ECMWF Feature article

from Newsletter Number 133– Autumn 2012

METEOROLOGY

Global, non-hydrostatic, convection-permitting, medium-range forecasts: progress and challenges

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/cequld5s

This article appeared in the Meteorology section of ECMWF Newsletter No. 133– Autumn 2012, pp. 17–22.

Global, non-hydrostatic, convection-permitting, medium-range forecasts: progress and challenges

Nils P. Wedi, Mats Hamrud, George Mozdzynski,
Geir Austad, Sinisa Curic, Jean Bidlot

Numerical weather prediction (NWP) requires an answer in real time: after observations have been gathered, transmitted, received, processed and analysed there is a window of approximately one hour to run ECMWF's medium-range global forecast such that it can be delivered to the Member States in time for them to provide services to their customers. Moreover, forecasts are not based only on single realizations when assessing the fidelity of severe weather predictions. Ensembles of simulations and assimilations are run every day using the same dynamical model, suitably perturbed, to spread as much as possible across the range of uncertainty in the initial conditions, due to model and observation errors, and more fundamentally due to the error growth of the truncated solutions of the governing non-linear equations

While computational efficiency remains one of the most pressing needs of NWP, there is an open question about how to make the most efficient use of the affordable computer power that will be available over the next decades, while seeking the most accurate forecast possible. Fundamental to the latter is the development of a high-resolution assimilation, forecast and ensemble system that is highly scalable (i.e. can exploit massively parallel computers). Such a system would be able to resolve and describe organized cloud systems with deep moist convection in the atmosphere and small-scale orographic, land-cover and land-use features as well as waves and eddies in the ocean. All of these are important aspects of increasing the skill and fidelity of severe weather predictions in a changing climate. Equally, the randomness of probabilistic Earth-system models must be built around a meaningful and accurate prediction of the mean circulation so as not to render the assessment of uncertainty useless.

At the same time there are significant scientific challenges associated with further resolution increases: how to change the governing equations because the hydrostatic assumption is no longer satisfactory for the scales resolved and how to best represent sub-gridscale processes? The problems associated with both aspects and their interconnection have already been described in an earlier newsletter article (*Wedi & Malardel*, 2010). ECMWF plans to implement a global horizontal resolution of approximately 10 km by 2015 for its assimilation and high-resolution forecasts, and approximately 20 km for the ensemble forecasts, which is in line with its steady progress in resolution increases over the past thirty years. The scales resolved at these resolutions are still hydrostatic and the efficiency of the current hydrostatic, semi-Lagrangian, semi-implicit solution procedure using the spectral transform method is likely to remain a relevant benchmark.

In this article we describe a breakthrough in the acceleration of the spectral transform technique, the fast Legendre transform, and subsequently ECMWF's first ever global, convection-permitting, non-hydrostatic weather forecasts at up to T7999 or equivalently ~2.5 km horizontal gridlength. Several examples illustrate the great potential of ECMWF's new ultra-high resolution capabilities.

Fast Legendre Transform

The spectral transform method involves discrete spherical harmonic transformations between:

- Physical (gridpoint) space, where the (semi-Lagrangian) advection and the physical parametrizations are computed.
- Spectral (spherical harmonic) space, where the Helmholtz equation arising from the semi-implicit time-stepping scheme can be solved easily and horizontal gradients are computed accurately, particularly the Laplacian operator (second-order spatial derivative) that is so fundamental to the propagation of atmospheric waves.

A spherical harmonic transformation is a Fourier transformation in longitude and a Legendre transformation in latitude. The Fourier transform is computed numerically very efficiently by using the Fast Fourier Transform (FFT). However, due to the relative cost increase of the Legendre transforms compared to the gridpoint computations, spectral transform models were believed to become prohibitively expensive at very high resolution. However, to address this issue a Fast Legendre Transform (FLT) has now been successfully implemented into Cycle 38r1 of ECMWF's Integrated Forecasting System and is the default option for horizontal resolutions beyond the current operational resolution (T1279). More detailed information about the FLT is given in Box A.

doi:10.21957/cequld5s

Fast Legendre Transform (FLT)

A

Legendre transforms involve many sums of products between associated Legendre polynomials at given (Gaussian) latitudes and corresponding spectral coefficients of the particular field (such as temperature or vorticity at a given vertical level).

