The ECMWF model, progress and challenges

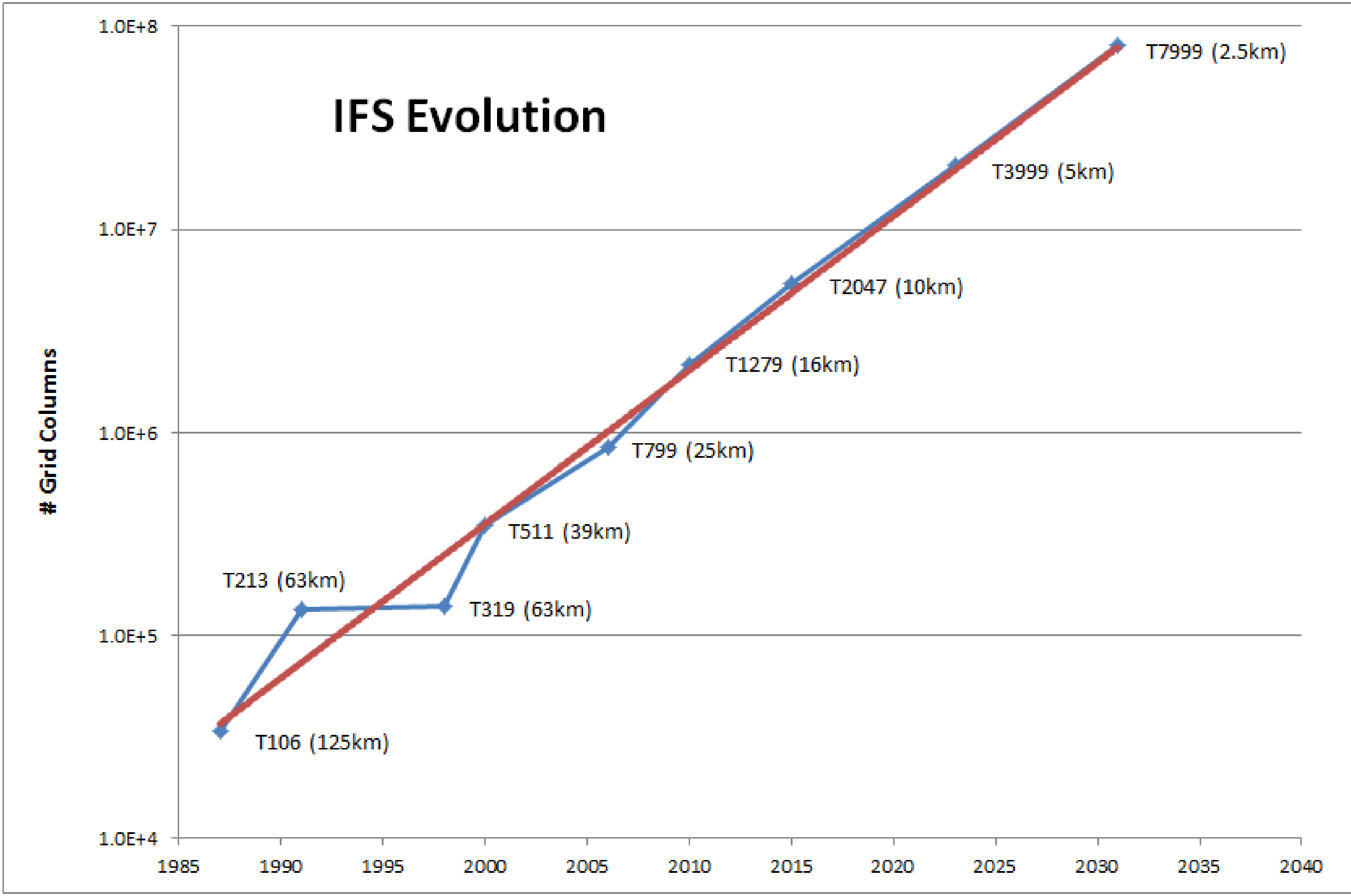
Nils Wedi*, Mats Hamrud, George Mozdzynski

*European Centre for Medium-range Weather Forecasts, wedi@ecmwf.int

Many thanks to Linus Magnusson and Jean Bidlot

Technology

- **4 foci:**
  - **Enabling technology for computations on massively parallel computers for science today**
  - **A cost-effective, low-energy consuming but highly accurate model and data assimilation system**
  - **Code resilience and understanding, what does each part of the model contribute, where is accuracy required, where not ?**
  - **Quantitative measures on how good the forecast is**


IFS Evolution
Grid Columns
1.0E+8
1.0E+7
1.0E+6
1.0E+5
1.0E+4
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
T7999 (2.5km)
T3999 (5km)
T2047 (10km)
T1279 (16km)
T799 (25km)
T511 (39km)
T213 (63km)
T319 (63km)
T106 (125km)

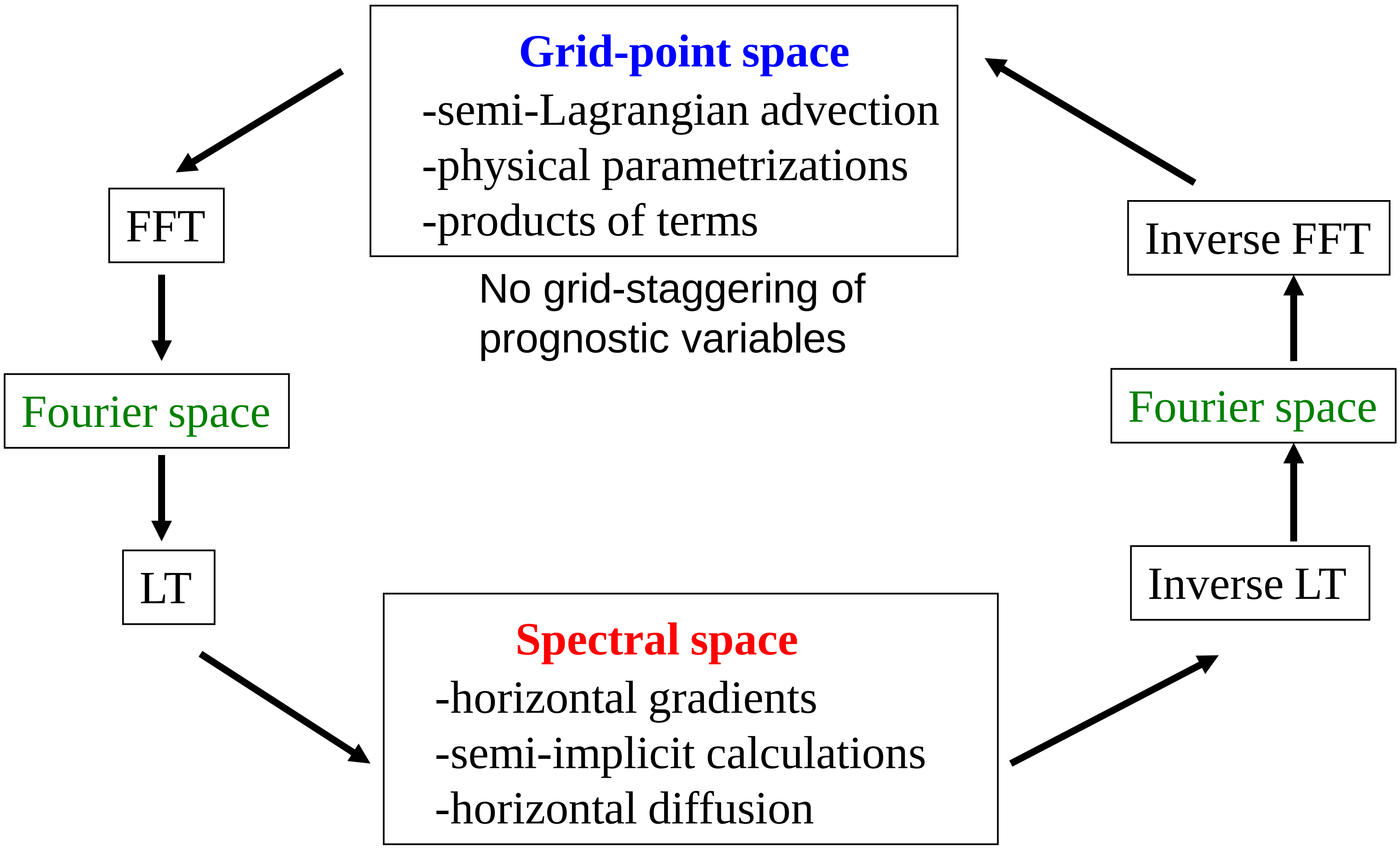
ECMWF

The Integrated Forecasting System (IFS)

technology applied at ECMWF for the last 30 years …

A spectral transform, semi-Lagrangian, semi-implicit (compressible) (non-)hydrostatic model

Schematic description of the spectral transform method in the ECMWF IFS model



FFT: Fast Fourier Transform, LT: Legendre Transform

Direct spectral transform (Forward)

Fourier transform:

$$\zeta_m(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \zeta(\lambda,\theta)e^{-im\lambda}d\lambda$$

FFT (fast Fourier transform)
using
$N_F \geq 2N+1$
points (linear grid)
(3N+1 if quadratic grid)

Legendre transform:

$$\zeta_n^m = \frac{1}{2}\int_{-1}^{1} \zeta_m \overline{P_n^m}(cos(\theta))dcos(\theta).$$

$$\zeta_n^m = \sum_{k=1}^{K} w_k \zeta_m(x_k)\overline{P_n^m}(x_k)$$

Direct Legendre transform
by Gaussian quadrature
using $N_L \geq (2N+1)/2$
"Gaussian" latitudes (linear grid)
((3N+1)/2 if quadratic grid)

$$w_k = \frac{2N+1}{[\overline{P_N^{m=1}}(x_k)]^2}.$$

(normalized) associated Legendre polynomials

Inverse spectral transform (Backward)

Inverse Legendre transform

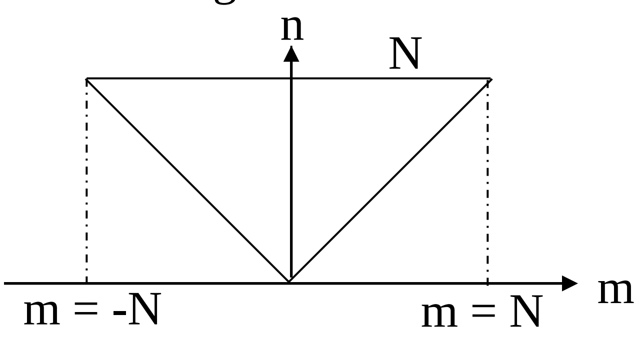
$$\zeta(\theta,\lambda) = \sum_{m=-N}^{N} e^{im\lambda} \sum_{n=|m|}^{N} \zeta_n^m \overline{P_n^m}(cos(\theta)),$$

$$\zeta(\lambda,\mu,\eta,t) = \sum_{m=-N}^{N} \sum_{n=|m|}^{N} \zeta_n^m(\eta,t) Y_n^m(\lambda,\mu)$$

Triangular truncation (isotropic)

Spherical harmonics

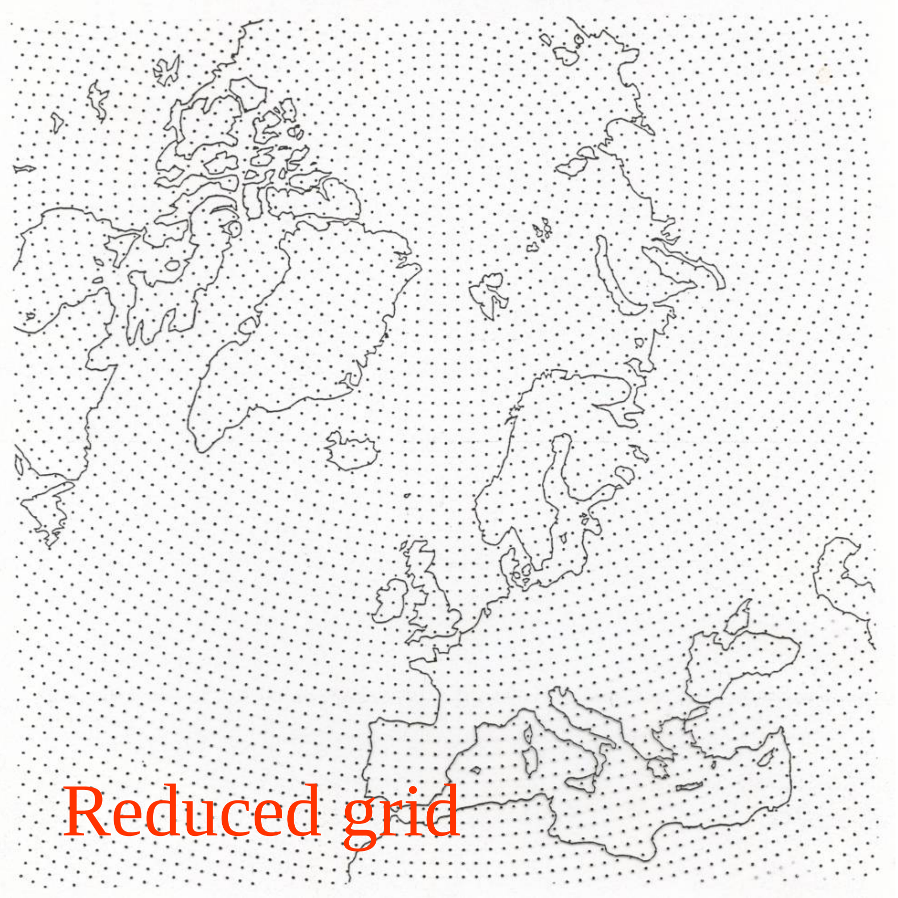
Triangular truncation:



ECMWF

The Gaussian grid

About 30% reduction in number of points



Reduction in the number of Fourier points at high latitudes is possible because the associated Legendre polynomials are very small near the poles for large *m*.

Note: number of points nearly equivalent to quasi-uniform icosahedral grid cells of the ICON model.

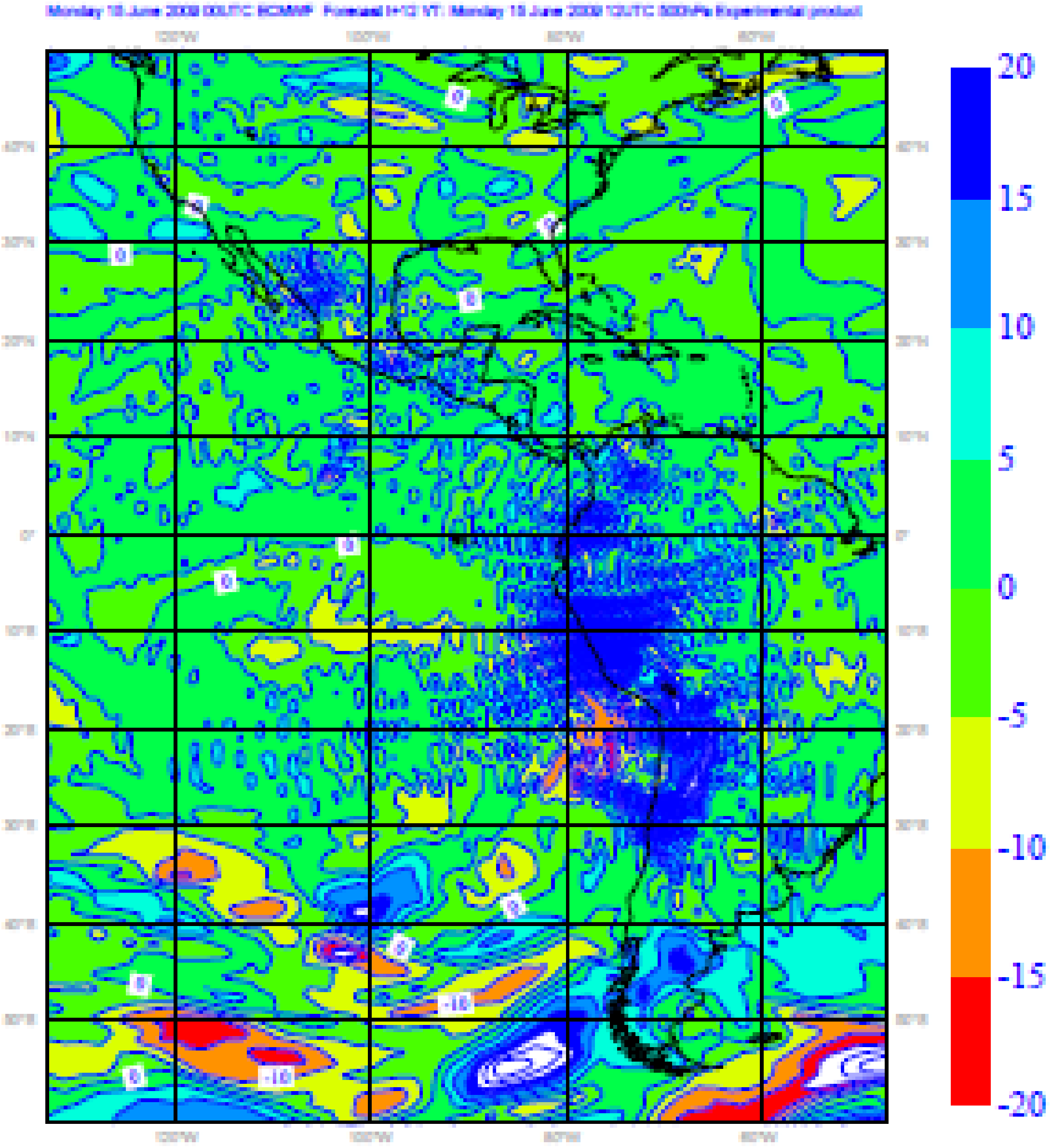
Aliasing

- **Aliasing of quadratic terms on the linear grid (2N+1 gridpoints per N waves), where the product of two variables transformed to spectral space cannot be accurately represented with the available number of waves (as quadratic terms would need a 3N+1 ratio).**
- **Absent outside the tropics in E-W direction due to the design of the reduced grid (obeying a 3N+1 ratio) but present throughout (and all resolutions) in N-S direction.**
- ***De-aliasing in IFS*: By subtracting the difference between a specially filtered and the unfiltered pressure gradient term at every time-step the stationary noise patterns can be removed at a cost of approx. 5% at T1279 (2 extra transforms).**

De-aliasing

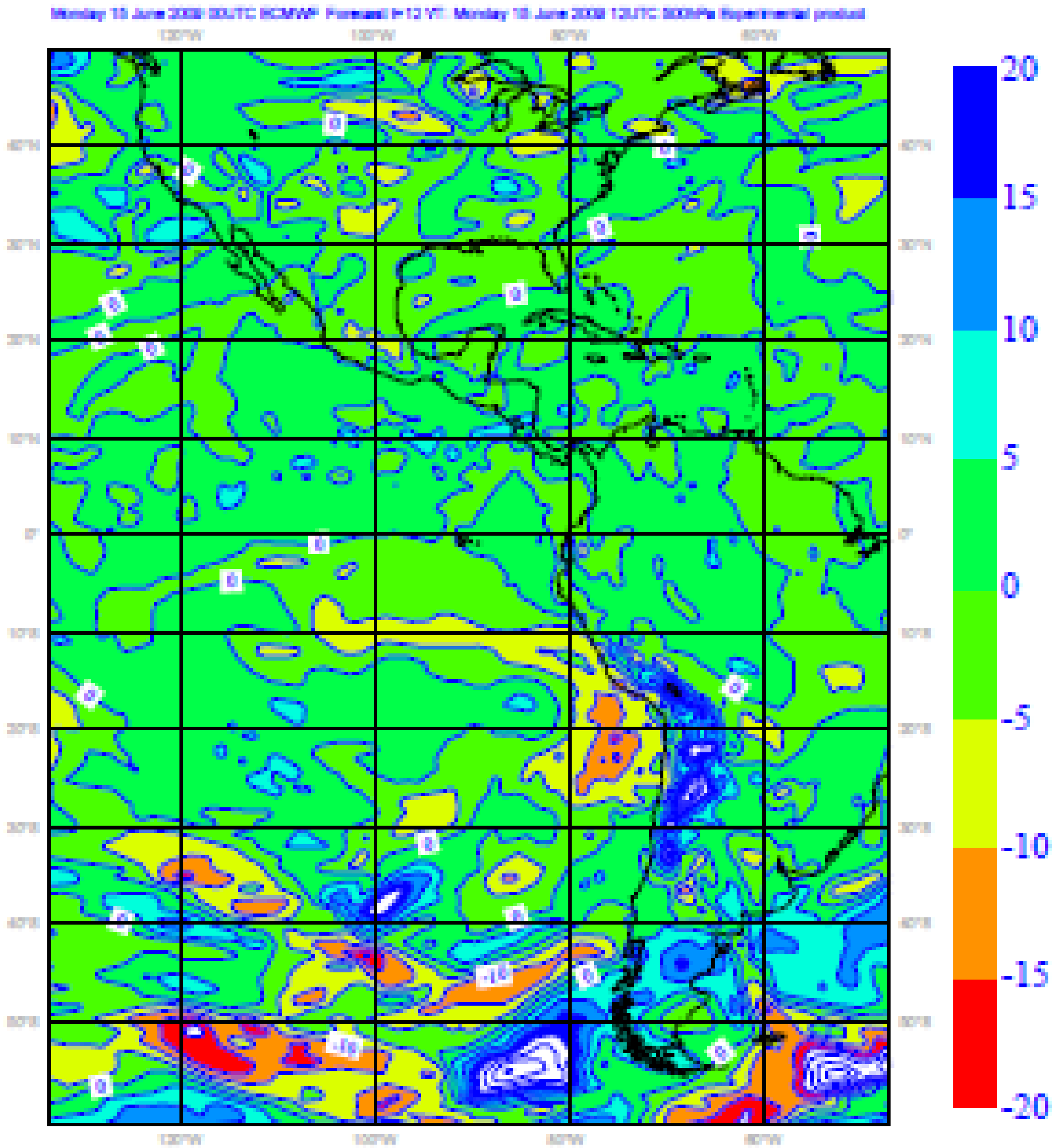
E-W

500hPa adiabatic zonal wind tendencies (T159)



