ECMWFDATA V2.0 update
Final report of Consultant

Univ. Ass. Dr. Leopold Haimberger

Institut für Meteorologie und Geophysik
Universität Wien
Althanstrasse 14, 1090 Wien

Wien, 29th January 2007

1 Introduction

The ECMWFDATA software prepares input for the trajectory models FLEXTRA, FLEXPART and HYSPLIT from ECMWF model output. Most fields can be retrieved and interpolated with the ECMWF MARS software. Only the vertical velocity $\dot{\eta}(\partial p/\partial\eta)$ in the hybrid coordinate system used by ECMWF (Simmons and Burridge 1981) must be calculated in an extra step from the archived wind and pressure fields.

This protocol documents the rationale for changes in codes and scripts of the TOR ECMWF-DATA software package between versions 1.0 and 2.0.

2 The continuity equation and its discretization in the ECMWF model

In order to understand the subtleties of calculating vertical velocities from ECMWF data, the continuity equation used for ECMWF postprocessing needs to be carefully examined.

The continuity equation in a pressure based hybrid coordinate system as used at ECMWF reads as follows:

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial \eta}\right) + \nabla\cdot\left(\mathbf{v}\frac{\partial p}{\partial \eta}\right) + \frac{\partial}{\partial \eta}\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right) = 0. \tag{2.1}$$

since $\dot{\eta}$ vanishes at $\eta = 0$ and $\eta = 1$, one can use (2.1) for calculating the surface pressure tendency:

$$\frac{\partial p_s}{\partial t} = -\int_0^1 \nabla\cdot(\mathbf{v}\frac{\partial p}{\partial \eta})d\eta, \tag{2.2}$$

In discretized form this yields (Simmons and Burridge 1981):

$$\frac{\partial p_s}{\partial t} = -\sum_{k=1}^{NLEV}\left[\Delta p_k \nabla\cdot\mathbf{v}_k + (\mathbf{v}_k\cdot\nabla p_s)\Delta B_k\right] \tag{2.3}$$

$NLEV$ is the number of model levels, $\Delta p_k = p_{k+\frac{1}{2}} - p_{k-\frac{1}{2}}$, $\Delta B_k = B_{k+\frac{1}{2}} - B_{k-\frac{1}{2}}$. k is the model level index die B_k and A_k are parameters that specify the hybrid vertical coordinate.

One can use the continuity equation to calculate the vertical velocity in hybrid coordinates as follows:

$$\dot{\eta}\frac{\partial p}{\partial \eta} = -\frac{\partial p}{\partial t} - \int_0^{\eta} \nabla\cdot\left(\mathbf{v}\frac{\partial p}{\partial \eta}\right) d\eta, \tag{2.4}$$

In discretized form, one gets a formula for $(\dot{\eta}\partial p/\partial\eta)_{k+\frac{1}{2}}$ on "half levels":

$$\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+\frac{1}{2}} = -B_{k+\frac{1}{2}}\frac{\partial p_s}{\partial t} + \sum_{j=1}^{k}\left[\Delta p_j \nabla\cdot\mathbf{v}_j + (\mathbf{v}_j\cdot\nabla p_s)\Delta B_j\right], \tag{2.5}$$

To calculate $\dot{\eta}$ one needs u, v and the surface pressure gradient. These quantities can be calculated exactly in spectral space from vorticity and divergence and from the surface pressure. In the forecast model a diagnostic equation is solved during initialization that ensures that the continuity equation is fulfilled and that the surface pressure tendency at each time step yields a realistic

Table 2.1: Steps performed during exact (V2.0 on full-resolution Gaussian grid), truncated (V2.0 on lower-resolution Gaussian grid) and fast (V2.0 or V1.0 on lat-lon grid) calculation of $\dot{\eta}$ and ω.

	Spectral space	Gaussian grid	Spectral space	Lat-Lon grid
Exact:	Full-resolution $\vec{v}, \nabla \cdot \vec{v}$, $\log p_s$ from MARS. Calculate $\nabla \cdot \log p_s$	Calculate $\dot{\eta}(\partial p/\partial\eta)$ and ω using (2.5) and (2.6)	Truncate $\dot{\eta}(\partial p/\partial\eta), \omega$, $\vec{v}$ to spectral resolution consistent with output grid resolution. This step is referred to as "smmothing" in the text.	Write $\dot{\eta}(\partial p/\partial\eta), \omega$ and input fields on disk
Truncated:	As exact, but with truncated $\vec{v}, \nabla \cdot \vec{v}$, $\log p_s$ from MARS.	As exact	As exact	As exact
Fast:	$\vec{v}$, $\nabla \cdot \vec{v}$, $\log p_s$ from MARS. Resolution must be consistent with output grid resolution. Calculate $\nabla \cdot \log p_s$	-	-	Calculate $\dot{\eta}(\partial p/\partial\eta)$ and ω using (2.5) and (2.6), Write $\dot{\eta}(\partial p/\partial\eta)$, ω and input fields on disk

value. This important balance is destroyed easily by any form of interpolation. It is therefore best to calculate the vertical velocity with the original discretization scheme and with the same resolution as the forecast model.

The exact calculation is essential for the accuracy of the resulting vertical velocity. Trying to discretize $\nabla \cdot \mathbf{v}$ or ∇p_s in (2.3) und (2.5) with a finite difference scheme in gridpoint space leads to intolerably large errors.

In Version 1.0 the divergence and the surface pressure gradients have been calculated exactly in spectral space but then have been transformed not to the associated Gaussian grid but to a regular lat-lon grid. This leads to inaccuracies since the 3D wind field is balanced only on the Gaussian grid. The inaccuracy (10-15%) has been found to be tolerable.

However, in recent years the spectral fields divergence and the pressure gradient have been truncated before transformation in gridpoint space for efficiency reasons. This can potentially destroy the mass consistency, which is only guaranteed at full vertical resolution. If equation (2.3) were linear, there would be no problem with the truncation. There is, however, the nonlinear term $(\mathbf{v}_k \cdot \nabla p_s)$, which may contain scale interaction, especially near the surface, that is not resolved by the truncated fields but affects lower wavenumbers. A quantification of this effect is therefore desirable.

In order to estimate the inconsistencies, $\dot{\eta}(\partial p/\partial\eta)$-fields and fields of the pressure vertical velocity ω are compared below for various resolutions and truncations. ω is an optional input field for HYSPLIT and (in contrast to $\dot{\eta}(\partial p/\partial\eta)$), it is saved in MARS. ω is calculated from the continuity equation as follows:

$$\omega = \int_{0}^{\eta} \nabla \cdot \left(\mathbf{v}_h \frac{\partial p}{\partial \eta} \right) d\eta + \mathbf{v}_h \cdot \nabla p \tag{2.6}$$

The discretized form of this equation is given in Simmons and Burridge (1981). The same quantities are involved as for the $\dot{\eta}(\partial p/\partial \eta)$. The calculation of ω is even more dependent on the accurate representation of the nonlinear term $\vec{v} \cdot \nabla p$, since it is not weighted with ΔB_k, as in (2.5) but is weighted fully. Therefore ω from MARS can be used as reference for the accuracy of the calculations below.

3 Problems with V1.0

The Version 1.0, implemented in 2003, was stable and reliable using the operational data available at that time. Since then, the ECMWF Integrated Forecast System (IFS) has undergone significant changes. In particular the horizontal and vertical resolution has been enhanced a lot. This resulted in two problems:

- Version 1.0 failed when the highest available spectral resolution (T799) was desired. Input fields had to be truncated to lower resolution (T319). The truncation and the calculation of the vertical velocities on a lat-lon grid (not on a Gaussian grid) may cause large errors in $\dot{\eta}(\partial p/\partial \eta)$.
- A bug was detected in the calculation of the pressure gradients.
- The high spectral resolution in spectral space makes sense only with high resolution in gridpoint space. This makes it desirable to specify sub-areas for data retrieval in order to reduce the transferred data amount. Version 1.0 works only with global fields, however.

4 Update of spectral transformation routines

The spectral transformation routines have been updated and extended to accommodate the new requirements.

- The whole calculations are now double precision (real*8) since this is necessary for the long summations. As a consequence, the double precision version of the `EMOSLIB` (`-lemos.R64.D64.I32`) needs to be called. The conversion software uses the `PBIO` routines, `GRIBEX`, and the fast fourier transforms `SET99, FFT99` from `EMOSLIB`
- The ordering of the spectral coefficients used in the conversion software is now equal to the ordering used by ECMWF. This results in a speedup of the transformation routines by a factor of at least 3.
- The calculation of the Legendre polynomials has been updated and is numerically stable for spectral truncations beyond 1000. The smallest gridsize handled by the conversion routines is 0.01 degrees (36000 points around a latitide circle).

