Aladin-NH/AROME dynamical core: status and possible extension to IFS

Pierre Bénard

Météo-France CNRM/GMAP
42 Av. G. Coriolis
F-31057 TOULOUSE CEDEX
pierre.benard@meteo.fr

ABSTRACT

The problems linked to the use of a so-called ”constant-coefficient” linear approach for solving the system of Euler equations with a semi-implicit scheme are discussed. The solutions adopted in the limited-area Aladin-NH model to circumvent these problems are presented. The resulting dynamical core seems to reach the desired level of robustness and accuracy for NWP purposes, and has thus be chosen for the future operational nonhydrostatic application AROME at Météo-France. Cooperations between the Aladin community (including Météo-France) on one hand, and the HIRLAM group and ECMWF on the other hand for a possible extension of their NWP systems to the set of Euler equations are examined.

1 Introduction

The models currently used for operational NWP purposes at ECMWF and at Météo-France are:

- the global hydrostatic model IFS,
- the global stretched hydrostatic model ARPEGE,
- the limited-area hydrostatic model Aladin.

The first one is used at ECMWF and the two last ones at Météo-France. The Aladin model is developped through an international cooperation involving European and North-African countries, and also used operationally in these countries. The dynamical cores of these models share very large parts in terms of code, with only small variations in the way this code is used operationally. Many options are implemented in the dynamical part of this common code and many of them are identical for all operational applications. For instance, all these operational applications use a spectral horizontal representation with a semi-Lagrangian (SL) semi-implicit (SI) time discretisation. However, some options are chosen different, e.g. ARPEGE has a stretched horizontal spherical geometry and a finite-differences vertical discretisation whereas IFS has a uniform spherical geometry and a finite-elements vertical discretisation.

All these operational models are built on a hydrostatic primitive equations (HPE) dynamical core, formulated with a hybrid mass-based terrain-following vertical coordinate: in effect, pressure-based coordinates are in fact mass-based coordinates in the HPE context, due to the direct link between the mass-density and the pressure vertical variations. Laprise (1992) noticed that formulating Euler equations (EE) in mass-based coordinates leads to a form of the dynamical system that is very similar to the form obtained with HPE systems. For shallow-atmosphere EE systems, the corresponding mass-based coordinate was found therein to be the hybrid hydrostatic-pressure terrain-following coordinate (see Wood and Staniforth, 2003 for more details on these aspects). After the publication of Laprise's paper, it was argued in the NWP community that using such a mass-based coordinate in the EE context would allow the possibility of transforming existing HPE models into EE models at minimum scientific and development costs, with the advantage of a direct and clean diagnostic of nonhydrostatic effects, all the other parts of the model being exactly unchanged.

The choice made in 1994 by the international Aladin team to build the dynamical core of Aladin-NH (the nonhydrostatic version of Aladin) on this principle can be viewed as a test of these attractive ideas. However, it will be seen herein that the initial allegation of a ”small scientific and development cost” revealed itself quite over-optimistic, at least in view of an application to small-scale forecasts.

As mentioned above, all current operational applications at ECMWF and Météo-France are based on the HPE system. Beside these existing applications, a NWP project (AROME) is being developed at Météo-France with an operational application targeted in 2008 with an horizontal resolution of 2.5 km. The AROME system will be composed of the nonhydrostatic dynamical core of Aladin-NH, the physical parameterisation package issued from the Meso-NH research model (Bougeault et al., 2001), and a 3D variational analysis system able to assimilate various data compatible with the considered time and space scales..

Most operational NWP systems with HPE dynamical cores in the world use numerical algorithms based on the relatively simple and efficient SI SL technique. Due to the moderate non-linearity of the medium-scale HPE system, a simple variant of the SI technique, consisting in a so-called ”constant-coefficients” approach is mostly used for HPEs, and namely in the Aladin HPE operational model. In section 2 a discussion on numerical schemes will show that the application of this very simple constant-coefficients SI technique is not totally straightforward for the EE system. The problems encountered in the first implementations in Aladin-NH and the solutions adopted in order to maintain this constant coefficient SI technique robust enough for NWP purposes are presented in section 3. In section 4, some preliminary results with a full physics package are presented to illustrate the capability of the model to reproduce fine-scale results of research models. In section 5, the possibility of an adaptation of the global ARPEGE and IFS models to EE equations is discussed, and some conclusions are given in section 6.

2 Discussion on numerical schemes for EE systems

In its most general acceptation (differing from the original one, as proposed for meteorological applications by Robert et al., 1972, RHT72 hereafter), the SI technique refers to a class of schemes in which the source terms of the dynamical system are separated into an arbitrary linear part, treated in a centred-implicit way, while the residual non-linear part is treated explicitly. In the original acceptation of RHT72, the linear system was implicitly assumed to be the tangent-linear system around a so-called reference state which had the properties of being stationary and horizontally homogeneous. Although this reference state was clearly understood as being arbitrary, the resulting class of schemes was more restrictive than in the above general definition.

Besides, some classes of partly implicit schemes which cannot be straightforwardly be included in the above class of SI schemes were proposed more recently. Generally, the aim of these schemes is to achieve a ”more implicit” treatment of the system than allowed by the SI technique.

In the following sub-section, we try to summarize and categorize the various implicit schemes commonly used in meteorology, in order to prepare the discussion about the numerical schemes implemented in Aladin-NH. Other classifications may be proposed, but the one presented hereafter seems the most relevant for the subsequent discussions.

