A non-hydrostatic SI dynamical core :current-state, limitations and perspectives

Pierre Bénard

Centre National de Recherches Météorologiques
42, Av G. Coriolis, TOULOUSE, FRANCE
pierre.benard@meteo.fr

ABSTRACT

A description of the the current status and future challenges for the dynamical core operationally used at ECMWF, Météo-France, HIRLAM, and Aladin Centres is given here, with a special focus on time-discretization aspects.

The current status for this dynamical core is shown to be expectably viable for Global and Limited Area Model (LAM) applications for mid-term future, but several concerns are identified, which will need to be addressed within typically ten years, and may imply heavy revisions in the algorithms and the discretization strategies.

In LAM application, if the use of the Semi-Implicit (SI) technique is to be maintained, the very high resolution will require an implicit treatment of orographic forcing terms. In Global and LAM applications the increase of the resolution in combination with the SI scheme will also require the ability to use local methods in space. Further perspectives are given in view of abandoning the SI method, if it appears that this is required in a longer-term future.

1 Introduction

The dynamical core of the global IFS model is closely related to that of the global ARPEGE model used at Météo-France, and of the ALADIN, AROME and HARMONIE limited-area variants used in several Meteorological Centres mainly in Europe. All these models are based on the same general method, which can be summarized by: spectral, semi-implicit, semi-Lagrangian with a transform method applied to an unstaggered collocation grid. Furthermore, all these models share their dynamical core with the same source code, but used with different options for the various above mentioned applications. This code is developed jointly among the partners, and the community of Meteorological Services making use of it for operational applications has been increasing through the years.

The present paper is focused on the time-discretization aspects of this dynamical core. Two related discussions, one mostly oriented on space-discretization, and the other on the particular case of IFS may be found respectively in the contributions of M. Hortal and N. Wedi in the same volume.

Historically, the first cooperation on the dynamical core involved ECMWF and Météo-France from the late 1980's, in order to develop a global spectral semi-implicit Hydrostatic Primitive Equations (HPE) model, which rapidly evolved into a semi-Lagrangian three-time levels (Ritchie *et al.*, 1995), then two-time levels scheme (Simmons and Temperton, 1997). The capability of a variable horizontal resolution through a conformal transformation (Courtier and Geleyn, 1988) was introduced, and is used as a specificity of ARPEGE.

The LAM (HPE) version ALADIN was developed in parallel, as a cooperation between Météo-France and the ALADIN consortium during the 1990's. A first fully compressible non-hydrostatic version in Euler Equations (EE) was then developed (Bubnova *et al.*, 1995) for the Eulerian time-scheme, and was progressively extended to semi-Lagrangian three time-levels and two time-levels schemes during the 2000's. Besides, this non-hydrostatic version of the code was then extended to global uniform and variable resolution.

All versions coexist within the same code, and some of them currently have or have got an operational status: the global HPE version is currently used by IFS and ARPEGE, the LAM HPE version is used in some centres of the ALADIN consortium, and the LAM EE version is used in ALADIN and HIRLAM centres, as well as in Météo-France. These LAM EE versions have received various names according to the specific physics or Data Assimilation packages that are associated to them for the operational use: the applications in Météo-France, ALADIN and HIRLAM Services are respectively termed AROME (Seity *et al.*, 2010), ALARO and HARMONIE.

The global EE version has not received any operational status yet, mainly because the horizontal resolutions currently available operationally are not believed to be fine enough to expect a significant improvement in forecasts quality, compared to the HPE version. However, the use of the global EE version will become necessary when the horizontal grid mesh will reach about 2.5 km, that is, in a mid-term future.

The current status of the highest resolution EE operational application (AROME) is discussed in section 2, then the limitations of the current system will be detailed in section 3, the challenges and perspectives in section 4, and finally some further perspectives and conclusions will be given in section 5.

2 Current status of AROME

AROME is the high resolution LAM non-hydrostatic model used operationally at Météo-France and in some other countries of the ALADIN consortium. The model is coupled to ARPEGE for the lateral boundary conditions. The spectral technique for LAM models is made possible through a mathematical extension of physical fields into doubly-periodic functions of space. The extension is performed through cubic spline functions inside a so-called extension zone.