- While the results from the spectral transformation results are not bit-identical to the equivalent conversions in MARS, they agree with a relative accuracy of 10^{-8}.
- Several routines that perform the calculations on the Gaussian grid have been included.
- The most demanding routines (calculation of Legendre Polynomials, spectral transformations) have been parallelized with OpenMP. Up to 32 threads allow substantial speedups, especially for calculations on the Gaussian grid. Parallel jobs are currently supported only on the HPC facility at ECMWF, not on ecgate.

5 New output fields

The conversion software optionally calculates three new fields:

- ω from MARS. It is needed as input for HYSPLIT. This is not really a new field since it can be retrieved without the conversion software as well. It has parameter number `137` in the output file (fort.15), which is different from the usual number `135` for ω. The reason for this is that 135 is already used by the conversion software for $\dot{\eta}(\partial p/\partial\eta)$.
- ω from MARS minus ω calculated by the TOR conversion software. This difference is valuable for assessing the error made by the calculation of $\dot{\eta}(\partial p/\partial\eta)$. It has parameter number `138`
- $\partial p_s/\partial t$ from integration of the continuity equation. This surface pressure tendency should vanish in the long term mean (monthly or seasonal). The field is also intended mainly for quality control. Parameter number is `136`.

6 New input datasets

The operational high resolution analyses and forecasts put high demands on computer resources and provide much detail that often cannot be exploited by the trajectory models using the output of ECMWFDATA. Therefore it is desirable to truncate the spectral input data before calculating the vertical velocities. Some users may also be satisfied with lower vertical resolution of the input fields. Therefore, the general extraction script allows to use not only the operational high resolution analyses but also lower resolution fields.

In addition to the operational analyses and forecasts, the following ECMWF products can be used as input for the trajectory models:

- The ECMWF VarEPS high resolution control/validation forecasts (T399L62), which are available at every 12 hours in steps of three hours. Three-hourly resolution input for the trajectory models can be realized by using these forecasts at steps 00h, 03h, 06h, 09h. In contrast to the high resolution model, there are no analyses at 06h and 18h.
- The ECMWF VarEPS low resolution control/validation forecasts (T255L62). These are produced from a truncated and initialized T399L62 reanalysis.
- The ERA-Interim reanalyses (T255L60 or T255L91) provides climate reanalyses from 1989 onwards. These analyses are currently produced and should be available for the whole period 1989-present in a few months.

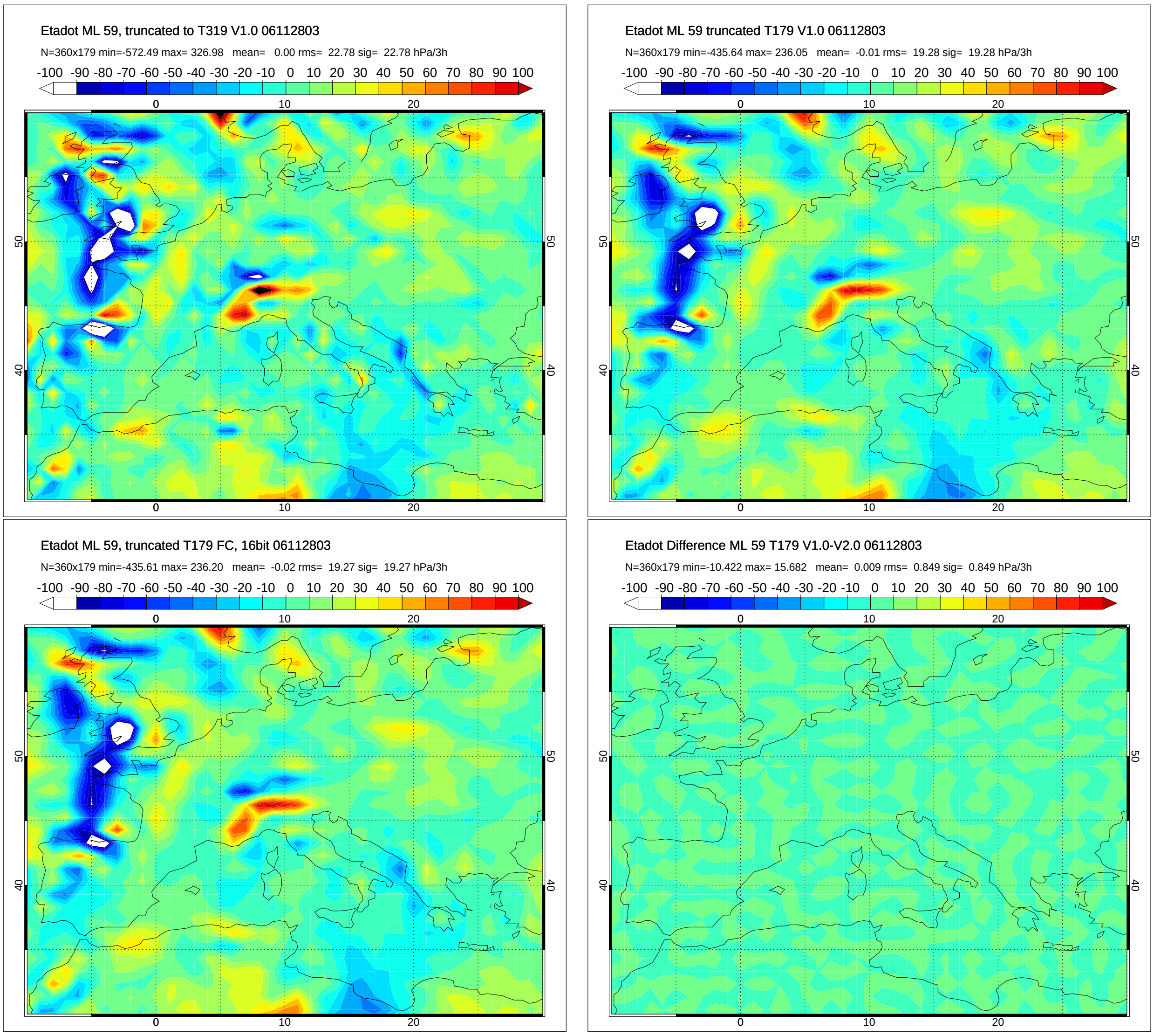


Figure 6.1: $\dot{\eta}(\partial p/\partial\eta)$ on 28 November 2006, 03GMT, a) from T319 V1.0, b) from T179 V1.0, c) from T179 V2.0 without Gaussian grid, d) Difference between b) and c). Difference between b) and c) is small but not zero because of a bug in the pressure gradient calculation in version 1.0.

- The operational T799L91 products and ERA-40 products can be used as in version 1.0. Note that the operational products can be extracted in a sensible way on ecgate only with truncations up to T399L91 and output grid sizes up to 720×360. Larger requests will fail there due to excessive memory requirements. They must be executed on the HPC facility.

7 Results

The main aim of the software update is to ensure optimal accuracy of the vertical velocities for different spectral resolutions and increased flexibility concerning spectral resolution and output grid spacing. This section shows how truncation, interpolation and smoothing affect the resulting vertical velocities.

