Stability of Constant and Variable Coefficient Semi-Implicit Schemes for the Fully Elastic System of Euler Equations in the Case of Steep Slopes

Thomas Burgot, Ludovic Auger, Pierre Bénard

Stability of constant and variable-coefficients semi-implicit schemes for the fully elastic system of Euler equations in case of steep slopes

Thomas Burgot,[a] Ludovic Auger,[a] Pierre Bénard,[a]

[a] *CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France*

Corresponding author: Thomas Burgot, thomas.burgot@meteo.fr

ABSTRACT: Constant and variable-coefficients schemes are studied to improve numerical stability on the steepest slopes of the relief encountered at hectometric scales in Numerical Weather Prediction models. Stability analyses are conducted on the Iterative Centered Implicit temporal scheme which approaches the Crank-Nicolson scheme. These analyses are led for the fully elastic system of Euler equations for different slopes and different thermal residuals. They are able to reproduce the maximum slopes currently encountered in real hectometric models for which the simulation is numerically stable. Because of the negligible price of these analyses, several strategies can thus be easily tested. No strategy among these considered for constant coefficient schemes improves significantly numerical stability without worsening efficiency or quality. Hence, constant coefficient schemes are probably not the most suitable schemes for high-resolution computing. A successful strategy consists of using the same features of constant coefficient schemes except for the orographic terms which are implicitly treated, resulting on a variable coefficient scheme. In this case, slopes up to 70° can be easily reached, even in case of a strong thermal residual. Since estimates on the condition number of the implicit problem containing orographic terms remains low even in case of steep slopes, the implicit problem should be easily inverted by an iterative solver.

1. Introduction

The dynamical core of a Numerical Weather Prediction (NWP) model is based on the temporal integration of the partial derivative equations of the form:

$$\frac{\partial \Psi}{\partial t} = \mathcal{M}(\Psi), \tag{1}$$

where $\mathcal{M}$ is a non-linear operator which represents dynamical equations, typically the fully elastic system of Euler equations, Ψ is the associated state vector containing the prognostic variables, and t is the time. To forecast small-scale phenomena, some operational Limited Area Models (LAM) now reach horizontal resolution around one kilometer. Some of them are HEVI (Horizontally Explicit Vertically Implicit) models like the ICON model (Zängl (2012)), others are semi-implicit semi-Lagrangian (SI-SL) like AROME of Météo-France (Seity et al. (2011)) or the UM of the Met-Office (Lean et al. (2008)). For forecasting phenomena of even finer scales, one of the objectives in the NWP domain in the coming years, is to continue refining the horizontal resolution by running models at hectometric resolution. At this resolution, the orography is better represented leading to an increase of the steepness of orographic slopes. These conditions can conduct to various numerical instabilities regardless of the class of atmospheric models considered. For example, in these conditions, the elliptic pressure solver fails to converge for the anelastic model Méso-NH (Lac et al. (2018)), and numerical instabilities occur in the AROME model for an horizontal resolution of 300 m corresponding of slopes around 50° over the Alps, leading to a useless forecast.

This study focuses on semi-implicit schemes and are now described. These schemes have been introduced by Robert et al. (1972). Today, most of the semi-implicit schemes approach the 2-TL (Time Levels) Crank-Nicolson's scheme through an Iterative Centered Implicit scheme (ICI) with one or two steps. Once discretized with an ICI scheme, the system (1) becomes:

$$\frac{\Psi^{+(n)} - \Psi^{0}}{\Delta t} = \frac{\mathcal{L}^{*}\left(\Psi^{+(n)} + \Psi^{0}\right)}{2} + \frac{\mathcal{R}(\Psi^{+(n-1)}) + \mathcal{R}(\Psi^{0})}{2}, \tag{2}$$

where n is the n-th step of the ICI scheme, Ψ^{0} and $\Psi^{+(n)}$ are respectively the state of the system at time t and the state of the system at time $t + \Delta t$ after n iterations of the corrector step, $\mathcal{L}^{*}$ is a linear operator, and $\mathcal{R} = \mathcal{M} - \mathcal{L}^{*}$ is the non-linear residual. This scheme can be initialized either

by a non-extrapolated scheme $\Psi^{+(0)} = \Psi^0$, or by an extrapolated one: $\Psi^{+(0)} = 2\Psi^0 - \Psi^-$ (where Ψ^- is the state of the system at time $t - \Delta t$). For example the current version of the AROME model uses a non-extrapolated scheme with two steps ($N_{iter} = 2$, sometimes called predictor-corrector scheme, and abbreviated ICI-2TL-PC in the following), while the previous version used only one step ($N_{iter} = 1$) with an extrapolation (abbreviated SI-2TL-E in the following).

The linear operator $\mathcal{L}^*$ treats the terms responsible of the propagation of the fastest waves by solving an implicit Helmholtz problem, while non-linear residual terms are treated explicitly. This bypasses some of the strongest numerical stability constraints, and the equations are thus integrated with a large time step. To guarantee the effectiveness of the entire procedure, the linear operator is chosen by ensuring:

1. numerical stability,
2. the invertibility of the implicit problem,
3. the efficient solving of the implicit problem,
4. the convergence in a few steps of the ICI scheme.

A common way to determine the linear operator is to define it from the linear tangent of the non-linear operator, linearised around a basic state that has to be specified. This method avoids the difficult problem of choosing the operator, for a simpler problem of choosing the basic state. Many strategies have been studied over the last few decades to design the best basic state and consequently the best linear operator satisfying these four conditions, and are now discussed.

The basic state chosen, is the one of the previous time step, in the dynamical core of the Unified Model (Davies et al. (2005)). Then, the basic state is updated at each time step, avoiding to prescribe it arbitrarily. However, this strategy leads to some issues to invert the implicit problem. An output criterion of the iterative solver has thus been implemented in case of low-convergence problems (Davies et al. (2005)). Moreover, the implicit problem considered is not sparse because of its stencil of 45 points, raising scalability and efficiency issues (Benacchio and Wood, 2016). To circumvent these problems, some terms, notably orographic, have been removed from the linear operator and are now treated explicitly, when moving from the 'New Dynamics' to the 'End GAME' dynamical core in the Unified Model (Walters et al. (2017)).

On the contrary, another strategy is to consider idealised basic states. For example, an isothermal atmosphere, dry, at rest with no orography has been chosen as a basic state by Bénard (2003). Additional degrees of freedom can even be added to the system to better control its stability. For example, Bénard (2004), introduces a 'cold' temperature in the vertical momentum equation, generating a linear operator different from the linear tangent. When the basic state is chosen such as the coefficients of the linear operator depends only on the vertical coordinate, the scheme is termed a 'constant coefficient' one (Bénard (2003)). However imposing a reference state to solve the implicit problem leads to a dependence of the forecast on this state as shown by Thuburn et al. (2010) in the case of a shallow water system. Some properties of transport schemes can thus be degraded but are not the subject of this paper.

Furthermore Bénard et al. (2005) have shown that numerical stability can be enhanced by using an appropriate prognostic variable for the linear part in the vertical momentum equation. Thus, the vertical derivative of the contravariant vertical velocity $\mathbb{D}$ is used, instead of the traditional vertical velocity w or its vertical derivative d. This variable expressed in the mass-based coordinate η designed by Laprise (1992) is defined as:

$$\mathbb{D} = \mathrm{d} + \frac{p}{mRT}\vec{\nabla}\phi \cdot \frac{\partial \vec{U}}{\partial \eta}, \tag{3}$$

where:

$$\mathrm{d} = -g\frac{p}{mRT}\frac{\partial w}{\partial \eta}. \tag{4}$$

where T is the temperature, $\vec{U} = [U, V]$ is the horizontal wind, where U and V are the zonal and meridional components, w is the vertical wind, ϕ is the geopotential, and the other notations are specified in the Appendix B. Using this special prognostic variable leads to introduce other terms in the right hand side of (2), like the 'cross term' mentioned in Bénard et al. (2010). This term can be treated optionally in the advection scheme leading to enhance numerical stability.

More recently, another strategy followed by the global GEM model in order to increase numerical stability, has been to replace the mass-based coordinate introduced by Laprise (1992) by a height-based coordinate (Husain et al. (2019)). Preliminaries results are promising but it requires several major changes in the code. A height-based coordinate has also been chosen for the new FVM

dynamical core, under development, as an alternative to the mass-based IFS of ECMWF (Kühnlein et al. (2019)).

Increasing the number of steps n of the ICI scheme (2) is another strategy to improve numerical stability. For example, when the resolution moved from 2.5 km to 1.3 km in the AROME model in 2015, the extrapolated ICI scheme with only one step (SI-2TL-E), was replaced by a non-extrapolated scheme with two steps (ICI-2TL-PC) (Brousseau et al. (2016)). The time step was kept roughly the same during this change (from 60 s at 2.5 km to 50 s at 1.3 km), thus the additional iteration of the ICI scheme did not result in a significant overcost.

Other simpler ways to improve numerical stability are explored in this study. Indeed, some of numerical instabilities come from the terms of the non-linear residual (2) which are treated explicitly and impose stability constraints. For example, for constant coefficient implicit schemes, orographic terms are treated explicitly, and are likely to trigger instability. The first aim of this study is to explore three strategies to reduce their contributions for a constant coefficient scheme which is isothermal with a specific 'cold' temperature for the vertical momentum equation, dry, at rest, and with no orography. These three strategies are: decreasing the cold temperature of the basic state, decreasing the time step, and increasing the number of steps of the ICI scheme.