ECMWF

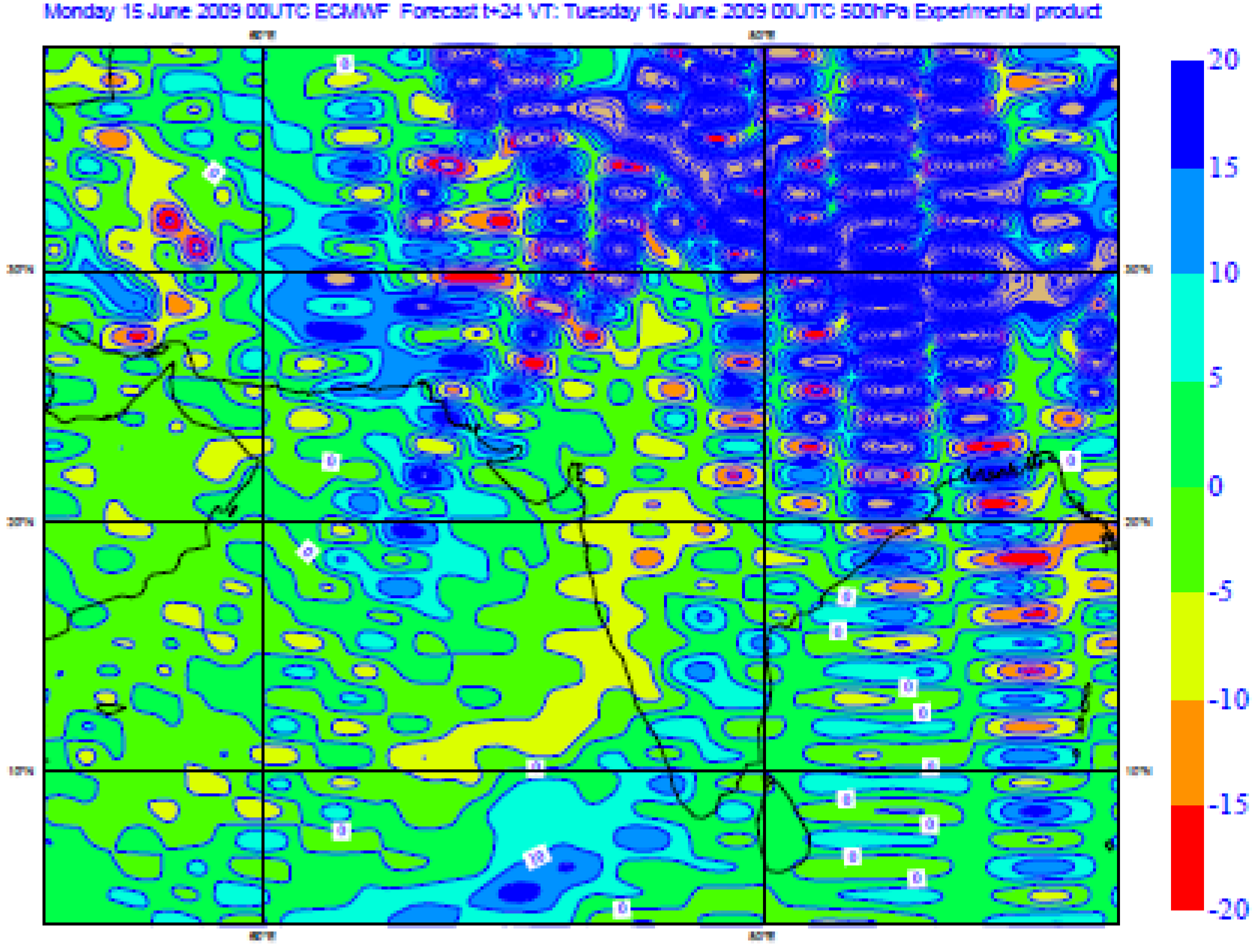
De-aliasing



De-aliasing

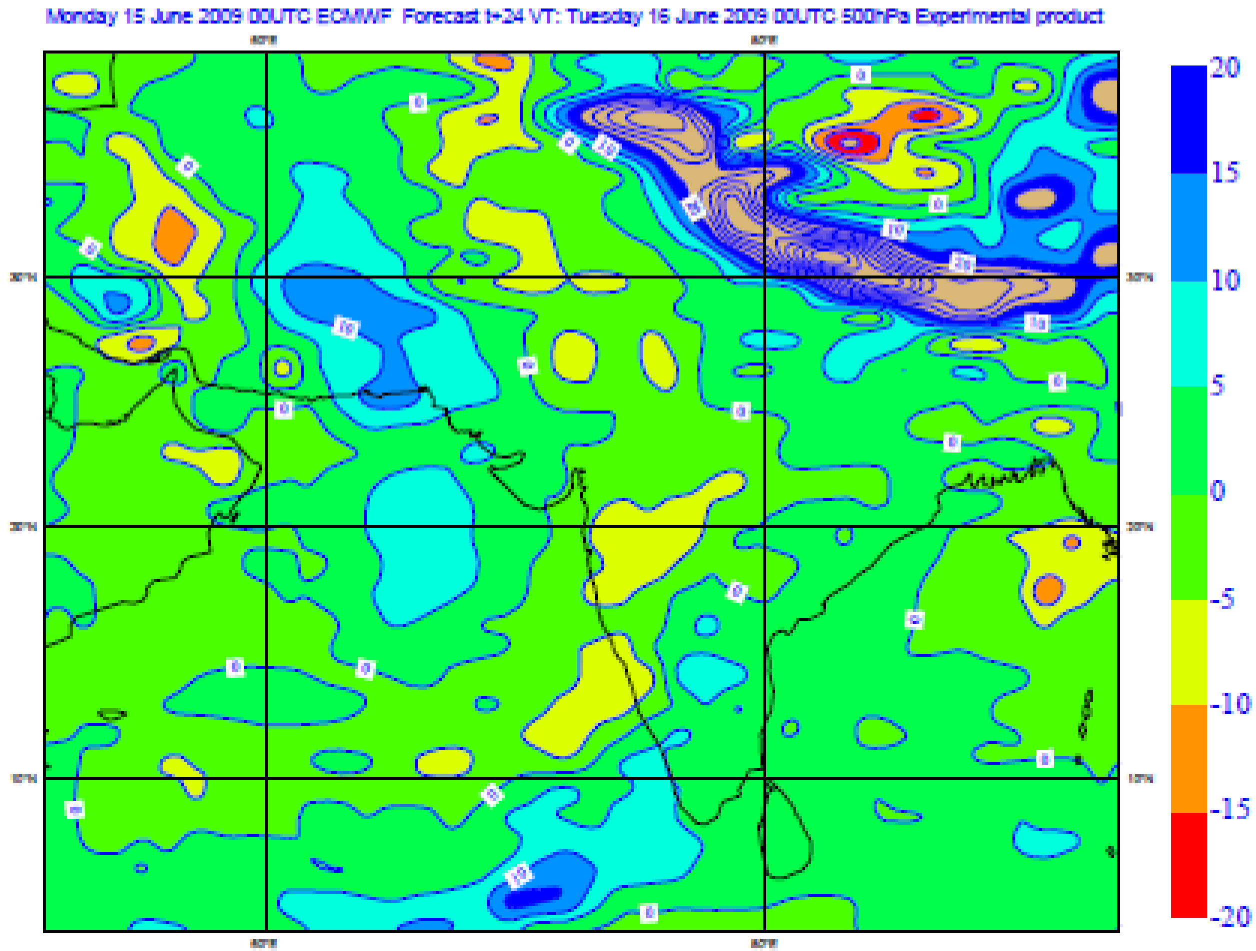
N-S

500hPa adiabatic meridional wind tendencies (T159)



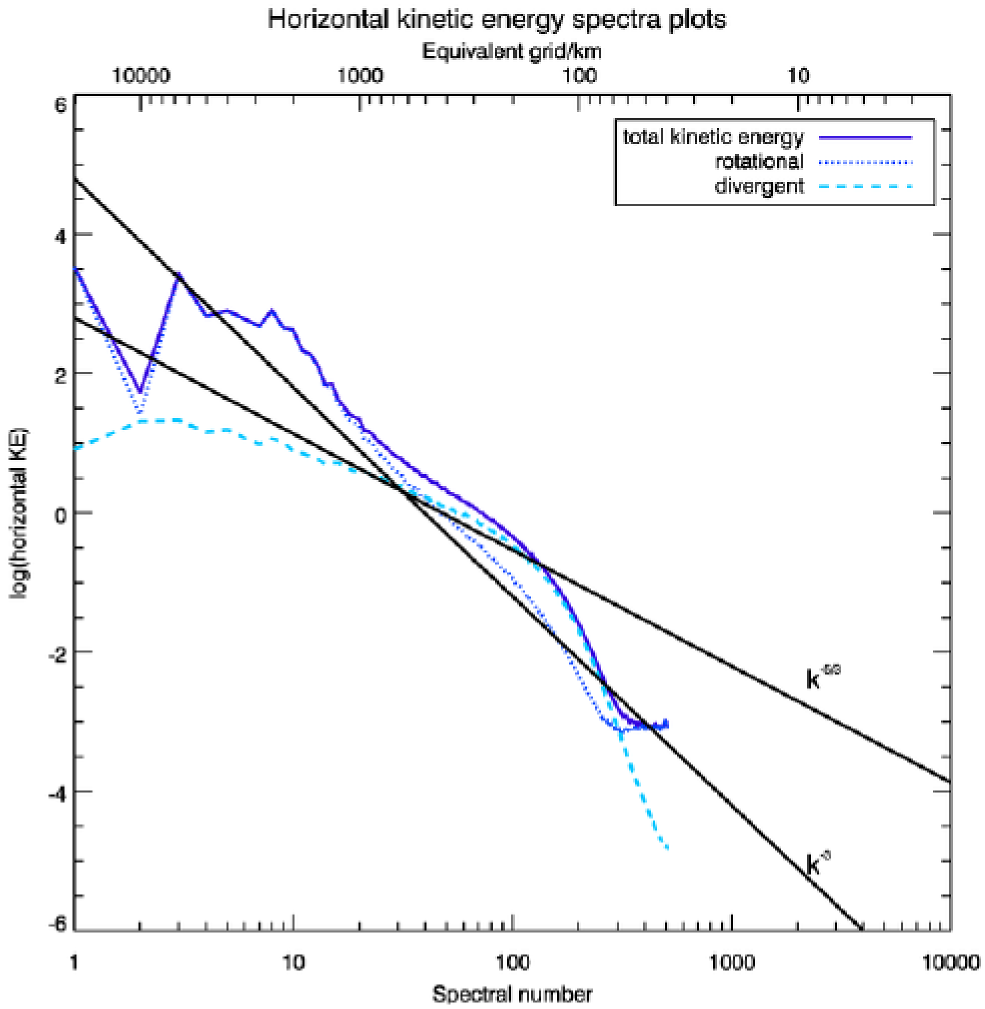
ECMWF

De-aliasing



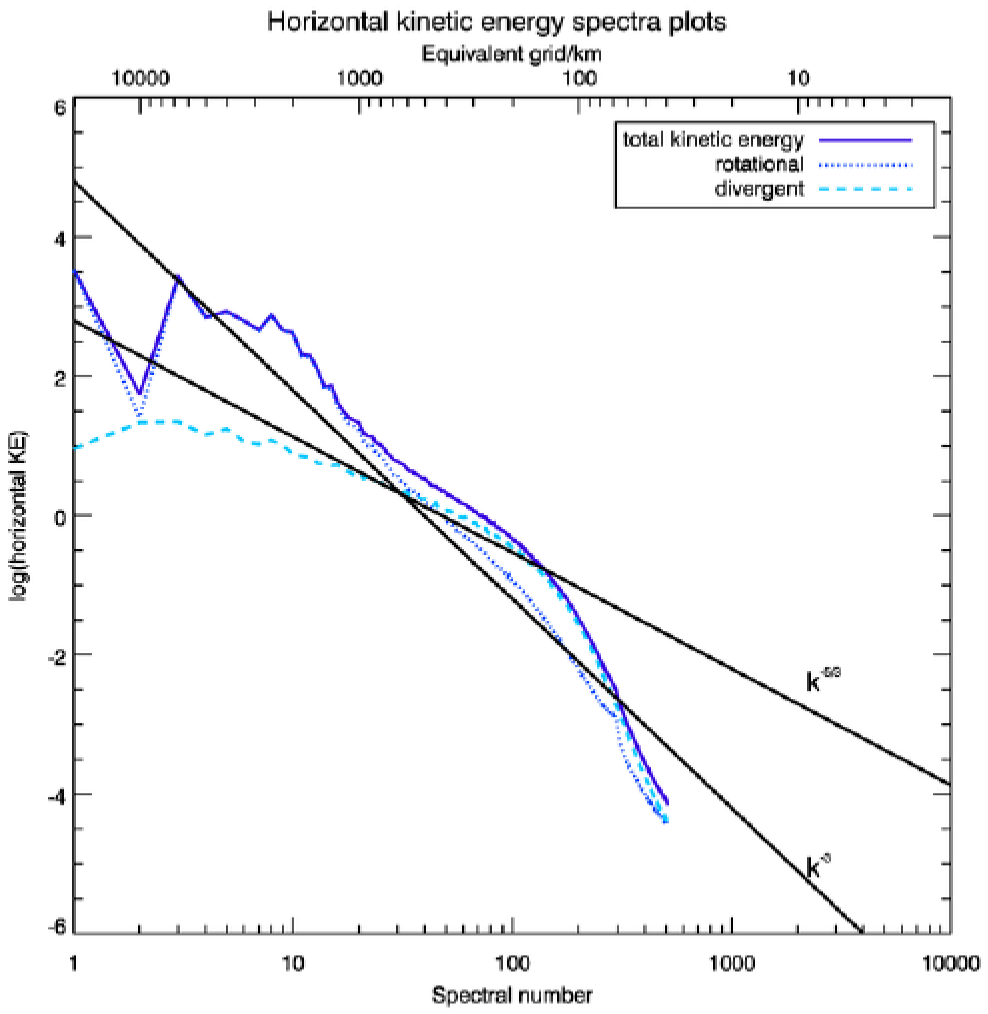
ECMWF

Kinetic Energy Spectra – 100 hPa



ECMWF

Kinetic Energy Spectra – 100 hPa



ECMWF

Fast Multipole Method (FMM) and spectral filtering *(Jakob-Chien and Alpert, 1997; Tygert 2008)*

$$f_j = \sum_{k=1}^{N} \frac{\beta_j P_k}{\tilde{\mu}_j - \mu_k}$$

For all j=1,..,J

FMM: We can do above sum for all points j in O(J+N) operations instead of O(J*N) !

Example: From Christoffel-Darboux formula for associated Legendre polynomials We can do a direct and inverse Legendre transform for a single Fourier mode as:

$$\tilde{\zeta}^m(\tilde{\theta}_j) = \epsilon^m_{N+1} \overline{P}^m_{N+1}(\mu_j) \sum_{i=1}^{J} \frac{\zeta^m(\theta_i) w_i \overline{P}^m_N(\mu_i)}{\tilde{\mu}_j - \mu_i}$$

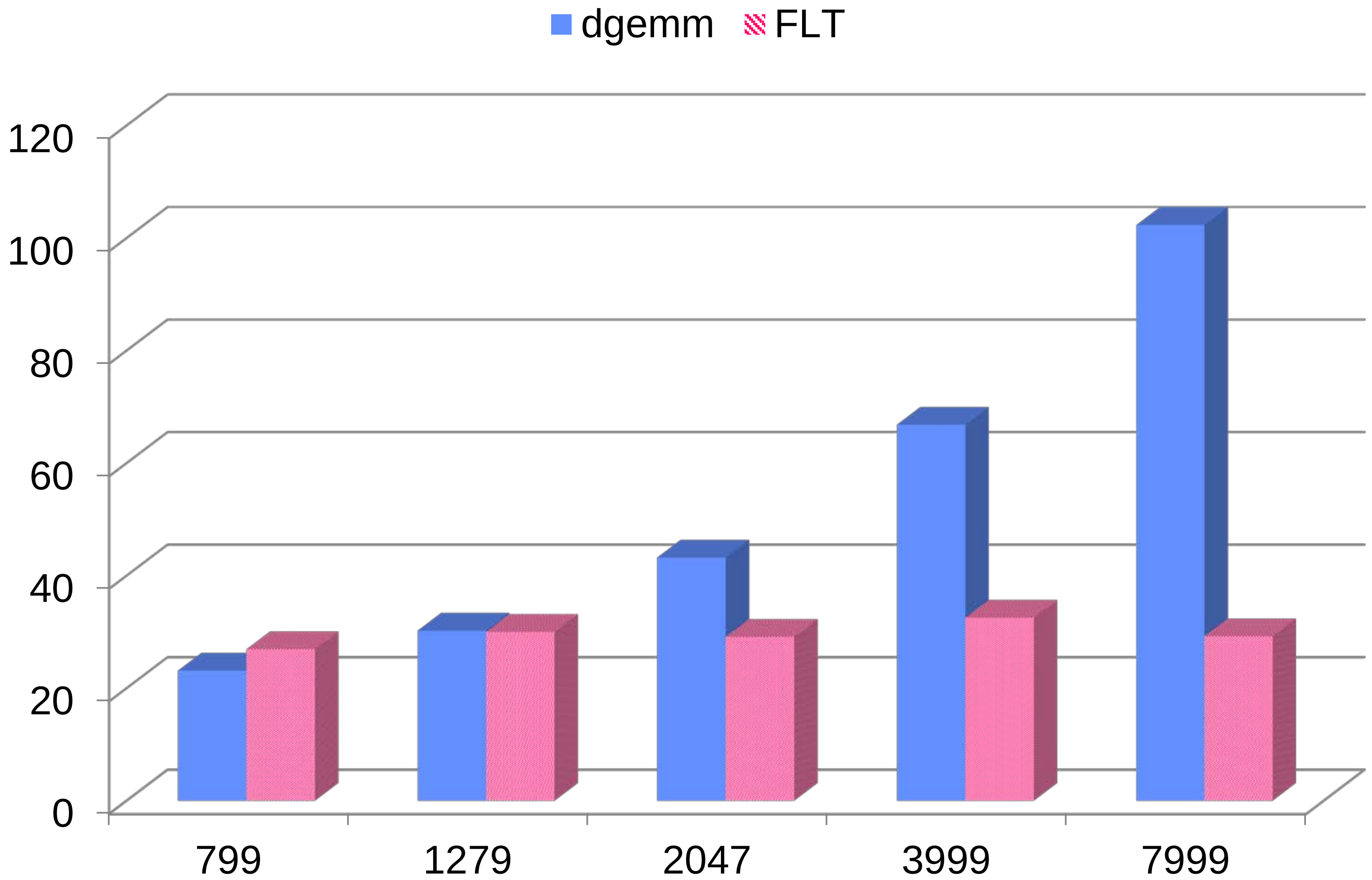
$$- \epsilon^m_{N+1} \overline{P}^m_N(\tilde{\mu}_j) \sum_{i=1}^{J} \frac{\zeta^m(\theta_i) w_i \overline{P}^m_{N+1}(\mu_i)}{\tilde{\mu}_j - \mu_i}$$

A fast Legendre transform (FLT)

(O'Neil, Woolfe, Rokhlin, 2009; Tygert 2008, 2010)

- **The computational complexity of the ordinary spectral transform is O(N^3) (where N is the truncation number of the series expansion in spherical harmonics) and it was therefore believed to be *not computationally competitive with other methods at very high resolution***
- **The FLT is found to be O(N^2 log N^3) for horizontal resolutions up to T7999** *(Wedi et al, 2013)*

Number of floating point operations for direct or inverse spectral transforms of a single field, scaled by $N^2 log^3 N$



ECMWF

Matrix-matrix multiply for each zonal wavenumber m

$$\begin{pmatrix} \zeta_{m,1}(x_1) & \cdots & \zeta_{m,l_{tot}}(x_1) \\ \vdots & \ddots & \vdots \\ \zeta_{m,1}(x_K) & \cdots & \zeta_{m,l_{tot}}(x_K) \end{pmatrix} = \begin{pmatrix} \overline{P}_1^m(x_1) & \cdots & \overline{P}_N^m(x_1) \\ \vdots & \ddots & \vdots \\ \overline{P}_1^m(x_K) & \cdots & \overline{P}_N^m(x_K) \end{pmatrix} \begin{pmatrix} \zeta_{1,1}^m & \cdots & \zeta_{1,l_{tot}}^m \\ \vdots & \ddots & \vdots \\ \zeta_{N,1}^m & \cdots & \zeta_{N,l_{tot}}^m \end{pmatrix}$$

$$\left[\zeta^m(x_k)_l \right] = \left[\overline{P_n^m}(x_k) \right] \left[\zeta_{n,l}^m \right]$$

Gaussian latitude

Field, vertical level

apply butterfly compression, this step is *precomputed* only once!

