Nonhydrostatic modelling with IFS: current status

Nils P. Wedi, Mats Hamrud and George Mozdzynski

ECMWF, Shinfield Park, Reading
RG2 9AX, United Kingdom
wedi@ecmwf.int

ABSTRACT

The article describes the current status of non-hydrostatic (global) modelling at ECMWF and the issues arising with the existing IFS modelling framework based on the spectral transform method when applied to ultra-high resolution simulations. Given the strategic importance of the future dynamical core of ECMWF's forecast model, some of the questions posed need to be addressed in the near future. However, considering the proven efficiency of the IFS on massively parallel computers for horizontal resolutions up to and including 10 km, considering the recent developments on fast Legendre transforms and given that a range of ECMWF applications will remain in the hydrostatic regime for some time, it strengthens the view to seek improvements to the dynamical core within the existing spectral code framework while continuing to review the progress of suitable alternatives.

1 Introduction

In 1987 ECMWF implemented an operational model and data assimilation system with T106 resolution (125 km). Since then the resolution has been steadily increased over the years as indicated in table 1. The Centre plans to implement a horizontal resolution of 10 km by 2015 for its assimilation and deterministic forecast system, beyond which a nonhydrostatic dynamical core will be required. The current dynamical core of the IFS model is based on the hydrostatic primitive equations and is likely to be of limited use at horizontal resolutions finer than about 10 km, where non-hydrostatic effects will become important (ECMWF, 2000; Wedi and Smolarkiewicz, 2009). Rather than developing a non-hydrostatic dynamical core for the Centre's model from scratch or investigate other existing formulations, it was decided to evaluate the non-hydrostatic formulation developed by the ALADIN group for regional numerical weather prediction (NWP) and made available by Météo-France in the global IFS/ARPEGE model framework Bénard et al. (2010). The aim was to assess whether this formulation is able to fulfil the requirements of high accuracy, efficiency and robustness imposed by ECMWF's various global operational applications and has the potential to form the basis of ECMWF's future non-hydrostatic dynamical core. The governing equations of this non-hydrostatic model are the unapproximated Euler equations for the (optionally) deep or shallow atmosphere.

The tests performed ranged from seasonal climate runs at T159 (125 km) resolution to medium-range forecasts up to and including T2047 (10 km) to assess the performance of the non-hydrostatic model in the hydrostatic regime, all the way to ultra-high resolution simulations in the non-hydrostatic regime (Wedi et al., 2009). Experiments with the T2047 horizontal resolution indicate that the differences between the hydrostatic and the non-hydrostatic simulations are still not significant at this resolution. Even the highest horizontal resolution at which the IFS can be run to date (T3999, 5 km) is still too coarse to fully resolve non-hydrostatic phenomena. Consequently, a testbed has been developed that enables testing of the global non-hydrostatic dynamical core at non-hydrostatic scales at an affordable computational cost. Rather than create a two-dimensional vertical slice model of the three-dimensional global model or develop a limited area version of the IFS, a testing framework more suited for the global code was considered. The testing framework is based on the idea of shrinking the radius of the planet

Table 1: Resolution increase of the operational forecasting system.

Year	Spectral truncation	smallest "resolved" half-wavelength
1987	106	188km
1991	213	95km
1998	319	63km
2000	511	39km
2006	799	25km
2010	1279	16km
2015?	2047	10km

such that, with an affordable number of grid-points covering the globe, the desired resolution resolving non-hydrostatic phenomena is achieved, without incurring the prohibitive cost associated with such a high resolution on the full-sized planet (Wedi and Smolarkiewicz, 2009). The size of the computational domain is reduced without changing the depth or the vertical structure of the atmosphere. The underlying assumption is that the essential flow characteristics remain unchanged when the ratio of horizontal to vertical scales is reduced. Consequently, the planetary radius is suitably reduced to capture non-hydrostatic phenomena without incurring the computational cost of actual simulations of weather and climate at non-hydrostatic resolution. This framework provided a direct comparison, both quantitatively and qualitatively, of non-hydrostatic simulations with analytic solutions and with large-eddy simulation (LES) benchmarks of limited-area models published in the literature.