The FLT algorithm is based on the fundamental idea that for a given zonal wavenumber all the values of the associated Legendre polynomials at all the Gaussian latitudes of the model grid have similarities that may be exploited in such a way that one does not have to compute all the sums. Rather, FLTs pre-compute a compressed (approximate) representation of the matrices involved in the original sums and then apply a compressed (reduced) representation instead of the full representation at every time-step of the model simulation. The cost increase involved in the additional pre-computation step is negligible (typically less than 0.1% of the total elapsed time of a 10-day forecast).

The success of the FLTs is best summarised in Figure 1 which shows the wall-clock time computational cost (in milliseconds scaled by N^2, where N denotes the truncation limit) of the spectral transforms during a one-hour T7999 simulation with 40 vertical levels with and without the fast transforms. The computational cost of the transforms at T7999 resolution is more than halved using the FLTs.

Figure 1 also shows results for a one-hour T3999 simulation with 40 vertical levels, indicating a much lower cost increase from T3999 to T7999 when the FLTs are used. Note that the number of spectral transforms performed in the T7999 simulations is twice as large because a time-step of 1 minute was used compared to 2 minutes in the T3999 simulations. All simulations used the non-hydrostatic high-resolution model. It is also worth noting that the saving is per simulated hour and therefore is also substantial at T3999.

A particular concern is the computational efficiency of the non-hydrostatic model formulation in part due to its increased use of spectral transforms. Additional runs using the hydrostatic version of the model confirm that the cost of the non-hydrostatic T7999 simulation with FLTs is reduced to a level below the cost of an equivalent hydrostatic T7999 simulation without FLTs. Notably, the hydrostatic model requires two prognostic three-dimensional variables less and does not need the use of an iterative scheme for stability, thus more than halving the number of transforms required compared to the corresponding non-hydrostatic simulation. The results suggest that the concern about the disproportionally growing computational cost of the Legendre transforms with increasing resolution has been mitigated. At T2047 (or equivalently ~10 km horizontal grid length), T3999 (~5 km) and T7999 (~2.5 km) the spectral transform computations take approximately 17%, 20% and 25% of the total elapsed model simulation time.

The cost and latency of the parallel communications of the spectral transforms and the communications within the spectral computations remain a concern. Spectral-to-gridpoint transformations require data-rich global communications at every timestep that may become too expensive on future massively parallel computers. This aspect is being investigated as part of the Collaborative Research into Exascale Systemware, Tools and Applications (CRESTA) project and preliminary results suggest a way forward through the use of modern computer language concepts (e.g. PGAS) to overlap computations and communications. Moreover, there is some evidence that further reductions in the wall-clock time computational cost of the spectral transforms may be achieved by use of GPU (graphics processor unit) and vector technology.

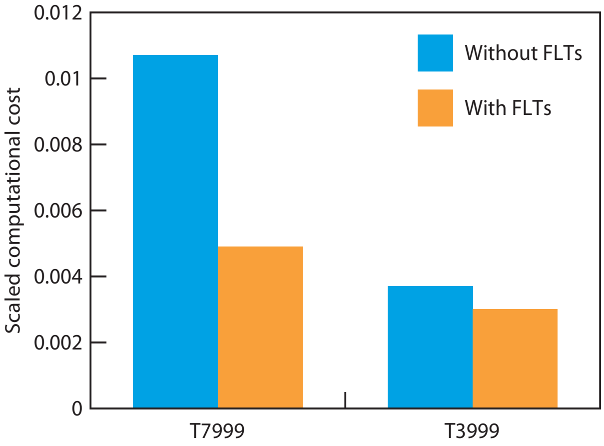


Figure 1 Wall-clock time computational cost (in milliseconds scaled by N^2, where N denotes the truncation limit) of all the spectral transforms during a one-hour simulation with N=7999 and N=3999 respectively, and with 40 vertical levels. A comparison is made between otherwise equivalent non-hydrostatic high-resolution simulations with (red) and without (blue) fast Legendre transforms (FLTs). The latter use the standard matrix-matrix multiply routine (DGEMM) instead. Since the number of spectral transforms performed in the T7999 simulations is twice as large, because a time-step of 1 minute was used compared to 2 minutes in the T3999 simulations, the T3999 values have been scaled accordingly.

