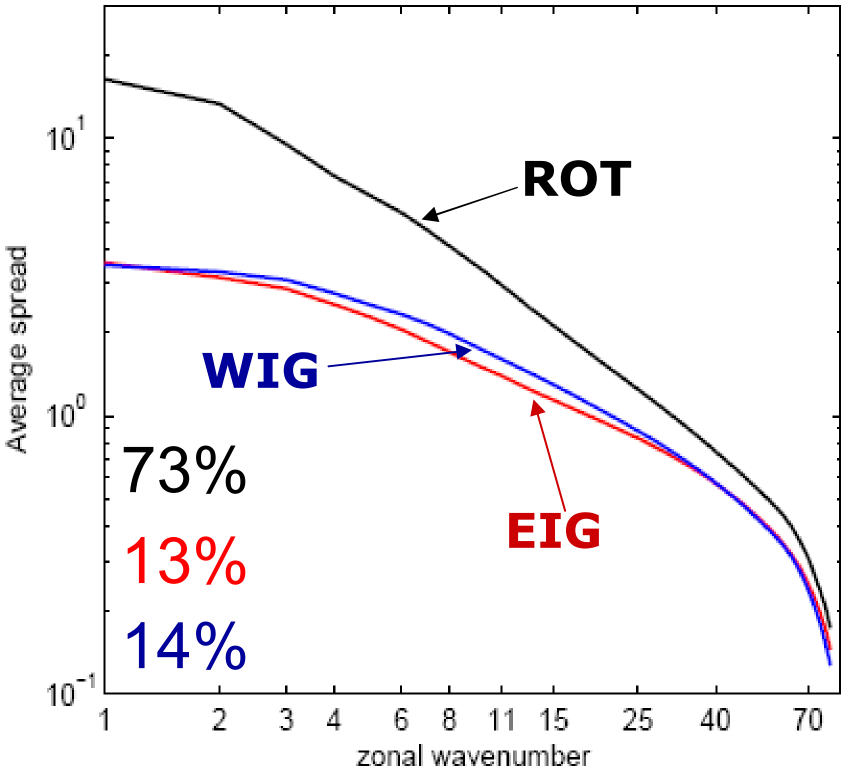
ECMWF: average 3-h fc spread in (m,n) space