Then, orographic terms will be added in the previous linear operator $\mathcal{L}^*$, resulting in a variable coefficient scheme. However, in contrast to some other variable coefficient schemes like the one of the 'New Dynamics' dynamical core, this problem is based on a prescribed idealised basic state. Furthermore, the strategy of treating implicitly the orographic terms, unlike the other strategies, requires to leave some beneficial properties of semi-implicit constant coefficient schemes, like the separability between the horizontal and vertical parts of the implicit problem. In developing this new scheme from a constant coefficient scheme, it would be necessary:

- to build a non-spectral discretization on the horizontal direction to make computations with a sparse linear matrix instead of a full one if a spectral discretization is used (a Fourier decomposition would lead to expensive convolution products between horizontal derivative operators and orographic terms);

- to add orographic terms in the linear operator and discretize them horizontally and vertically;

- to implement an efficient 3D iterative solver, like a Krylov one, with efficient preconditioners (see Müller and Scheichl (2014) for more details), in addition of choosing an appropriate initialisation and tuning its stopping criterion.

On the first point, not every semi-implicit constant coefficient model uses a spectral discretization (e.g Qaddouri and Lee (2010)). These generally tend to avoid spectral computations to circumvent the scalability problem due to all-to-all communications when using the Fast Fourier Transform algorithm. For example, an iterative Krylov solver (Saad and Schultz (1986)), has been used efficiently in a grid point version of the constant coefficient scheme for the AROME model (Burgot et al. (2021)). Finally, the last two points require an additional deep modification of the code, for which it is not *a priori* guaranteed that it indeed improves stability at an affordable cost in NWP as shown by the recent removing of orographic terms in the implicit problem in the new dynamical core of the UM model. Furthermore, as mentioned in Liesen and Tichỳ (2004), the convergence rate of a Krylov method depends partly on the condition number of the implicit problem to be inverted. Thus, the goal of this article is to compute the eigenvalues of the different operators to estimate numerical stability and the computational cost of the method, before undertaking any coding work.

In section 2 the general methodology to test the above-mentioned strategies is exposed. In section 3 governing equations are introduced under their continuous and temporal-discretized forms in the simple case of the σ coordinate and of a constant slope, with a specific set of prognostic variables. A common formalism is introduced to treat optionally orographic terms implicitly or expliclity. In section 4, the eigenvalues of the various operators are computed under the unified formalism previously introduced. Then, in section 5, the results of these analyses when the orography is explicitly treated are exposed for the three strategies previously mentioned (decreasing the cold temperature of the basic state, decreasing the time step, and increasing the number of iterations of the ICI scheme). In section 6, the results of the analyses are presented when the orographic terms are treated implicitly, while testing different configurations of the ICI scheme. In section 7, basic estimates of condition numbers are computed when orographic terms are treated implicitly or explicitly. A conclusion and some perspectives are presented in a final section.

2. Methodology

Performing stability analyses in a general framework of the ICI-2TL scheme (2) is out of reach. The main interest of the analyses lies in their simplicity. The first stability analyses for semi-implicit schemes were introduced by Simmons et al. (1978) for a filtered equations system and then have been extended by Bénard (2003) for fully compressible Euler equations in mass based coordinate. This article is based on Bénard (2003)'s methodology, applied to the different goals outlined above.

To carry out these analyses in a simple way, the non-linear operator $\mathcal{M}$ of (1), is replaced by its tangent linear $\bar{\mathcal{L}}$ linearised around a basic state $\bar{\Psi}$. Ideally, this basic state should be as realistic as possible, but this would make the analysis much more complex. In practice, a very simple basic state is used: isothermal of temperature $\bar{T}$, dry, hydrostatic, at rest and on a constant orographic slope G. In the following this state will be called the 'physical' state. The evolution of the disturbed state vector Ψ' is thus of the form:

$$\frac{\partial \Psi'}{\partial t} = \bar{\mathcal{L}}(\Psi'), \tag{5}$$

For more clarity, the primes notation are now dropped, and Ψ refers to the disturbed state vector in the following. The system (5) is discretised by the ICI scheme (2), where $\mathcal{L}^*$ is the linear operator chosen to improve numerical stability and described in details in the next paragraph. What we call the 'non-linear' residual is now: $\mathcal{R} = \bar{\mathcal{L}} - \mathcal{L}^*$.

Because of the mass-based coordinate, vertical operators are composed of integral operators (Appendix B). An operator l independant of time is applied to the system (6), transforming the operators $\mathcal{L}^*$ and $\bar{\mathcal{L}}$ into operators involving only vertical derivatives operators for which the spectra are easily computed. The continuous system (5) becomes:

$$l\frac{\partial \Psi}{\partial t} = l\bar{\mathcal{L}}(\Psi) \tag{6}$$

Discretizing the system (6) by an ICI-2TL scheme (2) gives:

$$\frac{l\Psi^{+(n)} - l\Psi^0}{\Delta t} = \frac{l\mathcal{L}^*\left(\Psi^{+(n)} + \Psi^0\right)}{2} + \frac{l\mathcal{R}\left(\Psi^{+(n-1)} + \Psi^0\right)}{2}. \tag{7}$$

As mentioned in the introduction and shown by Bénard (2004), the linear operator $\mathcal{L}^*$ can be freely chosen without necessarily being the linear tangent. The reference state around which

the $\mathcal{L}^*$ operator is linearized is identical to the one of the $\bar{\mathcal{L}}$ operator except that it allows two temperatures which are respectively T_e^* and T^* for the vertical momentum equation and for the other equations. The 'non-linear' residual is thus simply reduced to a thermal residual measured by the dimensionless number:

$$\alpha = \frac{\bar{T} - T^*}{T^*}, \tag{8}$$

The ratio between the T_e^* and T^* temperatures is defined as:

$$r = \frac{T_e^*}{T^*} \tag{9}$$

The operator $\mathcal{L}^*$ can also contain orographic terms if they are treated implicitly (option $\delta = 1$) or not if they are treated explicitly (option $\delta = 0$). In the next section, the different equations are detailed.

3. Governing equations

In this section, the vertical coordinate is introduced, and then the operators $\bar{\mathcal{L}}$, $\mathcal{L}^*$, $l\bar{\mathcal{L}}$, and $l\mathcal{L}^*$, are detailed. Structure equations associated to these operators are then derived for the different options tested.

a. Mass-based coordinate

As mentioned previously, the vertical coordinate η developed by Laprise (1992) is based on the hydrostatic pressure π and the surface hydrostatic pressure π_s linked by:

$$\pi(x, \eta, t) = A(\eta) + B(\eta)\pi_s(x, t),$$

where A and B are two functions chosen such as the coordinate is terrain-following at the bottom of the atmosphere and tends towards the hydrostatic pressure at the top.

In this study for simplicity, the vertical coordinate chosen is purely terrain-following by choosing $A = 0$ and $B = \sigma$, with $\sigma \in [0, 1]$. In this case, the coordinate is called a σ-coordinate.

b. Tangent-linear operator $\bar{\mathcal{L}}$

In this paragraph the equations of the tangent-linear operator (5) of the general system (1) are prescribed using the following prognostic variables:

$$\Psi = \begin{bmatrix} U \\ \mathbb{D} \\ T \\ \hat{q} \\ \pi_s \end{bmatrix}. \tag{10}$$

where U is the zonal wind component, $\mathbb{D}$ is the prognostic variable for the vertical momentum equation mentioned previously (3), T is the temperature, $\hat{q}$ is defined as $\ln(p/\pi)$ where p is the true pressure, π is the hydrostatic pressure, and π_s is the hydrostatic surface pressure. Equations are written closely to equations (56)-(60) in Bénard et al. (2005), except that the horizontal velocity U is used instead of the horizontal divergence D, and the pressure surface perturbation π_s is used instead of its logarithmic contribution ($\ln(\pi_s)$):

$$\begin{aligned} \frac{\partial U}{\partial t} =& - R\bar{\partial}'_x \mathcal{G} T + R\bar{T}\bar{\partial}'_x \mathcal{G}\hat{q} - R\bar{T}\bar{\partial}'_x \hat{q} \\ & - \frac{R\bar{T}}{\bar{\pi}_s}\left(\partial_x + \frac{G}{\bar{H}}\right)\pi_s, \end{aligned} \tag{11}$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{R\bar{T}}\tilde{\partial}(\tilde{\partial}+1)\hat{q} + (1-\delta_{SL})\dot{\tilde{\mathrm{X}}}, \tag{12}$$

$$\frac{\partial T}{\partial t} = -\frac{R\bar{T}}{C_v}\left(\partial_x U + \mathbb{D}\right), \tag{13}$$

$$\frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v}\left(\partial_x U + \mathbb{D}\right) + \mathcal{S}\bar{\partial}'_x U, \tag{14}$$

$$\frac{\partial \pi_s}{\partial t} = -\bar{\pi}_s \mathcal{N}\partial_x U + \bar{\pi}_s \frac{G}{\bar{H}}\mathcal{N}U, \tag{15}$$

where:

$$\bar{\partial}'_x = \partial_x + \frac{G}{\bar{H}}\tilde{\partial}, \tag{16}$$

and $\bar{H} = R\bar{T}/g$ is the characteristic height of the physical state. Operators $\mathcal{G}$, $\mathcal{S}$, $\mathcal{N}$, $\tilde{\partial}$ are defined in Appendix A, and other notations in Appendix B. As mentioned in the introduction using the $\mathbb{D}$ prognostic variable leads to form the cross term defined as:

$$\bar{\mathrm{X}} = \frac{G}{\bar{H}}\tilde{\partial}U, \tag{17}$$

whose time derivative is:

$$\dot{\bar{\mathrm{X}}} = \frac{G}{\bar{H}}\tilde{\partial}\frac{\partial U}{\partial t} = gG\bar{\partial}'_x\left(\frac{T}{\bar{T}} - (1+\tilde{\partial})\hat{q}\right). \tag{18}$$

