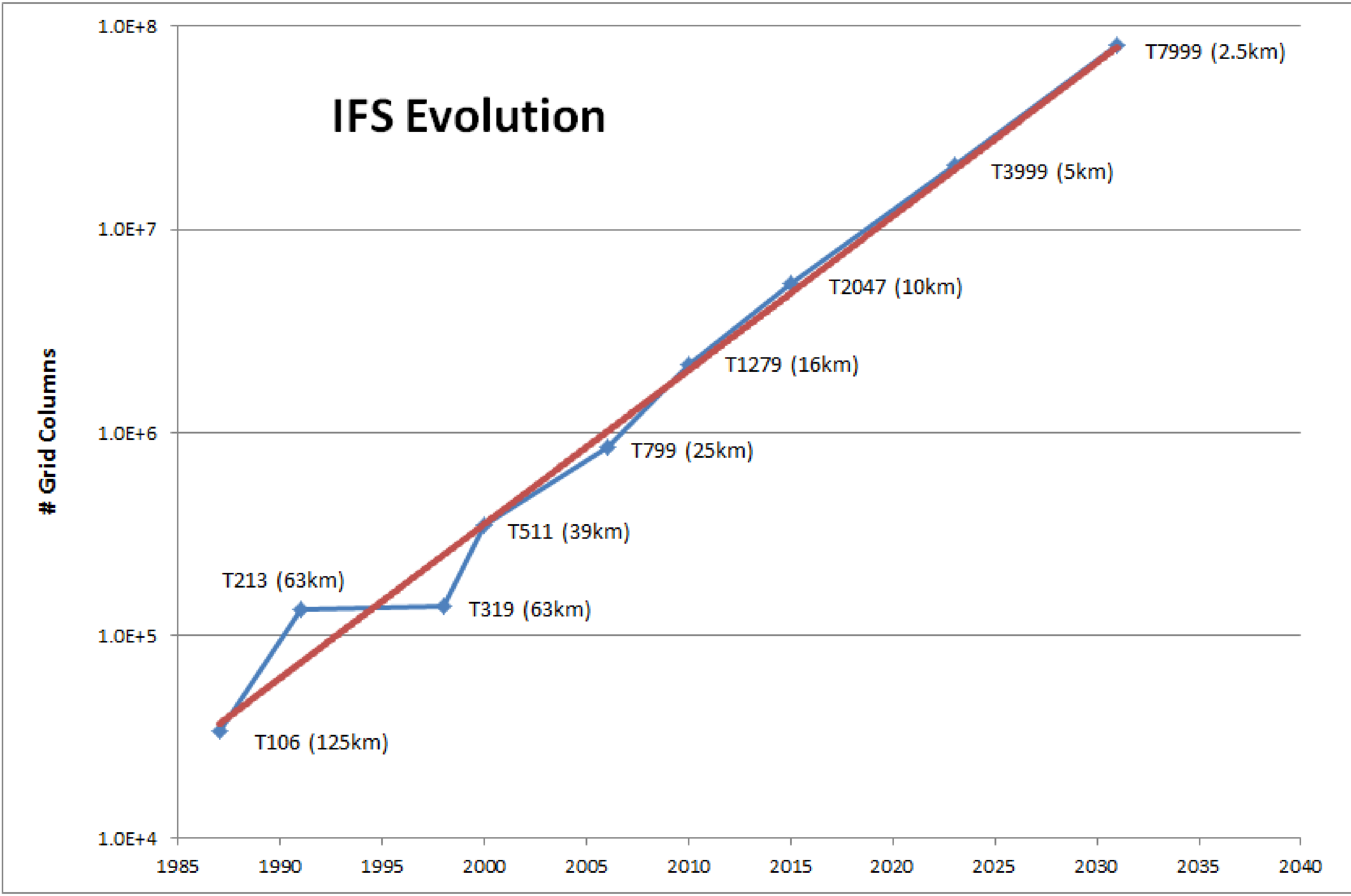
The spectral transform method

by Nils Wedi
European Centre for Medium-Range Weather Forecasts
wedi@ecmwf.int

ECMWF

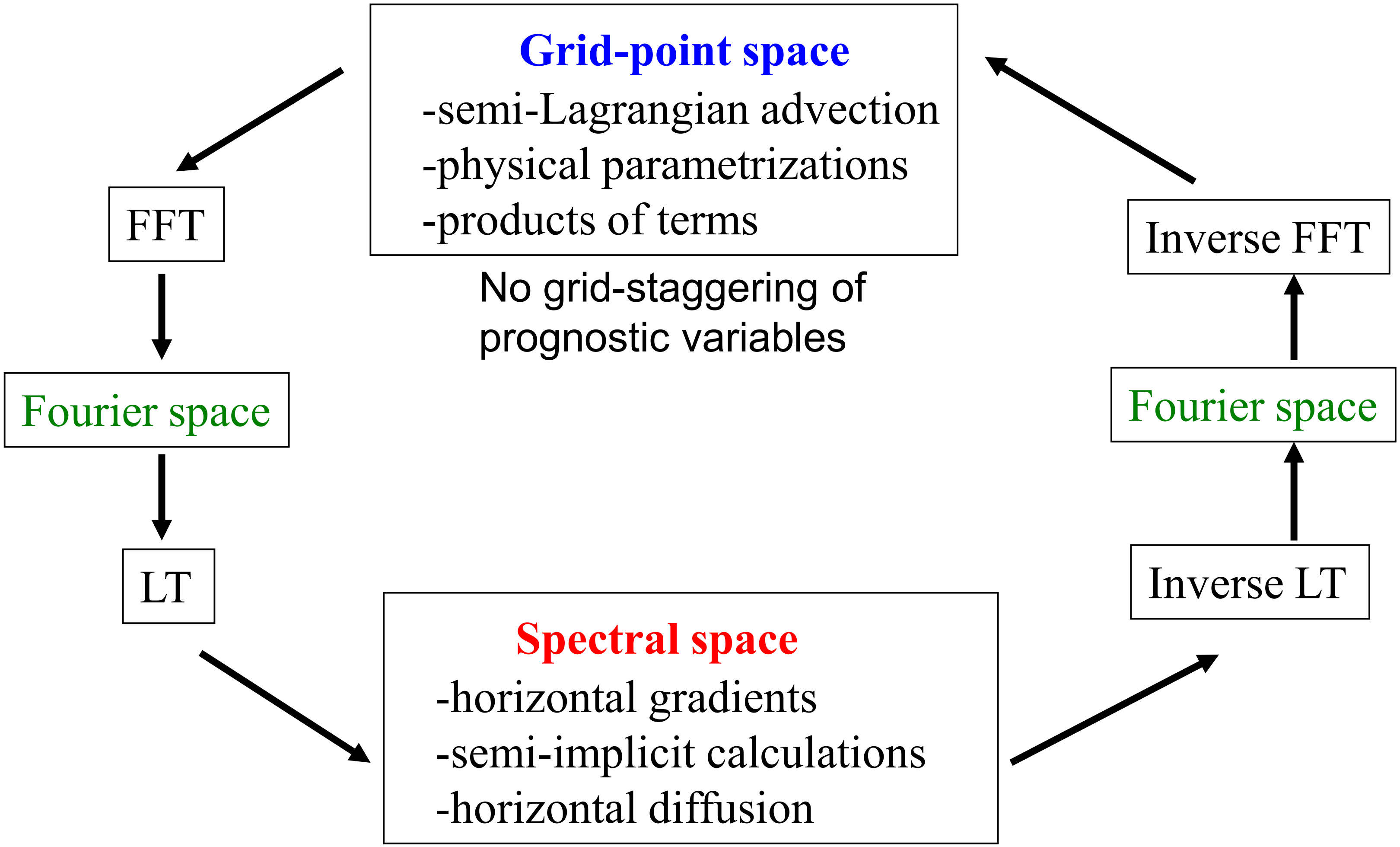

IFS Evolution
Grid Columns
1.0E+8
1.0E+7
1.0E+6
1.0E+5
1.0E+4
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
T7999 (2.5km)
T3999 (5km)
T2047 (10km)
T1279 (16km)
T799 (25km)
T511 (39km)
T213 (63km)
T319 (63km)
T106 (125km)

The Integrated Forecasting System (IFS)

technology applied at ECMWF for the last 30 years …

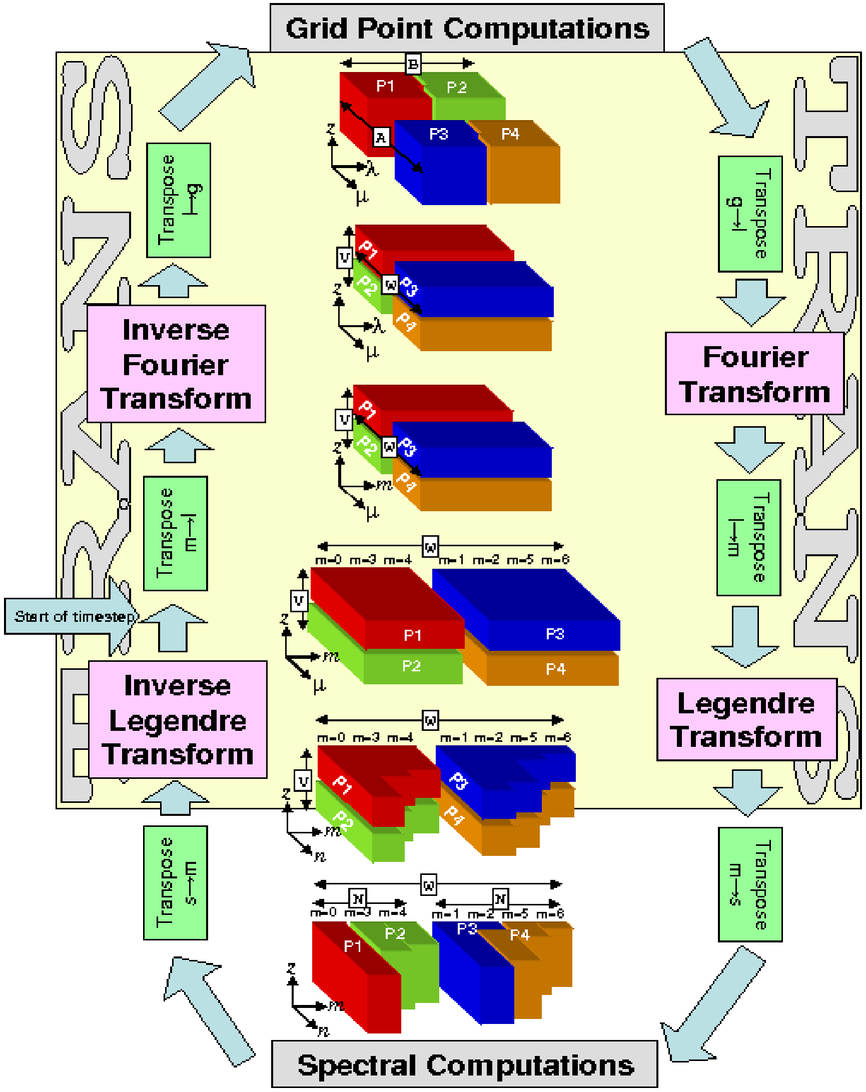
A spectral transform, semi-Lagrangian, semi-implicit (compressible) (non-)hydrostatic model

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



FFT: Fast Fourier Transform, LT: Legendre Transform

Several transpositions within the spectral transforms need to communicate, e.g. using MPI alltoallv



ECMWF

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:

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

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

(normalized) associated Legendre polynomials

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

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

Inverse spectral transform (Backward)

Inverse Legendre transform

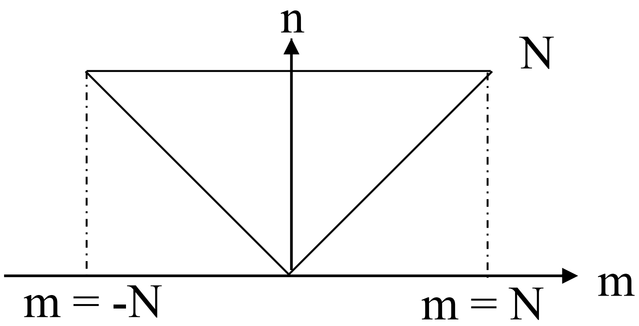
$$\zeta(\theta, \lambda) = \sum_{m=-N}^{N} e^{im\lambda}$$

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

Triangular truncation (isotropic)

Spherical harmonics

Triangular truncation:



A useful property of spherical harmonics

total wavenumber

$$\nabla^2 \zeta_n^m = -\frac{n(n+1)}{a^2} \zeta_n^m$$

Earth radius

Fast Multipole Method (FMM) and spectral filtering *(Boyd, 1992; 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) !

