



**HAL**  
open science

# Viabilité et performance des solveurs points de grille dans la résolution du problème implicite du noyau dynamique du modèle AROME

Thomas Burgot

► **To cite this version:**

Thomas Burgot. Viabilité et performance des solveurs points de grille dans la résolution du problème implicite du noyau dynamique du modèle AROME. Sciences de la Terre. Institut National Polytechnique de Toulouse - INPT, 2021. Français. NNT : 2021INPT0006 . tel-03836399

**HAL Id: tel-03836399**

**<https://theses.hal.science/tel-03836399>**

Submitted on 2 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THÈSE

En vue de l'obtention du

**DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

---

---

Présentée et soutenue le 19/01/2021 par :

**THOMAS BURGOT**

**Viabilité et performance des solveurs points de grille dans la résolution  
du problème implicite du noyau dynamique du modèle AROME**

---

---

LUDOVIC AUGER  
PIERRE BÉNARD  
XAVIER BLANC  
THOMAS DUBOS  
CHRISTINE LAC  
PIET TERMONIA

**JURY**  
CNRM  
CNRM  
Université de Paris  
École Polytechnique  
CNRM  
Université de Ghent

Membre du Jury  
Membre du Jury  
Membre du Jury  
Membre du Jury  
Présidente du Jury  
Membre du Jury

---

**École doctorale et spécialité :**

*SDU2E : Océan, Atmosphère et Surfaces Continentales*

**Unité de Recherche :**

*Centre National de Recherches Météorologiques - UMR 3589*

**Directeur(s) de Thèse :**

*Ludovic AUGER et Pierre BÉNARD*

**Rapporteurs :**

*Xavier BLANC, Thomas DUBOS et Piet TERMONIA*



---

# Remerciements

Je remercie tout d'abord Pierre Bénard, mon directeur de thèse, pour sa rigueur, ses conseils, et ses explications, m'évitant parfois de me noyer dans la complexité de la discipline. Je remercie Ludovic Auger, mon co-directeur de thèse, pour l'écriture d'un sujet de thèse passionnant et épuré. Son encadrement, sa gentillesse, et son écoute, nous ont permis de répondre à la plupart des questions posées il y a maintenant plus de trois ans et d'ouvrir quelques perspectives intéressantes. Je remercie également les Rapporteurs et Examineurs de ma thèse d'avoir pris le temps de lire, et commenter ce manuscrit.

Cette thèse s'inscrit dans la continuité de plusieurs années de formation de qualité que j'ai reçu depuis les petites classes jusqu'à l'Université. Je tenais à remercier plus particulièrement certains professeurs qui ont eu un rôle prépondérant dans la poursuite de mes études, en particulier Mmes Mandon et Cosmier, au lycée. Dans l'enseignement supérieur, je tenais à remercier Nathalie Rittemard, de m'avoir initié à la recherche via plusieurs stages, ce qui m'a donné l'envie de poursuivre.

Je remercie mes amis : certains empêtrés dans une thèse – quelle drôle d'idée ! – comme Damien, Marc, Mayeul, Thibault, Zied, et Olivier ; ceux qui sont restés fidèlement à Toulouse (ou qui en sont partis puis revenus) comme Adrien P., Alice, Axel, Cindy, Florian, Hélène, Hugo, Jérémy, Jordan, Julie, Lucas, Marie, Marine, Vincent ; ceux qui ont fui Toulouse pour l'autre bout de la France ou du Monde comme Quentin et Timothée ; ceux qui sont Bretons (et souvent exilés ailleurs) comme Arthur, Elisabeth, Irène, Jean-Baptiste, Maryline, Simon, Hugues et Maxime. Vous revoir le temps d'un midi, d'une randonnée, ou de quelques jours en vacances, est toujours un plaisir.

Je tenais également à souligner l'incroyable patience de mes co-bureaux du bureau 183 (Benoît, Charles, Marylis, Méryl, Adrien W., Martin, Yamina, et tant d'autres), bien qu'ils aient tous fini par le quitter prématurément pour une raison qui reste encore à déterminer. Je remercie également l'équipe RECYF, avec laquelle j'ai partagé de très bons moments notamment lors de nos pauses culinaires exceptionnelles. D'une manière plus générale, je remercie les gens du CNRM et de Météo-France : travailler avec eux est un réel plaisir.

Enfin, je remercie ma famille, d'avoir systématiquement œuvré pour que que je puisse étudier dans les meilleures conditions et pour son soutien précieux pendant ma thèse.



# Table des matières

<b>Introduction</b>	<b>7</b>
<b>1 État de l'art</b>	<b>11</b>
1.1 Prévion Numérique du Temps . . . . .	11
1.2 Moyens technologiques . . . . .	12
1.3 Problèmes scientifiques . . . . .	14
1.3.1 Problème multi-échelle, multi-physique . . . . .	14
1.3.2 Stabilité et instabilité physique . . . . .	16
1.4 Numérisation . . . . .	18
1.4.1 Convergence numérique . . . . .	19
1.4.2 Maillage et coordonnée . . . . .	21
1.4.3 Grille et discrétisation . . . . .	23
1.4.4 Filtrage spatial et diffusion numérique . . . . .	27
1.4.5 Discrétisation explicite ou implicite . . . . .	27
1.4.6 Modèle eulérien ou semi-lagrangien . . . . .	29
1.4.7 Discrétisation temporelle des modèles Eulériens . . . . .	31
1.4.8 Discrétisation temporelle des modèles semi-lagrangiens . . . . .	32
1.4.9 Opérateur linéaire . . . . .	34
1.4.10 Discrétisation spectrale . . . . .	35
1.4.11 Pas de temps . . . . .	38
1.4.12 Exemple de modèles utilisés . . . . .	40
1.4.13 Enjeux des prochaines années . . . . .	41
1.5 Conclusion . . . . .	42
<b>2 Noyau dynamique du modèle AROME</b>	<b>45</b>
2.1 Partie numérique du modèle AROME . . . . .	46
2.1.1 Présentation générale . . . . .	46
2.1.2 Coordonnées et grilles . . . . .	47
2.1.3 Équations du modèle . . . . .	47

2.1.4	Tangent linéaire . . . . .	49
2.1.5	Opérateur linéaire considéré . . . . .	51
2.1.6	Variables pronostiques . . . . .	53
2.1.7	Discrétisation spatiale . . . . .	55
2.1.8	Discrétisation temporelle ICI . . . . .	57
2.1.9	Inversion du problème implicite . . . . .	57
2.1.10	Schéma semi-Lagrangien . . . . .	59
2.1.11	Fonctionnement général du code actuel . . . . .	60
2.1.12	Pas de temps . . . . .	60
2.1.13	Diffusion et filtrage . . . . .	61
2.2	Limites de la version actuelle . . . . .	63
2.2.1	Scalabilité . . . . .	63
2.2.2	Fortes pentes . . . . .	64
2.2.3	Origine des problèmes . . . . .	64
2.3	Solutions étudiées dans la thèse . . . . .	65
2.3.1	Scalabilité . . . . .	65
2.3.2	Fortes pentes . . . . .	65
<b>3</b>	<b>Scalabilité d'AROME points de grille</b>	<b>67</b>
3.1	Version points de grille . . . . .	68
3.1.1	Problème implicite . . . . .	68
3.1.2	Dérivées horizontales . . . . .	68
3.1.3	Choix du solveur . . . . .	70
3.1.4	Solveur de Krylov . . . . .	71
3.1.5	Fonctionnement . . . . .	73
3.2	Méthodologie d'estimation de la scalabilité . . . . .	74
3.2.1	Mesure directe de la scalabilité . . . . .	75
3.2.2	Mesure indirecte de la scalabilité . . . . .	76
3.3	AROME 2D points de grille . . . . .	81
3.4	Article : évaluation de la scalabilité du modèle AROME 2D PG . . . . .	82
3.5	AROME 3D partiellement points de grille . . . . .	100
3.5.1	Présentation générale . . . . .	100

## TABLE DES MATIÈRES

---

3.5.2	Critère d'arrêt . . . . .	101
3.5.3	Estimation de la qualité . . . . .	102
3.6	Évaluation de la scalabilité du modèle AROME 3D PG . . . . .	102
3.6.1	Paramétrage du solveur de Krylov . . . . .	103
3.6.2	Validation sur plusieurs champs et niveaux verticaux . . . . .	105
3.6.3	Comparaison sur cas réel fortement précipitant . . . . .	106
3.6.4	Scores sur un mois . . . . .	107
3.7	Conclusion . . . . .	107
<b>4</b>	<b>Fortes pentes</b>	<b>113</b>
4.1	Nouvelles équations du système linéaire à coefficients variables . . . . .	114
4.1.1	Termes orographiques . . . . .	114
4.1.2	Équations en coordonnée $\eta$ . . . . .	114
4.1.3	Équations en coordonnée $\sigma$ . . . . .	115
4.1.4	Discrétisation spatiale . . . . .	117
4.1.5	Discrétisation temporelle . . . . .	118
4.2	Méthodologie d'estimation de la stabilité numérique . . . . .	119
4.2.1	Difficultés . . . . .	120
4.2.2	Analyse de stabilité . . . . .	120
4.2.3	Expériences numériques dans un cadre 2D . . . . .	123
4.2.4	En bref . . . . .	124
4.3	Article : analyse de différents schémas en présence de fortes pentes . . . . .	125
4.4	Expériences numériques dans un cadre 2D . . . . .	141
4.4.1	Configuration choisie . . . . .	141
4.4.2	Écoulement sur relief idéalisé . . . . .	141
4.4.3	Écoulement faiblement non-linéaire . . . . .	142
4.4.4	Écoulement non-linéaire . . . . .	145
4.4.5	Écoulement fortement non-linéaire sur relief réaliste . . . . .	146
4.4.6	Estimation de la vitesse de convergence . . . . .	149
4.5	Conclusion . . . . .	150
	<b>Conclusion</b>	<b>153</b>



<b>A</b>	<b>Discrétisation horizontale</b>	<b>157</b>
A.1	Discrétisation spatiale puis temporelle . . . . .	157
A.2	Discrétisation temporelle puis spatiale . . . . .	158
A.3	Analyse de Von-Neumann . . . . .	159
A.3.1	Discrétisation spatiale puis temporelle . . . . .	159
A.3.2	Discrétisation temporelle puis spatiale . . . . .	159
A.4	Analyse en modes normaux . . . . .	160
A.5	Conclusion . . . . .	161
<b>B</b>	<b>Analyse de stabilité du gradient conjugué</b>	<b>163</b>
	<b>Bibliographie</b>	<b>172</b>

# Introduction

La Prévision Numérique du Temps (PNT) est une discipline consistant à intégrer temporellement les équations qui régissent le comportement du système atmosphérique, en respectant une contrainte forte : fournir la meilleure prévision météorologique possible, à temps. Le modèle de PNT à aire limitée AROME répond à cet objectif en fournissant une prévision à une échéance de 48 heures, en seulement 70 minutes de simulation. Le noyau dynamique de ce modèle traite les phénomènes atmosphériques dont l'échelle caractéristique est supérieure à la taille de la maille, en résolvant les équations d'Euler compressibles non-hydrostatiques. Les phénomènes dont l'échelle caractéristique est inférieure sont paramétrisés, et leurs contributions sont fournies au noyau dynamique via des termes sources. Pour fournir une prévision à une résolution horizontale kilométrique, en respectant les délais contraints inhérents à la PNT, AROME tourne efficacement sur l'un des supercalculateurs massivement parallèles actuellement disponibles.

Une étape de numérisation est nécessaire en discrétisant spatialement et temporellement les équations du noyau dynamique, en l'absence de solution analytique connue. Cette discrétisation nécessite de satisfaire un ensemble de contraintes. Pour en affaiblir une partie des plus fortes, un schéma de transport semi-lagrangien est appliqué aux termes d'advection des équations du noyau dynamique. Il repose sur la reconstitution des trajectoires au cours d'un pas de temps, de chacune des particules associées aux différents points de grille, et permet de s'affranchir des contraintes CFL sur les termes d'advection au profit de contraintes plus faibles. De plus, les équations du noyau dynamique d'AROME sont intégrées avec une méthode semi-implicite, désormais basée sur un schéma ICI (Iterative Implicit scheme) à deux étapes, approchant le schéma de Crank-Nicolson. Un opérateur linéaire est ainsi introduit, et traite les termes responsables de la propagation des ondes les plus rapides en résolvant un problème implicite d'Helmholtz, alors que les autres termes qui constituent le résidu non-linéaire, sont traités explicitement. Cela permet de contourner une autre partie des contraintes CFL les plus fortes, et les équations sont ainsi intégrées avec un grand pas de temps, limitant le nombre d'étapes intermédiaires pour fournir une prévision à une échéance donnée. Un état de base particulièrement simple est prescrit pour l'opérateur linéaire, de sorte que le problème implicite est à coefficients constants. Les parties verticale et horizontale  $y$  sont séparées, autorisant ainsi plusieurs choix de discrétisation pour la partie horizontale. Dans la version actuelle du modèle AROME, une discrétisation en différences finies au second ordre est utilisée sur la verticale, alors que les champs sont décomposés en série de Fourier suivant l'horizontale via un algorithme de transformée de Fourier rapide. Dans cet espace spectral, le problème d'Helmholtz  $y$  est trivialement résolu.

Une partie des calculs étant effectuée en points de grille, alors que l'autre l'étant dans l'espace de Fourier (ou espace spectral), une transformée et son inverse, sont nécessaires pour passer de l'un à l'autre de ces espaces, à chaque étape du schéma ICI. Lors de ces transformées, le vecteur d'état est transposé, générant des communications globales peu adaptées aux machines massivement parallèles, et posant ainsi un problème de scalabilité.

De plus, les schémas à coefficients constants utilisés reposent notamment sur un traitement explicite des termes orographiques, soumis par conséquent à des contraintes de stabilité. Or, la contribution de ces termes augmente à mesure que la résolution du modèle s'affine, et donc que les plus fortes pentes des reliefs y sont notamment mieux représentées. Ainsi, pour une résolution horizontale de quelques centaines de mètres, le modèle AROME devient numériquement instable en présence de ces fortes pentes.

L'opérateur linéaire à coefficients constants actuellement utilisé est au cœur des deux problèmes rencontrés pour les évolutions futures du modèle : faible scalabilité, et instabilité aux fortes pentes. Il apparaît donc comme justifié d'entreprendre une révision à la fois de la forme et de la méthode d'inversion de l'opérateur dans une même étude, ce qui est précisément l'objet de ce travail de thèse.

Le premier objectif de la thèse est donc de développer une alternative scalable au modèle AROME où tous les calculs y sont effectués en points de grille, évitant ainsi les communications coûteuses des transformées. Les opérateurs de dérivées horizontales seront discrétisés en différences finies sur une grille co-localisée, et le problème implicite sera résolu par un solveur en points de grille. Pour cet objectif, l'approche à coefficients constants sera maintenue. Le second objectif de la thèse, est d'étudier différentes stratégies d'amélioration de la stabilité numérique en présence de fortes pentes. Des solutions simples seront testées en conservant tout d'abord un problème implicite à coefficients constants, puis en lui ajoutant les termes orographiques, le rendant ainsi à coefficients variables. Dans le cas d'un schéma à coefficients variables, les parties verticale et horizontale de l'opérateur linéaire n'étant plus séparables, la discrétisation spectrale ne peut être maintenue à un coût abordable. Comme pour le premier objectif, des approches de résolution en points de grille seront alors requises.

De nombreux autres modèles de PNT sont confrontés à des problèmes similaires. Dans le chapitre 1, nous exposerons les caractéristiques principales de plusieurs modèles de PNT à aire limitée, dont certaines d'entre elles sont communes avec AROME. Nous montrerons par exemple, que plusieurs modèles opérationnels semi-implicites semi-lagrangiens utilisent une discrétisation en différences finies sur une grille décalée, et un solveur itératif, préférentiellement de Krylov, pour résoudre le problème implicite. Certains d'entre eux utilisent également des schémas à coefficients constants, tandis que d'autres, utilisent des schémas à coefficients variables où l'état de base n'est pas prescrit et où les termes orographiques ont été parfois traités implicitement. Pour ces derniers, le traitement implicite des termes orographiques n'a pas significativement amélioré la stabilité numérique, et s'est même parfois traduit par une dégradation générale de la performance de la méthode. Une partie de ces éléments serviront à construire la version points de grille du modèle AROME, exposée dans le chapitre 3. Nous verrons également que les modèles Horizontalement Explicite / Verticalement Implicite (HEVI), soumis à de fortes contraintes de stabilité, sont néanmoins bien adaptés au paradigme de parallélisation actuel. Ils seront ainsi utilisés comme référence en terme de scalabilité dans la suite.

Dans le chapitre 2, nous exposerons en détail le noyau dynamique actuel du modèle AROME en nous focalisant plus particulièrement sur l'opérateur linéaire. Les limites de la version actuelle du modèle AROME y seront présentées en détail.

Dans le chapitre 3, une version en points de grille à coefficients constants du modèle AROME sera exposée. Les opérateurs de dérivée horizontale  $y$  seront discrétisés par des schémas en différences finies sur une grille co-localisée, et un solveur de Krylov inversera le problème implicite, comme la plupart des autres modèles résolvant des problèmes similaires. Néanmoins, aucun autre modèle parmi ceux exposés dans le premier chapitre ne possédant simultanément les mêmes caractéristiques que le modèle AROME, une étude spécifique est par conséquent requise. De plus, nous montrerons que le coût en communication de cette nouvelle version dépend fortement de la vitesse de convergence du solveur, permettant de réduire les estimations de scalabilité à de simples mesures d'efficacité. Ces différentes mesures seront comparées à celles des modèles HEVI exposés dans le premier chapitre et réputés particulièrement scalables. Nous vérifierons également que l'utilisation d'un solveur itératif non exact, ne dégrade pas la qualité par rapport à la version spectrale et aux observations, sur plusieurs cas, dont un épisode fortement précipitant à enjeux.

Dans le chapitre 4, des analyses seront menées pour déterminer si la stabilité numérique en présence de fortes pentes, peut être améliorée par la simple modification de quelques paramètres des schémas à coefficients constants et du modèle. Ensuite, nous ajouterons les termes orographiques dans le problème implicite, résultant ainsi en un problème à coefficients variables. Une version bidimensionnelle, sur un plan vertical du modèle AROME en points de grille, sera implémentée pour comparer la stabilité numérique de ces deux types de schémas. Différents écoulements sur des reliefs plus ou moins idéalisés seront testés, et permettront de quantifier les différents gains de stabilité obtenus pour chacun d'entre eux. Enfin, des estimations et mesures basiques de la vitesse de convergence seront menées.



# État de l'art

---

## Sommaire

---

<b>1.1</b>	<b>Prévision Numérique du Temps . . . . .</b>	<b>11</b>
<b>1.2</b>	<b>Moyens technologiques . . . . .</b>	<b>12</b>
<b>1.3</b>	<b>Problèmes scientifiques . . . . .</b>	<b>14</b>
1.3.1	Problème multi-échelle, multi-physique . . . . .	14
1.3.2	Stabilité et instabilité physique . . . . .	16
<b>1.4</b>	<b>Numérisation . . . . .</b>	<b>18</b>
1.4.1	Convergence numérique . . . . .	19
1.4.2	Maillage et coordonnée . . . . .	21
1.4.3	Grille et discrétisation . . . . .	23
1.4.4	Filtrage spatial et diffusion numérique . . . . .	27
1.4.5	Discrétisation explicite ou implicite . . . . .	27
1.4.6	Modèle eulérien ou semi-lagrangien . . . . .	29
1.4.7	Discrétisation temporelle des modèles Eulériens . . . . .	31
1.4.8	Discrétisation temporelle des modèles semi-lagrangiens . . . . .	32
1.4.9	Opérateur linéaire . . . . .	34
1.4.10	Discrétisation spectrale . . . . .	35
1.4.11	Pas de temps . . . . .	38
1.4.12	Exemple de modèles utilisés . . . . .	40
1.4.13	Enjeux des prochaines années . . . . .	41
<b>1.5</b>	<b>Conclusion . . . . .</b>	<b>42</b>

---

Dans ce chapitre, une introduction générale de la Prévision Numérique du Temps est exposée et une attention particulière est portée sur la partie dynamique.

## 1.1 Prévision Numérique du Temps

Soit  $\mathcal{N}$  un opérateur non-linéaire régissant le comportement du système atmosphérique, s'appliquant à un vecteur d'état  $\Psi$ . On appelle Prévision Numérique du Temps (PNT) la discipline qui intègre temporellement de  $t = 0$  à  $t = t_f$  le système dynamique

$$\frac{\partial \Psi}{\partial t} = \mathcal{N}(\Psi), \tag{1.1}$$

par des moyens numériques et qui trouve le meilleur compromis parmi des composantes clés identifiées, pour répondre à un objectif donné. Ces composantes (C1)–(C5) ont été listées par Mengaldo *et al.* (2018) :

- le temps à calculer la solution, (C1)
- l'énergie nécessaire à calculer la solution, (C2)
- la qualité de la solution, (C3)
- la robustesse du modèle, (C4)
- la quantification des incertitudes. (C5)

La connaissance de l'état initial  $\Psi(t = 0, \vec{x})$  est nécessaire et estimée par une discipline nommée *assimilation de données* (Carrassi *et al.*, 2018). La discipline *prévision d'ensemble* valorise les incertitudes portant sur l'état initial et l'opérateur  $\mathcal{N}$ , et traite notamment la composante (C5).

Lorsque le domaine géographique est la planète entière, le modèle est dit global. A contrario, lorsqu'il est restreint à une partie seulement, le modèle est dit à aire limitée (LAM) et nécessite l'imposition de conditions aux limites latérales fournies par un modèle global. Un modèle de surface impose la condition à la limite inférieure.

## 1.2 Moyens technologiques

Les moyens technologiques utilisés pour résoudre le problème de PNT ont fortement progressé depuis la première expérience concluante menée par Charney *et al.* (1950). Depuis l'apparition des premiers ordinateurs et pendant les plusieurs décennies qui s'ensuivirent, la puissance de calcul augmentait rapidement suivant la loi de Moore. Depuis quelques années néanmoins, la loi de Moore est de moins en moins respectée (Asanović *et al.*, 2006) et l'augmentation de la puissance de calcul passe d'une part par une augmentation du nombre de cœurs par processeur (CPU), et d'autre part par une augmentation du nombre de CPU travaillant en parallèles, ou par l'utilisation de nouvelles technologies pour le calcul scientifique, comme les cartes graphiques (GPU) *a priori* plus économes en énergie (C2) (Michalakes et Vachharajani, 2008, Müller *et al.*, 2013).

Dans cette étude, on modélise un supercalculateur comme un ensemble de nœuds de calculs reliés en réseau. On appelle coût en communication le coût d'un échange de données entre les différents nœuds de calcul. On considère que ce coût est régi : par la bande passante, c'est-à-dire le volume de communication qui peut transiter via le réseau de communication et par le temps de latence, c'est-à-dire le temps que prend le nœud pour recevoir une information depuis un autre nœud. Chaque nœud de calcul contient un certain nombre de processeurs reliés par un réseau interne et pour lequel le coût en communication est cette fois-ci considéré comme négligeable. Ces notions sont abordées plus en détail par Zheng et Marguinaud (2018). Une tâche est assignée à chaque nœud de calcul avec le protocole de parallélisation MPI (Barros *et al.*, 1995 donnent des détails de l'implémentation de ce protocole dans le modèle IFS). Le paradigme de parallélisation en point de grille actuel, découpe le domaine géographique de

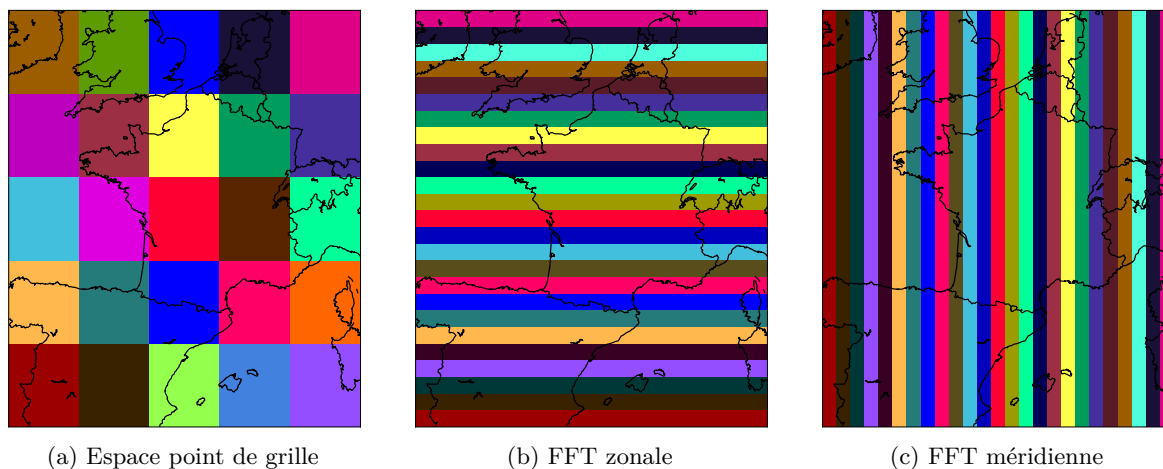


FIGURE 1.1 – Schéma du pavage de parallélisation du domaine. Chaque rectangle de couleur est traité par un nœud de calcul spécifique. Dans la pratique, 300 nœuds de calcul sont utilisés ; pour plus de lisibilité, seulement 25 nœuds sont représentés ici.

sorte que chaque nœud de calcul traite une partie du domaine géographique horizontal et les colonnes verticales correspondantes (Figure 1.1a).

Sur les machines actuelles, la performance totale d'un modèle dépend particulièrement de sa capacité à tourner dans un environnement très parallèle. On distingue deux types de scalabilité : la faible et la forte. La scalabilité faible consiste à mesurer le temps d'exécution d'un modèle lorsqu'on le fait tourner avec une configuration plus avancée (une résolution plus fine par exemple), et ce avec le même nombre de nœuds de calcul. La scalabilité forte consiste à mesurer le temps d'exécution du modèle pour une version donnée, en augmentant le nombre de nœuds de calcul. Des expériences de scalabilité sont régulièrement menées par de nombreux centres météorologiques en utilisant des supercalculateurs différents, comme par exemple par Müller et Scheichl (2014) et Müller *et al.* (2018).

Les modèles de PNT se complexifient régulièrement au gré des avancées de la recherche. Les plus fortes demandes en puissance de calcul sont généralement dues à un accroissement de la résolution, doublant environ tous les 5 ans (Wedi, 2014) et correspondant aussi à la période caractéristique du renouvellement d'un supercalculateur (Figure 1.2). Cela pose alternativement des problèmes de scalabilité faible (lorsque l'on double la résolution du modèle sur une machine donnée) et forte (lorsque l'on change de supercalculateur pour une version donnée du modèle). Par conséquent, nous visons à améliorer ces deux types de scalabilité et par la suite, on utilisera le terme scalabilité sans préciser s'il s'agit de la faible ou de la forte.

La plupart des centres météorologiques actuels disposent de puissants supercalculateurs. Par exemple, en 2020, Météo-France se dote de deux nouveaux supercalculateurs nommés Belenos et Taranis, disposant chacun d'un peu moins de 300 000 cœurs de calcul et développant une puissance maximale de plus de 10 PétaFlops. Cela représente un gain d'un facteur 5 en puissance de calcul par rapport aux précédents supercalculateurs (Figure 1.2). En juin 2020, Belenos occupe la 29<sup>ème</sup> place dans le Top500 qui répertorie les 500 supercalculateurs les plus



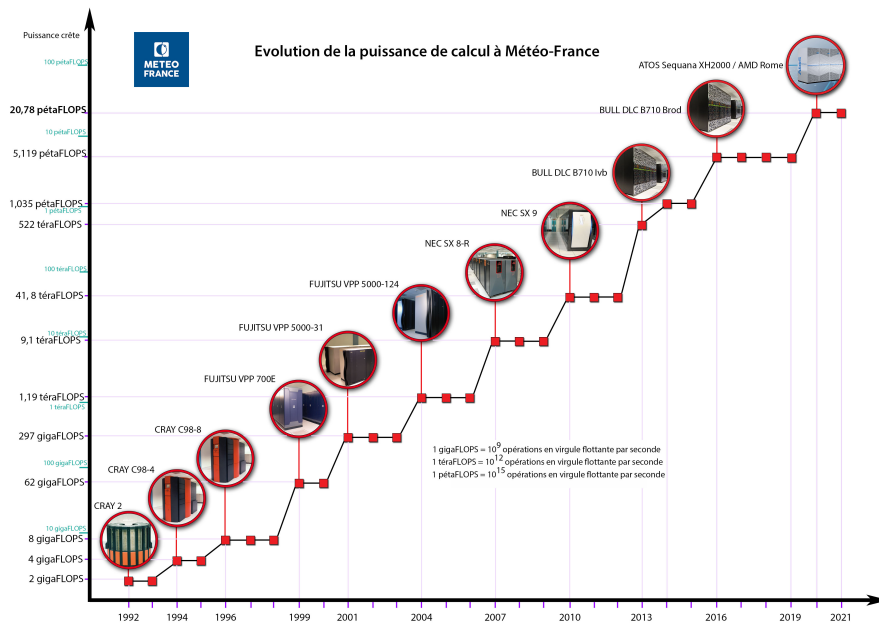


FIGURE 1.2 – Évolution de la puissance crête de calcul en fonction du temps à Météo-France : <http://www.meteofrance.fr/nous-connaître/activites-et-metiers/les-supercalculateurs-de-meteo-france>.

puissants au monde (<https://top500.org/>).

## 1.3 Problèmes scientifiques

### 1.3.1 Problème multi-échelle, multi-physique

L'étude du système atmosphérique fait intervenir de nombreuses disciplines de la physique à de multiples échelles : mécanique des fluides, thermodynamique, micro-physique, rayonnement, etc. Bauer *et al.* (2015) montrent que les modèles de PNT se sont améliorés par l'étude de chacune des composantes atmosphériques dont nous donnons ici un bref aperçu.

Comme nous le verrons ensuite, les équations sont résolues à une certaine résolution spatiale. Cette résolution sépare la partie des phénomènes résolus par le modèle, qui est appelée « dynamique », de la partie des phénomènes sous-maille qui sont paramétrisés (c'est-à-dire dont seul leur effet moyen dans la maille est considéré) via les « paramétrisations physiques ». Le raccord entre ces deux parties est fait pour garantir la « cascade d'énergie » discutée par exemple par Skamarock (2004). L'opérateur non-linéaire  $\mathcal{N}$  du système général (1.1) est alors décomposé en la somme d'une partie dynamique  $\mathcal{D}$  et d'une partie de paramétrisation physique  $\mathcal{P}$  :

$$\frac{\partial \Psi}{\partial t} = \mathcal{D}(\Psi) + \mathcal{P}(\Psi). \quad (1.2)$$

### 1.3. PROBLÈMES SCIENTIFIQUES

---

L'amélioration des descriptions de  $\mathcal{D}$  et de  $\mathcal{P}$  ou de leur couplage permet d'améliorer la qualité (C3). Or, certains phénomènes, comme la convection profonde, sont difficiles à paramétriser. Une alternative est de les traiter dans la partie dynamique en affinant la résolution horizontale.

Ainsi, pour éviter le délicat exercice de paramétrisation de la convection profonde, la plupart des centres météorologiques affinent fortement la résolution pour que ces phénomènes soient traités dans la partie dynamique. Cette raison a été l'une des motivations à l'émergence du projet AROME (Seity *et al.*, 2011) dont la résolution de 2,5 km, lors de son lancement en 2008, a permis de se passer de cette paramétrisation. Récemment, pour les mêmes raisons, des expériences à des résolutions comparables ont été menées sur certains modèles globaux (Stevens *et al.*, 2019), même si à l'heure actuelle, compte tenu des ressources de calcul disponibles, aucun modèle global en fonctionnement opérationnel n'atteint des résolutions inférieures à 5 km.

Les ressources en calcul limitées actuelles (et très probablement pour de nombreuses années encore), ne nous permettent pas de résoudre les équations de Navier-Stokes dans la partie dynamique  $\mathcal{D}$  et les petits tourbillons qui assurent la cascade d'énergie sont paramétrisés. L'enjeu est donc de choisir un système d'équations en phase avec la technologie de son époque. La première prévision numérique réussie menée par Charney *et al.* (1950) utilisait le système de conservation du tourbillon barotrope qui filtrait notamment les ondes de gravité et ne modélisait que les ondes de Rossby, en adéquation avec la technologie de l'époque. Depuis, pour éviter de filtrer des ondes d'intérêt météorologique, les équations d'Euler sont désormais résolues. Néanmoins, elles autorisent la propagation des ondes acoustiques, peu utiles en météorologie et de plus, fortement pénalisantes numériquement. Ainsi, des systèmes d'équations intermédiaires ont émergé, pour filtrer ces ondes, comme le système anélastique (une formulation est par exemple donnée par Lipps et Hemler (1982)). Cette approximation ralentit les ondes de Rossby (Davies *et al.*, 2003) et dégrade la qualité de la simulation à l'échelle synoptique mais elle est cependant parfois encore utilisée dans les modèles de méso-échelle (Lac *et al.*, 2018). Depuis, des systèmes d'équations plus sophistiqués, filtrant seulement les ondes acoustiques sans subir les effets indésirables de l'approximation anélastique, sont étudiés (voir par exemple Voitus *et al.* (2019)).

Que le système soit filtré ou non, il est possible d'effectuer l'approximation hydrostatique (H) pour les phénomènes dont l'échelle caractéristique horizontale est supérieure à la hauteur caractéristique de l'atmosphère (soit environ 10 km). Lorsque cette approximation est effectuée, la vitesse verticale est déduite des autres variables du modèle. La plupart des modèles globaux actuels dont la résolution horizontale est de cet ordre de grandeur, résolvent les équations d'Euler en faisant cette approximation. L'approximation hydrostatique filtre notamment les ondes acoustiques (à l'exception de l'onde de Lamb qui se propage uniquement suivant l'horizontale) et simplifie considérablement la résolution numérique de la partie dynamique. Pour une résolution horizontale plus fine que 10 km, quelques différences apparaissent entre un modèle faisant l'approximation hydrostatique et un modèle ne la faisant pas. Par exemple, à une résolution de 8 km, quelques différences sur les champs de pluie sont perceptibles suite aux expériences menées par Janjic *et al.* (2001). Pour des résolutions encore plus fines, de l'ordre du kilomètre, les vitesses verticales sont sur-estimées dans les zones de convection profonde

lorsque cette approximation est effectuée en l'absence de paramétrisation (Wedi et Malardel, 2010). D'autres expériences sur des cas idéalisés ont montré que cette approximation filtre une partie des ondes orographiques dans les atmosphères stratifiées (Wedi et Smolarkiewicz, 2009). Toutefois, comme nous le montrerons dans le paragraphe 1.4.11, la qualité de cette représentation dépend également de la résolution temporelle : un pas de temps trop grand conduit à ne représenter que partiellement la contribution linéaire Non-Hydrostatique (NH).

La mauvaise représentation des ondes de gravité d'une part, et des vitesses verticales d'autre part, dégradent la qualité des prévisions des phénomènes à enjeux comme les tempêtes de pentes aval ou la convection profonde, pour lesquels la contribution non-hydrostatique est majeure. Ainsi, contrairement aux modèles globaux, les modèles à aire limitée (LAM) sont majoritairement non-hydrostatiques. Le passage des équations, faisant l'approximation hydrostatique, appelées aussi équations primitives (HPE en anglais), aux équations NH a probablement été l'un des enjeux principal de l'amélioration de la partie dynamique ces dernières années pour les LAM et le sera aussi probablement pour les modèles globaux, pour lesquels les premières expériences en NH ont été menées par exemple lors du projet DYAMOND (Stevens *et al.*, 2019).

### 1.3.2 Stabilité et instabilité physique

#### Présentation générale

L'un des objectifs d'un modèle de PNT est de prévoir l'évolution spatio-temporelle de différentes perturbations de l'atmosphère, par rapport à un état moyen. Certaines de ces perturbations vont croître (et sont dites instables) et d'autres décroître (et sont dites stables) en fonction des propriétés physiques ou dynamiques de l'écoulement de base. De plus, elles sont limitées dans le temps et en amplitude. L'une des instabilités les plus connues en météorologie, est l'instabilité barocline, à l'origine du creusement des dépressions des moyennes latitudes. Ces questions ont été abondamment étudiées au sein de la discipline *météorologie dynamique*. La notion d'instabilité est toutefois dépendante de la norme considérée (enstrophie, énergie cinétique, etc) pour quantifier les taux de croissance des instabilités. Skamarock (2004) introduit par exemple les spectres d'énergie cinétique pour quantifier les échanges d'énergie entre les échelles.

#### Formalisme

Une définition plus formelle de la notion de stabilité est donnée ci-dessous en considérant uniquement la contribution de la partie dynamique du système (1.2). Soit  $\mathcal{L}_D$  le tangent-linéaire de l'opérateur non-linéaire  $\mathcal{D}$ , linéarisé autour d'un état de base  $\bar{\Psi}$  stationnaire, le système (1.2) devient alors :

$$\frac{\partial \Psi'}{\partial t} = \mathcal{L}_D \Psi' \quad (1.3)$$

### 1.3. PROBLÈMES SCIENTIFIQUES

---

où  $\Psi'$  est un vecteur d'état qui contient les  $p$  variables pronostiques du système. Décomposons le champ  $\Psi'$  comme dans Bénard (2003), c'est-à-dire sous la forme :

$$\Psi'(\vec{r}, t) = \widehat{\Psi}'(t)f(\vec{r}), \quad (1.4)$$

où  $f$  et  $\widehat{\Psi}$  sont respectivement la structure et la polarisation du mode,  $t$  et  $\vec{r}$  désignent respectivement la dépendance temporelle et spatiale. Le système (1.3) se réécrit alors sous la forme :

$$\frac{d\widehat{\Psi}'}{dt} = \widehat{\mathcal{L}}_D(\widehat{\Psi}'), \quad (1.5)$$

où  $\widehat{\mathcal{L}}_D$  est exprimé dans base de  $f$  et en le diagonalisant, on obtient :

$$\frac{d\widetilde{\Psi}'_j}{dt} = \omega_j \widetilde{\Psi}'_j, \quad (1.6)$$

où  $\omega_j$  est la  $j^{\text{ème}}$  valeur propre de l'opérateur  $\widehat{\mathcal{L}}_D$  et  $\widetilde{\Psi}$  est la projection de  $\widehat{\Psi}$  dans  $\widehat{\mathcal{L}}_D$ .

Un mode est physiquement stable si et seulement si

$$\forall j \in [1, p], \omega_j \in i\mathbb{R} \quad (1.7)$$

où  $i$  est le nombre imaginaire tel que  $i^2 = -1$ . De plus, un mode est instable si  $\Re(\omega_j) > 0$  ou amorti si  $\Re(\omega_j) < 0$ .

Les modes stables ont été abondamment étudiés car leurs fréquences d'oscillations temporelles (données par la pulsation  $\omega$ ) sont reliées à ses fréquences d'oscillations spatiales (donnée par ses nombres d'onde horizontal  $k$  et vertical  $m$ ) via une relation de dispersion, dont nous donnons dans le paragraphe suivant une liste non exhaustive.

#### Relations de dispersion usuelles

Les équations d'Euler admettent la propagation de plusieurs ondes couplées entre elles, comme nous le verrons dans le chapitre 2. Ici, on donne seulement quelques exemples d'ondes non couplées pour la plupart, en les listant de la plus lente à la plus rapide :

- les ondes de Rossby dont la relation de dispersion dans une atmosphère de vent uniforme  $U$  est de la forme :

$$\omega - iUk = -i\beta \frac{k}{k^2 + m^2} \quad (1.8)$$

où  $k$  et  $m$  sont respectivement les nombres d'onde zonal et vertical et  $\beta$  est le paramètre de Rossby et est défini comme  $\beta = \frac{\partial f}{\partial y}$  où  $f = 2\Omega \sin \phi$  est le paramètre de Coriolis,  $\Omega$  est la vitesse angulaire de rotation de la Terre sur elle-même et  $\phi$  est la latitude ;

- les ondes d'inertie-gravité dont la relation de dispersion dans une atmosphère de hauteur caractéristique  $H_0$ , de fréquence de Brunt-Väisälä uniforme  $N_b$ , de vent uniforme  $U$ , est

de la forme :

$$\omega - iUk = i\sqrt{f^2 + \frac{N_b^2}{k^2 + m^2 + 1/(4H_0^2)}}k^2, \quad (1.9)$$

qui peut se simplifier pour une onde de gravité pure en :

$$\omega - iUk = i\frac{N_b k}{\sqrt{k^2 + m^2 + 1/(4H_0^2)}}, \quad (1.10)$$

et pour une onde de gravité pure avec approximation hydrostatique :

$$\omega - iUk = i\frac{N_b k}{\sqrt{m^2 + 1/(4H_0^2)}}, \quad (1.11)$$

où  $H_0$  est la hauteur caractéristique de l'atmosphère ;

- les ondes acoustiques dont la relation de dispersion dans une atmosphère de vent uniforme  $U$  de température  $T$  est de la forme :

$$\omega - iUk = i\sqrt{\gamma RT(k^2 + m^2)} \quad (1.12)$$

où  $\gamma = C_p/C_v$ , où  $C_p$  et  $C_v$  désignent respectivement la capacité calorifique de l'air à pression et volume constants.

Notons que les ondes les plus rapides sont les ondes acoustiques dont la vitesse de phase  $c$  est :

$$c = \sqrt{\gamma RT}. \quad (1.13)$$

Notons également que certains modes « externes » des ondes de gravité peuvent atteindre des vitesses semblables ( $c \simeq \sqrt{gH_0}$  où  $H_0$  est la hauteur caractéristique de l'atmosphère).

## Séparation en processus lents et rapides

L'atmosphère fait donc intervenir des ondes qui se propagent lentement, comme les ondes de Rossby, et d'autres qui se propagent rapidement comme les ondes acoustiques et les modes les plus externes des ondes de gravité. Par conséquent, le problème (1.2) peut être décomposé en une composante lente et une composante rapide. Dans l'étape de numérisation, nous verrons que cette séparation est abondamment exploitée pour construire les schémas numériques les plus efficaces possibles.

## 1.4 Numérisation

Les moyens numériques permettent de résoudre le problème (1.2) en utilisant efficacement la technologie. L'étape de numérisation est soumise à un ensemble de contraintes que nous exposons. Cette étape consiste tout d'abord à se doter d'un maillage et d'un système de coor-

## 1.4. NUMÉRISATION

---

données pour discrétiser spatialement et temporellement le système à résoudre, en respectant au mieux les composantes clés (C1)–(C5) d’un modèle de PNT.

Les modèles à aire limitée AROME, UK-UM, WRF-HRRR, COSMO, GEM et ASUCA illustreront les différents choix numériques exposés ci-après. Les informations sont issues de données publiées ou rendues publiques avant fin 2020 :

- pour le modèle AROME : Seity *et al.* (2011), Brousseau *et al.* (2016) ;
- pour le UK-UM : Wood *et al.* (2014), Benacchio et Wood (2016), Walters *et al.* (2017), Bush *et al.* (2020) ;
- pour le modèle GEM-LAM : Qaddouri et Lee (2010), Qaddouri et Lee (2011) ;
- pour le modèle HRRR utilisant le noyau dynamique du modèle WRF : Skamarock *et al.* (2019) ;
- pour le modèle COSMO utilisé notamment par le DWD et Meteo-Swiss : Doms *et al.*, 2011
- pour le modèle ASUCA : Saito *et al.* (2006) et plus récemment par Ishida *et al.* (2010).

### 1.4.1 Convergence numérique

#### Consistance et stabilité

Les équations généralement utilisées en PNT (1.2) ne sont pas intégrables analytiquement. On procède donc à une étape de numérisation, en discrétisant les équations du système continu, spatialement et temporellement. Or, comme nous l’illustrerons dans le paragraphe (1.4.3) pour la discrétisation spatiale, et dans le paragraphe (1.4.5) pour la discrétisation temporelle, différents schémas peuvent être utilisés. L’une des propriétés à assurer est que la solution du système discret converge vers la solution du système continu. Ainsi certains schémas ou certaines combinaisons de schémas, ne convergent pas et aboutissent à une solution discrète inexploitable et, dans le cas qui nous intéresse, à un échec de la prévision météorologique. Plusieurs schémas numériques et leur condition d’utilisation sont présentés par exemple par Durran (2010).

Dans la pratique, la convergence d’un schéma est une mesure globale, et, est par conséquent difficile à évaluer. Toutefois, Lax et Richtmyer (1956) ont montré que pour un problème bien posé, un schéma numérique est convergent si et seulement s’il est stable et consistant. Comme indiqué par exemple par Allaire (2005), un schéma est consistant si lorsque les pas de discrétisation spatiale  $\Delta x$  et temporelle  $\Delta t$  tendent vers zéro, l’erreur entre les solutions des équations continues et discrètes tend vers zéro. Un schéma est numériquement stable, si pour un système physiquement stable et une norme donnés, l’état du système ne s’amplifie pas (au sens de cette norme) entre les instants  $t$  et  $t + \Delta t$ . L’analyse de Von Neumann est un cadre simple d’évaluation de la stabilité numérique, en ne considérant qu’un seul mode physiquement stable, et en trouvant les conditions pour lesquelles son facteur d’amplification est inférieur à l’unité. En PNT, cette analyse a été menée par Charney *et al.* (1950) pour l’équation du tourbillon barotrope. La contrainte de stabilité numérique la plus connue est

probablement la contrainte CFL qui émerge notamment pour la classe des schémas explicites (Courant *et al.*, 1967) dont nous donnons quelques exemples dans le paragraphe suivant.

### Contraintes CFL

Lorsque les équations sont discrétisées par des schémas explicites, des contraintes CFL peuvent être dérivées à partir des relations de dispersion (1.8)–(1.12). Les phénomènes les plus contraignants numériquement sont ceux dont la vitesse caractéristique est la plus forte : les modes externes des ondes de gravité ou les ondes acoustiques (de vitesse de phase  $c$ ), les jets d'altitude les plus forts (de vitesse horizontale de vent  $U_{max}$ ), ou les zones de convection les plus intenses (de vitesse verticale de vent  $W_{max}$ ). Cela conduit à définir les quatre contraintes de stabilité suivantes, classées de la moins contraignante à la plus contraignante :

$$\frac{U_{max}\Delta t}{\Delta x} \leq \alpha_1, \quad (1.14)$$

$$\frac{c\Delta t}{\Delta x} \leq \alpha_2, \quad (1.15)$$

$$\frac{W_{max}\Delta t}{\Delta z} \leq \alpha_3, \quad (1.16)$$

$$\frac{c\Delta t}{\Delta z} \leq \alpha_4. \quad (1.17)$$

Les facteurs  $\alpha_i$  dépendent de la grille, du maillage, des schémas de discrétisations utilisés, etc. Pour fixer les idées dans un premier temps, on peut considérer que  $\alpha_i \simeq 1$ . Les contraintes (1.14) et (1.16) qui portent sur le vent maximal sont appelées « CFL-vent » alors que les contraintes (1.15) et (1.17) sont appelées « CFL-ondes ».

Le maillage utilisé est en général fortement anisotrope entre les directions horizontale et verticale. En 2020, la plupart des LAM atteignent une résolution horizontale kilométrique  $\Delta x \simeq 1000$  m qui est significativement plus grande que la résolution verticale de sa plus petite résolution verticale (la hauteur de son premier niveau dans la pratique)  $\Delta z \simeq 5$  m. De plus, les vents les plus forts rencontrés sur Terre ( $U_{max} \simeq 100$  m s<sup>-1</sup>), principalement en haut de la troposphère, sont environ 3 fois plus faibles que la vitesse des ondes acoustiques et de gravité ( $c \simeq 350$  m s<sup>-1</sup>). Les vitesses verticales les plus fortes dans l'atmosphère sont généralement atteintes dans les zones de convection profonde, et sont dans les cas les plus extrêmes, de l'ordre de  $W_{max} \simeq 50$  m s<sup>-1</sup>. Ainsi, pour les résolutions actuelles des LAM, les ordres de grandeur des pas de temps maximum admis sont les suivants :

$$\Delta t \leq \alpha_1 * 10 \text{ s}, \quad (1.18)$$

$$\Delta t \leq \alpha_2 * 2,9 \text{ s}, \quad (1.19)$$

$$\Delta t \leq \alpha_3 * 0,10 \text{ s}, \quad (1.20)$$

$$\Delta t \leq \alpha_4 * 0,014 \text{ s}, \quad (1.21)$$

où  $*$  désigne la multiplication traditionnelle. Ces quatre contraintes sont classées par ordre croissant de sévérité : la dernière d'entre elles (1.17) est la plus pénalisante pour la discrétisation. Néanmoins, on note que dans certains modèles, la convection profonde peut être mal représentée, de sorte que les vitesses verticales modélisées  $W_{max}$  sont beaucoup plus faibles que celles observées en réalité. Par conséquent, la contrainte CFL sur le vent vertical (1.16) peut finalement être moins contraignante que celle sur la propagation des ondes rapides sur l'horizontale (1.15).

### Opérateurs non-linéaires

On note que la consistance et la stabilité numérique sont généralement étudiées dans le cadre d'équations linéaires, alors que les équations résolues en PNT sont non-linéaires (1.2). Les études de consistance ne peuvent pas être menées facilement dans ce cadre non-linéaire. De plus, la restriction au cas linéaire, conduit à n'étudier qu'une partie de la stabilité numérique : la stabilité numérique linéaire. Cette instabilité est probablement la plus pénalisante en PNT car les perturbations numériques linéaires croissent exponentiellement en temps et masquent le signal physique : il s'agit d'une « explosion numérique ». Cela aboutit finalement à une interruption de la simulation (due à une erreur arithmétique) après seulement quelques pas de temps en général. Notons toutefois que dans un modèle 3D, les différents filtrages, la diffusion numérique et les paramétrisations physiques, peuvent masquer ou retarder l'apparition de certaines instabilités. D'autres instabilités numériques peuvent survenir lors du traitement de certains termes non linéaires, comme les phénomènes d'aliasing, dont le filtrage est discuté dans le paragraphe 1.4.4. Notons que la stabilité numérique d'un modèle est une composante essentielle pour assurer sa robustesse (C4), et sera le point principal de cette thèse.

### 1.4.2 Maillage et coordonnée

#### Sur l'horizontale

Plusieurs stratégies de maillage peuvent être considérées pour la direction horizontale : maillage structuré, ou non structuré. Les maillages structurés ont l'avantage d'être particulièrement simples à mettre en œuvre mais ne sont pas toujours adaptés à la géométrie à mailler. Ainsi lorsque la géométrie est une sphère, la grille (structurée) Gaussienne présente un maillage nettement plus fin au voisinage des pôles, conduisant à un surcout de calcul d'une part, et à un problème de singularité pour calculer certains opérateurs de dérivée horizontale, d'autre part. Ainsi, la grille Gaussienne a été progressivement remplacée par une grille réduite (semi-structurée), moins coûteuse au voisinage des pôles (Hortal et Simmons, 1991). De plus, des travaux ont été entrepris par Bénard et Glinton (2019) pour traiter le problème de singularité des opérateurs de dérivée horizontale aux pôles.

Le maillage est une étape qui conditionne fortement certains grands choix de discrétisation. Par exemple, un maillage non structuré sera rarement associé à une discrétisation en différence finies, qui ne gère pas facilement les changements irréguliers de la taille des mailles. La



discrétisation en différence finies se fait généralement sur des maillages structurés pour lesquels les opérateurs de dérivée horizontale peuvent être discrétisés à ordre élevé et éviter certains comportements pénalisants qui apparaissent aux ordres faibles (paragraphe 1.4.3).

A contrario, la discrétisation en volumes finis est compatible avec une grande variété de maillages. Le nouveau noyau dynamique FV3 du modèle GFS, maille la sphère selon la technique de la « Cubed-Sphere grids », permettant de se ramener à 6 géométries planes (les 6 faces du cube) mais générant des problèmes de raccordement aux extrémités de chacun des plans (Putman et Lin, 2007). D'autres ont préféré une grille icosahédrique comme, par exemple, les modèles ICON (Zängl *et al.*, 2014) et DYNAMICO (Dubos *et al.*, 2015) ou encore une grille Yin-Yang pour le modèle GEM (Qaddouri et Lee, 2011). Le choix de ces différentes grilles est, par exemple, résumé par Staniforth et Thuburn (2011). Pour amoindrir la dépendance de la discrétisation au maillage, le modèle en cours de développement FVM du CEPMMT est construit pour fonctionner avec n'importe quel maillage (Kühnlein *et al.*, 2019).

Dans le cas d'un modèle à aire limitée sur une grille plane, le choix du maillage est plus immédiat et un maillage cartésien est souvent utilisé.

## Sur la verticale

Le paradigme de parallélisation actuel privilégie les maillages structurés pour mailler la direction verticale. La plupart des modèles opérationnels optent pour l'utilisation de ce type de maillage, y compris les modèles en cours de développement comme le modèle FVM (Kühnlein *et al.*, 2019).

Dans la pratique, la vraie hauteur appelée  $z$  n'est pas utilisée comme coordonnée verticale. Celle-ci conduirait à imposer difficilement des conditions aux limites. On lui préfère une coordonnée qui épouse la forme du relief, on dit aussi qu'elle « suit le terrain », que nous appellerons  $\bar{z}$ . En appliquant les règles de transformation de coordonnée définies par Kasahara (1974), la relation suivante peut être établie :

$$(\nabla\Psi)_z = (\nabla\Psi)_{\bar{z}} + G \frac{\partial\Psi}{\partial\bar{z}} \quad (1.22)$$

Par la suite, le terme :  $G = [\partial_x\bar{z}, \partial_y\bar{z}]^T$  est un terme orographique dont le traitement sera abondamment discuté. Lorsqu'un relief purement linéaire est considéré,  $G$  est la pente. Les modèles globaux ont majoritairement adopté la pression comme coordonnée verticale. Laprise (1992) a depuis étendu cette coordonnée aux systèmes d'équations non hydrostatiques, pour lesquels seule la composante hydrostatique de la pression sert à définir la coordonnée verticale, et est utilisée pour les modèles AROME, WRF, GEM-LAM. D'autres (UK-UM, COSMO, ASUCA) utilisent une coordonnée hauteur de type Gal-Chen et Somerville (1975). Notons que ces différentes coordonnées verticales sont des coordonnées hybrides : elles épousent l'orographie en bas de l'atmosphère et, leurs fortes variations horizontales dues au relief s'atténuent avec l'altitude.

### 1.4.3 Grille et discrétisation

Une fois le maillage choisi, l'endroit où sont effectués les calculs reste à déterminer : au centre de la maille, au milieu des arêtes de chacune des mailles, au sommet de chaque maille, etc. Un sommet d'une maille est appelé un « point de grille ». De plus, des endroits différents peuvent être considérés pour chacune des variables pronostiques. En particulier, lorsque les variables de flux sont calculées au milieu des arêtes alors que les variables de masse sont calculées sur les points de grille, on dit qu'on utilise une grille C (ou grille décalée). Cette grille est naturellement utilisée dans les approches à volumes finis où les variables pronostiques sont déterminées à partir des flux qui traversent les bords (et calculés en général au milieu des arêtes). A contrario, lorsque les variables pronostiques sont toutes calculées sur les points de grille, on dit qu'on utilise une grille A (ou grille co-localisée).

Les opérateurs de dérivée (première et seconde) horizontale dépendent de la grille de discrétisation. Un exemple de ces opérateurs est donné ci-dessous dans le cas d'un domaine 1D périodique et d'un maillage cartésien uniforme de résolution  $\Delta x$ . Une discrétisation via une décomposition en série de Fourier (appelée discrétisation « spectrale ») est tout d'abord exposée. Ensuite, une discrétisation en différences finies basée sur les développements en série de Taylor à l'ordre 2 pour une grille A et une grille C est présentée.

#### Discrétisation horizontale spectrale sur grille A

Lorsque le domaine est périodique (ou transformé en un domaine périodique par une technique appropriée), le champ  $\Psi$  peut alors être décomposé via une série de Fourier de la forme :

$$\Psi = \sum_{n_k=-K}^K \widehat{\Psi}_k e^{ikx}, \quad (1.23)$$

où  $k$  est le nombre d'onde zonal défini comme

$$k = \frac{2\pi n_k}{L_x}, \quad (1.24)$$

et  $K$  est la troncature spectrale du modèle (dont le choix est notamment discuté dans le paragraphe 1.4.10, pour éviter certains problèmes numériques),  $\widehat{\Psi}_k$  est le coefficient de Fourier associé au nombre d'onde  $k$ . Dans cette base, le calcul de la dérivée première est exact :

$$\partial_x \Psi = \sum_{n_k=-K}^K ik \widehat{\Psi}_k e^{ikx}, \quad (1.25)$$

tout comme celui de la dérivée seconde :

$$\partial_x^2 \Psi = \sum_{n_k=-K}^K -k^2 \widehat{\Psi}_k e^{ikx}. \quad (1.26)$$

La décomposition en série de Fourier sera utilisée ci-dessous pour évaluer les erreurs de la discrétisation en différences finies sur une relation de dispersion.

### Discrétisation horizontale en différences finies sur grille A

Dans ce paragraphe on discrétise les opérateurs de dérivée horizontale à l'ordre 2. La discrétisation d'une dérivée première est de la forme :

$$\delta_x \Psi = \frac{\Psi(x + \Delta x) - \Psi(x - \Delta x)}{2\Delta x} + O(\Delta x^2), \quad (1.27)$$

où les erreurs varient comme  $O(\Delta x^2)$ . La discrétisation d'une dérivée seconde que l'on note  $\delta_x \delta_x$  est de la forme :

$$\delta_x \delta_x \Psi = \frac{\Psi(x + 2\Delta x) - 2\Psi(x) + \Psi(x - 2\Delta x)}{4\Delta x^2} + O(\Delta x^2), \quad (1.28)$$

On note que dans ce cas particulier, le calcul de  $\delta_x \delta_x \Psi$  ne dépend pas de  $\Psi(x + \Delta x)$  et  $\Psi(x - \Delta x)$ . Les coefficients de ces différences finies peuvent être synthétisés sous la forme suivante :

$$\left( \frac{1}{4}, 0, -\frac{1}{2}, 0, \frac{1}{4} \right), \quad (1.29)$$

où le premier élément désigne le coefficient multiplicateur (normalisé par  $\Delta x^2$ ) de  $\Psi(x + 2\Delta x)$ , le second celui de  $\Psi(x + \Delta x)$ , etc.

### Discrétisation horizontale en différences finies sur grille C

Dans ce paragraphe on discrétise les opérateurs de dérivée horizontale à l'ordre 2. La discrétisation d'une dérivée première sur une grille C est de la forme :

$$\delta_x \Psi = \frac{\Psi(x + \Delta x/2) - \Psi(x - \Delta x/2)}{\Delta x} + O(\Delta x^2), \quad (1.30)$$

où les erreurs varient comme  $O(\Delta x^2)$ . La discrétisation d'une dérivée seconde sur une grille C que l'on note  $\delta_x^2$  est de la forme :

$$\delta_x^2 \Psi = \frac{\Psi(x + \Delta x) - 2\Psi(x) + \Psi(x - \Delta x)}{\Delta x^2} + O(\Delta x^2). \quad (1.31)$$

Les coefficients de ces différences finies peuvent être synthétisés sous la forme suivante :

$$(1, -2, 1). \quad (1.32)$$

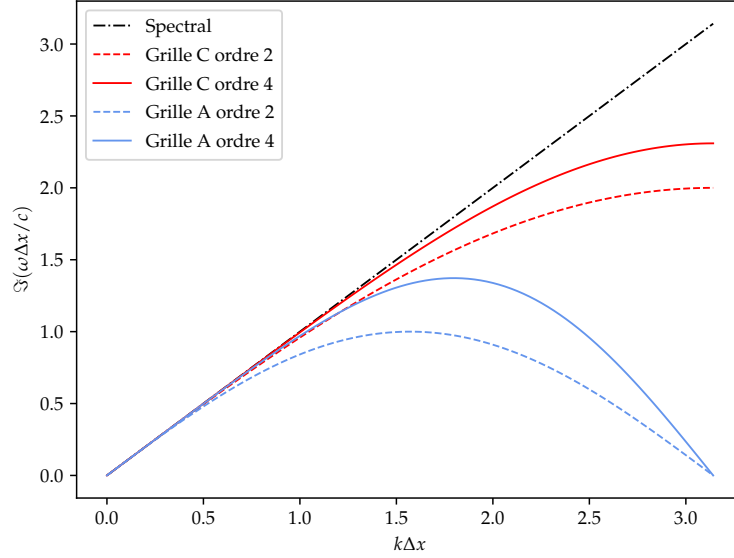


FIGURE 1.3 – Évolution de la pulsation normalisée en fonction du nombre d'onde normalisé  $k\Delta x$  pour différentes grilles horizontales et différents ordres de discrétisation.

### Répartition des erreurs de discrétisation

Pour fournir une prévision de qualité suffisante (C3), l'étape de numérisation doit dégrader le moins possible la solution du système continu. Un cadre simple pour quantifier cette détérioration est de mesurer les effets des différents schémas de discrétisation sur l'une des relations de dispersion précédemment exposées.

On considère ici un système shallow-water (avec seulement  $p = 2$  variables pronostiques, en général le vent et la pression) dont l'état de base est au repos ( $U = 0$ ) sur un domaine horizontal périodique et un maillage cartésien uniforme. Ce système admet la propagation d'un mode rapide (acoustique ou de gravité) de vitesse de phase  $c$  et qui satisfait la relation de dispersion (1.12) dans le cas où la discrétisation spatiale est exacte.

Lorsque la discrétisation spatiale n'est pas exacte, on peut calculer la relation de dispersion en calculant le spectre des opérateurs de discrétisation de dérivées horizontales et la relation (1.12) devient alors :

- pour une grille A (les deux variables sont calculées sur les points de grille) :

$$\omega = \pm ic \frac{\sin(k\Delta x)}{\Delta x}; \quad (1.33)$$

- pour une grille C (la pression est calculée sur les points de grille et le vent au milieu de la cellule) :

$$\omega = \pm ic \sqrt{2 \left( \frac{1}{\Delta x} - \frac{\cos(k\Delta x)}{\Delta x} \right)}. \quad (1.34)$$

Ces relations de dispersion sont représentées sur la Figure 1.3. Dans les deux cas, ces deux relations de dispersion tendent vers la relation de référence (1.12) traduisant un bon comportement aux grandes échelles (c'est-à-dire pour  $k\Delta x \rightarrow 0$ ). En revanche, le comportement est significativement différent à plus petite échelle. En particulier, il y a annulation de la vitesse de groupe ( $\partial_k \omega = 0$ ) pour  $k\Delta x = \pi/2$  (à l'ordre 2) lors de l'utilisation d'une grille A, empêchant ainsi la cascade d'énergie de s'effectuer et conduisant à des accumulations ponctuelles d'énergie nécessitant d'être filtrées aux échelles pour  $k\Delta x \in [\pi/2, \pi]$ . Toutefois, comme montré sur la Figure 1.3, ce problème est relégué en fin de spectre à mesure que l'ordre de discrétisation augmente et disparaît lorsque la discrétisation est exacte, comme vu précédemment, avec la discrétisation spectrale. La grille C ne présente pas ces problèmes y compris pour des ordres de discrétisation faibles. Par conséquent, alors que le modèle spectral AROME privilégie la grille A, la plupart des modèles en point de grille privilégient les grilles décalées.

### Discrétisations alternatives

Depuis des alternatives ont été explorées pour utiliser une grille co-localisée en évitant ce comportement potentiellement néfaste. Par exemple, l'approche dite en « grille Z », consiste à utiliser la divergence horizontale du vent à la place du vent horizontal pour éliminer les dérivées première, et ne considérer que des dérivées secondes dans le système d'équations. Les opérateurs (de dérivée seconde) peuvent ainsi être discrétisés comme pour une grille C (1.31), évitant ainsi les effets indésirables de la grille A (ces aspects sont par exemple approfondis par Caluwaerts (2016)). Toutefois dans un modèle complet, par exemple lors de l'appel au schéma d'advection, il est nécessaire de revenir au calcul du vent, faisant ainsi apparaître des opérateurs de dérivée première. Un traitement inconsistant peut être à l'origine de comportements numériques néfastes, comme nous le montrerons dans le paragraphe 1.4.3. Une alternative, respectant des propriétés de symétrie lors de la discrétisation des opérateurs, a ainsi été développée par Caluwaerts *et al.* (2014).

S'il est intéressant de représenter le mieux possible la relation de dispersion suite à la discrétisation, il ne faut pas perdre de vue que les plus petites ondes du spectre sont souvent filtrées dans un modèle complet, par exemple lors de l'utilisation d'un schéma de transport intrinsèquement diffusif, d'une diffusion numérique, d'un filtre spatial, etc. Ces différentes composantes sont discutées dans le paragraphe 1.4.4. Ainsi, le comportement indésirable inhérent à la grille A, mis en évidence ici, n'est peut être pas tant problématique, si un ordre de discrétisation suffisamment élevé est considéré (un ordre supérieur à 4 semble suffisant selon Bénard et Glinton (2019)).

Notons que l'approche en grille Z met en évidence le fait que certains comportements numériques sont dépendants des variables pronostiques. Les implications de cette propriété seront discutées dans le chapitre 2 pour le modèle AROME.

## Grille verticale

Sur la verticale, les grilles de Lorenz et Charney-Phillips sont principalement utilisées, dont une comparaison détaillée est, par exemple, donnée par Holdaway *et al.* (2013). Sur la grille de Lorenz, les calculs sont effectués sur des niveaux pleins pour toutes les variables pronostiques, à l'exception de la vitesse verticale du vent qui est calculée au niveau des interfaces. La grille de Lorenz est actuellement utilisée pour les modèles AROME, WRF, COSMO et ASUCA. Elle a été abandonnée récemment par le modèle GEM au profit d'une grille de Charney-Phillips, disposant de meilleures propriétés numériques selon Girard *et al.* (2014).

### 1.4.4 Filtrage spatial et diffusion numérique

Pour effectuer la transition entre les plus petits phénomènes résolus par la dynamique (de longueur d'onde  $2\Delta x$ ) et la partie paramétrisée, des filtres spatiaux ou une diffusion numérique sont implémentés dans les modèles. Cette étape intervient généralement à la fin du pas de temps, une fois que tous les autres calculs ont été effectués. Dans les grandes lignes, elle consiste à diffuser les différentes variables pronostiques selon une équation de la forme :

$$\frac{\partial \Psi}{\partial t} = D_\nu \nabla^r \Psi \quad (1.35)$$

où  $D_\nu$  est un coefficient de diffusion et  $r$  est l'ordre de la diffusion (souvent 2 ou 4). Toutefois, des subtilités peuvent être ajoutées en choisissant, par exemple, un coefficient de diffusion qui dépend de la longueur d'onde considérée, ou en traitant différemment les différentes variables pronostiques qui composent  $\Psi$ .