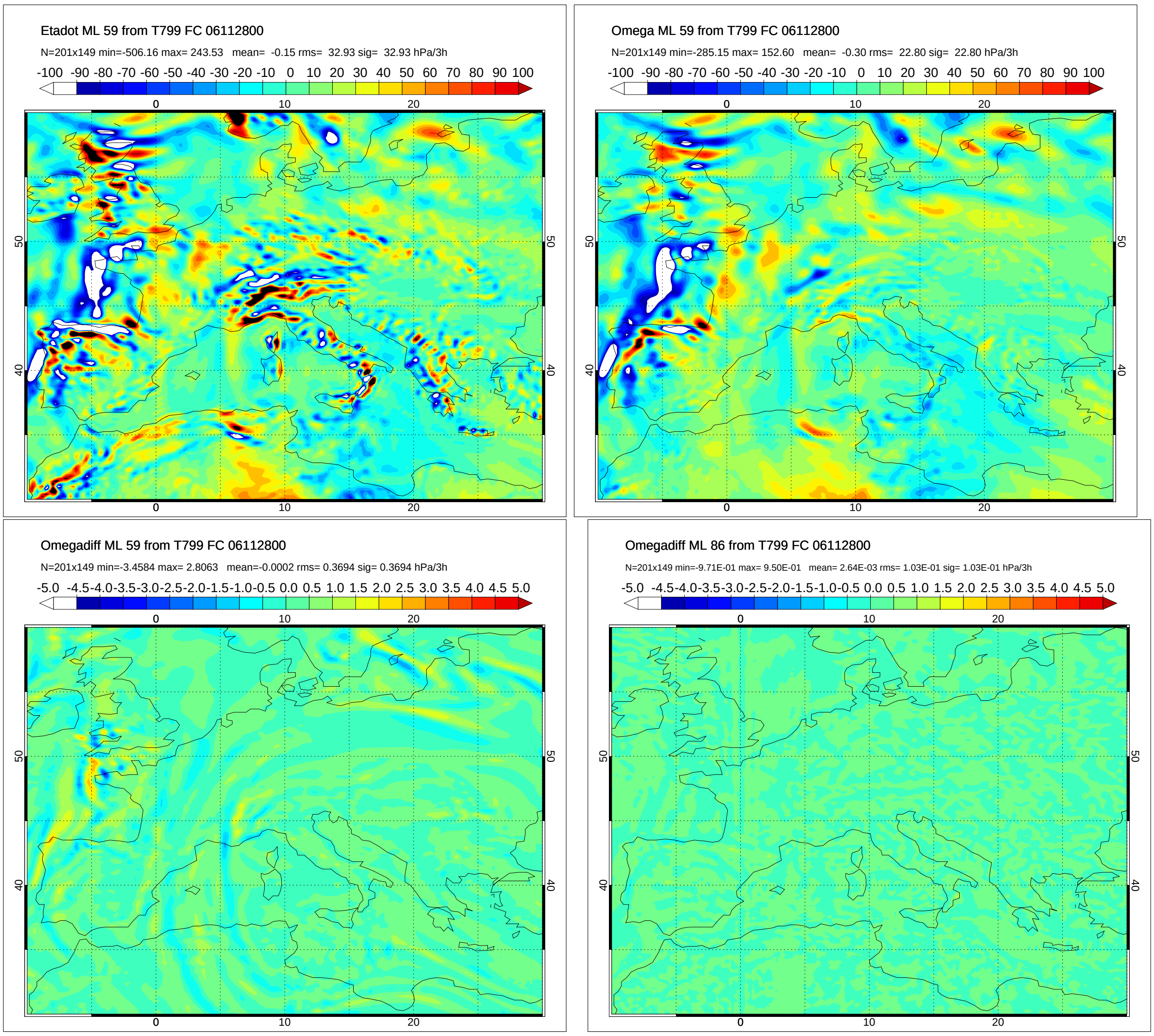


Figure 7.2: $\dot{\eta}(\partial p/\partial\eta)$, ω and ω(MARS)-ω(ECMWFDATA) calculated by TOR V2.0 software on the Gaussian grid at full resolution T799 and then transformed to a 0.2/0.2 lat/lon grid. Fields plotted at model level 59 (ca. 400 hPa) ω can reach more than 3 Pa/s (300 hPa/3 h). Panel d) shows fi eld on model level 86 (ca. 950 hPa). Note different scale for ω-difference: ω-difference is more than two magnitudes smaller than $\dot{\eta}(\partial p/\partial\eta)$ or ω

7.1 Comparison V1.0 vs. V2.0

To see the impact of migrating from V1.0 to V2.0 on a global 1.0/1.0 degree grid, Fig. 1 compares $\dot{\eta}(\partial p/\partial\eta$ over Europe. Comparison of a) and b) shows the effect of lowering the spectral resolution to T179, which is recommended for a 1.0/1.0 grid, since a 1.0/1.0 grid can only resolve 179 wavenumbers. Panel c) shows the same quantity from ECMWFDATA v2.0, using the old (Fast in table 1) computation method without the Gaussian grid. c) shows the difference between b) and c). It is nonzero (ca. 5% of signal) because of a bug in the pressure gradient calculation in v1.0 which is fixed in 2.0.

7.2 Validation strategy

In order to assess the accuracy of the $\dot{\eta}(\partial p/\partial\eta)$ calculation, the ω-field, which can be extracted routinely from MARS, has been calculated as an additional output field. The calculation of ω and $\dot{\eta}(\partial\eta/\partial p)$ is similar, although ω is even more prone to numerical errors. The ω-field on a regular lat/lon grid can be gained by three different methods:

- Retrieve from MARS (reference)
- Retrieve $u, v, \nabla \cdot v$ on lat/lon grid and $\log p_s$ as spherical harmonics coefficients. Calculate $\nabla \cdot \log p_s$ and interpolate to lat/lon grid. Then integrate continuity equation on lat/lon grid. This is the method used in Version 1.0 of ECMWFDATA and is also an option in ECMWFDATA v2.0. Inaccuracies may occur since the continuity equation is fulfilled exactly only by data on the Gaussian grid, not on the Lat/Lon grid
- Retrieve $\nabla \times v, \nabla \cdot v$ and $\log p_s$ as spherical harmonics coefficients. Calculate $\nabla \cdot \log p_s$ and interpolate to reduced Gaussian grid. Then integrate continuity equation on Gaussian grid. Transform $\dot{\eta}(\partial\eta/\partial p)$ and ω into spectral space and then back to lat/lon grid. If the archived spectral truncation is used, this calculation is practically identical to the practice at ECMWF. This should yield vertical velocity fields which are consistent with the horizontal wind field. Calculating the vertical velocity on the Gaussian grid should have the following additional advantages:
  - The spectral fields $(u, v, \nabla \cdot \vec{v}, \log(p_s))$ may be truncated and the etadot calculation may be performed on a coarser than operational reduced Gaussian grid. The errors due to this procedure are relatively small. While the truncated fields do not fulfil the discretized continuity equation with the same accuracy as operational forecasts, the inconsistencies are relatively small. In particular they are smaller than the inconsistencies observed when calculating etadot from analyses. For details see section on results.
  - The calculated $\dot{\eta}(\partial p/\partial\eta)$ field may be truncated or smoothed in spectral space before transformation on the output grid.
  - Valid resolutions for which a Gaussian grid exists are: T159, T255, T319, T399, T511, T639, T799, T1023. The spectral resolution should be chosen as 180/Δx or lower, where Δx is the grid spacing. Choosing a spectral resolution higher than that value violates the Nyquist criterion, especially on global grids. For a 1° output grid, either a spectral resolution T159 should be chosen or a spectral resolution 255 for the internal calculation with truncation to T179 just before transformation to the output grid (see parameter `M_SMOOTH` below).

7.3 Vertical velocities on 20061128

7.3.1 Comparison of ω with ω from MARS

Figure 7.2-a),b) shows $\dot{\eta}(\partial p/\partial\eta)$ an ω-field on the 59th model level on 20061128 00GMT over parts of Europe, as calculated by the ECMWFDATA software on the Gaussian grid and then transformed to a 0.2/0.2 lat/lon grid. This high resolution if required for resolvont T799 information. $\dot{\eta}(\partial p/\partial\eta)$ is more structured over orography at this level but otherwise, the two fields