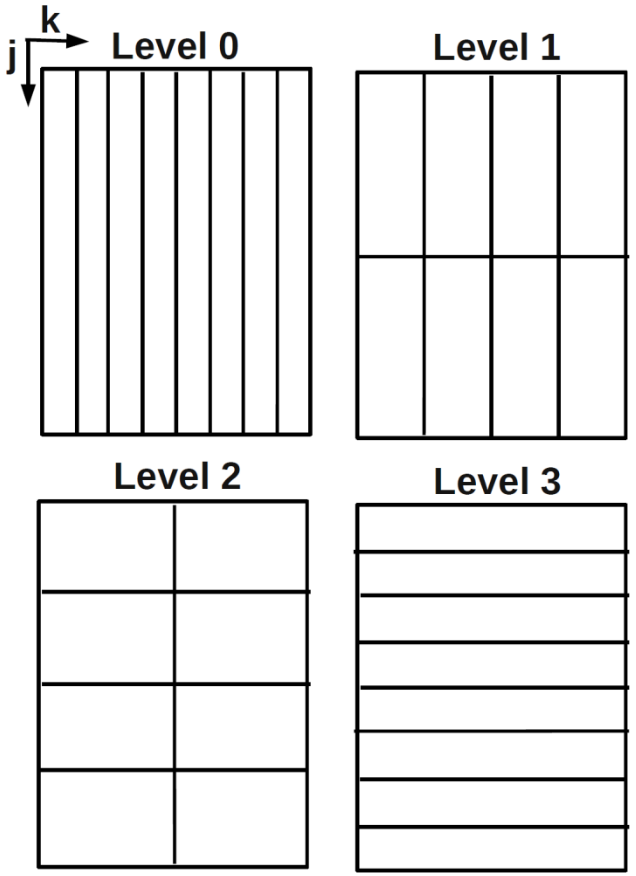
total wavenumber

zonal wavenumber

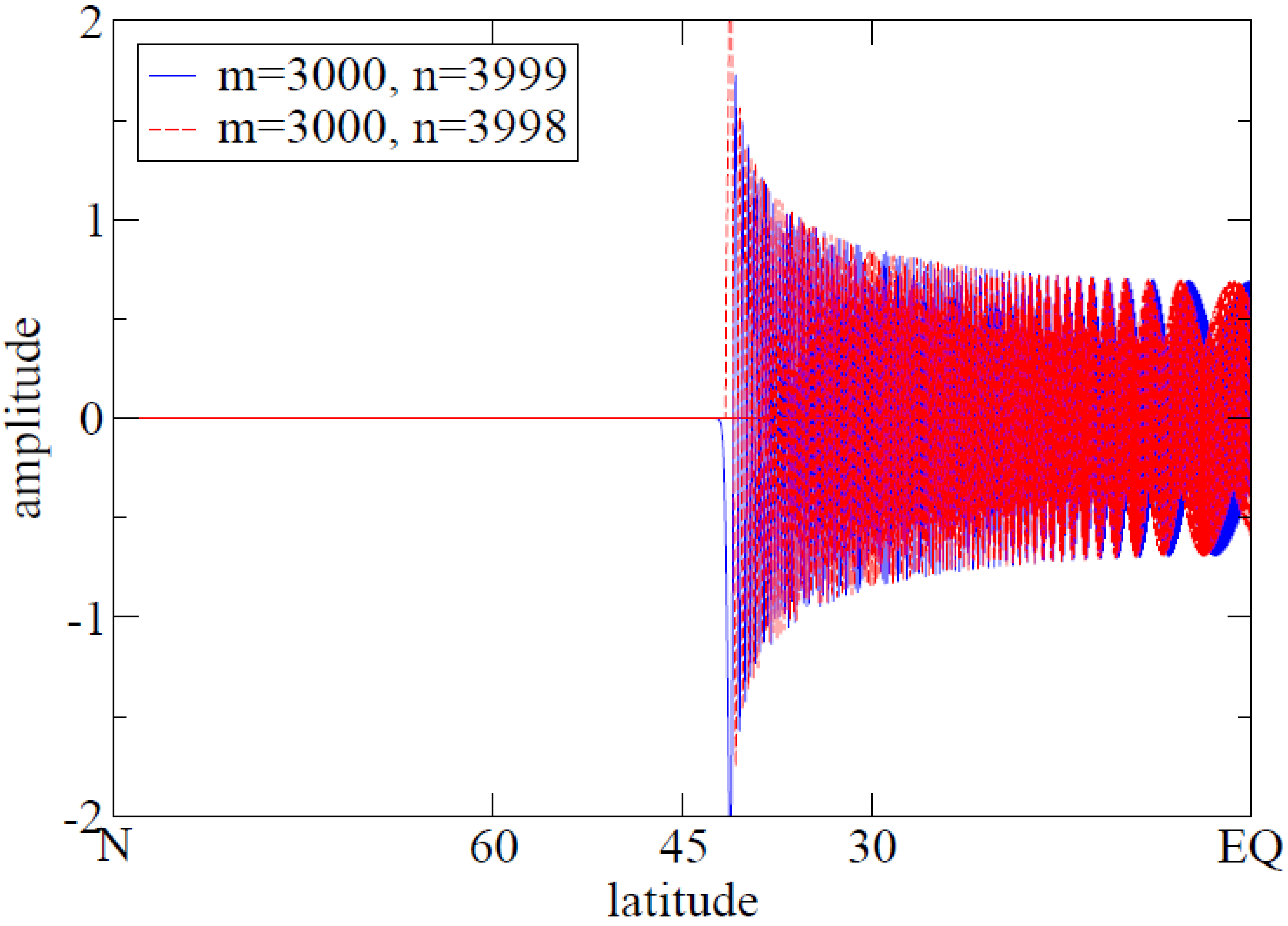
Butterfly algorithm: pre-compute $S_{rxs} \cong C_{rxk} A_{kxs}$

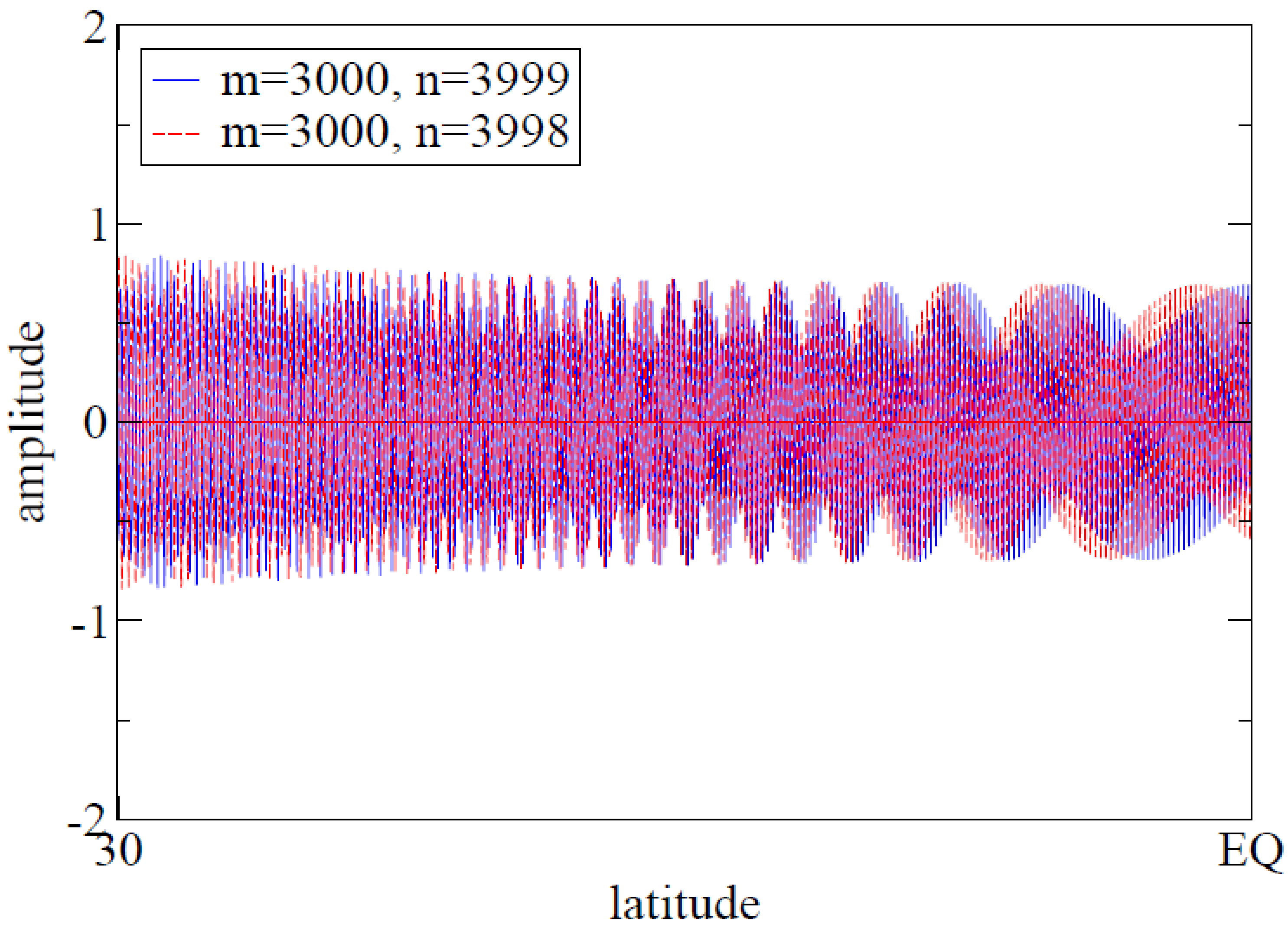
```
for l = 0 → L do
    for all j, k boxes do
        if l = 0 then
            S_{0,k} = extract_sub_matrix()
            compr_sub_matrix(S_{0,k}, A_{0,k})
            store A_{0,k}
        else
            S_{l,j,k} = comb_compr_l_and_r_neighb(C, l − 1)
            compr_sub_matrix(S_{l,j,k}, A_{l,j,k})
            store A_{l,j,k}
        end if
        if l = L then
            store C_{L,j}
        end if
    end for
end for
```

With each level l, double the columns and half the rows



ECMWF


m=3000, n=3999
m=3000, n=3998
amplitude
2
1
0
-1
-2
N
60
45
30
EQ
latitude



m=3000, n=3999
m=3000, n=3998
2
1
0
-1
-2
amplitude
30
EQ
latitude

Butterfly algorithm: apply $f = S\alpha$

for $l = 0 \rightarrow L$ do

  for all j, k boxes do

    if $l = 0$ then

      store $\beta_{0,k} = A_{0,k}\alpha_k$

    else

      store $\beta_{l,j,k}$

        $= A_{l,j,k} \times comb\_l\_and\_r\_neighb(\beta, l-1)$

    end if

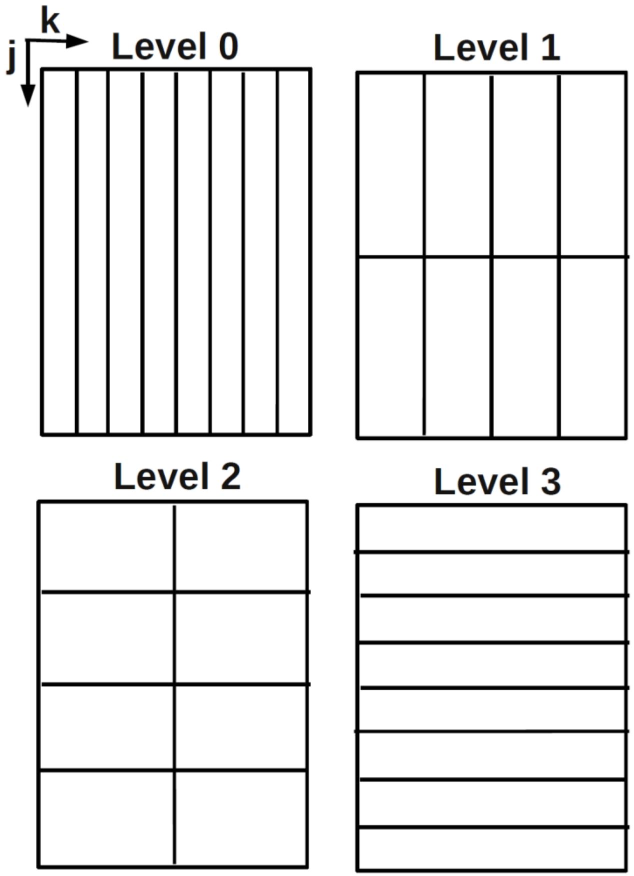
    if $l = L$ then

      store $f_{L,j} = C_{L,j}\beta_{L,j}$

    end if

  end for

end for



ECMWF

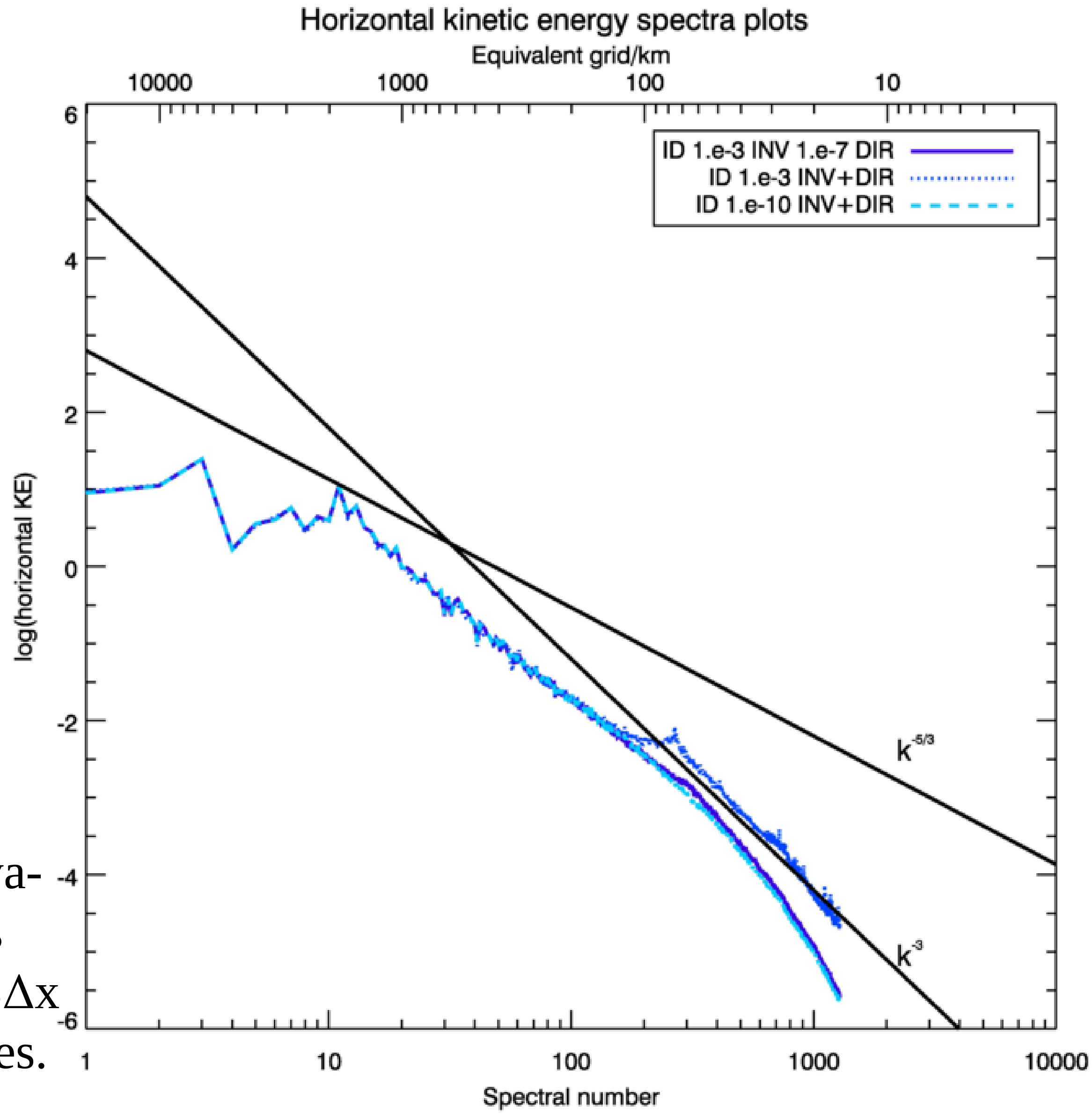
Interpolative Decomposition (ID)

- **The** *compression* **uses the interpolative decomposition (ID) described in** ***Cheng et al (2005).***
- **The** *r x s* **matrix** *S* **may be** *compressed* **such that**

$$\left\| S_{rxs} - C_{rxk} A_{kxs} \right\| \leq \varepsilon$$

With an *r x k* **matrix** *C* **constituting a subset of the columns of** *S* **and the** *k x s* **matrix** *A* **containing a** *k x k* **identity as a submatrix.** *k* **is the** *ε***-rank of the matrix** *S* *(see also e.g. Martinsson and Rokhlin, 2007).*

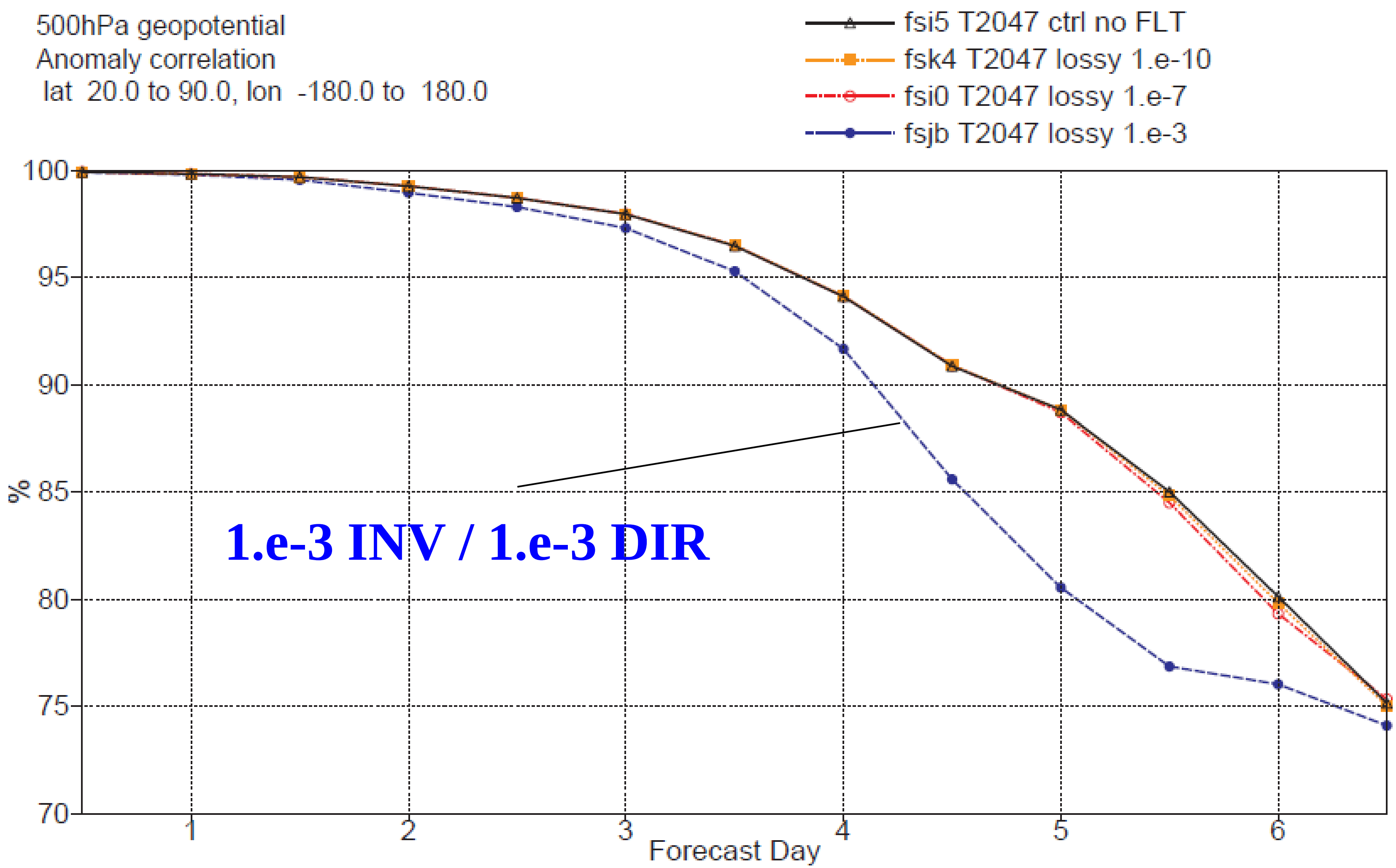
T1279 FLT using different ID epsilons for INV (1.e-3) + DIR (1.e-7)



A recent comparison with observational surface wind data suggests that IFS fully resolves scales at 8Δx and partially resolves 4Δx features.