2.1 Various classes of SI schemes

According to the above general definition of SI schemes, the evolution of a time-continuous system written symbolically as:

$$\frac{\partial X}{\partial t} = M.X \tag{1}$$

is time-discretized in the SI manner by:

$$\frac{\delta X}{\delta t} = (M - L^*).X + \overline{L^*.X}^t \qquad (2)$$

where X is the state vector, M is the system of source terms acting on X, L^* is the arbitrary linear system used for the SI partitioning, $(\delta X/\delta t)$ is the discrete time-derivative of X, and $\overline{[\ ]}^t$ is the implicit-centred time-average operator. The exact discrete meaning of $(\delta X/\delta t)$ and X, is not further detailed here, in order that the above formulation remains valid for both three time-levels (3-TL) and two time-levels (2-TL) schemes. In this formalism, the vector X and spatial operators M, L^* can indifferently be understood as spatially-continuous (for theory) or spatially-discretized (for applications).

The first conceptual distinction that can be made between SI schemes refers to whether the SI linear system L^* can be obtained as the tangent-linear operator of the initial system M around a given reference-state, noted X^*, or not. This is of course the case, by construction, in the original scheme introduced by RHT72, as described above. This is also the case for most of the SI schemes described in the literature, mainly because this provides an easy and efficient procedure to build a linear system L^* which is reasonably "close" to the initial system M. However, it will be shown below that this strategy, referred to as "standard linearization" procedure hereafter, is not necessarily always optimal.

A second important distinction between SI schemes refers to the nature of the coefficients of the linear system L^*. Three main classes are distinguished for the purpose of the present paper:

(i) The coefficients of L^* are constant in time and along horizontal directions

(ii) The coefficients of L^* are constant in time but not horizontally

(iii) The coefficients of L^* are nonconstant in time and horizontally

In the case of a standard linearization procedure, these three classes respectively correspond to a reference state X^* which is (i) stationary and horizontally homogeneous, (ii) stationary but non horizontally homogeneous, and (iii) nonstationary and non horizontally homogeneous.

The first class, which contains the original scheme of RHT72, but also most of SI schemes used for HPE applications is referred to as "constant-coefficient" SI (CCSI) schemes, the other classes being termed as non-CCSI schemes herein. The CCSI class of schemes was historically used (as in RHT72, and most subsequent HPE applications) because after separation of the vertical dependencies, it leads to a constant-coefficient 2D Helmholtz equation for each vertical eigenmode, which can be solved by very efficient direct methods, e.g. using standard spectral transforms.

SI schemes from the second and third class lead to nonregular Helmholtz equations which are more complicated to solve. The implicit equation cannot be solved by direct methods and iterative solvers must be used. A typical example of a SI scheme belonging to the class (ii) is given by the new form of the MC2 model, as described in Thomas et al. (1998). In this model, the surface geopotential of the reference state is taken as non horizontally homogeneous, but constant in time. Consequently, all metric terms involving the vertical coordinate become horizontally nonhomogeneous (although constant in time). This class of schemes requires iterative solvers as for the class (iii) because the resulting Helmholtz equation is not regular. A typical example of a SI scheme belonging to the class (iii) may be found in Skamarock et al. (1997). In this model, the potential temperature of the SI reference state is the one of the current state. The resulting Helmholtz equation thus has non-constant and horizontally nonhomogeneous coefficients, and must be solved by an iterative solver, as for the class (ii).

As a simple illustration of the difference between the CCSI and non-CCSI approaches, consider the following excerpt of the horizontal momentum equation:

$$\frac{\partial \mathbf{V}}{\partial t} = RT\nabla q \tag{3}$$

where q is the logarithm of the pressure and other notations are standard .

In the CCSI approach, the SI reference state could be chosen such as $[RT]_{\text{ref}} = [RT]^*$ = Const., and $\nabla q_{\text{ref}} = 0$. The tangent-linear model around this reference state writes:

$$\frac{\partial \mathbf{V}}{\partial t} = [RT]^*\nabla q \tag{4}$$

and the CCSI scheme writes:

$$\frac{\partial \mathbf{V}}{\partial t} = ([RT]^0 - [RT]^*)\nabla q^0 + [RT]^*\overline{\nabla q}^t \tag{5}$$

where the superscript ”0” refers to the current (explicitly known) state. The first rhs term is thus treated explicitly while the second rhs term is treated in a centred implicit way. The first rhs term represents a thermal non-linear residual for the selected partitioning.

In the non-CCSI approach, one could for instance, following Skamarock et al. (1997), chose the reference state such as $[RT]_{\text{ref}} = [RT]^0$ (the temperature in the current state) and $\nabla q_{\text{ref}} = 0$. The tangent-linear model around this reference state writes:

$$\frac{\partial \mathbf{V}}{\partial t} = [RT]^0\nabla q \tag{6}$$

and the non-CCSI scheme writes:

$$\frac{\partial \mathbf{V}}{\partial t} = [RT]^0\overline{\nabla q}^t \tag{7}$$

In this case there is no explicitly treated residual. It should be noticed however that even if there is no fully-explicit residual, this does not mean that the scheme is a centred implicit scheme: the scheme remains essentially semi-implicit because the factor $[RT]$ is still treated explicitly, due to the linear partition. By nature, for any product source term such as in this example, the SI technique results in terms with only one factor treated implicitly, due to the linearization procedure.

For a true centred implicit scheme, the time discretization would write:

$$\frac{\partial \mathbf{V}}{\partial t} = \overline{RT\nabla q}^t \tag{8}$$

2.2 Comparison of CCSI and non-CCSI schemes in view of the EE system

Of course, CCSI schemes result in simpler implicit problems, and therefore in possibly simpler and cheaper solution algorithms. Due to the separability of the problem in this case, the basis of vertical eigenmodes can be precomputed in the setup part of the model, and for each vertical eigenmode, the 2D regular Helmholtz equation can be solved trivially in the relevant spectral space (e.g. bi-Fourier harmonics for Cartesian geometries and spherical harmonics for spherical geometries). It thus appears that the spectral transform and the vertical eigenmodes decomposition act, in essence, as the (direct) solver for the implicit problem.

