



HAL
open science

Apport des infrasons pour l'assimilation de données dans un modèle global de prévision numérique du temps

Pierre Jacques Vanderbecken

► **To cite this version:**

Pierre Jacques Vanderbecken. Apport des infrasons pour l'assimilation de données dans un modèle global de prévision numérique du temps. Climatologie. Université Paul Sabatier - Toulouse III, 2020. Français. NNT : 2020TOU30328 . tel-03749296v2

HAL Id: tel-03749296

<https://theses.hal.science/tel-03749296v2>

Submitted on 10 Nov 2023

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 22/06/2020 par :
Pierre J. Y. VANDERBECKEN

**Apport des infrasons pour l'assimilation de données dans un
modèle global de prévision numérique du temps**

JURY

J.-P. CHABOUREAU
C. CLAUD
M. BOCQUET
A. HAUCHECORNE
P. MIALLE
J.-F. MAHFOUF
C. MILLET

CNRS - UMR 5560
CNRS - UMR 8539
CEREA
CNRS - UMR 8190
CTBTO - IDC
MÉTÉO-FRANCE - UMR 3589
CEA, DAM, DIF

Président
Rapporteure
Rapporteur
Rapporteur
Examineur
Directeur
Co-Directeur

École doctorale et spécialité :

SDU2E : Océan, Atmosphère, Climat

Unité de Recherche :

Centre National de Recherches Météorologiques/Météo-France (UMR 3589)

Directeur(s) de Thèse :

J.-F. MAHFOUF et C. MILLET

Rapporteurs :

C. CLAUD, M. BOCQUET et A. HAUCHECORNE

Remerciements

La réalisation de ce travail de recherche, résumé au sein de ce mémoire, n'a été possible que grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma gratitude. Je voudrais tout d'abord adresser toute ma reconnaissance à mes directeurs de thèse, Monsieur Jean-François MAHFOUF, responsable de l'équipe OBS à Météo-France/CNRM, et Monsieur Christophe MILLET, actuel directeur du laboratoire au CEA-DAM-DIF auquel j'appartenais. Je remercie Monsieur Jean-François MAHFOUF pour sa supervision éclairée tout au long de la thèse et pour toute l'aide qu'il m'a apportée. Je remercie Monsieur Christophe MILLET pour sa supervision de tous les jours et ses conseils avisés afin que je ne m'écarte pas trop de mes objectifs. Et je les remercie tous les deux pour leurs disponibilités et pour le temps qu'ils m'ont consacré.

Qui dit deux directeurs de thèse, dit deux équipes de recherche à remercier. Je remercie l'ensemble du laboratoire du CEA-DAM-DIF, Thomas, Jean-Yves, Yves, Jean-Noël, Françoise, Ludovic, Jean-Philippe, Christian, Olaf, Nicolas, Gaël et les Laurents, pour m'avoir fourni un environnement de travail stable et enrichissant. Vivement le prochain café des sciences !

Je remercie également l'ensemble de l'équipe OBS, et plus largement les membres du GMAP qui m'ont accueilli chaleureusement à chaque fois que je leur rendais visite. Je remercie Monsieur Loïk BERRE pour sa disponibilité et ses éclaircissements sur l'assimilation d'ensemble. Je remercie Monsieur Philippe CHAMBON, Monsieur Fabrice DURUISSEAU et Monsieur Alexandre MARY pour leurs aides lorsque je débutais avec le langage Python. Je remercie Monsieur Laurent DESCAMPS et Madame Carole LABADIE pour leur concours lors de la génération de l'expérience PEARP et leurs expertises.

Je tiens à remercier particulièrement Monsieur Alexis Le PICHON et Madame Elisabeth BLANC pour l'intérêt qu'ils ont porté à ma thèse, pour leurs précieux conseils et pour m'avoir fait profiter de leur réseau. Je remercie Monsieur Julien VERGOZ pour toute son aide, sa disponibilité et sa gentillesse à chaque fois que je venais frapper à sa porte pour réclamer de nouvelles données. Je remercie également mes collègues thésards, Florentin, Julien, Arthur, Yoan, Frédéric, Alexandre, Roxanne, Mayeul, Marine, Zied, Matthieu, Antoine, Annie, Marilys, et je leur souhaite tous mes vœux de réussite.

Je tiens à remercier ma sœur et mes amis pour m'avoir soutenu et avoir su me divertir à chaque fois que c'était nécessaire. Je tiens à exprimer ma plus grande gratitude à mon grand-père, Jacques MEUNIER, qui a grandement facilité la logistique de mes déplacements entre Toulouse et Paris. Je remercie aussi mes parents, Marc et Cécile, pour leur support à toute épreuve et leur amour de la langue de Molière. Et à l'instar de nombreux d'entre-nous, je remercie le personnel soignant de l'hôpital de Niort pour avoir su soigner mes parents et pour leur écoute alors que la COVID-19 le mobilisait.

Merci !

à mes grands-parents qui m'ont donné le goût des sciences et de l'érudition.

“Faithless is he that says farewell when the road darkens”, said Gimli.
“Maybe”, say Elrond, “but let him not vow to walk in the dark, who has not
seen the nightfall.”

“Yet sworn word may strengthen quaking heart”, said Gimli.

“Or break it”, said Elrond.

The Lord of the Rings, J.R.R. TOLKIEN, 1954.

Sommaire

Remerciements	iii
	Page
1 Introduction	1
1.1 Généralités sur les infrasons	1
1.2 L'observation de l'état de l'atmosphère	7
1.3 Les infrasons : traceurs de la dynamique de l'atmosphère	12
1.4 Objectifs de la thèse et plan du manuscrit	14
2 Modélisation de l'atmosphère	15
2.1 La circulation moyenne de l'atmosphère	16
2.1.1 Composition chimique et stratification verticale de l'atmosphère .	16
2.1.2 Circulation générale dans la troposphère et la moyenne atmosphère	16
2.2 Les modèles d'atmosphère	22
2.2.1 Modèles empiriques	22
2.2.2 Les modèles globaux de prévision numérique du temps	25
2.3 Assimilation de données	32
2.3.1 Formulation du problème variationnel	32
2.3.2 Matrice de covariances d'erreurs d'ébauche B et assimilation d'ensemble	34
3 Modélisation de la propagation des infrasons	37
3.1 Description de la propagation acoustique	37
3.1.1 Généralités sur la modélisation de la propagation	37
3.1.2 Mise en équation	38
3.2 Méthode numérique de résolution	40
3.2.1 Équation d'onde et approche modale	40
3.2.2 Acoustique géométrique et tracé de rayons	45
3.2.3 Conclusion	48
4 Études de sensibilité en vue de l'inversion des infrasons	51
4.1 Rivesaltes	52
4.1.1 Cas d'étude	52
4.1.2 Les résultats du code de tracé de rayons	54
4.1.3 Les résultats du modèle par décomposition modale	55
4.1.4 Influence d'une modification des profils atmosphériques sur les quantités d'intérêt et définition de la stratégie	57
4.1.5 Stratégie en vue de l'assimilation des signaux infrasons	59
4.2 Définition d'un cadre idéalisé	60
4.2.1 Atmosphère idéalisée à deux degrés de liberté	60
4.2.2 Extraction des quantités d'intérêt simulées	61

4.3	Problème inverse dans un cadre idéalisé	65
4.3.1	Optimisation locale	68
4.3.2	Expression de J sous différentes hypothèses	69
4.3.3	Résultats du problème inverse	72
4.4	Conclusions	76
5	Approche bayésienne pour l'inversion des données infrasons	77
5.1	Résumé et études préliminaires	78
5.1.1	Résumé	78
5.1.2	Points techniques révélés par l'étude	79
5.2	Bayesian selection	83
5.2.1	Abstract	83
5.2.2	Introduction	84
5.2.3	Infrasound propagation modeling	85
5.2.4	Infrasound-based atmospheric state selection	87
5.2.5	Application to the detections of Mount Etna	89
5.2.6	Discussion	95
5.2.7	Relevance of the best atmospheric states	97
5.2.8	Conclusions and Perspectives	101
5.3	Compléments à l'article	103
5.3.1	Annexe de l'article : <i>Numerical set-up of the Bayesian inversion over synthetic observations</i>	103
5.3.2	Utilisation des temps de trajet	107
5.3.3	Modèle de propagation <i>range-dependent</i>	107
5.3.4	Apport des prévisions à court terme	109
5.3.5	Cas d'une autre station	110
6	Conclusions et perspectives	113
6.1	Conclusions générales	113
6.2	Perspectives de la thèse	116
A	Extraction des données météorologiques	119
A.1	Les sorties d'un modèle de PNT	119
A.1.1	Extraction des champs d'intérêt	119
A.1.2	Construction de la grille verticale en pression	121
A.2	Construction des profils pour les modèles de propagation	125
A.2.1	Conversion altitude-pression	125
A.2.2	Données lacunaires et conditions opérationnelles	126
A.3	Calcul de la vitesse du son	127
A.3.1	Extrapolation dans la haute atmosphère et conditions de surface	131
B	Intérêt des prévisions d'ensemble	133
B.1	Les paramétrisations physiques pour le modèle ARPEGE et la PEARP	133
B.1.1	Les paramétrisations physiques	134
B.1.2	PEARP et perturbations par approche multi-physiques	135
	Références	139
	Table des figures	156
	Liste des tableaux	157

Glossaire	159
C Résumé-étendu	161

Chapitre 1

Introduction

Sommaire

1.1 Généralités sur les infrasons	1
1.2 L'observation de l'état de l'atmosphère	7
1.3 Les infrasons : traceurs de la dynamique de l'atmosphère	12
1.4 Objectifs de la thèse et plan du manuscrit	14

1.1 Généralités sur les infrasons

La plupart des phénomènes naturels de fortes énergies, tels que les tempêtes ou les séismes, émettent des ondes acoustiques, ou plus communément du son. Certaines de ces ondes acoustiques sont émises sous le seuil de l'audition humaine, dans une gamme de fréquences comprises entre environ 10 mHz et 20 Hz. Ce sont les infrasons. L'activité humaine est également source d'infrasons, ce qui fait que les émissions infrasonores sont aujourd'hui issues d'une grande variété de sources comme le montre l'inventaire au tableau 1.1. Ces sources sont soit impulsives comme les explosions, soit entretenues comme les chutes d'eau. Elles peuvent être ponctuelles comme les volcans, mobiles comme les lancements spatiaux ou encore étendues comme le bruit de la houle océanique, autrement appelé *microbarome*. Bien qu'aucune de ces sources ne soit globale, il est pourtant possible d'observer des infrasons à tout instant et en tout point du globe. Cela s'explique par la manière dont se propagent ces ondes.

Une fois émis, les infrasons se propagent dans le milieu qui les entoure, c'est-à-dire l'atmosphère terrestre. L'atmosphère terrestre est un milieu hétérogène dans toutes les directions. Elle modifie la trajectoire prise par les ondes tout au long de leur propagation et notamment par réfraction. L'importance de ces changements de direction imposés par l'atmosphère est proportionnelle à la valeur de l'indice de réfraction. L'indice de réfraction de l'atmosphère est proportionnel à la température et à la vitesse du vent ambiant. En particulier, lorsque la température (respectivement la vitesse du vent) croît avec l'altitude, l'indice de réfraction augmente et l'atmosphère infléchit la trajectoire prise par les ondes vers le sol. Une partie des ondes se retrouve alors confinée entre la surface et cette élévation de la température (et/ou de la vitesse du vent). L'atmosphère forme alors un guide d'ondes.

La structure synoptique verticale de l'atmosphère terrestre favorise naturellement la formation de tels guides d'ondes. Comme le montre le profil de température de la figure 1.1, il apparaît quatre strates distinctes dans l'atmosphère : la troposphère, la stratosphère, la mésosphère et la thermosphère. Ces couches sont délimitées par la surface

	Source	Amplitude enregistrée (Pa)	Fréquences (Hz)
Anthropogénique	essai nucléaire atmosphérique	>20	>0,002
	essai nucléaire sous-terrestre	~1	>1
	tir de mine	~5	>0,05
	explosion chimique	~10	>0,05
	structure vibrante (pont)	~0,2	>0,5
	lancement spatial	~5	>0,01
	rentrée atmosphérique	~1	>0,5
	vol subsonique	~2	>0,3
	vol supersonique	~10	>0,3
	aurore polaire	~2	>0,008
Naturelle	surf	~0,2	>1
	chute d'eau	~0,2	>0,5
	microbarome	~5	0,12-0,35
	éruption volcanique	>20	>0,002
Catastrophe naturelle	tempête	~0,5	0,01-0,1
	séisme	~4	~0,005-10
	incendie	~2	>2
	impact météorite	>10	>0,001
	éclair	~2	>0,5
	tornade	~0,5	>0,5
	tsunami	~0,1	~0,5-2
	vélage de la cryosphère	~1	~0,5-0,8
	éboulement, avalanche	~1	>0,1

Tableau 1.1 : Quelques propriétés des principales sources d'infrasons. En orange : les sources impulsives. En bleu : les sources dites entretenues. *Ce tableau est adapté des travaux de McKisic (1997) et Campus et Christie (2010).*

terrestre et par les extrema de température que sont la tropopause, la stratopause et la mésopause. Les accroissements de la température dans la stratosphère et la thermosphère auront tendance à confiner les ondes. Inversement, les décroissances de la température dans la troposphère et la mésosphère auront tendance à dévier les ondes vers le haut. Si l'on regarde maintenant le profil de vent moyen aux latitudes tempérées (38 °N) de l'hémisphère Nord en hiver, on peut également identifier des courants de jets qui auront aussi tendance à confiner les ondes. Le premier est troposphérique. Le second s'étale sur toute la stratosphère et la mésosphère, on parlera ici de jet de la moyenne atmosphère.

Ces guides d'ondes se déforment dans l'espace et le temps en suivant les changements dans les gradients de température et de vent. Sous ces conditions, la propagation devient anisotrope comme l'illustre la figure 1.2. Dans le cadre de cet exemple, une partie des ondes émises vers l'Est est confinée dans des altitudes inférieures à 40 km. Ce confinement résulte de l'élévation de la température dans la stratosphère conjuguée à la présence du courant du jet zonal¹ de la moyenne atmosphère. L'atmosphère forme alors un premier guide d'ondes dans la direction Est. Les trajectoires et l'empreinte au sol des ondes confinées par ce guide sont représentées en orange sur la figure 1.2. L'altitude maximale atteinte par ces ondes est habituellement appelée altitude de réfraction. Dans les directions Nord, Sud et Ouest, le vent dans la moyenne atmosphère n'est plus assez fort ou même contraire pour confiner une partie des ondes sous les 40 km d'altitude. Toutes les ondes continuent alors leurs propagations vers les hautes couches de l'atmosphère. Elles arrivent alors dans la thermosphère, où l'élévation forte de la température les réfracte vers le sol. L'atmosphère génère alors un guide d'ondes thermosphérique. Ces ondes confinées entre la thermosphère et la surface sont représentées en vert et bleu sur la figure 1.2. Les guides d'ondes stratosphérique et thermosphérique sont simultanément présents dans la direction Est. Cela implique qu'une partie des infrasons peut être piégée dans l'atmosphère entre ces deux guides.

La présence d'un guide d'ondes dans une direction spécifique peut être prédit simplement en comparant la célérité effective du milieu à celle au niveau de la surface. La

1. vent zonal = vent orienté Ouest-Est ; vent méridien = vent orienté Sud-Nord.

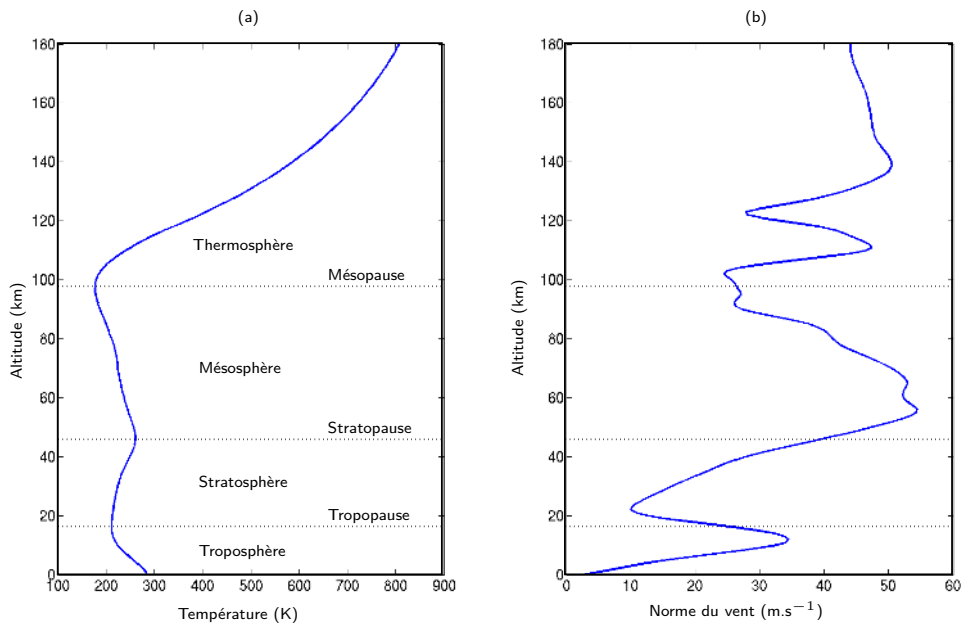


Figure 1.1 : Profils empiriques de température (a) et de norme du vent (b) à 38 degrés de latitude Nord.

Phase	Label
Troposphérique	<i>Phase W</i>
Stratosphérique	<i>Phase S</i>
Mésosphérique	<i>Phase M</i>
Thermosphérique	<i>Phase T</i>

Tableau 1.2 : Nomenclature des phases

célérité effective du milieu dans une direction donnée est la somme de la vitesse du son proportionnelle à la racine carrée de la température², et de la vitesse du vent dans cette direction. Si la célérité effective, notée c_e sur la figure 1.2, est supérieure à celle au niveau de la surface alors il y a un guide d'ondes. Ainsi, comme l'attestent les figures 1.2 (c, e, f, h), les guides d'ondes identifiés précédemment auraient pu être prédits simplement en examinant les profils de célérité effective.

Les ondes émises par une même source génèrent différentes arrivées au sol dans le temps et l'espace en raison de leurs multiples réfractions dans l'atmosphère. Ce même mécanisme de réfractions multiples est analogue à celui qui se déroule en sous-sol pour les ondes sismiques (Romanowicz, 2003). Par analogie, les différentes arrivées infrasonores sont alors appelées *phases*, et elles sont labellisées en fonction de leurs altitudes de réfraction. La première phase issue d'une réfraction dans la stratosphère et entendue par une station sera typiquement notée *Phase S1* en suivant la nomenclature au tableau 1.2.

L'atmosphère n'atténue les infrasons que faiblement, car ils ne sont notamment que très peu affectés par le mécanisme de diffusion moléculaire (Chatillon, 2011). Ce qui fait que, malgré l'atténuation liée à la distance de propagation, les infrasons peuvent être enregistrés à de très grandes distances. Un des plus anciens exemples provient du rapport de Symons *et al.* (1888) sur les conséquences de l'éruption du volcan Krakatoa³.

2. la vitesse du son c pour un gaz parfait est définie par $c^2 = \left. \frac{\partial p}{\partial \rho} \right|_S \propto T$ où p , ρ et S sont respectivement la pression, la masse volumique et l'entropie du gaz

3. volcan situé dans le détroit de la Sonde entre l'île de Java et de Sumatra

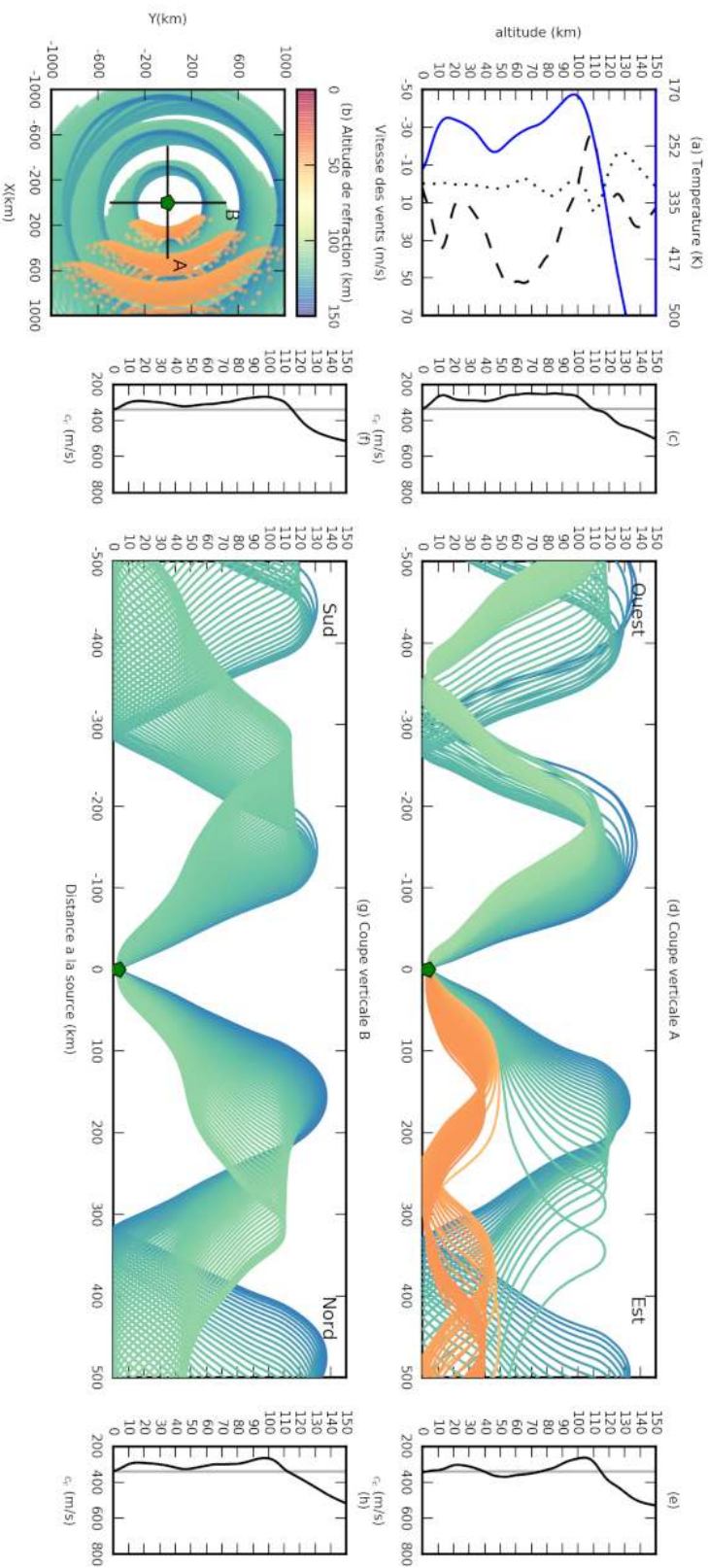


Figure 1.2 : Influence du vent sur la propagation des ondes infrasonores depuis le Mont Etna (14.99 ° E, 37.75 ° N) le 25/12/2018. (a) Profils empiriques de température (ligne bleue), de vent zonal (tirets noirs) et de vent méridien (pointillés noirs). (b) Empreinte au sol du champ d'ondes infrasonores prédit depuis le Mont Etna (pentagone vert). Chaque point est coloré en fonction de l'altitude maximale atteinte par les ondes qu'il représente. (d)-(g) Trajectoires des ondes infrasonores dans les plans verticaux correspondant aux sections A et B de la figure (b). (c)-(e)-(f)-(h) Profils de célérité effective du milieu respectivement dans la direction Ouest, Est, Sud et Nord, la célérité effective c_e étant l'addition de la vitesse du son et du vent dans la direction choisie. Les lignes grisées indiquent la valeur de c_e au sol. Lorsque la valeur de c_e devient supérieure à celle au sol, il y a un guide d'ondes.

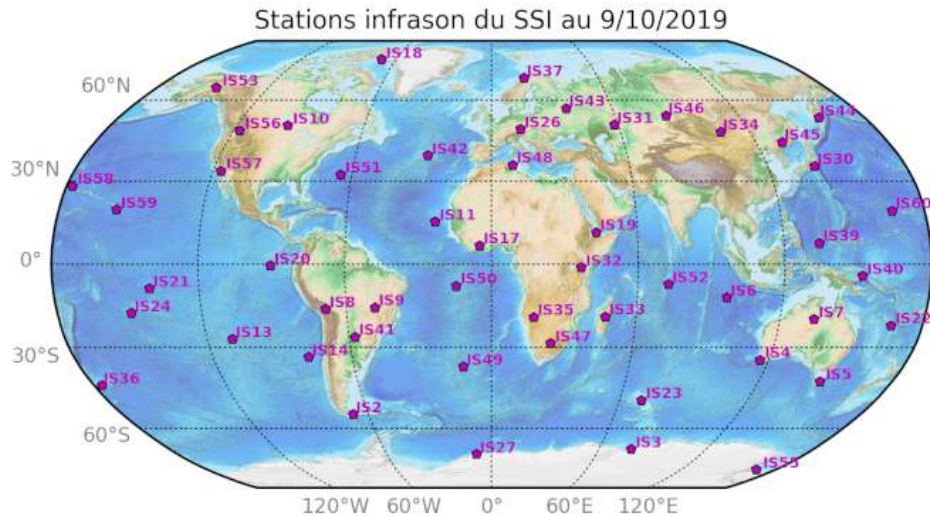


Figure 1.3 : Répartition des 51 stations infrasons du SSI certifiées par l'OTICE au 9 octobre 2019.

Il montre que le signal infrasonore émis par l'éruption volcanique avait pu être enregistré tout autour du globe (Aberdeen, Toronto,...). Certaines stations ont même vu le signal par deux fois signifiant que les ondes avaient fait le tour du globe. Le même constat a pu être fait pour d'autres cas recensés dans la synthèse historique de Evers et Haak (2010).

C'est pourquoi cette technologie a été vue comme un moyen efficace de détections à longue distance des essais nucléaires prohibés par le Traité d'Interdiction Complète des Essais nucléaires (TICE). L'Organisation du Traité d'Interdiction Complète des Essais nucléaires (OTICE) créée dans le cadre du TICE, développe et maintient de manière opérationnelle un Système de Surveillance International (SSI). Le SSI permet d'informer les instances signataires du traité sur les possibles infractions. Le SSI se compose de différents réseaux de stations réparties autour du globe dont 51 sont dédiées à la surveillance des infrasons. Les géolocalisations de ces 51 stations infrasons certifiées par l'OTICE sont présentées sur la figure 1.3. Ces stations écoutent et enregistrent en temps réel les émissions infrasonores autour du globe.

Par exemple, du 3 au 6 juin 2019, la station IS48 du SSI située en Tunisie, a notamment détecté des phases en provenance d'un azimuth autour des 65° Nord et d'autres autour des $300\text{-}350^\circ$ Nord. L'azimut de provenance des phases détectées sera par la suite usuellement appelé back-azimut pour ne pas confondre avec l'azimut de propagation⁴. Comme le montre la figure 1.4 (a), parmi les sources se trouvant dans le back-azimut 65° N, il y a le volcan du Mont Etna. Ce volcan était entré en éruption le 30 mai 2019 et resta actif durant les premières semaines de juin. Les phases issues de ce back-azimut peuvent donc être associées à l'éruption de l'Etna et non à une infraction du TICE. Celles détectées dans les back-azimuts pointant vers le golfe de Gascogne ($300\text{-}330^\circ$ N), résultent probablement, quant à elles, de la tempête Miguel, comme le confirme l'image satellite 1.4. Chacune des phases détectées se traduit par un point coloré de la figure 1.4 (c), la couleur indiquant la fréquence centrale de la phase détectée. On remarque que les phases en provenance des back-azimuts du volcan sont avec une fréquence centrale proche du hertz et que celles en provenance du back-azimut du golfe de Gascogne est de l'ordre de 0,1 Hz, ce

4. back-azimut = azimuth de propagation + 180° modulo 360°

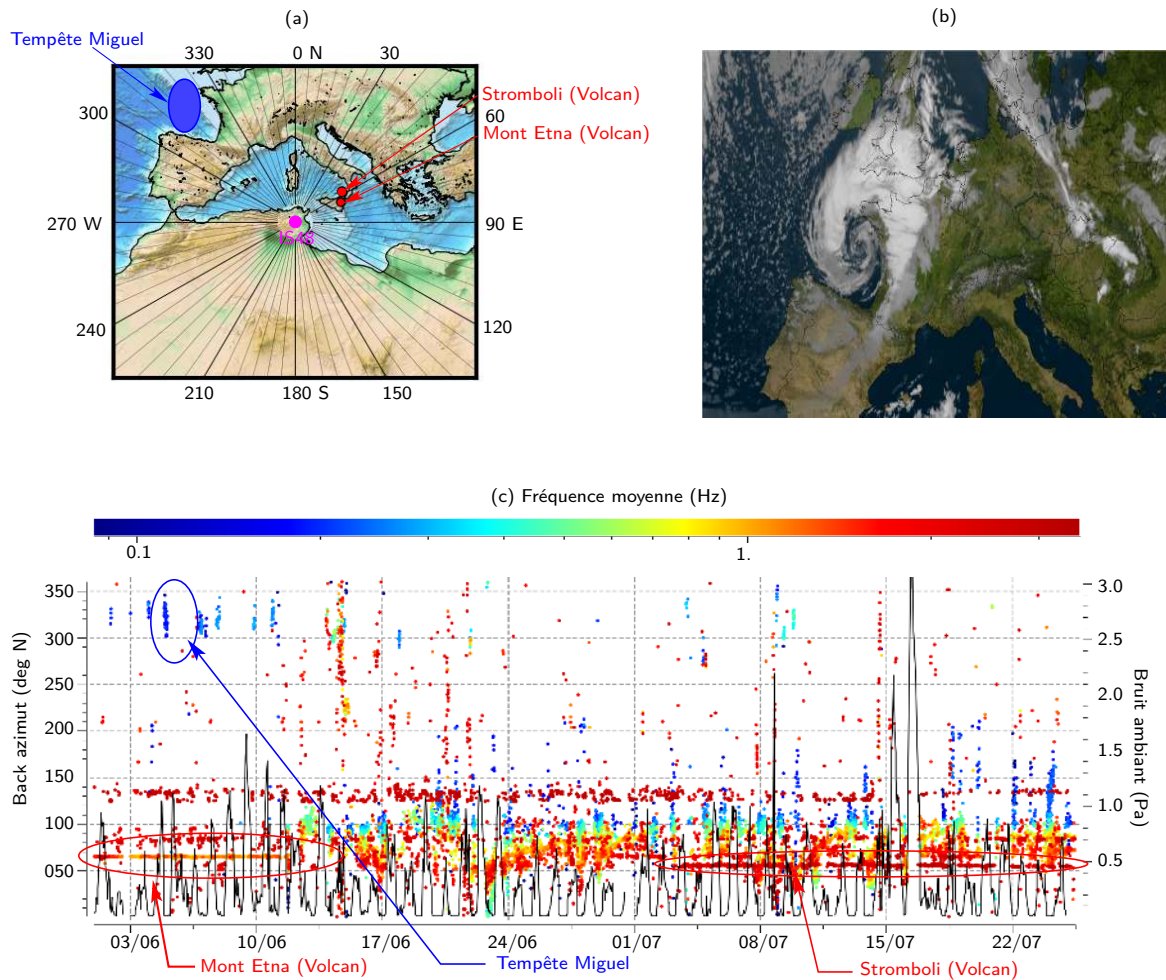


Figure 1.4 : Exemple illustratif des observations faites à la station IS48 (Tunisie)(a) Carte de navigation autour de la station IS48 (rond magenta).(b) Image satellite de la tempête Miguel au large de la France le 6/06/2019. *Image de Meteosat/FVALK* (c) Phases détectées par IS48 du SSI entre le 03/06/2019 et le 22/07/2019 en fonction de leurs azimuts de provenance (back-azimut échelle de gauche). Chaque point est une phase détectée. La couleur indique la fréquence moyenne du signal de la phase détectée. La courbe noire indique l'amplitude du bruit ambiant incohérent pouvant occulter les détections (échelle de droite).

qui correspond bien à une tempête d'après les caractéristiques évoquées au tableau 1.1. La station détecte également des phases en provenance du volcan Stromboli à partir de l'éruption qui a démarré le 3 juillet 2019.

Les phases sont détectées par la station à l'aide d'une grande « antenne » constituée de 7 microbaromètres. Ces capteurs sont sensibles à des fluctuations de pression de 1 mPa. La distance entre les capteurs peut être de l'ordre du kilomètre. Les variations du vent liées à la turbulence génèrent des fluctuations de pression qui parasitent le signal vu par les différents capteurs. L'antenne est en mesure de détecter les phases lorsque le rapport entre le signal de la phase et le bruit ambiant imposé par la turbulence est supérieur ou égal à 1. C'est pourquoi lorsque ce bruit ambiant généré autour de la station est important, comme c'était le cas le 17 juillet (cf figure 1.4), les capteurs peuvent ne pas détecter des phases. Cette série montre également l'influence d'un cycle diurne sur le bruit ambiant. La turbulence semble être plus faible la nuit.

Finalement au cours des 7 semaines présentées sur la figure 1.4 (c), IS48 a détecté 12217 phases (≈ 1 toutes les 6 minutes), ce qui représente une grande quantité de données à étudier. L'ensemble des données enregistrées par le SSI sont rassemblées et étudiées au sein du Centre International de Données (CID) à Vienne, où se trouve le siège de l'OTICE. Le CID est suppléé dans ses missions par des centres nationaux tels que celui hébergé par le Commissariat à l'Énergie Atomique et aux énergies alternatives (CEA).

Ce réseau global permet donc un suivi en temps réel des émissions des différentes sources d'infrasons, sous réserve de pouvoir localiser et caractériser ses sources à partir des différentes phases enregistrées par les différentes stations. La localisation et l'association des phases enregistrées peuvent être faites dès lors que plusieurs stations (typiquement 3 stations) ont détecté la même source. La caractérisation énergétique de la source, quant à elle, ne peut pas être effectuée uniquement avec les signaux enregistrés. C'est pourquoi dans le cadre opérationnel, on a recours à la simulation de la propagation pour quantifier l'énergie expliquant les phases détectées. Ces simulations requièrent une connaissance du milieu de propagation la plus précise possible afin de limiter les incertitudes sur la caractérisation de la source. Cependant il n'est pour le moment pas possible de simuler parfaitement numériquement ou analytiquement la circulation atmosphérique. Notre connaissance du milieu provient des modèles numériques des services météorologiques et des observations de l'atmosphère.

1.2 L'observation de l'état de l'atmosphère

La mise en place d'un réseau mondial d'observations coordonné par l'Organisation Météorologique Mondiale (OMM) depuis le début du XX^{ème} siècle permet le suivi en temps réel de l'atmosphère terrestre. Ce réseau est utilisé pour comprendre et prévoir les phénomènes météorologiques à diverses échelles de temps et d'espace. Il s'est enrichi au cours du temps pour fournir de nos jours un ensemble très varié de mesures documentant principalement la pression de surface, le vent, la température, l'humidité de l'air, ainsi que les nuages et la pluie. La figure 1.5 illustre la variété de ces mesures.