AROME uses a 'meso-scale' (column-wise) physics package partly issued from mesoscale research community. The operational applications (for which the horizontal mesh is always below 2.5 km) are used without any deep-convection parameterization. AROME also has its own assimilation cycle based on a 3D-VAR method with a 3 h cycle. The refinement over the 4D-VAR ARPEGE data assimilation mainly comes from the extra mesoscale observations assimilated (Radar reflectivities, Doppler wind, thiner satellite radiances...)

The current configuration of AROME operational at Météo-France has the following configuration. Horizontal mesh Δx = 2.5 km, 60 levels, 750*720 grid-points in the horizontal domain and time-step Δt = 60 s. The model delivers only short term forecasts, up to the 30 h range. One of the salient propertiy of this dynamical core is its very limited overcost with respect to an HPE version that would be used in exactly the same configuration (this can be accessed by logical switches in the code as mentioned above).

Besides this operational application, an experimental version has been built as a prototype for the next operational version which will be installed on Météo-France new computer and switched to operational status around summer 2014. This new machine has about 1000 nodes, subdivided in about a total number of about 25000 cores, and offers about 500 Tflops sustained capacity. This prototype AROME version has an horizontal mesh of Δx = 1.3 km, 90 levels, and a time-step Δt = 45 s. The horizontal grid, 1536*1440 points, results in a slightly larger domain than the current operational one. This prototype has run routinely at 00UTC every day, up to a range of 30 h, in dynamical adaptation mode (without Data Assimilation) on a smaller domain between June 2012 and July 2013, then on the nominal domain from August 2013. For this application, the semi-implicit (SI) scheme which was used from the origins, had to be replaced by a so-called Iterative Centred Implicit (ICI) scheme in order to guarantee a full robustness of the algorithm. The concept of the ICI scheme and its differences with the SI schemes are detailed in section 3. The ICI scheme induces an overcost of about 10% per model time-step, compared to the SI scheme. The behaviour of this application appeared to be robust (no failure was observed in the forecast model itself). Some improvements were observed in particular situations, as well as small improvements in statistical performances. It can be said that the kilometric resolution with horizontal grids up to 4 $\times 10^6$ points does not bring any significant problem in terms of operational exploitation with the current version of the dynamical core. The CPU overcost of this next operational version with respect to the current configuration is decomposed as follows:

Δx, domain	Δz	Δt	SI $\rightarrow$ ICI	Other (physics...)	Total
$\times$ 4.1	$\times$ 1.5	$\times$ 1.33	$\times$ 1.1	$\times$ 1.1	$\times$ 10

Another, more experimental version of AROME explores the sub-kilometric scales with an horizontal mesh Δx = 0.5 km, about 100 levels, and a time-step between 10 and 15 s. So far, this experimental version has been run only on isolated test cases (in dynamical adaptation) but not on a routine basis. The robustness seems to be compatible with an operational use in NWP, but the small sample of meteorological situations than only has been already explored is not enough to answer this question with full certainty.

3 Limitations of the current system

One of the most important identified concerns of the dynamical core of AROME (or IFS, ARPEGE,...) is the scalability on massively-parallel distributed-memory computing architectures. The code offers two nested levels of distribution for the computations, leading to column-wise sub-domains of arbitrarily small size (down to a single column) for grid-point space computations. However, a full efficiency cannot be reached with such extremely small sub-domains. The main reason is in the semi-Lagrangian scheme. The fast access to information located at the origin point of the semi-Lagrangian trajectory associated to any grid-point implies the existence of a necessary data halo for each sub-domain, all around the physically relevant part of the distributed sub-domain. The information in these halos is read-only, but must be duplicated from other neighbouring sub-domains. The area ratio of the physically relevant part of the sub-domain over the total sub-domain (including its halo) of course becomes adversely detrimental when the area of the physically relevant part of the subdomain reaches the limit of one column. Besides, the spectral transform technique implies one or several forth-and-back transforms between the physical grid-point space and the spectral space at each time-step, depending on the type of time-scheme (SI or ICI). For these transforms, a global transposition of data is required is necessary (in both global and LAM cases), and becomes a severe problem at large truncations.

On top of this, the distribution is two-dimensional while the integration domain is three-dimensional, therefore a full one-dimensional data-vector in any space (grid-point, Fourier or spectral) cannot be itself distributed on several separate computations cores. When the number of cores becomes larger than the product of the number of degrees of freedom of the two distributed dimensions in each space, no extra efficiency gain may of course be expected from any further increase of the number of computational cores. However, this latter problem (which corresponds to a very strong scalability problem) is not specific to the particular strategy of the dynamical core, but only witnesses the absence of a distribution on the third dimension in each computational space (as in most other models).