For non-CCSI schemes, the part of the evolution which is treated explicitly is potentially much smaller than for CCSI schemes, and an increased robustness is thus expected, but the implicit problem to be solved is more complicated, and requires more advanced mathematical techniques such as large non-symmetric solvers. It can then be argued than when CCSI schemes are suitable for a given practical purpose, the use of non-CCSI schemes for the same purpose appears as an undue complication, possibly resulting in undue overcosts. Although both pertaining to the class of SI schemes, CCSI and non-CCSI schemes lead in fact to very different algorithmic strategies, and changing a model from a constant-coefficients design to a nonconstant-coefficients design is quite a big work that would result in deep code changes.

The main disadvantage of CCSI schemes is that the magnitude of fully-explicit non-linear residuals may be quite large and then result in an intrinsically unstable behaviour. Moreover, Simmons et al.(1978) have outlined that when unstable, CCSI schemes fail to produce a forecast at a given range even by reducing the time-step, because the instability rate is roughly proportional to the length of the time-step in the small time-step limit, thus producing an unchanged amount of error at a given forecast range, regardless of the time-step length.

In spite of these limitations, CCSI schemes were successfully (and widely) adopted for NWP applications using HPE systems. A possible explanation for this success is that for scales in the domain of validity of the HPE system (that is, larger than roughly 10-50 km), horizontal variations were much smaller than vertical ones, and the choice of an horizontally homogeneous reference state was not so far from the reality at these scales. For instance, the typical magnitude of terrain slopes was less than one percent, resulting in very small metrics terms for terrain-following coordinates.

A way of thinking has emerged after Ikawa (1988), stating that the CCSI technique may not be suitable for EE system, namely due to the presence of large explicitly treated ”cross-terms” resulting from the orography. Côté et al. (1998) and Cullen (2001) among others, used this argument to explore and evaluate alternatives to CCSI schemes. We think that this statement, although true for very small scales, should be temperated when applied to mesoscales. For instance, the Aladin-NH model was found able to successfully operate real-case forecasts with a CCSI scheme down to 2.5 km resolutions. However, as outlined in Bénard (2003) and Bénard et al.(2005), we agree with previous authors that the ability of CCSI schemes to solve EE systems at mesoscales may not reasonably be extrapolated to significantly smaller scales (e.g. less than 1 km), because the problem of steep orography and associated metric terms would then become dominant. As a consequence, it is stated here, in agreement with the literature, that pure CCSI schemes may not be robust enough for fine scale NWP modeling with EE systems.

Therefore, alternative solutions must be found to overcome this problem. Two main types of solutions have been proposed in recent years: the first one is simply to use non-CCSI schemes as discussed above. Several models have been built on the non-CCSI strategy: Skamarock et al. (1997), the current UKMO unified model (Cullen et al., 1997), the current version of MC2 model (Thomas et al., 1998) and the NCSU model (Qian et al., 1998). In this approach, thermal and/or orographic terms are linearized around the actual state, and the resulting implicit problem is solved through a matrical solver for the whole 3D Helmholtz equation, generally based on generalized conjugate residual algorithms. The second solution for increasing the robustness but without abandoning the advantages of the ”constant coefficients” strategy, would be to apply alternative schemes belonging to the class of ”constant coefficients” iterative centred implicit (ICI) schemes instead of the traditional CCSI scheme. These constant-coefficients ICI schemes are described in more details in the next subsection.

2.3 Constant-coefficients ICI schemes

In ICI schemes, an arbitrary first guess for the (future) unknown state $X^+_{(0)}$ is chosen, then (1) is solved iteratively for the index k by:

$$\frac{\delta X_{(k+1)}}{\delta t} = \overline{M.X_{(k)}}^t + \frac{L^*}{2}.(X^+_{(k+1)} - X^+_{(k)}) \tag{9}$$

where $(\delta X_{(k+1)}/\delta t)$ is the time-discrete time-derivative computed using the state $X^+_{(k+1)}$ as the future state, and $\overline{M.X_{(k)}}^t$ is a time-average of the model sources using the state $X^+_{(k)}$ as the future state. This rewrites:

$$\frac{\delta X_{(k+1)}}{\delta t} = \overline{(M - L^*).X_{(k)}}^t + \overline{L^*.X_{(k+1)}}^t \tag{10}$$

with similar notations. When the convergence is achieved, the last term in (9) obviously vanishes, and the original system (1) is then solved in a fully centred-implicit way (also sometimes called "trapezoidal" scheme):

$$\frac{\delta X}{\delta t} = \overline{M.X}^t \tag{11}$$

The advance in time of the model is thus seen to consist in two nested loops: an "inner" loop on the index k, to find the next state vector X^+ at each time-step, and the classical temporal "outer" loop for the actual advance in time. The loop on the index k is stopped after a given number of iterations $k_{\max}$. The initial guess $X^+_{(0)}$ is generally taken as the state resulting from a classical semi-implicit scheme, and $k_{\max}$ is then the number of additional iterations subsequent to the classical SI scheme. In practice, it is found that for meteorological fine-scale applications with the EE system, a single additional iteration $k_{\max} = 1$ is sufficient to guarantee the robustness of the scheme. However, it will be seen below that such a fast convergence needs a certain care to be taken.

A further examination of (9) shows that the ICI scheme acts like a pre-conditionned fixed-point algorithm for iteratively approaching the target trapezoidal scheme. From this point of view, the operator L^* is just a pre-conditioner which allows or accelerates the convergence toward the target scheme. The output of the scheme is not very dependent on the particular choice of the operator L^* provided that the scheme is operated with a reasonable convergence.

If the L^* operator is constant in time and horizontally homogeneous, the ICI scheme is then referred to as a "constant-coefficients" ICI (CC-ICI) scheme similarly to the terminology adopted above for SI schemes. CC-ICI schemes have been advocated and used by the GEM model community (Côté et al., 1998). Further discussions on the behaviour of these schemes may be found e.g. in Cullen (2001), Bénard (2003), and Cordero et al. (2005).