On peut distinguer les mesures in-situ (dites « conventionnelles ») réalisées à partir de capteurs immergés dans le milieu : les données des stations météorologiques de surface sur terre ou en mer (bouées et bateaux), les données en provenance des avions de ligne commerciaux, ainsi que celles des ballons-sondes fournissant des profils verticaux (radiosondages) des principaux paramètres météorologiques depuis la surface jusque vers 35 km d'altitude (2 fois par jour pour un ensemble d'environ 500 stations réparties de manière hétérogène à la surface du globe, comme le montre la figure 1.6 (a)). Les autres mesures proviennent de la télédétection depuis le sol (radars) ou depuis l'espace

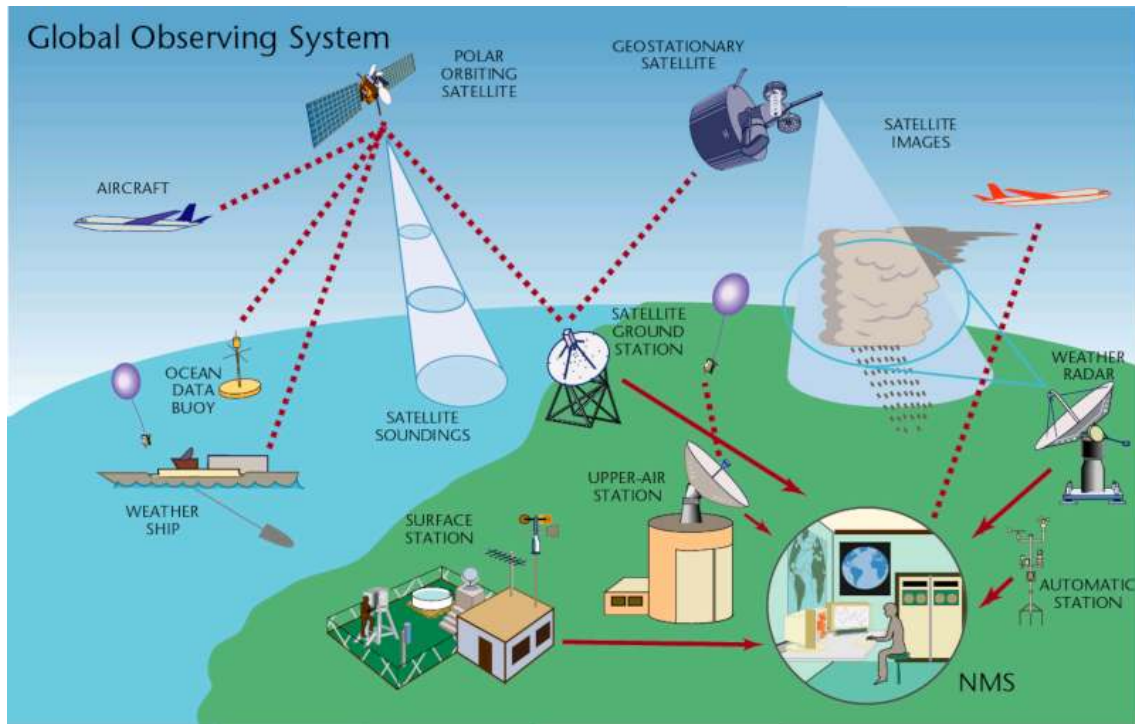


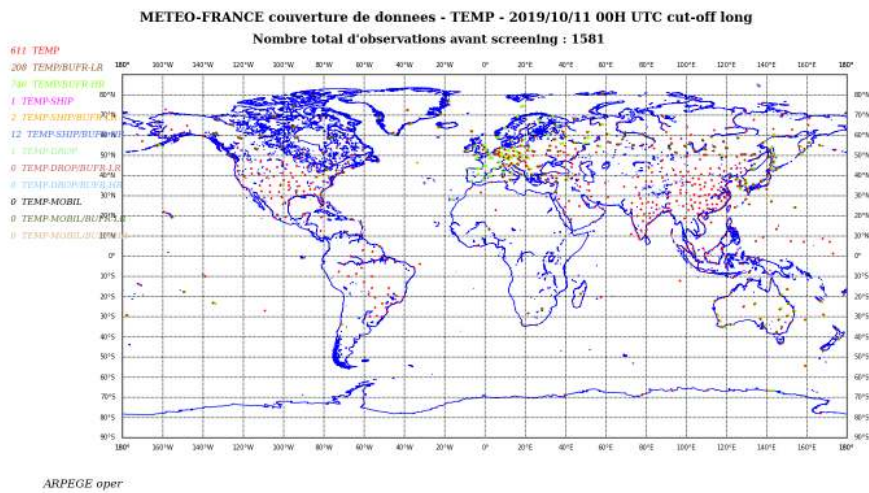
Figure 1.5 : Représentation schématique du système d'observation global de l'atmosphère.

(satellites). Ces informations sur l'atmosphère sont obtenues de manière indirecte car les instruments mesurent un rayonnement électromagnétique ayant interagi avec les particules présentes dans l'atmosphère (molécules, gouttelettes nuageuses, gouttes de pluie, aérosols, ...). La plupart des instruments de mesure depuis l'espace sont des radiomètres qui mesurent le rayonnement émis par l'atmosphère et la surface terrestre dans certaines parties du spectre électromagnétique, même si certains, moins nombreux, sont des instruments actifs qui fournissent une information sur le rayonnement rétrodiffusé comme les radars⁵ ou les lidars⁶. Tous ces instruments embarqués permettent d'avoir une couverture uniforme du globe. Les satellites en orbite basse couvrent la surface du globe en quelques heures (passant par les pôles, comme l'illustre la figure 1.6 (b)), et ceux en orbite géostationnaire observent une partie du globe en continu (comme sur la figure 1.6 (c)).

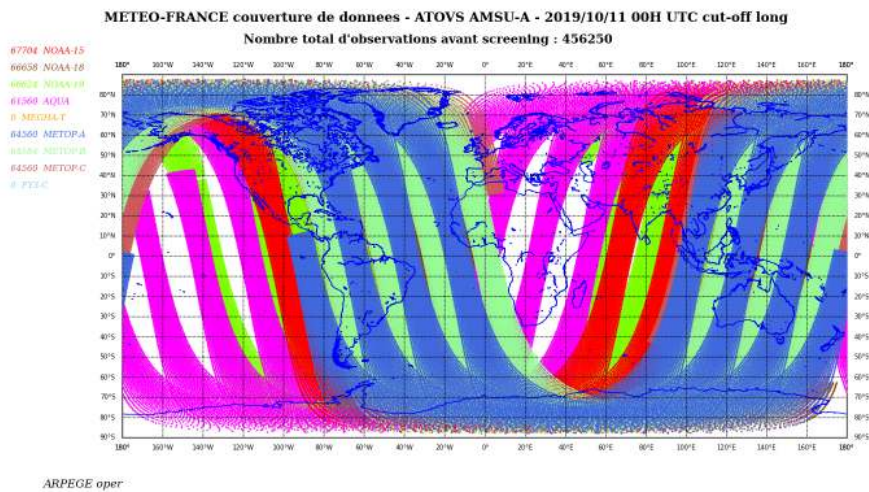
Ces données météorologiques sont utilisées pour la prévision numérique du temps (PNT). En effet, elles permettent de corriger les prévisions à courte échéance fournies par les modèles et ainsi fournir un état corrigé plus précis (nommé analyse). Cet état est utilisé comme condition initiale pour prévoir l'évolution de l'atmosphère à des échelles de temps variant entre quelques heures et plusieurs jours. À titre d'exemple, la figure 1.7 présente l'évolution des observations utilisées dans le modèle global de PNT de Météo-France ARPEGE sur une période de 18 ans. Au début de la période ce modèle utilisait 20 millions d'observations par mois alors que récemment ce nombre est proche de 850 millions. Cette augmentation est principalement due à une utilisation accrue des données de télédétection spatiale que l'on peut expliquer par deux facteurs. Le premier a été le déploiement en orbite basse à partir de 2002 des nouveaux sondeurs dans le domaine infra-rouge basés sur l'interférométrie que sont Atmospheric InfraRed Sounder (AIRS), Infrared Atmospheric Sounding Interferometer (IASI) et Cross-track Infrared Sounder (CrIS). Ces sondeurs fournissent des mesures à la fois plus précises et plus nombreuses que les instruments des

5. Radio Detection And Ranging

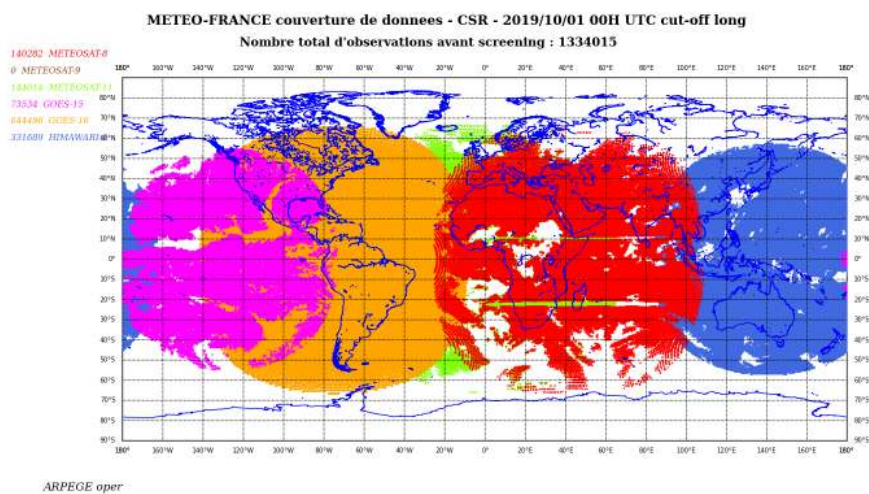
6. Light Detection And Ranging



(a) Radiosondages



(b) Satellites au nadir



(c) Satellites géostationnaires

Figure 1.6 : Exemple de distribution des observations directes et indirectes de l'atmosphère disponible au cours d'une période de 6 heures.

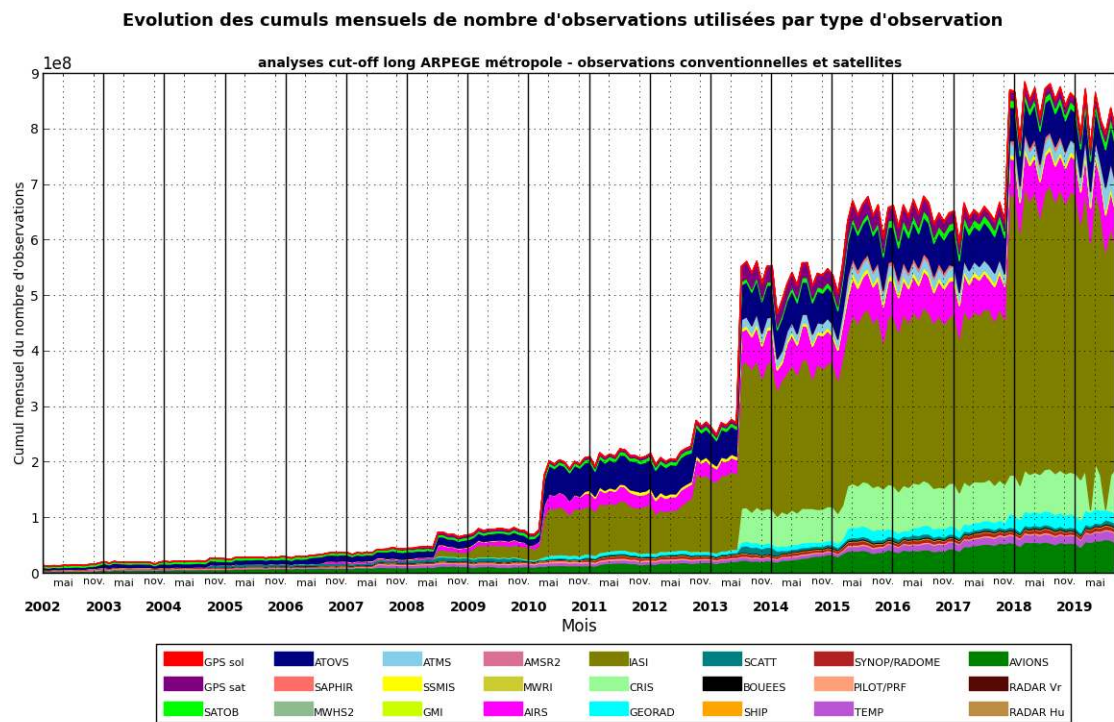


Figure 1.7 : Cumuls mensuels des observations in-situ et satellitaires utilisées par Météo-France pour initialiser le modèle de PNT globale ARPEGE. Tirée du rapport CNRM DirOP/COMPAS du 3/10/2019.

générations précédentes. Le second facteur a été l'amélioration des techniques permettant de générer les conditions initiales des modèles de PNT. Ces techniques sont maintenant capables de mieux exploiter l'information contenue dans ces mesures et d'en extraire le signal atmosphérique utile pour améliorer les prévisions numériques. Nous reviendrons plus en détails au chapitre 2 sur ces techniques dites d'« assimilation de données ».

L'inventaire précédent montre que les mesures autres que celles issues de la télédétection spatiale sont largement minoritaires, notamment celles qui renseignent sur la structure verticale de l'atmosphère malgré sa très forte stratification (cf figure 1.1). Les données depuis l'espace sont pour la plupart effectuées par des instruments au nadir⁷ qui fournissent des informations peu détaillées sur la structure verticale. La capacité de sondage des radiomètres s'appuie sur le phénomène d'absorption du rayonnement au voisinage d'une raie d'absorption d'un constituant atmosphérique interagissant dans le spectre infra-rouge ou micro-onde comme le dioxyde de carbone ou l'oxygène. En utilisant des mesures de plus en plus éloignées de la raie centrale, l'atmosphère devient de plus en plus transparente et permet à l'instrument de fournir des informations de plus en plus proches de la surface terrestre. Ces instruments dit « sondeurs » fournissent indirectement des profils de température et de vapeur d'eau, mais avec des résolutions verticales de plusieurs kilomètres. Des mesures mieux résolues selon la verticale sont disponibles depuis une quinzaine d'années grâce à des instruments avec une visée au limbe⁸ mais qui fournissent des informations peu détaillées sur la structure horizontale. Ils s'appuient sur une technique de radio-occultation à partir des signaux électromagnétiques des satellites de géonavigation (signaux dits *GPS GNSS-RO*). Des profils de température entre 10 et 50 km avec des résolutions comprises entre 250 m et 1 km peuvent être

7. orientation au nadir : l'instrument observe l'atmosphère directement en-dessous de lui.

8. orientation au limbe : l'instrument observe l'atmosphère à l'horizon.

obtenus. Un avantage de cette technique est d'être peu sensible à des biais instrumentaux contrairement aux sondeurs infra-rouge et micro-onde au nadir qui doivent être corrigés (notamment au moyen des radiosondages) avant leur utilisation dans les modèles de PNT.

Si la télédétection spatiale peut fournir des informations (même peu détaillées) sur des profils de température et de vapeur d'eau, elle est actuellement incapable de fournir des profils de vent (en dehors de la mission exploratoire AEOLUS de l'ESA avec un lidar Doppler ultra-violet lancée le 22 août 2018). Les informations depuis l'espace sont obtenues au moyen du déplacement de traceurs⁹, grâce à des images rapprochées dans le temps d'un radiomètre (souvent en orbite géostationnaire). Le vent est alors fourni au niveau de l'altitude de l'objet étudié. Pour les mesures de vent déduites du déplacement des nuages, on est souvent limité à des valeurs en basses couches vers 850 hPa (les premiers kilomètres de l'atmosphère) et en haute troposphère vers 200 hPa (tropopause). On dispose également à la surface des océans du signal rétrodiffusé par des radars spatiaux. Ces radars opèrent dans une bande de fréquence peu sensible à l'atmosphère mais très sensible aux états de mer (comme la rugosité de surface) que l'on peut relier à l'intensité du vent près de la surface. À titre d'exemple, ces données de vent par télédétection spatiale ne représentent que 2,5 % des données utilisées dans le modèle ARPEGE. Ce pourcentage est à peu près le même pour les données conventionnelles.

Depuis une dizaine d'années, la plupart des modèles globaux de PNT opérationnels s'étendent sur la verticale jusqu'à des altitudes comprises entre 50 et 80 km. Historiquement, ces modèles résolvaient essentiellement la troposphère qui est la zone où les phénomènes météorologiques affectent directement les activités humaines (orages, cyclones, brouillard, ...). La stratosphère était considérée comme une « zone tampon » dont le rôle était d'éviter la réflexion d'ondes au sommet du modèle vers la troposphère. L'augmentation de la puissance de calcul, une représentation plus réaliste des processus radiatifs et dynamiques, l'utilisation de sondeurs satellitaires sensibles à la stratosphère ainsi que le besoin d'étendre les échéances de prévision au-delà de quelques jours ont conduit les modélisateurs à vouloir décrire de manière plus réaliste la stratosphère et une partie de la mésosphère. Cette augmentation de l'altitude des modèles a l'avantage également d'écartier de la zone d'intérêt principal de modélisation (i.e. la troposphère) les problèmes numériques associés à la condition à la limite supérieure. Comme indiqué précédemment, si on dispose de mesures permettant de contrôler le comportement de la température dans ces régions lors de la prévision numérique, il n'en est pas de même pour les données de vent.

L'existence d'équilibres atmosphériques à grande échelle comme le géostrophisme et l'hydrostatisme permet de relier le champ de vent horizontal au champ de masse (température et pression). C'est ainsi que des observations sensibles à la température peuvent modifier le champ de vent dans ces régions lors du processus d'assimilation de données produisant l'état initial des modèles de PNT. Ces équilibres n'étant pas toujours valides dans la moyenne atmosphère, l'état de ces couches est modélisé par les PNT avec de fortes incertitudes. Or il a été montré par Baldwin *et al.* (2003), que si ces incertitudes sont limitées, la prévisibilité à long terme (au-delà de la semaine) des modèles de PNT s'améliore. De plus, l'étude récente de Vargin et Kiryushov (2019) montre que des anomalies dans ces couches peuvent induire des anomalies dans les basses couches. Il serait donc très utile de disposer de mesures de vent dans ces régions, pour améliorer la précision des prévisions des modèles de PNT. Aujourd'hui, différents moyens d'observation directs et indirects du vent ont pu être identifiés comme le montre la figure 1.8, qui recense huit techniques de mesure du vent ou de la température dans la moyenne atmosphère.

9. Ce sont des objets cohérents (comme les nuages).

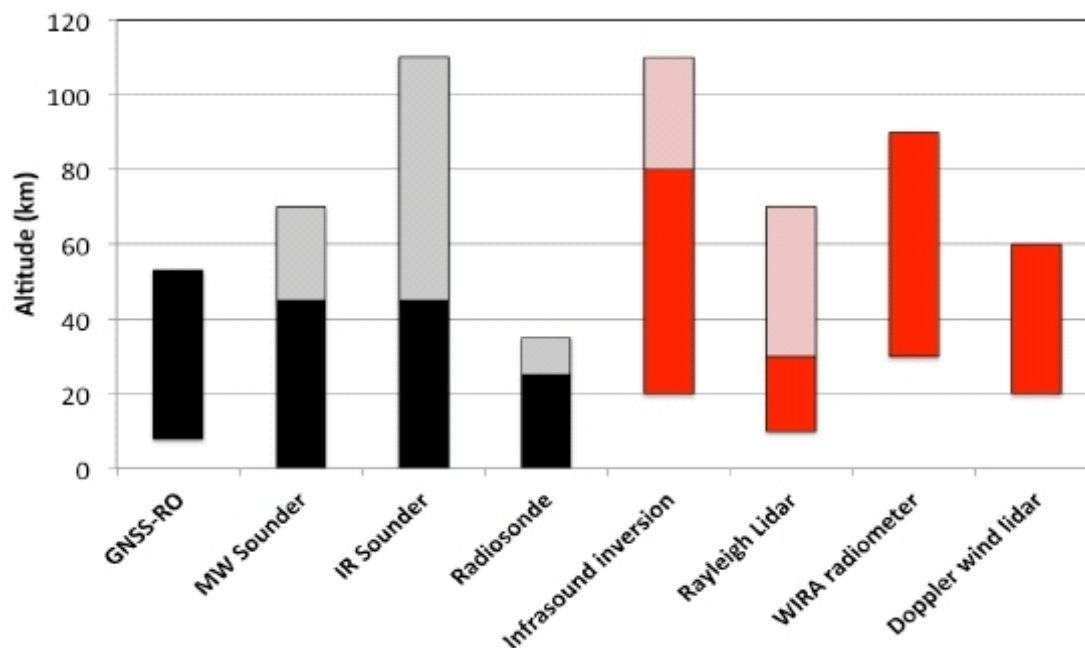


Figure 1.8 : Gammes d'altitudes sondées par les observations de température et de vent actuellement utilisées dans les modèles de PNT (en noir) et les nouvelles observations identifiées dans le cadre du projet européen d'infrastructure ARISE (en rouge). Les zones plus claires correspondent à des régions où l'observation est occasionnelle. *Adaptée du rapport du projet européen ARISE 2 (Grant 653980).*

Les techniques en noir sur la figure 1.8 sont celles aujourd'hui disponibles et fournissant opérationnellement des données aux services de prévision météorologique. Celles en rouge ont été identifiées dans le cadre du programme européen d'infrastructure Atmospheric dynamics Research InfraStructure in Europe (ARISE piloté par le CEA (Grant numéro 653980)). On y retrouve les mesures de température des instruments lidar réglés pour le domaine de la diffusion de Rayleigh (Khaykin *et al.*, 2015), les mesures de vent faites par instrument lidar à effet Doppler (Chanin *et al.*, 1989), les mesures de vent par des radiomètres (WInd RAdiometer for Campaigns) (Hagen *et al.*, 2018). Ces trois moyens de mesure assez directe depuis le sol sont toutefois peu déployés à l'échelle du globe. L'option supplémentaire identifiée par le programme ARISE est d'utiliser les infrasons. L'idée se base sur le principe qu'il existe des sources d'infrasons d'ores et déjà bien connues et caractérisées, comme certains volcans. Les infrasons émis par ces sources peuvent être alors considérés comme des « traceurs » de l'écoulement atmosphérique.

1.3 Les infrasons : traceurs de la dynamique de l'atmosphère

L'utilisation des infrasons comme moyen de sonder l'atmosphère a été étudiée dès les années 1970 (Donn & Rind, 1971 ; Rind *et al.*, 1973 ; Rind & Donn, 1975). Ces premières études mettaient en évidence l'intérêt des infrasons émis par la houle océanique pour obtenir de l'information sur le vent au-delà de la troposphère. Ces informations se basent sur les conditions de réfractions des ondes infrasonores. L'idée étant qu'en connaissant les coordonnées géographiques de la source et de l'antenne réceptrice, la concordance entre la simulation et les observations est informative sur les conditions atmosphériques. Cela revient à résoudre le problème inverse qui, à partir des signaux à la source et au

niveau du récepteur, permet de retrouver l'état de l'atmosphère. Ces études, pionnières dans le domaine, ont relevé trois limitations pour que les infrasons puissent donner de l'information à l'échelle globale sur les conditions atmosphériques. Une première limitation venait du manque de capteurs sur la surface du globe. La deuxième venait des limitations des modèles atmosphériques utilisés. L'information récupérée ne pouvait fournir que des tendances saisonnières (Rind & Donn, 1978), car les modèles atmosphériques ne pouvaient pas expliquer certaines variations dans les signaux infrasonores observés. La troisième venait de la modélisation de la propagation des ondes. Le fléchissement des trajectoires des ondes dans le plan horizontal n'était pas pris en compte dans leur modèle de propagation (Georges & Beasley, 1977).

Le déploiement des stations infrasons du SSI à partir de la fin des années 1990 a permis d'avoir accès à un réseau global de capteurs infrasons. Ce réseau a permis alors de vérifier que les infrasons pouvaient contraindre le vent de la moyenne atmosphère à l'échelle globale (Drob *et al.*, 2010). Les stations du SSI sont des antennes (plusieurs capteurs combinés), ce qui permet d'obtenir le back-azimut des phases et donc, si l'on connaît la source, d'en déduire le fléchissement dans le plan horizontal (Le Pichon *et al.*, 2005). Les modèles de propagation ont par ailleurs évolué pour pouvoir simuler ces comportements.

Depuis les années 1970, les capacités de calcul ont considérablement augmenté. Les modèles de prévision météorologique ont pu alors simuler l'état de l'atmosphère avec une meilleure résolution et ont pu enrichir la physique de leur modèle (Coiffier, 2000). Les modèles atmosphériques sont plus précis et permettent de mieux comprendre certaines des variations dans les signaux observés. Elles sont notamment dues à la propagation des ondes de gravité (Drob *et al.*, 2013) ou à des interactions avec le relief (Damiens *et al.*, 2018). Néanmoins, ces modèles ne résolvent pas explicitement l'ensemble des structures et font appel à des paramétrisations pour les échelles sous-mailles. Il existe donc une incertitude de modélisation de l'état atmosphérique et des variations dans les signaux qui sont inexpliquées.

Ces variations dans les signaux infrasons ont alors été vues comme un moyen de réduire les incertitudes liées à la non résolution explicite de certains phénomènes physiques par les modèles de PNT. Chunchuzov *et al.* (2015) ont alors développé une méthode par inversion qui se base sur une corrélation entre la variabilité dans les formes des signaux des phases détectées en station et les petites structures perturbant le profil vertical de vent dans la stratosphère. Au travers de cette méthode, ils ont pu montrer que le spectre des perturbations permettant d'expliquer les infrasons suit une loi de décroissance équivalente à celle des ondes de gravité. Les infrasons peuvent alors être vus comme un moyen de mieux régler les paramétrisations des ondes de gravité. D'autres méthodes développent directement des paramétrisations empiriques des phénomènes de grande et petite échelles influençant les profils verticaux de vent (Lalande & Waxler, 2016 ; Assink *et al.*, 2013). Ces paramétrisations se basent sur une décomposition en composantes principales des variations des profils verticaux de vent issus de leurs bases de données. Ils cherchent alors le profil perturbé optimal menant à la superposition des signaux détectés et ceux simulés. Ces études ont alors montré que, statistiquement, les profils verticaux de vent prévus par les modèles de PNT s'écartent systématiquement dans la stratosphère et la mésosphère (Assink *et al.*, 2014) des profils optimaux trouvés par la méthode inverse. Cependant, le réalisme des informations sur le vent résultant de ces méthodes inverses reste discutable. Les profils de vent ainsi trouvés ne sont plus contraints par les équations de la météorologie. De plus, ces méthodes font l'hypothèse que l'erreur n'est portée que par les profils de vent, simplifiant ainsi le problème inverse, mais du coup négligent le rôle du profil de température. Ces méthodes focalisent leur perturbation sur la moyenne

atmosphère considérant que les basses couches sont connues. Il en devient compliqué d'identifier la part de l'erreur due à la non résolution explicite de petites structures par le modèle de PNT, celle due aux erreurs de prévisions des structures explicitement résolues par le modèle de PNT, et celle due aux simplifications effectuées pour résoudre le problème inverse. Il est alors difficile d'estimer au cas par cas, la correction à effectuer sur le modèle de PNT. Il est également important de noter que les récents travaux de Smets *et al.* (2015) montrent que la simple considération de l'incertitude de prévision des modèles de PNT est suffisante pour statistiquement améliorer l'adéquation entre les résultats de simulation de la propagation des infrasons et les observations.

1.4 Objectifs de la thèse et plan du manuscrit

L'objectif de cette thèse est de répondre à la problématique suivante : comment fournir de l'information issue des détections infrasons qui puisse contraindre la température et la vitesse du vent dans la moyenne atmosphère et qui puisse être assimilées par un modèle de prévision numérique du temps ? L'enjeu serait d'améliorer ainsi la prévision météorologique dans la moyenne atmosphère, entraînant en retour une amélioration de la caractérisation des sources vues par le SSI.

Le manuscrit se décompose en six chapitres. Le deuxième chapitre introduit les différentes méthodes d'estimation de l'état de l'atmosphère utilisées dans le cadre de cette thèse. Dans le troisième chapitre, nous décrivons les méthodes rapides de simulation de la propagation des ondes infrasonores utilisées et qui doivent servir de base pour la construction d'un opérateur d'observation pour l'assimilation de données dédiés aux infrasons. Dans un premier temps, le quatrième chapitre présente les difficultés liées à la comparaison des observations réelles avec les résultats des différents modèles de propagation. On y évalue également l'effet qu'une modification du milieu a sur les signaux observés en station. Suite à ces résultats, il a été choisi dans un second temps d'adopter une nouvelle stratégie et d'examiner si des altérations du milieu sont effectivement identifiables à l'aide de méthodes d'inversion. L'idée n'est plus d'assimiler directement les infrasons mais d'assimiler des informations en termes de vent et de température obtenues par méthode inverse. Les résultats du quatrième chapitre montrent qu'il est important de réduire l'espace des possibles en considérant un état a priori réaliste de l'atmosphère pour faciliter la résolution du problème inverse. Il est alors choisi d'utiliser l'approche d'optimisation globale bayésienne. Cette approche est décrite et évaluée sur une base de détections réelles dans le cinquième chapitre. La base de détections est issue de l'activité éruptive du Mont Etna en mai 2016. Le dernier chapitre fournit des conclusions et des perspectives à ce travail de recherche.

Chapitre 2

Estimation de l'état atmosphérique

Sommaire

2.1	La circulation moyenne de l'atmosphère	16
2.1.1	Composition chimique et stratification verticale de l'atmosphère	16
2.1.2	Circulation générale dans la troposphère et la moyenne atmosphère	16
2.2	Les modèles d'atmosphère	22
2.2.1	Modèles empiriques	22
2.2.2	Les modèles globaux de prévision numérique du temps	25
2.3	Assimilation de données	32
2.3.1	Formulation du problème variationnel	32
2.3.2	Matrice de covariances d'erreurs d'ébauche \mathbf{B} et assimilation d'ensemble	34

Dans le cadre de l'étude de la propagation des infrasons sur des longues distances, il est nécessaire de disposer d'une estimation de l'état atmosphérique, principalement en termes de température T , de vent \mathbf{u} et de pression p . Une approche serait d'utiliser un modèle de prévision numérique de l'atmosphère qui fournirait un état atmosphérique de la surface jusqu'à la thermosphère. Cependant, aujourd'hui, cette approche reste très peu répandue dans les services météorologiques opérationnels. Il existe en revanche des modèles qui fournissent des estimations empiriques de l'état de l'atmosphère ainsi que des modèles de prévision numérique ne considérant que la troposphère et une partie de la moyenne atmosphère. L'objectif de ce chapitre est de recenser les sources d'informations sur l'état atmosphérique aujourd'hui disponibles. Une première partie résume les caractéristiques principales de la circulation générale de l'atmosphère. Dans une deuxième partie sont passées en revue les différentes sources d'informations sur l'état de l'atmosphère employées au cours de cette thèse que sont les modèles empiriques HWM et MSIS et les modèles globaux de prévision numérique du temps. Toujours dans cette seconde partie, le processus d'assimilation de données utilisé pour initialiser les modèles de PNT est décrit. Ce processus nécessite notamment la connaissance des erreurs de prévision. La troisième partie décrit l'approche ensembliste utilisée en météorologie pour obtenir ces erreurs de prévision.

2.1 La circulation moyenne de l'atmosphère

2.1.1 Composition chimique et stratification verticale de l'atmosphère

L'atmosphère terrestre est constituée d'eau et d'un mélange gazeux composé à 78.08% de molécules de diazote, à 20.95% de molécules de dioxygène, à 0.93% d'atomes d'argon, à 0.039% de molécules de dioxyde de carbone ainsi que d'autres molécules telles que l'ozone à l'état de trace. Ce mélange est appelé air. Le champ de gravité terrestre tend à maintenir proche de la surface les composants atmosphériques. Ce qui explique pourquoi la masse volumique de l'air et la pression atmosphérique décroissent avec l'altitude. L'évolution de la température avec l'altitude dans l'atmosphère neutre¹ (voir le chapitre 1), s'explique principalement au travers de trois phénomènes : le transfert de chaleur par turbulence et convection, les changements de phases de l'eau atmosphérique ainsi que l'absorption du rayonnement solaire dans les domaines de l'ultraviolet (UV) et du rayonnement dans le domaine infrarouge (IR) émis par l'atmosphère et la surface de la Terre. Entre la surface et une altitude voisine de 8 km aux pôles et de 17 km à l'équateur, c'est le transfert de chaleur par turbulence et convection qui l'emporte. Ces phénomènes imposent une diminution linéaire de la température de l'ordre de -6.5 K.km^{-1} . La décroissance de température est alors ralentie à cause de l'échauffement induit par l'absorption par photochimie du rayonnement UV conduisant à la formation de l'ozone par dissociation des molécules d'oxygène. La zone où le gradient vertical de température change de signe définit la tropopause. La hauteur de la tropopause varie entre 17 km au niveau de l'équateur (où les mouvements verticaux induits par la convection profonde² sont les plus intenses) et 8 km au niveau des pôles. Le gradient vertical positif de température au-dessus de la tropopause inhibe en partie les mouvements verticaux induits par la convection. C'est donc une zone stable, fortement stratifiée, qu'on appelle la stratosphère. La proportion d'ozone augmente dans la stratosphère jusqu'à environ 25 km. Vers 50 km, le gradient de température s'inverse une nouvelle fois, c'est la stratopause. Au-delà de celle-ci, la faible concentration en dioxygène fait que les réactions photochimiques précédentes ne réchauffent plus assez l'atmosphère pour compenser l'énergie rayonnée vers l'espace. La température diminue donc de nouveau, c'est la mésosphère. Vers environ 80 km, le gradient ralentit avant de s'inverser, c'est la mésopause. Au-delà de la mésopause, les fortes proportions d'ions d'oxygène et d'azote absorbent les hautes fréquences du rayonnement électromagnétique en provenance du Soleil et, de la même manière que dans la stratosphère, échauffent l'air. Le gradient vertical de température est alors fortement positif. C'est la thermosphère.

2.1.2 Circulation générale dans la troposphère et la moyenne atmosphère

L'inclinaison de la Terre conjuguée à ses révolutions font qu'à chaque instant, les différentes régions du globe ne reçoivent pas la même quantité de rayonnement solaire. La révolution de la Terre sur elle-même génère un cycle diurne. Son inclinaison fait que chacun des hémisphères reçoit alternativement plus de rayonnement solaire au cours d'une révolution autour du Soleil, lorsque leurs pôles respectifs pointent vers celui-ci. Cela

1. Atmosphère où les particules ionisées sont minoritaires.

2. En météorologie on fait la distinction entre la convection peu profonde qui implique un déplacement vertical des particules d'air sous l'isobare de 500 hPa et la convection profonde qui déplace les particules au-delà.