δ_{SL} is an option to apply a specific treatment of this term, and will be discussed in paragraph d. When this option is applied, the RHS (Right Hand Side) of (12) is reduced to:

$$\mathrm{RHS}_{\mathbb{D}} = -\frac{g^2}{R\bar{T}}\tilde{\partial}(\tilde{\partial}+1)\hat{q}. \tag{19}$$

c. Reference linear operator $\mathcal{L}^*$

The equations of the reference linear operator are similar to the system (11)-(15) but different reference temperatures are used (T^*, T_e^*):

$$\begin{aligned}\frac{\partial U}{\partial t} =& -R\partial'^*_x\mathcal{G}T + RT^*\partial'^*_x\mathcal{G}\hat{q} - RT^*\partial'^*_x\hat{q} \\ & - \frac{RT^*}{\pi_s^*}\left(\partial_x + \frac{G}{H^*}\right)\pi_s,\end{aligned} \tag{20}$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{rRT^*}\tilde{\partial}(\tilde{\partial}+1)\hat{q} + (1-\delta_{SL})\dot{\mathrm{X}}^*, \tag{21}$$

$$\frac{\partial T}{\partial t} = -\frac{RT^*}{C_v}\left(\partial_x U + \mathbb{D}\right), \tag{22}$$

$$\frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v}\left(\partial_x U + \mathbb{D}\right) + \mathcal{S}\partial'^*_x U, \tag{23}$$

$$\frac{\partial \pi_s}{\partial t} = -\pi_s^*\mathcal{N}\partial_x U + \pi_s^*\frac{G}{H^*}\mathcal{N}U, \tag{24}$$

where :

$$\partial_x'^* = \partial_x + \frac{G}{H^*}\tilde{\partial} \qquad \text{if } \delta = 1, \tag{25}$$

$$\partial_x'^* = \partial_x \qquad \text{if } \delta = 0, \tag{26}$$

and $H^* = RT^*/g$ is the characteristic height of the basic state. The cross term is defined as:

$$\mathrm{X}^* = \frac{G}{H^*}\tilde{\partial}U. \tag{27}$$

and its time derivative as:

$$\dot{\mathrm{X}}^* = \frac{G}{H^*}\tilde{\partial}\frac{\partial U}{\partial t} = gG\partial_x'^*\left(\frac{T}{T^*} - (1+\tilde{\partial})\hat{q}\right). \tag{28}$$

d. Specific treatment of the cross term

As mentioned in Bénard et al. (2005), the cross term can be treated optionally in the semi-Lagrangian advection scheme (option $\delta_{SL} = 1$) or in the semi-implicit scheme (option $\delta_{SL} = 0$). The semi-Lagrangian scheme is based on the computation of back trajectories of particles advected by the wind to find the origin points O from the end points F placed on the different grid points of the mesh. For the variable of the vertical wind divergence $\mathbb{D}$ at the first step ($n = 1$) of the ICI scheme:

$$\frac{\mathbb{D}_F^{+(1)} - \mathbb{D}_O^0}{\Delta t} = \mathrm{RHS}_{\mathbb{D}} + \frac{1}{\Delta t}(\mathrm{X}_F^{+(0)} + \mathrm{X}_O^{+(0)} - 2\mathrm{X}_O^0), \tag{29}$$

and for subsequent steps $n \geq 2$:

$$\frac{\mathbb{D}_F^{+(n)} - \mathbb{D}_O^0}{\Delta t} = \mathrm{RHS}_{\mathbb{D}} + \frac{1}{\Delta t}\left(\mathrm{X}_F^{+(n-1)} - \mathrm{X}_O^0\right), \tag{30}$$

where the RHS term (19) is discretized as the RHS of (7). When the SI-2TL-E scheme is used, only the equation (29) is computed with $X^{+(0)} = 2X^0 - X^-$.

e. Systems under analysis

A diagonal operator l is applied to the operators $\mathcal{L}^*$ and $\bar{\mathcal{L}}$ for which the diagonal terms are: $l_{11} = \tilde{\partial}$, $l_{22} = 1$, $l_{33} = 1$, $l_{44} = \tilde{\partial} + 1$, and $l_{55} = 1$.

Applying this operator to the equations (11)–(15) gives:

$$\tilde{\partial}\frac{\partial U}{\partial t} = R\bar{\partial}'_x T - R\bar{T}\bar{\partial}'_x \hat{q} - R\bar{T}\bar{\partial}'_x \tilde{\partial}\hat{q}, \tag{31}$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{R\bar{T}}\tilde{\partial}(\tilde{\partial}+1)\hat{q} + (1-\delta_{SL})\dot{\tilde{\mathrm{X}}}, \tag{32}$$

$$\frac{\partial T}{\partial t} = -\frac{R\bar{T}}{C_v}\left(\partial_x U + \mathbb{D}\right), \tag{33}$$

$$(\tilde{\partial}+1)\frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v}(\tilde{\partial}+1)\left(\partial_x U + \mathbb{D}\right) + \bar{\partial}'_x U. \tag{34}$$

$$\frac{\partial \pi_s}{\partial t} = -\bar{\pi}_s \mathcal{N}\partial_x U + \bar{\pi}_s \frac{G}{\bar{H}}\mathcal{N}U, \tag{35}$$

Applying the operator l to the equations (20)–(24) gives:

$$\tilde{\partial}\frac{\partial U}{\partial t} = R\partial'^*_x T - RT^*\partial'^*_x \hat{q} - RT^*\partial'^*_x \partial\hat{q}, \tag{36}$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{rRT^*}\tilde{\partial}(\tilde{\partial}+1)\hat{q} + (1-\delta_{SL})\dot{\mathrm{X}}^*, \tag{37}$$

$$\frac{\partial T}{\partial t} = -\frac{RT^*}{C_v}\left(\partial_x U + \mathbb{D}\right), \tag{38}$$

$$(\tilde{\partial}+1)\frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v}(\tilde{\partial}+1)\left(\partial_x U + \mathbb{D}\right) + \partial'^*_x U, \tag{39}$$

$$\frac{\partial \pi_s}{\partial t} = -\pi^*_s \mathcal{N}\partial_x U + \pi^*_s \frac{G}{H^*}\mathcal{N}U, \tag{40}$$

We notice that the RHS of equations (31)–(34) and (36)–(39) do not involved the surface pressure perturbation π_s. Thus in the following, the equation (35) and (40) are not considered, the term l_{55}

is ignored, and the state vector is reduced to:

$$\Psi = \begin{bmatrix} U \\ \mathbb{D} \\ T \\ \hat{q} \end{bmatrix}. \tag{41}$$

f. Structure equations

The structure equation associated to the operator $\bar{\mathcal{L}}$ is combined from (31)–(35) equations. When $\delta_{SL} = 0$, the structure equation is:

$$\left[-\frac{1}{\bar{c}^2}\frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2}\left(\bar{\partial}_x'^2 + \frac{\tilde{\partial}(\tilde{\partial}+1)}{\bar{H}^2} \right) + \bar{N}^2 \bar{\partial}_x'^2 \right] U = 0, \tag{42}$$

while when $\delta_{SL} = 1$ the structure equation is:

$$\begin{aligned} &\left[-\frac{1}{\bar{c}^2}\frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2}\left(\bar{\partial}_x'^2 + \frac{\tilde{\partial}(\tilde{\partial}+1)}{rH^{*2}} \right) + \bar{N}^2 \bar{\partial}_x'^2 \right] U = \\ &\left[\frac{G}{\bar{H}}\frac{R\bar{T}}{\bar{c}^2}\bar{\partial}_x' \frac{\partial^2}{\partial t^2} + \frac{G}{\bar{H}}\tilde{\partial}\bar{\partial}_x' \frac{\partial^2}{\partial t^2} \right] U, \end{aligned} \tag{43}$$

where $\bar{N}^2 = g^2/(C_p\bar{T})$, and $\bar{c}^2 = C_p/C_v R\bar{T}$ are respectively the Brunt-Väisälä frequency squared and the acoustic velocity squared of the physical state.

The structure equation associated to the operator $\mathcal{L}^*$ is combined from (36)–(40) equations. When $\delta_{SL} = 0$ the structure equation is:

$$\left[-\frac{1}{c^{*2}}\frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2}\left(\partial_x'^{*2} + \frac{\tilde{\partial}(\tilde{\partial}+1)}{rH^{*2}} \right) + \frac{N^{*2}}{r}\partial_x'^{*2} \right] U = 0, \tag{44}$$

while when $\delta_{SL} = 1$ the structure equation is:

$$\begin{aligned} &\left[-\frac{1}{c^{*2}}\frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2}\left(\partial_x'^{*2} + \frac{\tilde{\partial}(\tilde{\partial}+1)}{rH^{*2}} \right) + \frac{N^{*2}}{r}\partial_x'^{*2} \right] U = \\ &\left[\frac{G}{H^*}\frac{RT^*}{c^{*2}}\partial_x'^* \frac{\partial^2}{\partial t^2} + \frac{G}{H^*}\tilde{\partial}\partial_x'^* \frac{\partial^2}{\partial t^2} \right] U, \end{aligned} \tag{45}$$

where $N^{*2} = g^2/(C_p T^*)$, and $c^{*2} = C_p/C_v R T^*$ are respectively the Brunt-Väisälä frequency squared and the acoustic velocity squared of the basic state.

4. Stability analysis computations

In this section, the eigenvalues of the operators $\bar{\mathcal{L}}$ and $\mathcal{L}^*$ are computed for the different options studied. Physical and numerical growth rates are then introduced.

a. Different options

In this section, numerical stability is evaluated by an analysis for:

- the constant coefficient version ($\delta = 0$) when: decreasing the time step, increasing the number of steps of the ICI scheme, decreasing the cold temperature T_e^* ;
- the variable coefficient version ($\delta = 1$), when changing the number of steps of the ICI scheme.