Comme pour les équations aux dérivées partielles précédemment introduites, cette équation d'évolution est également soumise à une contrainte CFL si elle est traitée par un schéma explicite, limitant ainsi les choix possibles des coefficients de diffusion pour un pas de temps donné. Pour s'affranchir de cette contrainte de stabilité, elle peut être traitée via un schéma implicite, trivialement résolu lorsqu'une discrétisation spectrale est utilisée, comme nous le verrons dans le paragraphe 1.4.10. En revanche, comme mentionné dans Gustafsson et McDonald (1996), en point de grille la résolution de cette équation peut être coûteuse notamment dans le cas où un coefficient d'amortissement variable  $D_\nu$  est considéré. Par conséquent, les filtres spatiaux semblent plus prometteurs et ont été implémentés avec succès dans un modèle shallow-water par Bénard et Glinton (2019).

Dans la pratique, la diffusion numérique ou les filtres spatiaux servent également à atténuer ou filtrer une partie des ondes les plus difficiles à traiter numériquement.

### 1.4.5 Discrétisation explicite ou implicite

Il existe deux grandes classes de schémas pour intégrer le système (1.1) : les schémas explicites et les schémas implicites. Nous présentons ici leurs caractéristiques principales sans

s'attacher à ce stade aux traitements spécifiques qui s'appliquent aux différents termes des équations et conduisant à définir des schémas combinant les deux.

### Niveaux temporels

Plusieurs schémas numériques temporels ont été étudiés dans la littérature. Aujourd'hui, les schémas privilégiés sont seulement ceux à deux niveaux temporels ( $t + \Delta t, t$ ). Les schémas à trois niveaux temporels ( $t + \Delta t, t, t - \Delta t$ ) sont désormais jugés moins efficaces car ils nécessitent une étape d'évaluation supplémentaire de l'état du système par rapport aux schémas à 2 niveaux temporels, et accessoirement d'un filtre d'Asselin pour filtrer les modes computationnels (ce point est brièvement discuté par Bénard *et al.* (2010)).

### Ordre de discrétisation temporelle

Pour ne pas trop dégrader la qualité de la procédure de numérisation, il y a consensus pour utiliser des schémas temporels précis au moins au second ordre, a minima pour les phénomènes que l'on souhaite correctement représenter.

### Schémas explicites

Le système (1.1) peut être discrétisé selon un schéma explicite à 2 niveaux temporels. Connaissant l'état du système à un instant antérieur, on peut simplement déterminer l'état du système à l'instant ultérieur suivant un schéma Euler explicite :

$$\frac{\Psi^+ - \Psi^0}{\Delta t} = \mathcal{N}(\Psi^0), \quad (1.36)$$

où  $\Psi^+$  désigne l'état du système à l'instant  $t + \Delta t$  et  $\Psi^0$  désigne l'état du système à l'instant  $t$ . Ce schéma fait intervenir l'état du système à seulement deux instants :  $t$  et  $t + \Delta t$ , on dit qu'il est à 2 niveaux temporels. On peut montrer que ce système est inconditionnellement instable et il n'est donc pas utilisé dans la pratique. On lui a longtemps préféré le schéma leap-frog, du second ordre, conditionnellement stable et à trois niveaux temporels (nécessitant une évaluation du système aux instants  $t - \Delta t, t, t + \Delta t$ ) :

$$\frac{\Psi^+ - \Psi^-}{2\Delta t} = \mathcal{N}(\Psi^0) \quad (1.37)$$

où  $\Psi^+, \Psi^0$  et  $\Psi^-$  désignent respectivement l'état du système aux instants  $t + \Delta t, t$  et  $t - \Delta t$ . Toutefois, comme mentionné précédemment, du fait de l'évaluation de l'état du système à ces trois instants, des modes computationnels apparaissent et ce schéma n'est donc plus beaucoup utilisé en PNT. On lui préfère les méthodes de Runge-Kutta explicites à multi-étapes et à 2 niveaux temporels comme, par exemple, la méthode de Runge-Kutta à deux étapes (appelée

## 1.4. NUMÉRISATION

---

aussi méthode de Crank-Nicolson explicite) :

$$\frac{\Psi^+ - \Psi^0}{\Delta t} = \mathcal{N}\left(\Psi^0 + \frac{\Delta t}{2}\mathcal{N}(\Psi^0)\right) \quad (1.38)$$

### Schémas implicites

Le système peut également être discrétisé avec un schéma implicite à 2 niveaux temporels, comme par exemple avec un schéma Euler implicite :

$$\frac{\Psi^+ - \Psi^0}{\Delta t} = \mathcal{N}(\Psi^+), \quad (1.39)$$

où  $\Psi^+$  et  $\Psi^0$  désignent respectivement l'état du système aux instants  $t + \Delta t$  et  $t$ . Ce système peut être réécrit de façon équivalente sous la forme :

$$\Psi^+ - \Delta t\mathcal{N}(\Psi^+) = \Psi^0. \quad (1.40)$$

Contrairement à son homologue explicite, ce schéma est beaucoup plus stable mais nécessite l'inversion d'un problème, qui peut être très coûteuse notamment si le problème est fortement non-linéaire. Comme mentionné ci-dessus, l'ordre 1 est considéré insuffisant en PNT et on lui préfère ainsi un schéma d'ordre 2, comme le schéma de Crank-Nicolson à 2 niveaux temporels :

$$\frac{\Psi^+ - \Psi^0}{\Delta t} = \frac{\mathcal{N}(\Psi^+) + \mathcal{N}(\Psi^0)}{2}, \quad (1.41)$$

Cette fois encore, l'inversion du problème peut être particulièrement coûteuse. On montrera dans le paragraphe 1.4.8, les techniques qui ont été employées pour rendre cette inversion viable en PNT.

### 1.4.6 Modèle eulérien ou semi-lagrangien

En mécanique des fluides, deux manières de décrire le comportement d'un fluide sont généralement utilisées : une description eulérienne et une description lagrangienne.

#### Modèle eulérien

La description eulérienne consiste à définir un volume de contrôle et à quantifier les flux qui le traversent (au niveau de ses bords) pour déterminer l'évolution de l'état du système. Dans le cas qui nous intéresse, le système (1.2) peut être réécrit de façon équivalente sous la forme :

$$\frac{\partial \Psi}{\partial t} = \mathcal{A}(\Psi) + \mathcal{M}(\Psi) + \mathcal{P}(\Psi) \quad (1.42)$$



où  $\mathcal{A}$  sont les termes d'advection,  $\mathcal{M}$  est la partie dynamique sans les termes d'advection :  $\mathcal{D} - \mathcal{A}$ , et  $\mathcal{P}$  désigne la contribution des paramétrisations physiques. Comme nous l'avons vu, la discrétisation temporelle doit respecter les contraintes de stabilité (1.14) et (1.16) sur les termes d'advection  $\mathcal{A}$ , si un schéma explicite est utilisé. Par la suite, nous appelons modèle eulérien, un modèle qui résout les équations aux dérivées partielles écrites sous la forme (1.42).

### Modèle semi-lagrangien

La description lagrangienne consiste à prévoir l'évolution de l'état d'une particule quand on la suit dans son mouvement. La description semi-lagrangienne consiste à n'adopter le point de vue lagrangien que pendant l'application des schémas de transport, et à revenir à chaque pas de temps à une description eulérienne. Dans le cas qui nous intéresse, le système (1.2) peut être réécrit de façon équivalente sous la forme :

$$\frac{d\Psi}{dt} = \mathcal{M}(\Psi) + \mathcal{P}(\Psi) \quad (1.43)$$

où  $\mathcal{M}$  est la partie dynamique sans les termes d'advection :  $\mathcal{D} - \mathcal{A}$ , et  $\frac{d}{dt} = \frac{\partial}{\partial t} - \mathcal{A}$  désigne la dérivée particulaire (ou dérivée lagrangienne). Cette formulation est bien adaptée pour utiliser un algorithme semi-lagrangien pour traiter les termes d'advection et s'affranchir ainsi des contraintes (1.14) et (1.16). Toutefois, la contrainte CFL sur la propagation verticale des ondes rapides (1.17) est plus pénalisante. Par conséquent, ce schéma est très souvent associé à une discrétisation temporelle semi-implicite comme nous le verrons dans le paragraphe 1.4.8.

Ce schéma est utilisé à chaque pas de temps pour évaluer l'advection d'une quantité en :

- retrouvant le point origine à partir de chacun des points de grille considérés par un calcul de rétro-trajectoire,
- puis en calculant la quantité évaluée à chacun des points origines par interpolation des champs aux points de grille entourant le point de départ au pas de temps précédent.

Une description détaillée ainsi qu'un historique de ces schémas sont donnés par Staniforth et Côté (1991). L'utilisation de ce schéma de transport permet notamment d'affaiblir la contrainte CFL portant sur les termes d'advection (1.14) (1.16), au profit d'une contrainte de Lipschitz moins pénalisante, imposant aux trajectoires des particules de ne pas se croiser (Yessad, 2019b). Comme nous le verrons par la suite, certains termes de la partie linéaire peuvent être éventuellement traités dans ce schéma de transport, en particulier lorsque des variables pronostiques plus sophistiquées que les variables usuelles sont utilisées.

### Utilisation

Ces deux types de modèles sont couramment utilisés en PNT, bien que les modèles semi-lagrangiens soient davantage utilisés pour les modèles globaux opérationnels actuels. Toutefois, les modèles les plus récents, ou en cours de développement, ont tendance à abandonner la

formulation semi-lagrangienne pour une formulation eulérienne, présentant l’avantage notable d’être plus conservative. Par exemple, le modèle FVM en cours de développement au CEPMMT privilégie la formulation eulérienne en remplacement de la formulation semi-lagrangienne du modèle actuellement opérationnel IFS (Kühnlein *et al.*, 2019). La même tendance est observée pour FV3 le nouveau noyau dynamique du modèle américain GFS, optant pour une approche en volumes finis (Putman et Lin, 2007). Cette tendance s’explique en partie par la volonté de former des systèmes complets allant de la PNT jusqu’aux simulations climatiques, et pour lesquelles les aspects de conservativité jouent un rôle important.

Les modèles à aire limitée, qui sont pour la plupart apparus après les modèles globaux, ont parfois privilégié l’approche eulérienne (WRF, COSMO, ASUCA), parfois l’approche semi-lagrangienne (AROME, UM, GEM-LAM).

### 1.4.7 Discrétisation temporelle des modèles Eulériens

#### Méthode split-explicit

Les modèles eulériens ont dans l’ensemble privilégié une approche split-explicit (ou time splitting) consistant à intégrer les équations avec deux pas de temps différents : un grand pas pour les processus lents (notés  $\mathcal{S}$ ), et un plus petit pas de temps pour les processus rapides (notés  $\mathcal{F}$ ). Afin d’exploiter numériquement la forte anisotropie du maillage ( $\Delta z \ll \Delta x$ ), les processus rapides sont séparés en une composante horizontale  $\mathcal{F}_x$  traitée explicitement, et une composante verticale  $\mathcal{F}_z$  traitée implicitement. En ce sens, ces modèles font partie de la classe des modèles HEVI (Horizontalement-Explicite/Verticalement-Implicite). Le système (1.2) peut alors être réécrit sous la forme :

$$\frac{\partial \Psi}{\partial t} = \mathcal{F}_x(\Psi) + \mathcal{F}_z(\Psi) + \mathcal{S}(\Psi) + \mathcal{P}(\Psi). \quad (1.44)$$

Ces idées ont été introduites par Klemp et Wilhelmson (1978) qui utilisaient plusieurs schémas :

- un schéma leap-frog à 3 niveaux temporels pour intégrer les processus lents  $\mathcal{S}$ ,
- un schéma forward-backward pour les processus rapides se propageant suivant l’horizontale  $\mathcal{F}_x$ ,
- et un schéma implicite trapézoïdal pour les processus rapides se propageant suivant la verticale, conduisant à inverser un simple problème implicite tridiagonal.

De nombreux LAM opérationnels utilisent cette approche (WRF, COSMO, ASUCA). Ces modèles sont basés sur un schéma de Runge-Kutta à 3 étapes et à 2 niveaux temporels pour intégrer la composante lente  $\mathcal{S}$ .

Pour résumer, l’approche Split-Explicit (SE) a permis de s’affranchir de la contrainte CFL la plus pénalisante (1.17), d’intégrer seulement les termes responsables de la propagation horizontale des ondes rapides suivant la contrainte (1.15), et d’intégrer avec un plus grand pas de temps respectant les contraintes (1.14) et (1.16) toutes les autres composantes.

## Méthode IMEX

Récemment en météorologie, les méthodes IMEX (Implicite-Explicite) ont commencé à être explorées dans les modèles HEVI. Elles se basent sur le constat que désormais les deux contraintes CFL les plus pénalisantes sont celles qui portent sur la propagation verticale des ondes et du vent (1.17) et (1.16), alors que les contraintes sur la propagation horizontale (1.14) et (1.15) le sont moins. Par conséquent, il est peut être judicieux de séparer le problème en une partie horizontale  $\mathcal{H}$  traitée explicitement, et une partie verticale  $\mathcal{V}$  traitée implicitement. Une fois incorporée dans une méthode Runge-Kutta à un seul pas de temps et plusieurs étapes, elle définit deux tableaux de Butcher : l'un associé au traitement implicite et l'autre au traitement explicite. Le système (1.2) se réécrit alors :

$$\frac{\partial \Psi}{\partial t} = \mathcal{H}(\Psi) + \mathcal{V}(\Psi) + \mathcal{P}(\Psi). \quad (1.45)$$

Cette approche permet donc de s'affranchir des deux contraintes CFL les plus pénalisantes (1.17) et (1.16), contrairement à la méthode SE qui ne permettait de ne s'affranchir que de la contrainte (1.17). Récemment, Colavolpe *et al.* (2017) proposent, en supplément, de traiter séparément les termes d'advection et d'ajustement, résultant en quatre tableaux de Butcher et augmentant ainsi la taille des pas de temps.

### 1.4.8 Discrétisation temporelle des modèles semi-lagrangiens

#### Modèle semi-implicite

Comme nous l'avons vu précédemment, le schéma semi-lagrangien permet de s'affranchir des contraintes (1.14) et (1.16) mais pas de la contrainte la plus pénalisante sur la propagation verticale des ondes rapides (1.17). Pour pouvoir la contourner, en complément de la contrainte sur la propagation horizontale des ondes rapides (1.15), le schéma semi-implicite traite implicitement les termes responsables de la propagation de ces ondes rapides selon les idées de Robert (1981) tandis que les autres sont traités explicitement ou dans le schéma de transport semi-lagrangien pour les termes d'advection. Ces modèles utilisent majoritairement un schéma ICI à  $n$  étapes et à deux pas de temps, dont nous donnons une description ci-après.

#### Schéma ICI

Le schéma de Crank-Nicolson (1.41) se réécrit pour la formulation (1.43) sous la forme :

$$\frac{\Psi^+ - \Psi^0}{\Delta t} = \frac{\mathcal{M}(\Psi^+) + \mathcal{M}(\Psi^0)}{2} + \mathcal{P}(\Psi^\sim), \quad (1.46)$$

où  $\Psi^0$  et  $\Psi^+$  sont respectivement le vecteur d'état à l'instant  $t$  et  $t + \Delta t$ , et  $\mathcal{P}(\Psi^\sim)$  est la contribution de la partie physique, qui peut être évaluée au cours de la procédure (par exemple le modèle IFS du CEPMMT) ou en fin de procédure (par exemple le modèle AROME de

Météo-France).

Les outils numériques actuels ne permettent pas de résoudre très efficacement un problème non-linéaire et ce schéma n'est donc pas utilisé tel quel en PNT en général. La plupart des méthodes numériques actuelles reposent sur une étape de linéarisation qui conduit à inverser un problème généralement creux, tandis que la partie non-linéaire est traitée itérativement.

Le schéma ICI (Iterative Centered Implicit) à 2 niveaux temporels a par conséquent été introduit pour approcher le schéma de Crank-Nicolson (1.46) et peut être écrit sous les trois formes équivalentes suivantes :

$$\frac{\Psi^{+(n)} - \Psi^0}{\Delta t} = \frac{\mathcal{M}(\Psi^{+(n-1)}) + \mathcal{M}(\Psi^0)}{2} + \frac{\mathcal{L}^*(\Psi^{+(n)}) - \mathcal{L}^*(\Psi^{+(n-1)})}{2} + \mathcal{P}(\Psi^\sim), \quad (1.47)$$

$$\frac{\Psi^{+(n)} - \Psi^0}{\Delta t} = \frac{\mathcal{R}(\Psi^{+(n-1)}) + \mathcal{R}(\Psi^0)}{2} + \frac{\mathcal{L}^*(\Psi^{+(n)}) + \mathcal{L}^*(\Psi^0)}{2} + \mathcal{P}(\Psi^\sim), \quad (1.48)$$

$$\left(1 - \frac{\Delta t}{2}\mathcal{L}^*\right)\Psi^{+(n)} = \left(1 + \frac{\Delta t}{2}\mathcal{L}^*\right)\Psi^0 + \frac{\Delta t}{2}\mathcal{R}(\Psi^{+(n-1)}) + \frac{\Delta t}{2}\mathcal{R}(\Psi^0) + \Delta t\mathcal{P}(\Psi^\sim). \quad (1.49)$$

où  $n$  est l'étape courante du schéma ICI,  $\Psi^0$  et  $\Psi^+$  sont respectivement le vecteur d'état à l'instant  $t$  et  $t + \Delta t$ ,  $\mathcal{L}^*$  est un opérateur linéaire soigneusement choisi pour faire en sorte que la méthode ICI converge effectivement vers le schéma de Crank-Nicolson, et  $\mathcal{R} = \mathcal{M} - \mathcal{L}^*$  est le résidu non linéaire.

Notons qu'il reste à choisir comment initialiser le schéma ICI. Deux schémas sont principalement utilisés :

— un schéma non-extrapolé :

$$\Psi^{+(0)} = \Psi^0, \quad (1.50)$$

— un schéma extrapolé :

$$\Psi^{+(0)} = 2\Psi^0 - \Psi^- \quad (1.51)$$

où  $\Psi^-$  désigne l'état du système à l'instant  $t - \Delta t$ .

Par la suite, le nombre total d'étapes de ce schéma sera désigné par  $N_{iter}$ .

## Convergence

On dit que la méthode ICI a convergé vers le schéma de Crank-Nicolson s'il existe un certain nombre d'étapes  $n'$  à partir duquel  $\Psi^{+(n'+1)} \simeq \Psi^{+(n')}$ . Dans ce cas, le second terme du membre de droite de (1.47) s'annule et le schéma ICI converge bien vers le schéma de Crank-Nicolson (1.46).

La convergence n'est pas toujours garantie et dépend fortement de l'opérateur linéaire  $\mathcal{L}^*$ , du pas de temps d'intégration, de la discrétisation choisie, etc. Par exemple, prenons le cas

très particulier où  $\mathcal{L}^*$  est nul et donc annule le deuxième terme du membre de droite de (1.47). Dans ce cas, la méthode ICI se réduit à une méthode explicite à  $n$  étapes et donc soumise à des contraintes de stabilité. Augmenter le nombre d'étapes du schéma ICI peut ne pas être suffisant pour converger vers le schéma de Crank-Nicolson, si l'une des conditions de stabilité n'est pas respectée.

Le choix d'un opérateur linéaire assurant avant tout la convergence du schéma ICI vers un schéma de Crank-Nicolson est donc essentiel. Néanmoins, pour respecter les contraintes opérationnelles (C1)–(C5), cet opérateur doit également être choisi pour que :

- la méthode ICI converge suffisamment vite pour qu'elle soit compétitive, c'est-à-dire avec  $n'$  petit ;
- le problème implicite à inverser (l'opérateur du membre de gauche (1.49)) soit inversible, et inversé avec un solveur performant le cas échéant.

L'enjeu est donc de créer un opérateur linéaire  $\mathcal{L}^*$  tel que les phénomènes qui se propagent rapidement soient traités dans le problème implicite pour s'affranchir des contraintes de stabilité les plus pénalisantes d'une part, tout en garantissant une inversion performante du problème implicite d'autre part.

### 1.4.9 Opérateur linéaire

L'opérateur linéaire  $\mathcal{L}^*$  est choisi de sorte que la procédure d'intégration totale du système soit la plus efficace possible, en garantissant les quatre conditions suivantes pour un large choix de pas de temps :

- la stabilité numérique, (L1)
- l'inversibilité du problème implicite, (L2)
- la résolution efficace du problème implicite, (L3)
- un faible nombre d'étapes du schéma ICI. (L4)

De très nombreuses stratégies ont été implémentées ces dernières décennies pour choisir au mieux cet opérateur linéaire vis-à-vis des quatre conditions (L1)–(L4) et seront discutées dans la section 1.4.9. Au delà d'être une étape essentielle pour assurer la viabilité de la méthode ICI, le choix de l'opérateur linéaire conditionne également d'autres choix importants, comme la discrétisation des opérateurs de dérivée spatiale, qui eux-mêmes orientent les choix des différents filtrages numériques ou de diffusion. Ces aspects seront présentés dans la section 1.4.10 qui porte sur la discrétisation spectrale. De plus, cet opérateur linéaire s'applique à des variables pronostiques dont nous avons déjà entrevu dans la section 1.4.3 qu'elles pouvaient modifier certains comportements numériques. Nous verrons qu'il est ainsi possible de choisir un jeu particulier de variable pronostique pour améliorer la stabilité numérique (L1) dans le chapitre 2.

On discute désormais du choix de l'opérateur linéaire  $\mathcal{L}^*$ . Une manière de procéder pour le

déterminer est de le définir à partir du tangent linéaire de l'opérateur non-linéaire  $\mathcal{M}$ , linéarisé autour d'un état de base qui reste à déterminer. Cette méthode permet ainsi de remplacer le difficile problème du choix de l'opérateur, par un problème *a priori* plus simple du choix de l'état de base.

### Opérateur à coefficients variables

Pour ne pas avoir à déterminer arbitrairement cet état de base, le Met-Office reprend l'état du pas de temps précédent comme état de base du pas de temps courant (Davies *et al.*, 2005). Dans ce cas, l'état de base est actualisé à chaque pas de temps. Toutefois, cette stratégie conduit à des problèmes d'inversibilité du problème implicite. Un critère de sortie du solveur itératif a ainsi été implémenté en cas de problème de convergence (Davies *et al.*, 2005). De plus, le problème implicite considéré n'est pas très creux car possède un halo de 45 points posant ainsi également des problèmes de scalabilité et de coût de calcul (Benacchio et Wood, 2016).

Depuis 2014 et le remplacement de l'ancien noyau dynamique « New Dynamics » par le nouveau « End GAME », certains termes, notamment orographiques, ont été retirés de l'opérateur linéaire et sont désormais traités explicitement (Walters *et al.*, 2017), rendant ainsi le problème implicite plus creux et plus simple à inverser.

### Opérateur à coefficients constants

D'autres, au contraire, considèrent des états de base très idéalisés (isotherme, sans vent, sans orographie) en général obtenus à partir d'analyses de stabilité historiquement introduites par Simmons *et al.* (1978) pour les schémas semi-implicite à 3 niveaux temporels. Des degrés de liberté supplémentaires peuvent même être ajoutés au système pour mieux contrôler sa stabilité (Bénard, 2004), générant un opérateur linéaire différent du tangent linéaire, via l'introduction d'une température de référence spécifique pour l'équation de la quantité de mouvement sur la verticale. Dans ce cas, le schéma est par convention dit à « coefficients constants », bien que certains coefficients dépendent de la coordonnée verticale. Plus marginalement, l'imposition d'un état de référence pour résoudre le problème implicite conduit à une dépendance de la prévision à cet état comme montré par Thurnburn *et al.* (2010) dans le cas d'un système shallow-water. Certaines propriétés des schémas de transport peuvent ainsi être dégradées.

#### 1.4.10 Discrétisation spectrale

Dans cette partie nous montrons qu'une discrétisation spectrale sur la direction horizontale (comme celle introduite dans le paragraphe 1.4.3) ne peut être utilisée efficacement que pour les schémas à coefficients constants.

### Domaine non borné

Pour plus de simplicité, on restreint l'étude au cas d'un opérateur  $\mathcal{H}$  ne dépendant que de la coordonnée horizontale et dont l'équation d'évolution est de la forme :

$$\frac{\partial \Psi}{\partial t} = \mathcal{H}(\Psi). \quad (1.52)$$

Considérons de plus, un domaine horizontal non borné, on peut alors écrire le problème (1.52) comme :

$$\frac{\partial \Psi}{\partial t} = \mathcal{H} * \Psi \quad (1.53)$$

où  $*$  désigne le produit de convolution. Décomposons l'opérateur  $\mathcal{H}$  comme :

$$\mathcal{H} = \mathcal{H}^* + G \cdot \mathcal{H}_o^*, \quad (1.54)$$

où  $\mathcal{H}^*$  et  $\mathcal{H}_o^*$  sont deux opérateurs indépendants de la coordonnée horizontale,  $G$  est une fonction qui dépend de la coordonnée horizontale (comme par exemple le terme orographique de l'équation (1.22)) et «  $\cdot$  » est le produit terme à terme usuel. Le problème (1.53) peut alors être réécrit sous la forme :

$$\frac{\partial \Psi}{\partial t} = \mathcal{H}^* * \Psi + (G \cdot \mathcal{H}_o^*) * \Psi. \quad (1.55)$$

La transformée de Fourier spatiale du problème (1.55) aboutit à :

$$\frac{\partial \hat{\Psi}}{\partial t} = \hat{\mathcal{H}} \cdot \hat{\Psi} + \hat{G} * \hat{\mathcal{H}}_o^* \cdot \hat{\Psi}. \quad (1.56)$$

Dans le cas particulier où l'opérateur  $\mathcal{H}$  est indépendant de la coordonnée horizontale ( $G = 0$ ), la dernière équation se résout trivialement dans l'espace spectral et aucun produit de convolution n'est nécessaire. En revanche, lorsque l'opérateur dépend de la coordonnée horizontale, il faut alors calculer un produit de convolution qui peut s'avérer coûteux (deuxième terme du membre de droite de l'équation (1.56)).

Ainsi la discrétisation spectrale paraît adaptée uniquement pour traiter des termes strictement linéaires émergeant des schémas à coefficients constants et ne sera utilisée que pour traiter la partie linéaire  $\mathcal{L}^*$  des schémas ICI (1.47)–(1.49). Les autres termes seront ainsi calculés dans l'espace traditionnel en point de grille. L'un des enjeux est donc de gérer correctement la correspondance entre la résolution dans l'espace spectral (définie par exemple par la troncature zonale  $K$ ) et la résolution en point de grille (définie par exemple par la taille d'une maille  $\Delta x$ ).

### Domaine périodique

Dans la pratique, le domaine horizontal est naturellement périodique sur la sphère. Les fonctions sont ainsi décomposées en harmoniques sphériques, via une transformée de Fourier suivant la direction zonale et une transformée de Legendre suivant la direction méridienne.

## 1.4. NUMÉRISATION

---

Cette décomposition permet notamment de calculer avec une très grande précision les opérateurs de dérivée horizontale, y compris au voisinage des pôles. Des algorithmes « rapides », permettant de passer d'une complexité de  $\mathcal{O}(N^2)$  à  $\mathcal{O}(N \log N)$ , ont été développés pour calculer efficacement les séries de Fourier. Le plus connu d'entre eux est l'algorithme de FFT (Fast Fourier Transform) proposé par Cooley et Tukey (1965). Ce n'est que très récemment que le CEPMMT a implémenté un algorithme rapide pour les transformées de Legendre utilisées dans le modèle IFS (Wedi *et al.*, 2013).

Pour les modèles à aire limitée, le domaine n'est pas naturellement périodique mais peut être bi-périodisé selon les idées de Machenhauer et Haugen (1988). Le champ  $\Psi$  peut alors être décomposé via une double série de Fourier de la forme :

$$\Psi = \sum_{n_k=-K}^K \sum_{n_m=-M}^M \widehat{\Psi}_{k,m} e^{ikx} e^{imy}, \quad (1.57)$$

où  $k$  est le nombre d'onde zonal définit comme

$$k = \frac{2\pi n_k}{L_x}, \quad (1.58)$$

$m$  est le nombre d'onde méridien définit comme

$$m = \frac{2\pi n_m}{L_y}, \quad (1.59)$$

et  $K$  et  $M$  sont les troncatures spectrales zonale et méridienne du modèle (dont le choix est notamment discuté dans le paragraphe suivant, pour éviter certains problèmes numériques),  $\widehat{\Psi}_{k,m}$  est le coefficient de Fourier associé aux nombres d'ondes  $k$  et  $m$ .

### Troncature

Dans le cas d'un modèle spectral, la correspondance entre la troncature spectrale  $K$  et la résolution du modèle en point de grille  $\Delta x$  n'est pas triviale. Cette correspondance est choisie prioritairement pour éviter les phénomènes d'aliasing (repliement de spectre). Pour plus de clarté, on restreint cet exposé à un domaine à une seule dimension.

Pour éviter l'aliasing dû aux termes d'advection lorsqu'un modèle eulérien est utilisé, une grille quadratique est généralement considérée. Dans ce cas, le nombre de points du domaine  $N_x$  est relié à la troncature  $K$  par la relation :

$$N_x = 3K + 1. \quad (1.60)$$

Lorsqu'un schéma semi-lagrangien est utilisé, il ne reste plus les termes d'advection non-linéaires dans le reste du modèle. Dans ce cas, l'aliasing est moins important, et une simple grille linéaire est souvent utilisée, dont la relation est de la forme :

$$N_x = 2K + 1. \quad (1.61)$$



Toutefois, bien que le schéma semi-lagrangien traite les non-linéarités les plus fortes (quadratiques), un aliasing peut survenir pour les autres termes non-linéaires. Pour les éviter, les ondes de plus petite longueur d'onde sont généralement filtrées. Par exemple, les modèles AROME et ARPEGE filtrent les ondes dont la longueur d'onde est inférieure à  $3\Delta x$  pour utiliser cette grille linéaire. Ainsi, la partie du spectre comprise entre  $2\Delta x$  (qui est la plus petite onde représentable sur un modèle en grille A) et  $3\Delta x$  a d'abord été calculée puis amortie par le filtre. Pour ne pas avoir à calculer inutilement une partie du spectre pour ensuite la filtrer, les ondes comprises entre  $4\Delta x$  et  $2\Delta x$  sont posées nulles (et ne sont donc pas calculées) dans le modèle IFS du CEPMMT par l'utilisation d'une grille cubique (Wedi, 2014) :

$$N_x = 4K + 1 \quad (1.62)$$

Cette grille permet ainsi une économie de coût de calcul et ne nécessite quasiment pas de filtrage ou de diffusion supplémentaire. Toutefois le signal (météorologique) compris entre  $4\Delta x$  et  $3\Delta x$  n'est donc pas calculé, diminuant ainsi la résolution effective du modèle par rapport à un modèle à troncature linéaire qui filtre les ondes inférieures à  $3\Delta x$ .

### Diffusion implicite

Une fois discrétisée temporellement, l'équation (1.35) est trivialement résolue dans l'espace spectral, par simple division de coefficients qui dépendent du pas de temps et du coefficient de diffusion.

#### 1.4.11 Pas de temps

Faire une prévision à échéance donnée nécessite plusieurs évaluations intermédiaires dont le nombre est déterminé par la grandeur du pas de temps : plus le pas de temps est petit, plus ce nombre est grand. A contrario, un grand pas de temps diminue le nombre d'évaluations intermédiaires pour fournir une prévision à une échéance donnée.

En cherchant à contourner les contraintes numériques les plus pénalisantes, les centres météorologiques ont systématiquement opté pour des schémas numériques leur permettant d'augmenter le pas de temps d'intégration le plus possible, diminuant ainsi le nombre d'évaluations intermédiaires. Cela semble être encore le cas aujourd'hui avec les méthodes IMEX, nouvellement explorées en PNT, qui cherchent elles aussi à agrandir les pas de temps.

Ainsi selon ce comportement, il semblerait que pour la plupart des modèles le coût d'une évaluation en utilisant un grand pas de temps soit inférieur au coût de  $n$  évaluations utilisant un pas de temps  $n$  fois plus petit. Par exemple, si on estime que le coût de calcul d'un modèle spectral est principalement régi par le coût des transformées spectrales, intégrer les équations avec un pas de temps deux fois plus petit nécessite deux fois plus d'étapes de transformées spectrales, conduisant à doubler le coût de calcul, ce qui est difficilement viable en opérationnel.

## 1.4. NUMÉRISATION

---

Toutefois, l'agrandissement des pas de temps appauvrit la qualité de la solution. Par exemple, Pinty *et al.* (1995) montrent que les écoulements de petites échelles pilotés par les processus d'advection ne sont correctement résolus que pour des nombres CFL du vent ((1.14) et (1.16)) proches de l'unité. Ainsi les modèles HEVI actuels, présentant ce genre de pas de temps satisfont cet objectif en général, contrairement aux modèles semi-implicites semi-lagrangiens actuels pour lesquels les nombres CFL sur le vent sont bien supérieurs à l'unité.

Pour évaluer l'impact des grands pas de temps sur la propagation d'une onde pour un modèle semi-implicite, on procède de façon analogue à ce qui a été fait dans le paragraphe 1.4.3, en décomposant en série de Fourier la partie temporelle du problème pour déterminer la répartition spectrale des erreurs causées par la discrétisation. Pour un schéma semi-implicite à deux niveaux temporels, la pulsation discrète  $\omega'$  du système suite à la discrétisation temporelle devient :

$$\omega' = \frac{2}{\Delta t} \arctan\left(\frac{\Delta t \omega}{2i}\right) \quad (1.63)$$

où  $\omega$  désigne la pulsation du système continu. Ainsi, pour un système shallow-water admettant la propagation d'un mode rapide de vitesse de phase  $c$ , la relation de dispersion (1.12) suite à la discrétisation temporelle devient :

$$\omega' = \frac{2}{\Delta t} \arctan\left(\frac{\Delta t}{2}ck\right). \quad (1.64)$$

On note que lorsque le pas de temps tend vers 0, la relation de dispersion du système discrétisé temporellement tend bien vers celle du système continu (1.12). On note également que pour un pas de temps donné, les erreurs sur la pulsation sont d'autant plus faibles que les longueurs d'onde sont grandes (quand  $k \rightarrow 0$ ,  $\arctan(\Delta tck/2) \rightarrow \Delta tck/2$ ). Ainsi l'utilisation d'un grand pas de temps dégrade plus particulièrement le signal de petite échelle. En calculant la vitesse de phase correspondant à un schéma semi-implicite :

$$c_\phi = \frac{2}{k\Delta t} \arctan\left(\frac{\Delta t}{2}ck\right), \quad (1.65)$$

on retrouve le résultat bien connu que ce schéma ralentit d'autant plus les ondes rapides que le pas de temps est grand (quand  $\Delta t \rightarrow \infty$ ,  $c_\phi \rightarrow 0$ ). Dans le chapitre 2, nous montrerons via quelques simulations numériques que l'augmentation du pas de temps dégrade la représentation de certaines ondes orographiques.

Il semblerait donc qu'un choix doive être opéré entre efficacité (C1) et qualité (C3). En développant des approches agrandissant le plus possible les pas de temps, les centres météorologiques ont considéré jusqu'ici que la dégradation de la qualité suite à l'usage de grand pas de temps était largement compensée par le gain en efficacité. Néanmoins, les nouveaux noyaux dynamiques actuellement développés optent plutôt pour des intégrations temporelles avec de plus petits pas de temps. Par exemple au CEPMMT, le modèle FVM utilise un schéma d'advection conservatif MPDATA (soumis aux contraintes CFL sur le vent), appelé avec un plus petit pas de temps que le schéma semi-lagrangien utilisé dans le modèle IFS (Kühnlein *et al.*, 2019).

### 1.4.12 Exemple de modèles utilisés

Modèle	Dyn.	$\Delta x$	$\Delta t$	$c_x^*$	Gr.x	Gr.z	Co.z	Schéma	Discr.
AROME	NH	1,3 km	50 s	13.5	A	Lo.	$\pi$	SI*(2)-SL	SP
UK-UM	NH	1,5 km	60 s	14	C	Ch.	$z$	SI(2)-SL	DF
GEM-LAM	H	10 km	300 s	10.5	C	Ch.	$\pi$	SI*(2)-SL	SP/DF
WRF-HRRR	NH	3 km	20 s	2.3	C	Lo.	$\pi$	SE-RK3	DF
COSMO (DWD)	NH	2,2 km	20 s	3.2	C	Lo.	$z$	SE-RK3	DF
Meteo-Swiss		1,1 km	10 s	3.2					
ASUCA (JMA)	NH	2 km	16,7 s	2.9	C	Lo.	$z$	SE-RK3	VF

TABLE 1.1 – Synthèse des différents modèles à aire limitée utilisés en opérationnel en 2020. Les colonnes 2 à 10 spécifient respectivement : la dynamique (Dyn.) qui peut être Hydrostatique (H) ou non-hydrostatique (NH) ; la résolution horizontale  $\Delta x$  ; le (grand) pas de temps d'intégration  $\Delta t$  ; le nombre CFL sur les ondes horizontales  $c_x^*$  correspondant ; la grille horizontale (Gr.x) qui est soit co-localisée (A) soit décalée (C) ; la grille verticale (Gr.z) qui est soit une grille de Lorenz (Lo.) ou de Charney-Philipps (Ch.) ; la coordonnée verticale associée (Co.z) qui est soit une coordonnée hauteur ( $z$ ) ou une coordonnée masse ( $\pi$ ) ; le schéma numérique utilisé qui peut être semi-implicite semi-lagrangien (SI-SL) éventuellement à coefficients constants (SI\*-SL) ou Split-Explicit avec un schéma de Runge Kutta à 3 étapes (SE-RK3) ; discrétisation horizontale utilisée qui peut être spectrale (SP), en différences finies (DF) ou en volumes finis (VF).

Les principales caractéristiques de certains LAM sont listés dans la Table 1.1 avec les données publiées ou rendues publiques avant fin 2020.

Tous les modèles à aire limitée sont NH, à l'exception du modèle GEM, dont la résolution horizontale ne rend pas nécessaire pour le moment l'abandon de l'approximation hydrostatique. Cela s'explique certainement par le fait que le domaine horizontal à couvrir est significativement plus grand que ceux des différents pays européens. On note d'ailleurs que les résolutions horizontales les plus fines concernent très majoritairement des modèles dont la taille du domaine est assez restreinte (Suisse, Royaume-Uni, France).

Les modèles semi-implicites utilisent systématiquement un schéma ICI à deux itérations mais les opérateurs linéaires diffèrent : alors qu'il est à coefficients constants pour les modèles AROME et GEM, il est à coefficients variables pour le modèle UK-UM. Quoiqu'il en soit, ces modèles utilisent des grands pas de temps tels que les nombres CFL horizontaux sur les ondes les plus rapides sont supérieurs à 10. Les autres modèles utilisent un schéma SE dont l'intégration temporelle des processus lents est pilotée une méthode de Runge-Kutta à 3 étapes. A résolution horizontale équivalente, ces modèles utilisent donc des pas de temps environ 5 fois plus petits que les modèles SI.

Par ailleurs, le modèle AROME est le seul à utiliser une grille A, tout-à-fait adaptée pour une discrétisation spectrale. Les autres modèles utilisent une grille décalée C avec une discrétisation en différences finies. Enfin, le modèle japonais ASUCA a récemment abandonné

la discrétisation en différences finies pour une discrétisation en volumes finis.

### 1.4.13 Enjeux des prochaines années

#### Scalabilité

Comme nous l'avons évoqué dans la section 1.2, les architectures de calcul évoluent rapidement et se complexifient par l'utilisation de technologies non utilisées jusqu'ici pour le calcul scientifique, comme par exemple les cartes graphiques (GPU). L'un des enjeux dans les prochaines années en PNT, est notamment de faire fonctionner le code sur des architectures de plus en plus massivement parallèles et hybrides.

Or, comme nous l'approfondirons dans le chapitre 2, l'algorithme de Transformée de Fourier utilisé dans certains modèles fait intervenir de nombreuses communications globales lors de la transposition du vecteur d'état dans les directions zonale et méridienne. Cela nécessite de définir de nouveaux pavages spécifiques pour le protocole de parallélisation à chacune de ces étapes (Figure 1.1b) et (Figure 1.1c). Ces transpositions posent un problème de scalabilité sur les machines massivement parallèles utilisées.

Les nouveaux noyaux dynamiques développés pour les modèles globaux ont tendance à abandonner les transformées spectrales principalement utilisées dans les modèles semi-implicites semi-lagrangiens, au profit de méthodes plus conservatives et plus locales en Volumes Finis (Kühnlein *et al.*, 2019). Notons que le plupart des modèles HEVI, utilisent des schémas locaux suivant la direction horizontale, en accord avec le paradigme de parallélisation actuel. Ces modèles semblent donc particulièrement bien adaptés à l'évolution des architectures de calcul.

#### Fortes pentes

Comme nous l'avons vu dans le paragraphe 1.4.12, la plupart des modèles à aire limitée atteignent des résolutions horizontale de l'ordre de 1 km jusqu'à 10 km, en fonction de la taille du domaine considéré. Les résolutions les plus fines sont ainsi atteintes pour les modèles dont l'aire du domaine à couvrir est la plus faible, comme par exemple pour le modèle COSMO utilisé sur la Suisse, puis le modèle anglais UK-UM, puis le modèle français AROME.

L'un des axes d'amélioration des modèles de PNT pour le futur est de poursuivre l'affinement de la résolution. Pour le moment, aucun modèle opérationnel n'atteint des résolutions hectométriques sur la direction horizontale mais des premières expériences ont, par exemple, été menées à des résolutions de l'ordre de 500 m pour le modèle AROME. L'affinement de la résolution horizontale permet de mieux représenter l'orographie, y compris les plus fortes pentes des reliefs, dont le traitement pose de nombreux problèmes numériques. Par exemple, le solveur de pression du modèle anélastique Méso-NH ne converge plus pour des pentes trop fortes (Lac *et al.*, 2018), les opérateurs de discrétisation du gradient de pression ont été

modifiés dans le modèle ICON (Zängl, 2012), et finalement les pentes supérieures à 100% sont filtrées dans le modèle GEM (Husain *et al.*, 2019).

Notons que les problèmes numériques au voisinage des plus fortes pentes se manifestent plus particulièrement pour les modèles à aire limitée dont le domaine est petit, la résolution est fine et le relief est pentu. Ainsi, le modèle AROME, est l'un des premiers modèles à être confronté à ce problème. Cela fera l'objet d'un important développement dans le chapitre 2.

## 1.5 Conclusion

Dans ce chapitre, nous avons présenté le problème de la Prévision Numérique du Temps comme la résolution d'un système dynamique respectant cinq composantes clés (C1)–(C5). Nous avons montré que le problème de PNT résolvait dans la partie dynamique les phénomènes dont l'échelle caractéristique est supérieure à la taille de la maille, et en paramétrisant ceux dont la taille est inférieure. La résolution horizontale est désormais de l'ordre de quelques kilomètres pour la plupart des modèles à aire limitée. A cette résolution, la partie dynamique résout principalement les équations d'Euler non-hydrostatiques.

Pour pouvoir intégrer les équations de sorte à respecter les délais contraints inhérents à la PNT (C1), plusieurs méthodes numériques peuvent être utilisées. Nous avons illustré différents choix possibles via l'exemple de six modèles à aire limitée utilisés dans de nombreux pays. Les modèles opérationnels actuels sont ainsi séparés en deux classes : les modèles HEVI (Horizontalement Explicite Verticalement Implicite) qui utilisent très majoritairement des méthodes Split-Explicit et les modèles semi-implicites semi-lagrangiens qui utilisent majoritairement un schéma ICI (Iterative-Centered-Implicit scheme) pour approcher le schéma de Crank-Nicolson. Les modèles HEVI semblent particulièrement bien adaptés au paradigme de parallélisation actuel, mais sont soumis à des contraintes de stabilité sur le pas de temps d'intégration, l'obligeant ainsi à être significativement plus faible que le pas de temps caractéristique d'un modèle SI, et multipliant par conséquent le nombre d'étapes intermédiaires pour fournir une prévision à une échéance donnée. Sur un exemple simple nous avons néanmoins observé que l'agrandissement du pas de temps d'intégration dégradait une partie de la propagation de certaines ondes, notamment à petite échelle.

Une attention plus particulière a été notamment portée sur les modèles SI, en particulier dans le choix de l'opérateur linéaire. Nous avons montré que celui-ci devait être choisi pour garantir : la stabilité numérique (L1), l'inversibilité du problème implicite qui en résulte (L2), sa résolution efficace (L3) et la convergence rapide du schéma ICI vers le schéma de Crank-Nicolson (L4). Le choix de cet opérateur linéaire pilote d'autres choix importants comme celui de la discrétisation horizontale, de la diffusion ou des filtres nécessaires pour atténuer certains comportements numériques indésirables et assurer la cascade d'énergie.

Lorsque l'opérateur linéaire est à coefficients constants, les parties horizontale et verticale sont séparables et une discrétisation spectrale peut ainsi être utilisée sur l'horizontale. Toutefois, du fait des nombreuses communications qui en résultent, l'algorithme de transformée

## 1.5. CONCLUSION

---

de Fourier rapide ainsi utilisé n'est pas adapté aux machines massivement parallèles, posant un problème de scalabilité qui sera développé dans le chapitre suivant. Les modèles HEVI, en faisant intervenir des opérateurs essentiellement locaux sur la partie horizontale apparaissent, eux, plus scalables.

Sur l'horizontale, hormis pour les modèles spectraux, la grille décalée (C) est privilégiée pour la plupart des modèles et présente des propriétés de discrétisation spatiale *a priori* plus avantageuses que celles de la grille co-localisée (A), notamment aux petites échelles. Toutefois ce problème peut être, au moins partiellement, contourné en augmentant l'ordre de la discrétisation et doit de toute façon être replacé dans le contexte plus général où les plus petites longueurs d'onde sont filtrées par les mécanismes de diffusion, le filtrage, etc. Sur la direction verticale, la plupart des modèles utilisent une coordonnée qui suit le terrain de type hauteur ou masse, sur une grille de Lorenz ou de Charney-Phillips.

A mesure que la résolution s'affine, la plupart des modèles, indépendamment du fait qu'ils soient SI ou HEVI, sont confrontés au problème des fortes pentes apparaissant lorsque l'orographie  $y$  est finement représentée. Ce problème devrait concerner, dans un premier temps, les modèles à aire limitée dont le domaine horizontal est suffisamment petit pour qu'une résolution très fine soit utilisée en présence d'un relief chahuté, ce qui est le cas du modèle AROME aux résolutions hectométriques. Ce point sera également développé dans le prochain chapitre.

Dans le prochain chapitre, nous détaillerons le noyau dynamique du modèle AROME en précisant notamment le système d'équation et les variables pronostiques utilisés, l'état de base autour duquel est construit l'opérateur linéaire, et en précisant certaines autres caractéristiques évoquées dans ce chapitre. Nous évoquerons ensuite les limites que pose le noyau dynamique actuel pour s'adapter aux futures architectures de calcul, aborder le problème des fortes pentes aux résolutions hectométriques, et simuler à un coût raisonnable certaines ondes orographiques. Nous montrerons ainsi quelles alternatives peuvent être considérées pour contourner ces limites, au regard des pistes déjà explorées par les modèles à aire limitée étudiés dans ce chapitre.



# Noyau dynamique du modèle AROME

---

## Sommaire

---

<b>2.1</b>	<b>Partie numérique du modèle AROME . . . . .</b>	<b>46</b>
2.1.1	Présentation générale . . . . .	46
2.1.2	Coordonnées et grilles . . . . .	47
2.1.3	Équations du modèle . . . . .	47
2.1.4	Tangent linéaire . . . . .	49
2.1.5	Opérateur linéaire considéré . . . . .	51
2.1.6	Variables pronostiques . . . . .	53
2.1.7	Discretisation spatiale . . . . .	55
2.1.8	Discretisation temporelle ICI . . . . .	57
2.1.9	Inversion du problème implicite . . . . .	57
2.1.10	Schéma semi-Lagrangien . . . . .	59
2.1.11	Fonctionnement général du code actuel . . . . .	60
2.1.12	Pas de temps . . . . .	60
2.1.13	Diffusion et filtrage . . . . .	61
<b>2.2</b>	<b>Limites de la version actuelle . . . . .</b>	<b>63</b>
2.2.1	Scalabilité . . . . .	63
2.2.2	Fortes pentes . . . . .	64
2.2.3	Origine des problèmes . . . . .	64
<b>2.3</b>	<b>Solutions étudiées dans la thèse . . . . .</b>	<b>65</b>
2.3.1	Scalabilité . . . . .	65
2.3.2	Fortes pentes . . . . .	65

---

Dans ce chapitre, le noyau dynamique actuel du modèle AROME est présenté en détail. Nous exposons ensuite ses faiblesses pour appréhender les résolutions horizontales hectométriques du futur, et les architectures de calcul très parallèles.



## 2.1 Partie numérique du modèle AROME

### 2.1.1 Présentation générale

Le modèle à aire limitée opérationnel de Météo-France, AROME (Application of Research to Operations at MESoscale) a été mis en service fin 2008 (l'article fondateur est Seity *et al.* (2011)). Il a été développé en héritant d'une partie des composantes de son prédécesseur ALADIN (Aire Limitée Adaption dynamique Développement International). ALADIN est également le nom donné au consortium qui participe à son développement, et qui a été rejoint récemment par le consortium HIRLAM. Enfin, de nombreuses parties du code sont communes avec celles du modèle IFS du CEPMMT (Centre Européen pour les Prévisions Météorologiques à Moyen Terme).

Plusieurs domaines géographiques peuvent être considérés mais celui qui nous intéresse dans cette étude est celui centré sur la France métropolitaine (Fig 2.1). En 2020, AROME satisfait la condition (C1) en fournissant  $t_f \simeq 48$  heures de prévision en environ 70 min de simulation. Il dispose d'un algorithme d'assimilation de données 3D-Var (Gustafsson *et al.*, 2018) et sa prévision d'ensemble (Bouttier *et al.*, 2012) est composée de 16 membres. Il est forcé par le modèle global ARPEGE (Pailleux *et al.*, 2015) sur les bords latéraux et par le modèle SURFEX sur le bord inférieur (Masson *et al.*, 2013).

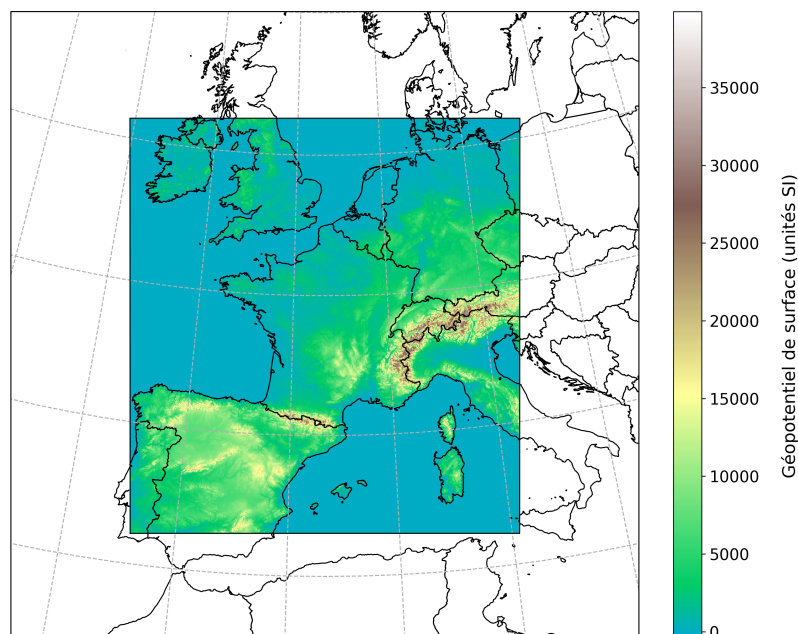


FIGURE 2.1 – Domaine du modèle opérationnel AROME. Source : M. Destouches.

### 2.1.2 Coordonnées et grilles

#### Sur la verticale

Comme mentionné dans le paragraphe précédent, le modèle AROME hérite de certaines parties du code du modèle ALADIN. Pour ne pas avoir à recoder ces routines dans un nouveau système de coordonnées, il a été choisi d'étendre la coordonnée masse utilisée dans le modèle hydrostatique ALADIN au modèle non-hydrostatique AROME, selon les idées de Laprise (1992). Une coordonnée hybride  $\eta$ , qui suit le terrain en bas de l'atmosphère et dont les effets s'annulent avec l'altitude, est ainsi définie comme :

$$\pi(x,y,\eta,t) = A(\eta) + B(\eta)\pi_s(x,y,t) \quad (2.1)$$

où  $\pi$  et  $\pi_s$  désignent la pression hydrostatique et la pression hydrostatique de surface,  $A$  et  $B$  sont les deux fonctions qui permettent de passer d'une coordonnée qui épouse la forme du terrain à une coordonnée pression et telles que :

- $A(0) = B(0) = 0$  au sommet,
- $A(1) = 0, B(1) = 1$  à la surface.

Dans le cas particulier où  $A(\eta) = 0$ , la coordonnée  $\eta$  est réduite à une coordonnée qui suit le terrain  $\sigma$ .

Sur la verticale, une grille de Lorenz est utilisée où toutes les variables pronostiques sont calculées sur des niveaux pleins, sauf la vitesse verticale  $w$  calculée aux interfaces.

#### Sur l'horizontale

Sur l'horizontale AROME utilise un système de coordonnées cartésiennes ( $x$  et  $y$ ) sur un maillage carré en utilisant une grille A co-localisée.

Notons que dans toute la suite, le gradient horizontal  $\nabla = [\partial_x, \partial_y]^T$  d'un champ est calculé en appliquant les règles de transformation de coordonnées définies par Kasahara (1974) :

$$(\nabla\Psi)_z = (\nabla\Psi)_\eta + G \frac{\partial\Psi}{\partial\eta} \quad (2.2)$$

Le terme  $G = [\partial_x\eta, \partial_y\eta]^T$  est un terme orographique dont le traitement sera discuté dans la suite.

### 2.1.3 Équations du modèle

L'opérateur non-linéaire  $\mathcal{N}$  du système (1.1) est décomposé en une partie dynamique  $\mathcal{D}$  et une partie de paramétrisation physique  $\mathcal{P}$  comme dans le système (1.2). On distingue les variables pronostiques associées à la partie physique (l'énergie cinétique turbulente, les

rappports de mélange : de la vapeur d'eau, de la neige, du graupel, de l'eau nuageuse et des cristaux de glace) et celles associées à la partie dynamique (le vecteur du vent horizontal  $\mathbf{U}$  qui comprend les composantes zonale  $U$  et méridienne  $V$ , la divergence verticale du vent  $\mathbb{D}$ , la température  $T$ , la pression non-hydrostatique normalisée  $\hat{q}$  et la pression hydrostatique de surface  $\pi_s$ ). Dans toute la suite, seule la partie dynamique sera décrite en détail en s'appuyant notamment sur la documentation de Bénard et Mašek (2013), les paramétrisations physiques étant considérées par l'ajout des termes sources ( $\boldsymbol{\nu}$ ,  $\mathcal{W}$ ,  $\mathcal{Q}$ ) dans les membres de droite des équations de la dynamique (2.6)–(2.10).

Le vecteur d'état est donc constitué des variables pronostiques :

$$\Psi = [\mathbf{U}, \mathbb{D}, T, \hat{q}, \pi_s]^T, \quad (2.3)$$

où la pression non-hydrostatique  $\hat{q}$  est définie à partir de la vraie pression  $p$  comme :

$$\hat{q} = \ln(p/\pi), \quad (2.4)$$

et la divergence verticale du vent  $\mathbb{D}$  comme :

$$\mathbb{D} = \frac{p}{mR_dT} \left( -g \frac{\partial w}{\partial \eta} + \nabla \phi \cdot \frac{\partial \mathbf{U}}{\partial \eta} \right). \quad (2.5)$$

où  $m$  désigne la métrique verticale (2.11),  $\phi$  le géopotentiel (2.13),  $g$  la constante de gravité, et  $R_d$  la constante des gaz parfaits de l'air sec. La divergence de la vitesse verticale du vent  $\mathbb{D}$  est utilisée à la place de la vitesse verticale du vent  $w$ . Cela permet d'utiliser une grille verticale où toutes les variables pronostiques sont calculées sur les niveaux pleins. De plus, comme nous l'évoquerons dans le paragraphe 2.1.6, cette variable pronostique possède de meilleures propriétés de stabilité numérique que la vitesse verticale du vent.

Les équations du modèle AROME sont écrites telles que la partie dynamique est présentée dans les membres de gauche et la partie de paramétrisation physique dans les membres de droite :

$$\frac{d\mathbf{U}}{dt} + \frac{RT}{p} \nabla p + \frac{1}{m} \frac{\partial p}{\partial \eta} \nabla \phi = \boldsymbol{\nu}, \quad (2.6)$$

$$\begin{aligned} \frac{d\mathbb{D}}{dt} + g^2 \frac{p}{mR_dT} \frac{\partial}{\partial \eta} \left( \frac{1}{m} \frac{\partial(p-\pi)}{\partial \eta} \right) - g \frac{p}{mR_dT} \frac{\partial \mathbf{U}}{\partial \eta} \cdot \nabla w \\ - (\mathbf{X} - \mathbb{D}) \left[ \frac{R_d}{R} \mathbb{D} + \left( 1 - \frac{R_d}{R} \mathbf{X} - \dot{\mathbf{X}} \right) \right] = -g \frac{p}{mR_dT} \frac{\partial \mathcal{W}}{\partial \eta}, \end{aligned} \quad (2.7)$$

$$\frac{dT}{dt} + \frac{RT}{C_v} (\nabla \cdot \mathbf{U} + \mathbb{D}) = \frac{\mathcal{Q}}{C_v}, \quad (2.8)$$

$$\frac{d\hat{q}}{dt} + \frac{C_p}{C_v} (\nabla \cdot \mathbf{U} + \mathbb{D}) + \frac{\dot{\pi}}{\pi} = \frac{\mathcal{Q}}{C_v T}, \quad (2.9)$$

$$\frac{\partial \pi_s}{\partial t} + \int_0^1 \nabla \cdot m \mathbf{U} d\eta = 0. \quad (2.10)$$

$C_p$  est la capacité calorifique de l'air à pression constante,  $C_v$  est la capacité calorifique de l'air à volume constant et  $R$  est la constante des gaz parfaits de l'air.

## 2.1. PARTIE NUMÉRIQUE DU MODÈLE AROME

---

Les relations suivantes sont des relations diagnostiques et portent sur le facteur métrique  $m$ , la vraie pression  $p$ , le géopotentiel  $\phi$  (le géopotentiel de surface est  $\phi_s$ ), la divergence 3D du vent  $D_3$ , la pseudo-vitesse verticale (en coordonnée  $\eta$ )  $m\dot{\eta}$ , la dérivée particulaire de la pression hydrostatique  $\dot{\pi}$ , le gradient horizontal de la vitesse verticale du vent, et le terme croisé  $X$  :

$$m = \frac{\partial \pi}{\partial \eta}, \quad (2.11)$$

$$p = \pi \exp(\hat{q}), \quad (2.12)$$

$$\phi = \phi_s + \int_{\eta}^1 \frac{mRT}{p} d\eta, \quad (2.13)$$

$$D_3 = \nabla \cdot \mathbf{U} + \frac{R_d}{R} \mathbb{D} + \left(1 - \frac{R_d}{R}\right) X, \quad (2.14)$$

$$m\dot{\eta} = B \int_0^1 \nabla \cdot m \mathbf{U} d\eta - \int_0^{\eta} \nabla \cdot m \mathbf{U} d\eta', \quad (2.15)$$

$$\dot{\pi} = \mathbf{U} \cdot \nabla \pi - \int_0^{\eta} \nabla \cdot m \mathbf{U} d\eta', \quad (2.16)$$

$$g\nabla w = g\nabla w_s + \int_{\eta}^1 \frac{mRT}{p} \nabla (\mathbb{D} - X) d\eta' + \int_{\eta}^1 (\mathbb{D} - X) \nabla \left(\frac{mT}{p}\right) d\eta', \quad (2.17)$$

$$X = \frac{p}{mRT} \nabla \phi \cdot \frac{\partial \mathbf{U}}{\partial \eta}. \quad (2.18)$$

Les dérivées particulières  $\frac{d}{dt}$  (ou dérivées lagrangiennes) sont traitées dans le schéma de transport semi-Lagrangien. Notons que le terme  $\dot{X} = \frac{dX}{dt}$  peut être soit traité dans le schéma semi-implicite, soit dans le schéma semi-Lagrangien, comme déjà mentionné par Bénard *et al.* (2005). L'option actuellement retenue est de le traiter dans le schéma semi-Lagrangien. Ce terme ne sera donc pas inclus dans l'opérateur linéaire par la suite.

### 2.1.4 Tangent linéaire

Comme mentionné dans le chapitre 1, l'intérêt de la méthode ICI repose notamment sur le choix de l'opérateur linéaire  $\mathcal{L}^*$  pour respecter les conditions (L1)–(L4). Une manière simple de le choisir est de calculer tout d'abord le tangent linéaire de l'opérateur non-linéaire  $\mathcal{M}$ , linéarisé autour d'un état de base  $\Psi^*$  à déterminer. Les analyses de stabilité menées par Bénard (2003) ont conduit à définir l'état de base  $\Psi^*$  comme étant : isotherme de température  $T^*$ , sans vent (au repos), sec (les phénomènes humides ne sont pas considérés), hydrostatique, homogène horizontalement, sur un terrain plat ( $G = 0$ ). La perturbation  $\Psi'$  par rapport à l'état de base  $\Psi^*$  peut être décomposée comme :

$$\Psi' = \Psi - \Psi^* \quad (2.19)$$

L'évolution des perturbations  $\Psi'$  par rapport à cet état de base est de la forme :

$$\frac{\partial \Psi'}{\partial t} = \mathcal{L}^* \Psi' \quad (2.20)$$

où  $\Psi'$  est le vecteur d'état contenant les variables pronostiques perturbées :

$$\Psi' = [\zeta', D', \mathbb{D}', T', \hat{q}', \pi'_s]^T \quad (2.21)$$

où  $\zeta'$  désigne la perturbation de vorticité horizontale et  $D'$  la perturbation de la divergence horizontale du vent définies comme :

$$D' = \frac{\partial U'}{\partial x} + \frac{\partial V'}{\partial y} \quad (2.22)$$

$$\zeta' = \frac{\partial V'}{\partial x} - \frac{\partial U'}{\partial y} \quad (2.23)$$

Notons que dans le modèle AROME, les effets de la courbure de la Terre sont considérés via une métrique spécifique, non mentionnée ici pour plus de simplicité, mais dont les détails sont par exemple exposés dans Bénard et Mašek (2013). Dans le détail, le système (2.20) se réécrit comme :

$$\frac{\partial \zeta'}{\partial t} = 0 \quad (2.24)$$

$$\frac{\partial D'}{\partial t} = -R_d \mathcal{G}^* \Delta T' + R_d T^* \mathcal{G}^* \Delta \hat{q}' - R_d T^* \Delta \hat{q}' - \frac{R_d T^*}{\pi_s^*} \Delta \pi'_s \quad (2.25)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{H^*} \mathcal{L}_v^* \hat{q}' \quad (2.26)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (D' + \mathbb{D}') \quad (2.27)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (D' + \mathbb{D}') + \mathcal{S}^* D' \quad (2.28)$$

$$\frac{\partial \pi'_s}{\partial t} = -\pi_s^* \mathcal{N}^* D' \quad (2.29)$$

où les opérateurs intégraux sont définis comme :

$$\mathcal{G}^* \Psi = \int_{\eta}^1 \frac{m^*}{\pi^*} \Psi d\eta', \quad (2.30)$$

$$\mathcal{S}^* \Psi = \frac{1}{\pi_s^*} \int_0^{\eta} m^* \Psi d\eta', \quad (2.31)$$

$$\mathcal{N}^* \Psi = \frac{1}{\pi_s^*} \int_0^1 m^* \Psi d\eta, \quad (2.32)$$

$$\tilde{\partial} \Psi = \frac{\pi^*}{m^*} \frac{\partial \Psi}{\partial \eta}. \quad (2.33)$$

$$\mathcal{L}_v^* \Psi = \tilde{\partial}(\tilde{\partial} + 1)\Psi \quad (2.34)$$

et où  $H^*$  est la hauteur caractéristique de l'atmosphère,  $c^*$  est la vitesse du son et  $N_b^*$  est la fréquence de Brunt-Väisälä :

$$H^* = \frac{R_d T^*}{g}. \quad (2.35)$$

$$c^{*2} = R_d T^* \frac{C_{pd}}{C_{vd}}, \quad (2.36)$$

$$N_b^{*2} = \frac{g^2}{C_{pd} T^*}, \quad (2.37)$$

### 2.1.5 Opérateur linéaire considéré

Des analyses de stabilité complémentaires menées par Bénard (2004) montrent que la stabilité numérique peut être encore améliorée en agissant sur les termes responsables de la propagation verticale des ondes acoustiques via le choix d'une température spécifique pour l'équation de la quantité de mouvement verticale (2.26). La stabilité numérique est améliorée si cette température  $T_e^*$  est inférieure à celle par ailleurs utilisée  $T^*$  et dont le ratio s'écrit comme :

$$r = \frac{T_e^*}{T^*} \quad (2.38)$$

Les températures actuellement choisies en opérationnel sont :  $T^* = 350$  K et  $T_e^* = 100$  K.

Les équations (2.24)–(2.29) se réécrivent alors sous la forme :

$$\frac{\partial \zeta'}{\partial t} = 0 \quad (2.39)$$

$$\frac{\partial D'}{\partial t} = -R_d \mathcal{G}^* \Delta T' + R_d T^* \mathcal{G}^* \Delta \hat{q}' - R_d T^* \Delta \hat{q}' - \frac{R_d T^*}{\pi_s^*} \Delta \pi_s' \quad (2.40)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{r H^*} \mathcal{L}_v^* \hat{q}' \quad (2.41)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (D' + \mathbb{D}') \quad (2.42)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (D' + \mathbb{D}') + \mathcal{S}^* D' \quad (2.43)$$

$$\frac{\partial \pi_s'}{\partial t} = -\pi_s^* \mathcal{N}^* D' \quad (2.44)$$

On note que dès lors que  $r \neq 1$ , l'opérateur linéaire est différent du tangent linéaire.

Ainsi l'état de base choisi dans AROME est :

— isotherme de température  $T_e^*$  et  $T^*$ , (H1)

— sans vent (au repos), (H2)

— sec, (H3)

— hydrostatique, (H4)

— homogène horizontalement sur un terrain plat ( $G = 0$ ). (H5)

L'opérateur linéaire ne dépend pas de la coordonnée horizontale mais seulement de la coordonnée verticale via les opérateurs verticaux  $\tilde{\partial}$ ,  $\mathcal{G}^*$ ,  $\mathcal{S}^*$ ,  $\mathcal{N}^*$ . Par conséquent, on considère qu'il est « à coefficients constants ».

