Quantification of uncertainties in atmospheric analyses and forecasts by using normal modes

Nedjeljka Žagar

in collaboration with J. Tribbia, J. Anderson, K. Raeder

National Center for Atmospheric Research
Advanced Study Program

NCEP, 3 June 2008

Outline: certainty and uncertainty

- Motivation for the revival of normal mode expansion with emphasis on large-scale tropical motions
- Derivation of normal modes for various datasets
- Quantification of energy in three analysis datasets: CAM/DART, ECMWF and NCEP
- Analysis of time averaged analysis increments in terms of various motions and scales
- Quantification of short-range forecast uncertainties in the ensemble system DART/CAM
- Conclusions

Normal mode functions

Kasahara and Puri, 1981

• Linearization around the mean state (vertically stratified in N_z σ levels and at rest)

• New mass variable P $\quad P = gz + RT_0(\sigma)q \quad$ q=ln(p_s)

$$\frac{\partial u'}{\partial t} - 2\Omega v' \sin\phi = -\frac{\partial P'}{a \cos\phi \partial\lambda},$$

$$\frac{\partial v'}{\partial t} + 2\Omega u' \sin\phi = -\frac{\partial P'}{a\,\partial\phi},$$

$$\frac{\partial}{\partial t}\left[\frac{\partial}{\partial\sigma}\left(\frac{\sigma}{R\Gamma_0}\frac{\partial P'}{\partial\sigma}\right)\right] - \nabla\cdot\mathbf{V}' = 0.$$

$$\longrightarrow$$

$$\frac{\partial\tilde{u}}{\partial t} - 2\Omega \sin\phi\tilde{v} = -\frac{g}{a \cos\phi}\frac{\partial\tilde{h}}{\partial\lambda},$$

$$\frac{\partial\tilde{v}}{\partial t} + 2\Omega \sin\phi\tilde{u} = -\frac{g}{a}\frac{\partial\tilde{h}}{\partial\phi},$$

$$\frac{\partial\tilde{h}}{\partial t} + D\nabla\cdot\tilde{\mathbf{V}} = 0,$$

assume separation of variables by new vertical dependence function Ψ

$$\left.\begin{aligned} u' &= \tilde{u}\Psi(\sigma) \\ v' &= \tilde{v}\Psi(\sigma) \\ P' &= g\tilde{h}\Psi(\sigma) \end{aligned}\right\} .$$

$$\int_0^1 \Psi_i(\sigma)\Psi_j(\sigma)d\sigma = \delta_{ij},$$

$$\frac{d}{d\sigma}\left(\frac{\sigma g}{R\Gamma_0}\frac{d\Psi}{d\sigma}\right) + \frac{1}{D}\Psi = 0.$$

$\Gamma_0 = \frac{\kappa T_0}{\sigma} - \frac{dT_0}{d\sigma}$ Stability parameter

D - separation constant of dimension length - “equivalent depth”

System of equations for the horizontal structure of modes

$$\frac{\partial \tilde{u}}{\partial t} - 2\Omega \sin\phi \tilde{v} = -\frac{g}{a\cos\phi}\frac{\partial \tilde{h}}{\partial \lambda},$$

$$\frac{\partial \tilde{v}}{\partial t} + 2\Omega \sin\phi \tilde{u} = -\frac{g}{a}\frac{\partial \tilde{h}}{\partial \phi}, \qquad (*)$$

$$\frac{\partial \tilde{h}}{\partial t} + D\nabla\cdot\tilde{\mathbf{V}} = 0,$$

$$(\tilde{u},\tilde{v},\tilde{h})^{\mathrm{T}} = \mathbf{S}_n \mathbf{H}_r^s(\lambda,\phi;\, n)\exp(-i\sigma_r^s t),$$

$$\mathbf{S}_n = \begin{pmatrix} (gD_n)^{1/2} & 0 & 0 \\ 0 & (gD_n)^{1/2} & 0 \\ 0 & 0 & D_n \end{pmatrix}$$

Hough functions

$$\mathbf{H}_r^s(\lambda,\phi;\, n) = \mathbf{H}_r^s(\phi;\, n)e^{is\lambda}$$

$$\frac{1}{2\pi}\int_0^{2\pi}\int_{-1}^{1} \mathbf{H}_{r'}^{s'}\cdot(\mathbf{H}_r^s)^* d\mu\, d\lambda = \delta_{rr'}\delta_{ss'},$$

n - vertical mode index

s - zonal mode index

r - meridional mode index

σ – eigen frequency

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

Motivation for the present research

• Previous applications of normal modes indicated a negligible amount of energy in divergent motions and a dominance of the first vertical mode in the energy spectra (Tanaka et al., J. Met. Soc. Japan)

• In the NWP model applications normal mode functions have been primarily used for the initialization purposes

• State of the art NWP and climate models with good physical parameterizations and high resolution represent divergent motions much better.

• Large-scale equatorial waves in recent years have been diagnosed from different mass-field observations. Quantification of their variance and dynamical relevance still puzzling.

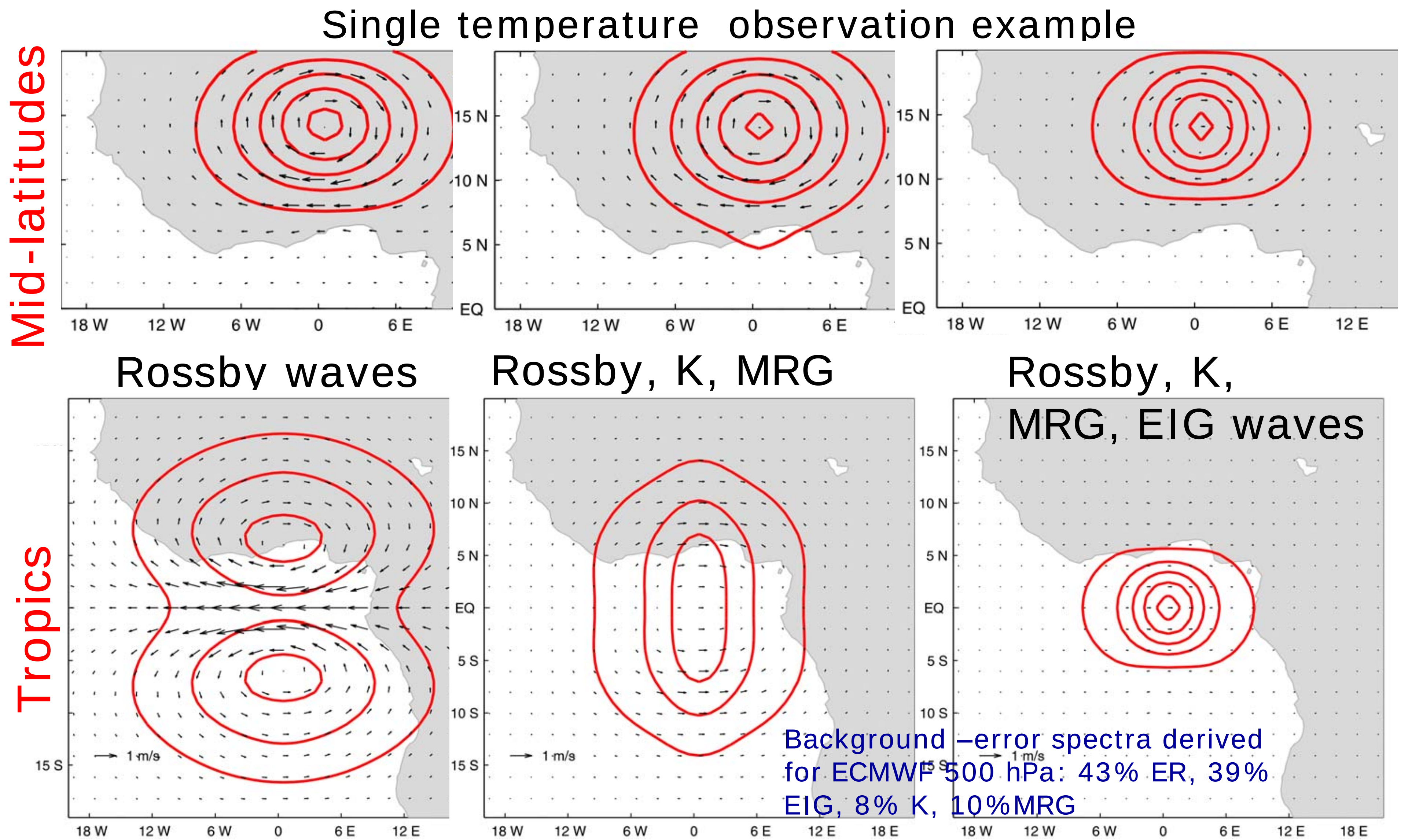
• Divergent tropical circulations crucial, but unreliable from present (re)analysis

Tropics

Region with largest uncertainties in the existing (re)analysis datasets, because of

- Lack of direct observations of the wind field, especially wind profiles
- Difficult task of the tropical data assimilation due to balance issue

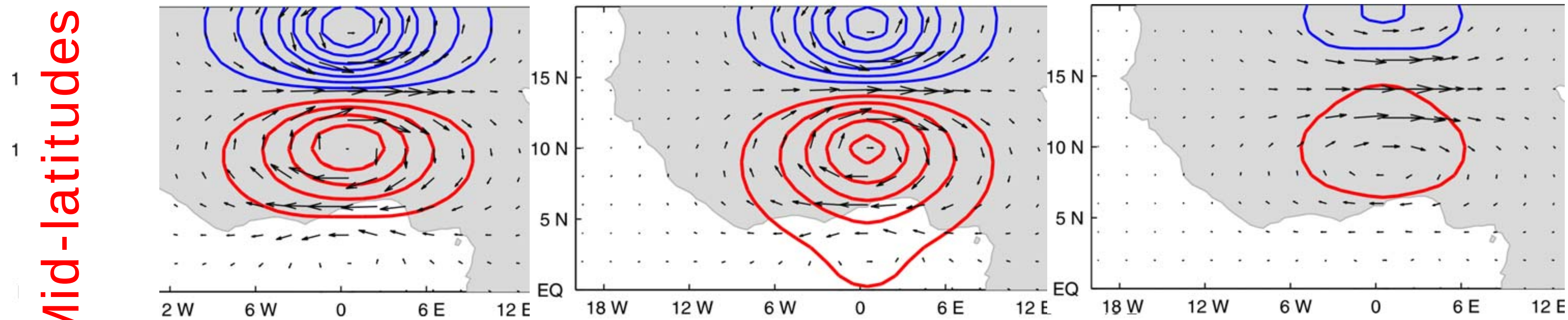
Uncertainty concerning the role of divergent motions: static bkg-error covariance



Uncertainty concerning the role of divergent motions: static bkg-error covariance

Single zonal wind observation

Mid-latitudes

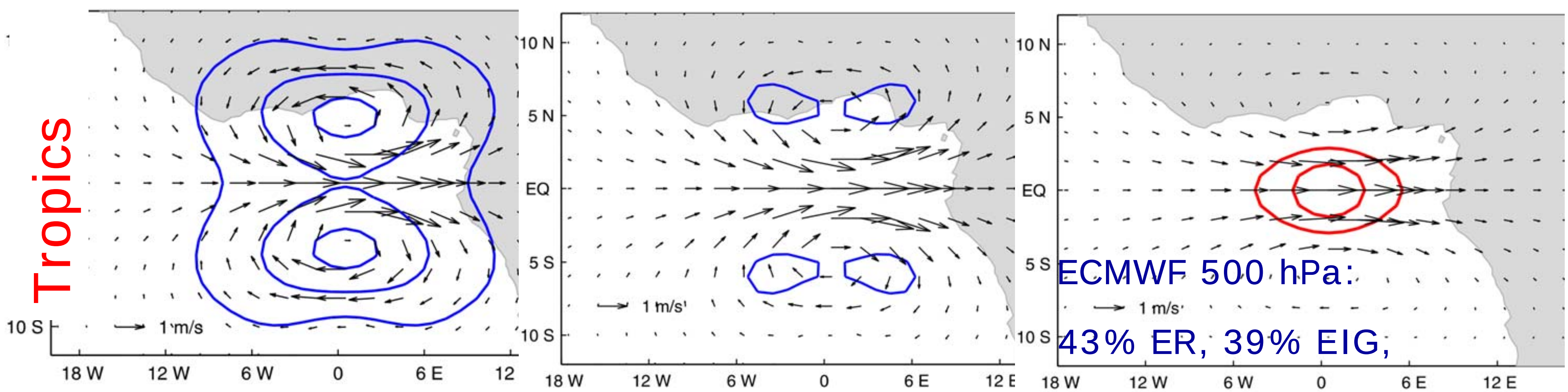


Rossby waves

Rossby, K, MRG

Rossby, K, MRG, EIG waves

Tropics



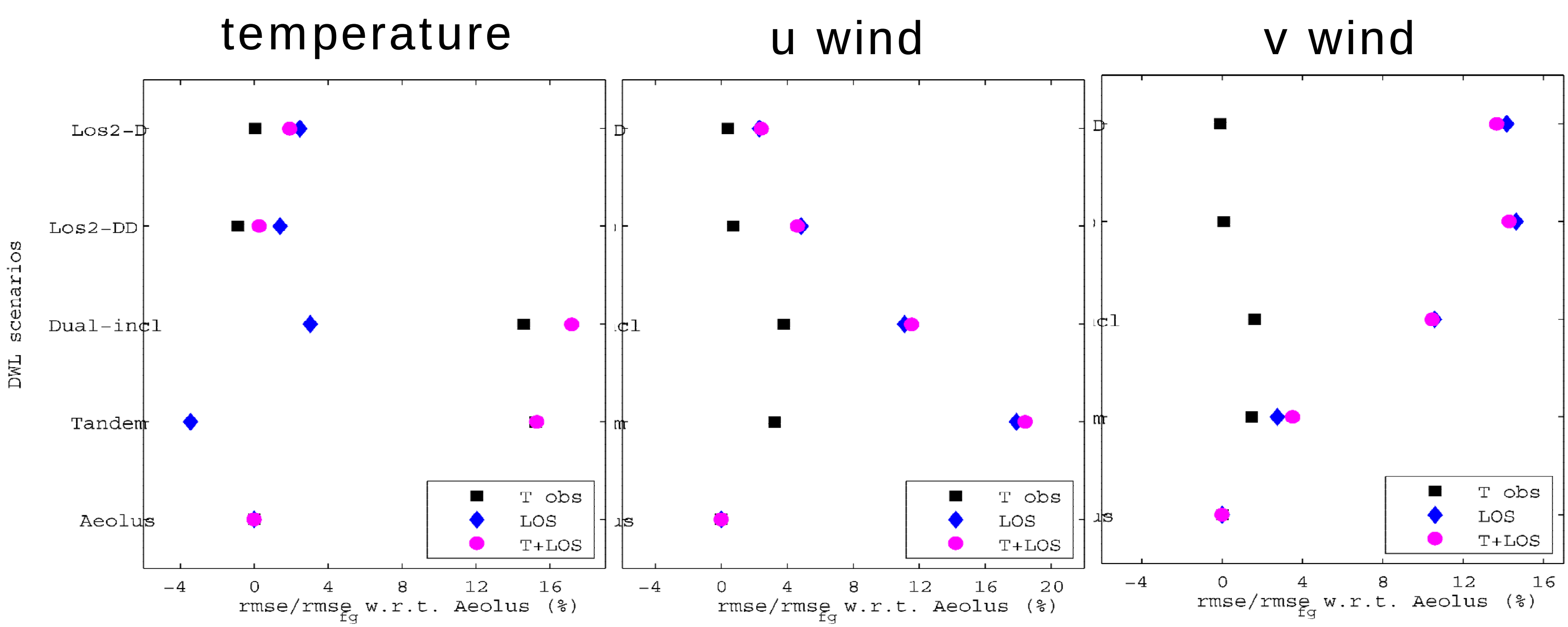
ECMWF 500 hPa:

43% ER, 39% EIG,

8% K, 10%MRG

Sensitivity experiments with 3D-Var

Reliable B matrix, simple tropical model: result from a study about added value by the second satellite with respect to ADM-Aeolus DWL satellite in terms of the reduction % of the first-guess error (Žagar et al., MWR, 2008)



3D-Var assimilation acts as a univariate analysis !

Tropics

Region with largest uncertainties in the existing (re)analysis datasets, because of

• Lack of direct observations of the wind field, especially wind profiles

• Difficult task of the tropical data assimilation due to balance issue

Remedies

• Improved global observing system

• More advanced data assimilation procedures

• Improvements of the models, especially convective parameterizations and resolution

Questions

Tropics

How much of the large-scale tropical circulation is made up by the Kelvin wave, mixed Rossby-gravity wave, other inertio-gravity waves?

How is this dependent on the model resolution, physics, biases?

What is the spectra of forecast errors in the tropics like? How are the tropical forecast errors spread across the scales and motion types? What modes do the biases project onto?

Related data assimilation issues

How important are special tropical waves (Kelvin, mixed Rossby-gravity, large-scale IG) for the data assimilation?

What is the real potential of the EnKF in the tropics due to flow-dependent background-error covariances in comparison to 4D-Var?