In addition, for each of these two versions, numerical stability is evaluated when the cross term X previously mentioned, is treated by the semi-Lagrangian transport scheme ($\delta_{SL} = 1$), or in the implicit problem ($\delta_{SL} = 0$).

b. Eigenvalues computation

The modes of the system (6) are looked for under:

$$\Psi(x,\sigma) = \widehat{\Psi} \exp(ikx)\sigma^{i\nu - 1/2} \tag{46}$$

The horizontal wave number k is defined as:

$$k = \frac{2\pi}{L_x} n_k, \tag{47}$$

with $n_k \in [-K, K]$ where K is the horizontal truncation chosen such as the largest wave number is $k_{max} = 2\pi K/L_x = \pi/\Delta x$. The smallest wave number is $k_{min} = 0$.

The vertical dimensionless wave number ν is linked to its dimensional counterpart m defined as:

$$m = \frac{\nu}{\bar{H}}, \tag{48}$$

with ν chosen such as $m \in [m_{min}, m_{max}]$ where $m_{min} = 2\pi/\bar{H}$, and $m_{max} = \pi/\Delta z$.

The eigenvalues of the operators become:

$$\widehat{\partial_x} = ik, \tag{49}$$

$$\widehat{\bar{\partial}'_x} = ik + i\frac{G\nu}{\bar{H}} - \frac{G}{2\bar{H}} = \bar{k}', \tag{50}$$

$$\widehat{\partial'^*_x} = ik + i\frac{G\nu}{H^*} - \frac{G}{2H^*} = k'^*, \tag{51}$$

$$\widehat{\bar{\partial}} = i\nu - 1/2. \tag{52}$$

Once the modes expressed as (46), the system (6) becomes:

$$\frac{d\widehat{\Psi}}{dt} = \widehat{\bar{\mathcal{L}}}\widehat{\Psi} \tag{53}$$

and (41) becomes:

$$\widehat{\Psi} = \begin{bmatrix} \widehat{U} \\ \widehat{\mathbb{D}} \\ \widehat{T} \\ \widehat{\hat{q}} \end{bmatrix}. \tag{54}$$

and the coefficients of the operator l become: $\widehat{l_{11}} = \xi_1$, $\widehat{l_{22}} = 1$, $\widehat{l_{33}} = 1$, $\widehat{l_{44}} = \xi_4$.

The operator $\widehat{\bar{\mathcal{L}}}$ is now defined as:

$$\widehat{\bar{\mathcal{L}}} = \begin{bmatrix} 0 & 0 & \bar{\mu}_{13}/\xi_1 & \bar{\mu}_{14}/\xi_1 \\ 0 & 0 & \bar{\mu}_{23} & \bar{\mu}_{24} \\ \bar{\mu}_{31} & \bar{\mu}_{32} & 0 & 0 \\ \bar{\mu}_{41}/\xi_4 & \bar{\mu}_{42}/\xi_4 & 0 & 0 \end{bmatrix}. \tag{55}$$

The coefficients of $\widehat{\bar{\mathcal{L}}}$ are:

$$\bar{\mu}_{13} = R\bar{k}',$$

$$\bar{\mu}_{14} = -R\bar{T}\bar{k}'(i\nu + 1/2),$$

$$\bar{\mu}_{23} = (1-\delta_{SL})\frac{gG}{\bar{T}}\bar{k}',$$

$$\bar{\mu}_{24} = \frac{g^2}{R\bar{T}}(\nu^2+1/4) - (1-\delta_{SL})gG(i\nu+1/2)\bar{k}',$$

$$\bar{\mu}_{31} = -\frac{R\bar{T}}{C_v}ik,$$

$$\bar{\mu}_{32} = -\frac{R\bar{T}}{C_v},$$

$$\bar{\mu}_{41} = \bar{k}' - \frac{C_p}{C_v}ik(i\nu+1/2),$$

$$\bar{\mu}_{42} = -\frac{C_p}{C_v}(i\nu+1/2).$$

The ξ_1 and ξ_4 coefficients are:

$$\xi_1 = i\nu - 1/2,$$

$$\xi_4 = i\nu + 1/2.$$

Following the same formalism, the linear operator of the implicit problem becomes:

$$\widehat{\mathcal{L}^*} = \begin{bmatrix} 0 & 0 & \mu^*_{13}/\xi_1 & \mu^*_{14}/\xi_1 \\ 0 & 0 & \mu^*_{23} & \mu^*_{24} \\ \mu^*_{31} & \mu^*_{32} & 0 & 0 \\ \mu^*_{41}/\xi_4 & \mu^*_{42}/\xi_4 & 0 & 0 \end{bmatrix}. \tag{56}$$

Coefficients of $\widehat{\mathcal{L}^*}$ are:

$$\mu^*_{13} = R\left[(1-\delta)ik + \delta k'^*\right],$$

$$\mu^*_{14} = -RT^*(i\nu+1/2)\left[(1-\delta)ik + \delta k'^*\right],$$

$$\mu^*_{23} = (1-\delta_{SL})\frac{gG}{T^*}\left[\delta k'^*\right],$$

$$\mu^*_{24} = \frac{g^2}{rRT^*}(\nu^2+1/4) - (1-\delta_{SL})gG(i\nu+1/2)\left[\delta k'^*\right],$$

$$\mu_{31}^{*} = -\frac{RT^{*}}{C_v} ik,$$

$$\mu_{32}^{*} = -\frac{RT^{*}}{C_v},$$

$$\mu_{41}^{*} = [(1-\delta)ik + \delta k'^{*}] - \frac{C_p}{C_v} ik(i\nu + 1/2),$$

$$\mu_{42}^{*} = -\frac{C_p}{C_v}(i\nu + 1/2).$$

The ξ_1 and ξ_4 coefficients are:

$$\xi_1 = i\nu - 1/2,$$

$$\xi_4 = i\nu + 1/2.$$

c. Physical instability or damping

By diagonalizing the operator $\widehat{\bar{\mathcal{L}}}$ and integrating it temporally from t to $t + \Delta t$ the system (6), gives:

$$\widetilde{\widehat{\Psi}}(t + \Delta t) = \widetilde{\widehat{\Psi}}(t) \exp(\bar{\omega} \Delta t), \tag{57}$$

where $\widetilde{\widehat{\Psi}}$ is the projection of $\widehat{\Psi}$ in the eigenspace of $\widehat{\bar{\mathcal{L}}}$ and $\bar{\omega}$ are the four eigenvalues associated to $\bar{\mathcal{L}}$. The eigenvalues $\bar{\omega}$ of $\bar{\mathcal{L}}$ are also solutions of the structure equation (42):

$$-\frac{1}{\bar{c}^2}\bar{\omega}^4 + \bar{\omega}^2 \left[\bar{k}'^2 - \frac{1}{\bar{H}^2}\left(\nu^2 + \frac{1}{4}\right)\right] + \bar{N}^2 \bar{k}'^2 = 0, \tag{58}$$

where of the four solutions, two of them are gravity modes and the two others are acoustic ones. The system is physically:

- unstable when $\Re(\bar{\omega}) > 0$,

- damped when $\Re(\bar{\omega}) < 0$,

where $\mathfrak{R}$ designed the real part. The growth rate of each combination (k, ν) is:

$$\bar{\Gamma}(k,\nu) = \frac{\left|\widetilde{\widehat{\Psi}}(t+\Delta t)\right|}{\left|\widetilde{\widehat{\Psi}}(t)\right|} = \exp\left(\mathfrak{R}(\bar{\omega})\Delta t\right) \tag{59}$$

The largest growth rate is then deduced:

$$\bar{\Gamma}_{max} = \max_{k,\nu}\left(\bar{\Gamma}(k,\nu)\right). \tag{60}$$

d. Temporal discretization

Once the system expressed as (46), the ICI-2TL scheme (7) becomes:

$$\mathcal{H}\widehat{\Psi}^{+(n)} = \mathcal{E}\widehat{\Psi}^{0} + \frac{\Delta t}{2}\widehat{\mathcal{R}}\left(\widehat{\Psi}^{0} + \widehat{\Psi}^{+(n-1)}\right), \tag{61}$$

where

$$\mathcal{H} = \left(\mathcal{I} - \frac{\Delta t}{2}\widehat{\mathcal{L}^{*}}\right), \tag{62}$$

$$\mathcal{E} = \left(\mathcal{I} + \frac{\Delta t}{2}\widehat{\mathcal{L}^{*}}\right), \tag{63}$$

and where $\mathcal{I}$ designed the identity matrix.

As shown in the section 7, the implicit operator $\mathcal{H}$ is invertible for the range of slope considered and thus (61) becomes:

$$\widehat{\Psi}^{+(n)} = \mathcal{H}^{-1}\mathcal{E}\widehat{\Psi}^{0} + \frac{\Delta t}{2}\mathcal{H}^{-1}\widehat{\mathcal{R}}\left(\widehat{\Psi}^{0} + \widehat{\Psi}^{+(n-1)}\right), \tag{64}$$

where $\mathcal{H}^{-1}$ is expected to be inverted by a solver in a numerical model. The growth rate of the ICI-2TL scheme for each combination (k, ν) is:

$$\Gamma^{*}(k,\nu) = \frac{\left|\widehat{\Psi}^{+(n)}\right|}{\left|\widehat{\Psi}^{0}\right|}. \tag{65}$$

The largest growth rate of the ICI-2TL scheme is then deduced:

$$\Gamma^*_{max} = \max_{k,\nu} \left(\Gamma^*(k,\nu) \right). \tag{66}$$

e. Normalized growth rate

When a physically unstable system is considered, in order to not attribute a physical instability to a numerical instability, the largest growth rate of the numerical scheme is normalized by the largest growth rate of the physical system. Hence, in the following, the growth rate is defined as:

$$\Gamma = \frac{\Gamma^*_{max}}{\bar{\Gamma}_{max}}. \tag{67}$$

In this study, stable numerical schemes are primarily looked for, i.e. these where $\Gamma \leq 1$.

f. Limits of this analysis

This study is based on several hypotheses that might overestimate the numerical stability of the problem. For example, effects of vertical and horizontal discretization with finite-difference schemes are not considered in this article. The different choice of the prognostic variable for the vertical momentum equation (w versus $\mathbb{D}$) between the linear and the non-linear part, can lead to some inconsistencies and are neither considered in this analysis, etc.