Comme montré par Bénard et Mašek (2013), les équations (2.39)–(2.44) peuvent être combinées entre-elles pour former l'équation de structure :

$$\left[ -\frac{1}{c^{*2}} \frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2} \left( \Delta + \frac{1}{rH^{*2}} \mathcal{L}_v^* \right) + \frac{N_b^{*2}}{r} \Delta \right] \mathbb{D} = 0 \quad (2.45)$$

On souhaite désormais retrouver la relation de dispersion du modèle AROME. Pour ce faire, on considère un système physiquement stable, on se place en coordonnée  $\sigma$ , et on suit la procédure fournie dans le paragraphe 1.3.2. On décompose tout d'abord le vecteur d'état  $\Psi$  sous la forme (1.4), c'est-à-dire :

$$\Psi = \hat{\Psi} \exp(ikx) \sigma^{i\nu - \frac{1}{2}}, \quad (2.46)$$

où  $k = 2\pi/\lambda_x$  est le nombre d'onde horizontal de longueur d'onde  $\lambda_x$  et  $\nu$  est le nombre d'onde vertical adimensionnel  $\nu = 2\pi H^*/\lambda_z$  de longueur d'onde  $\lambda_z$ . On calcule ensuite les valeurs propres  $\omega^*$  de l'opérateur linéaire  $\mathcal{L}^*$  qui sont également solutions de l'équation de structure (2.45), et on obtient la relation de dispersion :

$$\omega_{\pm}^{*2} = i \frac{c^{*2}}{2} (k^2 + k_z^2) \pm i \frac{c^{*2}}{2} \sqrt{(k^2 + k_z^2)^2 - 4 \frac{N_b^{*2}}{r} \frac{k^2}{c^{*2}}}, \quad (2.47)$$

où

$$\left( k^2 + k_z^2 \right)^2 - 4 \frac{N_b^{*2}}{r} \frac{k^2}{c^{*2}} \geq 0, \quad (2.48)$$

et

$$k_z^2 = \frac{1}{rH^{*2}} \left( \nu^2 + \frac{1}{4} \right). \quad (2.49)$$

La solution  $\omega_+^{*2}$  est le mode dit « acoustique ». En calculant la racine carré de ce mode, on obtient un mode acoustique positif noté  $\omega_{++}^*$  et un mode acoustique négatif noté  $\omega_{+-}^*$ . La solution  $\omega_-^{*2}$  est le mode dit « de gravité ». En calculant la racine carré de ce mode, on obtient un mode de gravité positif noté  $\omega_{-+}^*$  et un mode de gravité négatif noté  $\omega_{--}^*$ .

## 2.1. PARTIE NUMÉRIQUE DU MODÈLE AROME

---

Le mode acoustique  $\omega_{++}^*$  et le mode de gravité  $\omega_{-+}^*$  sont représentés en traits continus sur la Figure 2.2, pour un mode de longueur d'onde verticale égale à la hauteur caractéristique de l'atmosphère  $H^*$  (correspondant à  $\nu = 2\pi$ ), et un mode « plus interne » dont la longueur d'onde  $\lambda_z \simeq 50$  m (correspondant à  $\nu = 1000$ ).

### 2.1.6 Variables pronostiques

On note tout d'abord que le système linéaire est réécrit en considérant les perturbations de divergence et de vortacité (2.21) et non les composantes zonale et méridienne du vent (2.3) : il s'agit d'une approche en grille Z. Le choix de cet état de base ne crée ainsi pas de perturbation de vortacité, ce qui annule le membre de droite de l'équation (2.24) et simplifie par conséquent le système d'équations.

Toutefois à chaque pas de temps, il est nécessaire de reconstituer le champ de vent, par exemple lors de l'appel du schéma semi-Lagrangien. Pour retrouver le champ de vent  $(U', V')$  à partir du champ de vortacité/divergence  $(\zeta', D')$ , on détermine dans un premier temps les fonctions potentiel de vitesse  $\chi'$  et de Courant  $\psi'$  en résolvant avec des solveurs les deux problèmes de Poisson :

$$D' = \Delta\chi', \quad (2.50)$$

$$\zeta' = \Delta\psi', \quad (2.51)$$

puis on retrouve les composantes zonale et méridienne du vent en dérivant spatialement comme :

$$U' = \frac{\partial\chi'}{\partial x} - \frac{\partial\psi'}{\partial y}, \quad (2.52)$$

$$V' = \frac{\partial\psi'}{\partial x} + \frac{\partial\chi'}{\partial y}. \quad (2.53)$$

Ainsi, les variables pronostiques utilisées dans l'opérateur linéaire peuvent être éventuellement différentes de celles du modèle en général, ici pour simplifier certains calculs mais parfois aussi pour des raisons de stabilité numérique. Lorsqu'une discrétisation spectrale et donc exacte est utilisée, les problèmes elliptiques (2.50) et (2.51) sont résolus exactement (et trivialement) dans l'espace spectral. Dans ce cas, les approches en grille A et A sont équivalentes.

Sur la verticale, dans la partie linéaire, la divergence verticale du vent  $\mathbb{D}$  est utilisée à la place de la simple vitesse verticale  $w$  car cette variable présente de meilleures propriétés de stabilité numérique d'après les analyses de Bénard (2003). De plus, comme mentionné précédemment, une grille de Lorenz est utilisée sur la verticale, où toutes les variables pronostiques sont calculées sur les niveaux pleins, à l'exception de la vitesse verticale  $w$  calculée sur les interfaces pour faciliter l'imposition des conditions aux limites verticales. Contrairement à la direction horizontale, une discrétisation (non exacte) en différences finies est utilisée sur la verticale. Le passage de  $\mathbb{D}$  (variable pronostique de la partie linéaire), à  $w$  (variable pronostique de la partie non-linéaire) peut néanmoins être à l'origine d'inconsistances, dégradant la qualité



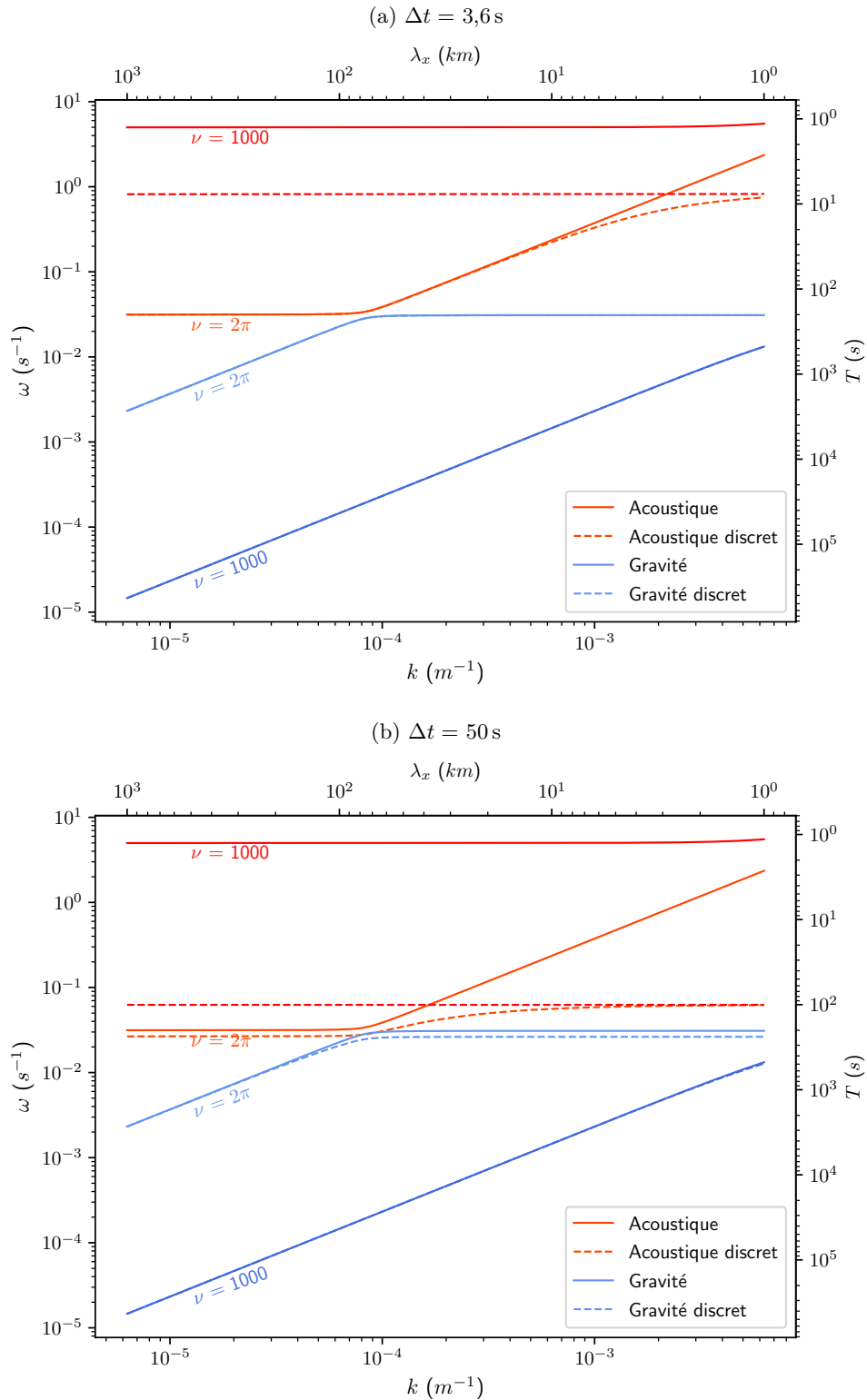


FIGURE 2.2 – Relation de dispersion de l'opérateur linéaire du modèle AROME  $\Delta x = 1300 \text{ m}$  pour un mode vertical très interne ( $\nu = 2\pi$ ) et un mode plus externe ( $\nu = 1000$ ). Les modes de gravité et acoustique sont représentés exactement (traits continus) et suite à la discrétisation temporelle (tiretés).

de la simulation, comme mentionné par Klemp *et al.* (2003).

### 2.1.7 Discrétisation spatiale

Sur la partie horizontale, les opérateurs sont décomposés suivant un algorithme de transformée de Fourier rapide, déjà exposé dans le paragraphe 1.4.10, une fois le domaine horizontal bi-périodisé.

La partie verticale est discrétisée pour conserver l'énergie et la partie hydrostatique du moment cinétique selon Bénard et Mašek (2013). Les différents opérateurs sont discrétisés suivant un schéma de différences finies à l'ordre 2 sur la grille de Lorenz précédemment mentionnée.

Sur cette grille, on distingue les niveaux pleins d'index  $l$  et les interfaces d'index  $\tilde{l}$ . Les interfaces sont indexées de  $\tilde{0}$  au sommet du modèle, jusqu'à  $\tilde{L}$  à la condition à la limite inférieure. Comme mentionné précédemment, les variables pronostiques sont toutes calculées sur les niveaux pleins, alors que les flux sont calculés sur les interfaces. Par exemple, la divergence verticale du vent  $\mathbb{D}$  est calculée sur les niveaux pleins, alors que la vitesse verticale est calculée sur les interfaces.

Une fois discrétisée verticalement, la coordonnée verticale de l'état de base (2.1) devient :

$$\pi_{\tilde{l}}^* = A_{\tilde{l}} + B_{\tilde{l}}\pi_{\tilde{L}}^* \quad (2.54)$$

La pression hydrostatique de l'état de base de la couche  $l$  est :

$$\delta\pi_l^* = \delta\pi_{\tilde{l}}^* - \pi_{\tilde{l}-1}^* \quad (2.55)$$

Les opérateurs verticaux (2.30), (2.31), (2.32), (2.34) se discrétisent respectivement comme :

$$(\mathbf{G}^*\Psi)_l = \left[ \sum_{k=l+1}^L \delta_k^*\Psi_k + \alpha_l^*\Psi_l \right] \quad (2.56)$$

$$(\mathbf{S}^*\Psi)_l = \left[ \frac{1}{\pi_l^*} \sum_{k=1}^{l-1} \delta\pi_k^*\Psi_k + \beta_l^*\Psi_l \right] \quad (2.57)$$

$$(\mathbf{N}^*\Psi)_l = \frac{1}{\pi_{\tilde{L}}^*} \sum_{k=1}^L \delta\pi_k^*\Psi_k \quad (2.58)$$

$$\mathbf{L}_v^*\Psi_l = A_l^*(\Psi_{l-1} - \Psi_l) + C_l^*(\Psi_{l+1} - \Psi_l) \quad (2.59)$$

où  $\pi_k^*$  est la pression hydrostatique de l'état de base dans la couche  $k$ , et  $\delta_k$  est la forme discrète de  $d(\log \pi^*)$  dans la couche  $k$ . Les quantités  $\alpha_l^*$  et  $\beta_l^*$  sont des incréments ajoutés aux intégrales discrètes définies sur les interfaces, pour les définir sur des niveaux pleins. Bénard

et Mašek (2013) montrent que les contraintes de conservation imposent que :

$$\alpha_l^* = \beta_l^* = 1 - \sqrt{\frac{\pi_{l-1}^*}{\pi_l^*}} \quad (2.60)$$

$$\pi_l^* = \sqrt{\pi_{l-1}^* \pi_l^*} = \frac{\delta \pi_l^*}{\delta_l^*} \quad (2.61)$$

$$\delta_l^* = \frac{\delta \pi_l^*}{\pi_l^*} \quad (2.62)$$

$$A_l^* = \frac{\pi_{l-1}^*}{\delta_l^* (\pi_l^* - \pi_{l-1}^*)} \quad (2.63)$$

$$B_l^* = -\frac{1}{\delta_l^*} \left( \frac{\pi_l^*}{\pi_l^* - \pi_{l-1}^*} + \frac{\pi_l^*}{\pi_{l+1}^* - \pi_l^*} \right) \quad (2.64)$$

$$C_l^* = \frac{\pi_{l+1}^*}{\delta_l^* (\pi_{l+1}^* - \pi_l^*)} \quad (2.65)$$

Au sommet de l'atmosphère, ces valeurs deviennent :

$$\alpha_1^* = \beta_1^* = 1 \quad (2.66)$$

$$\delta_1^* = 1 + \frac{C_{pd}}{R_d} \quad (2.67)$$

$$\pi_1^* = \frac{\delta \pi_1^*}{\delta_1^*} \quad (2.68)$$

$$A_1^* = 0 \quad (2.69)$$

$$C_1^* = \frac{\pi_2^*}{\delta_1^* (\pi_2^* - \pi_1^*)} \quad (2.70)$$

$$B_1^* = -C_1^* \quad (2.71)$$

En bas de l'atmosphère :

$$C_L^* = 0 \quad (2.72)$$

$$A_L^* = \frac{\pi_{L-1}^*}{\delta_L^* (\pi_L^* - \pi_{L-1}^*)} \quad (2.73)$$

$$B_L^* = \frac{\pi_L^*}{\delta_L^* (\pi_L^* - \pi_{L-1}^*)} \quad (2.74)$$

Une fois discrétisées spatialement, les équations (2.39)–(2.44) se réécrivent alors sous la forme :

$$\frac{\partial \zeta'}{\partial t} = 0 \quad (2.75)$$

$$\frac{\partial D'}{\partial t} = -R_d \mathbf{G}^* \Delta T' + R_d T^* \mathbf{G}^* \Delta \hat{q}' - R_d T^* \Delta \hat{q}' - \frac{R_d T^*}{\pi_s^*} \Delta \pi_s' \quad (2.76)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{r H^*} \mathbf{L}_v^* \hat{q}' \quad (2.77)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (D' + \mathbb{D}') \quad (2.78)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (D' + \mathbb{D}') + \mathbf{S}^* D' \quad (2.79)$$

$$\frac{\partial \pi_s'}{\partial t} = -\pi_s^* \mathbf{N}^* D' \quad (2.80)$$

La partie non-linéaire est ensuite discrétisée de sorte à assurer la consistance avec la partie linéaire (Bénard et Mašek, 2013), et n'est pas détaillée ici.

### 2.1.8 Discrétisation temporelle ICI

Le schéma ICI utilisé dans AROME suit la présentation donnée dans le paragraphe 1.4.8 et dont nous rappelons ici la formulation :

$$\frac{\Psi^{+(n)} - \Psi^0}{\Delta t} = \frac{\mathcal{M}(\Psi^{+(n-1)}) + \mathcal{M}(\Psi^0)}{2} + \frac{\mathcal{L}^*(\Psi^{+(n)}) - \mathcal{L}^*(\Psi^{+(n-1)})}{2} + \mathcal{P}(\Psi^\sim), \quad (2.81)$$

La version actuelle du modèle utilise un schéma à deux étapes ( $N_{iter} = 2$ ), appelé parfois aussi schéma prédicteur-correcteur, et n'est pas extrapolé temporellement. Notons que la précédente version utilisait un schéma à une seule itération ( $N_{iter} = 1$ ), extrapolé temporellement. Ce changement est intervenu en 2015 lors du passage de la résolution horizontale de 2,5 km à 1,3 km.

### 2.1.9 Inversion du problème implicite

A chacune des étapes du schéma ICI, le problème se réécrit sous la forme :

$$\left[ \mathcal{I} - \frac{\Delta t}{2} \mathcal{L}^* \right] \Psi^{+(n)} = \Psi^\bullet, \quad (2.82)$$

où :  $\mathcal{I}$  désigne la matrice identité,

$$\Psi^\bullet = X^0 + \frac{\Delta t}{2} \left[ \mathcal{R}(\Psi^{+(n-1)}) + \mathcal{R}(\Psi^0) \right] + \frac{\Delta t}{2} \mathcal{L}^*(\Psi^0), \quad (2.83)$$

et  $\mathcal{R} = \mathcal{M} - \mathcal{L}^*$  désigne le résidu non-linéaire. Le système (2.82) peut être réécrit de façon équivalente en utilisant un formalisme de matrice définies par blocs :

$$\left[ \begin{array}{c|c} \mathcal{I} & -\Delta t/2\mathcal{L}_A^* \\ \hline -\Delta t/2\mathcal{L}_B^* & \mathcal{I} - \Delta t/2\mathcal{L}_C^* \end{array} \right] \left[ \begin{array}{c} D^{+(n)} \\ \psi^{+(n)} \end{array} \right] = \left[ \begin{array}{c} D^\bullet \\ \psi^\bullet \end{array} \right]. \quad (2.84)$$

où  $\Psi = [D, \psi]^T$ ,  $\Psi^\bullet = [D^\bullet, \psi^\bullet]^T$ , et  $\psi = [\mathbb{D}, T, \hat{q}, \pi_s]^T$ . Les opérateurs  $\mathcal{L}_A^*$ ,  $\mathcal{L}_B^*$  et  $\mathcal{L}_C^*$  s'écrivent comme :

$$\mathcal{L}_A^* = \mathcal{V}_A \circ \Delta, \quad (2.85)$$

$$\mathcal{L}_B^* = \mathcal{V}_B \circ \Delta, \quad (2.86)$$

$$\mathcal{V}_A = \left[ \begin{array}{c|c|c|c} 0 & -R_d \mathbf{G}^* & R_d T^* \mathbf{G}^* - R_d T^* & -R_d T^* / \pi_s^* \end{array} \right], \quad (2.87)$$

$$\mathcal{V}_B = \left[ \begin{array}{c} 0 \\ \hline -R_d T^* / C_v \\ \hline \mathbf{S}^* - C_{pd} / C_{vd} \\ \hline -\pi_s^* \mathbf{N}^* \end{array} \right], \quad (2.88)$$

$$\mathcal{L}_C^* = \left[ \begin{array}{c|c|c|c} 0 & 0 & -g^2 / (RT_e^*) \mathbf{L}_v^* & 0 \\ \hline -RT^* / C_{vd} & 0 & 0 & 0 \\ \hline -C_{pd} / C_{vd} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right]. \quad (2.89)$$

On peut réduire algébriquement le problème sur la divergence horizontale du vent, le problème (2.84) devient alors :

$$\left[ \mathcal{I} - \frac{\Delta t^2}{4} \mathcal{B} \circ \Delta \right] D^{+(n)} = D^{\bullet\bullet} \quad (2.90)$$

où  $\mathcal{B}$  désigne l'opérateur vertical non-symétrique défini comme :

$$\mathcal{B} = \mathcal{V}_A \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C^* \right)^{-1} \mathcal{V}_B,$$

et

$$D^{\bullet\bullet} = D^\bullet - \frac{\Delta t}{2} \mathcal{L}_A^* \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C^* \right)^{-1} \psi^\bullet.$$

En projetant (2.90) dans l'espace propre de l'opérateur  $\mathcal{B}$ , on obtient :

$$\left[ 1 - \frac{\Delta t^2}{4} b_l \Delta \right] \tilde{D}_l^{+(n)} = \tilde{D}_l^{\bullet\bullet}, \quad (2.91)$$

## 2.1. PARTIE NUMÉRIQUE DU MODÈLE AROME

---

où  $b_l$  est la  $l$ -ème valeur propre de l'opérateur  $\mathcal{B}$  rangées par ordre décroissant.  $\tilde{D}_l^+$  et  $\tilde{D}_l^{\bullet\bullet}$  sont les projections des composantes  $D^+$  et  $D^{\bullet\bullet}$ , dans l'espace propre de l'opérateur  $\mathcal{B}$ .

On note que les  $N_{\text{lev}}$  problèmes implicites (membre de gauche de l'équation précédente) ont un opérateur identique à ceux d'un opérateur implicite d'un système shallow-water 1D (SW1D) de différentes vitesse de phase  $\sqrt{b_l}$ . Cette propriété sera réutilisée dans la suite, pour restreindre certaines analyses au cadre plus simple SW1D.

Il reste donc à résoudre  $N_{\text{lev}}$  problèmes horizontaux périodiques. En décomposant les différents champs en série de Fourier, on obtient  $K * K_y$  équations scalaires à inverser :

$$\left[ 1 + \frac{\Delta t^2}{4} b_l (k^2 + k_y^2) \right] \hat{D}_{l,k,k_y}^{+(n)} = \hat{D}_{l,k,k_y}^{\bullet\bullet}, \quad (2.92)$$

où  $\hat{D}_{l,k,k_y}^{+(n)}$  et  $\hat{D}_{l,k,k_y}^{\bullet\bullet}$  sont respectivement les coefficients de Fourier de  $\tilde{D}_l^{+(n)}$  et  $\tilde{D}_l^{\bullet\bullet}$  associés aux nombres d'ondes  $k$  et  $k_y$ . Notons que la discrétisation verticale est choisie pour faire en sorte que  $b_l > 0$ , par conséquent les problèmes (2.92) sont toujours inversibles, satisfaisant la condition (L2).

### 2.1.10 Schéma semi-Lagrangien

Dans ce paragraphe, nous précisons dans le cas du modèle AROME, les éléments exposés dans le paragraphe 1.4.6. En particulier, on s'intéresse à la façon dont est interfacé le schéma d'advection semi-Lagrangien avec le schéma ICI décrit dans le paragraphe 2.1.8.

La position du point origine à l'étape  $n = 0$  du schéma ICI, notée  $O^{(0)}$ , est donnée en résolvant l'équation :

$$O^{(0)} F = \frac{\Delta t}{2} \left( \mathbf{U}_F^0 + \mathbf{U}_{O^{(0)}}^0 \right) \quad (2.93)$$

où  $F$  sont les différents points « finaux » placés sur chacun des points de grille,  $\mathbf{U}$  désigne le vecteur vitesse du vent. Le schéma ICI non extrapolé à l'itération  $n = 0$  est ensuite appliqué, et on obtient  $\Psi_F^{+(0)}$  :

$$\frac{\Psi_F^{+(0)} - \Psi_{O^{(0)}}^0}{\Delta t} = \frac{\mathcal{M}(\Psi_F^0) + \mathcal{M}(\Psi_{O^{(0)}}^0)}{2} + \frac{\mathcal{L}^*(\Psi_F^{+(0)}) - \mathcal{L}^*(\Psi_F^0)}{2} \quad (2.94)$$

L'algorithme suivant est ensuite appliqué une seule fois, dans la version actuelle du modèle :

$$O^{(n)} F = \frac{\Delta t}{2} \left( \mathbf{U}_F^{+(n-1)} + \mathbf{U}_{O^{(n)}}^0 \right) \quad (2.95)$$

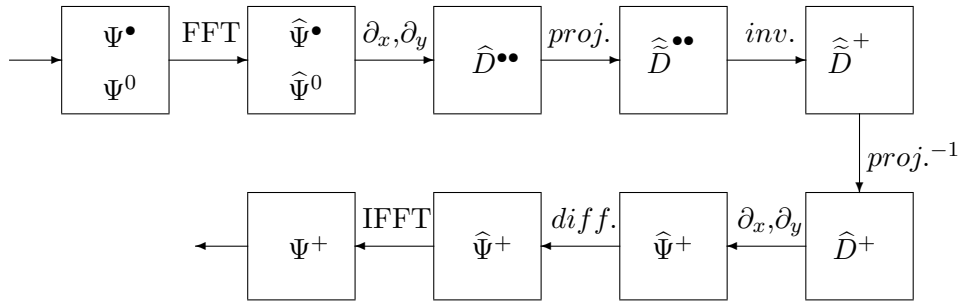
$$\frac{\Psi^{+(n)} - \Psi^0}{\Delta t} = \frac{\mathcal{M}(\Psi_F^{+(n-1)}) + \mathcal{M}(\Psi_{O^{(0)}}^0)}{2} + \frac{\mathcal{L}^*(\Psi_F^{+(n)}) - \mathcal{L}^*(\Psi_F^{+(n-1)})}{2} \quad (2.96)$$

Il existe plusieurs variantes à ces schémas, dont l'une par exemple consiste à ne pas

recalculer la position du point origine (2.95) pour économiser du coût de calcul. Pour plus d'informations sur ces différentes options utilisées dans AROME, le lecteur peut se référer à Yessad (2019b).

### 2.1.11 Fonctionnement général du code actuel

Schématiquement, le fonctionnement d'une partie du code actuel est de la forme :



FFT et IFFT désignent respectivement les appels à l'algorithme de transformée de Fourier rapide et son inverse;  $\partial_x, \partial_y$  indique des étapes de dérivation horizontales; *proj.* et *proj.*<sup>-1</sup> désignent respectivement la projection dans l'espace propre de l'opérateur  $\mathcal{B}$  et son inverse, *inv.* désigne l'inversion du problème implicite (dans l'espace spectral, il ne s'agit que de simples divisions); *diff.* désigne l'ajout implicite de la diffusion numérique.

### 2.1.12 Pas de temps

En 2020, la résolution spatiale du modèle AROME est de 1,3 km et le pas de temps choisi est de 50 s, correspondant à un nombre CFL sur les ondes rapides suivant la direction horizontale, de  $c_x^* \simeq 14$ .

Comme nous l'avons déjà montré, ces grands pas de temps ne sont utilisés que pour les modèles semi-implicites semi-Lagrangiens. Toutefois, comme montré dans le paragraphe 1.4.11 sur un simple système shallow-water en 1D, les erreurs de propagation des ondes suite à la discrétisation temporelle, augmentent avec le pas de temps. Dans un premier temps, on mène une étude comparable mais sur la relation de dispersion du problème linéaire du modèle AROME (2.47). Les résultats sont présentés sur la Figure 2.2, dans le cas où un CFL de l'ordre de l'unité est utilisé et dans la configuration actuelle où  $c_x^* \simeq 14$ . On retrouve le résultat bien connu que la méthode semi-implicite ralentit d'autant plus les ondes rapides, que le pas de temps est grand. Lorsque  $\Delta t = 50$  s, seules les ondes de gravité de grande échelle sont correctement représentées, les ondes de gravité de plus petite échelle sont ralenties et les ondes acoustiques le sont encore plus.

L'un des objectifs du modèle AROME est pourtant de prévoir la propagation des ondes de

petites échelles. Pour illustrer dans quelle mesure ce grand pas de temps dégrade la représentation de l'écoulement à petite échelle, des résultats de simulations numériques sont désormais exposés dans le cadre d'une version 2D en plan vertical du modèle AROME, exposée plus en détail dans le chapitre suivant (dans la section 3.3). Considérons pour cela un écoulement de vent horizontal uniforme (de vitesse  $\bar{U} = 10 \text{ m s}^{-1}$ ) dans une atmosphère stratifiée de fréquence de Brunt-Väisälä uniforme  $\bar{N}_b = 0,01 \text{ s}^{-1}$ , franchissant un relief de faible hauteur ( $h_{max} = 25 \text{ m}$ ) et dont la forme est celle donnée par Schär *et al.* (2002). Cette perturbation étant très petite, l'écoulement est régi par des mécanismes essentiellement linéaires. Un résultat connu établi dans ce cadre, est que dans une atmosphère stratifiée, les ondes se propagent uniquement suivant la direction verticale lorsque l'approximation hydrostatique est effectuée, et de façon inclinée par rapport à la verticale lorsque cette approximation n'est pas faite.

La Figure 2.3, représente une coupe verticale du champ de vent horizontal à une résolution horizontale  $\Delta x = 100 \text{ m}$ , et l'inclinaison du train d'onde principal par rapport à la verticale, est calculée pour quatre pas de temps différents à :  $\Delta t = 0,6 \text{ s}$  ( $c_x^* = 2.1$ ),  $\Delta t = 1 \text{ s}$  ( $c_x^* = 3.5$ ),  $\Delta t = 3 \text{ s}$  ( $c_x^* = 11.5$ ) et  $\Delta t = 6 \text{ s}$  ( $c_x^* = 21$ ). Les résultats montrent que plus le pas de temps est grand, moins le train d'onde est incliné par rapport à la verticale, c'est-à-dire plus la contribution linéaire non-hydrostatique est mal représentée. Par exemple pour un pas de temps de 6 s, le train d'onde principal est incliné de seulement  $16,8^\circ$  par rapport à la verticale, alors qu'il l'est de  $38,4^\circ$  pour un pas de temps dix fois plus petit.

Par conséquent, certains bénéfices apportés par la suppression de l'approximation hydrostatique s'amenuisent à mesure que le pas de temps augmente. Dans le cas limite où un pas de temps suffisamment grand est considéré tel que les ondes se propagent uniquement sur la verticale, la solution non-hydrostatique reproduit le même comportement que la solution hydrostatique, mais à un coût environ deux fois supérieur (une version NH nécessite environ deux fois plus de calculs qu'une version hydrostatique).

Notons que cette expérience se focalise sur un écoulement particulier (stationnaire, linéaire, etc) et n'étudie qu'une partie des contributions non-hydrostatiques. Les résultats obtenus ici ne se généralisent donc pas forcément à des écoulements plus réalistes, davantage non-linéaires. Enfin, si une représentation correcte des ondes de gravité est souhaitable, celle-ci ne doit pas être atteinte au détriment de la performance totale du modèle, dans un contexte où les contraintes opérationnelles sont particulièrement fortes (C1).

### 2.1.13 Diffusion et filtrage

Dans la version actuelle du modèle AROME, les ondes dont la longueur d'onde est inférieure à  $3\Delta x$  sont filtrées pour :

- éviter une accumulation d'énergie au bout du spectre résolu,
- éviter les phénomènes d'aliasing en présence de termes quadratiques.

Dans le cas d'AROME, elle est appliquée dans l'espace spectral et est proportionnelle à



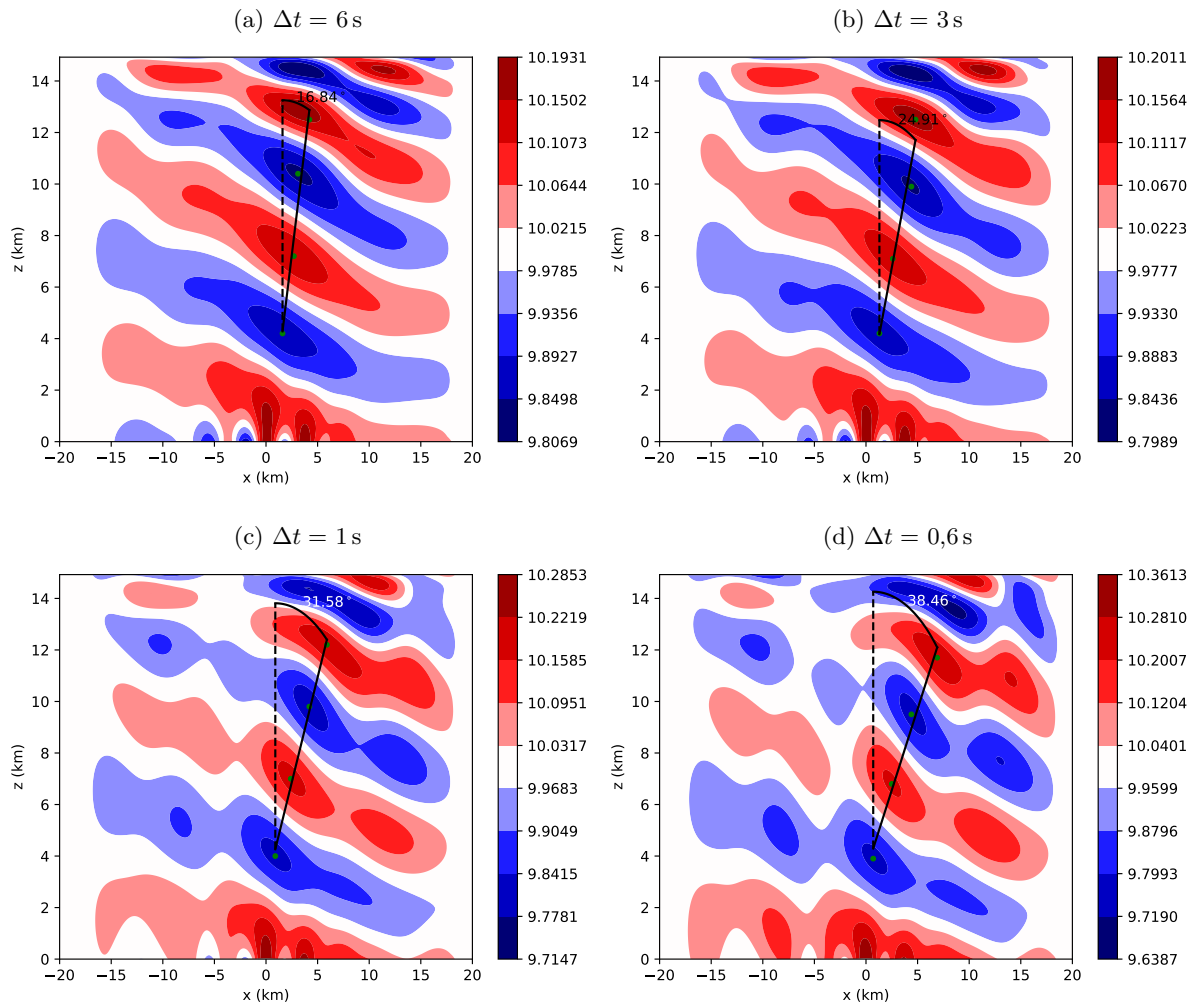


FIGURE 2.3 – Coupe verticale, du champ de vent horizontal  $U$  pour différents pas de temps, avec une résolution horizontale de  $\Delta x = 100$  m. Représentation du train d'onde principal par le segment noir. L'angle entre la verticale (tirets noirs) et le train d'onde est indiqué en degrés.

## 2.2. LIMITES DE LA VERSION ACTUELLE

---

un bi-laplacien. Pour chaque nombre d'onde  $(k, k_y)$  considéré :

$$\widehat{\Psi}_{(k, k_y)}^+ = \frac{1}{1 + D_\nu(k, k_y)\Delta t} \widehat{\Psi}_{(k, k_y)}^- \quad (2.97)$$

où  $D_\nu(k, k_y)$  est un paramètre d'amortissement qui dépend du mode horizontal considéré.

On note que d'autres mécanismes de diffusion sont également considérés (éponge au sommet du domaine, friction de Rayleigh, filtrage du terme de gradient de pression, etc). Pour plus de précision sur ces aspects dans les modèles ARPEGE, IFS et AROME, le lecteur peut se référer à Yessad (2019a). Enfin, du fait des interpolations nécessaires à chaque pas de temps, le schéma semi-Lagrangien est responsable d'une diffusion non contrôlée dans les simulations. Une alternative (SLHD) a ainsi été développée pour mieux la contrôler par Vána *et al.* (2008).

## 2.2 Limites de la version actuelle

### 2.2.1 Scalabilité

Pour que la complexité de l'algorithme de transformée de Fourier soit en  $\mathcal{O}(N_\Psi \log N_\Psi)$  (où  $N_\Psi$  est la taille du vecteur d'état), une transposition du vecteur d'état dans chacune des directions horizontale est nécessaire. Cela conduit à changer la répartition des tâches du protocole MPI selon un pavage différent (montré sur les Figures 1.1b et 1.1c) de celui considéré dans l'espace « points de grille » (Figure 1.1a). Cette étape de transposition génère des communications globales, peu adaptées aux machines massivement parallèles, pour lesquelles un modèle est d'autant plus efficace que les communications entre les nœuds sont rares. Cela pose un problème de scalabilité, devenant de plus en plus préoccupant à mesure que les supercalculateurs sont de plus en plus parallèles.

Des expériences de scalabilité ont été menées à l'occasion du projet DYAMOND (Stevens *et al.*, 2019) pour le modèle global ARPEGE avec le noyau dynamique non-hydrostatique du modèle AROME, en le faisant tourner sur 300 nœuds de calcul, à une résolution très fine de 2,5 km sur la planète entière (la version actuelle tourne sur une grille étirée à une résolution de 5 km au voisinage de la France et à 20 km aux antipodes). Bien que le modèle reste compétitif, environ 60 % du coût de calcul total est dû aux transformées spectrales. Toutefois, la géométrie sphérique impose que des transformées de type Legendre soient utilisées suivant la direction méridienne, alors qu'elles sont de type Fourier dans AROME. Or, il est connu que l'algorithme de transformée de Legendre est plus coûteux que l'algorithme de Fourier. Les conclusions obtenues avec le modèle global ARPEGE, surestiment donc probablement l'importance des transformées spectrales par rapport au modèle AROME.

Enfin, plus marginalement, comme déjà mentionné dans le paragraphe 1.4.11, le coût relativement important des transformées spectrales sur les machines parallèles, conduit à en faire le moins possible en utilisant un grand pas de temps d'intégration. Or, comme nous l'avons vu dans le paragraphe 2.1.12, les contributions linéaires non-hydrostatiques sont mal

représentées lorsqu'un grand pas de temps est utilisé. Cela aboutit potentiellement à une solution assez dégradée et peu différente de celle produite par un modèle moins coûteux faisant l'approximation hydrostatique. Pour mieux représenter la contribution linéaire non-hydrostatique, une intégration avec un plus petit pas de temps (sans forcément atteindre ceux utilisés dans les modèles HEVI) semblerait bénéfique, ce que ne permet pas la version spectrale utilisée dans un cadre opérationnel.

### 2.2.2 Fortes pentes

L'un des objectifs de la PNT dans les prochaines années est de faire tourner des modèles à résolution hectométrique. Différentes expériences ont ainsi été menées avant le début de cette thèse à des résolutions plus fines pour le modèle AROME. Ainsi pour des résolutions plus fines que 500 m, certaines instabilités numériques peuvent conduire à l'explosion numérique de la simulation.

Il a été montré que diminuer la température froide de l'opérateur linéaire  $T_e^*$ , en conservant les caractéristiques du modèle par ailleurs, permettait d'augmenter la stabilité numérique et d'aborder des pentes un peu plus fortes. Toutefois cette stratégie dégrade les scores quand  $T_e^* < 100$  K et ne peut donc être utilisée qu'avec parcimonie.

D'autres expériences ont montré que la stabilité numérique peut être améliorée en filtrant la partie de l'orographie la plus chahutée, c'est-à-dire en modifiant les propriétés spectrales du facteur orographique  $G$  de l'équation (2.2). Ce filtrage aboutit à une représentation moins réaliste du relief et annule donc en partie l'intérêt d'utiliser une résolution plus fine, et ne peut donc être utilisé qu'avec modération. Or, comme nous l'avons vu, l'état de base considéré ne contient pas les termes orographiques (H5) ( $G = 0$ ). Par conséquent ces termes sont traités dans le résidu non-linéaire de façon explicite, et sont donc soumis à des contraintes de stabilité. Il est suspecté que ce traitement explicite soit à l'origine d'instabilités numériques lorsque la résolution horizontale s'affine et que la contribution de ce terme augmente en conséquence.

### 2.2.3 Origine des problèmes

Le problème de scalabilité est principalement imputable aux transformées spectrales utilisables par le choix d'un opérateur linéaire à coefficients constants. De plus, nous pouvons supposer que le problème de stabilité numérique au voisinage des fortes pentes est également imputable à l'opérateur linéaire à coefficients constants (qui ne contient pas les termes orographiques). L'opérateur linéaire étant au cœur de ces deux problèmes, il paraît justifié de le traiter au sein d'un même travail.

### 2.3 Solutions étudiées dans la thèse

#### 2.3.1 Scalabilité

Pour contourner le problème de scalabilité, on étudie dans cette thèse l’abandon de l’algorithme de transformée de Fourier rapide (et donc des communications globales qui en résultent). Une alternative où tous les calculs sont effectués en points de grille est exposée et où le problème implicite est inversé par un solveur itératif. Les autres caractéristiques du modèle énoncées dans ce chapitre sont en revanche conservées.

Comme nous l’avons vu dans le chapitre 1, d’autres modèles opérationnels SI-SL utilisent avec succès des méthodes en points de grille (UK-UM, GEM). Toutefois certaines propriétés sont spécifiques à AROME : par exemple aucun des modèles exposés dans le chapitre 1 n’utilise une grille co-localisée (au vu des propriétés néfastes en fin de spectre mentionnées dans le paragraphe 1.4.3). De plus, les problèmes implicites sont très différents entre les modèles SI : certains d’entre eux sont à coefficients constants, d’autres à coefficients variables, etc. Aucun modèle ne combine simultanément toutes les caractéristiques du modèle AROME, nécessitant une étude spécifique qui sera l’objet du chapitre 3.

Comme nous le verrons également, l’abandon des transformées spectrales, permet *a priori* d’utiliser des pas de temps d’intégration plus faibles, pour par exemple mieux représenter certaines contributions non-hydrostatiques. Contrairement à la version spectrale, l’utilisation d’un pas de temps  $n$  fois plus petit ne nécessite pas  $n$  fois plus d’itérations du solveur. Ainsi, par la suite, certaines expériences seront menées en utilisant différents pas de temps.

#### 2.3.2 Fortes pentes

Comme nous l’avons vu dans le chapitre 1, AROME est l’un des modèles à aire limitée dont la résolution horizontale est la plus fine. De plus, son domaine, centré sur la France métropolitaine, contient des reliefs présentant de fortes pentes (comme les Alpes). Compte tenu de la géographie du territoire et de la volonté de faire tourner des modèles à résolution hectométrique, AROME est l’un des premiers modèles opérationnels à être confronté à des pentes aussi fortes.

Pour contourner le problème de stabilité numérique au voisinage de ces pentes, on étudie dans le chapitre 4, plusieurs alternatives nécessitant : peu de modifications du code dans le cas où l’approche à coefficients constants est maintenue, et des modifications plus importantes dans le cas contraire.

Dans un premier temps, à partir d’une version (soit spectrale, soit points de grille développée dans le chapitre 3) d’un modèle à coefficients constants, on change les paramètres de base qui influencent la stabilité numérique comme : le pas de temps  $\Delta t$ , la température froide de l’état de base  $T_e^*$ , et le nombre d’itérations du schéma ICI  $n'$ . Nous verrons que bien que certains de ces paramètres aient permis d’améliorer la stabilité numérique par le passé, aucun

d'entre eux ne le permet significativement en présence des fortes pentes des reliefs atteints aux échelles hectométriques.

Pour diminuer la contribution des termes orographiques traités explicitement dans l'approche à coefficients constants, une alternative est de les traiter dans le problème implicite, contournant les contraintes de stabilité numérique portant sur les termes orographiques. Il s'agirait de relâcher uniquement l'hypothèse (H5) en conservant les autres hypothèses (H1)–(H4), résultant en un problème à coefficients variables. En plus d'abandonner la discrétisation spectrale pour les raisons énoncées dans le paragraphe 1.4.10, plusieurs modifications profondes sont nécessaires et exposées dans le chapitre 4.

# Scalabilité d'AROME points de grille

---

## Sommaire

---

<b>3.1</b>	<b>Version points de grille . . . . .</b>	<b>68</b>
3.1.1	Problème implicite . . . . .	68
3.1.2	Dérivées horizontales . . . . .	68
3.1.3	Choix du solveur . . . . .	70
3.1.4	Solveur de Krylov . . . . .	71
3.1.5	Fonctionnement . . . . .	73
<b>3.2</b>	<b>Méthodologie d'estimation de la scalabilité . . . . .</b>	<b>74</b>
3.2.1	Mesure directe de la scalabilité . . . . .	75
3.2.2	Mesure indirecte de la scalabilité . . . . .	76
<b>3.3</b>	<b>AROME 2D points de grille . . . . .</b>	<b>81</b>
<b>3.4</b>	<b>Article : évaluation de la scalabilité du modèle AROME 2D PG . . .</b>	<b>82</b>
<b>3.5</b>	<b>AROME 3D partiellement points de grille . . . . .</b>	<b>100</b>
3.5.1	Présentation générale . . . . .	100
3.5.2	Critère d'arrêt . . . . .	101
3.5.3	Estimation de la qualité . . . . .	102
<b>3.6</b>	<b>Évaluation de la scalabilité du modèle AROME 3D PG . . . . .</b>	<b>102</b>
3.6.1	Paramétrage du solveur de Krylov . . . . .	103
3.6.2	Validation sur plusieurs champs et niveaux verticaux . . . . .	105
3.6.3	Comparaison sur cas réel fortement précipitant . . . . .	106
3.6.4	Scores sur un mois . . . . .	107
<b>3.7</b>	<b>Conclusion . . . . .</b>	<b>107</b>

---

Dans ce chapitre, on introduit une version du modèle AROME où les transformées spectrales sont abandonnées au profit de méthodes en points de grille. La méthodologie d'évaluation de la scalabilité est ensuite donnée. Une estimation de la scalabilité de cette nouvelle version dans un modèle simplifié sur un plan vertical faisant l'objet d'un article est ensuite exposée, puis dans un modèle partiellement porté en points de grille en 3D.

### 3.1 Version points de grille

L'approche consistant à choisir un opérateur linéaire à coefficients constants étant maintenue, la description générale faite dans le chapitre précédent reste valable ici, à l'exception de la partie concernant la discrétisation horizontale spectrale et de la résolution du problème implicite dans l'espace spectral (2.92). La version points de grille (PG) diffère de la version spectrale (SP) par :

- la discrétisation en différences finies des opérateurs de dérivées horizontales,
- l'utilisation d'un solveur de Krylov pour inverser le problème implicite (2.91).

#### 3.1.1 Problème implicite

Pour plus de lisibilité on écrit le  $l$ -ème problème (2.91) sous la forme :

$$\mathcal{H}x = y \tag{3.1}$$

où :  $x = \tilde{U}_l^{+(n)}$ ,  $y = \tilde{U}_l^{\bullet\bullet}$  et  $\mathcal{H}$  désigne l'opérateur d'Helmholtz du  $l^{\text{ème}}$  mode discrétisé spatialement et est de la forme :

$$\mathcal{H} = \left[ 1 - \frac{1}{4}c_l^{*2}\delta^2 \right] \tag{3.2}$$

où  $c_l^*$  désigne le CFL du  $l^{\text{ème}}$  mode considéré, c'est-à-dire :

$$c_l^* = \frac{\sqrt{b_l}\Delta t}{\Delta x}, \tag{3.3}$$

et  $\delta^2$  est une matrice contenant les coefficients des différences finies (introduits dans le paragraphe 1.4.3). Cette matrice sera illustrée sur la Figure 3.1a, dans le cas d'une discrétisation suivant la direction zonale ( $\delta^2 = \delta_x^2$ ) à l'ordre 2 pour les coefficients (1, -2, 1) (c'est-à-dire une matrice creuse, composée uniquement de -2 sur la diagonale principale, et de 1 sur les *sur*- et *sous*-diagonales, et au niveau des bords pour assurer la périodicité).

#### 3.1.2 Dérivées horizontales

Comme mentionné dans le paragraphe 1.4.3, contrairement à la version spectrale, il existe plusieurs manières de discrétiser les opérateurs de dérivée horizontale rencontrés dans le modèle :  $\partial_x$ ,  $\partial_y$ ,  $\Delta$ , etc. Toutefois pour toutes les raisons évoquées dans le chapitre 2, on souhaite maintenir l'utilisation d'une grille co-localisée, en utilisant :

- soit une grille A, c'est-à-dire avec les composantes du vent  $U$  et  $V$  comme variables pronostiques,
- soit une grille Z, c'est-à-dire avec la vorticit   $\zeta$  et la divergence  $D$  du vent comme variables pronostiques pour la partie linéaire.

### Grille Z

Les équations du chapitre précédent sont exposées sur une grille Z, c'est-à-dire en utilisant la divergence et la vorticité comme variables pronostiques de la partie linéaire (2.75)–(2.76). Dans ce cas, les dérivées horizontales ne sont que des dérivées secondes ( $\Delta$ ). Si la grille Z est privilégiée dans le cadre d'analyses menées sur de simples équations d'ondes (paragraphe 1.4.3), les tests conduits dans l'intégralité du modèle AROME par Caluwaerts (2016) montrent, au contraire, qu'elle présente des comportements numériques néfastes. L'origine de ces comportements, provient probablement des inconsistances entre les dérivées secondes calculées dans l'opérateur linéaire et les dérivées premières calculées dans le résidu non-linéaire. En considérant uniquement les dérivées première et seconde dans la direction zonale :  $\delta_x$  et  $\delta_x^2$ , les inconsistances se traduisent par le fait que pour un champ  $\psi$  quelconque :  $\delta_x \delta_x \psi \neq \delta_x^2 \psi$ . Pour éviter ces comportements, Caluwaerts *et al.* (2014) montrent dans le cadre d'un système shallow-water, qu'une grille Z préservant certaines propriétés de symétrie peut être utilisée de façon viable mais à un coût additionnel non négligeable.

Notons de plus que l'approche en grille Z nécessite la résolution non triviale de deux problèmes de Poisson (2.50)–(2.51), pour reconstituer le champ de vent à partir du champ de divergence et de vorticité, générant un coût de calcul supplémentaire.

### Grille A

Lorsqu'une grille A est utilisée, les équations du système linéaire (2.75)–(2.80) deviennent :

$$\frac{\partial U'}{\partial t} = -R_d \mathbf{G}^* \partial_x T' + R_d T^* \mathbf{G}^* \partial_x \hat{q}' - R_d T^* \partial_x \hat{q}' - \frac{R_d T^*}{\pi_s^*} \partial_x \pi_s' \quad (3.4)$$

$$\frac{\partial V'}{\partial t} = -R_d \mathbf{G}^* \partial_y T' + R_d T^* \mathbf{G}^* \partial_y \hat{q}' - R_d T^* \partial_y \hat{q}' - \frac{R_d T^*}{\pi_s^*} \partial_y \pi_s' \quad (3.5)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{r H_*} \mathbf{L}_v^* \hat{q}' \quad (3.6)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (\partial_x U' + \partial_y V' + \mathbb{D}') \quad (3.7)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (\partial_x U' + \partial_y V' + \mathbb{D}') + \mathbf{S}^* (\partial_x U' + \partial_y V') \quad (3.8)$$

$$\frac{\partial \pi_s'}{\partial t} = -\pi_s^* \mathbf{N}^* (\partial_x U' + \partial_y V') \quad (3.9)$$

Ainsi, lorsqu'une grille A est utilisée, seules des dérivées premières apparaissent ( $\partial_x$  et  $\partial_y$ ) dans les équations à résoudre, contrairement à la grille Z. Toutefois, dans la pratique, le problème peut être réduit algébriquement sur les composantes  $U'$  et  $V'$  du vent, faisant apparaître des dérivées secondes ( $\partial_x^2$ ,  $\partial_{xy}^2$ ,  $\partial_y^2$ ) dans l'opérateur linéaire (non montré). Cela laisse ainsi plusieurs choix possibles pour les discrétiser. Dans l'Annexe A, nous montrons que dans le cadre d'un système shallow-water en 1D, il vaut mieux privilégier une approche où



les opérateurs sont consistants entre eux (c'est-à-dire des opérateurs tels que  $\delta_x \delta_x \psi = \delta_x^2 \psi$ ). Cette conclusion est en accord avec les comportements numériques néfastes observés dans un modèle complet en grille Z lorsque  $\delta_x \delta_x \psi \neq \delta_x^2 \psi$ .

Notons également que la réduction algébrique sur les composantes du vent, conduit à inverser un opérateur deux fois plus grand que lorsque le problème est réduit sur la divergence du vent. Toutefois ce problème ne nécessite pas de résoudre les deux problèmes de Poisson.

## Bilan

Notons que l'approche qui nous semble être la plus viable, est celle de réduire algébriquement le problème sur la divergence du vent, puis de reconstituer le champ de vent à partir des équations classiques de la quantité de mouvement (3.4) (3.5). Cela permettrait d'une part d'inverser un problème ne portant que sur une variable pronostique, et d'autre part ne nécessiterait pas de résoudre les deux problèmes de Poisson.

Néanmoins comme nous le verrons ci-après, certains modèles seront uniquement en deux dimensions sur un plan vertical, simplifiant considérablement les problèmes mentionnés dans ce paragraphe. Des tests seront également menés en 3D, mais sur une version où seule une partie du code aura été portée en points de grille. Par conséquent, cette solution n'a pas pu être testée et est laissée en perspectives.

### 3.1.3 Choix du solveur

Les solveurs itératifs sont préférés aux solveurs directs s'ils convergent vers une solution de qualité acceptable en peu d'itérations. Le choix du solveur dépend du problème considéré et sont, en météorologie, principalement de deux types :

- Helmholtz pour les modèles SI,
- Poisson pour les modèles anélastiques.

S'il s'agit dans les deux cas d'inverser un problème, les propriétés diffèrent entre les deux.

Le conditionnement d'un opérateur mesure le ratio entre la plus grande et la plus petite des valeurs singulières (en valeur absolue). Lorsque la matrice est symétrique, il se réduit au ratio entre la plus grande et la plus petite des valeurs propres (en valeur absolue). Lorsqu'un grand domaine horizontal est considéré (de longueur  $L_x$ ), le conditionnement pour chaque problème, devient :

- pour le  $l^{\text{ème}}$  problème d'Helmholtz obtenu suite à une approche à coefficients constants :

$$C_l \simeq \frac{1 + b_l \Delta t^2 \pi^2 / (4 \Delta x^2)}{1 + b_l \Delta t^2 \pi^2 / L_x^2} \underset{L_x \gg \Delta x}{\simeq} 1 + \frac{\pi^2}{4} c_l^{*2}; \quad (3.10)$$

### 3.1. VERSION POINTS DE GRILLE

---

— pour un problème de Poisson :

$$C' \simeq \frac{k_{max}^2}{k_{min}^2} = \frac{L_x^2}{4\Delta x^2} \simeq \frac{N_x^2}{4}. \quad (3.11)$$

où  $N_x$  est le nombre de points suivant la direction zonale.

Dans le cas d'un problème de Poisson, le conditionnement  $C'$  se dégrade quand la résolution s'affine, nécessitant des solveurs appropriés pour traiter les fines résolutions du futur. Ainsi, les méthodes multi-grilles apparaissent particulièrement bien adaptées pour résoudre des problèmes dont le rapport d'échelle est important (Fulton *et al.*, 1986). En revanche, dans le cas d'un problème d'Helmholtz formé à partir d'un opérateur linéaire à coefficients constants, l'affinement de la résolution horizontale ne dégrade pas le conditionnement tant que le pas de temps est diminué de sorte à maintenir un CFL constant (ce qui a été souvent le cas à chaque changement de résolution effectué jusqu'ici). Si les méthodes multi-grilles peuvent effectivement être utilisées pour accélérer la convergence, leur utilisation semble moins indispensable que pour un problème de Poisson et sont généralement utilisées comme préconditionneur d'une méthode de Krylov (Degrauwe *et al.*, 2020).

Notons que lorsque le pas de temps  $\Delta t$  tend vers l'infini, le conditionnement du problème d'Helmholtz tend vers celui d'un problème de Poisson :

$$C_l \underset{\Delta t \rightarrow \infty}{\simeq} \frac{L_x^2}{4\Delta x^2} = C'. \quad (3.12)$$

Cette estimation reflète la nature différente des deux problèmes considérés : le problème de Poisson est elliptique alors que celui d'Helmholtz est hyperbolique. Le conditionnement du problème de Poisson peut être vu comme le conditionnement d'un problème d'Helmholtz non borné temporellement. Dans la pratique, bien que grand, le pas de temps utilisé dans un modèle SI est fini et l'usage d'un préconditionneur multi-grille est restreint aux modes dont le CFL  $c_l^*$  est le plus fort (Degrauwe *et al.*, 2020).

Dans le cadre de cette étude, un solveur de Krylov non préconditionné sera utilisé. Nous montrerons qu'une convergence très rapide peut quand même être atteinte grâce à un paramétrage minutieux du solveur de Krylov. La nature hyperbolique du problème d'Helmholtz sera notamment exploitée dans la section 3.5 pour définir un critère d'arrêt.

#### 3.1.4 Solveur de Krylov

Dans l'article, nous montrerons l'intérêt de maintenir la projection sur les modes propres de l'opérateur vertical  $\mathcal{B}$  dans la version PG. Il faut ainsi résoudre  $N_{lev}$  problèmes horizontaux, tous symétriques et définis positifs (2.91). Pour exploiter la définie positivité des problèmes, on utilise comme méthode de Krylov, un algorithme de gradient conjugué décrit en détail ci-dessous. L'utilisation d'un solveur itératif impose de l'initialiser et de déterminer un critère d'arrêt pertinent.

### Initialisation

Une approche intéressante pour diminuer le résidu initial, est d'initialiser le solveur par une estimation de l'état de l'atmosphère  $x_0$  aussi proche que possible de la solution recherchée. Le meilleur moyen que nous ayons trouvé est de prendre le champ du pas de temps précédent pour initialiser le solveur lors de la première étape du schéma ICI, et le champ obtenu suite à la première étape du schéma ICI pour initialiser la seconde. Dans la section 3.6, nous montrerons l'intérêt d'une telle initialisation, par rapport au cas répandu où le solveur est initialisé par un champ nul. La solution de chacun des  $N_{\text{lev}}$  problèmes est alors cherchée dans les sous-espaces de taille  $N$  (où  $N$  dépend du  $l^{\text{ème}}$  problème considéré) :

$$\mathcal{K}_N = \{r_0, \mathcal{H}r_0, \mathcal{H}^2r_0, \dots, \mathcal{H}^N r_0\} \quad (3.13)$$

où  $r_0 = y - \mathcal{H}x_0$  est le résidu initial.

### Critère d'arrêt

Le système linéaire étant symétrique, alors selon la Proposition 4 de Saad et Schultz (1986) l'inégalité suivante est valide :

$$\frac{\|r_{n_i}\|}{\|r_0\|} \leq \left(1 - \frac{1}{C_l^2}\right)^{n_i/2}, \quad (3.14)$$

où le résidu à l'itération  $n_i$  est  $r_{n_i} = y - \mathcal{H}x_{n_i}$ . Cette inégalité représente le pire cas de convergence et dans la plupart des cas, la convergence est bien plus rapide (Axelsson, 1996). Une manière évidente et répandue pour imposer le critère d'arrêt, est de sortir du solveur lorsque le résidu est réduit d'un certain facteur par rapport à la norme du membre de droite :

$$\text{Arrêt si } \frac{\|r_{n_i}\|}{\|y\|} \leq \varepsilon, \quad (3.15)$$

où  $\varepsilon < 1$  est un seuil limite fixé par l'utilisateur. Dans toute la suite ce seuil sera appelé « critère d'arrêt ». Ce critère d'arrêt sera utilisé dans la section 3.4 pour AROME 2D en le considérant identique pour les  $N_{\text{lev}}$  problèmes.

Dans un cadre à visée plus opérationnelle, exposé dans la section 3.6, on préférera imposer le nombre maximum d'itérations à l'avance pour que la prévision soit fournie en un temps fixe. Pour le régler le plus finement possible, ce critère dépend du  $l^{\text{ème}}$  problème considéré et devient alors :

$$\text{Arrêt si } n_i \geq N_{\text{max},l}, \quad (3.16)$$

Dans la section 3.6, nous exposerons la façon dont nous avons réglé  $N_{\text{max},l}$ .

#### Algorithme

L'algorithme de gradient conjugué est exposé ci-dessous :

```

1 Calculer  $r_0 = y - \mathcal{H}x_0$  et  $p_0 = r_0$ ;
2 tant que critère faire
3    $\alpha_{n_i} = r_{n_i}^T r_{n_i} / (p_{n_i}^T \mathcal{H} p_{n_i})$ ;
4    $x_{n_i+1} = x_{n_i} + \alpha_{n_i} p_{n_i}$ ;
5    $r_{n_i+1} = r_{n_i} - \alpha_{n_i} \mathcal{H} p_{n_i}$ ;
6    $\beta_{n_i} = r_{n_i+1}^T r_{n_i+1} / (r_{n_i}^T r_{n_i})$ ;
7    $p_{n_i+1} = r_{n_i+1} + \beta_{n_i} p_{n_i}$ ;
8 fin

```

où « *critère* » désigne :

- «  $r_{n_i} < \varepsilon$  », dans le cas d'un critère d'arrêt imposé sur l'écart de la solution courante à la solution convergée;
- «  $n_i < N_{max,l}$  », dans le cas où un critère basé sur le nombre maximum d'itérations est imposé.

#### Communications

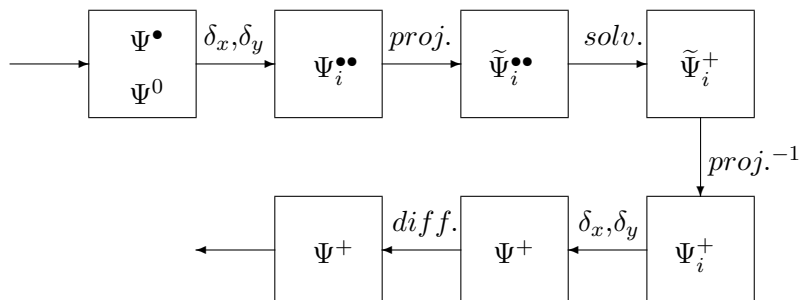
Les opérations susceptibles de faire intervenir des communications dans l'algorithme de gradient conjugué sont :

- lignes 3 et 5 : un calcul de produit-matrice vecteur;
- ligne 3 : un calcul d'un produit scalaire;
- ligne 6 : un autre calcul d'un produit scalaire.

Le calcul des produits scalaires génère des communications all-reduce lorsque le protocole de parallélisation MPI est utilisé. Les produits matrice-vecteur font intervenir des communications locales qui dépendent notamment du halo de la matrice considérée. On montre ainsi que la quantité de communication dépend du nombre d'itérations du solveur. Une estimation précise de la convergence du solveur est essentielle pour évaluer le coût en communication et donc estimer la scalabilité.

#### 3.1.5 Fonctionnement

Le fonctionnement d'une partie du noyau dynamique du code exposée dans le paragraphe 2.1.11 pour la version spectrale, devient pour la version points de grille :



$\Psi_i$  est l'une des variables pronostiques sur laquelle le problème est réduit algébriquement (en général la divergence horizontale du vent  $D$  ou le vent  $\mathbf{U}$ ). Les notations utilisées sont identiques à celles du paragraphe 2.1.11, sauf que *solv.* désigne désormais l'application d'un solveur en points de grille pour inverser séquentiellement les  $N_{\text{lev}}$  problèmes implicites projetés dans l'espace propre de l'opérateur vertical  $\mathcal{B}$ . Comme nous le verrons notamment dans l'article, un traitement séquentiel des modes est privilégié pour que les modes qui ne requièrent que peu d'itérations à être inversés n'attendent pas ceux qui en requièrent plus. L'utilisation du solveur a été exposée dans le paragraphe 3.1.4.  $\delta_x, \delta_y$  désigne l'étape d'application des opérateurs de dérivées horizontales discrétisés en différences finies dont le traitement a été exposé dans le paragraphe 3.1.2. Enfin, contrairement à la version spectrale, la diffusion numérique (*diff.*), n'est généralement pas ajoutée implicitement dans une version points de grille. Des filtres spatiaux sont souvent utilisés à la place, comme évoqué dans le paragraphe 1.4.4.

Malgré l'apparente simplicité du schéma ci-dessus, le codage d'une nouvelle version entièrement en points de grille dans le modèle AROME est une tâche ambitieuse qui nécessite un important travail de développement.

Avant d'entreprendre un tel travail, on souhaite tout d'abord déterminer si l'abandon de la version spectrale améliore effectivement la scalabilité sur des cas simplifiés. La méthodologie utilisée pour estimer la scalabilité est exposée dans la section 3.2 et comparée aux estimations traditionnellement menées. Ces estimations se feront dans une version simplifiée en 2D du modèle AROME décrite dans la section 3.3 et dont les résultats sont présentés dans un article qui constitue la section 3.4. Les résultats encourageants exposés dans cet article ont poussé à l'extension des tests dans un cadre plus réaliste, en 3D. Ainsi, à partir d'un cycle récent du modèle opérationnel AROME, une version partiellement en points de grille a été développée et est exposée dans la section 3.5. Dans la section 3.6, des comparaisons sur des cas réels ont ainsi pu être menées et des estimations de scalabilité ont été déduites dans un cadre plus réaliste.

## 3.2 Méthodologie d'estimation de la scalabilité

Mesurer directement la scalabilité d'un modèle sur une machine permet d'identifier les problèmes de performance éventuels et d'ajuster ainsi les traitements numériques à la tech-

## 3.2. MÉTHODOLOGIE D'ESTIMATION DE LA SCALABILITÉ

---

nologie et réciproquement. Pour obtenir des mesures de qualité, il faut disposer d'un modèle complet et parallélisé, ce qui n'est précisément pas le cas lorsque l'étude de viabilité porte sur l'intérêt de coder un tel modèle, comme dans cette thèse.

Une alternative serait donc de mesurer la scalabilité sur quelques routines bien ciblées, comme par exemple le solveur du problème implicite pour un modèle SI-SL. Nous montrerons néanmoins que cette démarche ne retranscrit pas toujours fidèlement la scalabilité d'un modèle complet.

De plus, comme nous l'avons vu précédemment, les modèles possédant de nombreuses caractéristiques en commun, on pourrait s'attendre à ce que les mesures de scalabilité menées sur les uns puissent être informatives pour les autres. Toutefois, dans le paragraphe suivant, nous montrerons que ces estimations dépendent de la machine sur laquelle les mesures ont été faites, et du modèle en lui-même.

Pour toutes ces raisons, la scalabilité de la version points de grille du modèle AROME sera estimée par une méthode simple et indirecte dans cette thèse.