Application of normal modes to CAM, NCEP, and ECMWF data

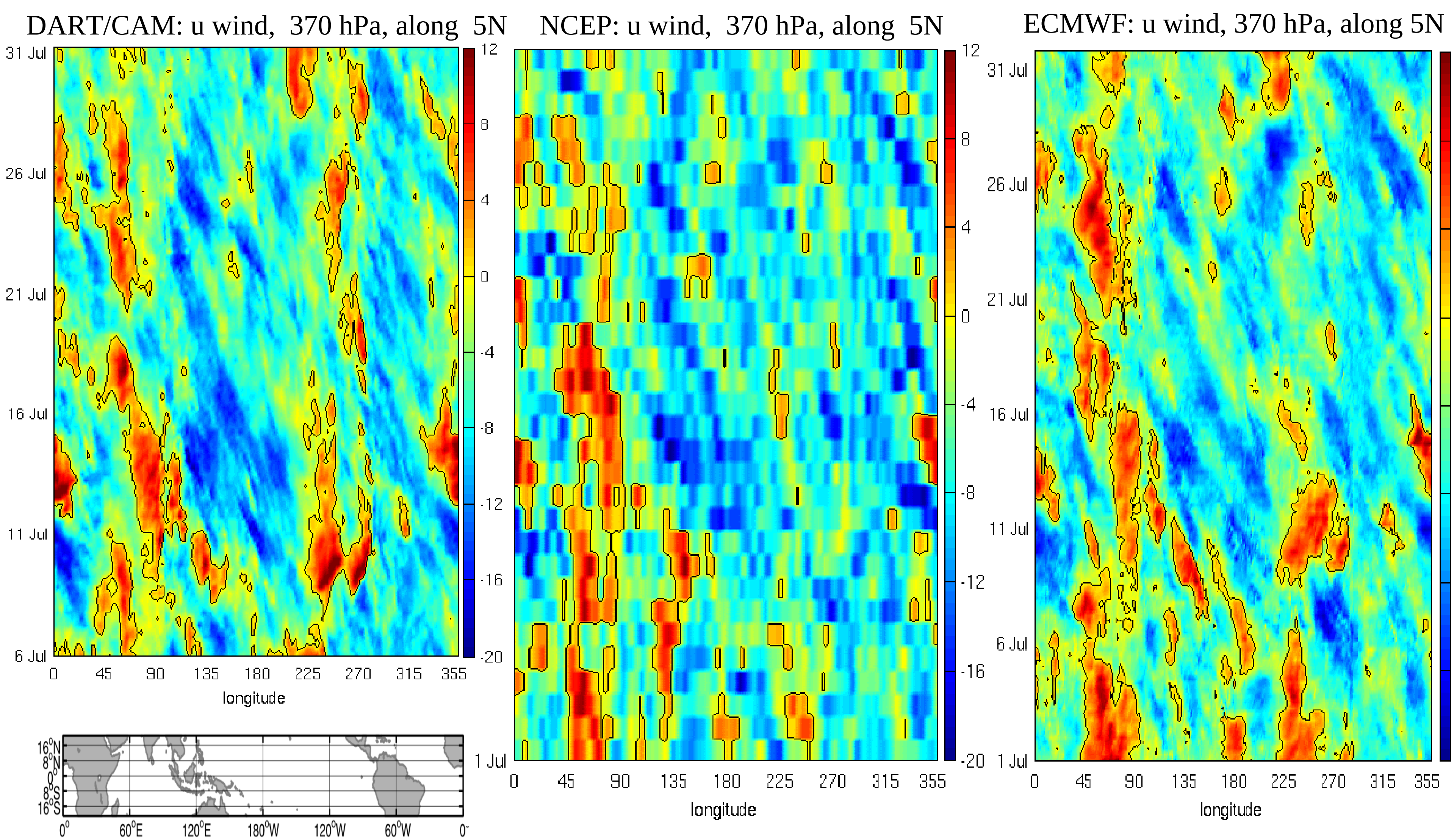
Three analysis dataset for July 2007, global fields every 6 hours

DART/CAM: ensemble mean from the DART system, version 3.1, T85 horizontal resolution, 26 vertical levels up to 3.5 hPa. Limited number of observations (conventional observations and AMVs).

ECMWF: operational analyses, 12-hour 4D-Var system, Cycle 32r2, T799 interpolated to N64 grid, 91 vertical level up to 1 Pa. Large amounts of satellite observations.

NCEP-NCAR reanalyses from NCAR mass archive: 3D-Var system, T62 horizontal resolution, 28 vertical levels up to 2.7 hPa. The assimilation system not the recent one.

Tropical winds in 3 analysis datasets in July 2007 at 370 hPa



Normal mode expansion

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

common expansion coefficient

3D normal mode functions

$$\boxed{\mathbf{X}(\lambda,\varphi,z,t)=\sum_{m=1}^{N_m}\sum_{n_{i=1,2,3}=0}^{N_n-1}\sum_{k=-N_k}^{N_k}\chi_{knm}(t)\mathbf{S}_{\mathbf{m}}\Pi_{knm}(\lambda,\varphi,z)}$$

input data vector

normalization matrix

$$\mathbf{X}=(u,v,{}^{P}\!/_{g})^T$$

N_m – no. vertical modes, index m

N_n – no. meridional modes per wave type, index n

N_k – no. zonal waves, index k

$$\Pi_{knm}(\lambda,\varphi,z)=\Phi_m(z)\cdot \mathrm{H}_{knm}$$

vertical normal modes

Hough functions

$$\mathbf{S}_{\mathbf{m}}=\begin{pmatrix}(gH_{eq})^{1/2} & 0 & 0\\ 0 & (gH_{eq})^{1/2} & 0\\ 0 & 0 & gH_{eq}\end{pmatrix}$$

$$\langle \Pi_{knm},\Pi_{k'n'm'}\rangle=\delta_{kk'}\delta_{nn'}\delta_{mm'}$$

orthogonal 3D expansion basis

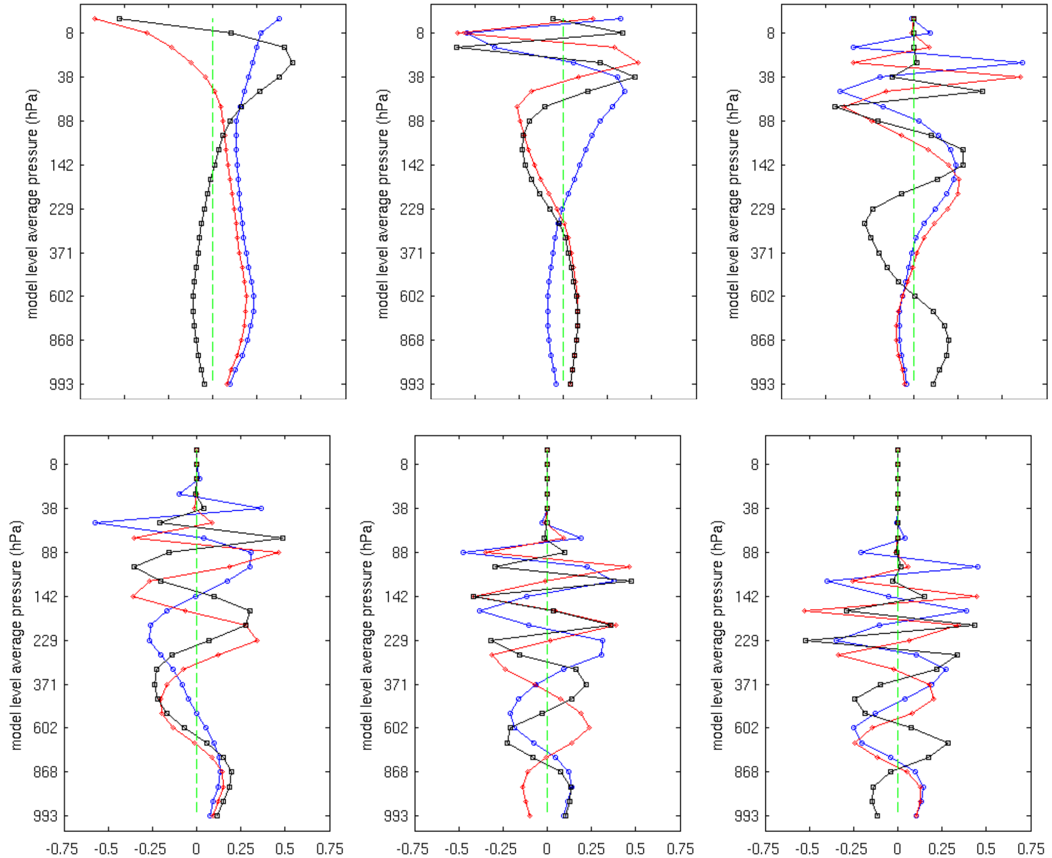
Vertical eigenstructures for CAM

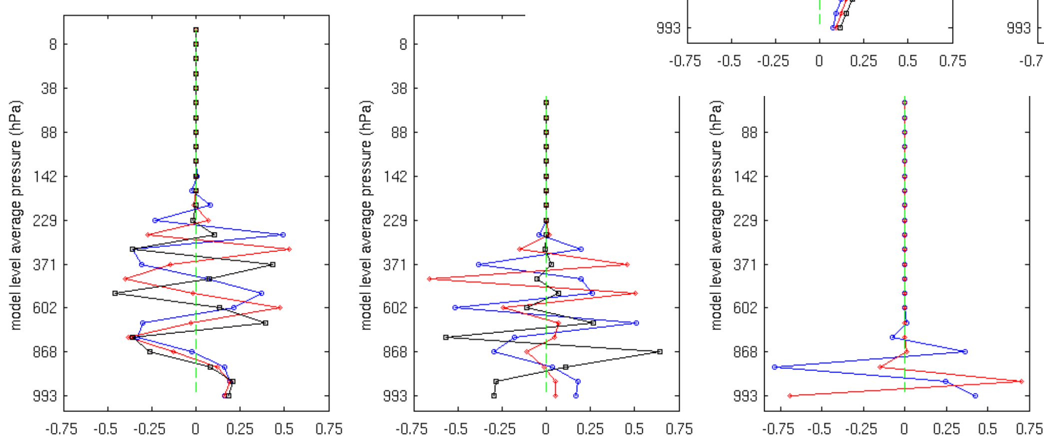
Input information:
vertical discretization,
temperature profile,
stability profiles

H_{eq} from 10 km to 0.3 m

10 km, 6.2 km, 2.2 km,
985 m, 572 m, 379 m,
250 m, 162 m, 107 m

Modes 10-26 have H_{eq}
below 100 m





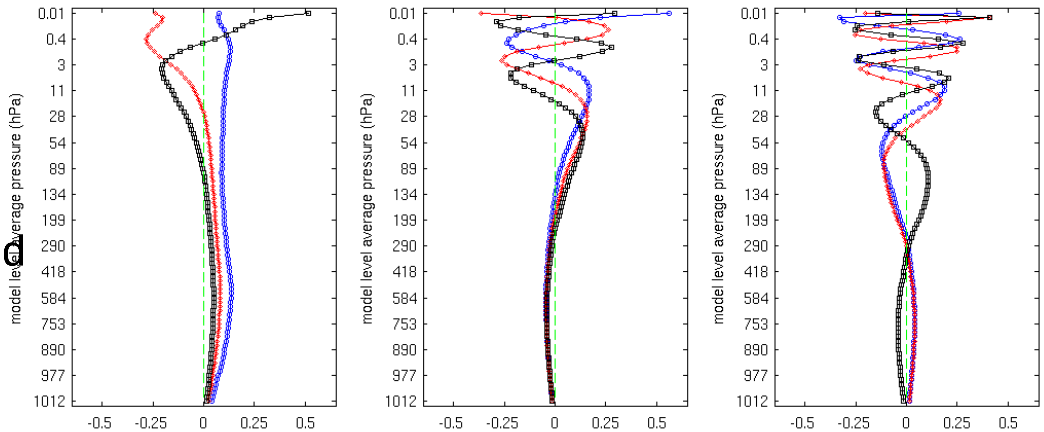
Vertical eigenstructures for ECMWF

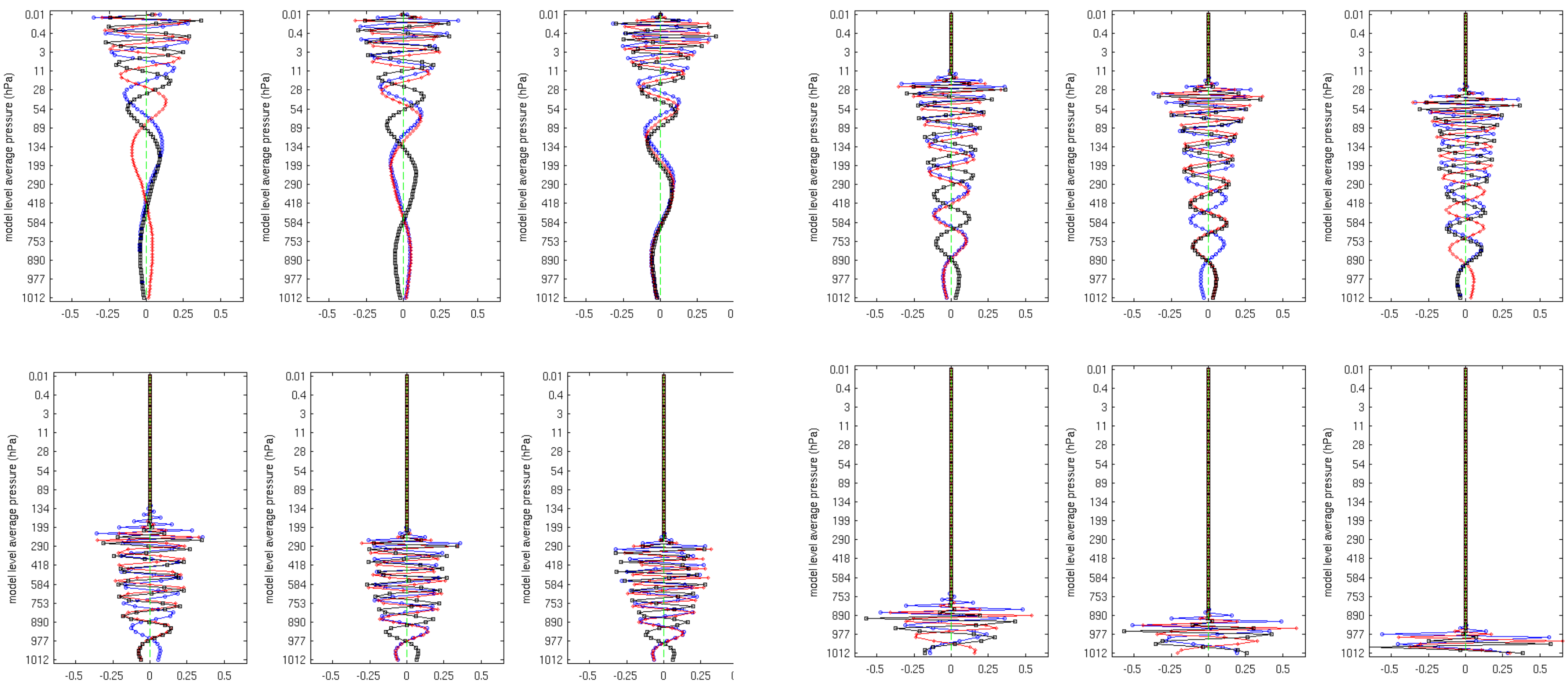
A difficult one:

H_{eq} from 10 km to 8 mm

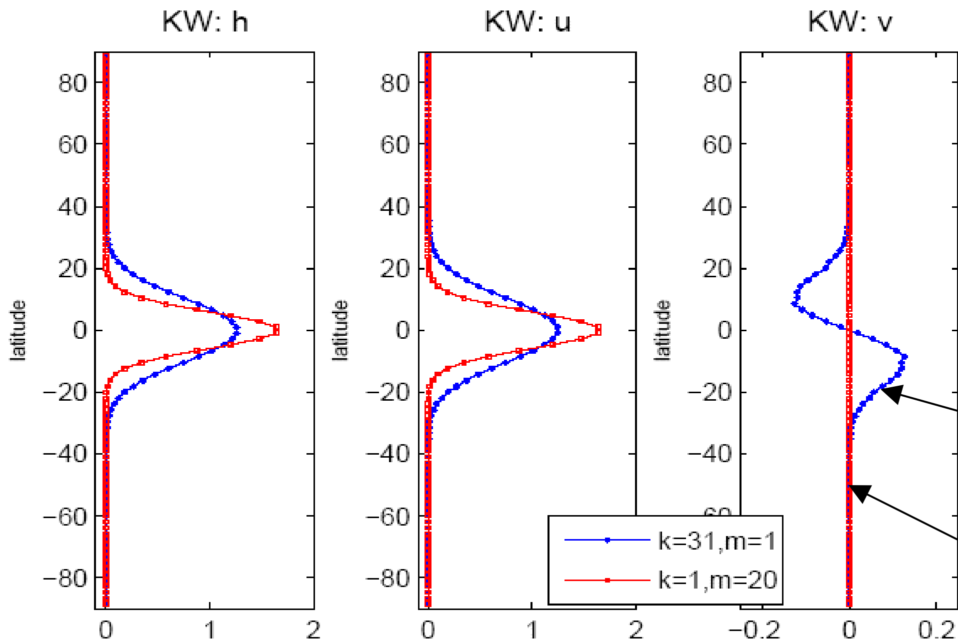
First 18 with H_{eq} > 100 m

Modes 19-38 between 100 m and 10 m, 39-66 between 10 and 1 m, and 66-91 below 1 m.