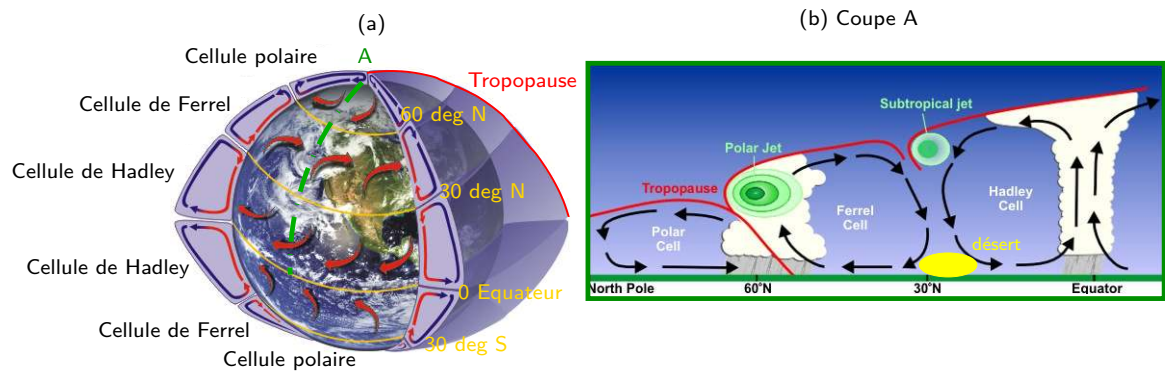


Figure 2.1 : Schéma illustratif de la circulation générale dans la troposphère. (a) Distribution des cellules (b) Coupe longitudinale. Ces figures ont été extraites du site de la NOAA <https://www.weather.gov/jetstream/> et complétées pour cette thèse.

induit un cycle saisonnier. Ces différences dans la distribution du rayonnement solaire génèrent des déséquilibres thermiques entre les différentes régions du globe. Les processus de transport tendent alors à redistribuer l'énergie pour revenir à l'équilibre. Cela induit une circulation atmosphérique qui doit alors composer avec la force de gravitation, la force de gradient de pression, la force d'entraînement sur l'atmosphère engendré par la rotation de la Terre et les forces de friction.

La circulation dans la troposphère

La circulation méridienne (i.e. selon un axe Nord-Sud) dans la troposphère peut se décomposer en cellules comme le montre le schéma 2.1 (a).

- Les cellules de Hadley, découvertes par Hadley au 18^{ème} siècle, couvrent les latitudes équatoriales 0-30 °N et 0-30 °S. Elles sont les régions recevant le plus de rayonnement du Soleil. L'air au niveau de l'équateur est donc plus chaud et s'élève. Ce mouvement ascendant, entraînant l'air, génère des vents à la surface qui convergent vers l'équateur. Ces vents sont déviés par la force de Coriolis vers la droite pour ceux en provenance du Nord et vers la gauche pour ceux en provenance du Sud. Ce sont les alizés à l'origine du courant d'Est équatorial. Les alizés transportent un air qui s'est chargé d'humidité, au contact des océans. Cet air chaud, chargé en humidité monte dans la troposphère où il se refroidit. En se refroidissant, la vapeur d'eau se condense, forme des nuages qui génèrent des précipitations. L'air asséché restant est alors transporté vers les pôles. La déviation vers l'Est induite par la force de Coriolis fait que cet air redescend au niveau des latitudes 30 °N et 30 °S. Cet air sec au niveau de ces latitudes assèche la surface et conduit à la formation de déserts.
- Les cellules polaires résultent des mêmes phénomènes. Au niveau des pôles, l'air froid et sec de la tropopause descend et est dévié en surface par la force de Coriolis. Cet air se réchauffe au niveau des latitudes tempérées et remonte vers les latitudes 60 °N et 60 °S.
- Les cellules de Ferrel, sont les cellules délimitées par les latitudes extrêmes des cellules de Hadley (± 30) et des cellules polaires (± 60). Leurs circulations sont entraînées par la circulation de leurs cellules voisines.

Au niveau des limites des cellules de Ferrel et sous la tropopause apparaissent des courants de jets soufflant vers l'Est et représentés sur le schéma 2.1 (b). Le courant de jet polaire est rapide et instable. Le courant de jet subtropical est, quant à lui, plus lent et plus stable. En raison des déplacements saisonniers du zénith, de l'alternance

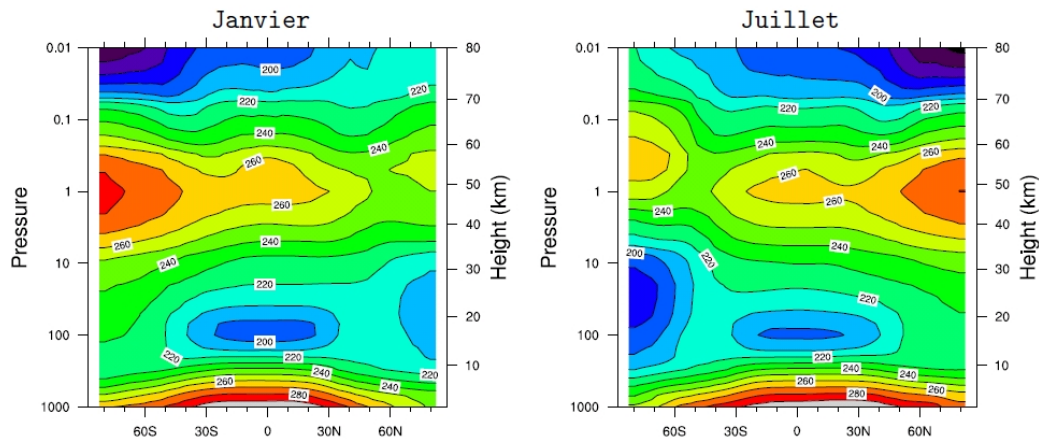


Figure 2.2 : Coupes latitude-pression de la moyenne mensuelle de la température (K) et moyennées en longitude d'après la climatologie SPARC (Randel *et al.*, 2004) pour le mois de janvier (à gauche) et de juillet (à droite). L'intervalle des contours est de 10 K. Cette figure est tirée de la thèse Saint-Martin (2010).

entre les océans sources d'évaporation intense, et les continents plus secs, mais aussi les instabilités non décrites par ce modèle simplifié, cette circulation est en réalité soumise à de fortes fluctuations.

La circulation dans la moyenne atmosphère

D'après Andrews *et al.* (1987), la moyenne atmosphère correspond à la couche de l'atmosphère comprise entre 10 km et 100 km d'altitude. Dans cette thèse, on réduira cette définition à la couche allant de la tropopause jusqu'à la mésopause. La structure méridienne de la moyenne atmosphère est bien différente de celle de la troposphère. Comme le montre la figure 2.2, la température dans la basse stratosphère (en dessous de 20 km) est minimale à l'équateur et atteint sa valeur maximale aux moyennes latitudes, dans l'hémisphère en été. Au-dessus de 25 km (30 hPa), la température décroît quasi-uniformément du pôle d'été vers le pôle d'hiver. Cette distribution de la température ne résulte pas seulement du chauffage causé par l'absorption de rayonnement solaire, comme le montre la figure 2.3. Il faut également considérer le transfert de quantité de mouvement et d'énergie entre cette région et les autres couches de l'atmosphère (Holton, 2004). Sans ces échanges, l'atmosphère moyenne serait en situation dite d'équilibre radiatif et son évolution suivrait uniquement le cycle solaire comme sur la figure 2.3.

L'écart entre la température réelle et la température en situation d'équilibre radiatif présenté sur la figure 2.3, trouve son origine dans l'interaction entre l'écoulement moyen principalement zonal et la propagation d'ondes. Ces ondes trouvent leurs origines dans les mouvements des particules d'air, qui, une fois écartées de leur position d'équilibre, se trouvent soumises à une force de rappel. Dans le cadre d'un déplacement dans la direction longitudinale, la force de rappel est due à la variation de la force de Coriolis avec la latitude. Ces ondes peuvent être émises par interaction entre l'écoulement atmosphériques et les grands systèmes orographiques (les Rocheuses, les Andes, ...) ou par les déséquilibres thermiques imposés par les contrastes de température de surface entre les surfaces continentales et les océans. Les ondes planétaires qui en résultent sont nommées *ondes de Rossby* et sont d'échelle caractéristique de l'ordre de plusieurs milliers de kilomètres sur l'horizontale et de plusieurs kilomètres sur la verticale. Lorsque le déplacement est vertical, c'est la gravité qui agit comme force de rappel. On parle alors d'*ondes de gravité*. Les

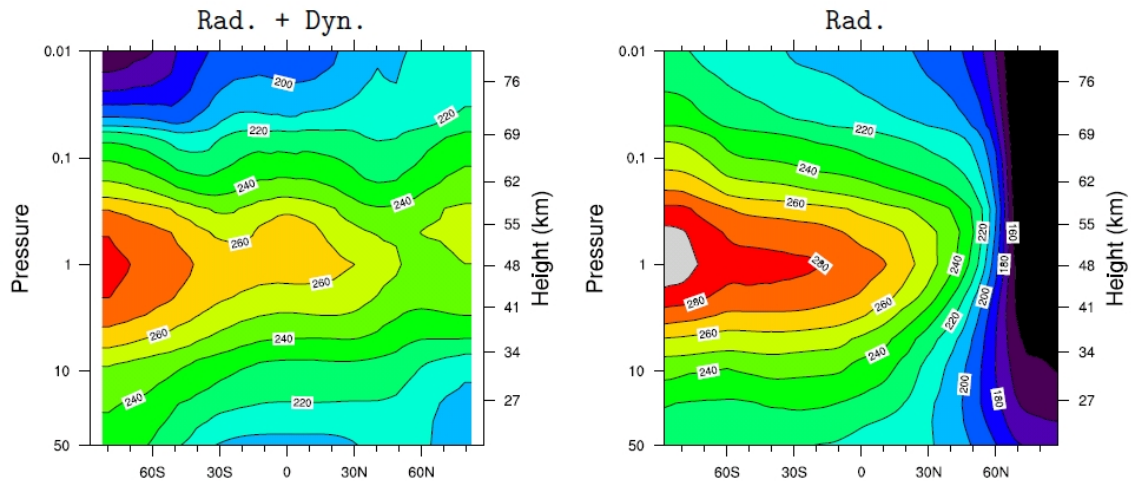


Figure 2.3 : Coupes latitude-pression de la moyenne mensuelle de la température (K) pour le mois de janvier et moyennées en longitude : à gauche celle d'après la climatologie SPARC (Randel *et al.*, 2004) ; à droite celle résultant de l'équilibre radiatif de la Terre d'après Holton (2004). L'intervalle des contours est de 10 K. Cette figure est tirée de la thèse Saint-Martin (2010).

ondes de gravité, quant à elles, sont d'échelle caractéristique plus réduite et sont créées par tous les phénomènes pouvant induire un déplacement vertical rapide des parcelles d'air (orographie, convection profonde, ...).

Au niveau de la mésopause, l'air du pôle d'été monte et est transporté vers le pôle d'hiver où il plonge. Dans les régions de subsidence (l'air plonge), l'air se réchauffe par compression adiabatique. Dans les régions d'ascendance (l'air monte), l'air se refroidit par détente adiabatique. Ce phénomène se traduit par une mésopause au niveau du pôle d'été plus froide que celle prévue par l'équilibre radiatif et une mésosphère au niveau du pôle d'été plus chaude. Ce mouvement est altéré par le déferlement des ondes de gravité. Le déferlement des ondes induit un dépôt de quantité de mouvement qui accélère le mouvement entre le pôle d'été et l'équateur et freine le mouvement entre l'équateur et le pôle d'hiver.

Dans la stratosphère, la circulation méridienne se caractérise par 2 cellules (à peu près symétriques par rapport à l'équateur). La forte convection au niveau de la région équatoriale fait que des masses d'air passent la tropopause et entrent dans la stratosphère. Ces masses d'air sont alors transportées vers le pôle d'hiver. Ces masses d'air plongent au niveau des hautes latitudes. Cette circulation porte le nom de *Brewer-Dobson*. Elle permet d'expliquer comment l'ozone émis par les orages équatoriaux est transporté jusqu'aux pôles d'hiver. Au cours de leur propagation, ces masses d'air sont ralenties par le déferlement des *ondes de Rossby*. Cette circulation est très lente et change de sens à chaque équinoxe. Cette alternance est visible sur les figures 2.4(a-c).

Oscillations caractéristiques de la moyenne atmosphère

Au niveau de la région équatoriale, la partie haute de la moyenne atmosphère est le siège d'une oscillation du vent zonal semi-annuels (*Semi Annual Oscillation* ou SAO). Comme le montre la figure 2.5 (b), l'oscillation est maximale au niveau de la stratopause et près de la mésopause. Le maximum de vent d'Ouest se produit juste après l'équinoxe, celui de vent d'Est juste après le solstice. L'accélération d'Ouest débute à la stratopause et se propage vers le bas à une vitesse de 10 km par mois. L'accélération d'Est, quant à

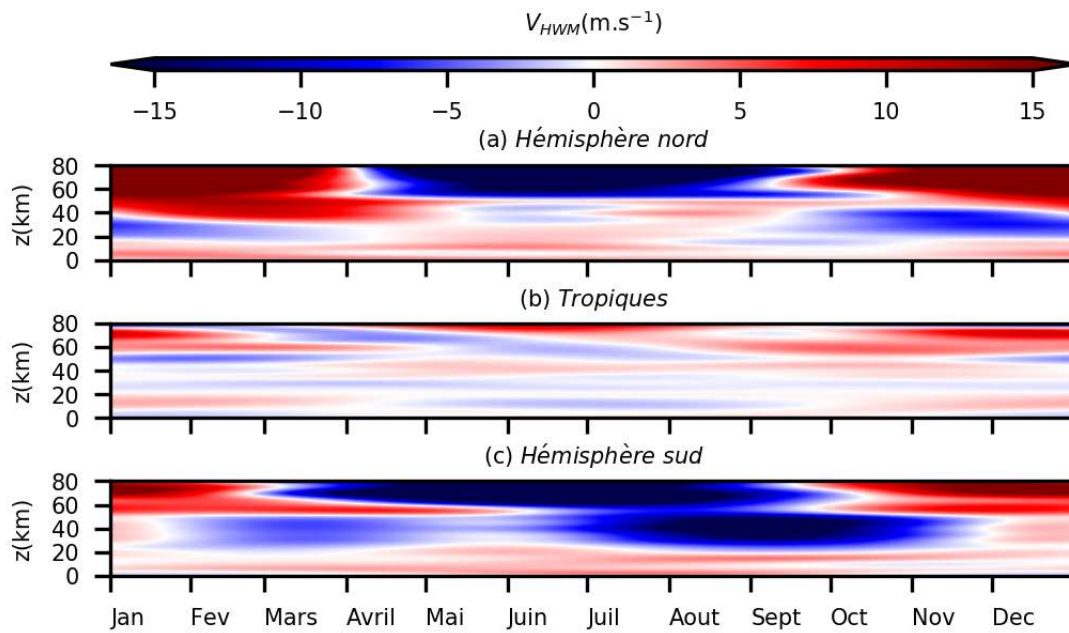


Figure 2.4 : Distribution moyenne du vent méridien au cours d'une année en trois points du globe. (a) 60 °N, 0 °W. (b) 0 °N, 0 °W. (c) 60 °S, 0 °W. Les données sont issues du modèle semi-empirique HWM (Drob *et al.*, 2008).

elle, se produit simultanément sur l'ensemble de la couche.

La partie basse de la moyenne atmosphère équatoriale est le siège d'une autre oscillation du vent zonal. Cette oscillation est quasi-biennale (*Quasi Biennial Oscillation* ou QBO) en référence à sa période moyenne qui va de 22 à 34 mois. L'amplitude maximale de l'oscillation est de l'ordre de 20 à 30 $m \cdot s^{-1}$ vers 30 hPa. Cette oscillation a été découverte simultanément par Reed *et al.* (1961) et Veryard et Ebdon (1961). Elle a fait depuis l'objet de nombreuses publications (voir Baldwin *et al.* (2001) pour une revue détaillée). Bien qu'étant un phénomène tropical, cette oscillation a un impact non-négligeable sur les régions extra-tropicales, notamment à travers la modulation des vortex polaires.

La circulation zonale dans les zones extra-tropicales est caractérisée par de forts jets zonaux (figure 2.5 (a-c)) délimitant les vortex polaires. Ces vents résultent des différences de température entre l'équateur et les pôles que l'on peut observer sur la figure 2.6 (a-c). Ces vents sont dits « thermiques ». Les équilibres hydrostatique (gradient vertical de pression et force de gravité) et géostrophique (force de Coriolis et gradient horizontal de pression) font que, si l'on a un gradient méridien de température, il engendrera un gradient vertical du vent zonal. En hiver, la température est plus faible au pôle qu'à l'équateur, et cela génère un vent thermique d'Ouest (qui va de l'Ouest vers l'Est) pour l'hémisphère Nord et Est pour l'hémisphère Sud. En été, la température est plus forte au pôle qu'à l'équateur, et cela génère un vent thermique d'Est dans l'hémisphère Nord et d'Ouest dans l'hémisphère Sud. Ces jets sont maximaux au moment des solstices et alternent aux équinoxes.

En hiver, la circulation zonale peut être altérée par le déferlement d'ondes planétaires d'amplitude particulièrement larges (Matsuno, 1971). Le déferlement de ces ondes freine le vortex polaire et le réchauffe. On parle alors de *réchauffement stratosphérique soudain* (SSW en anglais pour *Sudden Stratospheric Warming*, en référence à leur vitesse d'apparition). Ces réchauffements peuvent être suffisamment forts pour déplacer le vortex polaire sur les régions tempérées ou le scinder. On parle alors de réchauffement stratosphérique soudain majeur. Ces réchauffements peuvent être identifiés par une inversion de la direction du vent zonal à 10 hPa et au 60^{ème} °N (Charlton & Polvani,

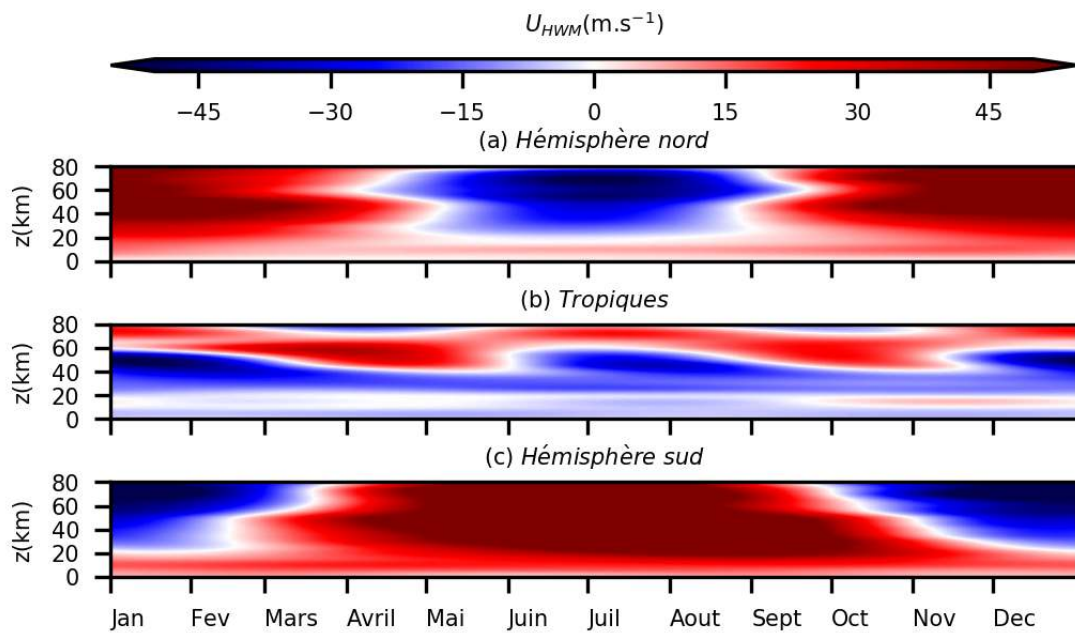


Figure 2.5 : Distribution moyenne du vent zonal au cours d'une année en trois points du globe. Même points et données que sur la figure 2.4.

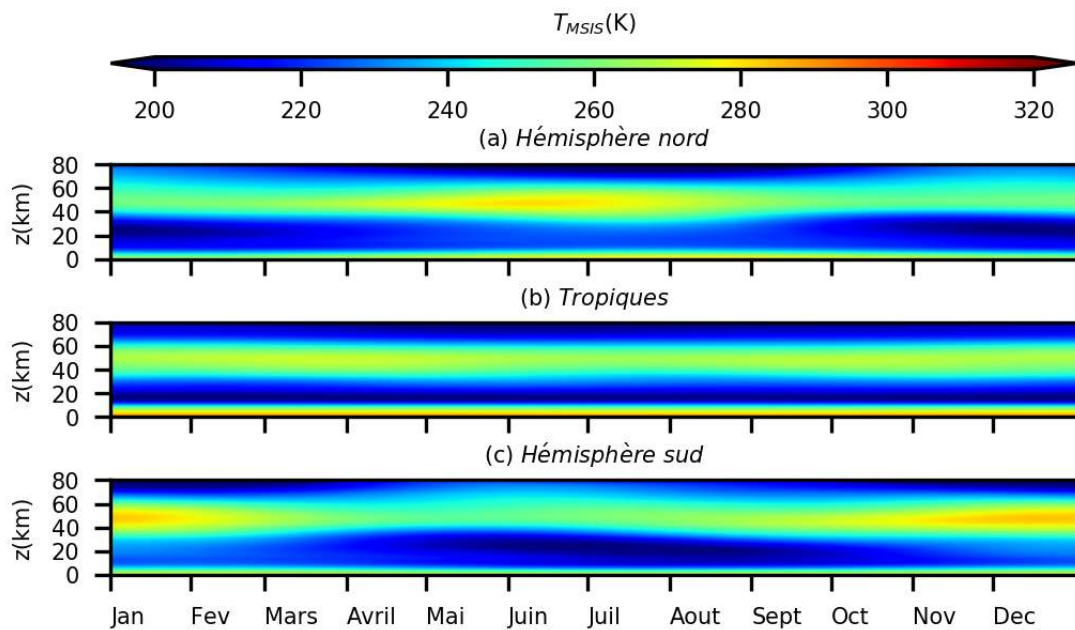


Figure 2.6 : Distribution moyenne de la température au cours d'une année en trois points du globe. Les points sont les mêmes que ceux de la figure 2.4. Les données sont issues du modèle semi-empirique MSIS (Picone *et al.*, 2002).

	Longueur caractéristique (L)	Échelle climatologique	Échelle synoptique	Méso-échelle	Micro-échelle
Temps caractéristique (T)		T > 10 jours	T > 1/2 jours	T > 1 heure	T ≤ 1 heure
Macroéchelle	L > 2000 km	Ondes de Kelvin	Ondes de Rossby Cyclones tropicaux	Orages	Turbulence
Mésoéchelle	L > 2 km				
Microéchelle	L ≤ 2 km				

Tableau 2.1 : Exemple de phénomènes de différents échelles présents dans l'atmosphère.

2007). Le réchauffement peut être plus modéré et n'induire qu'une inversion du gradient méridien de température entre 60 °N et le pôle Nord (Gao *et al.*, 2011). Ces anomalies apparaissent plus fréquemment dans l'hémisphère Nord que dans l'hémisphère Sud, car moins d'ondes planétaires y sont présentes. Elles impactent drastiquement la circulation dans la moyenne atmosphère, et on montre aujourd'hui qu'elles peuvent conduire à des anomalies de température de surface en Europe de l'Ouest comme cela a été le cas en février 2018 (Vargin & Kiryushov, 2019).

2.2 Les modèles d'atmosphère

L'atmosphère est le siège de phénomènes caractérisés par une large gamme d'échelles spatiales et temporelles, allant des ondes planétaires de Rossby de longueurs d'onde de l'ordre de 10^7 m et de période de l'ordre du mois, à la turbulence d'échelle caractéristique de l'ordre de 10 m et des échelles de temps de l'ordre de la seconde comme indiqué au tableau 2.1.

Le tableau 2.1 ne propose qu'un exemple par catégorie d'échelles. Il s'inspire d'un tableau plus exhaustif présenté par Coiffier (2000). Une première approche pour estimer l'état atmosphérique est de considérer que, comme son état en un point du globe est fortement dépendant des apports énergétiques variant avec la rotation terrestre autour de son axe et autour du Soleil, il suffit de connaître le jour solaire pour obtenir l'état de l'atmosphère. Cette approche permet alors de définir des modèles empiriques dépendant uniquement de la localisation sur le globe, de l'heure solaire et du jour de l'année. Un autre moyen de décrire l'atmosphère est de résoudre numériquement les équations de la mécanique des fluides qui régissent son évolution.

2.2.1 Modèles empiriques

Les modèles empiriques présentés ici ont été développés initialement pour l'étude de la circulation thermosphérique et au-delà, ainsi que de sa composition fortement ionisée. Bien que le comportement dominant de la circulation générale dans la thermosphère soit aujourd'hui bien compris (Rishbeth, 1972 ; Killeen, 1987), les mécanismes associés aux interactions entre les ondes et l'écoulement, ainsi que le forçage de la moyenne atmosphère induit sur la thermosphère, ne sont pas encore bien connus. La documentation de la dynamique de la thermosphère est principalement déduite des observations. Les observations montrent que la thermosphère possède une variation annuelle et semestrielle du vent dans la direction zonale, due à la propagation des ondes planétaires stationnaires. Il existe également une variabilité journalière imputable au cycle diurne et une variabilité semi-diurne liée aux marées solaires.

Le modèle empirique *Horizontal-Wind-Model* (HWM), développé par Hedin *et al.* (1988), propose une estimation du vent dans la thermosphère en fonction de l'heure locale, du jour de l'année, de la localisation sur le globe. Le principe de ce modèle

consiste à décomposer les champs horizontaux de vent³ à certaines altitudes spécifiques en harmoniques sphériques de surface et à ajuster les coefficients à partir des mesures. Le nombre d'harmoniques calculés est alors dépendant du nombre de données disponibles. Ces champs sont interpolés verticalement à l'aide d'une *B-spline* cubique pour conduire à des profils verticaux continus.

Dans sa dernière version HWM14 (Drob *et al.*, 2015), le modèle utilise pour estimer ces harmoniques près de 80×10^6 mesures du vent atmosphérique effectuées par 40 instruments. Ces données couvrent une période d'environ 60 ans comme le montre l'inventaire présenté au tableau 2.2. Toutes les données sont alors classées en fonction de leur longitude, de leur latitude, de l'heure du jour, du jour de l'année et de la gamme d'altitudes. Le modèle a été étendu successivement vers la moyenne atmosphère puis la troposphère. Pour assurer une tendance réaliste de la circulation générale dans la troposphère et la stratosphère, HWM utilise 5 années⁴ d'analyses du modèle de PNT *Global Forecast System* (GFS) de la *National Center Environment Prediction* (NCEP) et GEOS-4 de la *National Aeronautics and Space Administration* (NASA). Ce modèle se limite aux premiers harmoniques sphériques de surface et donc ne décrit pas de modes de variation de petite échelle. Comme la propagation des infrasons est sensible notamment aux petites échelles, ce modèle n'est, dans ce contexte précis, utilisable que pour estimer le vent d'au-delà de la moyenne atmosphère. Dans le cadre de cette thèse, il sera utilisée une version plus ancienne du modèle HWM, celle de 2007 (HWM07) (Drob *et al.*, 2008).

Pour la température et la masse volumique, Hedin (1983) propose de calculer les harmoniques sphériques de surface à partir des mesures des spectromètres de masse et d'interférences incohérentes réalisées au cours de la période 1965-1983. Le même type de méthode que celle pour HWM est employé. Dans sa version initiale, le modèle ne décrivait pas l'atmosphère en dessous de 85 km d'altitude. Comme pour HWM, le modèle a été étendu vers les couches de l'atmosphère plus basses par Hedin (1991). Pour retrouver la climatologie de la moyenne atmosphère, il utilise les moyennes zonales fournies par le *Middle Atmosphere Program* (Barnett & Corney, 1985). On notera que les estimations des profils verticaux sont faites de manière à ce qu'ils respectent l'équilibre hydrostatique. La dernière version en date, NRL-MSIS-00 (Picone *et al.*, 2002), intègre aujourd'hui encore plus de données.

Des exemples des cycles annuels du vent et de la température issus de ces modèles sont présentés sur les figures 2.4, 2.5 et 2.6. Si l'on compare ces estimations avec la situation météorologique réelle d'un jour particulier, le modèle MSIS diffère fortement des champs analysés par des modèles de prévision numérique du temps (PNT) comme sur la figure 2.7 avec les modèles de PNT Action de Recherche Petite Échelle Grande Échelle (ARPEGE) de Météo-France (MF) et *Integrated Forecasting System* (IFS) du Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT). Cette figure représente la température à 50 hPa (21 km d'altitude), soit dans la basse stratosphère. Le modèle MSIS-90 propose une moyenne saisonnière pour ce jour de l'année, or en 2018, un réchauffement stratosphérique soudain a affecté la dynamique de l'hémisphère Nord autour du 15 février. On remarque que les deux modèles de PNT globaux décrivent la même anomalie même si la température est légèrement plus élevée dans l'analyse d'IFS (figure 2.7 (b)).

Aujourd'hui, il existe des modèles allant de la surface à la thermosphère comme le modèle G2S développé par Drob *et al.* (2003) et modifié par Schwaiger *et al.* (2019). Ils combinent les champs des modèles de PNT analysés avec ceux proposés par les modèles empiriques HWM et MSIS.

3. En réalité ce sont les champs de divergence et de rotationnel.

4. Analyse de 2002 à 2007. GFS est utilisé entre 0 et 35 km et GEOS-4 entre 0 et 55 km d'altitude).

Nom de code	Coordonnées des observations (lon, lat)	Altitude min (km)	Altitude max (km)	Années concernées	Jour/Nuit	Jours de mesure	Nombre de points
Fabrice-Pérot Interferometer ($\lambda=630$ nm)							
Arecibo	18.4°N, 66.8°W	250	250	1980-1999, 2012-2013	Nuit	901	43632
Arequipa	16.2°S, 71.4°W	250	250	1983-2001, 2007-2013	Nuit	1308	48685
Arrival Heights	77.8°S, 116.7°E	250	250	2002-2005	Nuit	535	54214
Halley Bay	75.5°S, 26.6°W	250	250	1988-1998	Nuit	799	82614
Millstone Hill	42.6°N, 71.5°W	250	250	1989-2002	Nuit	1770	66333
Mount John	44.0°S, 170.4°E	89	250	1991-1996	Nuit	560	2660
Sonderstrom	67.0°N, 51.0°W	250	250	1984-2004	Nuit	1223	69734
South pole	90.0°S	86	250	1989-1999	Nuit	1091	361604
Svalbard	78.2°N, 15.6°E	250	250	1980-1983	Nuit	44	7472
Thule	76.5°N, 68.4°W	250	250	1987-1989	Nuit	172	21500
Resolute Bay	74.7°N, 94.9°E	250	250	2003-2005	Nuit	166	5299
Watson Lake	60.1°N, 128.6°W	250	250	1991-1992	Nuit	135	28000
Jicamarca	11.96°S, 76.86°W	250	250	2009-2013	Nuit	318	10056
Movil	14.97°S, 74.89°W	250	250	2011-2013	Nuit	293	10412
PARI	35.2°N, 82.85°W	250	250	2011-2013	Nuit	166	12610
Poker Flat	65.1°N, 147.5°W	250	250	2009-2011	Nuit	297	5983090
RENOIR	6.89°S, 38.56°W	250	250	2009-2012	Nuit	637	37301
Satellite							
A-E-NATE	± 18°N	220	400	1975-1979	Tout	799	200500
DE-2 WATS	± 89°N	200	600	1981-1983	Tout	536	391500
DE-2 FPI	± 89°N	250	250	1981-1983	Tout	308	47600
UARS HRDI	± 72°N	50	115	1993-1994	Jour	834	30100000
UARS WIND2 (557.7 nm)	± 72°N	90	300	1991-1996	Jour	949	24672000
UARS WIND2 (630.0 nm)	± 42°N	200	300	1991-1996	Nuit	243	2237942
GOCE	± 83.4°N	253	295	2009-2012	crépuscule	813	6613172
Rocket							
Falling Sphere	8°S-60°N	8	98	1969-1991	Tout	1186	96205
Rocketsonde	38°S-70°N	2	90	1969-1991	Tout	5082	843000
TMA	31°S-70°N	59	277	1956-1998	Tout	276	92792
Incoherent Scatter Radar							
Arecibo	18.3°N, 66.8°W	100	170	1974-1987	Jour	149	30600
Chatanika	65.1°N, 147.4°W	90	130	1976-1982	Jour	97	38721
EIS	69.1°N, 19.2°E	100	120	1985-1987	Jour	29	2900
Millstone Hill	42.6°N, 71.5°W	120	400	1983-1987	Tout	142	23536
Sondrestrom	67.0°N, 50.9°W	150	400	1983-1987	Tout	146	19600
St. Santin	44.6°N, 2.2°E	90	165	1973-1985	Jour	256	18382
Medium-Frequency radar							
Adelaide	34.5°S, 138.5°E	60	98	2001-2004	Tout	834	481634
Birbe Island	28.0°S, 153.0°W	60	98	1995	Tout	280	184176
Davis	68.6°N, 147.5°W	44	108	2001-2004	Tout	730	526160
Poker Flat	65.1°N, 147.5°W	44	108	1979-1985	Tout	1857	2746684
Wakkanai	45.4°N, 141.8°E	50	108	1996-2003	Tout	1538	1874672
Yamagawa	31.2°N, 130.6°E	60	98	1998-2003	Tout	1593	1040042
Wind and temperature LIDAR							
Fort Collins	40.6°N, 105.1°W	75	115	2002	Tout	244	93288

Tableau 2.2 : Base de données utilisée par HWM. En bleu celles modifiées par la version de 2014 (Drob et al., 2015). Ce tableau provient de Drob et al. (2008).

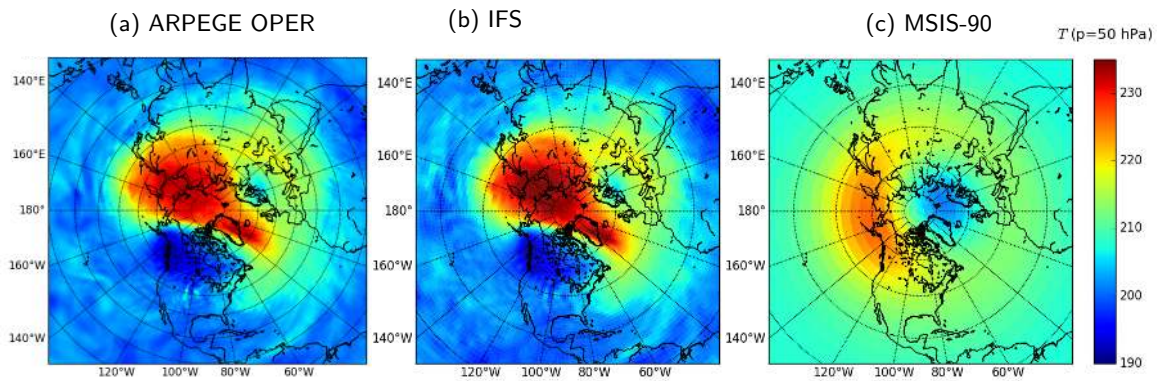


Figure 2.7 : Projection stéréographique polaire du champ de température (K) à 50 hPa (≈ 21 km) 15/02/2018 00 UTC. (a) Niveau 21 (≈ 50 hPa) de l'analyse haute-résolution d'ARPEGE. (b) Niveau 48 (≈ 50 hPa) de l'analyse de IFS du CEPMMT. (c) Climatologie fournie par le modèle MSIS-90 à 21 km d'altitude (≈ 50 hPa) pour le 15 février.