### 3.2.1 Mesure directe de la scalabilité

#### Une approche dépendante de la machine

Les expériences de scalabilité sont généralement menées sur les machines actuelles et parfois sur les plus puissantes, dont on estime qu'elles deviendront la norme dans le futur. Par exemple, le CEPMMT a mené récemment des expériences de scalabilité de leur modèle IFS sur les machines les plus puissantes actuellement disponibles : le supercalculateur *Piz Daint* du constructeur Cray et *Summit* de IBM. Si les architectures du futur seront en effet probablement massivement parallèles, de nombreuses incertitudes demeurent quant à leurs performances et à leurs caractéristiques principales (proportion de GPU/CPU, etc). Ainsi les conclusions dressées suite à ces expériences ne peuvent être extrapolées aux machines du futur qu'avec précaution. Notons également que le choix d'un supercalculateur repose sur un appel d'offre entre différents supercalculateurs, certes tous massivement parallèles, mais aux caractéristiques parfois assez différentes. Ainsi quelques degrés de liberté sont laissés au centre météorologique lors de l'appel d'offre pour choisir le supercalculateur le plus adapté à son modèle de PNT.

#### Une approche dépendante du problème considéré

Des mesures de scalabilité ont déjà été menées pour plusieurs modèles SI en points de grille. Toutefois, deux modèles SI-PG peuvent présenter des vitesses de convergence radicalement différentes pour inverser le problème implicite en fonction de l'opérateur linéaire considéré (coefficients constants versus coefficients variables, etc). Par exemple, certains problèmes d'Helmholtz sont mal conditionnés voire non inversibles comme ceux du noyau dynamique

« New dynamics » utilisé par le modèle UK-UM (Davies *et al.*, 2005), alors que d'autres, comme ceux du modèle AROME, sont symétriques et définis positifs et bien conditionnés (3.10). Ces différences se répercutent sur la vitesse de convergence du solveur.

La scalabilité étant notamment dépendante de la vitesse de convergence, les mesures de scalabilité obtenues pour les uns ne sont pas directement transposables aux autres. Une étude spécifique est donc requise pour la version points de grille du modèle AROME.

### Une approche isolée du reste du modèle

Comme mentionné par Liesen et Tichý (2004), la vitesse de convergence du solveur ne dépend pas uniquement du conditionnement du problème à inverser, mais également de son initialisation ainsi que des propriétés spectrales du membre de droite. Par simplicité, les tests de scalabilité sont parfois menés sur une partie du code ne contenant que le seul solveur, en lui prescrivant plus ou moins arbitrairement un membre de droite de l'équation à résoudre.

Toutefois, comme nous le montrerons dans la section 3.6, la vitesse de convergence peut être significativement améliorée lorsque des champs précédemment calculés dans le reste du modèle servent à initialiser le solveur. Ceci n'est pas possible lorsque le solveur est évalué isolément du reste du modèle où les vitesses de convergence obtenues peuvent être sous-estimées. Par exemple, pour évaluer le coût en communication Zheng et Marguinaud (2018) prescrivent 20 itérations du solveur de Krylov pour inverser le problème dans leur simulateur de performance. Nous montrerons dans la suite qu'une convergence plus rapide peut être atteinte, impactant ainsi les conclusions de l'étude.

### 3.2.2 Mesure indirecte de la scalabilité

Comme mentionné précédemment, l'intérêt de cette étude de viabilité est d'estimer la scalabilité sans avoir à développer une version en points de grille coûteuse en développement, pour estimer s'il est viable d'entreprendre un tel codage dans l'intégralité du modèle AROME. Par conséquent, des tests de scalabilité avancés dans une version en points de grille ne sont pas à l'ordre du jour.

#### Lien entre vitesse de convergence et coût en communication

L'intérêt principal d'une méthode itérative est de n'inverser que partiellement le problème jusqu'à une certaine précision en limitant le nombre d'itérations. Si cette méthode converge en un nombre important d'itérations (c'est-à-dire en un nombre égal à la taille du problème) il est peu probable que cette méthode soit compétitive par rapport à un algorithme direct, comme par exemple l'algorithme de transformée de Fourier rapide.

Comme nous l'avons vu à la fin du paragraphe 3.1.4, le coût en communication dépend

### 3.2. MÉTHODOLOGIE D'ESTIMATION DE LA SCALABILITÉ

---

fortement de la vitesse de convergence. Si une convergence rapide est atteinte, peu d'applications d'opérateurs et de calculs de produits scalaires sont nécessaires, et donc peu de communications ont lieu.

Comme exposé dans le chapitre 1, dans le paradigme de parallélisation actuel, un nœud de calcul gère la partie du domaine horizontal qui lui est allouée. Ainsi lorsqu'un produit matrice-vecteur est effectué dans l'algorithme de Krylov, des communications ne vont avoir lieu que sur les bords du domaine géré par le nœud. Dans toute la suite, on considère que les produits scalaires qui engendrent des communications all-reduce ont un coût en communication négligeable ( $\mathcal{C}(\langle, \rangle) \simeq 0$ ). Cette hypothèse repose par exemple sur les travaux de Zheng et Marguinaud (2018), où il est montré que le coût des communications all-reduce est faible, tant que moins de  $10^5$  nœuds de calculs sont utilisés. Cela devrait être encore le cas pendant plusieurs années. Toutefois si ces calculs venaient à être trop coûteux dans le futur, Degrauwe *et al.* (2020) suggèrent d'utiliser un solveur de Richardson préconditionné à la place du solveur de Krylov.

Supposons un domaine horizontal périodique uniquement 1D composé de 3 nœuds de calcul se partageant équitablement les calculs des 18 points de grille. Considérons de plus, un laplacien 1D discrétisé à l'ordre 2 de sorte que ces coefficients de différences finies soient (selon le formalisme introduit dans le paragraphe 1.4.3) :

$$(1, -2, 1)$$

Lorsqu'un problème d'Helmoltz à coefficients constants  $\mathcal{H}$  est considéré avec une discrétisation à 2 niveaux temporels, les coefficients deviennent alors :

$$\left( -\frac{1}{4}c^{*2}, 1 + \frac{2}{4}c^{*2}, -\frac{1}{4}c^{*2} \right)$$

Ces coefficients sont représentés par des couleurs différentes entre la diagonale principale et les sur et sous diagonales sur les Figures 3.1a et 3.1c. Les zéros y sont représentés par des carrés blancs et remplissent la majorité des figures, montrant ainsi le caractère creux de la matrice.

Le partage entre les différents nœuds est symbolisé par des traits verticaux gris. Ainsi, lors d'un produit matrice-vecteur, le premier tiers du produit sera effectué par le premier nœud, le second tiers par le second nœud, et ainsi de suite. Les communications inter-nœuds n'ont lieu qu'aux bords de chacun des domaines et sont symbolisés par un carré noir sur les Figures 3.1b et 3.1d. Plus les carrés noirs sont nombreux, plus le nombre de communications est important.

Lorsque 3 itérations sont considérées, l'opérateur qui en résulte,  $\mathcal{H}^3$ , est l'application de trois fois l'opérateur  $\mathcal{H}$ . Le halo est ainsi plus étendu et la matrice est désormais moins creuse (Figure 3.1c). Le nombre de communications augmente en conséquence (Figure 3.1d).

Considérons désormais un laplacien discrétisé à l'ordre 4 tel que ces coefficients en diffé-



rences finies soient :

$$\left(-1/12, 4/3, -5/2, 4/3, -1/12\right),$$

Comme montré sur la Figure 3.2, le halo est plus étendu, résultant en un nombre plus important de communications. La combinaison d'un ordre élevé et d'un nombre d'itérations plus important accroît encore le nombre de communications.

Dans cette thèse, nous faisons l'hypothèse que le coût en communication à un ordre donné, dépend uniquement du nombre de fois où les opérateurs sont appliqués, c'est-à-dire de la vitesse de convergence du solveur.

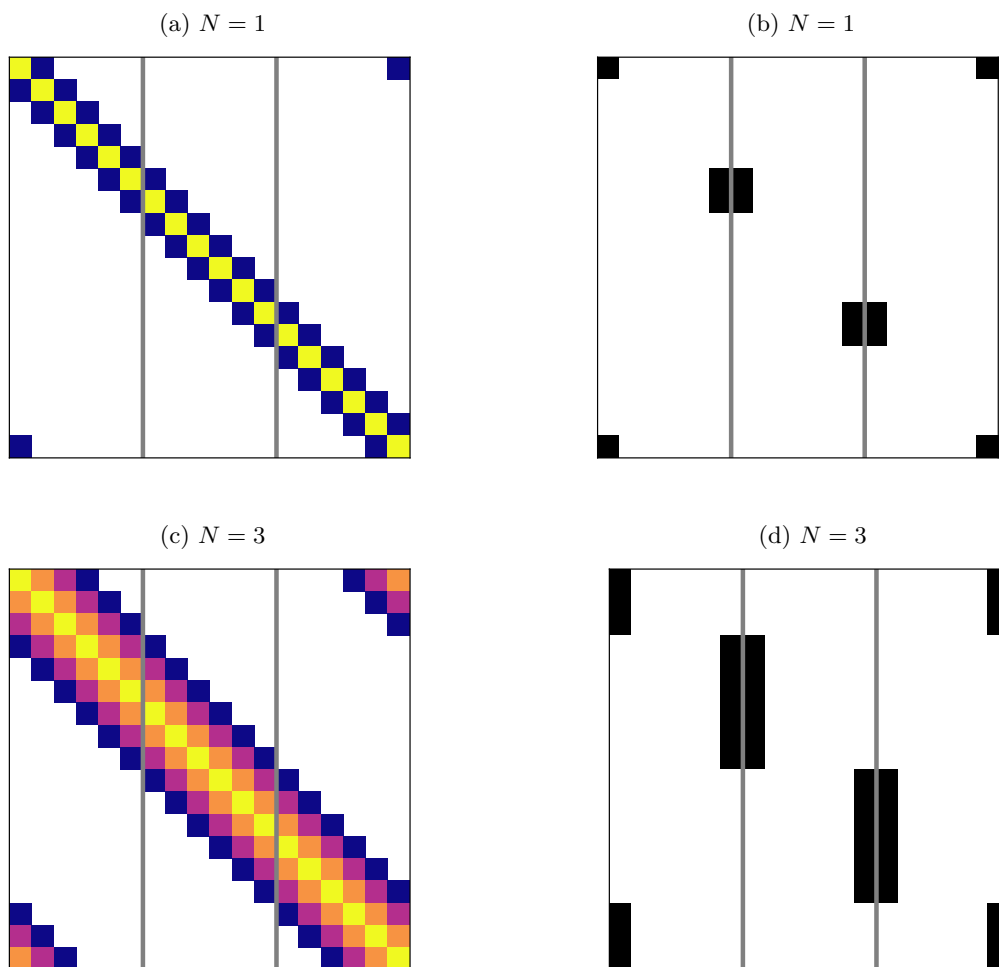


FIGURE 3.1 – Matrice  $\mathcal{H}$  (à gauche) et graphique de communication (à droite) pour une dérivée seconde discrétisée à l'ordre 2.

### Communications d'un modèle SI

On appelle  $\mathcal{C}(\mathcal{H})$  le coût de communication d'une application de l'opérateur  $\mathcal{H}$ . Le halo étant identique pour chacun des  $N_{\text{lev}}$  problèmes de Helmholtz, le coût d'une seule application

### 3.2. MÉTHODOLOGIE D'ESTIMATION DE LA SCALABILITÉ

de l'opérateur  $\mathcal{H}$  ne dépend pas du mode considéré. En revanche, tous les modes ne requièrent pas un nombre d'applications identique. Le coût de communication du  $l^{\text{ème}}$  problème pour faire une prévision après un pas de temps  $\Delta t$  est donc :  $N_l \mathcal{C}(\mathcal{H})$  où  $N_l$  est le nombre d'itérations pour inverser le  $l^{\text{ème}}$  problème. Or, les  $N_{\text{lev}}$  problèmes sont inversés séquentiellement. Par conséquent, le coût total en communication moyenné sur les  $N_{\text{lev}}$  niveaux verticaux après un pas de temps  $\Delta t$ , noté  $\mathcal{C}_t$ , devient pour un modèle SI :

$$\mathcal{C}_t(\mathcal{H}) = \bar{N} \mathcal{C}(\mathcal{H}) \quad (3.17)$$

où  $\bar{N} = 1/N_{\text{lev}} \sum_{l=1}^{N_{\text{lev}}} N_l$  désigne le nombre moyen d'itérations pour inverser les  $N_{\text{lev}}$  niveaux pour un pas de temps.

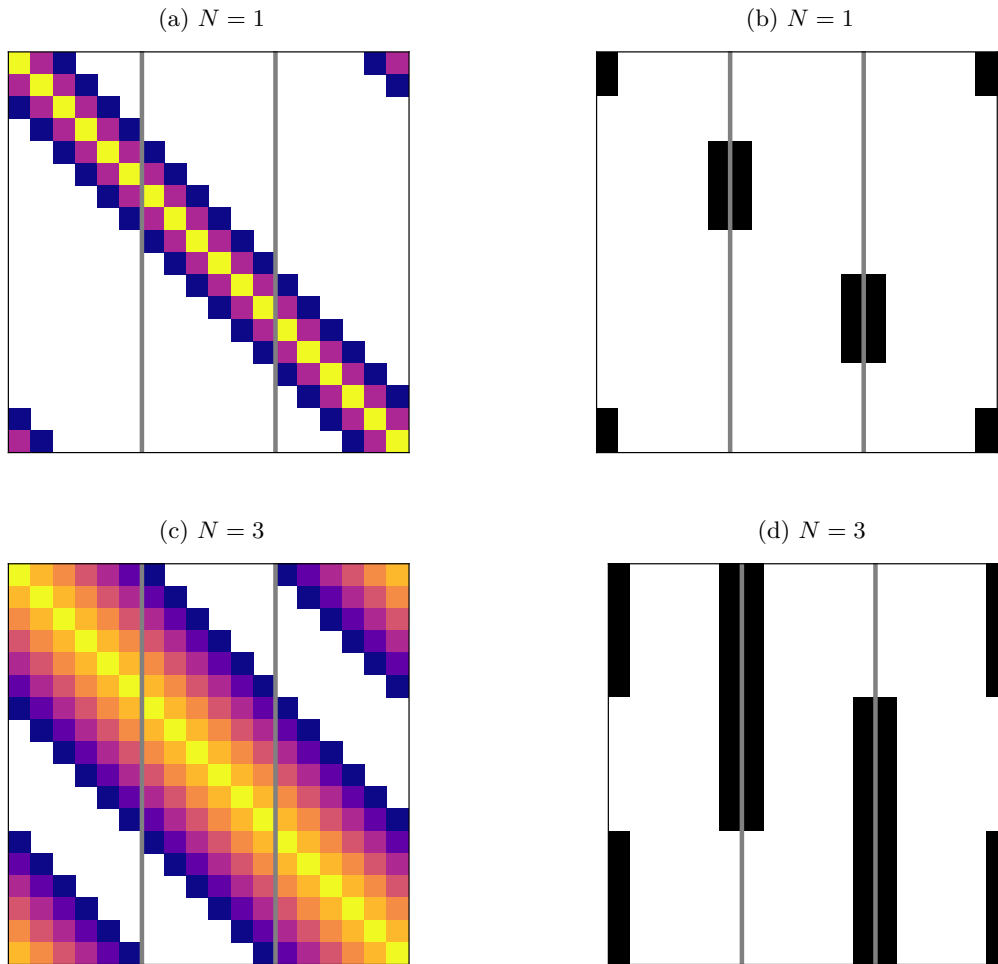


FIGURE 3.2 – Matrice  $\mathcal{H}$  (à gauche) et graphique de communication (à droite) pour une dérivée seconde discrétisée à l'ordre 4.

### Communications d'un modèle HEVI

Les modèles HEVI utilisent majoritairement des schémas split-explicit (SE), bien adaptés au paradigme de parallélisation du fait de l'application d'opérateurs locaux suivant la direction horizontale et d'opérateurs inverses suivant la direction verticale. Dans notre étude, ils seront considérés comme une référence en terme de scalabilité.

Sans trop entrer dans les détails, les termes responsables de la propagation horizontale des ondes les plus rapides, notés  $\mathcal{H}$ , sont traités avec un petit pas de temps respectant une contrainte CFL de la forme :

$$\frac{c_s \Delta \tau}{\Delta x} \leq \alpha \quad (3.18)$$

où le facteur  $\alpha \simeq 1/\sqrt{2}$  est donné à titre indicatif et est amené à varier en fonction des ordres de discrétisation spatial et temporel utilisés, de la grille considérée, etc. Les autres termes sont intégrés avec un plus grand pas de temps, pour économiser des évaluations. Par conséquent, durant l'intégration des termes responsables de la propagation des ondes rapides, les variables pronostiques sont réactualisées comme :

$$\Psi^{t+(n+1)\Delta\tau} = \mathcal{H}' \Psi^{t+n\Delta\tau} \quad (3.19)$$

où :  $\Psi^{t+n\Delta\tau}$  est le vecteur d'état à l'instant  $t + n\Delta\tau$ ,  $t$  est le précédent long pas de temps. Pour faire une prévision jusqu'à l'instant  $t + \Delta t$ , où  $\Delta t$  a la même valeur que le pas de temps du SI, plusieurs applications de (3.19) sont requises. Après  $N'$  petits pas de temps, la solution est :

$$\Psi^{t+N'\Delta\tau} = \mathcal{H}'^{N'} \Psi^t \quad (3.20)$$

Le coût total en communication  $\mathcal{C}_t$  pour faire une prévision après un pas de temps  $\Delta t$ , pour un seul niveau vertical, est donc pour un modèle HEVI-SE :

$$\mathcal{C}_t(\mathcal{H}') = N' \mathcal{C}(\mathcal{H}') \quad (3.21)$$

### Comparaison des coûts de communications

Les opérateurs  $\mathcal{H}$  et  $\mathcal{H}'$  régissent la propagation horizontale des ondes rapides. En considérant que ces deux opérateurs ont le même halo, alors le coût en communication  $\mathcal{C}$  d'une application de l'opérateur  $\mathcal{H}$  est similaire à celui d'une application de l'opérateur  $\mathcal{H}'$  :

$$\mathcal{C}(\mathcal{H}) \simeq \mathcal{C}(\mathcal{H}') \quad (3.22)$$

Par conséquent le coût total en communication après un pas de temps  $\Delta t$ , devient respectivement pour un modèle HEVI et SI :

$$\mathcal{C}_t(\mathcal{H}') \simeq N' \mathcal{C}(\mathcal{H}) \quad (3.23)$$

$$\mathcal{C}_t(\mathcal{H}) \simeq \bar{N} \mathcal{C}(\mathcal{H}) \quad (3.24)$$

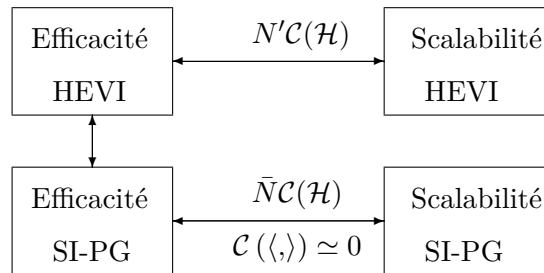
### 3.3. AROME 2D POINTS DE GRILLE

---

La comparaison du coût en communication se réduit ainsi à la comparaison des nombres  $\bar{N}$  et  $N'$ .

#### Résumé

Schématiquement cette comparaison peut se représenter comme :



Comme mentionné précédemment, la stratégie suivie ici est de ne pas mesurer directement la scalabilité. Notre approche est de comparer la version points de grille du modèle AROME à des modèles dont les communications sont du même type, comme les modèles HEVI-SE. En revanche, cette approche nous prive d'une comparaison avec la version spectrale actuellement utilisée dans AROME (présentant des communications globales all-to-all). Estimer la scalabilité de cette nouvelle version se réduit à comparer l'efficacité de la version points de grille du modèle AROME à celle d'un modèle HEVI.

Toutefois, en l'absence d'une version HEVI du modèle AROME, l'efficacité des modèles HEVI est prescrite par rapport aux modèles habituellement utilisés ailleurs, générant quelques incertitudes sur ces estimations. Seuls les ordres de grandeur seront par conséquent analysés dans toute la suite.

Ainsi sous ces hypothèses, une efficacité comparable obtenue entre la version AROME-PG et un modèle HEVI-SE, signifierait une scalabilité équivalente entre ces deux modèles. Ces derniers étant réputés particulièrement scalables dans le paradigme de parallélisation actuel, la scalabilité de la version PG serait améliorée par rapport à celle de la version spectrale actuelle.

### 3.3 AROME 2D points de grille

Avant le début de la thèse, L. Auger a développé une version simplifiée en 2D du modèle AROME sur un plan vertical. Ce modèle reprend les caractéristiques générales du modèle opérationnel (équations, schéma SI-SL, coordonnée masse, etc), sans paramétrer les phénomènes sous maille. Le noyau dynamique actuel y est fidèlement reproduit et permet ainsi d'étudier des écoulements idéalisés, largement utilisés dans la littérature. Ce code est capable de tourner

dans deux versions : une version spectrale (SP) et une version entièrement en points de grille (PG), facilitant les comparaisons. Ce code n'est pas parallélisé et ne permet donc pas de faire des mesures de scalabilité.

Une approche en grille A est considérée où le problème est réduit algébriquement sur le vent  $U$ . Les opérateurs sont codés de sorte à assurer la consistance, discutée dans le paragraphe 3.1.2.

### **3.4 Article : évaluation de la scalabilité du modèle AROME 2D PG**

Cette version de l'article n'est pas la version finale. La version finale peut être consultée à l'adresse : <https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.3976>



# Krylov solvers in a vertical-slice version of the semi-implicit semi-Lagrangian AROME model

Th. Burgot\*, L. Auger, and P. Bénard

*CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France*

\*Correspondence to: Thomas Burgot, CNRM/GMAP, 42 Avenue G. Coriolis, F-31057 Toulouse Cedex, France

To circumvent the scalability problem due to global communication involved in spectral transforms, a vertical-slice version of the dynamical core where all calculations are performed in grid-point space has been built for AROME, the Météo-France's operational limited-area model. It is shown in an idealized but nevertheless physically relevant framework that despite this major change, keeping the other main characteristics of the model (constant-coefficient semi-implicit scheme, semi-Lagrangian transport scheme, A-grid, mass based coordinate, etc.) is possible. A Krylov solver is used to solve the implicit problem. Using the solution given by the spectral model as a reference in terms of quality and required accuracy in an operational context, the chosen parameters of the Krylov solver are carefully tuned to maximize its convergence speed. The Krylov solver consists of several applications of a sparse operator, whose stencil is similar to the one of the operator applied during the small time step of split-explicit schemes traditionally used in HEVI (Horizontally-Explicit/Vertically-Implicit) models. HEVI models are considered particularly scalable in the current parallelization paradigm and are used as a reference for scalability in this study. Results show that a fast convergence can be reached, such as only a few applications of the sparse operator are required. This suggests an improvement in scalability compared to the spectral version. Experiments have been led regardless of the future strategy considered for temporal and spatial resolutions: increasing spatial resolution while either maintaining a high Courant number or reducing it.

Copyright © 0000 Royal Meteorological Society

*Key Words:* scalability, Krylov solver, semi-implicit, dynamical core, NWP

*Received ...*

*Citation: ...*

## 1. Introduction

The quality of Numerical Weather Prediction (NWP) models has significantly improved in recent decades (Bauer et al. 2015), partly due to the steady increase in spatial resolution, which itself was made possible by an increase of computing power. Operational limited-area models (LAM) commonly reach resolutions of about one kilometre or less in the horizontal and ten metres in the vertical. In this range

of resolutions, a formulation of forecast models in fully compressible Euler equations is desirable since reduced governing systems (as Hydrostatic Primitive Equations) generate detrimental errors in parts of the domain where non-hydrostatic conditions occur.

In the last few years, supercomputer architectures have become more and more massively-parallel and this is likely to continue in the future. Parallelisation techniques have led to split the three-dimensional (3D) integration domain

into many sub-domains, each owned by one computer node. It is generally chosen to split the domain horizontally, thereby defining a tessellation of the geographical domain, whilst no split occurs along the vertical. By doing so, data communications are required at each time step to exchange sub-domains information. Numerical methods that make the best use of parallelism are those that do communicate task data as little as possible with other tasks.

In order to cope with parallelization problems, some Meteorological Centres have chosen to use HEVI (Horizontally-Explicit/Vertically-Implicit) algorithms, whose implicit treatment only involves source terms responsible of the propagation of the fastest waves along the vertical, leading to the inversion of a one-dimensional (1D) problem along the vertical (see for example Satoh (2002) and Klemp et al. (2007)). Examples of operational NWP models using HEVI time stepping are ICON of DWD and MeteoSwiss (e.g. Zängl et al. (2014)), FV3 of NCEP (Putman and Lin (2007), Harris and Lin (2013)), or NICAM of JMA (Satoh et al. (2008)), etc. Other source terms –which include all terms involving discrete horizontal operators– are time-discretized in an explicit manner. The CFL constraint on the fast-wave part of horizontal motions thus imposes relatively small time steps. In the framework of the current massively-parallelized paradigm, these methods seem particularly well-adapted, and will be regarded in this study as a reference in terms of scalability.

Other Centres (e.g. UKMO (Davies et al. (2005)), Météo-France and ALADIN & HIRLAM consortia (Seity et al. (2011)), ECMWF (Wedi et al. (2015)), Environment Canada (Qaddouri and Lee (2010)), etc) made the choice to use a semi-implicit (SI) treatment of all terms potentially responsible for the fastest waves propagation. This approach is often associated with a semi-Lagrangian (SL) algorithm for the transport part of the evolution. This combination (SI-SL) has made possible to considerably alleviate stability constraints, and thus to use large time steps, at the price of solving a 3D ‘Helmholtz’ problem in space generated by the SI scheme. In the SI method, source terms are split into an appropriately chosen linear part and a so-called non-linear residual part. The system associated to linear terms allows a realistic propagation of fast-waves, and this part of the evolution is treated implicitly in time, while residuals are treated explicitly. When the coefficients of source terms in the linear system are horizontally-homogeneous and independent of time, the scheme is termed ‘constant-coefficient SI scheme’ (Bénard 2003).

In a ‘constant-coefficient’ SI scheme horizontal and vertical parts of the 3D-space problem to be solved are separated, making attractive a spectral-transform approach for the horizontal discretization, since direct methods may be used to solve the horizontal part of the problem. In the spectral-transform approach, (forth– and back–) transforms are carried out at each time step from the grid-point space, where non-linear (residual) terms are computed, to the spectral space, where several processes are performed: application of numerical diffusion, evaluation of differential operators, and solution of the –separated– 3D implicit problem by a direct method. This is the strategy adopted so far for the AROME model, Météo-France’s non-hydrostatic (NH) limited-area model (Seity et al. 2011).

However, a disadvantage of using time steps corresponding to large wave-CFL values is that the propagation of *all* fast waves –including gravity waves– is distorted. This implies namely that some distortion occurs in the response to orographic forcing. It may be argued that these wave-propagation errors may eventually exceed those generated by using a constant-coefficient approach instead of an approach retaining actual horizontal variations in linear terms, then leading to a non-separable 3D problem to be solved (e.g. Simmons et al. (1978)). Some studies even suggest that a fully-accurate representation of orographic effects can only be achieved for wind-CFL numbers below unity in SI-SL models (Pinty et al. 1995). Furthermore, spectral transforms require global communication of data among all computation nodes. Although spectral SI-SL schemes have shown an unchallenged efficiency and robustness, it is not sure whether this advantage will be maintained in the future, since costs of global communication could become prohibitive in operational contexts, finally raising an unsurmountable scalability problem. In current SI-SL applications, spectral transforms are therefore called as rarely as possible, by using very large time steps. These are carefully adjusted at each change in horizontal resolution in such a way to maintain the CFL number of fastest horizontal waves significantly greater than unity (about 14 currently for operational AROME).

Abandoning the spectral-transform SI-SL method for a SI-SL grid-point method with an iterative solver is the alternative strategy explored here. This move is likely to bypass the scalability problem and would allow to reconsider the paradigm of very large time steps without having to pay a high computational price for frequent spectral transforms.

However, the spectral-transform method did also provide a bunch of attractive features which must be kept in mind when building a grid-point alternative: (i) very high accuracy in the calculation of derivatives on a non-staggered horizontal grid (A-grid); (ii) trivial swap from differential vector fields (divergence, vorticity) to physical vector components (zonal and meridional velocities); (iii) trivial access to implicit numerical diffusion; and finally (iv) direct and highly-accurate solver for the implicit problem.

Some aspects related to these features in view of non-spectral models are now examined. Implications of using A-grids in numerical models based on grid-point algorithm have been extensively discussed in the literature where it has been noted that this choice leads to a spurious stationarity for shortest resolved gravity waves (see for example Mesinger (1979)). However, as mentioned in Bénard and Glinton (2019), this statement only applies to space-discretizations with low-order accuracy. For higher accuracy orders (e.g. 4 to 8) the problem is in practice relegated to the end of the resolved spectrum which, in any case, must be filtered out to avoid quadratic aliasing. Some current grid-point-based models appear to use an unstaggered A-grid without problems (e.g. Kühnlein et al. (2019)). Point (ii) may be disregarded in the first instance, when dealing only with idealised flows in the vertical plane: in this case, no swap between wind divergence and wind velocity is used (the model may be formulated in terms of wind components). In such an idealised model, the implementation of an implicit numerical diffusion is also not a critical issue; moreover, most of the cases presented

here pertain to linear regimes and do not require diffusion. Finally, the main advantage lost here with a grid-point SI model, is the existence of an efficient and accurate direct solver. The main focus of the study is therefore the convergence of the iterative solver, since this drives almost entirely the overall performance of the SI-SL grid-point method.

Because of the large dimension of the implicit space problem, typically involving  $10^9$  degrees of freedom, but thanks to its sparsity, several alternatives to direct solvers are extensively discussed in the literature (e.g. Müller and Scheichl (2014)). For instance, multigrid methods have been explored and may be slightly faster than FFT-based algorithm as shown in Hess and Joppich (1997) with an Helmholtz equation in spherical coordinates from a local NWP model. Krylov methods are also used for example in Smolarkiewicz and Margolin (1994) where the Generalized Conjugate Residual method (GCR) is used successfully in a density-stratified potential flow past a steep three-dimensional isolated hill on a plan or in Thomas et al. (1997) where a preconditioned Generalized Minimal Residual Method (GMRES) solver is successfully used in a global SI-SL model, very close to AROME model. More recently, Kühnlein et al. (2019) and Maynard et al. (2020) use preconditioned Krylov solvers in the context of, respectively, a conserving finite-volume model and a finite-element discretization. All these Krylov iterative methods seem particularly well suited since they mainly require local operations (as in HEVI schemes), the number of which depends on the accuracy required and the speed of convergence to achieve it. Some global calculations are still needed because of scalar products involved in the method. Some studies as e.g. Zheng and Marguinaud (2018) suggest that the communication cost of scalar products can be more penalizing than the local communications if many iterations are required in the solver and more than  $10^5$  computing nodes are used. However, current models are typically operated on  $10^2$  nodes. It can therefore be assumed that neglecting the cost of scalar products will be relevant for the next few years, though it might become an issue in some future. Since the overall communication cost is considerably alleviated if convergence is achieved with few iterations, the study of the actual speed of convergence in physically relevant contexts –as examined here– is a crucial aspect.

When only the most general characteristics of the problem are known (symmetry, definiteness etc), theoretical convergence bounds are then too loose to provide a usable estimate of the speed of convergence, but the knowledge of these characteristics may favour a particular Krylov method among the wide variety of those available (GMRES, GCR, etc). For instance, Steppeler et al. (2003) highlights the importance of the appropriate choice of the method for a non-symmetric linear problem.

Then, it becomes possible to conduct informative scalability tests or to test different preconditioners over a large variety of solvers and preconditioners like in Müller and Scheichl (2014). However, these results are not necessarily generalisable because most of the problems inverted in meteorology depend strongly on the model itself: Poisson problem for most filtered systems (e.g. for the anelastic approximation), Helmholtz problem for

implicit discrete systems, with variants depending on the assumptions made to construct the linear operator (presence of the orography or not,...), or the type of grid used, etc. Thus, convergence estimates obtained from a given framework are not necessarily transposable to another one, and require specific studies like in Skamarock et al. (1997) or in Qaddouri and Lee (2010). However, no operational NH model except AROME, simultaneously combines a SI-SL formulation using an A-grid with a mass coordinate, thus making necessary a specific study.

Once the stopping criterion has been defined, the critical threshold at which the solver loop is exited remains to be defined. This tuning is essential because an advantage of using a Krylov method is that it converges towards an acceptable solution (in a sense defined below) in only a few iterations in most cases. This threshold is usually defined in an academic framework, like this one, with test cases sufficiently relevant to be generically transposed to operational models. Using physically based stopping criteria like in Skamarock et al. (1997) and Smolarkiewicz and Margolin (1997) is an appropriate alternative way to do it. Two approaches can be considered to define it at a given time of a simulation, by measuring the deviation of the solution of the partially-converged solution computed by the solver from:

- the converged solution of the problem (with an error accuracy close to the finite precision of floating-point arithmetic);
- a reference solution.

On the last point, except for a few linear flows, analytical solutions are generally not known. Consequently, the reference solution chosen in this study is the one provided by the spectral version of the AROME model, a model that has now been validated for several years and will be regarded in this study as a reference in terms of quality. Since most of the studies are carried out with only one model, this second approach is particularly original and consolidates the approach.

As mentioned in Liesen and Tichý (2004) the convergence rate does not depend only on the eigenvalue spectrum of the problem, but also on : the initialization of the iterative process; and the spectral properties of the right-hand side (RHS): in our case the meteorological situation. All these points must be carefully studied to obtain the best convergence properties at a given quality.

The last two points strongly suggest to test convergence in a real model: this makes possible to initialize the solver from the solution at the previous time step, instead of starting from arbitrarily-prescribed meteorological fields.

These points lead to build a full grid-point version of AROME using a Krylov solver. The aim of this study is to assess convergence properties of this solver in a vertical-slice (2D) idealized but relevant meteorological framework by running an orographic test case and a convective one.

The orographic test flow considered in this study is relevant since it evaluates the ability of a model to simulate non-hydrostatic gravity waves excited by orography. Those waves are important atmospheric phenomena, they propagate at fast speed and consequently challenge the model dynamics and the solver.



The rising bubble test is another commonly used configuration. Here with a background horizontal wind, it involves more particularly other components of the model like the advection scheme or the non-linear thermal residual. Moreover, the shift from 3D to 2D does not modify the orders of magnitude of the condition numbers of the considered problems (as shown in the following). It is therefore expected that the convergence rate of the method will not differ much in 3D conditions. Consequently, it would not be surprising if the convergence results obtained in these particular cases apply to many other flow regimes. Simulations in 3D should be required to confirm it. In practice the 2D grid-point version of the model is derived from the pre-existing 2D (vertical plane) spectral version of AROME. This spectral version, giving the same response as would the operational model, is used henceforth to produce the aforementioned reference solutions.

Increasing the horizontal resolution might continue to improve forecast quality, provided adjusting the numerical techniques to the future computing architectures and by rightly handling forecast uncertainty (see e.g. Wedi (2014)). It is thus uncertain whether increasing resolution at the same rate as what has been done so far will be relevant, and other options could then be explored to improve forecasts, such as reducing the time step. Consequently, experiments in this study will then be carried out in two complementary situations: one where the time step is adjusted at each change of resolution to keep the ratio time step over horizontal resolution constant, and one where the time step decreases at a given horizontal resolution, to better represent orographic effects. All these experiments will enable efficiency estimates to be deduced in comparison to a HEVI model, regarded as a reference in terms of scalability.

In section 2, the equations and temporal discretization of the AROME model are written in a synthetic two-dimensional (vertical plane) formalism, reflecting the models used here to perform some idealized test case experiments extensively studied in the literature. Namely, the current spectral dynamic kernel (SP hereafter) is described there and the study of its properties allows to build its grid-point (GP hereafter) counterpart benefiting from the best eigenvalue properties of the implicit problem to be inverted. Test cases and numerical configurations used are presented in section 3. In section 4, a quality evaluation is led to validate the grid-point converged and spectral solutions. Then, the critical threshold at which the solver loop is exited is tuned so that the errors made compared to the spectral version are sufficiently small in comparison to those committed in an operational context. In section 5, a synthetic metric is defined to derive efficiency estimates relative to a HEVI model in the two complementary situations previously mentioned. Finally, a conclusion and some perspectives are presented in section 6.

## 2. Model formulation

### 2.1. Governing equations

As already mentioned, the 2D vertical plane model used in this study is derived from the operational 3D model AROME (Bénard et al. 2010): fully compressible Euler

Equations are formulated with a hybrid mass-based terrain-following coordinate  $\eta$  (Laprise 1992), and the transport scheme is based on the semi-Lagrangian technique.

The time evolution of the state-vector  $X$  gathering all prognostic variables may be written symbolically as:

$$\frac{dX}{dt} = \mathcal{M}(X) + \mathcal{F}(X), \quad (1)$$

where  $\mathcal{F}$  is the physical sources term,  $\mathcal{M}$  is the complete dynamical core model (see Appendix A and Appendix B). In this formalism,  $X$  is therefore a vector of differentiable functions in space and time. For further convenience, the state-vector  $X$  is split into two parts  $X = [U, \Psi]^T$ , with  $\Psi = [d, T, \hat{q}, \pi_s]^T$  and where  $U$  is the horizontal wind velocity,  $T$  the temperature,  $\pi_s$  the hydrostatic surface pressure,  $\hat{q}$  is the reduced non-hydrostatic pressure

$$\hat{q} = \ln(p/\pi), \quad (2)$$

where  $\pi$  is the hydrostatic pressure and  $p$  is the true pressure. The reduced vertical divergence  $d$ , in a dry atmosphere is defined as:

$$d = \frac{p}{mR_d T} \left( -g \frac{\partial w}{\partial \eta} + \frac{\partial \phi}{\partial x} \frac{\partial U}{\partial \eta} \right) \quad (3)$$

(see Appendix A and Appendix B for more details and for notations used). Unlike in the 3D model, the zonal wind  $U$  is used here instead of horizontal divergence of the wind  $D$ . Consequently some second horizontal derivatives in the original 3D version of  $\mathcal{M}$  are replaced here by first derivatives.

The linear system  $\mathcal{L}$  used for the implicit scheme is derived from the above complete system, through some linearization around a stationary reference state  $X^*$ , with no orography, resting, hydrostatically-balanced, isothermal and dry. The resulting linear system involves a reference hydrostatic surface-pressure  $\pi_s^*$  and two reference temperatures  $T^*$  and  $T_e^*$  (Bénard, 2004). The expression of  $\mathcal{L}$  is similar to its 3D counterpart (cf Bénard, et al. 2010):

$$\begin{aligned} \partial_t U &= -R_d \mathcal{G}^* \partial_x T + R_d T^* \mathcal{G}^* \partial_x \hat{q} - R_d T^* \partial_x \hat{q} \\ &\quad - \frac{R_d T^*}{\pi_s^*} \partial_x \pi_s, \\ \partial_t d &= -\frac{g^2}{R_d T_e^*} \mathcal{L}^* \hat{q}, \\ \partial_t T &= -\frac{R_d T^*}{C_{vd}} (\partial_x U + d), \\ \partial_t \hat{q} &= S^* \partial_x U - \frac{C_{pd}}{C_{vd}} (\partial_x U + d), \\ \partial_t \pi_s &= -\pi_s^* \mathcal{N}^* \partial_x U, \end{aligned} \quad (4)$$

where  $\mathcal{G}^*$ ,  $\mathcal{L}^*$ ,  $S^*$  and  $\mathcal{N}^*$  are vertical operators defined in Appendix C. Since all the coefficients in (4) are prescribed and kept constant in time and space, the implicit scheme is a ‘constant-coefficient’ one, in the sense discussed above.

## 2.2. Time discretization

As detailed in [Bénaud et al. \(2010\)](#), Euler equations are integrated by an Iterative Centered Implicit (ICI) scheme

$$\frac{X^{+(n)} - X^0}{\Delta t} = \frac{1}{2} \left[ \mathcal{R} \left( X^{+(n-1)} \right) + \mathcal{R} \left( X^0 \right) \right] + \frac{1}{2} \mathcal{L} \left( X^{+(n)} + X^0 \right) + \mathcal{F} \left( X^0 \right), \quad (5)$$

where  $\mathcal{R} = \mathcal{M} - \mathcal{L}$  is the non-linear residual,  $\Delta t$  is the time step and the superscript  $+(n)$  denotes the future time level obtained after the  $n$ -th iteration of the ICI scheme. Currently, in the operational context only 2 iterations are required. This configuration is sometimes called a ‘predictor-corrector’ scheme. At each iteration of the scheme, the implicit problem to be solved may be written as

$$\left[ \mathcal{I} - \frac{\Delta t}{2} \mathcal{L} \right] X^{+(n)} = X^\bullet, \quad (6)$$

where:

$$X^\bullet = X^0 + \frac{\Delta t}{2} \left[ \mathcal{R} \left( X^{+(n-1)} \right) + \mathcal{R} \left( X^0 \right) \right] + \frac{\Delta t}{2} \mathcal{L} \left( X^0 \right) + \Delta t \mathcal{F} \left( X^0 \right), \quad (7)$$

or equivalently, using a block matrix formalism for (4)

$$\left[ \begin{array}{c|c} \mathcal{I} & -\frac{\Delta t}{2} \mathcal{L}_A \\ \hline -\frac{\Delta t}{2} \mathcal{L}_B & \mathcal{I} - \frac{\Delta t}{2} \mathcal{L}_C \end{array} \right] \left[ \begin{array}{c} U^{+(n)} \\ \Psi^{+(n)} \end{array} \right] = \left[ \begin{array}{c} U^\bullet \\ \Psi^\bullet \end{array} \right]. \quad (8)$$

In (8),  $\mathcal{I}$  denotes the identity operator in the relevant space,  $\mathcal{L}_A = \mathcal{V}_A \circ \partial_x$ ,  $\mathcal{L}_B = \mathcal{V}_B \circ \partial_x$ , and

$$\mathcal{L}_C = \left[ \begin{array}{c|c|c|c} 0 & 0 & -\frac{g^2}{R_d T_s^*} \mathcal{L}^* & 0 \\ \hline -\frac{R_d T^*}{C_{vd}} & 0 & 0 & 0 \\ \hline -\frac{C_{pd}}{C_{vd}} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \quad (9)$$

with

$$\mathcal{V}_A = \left[ \begin{array}{c|c|c|c} 0 & -R_d \mathcal{G}^* & R_d T^* \mathcal{G}^* - R_d T^* & -\frac{R_d T^*}{\pi_s^*} \end{array} \right], \quad (10)$$

$$\mathcal{V}_B = \left[ \begin{array}{c} 0 \\ \hline -\frac{R_d T^*}{C_{vd}} \\ \hline \mathcal{S}^* - \frac{C_{pd}}{C_{vd}} \\ \hline -\pi_s^* \mathcal{N}^* \end{array} \right] \quad (11)$$

Operators  $\mathcal{L}_A$  and  $\mathcal{L}_B$  contain both vertical ( $\mathcal{V}_A$  and  $\mathcal{V}_B$ ) and horizontal ( $\partial_x$ ) operators and  $\mathcal{L}_C$  is a contribution of the non-hydrostatic vertical part, reduced to zero in the hydrostatic version. The system (8) may then be reduced

to a single variable equation by an appropriate algebraic elimination in favour of  $U$

$$\left[ \mathcal{I} - \frac{\Delta t^2}{4} \mathcal{B} \circ \partial_x^2 \right] U^{+(n)} = U^{\bullet\bullet} \quad (12)$$

where  $\mathcal{B}$ , referred to as the non-symmetric vertical operator, is defined by

$$\mathcal{B} = \mathcal{V}_A \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)^{-1} \mathcal{V}_B,$$

and

$$U^{\bullet\bullet} = U^\bullet - \frac{\Delta t}{2} \mathcal{L}_A \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)^{-1} \Psi^\bullet.$$

## 2.3. Space discretization

The time-discrete model described hitherto involved space-continuous variables and operators. A space-discretized version is then built, which detailed presentation is out of the scope of this paper. The model used here being indeed a 2D clone of the operational 3D version of AROME, the discretization follows [Bénaud et al. \(2010\)](#) except for the minor changes implied by the use of  $U$  instead of  $D$ . The only important feature to be highlighted here is the vertical discretization that ensures that the implicit problem created by (5) may still be reduced to a single variable problem as in (12). This requires some appropriate properties of vertical operators, always valid in the continuous context, to be also satisfied for vertically-discrete operators. In the vertical discretization of AROME, this constraint is satisfied (see [Bubnová et al. \(1995\)](#)), thereby permitting the above algebraic elimination procedure to hold in the space-discrete case.

For notation convenience, all symbols used so far to describe space-continuous variables and operators are kept the same to describe space-discretized variables and operators henceforth. An equation formally similar to (12) is therefore arrived at, but now involving space-discrete variables and operators. The aim of this paper is therefore to examine and compare various ways to solve this pivotal equation. Regarding this equation, the difference between the spectral and grid-point versions of the 2D model only lays in the concrete form of the discrete operator  $\partial_x^2$ .

## 2.4. Decomposition into the eigenspace of $\mathcal{B}$

The total number of levels in the vertical discretization is denoted  $N_{\text{lev}}$ . In a 2D framework, the space-discrete version of (12) implies a large-size 2D implicit problem that is strongly anisotropic. Furthermore, the current paradigm of parallelization allows to treat the vertical part within the same compute node, thus raising no communication problems.

Consequently, a specific method is usually used to invert directly the vertical part, while an iterative solver (possibly preconditioned) is usually used on the horizontal direction ([Skamarock et al. \(1997\)](#), [Thomas et al. \(1997\)](#)).

In the special case of the constant-coefficient scheme, a natural alternative approach to (12) is thus to exploit the separability of this equation along horizontal and vertical

directions, and to decompose this problem into a set of  $N_{lev}$  1D problems by projecting the space-discrete version of (12) in the eigenspace of the vertical operator  $\mathcal{B}$ . This directional separation in the implicit problem is a *raison d'être* of the spectral technique and is therefore always used in the spectral version of the model. For grid-point versions, both approaches could equally be considered, but the potential advantage of exploiting the separability is indeed used in this paper (see related discussion in section 2.6) because the projection here plays in some way the role of a vertically direct preconditionner.

For a given eigenmode  $\mathcal{B}_l$  of the operator  $\mathcal{B}$ , the 1D implicit equation to be solved is

$$\left[1 - \frac{\Delta t^2}{4} b_l \partial_x^2\right] \tilde{U}_l^{+(n)} = \tilde{U}_l^{\bullet\bullet}, \quad (13)$$

where  $b_l$  is the  $l$ -th eigenvalue of  $\mathcal{B}$  sorted in descending order, and  $(\tilde{U}_l^+, \tilde{U}_l^{\bullet\bullet})$  are the projected component of  $(U^+, U^{\bullet\bullet})$ , on the corresponding eigenmode  $\mathcal{B}_l$ .

$\mathcal{B}$  can be approximately seen as the product of an inverse operator on the vertical by the square of the sound speed. For the most external mode  $b_0$ , sometimes called the ‘barotropic’ mode, the propagation is essentially horizontal (and the part corresponding to the inverse operator tends towards identity) while for the internal modes, or ‘baroclinic’ modes, the propagation is essentially vertical (the part corresponding to the inverse operator tends towards zero).

The spectrum of  $\mathcal{B}$  is plotted in Fig. 1 for various time steps  $\Delta t$  in the non-hydrostatic (NH) and hydrostatic (H) cases, for an horizontal grid-mesh of 100 m and the vertical grid used in all experiments below. Eigenvalues are positive regardless of  $\Delta t$ , and rapidly decrease toward zero with the vertical mode index. This property will be strongly exploited in the following. It can be noted that the eigenvalues of  $\mathcal{B}$  are strongly dependent on the time step for the most internal modes but not for the most external ones where vertical propagation is weak. The representation of vertical wave propagation is improved as the time step decreases. On the contrary the eigen-spectrum of  $\mathcal{B}$  of the NH model tends towards which of the H one, as the time step increases. Thus, the improvements of the vertical wave propagation (especially gravity waves) sought during the implementation of an NH model in favour of an H model, are fully satisfied when the time step is small, as previously mentioned.

For a given resolution  $(\Delta x, \Delta t)$ , an horizontal wave-CFL number associated to the  $l$ -th mode may be introduced as

$$c_l^* = b_l^{1/2} \frac{\Delta t}{\Delta x}. \quad (14)$$

The maximum eigenvalue  $b_0$  is the square of the sound speed and the corresponding value  $c_0^*$  is the CFL number for horizontally-propagating sound waves. Since this is the most constraining value in the whole set of horizontal CFL numbers  $c_l^*$ ,  $c_0^*$  will be used as the reference wave-CFL number in each of the following experiments. For comparison to other models, the wind-CFL number  $c_{adv}^*$  is also computed as

$$c_{adv}^* = \frac{U_{max} \Delta t}{\Delta x}, \quad (15)$$

where  $U_{max}$  is the maximum horizontal wind over the domain.

## 2.5. Spectral version

In the 2D model, a Fourier representation of space-discretized variables along  $x$  is possible by making all fields periodic through an appropriate extension technique (Haugen and Machenhauer (1993)). The extended vectors  $U$  and  $U^{\bullet\bullet}$  may then be projected into the Fourier space using an efficient FFT algorithm, the decomposition  $\widehat{U}$  of a field  $U$  being termed the ‘spectral transform’ of  $U$ . Since Fourier components  $\exp(ikx)$  are the eigenmodes of the horizontal periodic Laplacian operator  $\partial_x^2$ , the implicit problem (12) projected in Fourier space is reduced to scalar inversions

$$\left[1 + \frac{1}{4} c_l^{*2} k'^2\right] \widehat{U}^{+(n)} = \widehat{U}^{\bullet\bullet} \quad (16)$$

where  $k' = k\Delta x \in [-\pi, \pi]$  is the non-dimensional zonal wavenumber.

## 2.6. Grid-point version

### Choice of strategy

Once the horizontal operator  $\partial_x^2$  is discretized on the A-grid, one of the following strategies has to be chosen:

- i. either solving  $N_{lev}$  1D independent symmetric problems (13);
- ii. or solving the large 2D non-symmetric problem (12).

As mentioned previously, the convergence of a Krylov solver depends on the eigenvalues spread of the linear operator matrix, measured by its condition number: the higher it is, the slower the convergence is. When a sufficiently high order is used to discretize the one-dimensional Laplacian operator so that its response is close to the spectral one, considering that the vertical operator  $\mathcal{B}$  is almost symmetric and considering a large horizontal domain of size  $L_x$  ( $L_x \gg \Delta x$ ), the condition number can be easily computed:

- i. for the  $N_{lev}$  independant 1D symmetric problems (13):

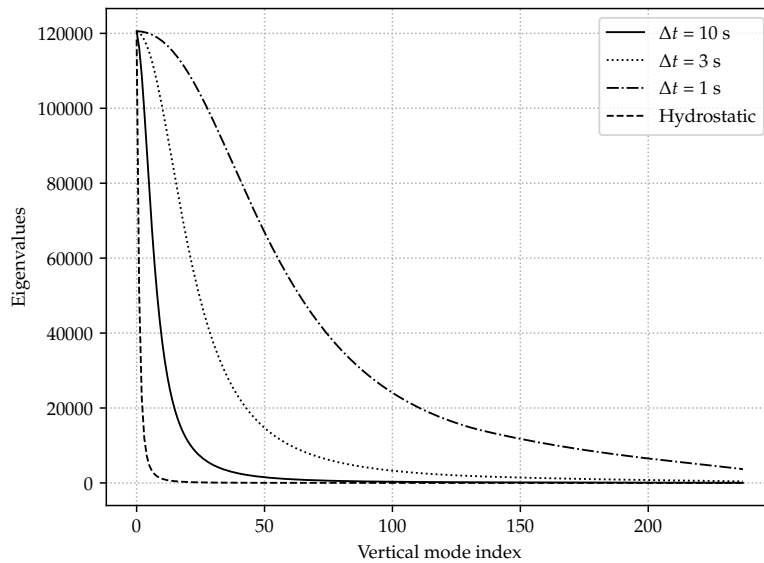
$$C_l \simeq \frac{1 + b_l \Delta t^2 \pi^2 / (4\Delta x^2)}{1 + b_l \Delta t^2 \pi^2 / L_x^2} \underset{L_x \gg \Delta x}{\simeq} 1 + \frac{\pi^2}{4} c_l^{*2}; \quad (17)$$

- ii. for the large 2D non-symmetric problem (12):

$$C \simeq \frac{1 + b_0 \Delta t^2 \pi^2 / (4\Delta x^2)}{1 + b_{N_{lev}-1} \Delta t^2 \pi^2 / L_x^2} \underset{L_x \gg \Delta x}{\simeq} 1 + \frac{\pi^2}{4} c_0^{*2}. \quad (18)$$

(that is to say the full 2D system is as ill-conditioned as the worst 1D equation)

The fastest vertical mode, although being the most difficult to solve, has a relatively low condition number,  $C_0 \simeq 500$  (corresponding to the current operational configuration), when compared to typical values encountered in recent literature (Ye (2017), Soleymani (2013)) with typical values of ill-conditioned matrices larger than



**Figure 1.** Eigen-spectrum of  $\mathcal{B}$  with different time step under an horizontal resolution of 100 m and the vertical grid configuration described section 3.3. The spectrum of the hydrostatic version which is independent of the time step is also plotted.

$10^{10}$ . The remaining  $N_{i,i=1\dots N_{\text{lev}}}$  problems are even better-conditioned and a very fast convergence using any iterative grid-point solver is expected for the majority of them, confirming the preconditioning effect of this vertical projection. In 3D, when a large domain and similar horizontal resolutions are considered in both directions, equations (17) and (18) become respectively  $C_l \simeq 1 + \frac{\pi^2}{2} c_l^{*2}$  and  $C \simeq 1 + \frac{\pi^2}{2} c_0^{*2}$ . These numbers are very close to those obtained previously in only 2D.

To exploit as best as possible the current paradigm of parallelization, one has a strong interest to treat the  $N_{\text{lev}}$  vertical modes sequentially instead of treating them in parallel and waiting for the barotropic mode to finish, as it would be the case by solving the large non-symmetric problem (12). Consequently solving the  $N_{\text{lev}}$  independant symmetric problems will be the chosen method. Because of the positivity of the  $b_l$  modes, all these problems are symmetric and positive definite which allows using specific algorithms like the conjugate gradient one. For the following, it is more convenient to write the  $l$ th problem under the form :

$$\mathcal{H}x = y \quad (19)$$

where :  $\mathcal{H} = [1 - \frac{1}{4}c_l^{*2}\delta_x^2]$ ,  $x = \tilde{U}_l^{+(n)}$  and  $y = \tilde{U}_l^{\bullet\bullet}$ .

#### Initialization

An interesting approach to improve the rate of convergence is to initialize the solver by an estimate of the atmospheric state  $x_0$  as close as possible to the sought solution  $y$ . The best way we found is by taking the field at the previous time step or at the previous iteration of the ICI scheme (Bénaud et al. (2010)) (not shown). The solution (of each of the  $N_{\text{lev}}$  problems) is then looked for in the subspaces of (different) size  $N$  :

$$\mathcal{K}_N = \{r_0, \mathcal{H}r_0, \mathcal{H}^2r_0, \dots, \mathcal{H}^N r_0\} \quad (20)$$

where  $r_0 = y - \mathcal{H}x_0$  is the initial residual.

#### Stopping criterion

The linear system being symmetric, thus, according to Proposition 4 in Saad and Schultz (1986) the following inequality holds:

$$\frac{\|r_n\|}{\|r_0\|} \leq \left(1 - \frac{1}{C_l^2}\right)^{n/2} \quad (21)$$

where the residual at iteration  $n$  is  $r_n = y - \mathcal{H}x_n$ . The most obvious and widely used stopping criterion with such methods is to halt the iterative process when the residual value is reduced by a certain amount compared to the norm of the RHS:

$$\text{Stop if } \frac{\|r_n\|}{\|y\|} \leq \varepsilon \quad (22)$$

where  $\varepsilon < 1$  is an a-priori reduction threshold fixed by the user. That inequality is the worst case convergence inequality and in practise the convergence rate can be much faster (Axelsson (1996)). In this study we assume  $\varepsilon$  does not depend on the vertical mode considered. From equation (21) it can be deduced that the number of iterations of the algorithm  $n$  varies as the logarithm of the residual parameter  $\varepsilon$  in (22). Consequently it is relevant to choose what we call ‘the residual reduction parameter’  $\varepsilon^{-1}$  with values of the form  $10^p$  in the numerical experiments to follow.

#### Additional remarks

The Helmholtz problems solved in this study are of a different nature than Poisson problems sometimes encountered in the literature for which a basic estimate of

the condition number  $C'$  is given by:

$$C' \simeq \frac{k_{max}^2}{k_{min}^2} = \frac{L_x^2}{4\Delta x^2} = \frac{N_x^2}{4} \quad (23)$$

For these problems, the condition number increases as the number of horizontal points  $N_x$ , raising concerns about a lack of convergence over larger domain or over finer resolutions. Using a preconditioner on the horizontal appears then necessary and multi-grid methods seem particularly well suited for this purpose. The interest of this method has been highlighted by performance analyses for large-scale parallel computers (see for example Fischer (2015)).

However for the Helmholtz problems considered in this study, using a finer resolution on the horizontal direction is usually balanced by decreasing the time step for stability reasons on the one hand, and using a larger domain does not worsen the condition number on the other hand (as seen in equations (17) and (18) when  $L_x \gg \Delta x$ ). Using an horizontal preconditioner seems less essential for Helmholtz equations than Poisson ones and will be not studied here, but could be considered in a second time to further optimize convergence.

### 3. Presentation of test cases

The following benchmark composed of two cases has been widely used for model validation purposes and thus results can be compared to existing simulations. It has been designed to be a challenge for meso-scale forecasting which intends to model convection and gravity waves accurately involving advection and adjustment terms respectively.

#### 3.1. Warm bubble test case

The warm bubble test case simulates the evolution of a buoyant thermal in a constant potential temperature environment of 300 K. The version used here is described in Carpenter Jr et al. (1990) where a background uniform speed of  $20 \text{ m}\cdot\text{s}^{-1}$  like in Wicker and Skamarock (1998) is added. A warm anomaly of radius 2 km and centred at 2 km height placed at the middle of the horizontal is given at the beginning. During the simulation the warm bubble rises and develops while being advected. The simulation ends after 1000 s.

#### 3.2. Orographic test case

The orographic test case used in Schär et al. (2002) simulates gravity waves in a stratified environment over a mountain profile given by:

$$h(x) = h_{\max} e^{-(x/r_a)^2} \cos^2(\pi x/r_b) \quad (24)$$

where  $h(x)$  is the orography profile,  $h_{\max} = 250 \text{ m}$  is the maximum orography height,  $r_a = 5000 \text{ m}$  is the overall mountain range width and  $r_b = 4000 \text{ m}$  is the length scale of the sub-orographic modulation. The initial condition is an homogeneous wind profile of  $10 \text{ m}\cdot\text{s}^{-1}$  with a ground temperature of 288 K in a constant Brunt-Väisälä frequency of  $0.01 \text{ s}^{-1}$  environment. The simulation ends after 8000 s, once a near stationary state is reached.

#### 3.3. Configuration

In both cases an horizontal domain size of 100 km and a vertical height of 10 km is given. Except when explicitly mentioned, the horizontal resolution  $\Delta x$  is 100 m and vertical levels are chosen such that for a reference vertical temperature profile and surface pressure, the vertical resolution is 100 m up to 10 km. Above, the vertical levels have a constant pressure depth up to the top of the domain that corresponds to a 0 hPa surface.

In the following experiments and for time steps larger or equal to 3 s, an horizontal numerical diffusion is applied at the end of the time step to overcome the ill-represented energy dissipation at the smallest scales. It is performed by a filtering applied to each wavenumber in spectral whereas a fourth order diffusion (such as  $\nabla^4$ ) is applied in grid-point.

In the following, a fourth-order finite difference derivative scheme has been used. Indeed, sensitivity experiments have proven that it is noticeably better than a second order scheme, whereas a sixth order scheme showed very little improvement. In addition, with a view to possibly solving a more general (non-symmetric) problem in the future, e.g. by including orographic terms in the implicit problem, experiments are conducted using the GMRES algorithm.

### 4. Quality evaluation

In this section, the converged solution of the grid-point model is compared to the spectral one and to the literature. It is shown they are very close to each other and thus allowing to define both of them as references. Then, the critical threshold  $\varepsilon$  is adjusted and the resulting solutions are compared to these references and lead to define a relevant threshold for the stopping criterion.

#### 4.1. Comparison with the spectral version and literature

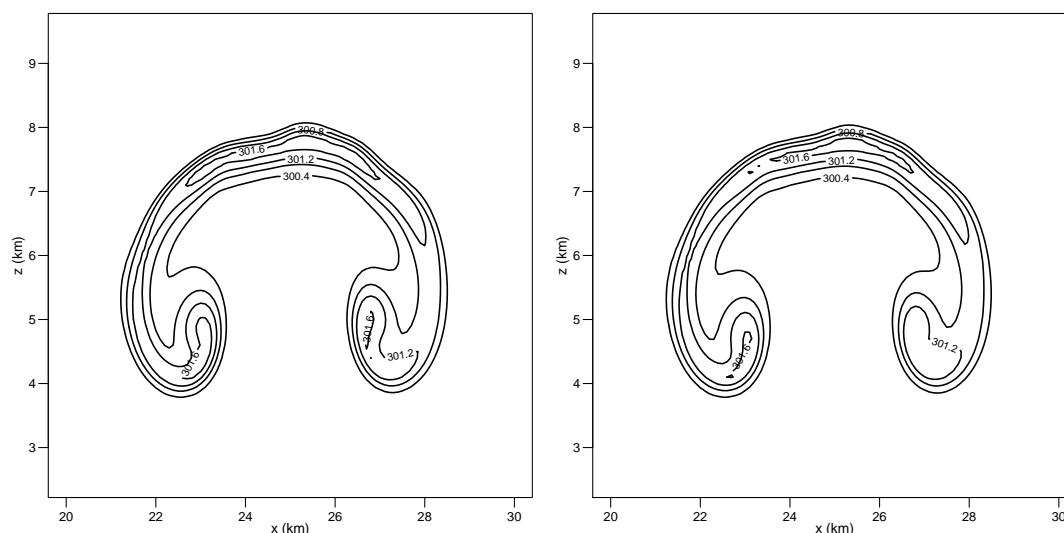
The spectral (SP) and grid-point (GP) model forecasts for the rising bubble case are shown here. The converged GP solution is obtained by setting a non optimal very large value for  $\varepsilon^{-1}$  leading to an error accuracy close to the finite precision of floating-point arithmetic. Results are presented for various CFL number ( $c_0^*$ ) values:

$$c_0^* = 60$$

Here  $\Delta t = 20 \text{ s}$ , and with this very large CFL number it is interesting to notice that the scheme is nevertheless stable, but as expected because of the long time step, the shape of the solution (not shown) is of poor quality with a maximum temperature of 301.64 K in good agreement with the  $\Delta t = 2 \text{ s}$  solution (see below the paragraph corresponding to  $c_0^* = 6$ ) but with a much too low maximum vertical velocity only of  $7.5 \text{ m}\cdot\text{s}^{-1}$ .

$$c_0^* = 9$$

The case with  $\Delta t = 3 \text{ s}$  corresponding to a  $c_0^* = 9$  value comparable to the one used in operational configuration has been reproduced. Figure 2 shows the comparison between SP and GP versions after 1000 s. As explained



**Figure 2.** Rising thermal simulation with  $\Delta x = 100$  m and  $\Delta t = 3$  s after 1000 s : (a) potential temperature for the GP version; (b) potential temperature for the SP version.

above, diffusion is treated differently in the two models, this is the largest source of difference in Figure 2. The shape of the bubble is also not perfectly symmetric to the median axis contrary to what would be expected. This is due to the background horizontal velocity field that makes the rising thermal more complex to simulate. Both GP and SP solutions are very close for the potential temperature and the vertical wind speed (not shown) with a maximum temperature of 301.83 K for the SP simulation and 301.77 K for the GP one and  $13.03 \text{ m}\cdot\text{s}^{-1}$  and  $12.39 \text{ m}\cdot\text{s}^{-1}$  respectively for the vertical wind.

$$c_0^* = 6$$

Figure 3 presents results of the same experiment but with a time step of 2 s. Now the solution is much more symmetric and similar to the numerical solution obtained with other models in literature using the same time step: the maximum vertical velocity in the rising thermal is  $13.3 \text{ m}\cdot\text{s}^{-1}$  to be compared with  $16.5 \text{ m}\cdot\text{s}^{-1}$  in Wicker and Skamarock (1998) and  $14 \text{ m}\cdot\text{s}^{-1}$  in Bryan and Fritsch (2002). The maximum potential temperature of 301.85 K, which is very similar to the values found in Wicker and Skamarock (1998) and Bryan and Fritsch (2002). Here when compared to Figure 2 since no diffusion is used, SP and GP simulations are closer to each other.

From the result presented in this section, it seems that the converged GP solution is very close to the SP one, whatever the CFL number  $c_0^*$  is. Both are also very close to what is obtained in the literature provided  $c_0^* \leq 6$ . This confirms the relevance of keeping the main characteristics of the model and especially the use of the A-grid with a sufficiently high-order space-accurate scheme. Moreover, the converged GP solution or the SP solution can both be defined as references in terms of quality.

#### 4.2. Stopping criterion

Contrary to the previous section, the quality of the solution is now degraded by changing the values of the residual

reduction  $\varepsilon^{-1}$  and deviation from references solutions are measured via the root mean square error (RMSE) on the potential temperature and vertical wind speed fields according to the test case used. These results are plotted in Figure 4.