It should be noticed that CC-ICI schemes do not need any newer technology for solving the implicit problem at each iteration k than the one employed in classical CCSI schemes, and thus the conversion of a CCSI model to a CC-ICI model is quite straightforward from the code point of view. This approach seems to be better suited than the non-CCSI approach for spectral models, since, as outlined above, the horizontal solver is readily available through the spectral transform after the separation of vertical eigenmodes is achieved. It could also be argued that provided one or very few iteration are sufficient for a satisfactorily convergent result, CC-ICI schemes offer an efficient alternative to non-CCSI schemes, for which the non-symmetric solver requires a large number of iterations before convergence, possibly resulting in an undue overcost; however this key question is not yet fully resolved.

Both ICI and non-CCSI classes of schemes share the property of being iterative by nature. Although they are expected to be quite robust as discussed above, they are not unconditionally stable. In particular, if the time-step is chosen too big, the iterative solver of a non-CCSI scheme will not converge, and the model will fail to be a SI model. Similarly, for too big time-steps in a CC-ICI scheme, the iterative fixed-point algorithm will not converge and the model will fail to be a trapezoidal model. In both cases, perturbations can freely develop in this unperfect context, and lead to the divergence of the model forecast.

3 Background and status of Aladin-NH/AROME dynamical core

A first version of the "Aladin-NH" EE dynamical core was implemented in 1995 (Bubnová et al., 1995). It consisted of a CCSI Leap-Frog three-time levels (3-TL) Eulerian model. Since one of the cross-terms in the evolution sources was found responsible of a large instability in presence of orography, a "mixed" CCSI/CC-ICI was built. The principle of this mixed scheme is described here: the M model was partitioned into the unstable source term (noted $\mathbf{m}$) and the remaining part, $M' = M - \mathbf{m}$. The equation system to be solved (1) then rewrites symbolically:

$$\frac{\partial X}{\partial t} = M'.X + \mathbf{m}.X \tag{12}$$

This system was time-discretized with an iterative algorithm as follows:

$$\frac{\delta X_{(k+1)}}{\delta t} = (M' - L^*).X + \overline{L^*.X_{(k+1)}}^t + \overline{\mathbf{m}.X_{(k)}}^t \tag{13}$$

with similar notations as in (10). The first guess $X^+_{(k)}$ was given by the classical CCSI scheme, and only one additional iteration was performed after this first guess ($k_{\max} = 1$). It is seen from the above equation that after convergence is achieved, $(L^* + \mathbf{m})$ is treated implicitly while $(M' - L^*) = M - (\mathbf{m} + L^*)$ is treated explicitly.

The advantage of this scheme was that due to the small amount of terms in $\mathbf{m}$ (that is, only one), the iterative algorithm was very computationally cheap, and converged very rapidly. This procedure allowed to obtain an implicit treatment of the leading problematic cross-term $\mathbf{m}$ at the price of only a 20% overcost on the dynamical core, compared to the simple SI scheme. The 3-TL Eulerian version of the model was satisfactorily stabilized with this modification, at least for meso-scale applications (Δx larger than $\sim 5km$).

However, in 1998, an attempt to extend this dynamical core for SL schemes, with the corresponding increase of time-step length, was unsuccessful: the 3-TL SL version was moderately, but invariably unstable, while the 2-TL SL version was dramatically unstable. A detailed investigation of these problems was initiated, and, in parallel, the decision to build a complete CC-ICI scheme [in opposition to the "mixed" existing scheme described in (13)] was taken, since CC-ICI schemes are expected to be very robust.

The moderate instability of the 3-TL SL model was found to be present even for resting flows without orography, and thus lending itself quite easily to analysis; the general principle of the analysis is presented in Bénard (2003). The instability was shown to be similar in nature to the one described by Simmons et al. (1978): the scheme is unstable in presence of thermal explicitly-treated nonlinear residual. However, whereas the HPE system could be stabilized by an appropriate choice of the CCSI reference temperature, this was not the case for the EE system.

Besides, the behaviour of the instability was found to be highly dependent on the choice of the vertical coordinate and of the prognostic variables (Bénard et al., 2004), especially the additional nonhydrostatic prognostic variables needed for the solution of the EE system. This study showed that the initial choice for the nonhydrostatic prognostic variables was directly responsible for the moderate instability of the 3-TL SL version. A first proposal was to replace the original nonhydrostatic prognostic variables of Bubnová et al. (1995):

$$\hat{\mathscr{P}} = \frac{p - \pi}{\pi^*} \tag{14}$$

$$\hat{d} = -g \frac{\pi^*}{m^* R T^*} \frac{\partial w}{\partial \eta} \tag{15}$$

by:

$$\mathscr{P} = \frac{p-\pi}{\pi} \tag{16}$$

$$\mathsf{d} = -g\frac{\pi}{mRT}\frac{\partial w}{\partial \eta} \tag{17}$$

where p is the pressure, π the hydrostatic pressure and $m = (\partial\pi/\partial\eta)$ where η is the hybrid coordinate. The star denotes values in the CCSI reference state (T^* is a pure constant).

This modification did of course not remove the unstable thermal nonlinear residuals (which, as seen above, are inherent to the CCSI approach), but their structure was modified in such a way that the scheme becomes stable for an appropriate choice of the CCSI reference state temperature.

This modification allowed to stabilize the 3-TL SL SI version of the model, however it was experienced that the stability, although now compatible with successful forecasts, was poorer in the EE version than in the HPE version in presence of orography, and especially with steep orography. An analysis of this problem in a context of uniform slope (Bénard et al., 2005), showed one more time that the choice of prognostic variables could have an impact on the stability in presence of slope. A new prognostic variable $\mathbb{d}$ defined by:

$$\mathbb{d} = -g\frac{\pi}{mRT}\frac{\partial w}{\partial \eta} + \frac{p}{mRT}\frac{\partial \mathbf{V}}{\partial \eta}.\nabla\phi \tag{18}$$

was proposed in replacement of d, where $\mathbf{V}$ is the horizontal wind vector and ϕ is the geopotential. However, the analysis showed that this change of variable, although alleviating the instability linked to the orography, was not able to remove it completely in the considered context. This contrasts with the flat-terrain case, for which the replacement of $\hat{d}$ by d removed the corresponding instability in the analytical context. As a consequence, a change from CCSI to CC-ICI schemes was recommended, since ICI schemes are not subject to this instability, provided that convergence is achieved. Practical experiment confirmed these views by showing that with a 10 km resolution, the use of an ICI scheme is not actually required, but for a resolution of 2.5 km, the use of an ICI scheme, even with only one additional iteration, allows a significant extension of the time-step length.