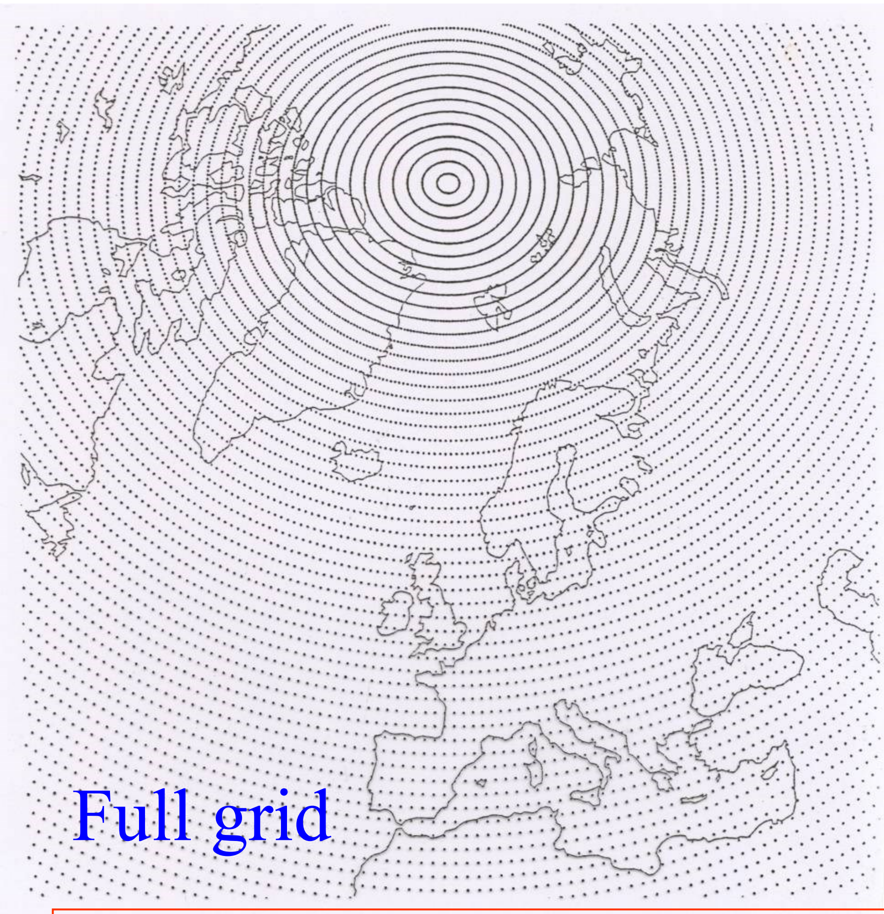
Example: From Christoffel-Darboux formula for associated Legendre polynomials
We can do a direct and inverse Legendre transform for a single Fourier mode as:

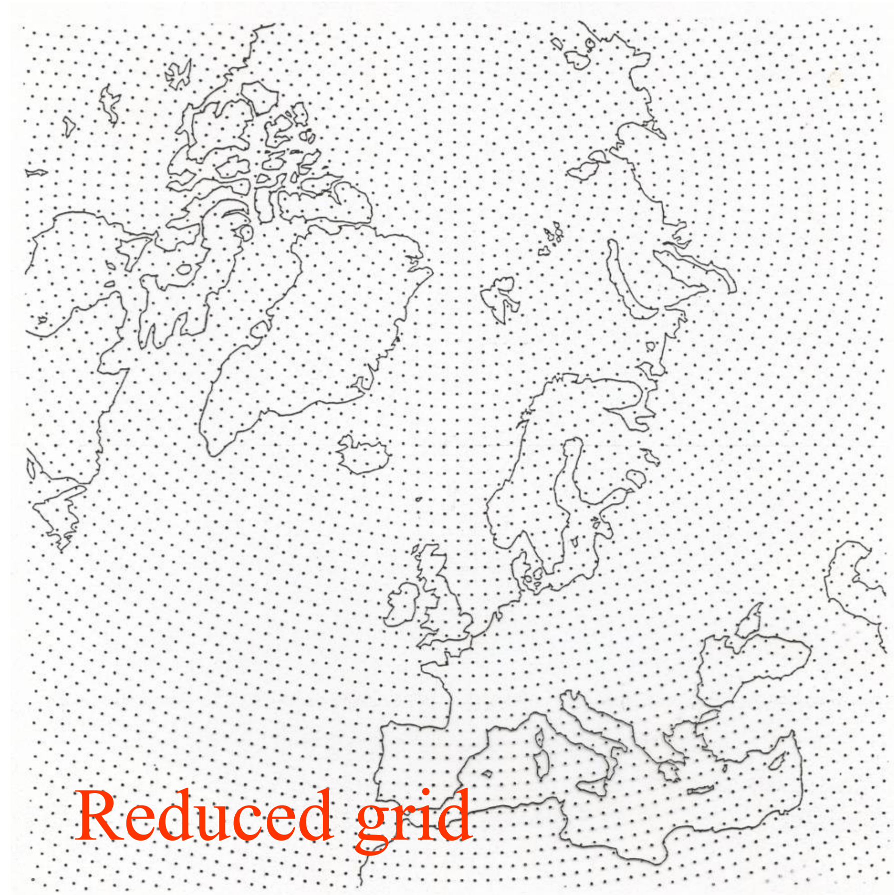
$$\tilde{\zeta}^m(\tilde{\theta}_j) = \epsilon_{N+1}^m \overline{P}_{N+1}^m(\mu_j) \sum_{i=1}^{J} \frac{\zeta^m(\theta_i) w_i \overline{P}_N^m(\mu_i)}{\tilde{\mu}_j - \mu_i}$$

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

The Gaussian grid

About 30% reduction in number of points



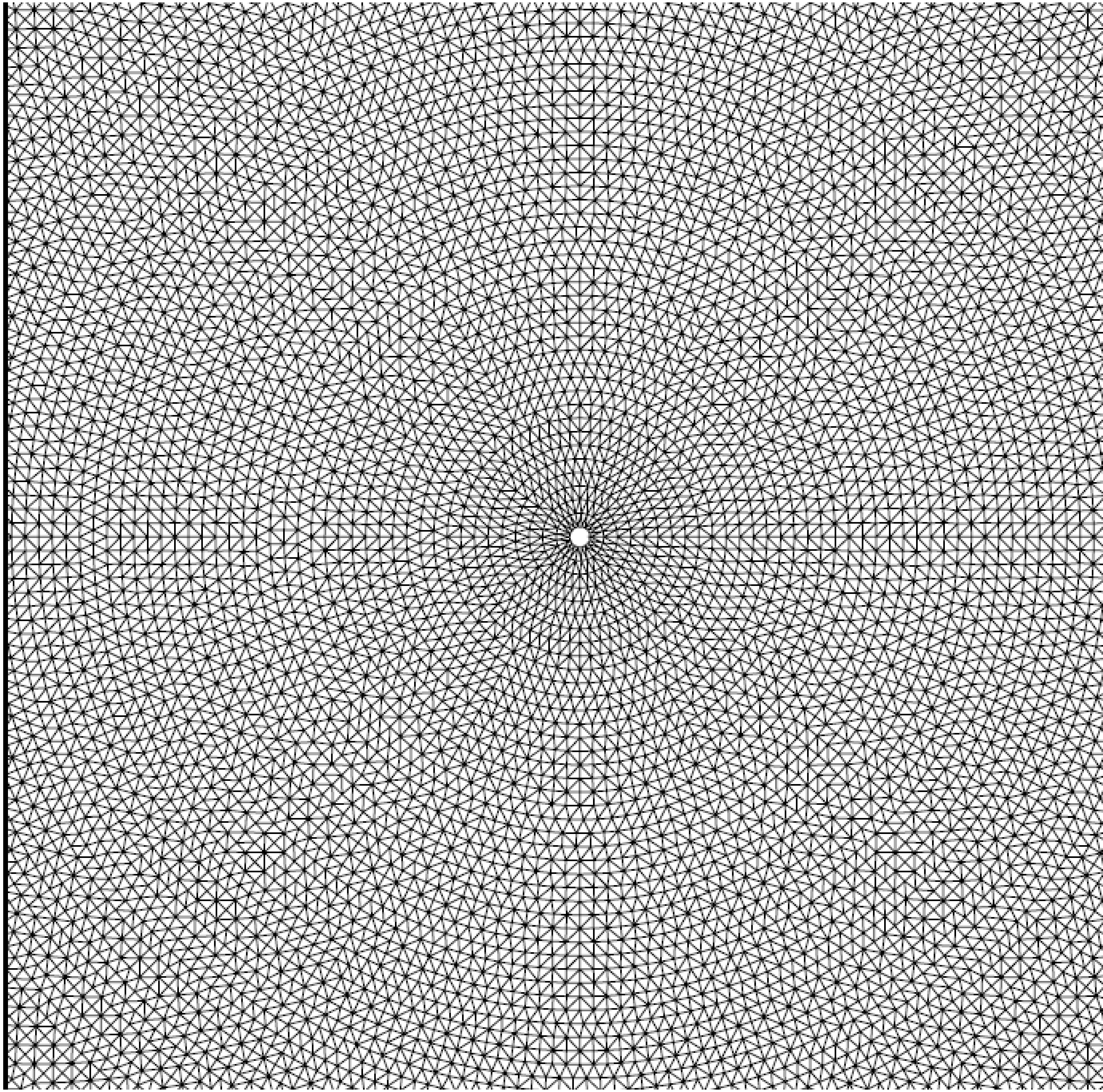


Reduction in the number of Fourier points at high latitudes is possible because the associated Legendre polynomials are very small near the poles for large m.

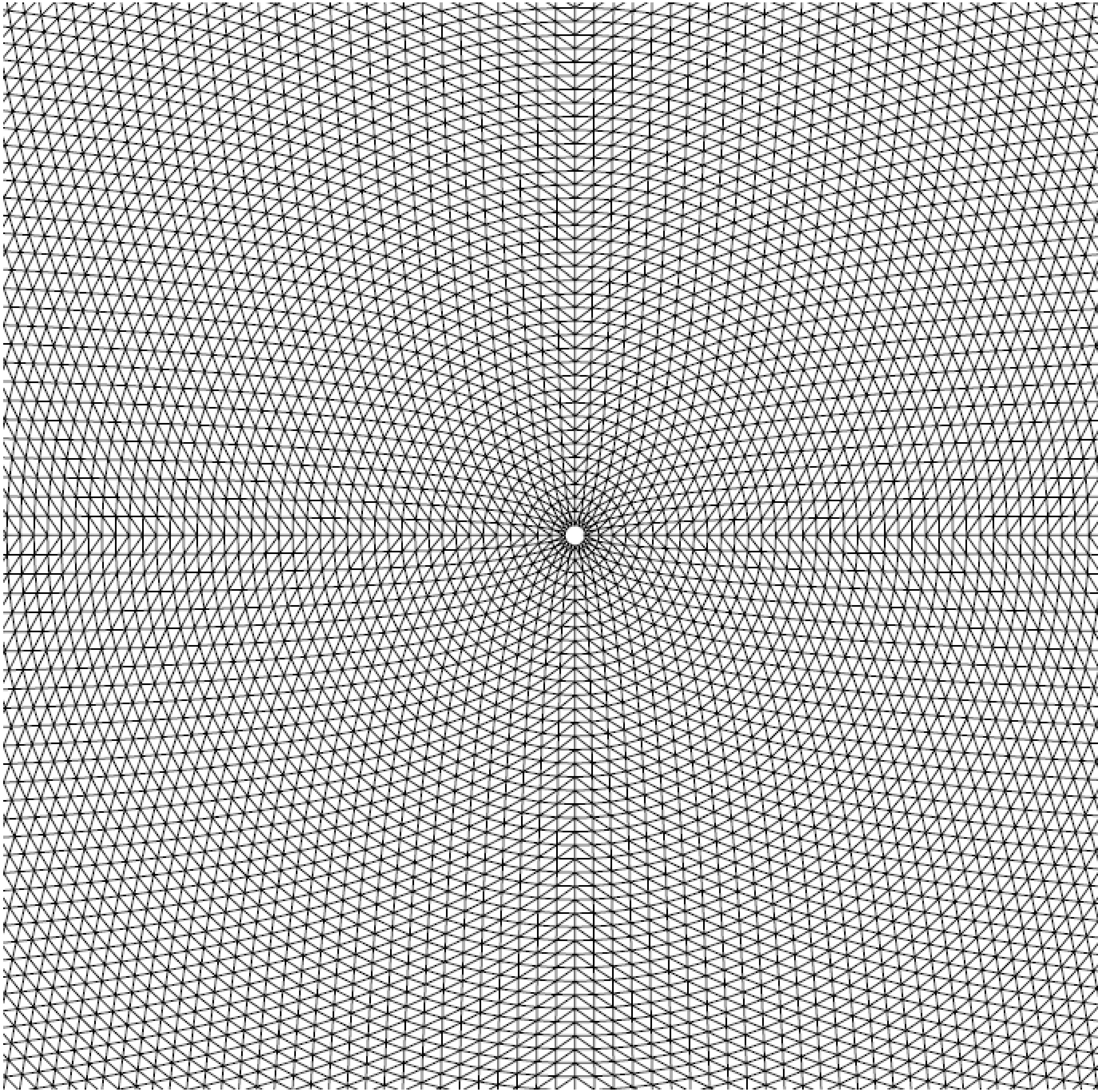
Note: number of points nearly equivalent to quasi-uniform icosahedral grid cells of the ICON model.