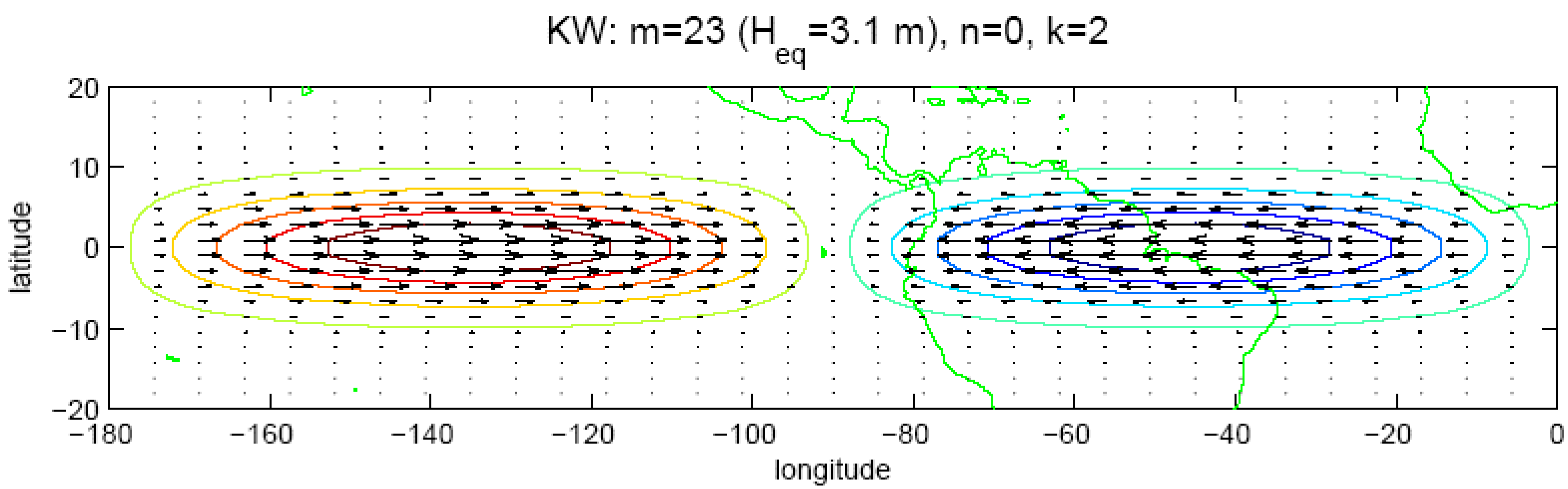
Horizontal structures: Kelvin wave example



Higher vertical modes and higher zonal wave numbers <=> horizontal structures more meridionally trapped

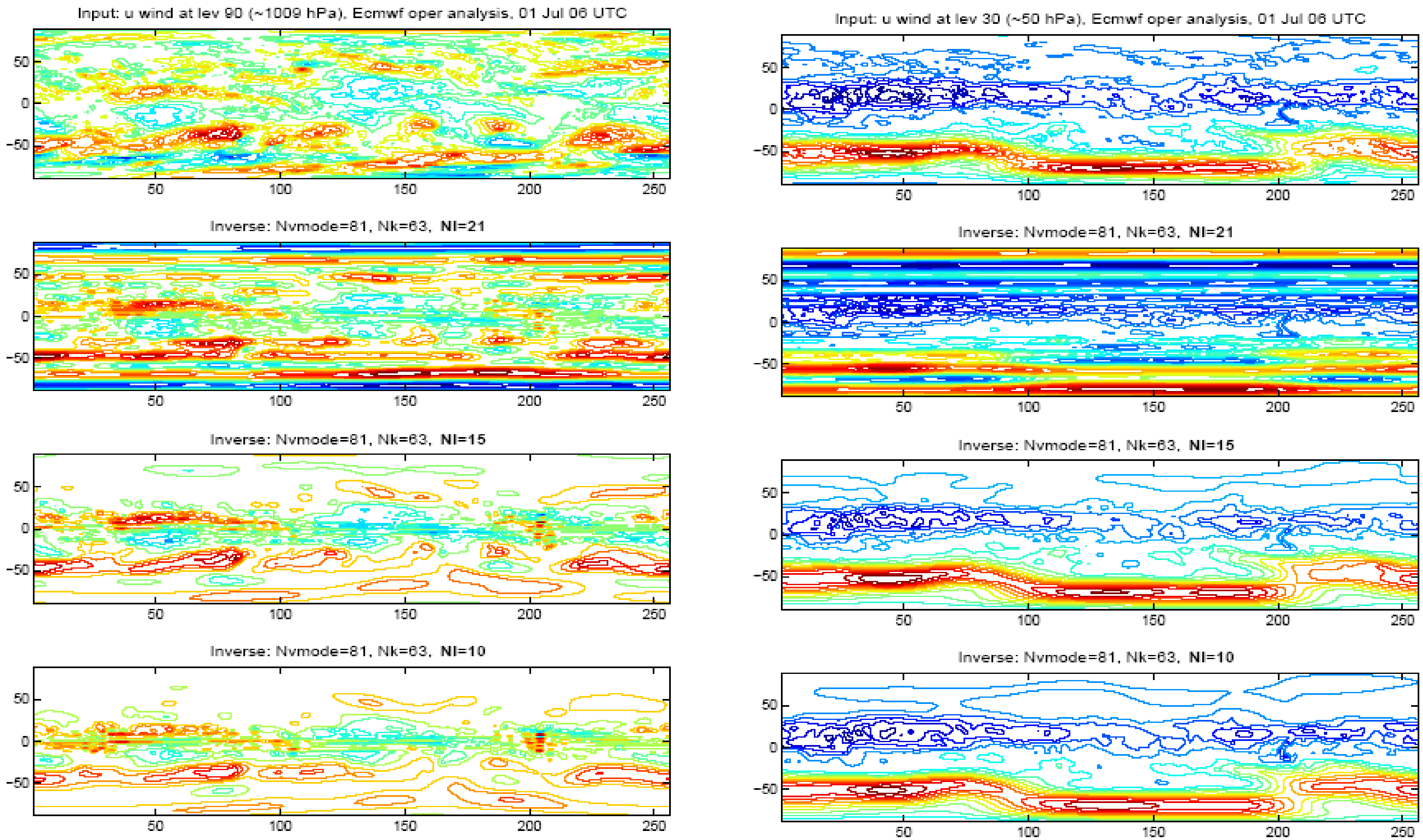
m=1,
k=31

m=20,
k=1

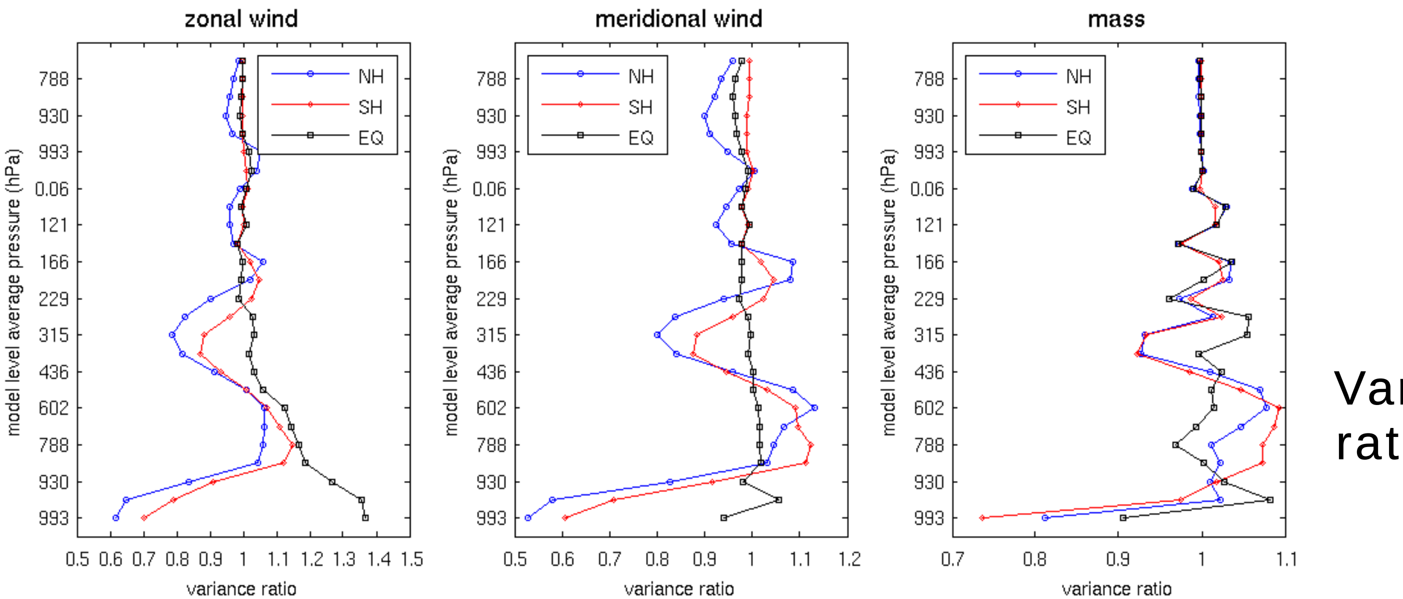


Tuning the expansion: an example of tuning N_n

Choosing a (N_k,N_n,N_m) combination that will represent the most of input data variance. A trafe-off between the desired fit, regions and variable of most interest



Verification of the expansion quality for CAM



$N_k = 80$

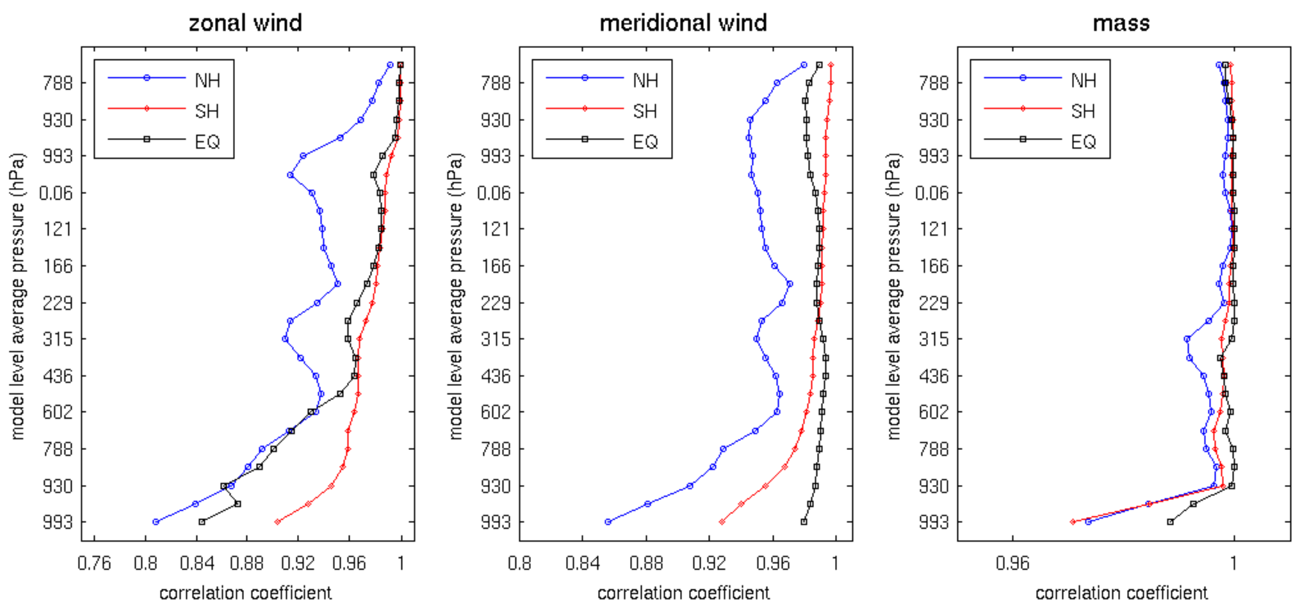
$N_n = 25$

$N_m = 25$

Variance ratio

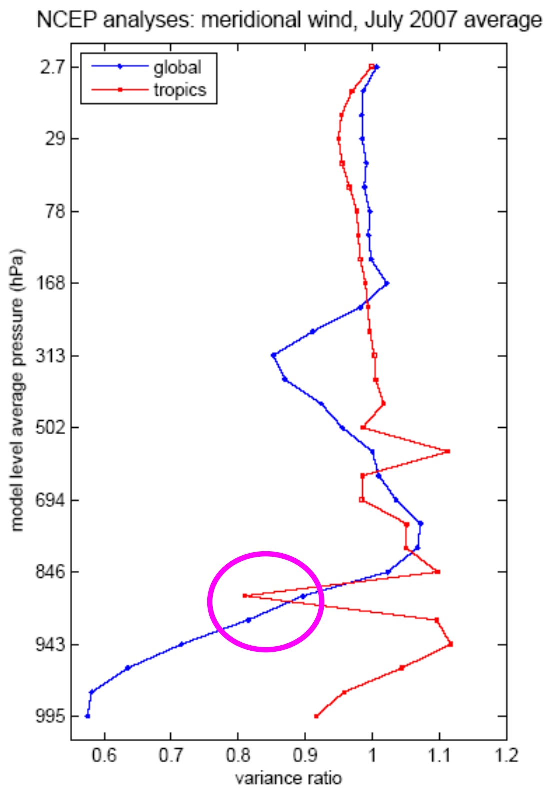
Below 900 hPa zonal wind variance overestimated in the tropics, and underestimated in the mid-latitudes

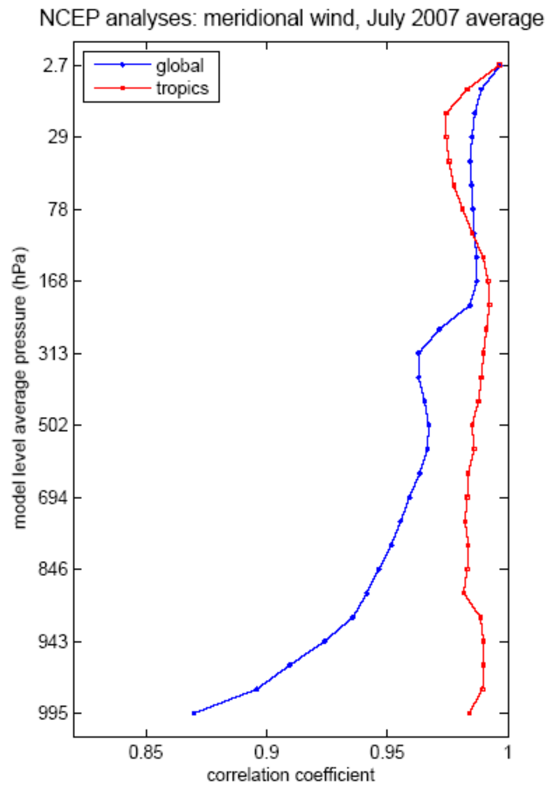
Mass-field variance poor close to the surface due to orography