#### Warm bubble test case

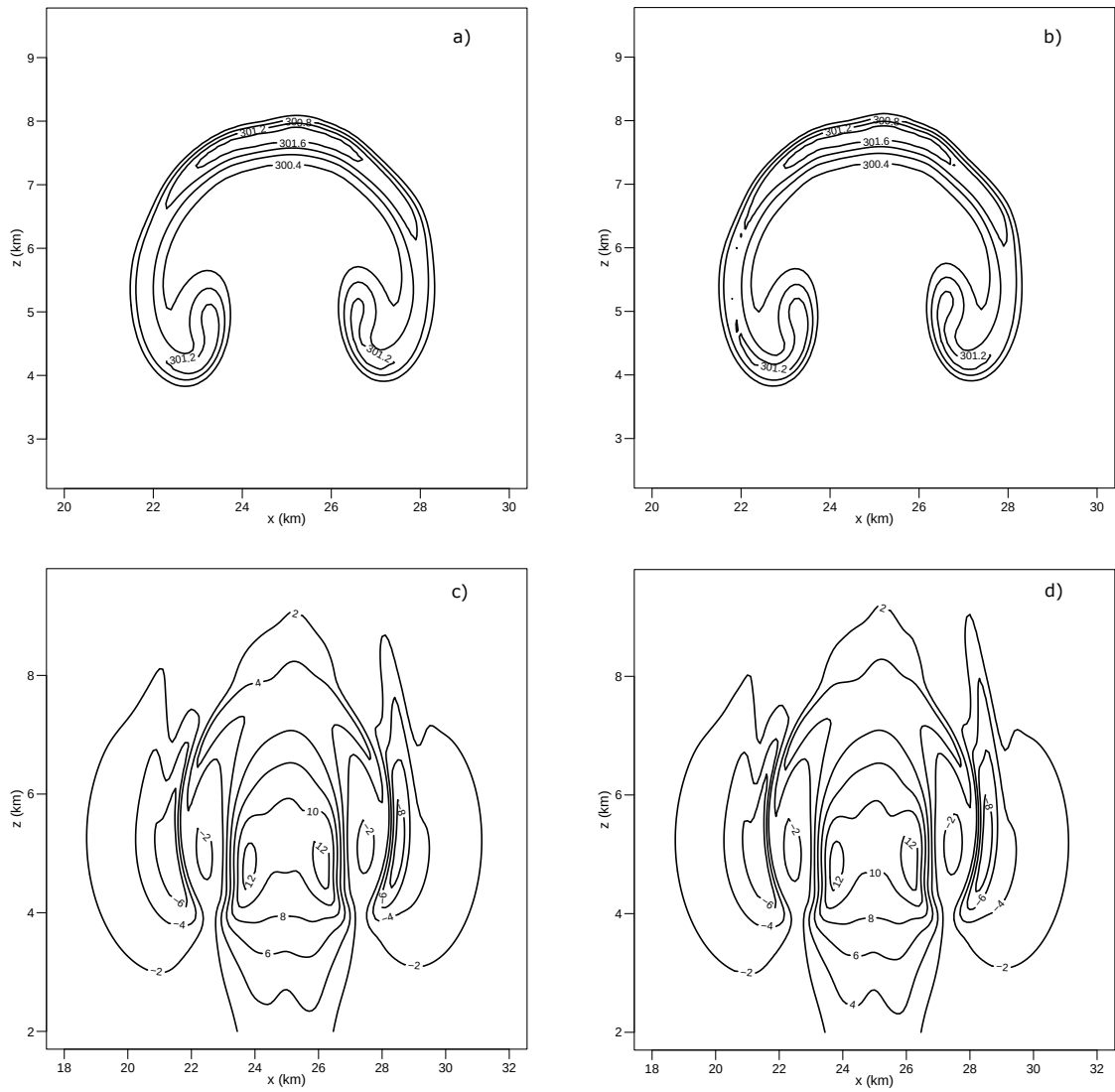
It can be shown first that as expected, the larger the residual reduction is, the closer the solution is from the exact grid-point solution (dotted line on Figure 4). However there is no plateau or inflection of the RMSE's curve that would be helpful to retrieve an optimal residual reduction  $\varepsilon^{-1}$  beyond which further iterations are useless.

The RMSE also decreases when the solution of GP is compared to SP for residual reduction up to  $10^3$ . Then, beyond that value, the error due to the use of an approximated iterative solution is less than the difference between the GP and the SP model so that the extra iterations do not bring the GP solution closer to the SP one. The RMSE between the converged GP solution and the SP one cannot therefore go below 0.036 K for the potential temperature and  $0.18 \text{ m}\cdot\text{s}^{-1}$  for the vertical wind (RMSE's corresponding to  $\varepsilon^{-1} = 10^3$ ).

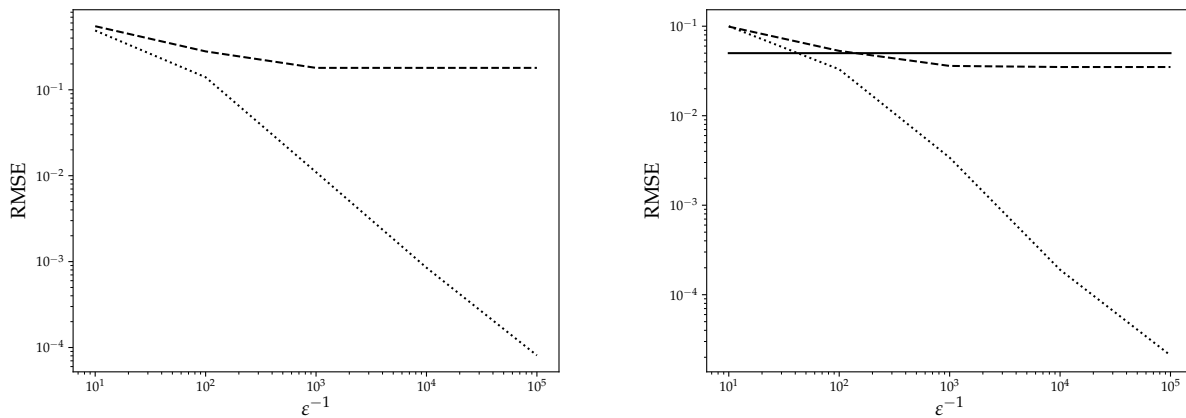
Furthermore, the order of magnitude of RMSEs for the current spectral version of AROME in an operational forecast is about  $1 \text{ m}\cdot\text{s}^{-1}$  for wind and 1 K for temperature even for short time ranges less than one hour (see for example Auger et al. (2015)). Therefore these RMSEs are significantly lower to other sources of model error. Considering a RMSE for temperature less than 0.05 K (lower by a factor about 20 compared to errors made in an operational forecast) is still an acceptable error, a residual reduction of  $\varepsilon^{-1} = 10^2$  is sufficient and will be taken as an acceptable threshold for the warm bubble test case.

#### Orographic test case

In a similar way, the maximum acceptable RMSE threshold for the different experiments was set to  $0.01 \text{ m}\cdot\text{s}^{-1}$  by trial



**Figure 3.** Rising thermal simulation with  $\Delta x = 100$  m and  $\Delta t = 2$  s after 1000 s: (a) potential temperature for the GP version; (b) potential temperature for the SP version; (c) vertical wind speed for the GP version; (d) vertical wind speed for the SP version.



**Figure 4.** RMSE compared to the SP version (dashed line), the exact GP version (dotted line) for the warm bubble test case as a function of the  $\varepsilon^{-1}$  criterion with  $\Delta x = 100$  m and  $\Delta t = 2$  s after 1000 s. Left: vertical velocity field; right: potential temperature field. Acceptable threshold for a typical temperature forecast is shown in continuous line.

and checking that the gravity waves patterns are reproduced as well as for the reference such as the ones to be seen on Figure 5.

To conclude, we managed to carefully define a threshold for the stopping criterion so that the errors committed by a solution that is not fully converged is significantly lower (at least 20 times) than the errors traditionally committed in an operational context. This estimate will be applied to the efficiency study in the following section.

## 5. Efficiency evaluation

In this section, we first define a synthetic criterion and then conduct several experiments in which we evaluate the efficiency of the SI method compared to the HEVI one.

As mentioned in the introduction, an appropriate, scalable alternative to the SI-SL approach is the HEVI time stepping method since it requires a minimum transfer of data (i.e. only a few layers surrounding the sub-domain owned by a computer node). We first show that HEVI models apply operators which can be considered as local as what is done when using a Krylov solver (the cost of dots products is ignored in this preliminary study). Consequently, a similar efficiency between the grid-point version of the SI model and a HEVI model might suggest an improvement in the scalability of the grid-point version compared to the spectral one. Several configurations of our grid-point model will be tested by changing: the critical threshold for the stopping criterion, the spatial resolution keeping the  $\Delta t/\Delta x$  ratio constant and then reducing it. In any cases, the same vertical grid is kept.

### 5.1. A measure of efficiency

It is interesting to evaluate the communication efficiency by looking at the amount of data exchanged in comparison to a typical HEVI scheme which will be briefly described in the following.

Recent HEVI models often combine an implicit solver for rapid waves on the vertical, an accurate explicit treatment of slow terms (e.g. Runge-Kutta integration) and a cheap explicit treatment for the horizontal propagation of fast waves (Gassmann (2013), Durran (2013)). Such a split-explicit scheme uses a short time step also named acoustic time step ( $\Delta\tau$ ) which, in order to ensure numerical stability, must satisfy the CFL constraint:

$$\frac{c_s \Delta\tau}{\Delta x} \leq \alpha \quad (25)$$

where  $c_s$  is the sound velocity,  $\alpha \simeq 1/\sqrt{2}$  is a factor that depends on the order of the temporal and spatial schemes used, the grid used, etc. Now the assumption will be made that, in an operational implementation, most of the inter-node communication is due to the small time step integration for the HEVI scheme, and due to the iterative semi-implicit inversion for AROME. Thus, during the acoustic time step integration the prognostic variables are updated following:

$$X^{t+(n+1)\Delta\tau} = \mathcal{H}' X^{t+n\Delta\tau} \quad (26)$$

where  $X^{t+n\Delta\tau}$  is the state vector at the small time step  $t + n\Delta\tau$ ,  $t$  is the previous long time step and  $\mathcal{H}'$  has a similar computational cost compared to  $\mathcal{H}$  in (19), representing also the terms controlling the horizontal propagation of the fast waves. Then to make a forecast until  $t + \Delta t$ , where  $\Delta t$  has the same value as a SI time step, several applications of iteration (26) are required and after  $N'$  small time steps, the solution is:

$$X^{t+N'\Delta\tau} = \mathcal{H}'^{N'} X^t \quad (27)$$

There are similarities between HEVI and Krylov methods used in a SI forecast model. They both involve applying several time linear operators that are similar in terms of computational cost but also in terms of exchange of data. As a matter of fact the solver might use additional communications for example for the calculation of the scalar product but its data exchange scheme involves only one variable (in the implicit problem of AROME, other variables are explicitly computed) instead of all prognostic variables for a typical HEVI scheme. Since most of the communication cost comes from the data exchanges for horizontal derivatives, the application of  $\mathcal{H}'$  might behave like  $\mathcal{H}$  scalability-wise.

Then an estimate of the efficiency of the problem can be reduced to the comparison between the values  $N$  and  $N'$  or equivalently by the effective time steps of both methods. We define the effective or equivalent explicit time step  $\Delta t_{eq}$  of a SI Krylov method as the time step divided by the overall number of iterations vertically averaged because of the eigenmode decomposition. An explicit model with this time step would perform similarly in terms of scalability. For example AROME which uses a predictor-corrector scheme to enable the use of longer time step, necessitates two stages at each time step leading to solve twice the implicit scheme using the solver. Consequently:

$$\Delta t_{eq} = \frac{\Delta t}{\bar{N}} \quad (28)$$

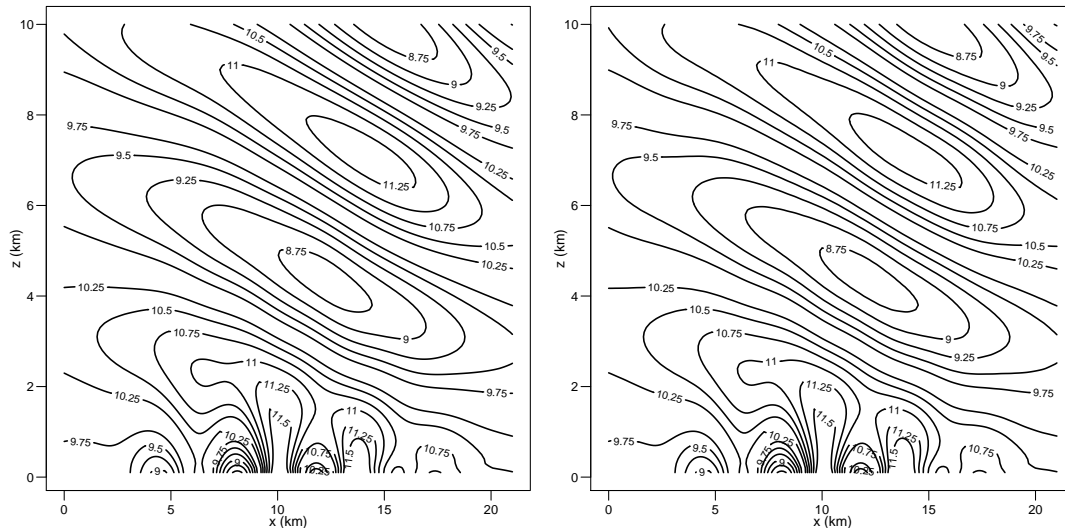
where  $\bar{N}$  is the average number of iterations required through all vertical levels for the predictor-corrector scheme. Then this equivalent time step can be compared to the acoustic time step of a HEVI model which can be guessed using the CFL constraint : for example in the context of the test cases presented here where  $\Delta x = 100$  m, considering a sound velocity  $c_s$  of  $350 \text{ m}\cdot\text{s}^{-1}$ , the reference small time step is then equal to  $\Delta\tau \simeq 0.2$  s. That number agrees with published results, for instance, in Wicker and Skamarock (1998) a warm bubble test case was used with  $\Delta x = 125$  m. Results shown were obtained with a second-order Runge-Kutta time integrator using a large time step of  $\Delta t = 2$  s and a number of steps  $N'$  equals to 12 leading to  $\Delta\tau \simeq 0.17$  s.

The efficiency comparison factor can be introduced as the ratio between the equivalent explicit time step and the HEVI acoustic time step:

$$f = \frac{\Delta t_{eq}}{\Delta\tau} \quad (29)$$

When  $f \gg 1$ , the SI method is more efficient than the HEVI one and vice versa. When  $f \simeq 1$ , it indicates a roughly similar efficiency between the grid-point version of the SI





**Figure 5.** Mountain waves simulation with  $\Delta x = 100$  m and  $\Delta t = 3$  s, horizontal wind after 8000 s focused on the middle part of the domain. Left: reference for the GP version; right: with a residual reduction  $\varepsilon^{-1}$  of 300.

model and a HEVI model and suggests an improvement in the scalability of the grid-point version compared to the spectral one.

### 5.2. Stopping criterion evaluation

We intend in this part to assess how explicit equivalent time step and  $f$  evolve as a function of the residual reduction with the two cases considered.

As mentioned in Table 1 when the residual reduction  $\varepsilon^{-1}$  is greater than  $10^3$ , the number of iterations of the solver increases such that the HEVI method ends up performing better than the SI-GP method ( $f < 1$ ). When  $\varepsilon^{-1} \simeq 10^2$ , as advocated previously, the number of iterations obtained is such that the SI-GP method has similar efficiency than the HEVI method ( $f \simeq 1.45$ ) for a quality approximately equivalent according to observations made in section 4.1. This result also confirms the stopping criterion adjustment influences the efficiency.

Similar results are obtained for the orographic wave test case as mentioned in Table 2. A simulation has been ran over a large horizontal domain usually encountered in operational models of 1000 km, with  $\Delta t = 4$  s and  $\Delta x = 200$  m with the same configuration otherwise. Results show that the convergence rate of the solver does not depend significantly on the size of the domain, according to the condition numbers (17) previously computed.

### 5.3. Future resolution keeping the ratio $\Delta t/\Delta x$ constant

In an operational context up to now at least, the horizontal resolution is improved trying to preserve the ratio  $\Delta t/\Delta x$  as much as possible for stability and accuracy reasons. Table 3 shows how the maximal and averaged number of iterations varies with a constant  $\Delta t/\Delta x$  ratio similar to the one used with the operational version. Each experiment starts with a different bubble size adapted to its resolution and consequently forecast outcome is different from one experiment to another. The maximal number of iterations

$\varepsilon^{-1}$	$N_0$	$\bar{N}$	$\Delta t_{eq}$	$f$
$10^1$	11.4	5.4	0.37	1.85
$10^2$	19.3	7.0	0.29	1.45
$10^3$	29.4	10.2	0.20	1.0
$10^4$	40.3	13.7	0.15	0.75
$10^5$	50.8	17.3	0.12	0.6

Table 1. Results of various experiments for the rising bubble experiment with  $\Delta t = 2$  s and  $\Delta x = 100$  m. The maximum and averaged number of iterations per time step ( $N_0$  and  $\bar{N}$ ) with the associated equivalent explicit time step  $\Delta t_{eq}$  and comparison factor  $f$  for different residual reduction parameters  $\varepsilon^{-1}$  are shown. For each of these experiments, the wind-CFL value  $c_{adv}^*$  is about 0.58.

$\Delta t$	$\Delta x$	$\varepsilon^{-1}$	$N_0$	$\bar{N}$	$\Delta t_{eq}$	$f$
2	100	$10^2$	15.5	6.2	0.32	1.6
4	200		16.8	5.2	0.77	1.9
2	100	$10^3$	19.5	6.7	0.30	1.5
4	200		19.0	5.6	0.71	1.8
2	100	$10^4$	20.7	7.6	0.26	1.3
4	200		20.2	6.3	0.63	1.6
2	100	$10^5$	25.3	9.8	0.20	1.0
4	200		25.2	7.9	0.51	1.3

Table 2. Results of various experiments for the orographic wave test case with  $\Delta t = 2$  s,  $\Delta x = 100$  m over a domain of 100 km and with  $\Delta t = 4$  s,  $\Delta x = 200$  m over a large domain of 1000 km (in italics). The maximum and averaged number of iterations per time step ( $N_0$  and  $\bar{N}$ ) with the associated equivalent explicit time step  $\Delta t_{eq}$  and comparison factor  $f$  for different residual reduction parameters  $\varepsilon^{-1}$  are shown. For each of these experiments, the wind-CFL value  $c_{adv}^*$  is about 0.24.

is obtained by respecting the RMSE threshold previously mentioned.

Table 3 shows that the average number of iterations is roughly of the same order of magnitude and the comparison factor  $f$  is close to the unity whatever the values ( $\Delta x, \Delta t$ )

$\Delta t$	$\Delta x$	$\varepsilon^{-1}$	$N_0$	$\bar{N}$	$\Delta t_{eq}$	$f$	$c_{adv}^*$
30	1000	29	26.2	4.7	6.4	1.60	0.83
15	500	19	17.3	4.4	3.4	1.70	0.99
6	200	90	26.6	6.3	0.95	1.19	1.02
3	100	90	26.2	7.9	0.38	0.95	0.88

Table 3. Impact of the time step value keeping  $\Delta t/\Delta x$  ratio constant in the warm bubble test case. The maximum and averaged number of iterations per time step ( $N_0$  and  $\bar{N}$ ) with the associated equivalent explicit time step  $\Delta t_{eq}$  and comparison factor  $f$  corresponding to the associated residual reduction under the RMSE threshold are shown. The wind-CFL is also given in the last column.

considered, thus confirming the good scalability. This also confirms the viability of keeping a constant  $\Delta t/\Delta x$  ratio when improving spatial resolution.

However, the slight increase in the average number of iterations  $\bar{N}$  (or equivalently the slight decrease of the comparison factor  $f$ ) as ( $\Delta x, \Delta t$ ) decreases may come from the use of slightly different initial conditions on the one hand, and a slight degradation of the condition number of each of the  $N_{lev}$  problems due to the change in eigenvalue distribution of  $\mathcal{B}$  when the time step decreases while the same vertical grid is kept (see for example Figure 1: eigenvalues are higher as the time step increases) on the other hand.

#### 5.4. Impact of time step reduction

We intend in this part to assess how the explicit equivalent time step evolves as a function of the time step at constant horizontal resolution, with the same experimental configuration as previously mentioned ( $\Delta x = 100$  m) with the warm bubble and orographic test cases. For each time step, the residual reduction value shown was adjusted by trial and error to obtain a solution close enough to the reference simulation just under the RMSE threshold imposed: see for example Figure 5 and Figure 6.

As mentioned in Table 4 and Table 5, the average number of iterations  $\bar{N}$  decreases as the time step decreases. In the orographic case, a substantially higher residual reduction is required but results are very similar to what is obtained with the warm bubble test case.

In all cases tested the comparison factor  $f$  is greater than unity, meaning that the efficiency of the method is maintained whatever the time step is. The slight decrease of the comparison factor  $f$  as the time step decreases comes from poorer condition numbers for each of the  $N_{lev}$  vertical problems, already mentioned in the previous paragraph. This is not a noticeable barrier since the comparison factor  $f$  remains larger than one even with smaller time step. Therefore, no obstacles have been identified from the efficiency point of view if a reduction in the time step were to be implemented.

Furthermore, we notice that in the GP version, a decrease by a factor of 3 of the time step (going from 3 s to 1 s for example) does not lead to an increase by a factor of 3 of the communication cost as it would be the case in a spectral model (because to make a forecast at a given time, 3 times more spectral transforms steps would then be necessary, whereas the GP model benefits from a decrease of iteration

$\Delta t$	$\varepsilon^{-1}$	$N_0$	$\bar{N}$	$\Delta t_{eq}$	$f$	$c_{adv}^*$
10	13500	87.2	18.2	0.55	2.75	1.20
5	3500	40.2	9.7	0.51	2.55	0.60
3	1100	27.7	7.6	0.39	1.95	0.36
2	300	19.1	6.6	0.30	1.50	0.24
1	100	9.4	5.0	0.2	1	0.12

Table 4. Impact of the time step value for the orographic wave test case. The maximum and averaged number of iterations per time step ( $N_0$  and  $\bar{N}$ ) with the associated equivalent explicit time step  $\Delta t_{eq}$  and comparison factor  $f$  corresponding to that residual reduction under the RMSE threshold are shown. The wind-CFL is also given in the last column.

$\Delta t$	$\varepsilon^{-1}$	$N_0$	$\bar{N}$	$\Delta t_{eq}$	$f$	$c_{adv}^*$
20	270	106.7	14.9	1.34	6.7	5.26
10	190	70.4	12.7	0.79	3.95	2.75
5	150	43.3	10.3	0.48	2.40	1.35
3	90	26.2	7.9	0.38	1.90	0.88
2	55	17.6	6.1	0.33	1.65	0.59
1	4	5.8	4.2	0.24	1.20	0.29

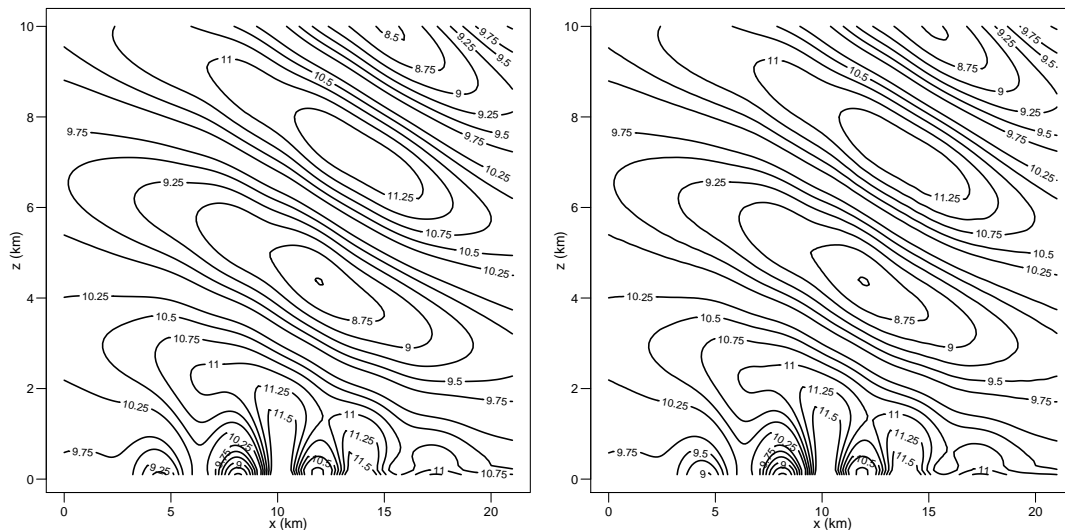
Table 5. Impact of the time step value for the warm bubble test case. The maximum and averaged number of iterations per time step ( $N_0$  and  $\bar{N}$ ) with the associated equivalent explicit time step  $\Delta t_{eq}$  and comparison factor  $f$  corresponding to the associated residual reduction under the RMSE threshold are shown. The wind-CFL is also given in the last column.

per time step). An increase only by a factor barely more than 2 is measured: the equivalent explicit time step goes from 0.39 s to 0.2 s and from 0.38 s to 0.24 s respectively in the orographic (Table 4) and warm bubble (Table 5) test cases. Therefore, the overhead that occurs when the time step decreases at a given resolution turns out to be lower in the GP version than in the SP one. This comment, in addition to those already mentioned, enhances the attractiveness of the grid-point method.

## 6. Conclusion

The objective of this paper was to present a more scalable alternative to the current spectral AROME model while keeping most of its main characteristics (constant coefficient SI scheme, SL transport scheme, A-grid, mass based coordinate, etc). This alternative relies on the removal of spectral transforms to perform all calculations in grid-point space including the implicit problem, solved here by a Krylov algorithm. This class of algorithms consists in iterating local operators with few communications, the number of iterations depending on the speed of convergence and the desired accuracy.

The known convergence results are often too general to lead to a sufficiently precise estimate of the speed of convergence, so that a specific study is required with all the characteristics of the model under study. This work has been led with a set of bidimensional non-linear flows which, though idealized, are reminiscent of realistic situations, and involve all source terms of the complete model. The



**Figure 6.** Mountain waves simulation with  $\Delta x = 100$  m and  $\Delta t = 10$  s, horizontal wind after 8000 s focused on the middle part of the domain. Left: reference for the GP version; right: with a residual reduction  $\varepsilon^{-1}$  of 13500.

simulations and developments were carried out with a 2D vertical-slice configuration of AROME.

The grid-point version developed here retains some beneficial features already present in the spectral version of the AROME model, namely the decomposition of the implicit problem on the vertical eigenmodes which plays the role of a vertical preconditioner. This is easily accessible through the choice of a ‘constant coefficient’ basic state.

The spectral version of the AROME model, which has been operational for several years, has been used here as a reference in terms of quality. In a first time, the possibility of producing forecasts of similar overall quality to spectral ones has been examined. It has been found that for the converged solution (i.e. with  $\varepsilon$  close to the machine precision), grid-point forecasts are very close to spectral ones with the same grid, provided a sufficiently high-order space-accurate scheme is used (fourth-order and sixth-order gave very similar results whereas second-order resulted in poorer solutions). This result thus validates the strategy of keeping as much as possible the main characteristics of the current AROME model.

After correctly initializing the solver, the critical threshold at which the solver loop is exited has been adjusted to ensure that the errors between the partially-converged solution (i.e. with  $\varepsilon$  far away from the machine precision) and the spectral version are significantly lower than those usually committed by the latter in an operational context. It has been found that between 4 and 7 iterations are thus necessary to invert the implicit problem per time step when a CFL equivalent to the one currently in the operational model is used, without any significant loss of quality compared to the spectral version.

In contrast to the spectral version of an SI model that requires global communications, HEVI and SI models involve essentially local operators of almost the same stencil, and are assumed in this study to have the same computational cost and exchange of data. The efficiencies of these two methods were then compared by measuring the necessary number of applications of these local operators,

and using a single synthetic metric based on the comparison of the effective time steps of the two methods.

Moreover, experiments have been carried out to be informative whatever the future strategy planned for the time step: keeping large time steps so as to keep the ratio of spatial resolution to temporal resolution constant, as has been done up to now, or gradually reducing this ratio to better represent some of the non-hydrostatic effects.

The results show that the implicit grid point method we used is effective regardless of the strategy considered. Although no direct measurement of scalability has been performed, data exchanges are expected almost as frequent as these involved during the small time step integration of the split-explicit scheme of a HEVI model. This suggests an improvement of the scalability of this new version in comparison to the spectral version. These results militate in favour of removing the spectral calculations in AROME while retaining its main characteristics elsewhere. Besides, the inevitable overhead that occurs when the time step decreases at a given resolution turns out to be relatively lower in the grid-point version than in the spectral one. In doing so, by bypassing the expensive spectral transform stage, the grid-point method appears to be more promising than the spectral one if a time step reduction strategy is adopted.

These promising results have led to the implementation of an almost entirely grid-point version in the operational 3D AROME model while numerical diffusion has been kept in spectral space. On average only 6 iterations per time step are needed to invert the implicit problem with the same quality as measured by scores to observations over a one-month period. These results, although partial, confirm the results obtained in this study and can be seen as supplementary material. Moreover, it has been checked that using a non-fully converged solver, does not result on additional significant lack of conservation, compared to the use of a semi-Lagrangian scheme.

Furthermore, in view of using models with hectometric resolution on the horizontal in the next few years, thus better representing orography and in particular steeper slopes, it

might be necessary to relax the assumption of flat terrain in the definition of the SI base state and to include some of the orographic terms in the implicit problem (such a problem can not be easily solved in spectral space). Preliminary studies show that the stability of the SI scheme in the grid-point version presented here may be significantly improved by doing so.

### Acknowledgements

We are grateful to the two anonymous reviewers for their constructive comments aimed at improve the paper.

### A. Governing equations

The vertical coordinate  $\eta$  is related to hydrostatic pressure  $\pi$  by :

$$\pi(x, \eta, t) = A(\eta) + B(\eta) \pi_s(x, t), \quad (30)$$

where  $A$  and  $B$  are functions used in the definition of the vertical coordinate  $\eta$ , and  $\pi_s(x, t)$  is the surface hydrostatic pressure (Laprise (1992)). Governing equations for the dry two-dimensional system are :

$$\frac{dU}{dt} + \frac{R_d T}{p} \frac{\partial p}{\partial x} + \frac{1}{m} \frac{\partial p}{\partial \eta} \frac{\partial \phi}{\partial x} = \nu,$$

$$\begin{aligned} \frac{dd}{dt} + g^2 \frac{p}{m R_d T} \frac{\partial}{\partial \eta} \left( \frac{1}{m} \frac{\partial(p - \pi)}{\partial \eta} \right) - g \frac{p}{m R_d T} \frac{\partial U}{\partial \eta} \frac{\partial w}{\partial x} \\ - d \left[ \frac{\partial U}{\partial x} - \left( \frac{\partial U}{\partial x} + d + \frac{p}{R_d T m} \frac{\partial \phi}{\partial x} \frac{\partial U}{\partial \eta} \right) \right] \\ = -g \frac{p}{m R_d T} \frac{\partial W}{\partial \eta}, \end{aligned}$$

$$\frac{dT}{dt} + \frac{R_d T}{C_{vd}} \left( \frac{\partial U}{\partial x} + d + \frac{p}{R_d T m} \frac{\partial \phi}{\partial x} \frac{\partial U}{\partial \eta} \right) = \frac{Q}{C_{vd}},$$

$$\frac{d\hat{q}}{dt} + \frac{C_{pd}}{C_{vd}} \left( \frac{\partial U}{\partial x} + d + \frac{p}{R_d T m} \frac{\partial \phi}{\partial x} \frac{\partial U}{\partial \eta} \right) + \frac{\dot{\pi}}{\pi} = \frac{Q}{C_{vd} T},$$

$$\frac{\partial \pi_s}{\partial t} + \int_0^1 \frac{\partial(mU)}{\partial x} d\eta = 0.$$

The following diagnostic relations are:

$$m = \frac{\partial \pi}{\partial \eta},$$

$$p = \pi e^{\hat{q}},$$

$$\phi = \phi_s + \int_{\eta}^1 \frac{m R T}{p} d\eta,$$

$$\dot{\pi} = U \frac{\partial \pi}{\partial x} - \int_0^{\eta} \frac{\partial(mU)}{\partial x} d\eta',$$

$$\begin{aligned} g \frac{\partial w}{\partial x} = g \frac{\partial w_s}{\partial x} + \int_{\eta}^1 \frac{m R_d T}{p} \frac{\partial d}{\partial x} d\eta' \\ + \int_{\eta}^1 R_d d \frac{\partial}{\partial x} \left( \frac{m T}{p} \right) d\eta'. \end{aligned}$$

### B. Notations used

We use the following notation:

- $p$  : local pressure
- $w$  : vertical velocity
- $\phi$  : geopotential
- $g$  : acceleration of gravity
- $R_d$  : 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
- $\nu, W, Q$  : physical contributions for  $U, w$  and  $T$  respectively

The underscript 's' refers to a surface field.

### C. Vertical operators of $\mathcal{L}$

$$\partial^* X = \frac{\pi^*}{m^*} \frac{\partial X}{\partial \eta}, \quad (31)$$

$$\mathcal{G}^* X = \int_{\eta}^1 \frac{m^*}{\pi^*} X d\eta, \quad (32)$$

$$\mathcal{S}^* = \frac{1}{\pi^*} \int_0^{\eta} m^* X d\eta, \quad (33)$$

$$\mathcal{N}^* X = \frac{1}{\pi_s^*} \int_0^1 m^* X d\eta, \quad (34)$$

$$\mathcal{L}^* X = \partial^* (\partial^* + 1) X, \quad (35)$$

where  $\pi^*(\eta) = A(\eta) + B(\eta) \pi_s^*$  is the hydrostatic pressure of the basic state and  $m^* = \partial \pi^* / \partial \eta$  is its vertical metric factor associated.

### References

- Auger, L., Dupont, O., Hagelin, S., Brousseau, P., and Brovelli, P. (2015). AROME–NWC: a new nowcasting tool based on an operational mesoscale forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 141(690):1603–1611.
- Axelsson, O. (1996). *Iterative Solution Methods*. Cambridge University Press.
- Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55.
- 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. and Glinton, M. R. (2019). Circumventing the pole problem of reduced lat–lon grids with local schemes. Part I: Analysis and model formulation. *Quarterly Journal of the Royal Meteorological Society*, 145(721):1377–1391.
- Bénard, P., Vivoda, J., Masek, J., Smolíková, P., Yessad, K., Smith, C., Brozková, R., and Geleyn, J.-F. (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.

- Bryan, G. H. and Fritsch, J. M. (2002). A benchmark simulation for moist nonhydrostatic numerical models. *Monthly Weather Review*, 130(12):2917–2928.
- Bubnová, R., Hello, G., Bénard, P., and Geleyn, J.-F. (1995). Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Monthly Weather Review*, 123(2):515–535.
- Carpenter Jr, R. L., Droegemeier, K. K., Woodward, P. R., and Hane, C. E. (1990). Application of the piecewise parabolic method (PPM) to meteorological modeling. *Monthly Weather Review*, 118(3):586–612.
- Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A., and Wood, N. (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.
- Durran, D. R. (2013). *Numerical methods for wave equations in geophysical fluid dynamics*, volume 32. Springer Science & Business Media.
- Fischer, P. F. (2015). Scaling limits for PDE-based simulation (invited). In *22nd AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics.
- Gassmann, A. (2013). A global hexagonal C-grid non-hydrostatic dynamical core (ICON-IAP) designed for energetic consistency. *Quarterly Journal of the Royal Meteorological Society*, 139(670):152–175.
- Harris, L. M. and Lin, S.-J. (2013). A two-way nested global-regional dynamical core on the cubed-sphere grid. *Monthly Weather Review*, 141(1):283–306.
- Haugen, J. E. and Machenhauer, B. (1993). A spectral limited-area model formulation with time-dependent boundary conditions applied to the shallow-water equations. *Monthly Weather Review*, 121(9):2618–2630.
- Hess, R. and Joppich, W. (1997). A comparison of parallel multigrid and a fast Fourier transform algorithm for the solution of the Helmholtz equation in numerical weather prediction. *Parallel Computing*, 22(11):1503–1512.
- Klemp, J. B., Skamarock, W. C., and Dudhia, J. (2007). Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. *Monthly Weather Review*, 135(8):2897–2913.
- Kühnlein, C., Deconinck, W., Klein, R., Malardel, S., Piotrowski, Z. P., Smolarkiewicz, P. K., Szmelter, J., and Wedi, N. P. (2019). FVM 1.0: a nonhydrostatic finite-volume dynamical core for the IFS. *Geoscientific Model Development*, 12(2):651–676.
- Laprise, R. (1992). The Euler equations of motion with hydrostatic pressure as an independent variable. *Monthly weather review*, 120(1):197–207.
- Liesen, J. and Tichý, P. (2004). Convergence analysis of Krylov subspace methods. *GAMM-Mitteilungen*, 27(2):153–173.
- Maynard, C., Melvin, T., and Muller, E. H. (2020). Multigrid preconditioners for the mixed finite element dynamical core of the LFRic atmospheric model. *Quarterly Journal of the Royal Meteorological Society*.
- Mesinger, F. (1979). Dependence of vorticity analogue and the Rossby wave phase speed on the choice of horizontal grid. *Bulletin (Académie serbe des sciences et des arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)*, pages 5–15.
- Müller, E. H. and Scheichl, R. (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.
- Pinty, J.-P., Benoit, R., Richard, E., and Laprise, R. (1995). Simple tests of a semi-implicit semi-Lagrangian model on 2D mountain wave problems. *Monthly Weather Review*, 123(10):3042–3058.
- Putman, W. M. and Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, 227(1):55–78.
- Qaddouri, A. and Lee, V. (2010). The elliptic solvers in the canadian limited area forecasting model GEM-LAM. In *Modeling Simulation and Optimization-Tolerance and Optimal Control*. IntechOpen.
- Saad, Y. and Schultz, M. H. (1986). GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3):856–869.
- Satoh, M. (2002). Conservative scheme for the compressible nonhydrostatic models with the Horizontally Explicit and Vertically Implicit time integration scheme. *Monthly Weather Review*, 130(5):1227–1245.
- Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., and Iga, S. (2008). Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *Journal of Computational Physics*, 227(7):3486–3514.
- Schär, C., Leuenberger, D., Fuhrer, O., Lüthi, D., and Girard, C. (2002). A new terrain-following vertical coordinate formulation for atmospheric prediction models. *Monthly Weather Review*, 130(10):2459–2480.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V. (2011). The AROME-France convective-scale operational model. *Monthly Weather Review*, 139:976–991.
- Simmons, A. J., Hoskins, B. J., and Burridge, D. M. (1978). Stability of the semi-implicit method of time integration. *Monthly Weather Review*, 106(3):405–412.
- Skamarock, W. C., Smolarkiewicz, P. K., and Klemp, J. B. (1997). Preconditioned conjugate-residual solvers for Helmholtz equations in nonhydrostatic models. *Monthly weather review*, 125(4):587–599.
- Smolarkiewicz, P. and Margolin, L. (1994). Variational solver for elliptic problems in atmospheric flows. *Appl. Math. Comp. Sci*, 4(4):527–551.
- Smolarkiewicz, P. K. and Margolin, L. G. (1997). On forward-in-time differencing for fluids: an Eulerian/semi-Lagrangian non-hydrostatic model for stratified flows. *Atmosphere-Ocean*, 35(sup1):127–152.
- Soleymani, F. (2013). A new method for solving ill-conditioned linear systems. *Opuscula Mathematica*, 33.
- Steffeler, J., Hess, R., Schättler, U., and Bonaventura, L. (2003). Review of numerical methods for nonhydrostatic weather prediction models. *Meteorology and Atmospheric Physics*, 82(1):287–301.
- Thomas, S. J., Malevsky, A. V., Desgagné, M., Benoit, R., Pellerin, P., and Valin, M. (1997). Massively parallel implementation of the mesoscale compressible

- community model. *Parallel Computing*, 23(14):2143–2160.
- Wedi, N., Bauer, P., Deconinck, W., Diamantakis, M., Hamrud, M., Kuehnlein, C., Malardel, S., Mogensen, K., Mozdzyński, G., and Smolarkiewicz, P. (2015). The modelling infrastructure of the integrated forecasting system: Recent advances and future challenges.
- Wedi, N. P. (2014). Increasing horizontal resolution in numerical weather prediction and climate simulations: illusion or panacea? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2018):20130289.
- Wicker, L. J. and Skamarock, W. C. (1998). A time-splitting scheme for the elastic equations incorporating second-order Runge–Kutta time differencing. *Monthly Weather Review*, 126(7):1992–1999.
- Ye, Q. (2017). Preconditioning for accurate solutions of linear systems and eigenvalue problems. *arXiv preprint arXiv:1705.04340*.
- Zheng, Y. and Marguinaud, P. (2018). Simulation of the performance and scalability of message passing interface (MPI) communications of atmospheric models running on exascale supercomputers. *Geoscientific Model Development*, 11(8):3409–3426.
- Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M. (2014). The ICON (ICOsahedral non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, 141(687):563–579.

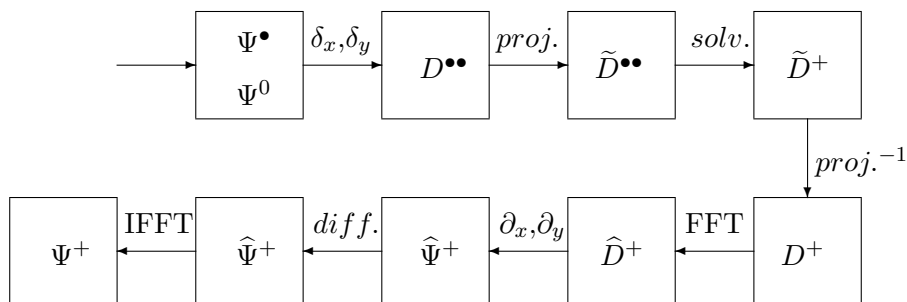
## 3.5 AROME 3D partiellement points de grille

### 3.5.1 Présentation générale

Les estimations de scalabilité exposées dans l'article ont été menées sur des écoulements idéalisés. On souhaite désormais mener des estimations en suivant la même méthodologie, pour des écoulements plus réalistes en 3D et en présence de paramétrisations physiques.

Comme mentionné précédemment, le portage de l'intégralité du code en points de grille est une tâche importante et qui n'a pas été effectuée pendant cette thèse. A la place, une version « partiellement points de grille » a été développée ici, où certains calculs sont effectués en points de grille tandis que d'autres demeurent dans l'espace spectral. Pour prendre en compte les derniers développements implémentés dans le modèle AROME, cette version a été développée à partir d'un cycle récent (cycle 46) du modèle opérationnel AROME.

Le fonctionnement schématique de cette version « partiellement points de grille » devient :



La différence principale de cette version par rapport à la version entièrement en points de grille exposée dans le paragraphe 3.1.5, est qu'une transformée de Fourier est appliquée au champ de divergence horizontale du vent  $D^{+}$ , une fois que le problème implicite est inversé via un solveur de Krylov. Les autres variables pronostiques sont déduites à l'instant ultérieur dans l'espace spectral ( $\hat{\Psi}^{+}$ ). Ainsi la reconstitution du champ de vent à partir du champ de divergence est faite dans cet espace spectral comme mentionné dans le paragraphe 2.1.6. De même, la diffusion est ajoutée implicitement dans l'espace spectral comme décrit dans le paragraphe 2.1.13. Pour plus de simplicité, la version « partiellement points de grille » sera dorénavant appelée version « points de grille ».

Les opérateurs  $(\delta_x, \delta_y)$  sont discrétisés en différences finies à un ordre élevé (jusqu'à l'ordre 10 dans certains cas) pour atténuer les inconsistances avec les opérateurs discrétisés dans l'espace spectral  $(\partial_x, \partial_y)$ . Cela atténue les effets numériques les plus néfastes discutés dans le paragraphe 3.1.2.

L'un des points principaux de cette thèse est d'évaluer la vitesse de convergence du solveur pour donner une estimation de la scalabilité. Le conditionnement du problème implicite n'étant pas très sensible à l'ordre de discrétisation utilisé, il est attendu que la vitesse de

convergence n'en dépende pas non plus. L'utilisation d'un ordre très élevé de discrétisation est donc valable uniquement pour ce travail de thèse (compte tenu des hypothèses de travail) et ne sera probablement pas utilisée dans un cadre opérationnel, une fois que tout le code sera porté en points de grille et que les opérateurs de dérivée horizontale seront tous consistants entre eux.

#### 3.5.2 Critère d'arrêt

La méthodologie d'estimation de la scalabilité est identique à celle utilisée pour la version AROME 2D exposée dans la section 3.2 et consiste à mesurer la vitesse de convergence qui dépend principalement du choix du critère d'arrêt et de l'initialisation du solveur.

Dans les tests qui suivront, plusieurs critères d'arrêt sont considérés. Le cadre idéalisé en 2D exposé précédemment était bien adapté pour utiliser un critère d'arrêt basé sur la mesure de l'écart de la solution courante à la solution convergée (3.15). De plus, un même critère d'arrêt était appliqué pour inverser les  $N_{lev}$  problèmes ( $\varepsilon$  était indépendant du mode considéré). En revanche, dans un cadre 3D à vocation opérationnelle, les critères basés sur le nombre d'itérations sont privilégiés pour fournir une prévision à un temps fixe, qui ne dépend pas de la vitesse de convergence du solveur. De plus, pour optimiser encore la vitesse de convergence, on souhaite explorer un critère d'arrêt qui dépend du mode vertical considéré (3.16).

La projection dans l'espace propre de l'opérateur vertical  $\mathcal{B}$ , conduit à inverser  $N_{lev}$  problèmes implicites (2.90) dont chacun est identique à celui d'un système shallow-water dont la vitesse de phase des ondes est  $\sqrt{b_l}$ . Dans un premier temps, on souhaite trouver le critère le plus économe possible et qui garantit la stabilité numérique. Le cadre shallow-water est bien adapté pour mener une analyse de stabilité simple du gradient conjugué où les paramètres de l'analyse sont la vitesse de phase du mode considéré  $\sqrt{b_l}$  et le nombre d'itération associé. Cette analyse est exposée dans l'Annexe B. Cette analyse ne permet de discriminer que deux cas pour garantir la stabilité numérique : celui où le nombre de CFL du mode considéré  $c_l^* = \frac{\sqrt{b_l}\Delta t}{\Delta x}$  est supérieur à l'unité (sur une grille A) et qui requière au moins deux itérations de l'algorithme, de celui où il est inférieur et qui ne requière qu'une seule itération.

Cette analyse est basée sur un système dont le membre de droite est très simplifié par rapport à celui du modèle AROME. Cela limite ainsi la portée de l'analyse mais aide à définir un critère d'arrêt qui maintienne la stabilité numérique. Toutefois, la stabilité numérique, bien que nécessaire n'est pas suffisante. Il faut s'assurer également que l'usage de ce critère d'arrêt ne détériore pas la qualité de la solution. Par conséquent, dans toute la suite, le critère d'arrêt est basé sur un nombre maximum d'itérations, proportionnel à la partie entière supérieure du nombre CFL de chacun des  $N_{lev}$  modes considéré, c'est-à-dire avec

$$N_{max,l} = \gamma \left\lceil \frac{\sqrt{b_l}\Delta t}{\Delta x} \right\rceil, \quad (3.25)$$



où  $\gamma$  est un entier à régler pour assurer la stabilité et une qualité suffisante. Le choix d'une référence pour estimer la qualité est désormais discuté.

### 3.5.3 Estimation de la qualité

De façon analogue à la démarche menée dans l'article, la version en points de grille peut être comparée à la version spectrale. L'utilisation de méthodes en différences finies conduit à une dégradation de la qualité de la discrétisation des opérateurs de dérivée horizontale par rapport à la version spectrale où les dérivées horizontales sont calculées très précisément. Par conséquent, un écart de la solution PG à la solution SP sera interprété comme une dégradation de la qualité de la simulation. Ces écarts seront mesurés par l'écart quadratique moyen (EQM) pour plusieurs champs régulièrement étudiés en météorologie : la température et les composantes horizontales du vent pour plusieurs échéances et à différents niveaux modèles. Cette comparaison sera principalement utilisée sur un domaine de petite taille (60 x 60 points sur l'horizontale et avec une grille verticale de 90 niveaux) qui facilite grandement les tests dans un cadre simple et non parallélisé. Ce domaine sera utilisé pour les expériences des paragraphes 3.6.1 et 3.6.2

Toutefois, comme nous l'avons vu, les deux versions (PG et SP) ne sont pas indépendantes l'une de l'autre. Des scores sont calculés en comparant les résultats de la version PG, avec des mesures indépendantes du modèle sur des radiosondages. Pour que cette évaluation de la qualité soit significative statistiquement, les expériences sont menées sur un grand nombre de cas (en l'occurrence de jours de simulation). La comparaison aux scores ne peut être utilisée que sur un code parallélisé, pour mener de nombreuses simulations le domaine opérationnel du modèle AROME. Le code a été parallélisé sous le protocole MPI par L. Auger. J'expose dans les paragraphes 3.6.3 et 3.6.4 les résultats obtenus suite à ce travail. De plus, comme précisé dans l'introduction, l'un des objectifs d'AROME est de modéliser les phénomènes à enjeux, comme les événements fortement précipitants qui ont lieux régulièrement autour de la Méditerranée chaque année en automne. La période analysée sera celle du 19 octobre 2019 au 16 novembre 2019, qui permet d'une part d'évaluer des scores sur une période assez longue, et d'autre part d'analyser un épisode à enjeu.

## 3.6 Évaluation de la scalabilité du modèle AROME 3D PG

Comme mentionné précédemment dans les cas en 2D, la vitesse de convergence du solveur dépend fortement du critère d'arrêt et de l'initialisation du solveur  $x_0$ . Une solution insuffisamment convergée peut conduire à :

- une instabilité numérique,
- une dégradation de la qualité de la solution.

Dans un premier temps dans cette partie, on règle le critère d'arrêt pour qu'il assure la stabilité numérique sans dégrader la qualité de la solution par rapport à la version spectrale sur les

premiers pas de temps d'intégration. Dans un second temps, on vérifie que le réglage choisi ne dégrade pas la qualité par rapport à la version spectrale actuelle sur les 200 premiers pas de temps, à différents niveaux, puis sur un évènement fortement précipitant à enjeu. Enfin, on vérifie également que ce réglage ne dégrade pas les scores par rapport à des observations de qualité (radiosondages) sur une période d'environ un mois.

### 3.6.1 Paramétrage du solveur de Krylov

Dans cette section on fait varier le critère d'arrêt pour assurer la stabilité sur les premiers pas de temps et fournir des solutions de qualité comparable (mesurée par l'écart quadratique moyen sur le champ de température, moyenné sur les  $N_{lev}$  niveaux du modèle). Le nombre moyen d'itérations  $\bar{N}$  (c'est-à-dire moyenné sur les  $N_{lev}$  niveaux du modèle) par pas de temps (c'est-à-dire pour les étapes de prédicteur et correcteur) est donné dans la Table 3.1 pour :

- différents critères portant sur l'écart à la solution convergée (3.15) ou sur le nombre d'itérations (3.16),
- différentes initialisations possibles : en prenant soit un champ de divergence nulle ( $x_0 = 0$ ), soit le champ de divergence calculée précédemment (c'est-à-dire le champ du pas de temps précédent pour le prédicteur, et le champ du prédicteur pour le correcteur ( $x_0 = D^{(n)}$ )).

$N_{max,l}$	$\varepsilon$	$x_0$	$\bar{N}$
$\left\lceil \frac{\sqrt{b_l} \Delta t}{\Delta x} \right\rceil$	–	$D^{(n)}$	3
–	$10^{-3}$	$D^{(n)}$	7
–	$10^{-3}$	0	10
–	$10^{-6}$	0	22
–	$10^{-10}$	0	35

TABLE 3.1 – Mesures du nombre d'itérations nécessaires pour inverser le problème à une qualité donnée et initialisé par  $x_0$  pour différents critères d'arrêts.

#### Sensibilité à l'initialisation

Lorsque le solveur est initialisé par un champ de divergence nulle, seules 7 itérations sont nécessaires pour inverser le problème à la précision ( $\varepsilon = 10^{-3}$ ) alors qu'il en faut en moyenne 10 lorsqu'il est initialisé par un champ à divergence horizontale nulle. Le coût additionnel de l'évaluation du champ initial est largement compensé par la vitesse de convergence plus rapide.

### Sensibilité au critère d'arrêt

Lorsque le solveur est initialisé par un champ de divergence nulle et qu'un critère d'arrêt restrictif est considéré ( $\varepsilon = 10^{-6}$ ), il faut alors en moyenne deux fois plus d'itérations que lorsqu'un critère plus permissif est considéré ( $\varepsilon = 10^{-3}$ ). La Figure 3.3 représente l'évolution du nombre d'itérations nécessaires pour chacun des  $N_{\text{lev}}$  problèmes, pour ces deux critères d'arrêt et pour chacune des étapes du schéma prédicteur-correcteur. On retrouve ainsi des résultats analogues à ceux obtenus dans l'article dans la version bidimensionnelle du modèle AROME, confirmant l'intérêt de régler précisément le critère d'arrêt.

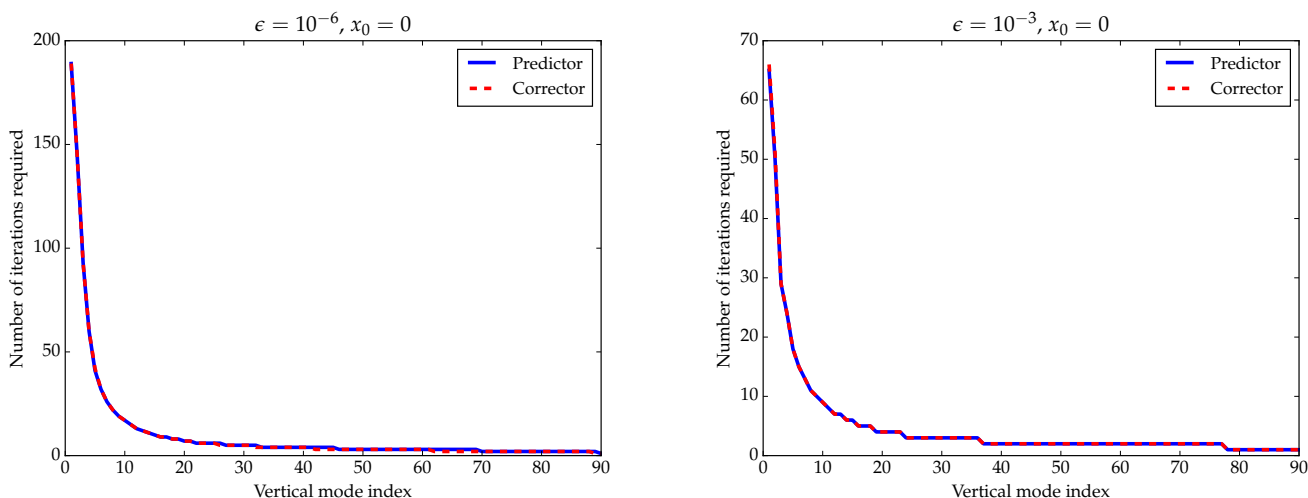


FIGURE 3.3 – Évolution du nombre d'itérations en fonction du mode considéré pour différentes initialisation  $x_0$  et différents critères d'arrêt pour l'étape de prédicteur (en bleu) et l'étape de correcteur (en rouge).

### Intérêt d'un critère d'arrêt basé sur le nombre d'itérations

En comparant les deux premières lignes de la Table 3.1, on note que lorsque le solveur est initialisé par le champ de divergence du pas de temps précédent, et qu'un critère d'arrêt imposant le nombre d'itérations égal au CFL de chacun des modes est posé (3.16) avec  $\gamma = 1$ , il faut moitié moins d'itérations que lorsqu'un critère (pourtant permissif) sur l'écart à la solution convergé est imposé ( $\varepsilon = 10^{-3}$ ). Dans toute la suite, ce genre de critère d'arrêt sera considéré.

La Figure 3.4 représente l'évolution du nombre d'itérations nécessaires pour chacun des  $N_{\text{lev}}$  problèmes, pour ces deux critères d'arrêt et pour chacune des étapes du schéma prédicteur-correcteur. Suivant cette expérience menée sur uniquement quelques pas de temps, 80 des 90 modes n'ont besoin que d'une seule itération du solveur pour être inversés. Par conséquent, utiliser un préconditionneur n'est pas nécessaire pour inverser la plupart des problèmes. On note également que le mode le plus coûteux requière 14 fois plus d'itérations que la

### 3.6. ÉVALUATION DE LA SCALABILITÉ DU MODÈLE AROME 3D PG

majorité des modes qui n'en requièrent qu'une seule. Cela confirme l'intérêt de les traiter séquentiellement, plutôt que simultanément. Comme mentionné dans l'article, en accélérant la convergence, la projection joue ici un rôle de préconditionneur, analogue à celui que jouerait un préconditionneur de type ADI (alternating-direction implicit) sur la verticale.

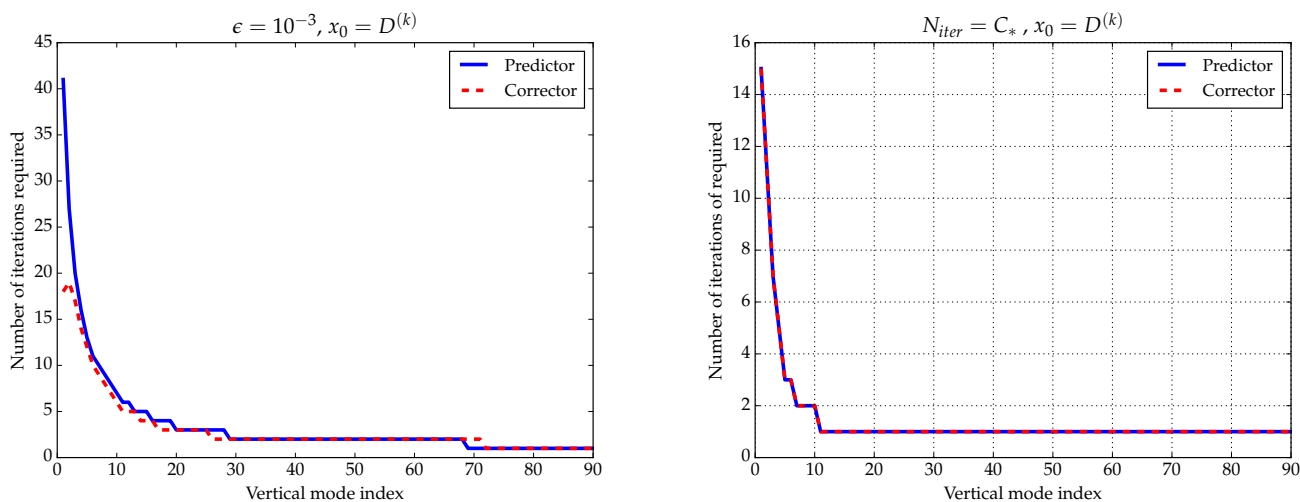


FIGURE 3.4 – Évolution du nombre d'itérations en fonction du mode considéré pour différentes initialisation  $x_0$  et différents critères d'arrêt pour l'étape de prédicteur (en bleu) et l'étape de correcteur (en rouge).

#### 3.6.2 Validation sur plusieurs champs et niveaux verticaux

On souhaite désormais vérifier ces résultats sur des simulations plus longues et sur des champs à différents niveaux pour un cas avec très peu d'itérations du solveur ( $\gamma = 1$ ) et un cas où beaucoup plus d'itérations sont utilisées ( $\gamma = 10$ ). Une discrétisation à l'ordre 4 sur une grille Z est considérée. Les niveaux choisis sont les niveaux d'intérêt en météorologie : 90, 61, 34, 13 qui désignent respectivement les niveaux au sol, à 850 hPa, 500 hPa et 200 hPa. Les champs analysés sont la température et les composantes zonale et méridienne du vent.

Les résultats sont exposés sur les Figures 3.5, 3.6 et 3.7. On note que les EQM sont de l'ordre de quelques dixième de degrés pour la température et de l'ordre du m/s pour les composantes du vent. Les changements apportés par le passage en points de grille sont donc perceptibles par rapport à la version spectrale sur la plupart des champs exposés ici. Dans les deux cas considérés (pour  $\gamma = 1$  et  $\gamma = 10$ ), les EQM pour chacun des critères, sont très proches. Par conséquent, le critère avec  $\gamma = 1$  semble suffisant pour assurer la stabilité numérique et une qualité acceptable, à faible coût. Toutefois cette observation nécessite d'être confirmée sur des simulations plus longues, sur d'autres champs, et sur d'autres cas.

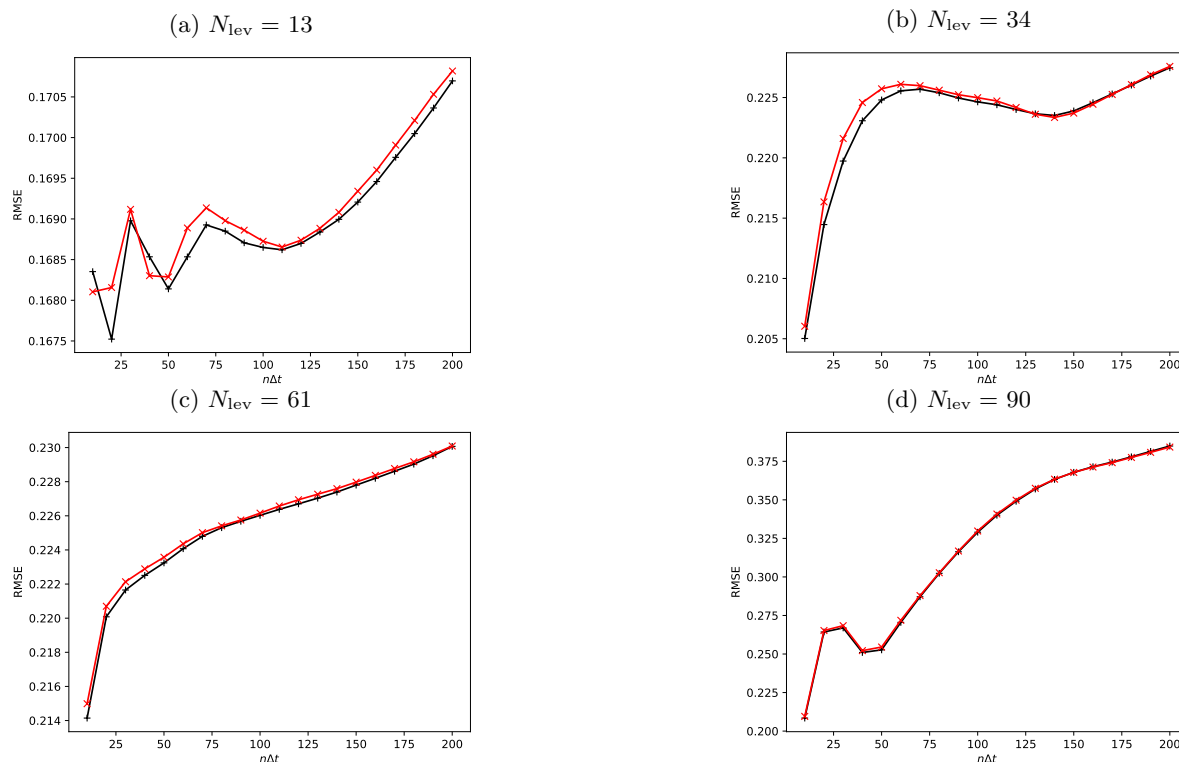


FIGURE 3.5 – EQM (par rapport à la version spectrale) pour la température à 4 niveau verticaux différents, en rouge (avec une croix x) pour  $\gamma = 10$ , en noir (avec un +) pour  $\gamma = 1$ .

### 3.6.3 Comparaison sur cas réel fortement précipitant

Les résultats obtenus dans cette section et la suivante, sont basés sur un code parallélisé par L. Auger.

Comme nous l'avons mentionné précédemment, l'un des objectifs du modèle AROME est notamment de prévoir la convection profonde. Ainsi un évènement fortement précipitant qui a eu lieu dans le sud de la France du 22/10/2019 à 0h au 23/10/2019 à 0h est modélisé avec la version spectrale et la version points de grille. Ces deux versions sont comparées aux observations via la lame d'eau Antilope (qui combine les observations radars et les relevés des pluviomètres). La version PG utilise un critère d'arrêt sur le nombre d'itérations (3.16) avec  $\gamma = 2$  (seulement 6 itérations sont effectuées en moyenne par le solveur par pas de temps) et une discrétisation à l'ordre 10 sur une grille Z est utilisée. On note que lorsque  $\gamma = 1$ , la qualité de la simulation est légèrement dégradée (non montré).

Les résultats des cumuls de précipitations sont montrés sur la Figure 3.8. Contrairement aux champs précédemment étudiés, le champ de cumul des précipitations est un champ diagnostic, dépendant de la partie dynamique et de la partie physique. Les différences entre les deux versions (SP et PG) sont très faibles, et bien inférieures aux différences entre chacune de ces versions avec la lame d'eau Antilope. Sur ce cas à enjeu, la version PG, n'a pas dégradé la prévision par rapport à la version spectrale. On vérifie désormais que ce résultat est également

### 3.7. CONCLUSION

---

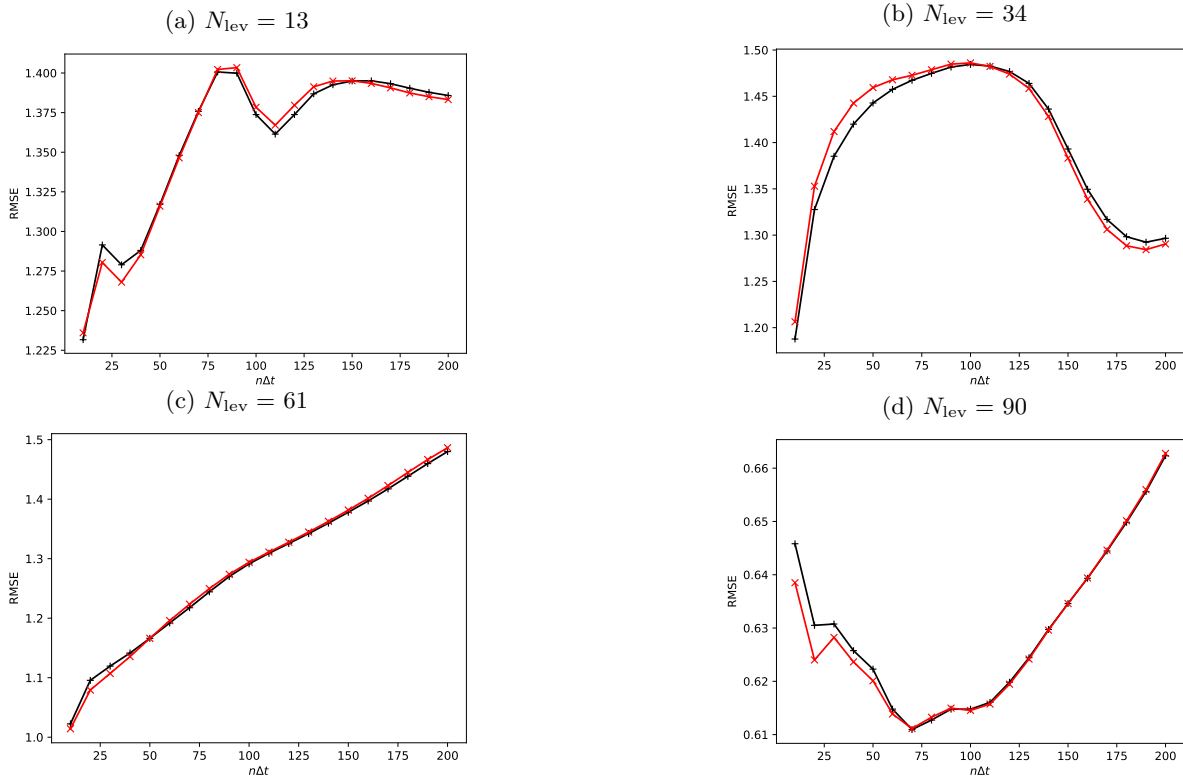


FIGURE 3.6 – EQM (par rapport à la version spectrale) pour la composante zonale du vent à 4 niveau verticaux différents, en rouge (avec une croix x) pour  $\gamma = 10$ , en noir (avec un +) pour  $\gamma = 1$ .

valable sur un plus grand nombre de cas.