ECMWF

Standard reduced Gaussian grid



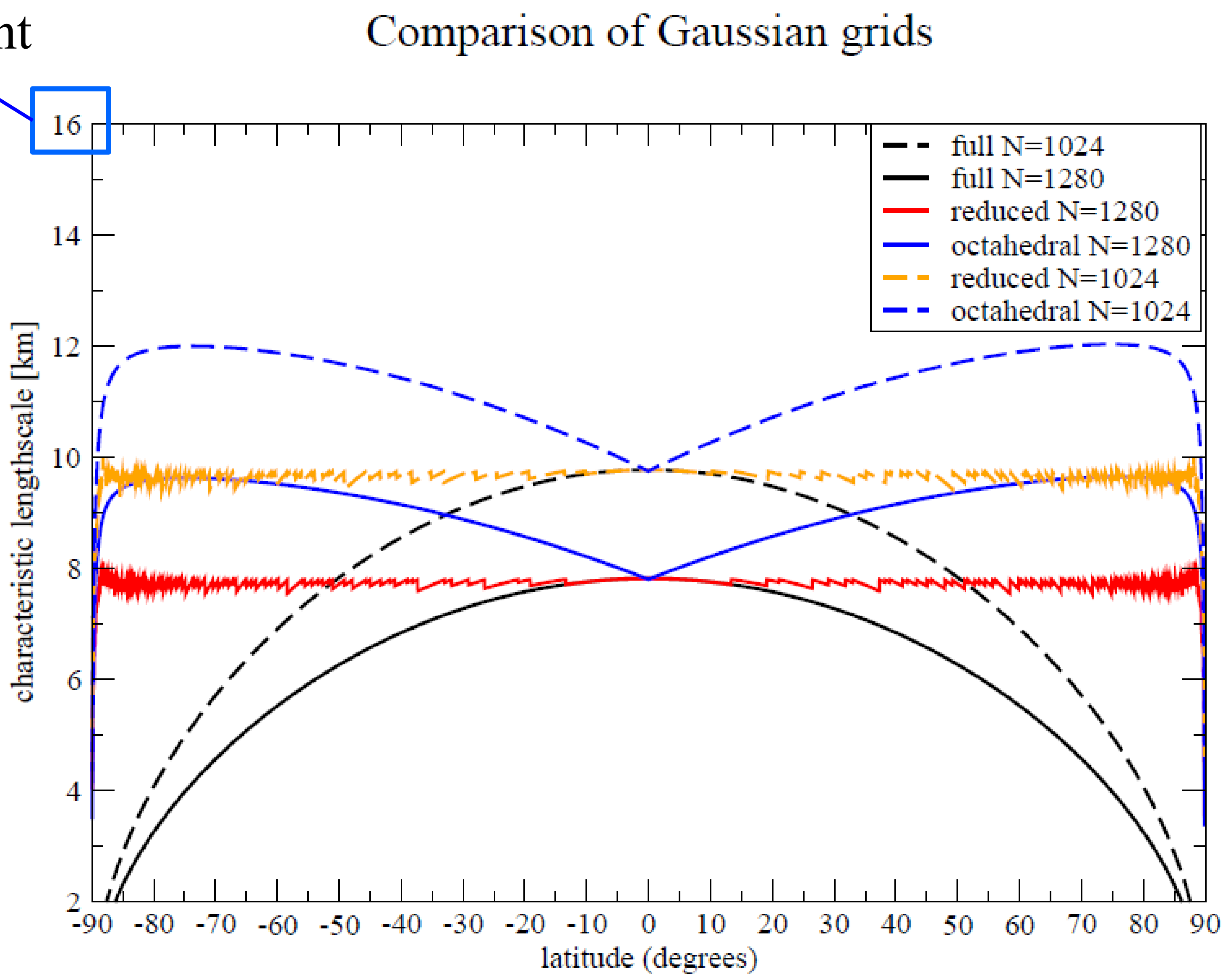
ECMWF



octahedral reduced Gaussian grid

ECMWF

current



Comparison of Gaussian grids
16
14
12
10
8
6
4
2
characteristic lengthscale [km]
-90 -80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90
latitude (degrees)
full N=1024
full N=1280
reduced N=1280
octahedral N=1280
reduced N=1024
octahedral N=1024

Spectral vs. physical space

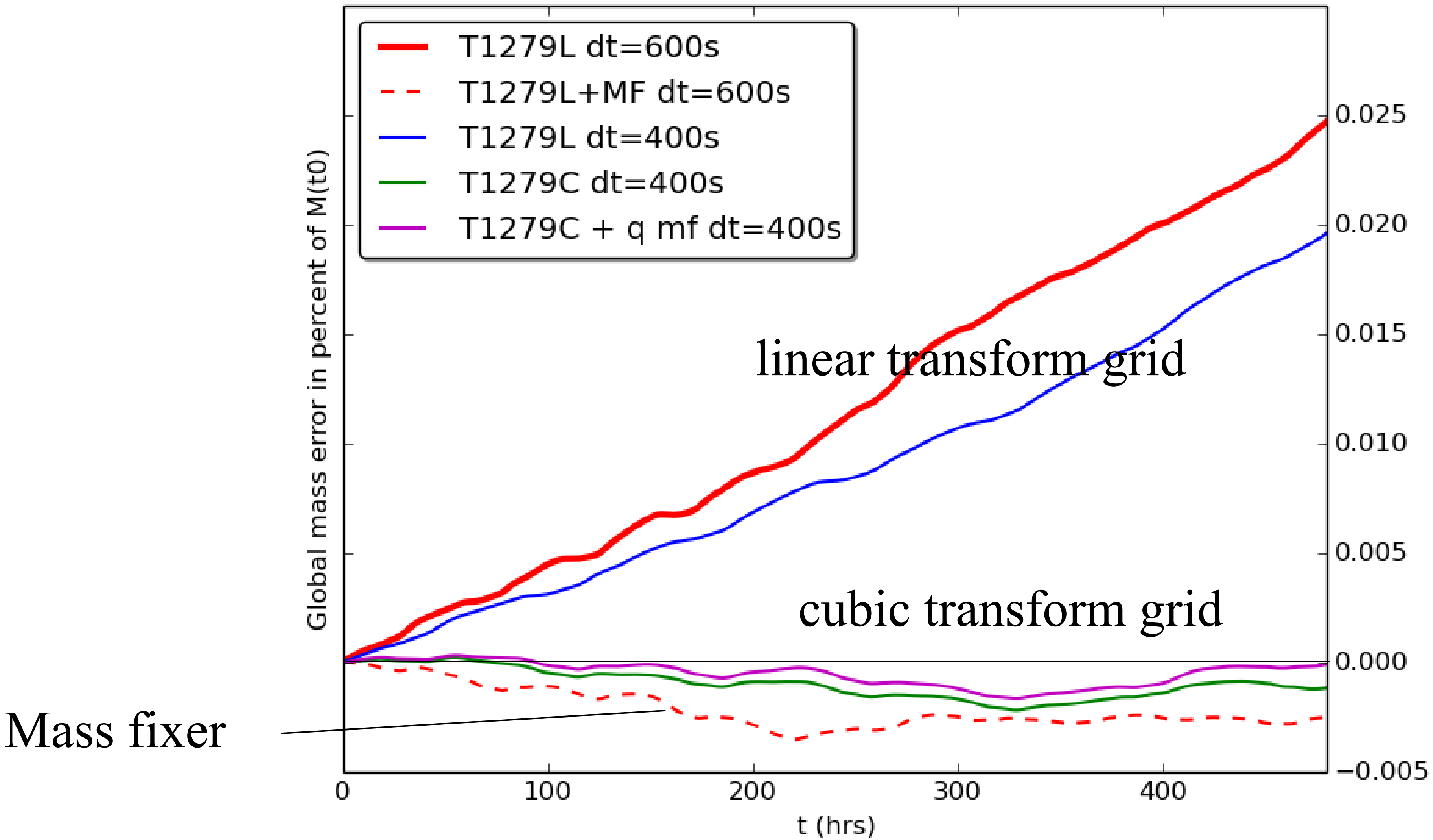
Orszag, J. Atmos. Sci., 28:1074, 1971, "On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components"

	Spatial filter range	
2N+1 gridpoints to N waves : linear grid	~ 1-2 Δ	
3N+1 gridpoints to N waves : quadratic grid	~ 2-3 Δ	
4N+1 gridpoints to N waves : cubic grid	~ 3-4 Δ	*(Wedi, 2014)*

Effective resolution of NWP models today : 6-8 Δ

(Abdalla et al, 2013)