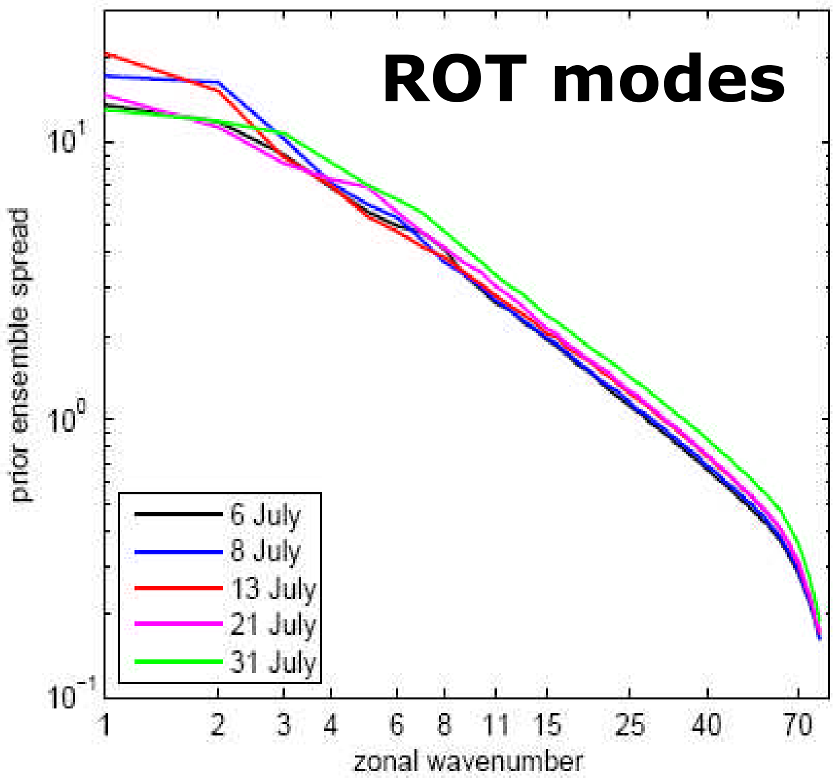






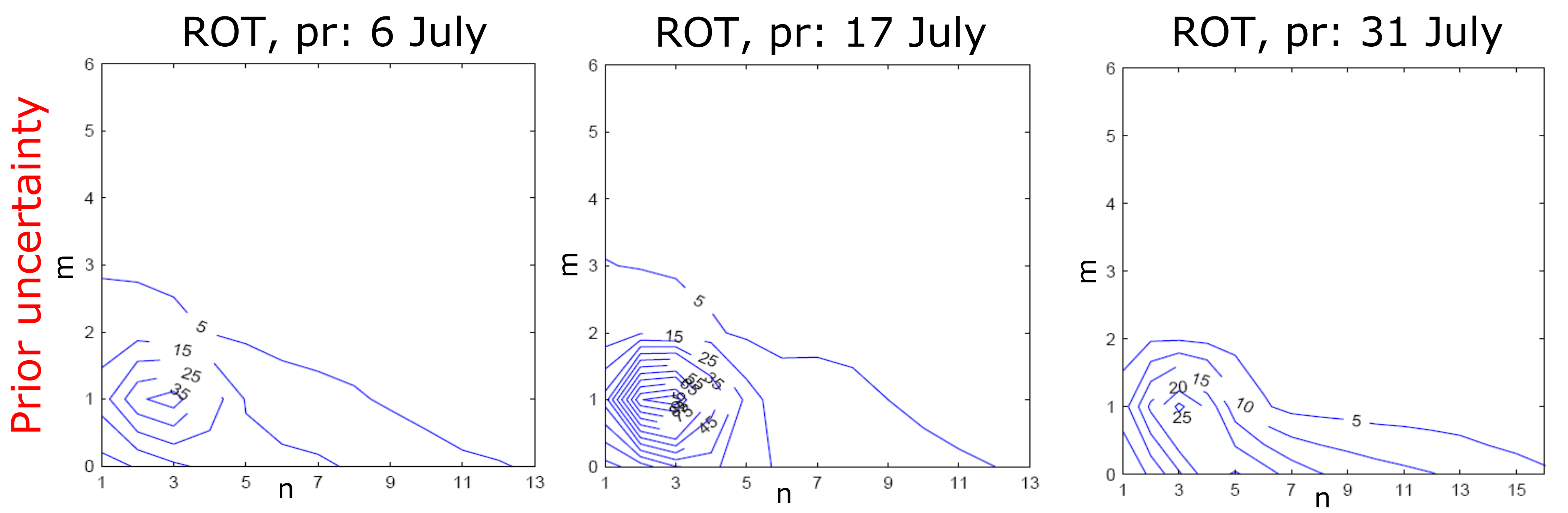
DART/CAM 6-h forecast ensemble: average spread vs. spread variability

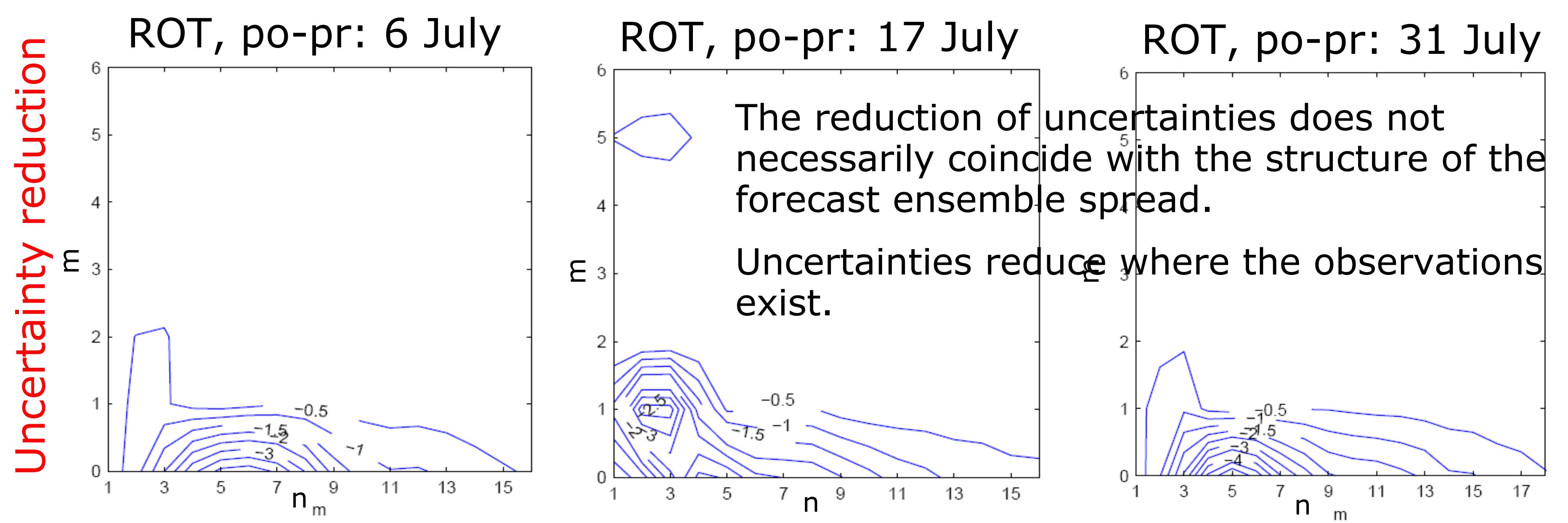




The ensemble spread is related to the impact of inflated covariances, the observation coverage and flow properties.