Furthermore, the complexity of the relief is reduced to a single measure of its slope, but other problems could arise in the case of a more realistic orography where orographic modes of different wave numbers interact non-linearly.

The optional treatment of the cross term ($\delta_{SL} = 1$), is studied here when the atmosphere of the basic state is at rest, which is often far from being the case in reality. Consequently, the choice of this option in this study leads to probably transfer some numerical stability problems in the SL scheme, not studied here, and thus to over-estimate the total numerical stability of the problem.

g. Numerical values

Except in the part dealing the ratio r, the reference temperatures are identical to these used in the AROME model: $T^* = 350\,\mathrm{K}$, $T^*_e = 100\,\mathrm{K}$. To assure numerical stability, the warmest temperature

T^* bas been chosen such as it is higher than $\bar{T}$, whatever the meteorological situation considered. The thermal residual factor (8) will be thus chosen in $\alpha \in [-0.95, 0]$ in the following.

Except explicitly mentioned, horizontal resolutions chosen in the following are these expected in a mid-term future for operations: an horizontal resolution of $\Delta x = 300\,\mathrm{m}$ and a large time step $\Delta t = 12\,\mathrm{s}$ corresponding to a large CFL: $C^* = c^*\Delta t/\Delta x \approx 14$. A vertical grid with a first level at $\Delta z \simeq 2\,\mathrm{m}$ is chosen. At these resolutions slopes are considered mostly in $G \in [0, 3]$.

Graphics of the following section are plotted in function of the slope G (in ordinate) and of the thermal residual (in abscissa). Only positive values of slope are shown in graphics but results can be easily recovered for negative values by applying a symmetry with the abscissa axis. A special focus will be given to the '$\Gamma = 1.000$' isoline which delimits the numerical stability domain.

As mentioned previously, this study focuses particularly on estimating instability due to the numerical treatment of orographic terms. Consequently, a special focus is given to the maximum slope for which the scheme is numerically stable when no thermal residual occurs ($\alpha = 0$). However, numerical instabilities can also occurred when the temperatures T^* and $\bar{T}$ are different, producing a thermal residual ($\alpha \neq 0$). Consequently, numerical stability must be analyzed in comparison between these both aspects by measuring the 'spread' of the stability domain (the area where $\Gamma \leq 1$).

5. Explicit treatment of orographic terms

In this section, orographic terms are treated explicitly ($\delta = 0$).

a. Sensitivity to the ratio r

When no specific temperature is used for the vertical momentum equation ($T_e^* = T^*$), $r = 1$. In this case, the maximum slope for which the scheme is numerically stable is around 45° for the configuration chosen (Fig 1). When a thermal residual is added, the stability strongly decreases even for small slopes.

On the contrary, when the temperature T_e^* is significantly lower than T^* ($r \simeq 0.3$) slopes over 68° can be achieved for $\alpha = 0$. Moreover, the stability domain is larger in general (i.e. for $\alpha \neq 0$ as shown in Fig 2). This confirms experiments carried out with the AROME model for which numerical instabilities appear for steeper slopes when r is smaller. This is also in agreement with

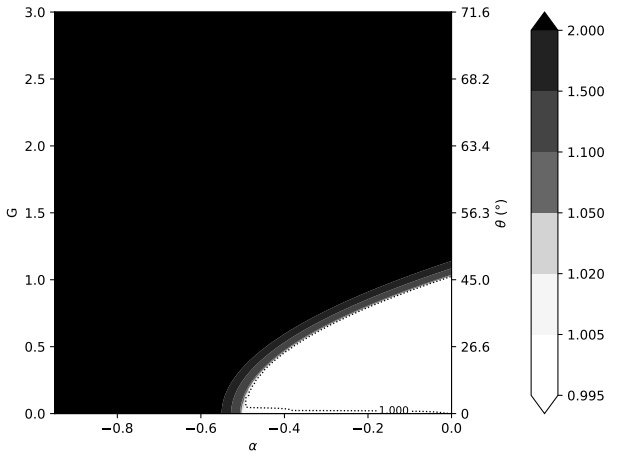


FIG. 1. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with $r = 1.0$, $\delta_{SL} = 0$, $\delta = 0$ and $C^* = 14$.

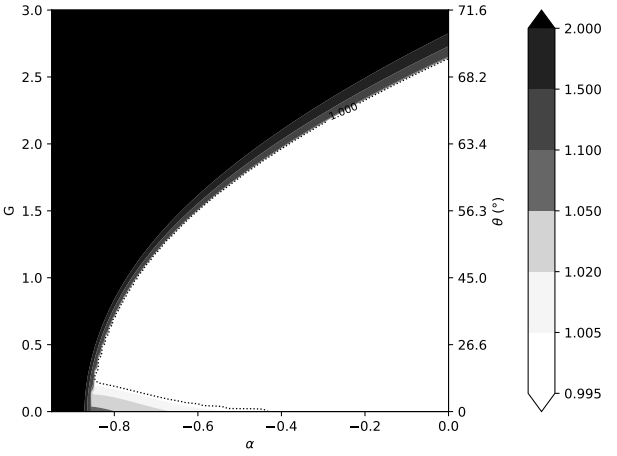
conclusions of Bénard (2004) showing the interest of adding degrees of freedom in the basic state to better control its stability.

The stability domain can be expanded by decreasing the value of the reference temperature T_e^*, and so the ratio value r (Fig 3 for $r \to 0$). However, an excessive decrease of this ratio r can worsen scores of a model and more particularly when $T_e < 100\,\mathrm{K}$ (not shown). That is why this strategy can barely be used and a value of $T_e = 100\,\mathrm{K}$ (corresponding to $r \simeq 0.3$) will be used in the following.

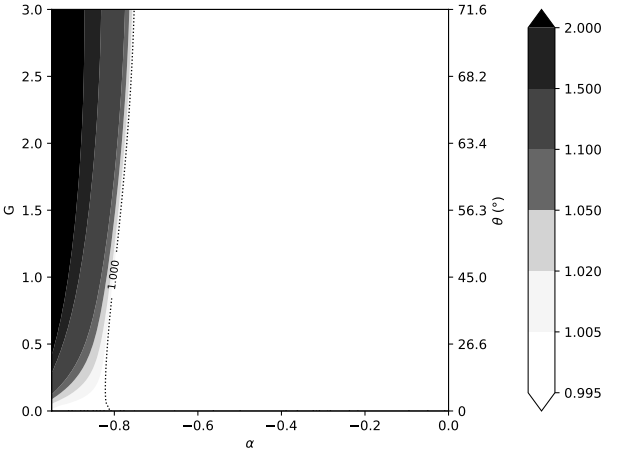
It is noted that despite the strong assumptions made to perform this analysis, the orders of magnitude of the maximum slopes obtained are close to those obtained when experiments are conducted in a real model. For example in the AROME model, numerical stability is achieved for slopes up to 50° in a realistic context (ie with a moderate thermal residual). This is in good agreement to what is given by this analysis with $\alpha \simeq -0.65$ in the Figure 2.

b. Specific treatment of the cross term

When a specific treatment of the cross term is applied, the stability domain is widely spread (Fig 4 compared to Fig 2). However, as mentioned previously, this analysis over-estimates probably the gain in stability, and this treatment will be not used any longer in the following stability analyses.



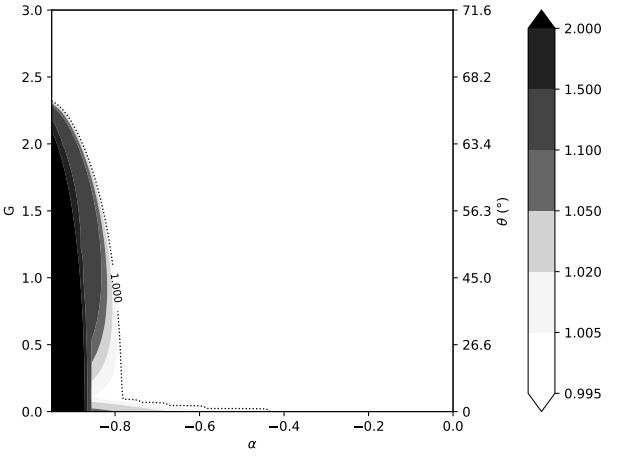
395 FIG. 2. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with
396 $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 14$.



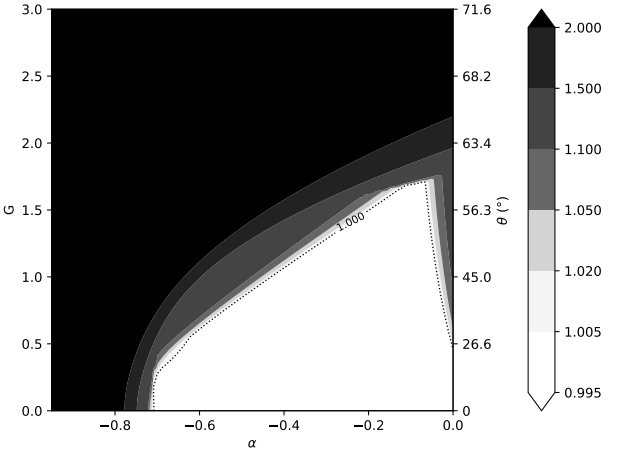
397 FIG. 3. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with
398 $r \simeq 0$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 14$.