Conservation of global mass



ECMWF

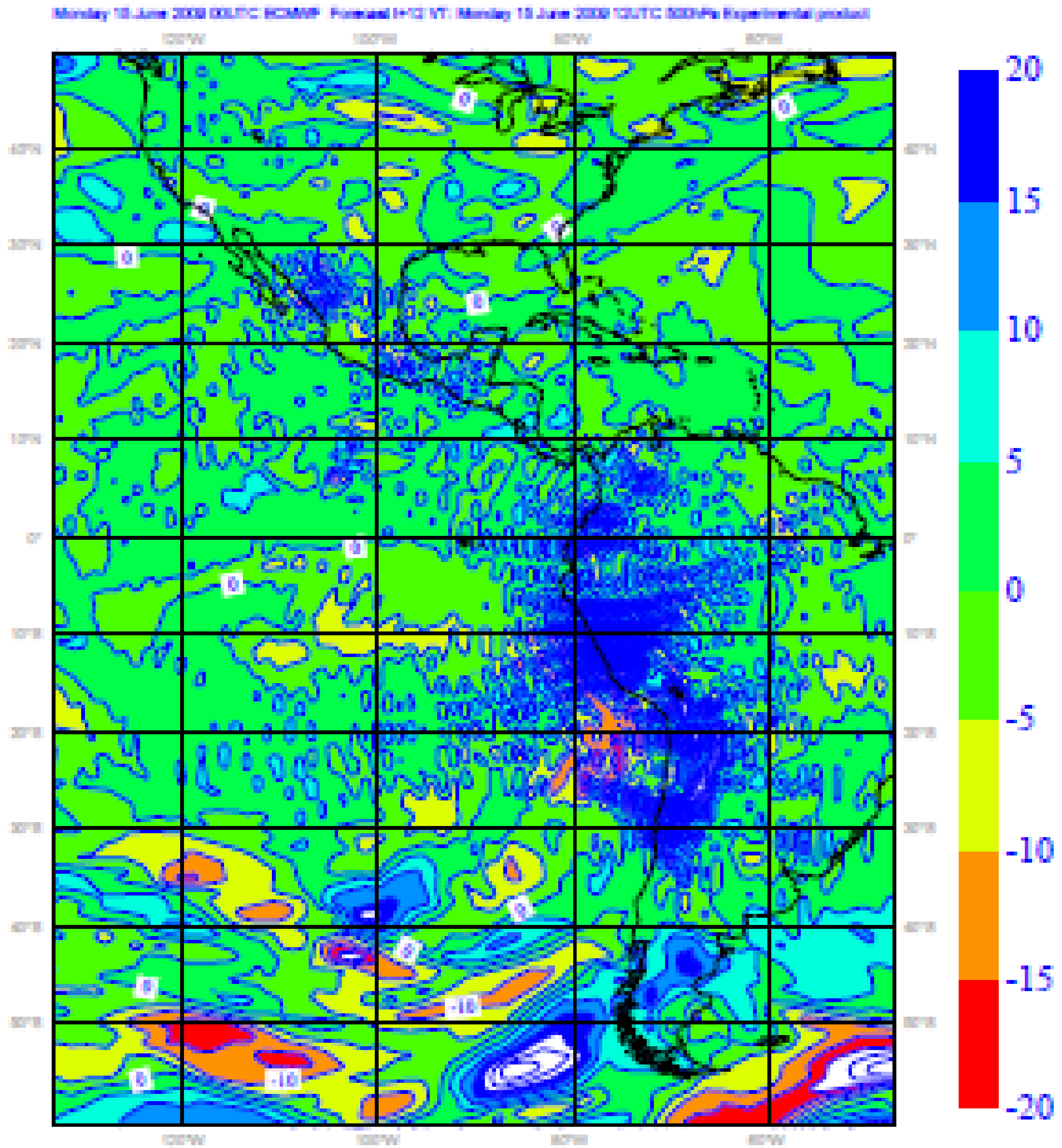
Aliasing

- 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

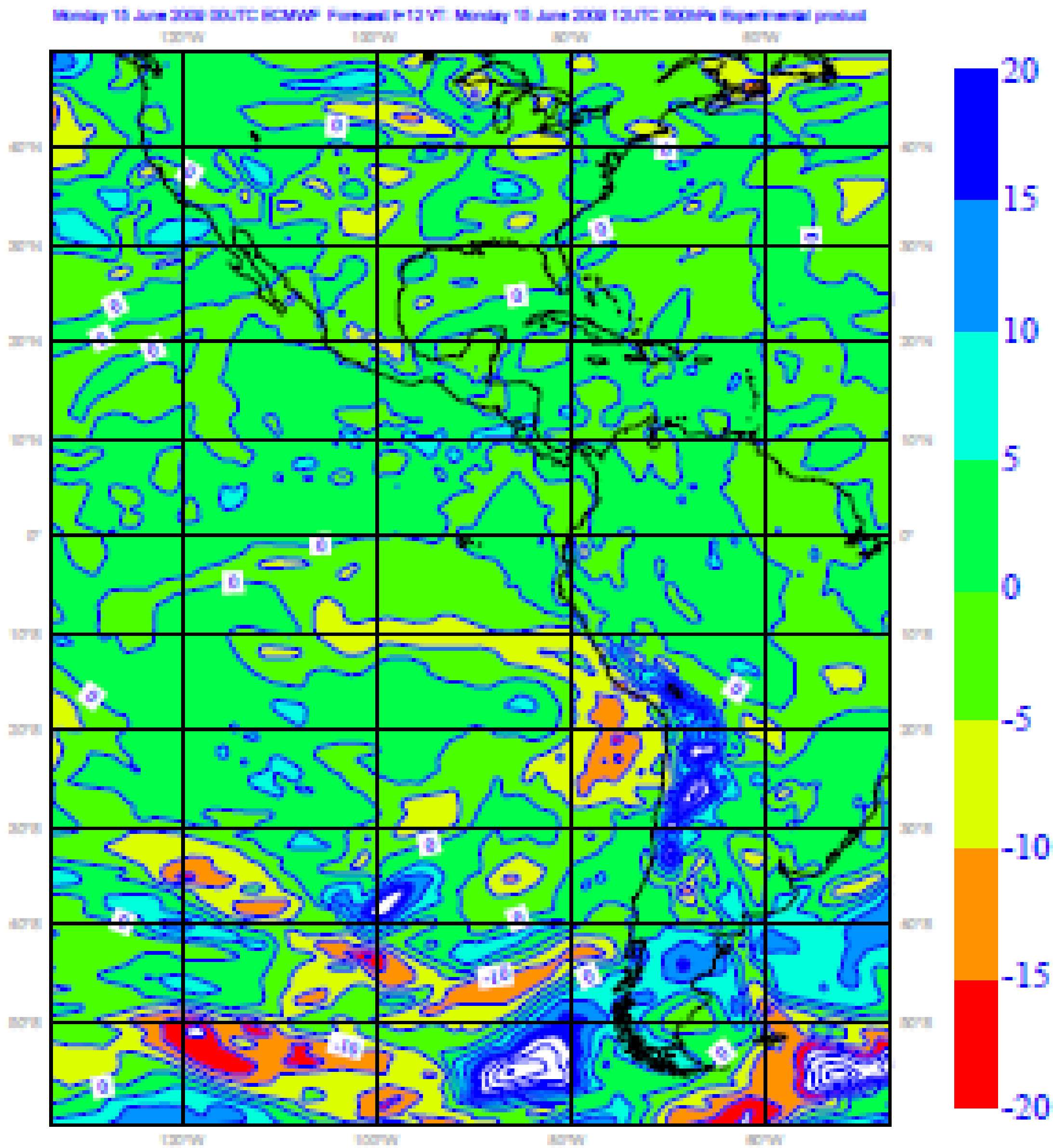
E-W

500hPa adiabatic zonal wind tendencies (T159)



ECMWF

De-aliasing

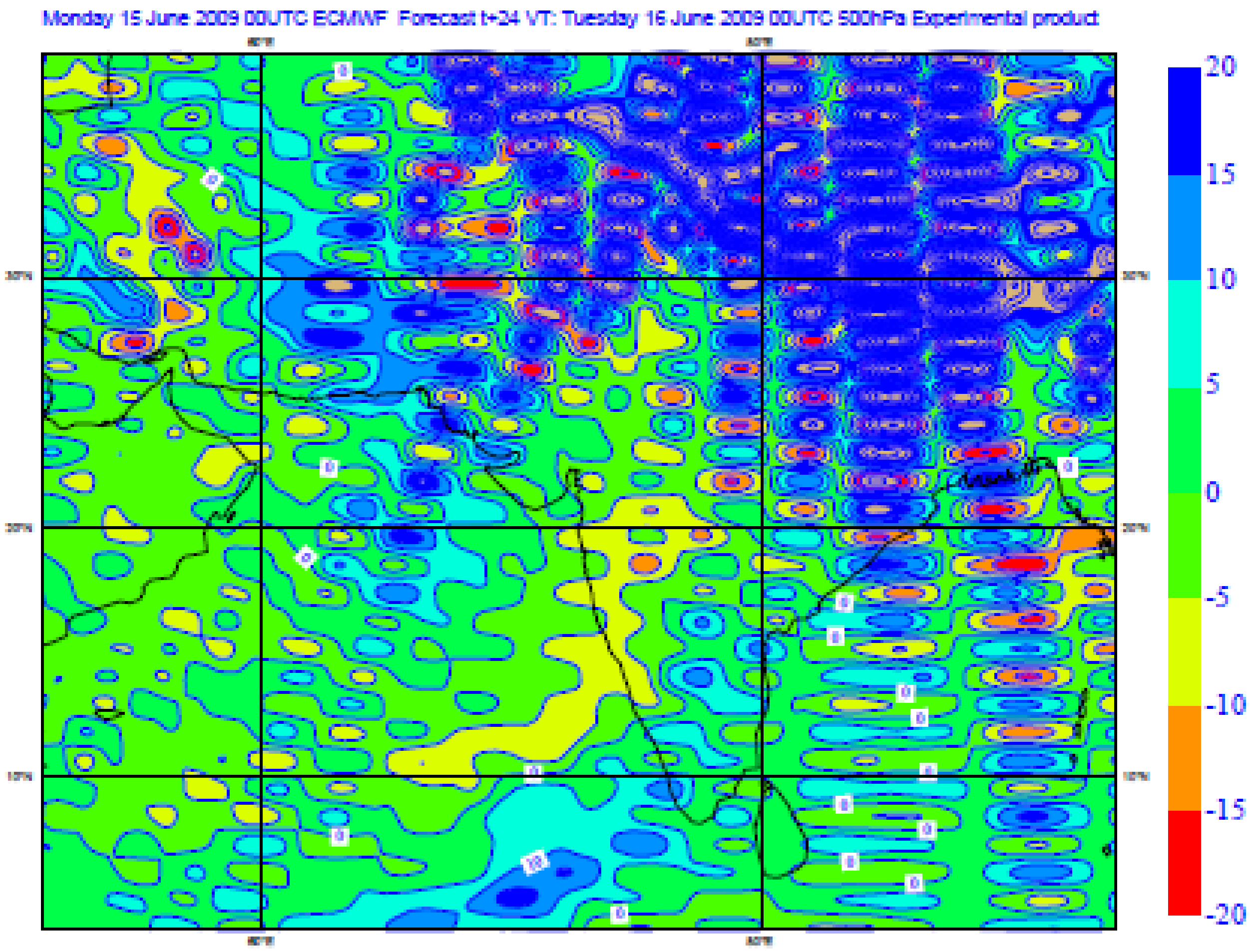


ECMWF

De-aliasing

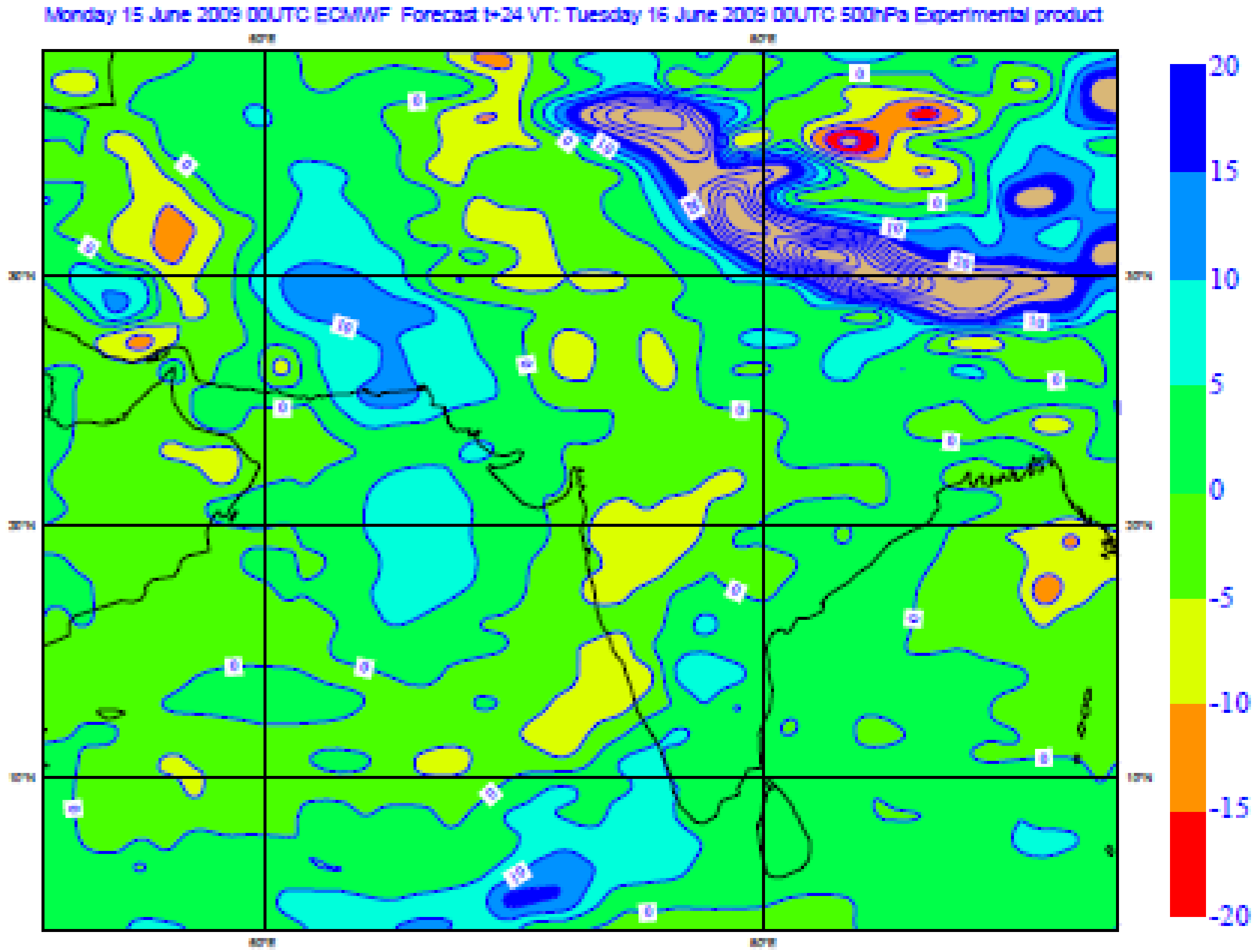
N-S

500hPa adiabatic meridional wind tendencies (T159)