ECMWF’s first ultra-high resolution, convection-permitting, global weather forecasts

With the FLTs in place, using ECMWF’s recently installed IBM Power 7 computer produced the first 10-day forecasts at T3999 (~5 km grid interval) with the operational 91 vertical levels. At this resolution each prognostic variable has approximately 21 million points per vertical level, and with 14,336 compute threads (i.e. independent and in parallel executed program instructions) about 23 forecast days per (wall-clock) day were achieved. In addition, the world’s first global T7999 12-hour forecast (~2.5 km grid interval with about 80 million points per vertical level) has been successfully completed. Using 12,800 compute threads, not surprisingly, the speed was only 2.7 forecast days per day, so just ahead of real time. The T7999 simulation was restricted to only 40 vertical levels due to the large memory requirements and the compute threads were physically spread over more than half of the total IBM Power 7 cluster.

To put these numbers in perspective, we would require about one million compute threads to achieve the operational requirement of 240 forecast days per day at T7999 resolution (and more with a corresponding increase in the number of vertical levels) and it remains an open question whether the relative cost of communications for such a computer will scale accordingly. Nevertheless, “that’s one small step…”. The ability to run such ultra-high resolution simulations and the evaluation of their results will greatly impact on our understanding of the multi-scale interactions, from convective to global scales, while feeding back into future weather and climate model developments.

The first integration of a global, non-hydrostatic, convection-permitting simulation at ECMWF, where non-hydrostatic effects start to be resolved, has been made possible not least because of the successful collaboration between Météo-France and ECMWF. Naturally, the handling of these ultra-high resolution datasets is challenging and specialised software had to be written to handle the latest satellite-derived surface datasets to be used as initial and boundary data in the T3999 and T7999 forecasts. ECMWF’s newly developed interpolation software ecRegrid has been used for creating and processing the climatological ultra-high resolution surface datasets. Here we would also like to thank Manuel Fuentes, Fernando Ii and Iain Russell for their help with archiving and plotting this data.

Error growth and kinetic energy spectra

An important point to make is that these resolutions are well beyond the spatial resolution of the globally-available observations, and our simulations indicate a faster (root mean square) error growth over the first 24 to 48 hours with substantial kinetic energy in the smaller scales cascading upwards, in particular over the first 12 hours. This effect is best illustrated with the T7999 simulation initialised from the T1279 (~16 km) analysis (suitably interpolated from low-to-high resolution) yet forced with the underlying climatological topographic and land-cover information representative of 2.5 km resolution.

Figure 2 shows the global horizontal kinetic energy spectra at 500 hPa for the first 12 hours of the T7999 simulation, where a relative maximum can be clearly identified at the smaller scales not represented by the analysis. But as time progresses the ‘energy gap’ is filled quickly, consistent with experience from