The problems listed above are therefore expected to successively occur at different granularities in the future. The spectral data transposition problem occurs at large truncations, independantly of the granularity; then, at increasing granularity, the problems which will successively be encountered are the bottleneck problem of the semi-Lagrangian halo, the distribution limit of the spectral transform computations, of spectral space computations, and finally of the grid-point space computations. It is believed that the three latter problems are not of practical interest in the foreseeable future, given the evolution of High Power Computing semi-conductor components. The problem of the semi-Lagrangian halo may be alleviated by specific techniques which may not require a deep overhaul of the general strategy for the dynamical core. However, the problem of the spectral transform data-transposition may become a handicap and strongly prompts to examinating alternative approaches. The problem is not reported as potentially severe at short term, but a change of method is a process that may take a significant time, and must be anticipated quite long in advance.

Some other limitations may be identified beside the various computational limitations listed above. First, the ability of the current dynamical core to maintain a high level of quality and robustness for severe high-resolution flows and steep orography is questioned. The limitations here seem to come mainly from the terrain-following vertical coordinate and in a smaller part from the semi-Lagrangian technique which assumes some smoothness of the transported variables. Second, the compatibility with new developments, as e.g. the change prognostic variables or the implementation of Vertical Finite Elements discretizations (VFE) is weak. Here a significant part of the difficulties comes from the spectral method, for which the discrete spatial operators are full matrices which cannot be controlled as easily as for more local methods. Moreover, the spectral method restricts the semi-implicit schemes to the class of 'constant-coefficients' semi-implicit schemes (see e.g. Bénard, 2004), thereby further restricting their adaptability to problems that have large horizontally-varying characteristics. For instance, the spectral technique used with terrain-following vertical coordinates does not allow any implicit treatment of orographic forcing terms, which become large horizontally-varying sources at fine resolutions..

The maximum slope in the AROME domain (which includes the steepest part of the European orography, in the Alps) is found to be 23° at Δx = 2.5 km, 38° at Δx = 1.3 km, and 53° at Δx = 0.5 km. The robustness of the SI scheme at Δx = 1.3 km was found weaker than at Δx = 2.5 km: at Δx = 2.5 km the forecast model was always stable for Δx = 60 s, but showed evidence of instability in some cases with Δx = 30 s at Δx = 1.3 km. This is why the SI scheme of the current application Δx = 2.5 km had to be changed for an ICI scheme for the prototype pre-operational version Δx = 1.3 km. This change suggests that some limit of the operational applicability of the SI scheme as it is currently is reached for kilometric resolutions.

The most severe of the concerns listed above involve directly or indirectly the current semi-implicit scheme and the spectral technique, and prompts the revision of the current spectral SI scheme. Three avenues may be considered for this. The first avenue is to change the spectral SI or ICI scheme into a non-spectral SI or ICI scheme (with a local horizontal discretization scheme). This avenue avoids the shortest-term computational bottleneck (global data transpositions), opens the way to SI or ICI schemes which are not based on 'constant-coefficient' linearisation, and might also improve the adaptability of the scheme to alternative prognostic variables or to VFE discretizations. The second avenue would consist in abandoning the implicit treatment of fast processes along the horizontal, and move to the class of Horizontally Explicit Vertically Implicit (HEVI) schemes. This would of course eliminate all the problems linked to the implicit treatment of horizontal high-frequency forcings, but it is not firmly established that the efficiency could compete with the one of implicit schemes, and therefore, a significant part of the computational effort might be unduly put on time computations. The third avenue would be to combine an horizontally explicit scheme with a governing equations system which filters fast vertically-propagating perturbations. For NWP, classical anelastic systems are not believed to be appropriate, because they distort the propagation of Rossby waves (Davies *et al.*, 2003), but the Quasi-Elastic system proposed by Arakawa and Konor, 2008 is free of this defect and may be considered for this avenue. In this case the scheme may become fully explicit with an efficiency comparable to the one of the second avenue (HEVI schemes). The option which is retained for the mid-term perspective is the first one, the two other ones being considered as rescue and longer-term solutions in case the first one is not able to provide satisfactory solutions.