Wedi et al. (2009) concluded that the non-hydrostatic dynamical core is a possible choice for future, globally-uniform high-resolution applications at ECMWF. However, there are issues, in particular with the computational efficiency, that need to be addressed before it is fit to be used as the dynamical core of the operational model at all resolutions. In particular, it appears necessary in the ECMWF framework that the semi-implicit (SI) time discretisation is augmented by an iterative-centred-implicit (ICI) scheme (Bénard, 2003; Bénard et al., 2010), where the prognostic variables used in the computation of the nonlinear explicit residual as well as in the semi-Lagrangian trajectory calculations are updated at every iteration. The combination of the iterative procedure together with the relatively high cost of the spectral transforms at every time-step limit the applicability of the new non-hydrostatic model. Notably both aspects appear to be a lesser problem in the limited-European-area operational context at Météo-France, where the ICI scheme is not used operationally and the Fourier transforms are relatively cheap compared to ECMWF's spectral transforms comprising of Fourier and Legendre transforms on the sphere.

The computational efficiency of the existing hydrostatic IFS model has been recently demonstrated in the Athena project, a NSF funded initiative to determine the feasibility of using dedicated supercomputing resources to rapidly accelerate progress in modelling climate variability out to decadal and longer time scales (Kinter et al., 2010). Over a six months period IFS was run for 3 x 47 years at T159 resolution, 1 x 47 years at T511, 3x 47 years at T1279, and an additional 19 years at T2047 resolution. All simulations used 91 vertical levels and were forced by prescribed sea surface temperatures (SSTs). Overall the time-step used was a factor 10-15 times larger compared to existing state-of-the-art non-hydrostatic models at equivalent resolution. The reduced grid ($\approx$ 30 percent less gridpoints compared to an equivalent latitude-longitude grid) and the relatively cheap direct solver in the semi-implicit scheme also contribute to the efficiency of the IFS.

However, when convection-permitting resolutions are reached, it is an open question if such a time-step advantage can be maintained and if the semi-Lagrangian advection algorithm remains competitive given the increased importance of micro-physical processes and the potential importance of their consistent, conservative, monotone and positive-definite advection (Grabowski and Smolarkiewicz, 1990). More-over, a further increase in resolution beyond 10 km requires a non-hydrostatic dynamical core which is more expensive in its current form (approximately three times more expensive at T3999 resolution). Some of these questions will have to be addressed in the future. In the following we concentrate on the question if the spectral transform method continues to be competitive beyond hydrostatic scales for future global ultra-high resolution NWP applications. Here we specifically deal with the computational effort required. However, it should be noted that the cost associated with the necessary global communications to transform the data from spectral space to physical space, or within spectral space (to vary the parallel distribution depending on the algorithmic requirements and their load balancing) make up 50 percent of the computational cost of the spectral transforms at T2047 on existing computer architectures and resolutions. These communications are relatively few but data rich, hence mostly limited by bandwidth. In this sense, the cost scales with the number of computing nodes as with more nodes each has to communicate less data. In contrast, communications in gridpoint models would be of much higher frequency with very little data to transfer, hence these are mostly limited by the latency of the communications. The latter also applies to (iterative) 3D elliptic solvers arising from semi-implicit discretizations of the governing equations.

2 Towards a fast Legendre transform

The spectral transform method has been successfully applied at ECMWF for approximately thirty years, with the first spectral model introduced at ECMWF in April 1983. Spectral transforms involve discrete spherical harmonics transformations between physical (gridpoint) space and spectral (spherical harmonics) space. The spectral transform method has been introduced to NWP following the work of Eliasen et al. (1970) and Orszaag (1970), pioneering the efficiency gain (underlying ECMWF's existing dynamical core) obtained by partitioning the computations, with a part in physical space (e.g. products of terms, the semi-Lagrangian advection and the physical parametrizations) and a part in spectral space (e.g. accurate horizontal derivatives, the Helmholtz equation arising from the semi-implicit scheme). The success in NWP in comparison to alternative methods has been comprehensively reviewed in Williamson (2007).