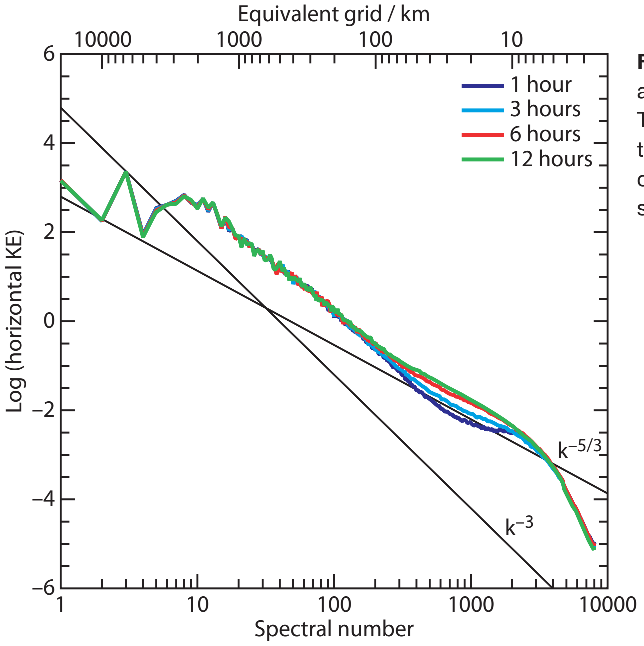
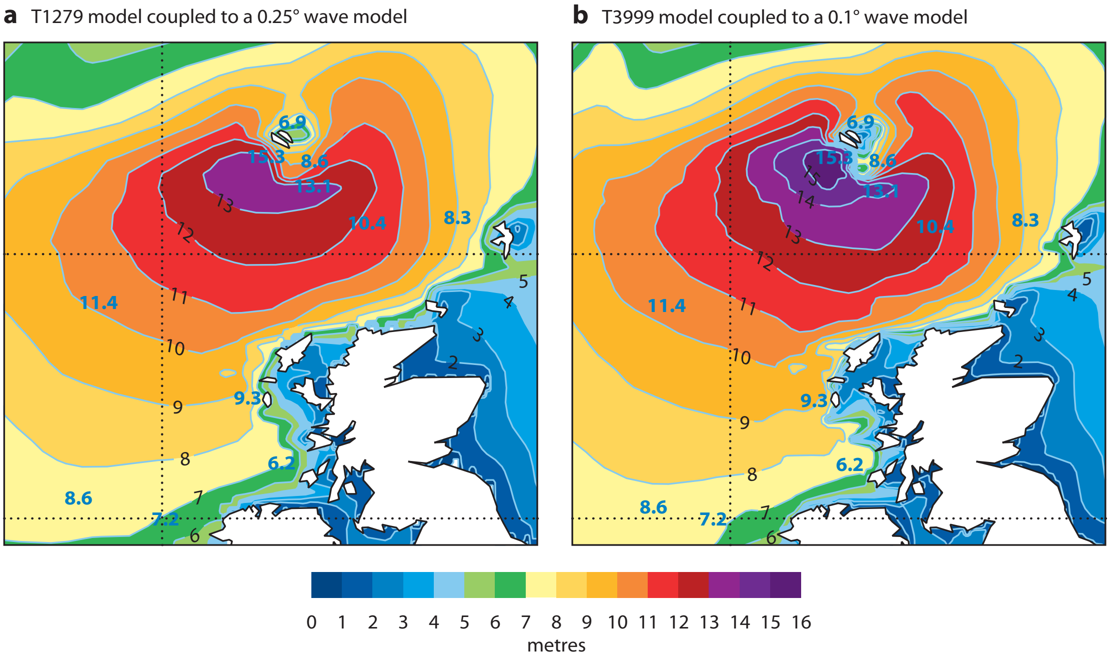


Figure 2 Global horizontal kinetic energy spectra at 500 hPa height for the first 12 hours of the T7999 (~2.5 km grid length) simulation against the total wavenumber (with the largest number corresponding to the truncation limit of the spherical harmonics series expansion, i.e. 7999).

limited-area model studies forced by global models, where the observed -5/3 energy spectrum is established very quickly (*Skamarock*, 2004). Notably, we find approximately equal contributions from divergent and rotational motions (not shown) to the well resolved part of the global -5/3 kinetic energy spectrum starting from about wavenumber 300 (~70 km).

The results concerning the kinetic energy spectrum are of significant theoretical importance (*Lindborg*, 2007). On the one hand this result provides bounds of validity to quasi-geostrophic theory, the fundamental explanation of the dynamic evolution of large-scale atmospheric flow, since the horizontal divergence spectrum is far from negligible as would be the case for geostrophic flow. On the other hand, the equipartition of divergent and rotational contributions over a wide range of wavenumbers (300 to 3,000), which is established within the first 6 to 12 hours of simulation and then maintained, rules out the dominance of divergent motions as the main reason for the shallower tilt of mesoscale energy spectra.

As pointed out by *Skamarock* (2004), the rapid adjustment of the model in the (uninitialized) mesoscale range may indicate that these ultra-high resolution forecasts provide added value in deterministically predicting mesoscale structures when forced with realistic external forcings (i.e. orography and land-cover) and perhaps also by removing problematic parametrizations (i.e. deep convection). Indeed it is not known if the initial rapid error growth due to the small time and length scales resolved by the model may diminish the predictive skill of the ultra-high resolution forecasts or not. However, the T3999 10-day forecasts beyond the initial period of 24 to 48 hours indicate a similar large-scale (root-mean-square) error evolution compared to the T1279 control forecasts. This suggests that the influence is not overwhelming and rather stresses (a) the important challenge for data assimilation in providing accurate initial conditions at these ultra-high resolutions and (b) substantial future opportunities for ensemble data assimilation and ensemble forecasting.



a T1279 model coupled to a 0.25° wave model

b T3999 model coupled to a 0.1° wave model

Figure 3 24-hour forecast for significant wave height for 00 UTC on 25 November 2011: (a) T1279 model coupled to a 0.25° global wave model (operational configuration) and (b) T3999 model coupled to a 0.1° global wave model. The blue numbers are the corresponding two-hourly averaged significant wave height observations (data from the UK Met Office, UK WaveNet Programme, Irish Marine Institute, and Faroese Office of Public Works).