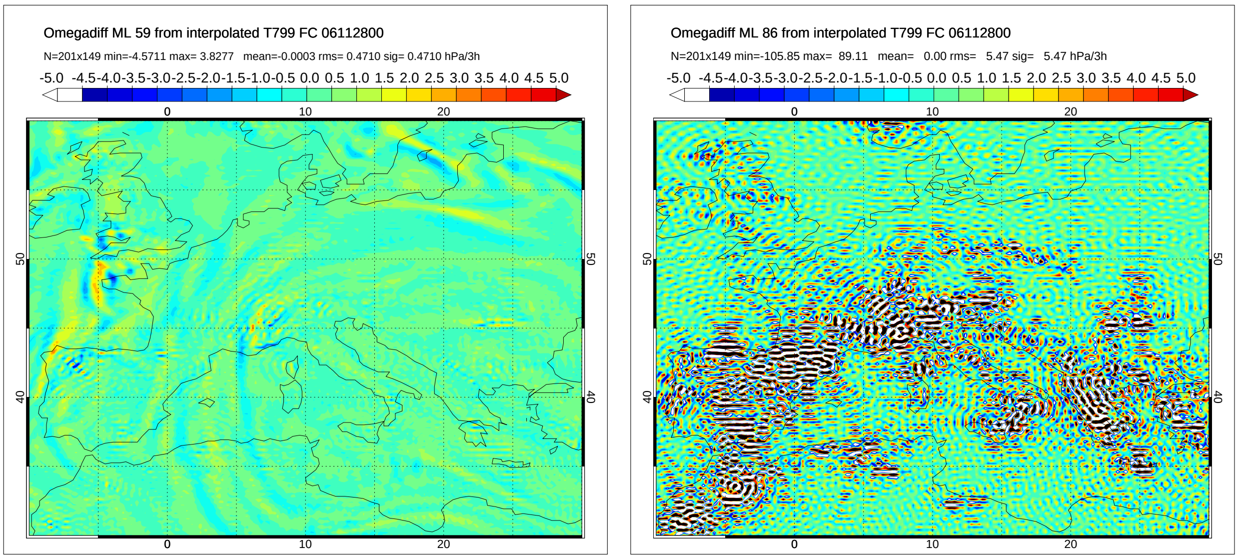


Figure 7.3: Difference between ω from MARS and ω calculated with fast method with input data on 0.2/0.2 lat/lon grid. Compare with panels c),d) of Fig. 7.2. Inconsistency of ω-fi elds is much larger near surface.

are fairly similar. Velocities are on the order of 100 hPa/3h, especially near mountains and fronts. Figure 7.2-c) shows the difference between ω calculated by ECMWFDATA and ω from MARS. These two should be equal and thus the difference should be near zero. It is found to be more than two orders of magnitude smaller than the actual vertical velocity fields. Thus it is negligilble compared to the uncertainties in ω or $\dot{\eta}(\partial p/\partial\eta)$.

This figure demonstrates that the new ECMWFDATA version is capable of processing T799L91 fields with high accuracy. It also demonstrates that a subregion can be extracted, since the the specified output grid is from 30N/10W to 60N/30W with 0.2° grid spacing.

Figure 7.3 shows the accuracy loss incurred if ω is calculated not on the Gaussian grid but on a high resolution lat/lon grid, as it is standard in the old ECMWFDATA version 1.0.

Since such high resolution calculations are often not feasible or not sensible for further processing, it is desirable to find an optimal calculation of the vertical velocity fields at lower spectral resolutions.

Fig. 7.4 shows vertical velocity fields gained from T799 input data truncated to T399 and calculated on a Gaussian grid suitable for this spectral resolution. Due the truncation the nonlinear terms $\vec{v}\cdot\nabla p$ is not correctly represented, which leads to differences between truncated ω from MARS and ω from truncated input data. The *rms* differences are on the order 10% for most levels but reach 30% at the lowest levels. The problem is even worse when ω is calculated from truncated data on a lat/lon grid (Fig. 799399L).

A way to resolve this is to use T399 or T255 control/validation forecasts from the EPS forecasting system. Fig. 7.6 shows vertical velocity fields gained from T399 EPS control forecast data and calculated on a Gaussian grid suitable for this spectral resolution. These data should be internally consistent since they have been treated by a T399 model and no truncation was necessary. The consistency of the ω-fields is indeed better than for the operational T799 input data truncated to T399. Note that this is model level 30 of the T399L62 EPS forecast system corresponds to level 59 of the T799L91 operational system.

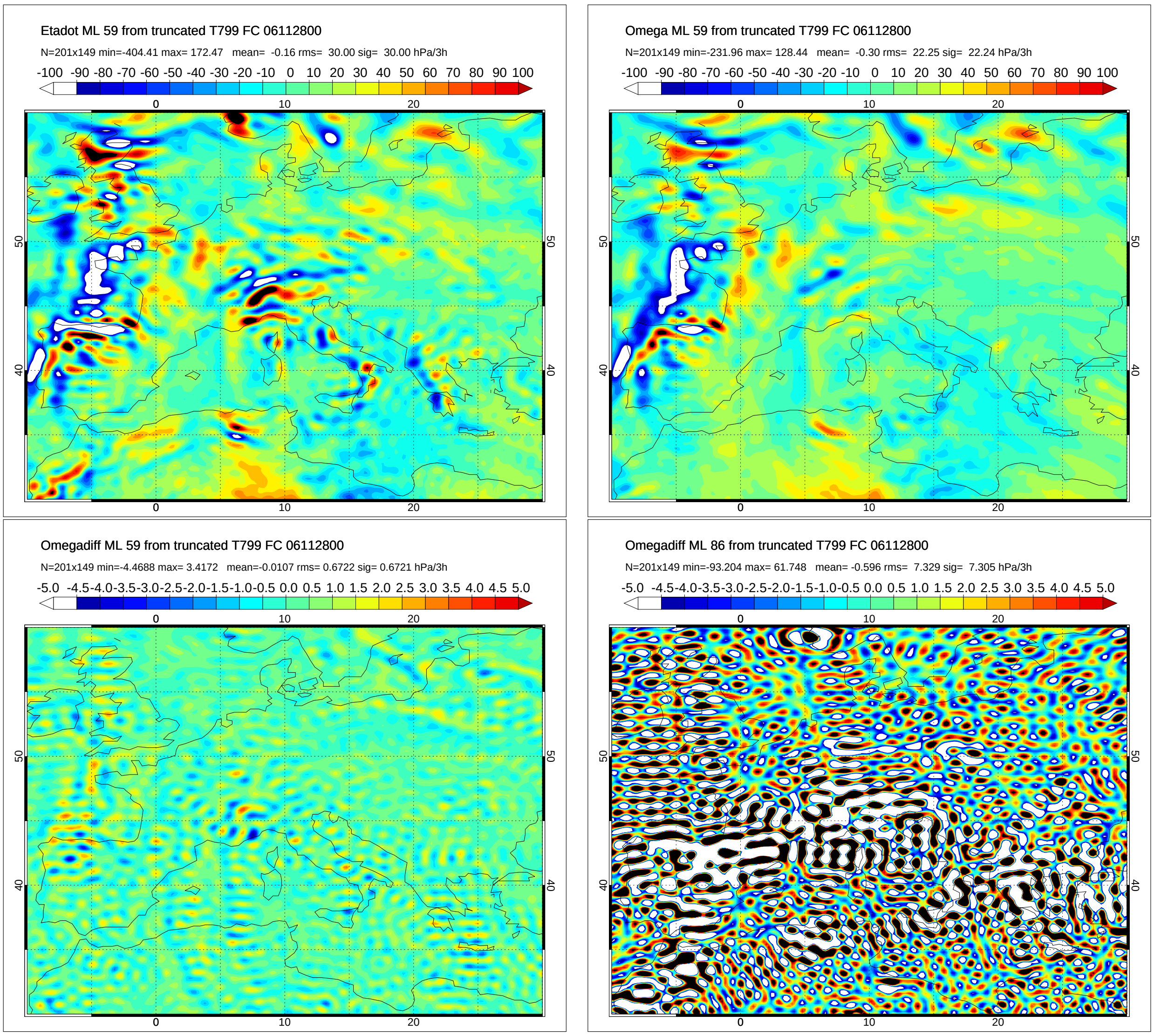


Figure 7.4: As Fig. 7.2, but calculated with input data truncated to T399 on the Gauusian grid. Inconsistency of ω-fi elds is large due to nonlinearity effects.

7.3.2 Comparison of $\dot{\eta}(\partial p/\partial \eta)$ calculated with different methods

Fortunately this dependence on the nonlinear term is much weaker for $\dot{\eta}(\partial p/\partial \eta)$. This is shown in Fig. 7.7 for T399 and in Figs 7.8 and 7.9 for T179. In panels 7.7-a) - 7.9-a) $\dot{\eta}(\partial p/\partial \eta)$ has been calculated from T799 fields on the Gaussian grid and then truncated to T399/179. This should give optimal results.

In panels 7.7-b) -7.9-b) $\dot{\eta}(\partial p/\partial \eta)$ has been calculated from T399/T255 data on the Gaussian grid. Panel c) shows difference between a) and b). Panels d) show difference between a) and $\dot{\eta}(\partial p/\partial \eta)$ computed from truncated input data transformed to a lat/lon/grid, as it was standard in V1.0. The *rms*-value of the difference in panel c) is only 1-2% of the signal, whereas the *rms* of the difference in panel d) is 10% of signal. This shows again that it is worthwile to perform the vertical velocity calculations on the Gaussian grid.