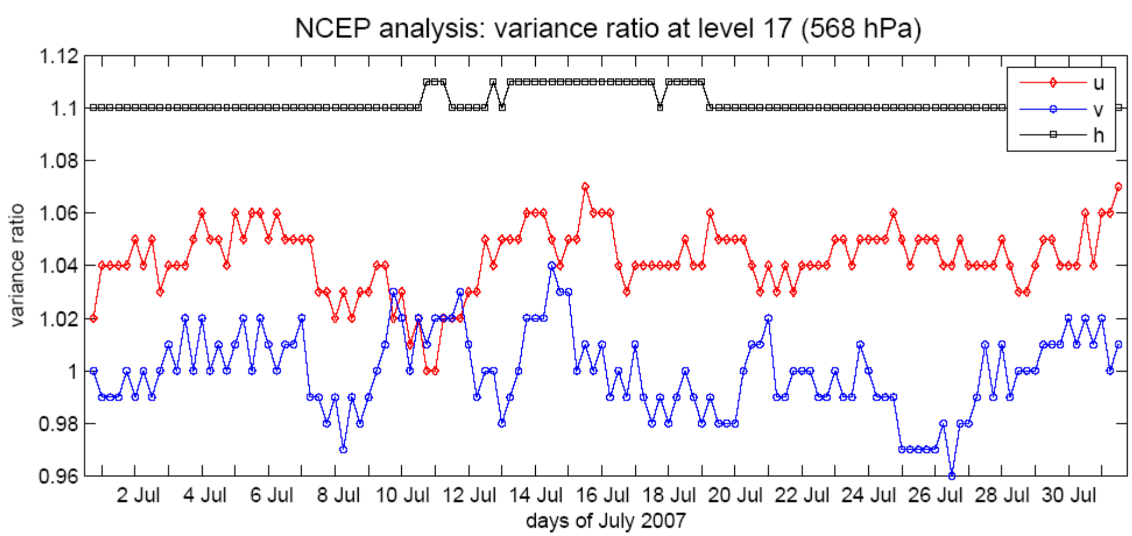


Correlation coefficient

Tuning the expansion: NCEP solution







$N_k = 46$

$N_n = 20$

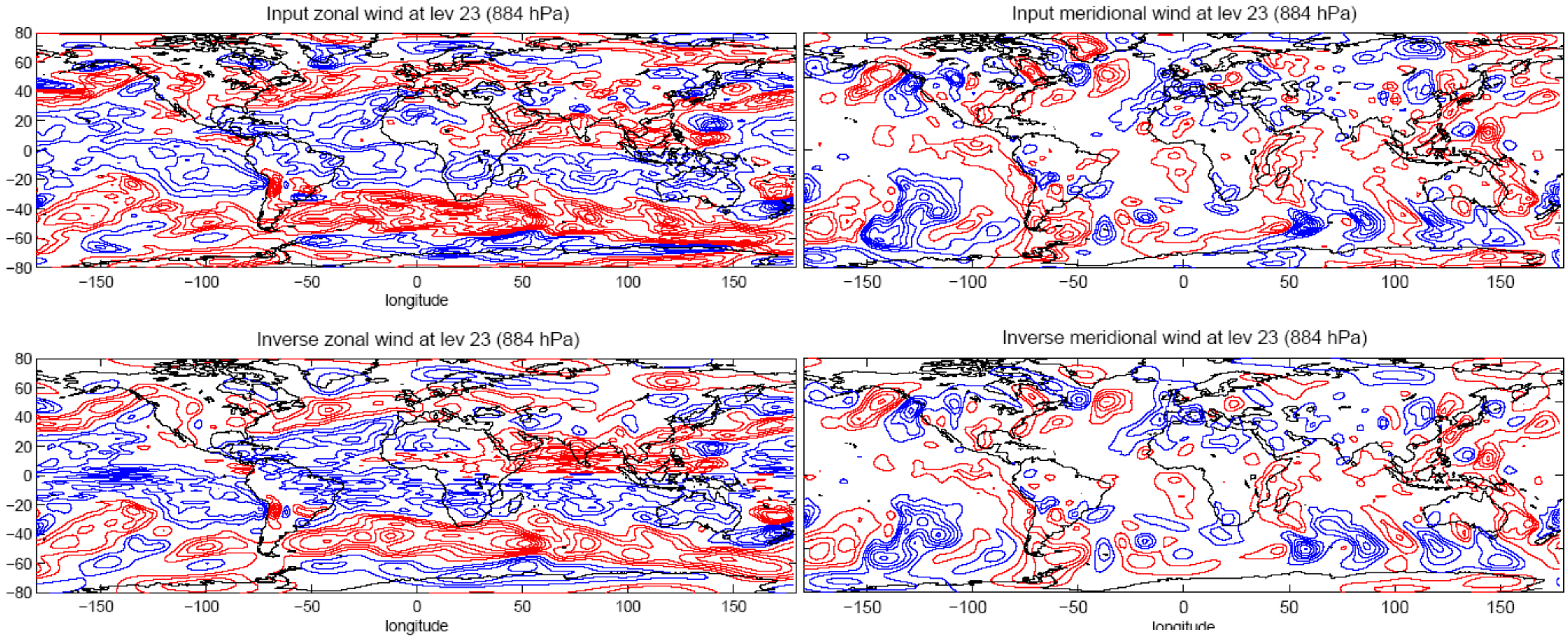
$N_m = 25$

Little variability of the mass field in the tropics

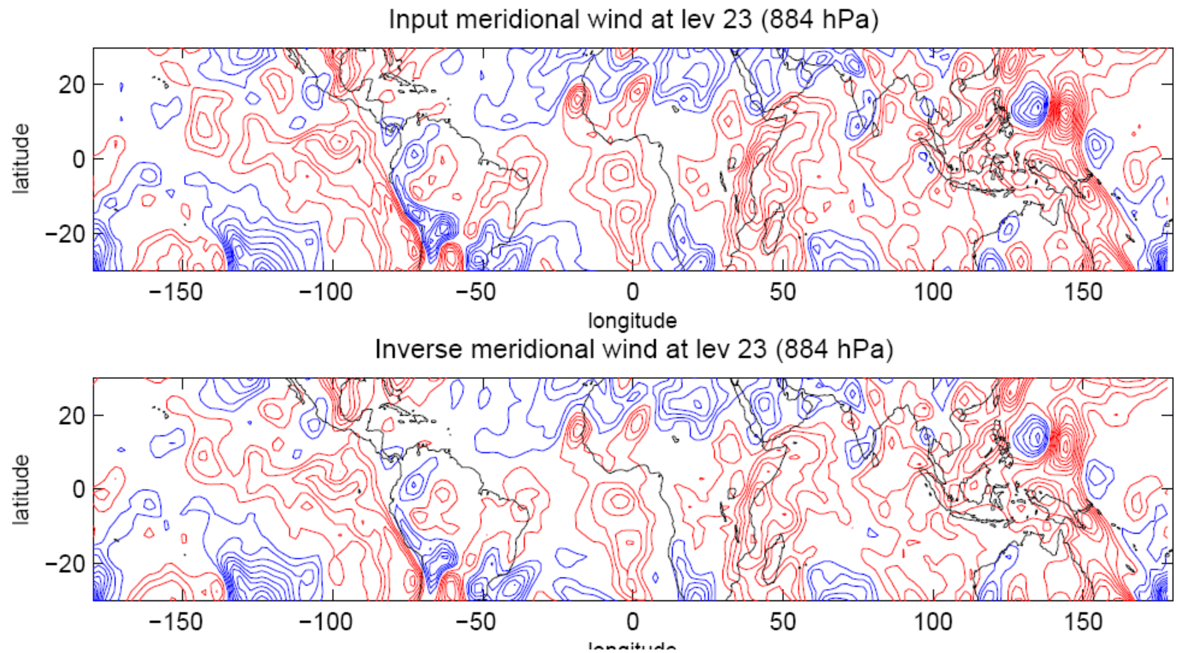
Fit worst at lowest levels

Variance of the tropical zonal wind overestimated at lowest levels

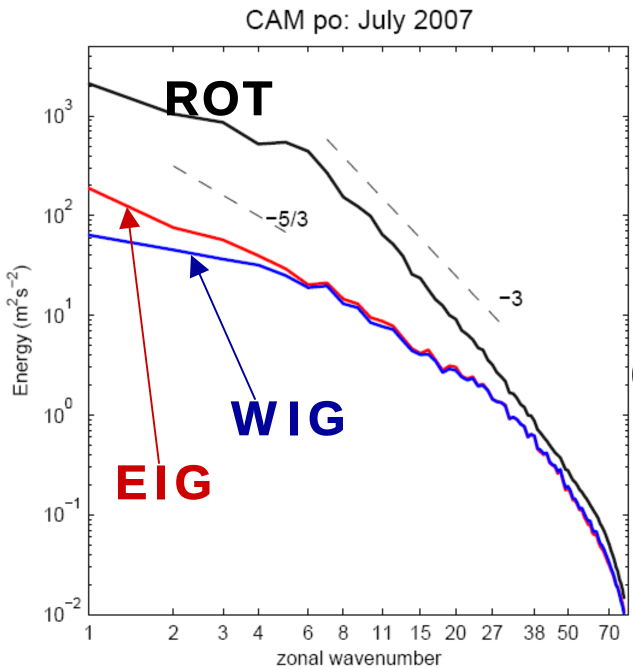
Temporal variation of the expansion quality do not vary significantly in time



Example of the projection quality for NCEP wind field at 884 hPa level



Energy distribution in CAM

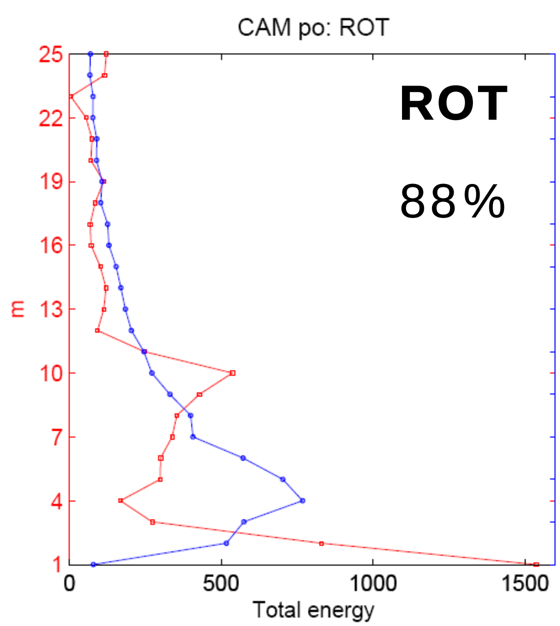


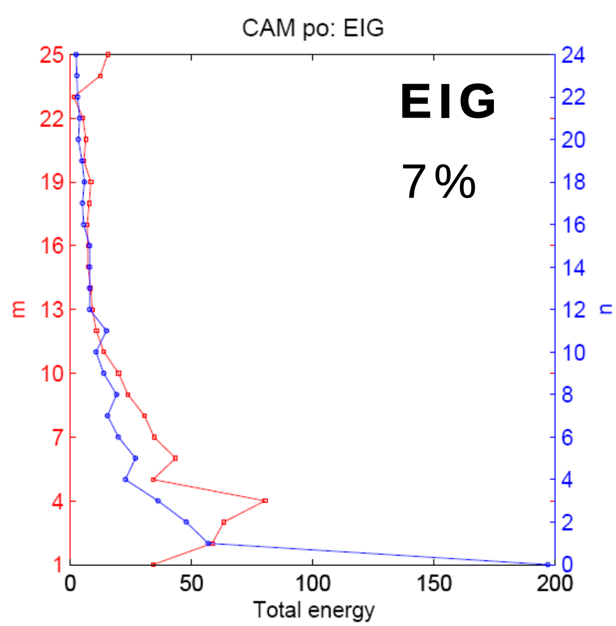
(m,n) Σ

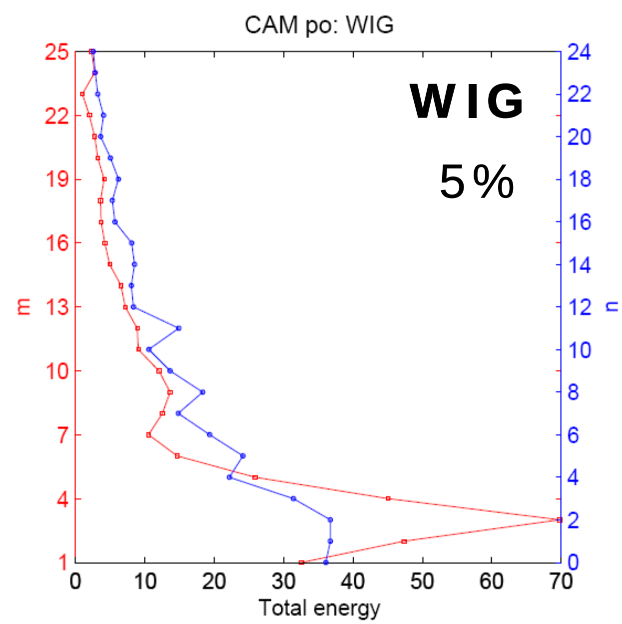
Posterior ensemble mean, average over 25-day period 6-31 July 2007

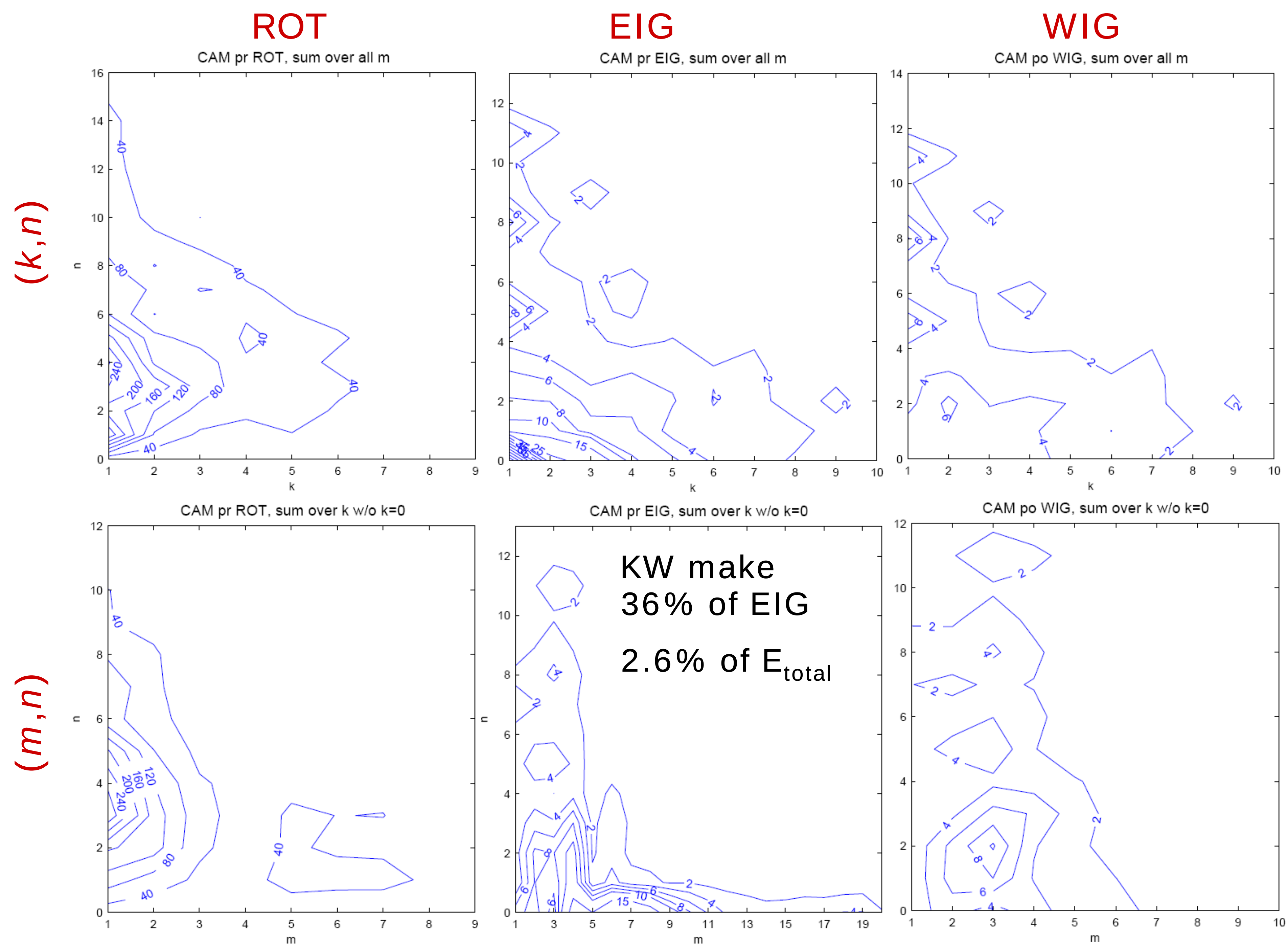
$$\sum_k \sum_n \sum_m gH_{eq} \left| \chi_{knm} \right|^2$$