Simulation of severe weather events

Improved realism can already be demonstrated with two examples of severe weather events. The first example evaluates the impact of ultra-high resolution on the predicted wave-height when the storm Xaver (also named Berit) hit the Faroe Islands on 25 November 2011. Figure 3 shows the 24-hour T1279 forecast coupled to the (operational) 0.25° wave model and the 24-hour T3999 forecast coupled to the higher resolution 0.1° wave model, both compared to buoy observations. The added value of the finer resolution is seen in the much more realistic prediction of monster waves approaching, as not only surfers but anybody at sea or on the shore will appreciate the difference, whether to expect a 10 or 15 metre swell!

The second example picks up on the often publicised shortcoming of operational forecast models at the time failing to predict the true intensity of 'Lothar' on 26 December 1999 with record wind speeds observed in France, Germany and Switzerland. A particular aspect of this storm was its rapid development and progression from the Atlantic across France to Germany making it difficult to capture with high-resolution limited-area models forced by low-resolution global boundary conditions. In this case the large-scale conditions triggered a meso-scale development of hurricane strength within just a few hours. Even as late as the evening before the storm affected large parts of Europe, one would not have easily spotted it on the satellite picture.

Figure 4 shows the expected wind gusts and surface winds predicted (just) 11 hours in advance by the T1279 (~16 km) forecast compared with the T7999 (~2.5 km) forecast, neither of which was available at the time. While both forecasts much improve on the predicted level of excessive wind speeds compared to the actual forecasts available at the time, it is the ultra-high resolution that matches much closer the reality of extraordinary high wind speeds, especially for the Alps and for the lower mountain ranges throughout Western Europe.