In parallel to this work, the dramatic instability of the CCSI 2-TL scheme for the EE system was examined. Facing this instability, Semazzi et al. (1995) and Qian et al. (1998) had recourse to a very large time-decentering, but, as they mentioned, this stabilizing process resulted in a loss of accuracy incompatible with exploitable forecasts. The origin of the problem was identified in Bénard (2004): 2-TL CCSI schemes have, by nature, a domain of stability which is inherently smaller than their 3-TL counterparts. Unlike what could intuitively expected, this reduced domain of stability is not related to the time-extrapolation usually employed in 2-TL scheme. This can be checked through the fact that nonextrapolating 2-TL schemes (where the extrapolation is replaced by a simple, first-order accurate in time, explicit evaluation) exhibit the same instability than extrapolating schemes. This reduced stability domain for 2-TL schemes was already pointed out for the HPE system in Simmons and Temperton (1997), and the solution proposed at this time was to choose a warmer CCSI reference temperature T^* for 2-TL schemes than for 3-TL schemes. However this simple solution does not extend to the EE system, because the domains of stability for vertical acoustic waves and for gravity waves do simply not overlap when a 2-TL scheme is used, as shown in Bénard (2004). However, the analysis suggested another solution. This consists in choosing a reference value T_e^* (different from T^*) for the linearisation of the terms responsible of vertically-propagating elastic waves (the others terms being still linearised with T^*). When this solution is applied, the resulting CCSI scheme no longer belongs to the class of SI schemes with a "standard linearization procedure" (in the sense introduced in section 2) because of the existence of two reference temperatures (T^*, T_e^*) instead of a single one. In this particular case, relaxing the constraint of a standard linearization procedure thus allows to dramatically increase the robustness of the scheme. It is worth noticing however that this very simple modification does not result in any overcost or loss of accuracy.

As discussed in section 2.3, the linear system L^* used for the CCSI scheme is used identically (at least in the version implemented in Aladin model) as a pre-conditioner for the corresponding CC-ICI scheme, i.e., when

k_{max} is chosen non-zero. All our abovementioned analyses in academic contexts show that the robustness of the CC-ICI scheme for small values of k_{max} is directly linked to the robustness of the corresponding CCSI corresponding scheme (with $k_{max} = 0$). Moreover, CCSI schemes with a poor robustness have also been found to exhibit a poor convergence of the inner iterative algorithm when the same linear system is used to build a CC-ICI scheme. As a consequence, we believe that even when a CC-ICI approach is chosen for the design of a EE model, the effort needed to prealably build a consistent and quite robut CCSI scheme is worth being done, since this effort will finally be benefcial for the CC-ICI system itself. Although this statement cannot of course be formally proved in the case of a general nonlinear system, its relevance was experienced empirically in all our validations.

Finally, the approach proposed for the operational AROME target in 2008 is to use a 2-TL SL CC-ICI ($k_{max} = 1$) with the nonhydrostatic prognostic variables $(\mathscr{P}, \mathbb{d})$ and relaxing the standard linearization procedure as decribed above. This strategy allows, in our opinion, a robust formulation together with simple and well-controlled technologies and a very competitive efficiency.

4 Preliminary results with full physics

Various academic tests have been performed in order to validate the robustness and the accuracy of the dynamical core of Aladin-NH. For these academic tests, an adiabatic version of the Aladin-NH model was used in three-dimensional (3D) and two-dimensional (2D) vertical plane frameworks. For non-linear flows, a horizontal diffusion is added, in order to prevent an accumulation of energy in the smallest resolved scales. Since the semi-Lagrangian dynamics itself had been extensively validated in the HPE context, the focus for these validations has been put on the robustness of the temporal scheme and the accuracy of the response for orographically forced flows in various regimes. These academic tests included:

- resting flows over an orography in presence of large thermal residuals,
- classical 2D and 3D stationary flows over an orography in a stably stratified atmosphere with a uniform background flow,
- 3D orographic flows with rotation, over a real steep orography in the Alpine region at resolutions up to 1.25 km.

Besides, semi-academic cases with physics also have been performed. The aim of these tests was to evaluate the capability of the dynamical core to reproduce phenomena for which the dynamics and the physics play an equally important role. These tests included:

- non-stationary orographically-induced 2D trapped lee waves in an atmosphere with a critical level (Keller, 1994). This test requires a ”dry” physics to be included via a parameterization of vertical turbulent exchanges.
- idealised 2D squall line triggered by an initial perturbation in a tropical profile, with a ”minimal” physics (turbulent exchanges, microphysics, and prescribed surface fluxes).

For all these tests, an analytical or a numerical reference existed and was to be reproduced. The results for these tests were successful from this point of view and thus are not discussed in more details here.

To illustrate concretely the capability of the dynamical core of Aladin-NH for the AROME operational target to give competitive forecasts, we present some results of a flash flood episod in September 2002 over the Nîmes area (South-East part of France), obtained with the AROME forecast model prototype (i.e. Aladin-NH dynamical core and Meso-NH physics package). The forecast starts from the Aladin operational analysis of 8

September 2002 at 1200Z, the coupling information at lateral boundaries is provided every 3 hours (and linearly interpolated in time in between). The results are examined after a 12 hours forecast.