#### 3.6.4 Scores sur un mois

Dans cette partie, les deux versions sont comparées aux mesures des radiosondages pour le vent et la température. Le biais et l'erreur quadratique moyenne sont calculés sur un période de quasiment un mois. En reprenant le même critère d'arrêt et le même ordre de discrétisation, les scores aux radiosondages ne sont pas dégradés dans la version points de grille par rapport à la version spectrale. En revanche, lorsqu'un critère d'arrêt (3.16) avec  $\gamma = 1$  est choisi, les scores sont légèrement dégradés (non montré).

### 3.7 Conclusion

Dans ce chapitre, nous avons présenté une version points de grille à coefficients constants du modèle AROME héritant de la plupart des propriétés de la version spectrale, à l'exception du calcul des opérateurs de dérivées horizontales, de la résolution du problème implicite, et du traitement implicite de la diffusion numérique. Après avoir exposé brièvement les caractéristiques principales de cette version et introduit les solveurs de Krylov pour résoudre

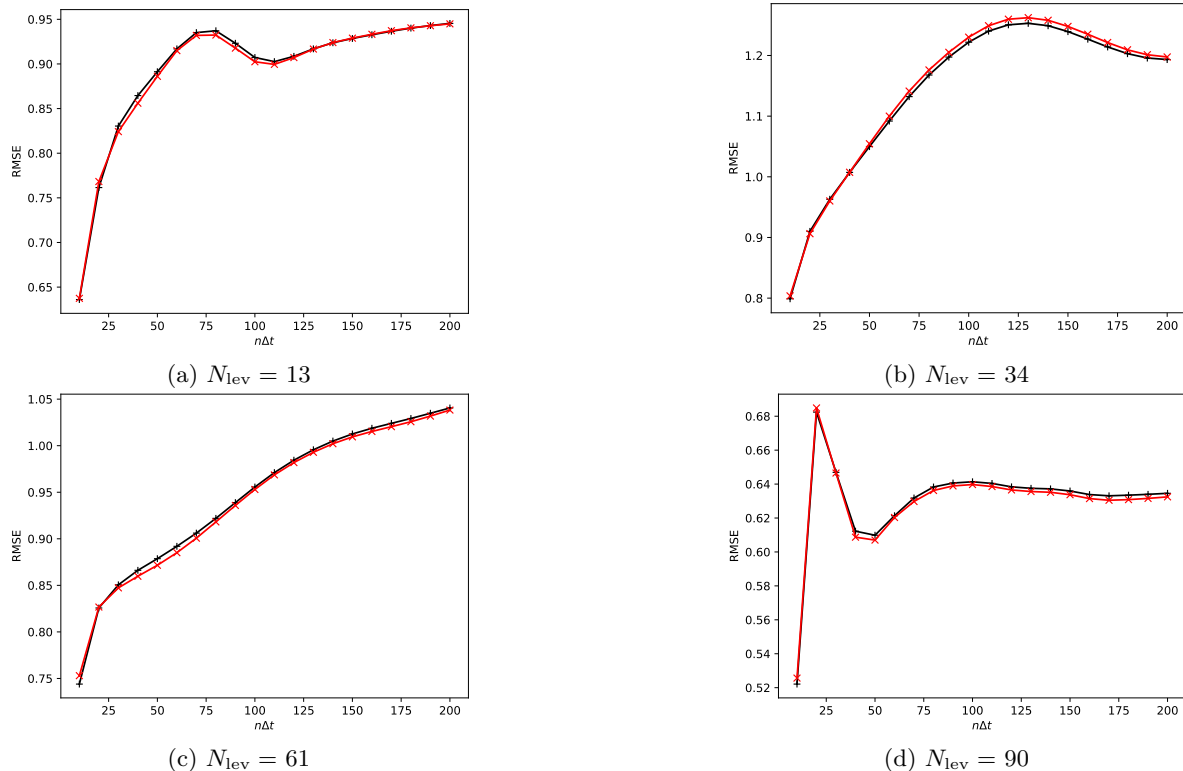


FIGURE 3.7 – EQM (par rapport à la version spectrale) pour la composante méridienne du vent à 4 niveaux verticaux différents, en rouge (avec une croix x) pour  $\gamma = 10$ , en noir (avec un +) pour  $\gamma = 1$ .

le problème implicite, on a souligné le lien étroit entre scalabilité et vitesse de convergence. Les questions relatives au filtrage spatial et à la diffusion numérique n'ont en revanche pas été abordées.

En utilisant plusieurs modèles portés en points de grille, soit intégralement (comme le modèle AROME 2D en plan vertical), soit partiellement (comme le modèle AROME 3D sur petit et grand domaine), nous avons souligné l'importance de régler le plus précisément possible le critère d'arrêt et l'initialisation du solveur de Krylov pour accélérer sa vitesse de convergence, sans détériorer la qualité de la prévision. En l'absence de réglage, le solveur fait davantage d'itérations, sans que la qualité de la prévision en soit pour autant améliorée. Les résultats montrent qu'une convergence très rapide peut être atteinte à une qualité acceptable, même sans préconditionnement, à condition de traiter séquentiellement les problèmes projetés dans l'espace propre de l'opérateur vertical du modèle.

Ainsi, dans le modèle 2D en plan vertical, nous avons montré qu'il était possible d'atteindre une efficacité comparable, voire supérieure, à celle des schémas split-explicit utilisés dans la plupart des modèles HEVI. De même, sur des cas réalistes en 3D, seules 3 à 6 itérations sont en moyenne nécessaires par pas de temps pour inverser le problème implicite, sans dégrader significativement la qualité de la solution par rapport à la version actuelle du modèle ou aux observations. Ces résultats demandent toutefois à être confirmés, une fois que l'intégralité du modèle aura été portée dans la version points de grille et que tous les opérateurs seront

### 3.7. CONCLUSION

---

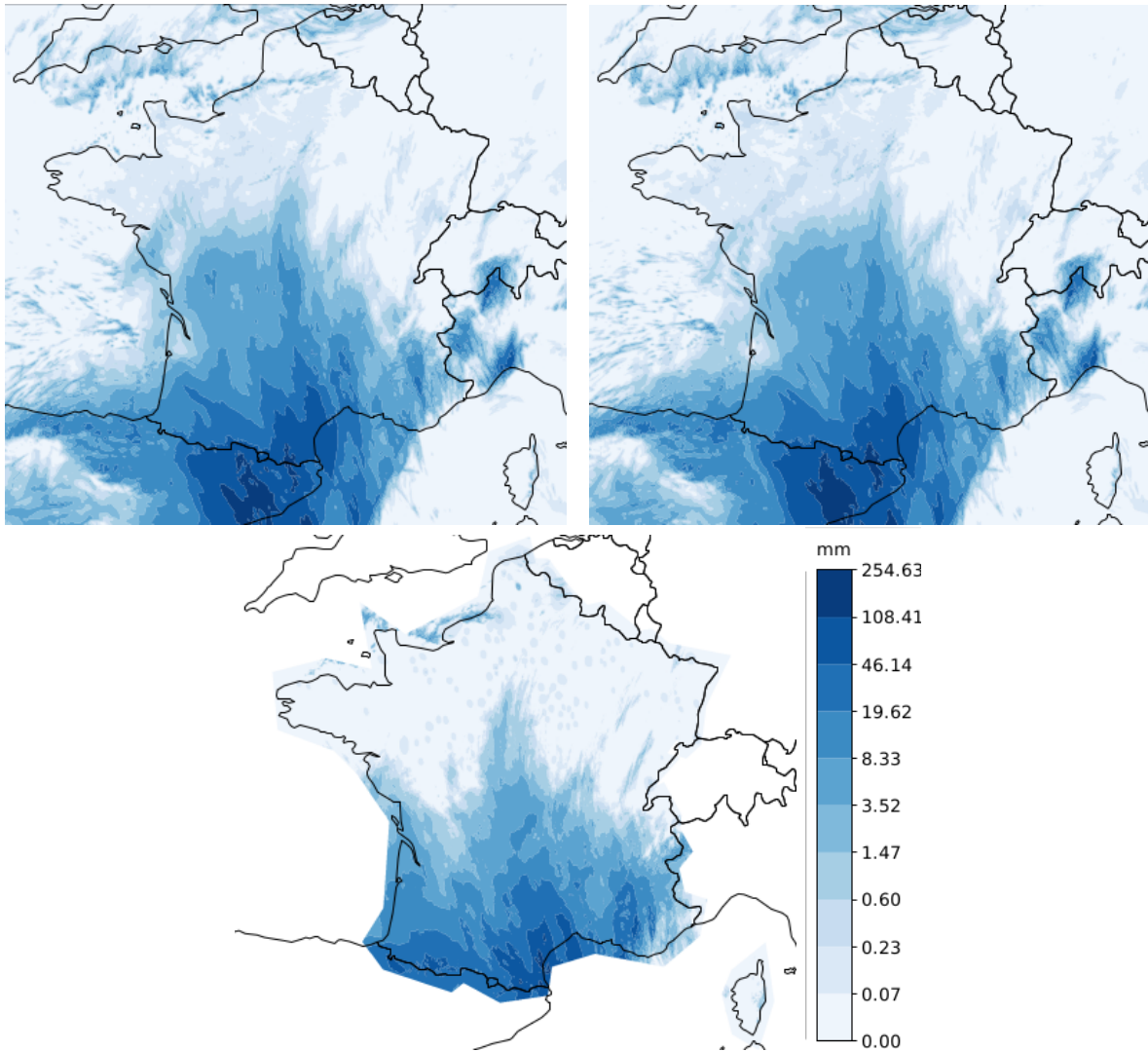


FIGURE 3.8 – Cumul de précipitations en *mm* du 22/10/2019 à 0h au 23/10/2019 à 0h pour la version PG du modèle AROME (en haut à gauche), la version SP du modèle AROME (en haut à droite) et la lame d'eau Antilope (en bas).

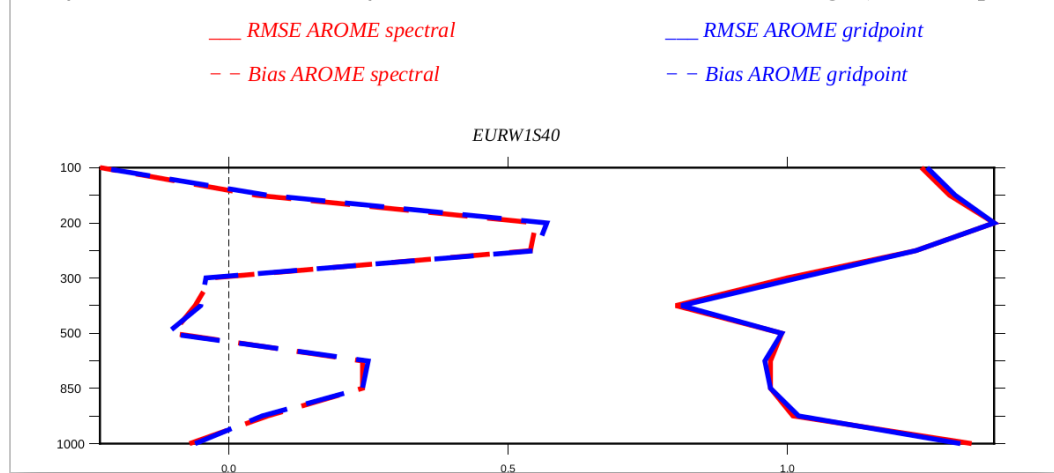
discrétisés de façon cohérente entre eux.

Le point central de ce chapitre se base sur le constat que le coût des communications d'un solveur de Krylov dépend de sa vitesse de convergence. Plus la vitesse de convergence est rapide, plus le coût en communication est faible, et donc plus le modèle est scalable. Par conséquent, dans cette étude, on a réduit l'estimation de la scalabilité à une estimation d'efficacité du solveur. La vitesse de convergence du solveur étant particulièrement rapide, une bonne scalabilité de la version points de grille est suggérée.

Toutefois, une partie seulement des communications ont été considérées (les communications dues au schéma de transport semi-Lagrangien n'ont pas été comptabilisées par exemple)



24h-forecast AROME scores for TEMPERATURE to radiosoundings (1 month period)



24h-forecast scores for WIND compared to radiosoundings (1-month period)

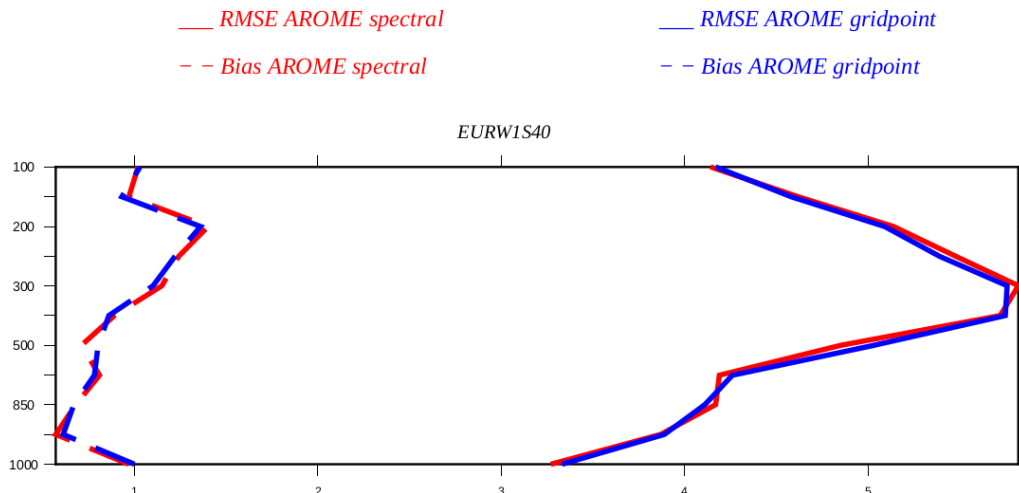


FIGURE 3.9 – Scores des versions SP et PG par rapport aux radiosondages pour une prévision de température (en haut) et de vent (en bas) à 24 heures. L'axe des ordonnées indique l'altitude exprimée en  $hPa$  et l'axe des abscisses l'erreur quadratique moyenne (EQM) et le biais.

et certaines entre-elles ont été négligées (les communications all-reduce lors du calcul des produits scalaires par exemple). De plus, en l'absence de modèle complet entièrement porté en points de grille, aucune mesure directe de scalabilité n'a pu être menée. Nous avons montré que l'alternative consistant à ne mesurer la scalabilité que sur seulement quelques parties présentait également des limites. Par exemple, isoler le solveur du reste du modèle empêche de régler correctement le critère d'arrêt et d'initialiser le solveur, influant la vitesse de convergence et

### 3.7. CONCLUSION

---

donc la scalabilité.

Les résultats encourageants exposés dans ce chapitre, militent pour porter l'intégralité du modèle AROME en points de grille et disposer d'un modèle où tous les opérateurs de dérivées horizontales seraient désormais discrétisés en différences finies, et le filtrage opéré en points de grille. La scalabilité de cette nouvelle version pourrait ainsi être directement mesurée et comparée à la version spectrale actuelle. Si, comme le suggère cette étude, la scalabilité était effectivement améliorée sans perte de qualité par rapport à la version spectrale, la version points de grille pourrait la remplacer dans un contexte opérationnel.



# Fortes pentes

---

## Sommaire

---

<b>4.1 Nouvelles équations du système linéaire à coefficients variables . . . .</b>	<b>114</b>
4.1.1 Termes orographiques . . . . .	114
4.1.2 Équations en coordonnée $\eta$ . . . . .	114
4.1.3 Équations en coordonnée $\sigma$ . . . . .	115
4.1.4 Discrétisation spatiale . . . . .	117
4.1.5 Discrétisation temporelle . . . . .	118
<b>4.2 Méthodologie d'estimation de la stabilité numérique . . . . .</b>	<b>119</b>
4.2.1 Difficultés . . . . .	120
4.2.2 Analyse de stabilité . . . . .	120
4.2.3 Expériences numériques dans un cadre 2D . . . . .	123
4.2.4 En bref . . . . .	124
<b>4.3 Article : analyse de différents schémas en présence de fortes pentes .</b>	<b>125</b>
<b>4.4 Expériences numériques dans un cadre 2D . . . . .</b>	<b>141</b>
4.4.1 Configuration choisie . . . . .	141
4.4.2 Écoulement sur relief idéalisé . . . . .	141
4.4.3 Écoulement faiblement non-linéaire . . . . .	142
4.4.4 Écoulement non-linéaire . . . . .	145
4.4.5 Écoulement fortement non-linéaire sur relief réaliste . . . . .	146
4.4.6 Estimation de la vitesse de convergence . . . . .	149
<b>4.5 Conclusion . . . . .</b>	<b>150</b>

---

Dans ce chapitre, suite aux solutions proposées dans le paragraphe 2.3.2 pour contourner le problème de stabilité numérique au voisinage des plus fortes pentes, on introduit un nouveau système d'équations linéaires à coefficients variables, contenant des termes orographiques. La méthodologie d'évaluation de la stabilité numérique en présence notamment de fortes pentes est ensuite exposée, pour ce nouveau schéma et pour celui à coefficients constants du chapitre 3. Elle consiste en une analyse de stabilité puis en différentes expériences numériques sur un relief idéalisé puis réaliste. Les résultats obtenus suite à l'analyse de stabilité sont exposés dans un article qui devrait être soumis prochainement et qui constitue la section 4.3. Les résultats obtenus suite aux expériences numériques sont exposés dans la section 4.4.

## 4.1 Nouvelles équations du système linéaire à coefficients variables

Comme mentionné dans le paragraphe 2.3.2, l'une des stratégies d'amélioration de la stabilité numérique au voisinage des fortes pentes est de traiter implicitement les termes orographiques. Cette stratégie nécessite de réécrire un nouveau système d'équations linéaires. Dans cette section, on dérive ce nouveau système en décomposant le vecteur d'état  $\Psi$  en un état de base  $\Psi^*$  et une perturbation  $\Psi'$ . L'état de base est identique à celui de la version à coefficients constants (H1)–(H4), sauf que l'hypothèse d'un terrain plat (H5) est relâchée.

### 4.1.1 Termes orographiques

L'orographie est introduite dans l'état de base par le géopotentiel de surface défini comme :

$$\phi_s^* = \phi_s = gz_s \quad (4.1)$$

où  $z_s$  est l'orographie. L'état de base étant à l'équilibre hydrostatique (H4) et isotherme (H1), on en déduit la relation :

$$m^* = \frac{\partial \pi^*}{\partial \eta} = -\frac{\pi^*}{R_d T^*} \frac{\partial \phi^*}{\partial \eta} \quad (4.2)$$

En considérant que la pression hydrostatique  $\pi^*$  ne s'annule pas et en intégrant entre la surface et le niveau de la mer ( $z_s = 0$ ), qui est à la pression  $\pi_0$  que l'on considère comme une constante fixée, on obtient :

$$\pi_s^* = \pi_0 \exp\left(-\frac{\phi_s}{R_d T^*}\right) \quad (4.3)$$

### 4.1.2 Équations en coordonnée $\eta$

Lorsque les termes orographiques sont inclus dans l'opérateur linéaire, l'équation de la vorticit   a un membre de droite non nul, contrairement    l'  quation (2.75) pour un sch  ma    coefficients constants. Par cons  quent,   crire les   quations en utilisant la divergence  $D$  et la vorticit    $\zeta$  comme variables pronostiques n'apporte pas de b  n  fices par rapport    une formulation plus classique sur le vecteur vent  $\mathbf{U}$  comprenant les composantes zonale et m  ridienne du vent  $U$  et  $V$ . Dor  navant, les   quations seront   crites avec ces derni  res variables pronostiques. On suppose de plus, que le terme crois   X (2.18) est trait   dans le sch  ma de transport semi-Lagrangien. Les perturbations par rapport    ce nouvel   tat de base s'  crivent alors comme :

$$\begin{aligned} \frac{\partial \mathbf{U}'}{\partial t} = & -R_d T^* \nabla \hat{q}' - \frac{R_d T^*}{\pi_s^*} \nabla \pi_s' - R_d \int_{\eta}^1 \frac{m^*}{\pi^*} \nabla T' d\eta + R_d T^* \int_{\eta}^1 \frac{m^*}{\pi^*} \nabla \hat{q}' d\eta + \frac{R_d T^*}{\pi_s^{*2}} \nabla \pi_s^* \pi_s' \\ & + \frac{1}{T^*} \nabla \phi^* T' - \nabla \phi^* \hat{q}' - \frac{\pi^*}{m^*} \nabla \phi^* \frac{\partial \hat{q}'}{\partial \eta} + R_d T^* \int_{\eta}^1 \nabla \left( \frac{m^*}{\pi^*} \right) \hat{q}' d\eta - R_d \int_{\eta}^1 \nabla \left( \frac{m^*}{\pi^*} \right) T' d\eta \end{aligned}$$

#### 4.1. NOUVELLES ÉQUATIONS DU SYSTÈME LINÉAIRE À COEFFICIENTS VARIABLES

---

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{rH^*} \tilde{\delta}(\tilde{\delta} + 1) \hat{q}'$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (\nabla \cdot \mathbf{U}' + \mathbb{D}')$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (\nabla \cdot \mathbf{U}' + \mathbb{D}') + \frac{1}{\pi^*} \int_0^\eta m^* \nabla \cdot \mathbf{U}' d\eta - \frac{\mathbf{U}'}{\pi^*} \cdot \nabla \pi^* + \frac{1}{\pi^*} \int_0^\eta \mathbf{U}' \cdot \nabla m^* d\eta$$

$$\frac{\partial \pi'_s}{\partial t} = -\int_0^1 m^* \nabla \cdot \mathbf{U}' d\eta - \int_0^1 \mathbf{U}' \cdot \nabla m^* d\eta$$

Les nouveaux termes supplémentaires par rapport aux équations (2.40)–(2.44) sont ajoutés en bleu. Quand  $z_s = 0$ , tous ces termes s'annulent et on retrouve le système à coefficients constants pour une grille A.

#### 4.1.3 Équations en coordonnée $\sigma$

Par la suite, on se restreint au cas d'une coordonnée  $\sigma$ , c'est-à-dire telle que  $A(\sigma) = 0$  et  $B(\sigma) = \sigma$ . Les métriques deviennent alors :

$$\pi = B(\sigma) \pi_s(x, y) = \sigma \pi_s(x, y) \quad (4.4)$$

$$m = \frac{\partial B(\sigma)}{\partial \sigma} \pi_s(x, y) = \pi_s(x, y) \quad (4.5)$$

Dans ce cas le ratio  $\pi/m$ , ne dépend plus de la coordonnée verticale :

$$\frac{\pi}{m} = \frac{B(\sigma)}{\partial_\sigma B(\sigma)} = \sigma. \quad (4.6)$$

Cette propriété sera exploitée dans la suite pour simplifier la résolution des équations. On note également que le gradient horizontal de ce ratio est nul, ce qui annule certains termes des équations précédemment exposées :

$$\nabla \left( \frac{m^*}{\pi^*} \right) = 0. \quad (4.7)$$

En coordonnée  $\sigma$ , les relations suivantes sont déduites pour l'état de base :

$$\begin{aligned} \pi^* &= \sigma \pi_s^* \\ m^* &= \pi_s^* \\ \nabla \phi^* &= -\frac{RT^*}{\pi^*} \nabla \pi^* = -\frac{RT^*}{\pi_s^*} \nabla \pi_s^* \end{aligned}$$

D'où :

$$\begin{aligned}\nabla\phi^* &= g\nabla z_s = gG \\ \frac{1}{\pi^*}\nabla\pi^* &= \frac{1}{\pi_s^*}\nabla\pi_s^* = -\frac{G}{H_*}\end{aligned}$$

Les opérateurs verticaux (2.30)–(2.34) en coordonnée  $\sigma$  se réécrivent comme :

$$\mathcal{G}^*\Psi = \int_{\sigma}^1 (\Psi/\sigma')d\sigma' \quad (4.8)$$

$$\mathcal{S}^*\Psi = (1/\sigma) \int_0^{\sigma} \Psi d\sigma' \quad (4.9)$$

$$\mathcal{N}^*\Psi = \int_0^1 \Psi d\sigma \quad (4.10)$$

$$\tilde{\partial}\Psi = \sigma \frac{\partial\Psi}{\partial\sigma} \quad (4.11)$$

Les opérateurs ne sont pas tous commutatifs et respectent les propriétés suivantes :

$$\tilde{\partial}\mathcal{G}^*\Psi = -\Psi \quad (4.12)$$

$$\mathcal{G}^*\tilde{\partial} = \Psi(\sigma = 1) - \Psi \quad (4.13)$$

$$(\tilde{\partial} + 1)\mathcal{S}^*\Psi = \mathcal{S}^*(\tilde{\partial} + 1)\Psi = \Psi \quad (4.14)$$

$$\nabla\mathcal{G}^*\Psi = \mathcal{G}^*\nabla\Psi \quad (4.15)$$

$$\mathcal{S}^*G\Psi = G\mathcal{S}^*\Psi \quad (4.16)$$

Par la suite, nous utiliserons les notations suivantes :

$$U_1 = \frac{R_d T^*}{\pi_s^{*2}} \nabla \pi_s^* \pi_s' = -\frac{R_d T^*}{\pi_s^*} \frac{G}{H_*} \pi_s' \quad (4.17)$$

$$U_2 = -\frac{R_d}{\pi^*} \nabla \pi^* T' = R_d \frac{G}{H_*} T' \quad (4.18)$$

$$U_3 = -\nabla \phi^* \hat{q}' = -R_d T^* \frac{G}{H_*} \hat{q}' \quad (4.19)$$

$$U_4 = -\frac{\pi^*}{m^*} \nabla \phi^* \frac{\partial \hat{q}'}{\partial \eta} = -R_d T^* \frac{G}{H_*} \tilde{\partial} \hat{q}' \quad (4.20)$$

$$Q_1 = -\frac{1}{\pi^*} \frac{\partial \pi^*}{\partial x} U' = \frac{G}{H_*} U' \quad (4.21)$$

$$Q_2 = \frac{1}{\pi^*} \int_0^{\sigma} U' \frac{\partial m^*}{\partial x} d\sigma' = -\frac{G}{H_*} \mathcal{S}^* U' \quad (4.22)$$

$$\Pi_S = -\int_0^1 U' \frac{\partial m^*}{\partial x} d\sigma = \pi_s^* \frac{G}{H_*} \mathcal{N}^* U' \quad (4.23)$$

## 4.1. NOUVELLES ÉQUATIONS DU SYSTÈME LINÉAIRE À COEFFICIENTS VARIABLES

---

### Équations en 3D

En réécrivant les équations précédentes en coordonnée  $\sigma$  et en utilisant les notations précédentes, on obtient :

$$\frac{\partial \mathbf{U}'}{\partial t} = R_d \left( -\mathcal{G}^* \nabla + \frac{G}{H_*} \right) T' + R_d T^* \left( \mathcal{G}^* \nabla - \frac{G}{H_*} - \nabla - \frac{G}{H_*} \tilde{\partial} \right) \hat{q}' - \frac{R_d T^*}{\pi_s^*} \left( \nabla + \frac{G}{H_*} \right) \pi_s' \quad (4.24)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{r H_*} \tilde{\partial} (\tilde{\partial} + 1) \hat{q}' \quad (4.25)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (\nabla \cdot \mathbf{U}' + \mathbb{D}') \quad (4.26)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (\nabla \cdot \mathbf{U}' + \mathbb{D}') + \mathcal{S}^* \nabla \cdot \mathbf{U}' + \frac{G}{H_*} \cdot \mathbf{U}' - \frac{G}{H_*} \cdot \mathcal{S}^* \mathbf{U}' \quad (4.27)$$

$$\frac{\partial \pi_s'}{\partial t} = -\pi_s^* \mathcal{N}^* \nabla \cdot \mathbf{U}' + \pi_s^* \frac{G}{H_*} \cdot \mathcal{N}^* \mathbf{U}' \quad (4.28)$$

### Équations en 2D

Par la suite, on se restreint à des écoulements en seulement deux dimensions (sur un plan vertical). L'opérateur  $\nabla$  est désormais remplacé par  $\partial_x$  et le vecteur  $\mathbf{U}$  par la seule composante  $U$ . Les équations (4.24)–(4.28) deviennent alors :

$$\frac{\partial U'}{\partial t} = R_d \left( -\mathcal{G}^* \partial_x + \frac{G}{H_*} \right) T' + R_d T^* \left( \mathcal{G}^* \partial_x - \frac{G}{H_*} - \partial_x - \frac{G}{H_*} \tilde{\partial} \right) \hat{q}' - \frac{R_d T^*}{\pi_s^*} \left( \partial_x + \frac{G}{H_*} \right) \pi_s' \quad (4.29)$$

$$\frac{\partial \mathbb{D}'}{\partial t} = -\frac{g}{r H_*} \tilde{\partial} (\tilde{\partial} + 1) \hat{q}' \quad (4.30)$$

$$\frac{\partial T'}{\partial t} = -\frac{R_d T^*}{C_{vd}} (\partial_x U' + \mathbb{D}') \quad (4.31)$$

$$\frac{\partial \hat{q}'}{\partial t} = -\frac{C_{pd}}{C_{vd}} (\partial_x U' + \mathbb{D}') + \mathcal{S}^* \partial_x U' + \frac{G}{H_*} U' - \frac{G}{H_*} \mathcal{S}^* U' \quad (4.32)$$

$$\frac{\partial \pi_s'}{\partial t} = -\pi_s^* \mathcal{N}^* \partial_x U' + \pi_s^* \frac{G}{H_*} \mathcal{N}^* U' \quad (4.33)$$

#### 4.1.4 Discrétisation spatiale

Les opérateurs verticaux sont discrétisés en différences finies à l'ordre 2, de façon analogue à la partie précédente. Les équations (4.29)–(4.33) se réécrivent en remplaçant les opérateurs  $\mathcal{G}^*$ ,  $\mathcal{S}^*$ ,  $\mathcal{N}^*$  et  $\mathcal{L}_v^*$  par leurs versions discrétisées verticalement :  $\mathbf{G}^*$ ,  $\mathbf{S}^*$ ,  $\mathbf{N}^*$  et  $\mathbf{L}_v^*$ .

Les opérateurs de dérivées horizontales sont discrétisés en différences finies comme men-



tionné dans le paragraphe 3.1.2 et sont désormais symbolisés par  $\delta_x$  au lieu de  $\partial_x$ .

#### 4.1.5 Discrétisation temporelle

La discrétisation est effectuée comme dans la version à coefficients constants actuelle énoncée dans le paragraphe 2.1.9. On réécrit la formulation par blocs (2.84), sauf qu'ici, la composante zonale du vent est considérée à la place de la divergence horizontale du vent comme variable pronostique :

$$\left[ \begin{array}{c|c} \mathcal{I} & -\Delta t/2\mathcal{L}_A \\ \hline -\Delta t/2\mathcal{L}_B & \mathcal{I} - \Delta t/2\mathcal{L}_C \end{array} \right] \left[ \begin{array}{c} U^{+(n)} \\ \psi^{+(n)} \end{array} \right] = \left[ \begin{array}{c} U^\bullet \\ \psi^\bullet \end{array} \right]. \quad (4.34)$$

où  $\mathcal{I}$  désigne la matrice identité,  $\Psi = [U, \psi]^T$ ,  $\Psi^\bullet = [U^\bullet, \psi^\bullet]^T$  et  $\psi = [\mathbb{D}, T, \hat{q}, \pi_s]^T$ . Les opérateurs  $\mathcal{L}_A$ ,  $\mathcal{L}_B$  et  $\mathcal{L}_C$  s'écrivent désormais comme :

$$\mathcal{L}_A = \left[ 0 \left| R_d \left( -\mathbf{G}^* \delta_x + \frac{G}{H_*} \right) \right| R_d T^* \left( \mathbf{G}^* \delta_x - \frac{G}{H_*} - \delta_x - \frac{G}{H_*} \tilde{\delta} \right) \left| -\frac{R_d T^*}{\pi_s^*} \left( \delta_x + \frac{G}{H_*} \right) \right. \right], \quad (4.35)$$

$$\mathcal{L}_B = \left[ \begin{array}{c} 0 \\ \hline -R_d T^* / C_{vd} \delta_x \\ \hline \mathbf{S}^* \delta_x - C_{pd} / C_{vd} \delta_x + G / H_* - G / H_* \mathbf{S}^* \\ \hline -\pi_s^* \mathbf{N}^* \delta_x + \pi_s^* G / H_* \mathbf{N}^* \end{array} \right] \quad (4.36)$$

$$\mathcal{L}_C = \left[ \begin{array}{c|c|c|c} 0 & 0 & -g^2 / (R_d T_e^*) \mathbf{L}_v^* & 0 \\ \hline -R_d T^* / C_{vd} & 0 & 0 & 0 \\ \hline -C_{pd} / C_{vd} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \quad (4.37)$$

Comme précédemment pour l'équation (2.90), le problème est ensuite réduit algébriquement, cette fois-ci sur la vitesse horizontale du vent  $U$  (au lieu de la divergence horizontale du vent  $D$ ) :

$$\left[ \mathcal{I} - \frac{\Delta t^2}{4} \mathcal{B} \right] U^{+(n)} = U^{\bullet\bullet}, \quad (4.38)$$

où :

$$U^{\bullet\bullet} = U^\bullet - \frac{\Delta t}{2} \mathcal{L}_A \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)^{-1} \psi^\bullet, \quad (4.39)$$

$$\mathcal{B}' = \mathcal{L}_A \left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)^{-1} \mathcal{L}_B. \quad (4.40)$$

Notons que cette fois-ci le problème n'est pas séparable : c'est-à-dire ne peut pas être décomposé comme (2.85) et (2.86). Par conséquent, le problème ne peut être facilement projeté dans l'espace propre d'un opérateur vertical ( $\mathcal{B}' \neq \mathcal{B} \circ \Delta$ ).

Dans le cas particulier de la coordonnée  $\sigma$ , l'opérateur  $\left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)$  est un opérateur purement vertical, inversé par une méthode directe, peu coûteuse en raison de la faible taille du problème. En revanche, en coordonnée  $\eta$ , cet opérateur dépend également de la coordonnée horizontale. Un second solveur itératif est alors nécessaire pour l'inverser, en plus de celui résolvant l'équation (4.38). Ce problème n'est pas étudié dans cette thèse et est laissé en perspective.

De façon analogue, lorsque le terme croisé  $X$  est traité dans l'opérateur linéaire en présence des termes orographiques, l'opérateur  $\left( 1 - \frac{\Delta t}{2} \mathcal{L}_C \right)$  dépend également de la coordonnée horizontale, et deux solveurs sont alors nécessaires pour inverser le problème (non montré). Les équations du système linéaire exposées jusqu'ici, sont écrites lorsque ce terme est traité dans le schéma de transport semi-Lagrangien et non dans l'opérateur implicite. Dans la section 4.3, une option permettra de traiter ce terme dans l'opérateur implicite, dans le cas particulier d'un système non borné.

## 4.2 Méthodologie d'estimation de la stabilité numérique

Dans cette étude, la stabilité numérique est tout d'abord évaluée via une analyse pour :

- la version à coefficients constants (notée  $\delta = 0$ ) en points de grille du chapitre 3 en : diminuant le pas de temps, augmentant le nombre d'étapes du schéma ICI, diminuant la température froide  $T_e^*$  ;
- la version à coefficients variables (notée  $\delta = 1$ ) en points de grille de la section 4.1.

De plus, pour chacune de ces deux versions, la stabilité numérique est évaluée lorsque le terme croisé  $X$  est traité par le schéma de transport semi-lagrangien (notée  $\delta_{SL} = 1$ ), ou dans l'opérateur linéaire (notée  $\delta_{SL} = 0$ ).

Mener une analyse de stabilité dans un cadre général est hors de portée. Pour que cette analyse soit menée avec des outils relativement simples, de fortes hypothèses sont établies et sont exposées dans le prochain paragraphe. Les estimations de stabilité sont données pour chacune des deux versions dans un contexte idéalisé, dont la portée des conclusions est par conséquent limitée mais est informative pour guider les futurs développements qui, eux, nécessitent un travail plus conséquent. Une fois l'analyse menée, les options les plus prometteuses pour améliorer la stabilité numérique sont implémentées dans la version 2D du modèle AROME, déjà utilisée dans le chapitre précédent. Des expériences numériques en présence d'un relief idéalisé puis réaliste sont ensuite menées en choisissant des cas tests pertinents pour analyser la stabilité numérique en présence de fortes pentes.

### 4.2.1 Difficultés

Dans cette section, nous visons à accroître la stabilité numérique en présence de fortes pentes (L1). Toutefois, comme mentionné dans le paragraphe 1.4.9, cela ne doit se faire ni au détriment de l'inversibilité du problème implicite (L2), ni au détriment de sa résolution efficace (L3) en peu d'étapes du schéma ICI (L4). Dans la suite nous veillerons à maximiser la stabilité numérique sans dégrader ces trois contraintes.

De plus, l'un des objectifs d'un modèle de PNT est de prévoir l'évolution spatio-temporelle des perturbations physiques de l'atmosphère. Nous veillerons également à ce qu'une instabilité physique ne soit pas attribuée à une instabilité numérique et réciproquement. Pour éviter cette confusion, les expériences ou analyses sont généralement menées en considérant des systèmes stables physiquement. La présence d'une instabilité est alors directement imputable à la partie numérique. Toutefois, cette démarche restreint fortement la classe des écoulements étudiés. Pour considérer une plus grande variété d'écoulements, nous étudierions dans cette thèse, des systèmes physiquement instables. On considèrera qu'un système est stable numériquement si son taux de croissance suite à l'étape de numérisation est inférieur à son taux de croissance physique.

Enfin, les fortes pentes génèrent de fortes contributions des termes orographiques  $G$ , et donc des perturbations de fortes amplitudes. Cela conduit à étudier des écoulements non-linéaires, plus difficiles à traiter que les écoulements linéaires.

### 4.2.2 Analyse de stabilité

#### Hypothèses simplificatrices

Mener des analyses de stabilité dans un cadre général est hors de portée. L'intérêt principal des analyses réside dans leur simplicité. Ainsi, pour cette étude, on considère notamment un système sans paramétrisation physique et tel que l'opérateur non-linéaire est remplacé par son tangent-linéaire ( $\mathcal{P} = 0$  et  $\mathcal{M} = \bar{\mathcal{L}}$  dans l'équation (1.43)).

Les premières analyses de stabilité pour les schémas semi-implicites ont été introduites par Simmons *et al.* (1978) pour un système d'équations filtrées, puis étendues par Bénard (2003) pour les équations d'Euler pleinement compressibles en coordonnée masse. On considère dans notre étude un système  $\bar{\mathcal{L}}$  non discrétisé tel que son état de base associé  $\bar{\Psi}$  est : isotherme de température  $\bar{T}$ , sec, hydrostatique, au repos, et de pente constante  $G$ . Sous ces hypothèses, le système à résoudre est de la forme :

$$\frac{\partial l\Psi}{\partial t} = l\mathcal{L}^*(\Psi) + l(\bar{\mathcal{L}} - \mathcal{L}^*)(\Psi) \quad (4.41)$$

où  $l$  est un opérateur transformant les opérateurs  $\mathcal{L}^*$  et  $\bar{\mathcal{L}}$  en des opérateurs non bornés spatialement (comme dans Bénard *et al.* (2005)). Les équations (4.24)–(4.28), se simplifient sous ces hypothèses et seront présentées dans l'article (section 4.3).

### Opérateur linéaire du problème implicite

L'opérateur linéaire du problème implicite  $\mathcal{L}^*$  est choisi librement, comme montré dans le chapitre 2, et n'est pas nécessairement le tangent linéaire. Dans cette étude, il ne diffère de  $\bar{\mathcal{L}}$  que par :

- la température de référence  $T^*$ , différente de  $\bar{T}$  et dont l'écart est mesuré synthétiquement par le facteur  $\alpha$  :

$$\alpha = \frac{\bar{T} - T^*}{T^*}, \quad (4.42)$$

- la température de référence pour l'équation de la quantité de mouvement verticale  $T_e^*$  dont le ratio est donné par :

$$r = \frac{T_e^*}{T^*}, \quad (4.43)$$

- l'absence éventuelle des termes orographiques (si  $\delta = 0$ ).

### Domaine de stabilité

Dans cette analyse, le résidu dit « non linéaire » contient des termes orographiques (si  $\delta = 0$ ) et un résidu thermique (le facteur  $\alpha$ ). Les analyses de stabilité sont menées pour différentes valeurs de résidu thermique  $\alpha$  et de pente  $G$ . On appelle « domaine de stabilité », les valeurs de  $G$  et de  $\alpha$  telles que le taux de croissance des instabilités numériques est inférieur au taux de croissance des instabilités physiques.

### Autres contraintes

Comme nous l'exposerons en détail dans l'article, le calcul des taux de croissance repose sur des calculs de valeurs propres des différents opérateurs. Ces calculs peuvent également servir à vérifier que les contraintes (L2)–(L4) restent satisfaites. Pour vérifier que le problème est inversible (L2), on vérifiera sous quelles conditions aucune valeur propre du problème implicite n'est nulle. Pour vérifier que la résolution du problème implicite reste efficace (L3), on vérifiera que le conditionnement de son opérateur n'est pas trop dégradé. Enfin, le nombre d'étapes du schéma ICI (L4) est un paramètre de cette étude.

### Discussion des hypothèses

Parmi les nombreuses hypothèses mentionnées ci-dessus, quatre d'entre-elles retiennent plus particulièrement notre attention :

1. considérer un état de base  $\bar{\Psi}$  au repos, c'est-à-dire sans vent ;
2. considérer un relief purement linéaire, alors que dans la réalité il est beaucoup plus chahuté ;

3. réduire l'opérateur linéaire  $\mathcal{M}$  à son simple tangent linéaire ;
4. considérer un système non borné.

La plupart des quatre hypothèses mentionnées ci-dessus sur-estiment le domaine de stabilité numérique. Il est par exemple connu que l'absence de vent stabilise numériquement le problème. Par conséquent, la première hypothèse sur-estime probablement la stabilité numérique en particulier quand le terme croisé  $X$  est traité dans le schéma de transport semi-Lagrangien (si  $\delta_{SL} = 1$ ) en l'absence de vent.

La seconde hypothèse implique de traiter des termes orographiques constants, permettant ainsi de mener l'analyse de stabilité sans calculer les coûteux produits de convolution entre les termes orographiques et les variables pronostiques (comme dans (1.56)). Cette approche permet finalement de conserver une approche à coefficients constants dans cette analyse, même en traitant les termes orographiques implicitement. Cette hypothèse sur-estime probablement la stabilité car les interactions non-linéaires entre les différentes longueurs d'onde dues à la présence du relief ne sont pas prises en compte dans le cas d'un relief linéaire.

La troisième hypothèse consiste à traiter le résidu non-linéaire par des opérateurs linéaires : l'opérateur non linéaire  $\mathcal{M}$  est remplacé par son tangent linéaire  $\tilde{\mathcal{L}}$ , linéarisé autour d'un état  $\bar{\Psi}$  très simple, assez éloigné de la réalité (par exemple l'atmosphère est loin d'être isotherme dans la réalité, contrairement à ce qui est supposé dans l'analyse).

Enfin, en ne prenant pas en compte les conditions aux limites dans l'analyse, la quatrième hypothèse aboutit à calculer des modes différents entre les systèmes bornés et non bornés. Staniforth et Wood (2005), le mettent en évidence dans le cas d'un système 1D acoustique : le système non borné admet deux modes acoustiques propagatifs alors que le système borné n'en admet qu'un seul non propagatif en complément d'un mode computationnel (dans le cas continu spatialement) ou d'un mode qui s'y apparente (dans le cas discret spatialement). Cela aboutit à calculer des taux de croissance différents entre les deux systèmes (borné et non borné). Relâcher cette quatrième hypothèse conduirait à complexifier l'analyse dans le cas où le système d'équation ne serait plus un simple système acoustique 1D mais les équations d'Euler pleinement compressibles, nous éloignant ainsi de l'objectif de mener une analyse dans un cadre simple.

Ces analyses bien que reposant sur des hypothèses fortes, ont guidé vers des choix qui se sont avérés tout à fait viables lorsqu'ils ont finalement été implémentés dans un modèle : la variable pronostique dérivée  $\mathbb{D}$  est désormais utilisée à la place de  $d$  dans le modèle opérationnel AROME, conformément aux résultats des analyses menées par Bénard *et al.* (2005) pour un système non borné. Par conséquent, les analyses dans cette étude seront menées suivant toutes ces hypothèses. Conscients des limites fortes qu'elles imposent, les analyses de stabilité ne nous permettront pas d'évaluer exactement la stabilité du système mais de guider le choix des futurs développements pour traiter le problème des fortes pentes.

### 4.2.3 Expériences numériques dans un cadre 2D

Les analyses de stabilité exposées dans la partie précédente nous permettront tout d'abord d'identifier les causes de ces instabilités numériques et des stratégies vers lesquelles il semble plus viable de se diriger pour aborder le problème des fortes pentes et guider ainsi les futurs développements. Les analyses de stabilité reposent sur des hypothèses fortes, que l'on souhaitera remettre ensuite en question, en considérant un cadre plus réaliste. Toutefois, le cadre AROME 3D reste hors de portée tant les modifications à effectuer peuvent être importantes. En revanche le cadre AROME 2D, en l'absence de paramétrisation physique ( $\mathcal{F} = 0$ ) est un cadre adapté pour mener des expériences de stabilité dans un contexte plus réaliste et assez peu coûteux en développements.

#### Hypothèses

Comme évoqué précédemment, le remplacement de la diffusion numérique implicite appliquée dans l'espace spectral, n'a pas encore été étudié, par un équivalent en points de grille. En plus d'assurer son rôle de cascade d'énergie entre les paramétrisations physiques et la partie dynamique, la diffusion peut masquer ou retarder l'apparition de certaines instabilités numériques. Or l'objectif de cette partie est d'évaluer précisément ces instabilités numériques sans chercher à les masquer ou à les retarder. Hormis la diffusion numérique intrinsèque au schéma semi-Lagrangien intervenant lors du calcul des rétro-trajectoires, aucune diffusion numérique n'est ajoutée pour les expériences qui suivront.

De plus, restreindre l'étude à des écoulements uniquement 2D génère des écoulements plus turbulents qu'en 3D : le vent incident au relief ne peut le contourner dans la direction transverse.

Enfin, l'absence de paramétrisation physique empêche la dissipation d'énergie via les petits tourbillons provoquant une accumulation d'énergie lorsque des perturbations suffisamment grandes pour qu'elles ne puissent plus être supposées infinitésimales sont considérées. Cette accumulation d'énergie peut aboutir à des déferlements (Durrant, 1990) qui conduisent à une interruption de la simulation, sans pour autant être imputables à la stabilité numérique. L'enjeu dans toute la suite est donc de choisir des expériences permettant d'étudier la stabilité numérique plutôt que la stabilité physique.

#### Domaine de stabilité

Contrairement aux analyses de stabilité, ces trois hypothèses conduisent, au contraire, à sur-estimer très probablement l'instabilité numérique. Autrement dit, si une simulation est stable numériquement dans ce cadre, nous sommes confiants dans le fait qu'elle le sera dans un cadre encore plus réaliste en 3D dans un modèle complet avec des paramétrisations physiques et de la diffusion.

### Choix des expériences

Par conséquent, dans toute la suite, l'enjeu sera tout d'abord de choisir des expériences dans lesquelles la perturbation due à la présence du relief est suffisante pour reproduire le problème numérique en présence de fortes pentes, mais pas trop forte pour ne pas générer de perturbations physiques conduisant à des déferlements en l'absence de mécanismes de diffusion d'énergie. Pour générer des fortes pentes deux leviers peuvent être activés : augmenter la hauteur à une largeur de relief fixée ou diminuer la largeur à une hauteur de relief fixée.

Dans un premier temps, une montagne de faible hauteur mais très étroite sera considérée. La faible hauteur du relief permet de reproduire des écoulements assez laminaires, ne conduisant pas à des déferlements prématurés. Des simulations sur de longues périodes ( $T_{simu}$ ) pourront ainsi être menées. La stabilité numérique de chacune des deux versions (avec ou sans traitement implicite des termes orographiques) sera mesurée par la valeur de la pente maximale telle que la stabilité est maintenue pour un pas de temps  $\Delta t$  donné.

Ensuite, les fortes pentes seront représentées en augmentant la hauteur du relief, s'approchant ainsi des cas plus réalistes. Dans ce cas toutefois, la nature de l'écoulement change en devenant rapidement turbulent. Les instabilités physiques croissent rapidement, conduisant à une interruption plus rapide de la simulation. En l'absence de solution de référence d'une part et de mécanismes de diffusion de l'énergie d'autre part, il n'est pas évident de savoir si la simulation s'interrompt à cause d'une instabilité numérique ou à cause d'une instabilité physique. Toutefois, dans notre étude, une information instructive est de savoir si l'une des deux simulations s'interrompt significativement avant ou après l'autre. Les mécanismes de diffusion n'étant pas plus importants dans l'une ou l'autre des versions, la simulation qui s'interrompt après l'autre est considérée plus stable. Ainsi, même si la stabilité absolue n'est pas estimée dans ce cas, une information sur la stabilité relative entre les deux versions considérées peut être obtenue.

### Autres contraintes

Dans cette partie, un schéma ICI à deux étapes sera utilisé, satisfaisant la condition (L4). Dans le cadre de l'utilisation d'un solveur itératif de Krylov, un problème non inversible se manifeste par une absence de convergence du solveur (nombre infini d'itérations). Ainsi, mesurer le nombre d'itérations du solveur à une précision donnée, permettra de vérifier son inversibilité (L2) d'une part, et sa capacité à converger rapidement (L3) d'autre part.

#### 4.2.4 En bref

Les analyses de stabilité exposées dans la prochaine section :

- sur-estiment les domaines de stabilité,
- sont peu réalistes,

#### 4.3. ARTICLE : ANALYSE DE DIFFÉRENTS SCHÉMAS EN PRÉSENCE DE FORTES PENTES

---

— sont simples à mener.

Les expériences numériques exposées dans la dernière section :

- sous-estiment les domaines de stabilité,
- sont plus réalistes,
- ne donnent qu'une estimation relative de la stabilité.

### **4.3 Article : analyse de différents schémas en présence de fortes pentes**



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

TH. BURGOT <sup>\*†</sup>

L. AUGER

P. BÉNARD

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 class of Iterative Centered Implicit temporal schemes whose approaches the Crank-Nicolson scheme. These analyses are lead for the fully elastic system of Euler equations for different slopes and different non-linear residuals. They are able to reproduce the maximum slopes currently encountered in real hectometric models, for which the simulation is numerically stable, resulting in a strong interest. Because of the negligible price of these analyses, several strategies can thus be easily tested. No strategy among those considered for constant coefficient schemes improves 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 variable coefficient schemes, with the same features of constant coefficient ones, except for orographic terms which are implicitly treated. In this case, slope up to 70° can be easily reached, even in case of a strong non-linear 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:

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

where  $\mathcal{M}$  is a non-linear operator which represents dynamical equations, typically the fully elastic system of Euler equations,  $\Psi$  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 kilometric resolutions. Some of them are HEVI (Horizontally Explicit Vertically Implicit) models like COSMO, 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 scale, 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 of the relief is better represented which leads 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 an unusable 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:

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

where  $n$  is the  $n$ -th step of the ICI scheme,  $\Psi^0$  and  $\Psi^+$  are respectively the state of the system at times  $t$  and  $t + \Delta t$ ,  $\mathcal{L}^*$  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  $\Psi^-$  is the state of the system at time  $t - \Delta t$ ). Most of SI models use a non-extrapolated scheme with two steps ( $N_{iter} = 2$ ) (sometimes

<sup>\*</sup>Météo-France, Toulouse, France

<sup>†</sup>Corresponding author: Thomas Burgot, thomas.burgot@meteo.fr

called predictor-corrector scheme, and abbreviated ICI-TL-PC in the following), while some of them use 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. 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 depend only on the vertical coordinate, the scheme is termed a ‘constant coefficient’ one (Bénard (2003)). However the imposition of 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. Instead of using the traditional vertical velocity  $w$  or the vertical divergence of the wind  $d$ , the gradient of the orography (contained in the horizontal gradient of the geopotential  $\vec{\nabla}\phi$  in two dimensions) is added into the prognostic variable  $\mathbb{D}$ :

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

where:

$$d = -g \frac{p}{mRT} \frac{\partial w}{\partial \eta}. \quad (4)$$

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, and the other notations are specified in the Appendix B. Using a special ‘derived’ prognostic variables leads to introduce other terms, 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)). A height-based coordinate has also been chosen for the new FVM dynamical core, under development, as a potential successor of the mass-based IFS of ECMWF (Kühnlein et al. (2019)). Preliminary results are promising but at the price of several major changes in the code.

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), has been replaced by a non-extrapolated scheme with two steps (ICI-TL-PC) (Brousseau et al. (2016)). However, this strategy can only be applied sparingly to satisfy the condition ‘4.’ above-mentioned.

To avoid too major changes in the code, we explore in this study other simpler ways to improve numerical stability. For example, some of numerical instabilities come from the terms of the non-linear residual (2) which are treated explicitly and are therefore subject to stability constraints. For constant coefficient schemes, orographic terms are treated explicitly and are therefore subject to stability constraints. The first aim of this study is to explore three strategies to reduce their contributions for the constant coefficient schemes (ie isothermal with a specific

‘cold’ temperature for the vertical momentum equation, dry, at rest, 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, resulting in a variable coefficient scheme. However, in contrast to some variable coefficients scheme like the ‘New Dynamics’ dynamical core, this problem is based on a prescribed idealised basic state. This 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. It would only be used if the first three fail. 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 algebra 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, all semi-implicit constant coefficient models do not use a spectral discretization (e.g Qaddouri and Lee (2010)). Those who still use them are anyway generally removing it 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 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: the recent removing of orographic terms in the implicit problem in the new dynamical core of the UM model must be kept in mind. Furthermore, as mentioned in Liesen and Tichý (2004), the convergence rate of a Krylov method depends partly on the condition number of the implicit problem to be inverted. Therefore, before undertaking any coding work, a simple study in which the numerical stability and an evaluation of the cost of the method would be given by studying the operators’ eigenvalues could be useful.

In section 2 the general methodology to test the four above-mentioned strategy is exposed. In section 3 governing equations are introduced under their continuous and temporal-discretized forms in the simple case of the  $\sigma$  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 explicitly. 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 all or some of the orographic terms are treated implicitly, while testing different configurations of the ICI-2TL 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 coordinate. This article is based on Bénard (2003)’s methodology, applied to the different options outlined below.

To carry out these analyses in a simple way, the non-linear operator  $\mathcal{M}$  of (1), is replaced by its tangent linear  $\tilde{\mathcal{L}}$  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 practise, a very simple state is used: isothermal of temperature  $\bar{T}$ , dry, hydrostatic, at rest and on a constant orographic slope. In the following this state will be called the ‘physical’ state. The evolution of the disturbed state vector  $\Psi'$  is thus of the form:

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

For more clarity, the primes notation are now dropped, and  $\Psi$  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:  $\mathcal{R} = \tilde{\mathcal{L}} - \mathcal{L}^*$ .

Then, the references to vertical boundaries are removed like in Bénard (2003). An operator  $l$  is applied, transforming the operators  $\mathcal{L}^*$  and  $\tilde{\mathcal{L}}$  into spatially unbounded

operators. The continuous system (5) becomes:

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

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

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

As mentioned in the introduction and shown by B enard (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 identical to 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 other equations. This state will be called the ‘basic’ or ‘reference’ state in the following. This operator can also contain orographic terms if they are treated implicitly (option  $\delta = 1$ ) or not if they are treated explicitly (option  $\delta = 0$ ). The operators  $\mathcal{L}^*$  and  $\bar{\mathcal{L}}$  thus differ only by:

- the reference state temperature  $T^*$ , different from  $\bar{T}$ , whose deviation is measured for convenience by the dimensionless ‘non-linear’ residual:

$$\alpha = \frac{\bar{T} - T^*}{T^*}, \quad (8)$$

- the reference cold temperature  $T_e^*$  for the vertical momentum equation:

$$r = \frac{T_e^*}{T^*}, \quad (9)$$

- the occurrence of orographic terms if  $\delta = 1$ .

In the next section, the different equations are detailed.

### 3. Governing equations

In this section, bounded operators  $\bar{\mathcal{L}}$ ,  $\mathcal{L}^*$ , and their unbounded counterpart  $l\bar{\mathcal{L}}$ ,  $l\mathcal{L}^*$ , are detailed.

#### a. Mass-based coordinate

The vertical coordinate  $\eta$  is the mass-coordinate developed by Laprise (1992):

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

where  $\pi_s$  is the surface hydrostatic pressure,  $\pi$  is the hydrostatic pressure at level  $\eta$ , and  $A$  and  $B$  are two functions chosen to tune the vertical coordinate. In this study the vertical  $\eta$  coordinate is called the  $\sigma$  coordinate, and in this case:  $A = 0$  and  $B = \sigma$ .

#### b. Bounded tangent-linear operator $\bar{\mathcal{L}}$

In this paragraph, the equations of the system (5) are specified when the different prognostic variables are:

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

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

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

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

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

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

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

where:

$$\bar{\delta}'_x = \partial_x + \frac{G}{\bar{H}} \bar{\delta}, \quad (16)$$

and  $\bar{H} = R\bar{T}/g$  is the characteristic height of the physical state. Operators  $\mathcal{G}$ ,  $\mathcal{S}$ ,  $\mathcal{N}$ ,  $\bar{\delta}$  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:

$$\bar{X} = \frac{G}{\bar{H}} \bar{\delta} U, \quad (17)$$

whose time derivative is:

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

$\delta_{SL}$  is an option to apply a specific treatment to the cross term, and will be discussed in the paragraph d.

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

The equations of the linear operator are:

$$\begin{aligned} \frac{\partial U}{\partial t} = & -R\partial_x'^* \mathcal{G}T + RT^* \partial_x'^* \mathcal{G}\hat{q} - RT^* \partial_x'^* \hat{q} \\ & - \frac{RT^*}{\pi_s^*} \left( \partial_x + \frac{G}{H^*} \right) \pi_s, \end{aligned} \quad (19)$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{rRT^*} \tilde{\partial}(\tilde{\partial} + 1)\hat{q} + (1 - \delta_{SL})\dot{X}^*, \quad (20)$$

$$\frac{\partial T}{\partial t} = -\frac{RT^*}{C_v} (\partial_x U + \mathbb{D}), \quad (21)$$

$$\frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v} (\partial_x U + \mathbb{D}) + \mathcal{S}\partial_x'^* U, \quad (22)$$

$$\frac{\partial \pi_s}{\partial t} = -\pi_s^* \mathcal{N}\partial_x U + \pi_s^* \frac{G}{H} \mathcal{N}U, \quad (23)$$

where :

$$\partial_x'^* = \partial_x + \frac{G}{H^*} \tilde{\partial} \quad \text{if } \delta = 1, \quad (24)$$

$$\partial_x'^* = \partial_x \quad \text{if } \delta = 0, \quad (25)$$

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

$$X^* = \frac{G}{H^*} \tilde{\partial} U. \quad (26)$$

and its time derivative as:

$$\dot{X}^* = \frac{G}{H^*} \tilde{\partial} \frac{\partial U}{\partial t} = gG\partial_x'^* \left( \frac{T}{T^*} - (1 + \tilde{\partial})\hat{q} \right). \quad (27)$$

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:

$$\mathbb{D}_F^{+(1)} = \mathbb{D}_O^0 + \Delta t \frac{d\mathbb{D}}{dt} + (X_F^{+(0)} + X_O^{+(0)} - X_O^0), \quad (28)$$

and for subsequent steps  $n \geq 2$ :

$$\mathbb{D}_F^{+(n)} = \mathbb{D}_O^0 + \Delta t \frac{d\mathbb{D}}{dt} + \left( X_F^{+(n-1)} - X_O^0 \right). \quad (29)$$

As mentioned earlier,  $d$  is the classical vertical divergence of the wind. In this study, a physical basic state at rest has

been considered. In this case without advection, points F and O are confused and (28) and (29) become:

$$\mathbb{D}^{+(1)} = \mathbb{D}^0 + \Delta t \frac{d\mathbb{D}}{dt} + 2(X^{+(0)} - X^0), \quad (30)$$

and for subsequent steps  $n \geq 2$ :

$$\mathbb{D}^{+(n)} = \mathbb{D}^0 + \Delta t \frac{d\mathbb{D}}{dt} + \left( X^{+(n-1)} - X^0 \right). \quad (31)$$

where indices F and O have been dropped for clarity.

e. *Systems under analysis*

In this paragraph, the operator  $l$  making the operator  $\tilde{\mathcal{L}}$  unbounded of the equation (6) is defined as:

$$l = \begin{bmatrix} \tilde{\partial} \\ 1 \\ 1 \\ \tilde{\partial} + 1 \end{bmatrix}, \quad (32)$$

and the state vector becomes:

$$\Psi = \begin{bmatrix} U \\ \mathbb{D} \\ T \\ \hat{q} \end{bmatrix}. \quad (33)$$

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

$$\tilde{\partial} \frac{\partial U}{\partial t} = R\tilde{\partial}'_x T - R\tilde{T}\tilde{\partial}'_x \hat{q} - R\tilde{T}\tilde{\partial}'_x \tilde{\partial} \hat{q}, \quad (34)$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{R\tilde{T}} \tilde{\partial}(\tilde{\partial} + 1)\hat{q} + (1 - \delta_{SL})\dot{X}, \quad (35)$$

$$\frac{\partial T}{\partial t} = -\frac{R\tilde{T}}{C_v} (\partial_x U + \mathbb{D}), \quad (36)$$

$$(\tilde{\partial} + 1) \frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v} (\tilde{\partial} + 1) (\partial_x U + \mathbb{D}) + \tilde{\partial}'_x U. \quad (37)$$

Applying the operator (32) to the equations (19)–(23) gives:

$$\tilde{\partial} \frac{\partial U}{\partial t} = R\partial_x'^* T - RT^* \partial_x'^* \hat{q} - RT^* \partial_x'^* \partial \hat{q}, \quad (38)$$

$$\frac{\partial \mathbb{D}}{\partial t} = -\frac{g^2}{rRT^*} \tilde{\partial}(\tilde{\partial} + 1)\hat{q} + (1 - \delta_{SL})\dot{X}^*, \quad (39)$$

$$\frac{\partial T}{\partial t} = -\frac{RT^*}{C_v} (\partial_x U + \mathbb{D}), \quad (40)$$

$$(\tilde{\partial} + 1) \frac{\partial \hat{q}}{\partial t} = -\frac{C_p}{C_v} (\tilde{\partial} + 1) (\partial_x U + \mathbb{D}) + \partial_x'^* U, \quad (41)$$

### f. Structure equations

The structure equation associated to the operator  $\tilde{\mathcal{L}}$  is combined from (34)–(37) equations:

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

where  $\tilde{N}^2 = g^2/(C_p \tilde{T})$ , and  $\tilde{c}^2 = C_p/C_v R \tilde{T}$  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 (38)–(41) equations. When  $\delta_{SL} = 0$  the structure equation is:

$$\left[ -\frac{1}{c^{*2}} \frac{\partial^4}{\partial t^4} + \frac{\partial^2}{\partial t^2} \left( \partial_x'^2 + \frac{\tilde{\partial}(\tilde{\partial} + 1)}{r H^{*2}} \right) + \frac{N^{*2}}{r} \partial_x'^2 \right] U = 0, \quad (43)$$

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

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

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

### a. Different options

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

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

In addition, for each of these both 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

For the system (6), the modes are looked for under:

$$\Psi_j(x, \sigma) = \widehat{\Psi}_j \exp(ikx) \sigma^{i\nu-1/2}, \quad (45)$$

where  $j$  is the  $j$ -th mode considered, and  $k$  and  $\nu$  are the horizontal and vertical wave numbers which are in:

$$k \in [-k_{max}, -k_{min}] \cup [k_{min}, k_{max}], \quad (46)$$

$$\nu \in [-\nu_{max}, -\nu_{min}] \cup [\nu_{min}, \nu_{max}], \quad (47)$$

and where maximum absolute values are defined as:

$$k_{max}^2 = \frac{\pi^2}{\Delta x^2}, \quad (48)$$

$$\frac{\nu_{max}^2}{H^{*2}} = \frac{\pi^2}{\Delta z^2}, \quad (49)$$

and minimal absolute values as:

$$k_{min}^2 = \frac{4\pi^2}{L^2} \approx 0, \quad (50)$$

$$\nu_{min}^2 = 0. \quad (51)$$

The eigenvalues of the operators become:

$$\widehat{\partial}_x = ik, \quad (52)$$

$$\widehat{\partial}_x' = ik + i \frac{G\nu}{\tilde{H}} - \frac{G}{2\tilde{H}} = \bar{k}', \quad (53)$$

$$\widehat{\partial}_x'^* = ik + i \frac{G\nu}{H^*} - \frac{G}{2H^*} = k'^*, \quad (54)$$

$$\tilde{\partial} = i\nu - 1/2. \quad (55)$$

Then, the system (6) becomes:

$$\frac{d\widehat{\Psi}}{dt} = \widehat{\mathcal{L}}\widehat{\Psi} \quad (56)$$

and (32) and (33) become:

$$l = \begin{bmatrix} \xi_1 \\ 1 \\ 1 \\ \xi_4 \end{bmatrix}, \quad (57)$$

$$\widehat{\Psi} = \begin{bmatrix} \widehat{U} \\ \widehat{D} \\ \widehat{T} \\ \widehat{q} \end{bmatrix}. \quad (58)$$

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

$$\widehat{\mathcal{L}} = \begin{bmatrix} 0 & 0 & \bar{\mu}_{13}/\xi_1 & \bar{\mu}_{14}/\xi_1 \\ 0 & 0 & \bar{\mu}_{23} & \bar{\mu}_{24} \\ \bar{\mu}_{31} & \bar{\mu}_{32} & 0 & 0 \\ \bar{\mu}_{41}/\xi_4 & \bar{\mu}_{42}/\xi_4 & 0 & 0 \end{bmatrix}. \quad (59)$$

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

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

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

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

$$\bar{\mu}_{24} = \frac{g^2}{RT}(v^2 + 1/4) - (1 - \delta_{SL})gG(iv + 1/2)\bar{k}',$$

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

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

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

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

Coefficients  $\xi_1$  and  $\xi_4$  are:

$$\xi_1 = iv - 1/2,$$

$$\xi_4 = iv + 1/2.$$

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

$$\widehat{\mathcal{L}}^* = \begin{bmatrix} 0 & 0 & \mu_{13}^*/\xi_1 & \mu_{14}^*/\xi_1 \\ 0 & 0 & \mu_{23}^* & \mu_{24}^* \\ \mu_{31}^* & \mu_{32}^* & 0 & 0 \\ \mu_{41}^*/\xi_4 & \mu_{42}^*/\xi_4 & 0 & 0 \end{bmatrix}. \quad (60)$$

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

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

$$\mu_{14}^* = -RT^*(iv + 1/2)[(1 - \delta)ik + \delta k'^*],$$

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

$$\mu_{24}^* = \frac{g^2}{rRT^*}(v^2 + 1/4) - (1 - \delta_{SL})gG(iv + 1/2)[\delta k'^*],$$

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

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

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

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

Coefficients  $\xi_1$  and  $\xi_4$  are:

$$\xi_1 = iv - 1/2,$$

$$\xi_4 = iv + 1/2.$$

### c. Physical instability or damping

By diagonalizing the operator  $\widehat{\mathcal{L}}$  and integrating it temporally from  $t$  to  $t + \Delta t$ , the  $j$ -th mode considered is:

$$\widetilde{\Psi}_j(t + \Delta t) = \widetilde{\Psi}_j(t) \exp(\bar{\omega}_j \Delta t), \quad (61)$$

where  $\widetilde{\Psi}$  designed the projection of  $\widehat{\Psi}$  in  $\widehat{\mathcal{L}}$  and  $\bar{\omega}_j$  are the eigenvalues associated to  $\widehat{\mathcal{L}}$ , which are also the solutions of the structure equation (42):

$$-\frac{1}{c^2}\bar{\omega}^4 + \bar{\omega}^2 \left[ \bar{k}'^2 - \frac{1}{H^2} \left( v^2 + \frac{1}{4} \right) \right] + N^2 \bar{k}'^2 = 0, \quad (62)$$

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

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

where  $\Re$  designed the real part. The growth rate of the  $j$ -th mode is:

$$\bar{\Gamma}_j = \frac{|\widetilde{\Psi}_j(t + \Delta t)|}{|\widetilde{\Psi}_j(t)|} = \exp(\Re(\bar{\omega}_j)\Delta t) \geq 1 \quad (63)$$

### d. Temporal discretization

When an ICI-2TL scheme is used, (7) becomes:

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

where

$$\mathcal{H} = \left( \mathcal{I} - \frac{\Delta t}{2}\widehat{\mathcal{L}}^* \right), \quad (65)$$

$$\mathcal{E} = \left( \mathcal{I} + \frac{\Delta t}{2}\widehat{\mathcal{L}}^* \right), \quad (66)$$

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 (64) becomes:

$$\widehat{\Psi}^{+(n)} = \mathcal{H}^{-1}\mathcal{E}\widehat{\Psi}^0 + \frac{\Delta t}{2}\mathcal{H}^{-1}\widehat{\mathcal{R}}(\widehat{\Psi}^0 + \widehat{\Psi}^{+(n-1)}), \quad (67)$$

where  $\mathcal{H}^{-1}$  will be inverted by a solver.