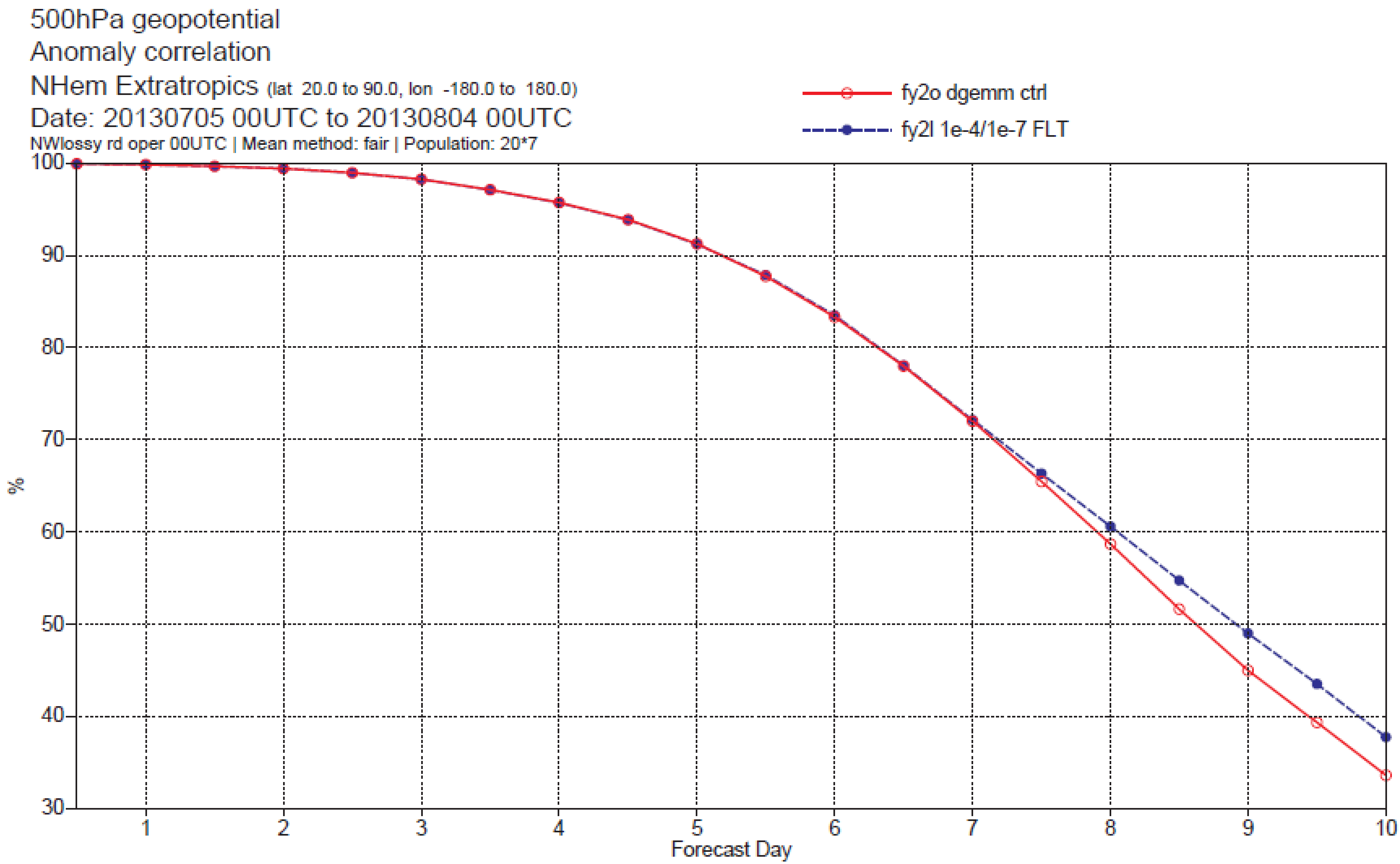
(Abdalla et al, 2013)

ECMWF


500hPa geopotential
Anomaly correlation
lat 20.0 to 90.0, lon -180.0 to 180.0
fsi5 T2047 ctrl no FLT
fsk4 T2047 lossy 1.e-10
fsi0 T2047 lossy 1.e-7
fsjb T2047 lossy 1.e-3
100
95
90
85
80
75
70
%
1
2
3
4
5
6
Forecast Day
1.e-3 INV / 1.e-3 DIR

ECMWF

Comparison T1279 1.e-4 INV / 1.e-7 DIR

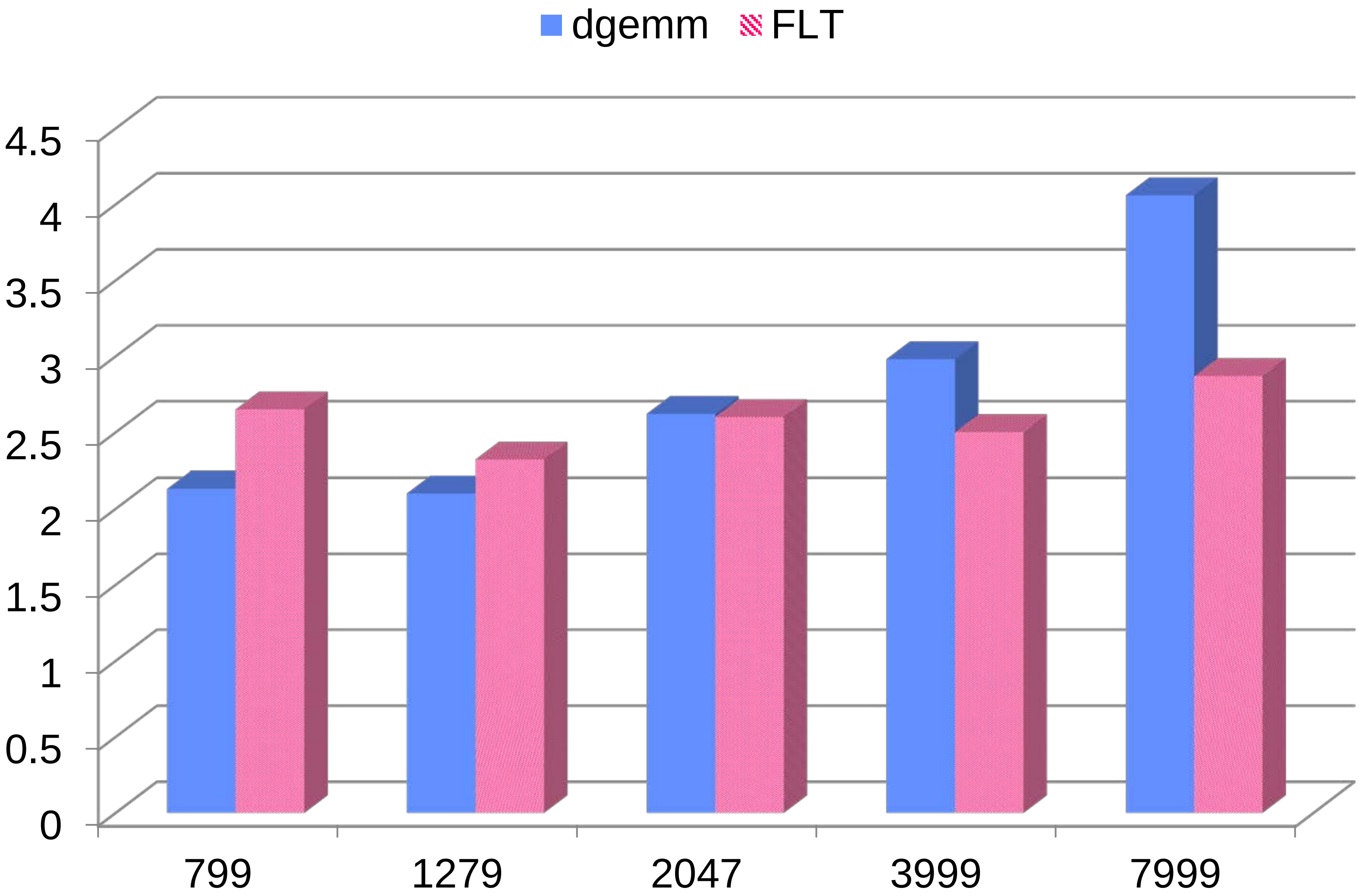


ECMWF

The FLT in a nutshell – $O(N^2 log^3 N)$

- **Speed-up the sums of products between associated Legendre polynomials at all Gaussian latitudes and the corresponding spectral coefficients of a field (e.g. temperature on given level)**
- **The essence of the FLT:**
  - **Exploit similarities of associated Legendre polynomials at all (Gaussian) latitudes but different total wave-number**
  - **Pre-compute (once, 0.1% of the total cost of a 10 day forecast) a compressed (approximate) representation of the matrices (for each m) involved**
  - **Apply the compressed (reduced) representation at every time-step of the simulation.**

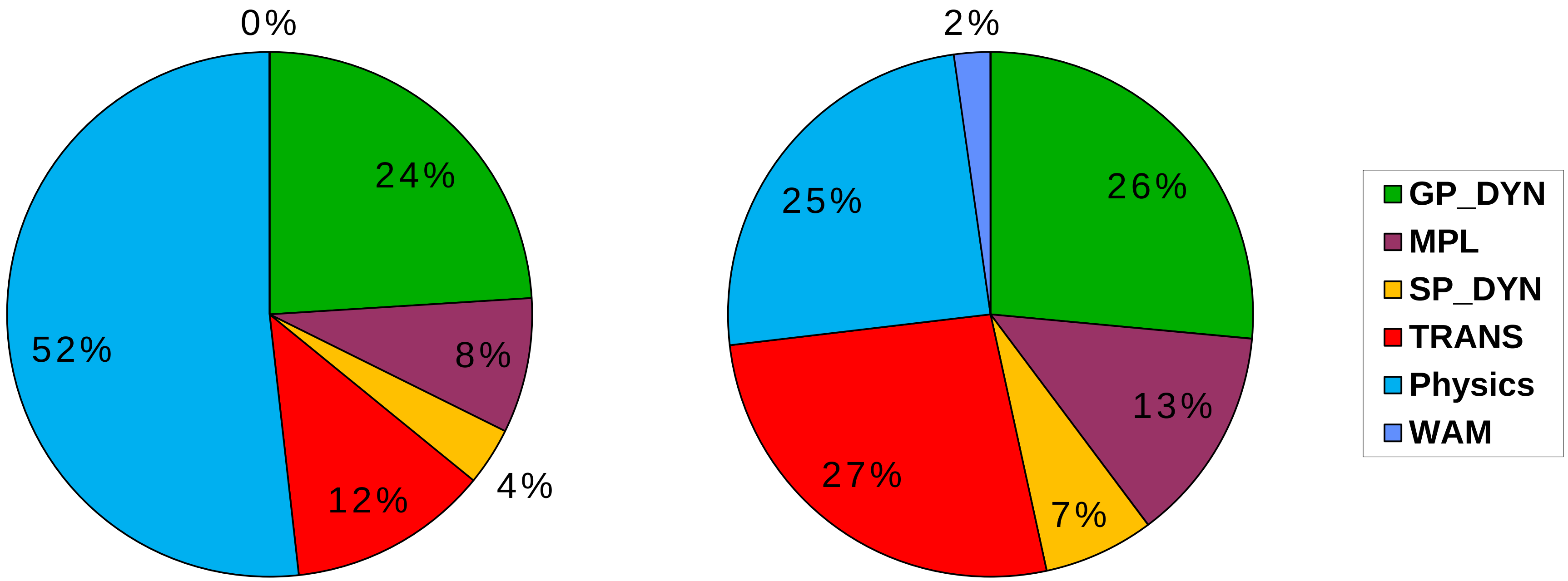
Average wall-clock time compute cost of 10^7 spectral transforms scaled by $N^2 log^3 N$



ECMWF

Computational Cost:
NH at T3999 vs. H T2047

SP_DYN was 23 percent for this model configuration, and is now 7 percent. Improvement due to exposing 'greater OpenMP parallelism' from 4K threads to a maximum of 4K * 91 threads ; in this case 16K threads.



H T_L2047 L91

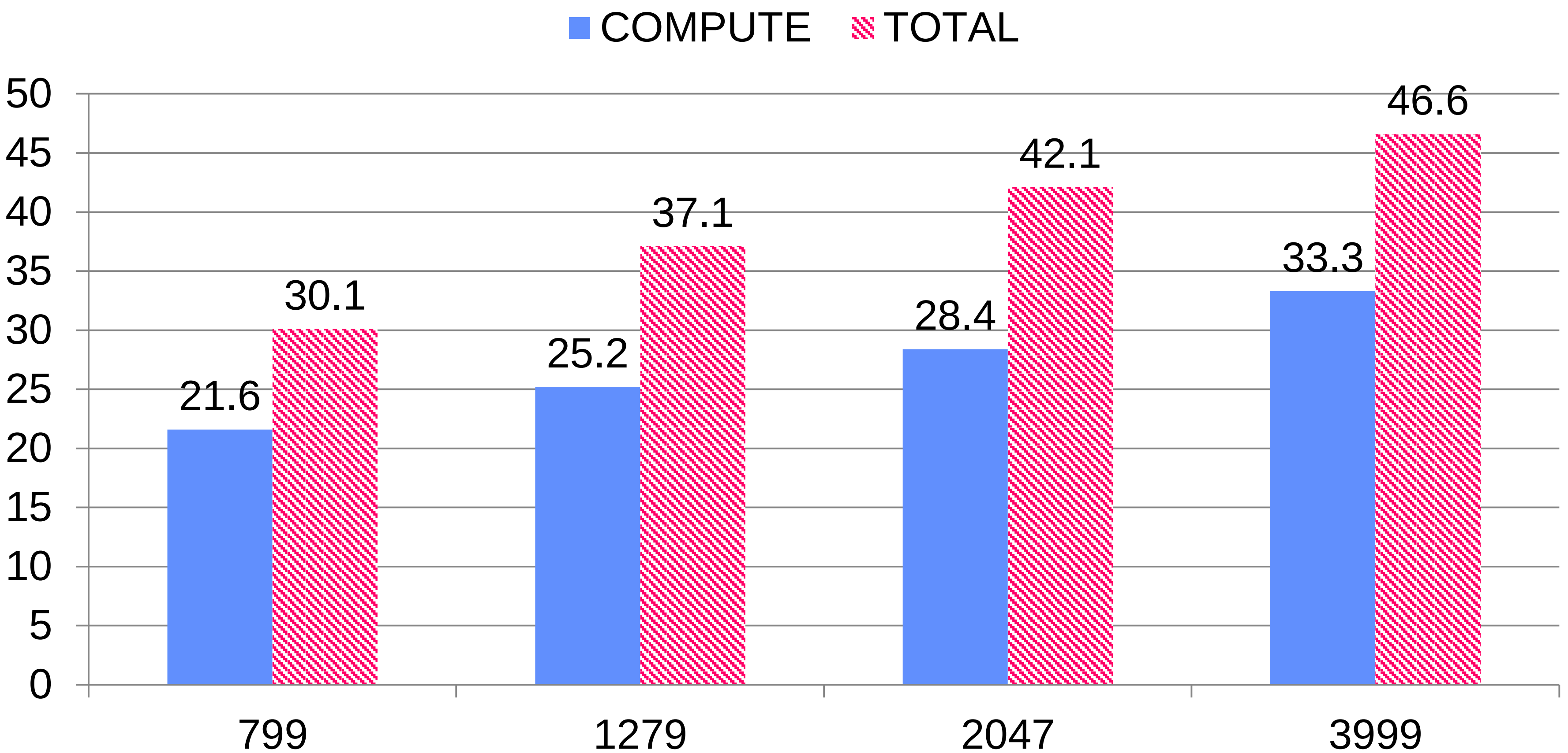
Tstep=450s, 0.8s/iteration
With 896x16 ibm_power7

NH T_L3999 L91

Tstep=180s, 3.1s/iteration
With 1024x16 ibm_power7

ECMWF

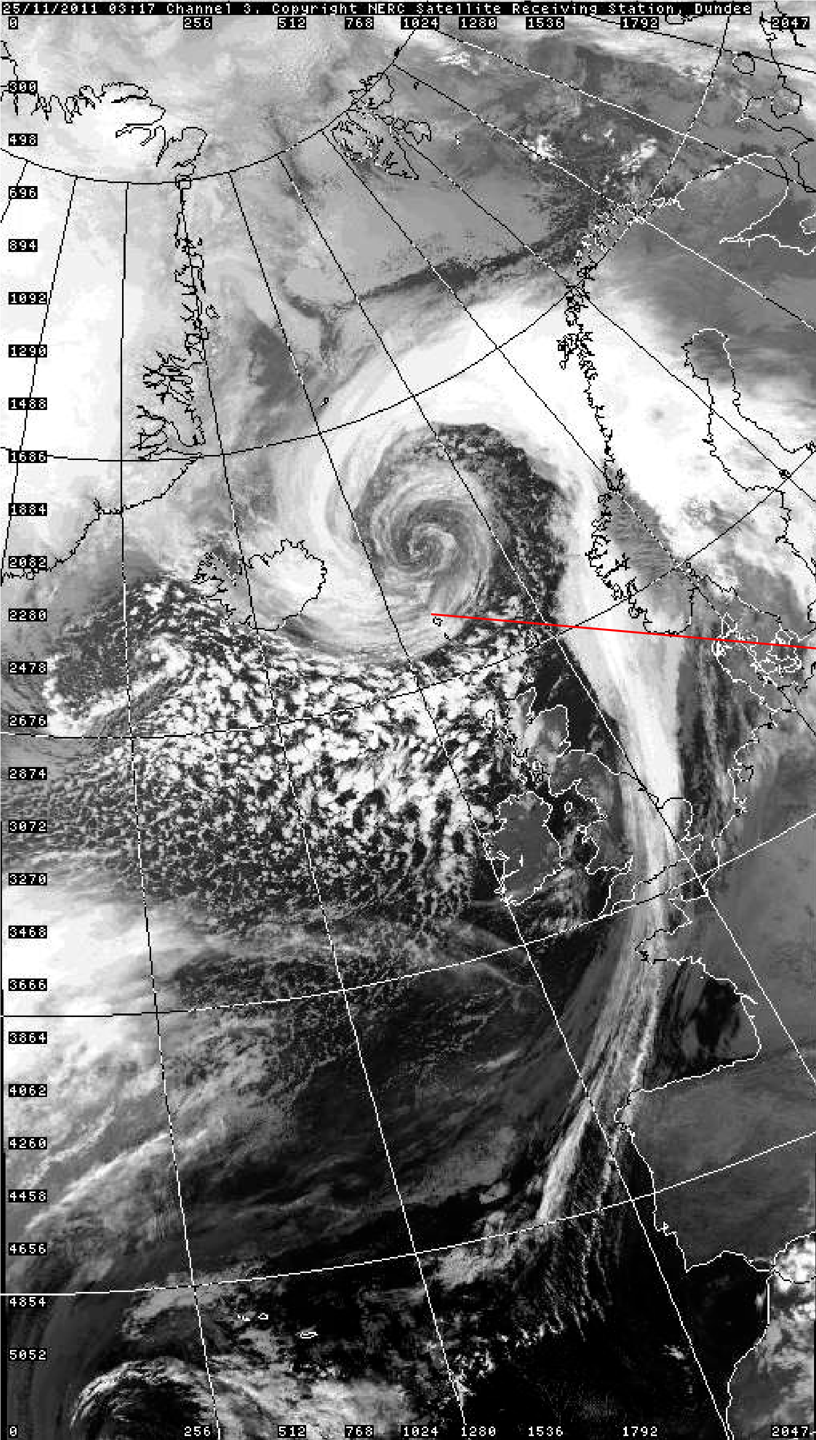
% (of total execution time) cost of spectral part of the model on IBM Power7 (all L91, all NH for comparison); Total includes communications



We expect significant reductions in future cores -> vector instr. / GPU

All these can be run with hydrostatic code == ½ of above numbers !

Norwegian storm Berit(Xaver) 22-27th November 2011



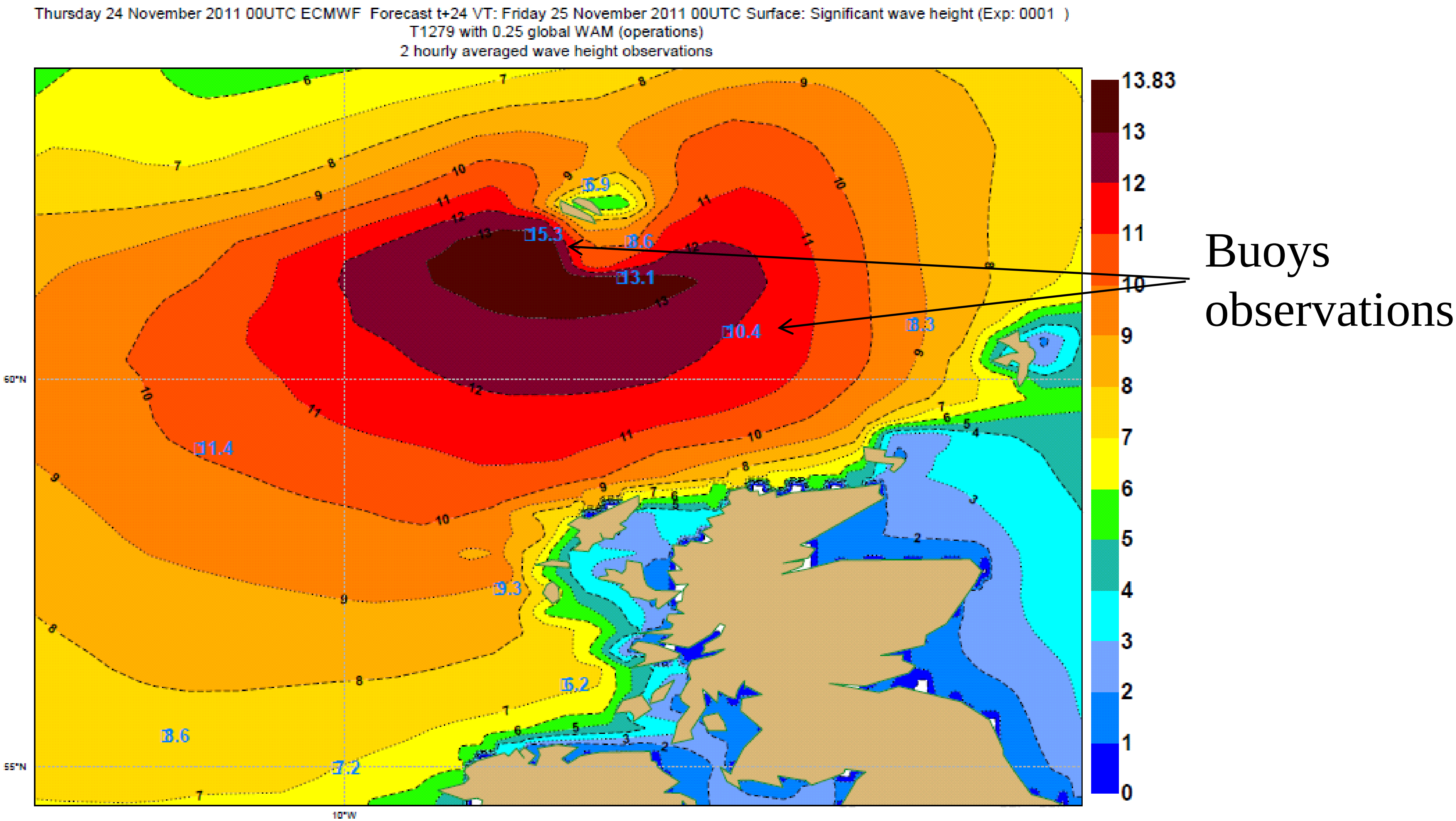
Faroe Islands

Monster waves battered the shores of the Faroe Islands, and Norway later, both reporting severe structural damage from the storm(s) during this period.