The resolution is 2.5 km, and the domain contains 180 $\times$ 180 grid points and 41 vertical levels. The results of the AROME prototype are compared to the validated results of the Meso-NH research model. The cumulated rainfall over the 12 hours of the forecast is depicted in Fig.1 for the AROME prototype and in Fig.2 for the reference forecast of the Meso-NH research model. The two forecasts are qualitatively similar and the maximum amount of precipitation compares well, in intensity and location. Both models also compare well to the observed evolution which led to a quasi-stationary convective line with heavy rains above the orography. The maximum cumulated rainfall is 239 mm for AROME, and 246 mm for Meso-NH. Other fields are also found to be similar between the two forecasts and to compare well to the observation (not shown).

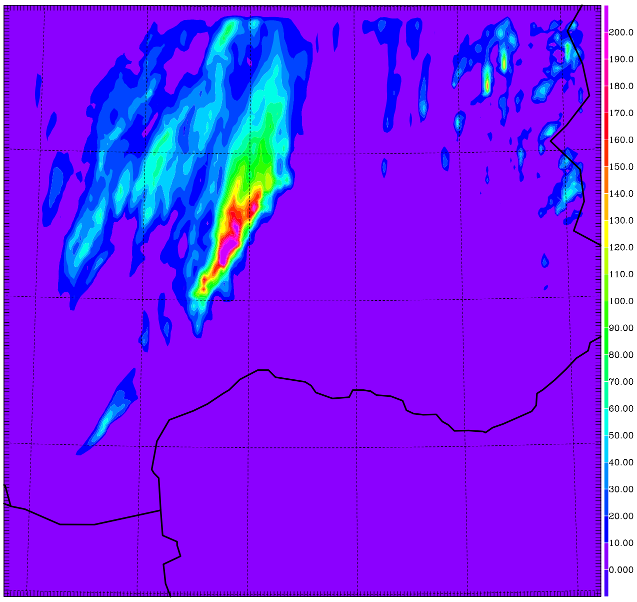

Fig 1: Cumulated rainfall over 12 hours for the real-case described in the text, with the AROME prototype (using the Aladin-NH dynamical core). Coastlines and borders are drawn in solid line.

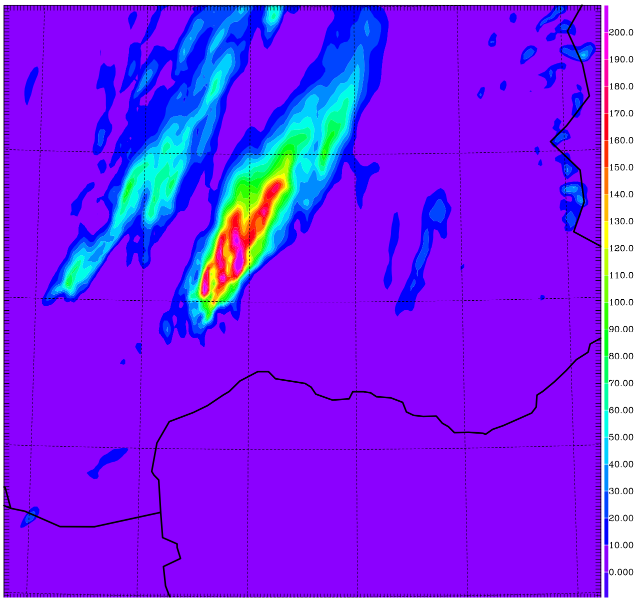

Fig 2: Cumulated rainfall over 12 hours for the real-case described in the text, with the Meso-NH research model. Coastlines and borders are drawn in solid line.

The time scheme used in the AROME prototype for this forecast is a 2-TL ICI scheme with 1 extra iteration after the SI first guess (i.e. $k_{max} = 1$ in the above notation). The computational cost of the dynamical part of the forecast is therefore twice as expensive as for a pure SI scheme. However, the practical impact of this overcost must be temperated by the fact that the computations for the parameterized physics need not be iterated. Since it is anticipated that in the future, a large part of the computational cost of the forecast model will be devoted to the parameterized physics, we believe that the increased robustness brought by using an ICI scheme instead of a SI scheme is worth the resulting computational overcost.

The combination of an ICI time discretisation with a SL scheme allows an efficient integration of the equations. As an illustration, the computational time for this forecast with the Meso-NH research model (anelastic system, 3-TL explicit Eulerian scheme) was 24.33 hours, whilst for the AROME prototype (EE system, 2-TL ICI SL scheme) it was reduced to 2.5 hours, with the same physics package. This difference in efficiency can be explained by the length of the time step: the maximum time step allowed in Meso-NH for this forecast was 4s, but could be increased up to 60s in the AROME protoype, for a similar level of quality.

5 Extension of IFS to the EE system

The EE dynamics of Aladin-NH is currently implemented only in the limited-area context. However, with the resolutions allowed by the increasing computational power, the relaxation of the hydrostatic approximation in global models will soon have to be considered. The stretched version of the French operational global model ARPEGE will probably be concerned first, but the uniform-resolution global model IFS of the ECMWF will inevitably follow sooner or later.

From a pure theoretical point of view, the extension of the EE system from limited-area to global models is not likely to cause serious problems: global models usually have a smoother orography and more regular fields than their limited area counterparts, due to the difference in resolutions.