### e. Normalized growth rate

When a physically damped or unstable system is considered, the growth rate of the numerical scheme is normalized by the growth rate of the physical system. Hence, in the following, the growth rate is defined as:

$$\Gamma = \frac{\max_j(|\Gamma_j|)}{\max_{j'}(|\bar{\Gamma}_{j'}|)} \quad (68)$$

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

### f. Limits

This study is based on several hypotheses that probably overestimate the numerical stability of the problem. In practice, any possible errors or inconsistencies in the vertical boundary conditions are not taken into account in this analysis where the equations are unbounded, probably leading to an overestimation of the numerical stability. Effects of vertical and horizontal discretization finite-difference schemes are neither considered in this article. 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 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

Excepted in the part dealing the ratio  $r$ , reference temperatures are identical to those used in AROME:  $T^* = 350$  K,  $T_e^* = 100$  K. To assure numerical stability, the warmest temperature  $T^*$  chosen is such as it is higher than  $\bar{T}$ , whatever the meteorological situation considered. The non-linear residual factor will be thus chosen in  $\alpha \in [-0.95, 0]$ .

Excepted explicitly mentioned, horizontal resolutions chosen in the following are these expected in a mid-term future: an horizontal resolution of  $\Delta x = 300$  m and a large time step  $\Delta t = 12$  s corresponding to a CFL:  $C^* = c^* \Delta t / \Delta x \approx 14$ . A vertical grid with spaced levels of 2 m is chosen. At these resolutions, slopes considered are usually 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) for the most unstable combination of wave numbers. These wave numbers are looked for in the intervals (46) and (47) where negative values are accepted. In practice, the most unstable combination is usually obtained for the smallest wave numbers in absolute value ( $|k_{max}|$  and  $|v_{max}|$ ). Graphics represent only positive values of slope 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 stability domain.

## 5. Explicit treatment of orographic terms

As mentioned previously, this study focuses on estimating instability due to the explicit treatment of orographic terms. Consequently, a special focus is given to the maximum slope without any thermal residual ( $\alpha = 0$ ). However,

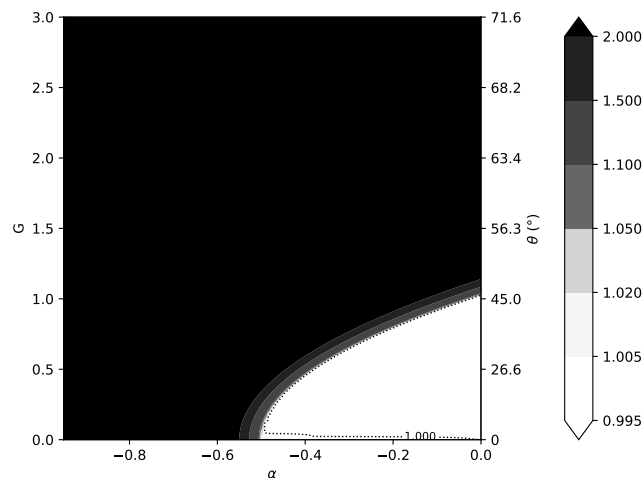


FIG. 1. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r = 1.0$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$  and  $C^* = 14$ .

numerical instabilities can of course come from thermal residual ( $\alpha \neq 0$ ). Hence, numerical stability must be analyzed in comparison between these both parameters by analysing the 'spread' of the stability domain too (the area where  $\Gamma \leq 1$ ).

### a. Sensitivity to the ratio $r$

When  $\alpha = 0$  and  $r = 1$ , the maximum slope is around  $45^\circ$  for the configuration chosen with a large CFL number (Fig 1). When a thermal residual is added, the stability strongly decreases even for small slopes.

On the contrary, when several reference temperatures are considered ( $T_e^* \neq T^*$ ), slopes over  $68^\circ$  can now be achieved when  $\alpha = 0$ . Moreover, the stability domain is larger in general (for  $\alpha \neq 0$ ) (Fig 2 for  $r \approx 0.3$ ). This confirms experiments carried out with the AROME model for which numerical instabilities appear for steeper slopes when  $r$  is weaker. This is also agree with conclusions of Bénard (2004) who shows this 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  $r \rightarrow 0$ ). However, an excessive decrease of this ratio  $r$  can worsen scores of a model and more particularly when  $T_e < 100$  K (not shown). That is why this strategy can only be used barely and a value of  $T_e = 100$  K (corresponding to  $r \approx 0.3$ ) 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 very close to those obtained when experiments are conducted in a real model.



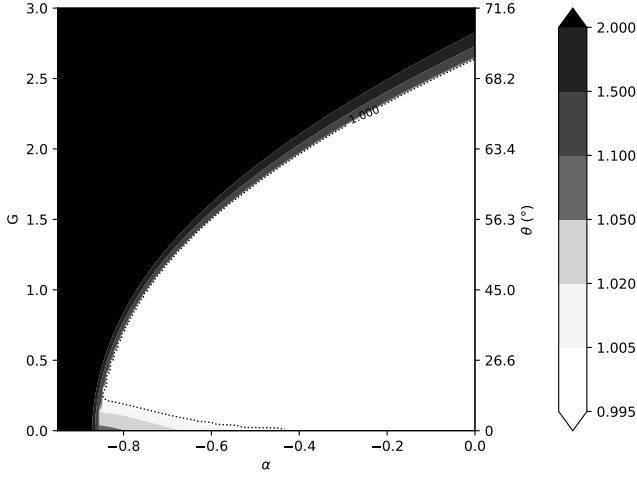


FIG. 2. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 14$ .

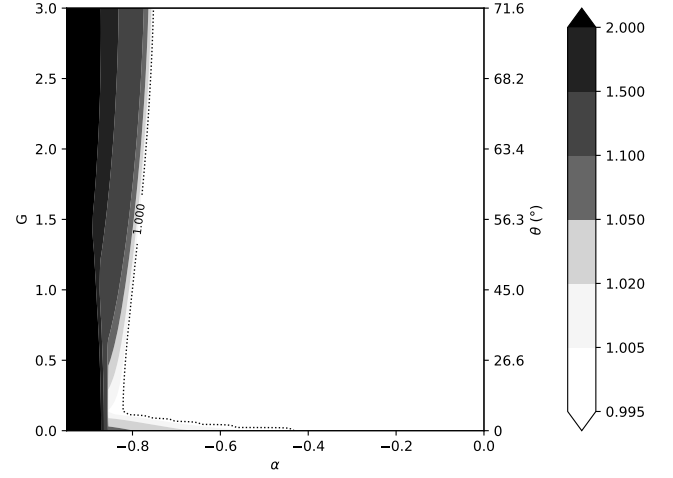


FIG. 4. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 1$ ,  $\delta = 0$ ,  $C^* = 14$ .

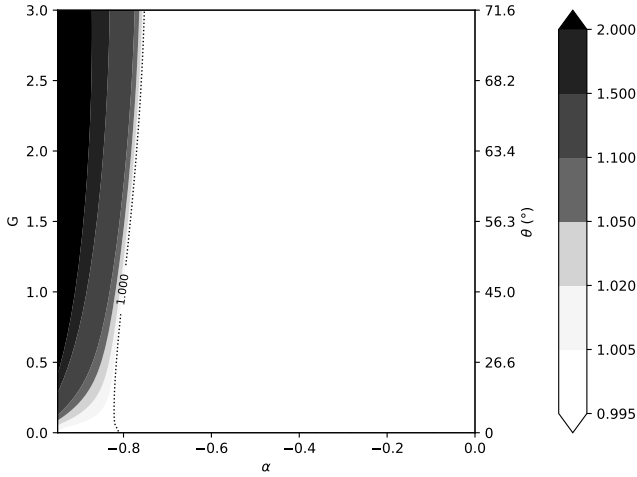


FIG. 3. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r \approx 0$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 14$ .

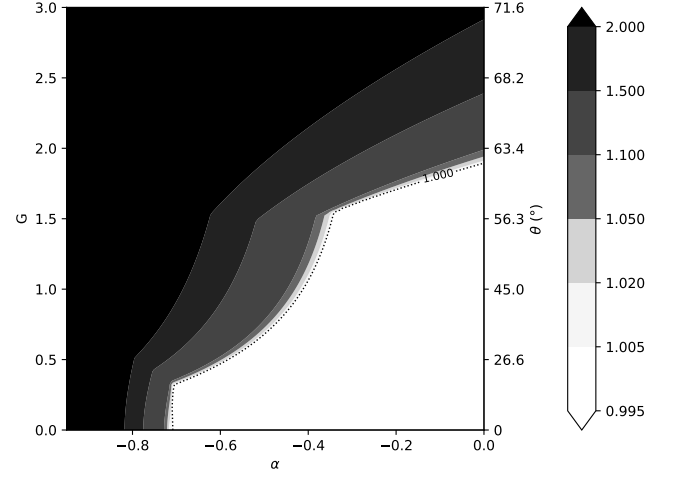


FIG. 5. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the SI-2TL-E scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 14$ .

### 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). 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.

### c. Changing temporal scheme

The stability domain of the SI-2TL-E scheme (Fig 5) is narrower than the ICI-2TL-PC one, more particularly for  $\alpha \in [-0.7, -0.4]$ . This is in good agreement with the shifting from the SI-2TL-E scheme to the ICI-2TL-PC

scheme for the AROME model in 2015 when the resolution changes from 2.5 km to 1.3 km (see Brousseau et al. (2016)).

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

### d. Time-step reduction

When CFL numbers close to unity are used, the stability domain is weakly modified compared to the case when a

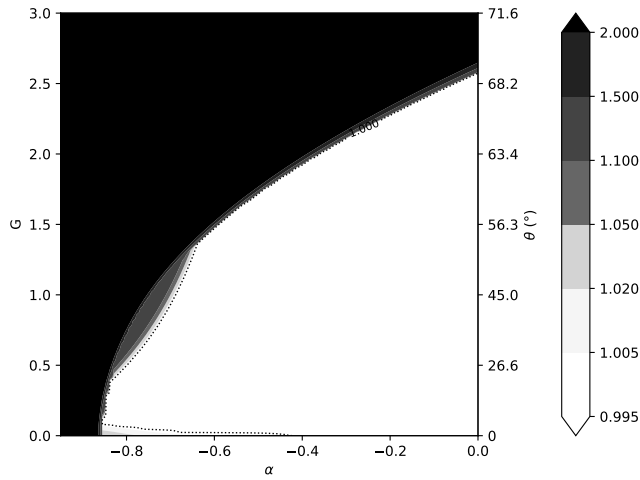


FIG. 6. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL scheme with  $N_{iter} = 4$ ,  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 14$ .

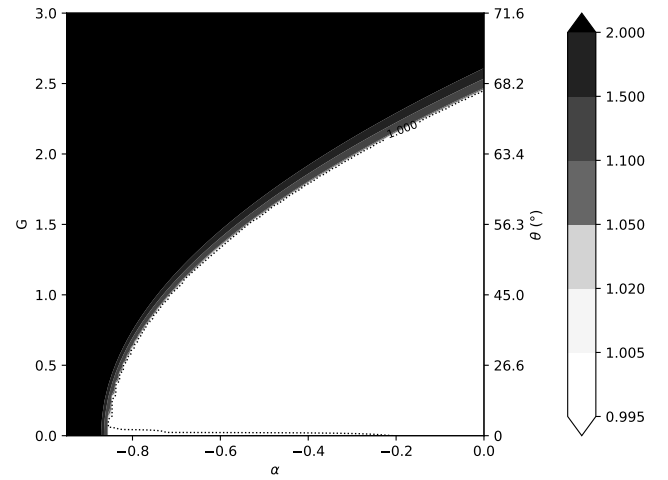


FIG. 8. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 1.0$ .

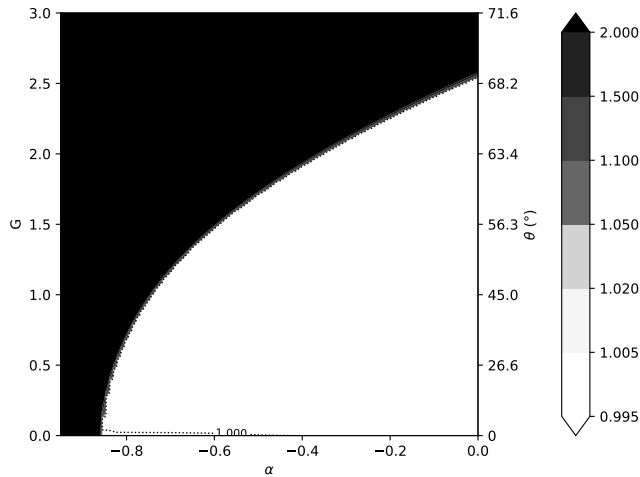


FIG. 7. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL scheme with  $N_{iter} = 8$ ,  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 0$ ,  $C^* = 14$ .

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 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 down to  $r \approx 0.3$  the ratio  $r$ , 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 only way to improve numerical stability and is discussed in the next section.

## 6. Implicit treatment of orographic terms

When all terms are implicitly treated and no thermal residual is added, the stability is reached unconditionally (whatever the slope is) because the total non-linear residual is null. In practise however, a thermal residual is always added, as in everyday meteorological situations, and justifies the following tests.

When an implicit treatment is applied on the orographic terms, the domain of stability is significantly larger (Fig 5) compared to what it is when terms are explicitly treated (Fig 2).

To reduce the computational cost, a solution would be to use again the SI-2TL-E scheme, roughly half the computation cost of the ICI-2TL-PC scheme. In this case, the stability domain is only smaller for  $\alpha < -0.5$  compared to the ICI-2TL-PC one. This alternative could be sufficient in a real model but need to be checked with real experiments.

By increasing the number of iterations of the ICI-2TL scheme up to 4 iterations, the stability domain is slightly wider (Fig 11) than with only 2 iterations, but benefits in

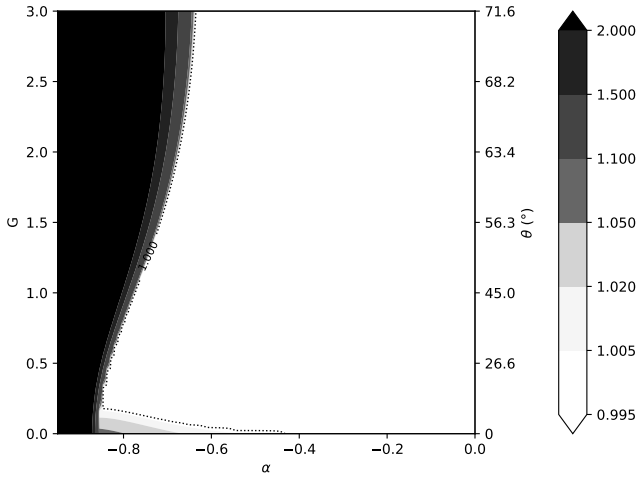


FIG. 9. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL-PC scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 1$ ,  $C^* = 14$

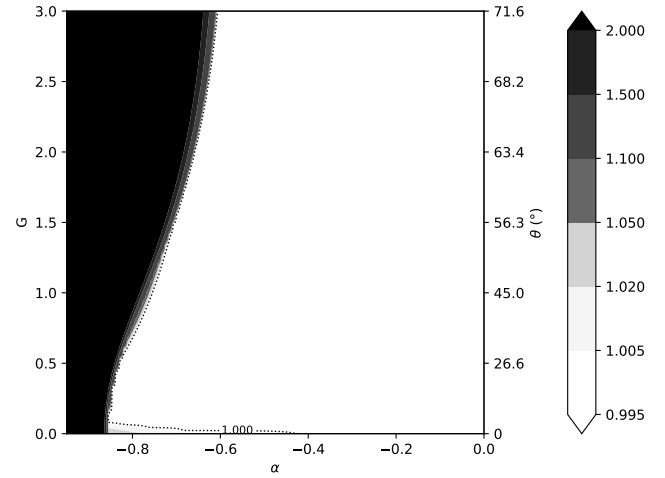


FIG. 11. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the ICI-2TL scheme with  $N_{iter} = 4$ ,  $r \approx 0.3$ ,  $\delta_{SL} = 1$ ,  $\delta = 0$ ,  $C^* = 14$ .

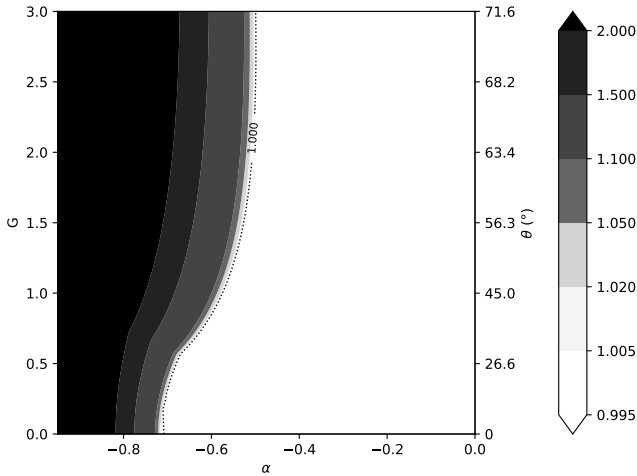


FIG. 10. Growth rate as a function of the non-linearity parameter  $\alpha$  and the slope  $G$  for the SI-2TL-E scheme with  $r \approx 0.3$ ,  $\delta_{SL} = 0$ ,  $\delta = 1$ ,  $C^* = 14$ .

stability are weak compared to the over-cost of this scheme. Increasing again the number of iterations does not improve the numerical stability in general (not shown).

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

## 7. Invertibility and condition numbers

Several methods can be used to invert the implicit problem. There are mainly two classes of methods to invert this problem: direct methods and iterative methods.

Direct methods invert the problem in a single iteration, are exact, and are not sensitive to the condition number. Nevertheless, they often suffer from a scalability problem when different MPI (Message Passing Interface) tasks distributed on different compute nodes communicate. The 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, such as the FFT algorithm, are generally being replaced by iterative methods (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 their condition number of the problem to be inverted (Liesen and Tichý (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 allows us to verify if the strategies previously studied to improve numerical stability in the presence of steep slopes are not too expensive numerically to be used in a future operational framework.

Therefore, in this section, the eigenvalues of the implicit problem are computed analytically to derive condition number estimates. The condition number is defined here as the ratio between the largest and the smallest eigenvalue in absolute value. It is computed as a function of the considered slope  $G$ . Moreover, invertibility conditions are also extracted and allow to find the maximum slope beyond which the problem is no longer invertible, by computing the maximum slope for which the real and imaginary parts are simultaneously zero.

These estimates are given on an implicit problem which has been algebraically reduced (noted  $\mathcal{H}_r$ ) 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.

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

#### a. Explicit treatment of orographic terms

When orographic terms are treated explicitly, the eigenvalues of the implicit problem reduced algebraically ( $\mathcal{H}_r$ ) on a single variable are:

$$\mathcal{H}_r = 1 + \frac{\Delta t^2}{4} b k^2 \quad (69)$$

where:

$$b = \left(1 + \frac{\Delta t^2}{4} \frac{c^{*2}}{r H^{*2}} (v^2 + 1/4)\right)^{-1} \left(1 + \frac{\Delta t^2}{4} \frac{N^{*2}}{r}\right) c^{*2} \quad (70)$$

are the eigenvalues of the product between a ‘vertical part’ and some constants. The condition number of the problem (69) is:

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

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 well 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 (69) are all greater than 1.

We note from the first factor of (70), that a ‘vertical’ problem has to be inverted. Its condition number can be easily estimated as:

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

Since  $\Delta z \ll \Delta x$ , the vertical problem is significantly less well conditioned than the full problem (69) ( $C_{vert} \gg C$ ). 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 (70). Thereafter, we consider that the vertical part is inverted by a specific method, not sensitive to its condition number.

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

When orographic terms are treated implicitly, the eigenvalues of the implicit problem reduced algebraically are:

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

where the vertical part  $b$  remains the same as (70). For the range of slope considered, the condition number is quite similar to the case 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 well convergence of the solver can hence be expected.

Furthermore, the problem is invertible only if the slope is:

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

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).

#### c. 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:

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

where:

$$b' = \left(1 + \frac{\Delta t^2}{4} \frac{c^{*2}}{r H^{*2}} (v^2 + 1/4)\right)^{-1} \quad (76)$$

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:

$$|G| \leq \sqrt{\frac{\frac{4}{\Delta t^2} + \frac{c^{*2}}{4 r H^{*2}} + \frac{c^{*2}}{r H^{*2}} v^2}{\frac{g}{2 H^*} + c^{*2} \frac{1}{H^{*2}} \beta (1 - \beta) v^2}} \quad (77)$$

with:

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

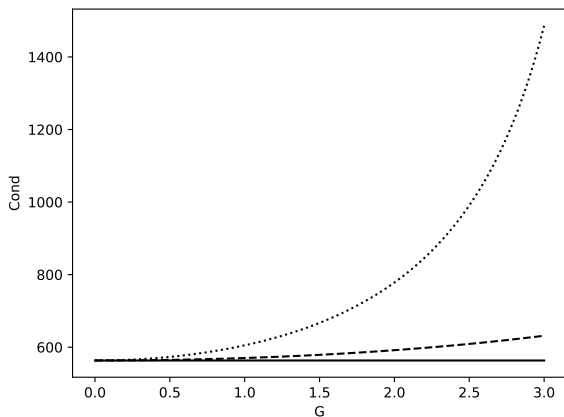


FIG. 12. Condition number of the implicit problem when  $r \approx 0.3$ ,  $C^* = 14$  with : implicit treatment of orographic terms  $\delta = 1$  and no specific treatment of the cross term  $\delta_{SL} = 0$  (dashed line), implicit treatment of orographic terms  $\delta = 1$  and specific treatment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

For large values of  $\nu$ , the invertibility condition becomes:

$$|G| \leq \frac{1}{\sqrt{\beta(1-\beta)r}} \approx \frac{2}{\sqrt{r}} \approx 3.7 \quad (79)$$

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 ( $r \approx 0.3$ ).

#### d. 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.

In addition, the vertical part of the problem is well suited to be inverted by a direct method, not sensitive to the condition number.

## 8. Conclusion

The stability analysis introduced in this article treats the ICI-2TL scheme with the fully elastic system of Euler equations using a mass coordinate and 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, including the orders of magnitude of the maximum slopes, which seem particularly realistic. For example, the limit slope (under which no numerical stability appears) of about  $45^\circ$ , 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 a colder temperature  $T_e^*$ ) without any thermal residual ( $\alpha = 0$ ), or when a thermal residual is added ( $\alpha \approx -0.7$ ) with an ICI scheme with two iterations and a cold temperature  $T_e^* = 100$  K is chosen. 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 orography is 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, provided the time step is reduced drastically, which is not viable in an operational context;
- decreasing the temperature  $T_e^*$  leads to a slight stabilization but at the cost of degrading 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 significantly reducing performance. Implicit treatment of orographic terms appears to be one of the only viable options and conclusions are now discussed on this point.

First, it was shown that the implicit problem was invertible in the range of slopes considered in this study, i.e. for slopes up to  $70^\circ$  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, excepted 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  $70^\circ$  with a moderate thermal residuals ( $\alpha \approx -0.5$ ) 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 ( $\alpha \approx -0.65$ ).

Condition number estimates show that it remains low, even for steep slopes (up to 3) whatever the configuration used: with or without implicit treatment of orographic terms. This suggests a well 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 favour of a variable scheme, that implicitly contain orographic terms, to tackle hectometric resolutions and the resulting steep slopes. It also abounds to remove the spectral discretization that can still be used in some SI models, in addition to the scalability problem raised by the Fast Fourier Transform algorithm on massively parallel machines. Nevertheless, a lot of work is needed to implement such a method:

- to discretize the new operators present in the implicit problem by ensuring consistency with the non-linear residual part;
- to implement and tune a Krylov solver, to solve the implicit problem containing the orographic terms, and possibly to add a preconditioner to accelerate convergence;
- to make test cases on non-linear flows in the presence of a more realistic (non-linear) relief and steep slopes.

## APPENDIX A

### Vertical operators

Les opérateurs verticaux sont définis comme:

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

$$\mathcal{S}\Psi = (1/\sigma) \int_0^{\sigma} \Psi d\sigma'$$

$$\mathcal{N}^*\Psi = \int_0^1 \Psi d\sigma$$

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

## APPENDIX B

The following notations are used:

- $m$  : the metric factor (equals to  $\partial_{\eta}\pi$ )
- $p$  : local pressure
- $w$  : vertical velocity
- $\phi$  : 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.

*Acknowledgments.*

*Data availability statement.*

## 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, doi: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, doi: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, doi: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, doi: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, doi: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, doi: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 dynamical core for the ifs. *Geoscientific Model Development*, **12** (2), 651–676, doi: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. doi: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, doi:10.1175/2008mwr2332.1.
- Liesen, J., and P. Tichý, 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, doi: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, doi: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, doi:10.5194/gmd-10-1487-2017.

## 4.4 Expériences numériques dans un cadre 2D

Dans cette partie, on étudie la stabilité numérique en présence de fortes pentes pour les deux versions : avec ( $\delta = 1$ ) et sans ( $\delta = 0$ ) traitement implicite des termes orographiques. Dans un premier temps, un relief idéalisé est considéré où seuls deux degrés de liberté permettent de régler la valeur de la pente. Dans un second temps, un relief réaliste est obtenu à partir d'une coupe transversale des Alpes. Dans ce cas, la valeur de la pente est augmentée par l'affinement de la résolution horizontale.

### 4.4.1 Configuration choisie

L'analyse de stabilité menée dans l'article suggérait de traiter tous les termes orographiques implicitement pour maximiser la stabilité numérique. Toutefois, le cadre non borné de cette analyse n'a pas permis d'étudier le terme  $\Pi_s$  (4.23), pour lequel des expériences ont montré une nette dégradation de la stabilité numérique lorsqu'il est traité implicitement (non montré). Ce comportement néfaste reste pour le moment inexpliqué mais est peut-être lié à un problème dans la discrétisation verticale. Dans une moindre mesure, il semblerait que le terme  $U_2$  (4.18) dégrade légèrement la stabilité numérique ou ne l'améliore pas significativement (non montré). Par conséquent, par la suite, ces deux termes ne seront pas traités dans le problème implicite. Enfin, le terme croisé sera traité dans le schéma semi-Lagrangien ( $\delta_{SL} = 1$ ).

### 4.4.2 Écoulement sur relief idéalisé

Les écoulements sur des reliefs idéalisés étudiés ici consistent à imposer un vent horizontal par les bords  $U = 10$  m/s, dans un environnement stratifié de fréquence de Brunt-Väisälä  $\bar{N}_b = 0,012$  s<sup>-1</sup>. On considère un relief -très idéalisé- de type Agnesi, de la forme :

$$z_s = h_{max} \frac{a^2}{a^2 + x^2} \quad (4.44)$$

où  $h_{max}$  est la hauteur du relief,  $a$  est la demi-largeur. Dans la plupart des études considérées, pour que l'écoulement reste laminaire, la perturbation due à la présence de ce relief est choisie suffisamment petite pour qu'elle puisse être supposée infinitésimale. Par conséquent, des solutions analytiques peuvent être facilement calculées et les approches phénoménologiques sont pertinentes. Néanmoins, aborder le problème des fortes pentes conduit à considérer des perturbations de fortes amplitudes qui ne peuvent être supposées infinitésimales. Cette approche ne pourra donc pas être considérée par la suite.

Pour étudier la stabilité numérique en présence de fortes pentes sur ce relief, on peut :

- augmenter la hauteur maximale du relief  $h_{max}$  ou,
- diminuer la valeur du paramètre  $a$ .

Ces deux possibilités se heurtent toutefois à de nombreuses difficultés supplémentaires de simulation. Lorsque le paramètre  $a$  est fixé, une augmentation trop importante de  $h_{max}$ ,



conduit à basculer d'un écoulement laminaire à un écoulement turbulent. En l'absence de paramétrisation physique turbulente ou même de mécanismes de diffusion, l'énergie s'accumule aux petites échelles et conduit généralement à une interruption rapide de la simulation due par exemple à un déferlement d'une onde (Durrán, 1990). A contrario, lorsque la hauteur du relief  $h_{max}$  est fixée, une diminution trop importante du paramètre  $a$  conduit à générer des champs atmosphériques assez bruités. On estime qu'il génère un forçage sur sur les champs atmosphériques de longueur d'onde environ égale à  $a$ . Or une trop forte diminution de la valeur de  $a$  jusqu'à des valeurs inférieures à  $4\Delta x$ , peut conduire à des instabilités numériques dues par exemple à l'aliasing, en l'absence de diffusion numérique.

L'enjeu dans toute la suite est donc d'une part, d'augmenter suffisamment  $h_{max}$  pour générer des fortes pentes mais pas trop pour ne pas basculer dans un écoulement trop turbulent, et d'autre part de diminuer la valeur du paramètre  $a$  mais pas trop pour ne pas générer d'instabilité numérique due à l'absence de filtrage ou de diffusion.

Dans un premier temps, on prend un relief d'amplitude maximale  $h_{max} = 500$  m et on fait varier le paramètre  $a$  jusqu'à produire des pentes suffisamment fortes pour générer des instabilités numériques. Ensuite, on considère des écoulements plus turbulents en prenant des amplitudes de relief plus fortes :  $h_{max} = 1000$  m puis  $h_{max} = 2000$  m.

### 4.4.3 Écoulement faiblement non-linéaire

Dans cette partie, on considère que la simulation est stable numériquement si elle parvient à son terme au bout de 8000 s, sans exploser numériquement. Dans un premier temps, un écoulement faiblement non linéaire est considéré avec une hauteur de relief  $h_{max} = 500$  m, puis on diminue la valeur du paramètre  $a$  (et donc on augmente la pente) jusqu'à ce que les instabilités numériques apparaissent et fassent exploser la simulation. Une pente maximale en est ainsi déduite.

#### Validation de la version codée

Dans un premier temps, on montre que les deux versions ( $\delta = 0$  et  $\delta = 1$ ) produisent des résultats similaires pour une valeur de  $a = 350$  m correspondant à une pente de 92%. Les champs de vent horizontaux sont très proches entre les Figures 4.1b et 4.1d lorsqu'un pas de temps de  $\Delta t = 3$  s est considéré. La diminution du pas de temps avec  $\Delta t = 1$  s conduit également à des résultats similaires (Figures 4.1a et 4.1c).

Dans la version où les termes orographiques sont traités implicitement, le paramètre  $a$  peut être diminué jusqu'à une valeur de 200 m, correspondant à des pentes plus fortes de l'ordre de 160% (soit un angle correspondant d'environ 58°). On note également que malgré la présence de ces très fortes pentes, l'écoulement reste laminaire avec une superposition d'ondes évanescentes et propagatives (Figure 4.2).

#### 4.4. EXPÉRIENCES NUMÉRIQUES DANS UN CADRE 2D

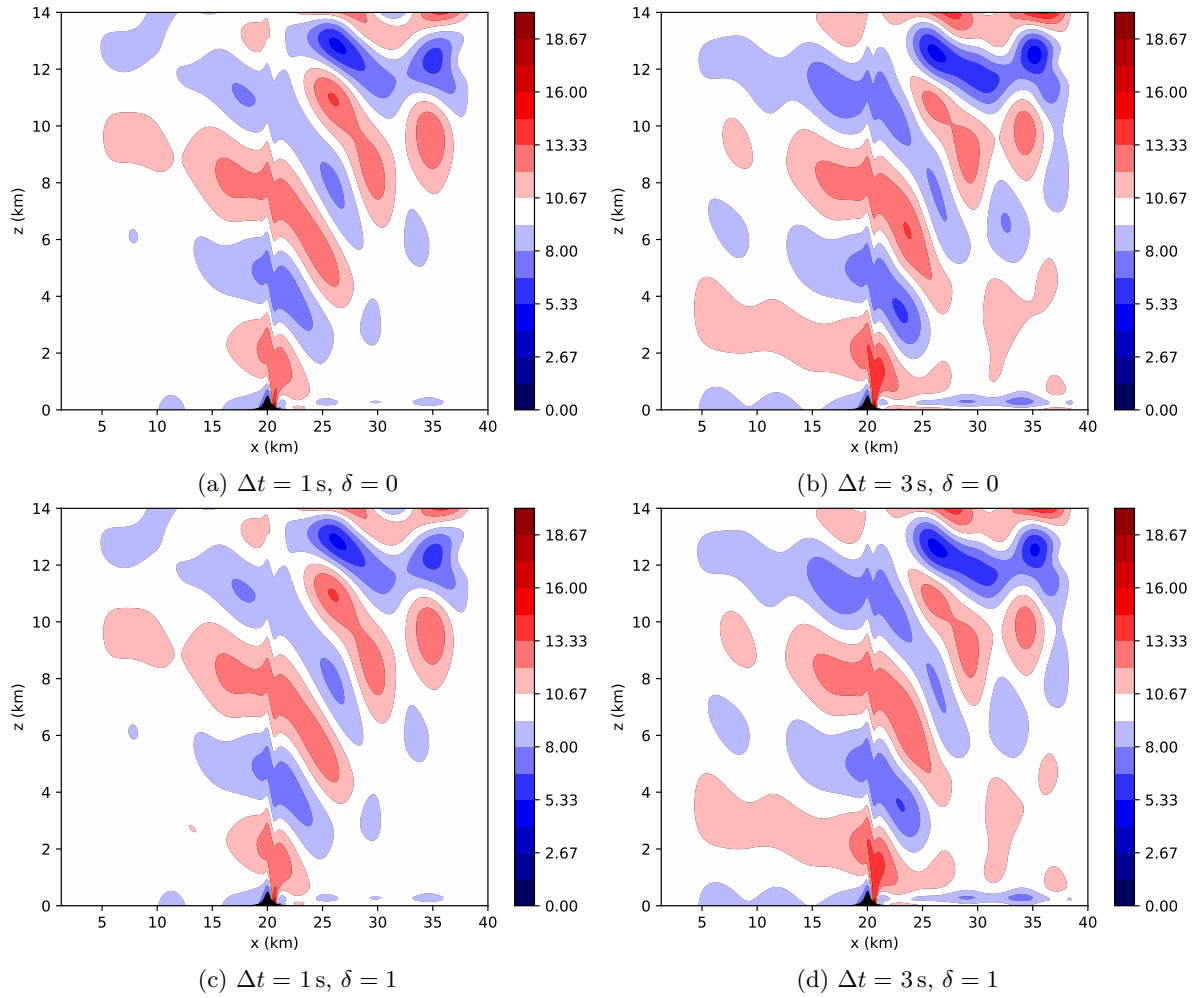


FIGURE 4.1 – Champ de vent horizontal  $U$  pour  $h_{max} = 500$  m,  $a = 350$  m après 7800 s.

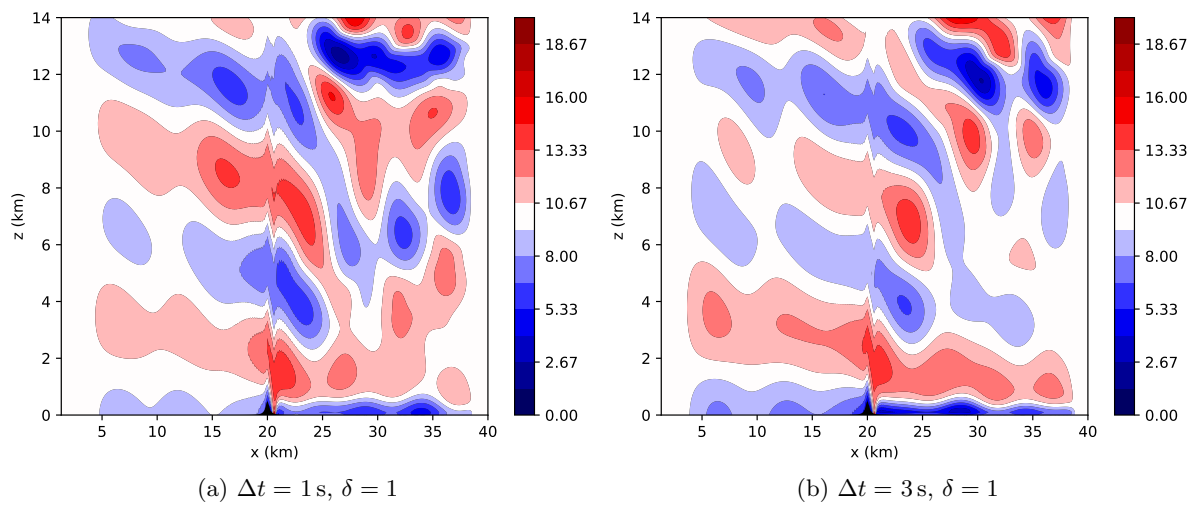


FIGURE 4.2 – Champ de vent horizontal pour  $h_{max} = 500$  m,  $a = 200$  m après 7800 s.

**Pentes maximales admises**

$\Delta t$	max	$\delta = 0$	$\delta = 1$
6 s	a	400 m	350 m
	G	0.80 ( <i>38.7°</i> )	0.92 ( <i>42.9°</i> )
3 s	a	350 m	200 m
	G	0.92 ( <i>42.9°</i> )	1.60 ( <i>58.0°</i> )
1 s	a	350 m	200 m
	G	0.92 ( <i>42.9°</i> )	1.60 ( <i>58.0°</i> )

TABLE 4.1 – Valeur maximale du paramètre  $a$  garantissant la stabilité pendant au moins 8000 s avec  $h_{max} = 500$  m et pour plusieurs pas de temps considérés. Dédution des pentes maximales admises  $G$  (et valeur de l’angle correspondant en degrés en italique).

Les résultats de la Table 4.1, ont pu être établis par essai-erreur en cherchant la pente maximale garantissant la stabilité numérique au bout de 8000 s. Ces résultats indiquent un gain de l’ordre de 10% sur la pente maximale lorsqu’un très grand pas de temps est considéré ( $\Delta t = 6$  s,  $c_0^* = 21$ ) et pouvant atteindre jusqu’à 70% de pente supplémentaire lorsque des plus petits pas de temps sont considérés ( $\Delta t = 3$  s ou  $\Delta t = 1$  s) dans la version où  $\delta = 1$ , par rapport à celle où  $\delta = 0$ .

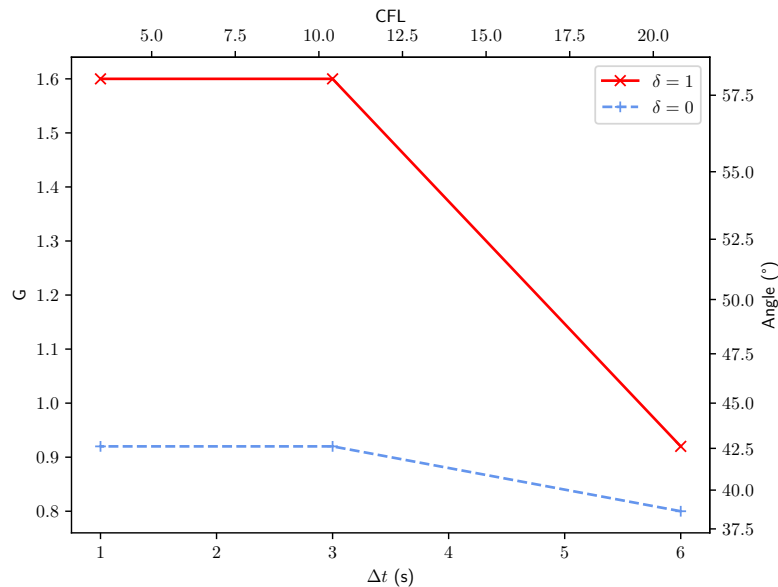


FIGURE 4.3 –  $h_{max} = 1000$  m,  $\delta = 1$

Dans la partie suivante, on analyse dans quelle mesure ces conclusions sont transposables à des écoulements plus turbulents en présence de reliefs plus hauts comme ceux rencontrés

dans les modèles actuels.

#### 4.4.4 Écoulement non-linéaire

On considère dans un premier temps un relief de hauteur maximale  $h_{max} = 1000$  m. Après plus de deux heures de simulation, l'écoulement conserve la propagation d'ondes évanescentes et propagatives (Figure 4.4a). En revanche pour une hauteur maximale  $h_{max} = 2000$  m, l'écoulement devient plus franchement turbulent (Figure 4.4b).

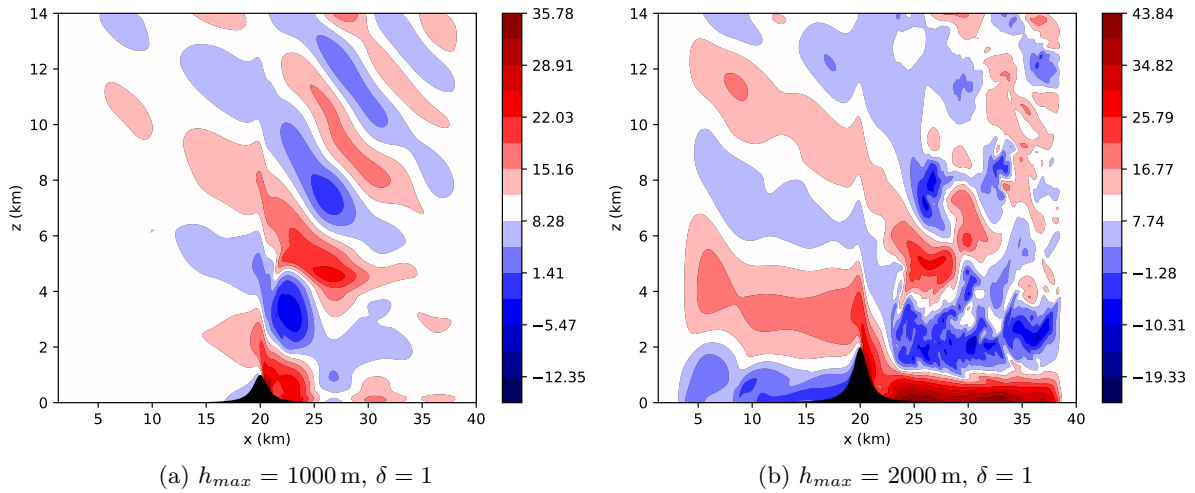


FIGURE 4.4 – Champ de vent horizontal pour  $a = 800$  m.

Dans toute la suite, on appelle « temps limite », et est noté  $T_{lim}$ , la durée maximale d'une simulation avant son interruption due à une explosion numérique. Lorsque l'écoulement devient très turbulent, le temps limite de la simulation est plus faible que celui des écoulements laminaires précédemment étudiés. Plusieurs causes peuvent être à l'origine de ce phénomène. Il peut s'agir d'une cause physique due au déferlement des ondes de gravité en l'absence de paramétrisation de la turbulence, ou bien d'une instabilité numérique. Dans la pratique, il n'est pas facile d'identifier la cause à l'origine de cette interruption.

Les deux versions étudiées ( $\delta = 0$  et  $\delta = 1$ ) résolvent les mêmes équations non-linéaires au sens du continu et ne diffèrent que par leur traitement des termes orographiques. Ainsi les différences entre ces deux versions sont d'origine numérique. Dans toute la suite, on considère que si les deux simulations présentent des temps limites significativement différents, alors celle dont le temps limite est le plus grand est numériquement plus stable que l'autre.

Dans toute la suite, pour une amplitude de relief donnée ( $h_{max} = 1000$  m ou  $h_{max} = 2000$  m), on comparera donc les durées maximales de simulation, en faisant varier le paramètre  $a$ , et aborder ainsi différentes valeurs de pente. La Figure 4.5 représente la durée maximale de la simulation avant son interruption (due à une erreur du code), en fonction de la pente maximale. Comme attendu, les temps limites sont plus faibles dans le cas où la montagne est plus haute. Par exemple, la simulation s'interrompt au bout d'environ 6000 s pour une pente

de l'ordre de l'unité et une hauteur  $h_{max} = 1000$  m, contre moins de 4000 s pour une pente du même ordre de grandeur et un relief deux fois plus haut ( $h_{max} = 2000$  m).

Les résultats montrent également que la simulation s'interrompt après seulement quelques pas de temps, c'est-à-dire pour  $T_{lim} < 500$  s pour :

- des pentes légèrement supérieures à 90% (40°) pour la version sans traitement implicite des termes orographiques ( $\delta = 0$ );
- des pentes supérieures à 220% (65°) pour la version avec traitement implicite des termes orographiques ( $\delta = 1$ ).

Par conséquent, le traitement implicite des termes orographiques permettrait de plus que doubler la pente maximale par rapport à la version où ils sont traités explicitement, ou de façon équivalente de gagner environ 25° sur la pente maximale. Ces résultats sont confirmés pour les deux valeurs de reliefs testées ( $h_{max} = 1000$  m (Figure 4.5a) et  $h_{max} = 2000$  m (Figure 4.5b)) et pour deux pas de temps différents ( $\Delta t = 3$  s et  $\Delta t = 1$  s).

Ces estimations doivent néanmoins être interprétées avec précaution, car elles dépendent fortement du seuil de durée maximale à partir duquel on considère que la simulation est systématiquement instable. Ici nous avons considéré un temps limite de seulement 500 s, mais une définition différente changerait les estimations de la pente maximale dans les deux versions.

On note que des pentes maximales de « seulement » 40° peuvent être atteintes dans le cas avec traitement explicite des termes orographiques alors qu'elles sont plutôt de l'ordre de 50° dans les expériences menées dans le modèle AROME en 3D. Plusieurs explications peuvent être avancées pour expliquer ces différences. Dans notre version 2D, contrairement à la version opérationnelle en 3D :

- l'écoulement ne peut contourner le relief dans la direction transverse à l'axe horizontal, et est par conséquent forcé à passer au-dessus du relief, provoquant éventuellement des déferlements prématurés;
- il n'y a pas de diffusion pour atténuer les différentes instabilités numériques;
- il n'y a pas de paramétrisation physique pour dissiper l'énergie, conduisant à une accumulation de l'énergie aux plus petites longueurs d'onde.

Ainsi, il est probable que la simulation d'écoulements en 3D soit numériquement plus stable que ceux en 2D, pour les deux versions étudiées ici ( $\delta = 0$  et  $\delta = 1$ ).

#### 4.4.5 Écoulement fortement non-linéaire sur relief réaliste

On considère désormais un relief réaliste obtenu par une coupe transversale nord-sud des Alpes à partir de la base orographique d'une résolution proche de 92 m (Reuter *et al.*, 2007). Un filtrage spectral est ensuite effectué pour interpoler cette orographie sur une grille à une résolution  $\Delta x$  différente. Un filtrage du relief pour les longueurs d'onde inférieures à  $4\Delta x$  est appliqué. Ce filtrage est plus important que celui appliqué habituellement dans les modèles opérationnels (de l'ordre de  $3\Delta x$ ), évitant ainsi une partie des instabilités numériques en l'absence de diffusion numérique. Pour les besoins de l'expérience, l'orographie est ramenée

#### 4.4. EXPÉRIENCES NUMÉRIQUES DANS UN CADRE 2D

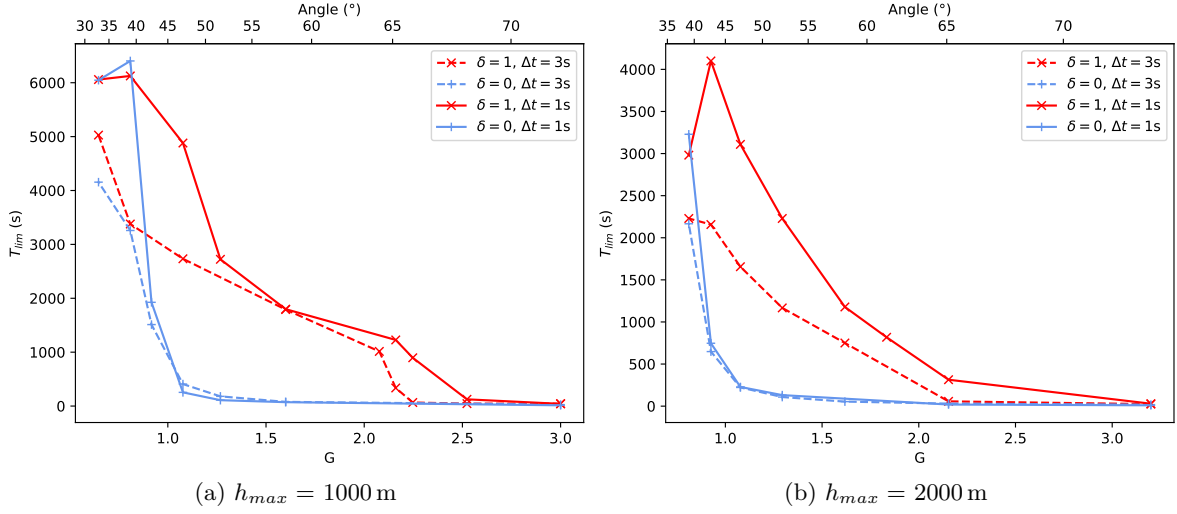


FIGURE 4.5 – Durée maximale de simulation avant interruption ( $T_{lim}$ ) en fonction de la pente maximale  $G$  pour les versions sans traitement ( $\delta = 0$ ) et avec traitement ( $\delta = 1$ ) des termes orographiques et pour  $\Delta t = 1$  s et  $\Delta t = 3$  s. Sur un second axe des abscisses en haut de la figure, l'angle correspondant en degré est également indiqué.

progressivement vers zéro sur les bords latéraux. Trois résolutions différentes sont testées et le pas de temps est ajusté pour chacune d'entre elles de sorte à maintenir le CFL constant et environ égal à  $c_x^* \simeq 14$  :

1.  $\Delta x = 1300,0$  m,  $\Delta t = 50$  s
2.  $\Delta x = 650,0$  m,  $\Delta t = 25$  s
3.  $\Delta x = 325,0$  m,  $\Delta t = 12,5$  s

Comme montré sur la Figure 4.6, l'orographie est de mieux en mieux décrite à mesure que la résolution spatiale est fine et représente des pentes de plus en plus fortes. Les valeurs de pentes maximales rencontrées pour ces trois résolutions sont :

1.  $G = 36\%$  pour  $\Delta x = 1300,0$  m,
2.  $G = 50\%$  pour  $\Delta x = 650,0$  m
3.  $G = 75\%$  pour  $\Delta x = 325,0$  m

Notons que les différents filtrages orographiques de cette étude atténuent les valeurs des pentes les plus fortes et sont ainsi bien plus faibles que celles rencontrées dans la version complète du modèle AROME quand des expériences sont conduites à ces résolutions. On note également que les valeurs de pente sont plus faibles que celles de la section précédente mais les hauteurs maximales des reliefs atteignent désormais les 3000 m environ.

La configuration testée a été choisie pour faire en sorte de retarder le plus possible l'apparition des instabilités physiques en prenant un vent faible :  $U = 2 \text{ m s}^{-1}$  et une assez forte stabilité thermique verticale  $\bar{N}_b = 0,012 \text{ s}^{-1}$ . Comme précédemment, une comparaison relative est effectuée entre les deux versions avec ( $\delta = 1$ ) ou sans ( $\delta = 0$ ) traitement implicite des termes orographiques. Les résultats de ces simulations sont exposés dans la Table 4.2.

Le champ de vent horizontal est tracé sur la Figure 4.7 pour les deux versions pour une résolution de 1300 m. On note que les résultats sont assez différents entre les deux simulations sur l'intensité des valeurs maximales de vent, sans qu'il ne soit possible de distinguer celle qui est de meilleure qualité. A la résolution fine de  $\Delta x = 325$  m, les champs de vent sont très bruités, tout comme le relief présent juste au-dessous (Figure 4.8). Sur cette figure, seule le champ de vent horizontal pour la version  $\delta = 1$  est représenté, à deux instants différents, l'autre version n'ayant pas aboutie à cause des instabilités numériques.

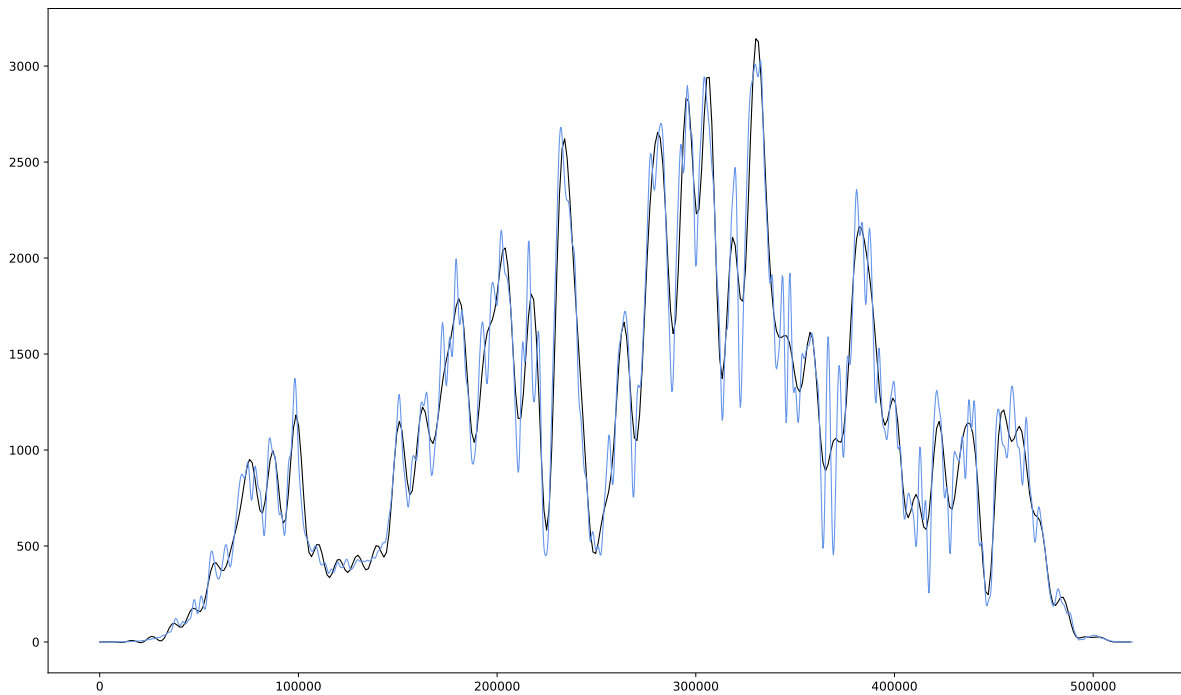


FIGURE 4.6 – Coupe transversale Nord-Sud du relief des Alpes à une résolution de  $\Delta x = 1300$  m (en noir) et à une résolution de  $\Delta x = 325$  m (en bleu clair) après filtrages orographiques.

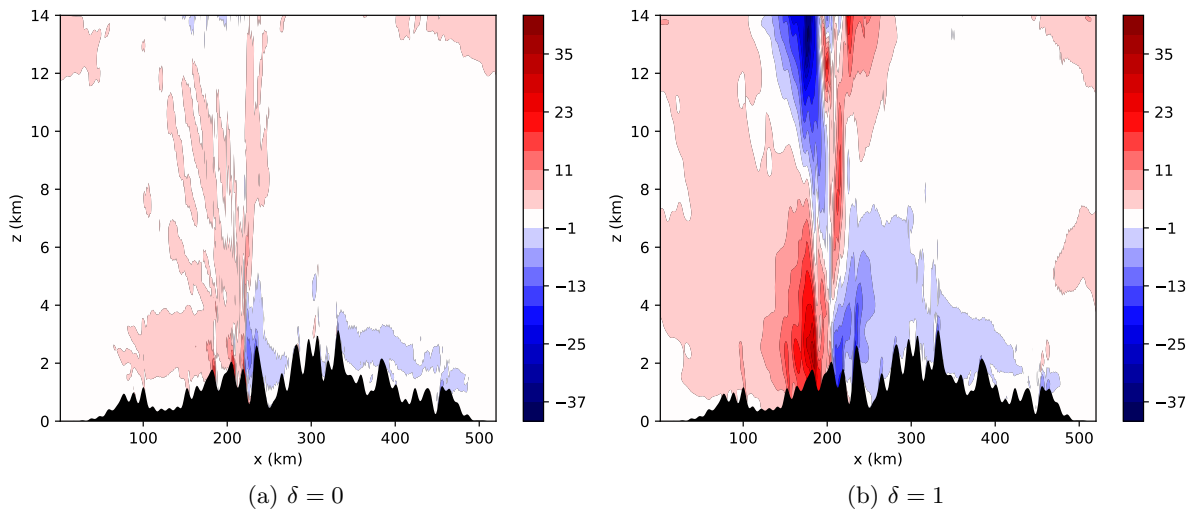


FIGURE 4.7 – Champ de vent horizontal avec  $\Delta x = 1300$  m,  $\Delta t = 50$  s après  $T = 15\,000$  s

#### 4.4. EXPÉRIENCES NUMÉRIQUES DANS UN CADRE 2D

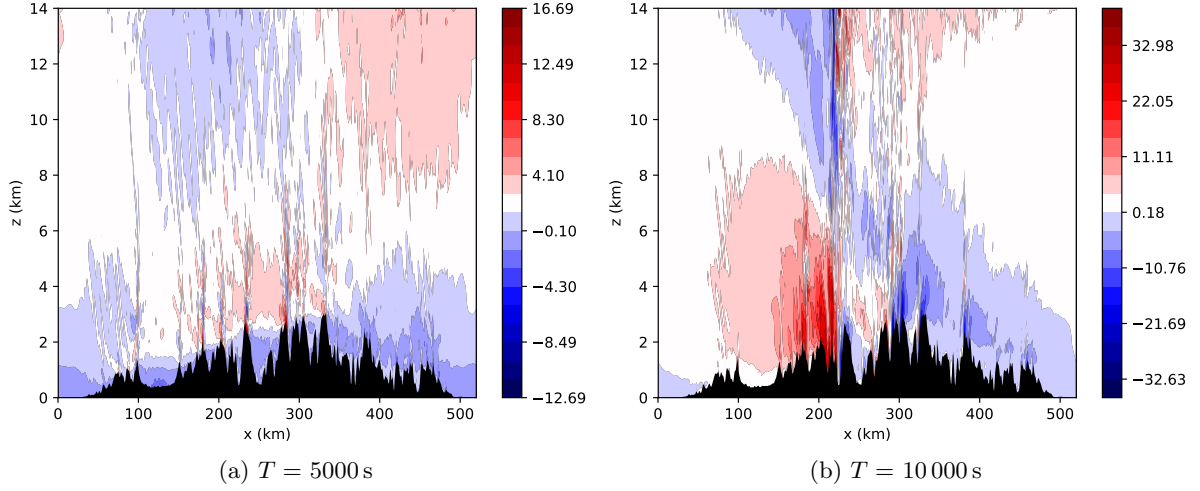


FIGURE 4.8 – Champ de vent horizontal avec  $\Delta x = 325$  m  $\Delta t = 12,5$  s pour  $\delta = 1$ .

$\Delta t$	$\Delta x$	$G_{max}$		$\delta = 0$	$\delta = 1$
50 s	1300 m	0.36	$T_{lim}$	15 900 s	18 600 s
			$N$	73	73
25 s	650 m	0.50	$T_{lim}$	8900 s	10 000 s
			$N$	75	75
12,5 s	325 m	0.75	$T_{lim}$	6363 s	10 375 s
			$N$	77	77

TABLE 4.2 – Temps limites  $T_{lim}$  et nombre d’itérations moyen  $N$  pour inverser le problème (non projeté sur les modes propres de l’opérateur vertical) à une précision  $\varepsilon = 10^{-8}$ , pour les écoulements sur le relief des Alpes. La pente maximale  $G_{max}$  (troisième colonne) est rappelée pour chacune des trois configurations.

#### 4.4.6 Estimation de la vitesse de convergence

Les expériences de la partie précédente ont été menées en prenant un critère d’arrêt strict ( $\varepsilon = 10^{-8}$ ) pour que les différents comportements observés soient attribués uniquement à l’étude de la stabilité numérique, et non à celle de la convergence du solveur. Aucune étude poussée n’a donc été menée pour régler finement le critère d’arrêt jusqu’ici. Ainsi avec ce critère, plus de 70 itérations sont nécessaires, alors qu’il en fallait significativement moins d’après les expériences menées dans le chapitre 3, lorsqu’un critère d’arrêt plus permissif était utilisé.

Quelques mesures du nombre d’itérations sont données dans la Table 4.2. On observe que ce nombre est identique entre les deux versions. Par conséquent, la faible dégradation du



conditionnement évoquée dans l'article de la version  $\delta = 1$  par rapport à la version  $\delta = 0$ , n'est pas perceptible pour les pentes considérées ici en présence d'un relief plus réaliste. Notons toutefois, que bien que plus réaliste qu'un relief linéaire, ce relief a été fortement filtré. De plus, à mesure que la pente augmente, le nombre d'itérations nécessaires pour respecter le critère  $\varepsilon = 10^{-8}$ , augmente également, comme suggéré par l'étude du conditionnement de l'article.

## 4.5 Conclusion

Dans ce chapitre on a étudié des alternatives aux problèmes numériques qui apparaissent lorsque la version actuelle (à coefficients constants) du modèle AROME est confrontée à des pentes supérieures à l'unité. Pour ce faire, nous avons mené des analyses de stabilité et des expériences numériques sur le modèle AROME 2D déjà présenté dans le chapitre précédent.

Malgré les hypothèses fortes sur lesquelles reposent ces analyses de stabilité, nous avons retrouvé pour la version à coefficients constants, l'ordre de grandeur de la pente maximale pour lequel le modèle est numériquement stable. L'une des raisons avancée pour l'expliquer, est que nous avons considéré des modes de stabilité physique quelconque (proches de ceux rencontrés dans la réalité), contrairement à la plupart des analyses qui se restreignent à des modes physiquement stables. Cette analyse nous a également montré que : diminuer le pas de temps, augmenter le nombre d'étapes du schéma ICI, ou changer la température froide de l'opérateur linéaire, ne permettait pas d'améliorer significativement la stabilité numérique. Ainsi, les résultats issus de ces analyses suggèrent que la version à coefficients constants ne permet pas d'aborder les fortes pentes qui apparaissent aux échelles hectométriques. Notons que ces analyses reposent sur des hypothèses fortes, qui par conséquent, limitent la portée de certaines conclusions. Par exemple les bénéfices ou inconvénients sur la stabilité numérique apportés par le traitement du terme croisé X dans l'opérateur implicite n'ont pas pu être tranchés et sont laissés en perspectives.

De plus, nous avons développé une version en points de grille à coefficients variables du modèle AROME, construite à partir de la version points de grille à coefficients constants du chapitre 3. Dans cette nouvelle version, les termes orographiques sont désormais traités dans l'opérateur linéaire, tout en conservant les autres hypothèses des schémas à coefficients constants. Le problème implicite associé est inversé à l'aide d'un seul solveur de Krylov dans le cas particulier où une coordonnée  $\sigma$  est utilisée et que le terme croisé X est traité dans le schéma de transport semi-Lagrangien.

Les analyses indiquent que ces schémas à coefficients variables permettent d'augmenter significativement la stabilité numérique, sans poser de problèmes d'inversibilité pour une large gamme de pentes, et sans détériorer trop fortement le conditionnement du problème implicite à inverser. Les expériences numériques ont confirmé ces résultats : un gain de l'ordre d'une dizaine de degrés sur la pente maximale par rapport à la version à coefficients constants est atteint, sans augmentation significative du nombre d'itérations du solveur, y compris pour des pentes très fortes. Des cas tests sur des reliefs réalistes ont confirmé les résultats de ceux obtenus sur des reliefs idéalisés. Notons que la vitesse de convergence du solveur pourrait

## 4.5. CONCLUSION

---

probablement être améliorée en le paramétrant précisément comme nous l'avons fait dans le chapitre 3. Notons également qu'aucun préconditionneur sur la verticale n'est utilisé, pour pallier l'absence de projection sur les modes propres de l'opérateur vertical utilisée dans les schémas à coefficients constants.

Ainsi, le traitement implicite des termes orographiques ne semble pas générer les comportements numériques néfastes rencontrés parfois dans la littérature (problème d'inversibilité dans le noyau dynamique « New dynamics » (Davies *et al.*, 2005)) lorsque les hypothèses portant sur l'état de base (H1)–(H4) sont par ailleurs maintenues.

Notons également que ces résultats, bien que prometteurs ont été obtenus dans le cas où une coordonnée  $\sigma$  est utilisée. La généralisation à la coordonnée  $\eta$ , devrait conduire à quelques complexifications supplémentaires, mentionnées dans la section 4.1, et laissées en perspectives.



# Conclusion et perspectives

Les centres météorologiques se préparent à aborder les enjeux futurs de la Prévision Numérique du Temps : faire tourner d'ici quelques années, des modèles opérationnels de résolution hectométrique sur des machines massivement parallèles et hybrides. Pour atteindre cet objectif ambitieux, les méthodes numériques, qui se situent à l'interface entre la science et la technologie, doivent être adéquatement dimensionnées et ont été l'objet d'étude principal de cette thèse.