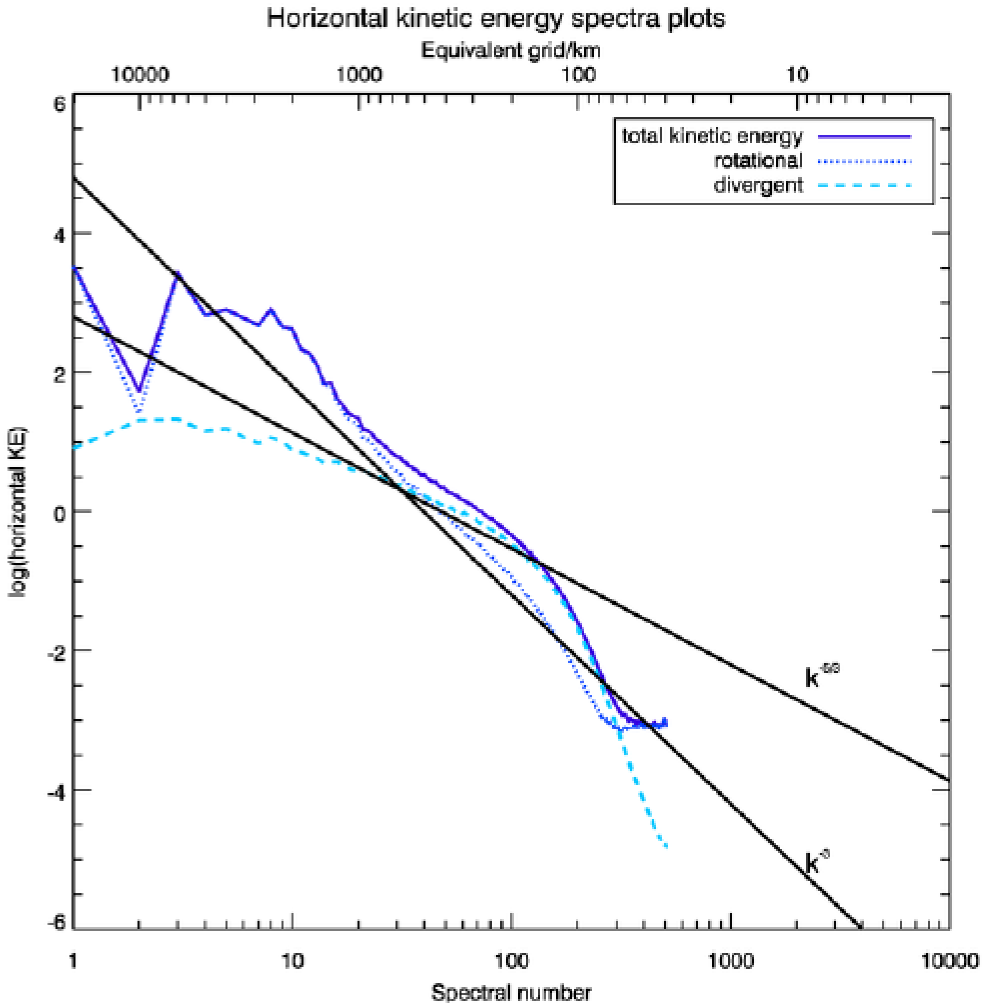
ECMWF

De-aliasing



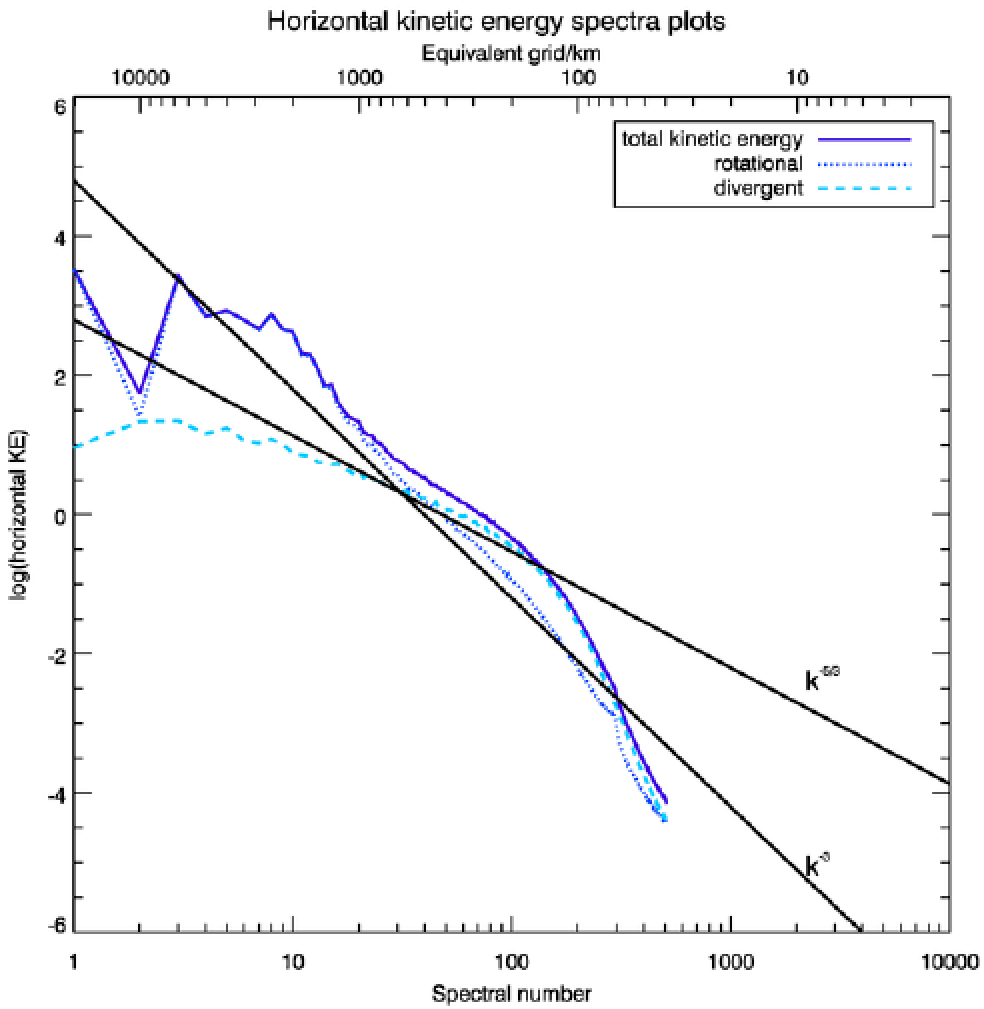
ECMWF

Kinetic Energy Spectra – 100 hPa



ECMWF

Kinetic Energy Spectra – 100 hPa



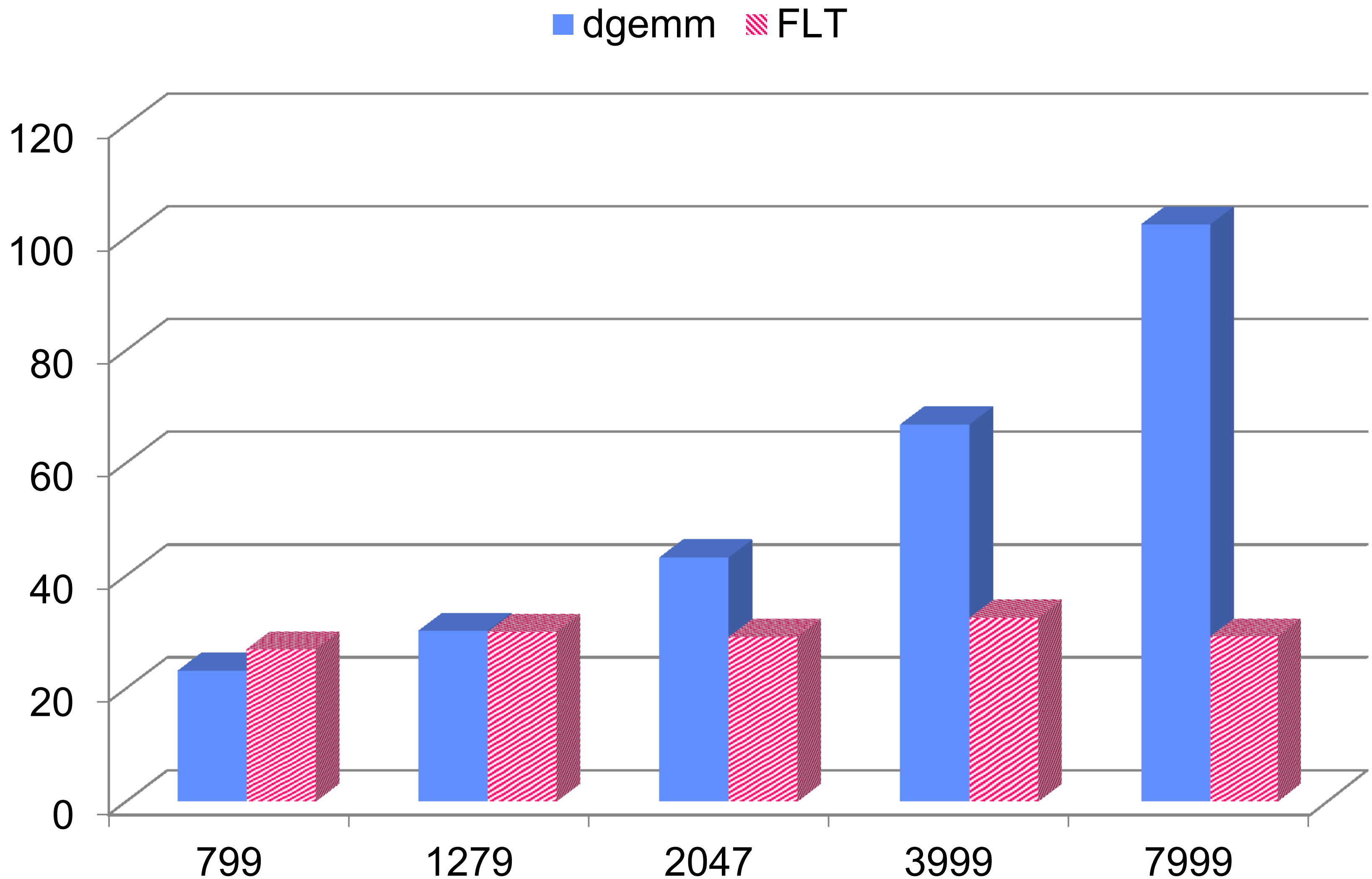
ECMWF

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 $l = 0 \rightarrow L$ 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