4 Challenges and perspectives

4.1 SI schemes

In a general formalism, the continuous governing equations of the system used in a forecast model may be written as:

$$\frac{\partial \mathscr{X}}{\partial t} = \mathscr{M}(\mathscr{X}) \tag{1}$$

where $\mathscr{X}$ is a set of continuous fields defined in the whole domain, and termed the 'prognostic variables'. The operator $\mathscr{M}$ is a spatial operator acting on $\mathscr{X}$. In a numerical forecast model, all the continuous functions or operators involved in (1) are discretized in space and time. Here we consider only leap-frog schemes for simplicity (but all reasonnings extend to other types of schemes). For an explicit-centred scheme, (1) becomes therefore:

$$\frac{X(t+\Delta t) - X(t-\Delta t)}{2\Delta t} = M[X(t)] \tag{2}$$

where X is now a state-vector (a finite list of numbers) and M an operator acting on X. For a centred-implicit scheme (also called Crank-Nicolson scheme) the evolution is discretized by:

$$\frac{X(t+\Delta t) - X(t-\Delta t)}{2\Delta t} = M\left[\frac{X(t+\Delta t) + X(t-\Delta t)}{2}\right] \tag{3}$$

The explicit scheme is of course very easy to implement, but has a very weak stability when M contains large coupled sources implying fast processes. On the other hand the centred-implicit scheme has a very good stability (may be unconditionnally stable in many cases), but is very difficult to implement as soon as M is non-linear, because the presence of the unknown $X(t+\Delta t)$ in the RHS implies the inversion of a large non-linear multivariate operator. In the concrete case, the inverse of the non-linear operator may only be approached, and this generally requires iterative algorithms. The SI method is based on a separation of M into a part L_* which is linear, and a part $(M - L_*)$ which does no longer contain fast coupled sources. The SI scheme then writes:

$$\frac{X(t+\Delta t) - X(t-\Delta t)}{2\Delta t} = (M - L*)[X(t)] + L_*\left[\frac{X(t+\Delta t) + X(t-\Delta t)}{2}\right] \tag{4}$$

The stability depends on the content of the explicitly treated residual $(M - L*)$, since the last RHS term in unconditionally stable. The choice of L_* is of course crucial for the stability of the scheme.

The first method proposed historically for finding the operator L_* associated to M was the linearization around a stationary state. In this method a stationary state X_* is chosen, then M is linearized around X_* to provide L_*, that is:

$$L_* = \left[\frac{\partial M}{\partial X}\right]_{(X=X_*)} \tag{5}$$

where the operator $\partial M/\partial X$ is the tangent linear to M (i.e. the Jacobian matrix of M when it is expressed in terms of the components of X). With this method, the difficult problem of choosing an operator L_* is reduced to the simpler problem of choosing a stationnary state X_*. However, there is no general rule for determining the most appropriate choices of X_*. In the case of Shallow-Water (SW) or Hydrostatic Primitive Equation (HPE) models with moderate resolution, this linearization strategy was found to be appropriate with very simple states (e.g. resting, isothermal, and hydrostatically balanced). For the non-hydostatic EE system, things become more complicated and this strategy sometimes fails. With the EE system, Bénard 2004 showed a very simple example (for a two-time levels scheme) where the linearization strategy fails, but a simple linear appropriate system L_* can still be found. In this case the linear system L_* may no longer be written in the form of (5) for any X_* state. This means that for complex and stiff systems, the linearization strategy around a simple stationary state is not always appropriate for building a stable SI scheme.

Ths SI scheme requires the inversion of the linear system L_* but this overcost (compared to explicit schemes) has always been, up to now, more than compensated by the much larger robustness in operational applications, giving access to long time-steps, and thereby, to a larger efficiency, especially when combined with a semi-Lagrangian approach. However, a disadvantage of the SI technique is that the results are not garanteed, and often must be obtained at the cost of tedious and adventurous predictive stability analyses (at least for complex governing systems as EE). It must be outlined here that if the explicit schemes as in (2) are always conditionally stable (in terms of Δt), SI schemes as in (4) may be unconditionally stable or unconditionally unstable (in Δt), according to the conditions and to their design. It is a false idea that 'adding some implicitness in the system will necessarily make it more stable, it may not hurt'. On contrary, when a SI scheme is ill-designed it is usually unconditionally unstable (in Δt), which means that there is no value of Δt leading a stable evolution, even for a fixed forecast range. Introducing some implicitness in the equations may therefore 'hurt' a lot, if made inappropriately.