2.2.2 Les modèles globaux de prévision numérique du temps

Nous présentons ici les principales caractéristiques des modèles globaux de Prévision Numérique du Temps (PNT) utilisés de manière opérationnelle par les grands centres météorologiques mondiaux et référencés au tableau 2.3.

Nous illustrerons notre propos avec deux modèles de PNT qui ont été exploités dans nos études : le modèle ARPEGE de Météo-France et le modèle IFS du CEPMMT. Les modèles globaux de PNT simulent l'évolution temporelle de l'atmosphère terrestre entre la surface et une altitude généralement comprise entre 50 et 80 km. L'air est considéré comme un fluide non-visqueux en rotation formé d'un mélange de gaz parfaits (azote, oxygène et vapeur d'eau principalement) et de particules comme les gouttelettes nuageuses ou les gouttes de pluie. Le fluide atmosphérique est gouverné par les lois d'évolution suivantes :

- La seconde loi de Newton dite de conservation de la quantité de mouvement ;
- L'équation de continuité (ou de conservation de la masse) ;
- L'équation d'état des gaz parfaits ;
- Le premier principe de la thermodynamique ou conservation de l'énergie ;
- L'équation de conservation de la masse pour l'eau atmosphérique.

À ces équations, il faut ajouter des conditions aux limites en surface et au sommet de la région de l'atmosphère que l'on considère.

Les modèles globaux résolvent des équations dites primitives⁵ dans lesquelles est faite l'approximation hydrostatique qui consiste à négliger l'accélération verticale dans l'équation de conservation de la quantité de mouvement. Cette approximation se traduit par l'équilibre (2.1) entre le gradient vertical de pression $\frac{\partial p}{\partial z}$ et la force de gravité ρg , où ρ est la masse volumique de l'air.

$$\frac{\partial p}{\partial z} = -\rho g. \quad (2.1)$$

Elle est justifiée par le fait que, dans l'atmosphère aux échelles dites synoptiques (plusieurs centaines de kilomètres) résolues par les modèles globaux, les mouvements horizontaux

5. Ce terme s'oppose aux équations filtrées ayant des hypothèses supplémentaires d'équilibre entre le champ de masse et le champ de vent et qui ont été utilisées pour développer les premiers modèles de prévision numérique.

Modèle	Origine	Résolution horizontale	Niveaux verticaux	Sommet du modèle
ARPEGE	Météo-France (France)	Spectrale T_L 1798 C2.2 (5 km sur la France et 24 km sur la Nouvelle-Zélande)	105 niveaux pression hybride	0.1 hPa (\approx 65 km)
IFS	CEPMMT (Europe)	Spectrale T_C 1279 (8 km sur le globe)	137 niveaux pression hybride	0.01 hPa (\approx 80 km)
GFS	NOAA (États-Unis d'Amérique)	Spectrale T 1534 (13 km sur le globe)	64 niveaux pression hybride	0.27 hPa (\approx 55 km)
ICON	DWD (Allemagne)	Icosaédrique (13 km sur le globe)	90 niveaux SLEVE (Leuenberger et al., 2010)	0.01 hPa (\approx 80 km)
UKMO	MetOffice (Royaume-Uni)	Point de grille (10 km sur le globe)	70 niveaux pression hybride	0.01 hPa (\approx 80 km)
GEM	CMC (Canada)	Point de grille Ying-Yang (Qaddouri & Lee, 2011)	81 niveaux pression hybride	0.1 hPa (\approx 80 km)
ACCESS-G	AGBU (Australie)	Point de grille (25 km sur le globe)	70 niveaux altitude hybride	76 km
JMA-GO	JMA (Japon)	Spectrale T_L 959 (20 km sur le globe)	100 niveaux pression hybride	0.01 hPa (\approx 80 km)

Tableau 2.3 : Modèles PNT globaux disponibles. Valeurs au 12 juillet 2019.

sont généralement d'un ordre de grandeur supérieur aux mouvements verticaux. Elle permet de filtrer les ondes sonores tout en gardant les ondes d'un intérêt météorologique comme les ondes de Rossby ou les ondes d'inertie-gravité.

Les équations primitives sont discrétisées dans le temps et dans l'espace pour permettre leur résolution numérique. Le choix des schémas de discrétisation est très important en PNT opérationnelle, car ils doivent être précis et rapides pour permettre la production au plus tôt des prévisions pour les besoins des prévisionnistes. Dans les modèles globaux, les équations sont écrites dans un système de coordonnées sphériques sur l'horizontale et un système de coordonnée verticale basé sur la pression plutôt que sur l'altitude. Si l'on suppose que l'écoulement respecte l'équilibre hydrostatique, alors la pression et l'altitude sont reliées au travers de l'hypsométrie (2.2). Cette équation fait intervenir une autre grandeur utilisée en météorologie qui est le géopotential Φ . Cette grandeur peut être approchée par le produit de la gravité g avec l'altitude z .

$$d\Phi = -\frac{R}{M_{air}} T d \log(p) \quad (2.2)$$

où $R = 8,314 \text{ J.K}^{-1}.\text{mol}^{-1}$ et M_{air} est la masse molaire de l'air. Dans les basses couches de l'atmosphère, on préfère utiliser un système de coordonnées dit « σ » qui suit le relief en utilisant le rapport de la pression à la pression de surface p_{gnd} ($\sigma = p/p_{gnd}$). C'est ainsi que, suivant la proposition de Simmons et Burridge (1981), un système de coordonnée verticale hybride est adopté dans les modèles ARPEGE et IFS avec une coordonnée sigma dans la troposphère se transformant de manière continue en une coordonnée pression dans les niveaux supérieurs. Un exemple de ces niveaux est présenté sur la figure 2.8. La disposition des variables sur la verticale dite de « Lorenz » est telle que le géopotential est estimé aux demi-niveaux et les autres quantités aux niveaux entiers.

La discrétisation horizontale la plus simple est basée sur l'utilisation de différences finies. Le choix des grilles de discrétisation est conditionné par des propriétés de conservation des principales quantités et de propagation des ondes. Toutefois pour les modèles globaux, la convergence des méridiens aux pôles pose un certain nombre de problèmes numériques, comme la diminution de la distance entre les points de calcul. C'est ainsi que des approches basées sur la méthode spectrale ont été proposées dès le début du développement des modèles globaux de PNT. L'utilisation de la méthode spectrale sur la sphère a été grandement facilitée par l'avènement des Transformées de Fourier Rapides (Fast Fourier Transforms : FFT) qui ont rendu cette méthode compétitive par rapport à la méthode en points de grille. C'est cette méthode de discrétisation spatiale qui est utilisée dans les modèles ARPEGE et IFS. Les champs atmosphériques sont représentés sous forme d'un développement en série sur une base de fonctions. Les fonctions choisies sont les harmoniques sphériques de surface qui sont des fonctions propres du laplacien sur la sphère.

$$Y_n^m(\lambda, \mu) = P_n^m(\mu) e^{im\lambda} \quad (2.3)$$

où μ est le sinus de la latitude et λ la longitude. Sur un cercle de latitude, $Y_n^m(\lambda, \mu)$ est périodique et m peut être interprété comme un nombre d'onde zonal (nombre de longueurs d'onde). P_n^m sont les fonctions de Legendre du premier ordre de degré n et d'ordre m . L'écriture d'un champ météorologique A sur une base d'harmoniques sphériques doit être tronquée. La troncature T exprime les nombres d'ondes n et m qui doivent être utilisés dans l'expression du champ :

$$A(\lambda, \mu) = \sum_{(n,m) \in T} A_n^m Y_n^m(\lambda, \mu). \quad (2.4)$$

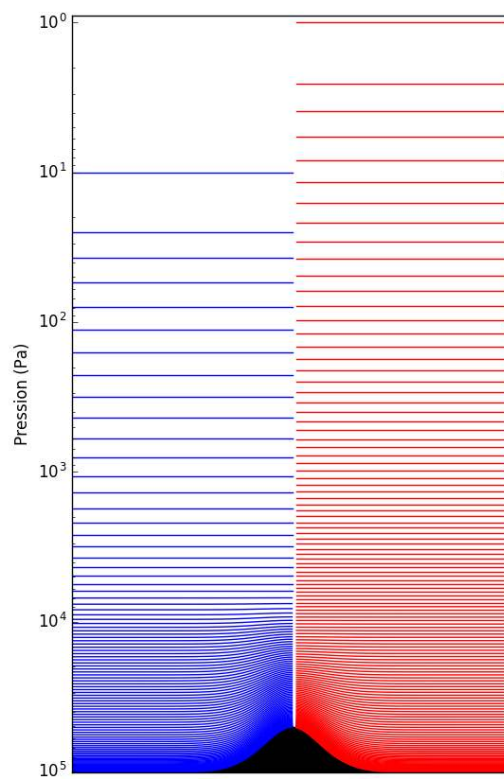


Figure 2.8 : Distribution des niveaux hybrides en pression au-dessus d'un même relief pour les 105 niveaux d'ARPEGE (bleu) et pour les 137 niveaux d'IFS (rouge).

Plusieurs types de troncature existent, on utilise pour les modèles ARPEGE et IFS une troncature dite « triangulaire » à l'ordre M définie par

$$T = \{(n, m), 0 \leq |m| \leq M, 0 \leq |n| \leq M\}. \quad (2.5)$$

La propriété principale de cette troncature est de permettre une résolution uniforme des champs sur la sphère. Avec cette propriété d'isotropie, il a été montré qu'elle permet de sélectionner les nombres d'ondes qui contribuent le plus à l'énergie cinétique du fluide atmosphérique (Machenhauer & Rasmussen, 1972 ; Baer, 1972).

Les propriétés des harmoniques sphériques permettent de calculer facilement et précisément les dérivées spatiales des champs. Lorsque des produits de termes doivent être calculés (comme les termes d'advection du type $U\partial A/\partial x$), on calcule la valeur dans l'espace physique de chaque terme sur une grille dite de transformation. On effectue le produit dans l'espace physique et on réalise une transformée spectrale du résultat. Ce retour dans l'espace spectral nécessite le calcul d'intégrales qui, pour être exact, utilise une quadrature trapézoïdale sur un cercle de latitudes avec au moins $3M + 1$ points et une quadrature de Gauss (latitudes μ_k et poids associés aux zéros du polynôme de Legendre P_k^0) entre le pôle Sud et le pôle Nord avec au moins $(3M + 1)/2$ points. Si cette condition n'est pas remplie, les nombres d'ondes au-delà de la troncature générés par les non-linéarités se projettent sur les nombres d'ondes résolus et conduisent au phénomène d'aliasing. Cette grille de transformation, dite *grille de Gauss quadratique*, est régulière sur un cercle de latitude. La convergence des méridiens vers les pôles conduit à une distance réduite entre deux points consécutifs sur un cercle de latitude. Hortal et Simmons (1991) ont montré qu'il était possible de réduire le nombre de points près des pôles, ce qui permet une réduction des coûts de calcul, sans perte en précision (car la contribution des nombres d'ondes les plus élevés aux coefficients spectraux devient extrêmement faible). Récemment, le CEPMMT a adopté pour le modèle IFS une nouvelle grille réduite octaédrique (méthode différente de décroissance des points le long des cercles de latitude en se rapprochant des pôles). La contrainte sur le nombre de points de la grille de Gauss permettant un calcul exact des termes quadratiques a pu être relâchée avec l'utilisation de schémas d'advection dits « Semi-Lagrangiens » (SL) où la tendance totale d'un champ dA/dt est calculée directement en un point de grille A_G en estimant le point origine A_O d'une trajectoire, soit : $(A_G(t + \Delta t) - A_O(t - \Delta t))/2\Delta t$ au lieu de la forme eulérienne $\partial A/\partial t + U\partial A/\partial x + V\partial A/\partial y$. L'avantage des schémas SL est d'autoriser des pas de temps plus longs pour le modèle de PNT tout en gardant une bonne précision sur les calculs. Il est alors possible d'utiliser une grille de calcul dite « linéaire » qui nécessite uniquement le traitement exact des termes linéaires, soit $2M + 1$ points sur un cercle de latitude et $(2M + 1)/2$ entre les deux pôles. C'est cette grille qui est utilisée dans le modèle ARPEGE alors que le modèle IFS est récemment passé à une grille de calcul permettant un calcul exact des termes d'ordre 3 (grille cubique) pour mieux réduire l'aliasing résiduel (nécessitant alors $(4M + 1)$ points par cercle de latitude). La grille de transformation permet de formuler, pour une troncature donnée M , la résolution physique correspondante avec une grille linéaire : $\Delta x = (2\pi a)/(2M + 1)$ où a est le rayon de la Terre.

Un autre avantage de la méthode spectrale pour les modèles globaux est de permettre de traiter facilement de manière implicite les termes linéaires (et linéarisés) dans les équations primitives (essentiellement, car une telle résolution conduit à la résolution d'un système d'équations elliptiques de type Helmholtz faisant intervenir des laplaciens qui sont faciles à calculer avec les harmoniques sphériques). Comme pour la méthode SL, cette méthode semi-implicite (car les termes non-linéaires sont calculés explicitement) permet d'allonger significativement le pas de temps du modèle sans perdre en précision et

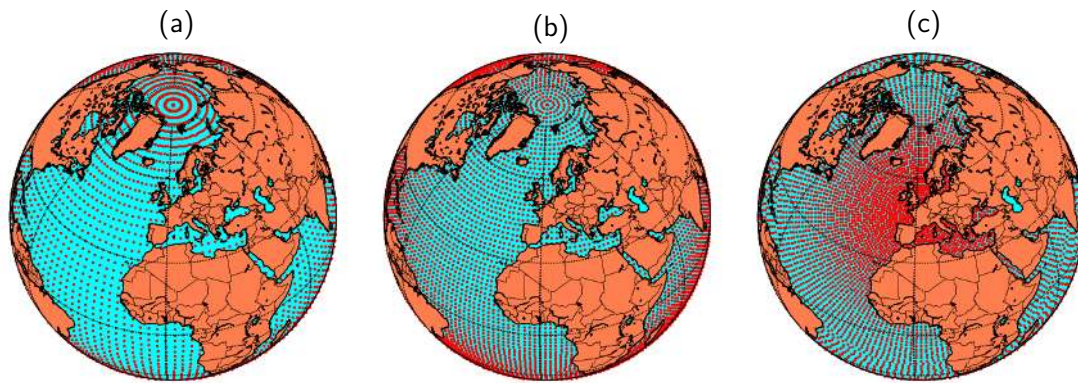


Figure 2.9 : Exemples de grilles horizontales dans une projection ortho centrée sur la France. (a) Une grille uniforme avec un point tout les 1.5 degrés (b) Une grille de Gauss linéaire réduite avec $T_L 479$. (c) Une grille basculée et étirée avec $T_L 798$ et un coefficient d'étirement $C = 2.4$. (Pour la lisibilité de la figure, tous les points de grilles ne sont pas représentés).

en stabilité numérique. Pour le modèle ARPEGE, les équations primitives sont modifiées grâce à une transformation dite « conforme » (rapport entre un déplacement sur la grille transformée et un déplacement sur la sphère ne dépendant pas de la direction du déplacement) sur la sphère décrite par Courtier et Geleyn (1988) qui propose un basculement du pôle au centre de la zone d'intérêt et une augmentation de la résolution de cette zone (avec un facteur d'étirement C) au détriment de la résolution aux antipodes comme le montre la figure 2.9.

La discrétisation spatiale ne permet pas de représenter explicitement certains processus physiques d'échelle inférieure à la plus petite longueur d'onde autorisée par la troncature. On peut citer les processus turbulents dans la couche limite planétaire près de la surface, les nuages et les précipitations associées, les ondes de gravité générées par le relief ou la convection, les interactions entre le rayonnement et les constituants atmosphériques (gaz et particules), les échanges d'énergie, de matière et de quantité de mouvement entre la surface et l'atmosphère. Tous ces processus non résolus affectent cependant les échelles résolues à travers la divergence des flux moyens (à l'échelle de la maille) qu'ils génèrent. Ils sont décrits au moyen de paramétrisations dans lesquelles on tente de représenter ces flux générés par des processus sous-mailles en fonction des variables résolues (il s'agit donc d'un problème de fermeture des équations comme en turbulence.). Les paramétrisations physiques représentent exclusivement des processus de transferts verticaux. Ces flux sont de ce fait calculés dans l'espace physique sous forme de colonnes indépendantes en chaque point de la grille de transformation. Ces modèles PNT globaux fournissent des estimations de l'atmosphère tout au long de l'année, comme le montrent les figures 2.10, 2.11 et 2.12. Elles représentent la température et les composantes horizontales du vent analysées au cours de l'année 2018 par le modèle IFS du CEPMMT. On remarque des structures cohérentes avec celles montrées précédemment avec les modèles HWM et MSIS (figures 2.4, 2.5 et 2.6.). La circulation est néanmoins éloignée de celle fournie par les modèles empiriques. On remarque notamment l'inversion du vent zonal liée au SSW de février 2018 sur la figure 2.11 (a).

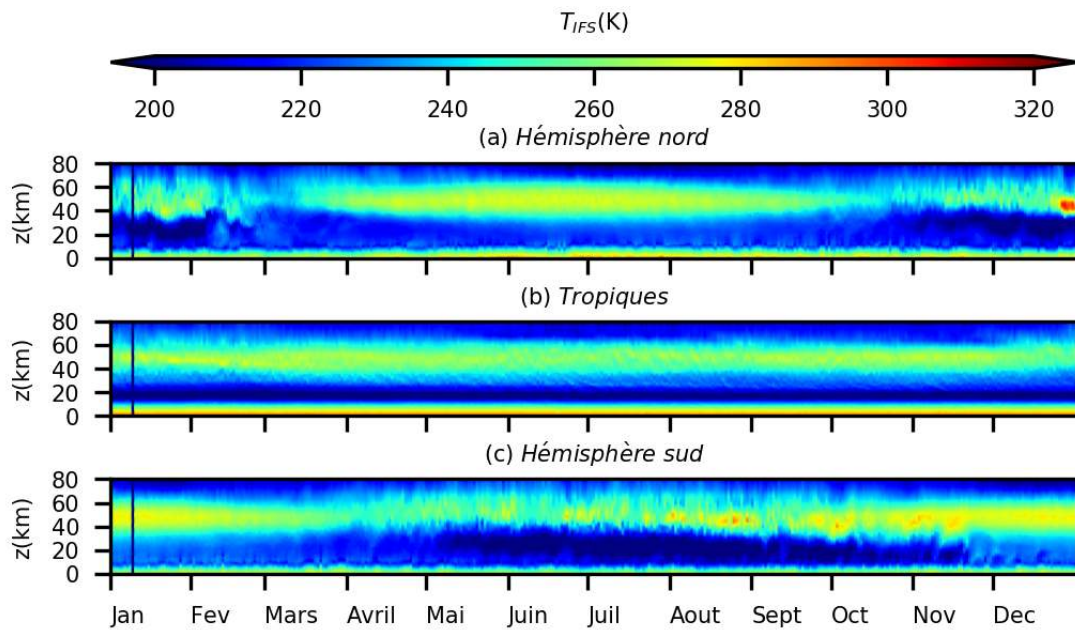


Figure 2.10 : Distribution de la température analysée par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018.

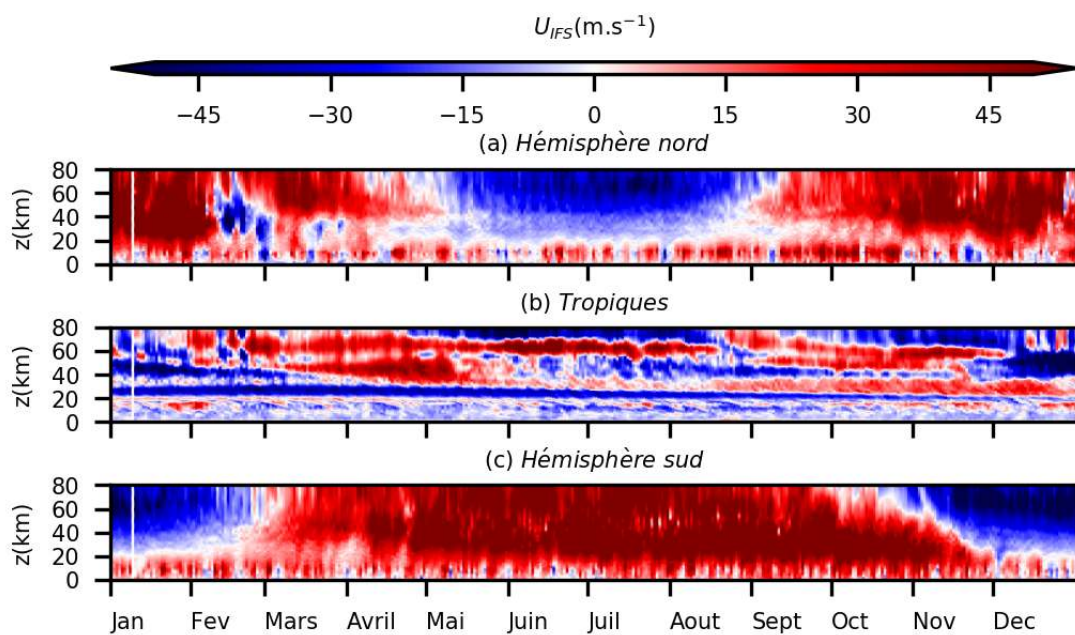


Figure 2.11 : Distribution du vent zonal analysé par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018.

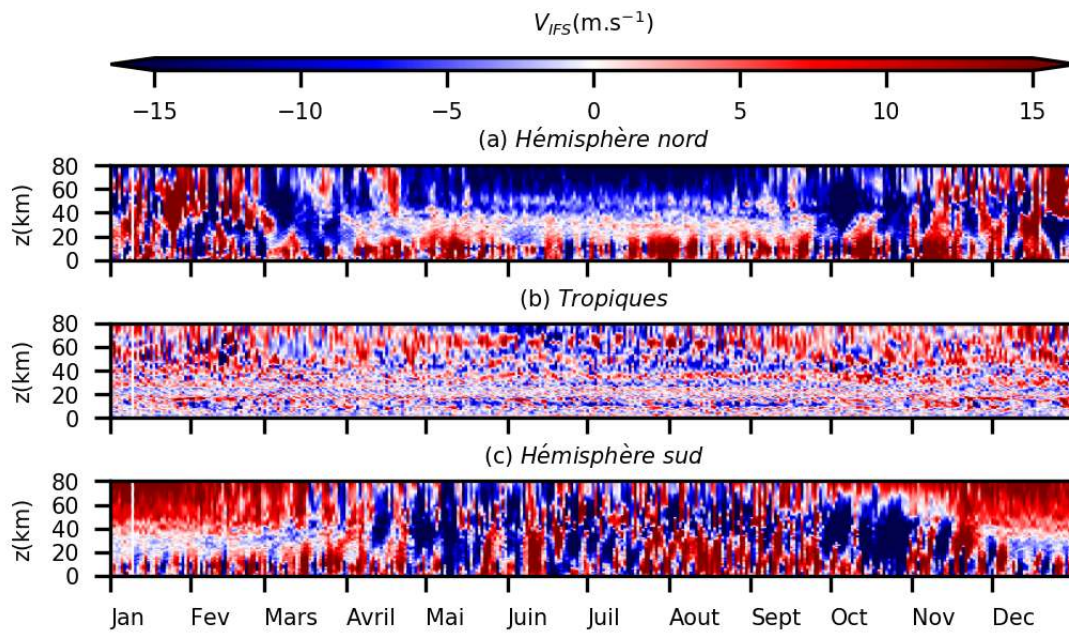


Figure 2.12 : Distribution du vent méridien analysé par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018.

2.3 Assimilation de données

2.3.1 Formulation du problème variationnel

La qualité de la prévision météorologique est fortement liée à la précision de l'état initial utilisé. Cet état initial est décrit par deux sources d'informations : une prévision à partir d'un état antérieur et des observations météorologiques. L'avantage de la prévision, c'est que l'état de l'atmosphère initial fourni respecte les équations de la dynamique de l'atmosphère, mais elle peut être éloignée de l'état réel. L'avantage des observations est de représenter l'information la plus précise que l'on ait sur l'état réel de l'atmosphère, mais elles se limitent aux domaines et aux champs qu'elles sondent. L'état initial le plus précis peut être considéré comme le meilleur compromis entre une prévision du modèle et les observations. Pour cela on se propose de résoudre le problème bayésien (2.6) et d'estimer la densité de probabilité \mathcal{P} d'un état atmosphérique \mathbf{X} prédit par le modèle PNT connaissant les observations \mathbf{Y} .

$$\mathcal{P}(\mathbf{X}|\mathbf{Y}) = \frac{\mathcal{P}(\mathbf{X})\mathcal{P}(\mathbf{Y}|\mathbf{X})}{\mathcal{P}(\mathbf{Y})}, \quad (2.6)$$

On considère alors que l'état maximisant la densité de probabilité $\mathcal{P}(\mathbf{X}|\mathbf{Y})$ est le meilleur estimateur a posteriori de l'état atmosphérique. Pour calculer cet estimateur, les centres de météorologie opérationnelle utilisent l'assimilation de données. Cette technique estime l'état du modèle qui s'accorde le mieux avec les observations disponibles sur une fenêtre temporelle. Il est important de noter que le terme « disponible » est essentiel ici. En effet, la quantité d'observations à assimiler dépend du temps que l'on s'accorde pour produire un état \mathbf{X} . L'information a priori est généralement donnée sous forme d'une prévision à court terme du modèle de PNT (trajectoire de l'état de l'atmosphère à partir d'un instant initial t_0 sur la fenêtre temporelle choisie). Cette première estimation nommée ébauche sera notée \mathbf{X}^b par la suite. Cette prévision couvre toute la fenêtre que l'on s'est donnée. On suppose que le modèle de PNT peut être modélisé comme un opérateur

discret permettant de passer d'un état à l'instant t à un état à l'instant $t + \Delta t$ et que l'on notera \mathbf{M} . Si ce modèle est parfait, alors toute la trajectoire \mathbf{X}^b peut être obtenue en appliquant suffisamment de fois le modèle \mathbf{M} (non-linéaire) à l'état initial que l'on a en t_0 et que l'on note \mathbf{X}_0^b . Donc, pour tous pas de temps $t_i = t_0 + i.\Delta t$ de la fenêtre, on a :

$$\mathbf{X}_i^b = \mathbf{M}^i(\mathbf{X}_0^b). \quad (2.7)$$

où $\mathbf{M}^i = \mathbf{M}.\mathbf{M}.\dots.\mathbf{M}$ est la i -ème composition successive de l'opérateur \mathbf{M} . Bien que le modèle soit considéré comme parfait, l'ébauche est quant à elle entachée d'incertitudes dont la matrice de covariances d'erreurs est notée \mathbf{B} . On suppose alors que la probabilité a priori de l'état atmosphérique \mathbf{X}_0 suit une loi gaussienne de covariances d'erreurs \mathbf{B} .

$$\mathcal{P}(\mathbf{X}_0) \propto \exp \left[-\frac{1}{2}(\mathbf{X}_0 - \mathbf{X}_0^b)^T \mathbf{B}^{-1}(\mathbf{X}_0 - \mathbf{X}_0^b) \right], \quad (2.8)$$

où l'exposant T représente l'opérateur transposition. Pour pouvoir comparer l'ébauche aux observations, il faut alors utiliser un opérateur dit d'observation, que l'on note usuellement \mathbf{H} . Au premier pas de temps t_1 , m observations différentes de l'atmosphère sont effectuées. Ces observations sont rassemblées dans un vecteur noté $\mathbf{Y}_1 = \bar{\mathbf{Y}}_1 + \xi_1$ de dimension m , où $\bar{\mathbf{Y}}_1$ est le vecteur comportant les valeurs des observations et ξ_1 le vecteur des erreurs d'observation. On définit la matrice $\mathbf{R}_1 = E(\xi_1 \xi_1^T)$ de covariances d'erreurs d'observation, où $E()$ représente l'espérance mathématique (la moyenne). Pour mesurer l'écart de l'atmosphère en \mathbf{X}_1^b à l'instant t_1 avec les observations, on calcule une distance quadratique entre le modèle et les observations pondérée par l'inverse de la matrice \mathbf{R}_1 . L'opérateur \mathbf{H} d'observation peut être non-linéaire en fonction du type d'observation considérée. En effet, les observations peuvent ne pas être des mesures des variables du modèle. Par exemple pour la température de brillance⁶, on utilise un modèle de transfert radiatif dans la colonne atmosphérique comme RTTOV (Saunders *et al.*, 2018) pour construire $\mathbf{H}(\mathbf{X})$. Pour les radiosondages, les valeurs mesurées sont directement des variables du modèle, mais il faut néanmoins les interpoler spatialement au point des observations. De la même manière que précédemment, on suppose que la probabilité $\mathcal{P}(\mathbf{Y}_1|\mathbf{X}_1)$ suit une loi gaussienne telle que :

$$\mathcal{P}(\mathbf{Y}_1|\mathbf{X}_1) \propto \exp \left[-\frac{1}{2}(\mathbf{Y}_1 - \mathbf{H}_1(\mathbf{X}_1^b))^T \mathbf{R}_1^{-1}(\mathbf{Y}_1 - \mathbf{H}_1(\mathbf{X}_1^b)) \right] \quad (2.9)$$

La trajectoire sera celle qui minimisera les écarts aux observations, tout en ne s'éloignant pas trop de l'ébauche. On formalise le problème de la manière suivante. La trajectoire finale devra voir son état en t_0 minimiser la fonction objective J suivante :

$$\left\{ \begin{array}{l} J(\mathbf{X}_0) = J^b + J^o = \frac{1}{2}(\mathbf{X}_0 - \mathbf{X}_0^b)^T \mathbf{B}^{-1}(\mathbf{X}_0 - \mathbf{X}_0^b) + \dots \\ \frac{1}{2} \sum_k (\mathbf{Y}_k - \mathbf{H}_k(\mathbf{X}_k))^T \mathbf{R}_k^{-1}(\mathbf{Y}_k - \mathbf{H}_k(\mathbf{X}_k)) \\ \mathbf{X}_k = \mathbf{M}^k(\mathbf{X}_0). \end{array} \right. \quad (2.10)$$

La minimisation de la fonction objective J de l'équation (2.10) correspond à ce qui est appelé la formulation 4D-Variationnelle (4D-Var) du problème d'assimilation. La trajectoire minimisant la fonction J est appelée analyse. Par la suite, l'analyse sera notée \mathbf{X}^a et sa matrice de covariances d'erreurs d'analyse sera notée \mathbf{A} . Dans le cas où le

6. La température de brillance est la température du corps noir équivalent émettant la luminance spectrale observée.

modèle et les opérateurs d'observation sont linéaires, alors l'analyse et sa matrice de covariances d'erreurs résolvent le problème bayésien (2.6) conduisant à :

$$\mathcal{P}(\mathbf{X}|\mathbf{X}^b, \mathbf{Y}) = \mathcal{N}[\mathbf{X}^a, \mathbf{A}], \quad (2.11)$$

où $\mathcal{N}[\mu, \sigma]$ désigne la loi normale de moyenne μ et d'écart type σ .

La recherche d'extrema locaux de J passe par l'annulation de son gradient par rapport à \mathbf{X}_0 . Cela implique un calcul explicite et la sauvegarde du gradient de $\nabla_{\mathbf{X}_0} J$ vis-à-vis de \mathbf{X}_0 . Les utilisations d'une méthode perturbative et d'un schéma aux différences finies nécessitent autant d'appels à la fonction J qu'il y a de composantes dans \mathbf{X}_0 , ce qui serait numériquement bien trop coûteux pour un modèle de PNT opérationnel. C'est pourquoi ce gradient est calculé par la méthode de l'adjoint (Talagrand, 1997), méthode plus efficace dès lors que la dimension est élevée (Lions, 1971). L'avantage est que la résolution numérique de ce type de problème via l'utilisation de l'adjoint ne nécessite que le stockage en mémoire des trajectoires du modèle, cependant celles-ci restent de très grande dimension. C'est pourquoi, pour résoudre ce problème, Courtier *et al.* (1994) ont proposé une méthode de résolution incrémentale qui est aujourd'hui la base de l'algorithme d'assimilation utilisé dans ARPEGE et IFS entre autres. Pour ce faire, on linéarise les opérateurs \mathbf{M} et \mathbf{H} autour de l'ébauche. En introduisant ces opérateurs linéaires tangents dans l'équation (2.10), on obtient alors une fonction de coût quadratique en l'incrément, c'est-à-dire l'écart à l'ébauche. Généralement, la minimisation de l'incrément est faite à faible résolution. On peut alors utiliser l'incrément optimal ainsi trouvé pour redéfinir une « ébauche » sur laquelle on relinéarise pour recommencer le processus. Cela permet ainsi de gérer les non-linéarités. Par exemple dans le cadre du modèle ARPEGE, une première minimisation est faite pour une grille T_L149 (≈ 130 km entre deux points), puis l'on refait une linéarisation pour effectuer une seconde minimisation à la résolution T_L479 (≈ 40 km entre deux points). Les phases de prévisions sont, quant à elles, effectuées avec l'opérateur non linéaire \mathbf{M} avec la résolution horizontale maximale.

2.3.2 Matrice de covariances d'erreurs d'ébauche \mathbf{B} et assimilation d'ensemble

L'assimilation de données requière une estimation des matrices de covariances d'erreurs de l'ébauche \mathbf{B} et des observations \mathbf{R} . Dans la pratique, l'erreur de mesure et l'erreur de représentativité permettent de construire la matrice \mathbf{R} . L'estimation de la matrice \mathbf{B} est plus complexe. Une première modélisation de la matrice de covariances d'erreurs \mathbf{B}_0 peut être construite à partir de la climatologie. On constitue alors un ensemble de N ébauches perturbées en suivant la formule :

$$\mathbf{X}_k^b = \mathbf{X}^b + \mathbf{B}_0^{1/2} \zeta_k, \text{ où } \zeta_k \in \mathcal{N}(0, 1). \quad (2.12)$$

où \mathbf{X}_k^b désigne la k -ème ébauche de l'ensemble, qui sont usuellement appelées membres. Par construction, les membres échantillonnent \mathbf{B}_0 . À partir de cet ensemble d'ébauches, on pourrait construire un ensemble d'analyses en effectuant une assimilation utilisant les mêmes données sur chacun des membres. Cependant, l'ensemble d'analyses qui en résulterait ne serait pas représentatif de l'erreur d'analyse régie par la matrice de covariances d'erreurs d'analyse \mathbf{A} . Il faut pour ce faire tenir compte des erreurs d'observation. On définit des jeux d'observations perturbées obtenus en modifiant le vecteur \mathbf{Y} initial aléatoirement selon sa matrice de covariances d'erreurs \mathbf{R} . Supposons que l'ensemble d'ébauches soit de taille N , alors les N jeux distincts d'observations sont assimilés pour chacune des N ébauches. Ainsi, on obtient un ensemble d'analyses perturbées, qui est ainsi

représentatif de l'erreur d'analyse régie par \mathbf{A} . On parle alors d'assimilation d'ensemble. A contrario, lorsque l'assimilation se fait avec un unique membre, elle est généralement désignée comme « déterministe ».