Nous avons illustré via l'exemple de six modèles à aire limitée, que la plupart d'entre eux utilisent aujourd'hui, soit des modèles Horizontalement Explicite Verticalement Implicite (HEVI), soit des modèles semi-implicites (SI), pour intégrer les équations d'Euler compressibles non-hydrostatiques. Les modèles HEVI exposés dans cette thèse, utilisent des schémas split-explicit pour intégrer les différents termes des équations avec des pas de temps spécifiques pour intégrer les processus lents et les processus rapides. Ces schémas sont soumis à de sévères contraintes de stabilité numérique sur ces différents pas de temps. Cela conduit à multiplier le nombre d'étapes intermédiaires pour fournir une prévision à une échéance donnée, et impacte finalement l'efficacité générale du modèle. Néanmoins, en ne faisant qu'appliquer des opérateurs locaux, les modèles HEVI sont bien adaptés au paradigme de parallélisation actuel, et ont été utilisés comme référence en terme de scalabilité dans cette étude. Les modèles semi-implicites exposés dans cette thèse, approchent le schéma de Crank-Nicolson via un schéma ICI (Iterative Centered Implicit). Un opérateur linéaire est introduit et traite les termes responsables de la propagation des ondes les plus rapides, conduisant à résoudre un problème implicite d'Helmholtz. Ce problème est inversé via un solveur, qui dans certains cas, génère de nombreuses communications, non adaptées aux machines massivement parallèles. Toutefois, cette méthode permet de s'affranchir des contraintes de stabilité les plus fortes, et autorise de grands pas de temps d'intégration en réduisant le nombre d'étapes intermédiaires pour fournir une prévision à une échéance donnée. Bien que tous les modèles SI exposés dans cette thèse privilégient de grands pas de temps d'intégration, nous avons montré qu'une partie du signal non-hydrostatique pouvait être détérioré par un tel choix.

Pour aborder les enjeux du futur, nous avons exposé les nouveaux schémas ou nouvelles discrétisations qui sont explorés et qui dépendent des objectifs des différents centres météorologiques. Par exemple, pour n'utiliser qu'un seul modèle couvrant une large variété d'applications (météorologie, climat, etc), les approches en volume finis sur des maillages moins structurés, avec des schémas numériques plus conservatifs, sont actuellement étudiées. D'autres études cherchent des alternatives aux schémas split-explicit utilisés dans les modèles HEVI, comme par exemple les schémas IMEX. Pour le modèle à aire limitée AROME, l'un des enjeux principal de ces prochaines années, est de le faire tourner efficacement à une résolution horizontale d'une centaine de mètres sur son domaine géographique actuel, c'est-à-dire en présence d'une orographie fortement chahutée. Cela n'est pas forcément possible avec la version actuelle pour deux raisons :

- les champs sont décomposés en série de Fourier par l'utilisation d'un algorithme de

transformée de Fourier rapide, qui génère de nombreuses communications globales lors de la transposition du vecteur d'état et pose un problème de scalabilité ;

- le problème implicite ne contient pas les termes orographiques, qui sont traités dans la partie explicite du schéma ICI, générant des instabilités numériques dans les zones où leur contribution est la plus forte, c'est-à-dire au voisinage des fortes pentes.

Dans les deux cas, l'opérateur linéaire à coefficients constants est à l'origine de ces deux problèmes : parce qu'il permet d'utiliser une discrétisation spectrale d'une part, et parce qu'il ne contient pas les termes orographiques d'autre part.

Dans le chapitre 3, nous avons ainsi dérivé un système d'équations en point de grille dans une version à coefficients constants, où les dérivées horizontales sont discrétisées via des schémas en différences finies sur une grille co-localisée. Le problème implicite est projeté dans l'espace propre de l'opérateur vertical, permettant ainsi de résoudre séquentiellement  $N_{lev}$  problèmes bidimensionnels symétriques et définis positifs (où  $N_{lev}$  désigne le nombre de niveaux verticaux du modèle). Ces problèmes sont inversés via un algorithme de Krylov, de gradient conjugué, qui génère principalement des communications locales, peu nombreuses quand le solveur converge rapidement. Une version partiellement portée en point de grille a ainsi été développée dans un cycle récent du modèle AROME mais où une partie seulement des opérateurs de dérivée horizontale  $y$  sont calculés en différences finies et où certaines autres étapes restent calculées dans l'espace spectral comme par exemple la diffusion numérique. Nous avons montré que la vitesse de convergence pouvait être considérablement accélérée en : initialisant astucieusement le solveur, et en lui prescrivant un critère d'arrêt spécifique basé sur un nombre d'itérations maximal à ne pas dépasser, plutôt que sur un critère basé sur une tolérance limite. Suite à ce paramétrage et grâce notamment à la projection sur les modes propres de l'opérateur vertical, une convergence très rapide a pu être atteinte, même sans préconditionneurs. Nous avons montré sur de nombreuses expériences, dont l'une portant sur un phénomène à enjeux, que cette version partiellement en point de grille, ne dégradait pas significativement la qualité de la prévision par rapport à la version spectrale ou par rapport aux observations.

Le coût en communications d'un solveur étant principalement régi par sa vitesse de convergence, nous avons pu dériver des mesures indirectes de scalabilité. Les résultats suggèrent qu'une scalabilité comparable à celle d'un modèle HEVI peut être atteinte pour cette nouvelle version. Toutefois des tests complémentaires par des mesures directes, sont nécessaires pour affiner ce résultat. Ces tests poussés de scalabilité ne pourront se faire que lorsque l'intégralité du code parallélisé aura été porté en point de grille, et que tous les calculs effectués dans l'espace spectral auront été supprimés. Une première étape a été franchie dans cette direction, lorsque les nouvelles routines implémentées dans la thèse ont été parallélisées sous le protocole MPI.

Néanmoins d'autres étapes nécessitent encore d'être franchies. L'une d'entre-elles est de trouver une alternative pour traiter la diffusion numérique. Une autre serait de trancher le choix de la discrétisation des opérateurs de dérivée horizontale sur la grille co-localisée, ce que n'ont pas permis les hypothèses de ce travail de thèse. La stratégie qui nous semble la plus viable, mais qui nécessite des tests complémentaires, est de conserver les variables pronostiques

de divergence et de vorticit  pour la partie lin aire, sur lesquelles le probl me implicite est r duit alg briquement, et de reconstituer le champ de vent   partir des  quations classiques de la quantit  de mouvement, sans inverser de probl mes de Poisson. Conserver une discr tisation consistante entre les d riv es premi re et seconde, nous para t  galement souhaitable.

En plus d'am liorer probablement la scalabilit  du mod le, l'abandon des transform es spectrales ouvre de nouvelles opportunit s. Ainsi, contrairement   la version spectrale, dont l'un des objectifs  tait de r duire le nombre des transform es en augmentant fortement le pas de temps, la version point de grille para t plus attractive pour int grer les  quations avec de plus petits pas de temps et ainsi am liorer la description des ph nom nes non-hydrostatiques de petite  chelle. Notons  galement que l'abandon des transform es spectrales ne n cessite plus de disposer d'un domaine bi-p riodique et ouvre de nouvelles possibilit s pour imposer les conditions aux limites lat rales. Enfin, suite aux modifications effectu es dans le code durant la th se, il nous semble que l'architecture du code est telle qu'une nouvelle version enti rement en point de grille pourrait  tre d velopp e sans n cessiter d'investissements majeurs. Une fois une telle version impl ment e, des comparaisons de scalabilit  pourraient  tre men es entre cette version et la version actuelle. En cas de r sultat concluant, cette nouvelle version pourrait remplacer la version spectrale dans un contexte op rationnel.

Des analyses ont ensuite  t  men es pour  tudier la stabilit  num rique des sch mas   coefficients constants. Malgr  leur simplicit , ces analyses sont parvenues   reproduire les instabilit s num riques principales rencontr es dans le mod le AROME en pr sence de pentes rencontr es aux r solutions horizontales hectom triques. Aucune des strat gies bas e sur des sch mas   coefficients constants, n'a am lior e significativement la stabilit  num rique en pr sence de fortes pentes dans ces analyses. En revanche, des gains significatifs ont  t  obtenus lorsque les termes orographiques ont  t  ajout s dans le probl me implicite, g n rant un probl me   coefficients variables.

Ces r sultats ont  t  confirm s pour diff rents  coulements bidimensionnels en pr sence d'un relief plus ou moins id alis . Un gain d'une dizaine de degr s sur la pente maximale admissible du relief a ainsi  t  obtenu par rapport   la version   coefficients constants. Ce gain en stabilit  num rique n'a pas g n r  de probl me d'inversibilit  ou de probl me de convergence du solveur de Krylov pour la large gamme de pentes test es, contrairement   ce qui a pu  tre rencontr  dans la litt rature lorsqu'un  tat de base trop r aliste est consid r . Ainsi, prescrire un  tat de base atmosph rique id alis  (isotherme, sans vent, hydrostatique) en pr sence d'une orographie r aliste, nous para t  tre une option cr dible pour am liorer la stabilit  num rique   un co t de calcul acceptable.

En maintenant une coordonn e  $\sigma$ , et un traitement du terme crois  X dans le sch ma de transport semi-Lagrangien, un codage de ces nouvelles modifications dans le mod le AROME, para t envisageable. Des tests en 3D pourraient ainsi  tre men s et les estimations de stabilit , affin es. La g n ralisation   la coordonn e  $\eta$ , bien que probablement surmontable, nous para t cependant plus d licate. De plus, si la version   coefficients constants ne n cessite pas de pr conditionneurs pour acc l rer l'inversion du probl me implicite, ces derniers semblent plus indispensables pour les sch mas   coefficients variables, o  la projection sur les modes propres des op rateurs verticaux n'est plus triviale.

Ainsi, les méthodes en point de grille apparaissent plus flexibles pour aborder les enjeux du futur, tant sur le plan scientifique, pour simuler des écoulements d'échelle hectométriques, que sur le plan technologique, pour utiliser des machines de calculs toujours plus complexes.

# Discrétisation horizontale

Dans cette partie, on s'intéresse à la discrétisation en différences finies des dérivées horizontales  $\partial_x$  et  $\partial_x^2$ . Le système le plus simple pour étudier les différentes discrétisations est le système shallow-Water 1D dont les équations sont de la forme :

$$\partial_t u = -c \partial_x p \quad (\text{A.1})$$

$$\partial_t p = -c \partial_x u \quad (\text{A.2})$$

où  $c$  est la vitesse des ondes rapides,  $u$  désigne la vitesse horizontale du vent, et  $p$  la pression.

Pour résoudre ce système avec un schéma semi-implicite sur une grille A et en réduisant algébriquement le problème, deux approches peuvent être considérées et aboutissent à deux problèmes différents. La première consiste à discrétiser temporellement puis spatialement le système alors que la seconde consiste à le discrétiser spatialement puis temporellement. Comme l'une ou l'autre de ces méthodes peuvent être utilisées, une étude spécifique est requise pour chacune d'entre elles.

## A.1 Discrétisation spatiale puis temporelle

On discrétise spatialement (A.1) et (A.2) avec un schéma en différences finies d'ordre 2 :

$$\partial_t u = -c \delta_x p$$

$$\partial_t p = -c \delta_x u$$

où  $\delta_x = [-0.5, 0, 0.5] / \Delta x$ . On discrétise ensuite temporellement avec un schéma semi-implicite (SI) :

$$u^+ = u^- - c \delta t \delta_x (p^+ + p^-)$$

$$p^+ = p^- - c \delta t \delta_x (u^+ + u^-)$$

où  $\delta t = \Delta t / 2$  pour un schéma à 2 niveaux temporels et  $\delta t = \Delta t$  pour un schéma à 3 niveaux temporels. On réduit algébriquement le problème sur l'une des deux variables :

$$u^+ - c^2 \delta t^2 \delta_x \delta_x u^+ = u^- + c^2 \delta t^2 \delta_x \delta_x u^- - 2c \delta t \delta_x p^-$$



où  $\delta_x \delta_x = [0.25, 0, -0.5, 0, 0.25] / \Delta x^2 := \delta_x^2$ . Dans cette approche l'opérateur de dérivée seconde est égal au produit des dérivées premières :

$$\delta_x^2 = \delta_x \delta_x \quad (\text{A.3})$$

On peut réécrire le système général sous la forme matricielle :

$$\left[ \begin{array}{c|c} 1 - c^2 \delta t^2 \delta_x \delta_x & 0 \\ \hline c \delta t \delta_x & 1 \end{array} \right] \begin{bmatrix} u^+ \\ p^+ \end{bmatrix} = \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \delta_x \delta_x & -2c \delta t \delta_x \\ \hline -c \delta t \delta_x & 1 \end{array} \right] \begin{bmatrix} u^- \\ p^- \end{bmatrix}$$

On considère un système non borné ou périodique spatialement et on effectue la transformée de Fourier, le système peut alors se réécrire sous la forme :

$$\left[ \begin{array}{c|c} 1 - c^2 \delta t^2 \tilde{k}_2^2 & 0 \\ \hline c \delta t \tilde{k}_2 & 1 \end{array} \right] \begin{bmatrix} \tilde{u}^+ \\ \tilde{p}^+ \end{bmatrix} = \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \tilde{k}_2^2 & -2c \delta t \tilde{k}_2 \\ \hline -c \delta t \tilde{k}_2 & 1 \end{array} \right] \begin{bmatrix} \tilde{u}^- \\ \tilde{p}^- \end{bmatrix}$$

où  $\tilde{k}_2 = i \sin(k \Delta x) / \Delta x$ , qui peut aussi s'écrire comme :

$$\begin{bmatrix} \tilde{u}^+ \\ \tilde{p}^+ \end{bmatrix} = \frac{1}{1 - c^2 \delta t^2 \tilde{k}_2^2} \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \tilde{k}_2^2 & -2c \delta t \tilde{k}_2 \\ \hline -2c \delta t \tilde{k}_2 & 1 + c^2 \delta t^2 \tilde{k}_2^2 \end{array} \right] \begin{bmatrix} \tilde{u}^- \\ \tilde{p}^- \end{bmatrix}$$

## A.2 Discrétisation temporelle puis spatiale

On discrétise temporellement avec un schéma SI :

$$u^+ = u^- - c \delta t \partial_x (p^+ + p^-)$$

$$p^+ = p^- - c \delta t \partial_x (u^+ + u^-)$$

On peut réduire algébriquement le problème sur l'une des deux variables :

$$u^+ - c^2 \delta t^2 \partial_x^2 u^+ = u^- + c^2 \delta t^2 \partial_x^2 u^- - 2c \delta t \partial_x p^-$$

Puis on discrétise les deux équations spatialement, et on obtient alors :

$$u^+ - c^2 \delta t^2 \delta_x^2 u^+ = u^- + c^2 \delta t^2 \delta_x^2 u^- - 2c \delta t \delta_x p^-$$

$$p^+ + c \delta t \delta_x u^+ = p^- - c \delta t \delta_x u^-$$

où  $\delta_x = [-0.5, 0, 0.5] / \Delta x$  et  $\delta_x^2 = [1, -2, 1] / \Delta x^2$ . Dans ce cas, on note que les opérateurs de dérivée première et seconde ne sont pas consistants entre eux : l'opérateur de dérivée seconde n'est pas égal au produit des dérivées premières :

$$\delta_x^2 \neq \delta_x \delta_x \quad (\text{A.4})$$

### A.3. ANALYSE DE VON-NEUMANN

---

On peut réécrire ce système sous la forme matricielle :

$$\left[ \begin{array}{c|c} 1 - c^2 \delta t^2 \delta_x^2 & 0 \\ \hline c \delta t \delta_x & 1 \end{array} \right] \left[ \begin{array}{c} u^+ \\ p^+ \end{array} \right] = \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \delta_x^2 & -2c \delta t \delta_x \\ \hline -c \delta t \delta_x & 1 \end{array} \right] \left[ \begin{array}{c} u^- \\ p^- \end{array} \right]$$

On considère un système non borné ou périodique spatialement et on effectue la transformée de Fourier, le système peut alors se réécrire sous la forme :

$$\left[ \begin{array}{c|c} 1 - c^2 \delta t^2 \tilde{k}_1^2 & 0 \\ \hline c \delta t \tilde{k}_2 & 1 \end{array} \right] \left[ \begin{array}{c} \tilde{u}^+ \\ \tilde{p}^+ \end{array} \right] = \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \tilde{k}_1^2 & -2c \delta t \tilde{k}_2 \\ \hline -c \delta t \tilde{k}_2 & 1 \end{array} \right] \left[ \begin{array}{c} \tilde{u}^- \\ \tilde{p}^- \end{array} \right]$$

qui peut aussi s'écrire comme :

$$\left[ \begin{array}{c} \tilde{u}^+ \\ \tilde{p}^+ \end{array} \right] = \frac{1}{1 - c^2 \delta t^2 \tilde{k}_1^2} \left[ \begin{array}{c|c} 1 + c^2 \delta t^2 \tilde{k}_1^2 & -2c \delta t \tilde{k}_2 \\ \hline -2c \delta t \tilde{k}_2 & 1 - c^2 \delta t^2 \tilde{k}_1^2 + 2c^2 \delta t^2 \tilde{k}_2^2 \end{array} \right] \left[ \begin{array}{c} \tilde{u}^- \\ \tilde{p}^- \end{array} \right]$$

## A.3 Analyse de Von-Neumann

### A.3.1 Discrétisation spatiale puis temporelle

Les valeurs propres sont :

$$\lambda_1 = \frac{(1 - c \delta t \tilde{k}_2)^2}{1 - c^2 \delta t^2 \tilde{k}_2^2}$$

$$\lambda_2 = \frac{(1 + c \delta t \tilde{k}_2)^2}{1 - c^2 \delta t^2 \tilde{k}_2^2}$$

On en déduit que :

$$|\lambda_1| = 1$$

$$|\lambda_2| = 1$$

On en déduit que le schéma est neutre en terme de stabilité, quel que soit le pas de temps considéré.

### A.3.2 Discrétisation temporelle puis spatiale

On peut montrer numériquement que quel que soit  $k \Delta x$  considéré (cf Figure A.1) :

$$|\lambda_{1,2}| \leq 1$$

Ainsi un amortissement des ondes a lieu quel que soit le pas de temps considéré.

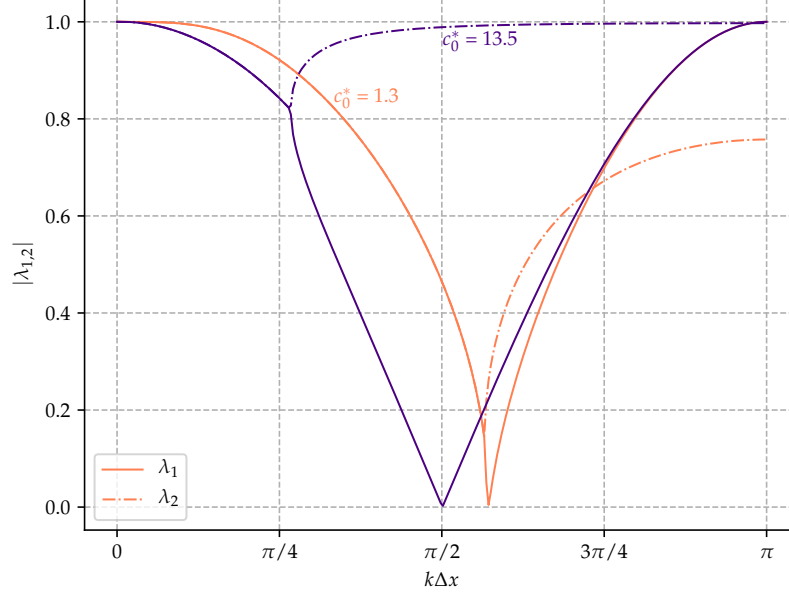


FIGURE A.1 – Facteur d’amplification de chacune des valeurs propres en fonction du nombre d’onde  $k\Delta x$  pour deux nombres de Courant  $c_0^*$  différents lorsque les opérateurs de dérivée première et seconde sont non consistants entre eux.

#### A.4 Analyse en modes normaux

On considère des modes périodiques temporellement, c’est-à-dire tels que :

$$\tilde{X} = \hat{X} \exp(i\omega t)$$

Par conséquent,

$$\tan(2\omega_{1,2}\delta t) = \frac{\Im(\lambda_{1,2})}{\Re(\lambda_{1,2})}$$

où  $\Im$  et  $\Re$  désignent respectivement les parties imaginaires et réelles. On en déduit que :

$$\omega\delta t = \frac{1}{2} \arctan\left(\frac{\Im(\lambda)}{\Re(\lambda)}\right)$$

$$\omega\delta t = \begin{cases} \frac{1}{2} \arctan\left(\frac{\Im(\lambda)}{\Re(\lambda)}\right), & \text{si } \frac{\Im(\lambda)}{\Re(\lambda)} > 0 \\ \frac{1}{2} \arctan\left(\frac{\Im(\lambda)}{\Re(\lambda)}\right) + \frac{\pi}{2}, & \text{sinon} \end{cases}$$

En particulier, un comportement que l’on souhaite reporter le plus possible en extrémité

## A.5. CONCLUSION

du spectre (c'est-à-dire pour les grands nombre d'onde) est le nombre d'onde  $k_c$  tel que :

$$\frac{\partial \omega(k = k_c)}{\partial k} = 0$$

Les résultats présentés en Figure A.2 et Figure A.3 sont tracés dans trois cas différents :

- SI 3TL en spectral (discrétisation spatiale exacte :  $\delta_x^2 = \partial_x \partial_x$ ) :
- SI 3TL discrétisé spatialement puis temporellement (opérateurs de dérivée première et seconde consistants entre eux :  $\delta_x^2 = \delta_x \delta_x$ )
- SI 3TL discrétisé temporellement puis spatialement (opérateurs de dérivée première et seconde non consistants entre eux  $\delta_x^2 \neq \delta_x \delta_x$ )

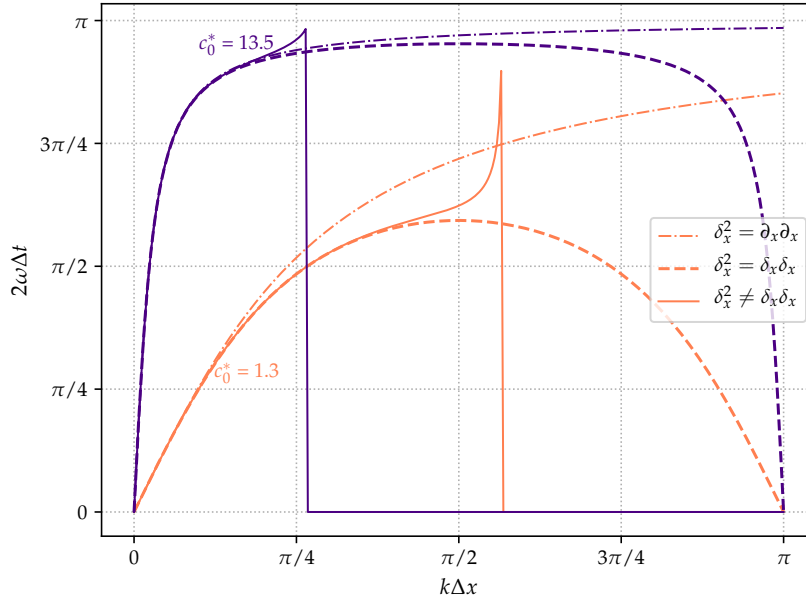


FIGURE A.2 – Pulsation normalisée en fonction du nombre d'onde  $k\Delta x$  pour deux nombres de Courant  $c_0^*$  différents.

La Figure A.3 montre que prendre des opérateurs non consistants entre eux reporte le problème d'annulation de la vitesse de groupe sur des plus grands nombre d'onde uniquement pour les nombres CFL  $c_0^* \leq 2$ . Or, les nombres de CFL couramment utilisés en PNT pour les schémas SI, sont supérieurs à 10. Par conséquent prendre des opérateurs consistants entre eux est une approche à privilégier.

## A.5 Conclusion

Pour éviter une annulation de la vitesse de groupe pour de trop petits nombres d'onde, ainsi qu'une atténuation trop forte du signal dans un contexte d'utilisation de grands pas de

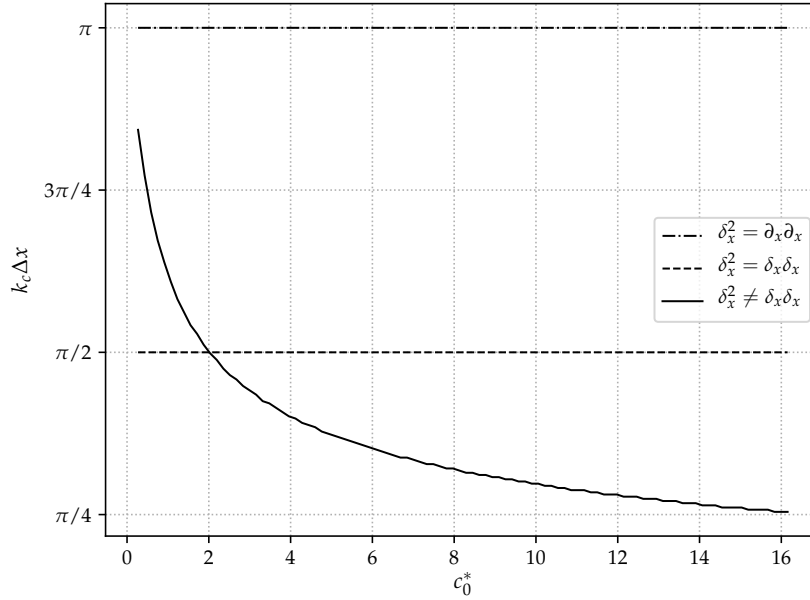


FIGURE A.3 – Nombre d'onde normalisé pour lequel la vitesse de groupe s'annule en fonction du nombre de Courant  $c_0^*$ .

temps, on privilégie l'approche consistant à discrétiser spatialement puis temporellement le système, c'est-à-dire consistant à choisir des opérateurs cohérents entre les dérivées première et seconde :

$$\delta_x = [-0.5, 0, 0.5] / \Delta x$$

$$\delta_x \delta_x = [0.25, 0, -0.5, 0, 0.25] / \Delta x^2$$

# Analyse de stabilité du gradient conjugué

Dans cette partie, on mène une analyse de stabilité pour la première itération de l'algorithme de gradient conjugué exposé dans le paragraphe 3.1.4, sur un système shallow-water en une seule dimension (A.1)(A.2) discrétisé temporellement avec un schéma semi-implicite à 2 niveaux temporels, et spatialement avec un schéma en différences finies au second ordre.

En réduisant algébriquement le problème sur l'une des variables pronostiques, comme par exemple la vitesse du vent, on obtient :

$$\left(1 - c^2 \delta t^2 \partial_x^2\right) u^+ = \left(1 + c^2 \delta t^2 \partial_x^2\right) u^- + 2\delta t c \partial_x p^- \quad (\text{B.1})$$

où  $\delta t = \Delta t/2$  pour un schéma à 2 niveaux temporels,  $u^+$  et  $u^-$  désignent respectivement le champ  $u$  aux instants  $t + \delta t$  et  $t$ . Pour plus de simplicité, considérons le cas particulier où le deuxième terme du membre de droite de l'équation précédente s'annule, on obtient alors :

$$\left(1 - c^2 \delta t^2 \partial_x^2\right) u^+ = \left(1 + c^2 \delta t^2 \partial_x^2\right) u^- \quad (\text{B.2})$$

En considérant que le domaine est périodique, on décompose le problème en série de Fourier et on obtient pour chacun des coefficients de Fourier  $\tilde{u}$  :

$$\left(1 + c^2 \delta t^2 \tilde{k}^2\right) \tilde{u}^+ = \left(1 - c^2 \delta t^2 \tilde{k}^2\right) \tilde{u}^- \quad (\text{B.3})$$

où  $-\tilde{k}^2$  est la valeur propre du laplacien discrétisé par un schéma en différences finies au second ordre.

— sur une grille Z :

$$-\tilde{k}^2 = \frac{2 \cos(k\Delta x) - 2}{\Delta x^2}, \quad (\text{B.4})$$

— sur une grille A :

$$-\tilde{k}^2 = \frac{\sin(k\Delta x)^2}{\Delta x^2}, \quad (\text{B.5})$$

dont les valeurs propres maximales sont :

— sur une grille Z :

$$-\tilde{k}_{max}^2 = \frac{-4}{\Delta x^2}, \quad (\text{B.6})$$

— sur une grille A :

$$-\tilde{k}_{max}^2 = \frac{-1}{\Delta x^2}. \quad (\text{B.7})$$

En introduisant le nombre sans dimension  $C_*$  comme :

$$C_*^2 = c^2 \delta t^2 \tilde{k}^2, \quad (\text{B.8})$$

Il devient, dans le cas d'une discrétisation au second ordre pour le nombre d'onde  $k_{max}$  :

— sur une grille Z :

$$C_*^2 = \left( c \frac{\Delta t}{\Delta x} \right)^2 \quad (\text{B.9})$$

— sur une grille A :

$$C_*^2 = 4 \left( c \frac{\Delta t}{\Delta x} \right)^2 \quad (\text{B.10})$$

Le système (B.3), se réécrit alors :

$$\left(1 + C_*^2\right) \tilde{u}^+ = \left(1 - C_*^2\right) \tilde{u}^- \quad (\text{B.11})$$

On note qu'en inversant le problème exactement (en utilisant par exemple un solveur direct et exact), on peut montrer (le résultat bien connu) que le carré du facteur d'amplification est inférieur à 1 quel que soit  $C_*$  :

$$\frac{|\tilde{u}^+|}{|\tilde{u}^-|} = \frac{|1 - C_*^2|}{|1 + C_*^2|} \leq 1. \quad (\text{B.12})$$

Dans ce cas très simplifié, le schéma est dit inconditionnellement stable.

Pour plus de lisibilité, on rappelle ici l'algorithme du gradient conjugué :

- 1 Calculer  $r_0 = y - \mathcal{H}x_0$  et  $p_0 = r_0$ ;
- 2 **tant que critère faire**
- 3     $\alpha_{n_i} = r_{n_i}^T r_{n_i} / p_{n_i}^T \mathcal{H} p_{n_i}$ ;
- 4     $x_{n_i+1} = x_{n_i} + \alpha_{n_i} p_{n_i}$ ;
- 5     $r_{n_i+1} = r_{n_i} - \alpha_{n_i} \mathcal{H} p_{n_i}$ ;
- 6     $\beta_{n_i} = r_{n_i+1}^T r_{n_i+1} / r_{n_i}^T r_{n_i}$ ;
- 7     $p_{n_i+1} = r_{n_i+1} + \beta_{n_i} p_{n_i}$ ;
- 8 **fin**

Pour plus de lisibilité, et en reprenant les notations de l'algorithme, on pose les notations suivantes :

$$x = u \quad (\text{B.13})$$

$$\mathcal{H} = 1 - c^2 \delta t^2 \partial_x^2 \quad (\text{B.14})$$

$$y = \left(1 + c^2 \delta t^2 \partial_x^2\right) u^- \quad (\text{B.15})$$

En initialisant le solveur par le champ du pas de temps précédent, le résidu initial est alors :

$$r_0 = y - \mathcal{H}x_0 \quad (\text{B.16})$$

L'itéré suivant s'écrit comme :

$$x_1 = x_0 + \alpha_0 r_0 \quad (\text{B.17})$$

Le domaine étant périodique,  $\mathcal{H}$  est une matrice circulante. Diagonalisons cette matrice :  $\mathcal{H} = PDP^{-1}$  où  $P$  est la matrice de passage, et  $D$  est la matrice diagonale qui contient les valeurs propres de  $\mathcal{H}$ .  $P$  est la matrice des coefficients de Fourier et  $P^{-1} = P^T = P$  est sa matrice inverse, égale à sa matrice de passage (matrice unitaire). En décomposant les équations (B.13), (B.14), (B.15), (B.16) et (B.17) en série de Fourier, on obtient :

$$\tilde{x} = \tilde{u}, \quad (\text{B.18})$$

$$\tilde{\mathcal{H}} = 1 + C_*^2, \quad (\text{B.19})$$

$$\tilde{y} = (1 + C_*^2) \tilde{x}_0, \quad (\text{B.20})$$

$$\tilde{r}_0 = (1 - C_*^2) \tilde{x}_0 - (1 + C_*^2) \tilde{x}_0 = -2C_*^2 \tilde{x}_0 \quad (\text{B.21})$$

$$\tilde{x}_1 = \tilde{x}_0 + \alpha_0 \tilde{r}_0 = \tilde{x}_0 - 2\alpha_0 C_*^2 \tilde{x}_0 \quad (\text{B.22})$$

On introduit le taux de croissance numérique  $\Gamma$  :

$$\Gamma = \frac{|\tilde{x}_1|}{|\tilde{x}_0|} = |1 - 2\alpha_0 C_*^2| \quad (\text{B.23})$$

Le système est dit stable numériquement si son taux de croissance numérique est inférieur à 1, c'est-à-dire si :

$$\alpha_0 C_*^2 \leq 1 \quad (\text{B.24})$$

Or, le coefficient  $\alpha_0$  de l'algorithme peut se réécrire comme :

$$\alpha_0 = \frac{r_0^T P P^{-1} r_0}{r_0^T P D P^{-1} r_0} \quad (\text{B.25})$$

ou bien comme :

$$\alpha_0 = \frac{(P^{-1} r_0)^T (P^{-1} r_0)}{(P^{-1} r_0)^T P D (P^{-1} r_0)}. \quad (\text{B.26})$$

En décomposant le résidu initial (B.16)  $r_0$  en série de Fourier dont les coefficients de Fourier sont  $\tilde{r}_0$ , on obtient :

$$\alpha_0 = \frac{\tilde{r}_0^T \tilde{r}_0}{\tilde{r}_0^T D \tilde{r}_0} \quad (\text{B.27})$$

Les valeurs propres de  $\mathcal{H}$  étant toutes réelles et supérieures à 1, on en déduit que :

$$\alpha_0 \in [0,1] \quad (\text{B.28})$$



On en déduit ainsi que (B.23) devient, pour le nombre d'onde  $k_{max}$  le plus contraignant pour la stabilité numérique :

— sur une grille Z :

$$C_* = \frac{c\Delta t}{\Delta x} \leq 1 \quad (\text{B.29})$$

— sur une grille A :

$$C_* = \frac{c\Delta t}{\Delta x} \leq 2 \quad (\text{B.30})$$

Ces deux contraintes sont les fameuses contraintes CFL. Cette analyse indique donc que la stabilité numérique est garantie en une seule itération de l'algorithme du gradient conjugué, lorsque le nombre CFL du problème est inférieur à 1 pour une grille A et 2 pour une grille Z.

Dans la pratique cette propriété est exploitée dans le paragraphe 3.5.2, pour les équations, plus complexes, du modèle AROME. En effet, de nombreux modes ont des vitesses de phase  $c = \sqrt{b_l}$ , telles que ces deux contraintes sont satisfaites. Pour ces modes, une seule itération du gradient conjugué serait ainsi nécessaire pour inverser le problème en garantissant la stabilité numérique. Enfin, notons que seule la stabilité numérique est évaluée ici, et non la qualité de la solution.

# Bibliographie

- Allaire, Grégoire (2005). *Analyse numérique et optimisation : Une introduction à la modélisation mathématique et à la simulation numérique*. Editions Ecole Polytechnique (cf. p. 19).
- Asanović, Krste, Ras Bodik, Bryan Christopher Catanzaro, Joseph James Gebis, Parry Husbands, Kurt Keutzer, David A. Patterson, William Lester Plishker, John Shalf, Samuel Webb Williams et Katherine A. Yelick (2006). *The Landscape of Parallel Computing Research : A View from Berkeley*. Rapp. tech. UCB/EECS-2006-183. EECS Department, University of California, Berkeley (cf. p. 12).
- Axelsson, O. (1996). *Iterative Solution Methods*. Cambridge University Press, p. 558. ISBN : 9780521555692 (cf. p. 72).
- Barros, S.R.M., D. Dent, L. Isaksen, G. Robinson, G. Mozdzynski et F. Wollenweber (1995). « The IFS model : A parallel production weather code ». In : *Parallel Computing* 21.10, p. 1621-1638. DOI : 10.1016/0167-8191(96)80002-0 (cf. p. 12).
- Bauer, Peter, Alan Thorpe et Gilbert Brunet (2015). « The quiet revolution of numerical weather prediction ». In : *Nature* 525.7567, p. 47-55. DOI : 10.1038/nature14956 (cf. p. 14).
- Benacchio, Tommaso et Nigel Wood (2016). « Semi-implicit semi-Lagrangian modelling of the atmosphere : a Met Office perspective ». In : *Communications in Applied and Industrial Mathematics* 7.3, p. 4-25. DOI : 10.1515/caim-2016-0020 (cf. p. 19, 35).
- 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 ». In : *Monthly Weather Review* 132.5, p. 1319-1324. DOI : 10.1175/1520-0493(2004)132<1319:otuoaw>2.0.co;2 (cf. p. 35, 51).
- Bénard, P., R. Laprise, J. Vivoda et P. Smolíková (2004). « Stability of Leapfrog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations : Flat-Terrain Case ». In : *Monthly Weather Review* 132.5, p. 1306-1318. DOI : 10.1175/1520-0493(2004)132<1306:solcss>2.0.co;2.
- Bénard, P., J. Mašek et P. Smolíková (2005). « Stability of Leapfrog Constant-Coefficients Semi-Implicit Schemes for the Fully Elastic System of Euler Equations : Case with Orography ». In : *Monthly Weather Review* 133.5, p. 1065-1075. DOI : 10.1175/mwr2907.1 (cf. p. 49, 120, 122).
- Bénard, P., J. Vivoda, J. Mašek, P. Smolíková, K. Yessad, Ch. Smith, R. Brožková et J.-F. Geleyn (2010). « Dynamical kernel of the Aladin-NH spectral limited-area model : Revised formulation and sensitivity experiments ». In : *Quarterly Journal of the Royal Meteorological Society* 136.646, p. 155-169. DOI : 10.1002/qj.522 (cf. p. 28).
- Bénard, Pierre (2003). « Stability of semi-implicit and iterative centered-implicit time discretizations for various equation systems used in NWP ». In : *Monthly weather review* 131.10, p. 2479-2491 (cf. p. 17, 49, 53, 120).
- Bénard, Pierre et Michael R. Ginton (2019). « Circumventing the pole problem of reduced lat-lon grids with local schemes. Part I : Analysis and model formulation ». In : *Quarterly Journal of the Royal Meteorological Society* 145.721, p. 1377-1391. DOI : 10.1002/qj.3509 (cf. p. 21, 26, 27).
- Bouttier, François, Benoît Vié, Olivier Nuissier et Laure Raynaud (2012). « Impact of Stochastic Physics in a Convection-Permitting Ensemble ». In : *Monthly Weather Review* 140.11, p. 3706-3721. DOI : 10.1175/mwr-d-12-00031.1 (cf. p. 46).
- Brousseau, Pierre, Yann Seity, Didier Ricard et Julien Léger (2016). « Improvement of the forecast of convective activity from the AROME-France system ». In : *Quarterly Journal of the Royal Meteorological Society* 142.699, p. 2231-2243. DOI : 10.1002/qj.2822 (cf. p. 19).
- Bush, Mike, Tom Allen, Caroline Bain, Ian Boutle, John Edwards, Anke Finnenkoetter, Charmaine Franklin, Kirsty Hanley, Humphrey Lean, Adrian Lock *et al.* (2020). « The first Met Office Unified Model-JULES Regional Atmosphere and Land configuration, RAL1 ». In : *Geoscientific Model Development* 13.4, p. 1999-2029 (cf. p. 19).
- Bénard, Pierre et Ján Mašek (2013). *Scientific Documentation for ALADIN-NH Dynamical Kernel*. Rapp. tech. Météo-France, CHMI (cf. p. 48, 50, 52, 55, 57).
- Caluwaerts, S. (2016). « Horizontal spatial discretization modularity within the spectral semi-implicit semi-Lagrangian ALADIN framework ». Thèse de doct. Ghent University (cf. p. 26, 69).
- Caluwaerts, Steven, Daan Degrauwe, Piet Termonia, Fabrice Voitus, Pierre Bénard et Jean-François Geleyn (2014). « Importance of temporal symmetry in spatial discretization for geostrophic adjustment in semi-

- implicit Z-grid schemes ». In : *Quarterly Journal of the Royal Meteorological Society* 141.686, p. 128-138. DOI : 10.1002/qj.2344 (cf. p. 26, 69).
- Carrassi, Alberto, Marc Bocquet, Laurent Bertino et Geir Evensen (2018). « Data assimilation in the geosciences : An overview of methods, issues, and perspectives ». In : *Wiley Interdisciplinary Reviews : Climate Change* 9.5, e535. DOI : 10.1002/wcc.535 (cf. p. 12).
- Charney, J. G., R. Fjortoft et J. Neumann (1950). « Numerical Integration of the Barotropic Vorticity Equation ». In : *Tellus* 2.4, p. 237-254. DOI : 10.1111/j.2153-3490.1950.tb00336.x (cf. p. 12, 15, 19).
- Colavolpe, Charles, Fabrice Voitus et Pierre Bénard (2017). « RK-IMEX HEVI schemes for fully compressible atmospheric models with advection : analyses and numerical testing ». In : *Quarterly Journal of the Royal Meteorological Society* 143.704, p. 1336-1350. DOI : 10.1002/qj.3008 (cf. p. 32).
- Cooley, James W. et John W. Tukey (1965). « An algorithm for the machine calculation of complex Fourier series ». In : *Mathematics of Computation* 19.90, p. 297-297. DOI : 10.1090/s0025-5718-1965-0178586-1 (cf. p. 37).
- Courant, R., K. Friedrichs et H. Lewy (1967). « On the Partial Difference Equations of Mathematical Physics ». In : *IBM Journal of Research and Development* 11.2, p. 215-234. DOI : 10.1147/rd.112.0215 (cf. p. 20).
- Davies, T, MJP Cullen, AJ Malcolm, MH Mawson, A Staniforth, AA White et N Wood (2005). « A new dynamical core for the Met Office's global and regional modelling of the atmosphere ». In : *Quarterly Journal of the Royal Meteorological Society* 131.608, p. 1759-1782 (cf. p. 35, 76, 151).
- Davies, Terry, Andrew Staniforth, Nigel Wood et John Thuburn (2003). « Validity of anelastic and other equation sets as inferred from normal-mode analysis ». In : *Quarterly Journal of the Royal Meteorological Society* 129.593, p. 2761-2775. DOI : 10.1256/qj.02.1951 (cf. p. 15).
- Degrauwe, Daan, Fabrice Voitus et Piet Termonia (2020). « A non-spectral Helmholtz solver for numerical weather prediction models with a mass-based vertical coordinate ». In : *Quarterly Journal of the Royal Meteorological Society*. DOI : 10.1002/qj.3902 (cf. p. 71, 77).
- Doms, G, J Förstner, E Heise, HJ Herzog, D Mironov, M Raschendorfer, T Reinhardt, B Ritter, R Schrodin, JP Schulz *et al.* (2011). « A description of the nonhydrostatic regional COSMO model ». In : *Part II : Physical Parameterization* 154 (cf. p. 19).
- Dubos, Thomas, Sarvesh Dubey, Marine Tort, Rashmi Mittal, Yann Meurdesoif et Frédéric Hourdin (2015). « DYNAMICO-1.0, an icosahedral hydrostatic dynamical core designed for consistency and versatility ». In : *Geoscientific Model Development* 8.10, p. 3131-3150 (cf. p. 22).
- Durrán, Dale R. (1990). « Mountain Waves and Downslope Winds ». In : *Atmospheric Processes over Complex Terrain*. American Meteorological Society, p. 59-81. DOI : 10.1007/978-1-935704-25-6\_4 (cf. p. 123, 142).
- Durrán, Dale R (2010). *Numerical methods for fluid dynamics : With applications to geophysics*. T. 32. Springer Science & Business Media (cf. p. 19).
- Fulton, Scott R., Paul E. Ciesielski et Wayne H. Schubert (1986). « Multigrid Methods for Elliptic Problems : A Review ». In : *Monthly Weather Review* 114.5, p. 943-959. DOI : 10.1175/1520-0493(1986)114<0943:mmfepa>2.0.co;2 (cf. p. 71).
- Gal-Chen, Tzvi et Richard C.J Somerville (1975). « On the use of a coordinate transformation for the solution of the Navier-Stokes equations ». In : *Journal of Computational Physics* 17.2, p. 209-228. DOI : 10.1016/0021-9991(75)90037-6 (cf. p. 22).
- Girard, Claude, André Plante, Michel Desgagné, Ron McTaggart-Cowan, Jean Côté, Martin Charron, Sylvie Gravel, Vivian Lee, Alain Patoine, Abdessamad Qaddouri *et al.* (2014). « Staggered vertical discretization of the Canadian Environmental Multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type ». In : *Monthly Weather Review* 142.3, p. 1183-1196 (cf. p. 27).
- Gustafsson, Nils et A. McDonald (1996). « A Comparison of the HIRLAM Gridpoint and Spectral Semi-Lagrangian Models ». In : *Monthly Weather Review* 124.9, p. 2008-2022. DOI : 10.1175/1520-0493(1996)124<2008:acothg>2.0.co;2 (cf. p. 27).
- Gustafsson, Nils, Tijana Janjić, Christoph Schraff, Daniel Leuenberger, Martin Weissmann, Hendrik Reich, Pierre Brousseau, Thibaut Montmerle, Eric Wattrelot, Antonín Bučánek, Máté Mile, Rafiq Hamdi, Magnus Lindskog, Jan Barkmeijer, Mats Dahlbom, Bruce Macpherson, Sue Ballard, Gordon Inverarity, Jacob Carley, Curtis Alexander, David Dowell, Shun Liu, Yasutaka Ikuta et Tadashi Fujita (2018). « Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres ». In : *Quarterly Journal of the Royal Meteorological Society* 144.713, p. 1218-1256. DOI : 10.1002/qj.3179 (cf. p. 46).

## BIBLIOGRAPHIE

---

- Holdaway, D, J Thurn et N Wood (2013). « Comparison of Lorenz and Charney–Phillips vertical discretisations for dynamics–boundary layer coupling. Part I : Steady states ». In : *Quarterly Journal of the Royal Meteorological Society* 139.673, p. 1073-1086 (cf. p. 27).
- Hortal, M. et A. J. Simmons (1991). « Use of Reduced Gaussian Grids in Spectral Models ». In : *Monthly Weather Review* 119.4, p. 1057-1074. DOI : 10.1175/1520-0493(1991)119<1057:uorggi>2.0.co;2 (cf. p. 21).
- Husain, Syed Zahid, Claude Girard, Abdessamad Qaddouri et André Plante (2019). « A New Dynamical Core of the Global Environmental Multiscale (GEM) Model with a Height-Based Terrain-Following Vertical Coordinate ». In : *Monthly Weather Review* 147.7, p. 2555-2578. DOI : 10.1175/mwr-d-18-0438.1 (cf. p. 42).
- Ishida, Junichi, Chiashi Muroi, Kohei Kawano et Yuji Kitamura (2010). « Development of a new nonhydrostatic model ASUCA at JMA ». In : *CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modelling* 40, p. 0511-0512 (cf. p. 19).
- Janjic, Z. I., J. P. Gerrity et S. Nickovic (2001). « An Alternative Approach to Nonhydrostatic Modeling ». In : *Monthly Weather Review* 129.5, p. 1164-1178. DOI : 10.1175/1520-0493(2001)129<1164:aaatnm>2.0.co;2 (cf. p. 15).
- Kasahara, Akira (1974). « Various Vertical Coordinate Systems Used for Numerical Weather Prediction ». In : *Monthly Weather Review* 102.7, p. 509-522. DOI : 10.1175/1520-0493(1974)102<0509:vvcsuf>2.0.co;2 (cf. p. 22, 47).
- Klemp, Joseph B. et Robert B. Wilhelmson (1978). « The Simulation of Three-Dimensional Convective Storm Dynamics ». In : *Journal of the Atmospheric Sciences* 35.6, p. 1070-1096. DOI : 10.1175/1520-0469(1978)035<1070:tsotdc>2.0.co;2 (cf. p. 31).
- Klemp, Joseph B, William C Skamarock et Oliver Fuhrer (2003). « Numerical consistency of metric terms in terrain-following coordinates ». In : *Monthly weather review* 131.7, p. 1229-1239 (cf. p. 55).
- Kühnlein, C., W. Deconinck, R. Klein, S. Malardel, Z. P. Piotrowski, P. K. Smolarkiewicz, J. Szmelter et N. P. Wedi (2019). « FVM 1.0 : a nonhydrostatic finite-volume dynamical core for the IFS ». In : *Geoscientific Model Development* 12.2, p. 651-676. DOI : 10.5194/gmd-12-651-2019 (cf. p. 22, 31, 39, 41).
- Lac, Christine, Jean-Pierre Chaboureau, Valéry Masson, Jean-Pierre Pinty, Pierre Tulet, Juan Escobar, Maud Leriche, Christelle Barthe, Benjamin Aouizerats, Clotilde Augros, Pierre Aumond, Franck Auguste, Peter Bechtold, Sarah Berthet, Soline Bielli, Frédéric Bosseur, Olivier Caumont, Jean-Martial Cohard, Jeanne Colin, Fleur Couvreur, Joan Cuxart, Gaëlle Delautier, Thibaut Dauhut, Véronique Ducrocq, Jean-Baptiste Filippi, Didier Gazen, Olivier Geoffroy, François Gheusi, Rachel Honnert, Jean-Philippe Lafore, Cindy Lebeaupin Brossier, Quentin Libois, Thibaut Lunet, Céline Mari, Tomislav Maric, Patrick Mascart, Maxime Mogé, Gilles Molinié, Olivier Nuissier, Florian Pantillon, Philippe Peyrillé, Julien Pergaud, Emilie Perraud, Joris Pianezze, Jean-Luc Redelsperger, Didier Ricard, Evelyne Richard, Sébastien Riette, Quentin Rodier, Robert Schoetter, Léo Seyfried, Joël Stein, Karsten Suhre, Marie Taufour, Odile Thouron, Sandra Turner, Antoine Verrelle, Benoît Vié, Florian Visentin, Vincent Vionnet et Philippe Wautelet (2018). « Overview of the Meso-NH model version 5.4 and its applications ». In : *Geoscientific Model Development* 11.5, p. 1929-1969. DOI : 10.5194/gmd-11-1929-2018 (cf. p. 15, 41).
- Laprise, René (1992). « The Euler equations of motion with hydrostatic pressure as an independent variable ». In : *Monthly weather review* 120.1, p. 197-207 (cf. p. 22, 47).
- Lax, P. D. et R. D. Richtmyer (1956). « Survey of the stability of linear finite difference equations ». In : *Communications on Pure and Applied Mathematics* 9.2, p. 267-293. DOI : 10.1002/cpa.3160090206 (cf. p. 19).
- Liesen, Jörg et Petr Tichý (2004). « Convergence analysis of Krylov subspace methods ». In : *GAMM-Mitteilungen* 27.2, p. 153-173 (cf. p. 76).
- Lipps, Franik B. et Richard S. Hemler (1982). « A Scale Analysis of Deep Moist Convection and Some Related Numerical Calculations ». In : *Journal of the Atmospheric Sciences* 39.10, p. 2192-2210. DOI : 10.1175/1520-0469(1982)039<2192:asaodm>2.0.co;2 (cf. p. 15).
- Machenhauer, Bennert et Jan E Haugen (1988). *Test of a Spectral Limited Area Shallow Water Model with Time-dependent Lateral Boundary Conditions and Combined Normal Mode/semi-Lagrangian Time Integration Schemes*. Dan. Meteorolog. Inst. (cf. p. 37).
- Masson, V., P. Le Moigne, E. Martin, S. Faroux, A. Alias, R. Alkama, S. Belamari, A. Barbu, A. Boone, F. Bouyssel, P. Brousseau, E. Brun, J.-C. Calvet, D. Carrer, B. Decharme, C. Delire, S. Donier, K. Essaouini, A.-L. Gibelin, H. Giordani, F. Habets, M. Jidane, G. Kerdraon, E. Kourzeneva, M. Lafaysse, S. Lafont,

- C. Lebeaupin Brossier, A. Lemonsu, J.-F. Mahfouf, P. Marguinaud, M. Mokhtari, S. Morin, G. Pigeon, R. Salgado, Y. Seity, F. Taillefer, G. Tanguy, P. Tulet, B. Vincendon, V. Vionnet et A. Voldoire (2013). « The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes ». In : *Geoscientific Model Development* 6.4, p. 929-960. DOI : 10.5194/gmd-6-929-2013 (cf. p. 46).
- Mengaldo, Gianmarco, Andrzej Wyszogrodzki, Michail Diamantakis, Sarah-Jane Lock, Francis X. Giraldo et Nils P. Wedi (2018). « Current and Emerging Time-Integration Strategies in Global Numerical Weather and Climate Prediction ». In : *Archives of Computational Methods in Engineering* 26.3, p. 663-684. DOI : 10.1007/s11831-018-9261-8 (cf. p. 12).
- Michalakes, John et Manish Vachharajani (2008). « GPU acceleration of numerical weather prediction ». In : *2008 IEEE International Symposium on Parallel and Distributed Processing*. IEEE. DOI : 10.1109/ipdps.2008.4536351 (cf. p. 12).
- Müller, Eike H et Robert Scheichl (2014). « Massively parallel solvers for elliptic partial differential equations in numerical weather and climate prediction ». In : *Quarterly Journal of the Royal Meteorological Society* 140.685, p. 2608-2624 (cf. p. 13).
- Müller, Andreas, Michal A Kopera, Simone Marras, Lucas C Wilcox, Tobin Isaac et Francis X Giraldo (2018). « Strong scaling for numerical weather prediction at petascale with the atmospheric model NUMA ». In : *The International Journal of High Performance Computing Applications* 33.2, p. 411-426. DOI : 10.1177/1094342018763966 (cf. p. 13).
- Müller, Eike, Xu Guo, Robert Scheichl et Sinan Shi (2013). « Matrix-free GPU implementation of a preconditioned conjugate gradient solver for anisotropic elliptic PDEs ». In : *Computing and Visualization in Science* 16.2, p. 41-58. DOI : 10.1007/s00791-014-0223-x (cf. p. 12).
- Pailleux, Jean, Jean-François Geleyn, Ryad El Khatib, Claude Fischer, Mats Hamrud, Jean-Noël Thépaut, Florence Rabier, Erik Andersson, Deborah Salmond, David Burridge, Adrian Simmons et Philippe Courtier (2015). « Les 25 ans du système de prévision numérique du temps IFS/Arpège ». In : *La Météorologie* 8.89, p. 18. DOI : 10.4267/2042/56594 (cf. p. 46).
- Pinty, Jean-Pierre, Robert Benoit, Evelyne Richard et René Laprise (1995). « Simple Tests of a Semi-Implicit Semi-Lagrangian Model on 2D Mountain Wave Problems ». In : *Monthly Weather Review* 123.10, p. 3042-3058. DOI : 10.1175/1520-0493(1995)123<3042:stoasi>2.0.co;2 (cf. p. 39).
- Putman, William M. et Shian-Jiann Lin (2007). « Finite-volume transport on various cubed-sphere grids ». In : *Journal of Computational Physics* 227.1, p. 55-78. DOI : 10.1016/j.jcp.2007.07.022 (cf. p. 22, 31).
- Qaddouri, Abdessamad et Vivian Lee (2010). « The elliptic solvers in the Canadian limited area forecasting model GEM-LAM ». In : *Modeling Simulation and Optimization-Tolerance and Optimal Control*. IntechOpen (cf. p. 19).
- (2011). « The Canadian global environmental multiscale model on the Yin-Yang grid system ». In : *Quarterly Journal of the Royal Meteorological Society* 137.660, p. 1913-1926 (cf. p. 19, 22).
- Reuter, H. I., A. Nelson et A. Jarvis (2007). « An evaluation of void-filling interpolation methods for SRTM data ». In : *International Journal of Geographical Information Science* 21.9, p. 983-1008. DOI : 10.1080/13658810601169899 (cf. p. 146).
- Robert, André (1981). « A stable numerical integration scheme for the primitive meteorological equations ». In : *Atmosphere-Ocean* 19.1, p. 35-46 (cf. p. 32).
- Saad, Youcef et Martin H Schultz (1986). « GMRES : A generalized minimal residual algorithm for solving nonsymmetric linear systems ». In : *SIAM Journal on scientific and statistical computing* 7.3, p. 856-869 (cf. p. 72).
- Saito, Kazuo, Tsukasa Fujita, Yoshinori Yamada, Jun ichi Ishida, Yukihiko Kumagai, Kohei Aranami, Shiro Ohmori, Ryoji Nagasawa, Saori Kumagai, Chiashi Muroi, Teruyuki Kato, Hisaki Eito et Yosuke Yamazaki (2006). « The Operational JMA Nonhydrostatic Mesoscale Model ». In : *Monthly Weather Review* 134.4, p. 1266-1298. DOI : 10.1175/mwr3120.1 (cf. p. 19).
- Schär, Christoph, Daniel Leuenberger, Oliver Fuhrer, Daniel Lüthi et Claude Girard (2002). « A new terrain-following vertical coordinate formulation for atmospheric prediction models ». In : *Monthly Weather Review* 130.10, p. 2459-2480 (cf. p. 61).
- Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac et V. Masson (2011). « The AROME-France convective-scale operational model ». In : *Monthly Weather Review* 139, p. 976-991 (cf. p. 15, 19, 46).

## BIBLIOGRAPHIE

---

- Simmons, A. J., B. J. Hoskins et D. M. Burridge (1978). « Stability of the Semi-Implicit Method of Time Integration ». In : *Monthly Weather Review* 106.3, p. 405-412. DOI : 10.1175/1520-0493(1978)106<0405:sotsim>2.0.co;2 (cf. p. 35, 120).
- Skamarock, William C. (2004). « Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra ». In : *Monthly Weather Review* 132.12, p. 3019-3032. DOI : 10.1175/mwr2830.1 (cf. p. 14, 16).
- Skamarock, William C., Joseph B. Klemp, Jimy Dudhia, David O. Gill, Zhiqian Liu, Judith Berner, Wei Wang, Jordan G. Powers, Michael G. Duda, Dale M. Barker et Xiang-Yu Huang (2019). « A Description of the Advanced Research WRF Model Version 4 ». en. In : DOI : 10.5065/1DFH-6P97 (cf. p. 19).
- Staniforth, Andrew et Jean Côté (1991). « Semi-Lagrangian Integration Schemes for Atmospheric Models—A Review ». In : *Monthly Weather Review* 119.9, p. 2206-2223. DOI : 10.1175/1520-0493(1991)119<2206:slisfa>2.0.co;2 (cf. p. 30).
- Staniforth, Andrew et John Thuburn (2011). « Horizontal grids for global weather and climate prediction models : a review ». In : *Quarterly Journal of the Royal Meteorological Society* 138.662, p. 1-26. DOI : 10.1002/qj.958 (cf. p. 22).
- Staniforth, Andrew et Nigel Wood (2005). « An Unsuspected Boundary-Induced Temporal Computational Mode in a Two-Time-Level Discretization ». In : *Monthly Weather Review* 133.3, p. 712-720. DOI : 10.1175/mwr-2889.1 (cf. p. 122).
- Stevens, Bjorn, Masaki Satoh, Ludovic Auger, Joachim Biercamp, Christopher S. Bretherton, Xi Chen, Peter Düben, Falko Judd, Marat Khairoutdinov, Daniel Klocke, Chihiro Kodama, Luis Kornbluh, Shian-Jiann Lin, Philipp Neumann, William M. Putman, Niklas Röber, Ryosuke Shibuya, Benoit Vanniere, Pier Luigi Vidale, Nils Wedi et Linjiong Zhou (2019). « DYAMOND : the DYNAMICS of the Atmospheric general circulation Modeled On Non-hydrostatic Domains ». In : *Progress in Earth and Planetary Science* 6.1. DOI : 10.1186/s40645-019-0304-z (cf. p. 15, 16, 63).
- Thuburn, J., M. Zerroukat, N. Wood et 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 ». In : *Quarterly Journal of the Royal Meteorological Society* 136.646, p. 146-154. DOI : 10.1002/qj.517 (cf. p. 35).
- Vána, Filip, Pierre Bénard, Jean-François Geleyn, André Simon et Yann Seity (2008). « Semi-Lagrangian advection scheme with controlled damping : An alternative to nonlinear horizontal diffusion in a numerical weather prediction model ». In : *Quarterly Journal of the Royal Meteorological Society* 134.631, p. 523-537. DOI : 10.1002/qj.220 (cf. p. 63).
- Voitus, Fabrice, Pierre Bénard, Christian Kühnlein et Nils P. Wedi (2019). « Semi-implicit integration of the unified equations in a mass-based coordinate : model formulation and numerical testing ». In : *Quarterly Journal of the Royal Meteorological Society* 145.725, p. 3387-3408. DOI : 10.1002/qj.3626 (cf. p. 15).
- Walters, David, Ian Boutle, Malcolm Brooks, Thomas Melvin, Rachel Stratton, Simon Vosper, Helen Wells, Keith Williams, Nigel Wood, Thomas Allen, Andrew Bushell, Dan Copsey, Paul Earnshaw, John Edwards, Markus Gross, Steven Hardiman, Chris Harris, Julian Heming, Nicholas Klingaman, Richard Levine, James Manners, Gill Martin, Sean Milton, Marion Mittermaier, Cyril Morcrette, Thomas Riddick, Malcolm Roberts, Claudio Sanchez, Paul Selwood, Alison Stirling, Chris Smith, Dan Suri, Warren Tennant, Pier Luigi Vidale, Jonathan Wilkinson, Martin Willett, Steve Woolnough et Prince Xavier (2017). « The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations ». In : *Geoscientific Model Development* 10.4, p. 1487-1520. DOI : 10.5194/gmd-10-1487-2017 (cf. p. 19, 35).
- Wedi, Nils et S. Malardel (2010). « Non-hydrostatic modelling at ECMWF ». In : DOI : 10.21957/RZ0JR98E (cf. p. 16).
- Wedi, Nils P. (2014). « Increasing horizontal resolution in numerical weather prediction and climate simulations : illusion or panacea? ». In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 372.2018, p. 20130289. DOI : 10.1098/rsta.2013.0289 (cf. p. 13, 38).
- Wedi, Nils P. et Piotr K. Smolarkiewicz (2009). « A framework for testing global non-hydrostatic models ». In : *Quarterly Journal of the Royal Meteorological Society* 135.639, p. 469-484. DOI : 10.1002/qj.377 (cf. p. 16).
- Wedi, Nils P., Mats Hamrud et George Mozdzynski (2013). « A Fast Spherical Harmonics Transform for Global NWP and Climate Models ». In : *Monthly Weather Review* 141.10, p. 3450-3461. DOI : 10.1175/mwr-d-13-00016.1 (cf. p. 37).
- Wood, Nigel, Andrew Staniforth, Andy White, Thomas Allen, Michail Diamantakis, Markus Gross, Thomas Melvin, Chris Smith, Simon Vosper, Mohamed Zerroukat *et al.* (2014). « An inherently mass-conserving

- 
- semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations ». In : *Quarterly Journal of the Royal Meteorological Society* 140.682, p. 1505-1520 (cf. p. 19).
- Yessad, Karim (2019a). *Horizontal diffusion computations in the cycle 46T1R1 of ARPEGE/IFS*. Rapp. tech. Météo-France, CNRM/GMAP/ALGO (cf. p. 63).
- (2019b). *Semi-Lagrangian computations in the cycle 46T1R1 of ARPEGE/IFS*. Rapp. tech. Météo-France, CNRM/GMAP/ALGO (cf. p. 30, 60).
- Zheng, Yongjun et Philippe Marguinaud (2018). « Simulation of the performance and scalability of message passing interface (MPI) communications of atmospheric models running on exascale supercomputers ». In : *Geoscientific Model Development* 11.8, p. 3409-3426 (cf. p. 12, 76, 77).
- Zängl, Günther (2012). « Extending the Numerical Stability Limit of Terrain-Following Coordinate Models over Steep Slopes ». In : *Monthly Weather Review* 140.11, p. 3722-3733. DOI : 10.1175/mwr-d-12-00049.1 (cf. p. 42).
- Zängl, Günther, Daniel Reinert, Pilar Rípodas et Michael Baldauf (2014). « The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M : Description of the non-hydrostatic dynamical core ». In : *Quarterly Journal of the Royal Meteorological Society* 141.687, p. 563-579. DOI : 10.1002/qj.2378 (cf. p. 22).

---

## Résumé —

AROME est le modèle à aire limitée de Prévision Numérique du Temps (PNT) de Météo-France. Son noyau dynamique résout les équations d'Euler pleinement compressibles non-hydrostatiques à une résolution horizontale kilométrique. Pour respecter les délais fortement contraints inhérents à la PNT, les méthodes numériques utilisent le mieux possible des supercalculateurs composés d'un grand nombre de nœuds travaillant en parallèle. Le noyau dynamique semi-implicite aujourd'hui utilisé, est basé sur un opérateur linéaire à coefficients constants qui traite implicitement les termes responsables de la propagation des ondes les plus rapides, contournant ainsi les contraintes de stabilité numérique les plus fortes et autorisant de grands pas de temps d'intégration. Cet opérateur autorise également l'utilisation d'une discrétisation spectrale des opérateurs de dérivée horizontale à l'aide d'un algorithme de transformée de Fourier rapide, peu scalable sur les machines parallèles du futur. De plus, en traitant les termes orographiques explicitement, donc soumis à des contraintes de stabilité, les résolutions hectométriques du futur représentant notamment mieux les plus fortes pentes des reliefs, ne peuvent être abordées avec ces schémas. Dans cette thèse, nous abandonnons les schémas à coefficients constants d'une part et la discrétisation spectrale d'autre part, pour améliorer la stabilité numérique au voisinage des plus fortes pentes pour les futures architectures de calcul.

**Mots clés :** Noyau dynamique, Prévision Numérique du Temps, scalabilité

---

**Abstract** — AROME is the Limited Area Model of Météo-France. Its dynamical core solves the fully compressible non-hydrostatic Euler equations at kilometeric horizontal resolution. To satisfy the highly constrained deadlines inherent in Numerical Weather Prediction, numerical methods make the best possible use of supercomputers composed of a large number of nodes working in parallel. The semi-implicit dynamical core used today is based on a constant coefficient linear operator that allows an implicit treatment of the terms responsible of the fastest waves propagation, thus bypassing the strongest numerical stability constraints and allowing large integration time steps. This operator also allows the use of spectral discretization of horizontal derivative operators using a fast Fourier transform algorithm, which is not very scalable on the parallel machines of the future. Moreover, by treating the orographic terms explicitly, and therefore subject to stability constraints, the hectometric resolutions of the future, representing in particular better the steepest slopes of the reliefs, cannot be approached with these schemes. In this thesis, we abandon constant coefficient schemes on the one hand and spectral discretization on the other hand, in order to improve numerical stability around the steepest slopes for the future computing architectures.

**Keywords :** Dynamical core, Numerical Weather Prediction, scalability