(m,k) Σ (n,k) Σ



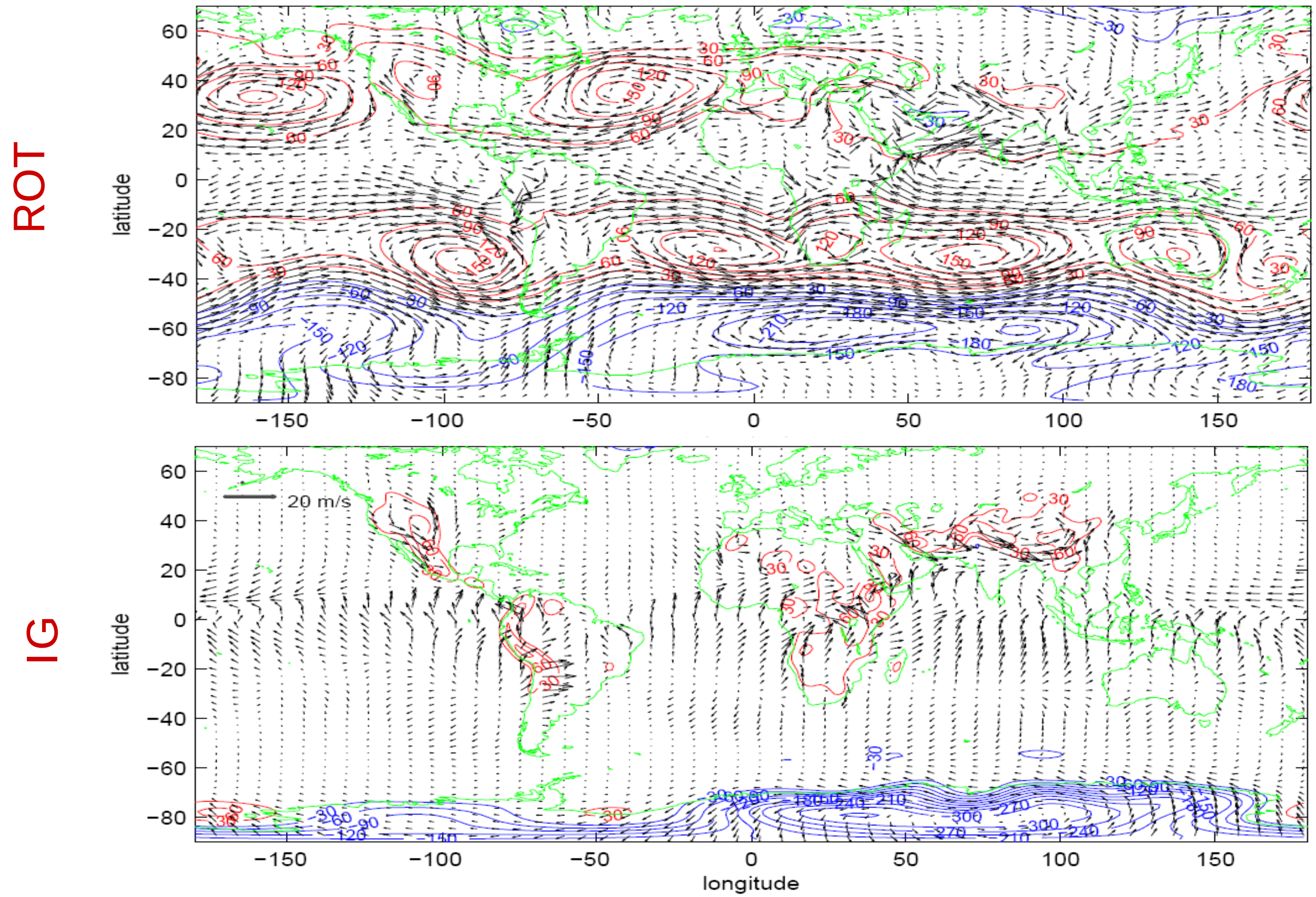






ROT
EIG
WIG
(k,n)
(m,n)
CAM pr ROT, sum over all m
CAM pr EIG, sum over all m
CAM po WIG, sum over all m
CAM pr ROT, sum over k w/o k=0
CAM pr EIG, sum over k w/o k=0
CAM po WIG, sum over k w/o k=0
KW make
36% of EIG
2.6% of E_total

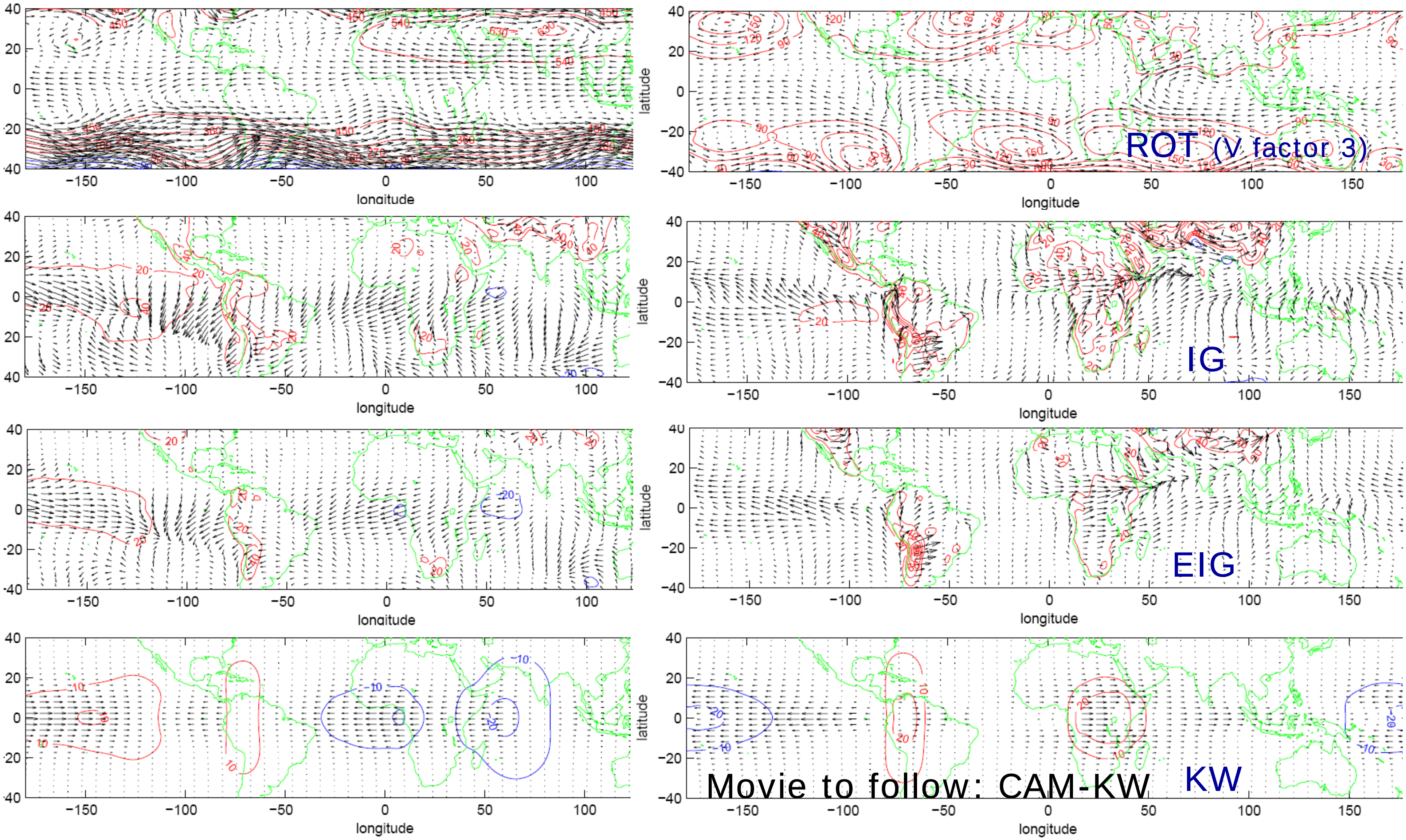
Mean low-level July circulation in CAM



Model level 24 (~929 hPa)

Tropics as envisaged by A. Gill (1980)

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



Kelvin wave evolution in CAM in July 2007

Movie is available at
//www.cgd.ucar.edu/cdp/nzagar/cam_kw.gif

Kelvin wave evolution in CAM: summary

• Reversed flow in the lower and upper troposphere

• Spatial discontinuity of the k=1 signal

• Stronger k=1 signal developed by the end of month, especially in the Pacific

• Oscillations of daily period due to observations

How reliable is this Kelvin wave evolution?

DART/CAM uses few observations in the tropics. The assimilation uses flow-derived (multivariate) background-error covariances

• Inter-comparison with other analyses

Kelvin wave evolution in July 2007 by NCEP

Movie is available at

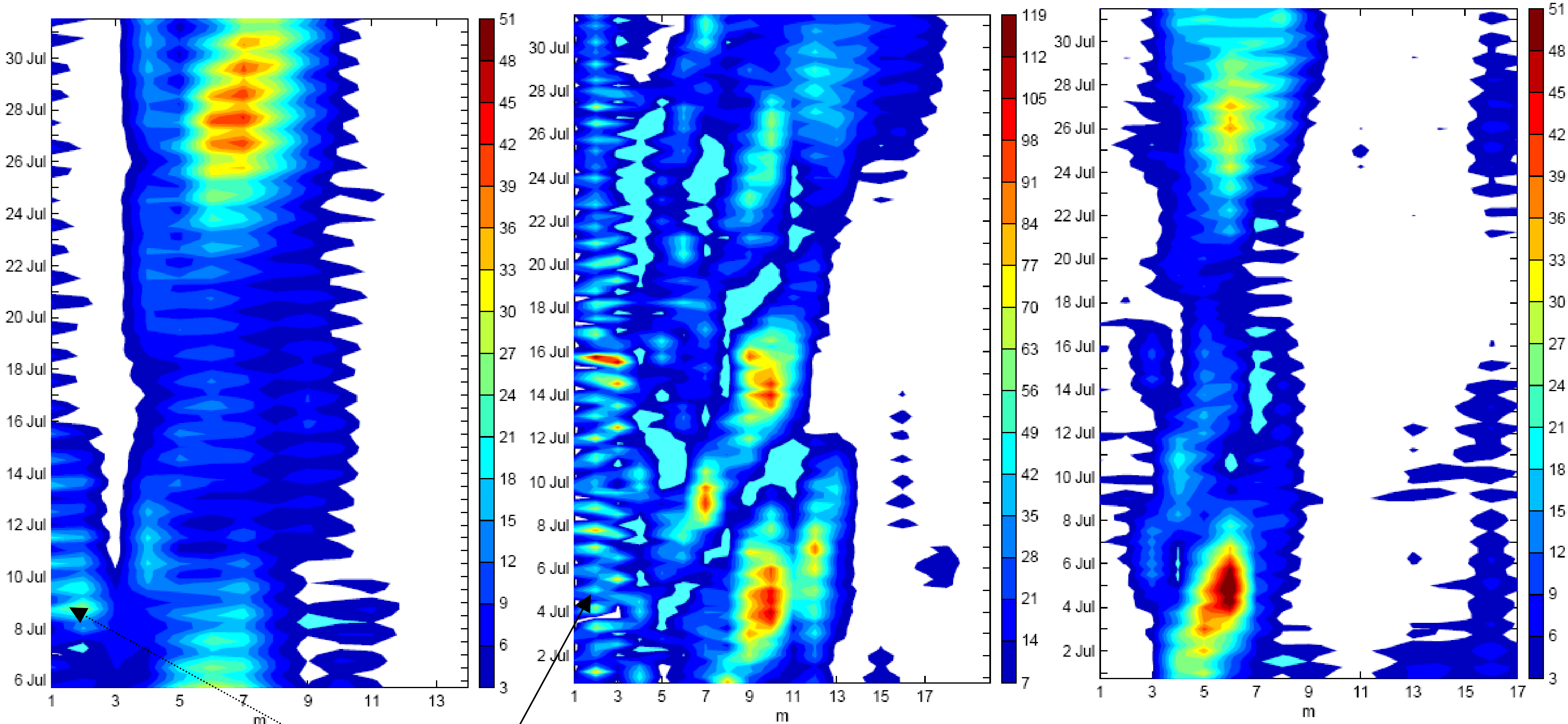
//www.cgd.ucar.edu/cdp/nzagar/ncep_kw.gif

Kelvin wave evolution in July 2007 by ECMWF

Movie is available at
//www.cgd.ucar.edu/cdp/nzagar/ecm_kw.gif

Temporal evolution of the KW, *k*=1 signal

CAM ECMWF NCEP



Tidal signal

$H_{eq(5\text{-}7)}$ = 570, 370, 250 m

$H_{eq(7\text{-}11)}$ = 700, 528, 413, 332, 271

$H_{eq(5\text{-}6)}$ = 500, 300 m

Kelvin wave evolution in CAM: summary

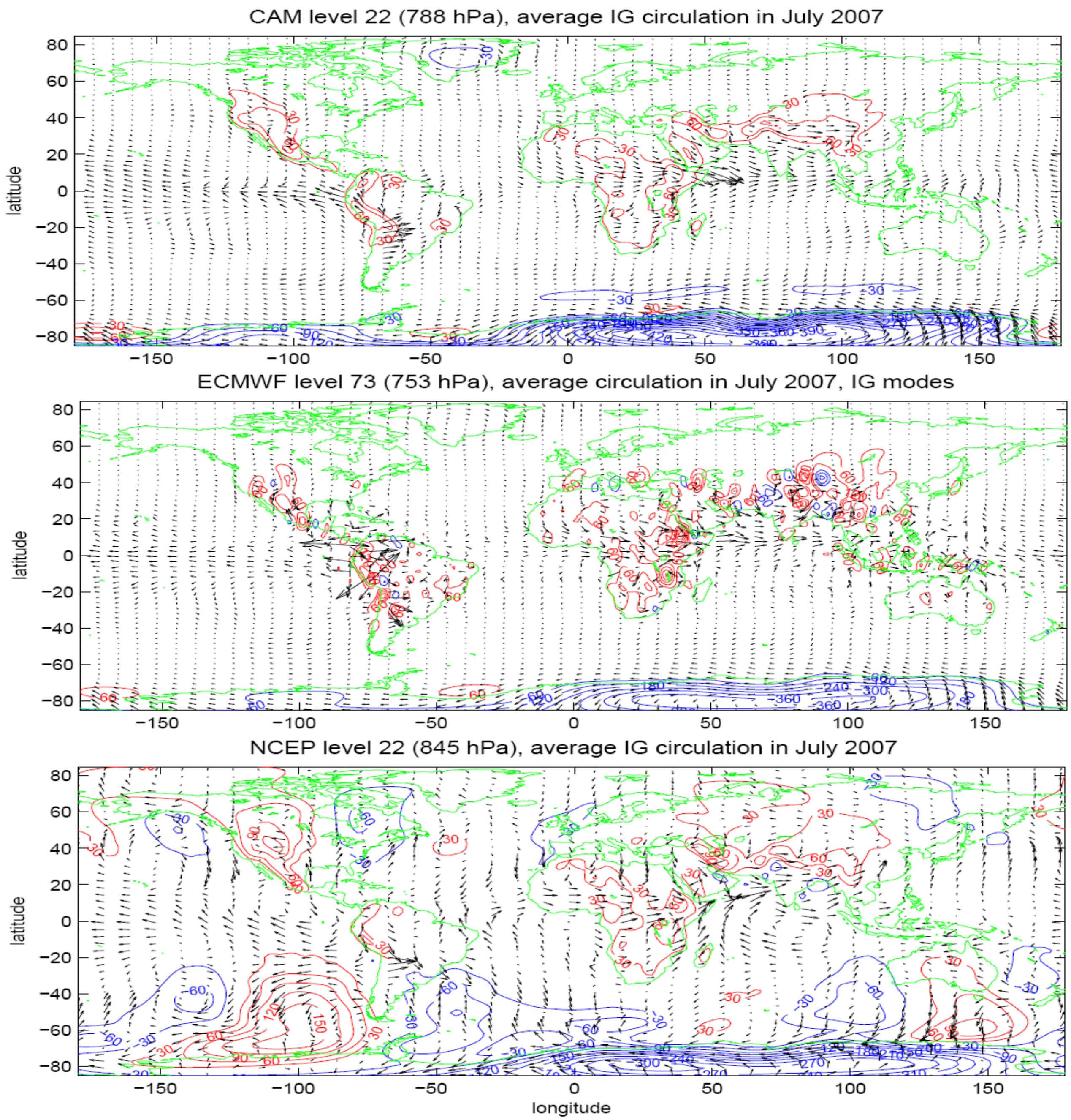
- Reversed flow in the lower and upper troposphere
- Spatial discontinuity of the k=1 signal
- Stronger $k=1$ signal developed by the end of month, especially in the Pacific
- Oscillations of daily period due to observations

How reliable is this Kelvin wave evolution?