ECMWF

T1279 coupled to 0.25 degree wave-model

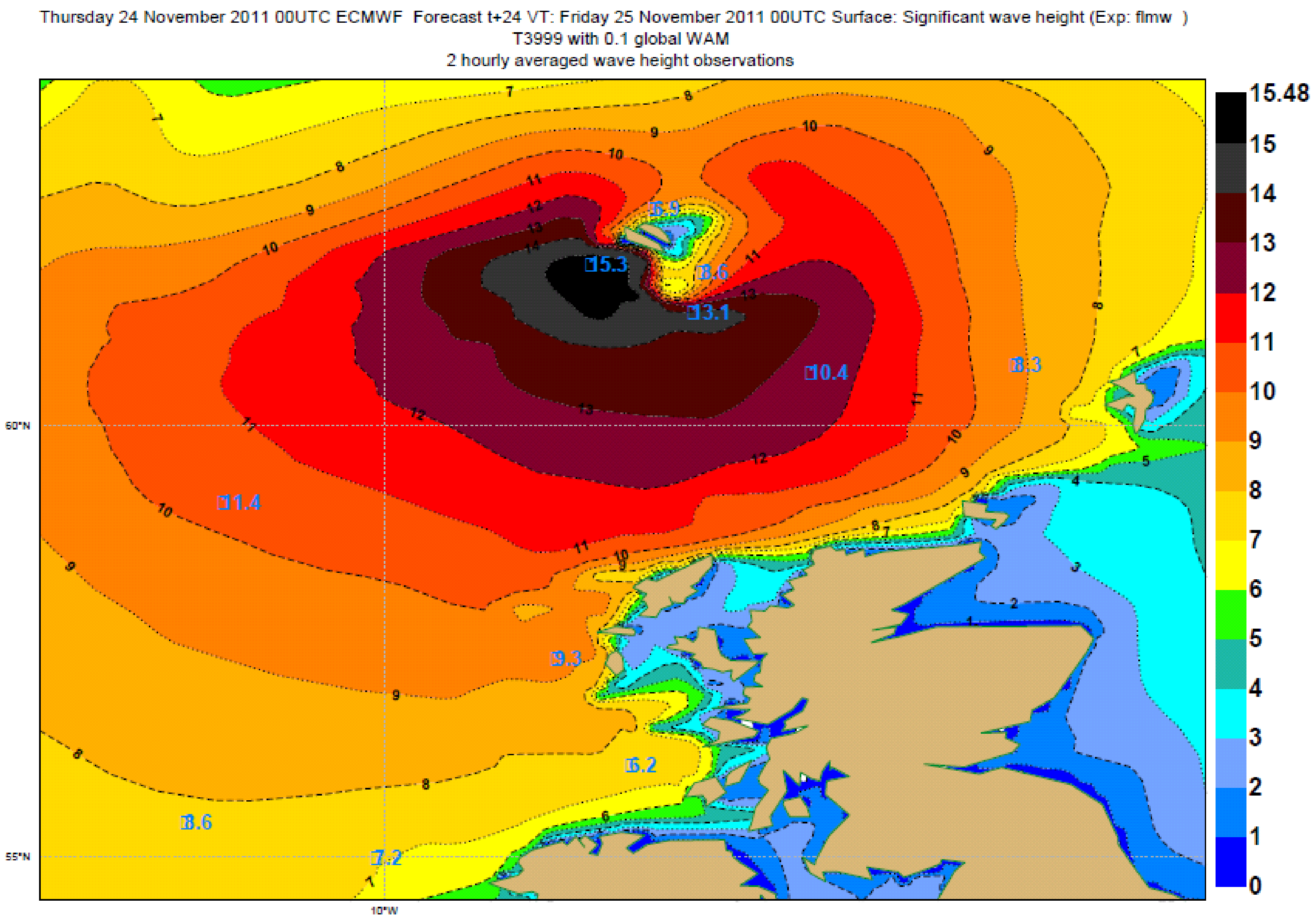
Faroe Islands hit by storm Xaver (Berit) on 25th November 2011



ECMWF

T3999 coupled to 0.1 degree wave model

Faroe Islands hit by storm Xaver (Berit) on 25th November 2011

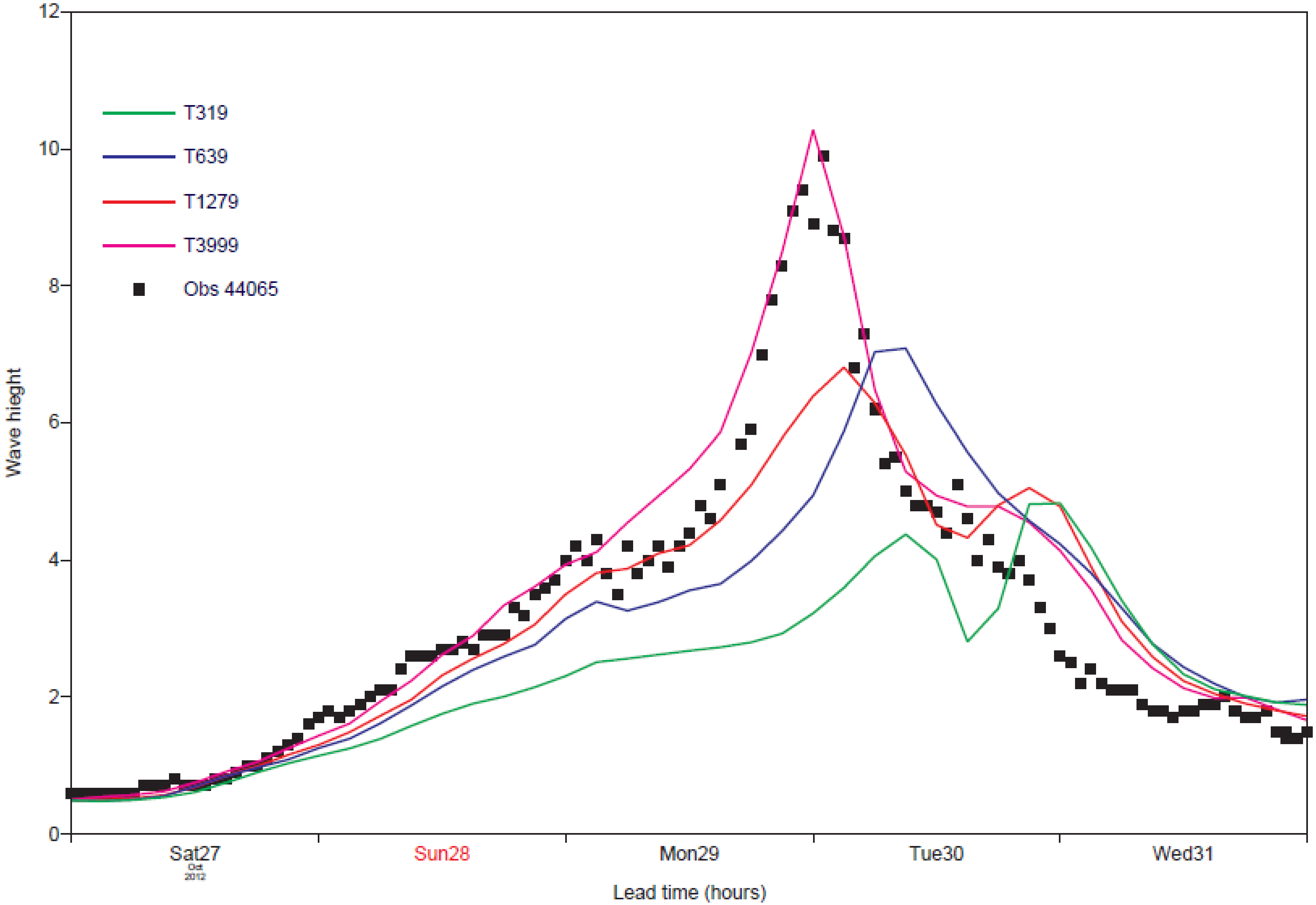


ECMWF

Extreme events: Hurricane Sandy

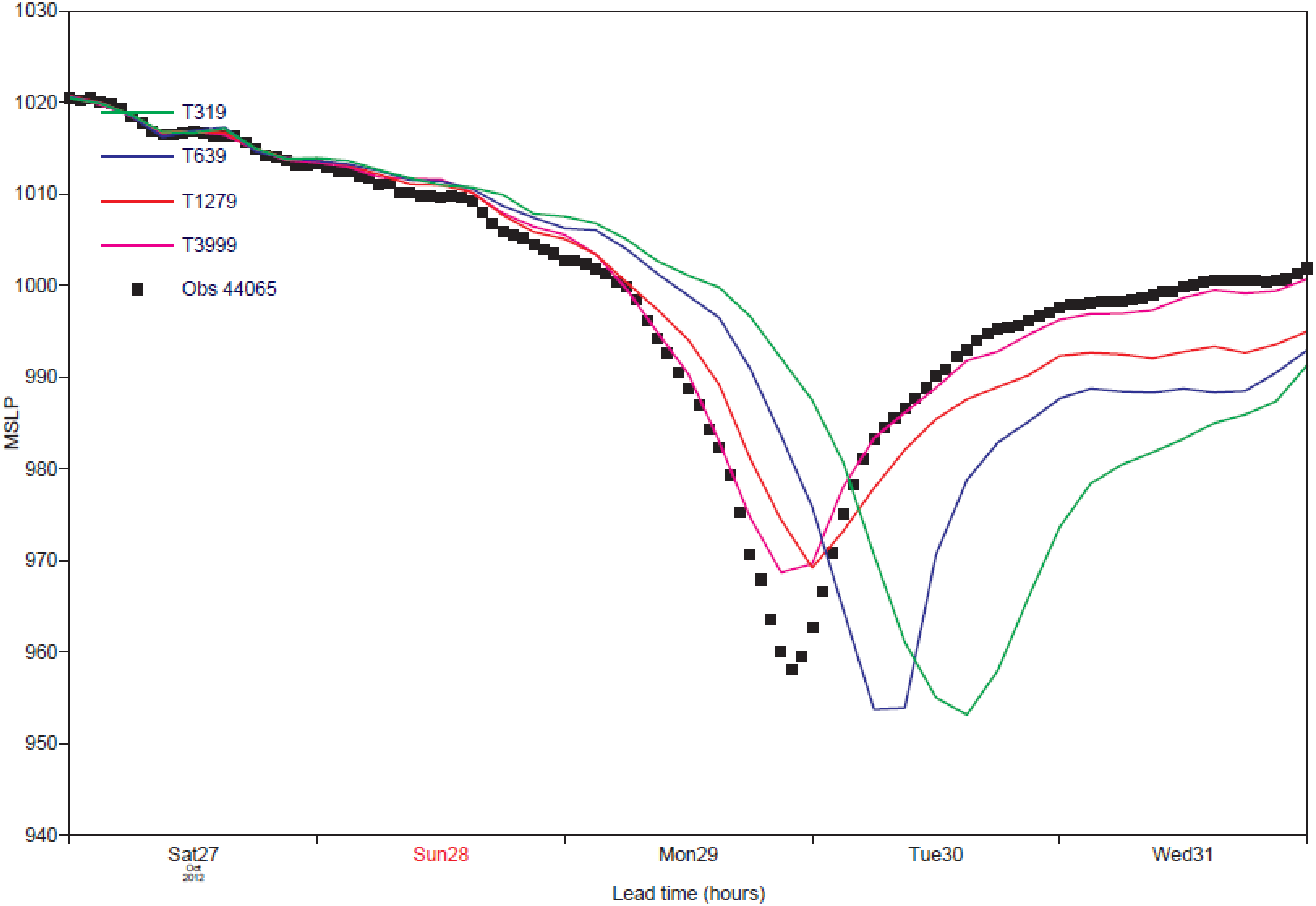
NASA Earth Observatory, 28th October 2012