The (triangular) truncated series expansion in spherical harmonics of a scalar variable ζ on the surface of a sphere is given by

$$\zeta(\theta,\lambda) = \sum_{m=-N}^{N} e^{im\lambda} \sum_{n=|m|}^{N} \zeta_n^m \overline{P_n^m}(cos(\theta)), \tag{1}$$

where θ, λ denote latitude and longitude, respectively, ζ_n^m are the spectral coefficients of the variable ζ, and the $\overline{P_n^m}$ are the normalised associated Legendre polynomials of degree n and order m as a function of latitude only; N symbolises the cut-off wavenumber in the spherical harmonics expansion and it is approximately half the number of gridpoints along the equator. Equation (1) represents the inverse spherical harmonics transform. A direct spherical harmonics transformation is accomplished by a Fourier transformation in longitude as

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

and a Legendre transformation in latitude for each m as

$$\zeta_n^m = \frac{1}{2}\int_{-1}^{1} \zeta_m \overline{P_n^m}(cos(\theta)) dcos(\theta). \tag{3}$$

The Fourier transform is computed numerically very efficiently by using the Fast Fourier Transform (FFT). The Legendre transforms require the accurate (discrete) computation of the integral in (3) which is accomplished by using Gaussian quadrature

$$\zeta_n^m = \sum_{k=1}^{K} w_k \zeta_m(x_k) \overline{P_n^m}(x_k) \tag{4}$$

at the $K = (2N+1)/2$ (linear grid) special quadrature points ("Gaussian latitudes") given by the roots of the ordinary Legendre polynomials $\overline{P_N^{m=0}}(x_k) = 0$ and the Gaussian weights are given by the formula

$$w_k = \frac{2N+1}{[\overline{P_N^{m=1}}(x_k)]^2}. \tag{5}$$

The asymptotic cost of the spectral transform method scales with N^3 (Tygert, 2008). Due to this relative cost increase (in particular the Legendre transforms) compared to the gridpoint computations, very high resolution spectral models may become prohibitively expensive. Up to a resolution of approximately $N = 2047$ (10 km) the very high level of optimisation achieved for the spectral computations masks the N^3 asymptotic cost of this part of the IFS model.

For ultra-high horizontal resolution — such as the T3999 simulations — the cost and the accuracy of the pre-computations also becomes important. In IFS the associated Legendre polynomials P_n^m at the roots of the ordinary Legendre polynomials $P_n^{m=0}$ are pre-computed and stored. The recurrence formula provided in Belousov (1962) is used for the computation of the associated Legendre polynomials P_n^m at the roots of the ordinary Legendre polynomials $P_n^{m=0}$. The Belousov recurrence relation is stable and accurate even at very large N but the cost of the computation is relatively high as the recurrence formula involves varying degrees and orders of the "intermediate" associated polynomials. The accuracy for large N can be improved by computing the polynomials in extended (quadrupole) precision. Alternatively, Schwarztrauber (2002) suggests a Fourier representation of the ordinary Legendre polynomials together with simple formulas to evaluate the corresponding coefficients. The resulting $P_n^{m=0}$ and $P_n^{m=1}$ are then used for finding the roots of the ordinary polynomials via Newton iterations as well as for the starting points in the Belousov formula. Such a procedure has been implemented in the IFS transform library and provides in fact improved accuracy at large N while keeping the computations in double precision as opposed to extended precision.