405 *c. Changing temporal scheme*

406 The stability domain of the SI-2TL-E scheme (Fig 5) is narrower than the ICI-2TL-PC one. The
407 clear benefit of having an additional iteration of the ICI scheme can be seen here. This behaviour
408 which has also been observed with the full non-linear model is in good agreement with the shifting
409 from the SI-2TL-E scheme to the ICI-2TL-PC scheme for the AROME model in 2015 when the
410 resolution changed from 2.5 km to 1.3 km (see Brousseau et al. (2016)).



403 FIG. 4. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with
404 $r \simeq 0.3$, $\delta_{SL} = 1$, $\delta = 0$, $C^* = 14$.



411 FIG. 5. Growth rate as a function of the thermal parameter α and the slope G for the SI-2TL-E scheme with
412 $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 14$.

413 Moreover, increasing the number of iterations N_{iter} of the ICI scheme beyond 2 barely changes the
414 stability domain: two examples are given with $N_{iter} = 4$ (Fig 6) and $N_{iter} = 8$ (Fig 7). Consequently,
415 increasing the number of iterations of an ICI scheme does not seem to be a viable strategy to improve
416 numerical stability around the steepest slopes.

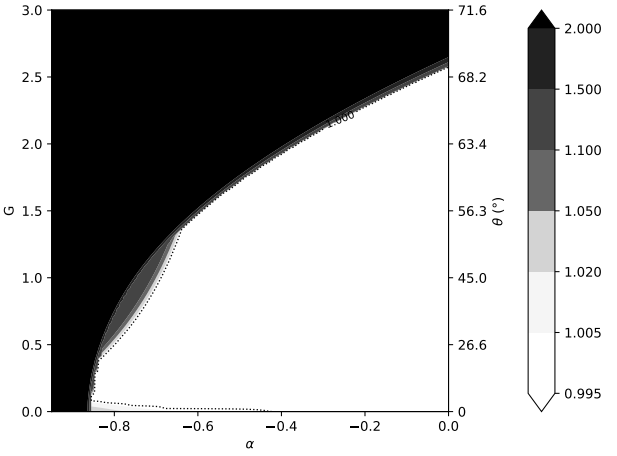


FIG. 6. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL scheme with $N_{iter} = 4$, $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 14$.

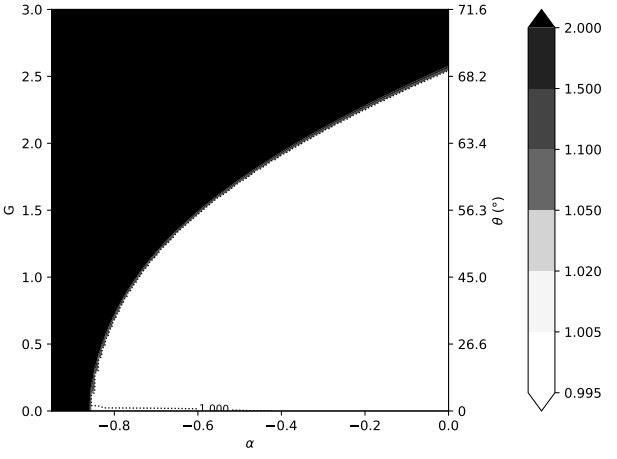


FIG. 7. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL scheme with $N_{iter} = 8$, $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 14$.

d. Time-step reduction

When CFL numbers close to unity are used, the stability domain is weakly modified compared to the case when a larger CFL number is used: an example is given when the time step is thus reduced by 14 (Fig 8).

Consequently, decreasing the time step does not increase numerical stability around the steepest slopes and is not a viable strategy. This is in good agreement with Bénard (2003), where numerical

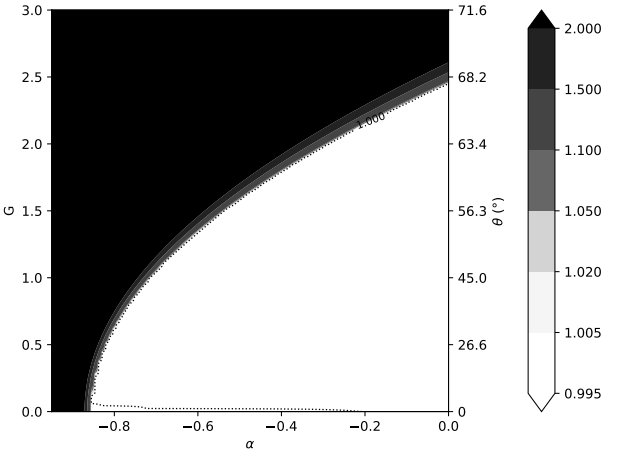


FIG. 8. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 0$, $C^* = 1.0$.

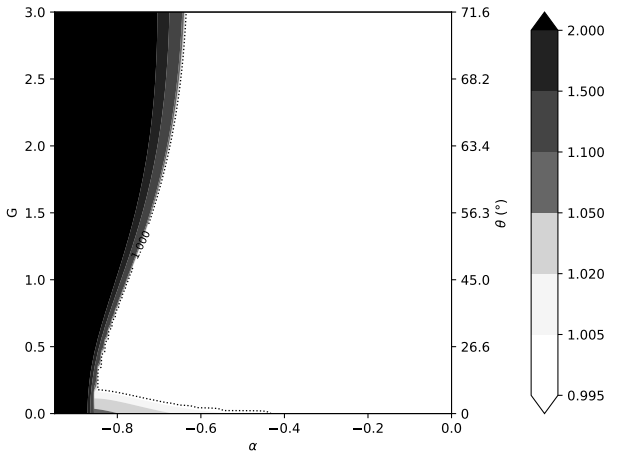
instabilities generated by the thermal residual does not depend on the time step in general. Steeper slopes can only be considered (when $\alpha = 0$) when only time steps are reduced by more than a 100 factor (not shown). Of course, these time steps cannot be used in an operational context.

e. Conclusion

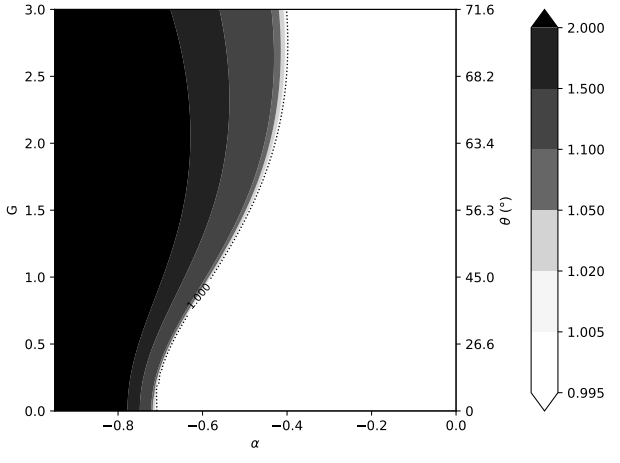
The strategy of treating orographic terms explicitly does not improve numerical stability in case of steep slopes even by: reducing down to the unity the CFL number, decreasing the ratio r down to $r \simeq 0.3$, or increasing the number of iterations of the ICI-2TL scheme. The only improvements in stability can be reached by worsening the quality of simulation (by taking too low values of r) or by degrading significantly the efficiency (by taking too low values of time steps). The implicit treatment of orographic terms appears to be the main way of improving numerical stability and is discussed in the next section.

6. Implicit treatment of orographic terms

In this section, orographic terms are treated implicitly ($\delta = 1$). When no thermal residual is added, the stability is reached unconditionally (whatever the slope is) because the total non-linear residual is null. In practice however, a thermal residual must be added like in the previous section. The



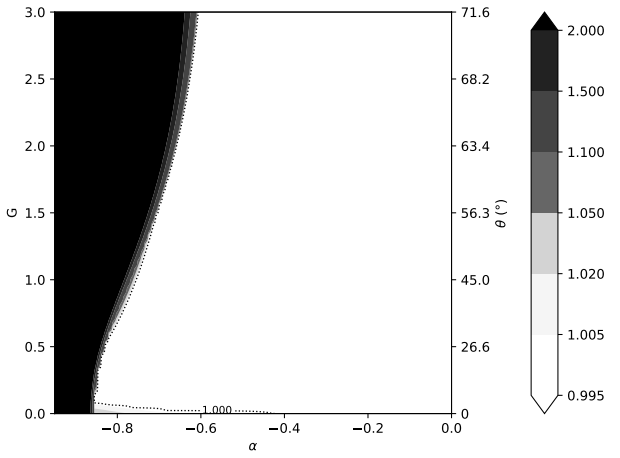
441 FIG. 9. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL-PC scheme with
442 $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 1$, $C^* = 14$



448 FIG. 10. Growth rate as a function of the thermal parameter α and the slope G for the SI-2TL-E scheme with
449 $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 1$, $C^* = 14$.

446 domain of stability is significantly larger (Fig 9) compared to what it is when terms are explicitly
447 treated (Fig 2).

450 When increasing the number of iterations of the ICI-2TL scheme up to 4 iterations, the stability
451 domain is slightly wider (Fig 11) than with only 2 iterations, but benefits in stability are weak
452 compared to the over-cost of this scheme. Increasing again the number of iterations does not
453 improve the numerical stability in general (not shown).



458 FIG. 11. Growth rate as a function of the thermal parameter α and the slope G for the ICI-2TL scheme with
459 $N_{iter} = 4$, $r \simeq 0.3$, $\delta_{SL} = 0$, $\delta = 1$, $C^* = 14$.