All figures together show that the vertical velocity fields depend mainly on the spectral truncation applied. The calculation of $\dot{\eta}(\partial p/\partial \eta)$ is more accurate on the Gaussian grid compared to the lat/lon grid but the advantage is not as dramatic as for the ω-field. Especially near the surface

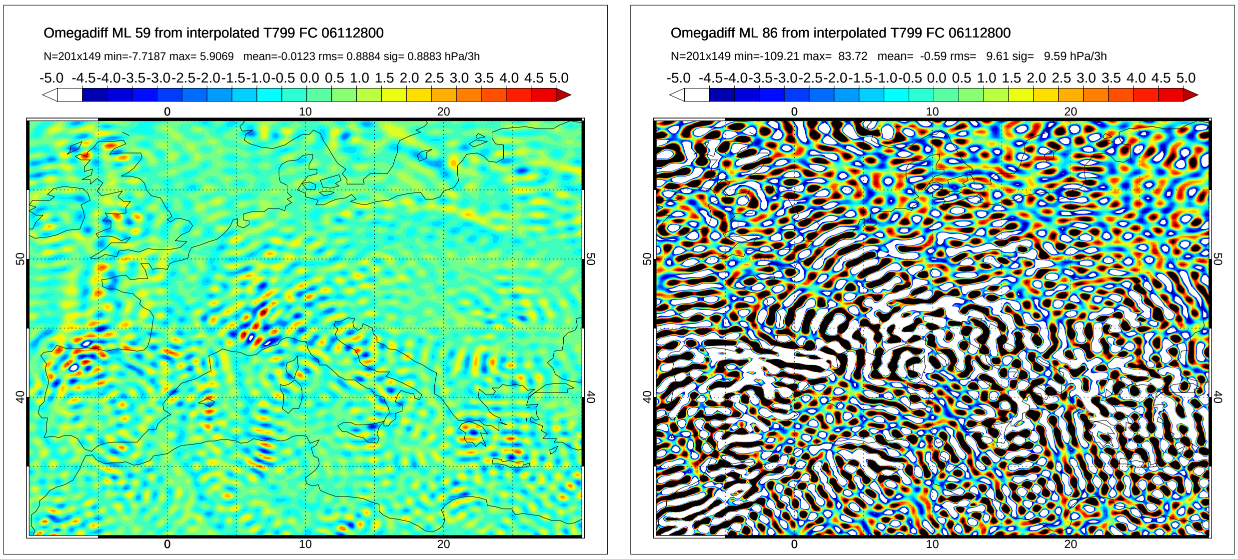


Figure 7.5: Difference between ω and ω from MARS as in Fig. 7.3, but calculated with input data truncated to T399 and then transformed to lat/lon grid. This is very close to the procedure used in ECMWFDATA version 1.0. Inconsistency of ω-fields is rather large, especially near the surface, and larger than in Fig. 7.4

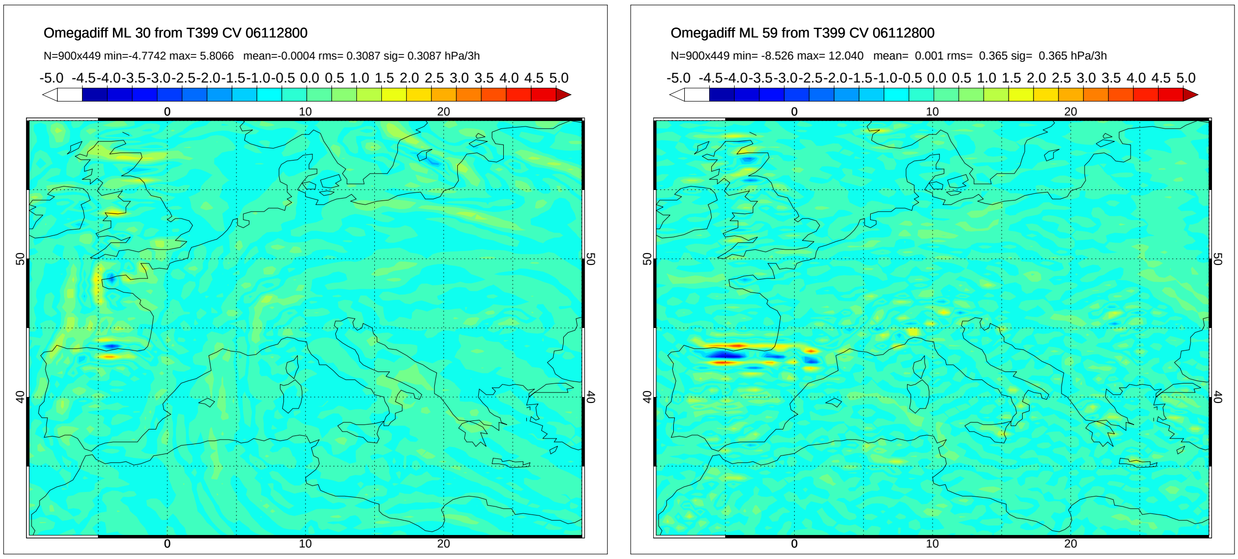


Figure 7.6: As Fig. 7.2-c) and d), but calculated with untruncated input data from the EPS control forecast on the Gausian grid. Inconsistency of ω-fields is very small, compared to T799 operational forecasts truncated to T399.

$\dot{\eta}(\partial p/\partial \eta)$ is much less dependent on numerical details than is ω.

7.4 Computer requirements

The high resolution of the ECMWF model output leads to significant computer demands, especially if calculations on the Gaussian grid are needed. For example, each field ($u, v, \nabla \cdot v$ and $\dot{\eta}$ at least) on the Gaussian grid for the currently operational T799L91 model requires 500MB and the auxiliary arrays for holding the spherical harmonics coefficients and Legendre polynomials need another 2GB). Table 7.2 gives an overview of the computer resources needed per analysis date.