As a matter of fact the robustness of these SI schemes, obtained through a very simple linearization approach, was found to be more and more difficult to maintain when the complexity of the systems increased, during the history of NWP. The strategy was fully appropriate for original SW and HPE systems, it could be maintained for EE systems with moderate resolution, but seems to fail for EE systems with high resolution (because the non-linearities become more and more dominant in proportion).

Two main solutions were proposed to solve this difficulty: the first approach can be termed 'TL-SI' and the second, 'ICI'.

4.2 TL-SI and ICI schemes

The underlying idea of the TL-SI approach is that the observed lack of robustness of SI schemes for modern NWP systems mainly comes from the large departure that may exist between the current actual state $X(t)$ and the simple stationary states chosen in the initial method. It is therefore natural to choose $X_* = X(t)$ and then to apply the linearization strategy to obtain L_*:

$$L_* = \left[\frac{\partial M}{\partial X}\right]_{(X=X(t))} \tag{6}$$

In this case, the magnitude of non-linear residuals is minimal, which in principle minimizes the risks of instability. The resulting scheme is very close to the Crank-Nicolson scheme. However, the linear operator L_* now becomes dependant on all time and space dimensions, which makes its computation and its inversion computationally expensive. Practical solutions to this consist in trying to appropriately determine the terms which are the most important in L_*, and to keep only these. In any case, the full non-separable dependency of L_* on all dimensions makes iterative algorithm required for its inversion. The instability of the scheme then usually results from a lack of convergence of this algorithm (since a limited fixed number of iterations must be specified to maintain an acceptable gain in efficiency with respect to explicit schemes).

In ICI schemes, a simple, time-independant, linear operator L_* is kept, but the time-scheme is iterated. The scheme writes (with the iterative index k):

$$\frac{X^{(k)}(t+\Delta t)-X(t-\Delta t)}{2\Delta t}=M\left[\frac{X^{(k-1)}(t+\Delta t)+X(t-\Delta t)}{2}\right]+L_*\left[\frac{X^{(k)}(t+\Delta t)-X^{(k-1)}(t+\Delta t)}{2}\right] \quad (7)$$

An initial condition ($k=0$) is required. If a simple time extrapolation is chosen:

$$X^{(0)}(t+\Delta t)=2X(t)-X(t-\Delta t) \quad (8)$$

then the first iteration ($k=1$) precisely amounts to the SI scheme (4):

$$\frac{X^{(1)}(t+\Delta t)-X(t-\Delta t)}{2\Delta t}=M[X(t)]+L_*\left\{\frac{X^{(1)}(t+\Delta t)-[2X(t)-X(t-\Delta t)]}{2}\right\}. \quad (9)$$

The subsequent iterations lead to successive schemes in which a larger part of the evolution is treated implicitly. In the case where the iterative algorithm converges the vanishing last term of (7) shows that the Crank-Nicolson scheme is reached after convergence. If L_* is time-independant, then the inversion of L_* may be performed once at the beginning of the forecast and the inverse stored instead of being recomputed at each time-step and iteration. The resulting scheme may therefore be made quite efficient (though the iteration decreases the total efficiency of course). However, the convergence of the iterative strongly depends on the choice of L_* and is not garanteed, and instability in this case occurs due to the divergent iterative algorithm.

These two methods (TL-SI and ICI) to go beyond the classical SI scheme are quite different. In the first method, the aim is to minimize $L_*-L_{X(t)}$, and the result is that the explicit residual is close to the true non-linear part of M; in the second method, the aim is to approach Crank-Nicolson scheme globally, and the result is that the explicit residual is made progressively smaller through iterations. The efficiency and robustness of the first scheme is linked to the fast convergence of the iterative solver for the time-dependant implicit problem, whereas for the second it is linked to the fast convergence of the pre-inverted iterative implicit algorithm.