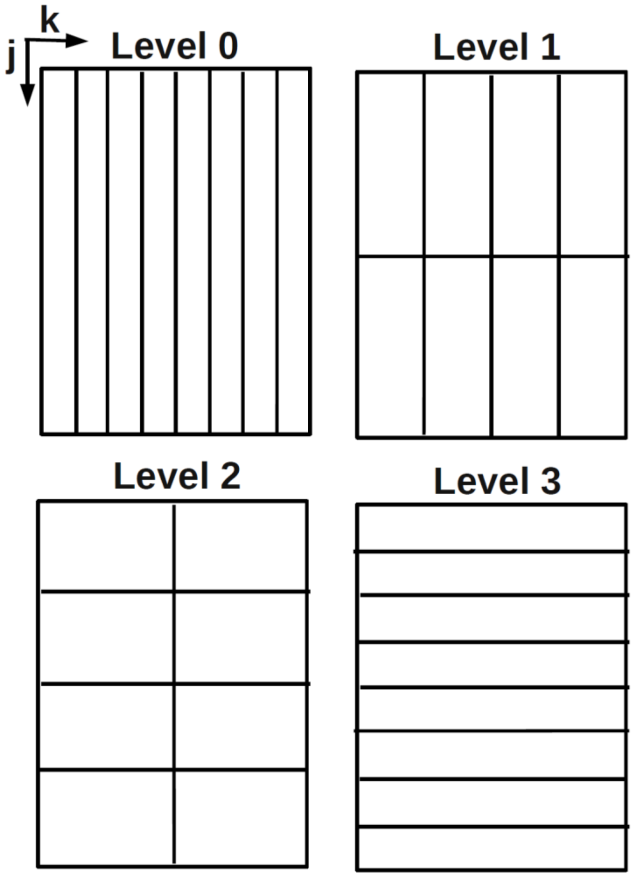
      $store\ C_{L,j}$

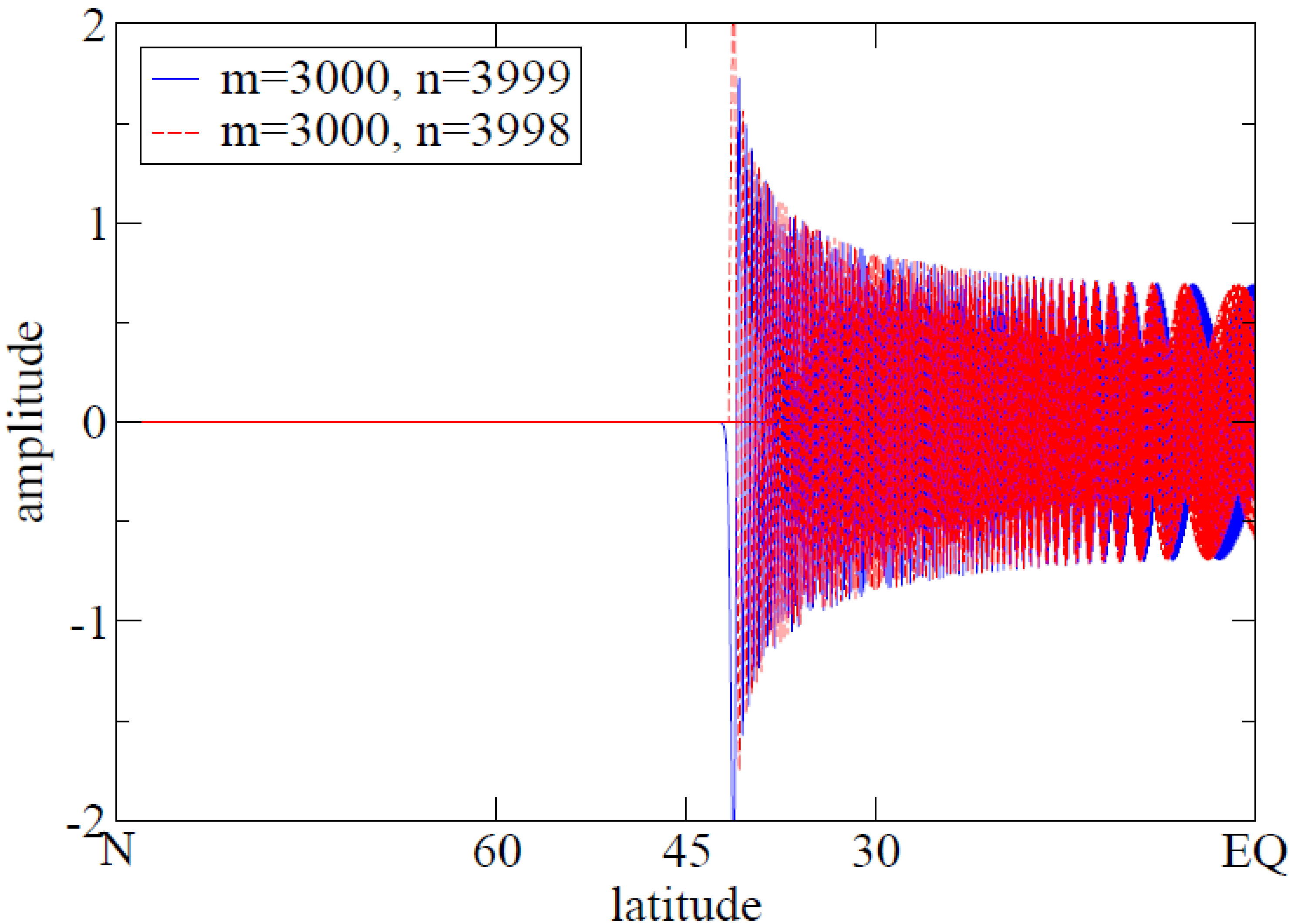
    end if

  end for

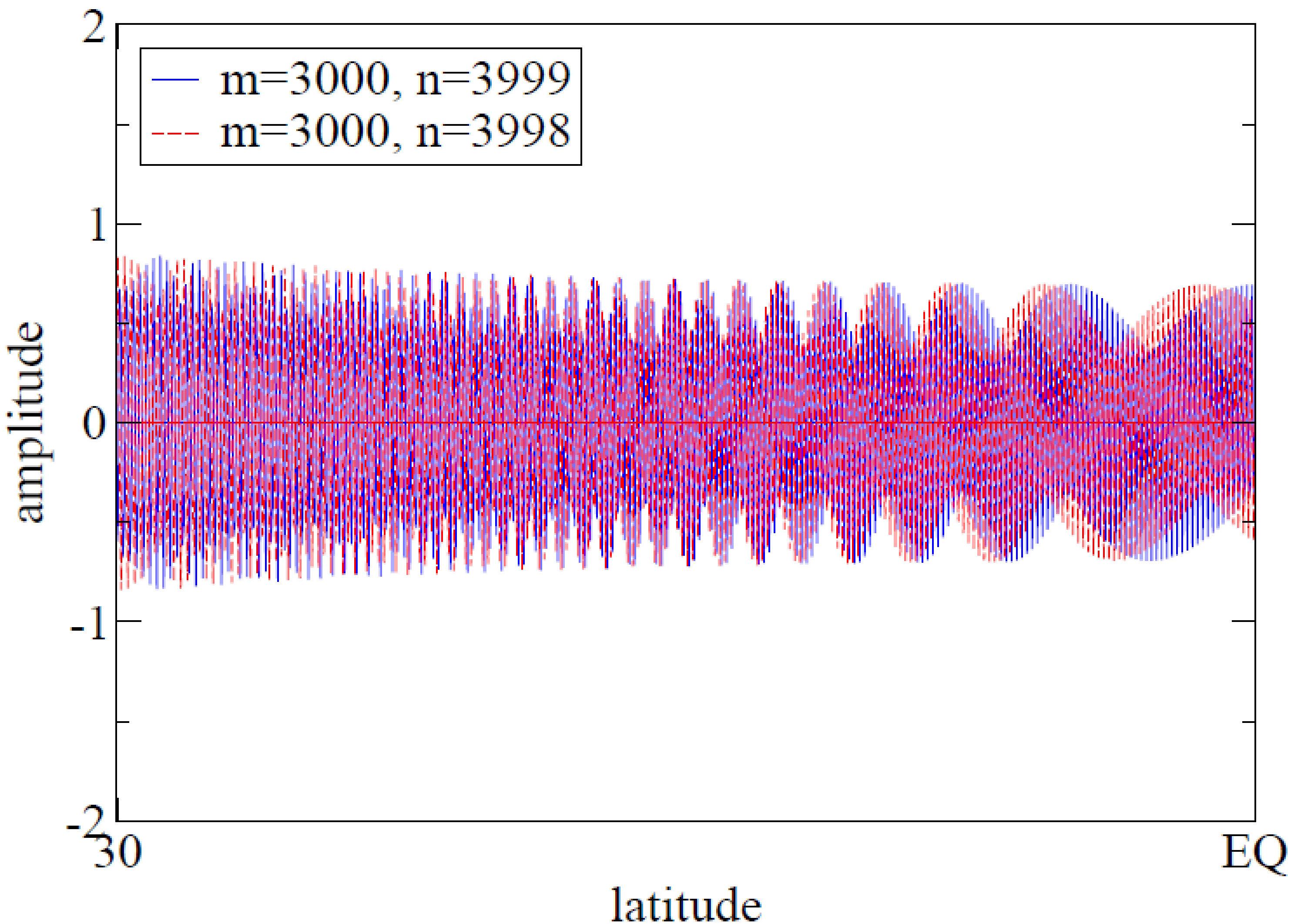
end for

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




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

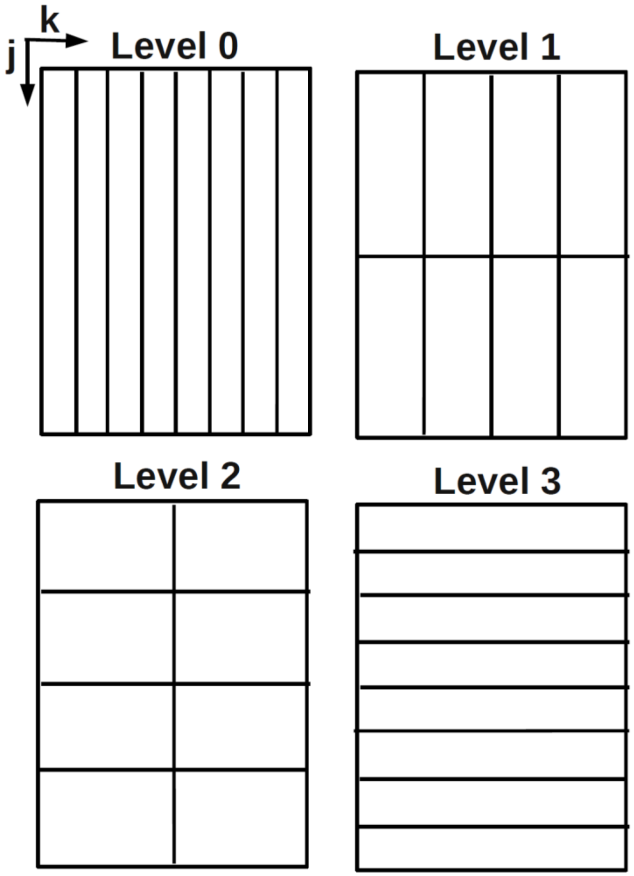
ECMWF



ECMWF

Butterfly algorithm: apply $f = S\alpha$

```
for l = 0 → L do
    for all j, k boxes do
        if l = 0 then
            store β_{0,k} = A_{0,k} α_k
        else
            store β_{l,j,k}
              = A_{l,j,k} × comb_l_and_r_neighb(β, l − 1)
        end if
        if l = L then
            store f_{L,j} = C_{L,j} β_{L,j}
        end if
    end for
end for
```



ECMWF

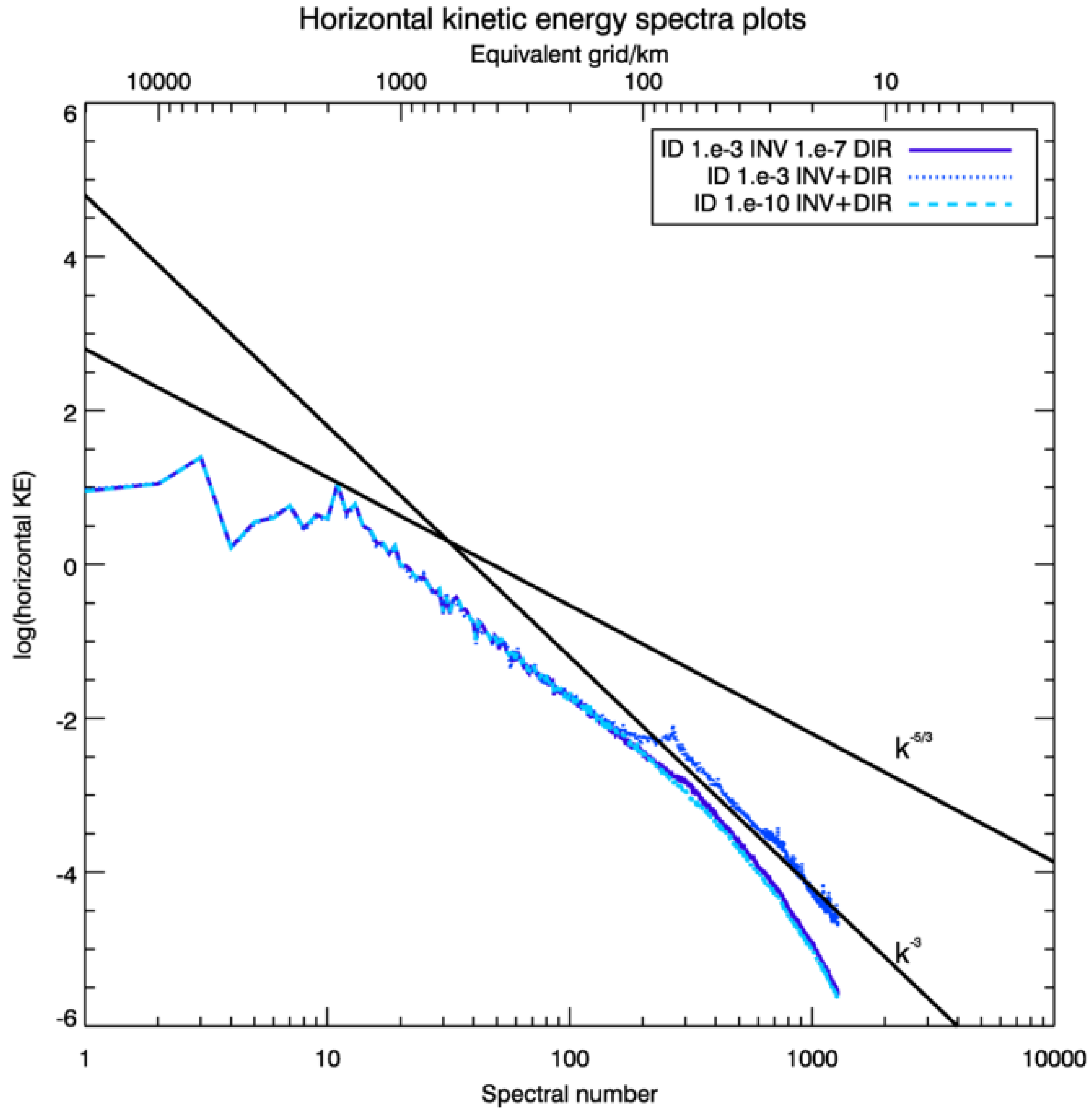
Interpolative Decomposition (ID)

- 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)