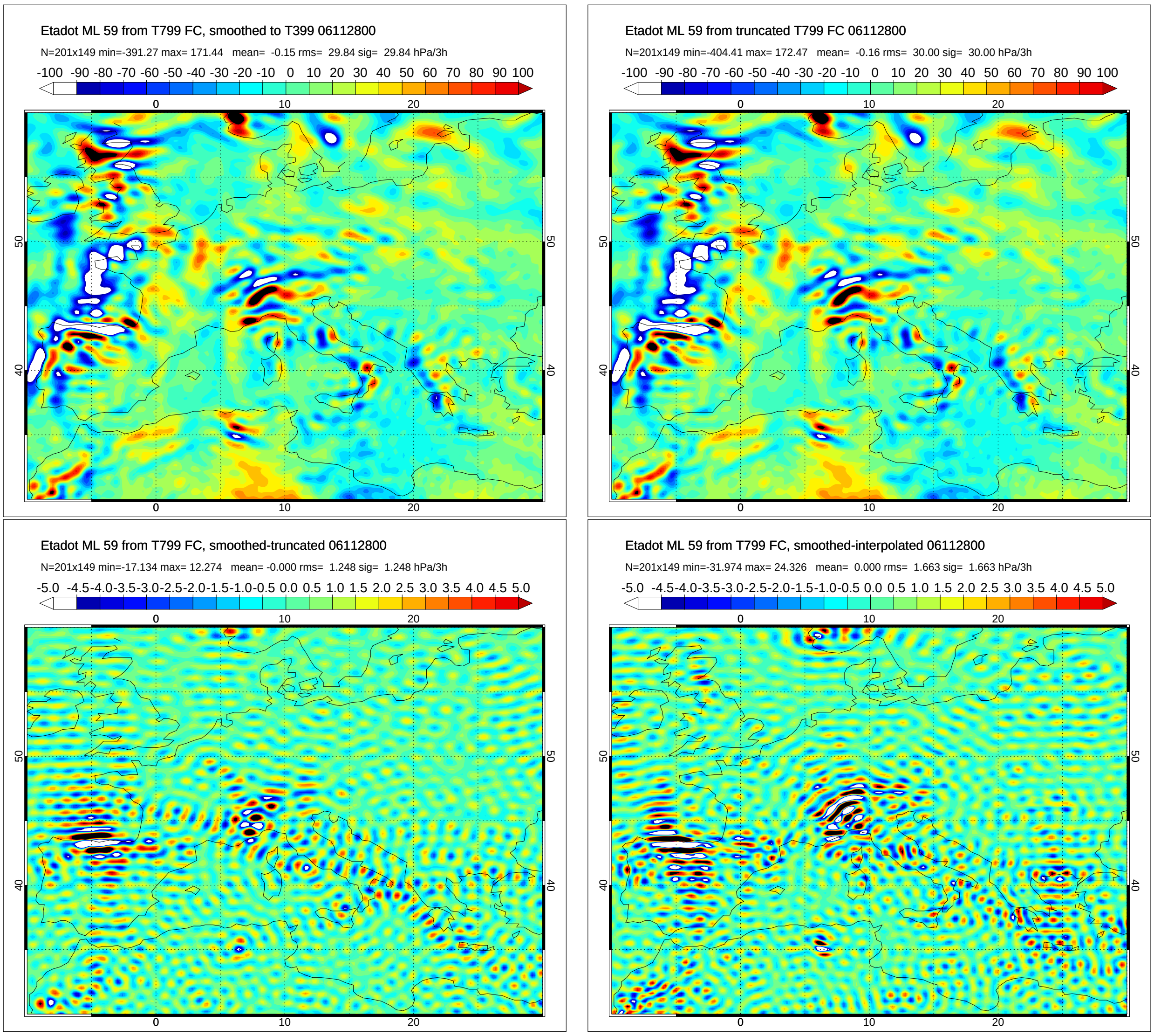


Figure 7.7: Comparison of $\dot{\eta}(\partial p/\partial\eta)$ (compare figures 7.4,7.3,7.6) a) from untruncated operational data, b) from operational data truncated at T399 and calculated on Gaussian grid, c) from operational data, calculated on lat/lon grid, d) from untruncated T399 control forecast. $\dot{\eta}(\partial p/\partial\eta)$ iin panel d) is slightly higher since level 55 or EPS forecasts is slightly higher than level 86 of operational forecast.

Most times are for serial execution, except T799G-32, which has been executed using 32 threads.

When dealing with T799 data, at least 30 elapsed minutes are used for (serial) interpolation within MARS. The actual etadot calculation on the lat/lon grid using T799 takes less than 5 minutes but more than 4 CPU hours on the Gaussian grid. With 32 threads, the elapsed time can be reduced to less than 30 minutes even for the calculation on the Gaussian grid at full resolution, however.

Therefore the T799 calculation on the Gaussian grid is technically possible for operational use, if it is performed as parallel job on HPC.

If HPC access is not available, as is the case for many organizations, one has two choices:

- Transfer the input data to local server and perform the T799 calculation there. This leads to a much higher (4.4 GB) transfer volume to the local server.

- Reduce the spectral resolution to T399 or lower and perform the calculations on ecgate.

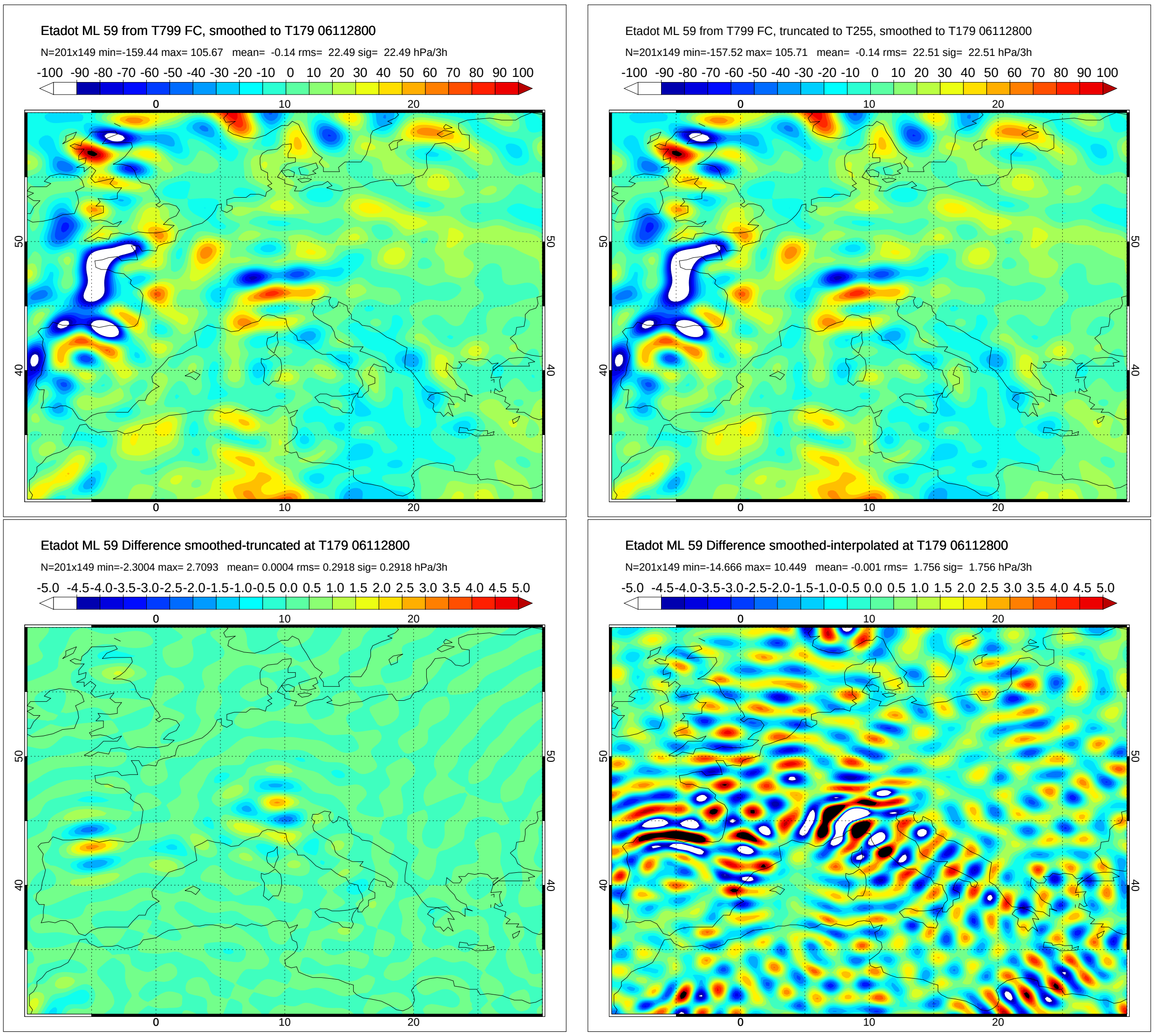


Figure 7.8: As in Fig. 7.7 but for spectral truncation/smoothing T179. Error in panel d) relative to signal in a), b) is 8%.

> For spectral resolutions up to T255 and output grids up to 1.0/1.0, the queue `normal` can be used on ecgate. For higher resolutions up to T399 the queue `long`, which permits the necessary memory requests up to 2GB, must be used on ecgate. The elapsed runtime of T255 calculations for a 1.0/1.0 output grid (4 EN output fields) is around 30 minutes. The elapsed runtime of T399 calculations for a 0.5/0.5 grid is at least 1 hour.

The data transfer volume to CTBTO does change little compared to V1.0 as long as the 1.0/1.0 output grid is kept. It is increased by 30% due if ω and the ω-error are added as additional fields.