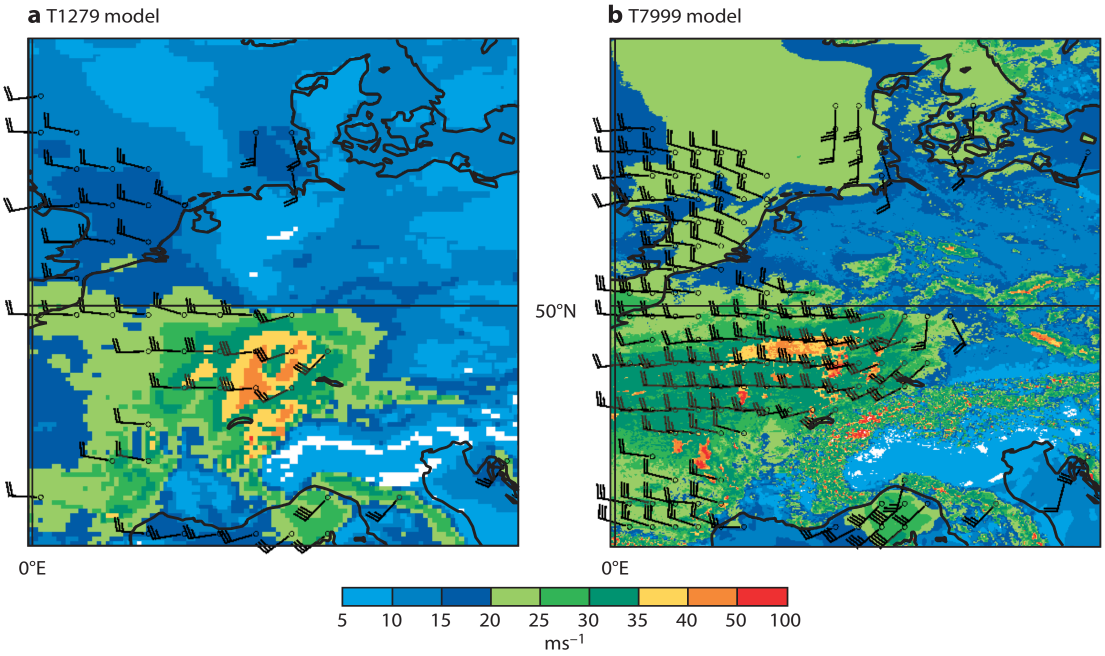


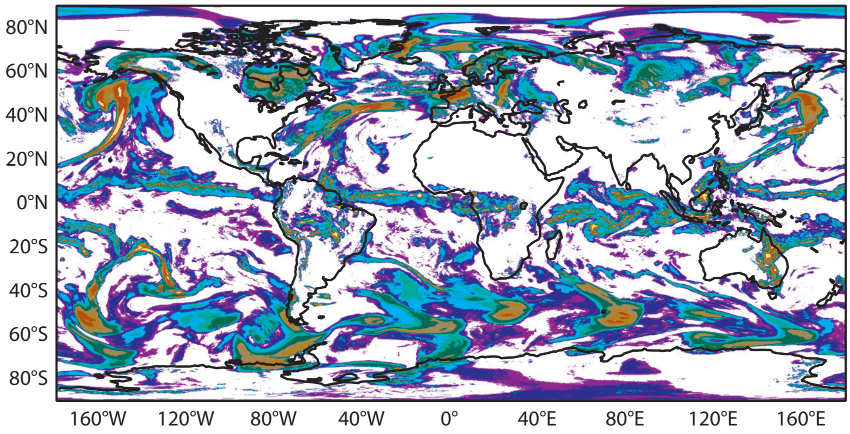
Figure 4 11-hour forecast of wind gusts and 10-metre surface winds for the Christmas storm 'Lothar' on 26 December 1999 from: (a) T1279 (~16 km) model and (b) T7999 (~2.5 km) model. During the storm some of the highest wind speeds ever recorded in Europe were observed (69 ms^{-1} on Jungfraujoch, Switzerland; 75 ms^{-1} on Hohentwiel, Singen, Germany). The trail of 'destruction' is clearly marked by the wind gust data. Note that the lower mountain ranges now prominently feature in the wind gust data for T7999.

Explicit representation of convection

A big step forward in the development of global models is the transition from a parametrized (where sub-gridscale effects are expressed in terms of grid-scale parameters) to an explicit representation of convective processes. This is a big and problematic transition because, with increasing resolution, the model explicitly resolves large thunderstorm cells, towering nimbostratus and cumulonimbus clouds, as well as cloud clusters and fronts with embedded deep convection, yet part of their early development (i.e. starting from a little cumulus cloud) is not. The T7999 simulations are also the first global high-resolution forecasts from ECMWF with an explicit representation of deep convection (i.e. with the deep convection parametrization switched off). Detailed studies of the data will hopefully provide us with much insight on this transition region.

Here a preview is given of the large-scale and convective precipitation from the T7999 model (Figure 5) compared with equivalent simulations using parametrized deep convection at the operational resolution of T1279 (Figure 6). At ultra-high horizontal resolutions the cloud microphysics and the vertical and horizontal transport of hydrometeors become very important and a more different picture may have perhaps been expected (*Wedi & Malardel*, 2010). However, it is reassuring that the large-scale precipitation is quite similar in the T1279 and T7999 simulations, with parametrized shallow convection still covering large parts of the globe in both, and with only few systematic (but positive) reductions in convective precipitation over some areas in the T7999 simulation. This comparison provides further evidence of the fundamentally deterministic predictability at larger scales albeit explicit representation of the smallest scales.