Recent progress in the development of fast spherical harmonics transforms by Tygert (2008) and Tygert (2010) have suggested the possibility of a stable algorithm that scales according to an asymptotic cost of $C * N^2 log(N)$ (where C is a constant). The new algorithm based on the butterfly scheme (Tygert, 2010, and references therein) has been incorporated into the IFS. In the butterfly scheme the rank of the matrices involved and thus the computations necessary to multiply them are reduced. A second step of the new algorithm requires interpolations from locations at the roots of the ordinary Legendre polynomials to the roots of associated Legendre polynomials and vice versa, which is accelerated using a simple fast-multipole method (FMM) (Yarvin and Rokhlin, 1999). Both steps require highly accurate and extensive pre-computations which are the subject of current research. In particular, the Belousov formula appears to be too expensive to be applied in the pre-computations for the algorithm in Tygert (2010) and the recurrence relation advocated in Tygert (2008) is used instead. The latter only depend on the degree n but is independent of previous orders m (which is advantageous as these are distributed over different processors). The recurrence relationship appears to be stable (even in double precision) for large N although the accuracy of the Gaussian weights (which depend on the value of two successive degrees of associated polynomials) is poor towards the equator. Equally, for large order m the accuracy of the small valued associated polynomials is poor. The roots of the associated Legendre polynomials are found by solving the Prüfer transformed ordinary second order linear differential equation which is satisfied by the normalised associated Legendre polynomials (see Tygert, 2008, for details). We use a fourth order Runge-Kutta method to solve the associated first order initial value problem. However, we found the accuracy of the corresponding Gaussian weights to be poor and we improve the accuracy by additional Newton iterations.

Preliminary results have now been obtained with IFS for resolutions up to T3999 (5 km). The left panel of Fig.1 compares the total number of floating point operations executed per model time-step during the inverse Legendre transform and as expected, the new algorithm shows less floating point operations with increased resolution compared to the conventional method. The right panel of Fig.1 shows the associated computational cost indicating the desired reduced cost compared to the conventional method with increasing resolution. Currently, the pre-computations at T3999 take approximately 20 minutes for the butterfly method. However, the end result of the computation of the associated polynomials at all the roots can be easily stored on file, which would substantially reduce that time.

The top panel in Fig.2 obtained with the conventional transform compared to the bottom panel obtained with the butterfly method in the inverse transform suggests a small loss of accuracy at T3999 resolution that is not seen for example at T2047 (not shown). Further investigation indicates that this is not related to the FMM interpolation step but likely related to the accuracy of the associated Legendre polynomials in the pre-computations with increased resolution and their subsequent use in the butterfly algorithm.

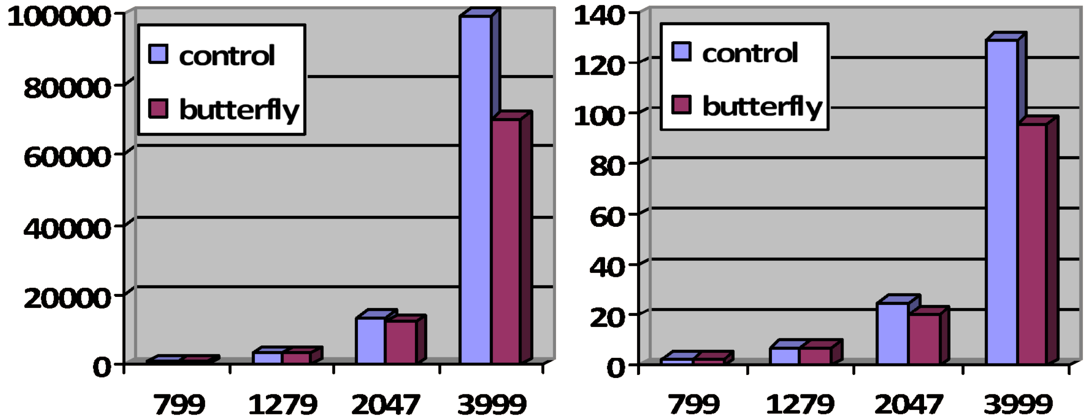

Figure 1: Panel a shows the number of floating point operations [Gflop] required per time-step for the inverse transform (from spectral to grid-point space) for four different resolutions comparing the conventional and the butterfly algorithm. Panel b shows the corresponding computational cost (in seconds) required.