8 Conclusions and recommendations

The TOR ECMWFDATA package has been updated in order to cope with high resolution global input data from the operational ECMWF forecast system and in order to assess the impact of

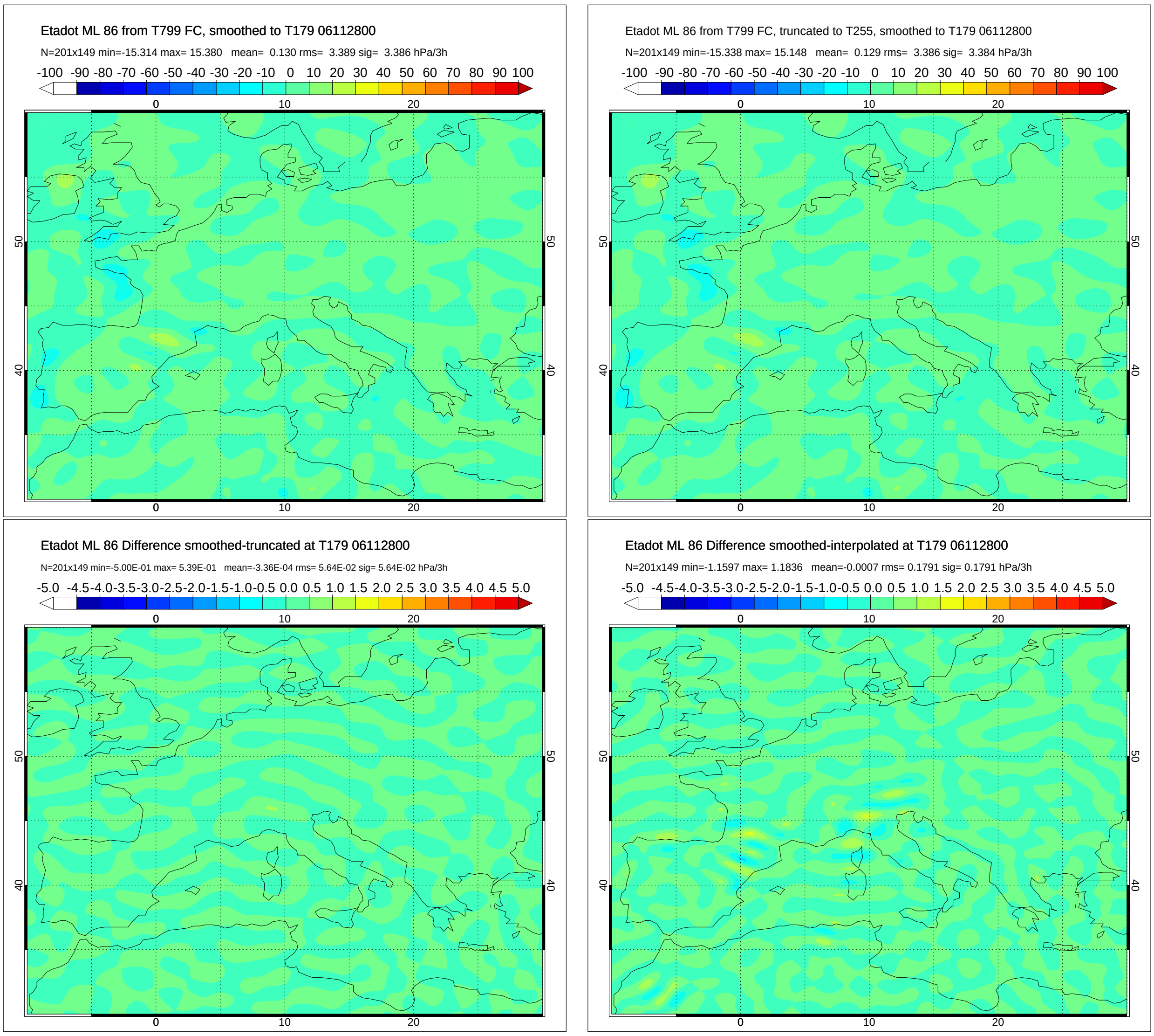


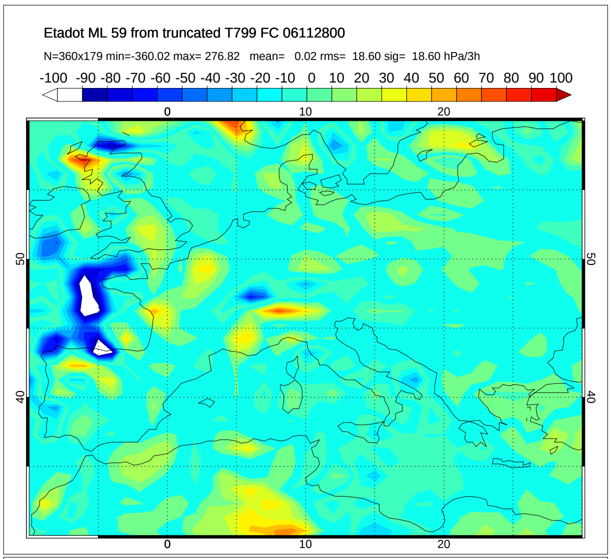
Figure 7.9: As in Fig. 7.7 but for spectral truncation/smoothing T179 on the 86th model level. Error in panel d) relative to signal in a), b) is 5%.

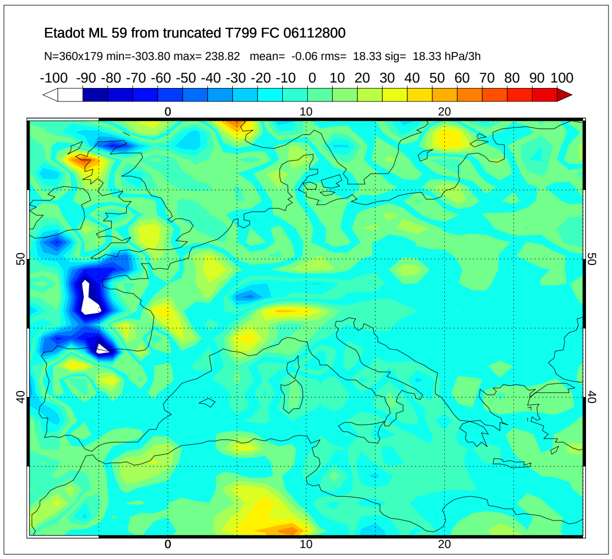
spectral truncation of the input data for the calculation of the vertical velocities. From the results gained above, the following recommendations can be given:

- For global output fields, spectral truncation of the input fields and calculation of the vertical velocity on the Gaussian grid is recommended. For 1.0 degree resolution, spectral truncation of 159 or 255 (with smoothing to T179) is recoommended, for 0.5 degree resolution, spectral truncation of 319 or 399 (with smoothing to 359) is recommended. For these resolutions the elapsed time on ecgate for MARS retrieval is comparable to the elapsed time for vertical velocity calculation.

  The full spectral resolution is necessary only for resolutions of 0.25 degrees or finer. Calculation of the vertical velocity at full resolution on the Gaussian grid is rather expensive and works only on the HPC. While it is quite accurate the costs may be prohibitive for operational implementation.

- Calculation of the vertical velocity from T799 input data for a region on a fine lat-lon grid is





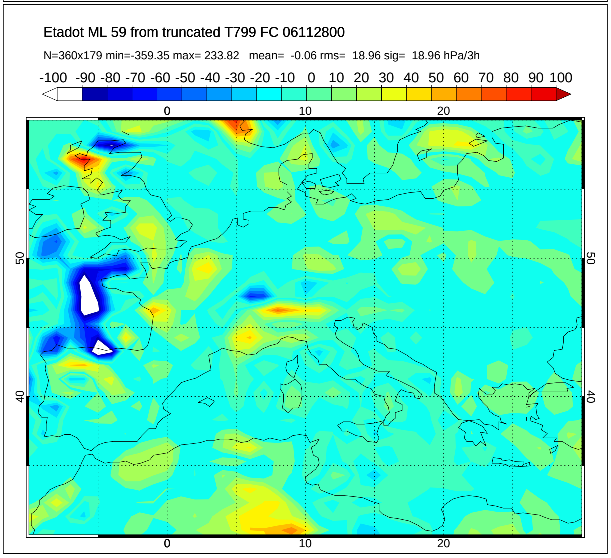


Figure 7.10: Vertical velocity ω in Pa/s, on 1.0/1.0 degree grid, calculated by TOR V2.0 software. a) calculated from T159 data on lat/lon grid, b) calculated from T159 data on Gaussian grid, c) calculated from T255 data on Gaussian grid, then truncated to T179 and transformed to 1.0/1.0 grid. This is the recommended procedure. Compare with higher resolution fi elds above.

Table 7.2: CPU, Disk and memory requirements for performing TOR ECMWFDATAv2.0 calculations for one day (8 EN fi elds) on ecgate and on the HPC facility. Numbers are approximate and vary with CPU/Disk load on HPC. T??? refers to spectral resolution, LL=$\dot{\eta}(\partial p/\partial \eta)$ calculated on Lat/Lon grid, G=$\dot{\eta}(\partial p/\partial \eta)$ calculated on Gaussion grid. T799G-32 means calculated in parallel on 32 processors. 0.2/0.2 E=Europe, 1.0/1.0 G=Global grid.