Wave height 72h forecast

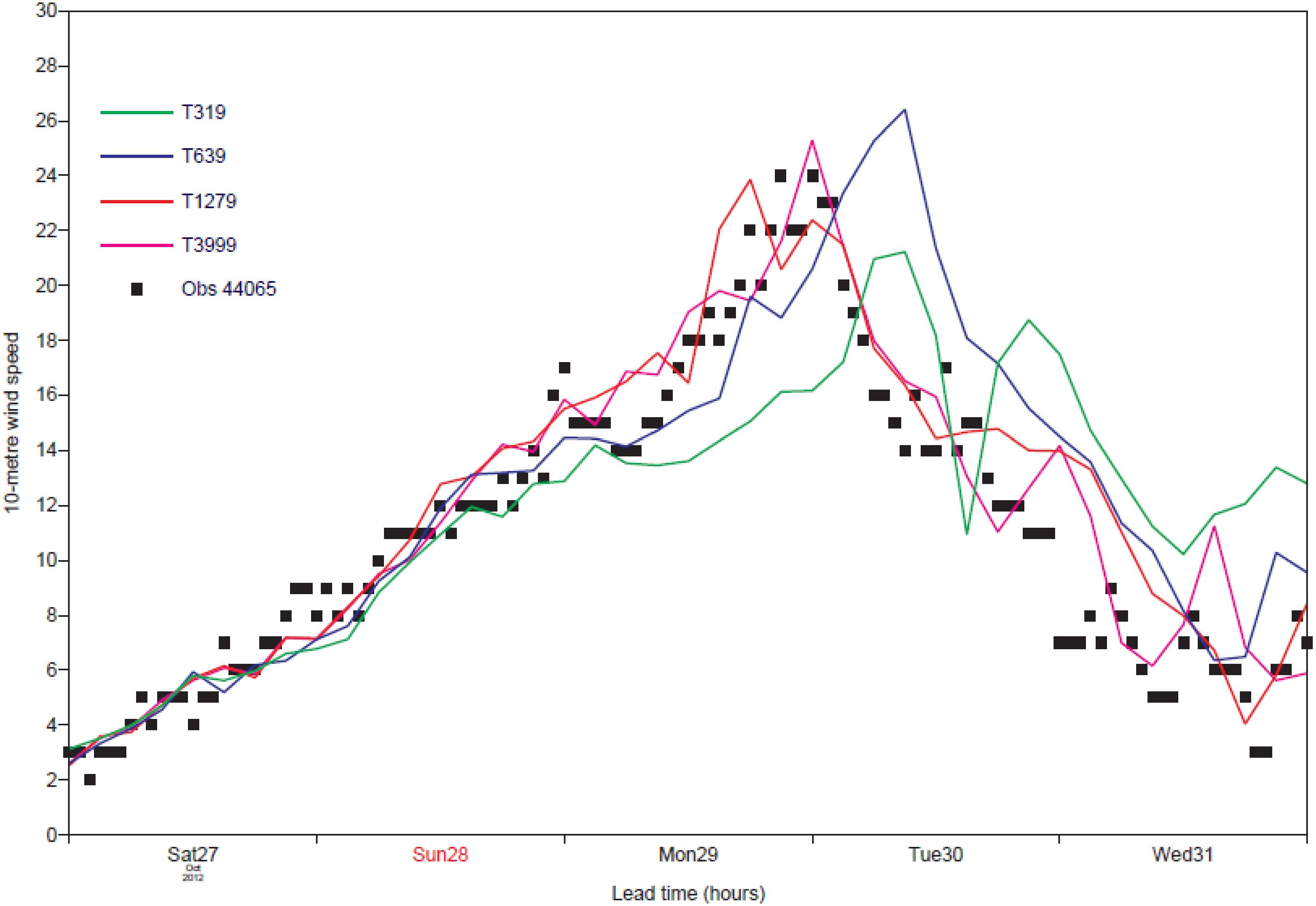


ECMWF

MSL pressure 72h forecast

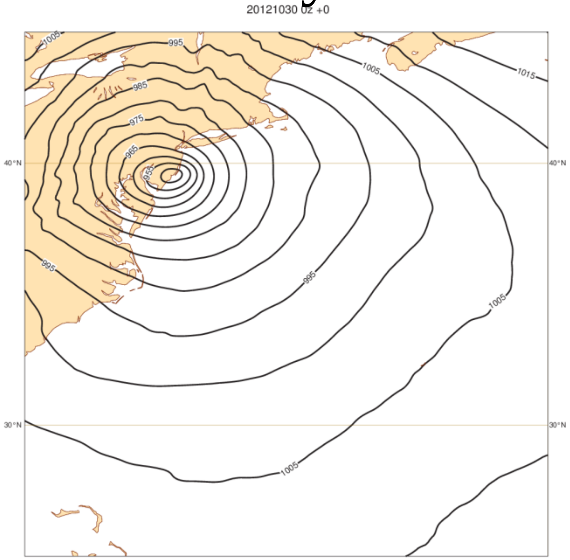


10m – wind speed 72h forecast

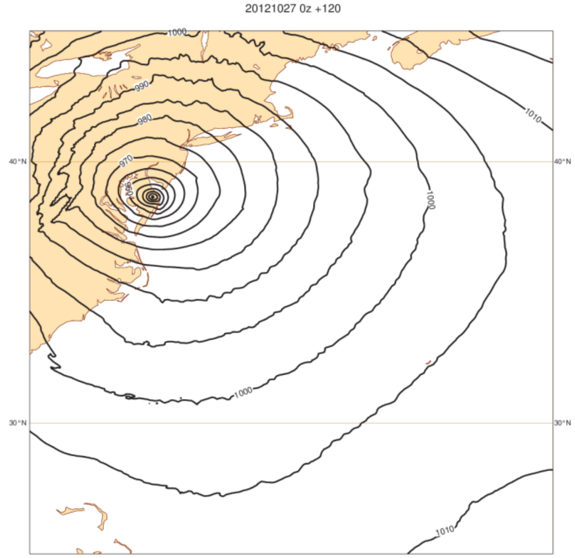


2012-10-25 00z +120

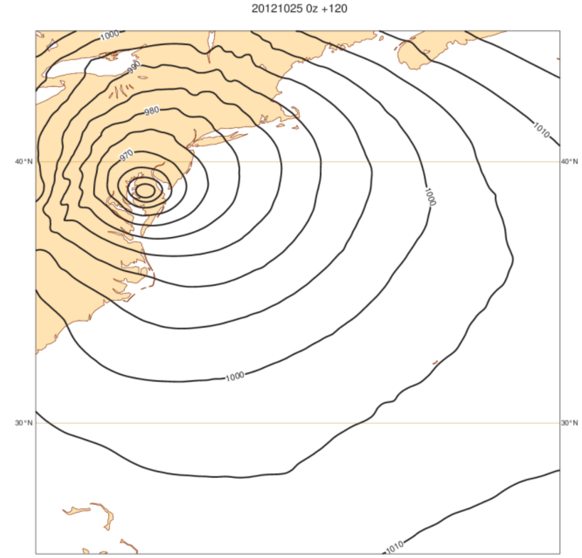
Analysis



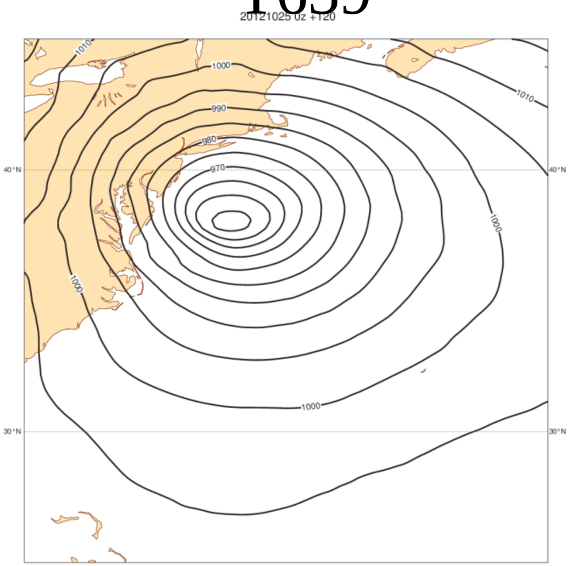
T3999



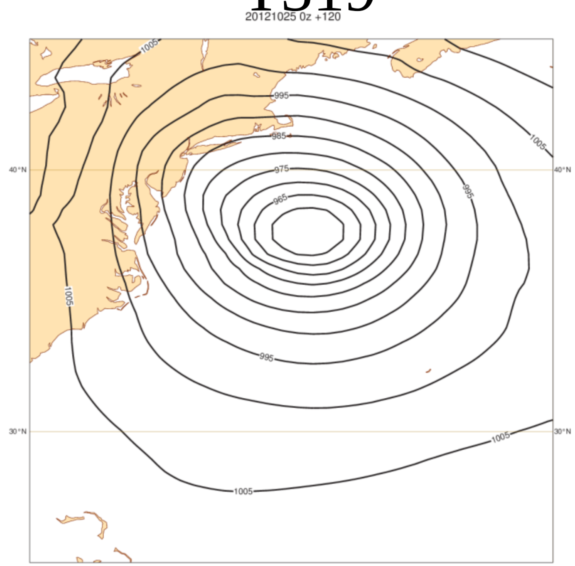
T1279



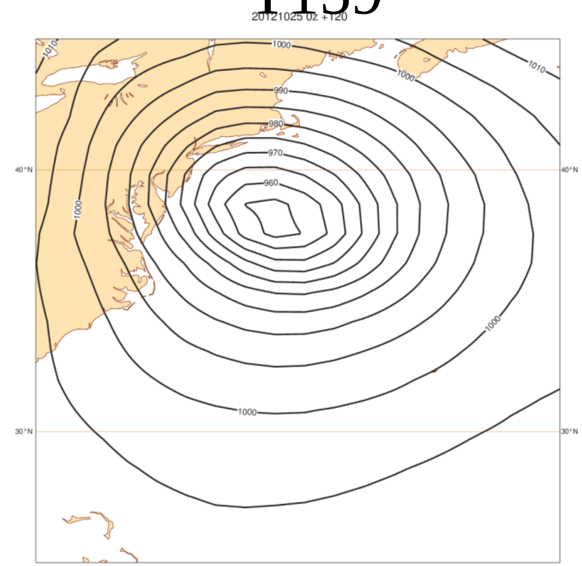
T639



T319



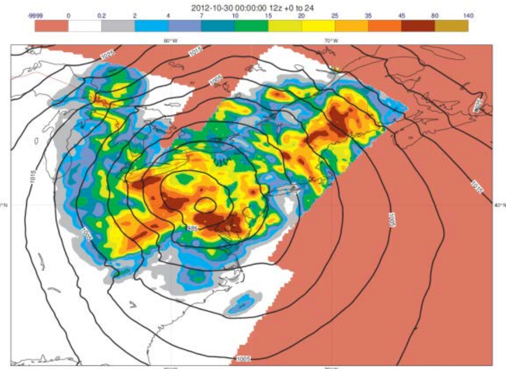
T159



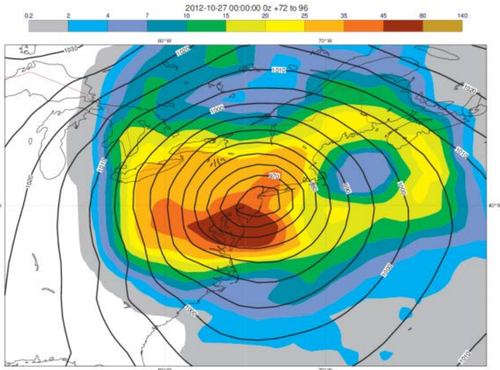
24-hour precipitation and MSLP

2012-10-27 00z, precip 72 to 96h, 84h for MSLP

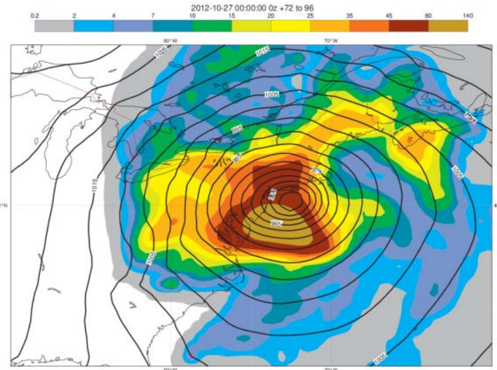
NEXRAD 24h precip
Oper analysis of MSLP



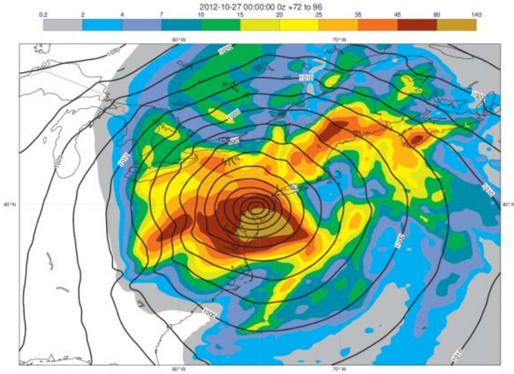
T159



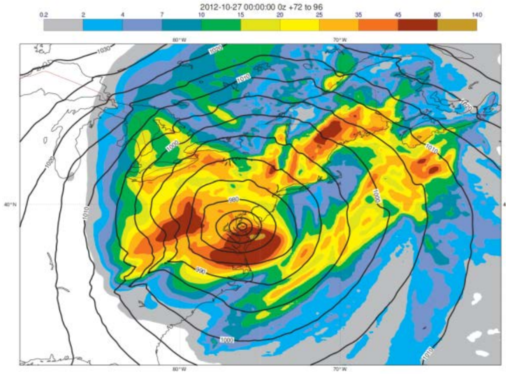
T319



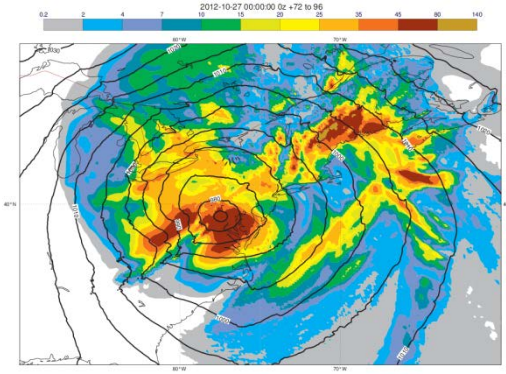
T639



T1279

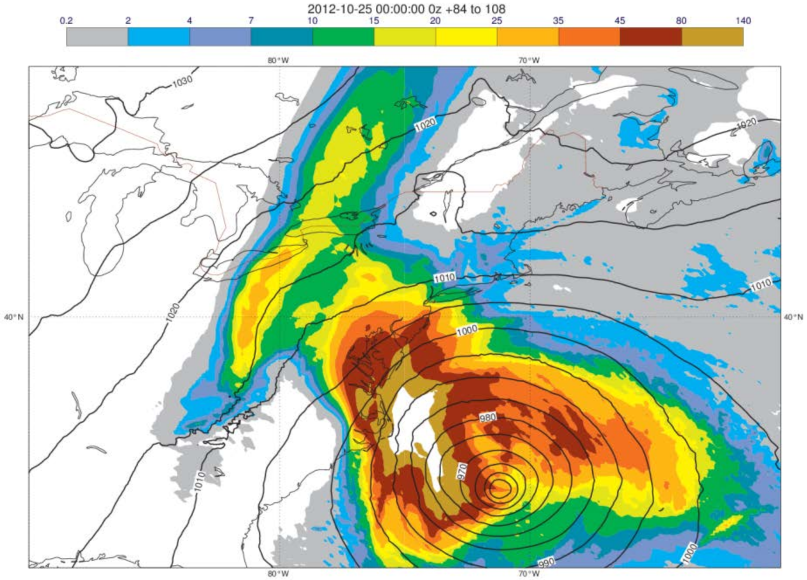


T3999

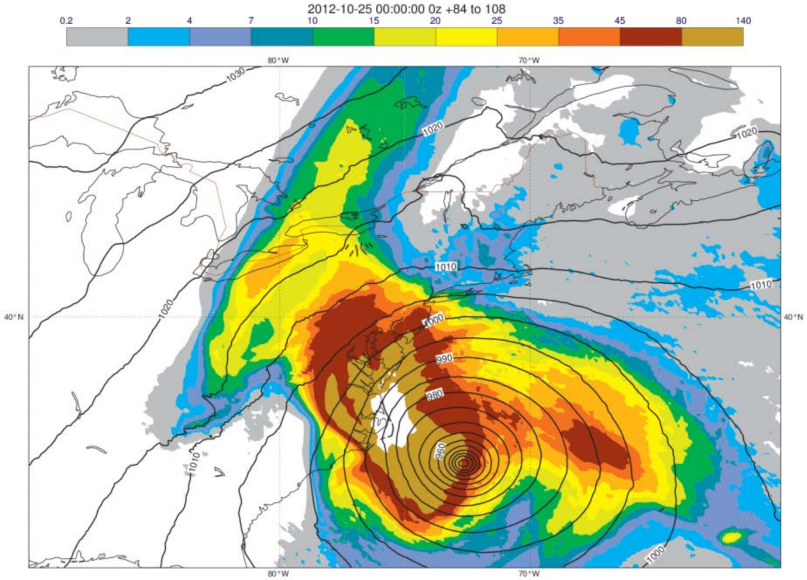


Hydrostatic vs Nonhydrostatic

Hydrostatic
Precip 84h to 108h



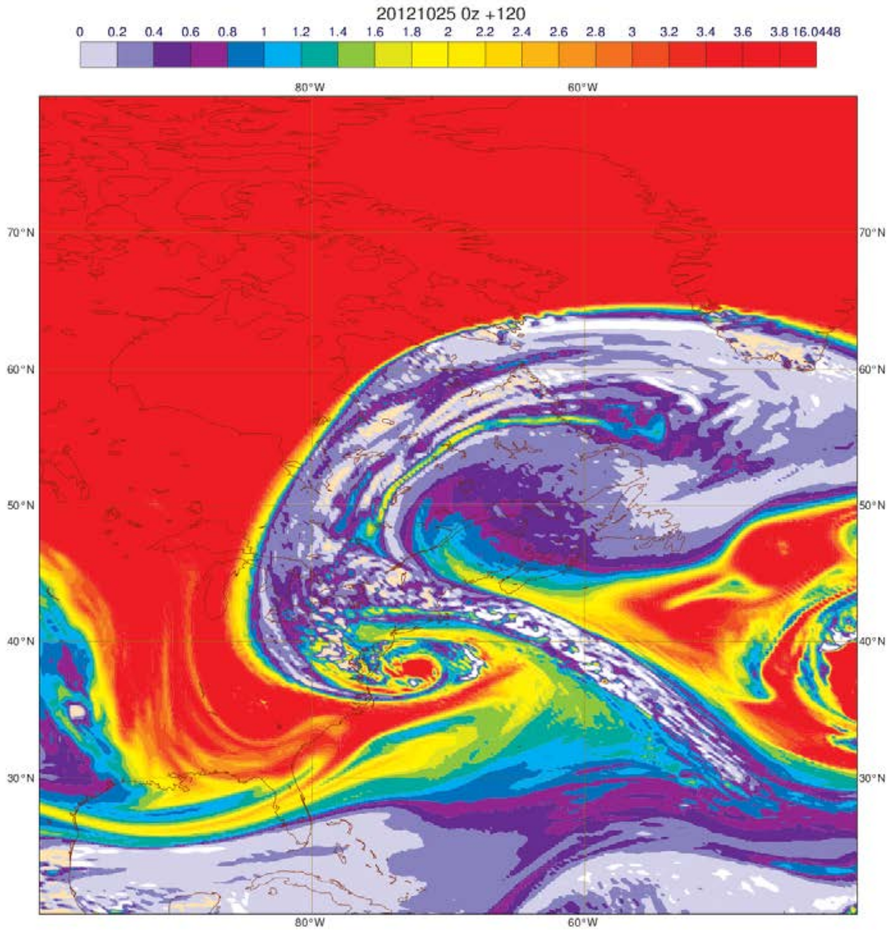
Non-Hydrostatic
Precip 84h to 108h



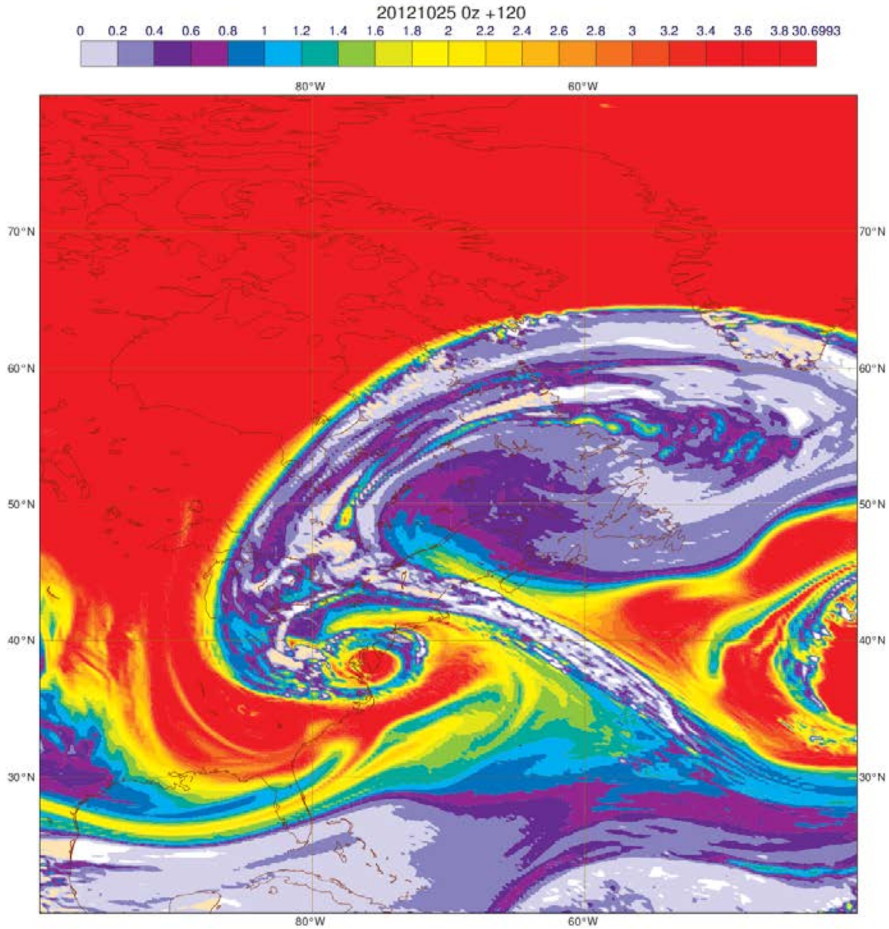
ECMWF

Hydrostatic vs Nonhydrostatic

Hydrostatic
PV330 +120

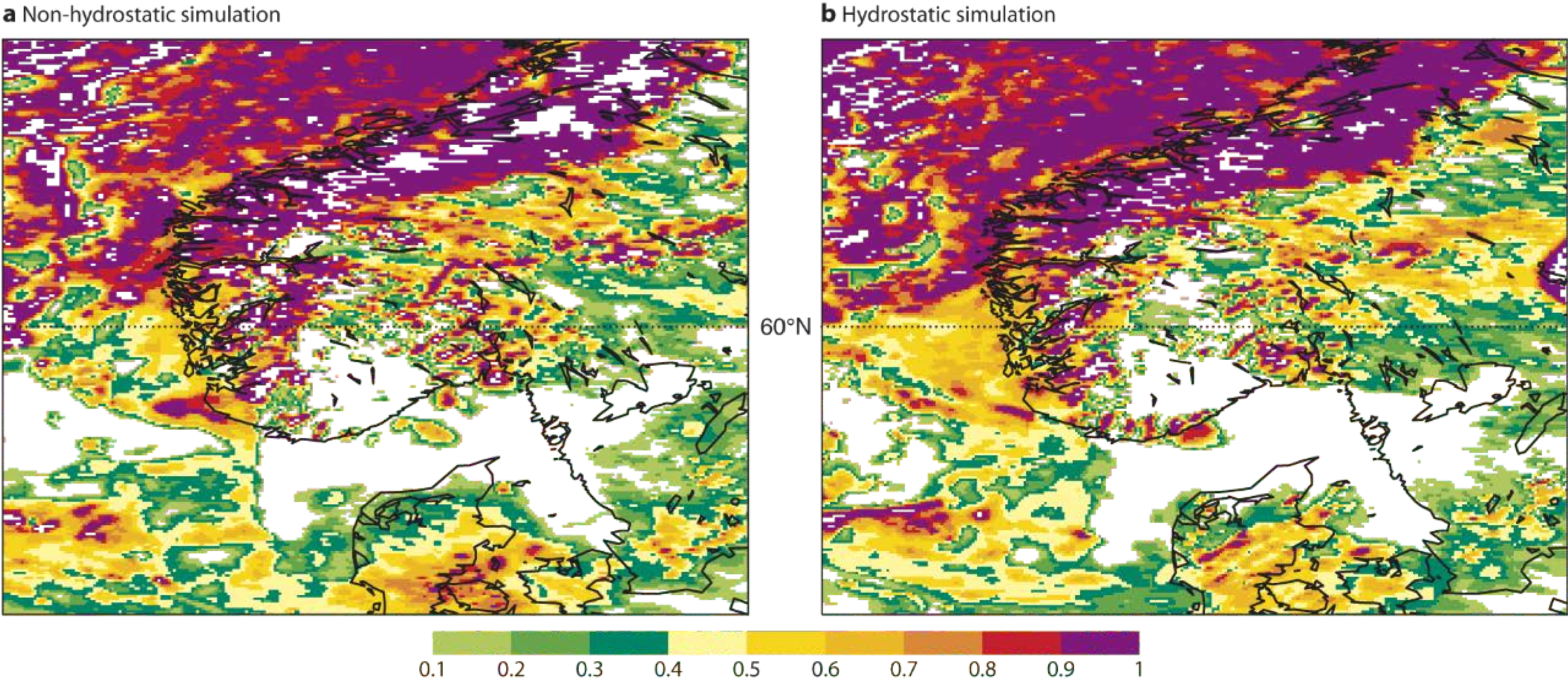


Non-Hydrostatic
PV330 +120



T3999 L91 (~5km)

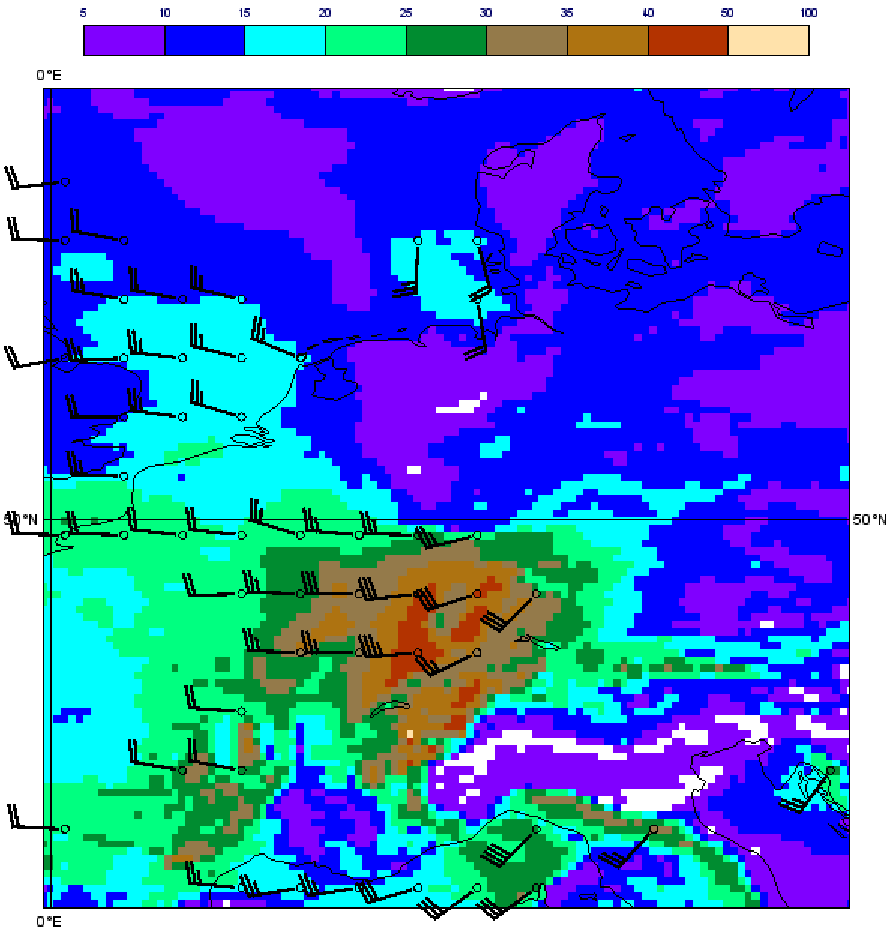
Hydrostatic vs Nonhydrostatic



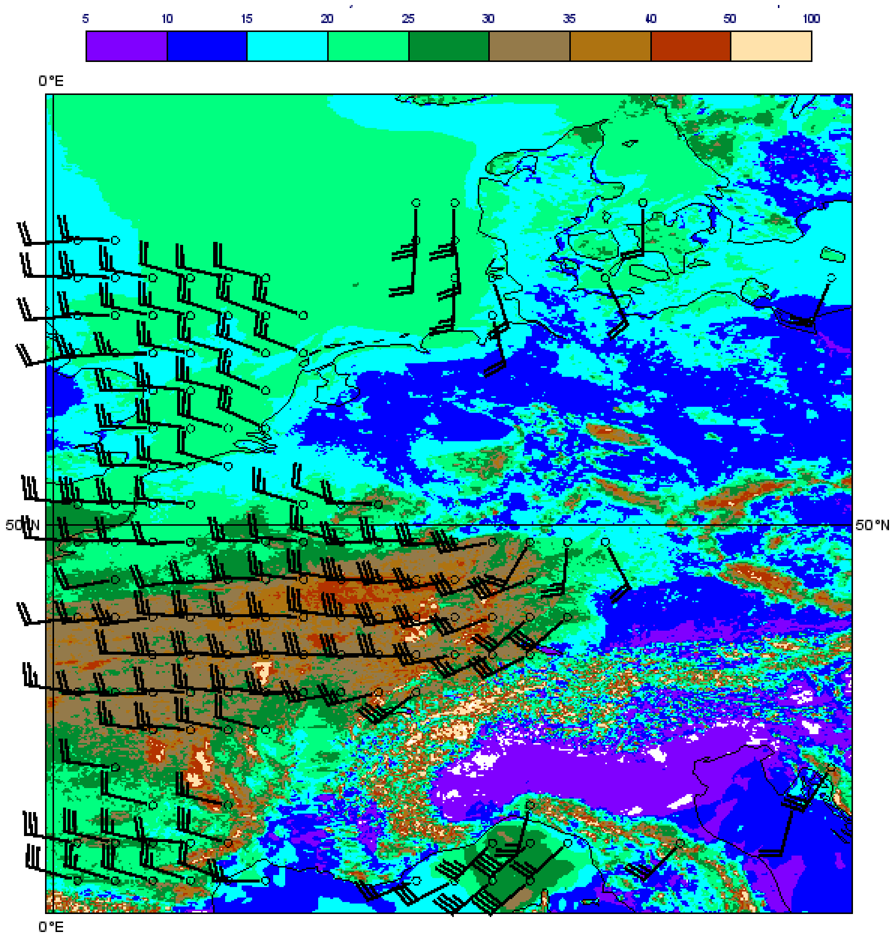
Cloud cover 24h forecast T3999 (~5km)

Era-Interim shows a wind shear with height in the troposphere over the region!