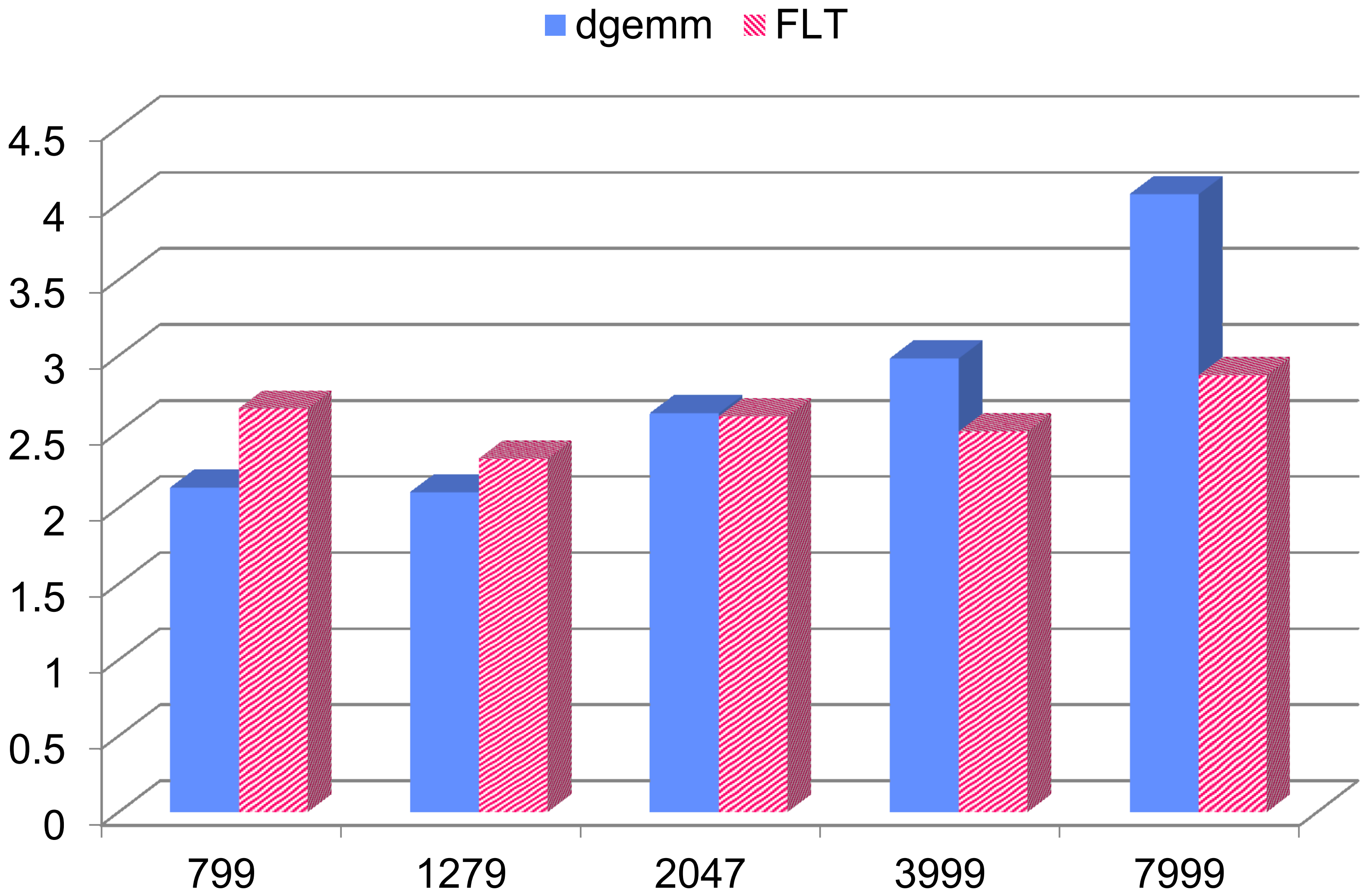


ECMWF

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 $N^2 log^3 N$

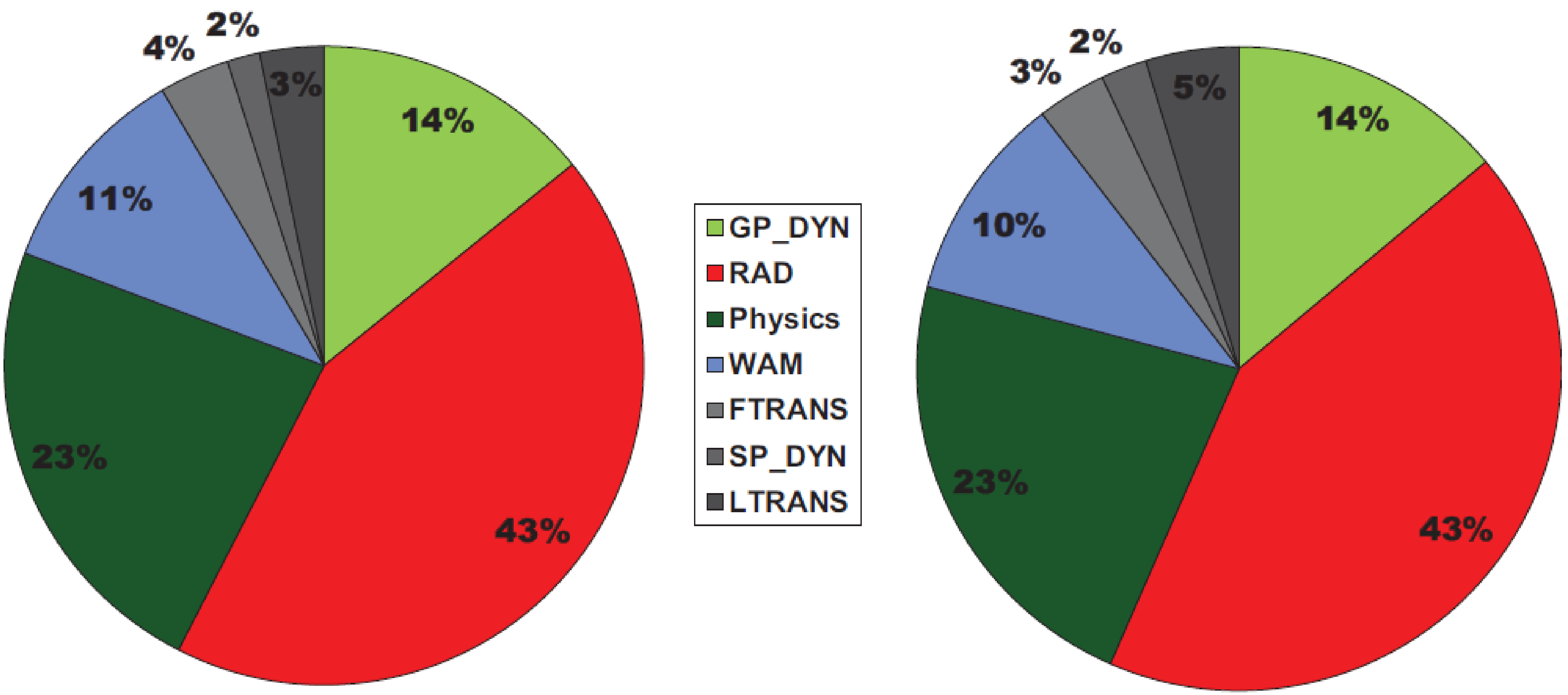


ECMWF

The quadratic or cubic grid *(Wedi, 2014)*

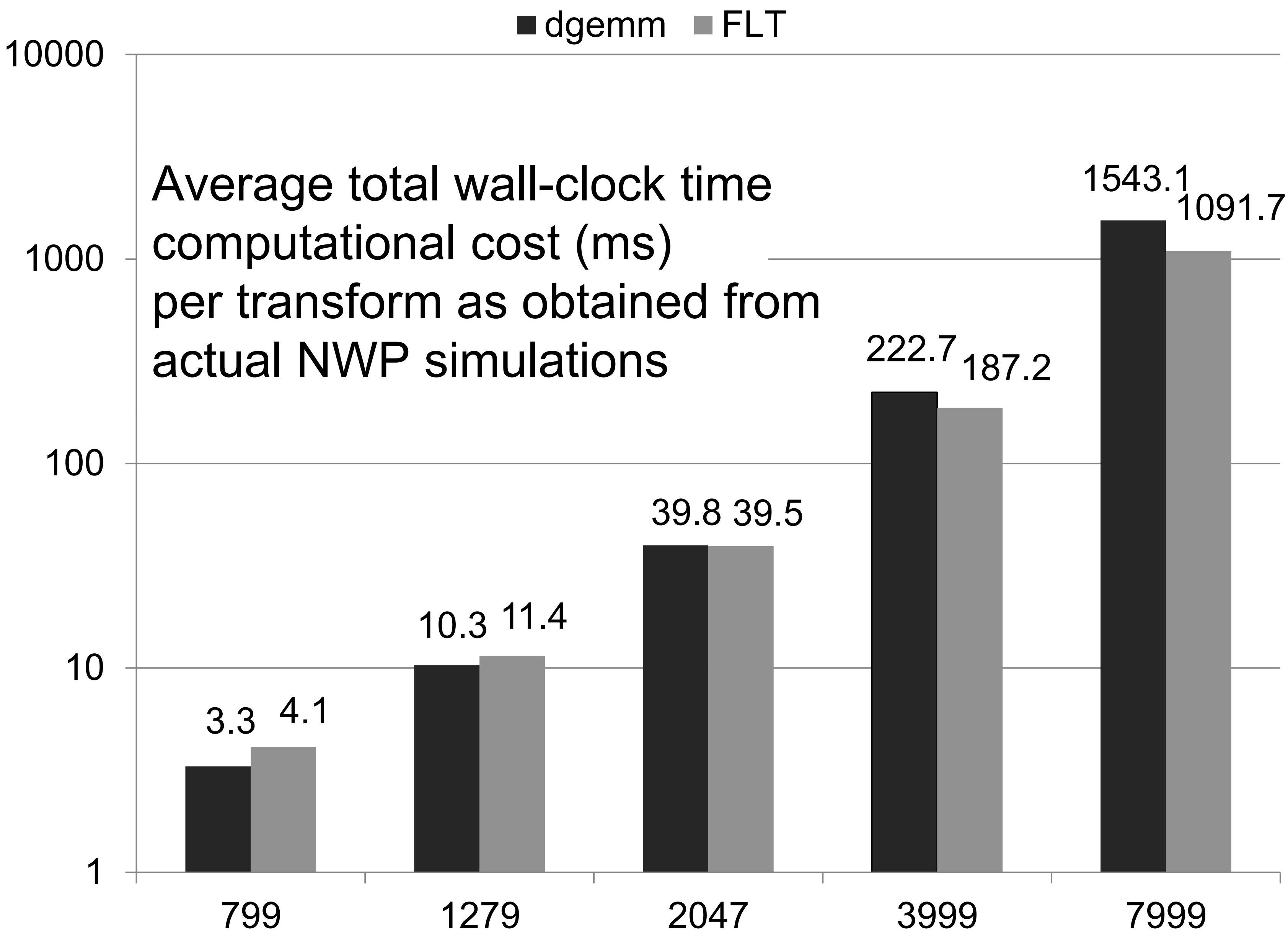
- Adjustment of formal accuracy/relative resolution in spectral and physical space.
- All nonlinear rhs forcings, advection, moist quantities, physical forcings and surface processes are computed on the higher resolution grid. All horizontal derivatives (T,vor/div, u/v,ln p) and the spectral computations are "filtered" to the cubic truncation wavenumber.