Si le modèle est parfait, l'ensemble de prévisions que l'on pourrait construire à partir de l'analyse d'ensemble échantillonnerait l'erreur de prévision et nous permettrait d'actualiser la matrice \mathbf{B} . En réalité, le modèle de PNT possède des erreurs (Daley, 1993) notamment dues aux approximations sur la dynamique et les paramétrisations physiques. Par la suite, on pose e^a l'erreur d'analyse, e^m l'erreur de modèle. L'erreur d'ébauche finale sera donnée par l'erreur de prévision $e^b = \tilde{\mathbf{M}}e^a + e^m$, où $\tilde{\mathbf{M}}$ est le modèle linéaire tangent. Pour compenser l'absence de prise en compte d'une erreur modèle, on amplifie alors les perturbations en suivant trois étapes (Raynaud *et al.*, 2012). Une première étape consiste à estimer la variance, notée $v()$, de l'erreur d'ébauche théorique à partir des écarts entre l'innovation⁷ réelle vue par l'analyse initiale et celle théorique (Desroziers & Ivanov, 2001 ; Chapnik *et al.*, 2004). La variance $v(e^b)$ est alors donnée par :

$$v[\tilde{\mathbf{M}}e^a + e^m] = s^b V_{sp} \quad (2.13)$$

où le terme V_{sp} correspond aux variances d'ébauche considérées pour le calcul de l'analyse précédente et s^b est le rapport entre le minimum de la fonctionnelle J trouvée pour l'analyse déterministe rapportée à sa valeur théorique. La seconde étape consiste à calculer l'amplification⁸ nécessaire à apporter pour que l'ensemble d'analyses perturbées ait une variance égale à celle de (2.13). On définit alors le coefficient d'amplification α tel que :

$$\alpha = \sqrt{\frac{v[\tilde{\mathbf{M}}e^a + e^m](x, y, z, t, T)^{xyz}}{v[\mathbf{M}e^a](x, y, z, t, T)^{xyz}}} \quad (2.14)$$

où \bar{v}^{xyz} représente une moyenne spatiale globale. La définition de α (2.14) est ici construite à partir de la variance moyenne que l'on a sur la température en un instant donné, mais cela peut être fait à partir d'autres variables. Généralement α est supérieur à l'unité, mettant en évidence une sous-estimation de l'incertitude de prévision. Une fois le facteur α déterminé, il est alors appliqué aux membres de l'ensemble. On pose $\bar{\mathbf{X}}^b$ la moyenne arithmétique des membres des prévisions perturbées mais dont les perturbations ne sont pas amplifiées. L'ensemble final des ébauches $\hat{\mathbf{X}}_k^b$ est alors construit de la manière suivante :

$$\begin{cases} \hat{\mathbf{X}}_k^b = \bar{\mathbf{X}}^b + \alpha(\mathbf{X}_k^b - \bar{\mathbf{X}}^b), \text{ pour } k = 1..N \\ v(e^b) \leftarrow \frac{1}{N-1} \sum_{k=1}^N (\hat{\mathbf{X}}_k^b - \bar{\mathbf{X}}^b)^T (\hat{\mathbf{X}}_k^b - \bar{\mathbf{X}}^b). \end{cases} \quad (2.15)$$

Les variances ainsi que les portées spatiales de la matrice \mathbf{B} sont alors actualisées à partir de cet ensemble d'ébauches. Les autres termes de la matrice proviennent des corrélations spatiales croisées entre les variables traduisant les équilibres atmosphériques (Chabot *et al.*, 2017).

Afin de contrôler l'incertitude de prévision et d'analyse, cette actualisation de la matrice de covariances d'erreurs de l'ébauche est ensuite utilisée lors de la prochaine étape d'assimilation de données. On parle alors de cycle d'assimilation. Les cycles d'assimilation « déterministe » et « ensembliste » évoluent suivant le schéma de la figure 2.13. On notera que l'ensemble d'ébauches du cycle « ensembliste » est représentatif de l'erreur de prévision à court terme, d'autres ensembles sont aussi nécessaires pour estimer l'erreur de prévision à plus long terme (cf annexe B). L'assimilation d'ensemble est numériquement

7. l'innovation est l'écart entre l'observation assimilée et la valeur simulée avec l'opérateur d'observation.

8. On parle également d'inflation de la perturbation.

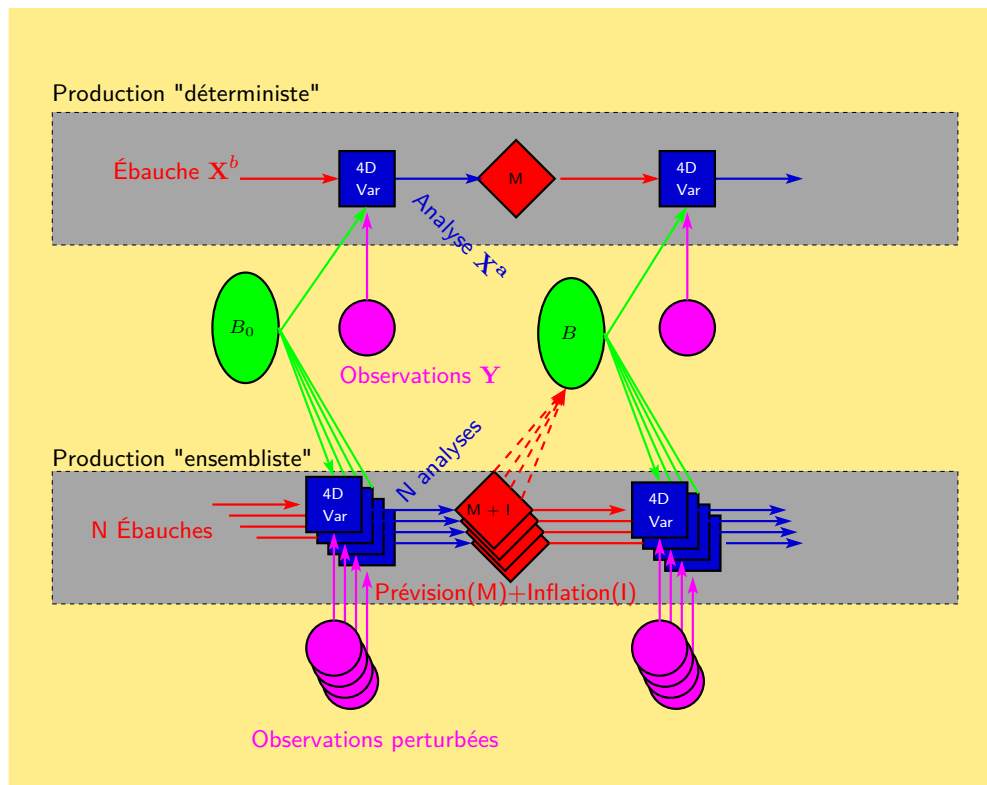


Figure 2.13 : Représentation schématique des liens entre le cycle d'assimilation « déterministe » et celui « ensembliste »

coûteuses. Les ensembles ne contiennent donc que peu de membres et avec une résolution plus faible que celle utilisée dans le cadre de l'assimilation « déterministe ». Les ensembles utilisés dans cette thèse sont décrits dans Berre *et al.* (2015), ils seront dénommés par la suite AEARP pour Assimilation d' Ensemble pour ARPEGE. Les ébauches et les analyses d'ensemble AEARP sont constituées de 25 estimations de l'état de l'atmosphère ou membres. Ces 25 membres⁹ de l'ensemble sont calculés sur une grille spectrale uniforme de troncature T_L479 durant l'étape de prévision, et une troncature T_L149 est utilisée durant la minimisation 4D-Var. On notera que ces deux grilles sont à des résolutions identiques à celles adoptées pour les étapes de minimisation de l'assimilation 4D-Var incrémentale déterministe pour faciliter l'utilisation de la matrice B issue de l'AEARP dans ce système. La grille verticale est la même qu'ARPEGE, soit 105 niveaux hybrides compris entre la pression de surface et l'isobare 0.1 hPa (≈ 65 km d'altitude pour les latitudes tempérées). Le cœur dynamique du modèle et les paramétrisations sont communs aux cycles déterministe et ensembliste, et les analyses sont disponibles toutes les 6h (00, 06, 12, 18 UTC).

9. Ces caractéristiques sont celles disponibles au moment de la thèse. Aujourd'hui, l'assimilation d'ensemble au CEPMMT et à Météo-France utilise des ensembles de 50 membres et avec des résolutions plus élevées.

Chapitre 3

Modélisation de la propagation des infrasons

Sommaire

3.1	Description de la propagation acoustique	37
3.1.1	Généralités sur la modélisation de la propagation	37
3.1.2	Mise en équation	38
3.2	Méthode numérique de résolution	40
3.2.1	Équation d'onde et approche modale	40
3.2.2	Acoustique géométrique et tracé de rayons	45
3.2.3	Conclusion	48

Dans ce chapitre, nous nous intéressons à la modélisation de la propagation des infrasons dans le milieu aléatoire qu'est l'atmosphère, et dont la distribution des hétérogénéités est à dominante verticale. L'objectif est d'identifier les forces et faiblesses des méthodes disponibles ainsi que leurs complémentarités afin de fournir une information qui puisse être comparée aux enregistrements du système de surveillance international (SSI). La première partie de ce chapitre est consacrée à la modélisation de la propagation des ondes acoustiques dans l'atmosphère, des grands principes aux équations qu'il faut résoudre. Dans la seconde partie, les deux méthodes numériques approchées de résolution utilisées dans le cadre de cette thèse sont détaillées.

3.1 Description de la propagation acoustique

3.1.1 Généralités sur la modélisation de la propagation

Dans l'atmosphère, les hétérogénéités de grande échelle imposées par l'évolution verticale de la température et du vent sont à l'origine du phénomène de réfraction des ondes acoustiques. Ce phénomène se traduit par la formation de guides d'ondes qui piègent l'onde dans des canaux de propagation. Ces canaux sont désignés par la couche dans laquelle les ondes sont réfractées. L'existence simultanée de ces guides d'ondes entraîne une division de front en phases, générant alors différentes arrivées au sol. Pour les sources au sol, les phases mises en jeu ont été activement étudiées et classifiées. Elles sont définies et discutées en détail dans la publication de Kulichkov (2010) et la thèse de Gainville (2008).

Le vent peut également transporter les ondes par advection. Selon les guides d'ondes où elles se propagent, les ondes sont plus ou moins déviées par le vent perpendiculaire à

leurs propagations comme un bateau qui dériverait. Les ondes atteignant alors le récepteur arrivent avec des directions différentes que celle de leur source.

Les hétérogénéités de l'atmosphère ne font pas que transporter et réfracter les ondes, elles les amplifient également. L'amplitude de l'onde évolue avec l'altitude de façon proportionnelle à l'inverse de la racine carrée de la masse volumique ρ . Ainsi, si l'onde se propage vers le haut, l'onde est amplifiée. Si elle se propage vers le sol, elle est atténuée. C'est pourquoi, pour s'affranchir de ces effets d'amplification, Bergmann (1946) propose de normaliser la perturbation en pression par $1/\sqrt{\rho_0}$. De plus, lors de leur propagation, ces phases sont soumises aux phénomènes d'absorption et de dispersion, qui mettent en jeu la viscosité de l'air, sa conductivité et la relaxation de l'énergie interne de rotation et de vibration des molécules (Sutherland & Bass, 2004). De manière classique, la modélisation de la propagation des ondes acoustiques part des équations conservatives du mouvement, bien que certains phénomènes liés à l'énergie de vibration moléculaire doivent être traités dans un cadre spécifique (Pierce, 1965 ; Lighthill, 1978 ; De Groot & Mazur, 1963). Dans le cadre de cette thèse, ces effets seront négligés en raison des faibles énergies considérées.

3.1.2 Mise en équation

L'onde acoustique est une perturbation de l'écoulement atmosphérique qui se propage au sein de celui-ci. On suppose alors que (i) la viscosité est négligeable, (ii) il n'y a ni source de chaleur, ni force de volume dans le domaine d'intérêt et (iii) le gaz est parfait et de conductivité thermique nulle. Sous ces hypothèses, les équations du mouvement fluide (dites équations d'Euler) régissant l'écoulement peuvent s'écrire sous la forme d'un système d'équations aux dérivées partielles (non linéaire). Les équations d'Euler font intervenir les champs météorologiques que sont la pression p , la masse volumique ρ , le vent \mathbf{u} , et la température T . L'onde ne rétroagit pas sur l'écoulement, et donc les champs météorologiques peuvent être considérés comme la superposition de la perturbation infinitésimale qu'est l'onde acoustique, à un écoulement de référence ambiant et non perturbé. On a alors la décomposition suivante :

$$\begin{aligned} p &= p_0 + p' \\ \rho &= \rho_0 + \rho' \\ \mathbf{u} &= \mathbf{u}_0 + \mathbf{u}', \end{aligned} \tag{3.1}$$

où les champs p_0 , ρ_0 et \mathbf{u}_0 sont les champs caractéristiques de l'écoulement ambiant non perturbé et permanent. Substituant les développements (3.1) dans les équations d'Euler, et ne retenant que les termes de premier ordre, le système d'équations linéarisées en (\mathbf{u}', p', ρ') qu'il faut résoudre s'écrit :

$$\begin{cases} \frac{\partial \rho'}{\partial t} + \mathbf{u}_0 \cdot \nabla \rho' + \mathbf{u}' \cdot \nabla \rho_0 + \rho' \nabla \cdot \mathbf{u}_0 + \rho_0 \nabla \cdot \mathbf{u}' = 0 \\ \frac{\partial \mathbf{u}'}{\partial t} + (\mathbf{u}' \cdot \nabla) \mathbf{u}_0 + (\mathbf{u}_0 \cdot \nabla) \mathbf{u}' + \frac{1}{\rho_0} \nabla(p') = 0, \end{cases} \tag{3.2}$$

où t désigne le temps et où les variables en gras sont des vecteurs. Dans les équations (3.2) et pour le reste de cette thèse, les opérateurs linéaires ∇ et $\nabla \cdot$ représentent respectivement les opérateurs gradient et divergence. Les équations d'Euler linéarisées (3.2) sont insuffisantes pour caractériser complètement la propagation des ondes. Si toutes deux font intervenir le champ de vitesse, la première fait intervenir le champ de masse

volumique et la seconde le champ de pression. Le processus de propagation étant supposé adiabatique, on en déduit :

$$p' = \rho' \left. \frac{\partial p}{\partial \rho} \right|_{S, \rho_0} \approx \frac{\rho'}{\rho_0 \kappa_S}. \quad (3.3)$$

La relation (3.3) fait intervenir le coefficient de compressibilité isentropique du milieu κ_S défini par l'équation :

$$\kappa_S = \frac{1}{\rho_0} \left. \frac{\partial \rho_0}{\partial p_0} \right|_S. \quad (3.4)$$

On cherche à simplifier la relation de fermeture (3.3). Comme l'air est un gaz parfait, on a le système suivant :

$$\begin{cases} p_0 \rho_0^{-\gamma} = \text{constante} \\ p_0 = \rho_0 \frac{R}{M} T_0. \end{cases} \quad (3.5)$$

Du système (3.5), il est alors possible de déduire que $\kappa_S = 1/(\gamma p_0)$. On introduit alors la vitesse du son c_0 , définie par $c_0 = \sqrt{1/(\kappa_S \rho_0)}$. En utilisant la définition simplifiée de κ_S , ainsi que la seconde équation du système (3.5) dans la définition de c_0 , on obtient :

$$c_0 = \sqrt{\gamma \frac{R}{M} T_0}, \quad (3.6)$$

La vitesse du son dans l'air est donc proportionnelle à la racine carrée de la température comme indiqué au chapitre 1. Pour les latitudes tempérées et au niveau du sol, la vitesse du son est de l'ordre de 350 m.s^{-1} (Des détails sur le calcul de la vitesse du son à partir des données météorologiques sont donnés en annexe A). Dans une atmosphère sans vent, les infrasons se propagent à cette vitesse c_0 . En introduisant c_0 dans l'équation (3.3), la relation devient $p' = c_0^2 \rho'$, ce qui permet de fermer le système d'équations régissant la propagation. Généralement, on transforme les dépendances en ρ' des équations (3.2) en dépendances en p' avec cette relation de fermeture pour obtenir le système linéaire suivant :

$$\begin{cases} \frac{Dp'}{Dt} + \mathbf{u}' \cdot \nabla p_0 + \gamma p' \nabla \cdot \mathbf{u}_0 + \rho_0 c_0^2 \nabla \cdot \mathbf{u}' = 0 \\ \frac{D\mathbf{u}'}{Dt} + (\mathbf{u}' \cdot \nabla) \mathbf{u}_0 + \frac{1}{\rho_0} \nabla(p') - \frac{p' \nabla p_0}{\gamma p_0 \rho_0} = 0 \\ p' = c_0^2 \rho', \end{cases} \quad (3.7)$$

où l'opérateur D/Dt est la dérivée particulière. Le système (3.7) peut être résolu numériquement (Sabatini *et al.*, 2016, 2019), mais au prix d'un coût calcul important. Or, dans le cadre opérationnel lié au TICE, on cherche à modéliser la propagation des infrasons rapidement sur des milliers de kilomètres. C'est pourquoi dans cette thèse, comme dans le cadre opérationnel, nous utiliserons des méthodes approchées et rapides pour modéliser la propagation des infrasons.

3.2 Méthode numérique de résolution

3.2.1 Équation d'onde et approche modale

Pour étudier la propagation dans une direction (azimut) donnée, on peut utiliser une approche modale pour estimer rapidement le champ de pression acoustique. Dans le cadre de cette approche, on suppose que les grandeurs en régime permanent ne dépendent que de z . On fait également l'hypothèse que la composante verticale du vent est nulle. La dérivée particulaire s'écrit alors :

$$\frac{D}{Dt} = \left[\frac{\partial}{\partial t} + \mathbf{u}_{0,H} \cdot \nabla_H \right], \quad (3.8)$$

où ∇_H est l'opérateur gradient dans le plan horizontal et $\mathbf{u}_{0,H}$ sont les composantes horizontales du vent en régime permanent. Il est démontré dans l'ouvrage de référence d' Ostashev *et al.* (2005) que l'évolution de la perturbation en pression dans un milieu stratifié peut être décrite par :

$$\frac{D}{Dt} \left[\frac{1}{c_0^2} \frac{D^2 p'}{Dt^2} - \nabla_H^2 p' - \rho_0 \frac{\partial}{\partial z} \left(\frac{1}{\rho_0} \frac{\partial p'}{\partial z} \right) \right] + 2 \left(\frac{d\mathbf{u}_{0,H}}{dz} \cdot \nabla_H \right) \frac{\partial p'}{\partial z} = 0, \quad (3.9)$$

où ∇_H^2 désigne le laplacien. On considère que $p'(x, y, z, t)$ peut s'écrire sous la forme d'un mode normal (une onde plane) :

$$p' \propto \hat{p}(z) \exp(i(\mathbf{k}_H \cdot \mathbf{x}_H - \omega t)), \quad (3.10)$$

où \mathbf{k}_H et \mathbf{x}_H sont respectivement les vecteurs d'onde et position dans le plan horizontal $z = cst$, ω étant la pulsation de l'onde. La grandeur \hat{p} est liée à l'amplitude de p' par la transformée de Fourier :

$$\hat{p}(k_H, \omega, z) = \frac{1}{2\pi} \int \left\{ \int p'(x, z, t) e^{-ik_H x} dx \right\} e^{-i\omega t} dt, \quad (3.11)$$

et la transformée inverse

$$p'(x, z, t) = \frac{1}{2\pi} \int \left\{ \int \hat{p}(k_H, z, \omega) e^{ik_H x} dk_H \right\} e^{i\omega t} d\omega, \quad (3.12)$$

Dans la suite, nous traiterons simultanément les problèmes en coordonnées cartésiennes et cylindriques (cela revient pratiquement au même car nous regardons uniquement dans une direction donnée). La différence de traitement n'apparaîtra finalement qu'au moment d'inverser la transformée $\hat{\psi}$ définie plus loin.

Pour plus de clarté, on supposera que l'atmosphère est homogène dans la direction horizontale et on se placera dans le plan vertical de propagation, bien que cette hypothèse ne soit pas indispensable (Bertin, 2014). Afin de simplifier encore le modèle, on considère que le gradient vertical de $\mathbf{u}_{0,H}$ peut être négligé, permettant de supprimer le terme en $\frac{d\mathbf{u}_{0,H}}{dz}$ de l'équation (3.9). On introduit alors la décomposition en onde plane (3.10) dans l'équation (3.9) et on l'exprime en fonction de la variable $\hat{\psi} = \hat{p}/(\Omega\sqrt{(\rho_0)})$ avec la notation $\Omega = \omega - \mathbf{k}_H \cdot \mathbf{u}_{0,H}$.

$$\left[\frac{d^2}{dz^2} + \frac{\Omega^2}{c_0^2} + \frac{Na^2}{c_0^2} - k_H^2 \right] \hat{\psi} = 0, \quad (3.13)$$

où Na désigne la fréquence de coupure acoustique caractéristique du milieu et définie par la relation :

$$Na = c_0 \sqrt{\frac{3}{4} \left(\frac{\nabla \rho_0}{\rho_0} \right)^2 - \frac{\nabla^2 \rho_0}{2\rho_0}}. \quad (3.14)$$

Les ondes acoustiques de fréquence inférieure ou égale à Na sont dispersées et non propagées par le milieu¹. Le terme lié à la fréquence de coupure Na dans l'équation (3.13) est négligeable pour les infrasons au-delà de 0,05 Hz (Assink *et al.*, 2017), et sera négligé par la suite car nous nous intéressons à des fréquences comprises entre 0,2 Hz et 20 Hz. Il est avantageux de simplifier l'équation (3.13) en éliminant la dépendance en k_H présente dans Ω^2/c_0^2 car elle permet de se ramener à un problème linéaire aux valeurs propres, où k_H est la valeur propre. Pour ce faire, on introduit le nombre de Mach $\mathcal{M} = \frac{\mathbf{k}_H}{c_0 \|\mathbf{k}_H\|} \cdot \mathbf{u}_{0,H}$ de l'écoulement. On suppose que le vecteur d'onde \mathbf{k} fait un angle faible avec l'horizontale et que le nombre de Mach est faible. On en déduit que $\frac{c_0 \|\mathbf{k}_H\|}{\omega} \approx 1$. Avec ces hypothèses, il est possible d'approximer le terme en $\frac{\Omega^2}{c_0^2}$ (Godin, 2002) de la manière suivante :

$$\frac{\Omega^2}{c_0^2} = \frac{\omega^2}{c_0^2} \left(1 - \frac{\mathbf{k}_H \cdot \mathbf{u}_{0,H}}{\omega} \right)^2 \approx \frac{\omega^2}{c_0^2} (1 - \mathcal{M})^2 \approx \frac{\omega^2}{c_{eff}^2}, \quad (3.15)$$

où $c_{eff} = c_0 + \mathbf{k}_H \cdot \mathbf{u}_{0,H}$ représente la célérité effective du son dans le milieu. En introduisant cette approximation dans l'équation (3.13), on obtient alors l'équation de Helmholtz :

$$\left\{ \frac{d^2}{dz^2} + \frac{\omega^2}{c_{eff}^2} - k_H^2 \right\} \hat{\psi} = \hat{\mathcal{L}}\hat{\psi} = 0, \quad (3.16)$$

où $\hat{\mathcal{L}}$ est l'opérateur de propagation dans l'espace spectral. Ce problème (3.16) est défini sur un domaine semi-infini, limité par la surface terrestre en $z = 0$ et la raréfaction de l'atmosphère en $z \rightarrow \infty$. Nous cherchons ici à obtenir un champ qui reste borné pour les plus hautes altitudes et donc non-exponentiellement croissant en $z \rightarrow \infty$. Il est alors supposé que la célérité effective du son tend vers une valeur fixée notée c_∞ quand $z \rightarrow \infty$. Les fonctions $\hat{\psi}$, solutions de l'équation (3.16), tendent alors vers des $\hat{\psi}_\infty$ telles que :

$$\hat{\psi}_\infty \propto \exp(\pm \alpha z) \text{ avec } \alpha^2 = k_H^2 - \omega^2/c_\infty^2. \quad (3.17)$$

Seuls les modes amortis sont physiquement acceptables, cela impose alors l'inégalité suivante :

$$-\pi \leq \arg \left(k_H^2 - \frac{\omega^2}{c_\infty^2} \right) < \pi. \quad (3.18)$$

Concernant le comportement au niveau de la surface, il est possible de considérer une impédance acoustique pour le sol (Attenborough, 1988). L'effet du sol peut alors être modélisé par la condition de Robin (Waxler, 2002). Cette impédance complexe, notée généralement Z , peut être modélisée par la condition de Robin (3.19) proposée dans Waxler (2002) :

$$\frac{d\hat{\psi}}{dz}(0) = -C(\omega)\hat{\psi}(0), \quad (3.19)$$

où le coefficient complexe $C(\omega)$ est lié à Z par la relation :

$$C(\omega) = \frac{i\omega\rho'(0)}{Z(\omega)}. \quad (3.20)$$

Dans la mesure où les phases étudiées dans le cadre de cette thèse n'interagissent que très peu avec la surface, on peut faire l'hypothèse d'une impédance du sol infinie. Cette

1. Une interprétation plus poussée de la relation de dispersion et de l'influence de Na est disponible au paragraphe 2.3.1 de la thèse de Hanique-Cockenpot (2011).

hypothèse revient à considérer une condition de Neumann homogène au niveau de la surface (i.e. $\hat{\psi}(z=0) = 0$).

En considérant une source s , de transformée de Fourier \hat{s} , l'équation (3.16) devient alors :

$$\left\{ \frac{d^2}{dz^2} + \frac{\omega^2}{c_{eff}^2} - k_H^2 \right\} \hat{\psi} = \hat{s}(\omega, k_H, z). \quad (3.21)$$

Le problème dont on cherche une solution peut s'écrire sous la forme condensée $\hat{\mathcal{L}}\hat{\psi} = \hat{s}$. On cherche à calculer le champ acoustique généré par un dirac δ placé au point de coordonnées (x, z_0) ($z_0 > 0$) à l'instant $t = 0$. La solution $\hat{\psi}$ dans l'espace spectral de l'équation doit satisfaire :

$$\hat{\mathcal{L}}\hat{\psi} = \frac{\partial^2 \hat{\psi}}{\partial z^2} + \left(\frac{\omega^2}{c_{eff}^2} - k_H^2 \right) \hat{\psi} = -2\Upsilon(k_H)\delta(z_0). \quad (3.22)$$

Cette expression reste vraie en coordonnées cylindriques, sous réserve de considérer la transformée de Hankel. La présence de la fonction de Heaviside Υ au second membre de (3.22) est liée à la condition de causalité, qui impose de réduire l'étude aux seuls nombres k_H d'onde positifs. Les formules de dérivation (au sens des distributions) de fonctions discontinues donnent :

$$\frac{\partial \hat{\psi}}{\partial z} = \left\{ \frac{\partial \hat{\psi}}{\partial z} \right\} + \sigma_0 \delta(z_0), \quad (3.23a)$$

$$\frac{\partial^2 \hat{\psi}}{\partial z^2} = \left\{ \frac{\partial^2 \hat{\psi}}{\partial z^2} \right\} + \sigma_0 \left\{ \frac{\partial \delta}{\partial z} \right\}(z_0) + \sigma_1 \delta(z_0), \quad (3.23b)$$

où les fonctions $\left\{ \frac{\partial \hat{\psi}}{\partial z} \right\}$ et $\left\{ \frac{\partial^2 \hat{\psi}}{\partial z^2} \right\}$ désignent les distributions représentées par les fonctions égales aux dérivées usuelles pour $z < z_0$ et $z > z_0$ et non définies pour $z = z_0$, et où σ_j est le saut de la dérivée j -ième pour $z = z_0$, de sorte que $\hat{\psi}$ vérifie l'équation :

$$\left\{ \frac{\partial^2 \hat{\psi}}{\partial z^2} \right\} + \sigma_0 \left\{ \frac{\partial \delta}{\partial z} \right\}(z_0) + \sigma_1 \delta(z_0) + \left(\frac{\omega^2}{c_{eff}^2} - k_H^2 \right) \hat{\psi} = -2\Upsilon(k_H)\delta(z_0), \quad (3.24)$$

ce qui conduit à

$$\left\{ \frac{\partial^2 \hat{\psi}}{\partial z^2} \right\} + \left(\frac{\omega^2}{c_{eff}^2} - k_H^2 \right) \hat{\psi} = 0, \quad (3.25)$$

à prendre au sens des fonctions, et aux relations de saut :

$$\sigma_0 = 0 \quad \text{et} \quad \sigma_1 = -2\Upsilon(k_H). \quad (3.26)$$

Ces conditions de saut forment un système linéaire pour les constantes d'intégration. Désignons par $\hat{\psi}_-$ et $\hat{\psi}_+$ des solutions de l'équation $\hat{\mathcal{L}}\hat{\psi} = 0$, respectivement pour $z < z_0$ et $z > z_0$, satisfaisant à des conditions restrictives convenables aux points $z = 0$ et $z = \infty$. En notant $z_0^+ = z_0 + \epsilon$ et $z_0^- = z_0 - \epsilon$, les conditions de saut peuvent encore s'écrire sous la forme d'un système linéaire (pour les constantes d'intégration qu'on note A et B) :

$$\sigma_1 = A\hat{\psi}'_+(z_0^+) - B\hat{\psi}'_-(z_0^-) = -2\Upsilon(k_H), \quad (3.27a)$$

$$\sigma_0 = A\hat{\psi}_+(z_0^+) - B\hat{\psi}_-(z_0^-) = 0, \quad (3.27b)$$

dont la solution n'est unique que si le déterminant, ici donné par le Wronskien :

$$W(z_0^\pm) = \hat{\psi}'_+(z_0^+) \hat{\psi}_-(z_0^-) - \hat{\psi}_+(z_0^+) \hat{\psi}'_-(z_0^-), \quad (3.28)$$

est non nul et où la notation f' définit une dérivation selon z au sens des distributions. En passant à la limite $\epsilon \rightarrow 0$, on arrive à :

$$A = -\frac{2\Upsilon \hat{\psi}_-(z_0)}{W(z_0)} \quad \text{et} \quad B = -\frac{2\Upsilon \hat{\psi}_+(z_0)}{W(z_0)}, \quad (3.29)$$

d'où, en rétablissant la dépendance en k_H et ω :

$$\hat{\psi}(z, k_H, \omega) = \begin{cases} -2\Upsilon(k_H) \frac{\hat{\psi}_-(z_0, k_H, \omega) \hat{\psi}_+(z, k_H, \omega)}{W(z_0, k_H, \omega)} & (z > z_0) \\ -2\Upsilon(k_H) \frac{\hat{\psi}_-(z, k_H, \omega) \hat{\psi}_+(z_0, k_H, \omega)}{W(z_0, k_H, \omega)} & (z < z_0). \end{cases} \quad (3.30)$$

Des propriétés mathématiques de l'opérateur $\hat{\mathcal{L}}$, on déduit que l'équation $\hat{\mathcal{L}}\hat{\psi} = 0$ a une solution non triviale, qui s'identifie au vecteur propre de $\hat{\mathcal{L}} + k^2 I$, où I désigne l'identité. Dans ce cas, les solutions $\hat{\psi}_-$ et $\hat{\psi}_+$ étant proportionnelles, il n'est plus nécessaire de conserver deux notations différentes. Dans la suite on notera $\hat{\psi}_j$ la fonction propre associée à la valeur propre k_j (on supprime l'indice H pour plus de clarté). Il reste à effectuer la transformée inverse de l'expression (3.30). Cela revient à calculer l'intégrale de (3.30) d'abord par rapport à la variable k , puis par rapport à ω . S'agissant de l'intégrale sur k , et se plaçant en coordonnées cylindriques, l'application du théorème des résidus permet d'arriver à l'expression :

$$\tilde{G}(r, z, \omega) = -i \sum_{j \in J} \frac{\hat{\psi}_j(z, \omega, k_j) \hat{\psi}_j(z_0, \omega, k_j) k_j(\omega) H_0^{(1)}(k_j r)}{\partial_k W(k_j, z_0, \omega)}, \quad (3.31)$$

où J est l'ensemble fini des indices j tels que les k_j soient solutions de :

$$W(k_j, z_0, \omega) = 0, \quad (3.32)$$

qui est l'expression la plus générale de la relation de dispersion. Autrement dit, pour une pulsation ω fixée, les nombres d'onde k_j^2 solutions de (3.32) sont les valeurs propres de l'opérateur $\hat{\mathcal{L}} + k^2 I$ et les $\hat{\psi}_j$ sont les fonctions propres associées. La seconde intégrale (i.e. sur ω) est classiquement estimée via des algorithmes de type FFT (Fast Fourier Transform). Il faut alors estimer numériquement les paires propres $(k_j, \hat{\psi}_j)$ pour un ensemble de fréquences ω_j afin de reconstruire les fonctions de Green G et obtenir le champ de pression acoustique. L'opérateur linéaire de l'équation (3.16) doit être discretisé pour pouvoir résoudre numériquement le problème aux valeurs propres $\hat{\mathcal{L}} + k^2 I$.

Une première méthode est d'estimer l'opérateur en différences finies. On la notera par la suite FD (pour *Finite Differences*). Elle consiste à définir une grille discrète en $(z_k)_{k=1, \dots, N_d}$ de maille uniforme h et à utiliser un schéma aux différences finies pour estimer les dérivées successives intervenant dans $\hat{\mathcal{L}}$. Les dérivées successives aux points de grille $(z_k)_k$ peuvent être estimées à partir d'un schéma centré d'ordre deux. Les dérivées sont alors données par les relations :

$$\begin{cases} \frac{d\hat{\psi}}{dz}(z_k) = \frac{\hat{\psi}(z_{k+1}) - \hat{\psi}(z_{k-1})}{2h}, \\ \frac{d^2\hat{\psi}}{dz^2}(z_k) = \frac{\hat{\psi}(z_{k+1}) - 2\hat{\psi}(z_k) + \hat{\psi}(z_{k-1}))}{h^2}. \end{cases} \quad (3.33)$$

En introduisant les relations (3.33) dans l'équation (3.16), on obtient un opérateur discret.