4.3 Challenges for high resolution modelling

The above versions of ICI and SI schemes are implemented in the dynamical core of the models considered here. The use of a spectral representation makes the solution of the implicit problem directly accessible in spectral space, and therefore the SI and ICI schemes are very efficient. However, as mentionned above the spectral technique imposes that the coefficients in the linear operator L_* have no horizontal dependencies. This limitation is an inconvenience at high resolution, because the orographic forcing becomes a very important and fast contribution, and is a large horizontally-varying linear term. This forcing cannot therefore be represented implicitly when the implicit solver is a spectral one, and must be represented explicitly.

When the resolution increases, the slope of the orography also increases, and there is an indication that the explicit treatment of orographic forcing terms becomes a handicap for mesh sizes below 1 km. For $\Delta x = 2.5$ km, the SI scheme appears to be robust with $\Delta t = 60$ s, whereas for $\Delta x = 1.3$ km it is not fully robust with $\Delta t = 30$ s. However, the stability is restored for the ICI with one iteration beyond the SI scheme, that is, $k = 2$ in (7): in this case, the scheme is fully robust for $\Delta t = 45$ s. Since the extra cost of the ICI scheme is marginal (about 10-15% of the forecast model total cost), the efficiency of the scheme at $\Delta x = 1.3$ km is maintained by the additional iteration (and even slightly improved), but this indicates that the explicit handling of orographic terms begins to be a problem for the stability of the SI at kilometric scales. The problem is overcome by the iteration of the scheme, but at higher resolution, the larger discrepancy between the linear and non-linear terms might prevent the convergence of the iterative process, and the stability of the ICI scheme itself.

In the following, it is assumed that the problem of the explicit treatment of orographic forcing terms is the most critical and the first that will be faced in the mid-term future for the SI scheme. To support this statement, fluid mechanics flows with smooth orography or no orography at very high resolutions (about 10 m) are simulated with a very satisfactory robustness by the current discretization schemes. The orographic forcing is linear in terms of the horizontal wind and the coefficient, while strongly space-dependent is time-independent. The implicit problem with orographic forcing included becomes non-homogeneous and non-separable, and therefore requires a non-direct algorithm to be solved (through a three-dimensional iterative solver). The best choice for the horizontal discretization is then a compact-support method as e.g. Finite Elements (FE) or Finite Differences. The choice of the solver therefore should make use of the sparsity of the matrix to be inverted, and possibly also of its time-invariance which in principle allows a method with pre-inversion, unless the memory requirements are too high.

The choice of local discretization schemes instead of the spectral discretization leaves two questions open: which prognostic variables and staggering on one hand; and which grid geometry for the global case on the other hand ?

The choice of prognostic variables and staggering is strongly linked to the propagation of short and fast modes, to the geostrophic adjustment and to the existence of computational modes (cf. Staniforth and Thuburn, 2012, and contribution of J. Thuburn in this volume). An advantage of non-staggered grids is that they preserve by nature the correct ratio between the number of horizontal wind components and pressure degrees of freedom (2:1). This prevents the existence of spurious computational modes. Moreover, with unstaggered grids, the discretization of the continuous problem is disconnected from the details of the grid geometry, since all data points have the same status. Therefore structured and unstructured grids may be used indifferently. However, it is known that a formulation of the equations through wind components (u, v) leads to a poor representation of the shortest resolved gravity waves propagation (and hence to a poor representation of the fine-scale geostrophic adjustment). This problem may be avoided when using vorticity and divergence (ζ, D) as prognostic variables (this choice is known as 'Z–grid'). However, the wind component are still needed in other parts of the forecast model (e.g. for the advection or the physics), and a transformation algorithm $(\zeta, D) \leftrightarrow (u, v)$ is required. This implies the solution of a set of two-dimensional Poisson equations, but here also, the coefficients of the problem are time-independent. This standard problem is not anticipated to weight significantly on the global efficiency.

However, in two or more dimensions, staggered C–grids, which lead to an optimal propagation of short gravity waves, have a poor representation of the fine-scale vortical mixing , whereas unstaggered grids are optimal regarding this point. It is therefore seen that both A– and C–grids have one advantage and one inconvenience for the representation of fine-scale processes (propagation of gravity waves and vortical mixing). At low resolution, vortical mixing is maybe not as important as the geostrophic adjustment, but at fine scales, this mixing becomes relatively more important, because it generates the fine-scale structures that will have to be subsequently handled by the subscale turbulent mixing parameterization, and therefore strongly participate to a correct energy cascade. It is therefore possible that the advantages of C–grids over A–grids are smaller at high resolution, and this question should be re-examined prior to choosing a strategy.