Hence the extension of the IFS/ARPEGE model to the EE system seems mostly a matter of technical work. The most significant part of this technical work will be to unify the formulation of the semi-implicit schemes in IFS/ARPEGE and Aladin: for ARPEGE and IFS models, the SI inversion is formulated as the product of three operators: the transform to the space of vertical eigenmodes $\mathbf{Q}$, the inversion of the set of separated horizontal 2D Helmholtz operators, matrically noted $\mathbf{H}$, then the inverse transform to the vertical physical space $\mathbf{Q}^{-1}$. The operators $\mathbf{Q}$, $\mathbf{H}$ and $\mathbf{Q}^{-1}$ are precomputed in the setup part of the model and stored separately. The inverse of the Helmholtz operator $\mathbf{H}^{-1}$ is computed at each time-step. The global SI inverse operator $\mathbf{Q}.\mathbf{H}^{-1}.\mathbf{Q}^{-1}$ is then recomposed at each time-step and applied to the explicit rhs vector of the implicit problem. For the Aladin model the global SI inverse operator $\mathbf{Q}.\mathbf{H}^{-1}.\mathbf{Q}^{-1}$ is precomputed in the setup part of the model, and stored as a whole. At each time-step, this global inverse operator is directly applied to the explicit rhs vector of the implicit problem. The method used in Aladin slightly minimizes the computations at each time-step, but leads to a larger storage than the method used in ARPEGE and IFS. Moreover, the method used in ARPEGE and IFS allows an exact treatment of the map-factor in the SI scheme (while the map-factor has to be linearized to a constant unit value in the SI formulation of Aladin). The exact treatment of the map factor in the SI scheme is an important ingredient for the robustness of the stretched ARPEGE model (Yessad and Bénard, 1996), and could become also beneficial for very large domains in Aladin-NH for which the map-factor would significantly depart from the unit value. The unification of the algorithms of IFS/ARPEGE and Aladin SI schemes would therefore consist in extending the method used in IFS/ARPEGE to the Aladin model, rather than the inverse solution.

The point which requires most of the scientific work for the extension of IFS to the EE system is the vertical discretisation: Aladin and ARPEGE are formulated in finite-differences while IFS uses a finite-elements discretisation. The scheme used in IFS is more accurate and has more realistic normal-modes, especially in the stratosphere (Untch and Hortal, 2004), hence it is thus wishable that the finite-elements vertical discretisation could be kept for the integration of the EE system. However changing the vertical discretisation of an EE system with a SI or ICI scheme is not a trivial operation, since some mathematical or physical constraints of the continuous system should be also fulfilled by the discrete system. The physical constraints of global energy or angular momentum conservation in the absence of sources can simply be ignored for a short or medium NWP model (and indeed they are ignored by the current finite-elements discretisation of IFS). One of the justifications for this is that the increased accuracy linked to using the finite-elements discretisation inherently improves the conservation of these dynamical global invariants. However the status of the mathematical constraints termed (C1) and (C2) in Bubnová et al. (1995) is less flexible: for instance, the constraint (C1) must be fulfiled by the discrete vertical integral operators in order that the SI scheme keeps its current algorithmic structure (algebraic elimination of all prognostic variables but one in the implicit linear system, and solution of the resulting Helmholtz equation for this variable, followed by a back-substitution to recover the other prognostic variables). Assuming that the structure of the SI scheme is left unchanged, the constraint (C1) must thus be fulfilled by the finite-elements discretisation, which is not the case with the finite-elements scheme of the current HPE system in IFS. Various solutions to achieve this goal may be considered, and the scientific task will consist in examining the advantages and disadvantages of these solutions in order to choose the best-suited final formulation. The status of the constraint (C2) will have to be examined in a similar way.

6 Concluding remarks

The initial version of Aladin-NH, such as documented in Bubnová et al. (1995) suffered from a lack of robustness which can be viewed as originating from the direct extension of the ”constant coefficients” approach for the SI scheme from the HPE system to the EE system. The CC-SI technique applied to the HPE system already suffered from a lack of robustness in some conditions involving the reference state, as shown by Simmons et al. (1978), but when this CCSI technique is applied to the EE system, the lack of robustness may come from more various origins (such as the choice of prognostic variables or the restricting choice of a ”standard linearization procedure” as described above), and may result in much more severe instabilities, as shown by Bénard (2003), Bénard et al. (2004), Bénard (2004) and Bénard et al. (2005).

However the latter series of papers also shows that the ”constant coefficient” approach can still be considered as suitable for solving the EE system, provided that a relevant choice is made for the prognostic variables, that a more general approach for choosing the implicit linear system is adopted, and that an iterative procedure is applied to achieve a ”more implicit” ICI scheme, in which even the non-linear source terms are treated implicitly (or at least with some degree of implicitness if small values of k_{max} are chosen). These studies also seemingly demonstrate that there is no special advantage in using a height-based coordinate instead of a mass-based coordinate for solving the EE system.

The dynamical core of Aladin-NH was modified according to these guidelines and was consequently found to have acquired a robustness largely compatible with NWP purposes at the kilometric scale. The changes involved by these modifications are quite deep and the resulting dynamical core may be viewed as a new dynamical core compared to the one described in Bubnová et al. (1995).

The dynamical core of Aladin-NH has been officially chosen as the dynamical core of the future Météo-France operational AROME model, coupled with the physical parameterisation package of the Meso-NH research model and a 3D variational analysis system able to assimilate various fine-scale data.

From the point of view of the project perspectives, the dynamical core of Aladin-NH is planned to be involved in two projects of international cooperations: the HIRLAM group would initiate a cooperation with Aladin in order to implement the Aladin-NH dynamical core in their forecast system, in view of future operational NWP using the EE system. However, the current geometry of Aladin (plane Stereo-Lambert or Mercator projection) is not identical to the geometry chosen in HIRLAM (spherical geometry with a rotated origin). This difference mainly comes from the fact that HIRLAM applications must be able to work in large domains containing one of the geographical poles, a constraint which was not essential in the initial design of the Aladin project.

Consequently, this cooperation implies to find a geometry which would be compatible with the constraints of the two systems. For this, it has been decided to chose a plane Mercator projection but with a rotated origin, which meets the requirements of being very close from both current geometries of HIRLAM and Aladin, and to allow an easy implementation in the two systems. Moreover, this geometry would have the advantage of allowing an exact inclusion of the map factor in the SI scheme (similarly to what was documented for the global context in Yessad and Bénard, 1996), an ingredient which could become necessary when perfoming forecast in large domains, as discussed above. The second possible cooperation would be with the ECMWF, and would consist in extending the EE system to the global model (stretched or not). As indicated above, the key point for this cooperation is the possibility to extend the current finite-differences vertical discretisation of Aladin-NH to the finite-elements vertical discretisation of IFS.