The computational cost distribution with full radiation and high-res wave model



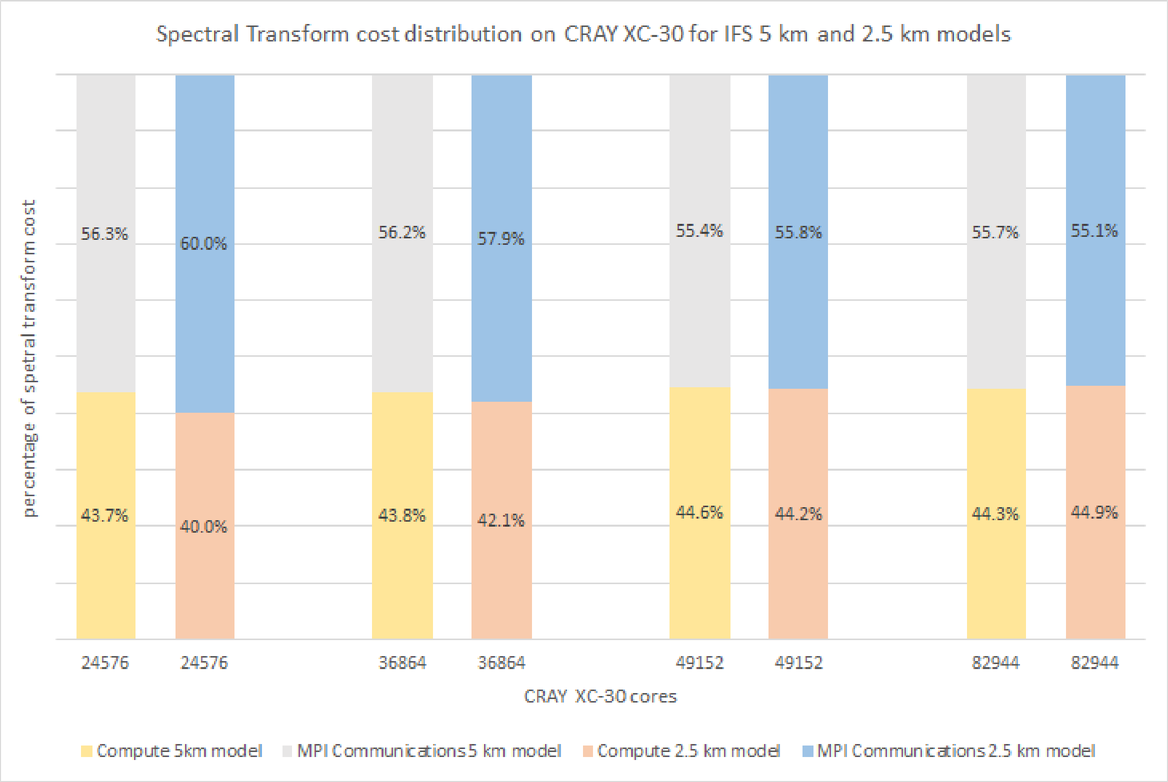
(T_q1364) (T_l2047)

ECMWF


dgemm
FLT
10000
1000
100
10
1
Average total wall-clock time
computational cost (ms)
per transform as obtained from
actual NWP simulations
3.3
4.1
10.3
11.4
39.8
39.5
222.7
187.2
1543.1
1091.7
799
1279
2047
3999
7999

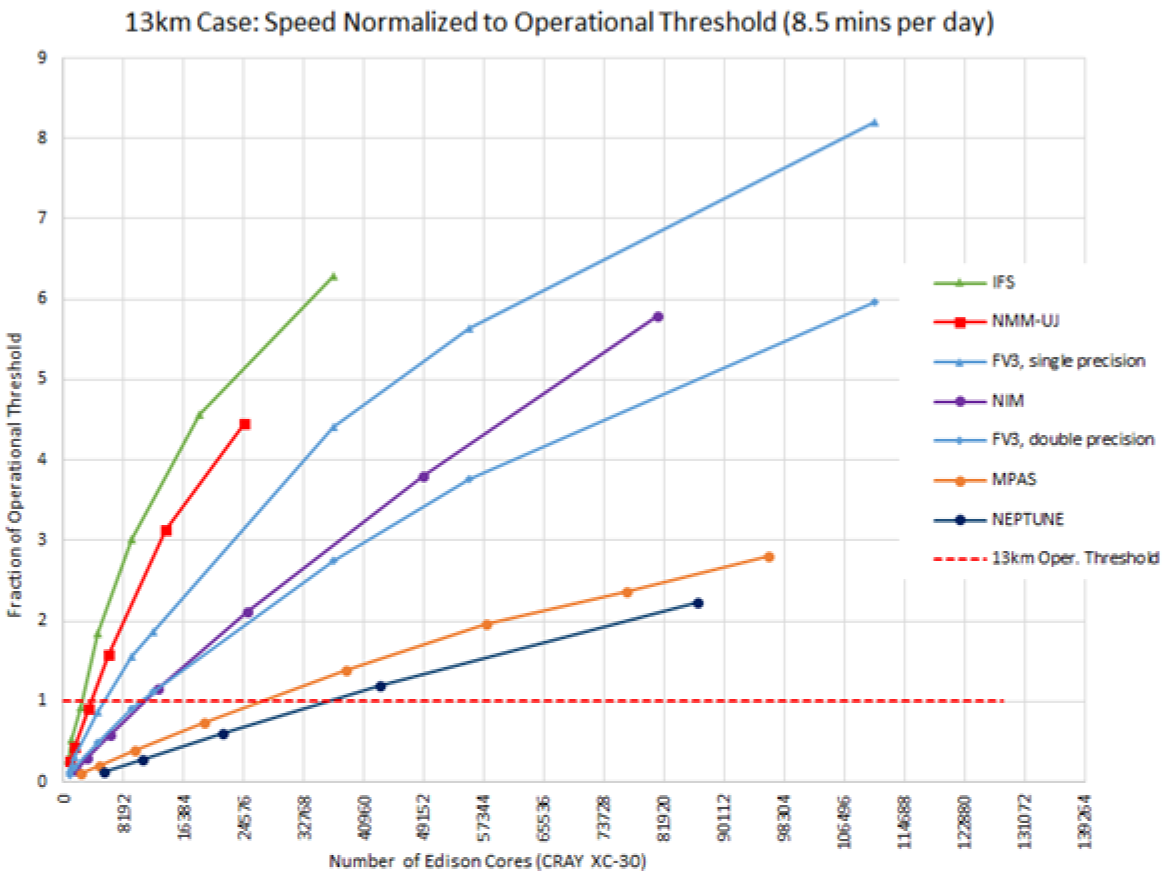
ECMWF

Cost of communication as percentage of spectral transform cost



SAC Report 2015

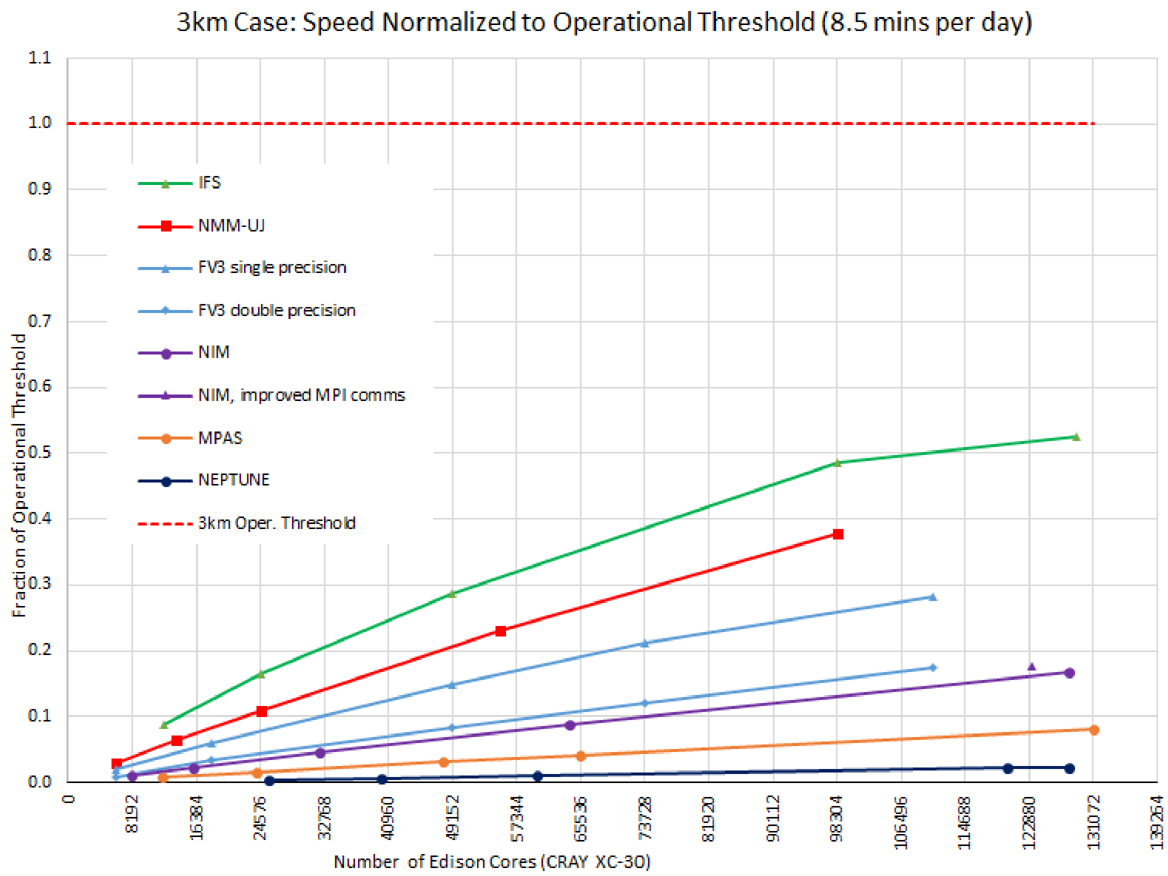
AVEC model intercomparison report 13km



SAC Report 2015

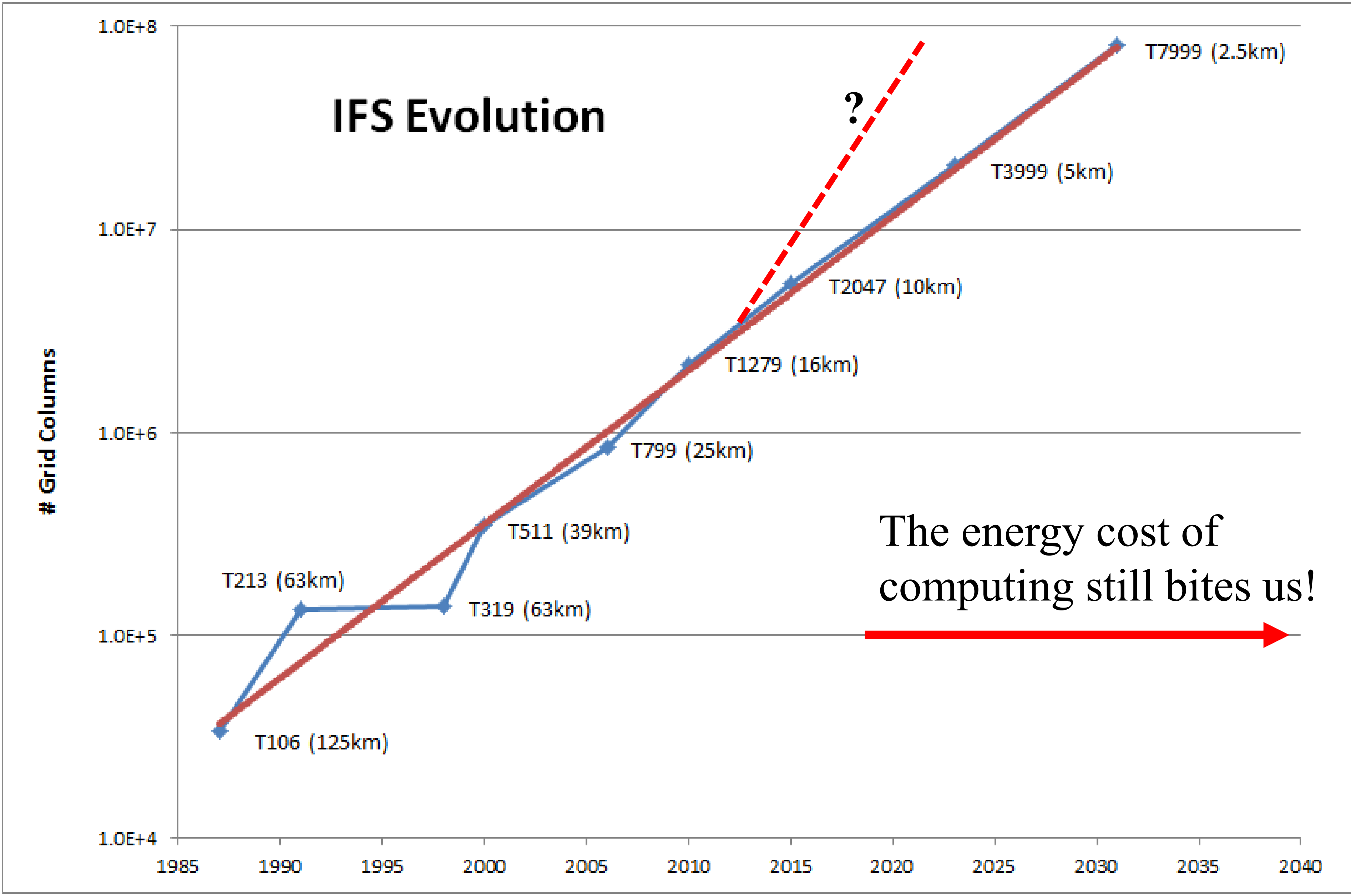
ECMWF

AVEC model intercomparison report 3km



SAC Report 2015

ECMWF


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
T106 (125km)
T213 (63km)
T319 (63km)
T511 (39km)
T799 (25km)
T1279 (16km)
T2047 (10km)
T3999 (5km)
T7999 (2.5km)
?
The energy cost of computing still bites us!

Additional slides