DART/CAM uses few observations in the tropics. The assimilation uses flow-derived (multivariate) background-error covariances

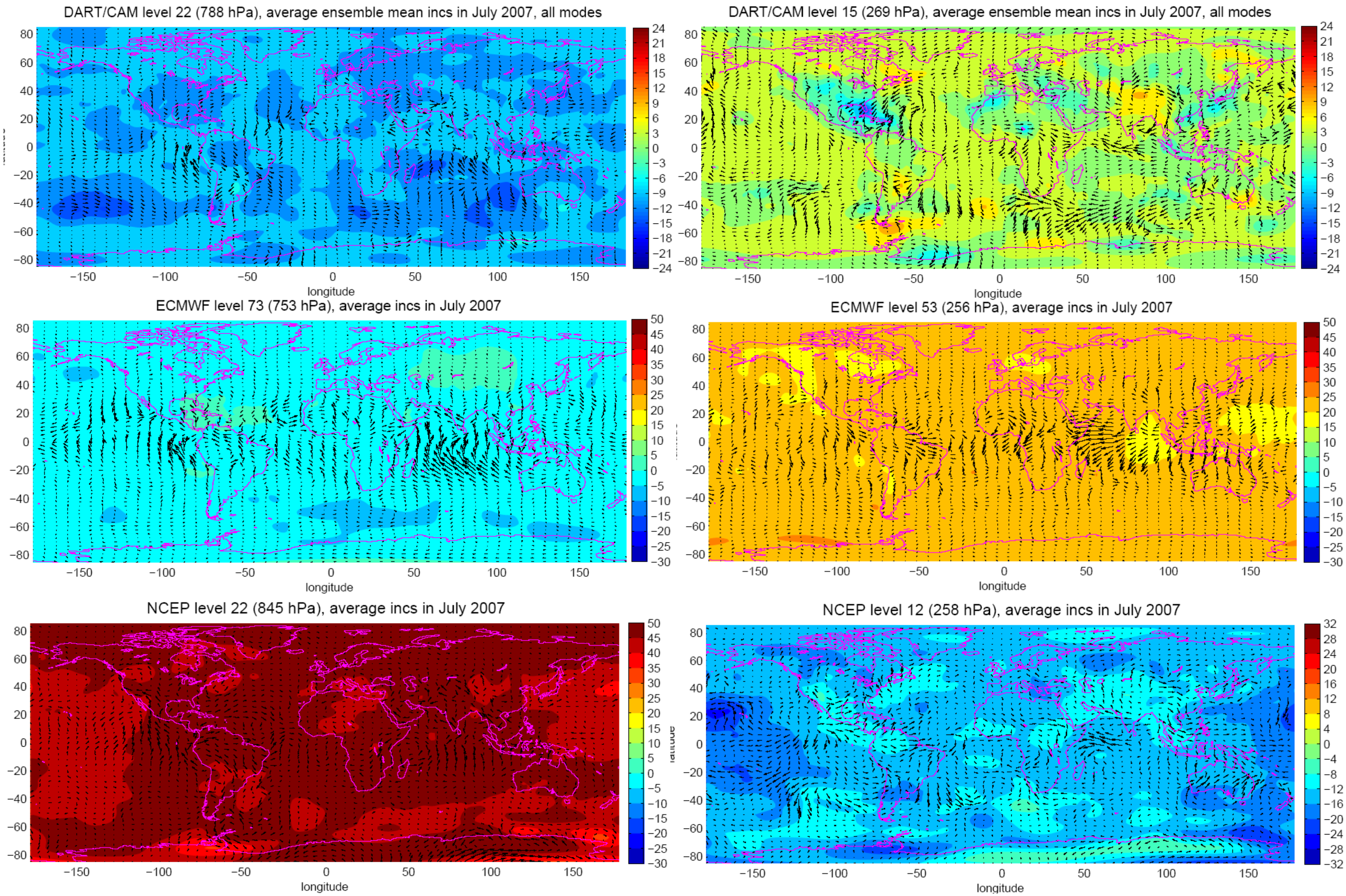
- Inter-comparison with other analyses
- Impact of models' biases
- Estimate of the analysis uncertainty

Average IG motions in July 2007 in the lower troposphere



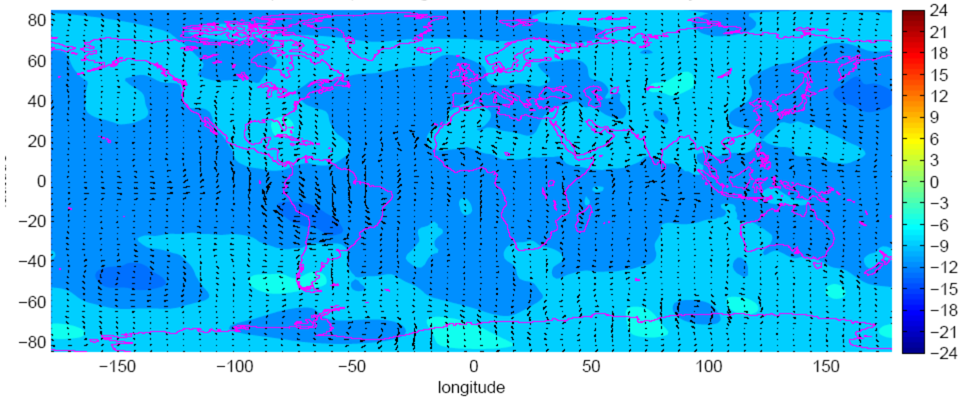
Quantitative comparison for the wind field

Time-averaged analysis increments ~ biases

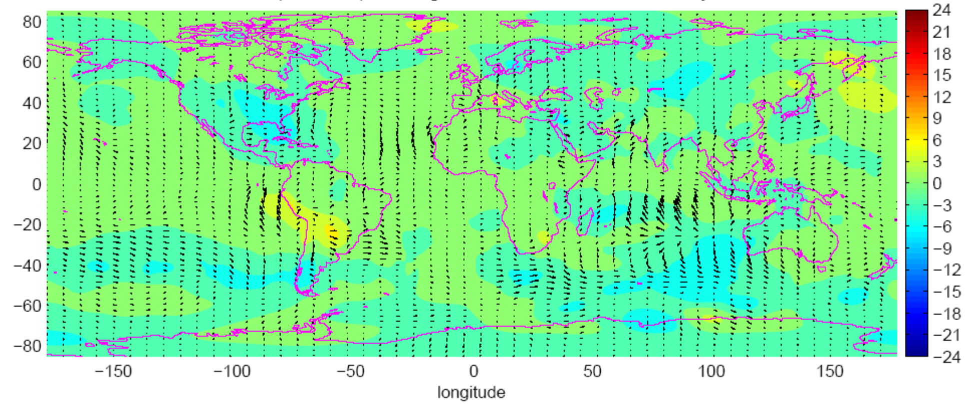


Biases split in ROT and IG modes

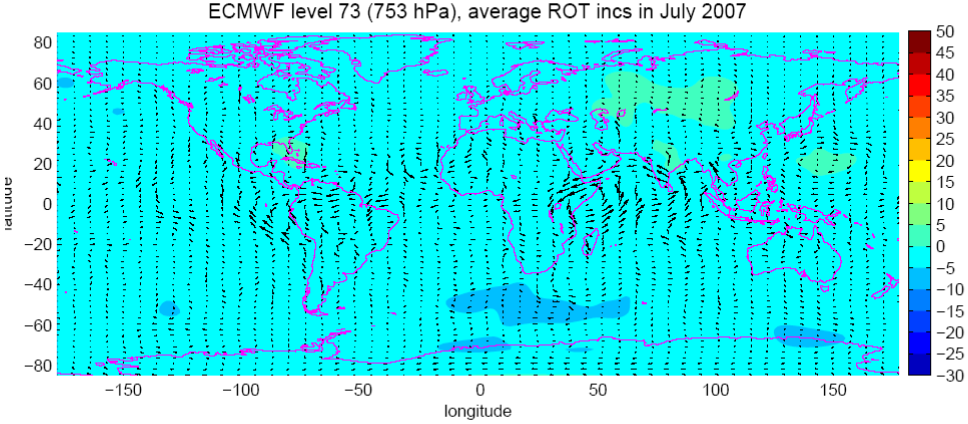
DART/CAM level 22 (788 hPa), average ensemble mean incs in July 2007, ROT modes



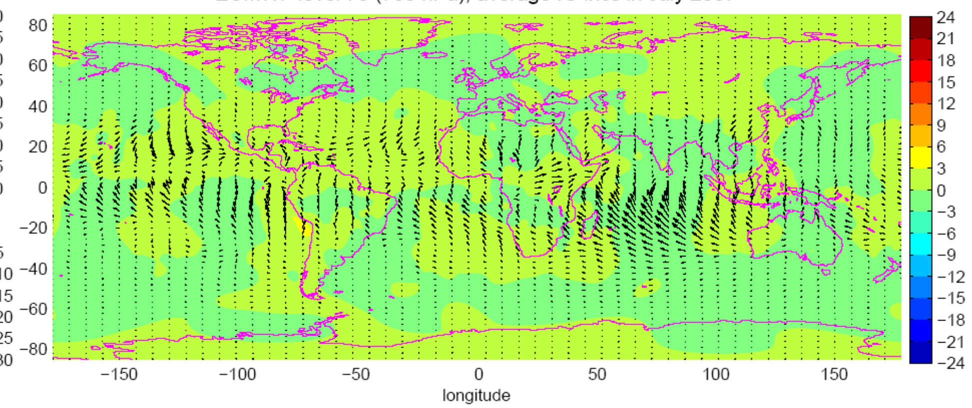
DART/CAM level 22 (788 hPa), average ensemble mean incs in July 2007, IG modes



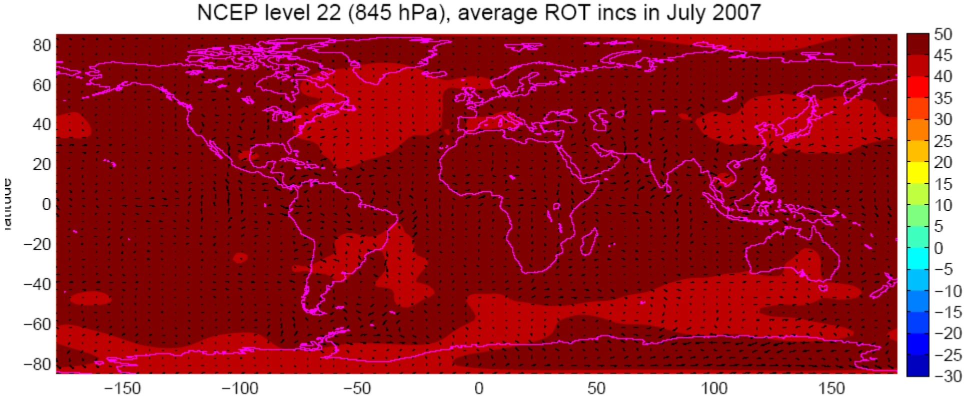
ECMWF level 73 (753 hPa), average ROT incs in July 2007



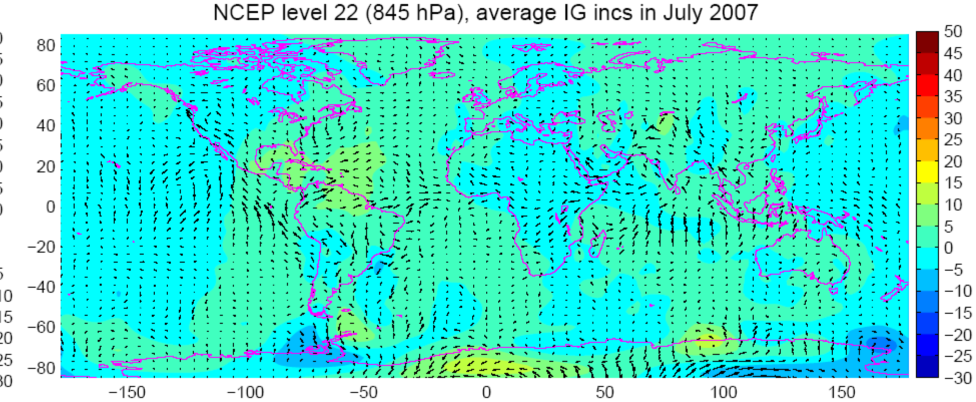
ECMWF level 73 (753 hPa), average IG incs in July 2007



NCEP level 22 (845 hPa), average ROT incs in July 2007

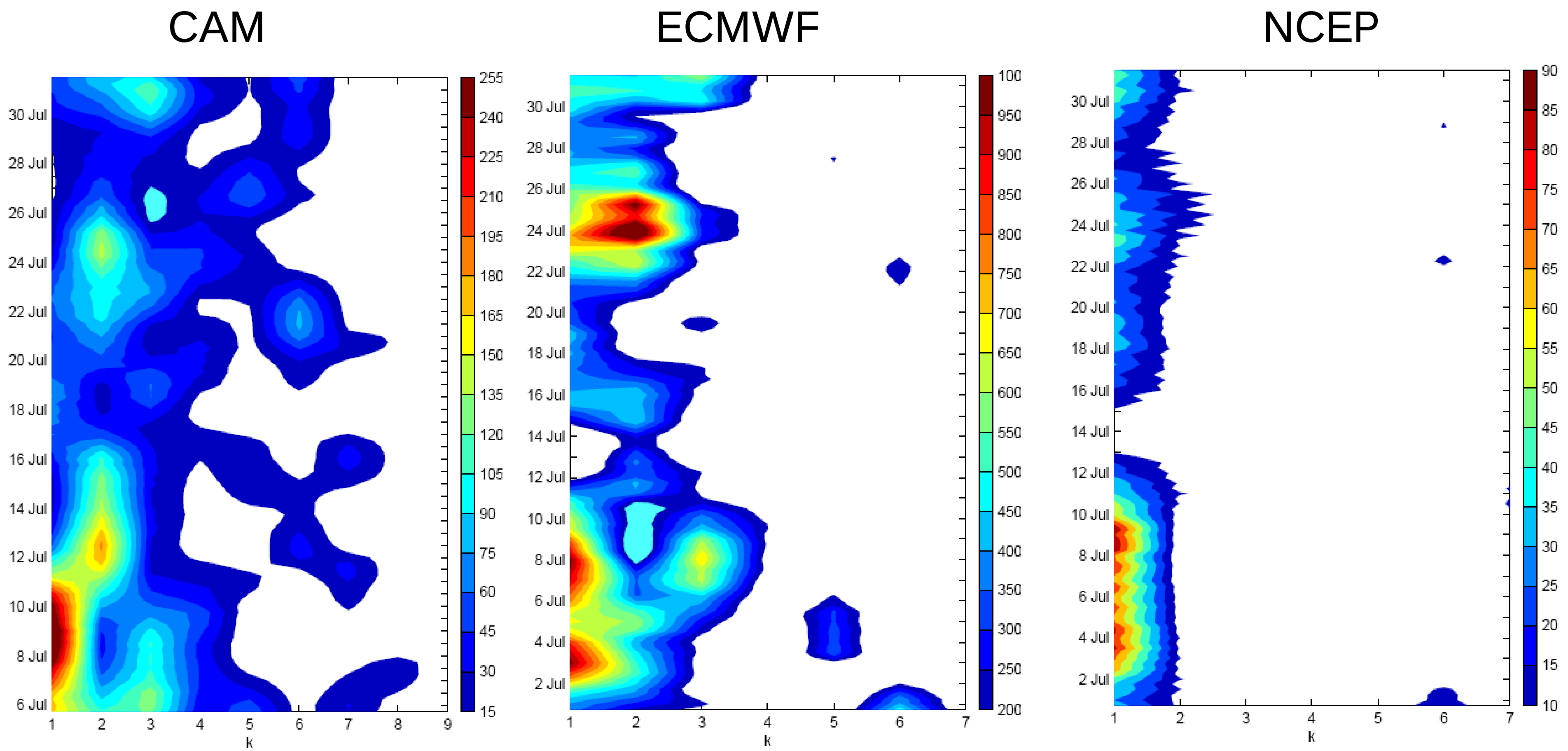


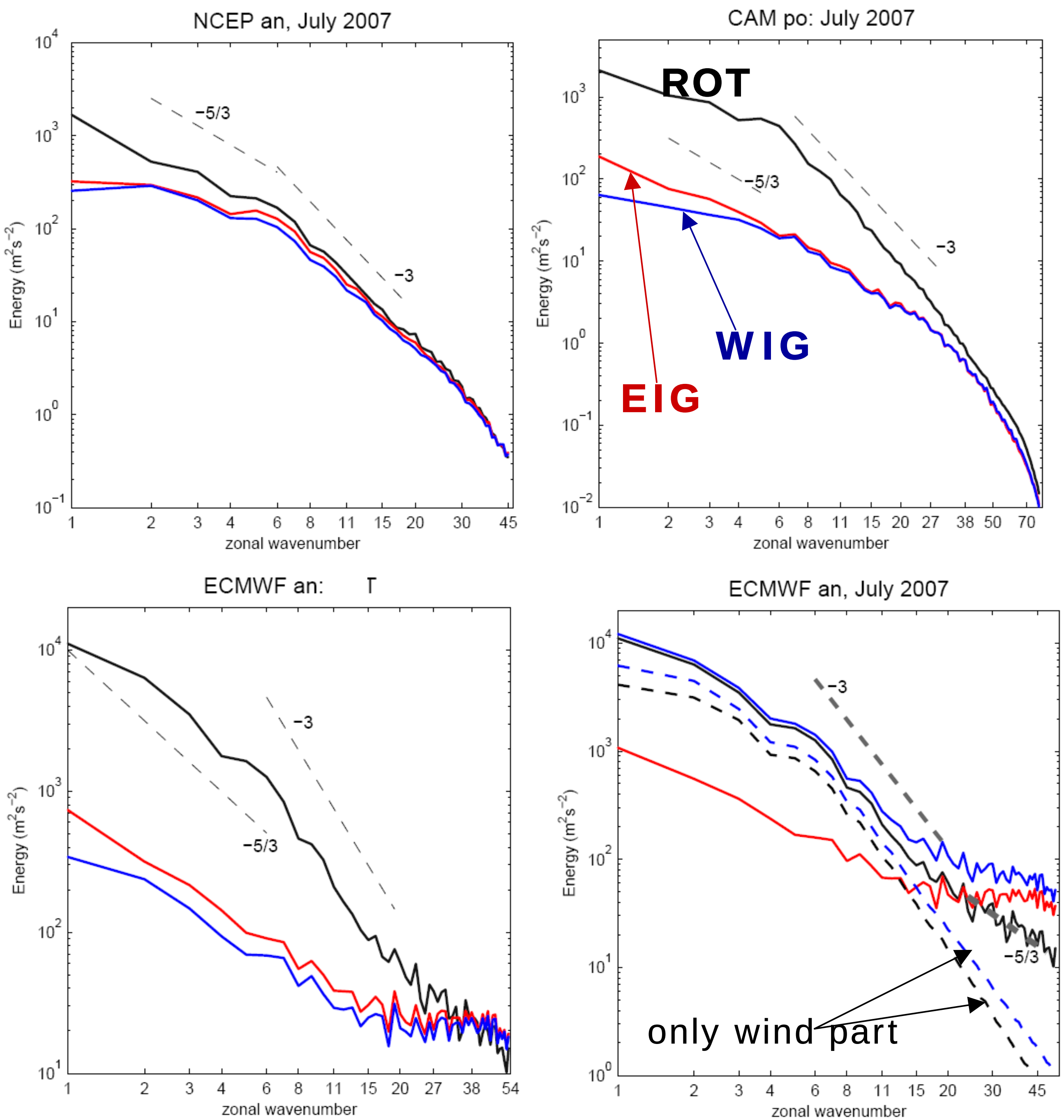
NCEP level 22 (845 hPa), average IG incs in July 2007