3 Conclusion

It can be concluded that substantial reductions in computational cost may be achieved for the spectral transforms by using the butterfly algorithm in ultra-high resolution simulations. Further work is required to implement the direct transforms and to optimise the pre-computations in terms of accuracy and cost. A more detailed analysis of the precise distribution of the matrix rank reductions in the butterfly algorithm may also lead to more optimal pre-computations for NWP applications.

The non-hydrostatic model will clearly benefit from the reduced cost of the spectral transforms. However, the cost could be halved if the ICI scheme was not required for stability. Further understanding is required why ECMWF's nonhydrostatic model is not stable without ICI scheme (unless severe reductions of the time-step are imposed compared to the hydrostatic IFS). An investigation into alternative formulations of the non-hydrostatic code framework (e.g. Arakawa and Konor, 2009) may equally lead to improved stability and accuracy for ultra-high global simulations and is pursued in future.

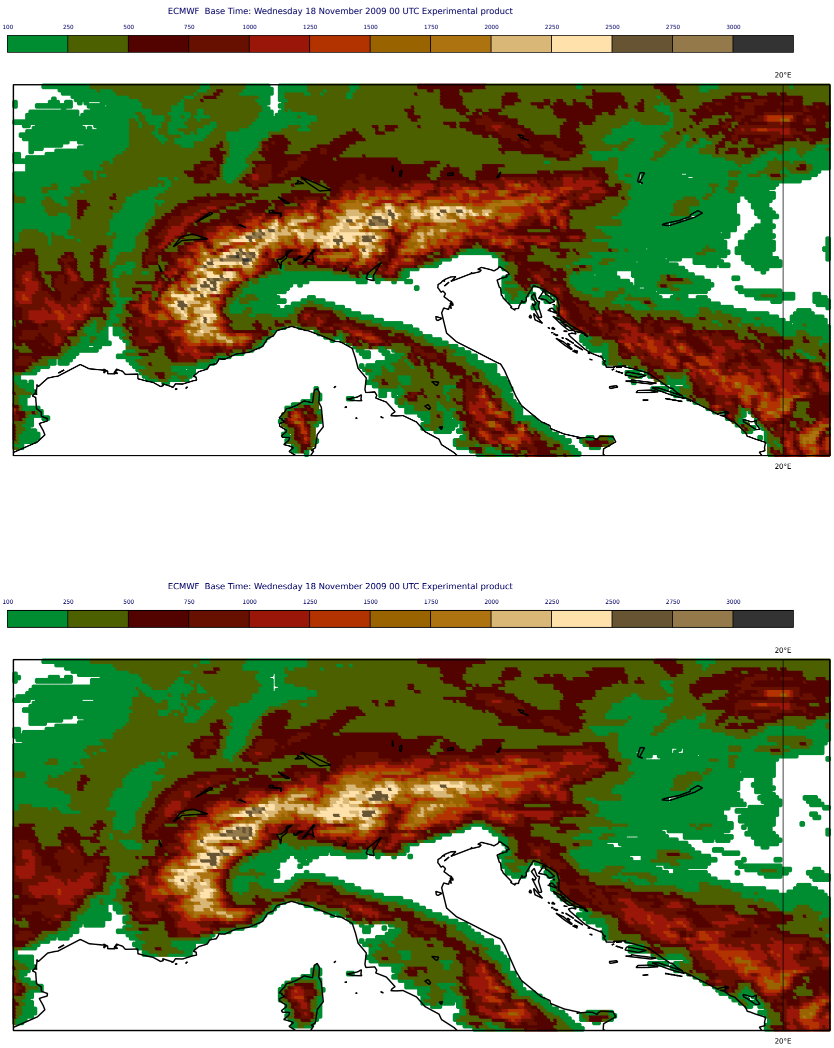

Figure 2: Panel a shows the orography of the Alps after a direct and an inverse transform with the conventional method at T3999 resolution. Panel b shows the same but where the inverse transform is done with the butterfly method.