Lothar "Christmas storm"



T1279 (~16km)

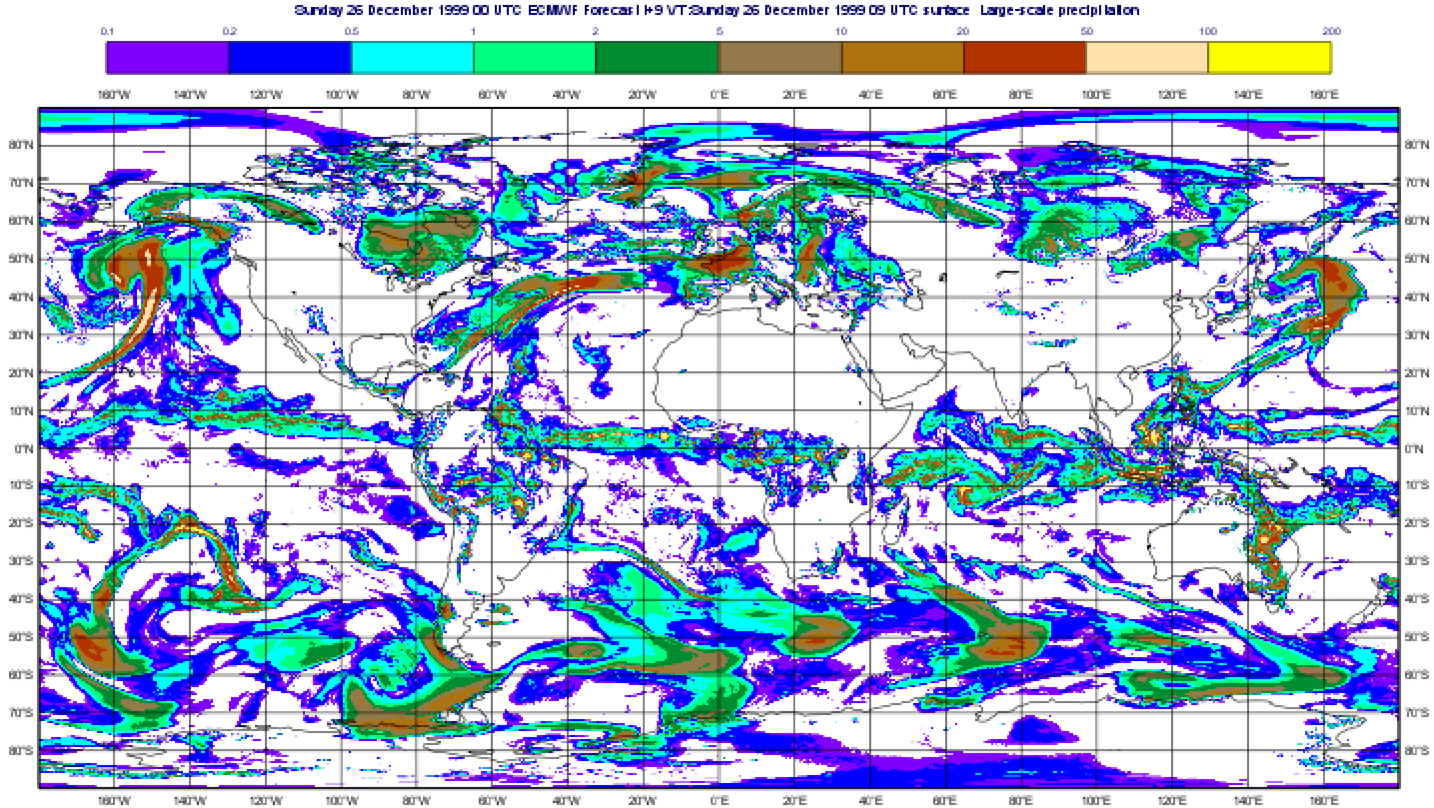


T7999 (~2.5km)

ECMWF

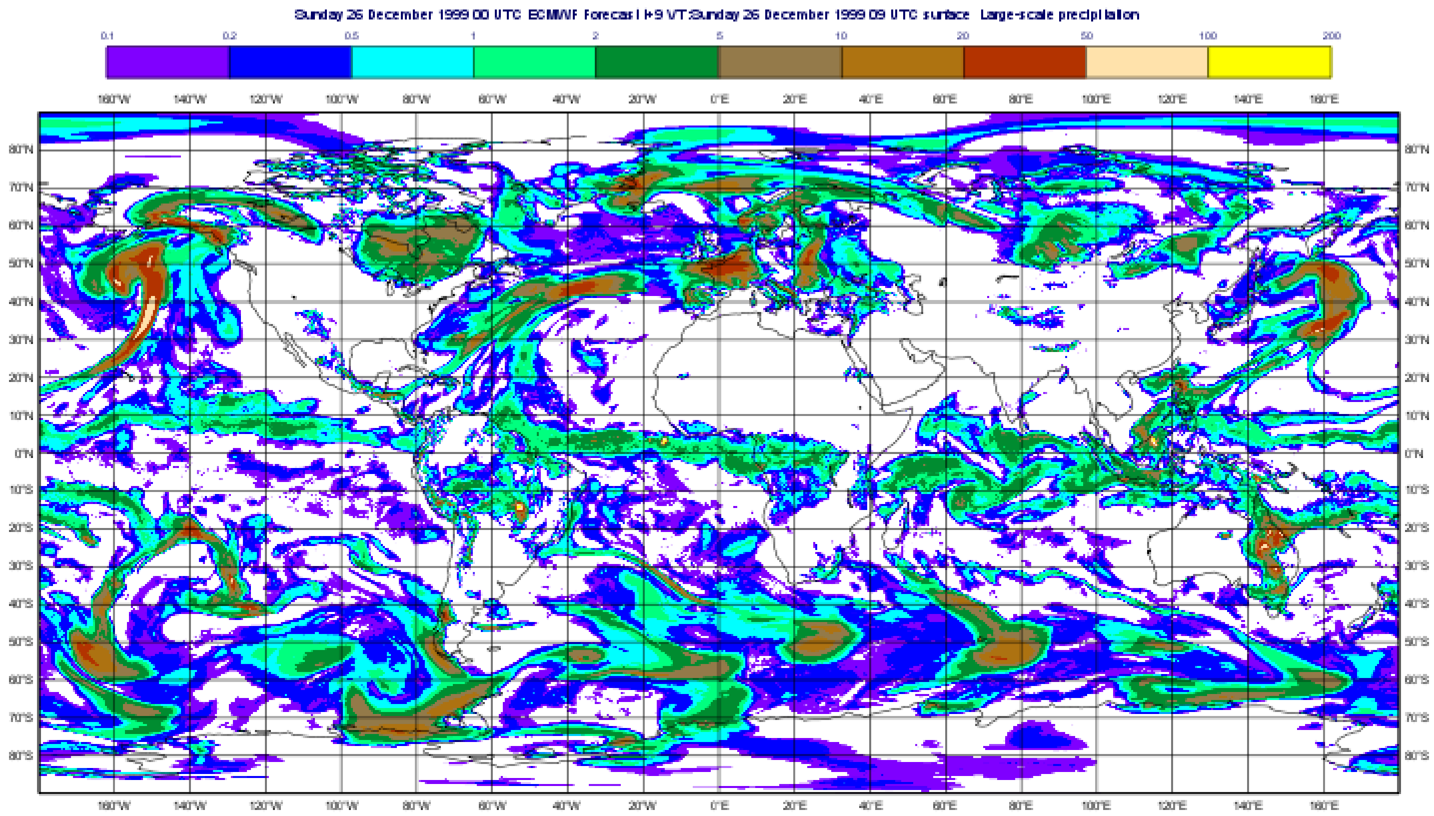
T7999 (~2.5km global) large-scale precipitation

(ran without deep convection parametrization)



ECMWF

T1279 large-scale precipitation



Identified Issues

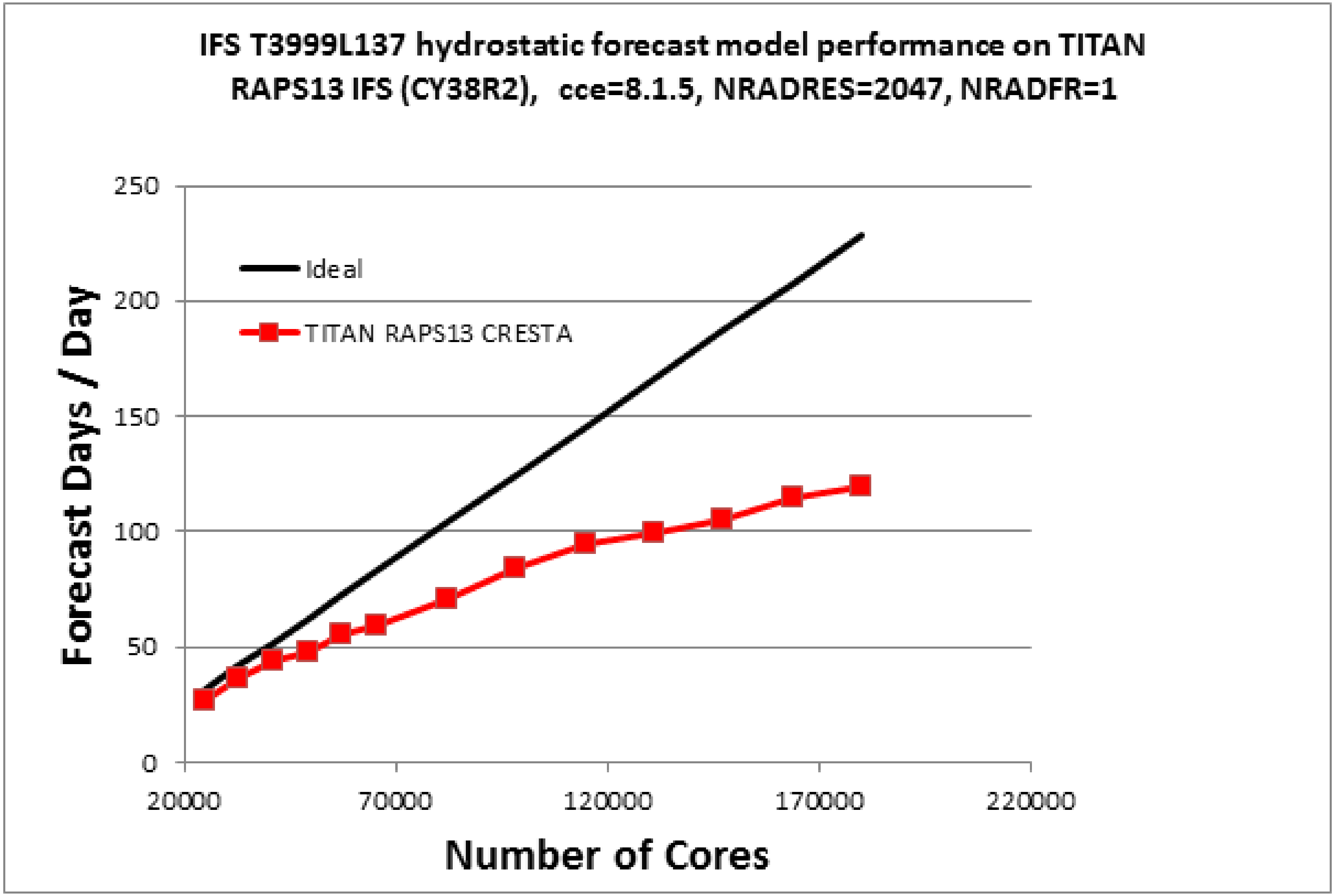
- **The overall scalability of the IFS and in particular the communication cost of (global) transpositions (gridpoint to spectral to gridpoint).**
- **Efficient exploitation of and lacking sufficient local parallelism (outside the spectral transforms) suitable for future computing architectures.**
- **The constant-coefficient, semi-implicit scheme is explicit in the boundary forcing (potentially a steep orography issue).**
- **Lacking the ability to compute local derivatives and fluxes and a satisfactory treatment of vertical derivatives.**
- **Lacking the ability to solve an elliptic equation that (nonlinearly) couples horizontal and vertical discretization.**
- **Coupling of moist physics and dynamics at non-hydrostatic scales.**

CRESTA project

- Parallelism too limited, need to explore new ways to unlock more parallelism
- Overlapping of communication and computations
- Reducing communications and data movement
- New data structures allowing alternative grids
- New approaches to data parallelism and I/O given that ultra-high resolution spectral transform models may be viable in ensemble mode

>> See George's presentation

IFS scalability



ECMWF

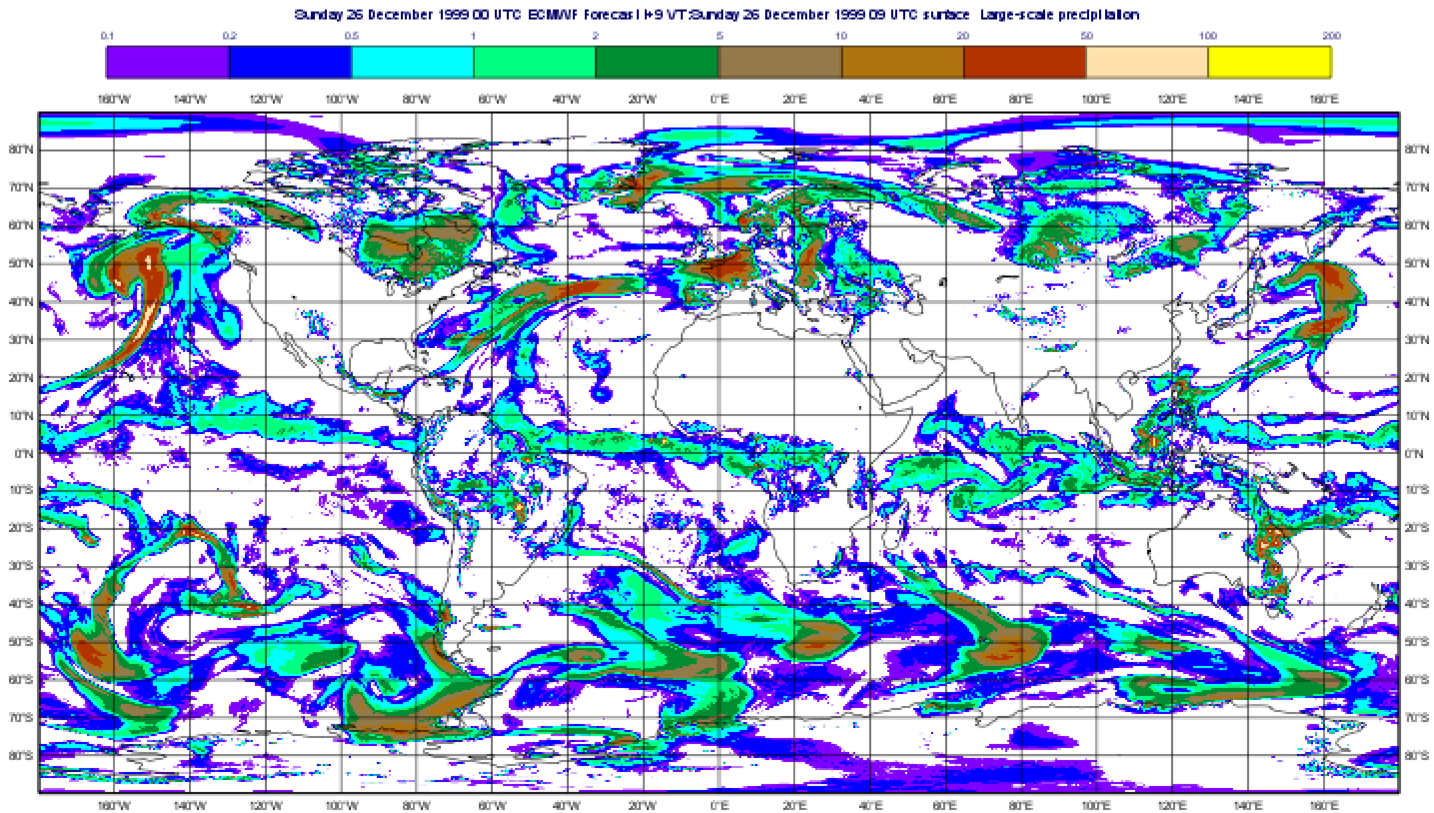
Panta Rhei Project - ERC Advanced Grant

- **Focus on alternative pathways for accurate modelling of multi-scale fluid flows**
- **Combining several disciplines of Earth System sciences via a unifying modelling framework with the initial focus on advancing a global nonhydrostatic forecast model, and finding an "optimal" equation set/time-stepping algorithm**
- **Step change in forecast skill, model accuracy and cost-effectiveness**
- **A pioneering numerical approach, where a non-hydrostatic global model is conditioned by global hydrostatic solutions within a single code framework via numerical procedures expressed in time-dependent generalized curvilinear coordinates**

>> see Piotr's presentation

Additional slides

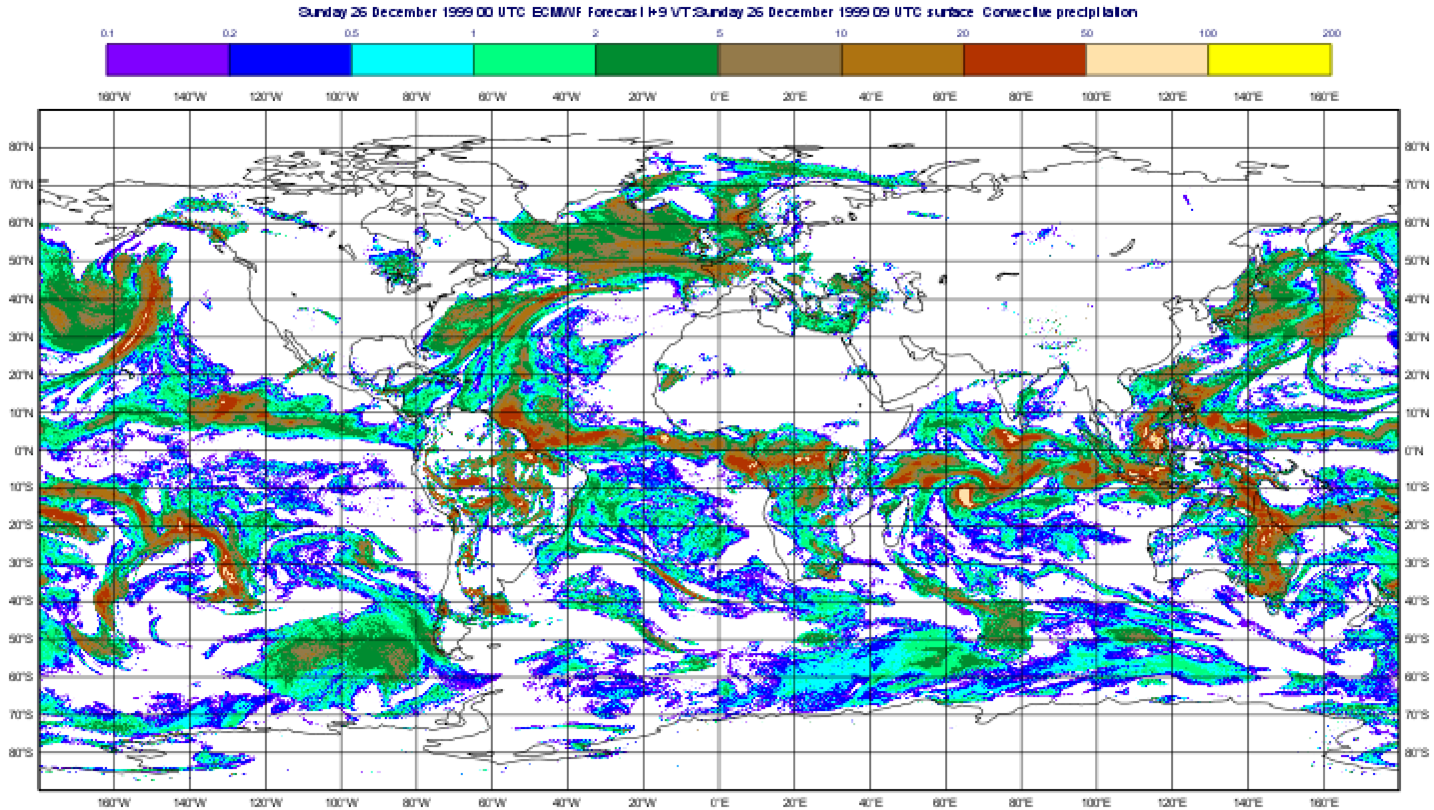
T7999 large-scale precipitation
(ran with deep convection parametrization)