a T7999 large-scale precipitation



b T7999 convective precipitation

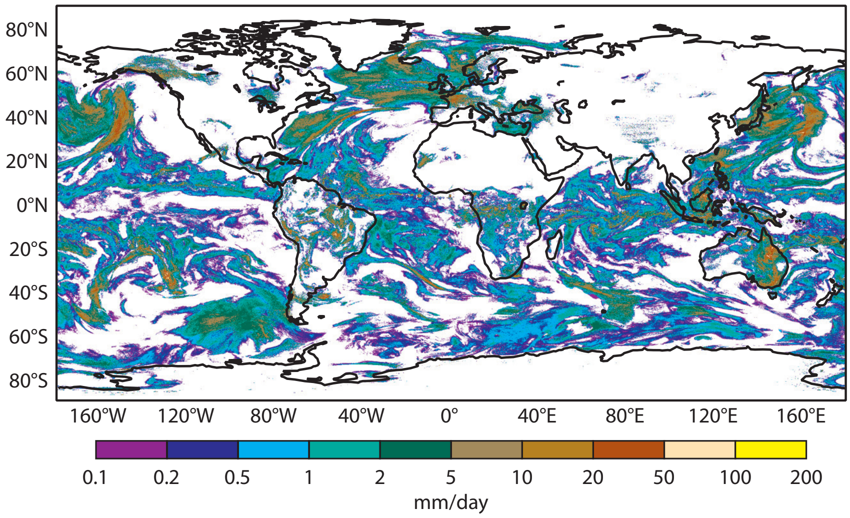
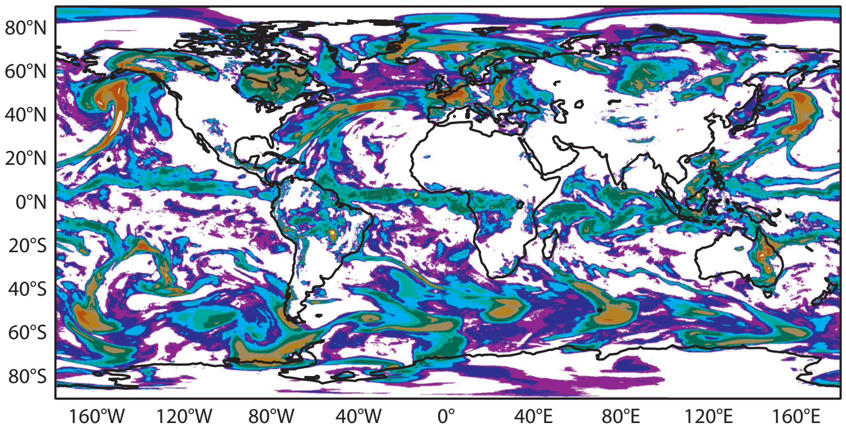


Figure 5 Global (a) large-scale precipitation and (b) convective precipitation for the T7999 (~2.5 km) simulation with an explicit representation of deep convection after 9 hours. Note that in this simulation convective precipitation originates only from the still parametrized shallow convection. The parameter 'large-scale precipitation' is not so large-scale anymore at this resolution as it includes all the explicitly resolved motions including deep convective systems.

doi:10.21957/cequld5s

a T1279 large-scale precipitation



b T1279 convective precipitation

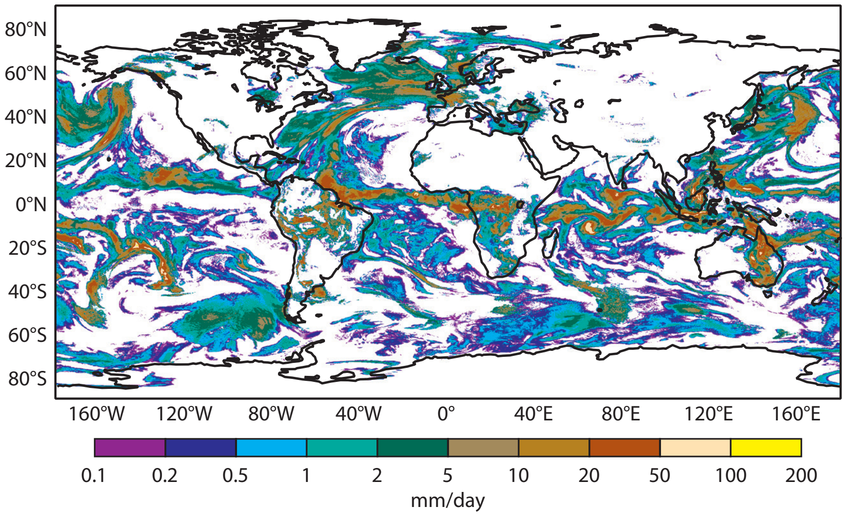


Figure 6 Same as Figure 5 but for the T1279 (~16 km). In this case both deep and shallow convection are parametrized.

What next?

Much work will be required to turn the glimpse of ultra-high resolution capabilities into an operational reality over the next 20 years, but undoubtedly, considerable opportunities for research lie ahead and ECMWF is now in a good position to explore these.

Further reading

Lindborg, E., 2007: Horizontal wavenumber spectra and vertical vorticity and horizontal divergence in the upper troposphere and lower stratosphere. *J. Atmos. Sci.*, **64**, 1017–1025.

Skamarock, W., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032.

Wedi, N. & **S. Malardel,** 2010: Non-hydrostatic modelling at ECMWF. *ECMWF Newsletter* **No. 125,** 17–21.

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England