Qualitative agreement in most of balanced modes

Example: ROT, m=1, n=3





Analyses inter-comparison

CAM

On average, smallest energy % in IG among the three datasets

ECMWF

n-mode symmetry in EIG-WIG,

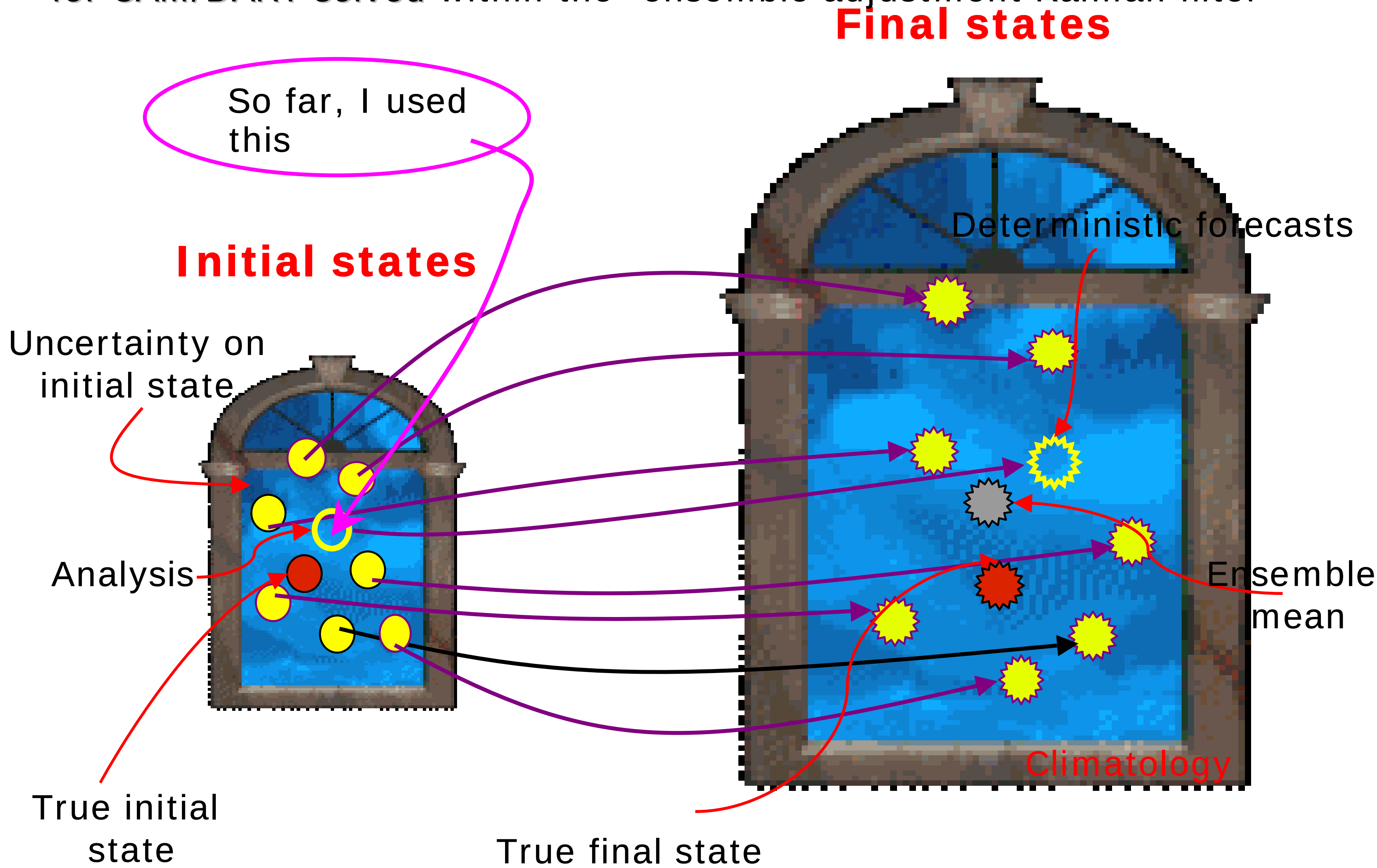
Lowest vertical mode dominant

NCEP

Significant energy % in IG modes also in mid-latitudes

Ensemble assimilation:

for CAM/DART solved within the “ensemble adjustment Kalman filter”



Quantifying uncertainties in CAM analyses

To analyse the uncertainty, each prior and posterior ensemble member projected.

To analyse equivalents of 6-hr forecast errors, departures from the ensemble mean fields projected.

$$\mathbf{X}(\lambda,\varphi,z,t)=(u,v,P)^T \qquad \mathbf{X}(\lambda,\varphi,z,t)=(u-\overline{u},v-\overline{v},P-\overline{P})^T$$

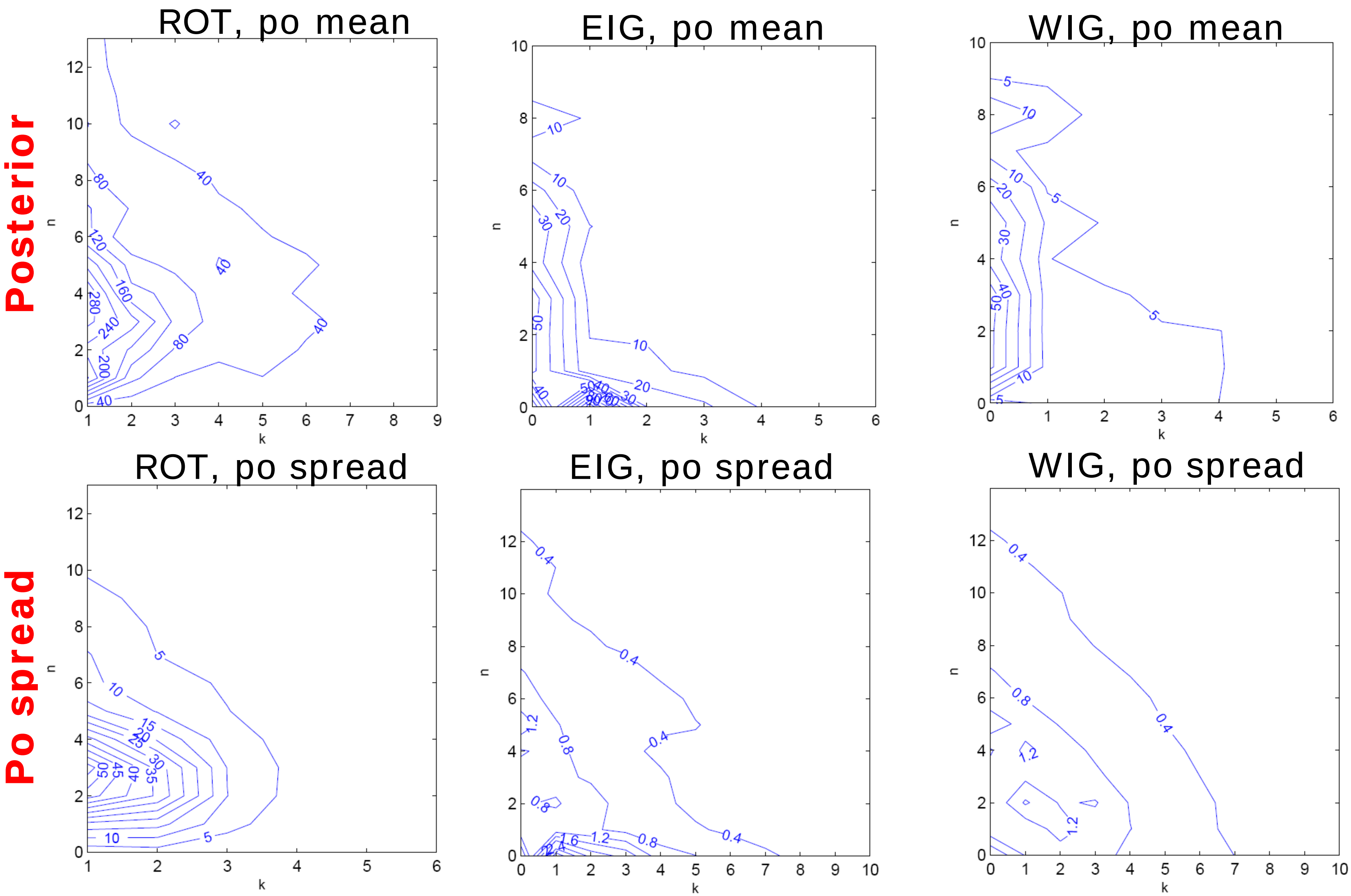
$$\mathbf{X}(\lambda,\varphi,z,t)=\sum_{m=1}^{N_m}\sum_{n_{i=1,2,3}=0}^{N_n-1}\sum_{k=-N_k}^{N_k}\chi_{knm}(t)\mathbf{S}_{\mathbf{m}}\Pi_{knm}(\lambda,\varphi,z)$$

Ensemble size problem accounted for by:

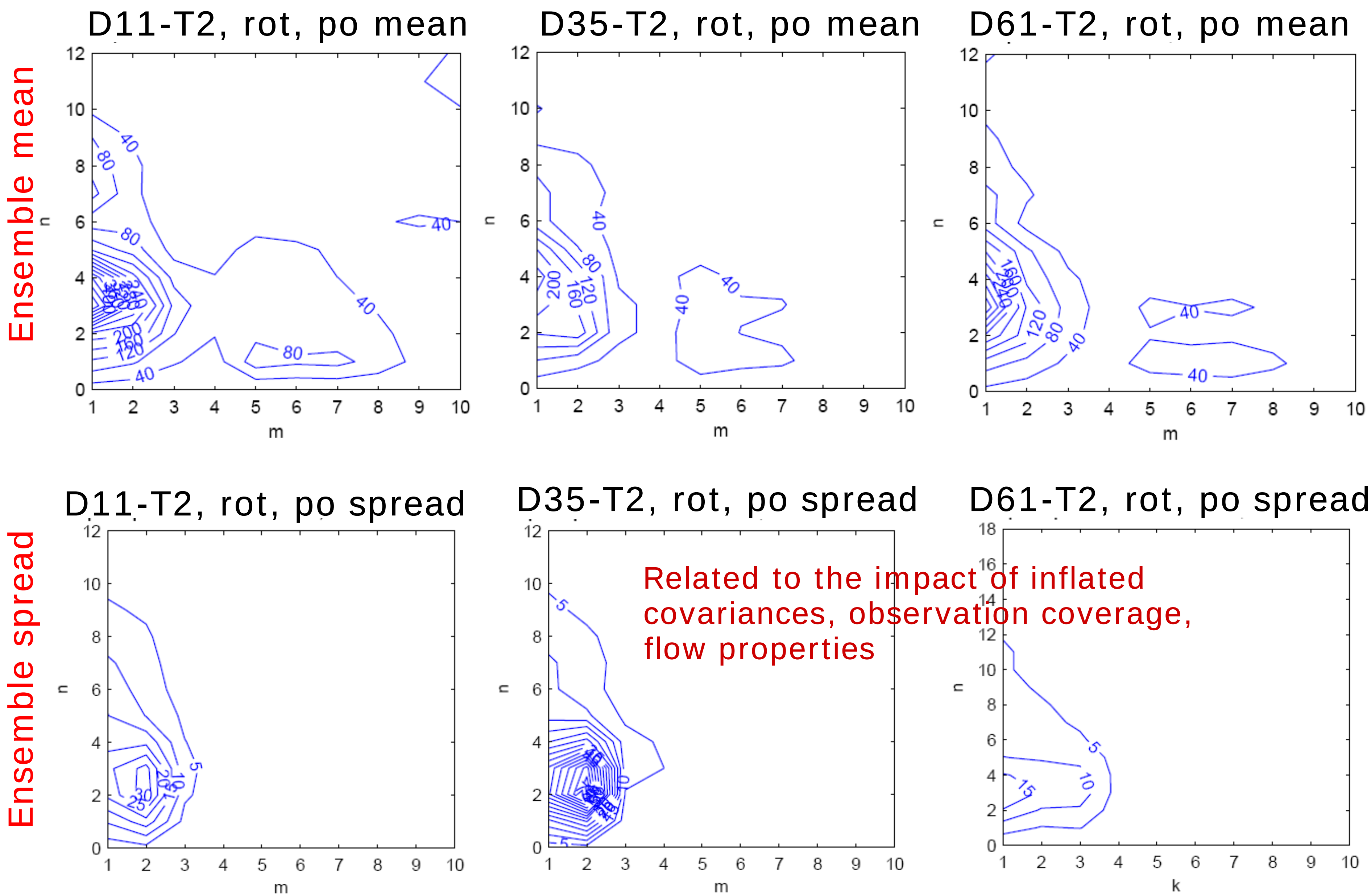
• Covariance localization – reduces the impact of an observation on a state variable by a factor which is a function of their physical distance.

• Covariance inflation – increases the prior ensemble spread leaving the mean and correlations between the variables unchanged (here used is a time constant, spatially varying inflation applied on posterior)