454 To reduce the computational cost, a solution would be to use the SI-2TL-E scheme, roughly half
455 the computation cost of the ICI-2TL-PC scheme. In this case, the stability domain is only smaller
456 for $\alpha < -0.5$ (Fig 10) compared to the ICI-2TL-PC one. This cheapest alternative could be used
457 for running a real model but needs to be checked in a more realistic context.

460 Finally, the stability is not significantly improved when only some terms are treated implicitly
461 while the others are treated in the non-linear residual (not shown). The larger stability domain is
462 obtained when all terms are treated implicitly.

463 This analysis has shown a real interest of treating implicitly the orographic terms to improve
464 numerical stability. In the next section, the additional cost of adding these terms in the implicit
465 problem will be estimated.

466 7. Invertibility and condition numbers

467 Several methods can be used to invert the implicit problem and are split into two main classes:
468 direct methods and iterative methods.

469 Direct methods invert the problem in a single iteration, are exact, and are not sensitive to the
470 condition number. Nevertheless, they often suffer from a scalability problem when different
471 MPI (Message Passing Inteface) tasks distributed on different compute nodes communicate. The
472 current parallelization paradigm leads to split the geographical domain horizontally into different

MPI tasks, while no splitting takes place vertically. Therefore, from a communication perspective, direct methods are well suited to invert only vertical problems. For problems with a horizontal part, direct methods use transforms algorithms such as FFT, and can be replaced by iterative methods for improving scalability (Burgot et al. (2021)). Iterative methods, such as the Krylov methods, are generally more scalable, but their speed of convergence depends on their initialization, their stopping criterion, and the condition number of the problem to be inverted (Liesen and Tichỳ (2004)).

Estimating the convergence speed of a method, by computing the condition number of the problem to be inverted, allows to roughly and simply evaluate the cost of a method. This enables to verify if the strategies previously studied to improve numerical stability in the presence of steep slopes, are not too expensive to be used for operations.

These estimates are computed for the configurations studied previously: when the orographic terms are treated explicitly, and when they are treated implicitly with or without the specific treatment of the cross term in the semi-Lagrangian scheme.

a. Methodology

In this section, the implicit problem is decomposed as in (46) and the eigenvalues of the implicit problem are thus analytically computed with $k_{min} = 0$, $k_{max} = \pi/\Delta x$, $m_{min} \simeq 0$, $m_{max} \simeq \pi/\Delta z$, and with the same numerical values for constants and parameters than in the paragraph g of the section 4.

Condition numbers are then estimated as the ratio between the largest and the smallest eigenvalue in absolute value for each considered slope G.

Moreover, invertibility conditions are also computed and allow to find the maximum slope beyond which the problem is no longer invertible, by computing the smallest slope for which the real and imaginary parts are simultaneously zero.

These estimates are given on an implicit problem which has been algebraically reduced on a single prognostic variable. This allows to reduce the size of the problem to be solved. In this study we will choose to reduce it on the horizontal wind speed U. The other prognostic variables are then deduced from this variable.

b. Explicit treatment of orographic terms

When the orographic terms are treated explicitly, the implicit problem (62) once algebraically reduced, has the following eigenvalues:

$$\mathcal{H}_r = 1 + \frac{\Delta t^2}{4} b k^2 \tag{68}$$

where:

$$b = \left(1 + \frac{\Delta t^2}{4} \frac{c^{*2}}{r H^{*2}} (\nu^2 + 1/4)\right)^{-1} \left(1 + \frac{\Delta t^2}{4} \frac{N^{*2}}{r}\right) c^{*2} \tag{69}$$

are the eigenvalues of the product between a 'vertical part' and some constants. For the numerical values considered, it can be noticed that

$$1 + \frac{\Delta t^2}{4} \frac{N^{*2}}{r} \simeq 1. \tag{70}$$

This approximation will be done in this paragraph and the following ones. The condition number of the problem (68) is thus:

$$C \simeq 1 + \frac{\Delta t^2}{4} c^{*2} \frac{\pi^2}{\Delta x^2} \tag{71}$$

because $b_{max} \simeq c^{*2}$, and $b_{min} \simeq 0$.

This condition number is plotted according to the slope G on the Figure 12. It is very low ($C \simeq 500$) compared to what is sometimes encountered in the literature. Consequently, a fast convergence of the Krylov solver can be reached, as shown by Burgot et al. (2021) in a spatially discretized context. Furthermore, the problem is always invertible whatever the slope is, because the eigenvalues (68) are all greater than 1.

We note from the first factor of (69), that a 'vertical' problem has to be inverted. Its condition number can be easily estimated as:

$$C_{vert} \simeq 1 + \frac{\Delta t^2}{4} c^{*2} \frac{\pi^2}{r \Delta z^2}. \tag{72}$$

Since in operational models the first vertical level is often lower than 10 meters ($\Delta z \ll \Delta x$), the vertical problem is thus significantly less well conditioned than the full problem (68) ($C_{vert} \gg C$). There is therefore a real interest in treating the vertical part differently from the rest of the problem.

For example, a direct method, which is not sensitive to the condition number, can be specifically used to invert the vertical part (69). Thereafter, we consider that the vertical part is inverted by a specific method, not sensitive to its condition number.

c. Implicit treatment of orographic terms with $\delta_{SL} = 0$

When orographic terms are treated implicitly and no specific treatment is applied for the cross term, the eigenvalues of the implicit problem reduced algebraically are:

$$\mathcal{H}_r = 1 - \frac{\Delta t^2}{4} b k'^{*2} \tag{73}$$

where the vertical part b remains the same as (69). For the range of slope considered, the condition number is quite similar to the one when all terms are treated explicitly (Fig 12) and remains low. The implicit treatment of orographic terms does not lead to worsen significantly the condition number of the implicit problem. A fast convergence of the solver can hence be expected.

Furthermore, from equation (73) and the null eigenvalue limit, it can be deduced that the problem is invertible only if the slope is:

$$|G| \leq \sqrt{\frac{16 H^{*2}}{c^{*2} \Delta t^2} + \frac{1}{r}} \simeq 9.3 \tag{74}$$

For the numerical values previously given, a slope up to 9.3 can be reached. This confirms that the problem is invertible for a wide range of slopes, and in any case for the slopes studied in this article (up to 3).

d. Implicit treatment of orographic terms with $\delta_{SL} = 1$

When orographic terms are treated implicitly, and a specific treatment is applied for the cross term, the eigenvalues of the implicit problem reduced algebraically are:

$$\begin{aligned} \mathcal{H}_r =& 1 - \frac{\Delta t^2}{4} c^{*2} k'^{*2} \left(\frac{\Delta t^2}{4} \frac{N^{*2}}{r} + 1 \right) b' \\ &+ \frac{\Delta t^2}{4} \frac{G}{H^*} k'^* \left(c^{*2} (i\nu - 1/2) + R T^* \right) b' \end{aligned} \tag{75}$$

where:

$$b' = \left(1 + \frac{\Delta t^2}{4}\frac{c^{*2}}{rH^{*2}}(\nu^2 + 1/4)\right)^{-1} \tag{76}$$

The condition number is greater than in the previous configurations, but remains low for the range of slope considered (Fig 12) and the configuration used ($r \simeq 0.3$). It remains well below to conditions numbers usually encountered in recent literature in applied mathematics (around 10^7).

Furthermore, the problem is invertible only if the slope respects approximately the following condition:

$$|G| \leq \sqrt{\frac{\dfrac{4}{\Delta t^2} + \dfrac{c^{*2}}{4rH^{*2}} + \dfrac{c^{*2}}{rH^{*2}}\nu^2}{\dfrac{g}{2H^*} + c^{*2}\dfrac{1}{H^{*2}}\beta\,(1-\beta)\,\nu^2}} \tag{77}$$

with:

$$\beta = g\left(\frac{1}{2}c^{*2}\frac{1}{H^*} + g\right)^{-1} = \frac{1}{1 + \dfrac{1}{2}\dfrac{C_p}{C_v}} \simeq 0.59 \tag{78}$$

For large values of ν, the invertibility condition becomes:

$$|G| \leq \frac{1}{\sqrt{\beta\,(1-\beta)\,r}} \simeq \frac{2}{\sqrt{r}} \simeq 3.7 \tag{79}$$

This condition is more restrictive than the one when no specific treatment is applied on the cross term (74), but remains acceptable for the range of slope considered in this article and the configuration used ($r \simeq 0.3$).

e. Conclusion

The condition number remains low for the three configurations studied above, suggesting a good convergence of the iterative solver used to invert the implicit problem. However, treating implicitly the orographic terms and the cross term in the semi-Lagrangian scheme leads to more restrictive invertibility conditions. Therefore this treatment is probably not the most appropriate in the context of steep slopes. The implicit treatment of orographic terms and the implicit treatment of the cross term is the preferred solution.

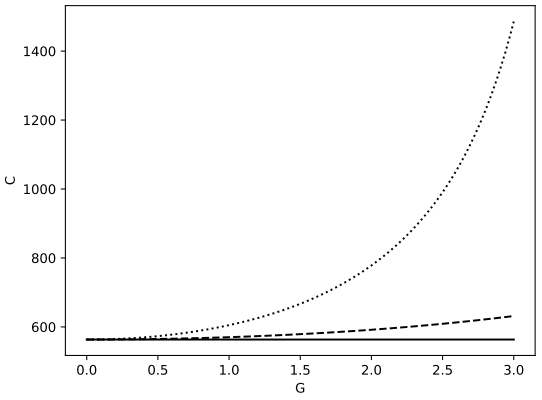


FIG. 12. Condition number C of the implicit problem when $r \simeq 0.3$, $C^* = 14$ with : implicit treatment of orographic terms $\delta = 1$ and no specific treatment of the cross term $\delta_{SL} = 0$ (dashed line), implicit treatment of orographic terms $\delta = 1$ and specific treatment of the cross term $\delta_{SL} = 1$ (dotted line), and explicit treatment of orographic terms $\delta = 0$ (continuous line).