Une autre méthode pour discretiser l'opérateur est de définir une grille de points de collocation et de choisir une base de polynômes d'interpolation pour représenter les fonctions $\hat{\psi}_j$ (Canuto *et al.*, 2006). On parle alors de méthode par collocation spectrale, que l'on note par la suite SC (pour *Spectral Collocation*). Dans le cadre de cette thèse, on utilise une grille de Gauss-Lobatto-Tchebychev définie par $(s_k)_{k=1,\dots,N_d} = \cos(k\pi/N_d)$ et reliée aux points z_k par le changement de variable :

$$z_j = \frac{1 - s_j}{2} z_{max}, \quad (3.34)$$

où z_{max} est le toit du domaine, déterminé de manière à ce que l'amplitude des ondes soit suffisamment petite. On décompose les fonctions propres $\hat{\psi}$ sur une base de polynômes de Tchebychev $T_k(s) = \cos(k \arccos(s))$. On a alors :

$$\begin{cases} \psi(s) = \sum_k a_k T_k(s), \text{ où} \\ T_0(s) = 0 \\ T_1(s) = s \\ T_{k+1}(s) = 2sT_k(s) - T_{k-1}(s) \end{cases} \quad (3.35)$$

Les coefficients a_k sont obtenus par application de la propriété d'orthogonalité des polynômes de Tchebychev. La dérivée se décompose également sur la base de polynômes précédente. En suivant les développements dans Gottlieb *et al.* (1984), les opérateurs de dérivation exprimés aux points de collocation $\frac{d^k}{ds^k}(s_j)$ se traduisent par l'application k fois de l'opérateur D dont la matrice est remplie ainsi :

$$[D_{jl}] \begin{cases} \frac{c_j(-1)^{j+l}}{c_l(s_j - s_l)}, \text{ pour } j \neq l, \\ -\frac{2(1 - s_l^2)}{s_l}, \text{ pour } 1 \leq j = l \leq N_d - 1, \\ \frac{2N_d^2 + 1}{6}, \text{ pour } j = l = 0, \\ -\frac{2N_d^2 + 1}{6}, \text{ pour } j = l = N_d. \end{cases} \quad (3.36)$$

Notons que D est définie sur l'intervalle $[-1, 1]$ et que pour revenir à la variable primitive (i.e z), il faut appliquer la jacobienne du changement de variable (3.34) (i.e. $-\frac{2}{z_{max}}$), où I est la matrice identité de taille N_d^2). Connaissant alors la matrice de dérivation, on peut construire la matrice dont on cherche les valeurs propres qui, contrairement à celle obtenue avec le schéma aux différences finies (3.33), est représentée par une matrice pleine. La figure (extraite des travaux de thèse Goupy A.) illustre les résultats obtenus à partir des deux méthodes de discrétisation (i.e. différences finies et collocation spectrale). Pour un nombre de points en z donné, la méthode SC permet d'obtenir des résultats plus précis, en raison des propriétés de convergence spectrale (Canuto *et al.*, 2006). Les calculs seront par la suite effectués à l'aide de l'opérateur discret obtenu par collocation spectrale.

Il est intéressant de noter qu'il existe différentes méthodes pour calculer les valeurs propres d'une matrice. Si un petit nombre de modes domine, alors il est possible de se limiter à leur calcul, en utilisant des méthodes itératives (Bertin *et al.*, 2014), ce qui permet une réduction significative du coût de calcul.

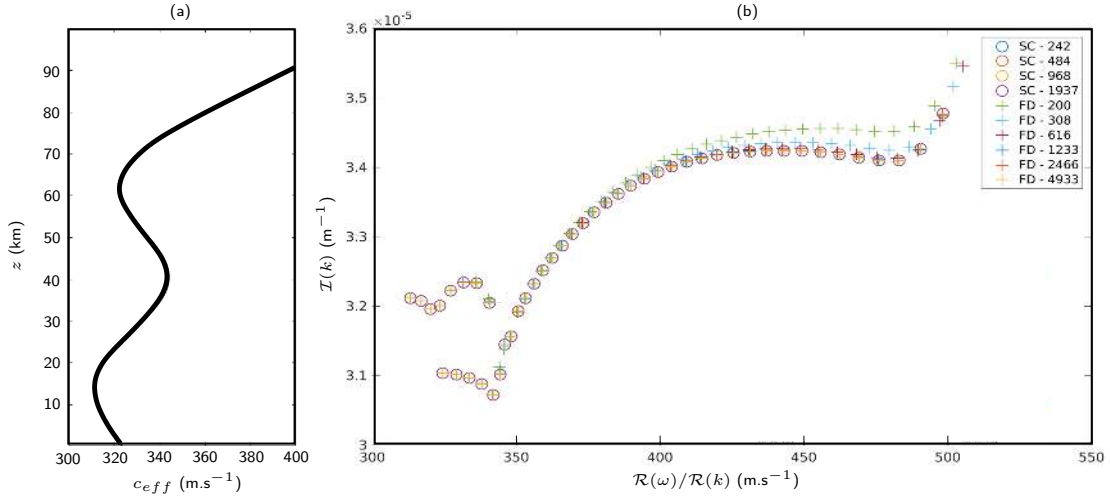


Figure 3.1 : Profils de c_{eff} (a) et spectre de l'opérateur $\hat{\mathcal{L}}$ pour une fréquence de 0.1 Hz (b) obtenu à partir des méthodes de discrétisation : par collocation spectrale (SC) et par différences finies (FD). Le nombre de points N_d de grille ou de collocation est donné en légende. $\mathcal{R}(x)$ et $\mathcal{I}(x)$ sont respectivement la partie réelle et la partie imaginaire de x . Cette figure est adaptée du manuscrit de thèse de Goupy A. .

3.2.2 Acoustique géométrique et tracé de rayons

Un autre moyen de résoudre le système (3.7) est d'utiliser l'acoustique géométrique. Dans ce cadre, on considère que les propriétés de l'atmosphère varient très peu sur une longueur d'onde λ , ce qui permet de définir un petit paramètre $\epsilon = 1/k$, où k désigne le nombre d'onde. Concrètement, si la fréquence de l'onde est de 0,1 Hz, cette méthode ne permet pas de simuler l'effet des hétérogénéités dont la taille serait inférieure à quelques dizaines de kilomètres. Mais pour les fréquences de l'ordre du Hertz, et si on omet l'effet des petites structures (turbulence, etc.), cette technique présente un grand intérêt : d'une part, parce qu'elle est très rapide, et, d'autre part, parce qu'elle propose une représentation particulièrement intuitive et simple du mécanisme de propagation (exemple du chapitre 1). Cette méthode peut facilement être implémentée et plusieurs codes *open-source* sont disponibles (e.g., plateforme *ncpaprop*², code *GeoAc*³). On considère l'onde infrasonore comme décomposable en ondes planes harmoniques. Les perturbations p' , ρ' et \mathbf{u}' peuvent alors s'écrire sous la forme :

$$(p', \rho', \mathbf{u}') = (\hat{p}, \hat{\rho}, \hat{\mathbf{u}}) \exp(ik\Theta), \quad (3.37)$$

où Θ est la phase et où les grandeurs $(\hat{p}, \hat{\rho}, \hat{\mathbf{u}})$ correspondent aux amplitudes des perturbations acoustiques. Cette hypothèse est valide lorsque l'amplitude de l'onde varie peu sur une longueur d'onde et si la courbure du front d'onde est très supérieure à la longueur d'onde. En substituant (3.37) dans le système (3.7) tout en négligeant les

2. lien GitHub *ncpa-prop* : <https://github.com/chetzer-ncpa/ncpaprop/blob/master/docs/ncpaprop-manual.pdf>

3. lien GitHub *GeoAc* : <https://github.com/LANL-Seismoacoustics/GeoAc>

termes proportionnels à $\nabla \cdot \mathbf{u}_0$ et ∇P_0 , on obtient le système suivant :

$$\begin{cases} ik \left(\rho_0 c_0^2 \hat{\mathbf{u}} \cdot \nabla \Theta + \hat{p} \frac{D\Theta}{Dt} \right) = -\frac{D\hat{p}}{Dt} - \rho_0 c_0^2 \nabla \cdot \hat{\mathbf{u}}, \\ ik \left(\hat{\mathbf{u}} \cdot \frac{D\Theta}{Dt} + \frac{\hat{p} \nabla \Theta}{\rho_0} \right) = -\frac{D\hat{\mathbf{u}}}{Dt} - (\hat{\mathbf{u}} \cdot \nabla) \mathbf{u}_0 - \frac{\nabla \hat{p}}{\rho_0}, \\ \hat{p} = c_0^2 \hat{\rho}. \end{cases} \quad (3.38)$$

Adoptant alors la notation $\epsilon = 1/ik$ proposée par Lalande (2012), les quantités \hat{p} , $\hat{\rho}$ et $\hat{\mathbf{u}}$ peuvent s'écrire sous la forme des développements asymptotiques suivants.

$$\begin{cases} \hat{p} = p_1 + p_2 \epsilon^1 + p_3 \epsilon^2 + \dots, \\ \hat{\rho} = \rho_1 + \rho_2 \epsilon^1 + \rho_3 \epsilon^2 + \dots, \\ \hat{\mathbf{u}} = \mathbf{u}_1 + \mathbf{u}_2 \epsilon^1 + \mathbf{u}_3 \epsilon^2 + \dots \end{cases} \quad (3.39)$$

En substituant les développements asymptotiques (3.39) dans le système (3.38), à l'ordre zéro, on obtient le système linéaire d'équations :

$$\begin{cases} p_1 \frac{D\Theta}{Dt} + \rho_0 c_0^2 \mathbf{u}_1 \cdot \nabla \Theta = 0, \\ \mathbf{u}_1 \frac{D\Theta}{Dt} + \frac{p_1 \nabla \Theta}{\rho_0} = 0. \end{cases} \quad (3.40)$$

et, à l'ordre en ϵ , on obtient le système homogène suivant :

$$\begin{cases} \rho_0 c_0^2 \mathbf{u}_2 \cdot \nabla \Theta + p_2 \frac{D\Theta}{Dt} = -\frac{Dp_1}{Dt} - \rho_0 c_0^2 \nabla \cdot \mathbf{u}_1, \\ \mathbf{u}_2 \frac{D\Theta}{Dt} + \frac{p_2 \nabla \Theta}{\rho_0} = -\frac{D\mathbf{u}_1}{Dt} - (\mathbf{u}_1 \cdot \nabla) \mathbf{u}_0 - \frac{\nabla p_1}{\rho_0}, \\ p_1 = c_0^2 \rho_1. \end{cases} \quad (3.41)$$

dont les solutions non triviales en $(\hat{p}_1, \hat{\mathbf{u}}_1)$ sont données par l'annulation du déterminant du système linéaire (3.40). On obtient alors une équation appelée *équation eikonale* :

$$(\nabla \Theta)^2 = \frac{1}{c_0^2} \left(\frac{D\Theta}{Dt} \right)^2. \quad (3.42)$$

Cette équation, à elle seule, caractérise la propagation des ondes en tout point \mathbf{x} du milieu. Si l'état atmosphérique ne dépend pas du temps, alors la phase peut s'écrire en fonction de la position \mathbf{x} et du temps de propagation de l'onde, ou temps de trajet $T_t(\mathbf{x})$, ce qui se traduit par les changements suivants dans l'équation eikonale :

$$(\nabla T_t)^2 = \frac{1}{c_0^2(\mathbf{x})} (1 - \mathbf{u}_0(\mathbf{x}) \cdot \nabla T_t)^2, \quad (3.43)$$

qui est l'expression obtenue par Blokhintzev (1946). Il est important de noter que la résolution de l'équation (3.43) est ici indépendante de la fréquence du signal.

La résolution de l'équation eikonale peut se faire via la méthode des caractéristiques. Comme Θ représente la phase en acoustique géométrique, les courbes $\Theta(\mathbf{x}, t) = \text{constante}$ sont les fronts d'onde. Ainsi, par analogie avec les ondes planes, la pulsation locale ω et le vecteur d'onde local \mathbf{k} sont définis par :

$$\begin{cases} \omega(\mathbf{x}, t) = -\frac{\partial \Theta}{\partial t}, \\ \mathbf{k}(\mathbf{x}, t) = \nabla \Theta. \end{cases} \quad (3.44)$$

Notons que, par définition, ces grandeurs locales vérifient la relation :

$$\frac{\partial \mathbf{k}}{\partial t} + \nabla \omega = 0. \quad (3.45)$$

L'équation eikonale (3.42) possède deux racines correspondant à deux modes de propagation acoustique qui s'effectuent dans deux directions opposées :

$$\omega = \mathbf{k} \cdot (c_0 \mathbf{n} + \mathbf{u}_0), \quad (3.46)$$

où $\mathbf{n}(\mathbf{x}, t) = \mathbf{k}/\|\mathbf{k}\|$ est le vecteur unitaire, normal au front d'onde. La relation de dispersion fait apparaître la vitesse de groupe de l'onde $\mathbf{c}_g = c_0 \mathbf{n} + \mathbf{u}_0$. Les lignes caractéristiques de cette équation sont encore appelées *rayons acoustiques*. On introduit la dérivée advectée le long de la caractéristique $\frac{D_g}{Dt} = \frac{\partial}{\partial t} + \mathbf{c}_g \cdot \nabla$. Les équations des rayons sont alors explicitées sous la forme :

$$\begin{cases} \frac{D_g \mathbf{x}}{Dt} = \mathbf{c}_g, \\ \frac{D_g \mathbf{k}}{Dt} = -(\mathbf{k} \cdot \mathbf{n}) \cdot \nabla c_0 - \nabla \mathbf{u}_0 \cdot \mathbf{k}, \\ \frac{D_g \omega}{Dt} = (\mathbf{k} \cdot \mathbf{n}) \frac{\partial c_0}{\partial t} + \mathbf{k} \cdot \frac{\partial \mathbf{u}_0}{\partial t}. \end{cases} \quad (3.47)$$

La relation de dispersion (3.46) étant toujours satisfaite, il n'est pas nécessaire d'intégrer l'équation en ω . L'évolution temporelle du milieu est prise en compte dans l'intégration du système (3.41). La trajectoire du rayon est alors donnée en intégrant numériquement ces équations à l'aide d'une méthode comme celle de Runge & Kutta standard du 4ème ordre à pas variable. La trajectoire prise par les rayons ne dépend que de l'état initial, c'est-à-dire de la position de la source et la direction initiale de la propagation décrite par l'angle d'incidence et l'azimut (Gainville, 2008). On peut obtenir le long des rayons le temps de trajet, son incidence par rapport au plan horizontal ainsi que son azimut de propagation. Pour la suite, on considérera que la réflexion en surface est totale et spéculaire. La réfraction des ondes acoustiques par le milieu est alors modélisée par la modification des trajectoires des rayons en altitude. Par réfraction, les rayons se focalisent, donnant naissance à ce que l'on appelle des *caustiques*⁴, ainsi qu'à des *zones d'ombre*. Les zones d'ombre sont les régions de l'espace où aucun rayon ne parvient, et, a priori, où le champ acoustique serait inexistant. En réalité, l'onde acoustique est diffractée par l'atmosphère. Pour obtenir une estimation de la pression acoustique transportée par les rayons, il faut alors intégrer les équations de transport (3.41).

4. une caustique est une singularité où la pression acoustique donnée par l'acoustique géométrique est infinie.

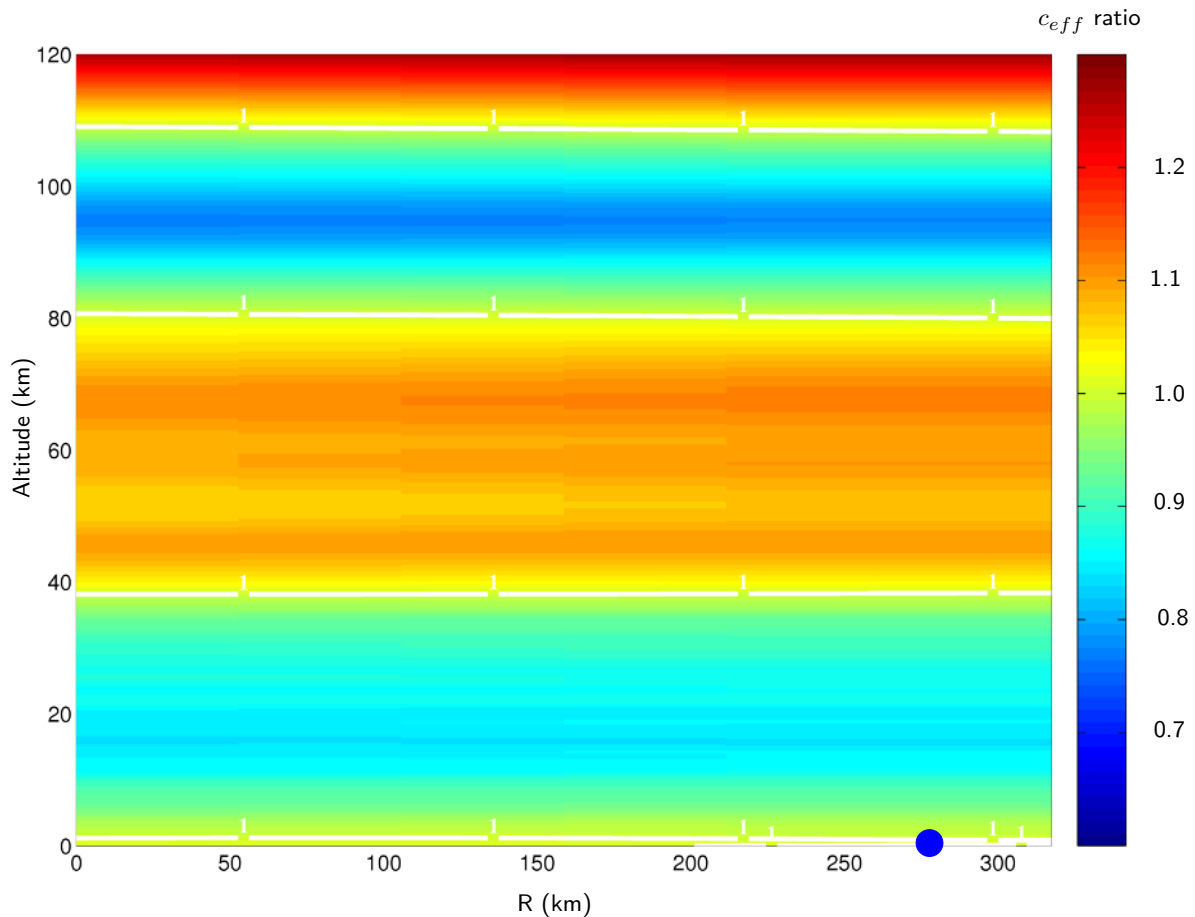


Figure 3.2 : Rapport de célérité effective dans le plan de propagation. R est la distance à la source et le rond bleu symbolise la station réceptrice. Conditions météorologiques déduites de l'analyse du CEPMMT du 6/11/2011 à minuit entre Rivesaltes et l'Observatoire de Haute Provence.

3.2.3 Conclusion

Pour visualiser les apports complémentaires de ces modèles, on se propose de traiter un exemple de propagation. Supposons que l'on souhaite vérifier qu'une source connue ait conduit à des arrivées au niveau d'une station réceptrice 268 km plus loin dans l'azimut $59,6^\circ\text{N}$. Si l'atmosphère est sans vent transverse (perpendiculaires à l'azimut $59,6^\circ\text{N}$), la station réceptrice enregistre des arrivées de back-azimut théorique égale à $241,7^\circ\text{N}$. Les conditions météorologiques lors de la propagation sont présentées sur la figure 3.2 par l'intermédiaire du c_{eff} ratio qui est défini comme étant le rapport entre la célérité effective en altitude et celle au sol. Lorsque ce rapport est supérieur à 1, les ondes peuvent être réfractées à partir de cette altitude. Ces régions sont indiquées par les lignes blanches sur la figure 3.2. Au vu de ces conditions de propagation, la réfraction est possible dans trois couches de l'atmosphère (en-dessous 1 km, entre 40 km et 80 km, au-dessus de 110 km d'altitude).

À partir de l'approche modale, on est en mesure d'estimer le champ de pression acoustique pour diverses fréquences et dans diverses directions et notamment celle dans l'azimut de la station. La figure 3.3 présente le logarithme du rapport entre la pression acoustique en champ lointain et celle émise à la source pour deux fréquences particulières 0,2 et 0,8 Hz. Plus ce rapport est faible et plus les ondes à cette fréquence ont été atténuées. On remarque que la station réceptrice est dans une zone où le champ de pression acoustique n'est que peu atténué signifiant qu'elle a dû enregistrer du signal. Si

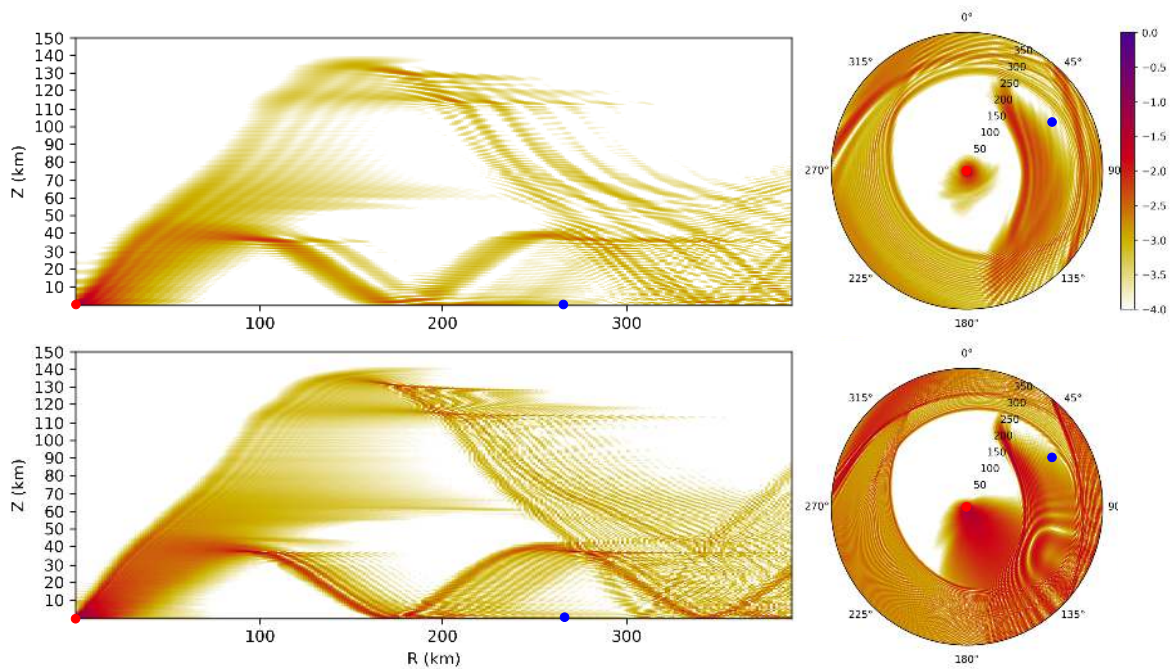


Figure 3.3 : Champs de pertes par transmission pour la fréquence 0.2 Hz (en haut) et la fréquence 0.8 Hz (en bas) dans le plan de propagation (colonne de gauche) et au sol (colonne de droite). La station réceptrice est représentée par un rond bleu et la source par le rond rouge.

l'on examine les coupes verticales, nous constatons que ce signal provient de réfractions au niveau de la stratosphère, les autres réfractions impactant la surface soit en amont de la station, soit en aval. La méthode modale permet d'identifier rapidement les zones recevant au sol du signal, de définir le temps de trajet des ondes ainsi que l'incidence du front d'onde au moment de son passage en station. Cependant, cette méthode ne considère pas le vent transverse à la propagation et les possibles effets sur celle-ci. Il manque ainsi une information sur les azimuts des signaux en station. À partir d'un code de tracé de rayons, on estime par le biais de leurs trajectoires, les altitudes de réfractions et des zones où l'on attend du signal. Dans ces zones, dites d'*arrivées géométriques*, les rayons fournissent une estimation précise du temps de trajet et de l'incidence que fait le front d'onde avec la station ainsi que de son azimut car ce modèle considère les effets de déviations induits par le vent transverse. Cependant, pour obtenir le champ de pression il faut intégrer en plus les équations de transport (3.41) sur les rayons, et, comme la diffraction est négligée, il existera des zones où la pression acoustique sera mal-estimée. Dans le cas de notre exemple, il s'avère que la station examinée sera en zone d'ombre et donc que le tracé de rayons ne pourra pas donner une estimation de la pression acoustique. Nous pouvons affirmer cela, car cet exemple s'inspire du cas réel « Rivesaltes » qui sera traité en détails dans le chapitre suivant.

En plus des deux méthodes qui viennent d'être décrites et qui seront utilisées par la suite, il existe d'autres approches comme celles basées sur l'approximation parabolique (Kulichkov *et al.*, 2017) ayant elle aussi ses propres limitations.

Dans le cadre de cette thèse, nous avons utilisé des codes de tracé de rayons linéaires afin d'obtenir une information sur la trajectoire du front d'onde en trois dimensions et un code utilisant la décomposition modale pour le calcul du champ de pression acoustique et des formes d'onde des signaux. Les codes de tracé de rayons sont ceux développés par P. Blom (Blom & Waxler, 2012 ; Blom, 2019) dans la plateforme *open source* GeoAc. Le code de décomposition modale est le code Fast Low Order Wave Simulation (FLOWS) (Millet, 2015), développé au CEA.

Méthode	Tracé de rayons linéaire	Tracé de rayons + transport	Décomposition modale sans réduction	Décomposition modale avec réduction	Equations paraboliques petit et grand angle
Milieu 2D (invariant selon x_H) ou 3D	Oui	Oui	\approx	Non	Oui
Propagation 3D	Oui	Oui	Non	Non	Non
Forme d'onde	Non	Oui	Oui	Oui	Oui
Diffraction	Non	Non	Oui	Oui	Oui
Champs nécessaires	$c(\mathbf{x}, t, z), \mathbf{u}(\mathbf{x}, t, z)$	$c(\mathbf{x}, t, z), \mathbf{u}(\mathbf{x}, t, z),$ $p(\mathbf{x}, t, z), \rho(\mathbf{x}, t, z)$	$c(\mathbf{x}_H, t, z), \mathbf{u}_H(\mathbf{x}_H, t, z),$ $p(\mathbf{x}_H, t, z), \rho(\mathbf{x}_H, t, z)$	$c(\mathbf{x}_H, t, z), \mathbf{u}_H(\mathbf{x}_H, t, z),$ $p(\mathbf{x}_H, t, z), \rho(\mathbf{x}_H, t, z)$	$c(\mathbf{x}_H, t, z), \mathbf{u}_H(\mathbf{x}_H, t, z),$ $p(\mathbf{x}_H, t, z), \rho(\mathbf{x}_H, t, z)$

Tableau 3.1 : Comparaison entre les codes de propagation acoustique basés sur le tracé de rayons et ceux utilisant la méthode de décomposition modale.

Chapitre 4

Études de sensibilité en vue de l'inversion des infrasons

Sommaire

4.1 Rivesaltes	52
4.1.1 Cas d'étude	52
4.1.2 Les résultats du code de tracé de rayons	54
4.1.3 Les résultats du modèle par décomposition modale	55
4.1.4 Influence d'une modification des profils atmosphériques sur les quantités d'intérêt et définition de la stratégie	57
4.1.5 Stratégie en vue de l'assimilation des signaux infrasons	59
4.2 Définition d'un cadre idéalisé	60
4.2.1 Atmosphère idéalisée à deux degrés de liberté	60
4.2.2 Extraction des quantités d'intérêt simulées	61
4.3 Problème inverse dans un cadre idéalisé	65
4.3.1 Optimisation locale	68
4.3.2 Expression de J sous différentes hypothèses	69
4.3.3 Résultats du problème inverse	72
4.4 Conclusions	76

Dans ce chapitre, on cherche à développer un opérateur d'observation \mathbf{H} pour pouvoir assimiler les informations sur l'atmosphère fournies par des mesures infrasons. Comme les modèles de PNT ne simulent pas la propagation des infrasons, elle doit être décrite au moyen d'un modèle spécifique, pouvant être considéré comme un opérateur d'observations (de manière similaire à un code de transfert radiatif pour assimiler les luminances satellitaires). Dans une première partie, on se propose d'étudier l'information mise à disposition par les modèles de propagation usuels au travers d'un exemple réaliste et de visualiser l'influence qu'une modification de l'atmosphère peut avoir sur les quantités d'intérêt des signaux simulés que l'on peut comparer à celles des détections des stations infrasons du SSI. La question qui se pose est de savoir si, par simple comparaison de certaines de ces quantités, il est possible par méthode inverse de retrouver une information pertinente sur les conditions atmosphériques. Cela fait l'objet de la deuxième partie de ce chapitre.

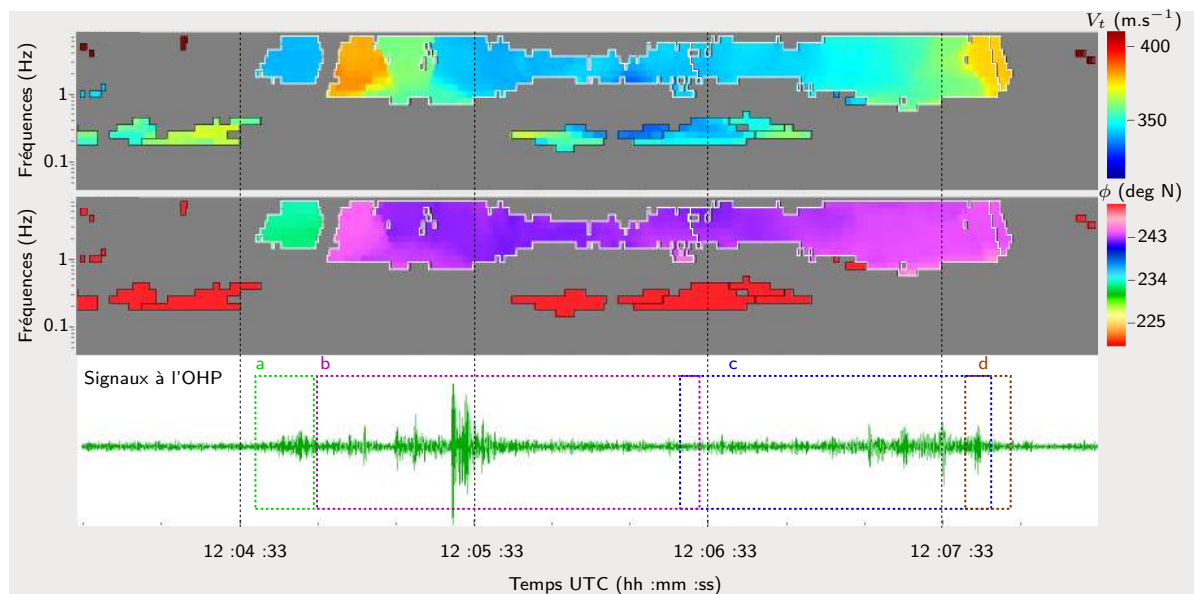


Figure 4.1 : Résultats d'analyse avec l'algorithme PMCC pour les signaux infrasons enregistrés à l'OHP le 06/11/2015 entre 1150 UTC et 1215 UTC. (haut) Vitesse de passage du signal cohérent dans l'espace temps-fréquence. (milieu) Back-azimut du signal cohérent arrivant dans l'espace temps-fréquence. (bas) Signal cohérent moyen.

4.1 Comparaison des quantités d'intérêt des détections infrasons aux sorties des modèles de propagation : application aux détections d'une source anthropique par l'OHP.

Dans cette section, on s'intéresse à l'information disponible en sortie des modèles de propagation des infrasons par décomposition modale de l'équation d'onde ou par tracés de rayons au travers d'un exemple de détections en France.

4.1.1 Cas d'étude

Au cours des dernières années, et à raison de plusieurs fois par semaine, un processus industriel de soudage par explosif a été employé par une usine située aux alentours de Rivesaltes (10 km au nord de Perpignan). Ces soudures ont généré des infrasons qui ont été enregistrés par une station située au niveau de l'observatoire de Haute Provence (OHP) lorsque les conditions atmosphériques étaient favorables. Cela a été le cas notamment le 06/11/2015. Les signaux enregistrés à l'OHP sont analysés en utilisant le traitement d'antenne opérationnel du CEA : PMCC¹ (Cansi, 1995). Ce traitement permet d'identifier des signaux cohérents, dans l'espace temps-fréquence, au sein du bruit enregistré entre 1205 UTC et 1210 UTC. Sur la figure 4.1, il est possible de constater que les azimuts de provenance (ou back-azimuts) notés ϕ , d'une partie des signaux pointent vers des directions proches de celle théorique de Rivesaltes ($\phi = 241.7^\circ \text{N}$) et que d'autres vers l'océan Atlantique ($\phi \geq 270^\circ$). Comme le signal de back-azimut pointant vers l'océan Atlantique se situe dans les basses fréquences, il peut être associé à une source microbarome Atlantique. Une fois ce signal basse fréquence filtré, on obtient le signal cohérent de la figure 4.1. Le signal cohérent haute fréquence dont le back-azimut

1. PMCC pour Progressive Multi-Channel Correlation

Détection	Début (hh :mm :ss) UTC	Fin (hh :mm :ss) UTC	Fréq. (Hz)	P_k (Pa)	ϕ (°N)	V_t (m.s ⁻¹)
a	12 :04 :37	12 :04 :53	3.55	0.0163	233.30	341
b	12 :04 :55	12 :06 :30	2.99	0.1079	242.93	352
c	12 :06 :27	12 :07 :46	3.26	0.0718	243.55	354
d	12 :07 :43	12 :07 :49	3.54	0.0333	244.01	378

Tableau 4.1 : Quelques quantités d'intérêt des détections de l'OHP associées à la source Rivesaltes de 1150 UTC. L'amplitude maximale est notée P_k . Les fréquences $Fréq.$, les back-azimuts ϕ et les vitesses de passage V_t des détections sont des valeurs moyennes sur la durée de la détection.

est proche de 242 °N peut être associé aux explosions à Rivesaltes effectuées vers 1150 UTC ce jour-là. Ainsi, les ondes ont parcouru les 268 km qui les séparent de l'OHP en 15-20 min. Si l'arrivée était directe, celle-ci aurait dû arriver environ deux minutes plus tôt si l'on considère que la vitesse du son est de 340 m.s⁻¹. De plus, les fortes vitesses de passage (≥ 370 m.s⁻¹) de ces signaux au travers de la station (cf. figure 4.1) nous indiquent qu'ils arrivent avec une forte incidence et donc proviennent de réfractions dans la moyenne atmosphère (20 à 80 km d'altitude). À partir de critères de proximité décrits dans Brachet *et al.* (2010), ce signal est décomposé en 4 détections successives (a, b, c, d) dont les portions des signaux sont indiquées par les rectangles pointillés sur la figure 4.1. Elles correspondent aux regroupements des pixels délimités par les lignes blanches présentes sur la même figure. Les quantités d'intérêt de ces détections sont données au tableau 4.1. On y retrouve les bornes temporelles des détections qui fournissent une estimation du temps de trajet T_t des ondes², la fréquence moyenne du signal associé à la détection, l'amplitude maximale du signal P_k , le back-azimut de la détection ϕ et sa vitesse de passage³ V_t (c'est la vitesse horizontale qu'ont les ondes cohérentes traversant l'antenne).