Another issue is the choice of the grid geometry for the global application. The current grid is a reduced longitude-latitude grid. In the case of non-spectral models, this grid is reported as being of poor quality because it has very large truncation errors near poles and this problem – *inter alia* – leads to a poor representation of Rossby waves propagation (through spurious meridional transports). However, it appears that these large truncation errors near pole and this poor representation only occur when local and inappropriate discretization schemes with low-order accuracy are used. High-order local schemes with an exact representation of zonal waves of low spatial frequency should not be subject to this problem. It may therefore be expected that the use of this type of grid in combination with appropriate local schemes could lead to viable solutions. We believe that a the properties of this grid should be re-examined in detail before considering a complete change of grid geometry for our global applications (IFS and ARPEGE). The advantages of the reduced lat-lon grid is that it has a zonally structured configuration, which makes the computation of horizontal derivatives potentially a simple and separable problem along zonal and meridional directions. The data inside zonal rows have a uniform longitude spacing, and zonal rows may also be separated by a uniform meridional spacing. The possibility of a staggered C-grid may also be considered with this grid-geometry, provided high-order zonal interpolations are used to accurately evaluate meridional wind components along a meridian, through meridional interpolations.

5 Further perspectives and conclusion

The viability of the improved SI and ICI approaches considered above is not garanteed. The convergence of the iterative ICI scheme, the cost of the non-direct implicit solver, and the solution of the set of two-dimensional Poisson equation might question the advantage of the implicit scheme over less robust and simpler methods which finally could become competitive in terms of cost. For this reason, the properties of Horizontally-Explicit Vertically-Implicit (HEVI) scheme will be examined for the EE system. By coupling split-explicit methods for the horizontal propagation of fast waves to an implicit scheme for the treatment of vertically propagating elastic disturbances, the limitation in time-steps (compared to SI or ICI) schemes would be compensated by a much more economical algorithm per time-step.

In an other domain, the limitations brought by the current terrain-following coordinates will be examined and the potential benefits of semi-destructured approaches with stepwise orography will be examined. This type of representation of the orography is, here also, independent of the grid geometry, and therefore may be applied with triangular, qualdrilateral or whatever grid. One of the possible justifications of representing a continuous real orography by a discontinuous step orography is that with increasing resolution the spurious response generated just above the steps would become potential at moderate wind velocities, and therefore this response, being evanescent, would not significantly contaminate the whole domain depth.

Acknowledgements

Fruitful discussions and collaborations with Fabrice Voitus, Charles Colavolpe, Steven Caluwaerts, Ludovic Auger, Jean-François Geleyn and many others are acknowledged.

References

P. Bénard, R. Laprise, J. Vivoda, P. Smolíková. (2004) Stability of Leapfrog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations: Flat-Terrain Case. *Monthly*

Weather Review, 132, 1306–1318

Bubnova R., G. Hello, P. Bénard et J.-F. Geleyn, 1995 : Integration of the fully-elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the Arpge/Aladin NWP system. *Mon. Wea. Rev.*, 123, 515–535.

Courtier, P. and Geleyn, J.-F. (1988), A global numerical weather prediction model with variable resolution: Application to the shallow-water equations. *Q.J.R. Meteorol. Soc.*, 114: 1321–1346.

Davies, T., Staniforth, A., Wood, N. and Thuburn, J. (2003), Validity of anelastic and other equation sets as inferred from normal-mode analysis. Q.J.R. Meteorol. Soc., 129: 2761–2775.

Ritchie, H., Temperton, C., Simmons, A., Hortal, M., Davies, T., Dent, D., and Hamrud, M. (1995). Implementation of the Semi-Lagrangian Method in a High-Resolution Version of the ECMWF Forecast Model. *Monthly Weather Review*, 123(2), 489–514.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F, , Lac, C., and Masson, V. (2011). The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3), 976–991.

Simmons, A. J., C. Temperton, 1997: Stability of a two-time-level semi-implicit integration scheme for gravity wave motion. *Mon. Wea. Rev.*, 125, 600–615.

Staniforth, A. and Thuburn, J. (2012), Horizontal grids for global weather and climate prediction models: a review. *Q.J.R. Meteorol. Soc.*, 138: 1–26.