In addition, the vertical part of the problem is well suited to be inverted by a direct method which is not sensitive to the condition number and where communications between vertical levels are not expensive in the current parallelization strategy.

8. Conclusion

The ICI-2TL scheme which is approaching the Crank-Nicolson scheme is studied for the fully elastic system of Euler equations using a mass coordinate, a derived prognostic variable for the vertical momentum equation $\mathbb{D}$, in case of a linear relief and a simple thermal residual.

The numerical instabilities associated with steep slopes usually encountered in NWP were thus faithfully reproduced by this study, including the orders of magnitude of the maximum slopes over which numerical instabilities occur. For example, the limit slope of about 45° encountered by Husain et al. (2019), has been found when: no degree of freedom is added to control the vertical propagation of fast waves (by using a colder temperature T_e^*) without any thermal residual ($\alpha = 0$), or when a strong thermal residual is added ($\alpha \simeq -0.7$) with an ICI-PC scheme and a cold temperature T_e^* =100 K is chosen for the vertical momentum equation. This analysis appears to be

an inexpensive way to test several dynamical core strategies to tackle the problem of steep slopes, without running an entire model at hectometric scales.

When orographic terms are explicitly treated, easy-to-implement strategies could be evaluated to expand the stability domain. This study leads to the following conclusions:

- decreasing the time step results in stability improvements only if it is reduced drastically, which is not viable in an operational context;
- decreasing the temperature T_e^* whilst keeping two iterations of the ICI scheme, leads to stabilize but also to worsen scores, and can therefore be considered sparingly;
- the specific treatment of the cross term X allows a significant improvement in stability but is probably overestimated due to the choice of a basic state at rest in this analysis;
- increasing the number of iterations of the ICI-2TL scheme does not expand the stability domain.

These conclusions show the limitations of constant coefficient schemes to handle the stability problem caused by steep slopes. Modest improvements can only be achieved by worsening scores or by significantly reducing efficiency. Implicit treatment of orographic terms appears to be one of the only viable options and conclusions are now given on this point.

First, it has been shown that the implicit problem is invertible in the range of slopes considered in this study, i.e. for slopes up to 70° with an hectometric horizontal resolution, in contrast to previous experiments with the older dynamical core of the UM model, where invertibility problems appeared for much lower slopes (Davies et al. (2005)). Thus, maintaining a horizontally homogeneous state independent of time, except for the orographic terms, seems viable to guarantee the invertibility of the implicit problem.

Results show that significant gains in stability can be achieved with slopes up to 70° with a moderate thermal residuals ($\alpha \simeq -0.5$) using the low-cost SI-2TL-E scheme. The use of an additional iteration of the ICI scheme allows to tackle these slopes for even stronger thermal residuals ($\alpha \simeq -0.65$).

Condition number estimates show that it remains low, even for steep slopes (up to $G = 3$) whatever the configuration used: with or without implicit treatment of orographic terms. This suggests a fast convergence of the solver in all configurations. However, more stringent invertibility conditions

and stronger condition numbers are expected when the cross term is treated in the semi-Lagrangian scheme compared to its treatment in the implicit scheme. Consequently, the treatment of the cross term in the implicit scheme instead of the semi-Lagrangian scheme appears more attractive.

This study argues for the replacement of the current constant coefficient semi-implicit scheme in favor of a variable scheme, that treats orographic terms in the implicit problem, to tackle hectometric resolutions and the resulting steep slopes. Nevertheless, a lot of work is needed to implement such a method, such as:

- building a non-spectral discretization on the horizontal direction if a spectral one is currently used;
- discretizing the new operators present in the implicit problem by ensuring consistency with the non-linear residual part;
- implementing and tuning a Krylov solver, to solve the implicit problem containing the orographic terms, and possibly to add a preconditioner to accelerate convergence;
- conducting test cases on non-linear flows in the presence of a more realistic (non-linear) relief and steep slopes.

Acknowledgments.

Data availability statement.

APPENDIX A

Vertical operators

Vertical operators are defined such as:

$$\mathcal{G}\Psi = \int_{\sigma}^{1} (\Psi/\sigma')\mathrm{d}\sigma'$$

$$\mathcal{S}\Psi = (1/\sigma)\int_{0}^{\sigma} \Psi\mathrm{d}\sigma'$$

$$\mathcal{N}\Psi = \int_{0}^{1} \Psi\mathrm{d}\sigma$$

$$\tilde{\partial}\Psi = \sigma\frac{\partial\Psi}{\partial\sigma}$$

APPENDIX B

The following notations are used:

- m : the metric factor (equals to $\partial_\eta\pi$)
- p : local pressure
- w : vertical velocity
- ϕ : geopotential
- g : acceleration of gravity
- R : gas constant of dry air
- C_p : specific heat capacity of dry air at constant pressure
- C_v : specific heat capacity of dry air at constant volume

The underscript ‘s’ refers to a surface field.

References

Bénard, P., 2003: Stability of semi-implicit and iterative centered-implicit time discretizations for various equation systems used in nwp. *Monthly weather review*, **131 (10)**, 2479–2491.

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. *Monthly Weather Review*, **132 (5)**, 1319–1324, https://doi.org/10.1175/1520-0493(2004)132<1319:otuoaw>2.0.co;2.

Bénard, P., J. Mašek, and P. Smolíková, 2005: Stability of leapfrog constant-coefficients semi-implicit schemes for the fully elastic system of euler equations: Case with orography. *Monthly Weather Review*, **133 (5)**, 1065–1075, https://doi.org/10.1175/mwr2907.1.

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. *Quarterly Journal of the Royal Meteorological Society*, **136 (646)**, 155–169, https://doi.org/10.1002/qj.522.

Brousseau, P., Y. Seity, D. Ricard, and J. Léger, 2016: Improvement of the forecast of convective activity from the AROME-france system. *Quarterly Journal of the Royal Meteorological Society*, **142 (699)**, 2231–2243, https://doi.org/10.1002/qj.2822.

Burgot, T., L. Auger, and P. Bénard, 2021: Krylov solvers in a vertical-slice version of the semi-implicit semi-lagrangian AROME model. *Quarterly Journal of the Royal Meteorological Society*, **147 (736)**, 1497–1515, https://doi.org/10.1002/qj.3976.

Davies, T., M. Cullen, A. Malcolm, M. Mawson, A. Staniforth, A. White, and N. Wood, 2005: A new dynamical core for the met office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **131 (608)**, 1759–1782.

Husain, S. Z., C. Girard, A. Qaddouri, and A. Plante, 2019: A new dynamical core of the global environmental multiscale (GEM) model with a height-based terrain-following vertical coordinate. *Monthly Weather Review*, **147 (7)**, 2555–2578, https://doi.org/10.1175/mwr-d-18-0438.1.

Kühnlein, C., W. Deconinck, R. Klein, S. Malardel, Z. P. Piotrowski, P. K. Smolarkiewicz, J. Szmelter, and N. P. Wedi, 2019: Fvm 1.0: a nonhydrostatic finite-volume dynami-

cal core for the ifs. *Geoscientific Model Development*, **12 (2)**, 651–676, https://doi.org/10.5194/gmd-12-651-2019, URL https://www.geosci-model-dev.net/12/651/2019/.

Lac, C., and Coauthors, 2018: Overview of the meso-NH model version 5.4 and its applications. https://doi.org/10.5194/gmd-2017-297.

Laprise, R., 1992: The euler equations of motion with hydrostatic pressure as an independent variable. *Monthly weather review*, **120 (1)**, 197–207.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the met office unified model for forecasting convection over the united kingdom. *Monthly Weather Review*, **136 (9)**, 3408–3424, https://doi.org/10.1175/2008mwr2332.1.

Liesen, J., and P. Tichỳ, 2004: Convergence analysis of krylov subspace methods. *GAMM-Mitteilungen*, **27 (2)**, 153–173.

Müller, E. H., and R. Scheichl, 2014: Massively parallel solvers for elliptic partial differential equations in numerical weather and climate prediction. *Quarterly Journal of the Royal Meteorological Society*, **140 (685)**, 2608–2624.

Qaddouri, A., and V. Lee, 2010: The elliptic solvers in the canadian limited area forecasting model gem-lam. *Modeling Simulation and Optimization-Tolerance and Optimal Control*, IntechOpen.

Robert, A., J. Henderson, and C. Turnbull, 1972: An implicit time integration scheme for baroclinic models of the atmosphere. *Monthly Weather Review*, **100 (5)**, 329–335.

Saad, Y., and M. H. Schultz, 1986: Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, **7 (3)**, 856–869.

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

Simmons, A. J., B. J. Hoskins, and D. M. Burridge, 1978: Stability of the semi-implicit method of time integration. *Monthly Weather Review*, **106 (3)**, 405–412, https://doi.org/10.1175/1520-0493(1978)106<0405:sotsim>2.0.co;2.

Thuburn, J., M. Zerroukat, N. Wood, and A. Staniforth, 2010: Coupling a mass-conserving semi-lagrangian scheme (SLICE) to a semi-implicit discretization of the shallow-water equations: Minimizing the dependence on a reference atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **136 (646)**, 146–154, https://doi.org/10.1002/qj.517.

Walters, D., and Coauthors, 2017: The met office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development*, **10 (4)**, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017.

Zängl, G., 2012: Extending the numerical stability limit of terrain-following coordinate models over steep slopes. *Monthly Weather Review*, **140 (11)**, 3722–3733.