Pour les simulations, le milieu de propagation est spécifié de la manière suivante :

- les profils de température, de vent, de pression et d'humidité spécifique sont extraits des analyses IFS du CEPMMT du réseau de 00 UTC sur 90 niveaux pression ;
- la conversion des niveaux hybrides en altitudes est faite en utilisant la méthode opérationnelle employée au CEA (cf Annexe pour plus de détails) ;
- Pour tenir compte de la courbure de la Terre⁴, on interpole bi-linéairement les profils le long du grand cercle liant Rivesaltes à l'OHP ;
- les profils météorologiques sont raccordés par une méthode de type spline cubique entre 80 et 90 km d'altitude aux profils HWM et MSIS (paragraphe 2.2.1) ;
- les profils raccordés sont ré-interpolés par des splines sur une grille uniforme en altitude ;
- l'impédance de la surface n'est pas considérée, la réflexion des ondes est donc totale.

Pour simplifier l'étude, nous faisons l'hypothèse commune que la moyenne atmosphère varie suffisamment peu le long du grand cercle et donc que le milieu peut être supposé homogène horizontalement. Les profils caractéristiques de notre milieu sont ceux extraits aux coordonnées de la source et sont représentés sur la figure 4.2 (a). On effectue alors des simulations à l'aide du code de tracé de rayons linéaires en coordonnées sphériques de

2. On peut également parler de temps de vol.

3. On peut également parler de vitesse apparente.

4. La courbure de la Terre modifie l'azimut de projection du vent utilisé pour le calcul de la c_{eff} .

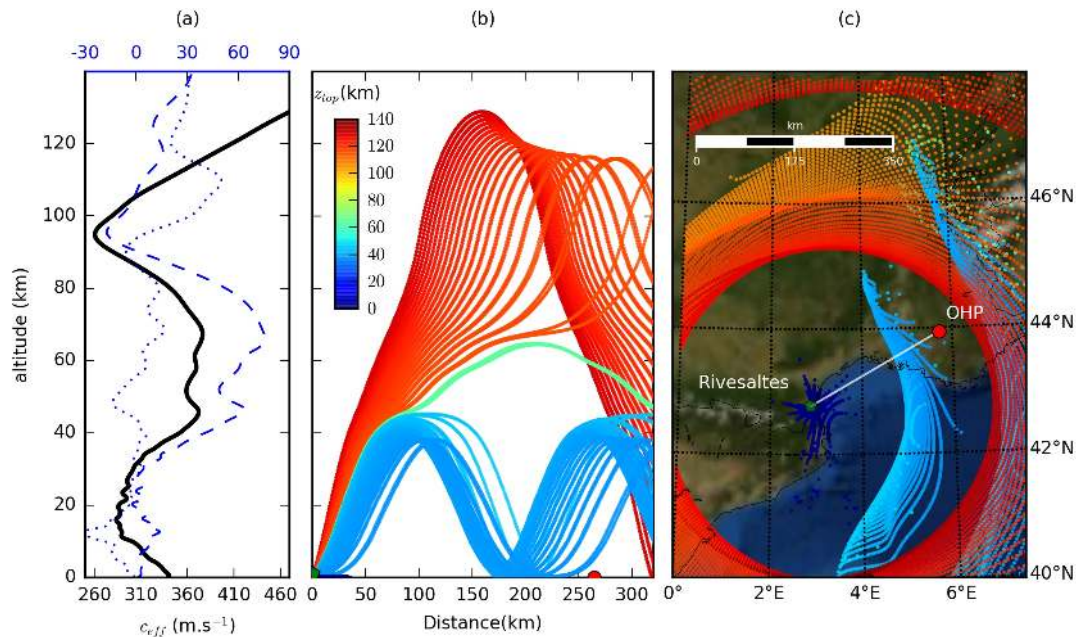


Figure 4.2 : Propagation d'infrasons à partir de Rivesaltes simulée par un modèle de tracé de rayons. (a) Profil de célérité effective dans la direction de l'OHP (axe horizontal inférieur) et de vent zonal (tiret bleu) et de vent méridien (pointillé bleu) (axe horizontal supérieur). (b) Trajectoires des rayons depuis Rivesaltes (pentagone vert) en direction de l'OHP (rond rouge). Les rayons sont colorés en fonction de leurs altitudes de réfraction z_{top} . (c) empreinte au sol donnée par les rayons.

GeoAc et d'un code utilisant la décomposition modale (FLOWS) pour les formes d'onde.

4.1.2 Les résultats du code de tracé de rayons

On se place aux coordonnées de la source supposée (Rivesaltes) et on calcule les rayons pour des azimuts variant entre 0 et 359 °N par pas de 1 degré et des incidences comprises entre 0 et 45 ° par pas de 1 degré. On obtient alors les trajectoires et l'empreinte au sol présentées sur la figure 4.2 (b-c). À partir de cette figure 4.2 (b), il est alors possible d'identifier les altitudes où ont lieu les réfractions z_{top} et le nombre de réflexions au sol qu'ont dû faire les ondes avant d'atteindre la station. On remarque alors quatre régions de réfractions distinctes ou guides d'ondes. Un premier guide d'ondes troposphérique à environ 2 km d'altitude (bleu foncé sur la figure 4.2), un guide stratosphérique vers 40 km d'altitude (cyan), un guide mésosphérique vers 60 km d'altitude (vert) et un dernier guide thermosphérique (rouge).

Au total, près de 16 200 rayons sont lancés⁵ En fait, seules les quantités d'intérêt des rayons qui parviennent à la station importent. C'est pourquoi, généralement, on ne s'intéresse qu'aux rayons qui relient la source à la station. Ils sont généralement appelés *rayons propres*. Comme le montre le schéma de la figure 4.3, il est alors possible de définir un back-azimut, un temps de trajet et une vitesse de passage des ondes.

Ces rayons propres peuvent être déterminés à l'aide d'un algorithme d'optimisation locale (Blom & Waxler, 2017), sous réserve de spécifier la zone où les rayons parviennent. Nous définissons ici la station comme un disque au sol, et seuls les rayons impactant ce disque sont qualifiés de propres. L'utilisation d'un algorithme d'optimisation locale pour

5. Ici le nombre de rayons tracés reste raisonnable. Dans certains cas, il est nécessaire de calculer plusieurs centaines de milliers de rayons, ce qui se traduit par un besoin significatif d'espace mémoire pour les sauvegarder et un coût calcul important.

Vitesse du son : $c = \sqrt{\gamma r T}$, où $\gamma = 1.4$ et $r = 287 \text{ J.kg}^{-1}.\text{K}^{-1}$.

Vitesse de passage : $V_t = \frac{c(z=0)}{\cos(\theta)}$

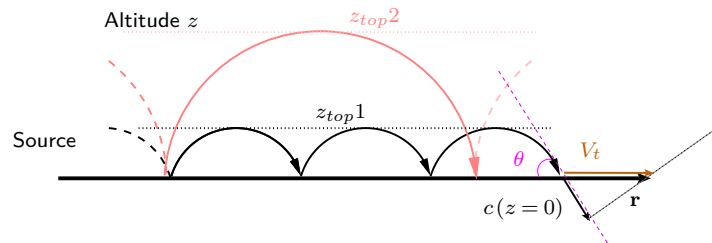


Figure 4.3 : Schéma illustratif de la relation entre l'incidence du rayon et la vitesse de passage.

trouver les rayons propres pose problème dans la mesure où celui-ci convergera vers un rayon représentatif d'une unique *phase* (arrivée cohérente en station issue d'une réfraction dans une région spécifique de l'atmosphère). Il est donc nécessaire de lancer l'algorithme d'optimisation pour un nombre suffisant de conditions initiales afin de s'assurer de ne pas omettre de phases. Il est donc nécessaire quoi qu'il arrive de calculer de nombreux rayons. C'est pourquoi dans un souci de simplicité, nous n'utilisons pas ces méthodes pour trouver les rayons propres mais lançons un grand nombre de rayons échantillonnant l'espace des possibles et sélectionnons les rayons propres qui atteignent la station. Ces rayons propres donnent alors un ensemble de quantités d'intérêt simulées, que l'on peut ensuite comparer à celles observées au tableau 4.1.

On remarque sur les figures 4.2 (c) qu'aucun des rayons calculés n'atteint le voisinage de la station OHP et donc que la comparaison aux observations est impossible. Si Rivesaltes est bien à l'origine de ces détections, il reste alors deux raisons plausibles à ce désaccord entre les résultats de simulation et les observations :

1. les détections résultent de la diffraction des ondes ;
2. les conditions météorologiques réelles sont significativement différentes de celles extraites des analyses IFS du CEPMMT.

Nous vérifions alors que ces phases ne résultent pas d'un phénomène de diffraction des ondes en utilisant un modèle de propagation capable de le représenter.

4.1.3 Les résultats du modèle par décomposition modale

Nous calculons les modes acoustiques pour une propagation dans les conditions atmosphériques décrites précédemment dans la direction de l'OHP depuis Rivesaltes pour une fréquence de 0,8 Hz. Une fois les modes calculés, on reconstruit la carte de pertes par transmission.

Les pertes par transmission

En reprenant les notations introduites dans le chapitre 3, les pertes par transmission sont définies par la formule suivante pour une fréquence donnée :

$$TL = 20 \log_{10} \left(\frac{|p'(r, z)|}{|p'(0, z_s)|} \right) \quad (4.1)$$

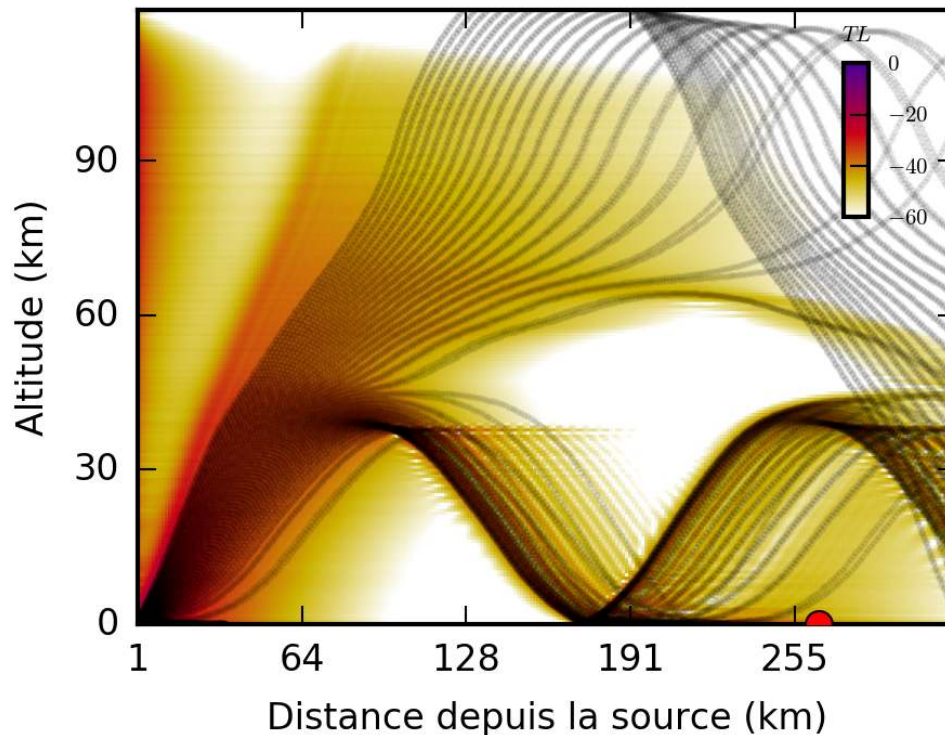


Figure 4.4 : Carte des pertes par transmission (dB) pour une fréquence de 0.8 Hz dans le cas d'une propagation d'ondes infrasonores entre Rivesaltes et l'OHP (rond rouge) pour le 06/11/2015 00 UTC. En gris, les trajectoires de la figure 4.2.

où $p'(r, z)$ représente le champ de pression acoustique en altitude z , $p'(0, z_s)$ le champ de pression acoustique au niveau de la source et r représente la distance horizontale par rapport à la source. En suivant la définition (4.1), une perte par transmission positive signifie que le signal est amplifié, alors que, plus la perte par transmission est négative, plus le signal est atténué. Par exemple, une perte par transmission de -60 dB implique une division d'un facteur 1000 du signal. Les cartes produites par FLOWS, comme celles de la figure 4.4, informent sur les régions où l'énergie acoustique, associée à une fréquence donnée, n'est que peu ou fortement atténuée par l'atmosphère. En comparant la figure 4.4 aux trajectoires des rayons de la figure 4.2, les zones de faibles pertes par transmission en valeur absolue correspondent aux zones de focalisation des rayons, c'est-à-dire aux *caustiques*⁶. On remarque que les pertes par transmission en valeur absolue restent faibles dans les zones d'ombre géométriques et notamment à l'OHP si bien qu'en définitive, et contrairement à ce que laissent penser les simulations par tracé de rayons, cette station est bien en mesure de détecter le signal émis par Rivesaltes grâce à la diffraction.

Le signal n'est pas mono-fréquentiel. Il faut donc calculer les formes d'ondes pour pouvoir les comparer aux observations. Les formes d'ondes sont obtenues par convolution entre les fonctions de Green (vues dans le chapitre précédent) associées au milieu de propagation et un modèle de source. Au lieu d'examiner les formes d'ondes dans toutes leurs complexités, on se limitera à l'enveloppe du signal au sol à l'aide de la norme de la transformée de Hilbert des formes d'ondes. Cela nous renseigne sur l'amplitude de la pression acoustique maximale perçue au sol comme le montre la figure 4.5. La simulation est faite pour des fréquences inférieures à 0,8 Hz car la majeure partie de l'énergie est dans les basses fréquences. Ainsi, pour avoir une estimation précise de

6. une caustique est une singularité où la pression acoustique donnée par l'acoustique géométrique est infinie.

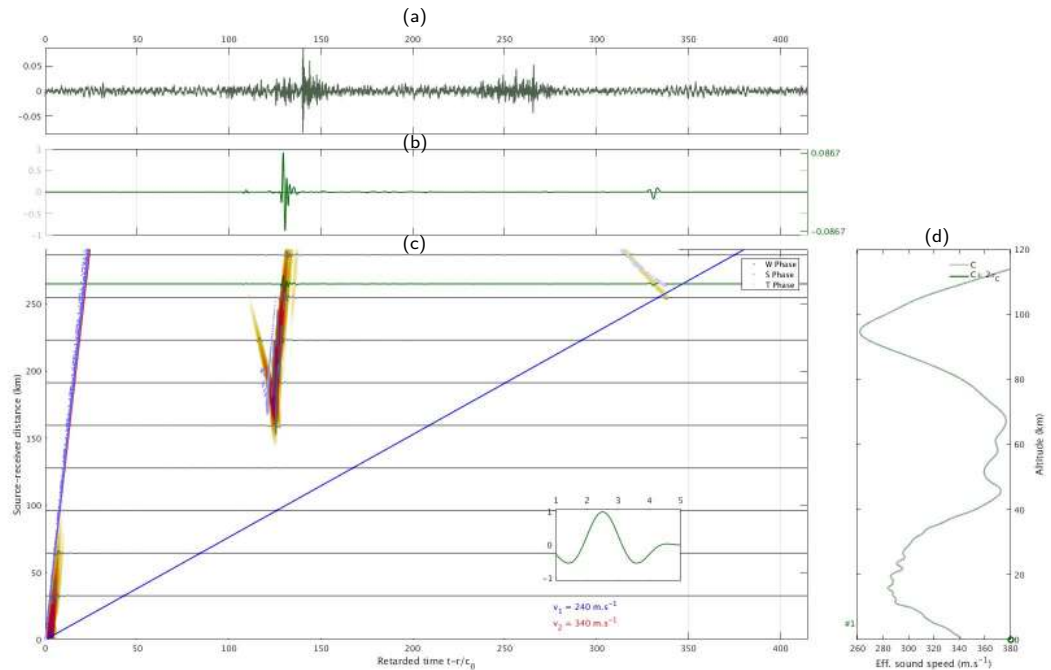


Figure 4.5 : Résultats de la modélisation par FLOWS de la propagation depuis Rivesaltes vers l'OHP pour le 06/11/2015. (a) Signaux cohérents moyens enregistrés par l'OHP dans la gamme de fréquence 0.5-7 Hz. (b) Signaux simulés à l'OHP dans une gamme de fréquence entre 0-0.8 Hz pour une source de fréquence caractéristique 0.4 Hz. (c) Distributions temporelle et spatiale de l'enveloppe du signal en fonction de la distance en km par rapport à la source r et du temps retardé ($t-r/c_0$, avec $c_0 = 0,35 \text{ km.s}^{-1}$). Les points sont les point d'impacts identifiés par un code de tracé de rayons. La couleur des points est fonction de l'altitude de réfraction (bleu pour les phases troposphériques W , violet pour les phases stratosphériques S et gris pour les phases thermosphériques T). (d) Profil de célérité effective du son utilisé.

l'amplitude, il n'est pas nécessaire de considérer les hautes fréquences. Ce choix permet également de réduire les coûts de calcul. Si l'on compare l'estimation obtenue aux observations, toutes deux représentées sur la figure 4.5, on remarque que seulement deux phases sont prédites au niveau de l'OHP. La forme en V de ces phases dans ce type de représentation est typique (Waxler *et al.*, 2015). La « branche » avancée du V est ce que l'on appelle une *arrivée rapide* et la seconde branche est l'*arrivée lente*. Ces deux arrivées proviennent de réfraction dans la même région de l'atmosphère. On cherche à identifier cette région en comparant les solutions proposées par le tracé de rayons et la transformée de Hilbert du signal simulé par FLOWS. On s'aperçoit que les phases simulées sont issues d'une réfraction dans la stratosphère. On remarque également qu'il n'y a pas de phase thermosphérique, ni de phase troposphérique se propageant jusque là. Ces arrivées simulées sont en accord, d'un point de vue temporel, avec les deux premières détections, mais n'expliquent pas les deux autres détections. Cela indique, soit l'existence d'une autre source, soit des erreurs dans la spécification du milieu de propagation. Dans la suite, on se propose d'étudier la seconde hypothèse.

4.1.4 Influence d'une modification des profils atmosphériques sur les quantités d'intérêt et définition de la stratégie

Les modèles météorologiques ne résolvent pas explicitement toutes les échelles spatiales et temporelles de la dynamique de l'atmosphère. On se propose ainsi de perturber la

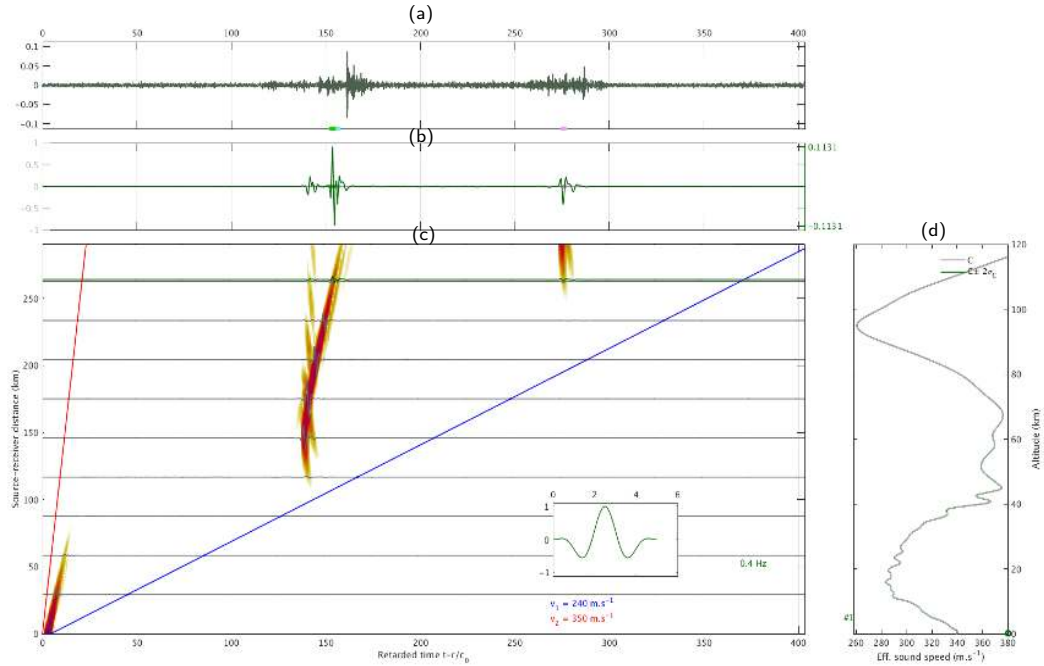


Figure 4.6 : Résultats de la modélisation de la propagation depuis Rivesaltes vers l'OHP dans le cas d'un membre spécifique de l'ensemble issu de la perturbation de l'état atmosphérique par une structure de grande échelle.

région autour de 40 km d'où proviennent les deux premières arrivées. Nous perturbons le profil de célérité effective $c_{eff,0}$ initial par une perturbation aléatoire ad hoc μ_{ceff} construite de la manière suivante :

$$\begin{cases} c_{eff}(z) = c_{eff,0}(z) (1 + \mu_{ceff}), \\ \mu_{ceff} = 0.2 \exp\left(-\frac{(z-40)^2}{2\sigma_z^2}\right) \cos(k_L z) \cos(k_p z) \\ k_L = \frac{2\pi}{3+2\mathcal{U}(0,1)}, \text{ (Grande échelle)} \\ k_p = \frac{2\pi}{0.5+\mathcal{U}(0,1)}, \text{ (Petite échelle)} \end{cases} \quad (4.2)$$

où z est l'altitude en kilomètre et $\mathcal{U}(0,1)$ représente une loi uniforme entre 0 et 1. Les tirages des nombres d'onde des grandes échelles k_L et des petites échelles k_p sont faits de façon indépendante. La perturbation de grande échelle permet de modéliser une structure que peut résoudre le modèle, mais qui a été, soit lissée par le manque de niveaux à ces altitudes, soit mal placée. La perturbation de petite échelle décrit les petites structures (ondes de gravité, turbulence) non résolues par le modèle de PNT.

On se propose de n'ajouter que la perturbation de grande échelle dans un premier temps. Pour l'une des 100 réalisations effectuées avec seulement une perturbation de grande échelle, on obtient les résultats présentés sur la figure 4.6. La perturbation de grande échelle, ici, a séparé la phase stratosphérique dite rapide de celle dite lente. Ces deux phases étaient d'ores et déjà simulées, mais cette réalisation a rapproché le second rebond stratosphérique. Ce second rebond concorde maintenant avec les deux autres détections enregistrées à l'OHP. On considère ce membre comme référence et on lui rajoute une perturbation de petite échelle. On obtient un nouvel ensemble d'atmosphères perturbées. Pour un membre particulier, on obtient les signaux reportés sur la figure 4.7. La perturbation de petite échelle tend à étendre la durée des paquets d'ondes simulés. Avec cette spécification du milieu de propagation, il est à présent possible d'expliquer les quatre arrivées observées. Cette perturbation n'est cependant pas physiquement

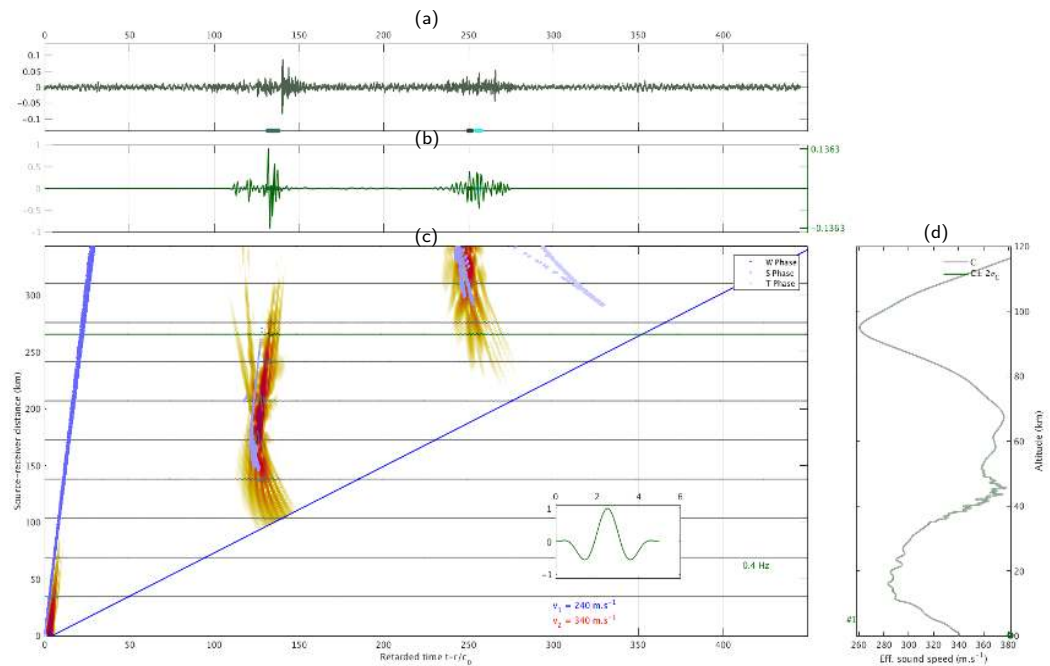


Figure 4.7 : Résultats de la modélisation de la propagation depuis Rivesaltes vers l'OHP dans le cas du membre perturbé de la figure 4.6 auquel l'on a ajouté une perturbation de petite échelle.

satisfaisante, dans le sens où l'état de l'atmosphère ainsi corrigé n'a aucune raison de satisfaire les équations de la météorologie. Le but de cette étude a été de mettre en évidence une sensibilité significative de la modélisation de la propagation des infrasons à deux types d'erreurs dans la spécification du milieu : une erreur de grande échelle du modèle de PNT et une erreur de petite échelle liée à des processus, comme les ondes de gravité, qui ne sont pas explicitement résolus par le modèle.

4.1.5 Stratégie en vue de l'assimilation des signaux infrasons

La comparaison aux quantités d'intérêt des bulletins est complexe, car elle met en jeu différents modèles de propagation complémentaires et sensibles aux changements de spécifications du milieu. De plus, il est nécessaire de connaître la source et son signal. Tout cela rend compliquée la construction d'un opérateur d'observation pour assimiler directement ces quantités d'intérêt des détections infrasons. On se propose alors de résoudre en amont le problème inverse en exploitant la sensibilité des résultats de simulations à une modification du milieu. On ne cherche plus à assimiler directement les quantités d'intérêt, mais à estimer les spécifications du milieu, en termes de profils verticaux de vent et de température, qu'il faudrait avoir pour que les résultats de simulation de la propagation et les détections s'accordent. Ce type d'approche en deux étapes est assez semblable avec ce qui est fait notamment pour l'assimilation des radiances pluvieuses (Duruiseau *et al.*, 2019) ou des réflectivités radar (Wattrelot *et al.*, 2014). Ce problème inverse est non-linéaire. Il est étudié depuis des décennies (Donn & Rind, 1971) et de nombreuses voies ont été explorées. Certaines utilisent le signal temporel complet (Chunchuzov *et al.*, 2015), d'autres utilisent les quantités d'intérêt des détections (Assink *et al.*, 2013) comme les méthodes développées ci-après. Les toits des modèles de PNT n'allant pas

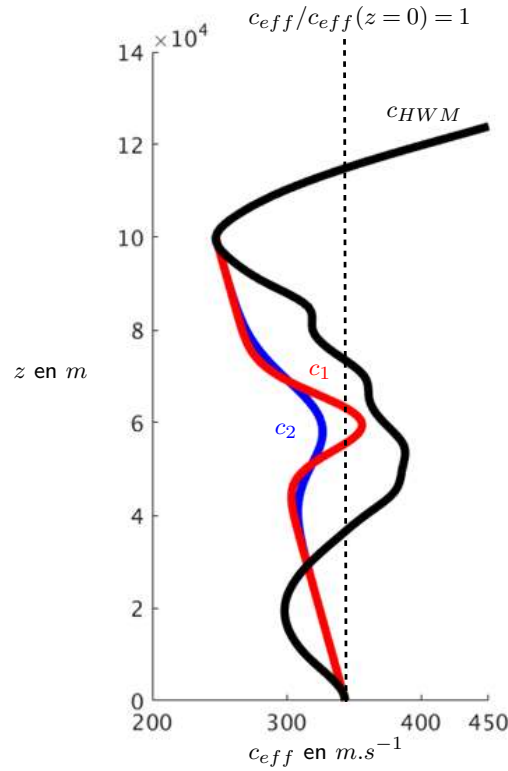


Figure 4.8 : Profils de célérité effective pour plusieurs atmosphères idéalisées (c_2 : jet de 50 m.s^{-1} , c_1 : jet de 75 m.s^{-1} .)

au-delà de la mésosphère, nous filtrerons⁷ les possibles détections issues de réfractions dans la thermosphère.

4.2 Définition d'un cadre idéalisé

4.2.1 Atmosphère idéalisée à deux degrés de liberté

En vue de définir une méthode d'inversion, nous avons choisi une atmosphère idéalisée afin de limiter le nombre de degrés de liberté à estimer. On suppose que l'atmosphère est homogène horizontalement et on se place dans le plan de propagation. Nous faisons l'approximation de célérité effective vue au chapitre 3. Le profil de célérité effective c_{eff} simplifié est défini par :

$$\begin{cases} c_{eff}(z) = c_e(z=0) + \frac{(c_e(z=100) - c_e(z=0))}{100} z + A_J \exp\left(-\frac{(z-60)^2}{2\sigma_J^2}\right) & , \text{ pour } z < 100\text{km} \\ c_{eff}(z) = c_e(z) & , \text{ pour } z \geq 100\text{km}. \end{cases} \quad (4.3)$$

(A_J, σ_J) sont respectivement l'amplitude et l'étalement du jet. Nous fixons arbitrairement le jet à une altitude de 60 km. Par la suite, nous utilisons le code de propagation FLOWS pour les simulations des signaux. Cette atmosphère idéalisée permet d'obtenir une grande diversité de signaux simulés en modifiant faiblement les caractéristiques du jet (cf figure 4.9). On peut obtenir différentes phases (ou arrivées) et visualiser la dissociation d'une paire d'arrivées rapide et lente dans le cas d'un fort jet. Ces deux phénomènes sont

7. Du côté des observations cela se traduit par l'élimination des détections dont les vitesses de passage dépassent un certain seuil.

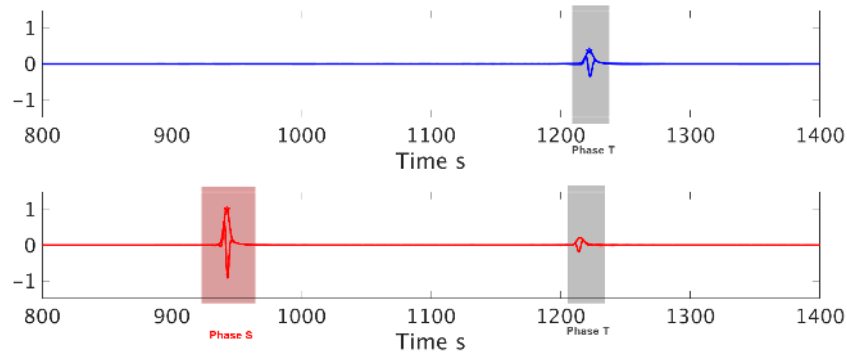


Figure 4.9 : Signaux normalisés obtenus avec les profils de célérité effective c_1 et c_2 à 285 km, présentés sur la figure 4.8. Les signaux sont normalisés par le maximum de leurs enveloppes respectives.

observés dans la réalité (Waxler *et al.*, 2015). Pour comprendre l'impact des méthodes d'inversion que nous allons utiliser, on simule avec le modèle FLOWS les signaux⁸ à 285 km pour un ensemble de 3600 (60x60) profils simplifiés distincts. Ces profils quadrillent uniformément l'espace des paramètres A_J et σ_J , les faisant respectivement varier entre 50 m.s^{-1} et 100 m.s^{-1} et entre 1000 m et 9000 m. Cet ensemble sera appelé \mathcal{Q} . Un modèle de source ad hoc est utilisé comme dans le cas « Rivesaltes ». Dans la suite on notera f_0 la fréquence caractéristique de la source, que l'on a prise égale à 0.1 Hz pour éviter de devoir modéliser de la propagation d'ondes de plus hautes fréquences (numériquement plus coûteuses). L'instant d'émission t_0 est en $t = 0$.

4.2.2 Extraction des quantités d'intérêt simulées

Comme indiqué précédemment, on ne comparera pas l'ensemble du signal simulé à celui de la détection réelle, mais seulement certaines quantités d'intérêt. Nous comparerons le temps de trajet T_t , l'amplitude maximale P_k et la durée que l'on notera par la suite Δ . Pour ce cas idéalisé, on ne s'intéresse qu'aux détections liées à des réfractions en dessous de 60 km car ce sont celles qui sont intéressantes pour la météorologie. On nommera ces phases spécifiques *Phases S*⁹. La première étape est donc d'identifier dans les signaux les arrivées cohérentes stratosphériques (*Phase S*) et de filtrer les arrivées cohérentes thermosphériques (*Phase T*).

Filtrage simple

La première approche est de considérer que la phase stratosphérique (*Phase S*) est celle de plus forte amplitude. Cette hypothèse se base sur le fait que, dans la réalité, les phases thermosphériques sont atténuées lors de leurs passages dans la thermosphère (Sutherland & Bass, 2004).

Néanmoins, pour un milieu donné, l'amplitude de la *Phase S* peut devenir inférieure à l'amplitude de la *Phase T* en fonction de la distance source-récepteur comme le montre la figure 4.10. Si l'on considère un récepteur à 166 km ou 285 km de la source, pour ces conditions idéalisées, ce critère sélectionne effectivement la *Phase S*. Cependant à 380 km, on voit apparaître la *Phase S rapide* et les deux *Phases S* se retrouvent

8. Les signaux simulés ont été résolus pour 128 fréquences réparties uniformément entre 0 et 0.25 Hz

9. les phases thermosphériques sont peu fiables, en raison des fortes incertitudes associées aux données HWM et MSIS.