DART/CAM: Uncertainty reduction in time

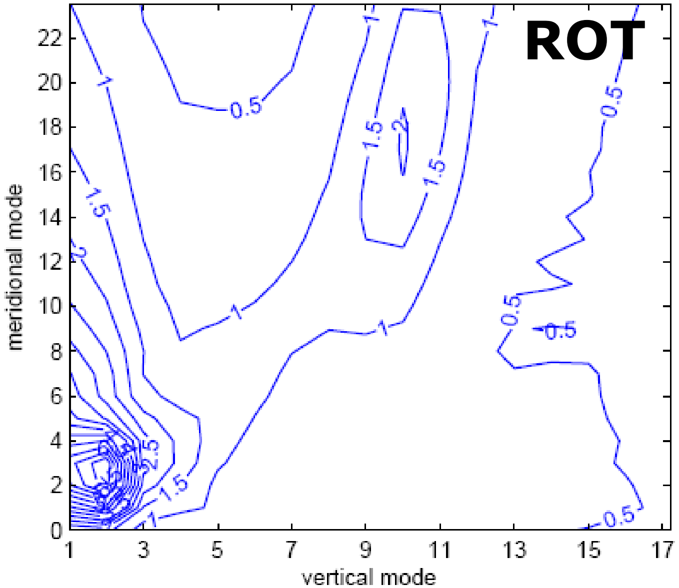


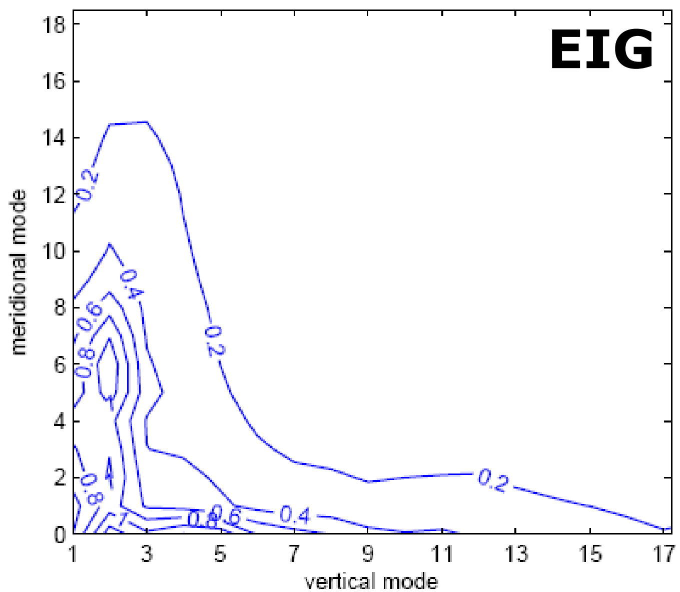


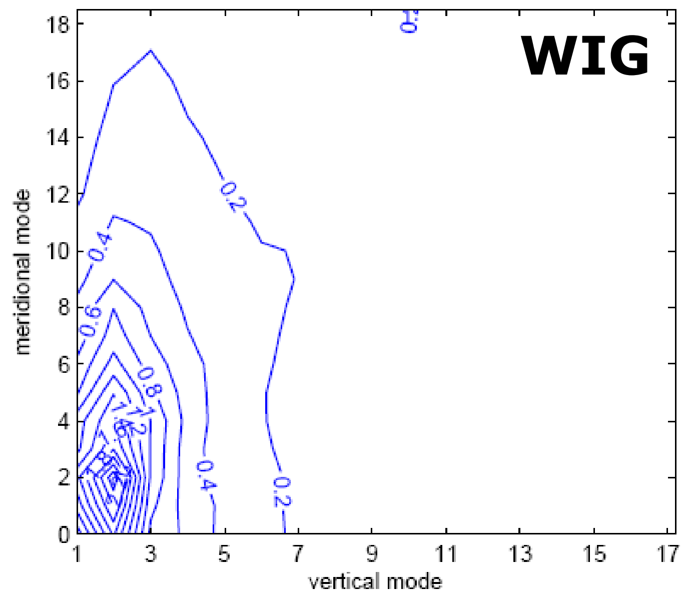
The reduction of uncertainties does not necessarily coincide with the structure of the forecast ensemble spread.

Uncertainties reduce where the observations exist.

DART/CAM: average 6-h fc spread in (m,n) space







Conclusions

- Short-range forecast-error variance contribution of IG motions is about 50% in the ECMWF system. The portion of IG motions in increment fields is about the same (different cycles in two cases).
- The fact that there is twice more IG motion in the increment fields than in the prior ensemble spread in the DART/CAM system is possibly an indication of the noise introduced by the assimilation step.
- ECMWF, NCEP, DART/CAM: levels of IG motions in increment fields far exceed their levels in full analyses fields.
- NCEP/NCAR: Increment fields contain more energy in EIG than in WIG motion, just like the full analyses fields.
- The magnitudes of upper-troposphere wind-field biases in July 2007 were in the range between 1 m/s and 2 m/s illustrating the importance of diagnosing analysis systems in the tropics where magnitudes of the large-scale variability is small. Among various IG motion, the greatest uncertainty is found in the Kelvin wave, the most energetic IG motion in each system.

Thank you very much for your attention!