From the point of view of other scientific development perspectives, the implementation of a deep-atmosphere version (thus relaxing the current shallow-atmosphere approximation) of Aladin-NH is planned. As shown by Wood and Staniforth(2003) this extension can be realized for the EE system in a clean and natural way in mass-based coordinates, contrasting with the HPE system, where this extension was less natural and required some approximations (White and Bromley, 1995)

Acknowledgements

The author is grateful to the whole international Aladin-NH team (especially Jean-François Geleyn, Radmila Brozková, Jozef Vivoda, Jan Mašek, Petra Smolíková, Alena Trojakova and Gwenaëlle Hello) for their appreciated cooperation in the design, development and validation of the Aladin-NH dynamical core. Yann Seity and Sylvie Malardel, from the AROME project team are also gratefully aknowledged for developing the AROME prototype, performing semi-academic and real-case forecasts including those mentioned in section4, and for providing the corresponding validation results.

References

Bénard, P., 2003: Stability of Semi-Implicit and Iterative Centred-Implicit Time Discretisations for Various Equation Systems Used in NWP. *Mon. Wea. Rev.*, **131**, 2479-2491.

Bénard, P., R. Laprise, J. Vivoda and P. Smolíková, 2004: Stability of the leap-frog constant-coefficients semi-implicit scheme for the fully elastic system of Euler equations. Flat terrain case. *Mon. Wea. Rev.*, **132**, 1303-1318.

Bénard, P., 2004: On the use of a wider class of linear systems for the design of constant-coefficients semi-implicit time-schemes in NWP. *Mon. Wea. Rev.*, **132**, 1319-1324.

Bénard, P., J. Mašek and P. Smolíková, 2005: Stability of Leap-Frog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations. Case with Orography. *Mon. Wea. Rev.*, **133**, (accepted).

Bougeault, Ph., and Co-authors, 2001: The Meso-NH Atmospheric Simulation System: Scientific Documentation. *Available on internet at:*
`http://www.aero.obs-mip.fr/mesonh/dir_doc/book1_14dec2001/`

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

Cordero, E., N. Wood and A. Staniforth, 2005: Impact of semi-Lagrangian trajectories on the discrete normal modes of a non-hydrostatic vertical column model. *Q. J. R. Meteorol. Soc.*, **131**, (accepted).

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998: The Operational CMC-MRB Global Environmental Multiscale (GEM) Model. Part I: Design Considerations and Formulation. *Mon. Wea. Rev.*, **126**, 1373-1395.

Cullen, M. J. P., T. Davies, M. H. Mawson, J. A. James and S. C. Coulter, 1997: An overview of numerical methods for the next generation UK NWP and climate model. *Numerical Methods in Atmospheric Modelling*, Canadian Meteorological and Oceanographical Society, 581 pp.

Cullen, M. J. P., 2001: Alternative implementations of the semi-Lagrangian semi-implicit schemes in the ECMWF model. *Q. J. R. Meteorol. Soc.*, **127**, 2787-2802.

Ikawa, M., 1988: Comparison of some schemes for nonhydrostatic models. *J. Meteorol. Soc. Jap.*, **66**, 753-776.

Keller, T. L., 1994: Implications of the hydrostatic assumption on atmospheric gravity waves. *J. Atmos. Sci.*, **51**, 1915-1929.

Laprise, R., 1992: The Euler equations of motion with hydrostatic pressure as an independent variable. *Mon. Wea. Rev.*, **120**, 197-207.

Qian, J.-H., F. H. M. Semazzi, and J. S. Scroggs, 1998: A global nonhydrostatic semi-Lagrangian atmospheric model with orography. *Mon. Wea. Rev.*, **126**, 747-771.

Robert, A. J., J. Henderson, and C. Turnbull, 1972: An implicit time integration scheme for baroclinic models of the atmosphere . *Mon. Wea. Rev.*, **100**, 329-335.

Semazzi, F. H. M., J. H. Qian, and J. S. Scroggs, 1995: A global nonhydrostatic semi-Lagrangian atmospheric model without orography. *Mon. Wea. Rev.*, **123**, 2534-2550.

Simmons, A. J., B. Hoskins, and D. Burridge, 1978: Stability of the semi-implicit method of time integration. *Mon. Wea. Rev.*, **106**, 405-412.

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.

Skamarock, W. C., P. K. Smolarkiewicz, and J. B. Klemp, 1997: Preconditioned conjugate-residual solvers for Helmholtz equations in nonhydrostatic models. *Mon. Wea. Rev.*, **125**, 587-599.

Thomas, S.J., C. Girard, R. Benoit, M. Desgagné, and P. Pellerin, 1998: A new adiabatic kernel for the MC2 model. *Atmos. Ocean*, **36 (3)**, 241-270.

Untch, A., and M. Hortal, 2004: A finite-element scheme for the vertical discretization of the semi-Lagrangian version of the ECMWF forecast model. *Q. J. R. Meteorol. Soc.*, **130**, 1505-1530.

White, A. A, and R. A. Bromley, 1995 : Dynamically consistent, quasi-hydrostatic equations for global models with a complete representation of the Coriolis force. *Q. J. R. Meteorol. Soc.*, **121**, 399-418.

Wood, N. and A. Staniforth, 2003: The deep-atmosphere Euler equations with a mass-based vertical coordinate. *Quart. J. Roy. Meteor. Soc.*, **129**, 1289-1300.

Yessad, K. and P. Bénard, 1996: Introduction of a local mapping factor in the spectral part of the Meteo-France global variable mesh numerical forecast model *Quart. J. Roy. Meteor. Soc.*, **122**, 1701-1720.