Averaged ens mean and its uncertainty

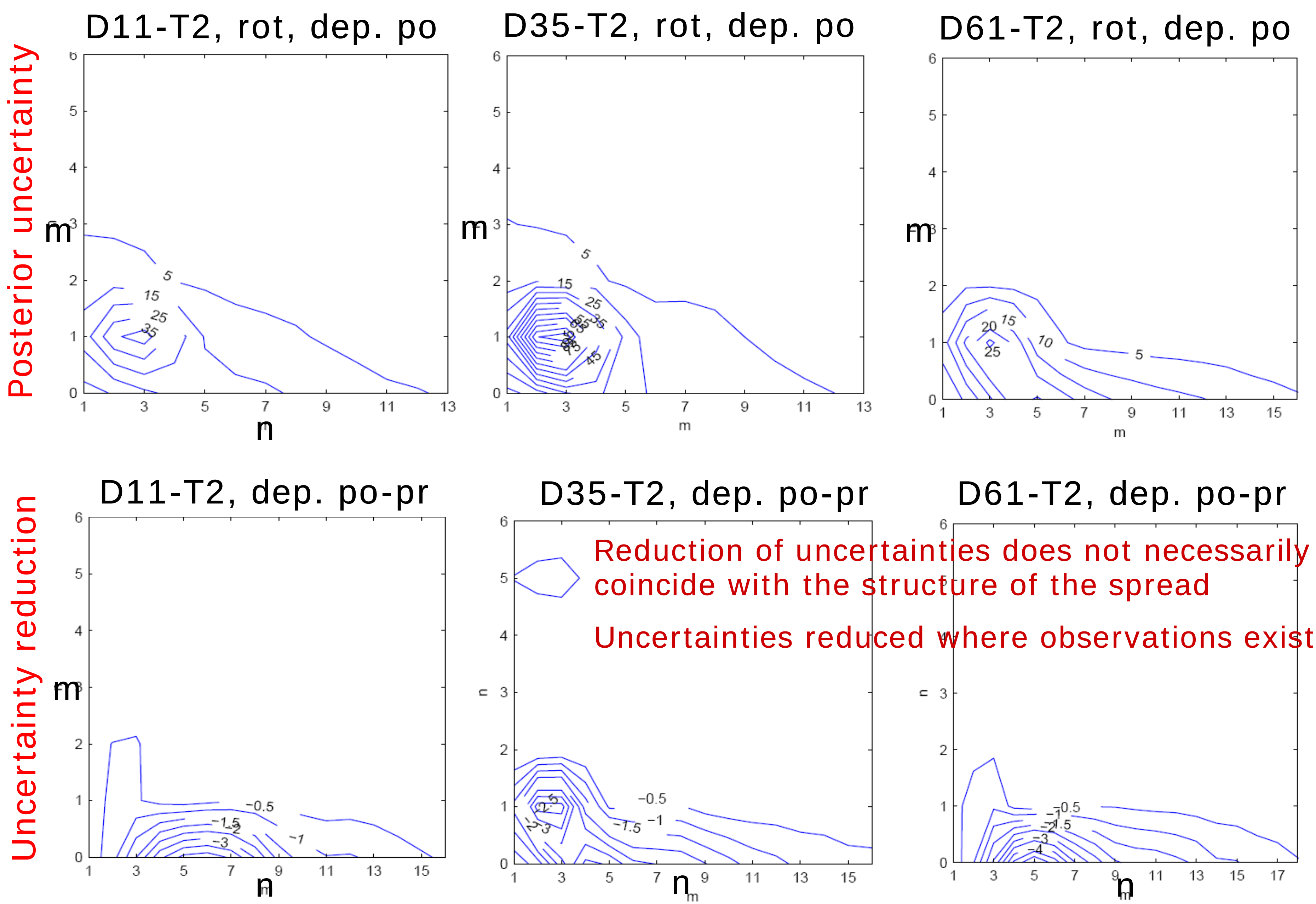


Analysis and its uncertainties: ROT modes

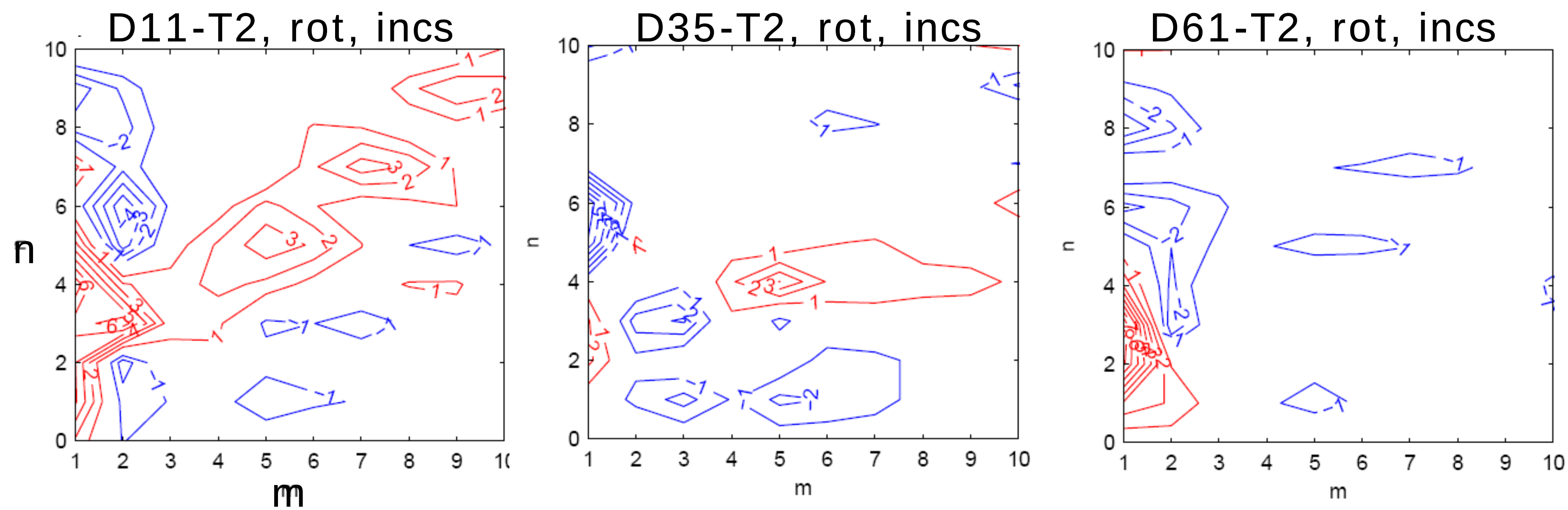


Related to the impact of inflated covariances, observation coverage, flow properties

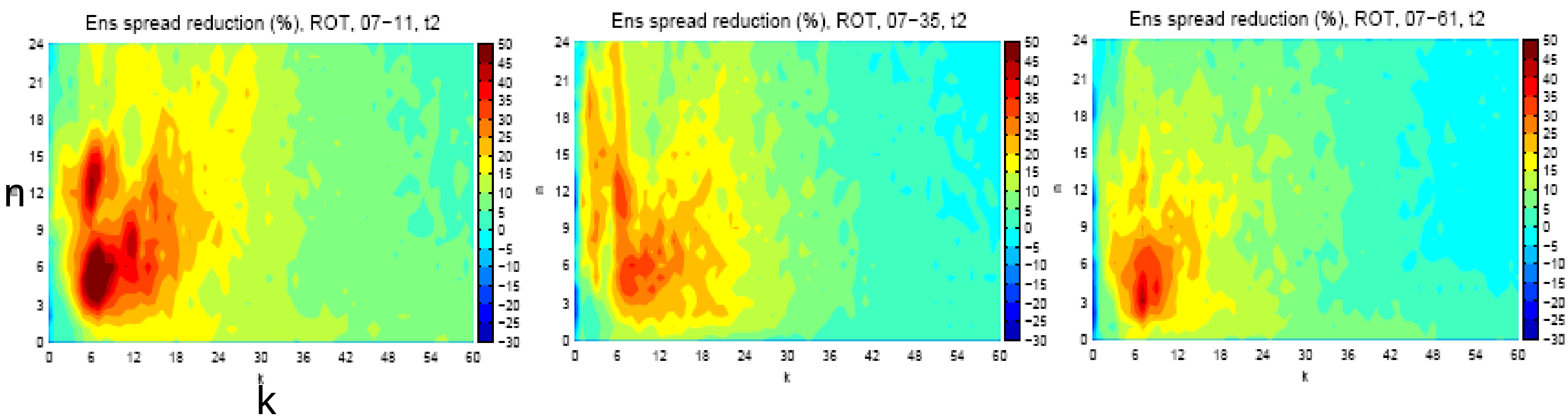
Uncertainty reduction in time



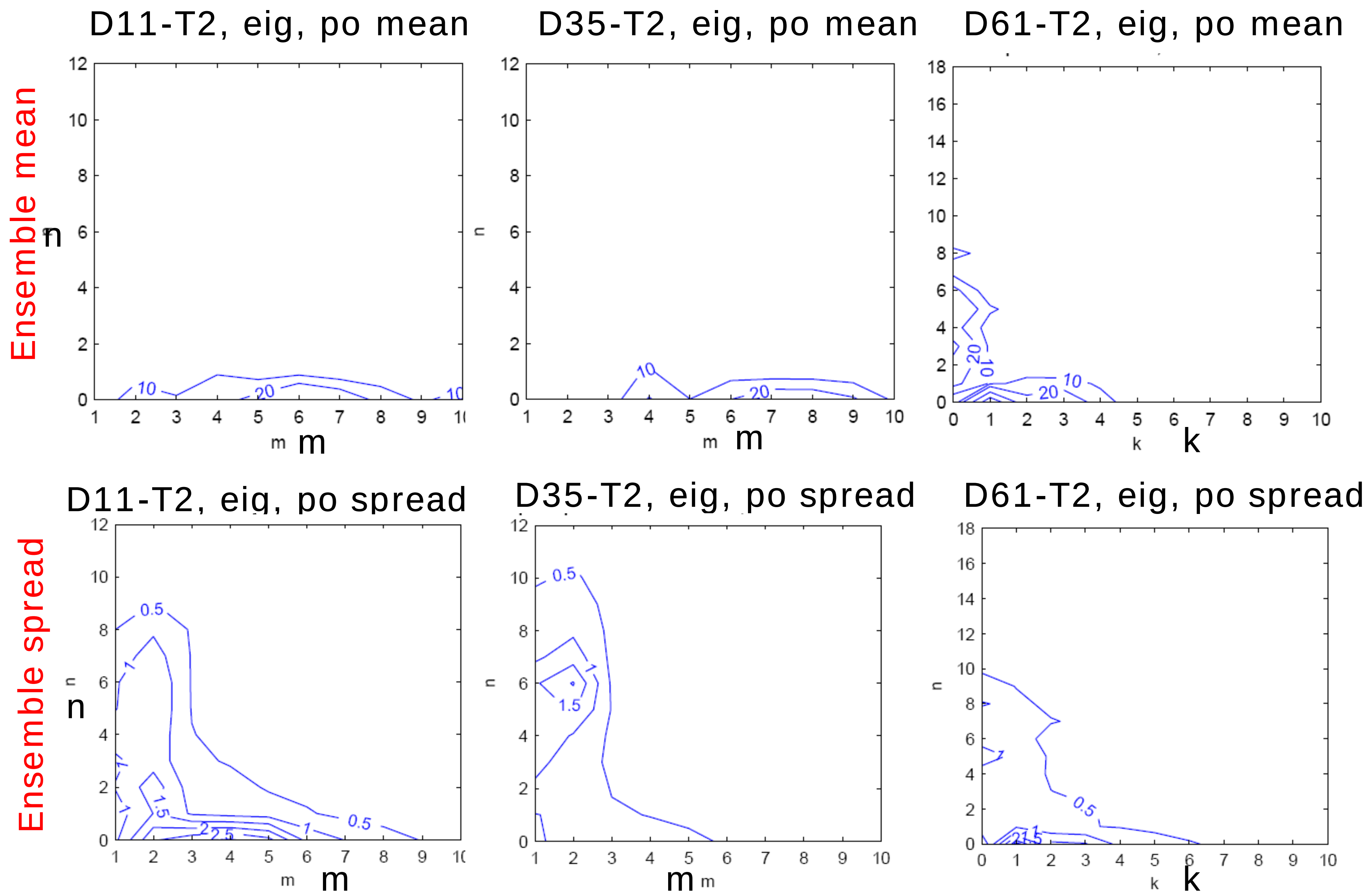
Impact of observations



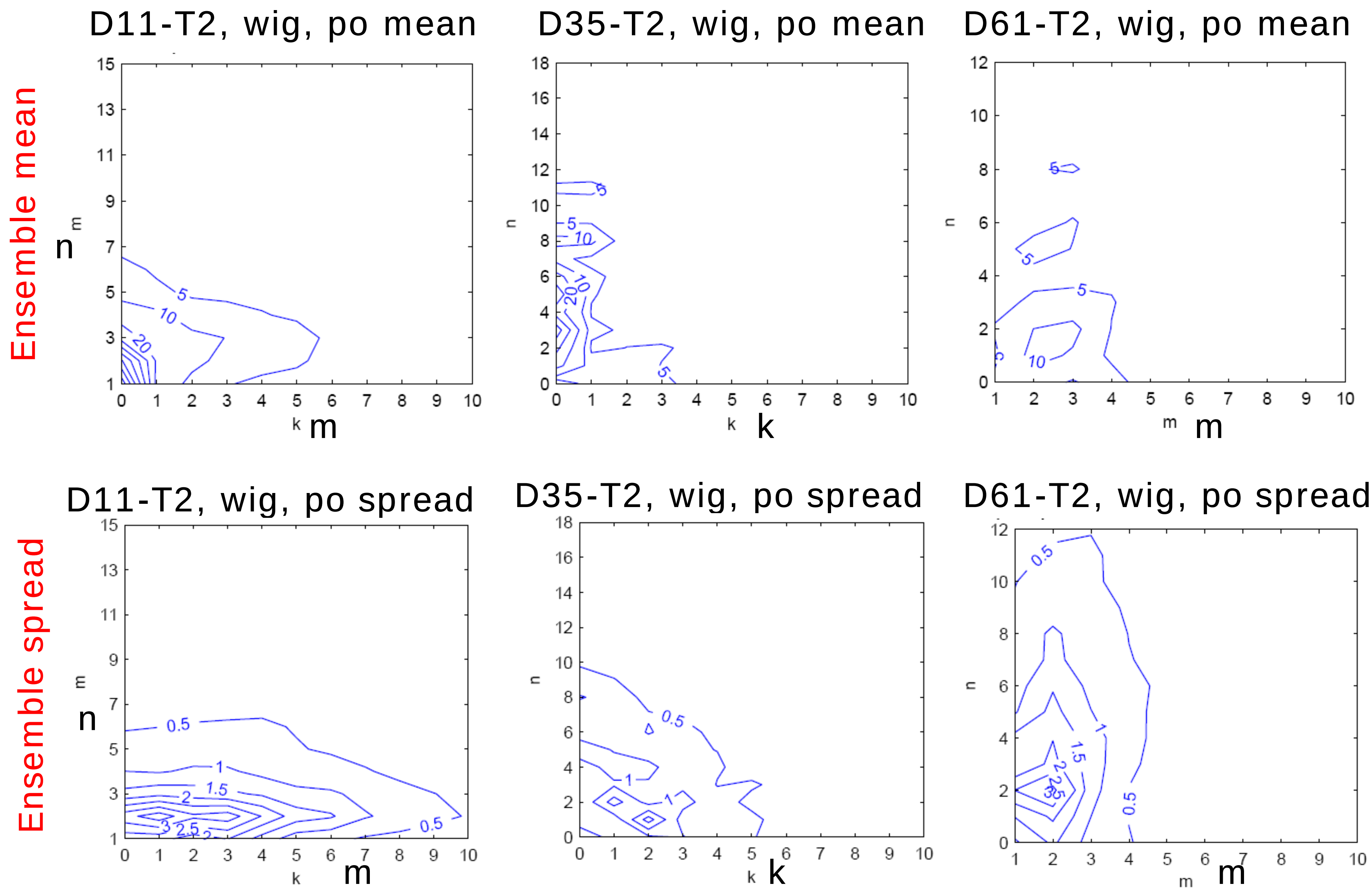
Reduction of the ensemble spread



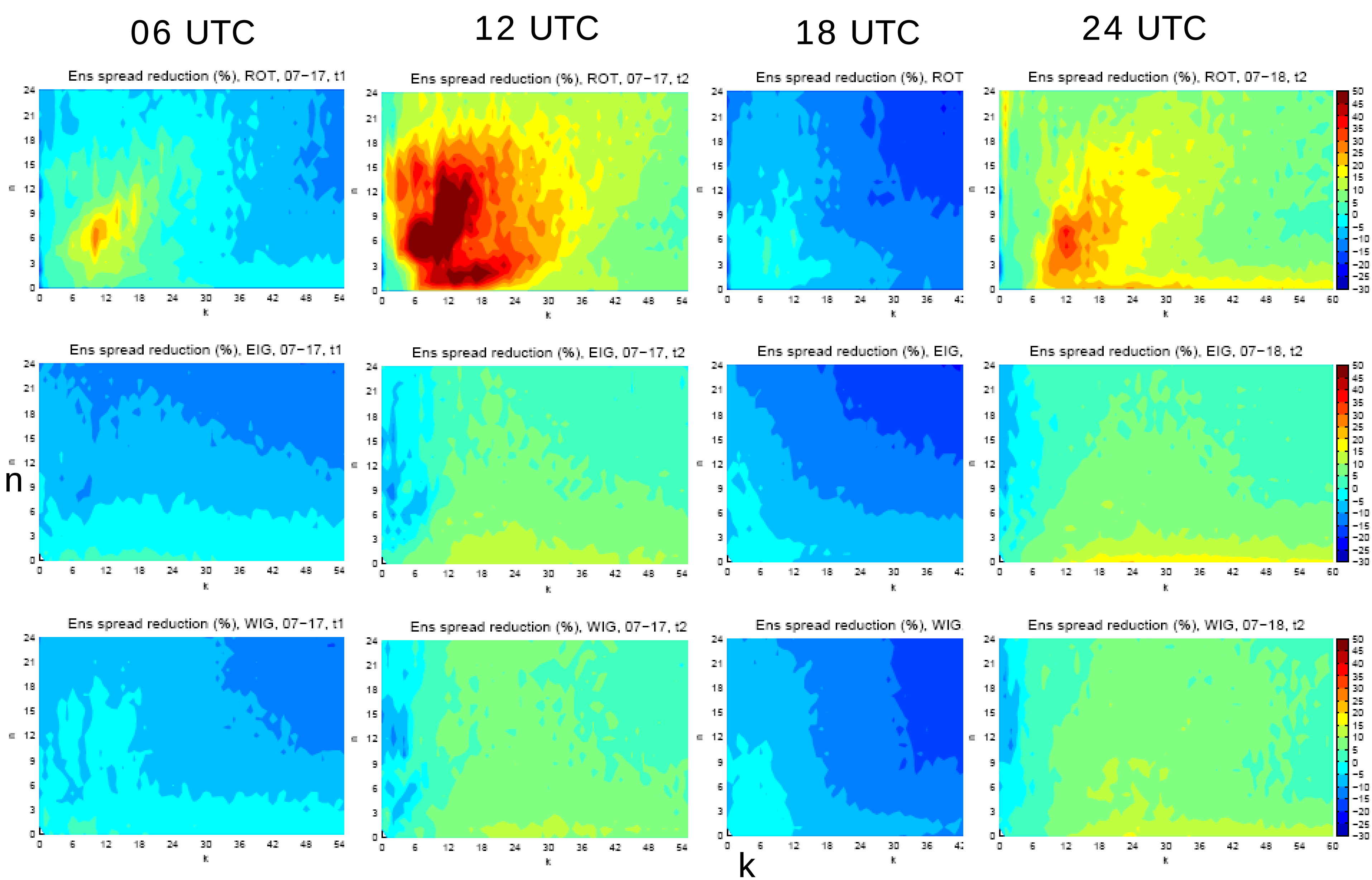
Mean energy and its uncertainties in EIG modes



Analysis uncertainties in WIG modes



Observation non-homogeneity and inflation



Summary

- Tropics are the area with largest uncertainties in existing analysis datasets. Tropics are the area with largest biases in three studied data assimilation systems

- Normal mode expansion allows to quantify energy in various motions and to modify traditional view of inertio-gravity motions as junk. With normal modes it is possible to quantify variance in various tropical divergent motions and its relevance for data assimilation.

- Application of normal modes offers a physically attractive approach to the quantification of uncertainties in analyses and forecasts. It points out the scales and motion types most affected by the inflation, localization, observations and model biases.

- Uncertainties vary in time and space, thus an argument for a flow-dependent covariance matrix for the forecast errors. The normal mode application may also help to address modeling aspects such as model-error covariances and the initialization.