Task		CPU time [h]	Elapsed time[h]	Memory [GB]	Units	Data ECMWF	Data transferred
V1.0, T319LL, 1.0/1.0 G	ecgate	0.1	0.5	0.3			8×62MB
V2.0, T799G, 0.2/0.2 E	hpce	4.5	5	8	59	4.4GB	8×125MB
V2.0, T799G-32, 0.2/0.2 E	hpce	4.5	1.0	8	59	4.4GB	8×125MB
V2.0, T799LL, 0.2/0.2 E	hpce	0.8	1.1	0.5	11	0.7GB	8×125MB
V2.0, T799LL, 0.2/0.2 E	ecgate	1.2	1.75	0.8		0.7	8×125MB
V2.0, T799G, 0.2/0.2 E	ecgate	does not work on ecgate					
V2.0, T179LL, 1.0/1.0 G	ecgate	0.1	0.5	0.3			8×125MB
V2.0, T255G, 1.0/1.0 G	ecgate	0.2	0.6	1.0		0.7GB	8×125MB
V2.0, T399G, 0.5/0.5 G	ecgate	1	2	2.0	1.2	-	8×500MB

less costly and feasible on both ecgate and HPC if the vertical velocities are not calculated on the Gaussian grid but on the lat-lon grid. The errors incurred by not using a Gaussian grid are on the order of 15%. This may be acceptable given the uncertainties involved in the vertical velocity calculations.

- In order to remove the nonlinearity problem of vertical velocity of truncated data, some form of initialization is necessary, which is normally performed within the IFS at ECMWF. It is presently unclear, if this initialization can be performed outside IFS. Currently it can be avoided only for T399 and T255 resolutions by using EPS control forecasts.

- In any case of changes in the numerical representation of the ECMWF forecast system it is recommended to set option `M_OMEGA=1`. It yields the *rms* difference between ω calculated in MARS and TOR ECMWFDATA, which is a good error estimate for $\dot{\eta}(\partial p/\partial\eta)$

References

Sardeshmukh, P.D., and B.J. Hoskins, 1984: Spatial smoothing on the sphere. *Mon. Wea. Rev.*, **112**, 2524–2529.

Simmons, A.J., and D.M. Burridge, 1981: An energy and momentum conserving finite difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758–766.

APPENDIX

A Update of scripts

The increased flexibility of the decoding software required several new options. These are either job parameters or parameters that can be specified in the control file (`CONTROL_OPS` or `CONTROL_ERA`).

A.1 Job parameters

- The scripts can be submitted to either `ecgate` or the ECMWF High Performance Computing (HPC) system (`hpce`). On ecgate the following should be noted:

- Jobs on ecgate are sent to queue `normal` or `long`. Queue long permits memory requests up to 2GB.

- The program `CONVERT2` has been compiled with `OpenMP` enabled. Note that on ecgate parallel execution is not supported. `OMP_NUM_THREADS` must be set to 1 in order to avoid crashes. Thus do not delete the corresponding line in the jobs.

- On `ecgate` the working directory is `$SCRATCH`.

On HPC system, we recommend the following.

- Jobs on `hpce` should be sent to queue `ns`. In the job header a line such as
  `#@ resources = ConsumableCpus(1) ConsumableMemory(8000MB)`
  is required for full resolution calculations on the Gaussian grid. Without this option the job may fail because not enough memory could be allocated.

  Queue `ns` is intended for serial jobs but permits execution of OMP parallel jobs as well. Queue `np` is not recommended since it leads to less efficient allocation of resources and higher unit consumption.

- The program `CONVERT2` has been compiled with `OpenMP` enabled. The number of threads used can be controlled with the `OpenMP` standard environment variable `OMP_NUM_THREADS`. 32 threads are the default for calculations on `hpce`. For the calculations on the lat/lon grid parallel execution has little impact. On `hpce`, one can choose between `$TEMP` and `$JTMP`. On `$JTMP` the data are automatically destroyed after the end of the job, even if it fails. On `$TEMP` the data are removed only after successful completion of the job.

A.2 New parameters in the control file

- `M_UPPER, M_LOWER, M_LEFT, M_RIGHT` to specify the rectangular subdomain to be exploited. The parameters need to be specified in hundredth of a degree (e.g. 3750 for 37.5°E) since the korn shell allows only integer arithmetic.

- `M_GRID` is the grid resolution, also in hundredth of a degree. Grid sizes down to 1/100 degrees are possible. Note that this is different to Version 1.0, where `M_GRID` was given in degrees.

- The output grid normally has
  `(M_RIGHT-M_LEFT+1)/M_GRID`×`(M_UPPER-M_LOWER+1)/M_GRID`
  elements. If `M_RIGHT-M_LEFT-M_GRID=36000`,
  the output grid is assumed cyclic in E-W direction and has
  `(M_RIGHT-M_LEFT)/M_GRID`
  elements in E-W direction.

- `M_LEVMIN, M_LEVMAX (1,M_LEVEL)` to specify the range of model levels to be extracted. Model level fields contain many stratospheric and mesospheric data which are normally not required for trajectory calculations. The parameters affect only model level fields in the output file `ENYYMMDDHH`. For the calculation of ω and $\dot{\eta}(\partial\eta/\partial p)$ all model level fields are still required.

- `M_ACCURACY (24)` to specify the accuracy of the GRIB encoding of the fields retrieved from MARS. 16 bit is the MARS default but 24 bit is more accurate and is necessary for high spectral resolutions. Note that the output fields ENyymmddhh from the vertical velocity calculation remain 16 bit.

- `M_GAUSS (0)` This option should be used if highest accuracy and special smoothing of the vertical velocity field is required. About 12GB are retrieved from MARS and the computation of the vertical velocity field takes about 10 times more time. Due to significant memory requirements this option should be specified only on the HPC facility at ECMWF, not on ecgate. This option should be used only with the highest available spectral resolution (T159 for ERA40, T255 for ERA-Interim, T799 for the operational model as of November 2006.

- `M_SMOOTH (0)` can be used to filter the highest wavenumbers the vertical velocity field with a spectral filter described by Sardeshmukh and Hoskins (1984) or to simply truncate the highest wavenumbers (which is the default). Specifying `M_SMOOTH=0` implies no filtering. Specifying e.g. 359 removes wavenumbers larger than 359. In order to enable the Sardeshmukh-filter, line 505 in the file phgrreal.f (i.e. line 18 of subroutine SPFILTER) must be uncommented and line 506 must be deleted.

- `M_OMEGA (0)` can be used to extract the vertical velocity ω available from MARS. The fields are written with resolution `M_SMOOTH` (or `M_RESOL` if `M_SMOOTH` is set to zero) to the file `OMEGAyymmddhh`

- `M_OMEGADIFF (0)` can be used to verify the pressure vertical velocity field (ω) calculated by the conversion software with the vertical velocity available from MARS. The difference between these fields is a measure for the accuracy of the $\dot{\eta}(\partial p/\partial\eta)$ calculation. The cost of the job is increased by about 20% if ω is required. The difference field and the surface pressure tendency are appended to file `OMEGAyymmddhh`.

- `SOURCECODE (ecgate:flex_extract_ecgate)` For easier maintenance, it is preferable not to use precompiled binaries but to compile the source code during the job. Further the HPC and ecgate use different compiler defaults (32bit addressing on ecgate, 64bit addressing on HPC). `SOURCECODE` points to a directory with the file `source.tar`, which contains all the Fortran sourcecode as well as the Makefiles. This file is copied from ecgate and the conversion programs `FLXACC2` and `CONVERT2` are compiled during the job.

- `FORCE_EI` forces the use of ERA-Interim reanalyses. These are available for 1989-1992 at the moment but will catch up with the present time in a few months. ERA-Interim uses T255 at 60 model levels and 12 hourly 4D-VAR. Three-hourly resolution is possibly by using 3h and 9h forecasts which are routinely saved. Experiment number 1111 is used by default. There is still experimentation going on in ERA-Interim and the experiment number may change.

- `FORCE_EPS` forces the use of EPS control forecasts instead of operational forecasts. These are available from 26 November 2006 onwards at resolutions T255L62 (low resolution forecasts) and T399L62 (high resolution control forecasts). This option has been included mainly for verification purposes but for some users the lower vertical resolution of the input fields may be attractive.

- `M_STREAM (OPER)` allows to choose some experimental analyses as input.

- `M_NUMBER (1)` This parameter allows to choose between high resolution EPS control forecast (1) and low resolution EPS control forecast (2)

- `M_EXPVER (1)` This parameter allows to choose research department experiments. This is necessary for accessing ERA-Interim products but may be useful for experimental analyses as well.