Acknowledgements

We would like to express our sincere thanks to Mark Tygert for his help and advice. We are also grateful to Mike Fisher for his help with the implementation of the Cuppen algorithm originally proposed in Tygert (2008). Finally, we would like to thank Agathe Untch for her helpful comments on an earlier version of the manuscript.

References

Arakawa, A. and C. Konor (2009). Unification of the anelastic and quasi-hydrostatic systems of equations. *Mon. Weather Rev. 137*, 710–726.

Belousov, S. L. (1962). *Tables of normalized associated Legendre polynomials*. New York: Pergamon Press.

Bénard, P. (2003). Stability of semi-implicit and iterative centered-implicit time discretizations for various equation systems used in NWP. *Mon. Weather Rev.*, 2479–2491.

Bénard, P., J. Vivoda, J. Mašek, P. Smolíková, K. Yessad, C. Smith, R. Brožková, and J.-F. Geleyn (2010). Dynamical kernel of the Aladin-NH spectral limited-area model: Revised formulation and sensitivity experiments. *Q.J.R. Meteorol. Soc. 136*, 155–169.

ECMWF (2000). In *Proc. 2000 Workshop on Developments in numerical methods for very high resolution global models*, Reading, UK. Eur. Cent. For Medium-Range Weather Forecasts.

Eliasen, E., B. Machenhauer, and E. Rasmussen (1970). On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Report 2, Institute for Teoretisk Meteorologi, University of Copenhagen.

Grabowski, W. W. and P. Smolarkiewicz (1990). Monotone finite difference approximations to the advection-condensation problem. *Mon. Weather Rev. 118*, 2082–2097.

Kinter, J., D. Achuthavarier, J. Adams, E. Altshuler, B. Cash, P. Dirmeyer, B. Huang, E. Jin, L. Marx, J. Manganello, C. Stan, T. Wakefield, T. Palmer, M. Hamrud, T. Jung, M. Miller, P. Towers, N. Wedi, M. Satoh, H. Tomita, C. Kodama, Y. Yamada, P. Andrews, T. Baer, M. Ezell, C. Halloy, D. John, B. Loftis, and K. Wong (2010). Revolutionizing climate modelling project Athena: A multi-institutional, international collaboration. *Bull. Am. Meterol. Soc.*. in prep.

Orszaag, S. A. (1970). Transform method for calculation of vector coupled sums: application to the spectral form of the vorticity equation. *J. Atmos. Sci. 27*, 890–895.

Schwarztrauber, P. N. (2002). On computing the points and weights for Gauss-Legendre quadrature. *SIAM J. Sci. Comput. 24*, 945–954.

Tygert, M. (2008). Fast algorithms for spherical harmonic expansions, II. *J. Comput. Phys. 227*, 4260–4279.

Tygert, M. (2010). Fast algorithms for spherical harmonic expansions, III. *J. Comput. Phys. 229*, 6181–6192.

Wedi, N. P. and P. K. Smolarkiewicz (2009). A framework for testing global nonhydrostatic models. *Q.J.R. Meteorol. Soc. 135*, 469–484.

Wedi, N. P., K. Yessad, and A. Untch (2009). The nonhydrostatic global IFS/ARPEGE: model formulation and testing. Technical Report 594, Eur. Cent. For Medium-Range Weather Forecasts, Reading, UK.

Williamson, D. L. (2007). The evolution of dynamical cores for global atmospheric models. *J. Meteor. Soc. Japan 85B*, 241–269.

Yarvin, N. and V. Rokhlin (1999). An improved fast multipole algorithm for potential fields on the line. *SIAM J. Numer. Anal. 36*, 629–666.