ECMWF

T7999 convective precipitation

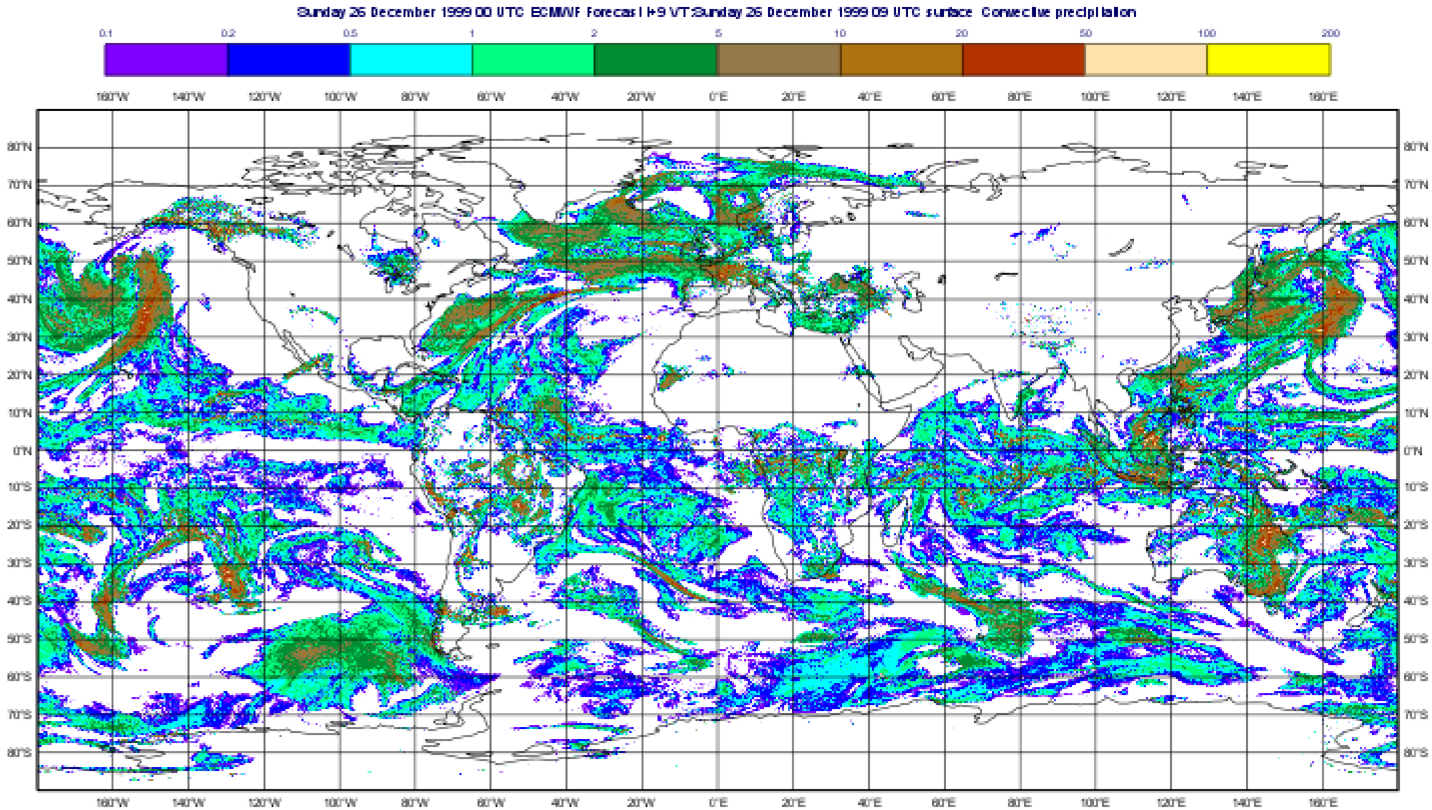
(ran with deep convection parametrization)



ECMWF

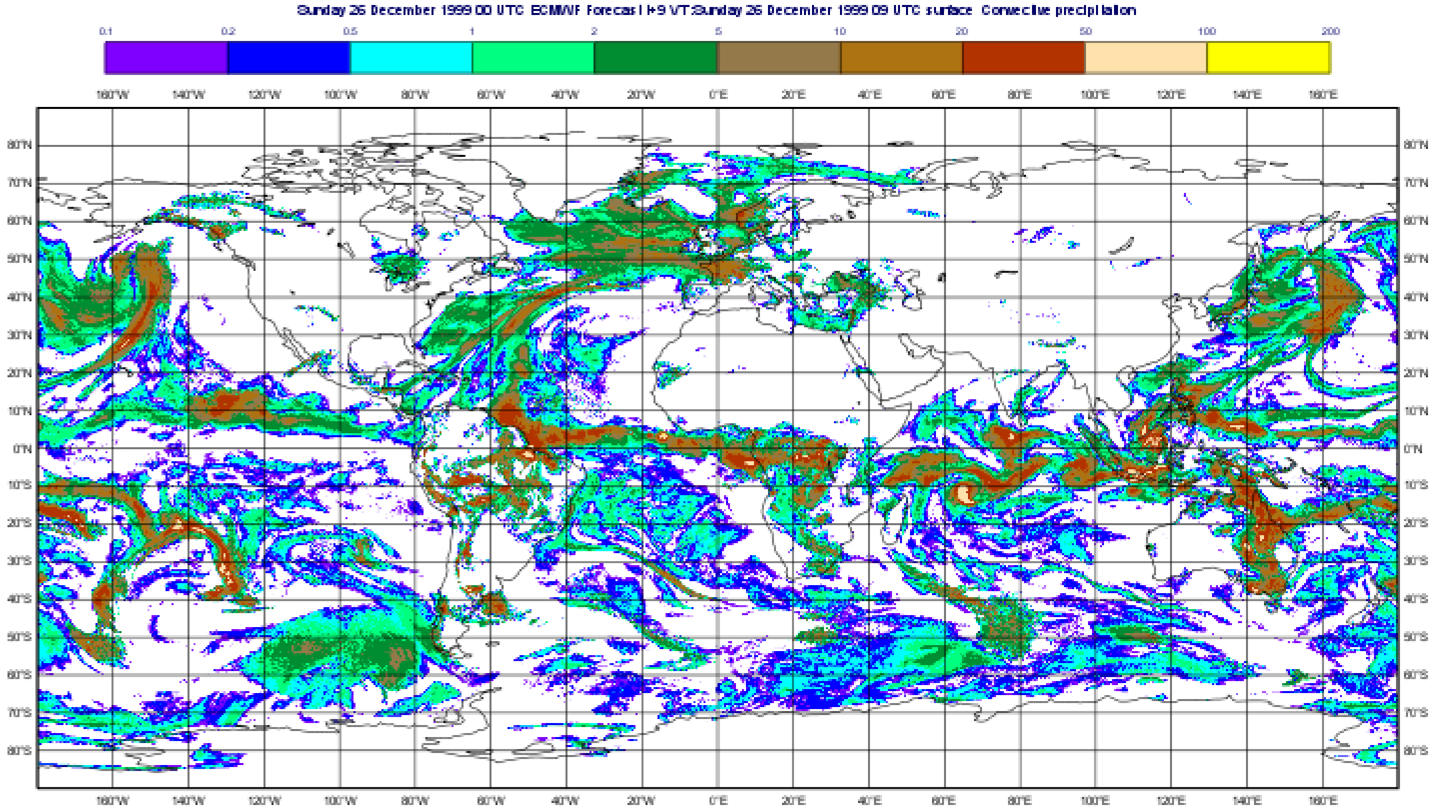
T7999 convective precipitation

(ran without deep convection parametrization)



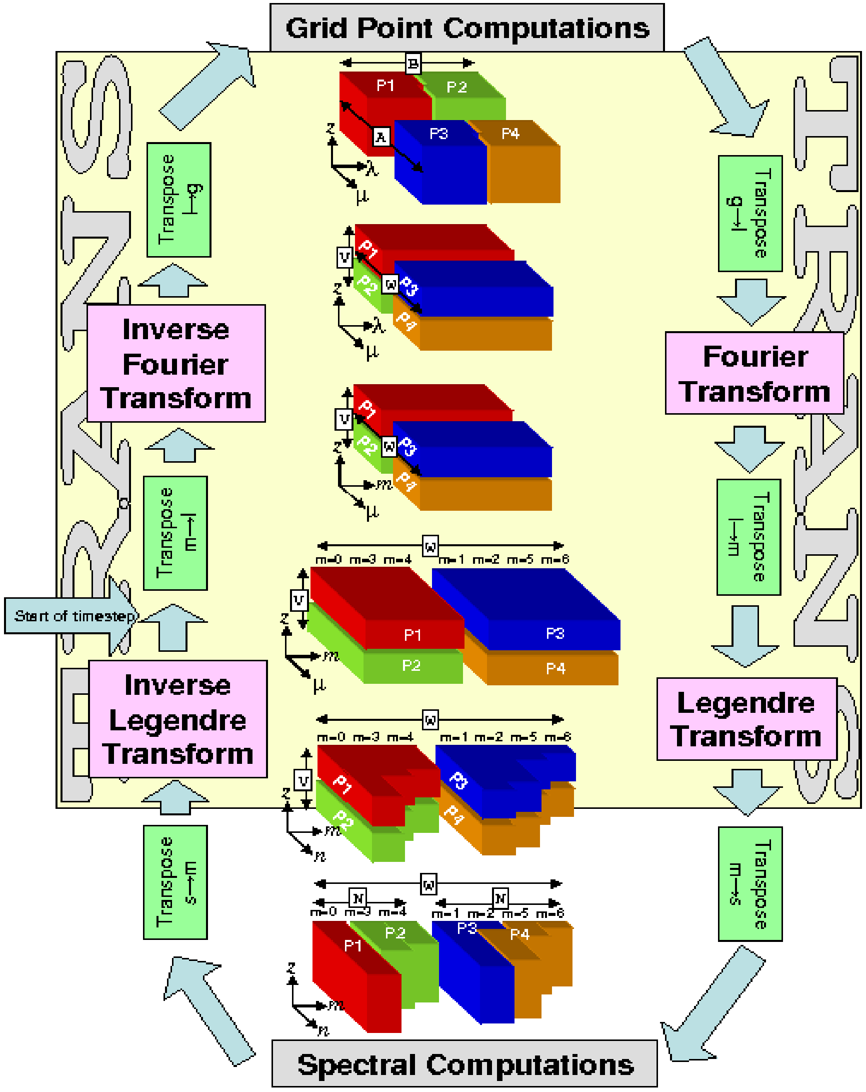
ECMWF

T1279 convective precipitation



ECMWF

Transpositions within the spectral transforms



ECMWF

Nonhydrostatic IFS (NH-IFS)

Bubnovă et al. (1995); Bénard et al. (2004), Bénard et al. (2005), Bénard et al. (2010), Wedi et al. (2009), Yessad and Wedi (2011)

- **Arpégé/ALADIN/Arome/HIRLAM/ECMWF nonhydrostatic dynamical core, which was developed by Météo-France and their ALADIN partners and later incorporated into the ECMWF model and also adopted by HIRLAM.**

Hydrostatic Primitive Equations (HPE)

$$\frac{d\mathbf{V}}{dt} = -2(\Omega \times \mathbf{V}) - \nabla\Phi - RT\nabla(\log\Pi) + \mathbf{P_V}$$

$$\frac{dT}{dt} = \frac{RT}{c_\mathrm{p}}\frac{\omega}{\Pi} + P_\mathrm{T}$$

$$\frac{\partial\Phi}{\partial\Pi} = -\frac{RT}{\Pi}$$

Non-hydrostatic shallow atmosphere (NHS)

Distinction between hydrostatic pressure and total pressure

$$\partial \Pi / \partial z = -\rho g$$

$$p = \rho R T$$

Introduce: $\hat{Q} \equiv \log(p/\Pi)$ 'Nonhydrostatic pressure departure'

Here hydrostatic pressure follows from the prognostic surface pressure equation as before !

Note that the geopotential is derived from

$$\frac{\partial \Phi}{\partial \Pi} = -\frac{RT}{p}$$

NHS – continued ...

$$\frac{d\mathbf{V}}{dt} = -2\Omega \times \mathbf{V} - \frac{\partial p}{\partial \Pi}\nabla\Phi - RT\frac{\nabla p}{p} + \mathbf{P_V}$$

$$\frac{dw}{dt} = g_0 \frac{\partial(p - \Pi)}{\partial \Pi} + P_{\mathrm{w}}$$

$$\frac{dT}{dt} = -\frac{RT}{c_{\mathrm{v}}}D_3 + \frac{c_{\mathrm{p}}}{c_{\mathrm{v}}}P_{\mathrm{T}}$$

$$\frac{d\hat{Q}}{dt} = -\frac{c_{\mathrm{p}}}{c_{\mathrm{v}}}D_3 - \frac{\omega}{\Pi} + \frac{c_{\mathrm{p}}}{c_{\mathrm{v}}T}P_{\mathrm{T}}$$

'Physics'

Projecting on temperature and horizontal velocities only, quasi-anelastic coupling ?!

NHS – For the solution we need the three-dimensional divergence

$$D_3 = D + \mathcal{X} + \frac{R_\mathrm{d}}{R} d$$

horizontal divergence

$$D \equiv \nabla_\eta \cdot \mathbf{V}$$

'vertical divergence'

$$d \equiv -(g_0 p / m R_\mathrm{d} T) \partial w / \partial \eta$$

'X-term residual'

$$\mathcal{X} = \frac{p}{mRT} \nabla \Phi \cdot \left(\frac{\partial \mathbf{V}}{\partial \eta} \right)$$

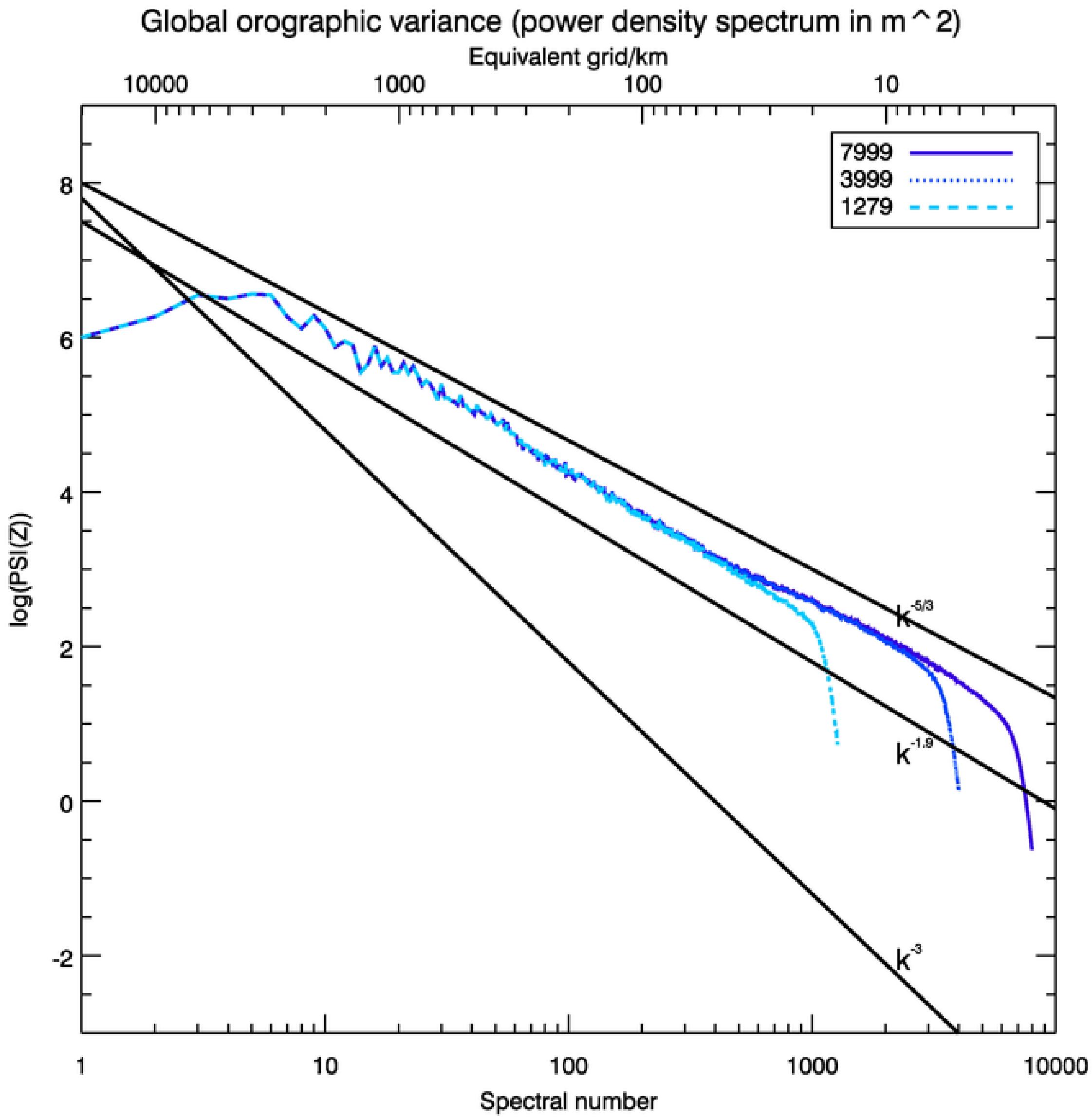
Arises due to the formulation of divergence in time-dependent curvilinear coordinates !

Numerical solution

- **Two-time-level, semi-implicit, semi-Lagrangian.**
- **Semi-implicit procedure with two reference states, with respect to gravity and acoustic waves, respectively.**
- **The resulting Helmholtz equation can be solved (subject to some constraints on the vertical discretization) with a *direct spectral method*, *that is, a mathematical separation of the horizontal and vertical part of the linear problem in spectral space, with the remainder representing at most a pentadiagonal problem of dimension NLEV^2. Non-linear residuals are treated explicitly (or iteratively implicitly)!***

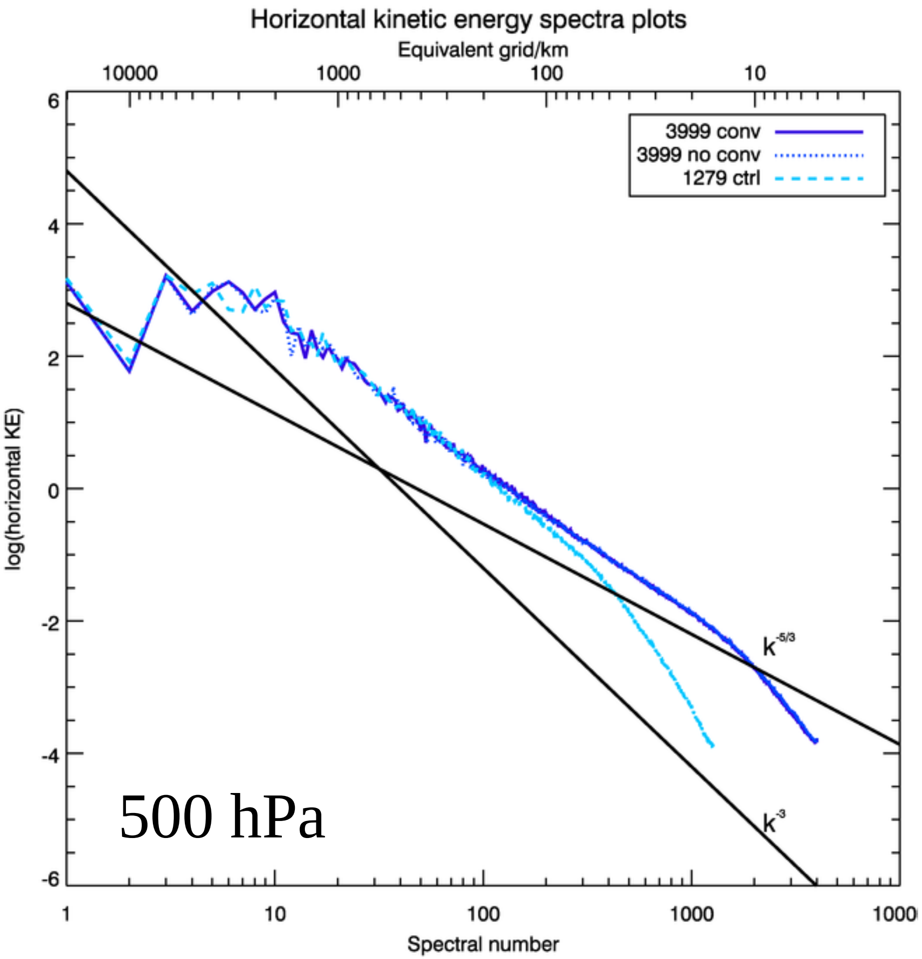
(Robert, 1972; Bénard et al 2004,2005,2010)

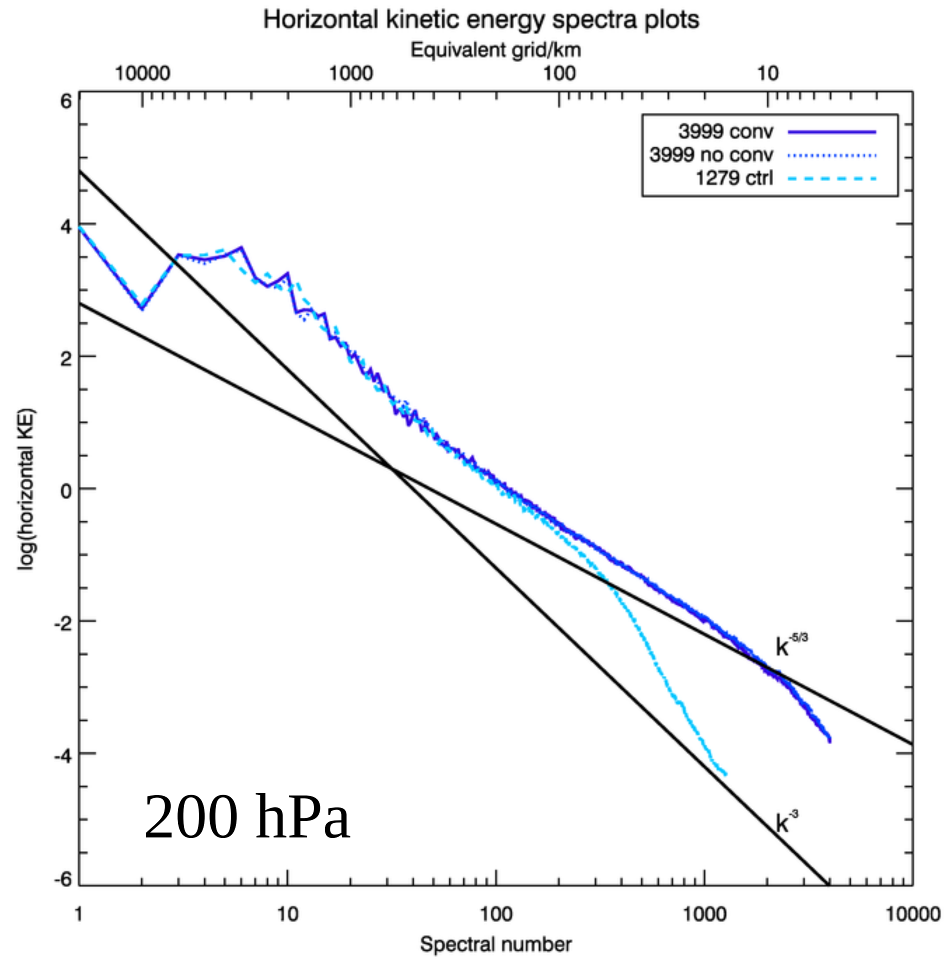
Orographic forcing



ECMWF

T3999 vs T1279 kinetic energy spectra





After 10 days

Global horizontal kinetic energy spectra at 500 hPa height for the first 12 hours of the T7999