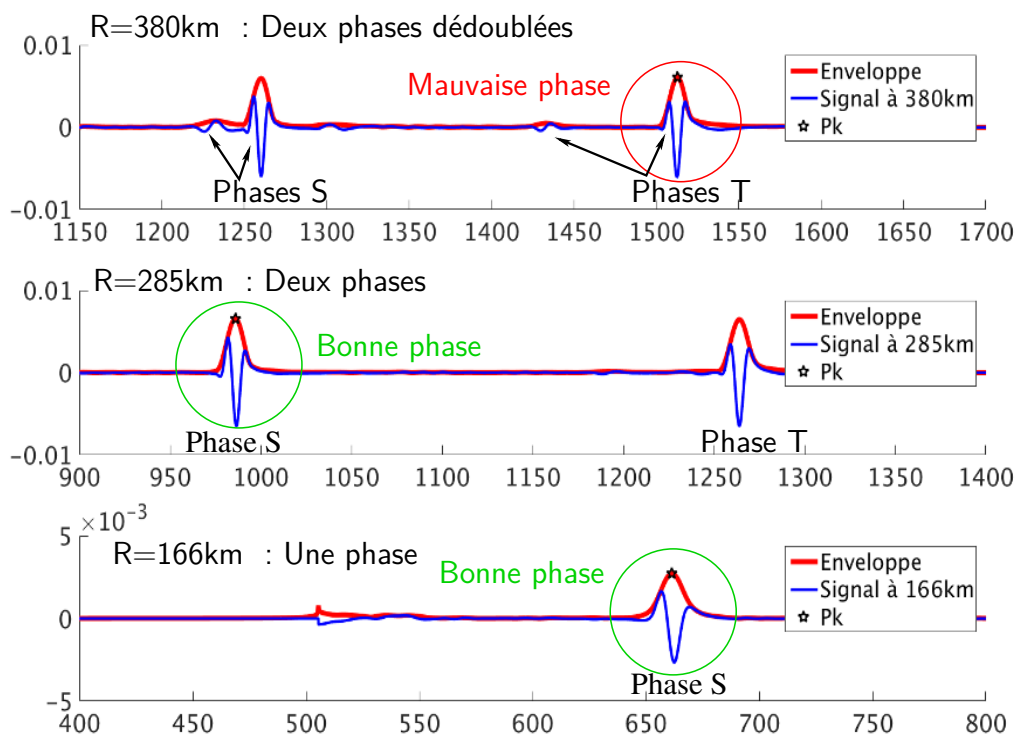


Figure 4.10 : Filtrage visant à éliminer les *Phases T* appliqué sur le signal simulé pour différentes distances source-récepteur (R). Les signaux sont obtenus pour une même atmosphère idéalisée.

d'amplitude inférieure à celle de la *Phase T*. Le critère sélectionne alors la *Phase T* qui peut également être d'intensité supérieure lorsque l'amplitude du jet stratosphérique n'est pas assez importante et que la majorité de l'énergie se réfracte dans la thermosphère. Ce filtrage présente de nombreuses limitations, il a cependant le mérite de ne nécessiter aucune hypothèse sur le temps de trajet des *Phases S*.

Filtrage par troncature du signal

Une autre approche est de supposer que l'on connaît approximativement le temps de trajet de la *Phase S* au niveau de la station et donc de ne faire l'extraction que sur la portion du signal comprise dans une fenêtre temporelle autour de cette estimation. Cette estimation du temps de trajet peut être obtenue par une simulation par tracé de rayons par exemple, mais implique des coûts numériques supplémentaires. Cette estimation fournit une information a priori sur le temps de trajet de la phase et sa durée. Cette information doit être précise sous peine de tronquer le signal d'intérêt ou d'englober d'autres phases. C'est pourquoi la dissociation de phases trop proches les unes des autres n'est pas résolue par cette méthode.

Une fois la *Phase S* identifiée pour chaque profil du quadrillage \mathcal{Q} , nous extrayons P_k , T_t et Δ des signaux enregistrés. Nous utiliserons par défaut, le filtrage simple pour identifier les *Phases S*.

Extraction de l'amplitude maximale P_k

L'amplitude maximale P_k est donnée par la valeur maximale prise par la norme de la transformée de Hilbert du signal à la station réceptrice. On remarque sur la figure 4.11 que pour les jets stratosphériques de faibles amplitudes ($A_J=55 \text{ m.s}^{-1}$), la valeur de P_k varie très faiblement quel que soit celle de σ_J . Ce qui se passe c'est que la phase de plus grande amplitude enregistrée à 285 km est la *Phase T* pour ces conditions. Le filtrage naïf sélectionne alors cette *Phase T* qui n'est plus que peu affectée par les caractéristiques du jet stratosphérique. C'est pourquoi l'amplitude maximale P_k ne varie que très peu pour des jets stratosphériques de faibles amplitudes. On remarque également que P_k présente un maximum sous ces conditions vers ($A_J = 80, \sigma_J = 9000$).

Extraction du temps de trajet T_t

On considère que le temps de trajet T_t correspond à la différence entre le temps d'arrivée du maximum d'amplitude P_k et le temps d'émission pris comme origine du temps. On constate alors que le filtrage simple de la phase engendre une discontinuité sur T_t pour les jets de faibles amplitudes, visible sur la figure 4.12. Cette discontinuité est due à la sélection de la *Phase T*, plus tardive, au lieu de la *Phase S* par le filtrage.

Extraction de la durée du signal Δ

La première définition prise pour la durée, que l'on note Δ_{RSB} , correspond à la portion du signal dont le rapport de l'amplitude à sa valeur maximale, qu'on note RSB, est supérieur à 10 % comme le montre la figure 4.13.

Cette définition rend Δ_{RSB} dépendante de la fréquence d'échantillonnage des signaux. Cette dépendance engendre un bruit numérique sur Δ_{RSB} , visible sur la figure 4.14, issu de l'incertitude que l'on a sur l'instant exact du dépassement du RSB. On constate également sur la figure 4.14 que cette définition conduit à deux discontinuités dans l'espace \mathcal{Q} . La première est, là encore, due à la sélection de la *Phase T* au lieu de la

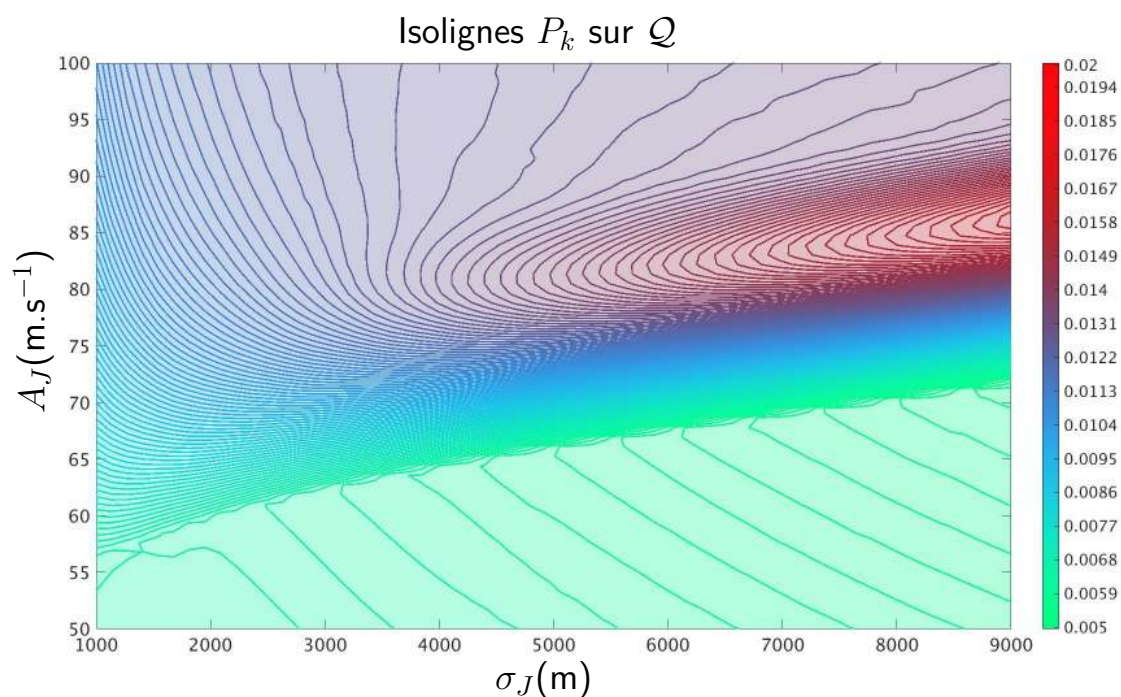


Figure 4.11 : Amplitude maximale P_k des signaux enregistrés à 285 km de la source en fonction des paramètres de jet stratosphérique A_J et σ_J . Les iso-valeurs vont de 5.0×10^{-3} Pa à 2.0×10^{-2} Pa.

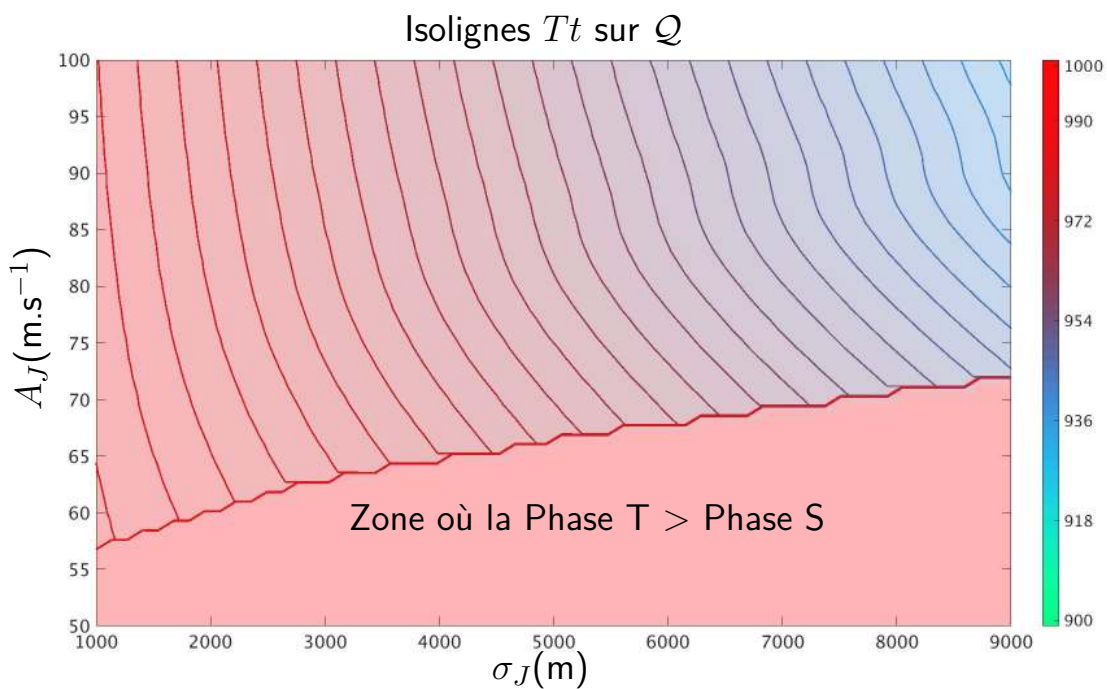


Figure 4.12 : Temps de trajet Tt des signaux enregistrés à 285 km en fonction de A_J et σ_J . Les iso-valeurs vont de 900 s à 1000 s et +.

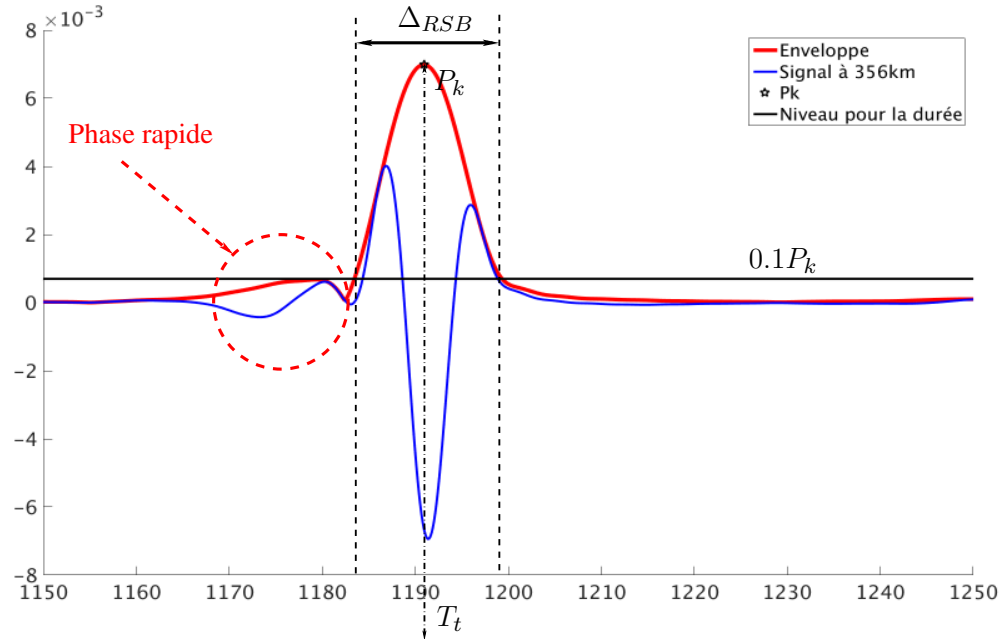


Figure 4.13 : Extraction des quantités d'intérêt simulées. L'amplitude des signaux est en Pascal.

Phase S par le filtrage. La seconde discontinuité est due à la séparation de la phase en une *Phase S* dite rapide et une dite lente.

Une autre définition de la durée est possible en considérant que l'enveloppe des signaux peut être approximée par une gaussienne permettant ainsi de l'exprimer en fonction de son écart-type. Avec cette approximation, on peut définir Δ_G comme la solution d'un problème de minimisation.

Cette définition nécessite un coût calcul supplémentaire lié à l'étape de minimisation. L'évolution de cette durée en fonction des paramètres du jet est présentée sur la figure 4.15. La discontinuité due à la *Phase T* est toujours présente et les isovalues sont encore bruitées.

La dernière définition proposée pour la durée est basée non plus sur le contenu temporel du signal mais sur son contenu fréquentiel. Afin de limiter l'implication de la *Phase T* dans le spectre, elle requiert la sélection de la *Phase S* par troncature du signal. L'algorithme recherche d'abord la centroïde ω_p définie par :

$$\omega_p = \int \omega \|\hat{s}(\omega)\|^p d\omega \left[\int \|\hat{s}(\omega)\|^p, d\omega \right]^{-1}, \quad (4.4)$$

où \hat{s} est la transformée de Fourier du signal et p est l'ordre de la centroïde.

On définit alors $\Delta_{\omega_p} = 1/\omega_p$. Cette définition est dépendante du nombre de fréquences résolues. Comme pour les autres observables, on a étudié son comportement en fonction des caractéristiques du jet. La figure 4.17 montre que l'utilisation d'une centroïde d'ordre 2 pour la définition de Δ conduit à une topologie plus régulière qu'avec une centroïde d'ordre 1 (4.16).

4.3 Problème inverse dans un cadre idéalisé

On considère que la réduction de l'écart entre les quantités d'intérêt des détections en station $\mathbf{Y}^t = (y_{i=1..3}^t)(P_k^t, T_t^t, \Delta^t)$ et celles simulées $\mathbf{Y}(\mathbf{X})$ indique que les spécifications

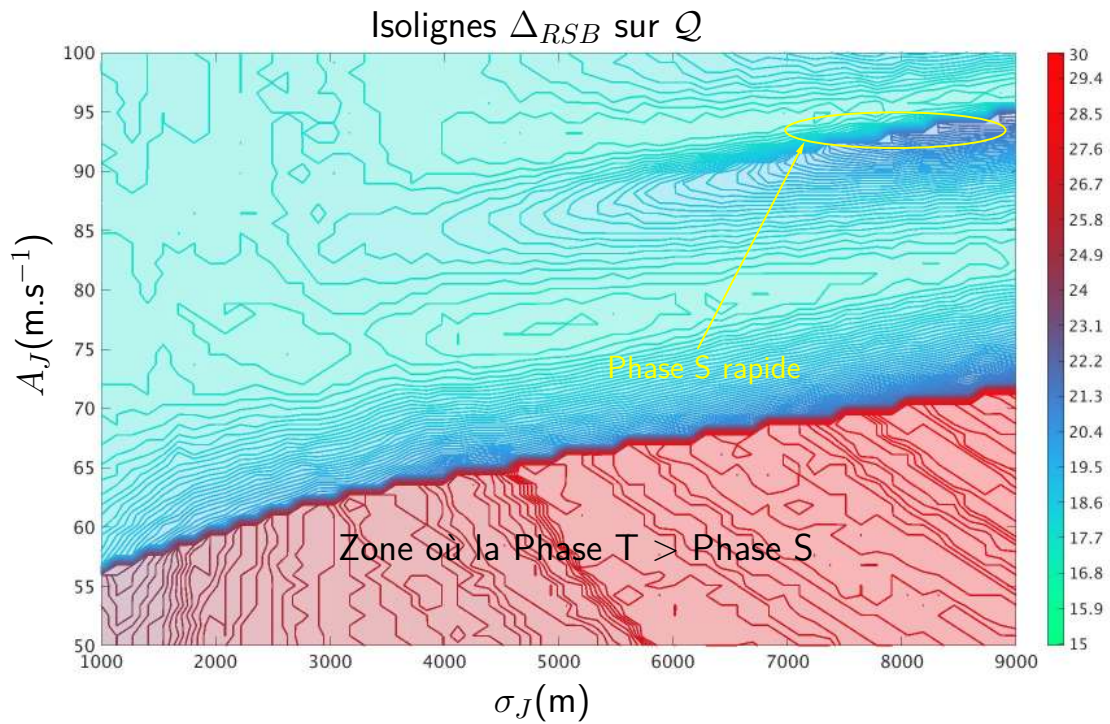


Figure 4.14 : Durée Δ_{RSB} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 15 s à 30 s et +.

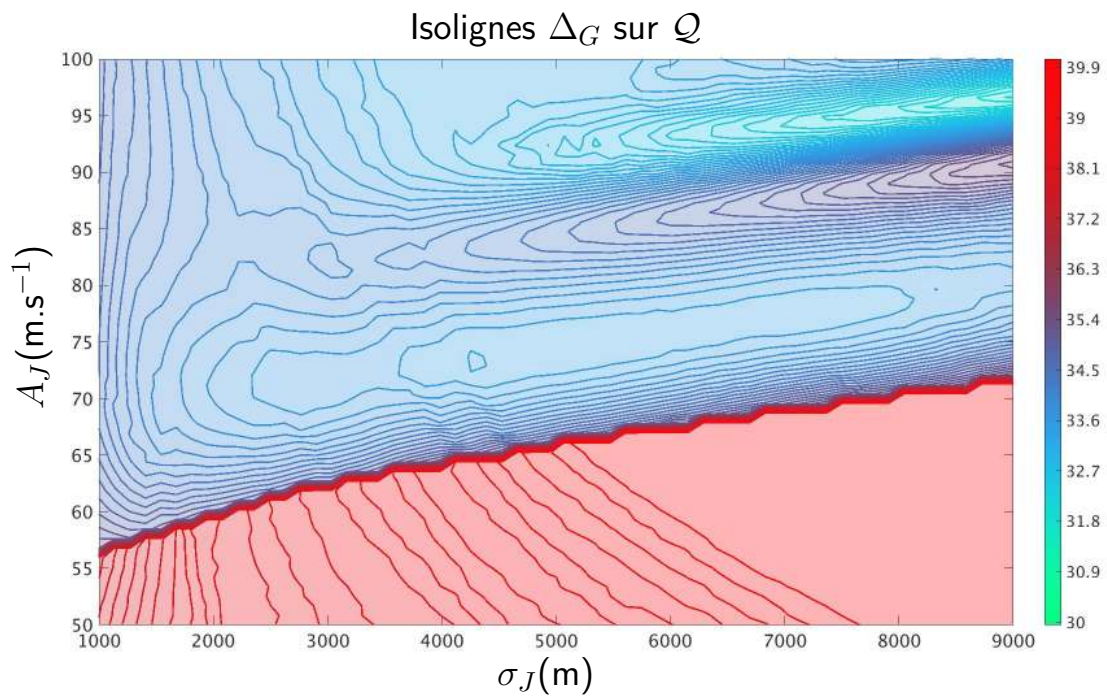


Figure 4.15 : Durée Δ_G (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 30 s à 40 s et +.

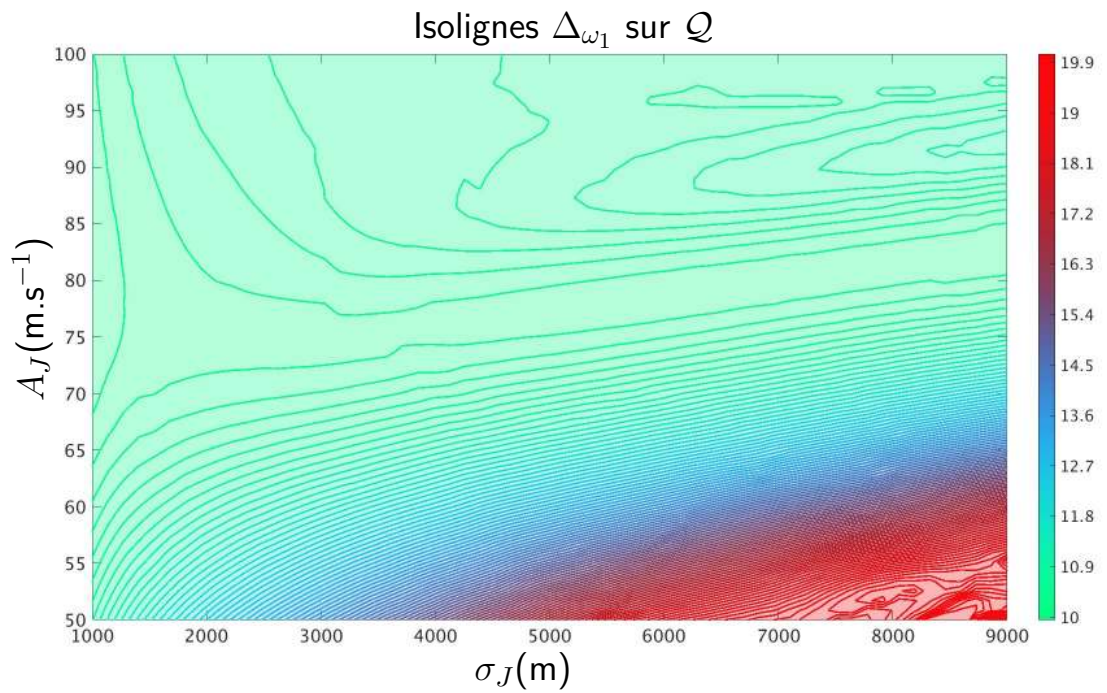


Figure 4.16 : Durée Δ_{ω_1} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 10 s à 20 s et +.

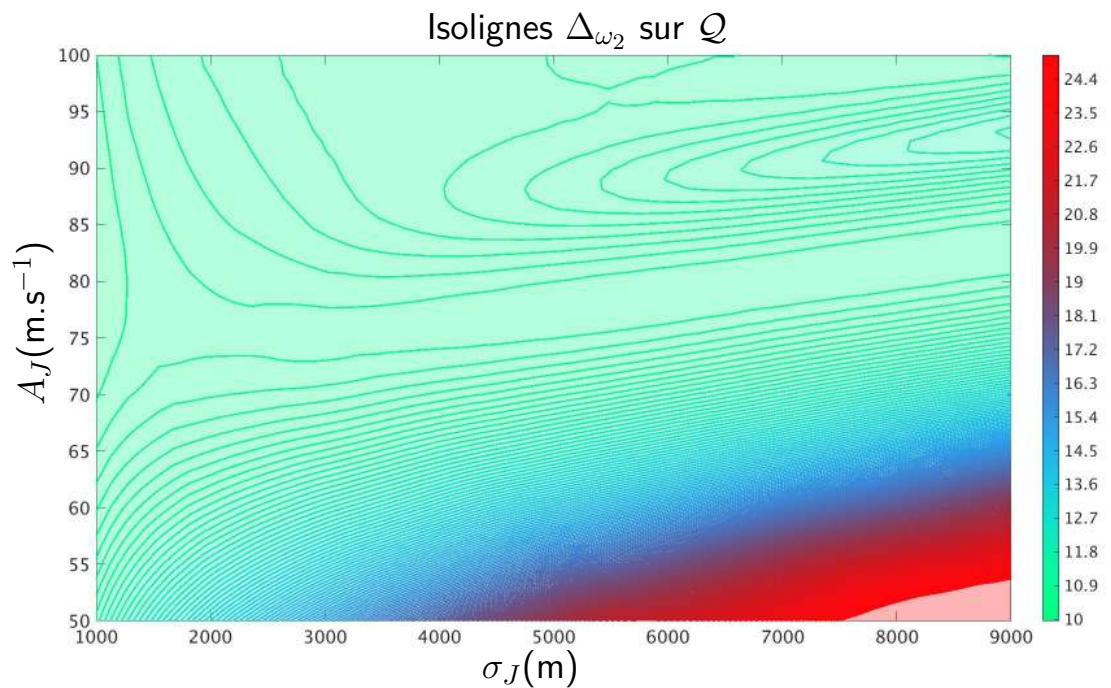


Figure 4.17 : Durée Δ_{ω_2} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 10 s à 25 s et +.

du milieu de propagation \mathbf{X} (ici réduites aux paramètres A_J, σ_J) se rapprochent de celles réellement rencontrées par les ondes infrasonores au cours de leur propagation. On définit alors une fonction objective basée sur cette différence entre quantités simulées $\mathbf{Y}(\mathbf{X})$ et celles observées \mathbf{Y}^t :

$$J(\mathbf{X}) = (\mathbf{Y}(\mathbf{X}) - \mathbf{Y}^t)^T \mathbf{R}^{-1} (\mathbf{Y}(\mathbf{X}) - \mathbf{Y}^t) \quad (4.5)$$

\mathbf{R} est une matrice pondérant les observations. Ce formalisme se rapproche du problème d'assimilation de données variationnelle du chapitre 2, à la différence que l'on ne considère aucun a priori sur l'état atmosphérique \mathbf{X} . On fait le choix de considérer \mathbf{R}^{-1} diagonale et constituée des $1/\|y_i^t\|^2$. Cela permet de travailler en écart relatif pour chaque quantité. On utilisera le filtrage simple de la *Phase S* et Δ suivra la définition basée sur le RSB.

En suivant ce formalisme, le vecteur \mathbf{X} minimisant globalement J sera l'état de l'atmosphère qui simulera au mieux les observations. On souhaite alors identifier ce minimum à moindre coût et pour ce faire, on se propose d'utiliser des algorithmes d'optimisation locale. Par la suite, on sélectionne une réalisation parmi celles de \mathcal{Q} que l'on choisie comme notre état réel. On considère les quantités de la *Phase S* simulées en station comme étant celles observées. Ainsi J possède au moins un minimum égal à zéro pour cette réalisation spécifique dans l'espace échantillonné par \mathcal{Q} . La réalisation que l'on considère comme l'état réel est celle avec un jet stratosphérique de caractéristiques $A_J = 88 \text{ m.s}^{-1}$ et $\sigma_J = 4500 \text{ m}$.

4.3.1 Optimisation locale

On pose alors le problème de minimisation suivant :

$$\begin{cases} \min_{\mathbf{X}} J(\mathbf{X}) \\ c_i(\mathbf{X}) = 0, \quad i \in \mathcal{E} \\ c_j(\mathbf{X}) \geq 0, \quad j \in \mathcal{I} \end{cases} \quad (4.6)$$

où J est notre fonction objective à minimiser, les $c(\mathbf{X})$ représentent les contraintes que l'on s'impose pour la minimisation. En l'absence de contrainte, on parle d'optimisation sans contrainte.

L'optimisation locale utilise les propriétés locales de J pour rechercher de manière itérative un de ses minima. Nous avons sélectionné ici des algorithmes de descente en gradient ou de Newton que l'on combinera ensuite à un estimateur de pas par «recherche en ligne». Ces algorithmes se basent sur le développement de Taylor de J en \mathbf{X} :

$$J(\mathbf{X}_k + \mathbf{p}_k) \approx J(\mathbf{X}_k) + \mathbf{p}_k^T \nabla J(\mathbf{X}_k). \quad (4.7)$$

La formule (4.7) implique l'hypothèse d'une certaine régularité de J pour que le gradient soit défini. Dans la suite, un problème linéaire non-contraint est examiné, mais le principe général de ces algorithmes demeure le même dans le cas avec contrainte moyennant quelques adaptations. Notre problème inverse est de petite dimension et le gradient peut être estimé par un schéma aux différences finies pour un moindre coût numérique. On se propose alors de calculer directement le gradient local via un schéma centré d'ordre 2.

Le choix de la pente doit engendrer une diminution de J . D'après l'équation (4.7), si l'on cherche à réduire le plus possible J , on trouve qu'il faut que $\mathbf{p}_k = -\nabla J(\mathbf{x}_k)$ pour maximiser la réduction. C'est la direction de plus grande pente et la base de l'algorithme de « descente en gradient ». Si l'on cherche également à tenir compte de la courbure locale de J , il faut alors développer J à l'ordre deux. La meilleure pente minimise alors :

$$\mathbf{p}_k = \min_{\mathbf{p}} J(\mathbf{X}_k) + \mathbf{p}^T \nabla J(\mathbf{X}_k) + \frac{1}{2} \mathbf{p}^T \nabla^2 J(\mathbf{X}_k) \mathbf{p}. \quad (4.8)$$

Si la hessienne $\nabla^2 J(\mathbf{x}_k)$ est définie positive au voisinage du point \mathbf{X}_k , alors la pente optimale \mathbf{p}_k est égale à :

$$\mathbf{p}_k = -(\nabla^2 J(\mathbf{X}_k))^{-1} \nabla J(\mathbf{X}_k). \quad (4.9)$$

C'est la pente de Newton. Le calcul exact de la hessienne par des différences finies est souvent compliqué et coûteux. Il a donc été développé des méthodes visant à l'approximer, exploitant le gradient calculé durant les itérations précédentes : ce sont les méthodes de Quasi-Newton. Pour cela, on définit une nouvelle matrice B_k de même dimension que la hessienne, avec comme propriété d'être symétrique définie positive. Le but est de faire en sorte que B_k , approxime $\nabla^2 J(\mathbf{X}_k)$ au voisinage de \mathbf{X}_k . La pente s'écrit alors :

$$\mathbf{p}_k = -B_k^{-1} \nabla J(\mathbf{X}_k), \quad (4.10)$$

où B_k est actualisé à chaque itération en tenant compte de l'information des itérations précédentes.

Une fois la pente choisie, la seconde étape consiste en la définition de la longueur du pas amenant à une progression acceptable. Le but étant d'atteindre le minimum avec le moins d'itérations possibles. Pour cela il faut optimiser la progression. On définit un paramètre α_k tel que $\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{p}_k$ qui maximise la progression ou, autrement dit, minimise J dans la direction donnée par la pente. On parle alors de recherche en ligne. Une première approche serait d'évaluer à l'aveugle $J(\mathbf{X}_{k+1})$ pour différentes valeurs de α_k et de prendre le α_k qui minimise J , mais cette stratégie simpliste nécessite un grand nombre d'évaluations de J . C'est pourquoi il est commun d'utiliser une « recherche en ligne inexacte ». Elle consiste à se donner des conditions à satisfaire pour α_k afin d'assurer la bonne progression de l'algorithme. Une première condition, qui généralement assure une réduction de J dans la direction \mathbf{p}_k , est donnée par :

$$J(\mathbf{X}_k + \alpha_k \mathbf{p}_k) - J(\mathbf{X}_k) \leq c_1 \alpha_k \nabla J(\mathbf{X}_k) \cdot \mathbf{p}_k, \quad (4.11)$$

où c_1 est une constante prise arbitrairement. On prend souvent $c_1 \ll 1$.

On peut ajouter à cette condition, une condition sur le gradient de $\mathbf{X}_k + \alpha_k \mathbf{p}_k$, comme :

$$\nabla J(\mathbf{X}_k + \alpha_k \mathbf{p}_k) \cdot \mathbf{p}_k \geq c_2 \nabla J(\mathbf{X}_k) \cdot \mathbf{p}_k, \quad (4.12)$$

où c_2 est une constante arbitraire comprise entre c_1 et 1. Ces deux conditions sont connues sous le nom de *conditions de Wolfe* dont l'influence sur le choix du pas d'optimisation est illustrée sur la figure 4.18. Si l'on utilise la valeur absolue de la condition de courbure (4.12), on parle alors de conditions de Wolfe fortes. Il existe également d'autres conditions pour définir un pas acceptable. Une description complète de ces méthodes d'optimisation par recherche en ligne est disponible dans Nocedal et Wright (2006).

4.3.2 Expression de J sous différentes hypothèses

On effectue l'extraction de la *Phase S* par l'approche dite simple sur \mathcal{Q} et on estime J sur cet espace. On peut noter sur la figure 4.19 deux discontinuités. L'une pour les jets de faibles amplitudes, correspondant au basculement de l'algorithme vers la phase thermosphérique (*Phase T*), et une seconde pour les jets de fortes amplitudes et épaisseurs correspondant à la séparation de la phase rapide et de la phase lente. On remarquera également les irrégularités qui brulent J et qui se traduisent par la complexité des

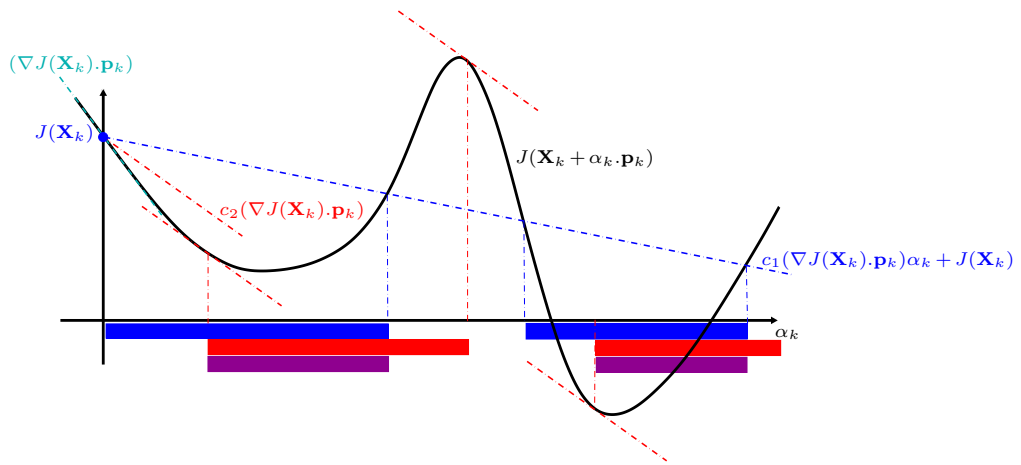


Figure 4.18 : Schéma de la « recherche en ligne » avec conditions de Wolfe pour la fonction J . En violet les zones où α_k satisfait les conditions (4.11) et (4.12). En rouge les zones où α_k vérifie uniquement la condition de courbure (4.12). En bleu les zones où α_k vérifie uniquement la condition de réduction (4.11). En cyan le gradient au point $J(\mathbf{X}_k)$. Figure inspirée de celle proposée dans Nocedal et Wright (2006).

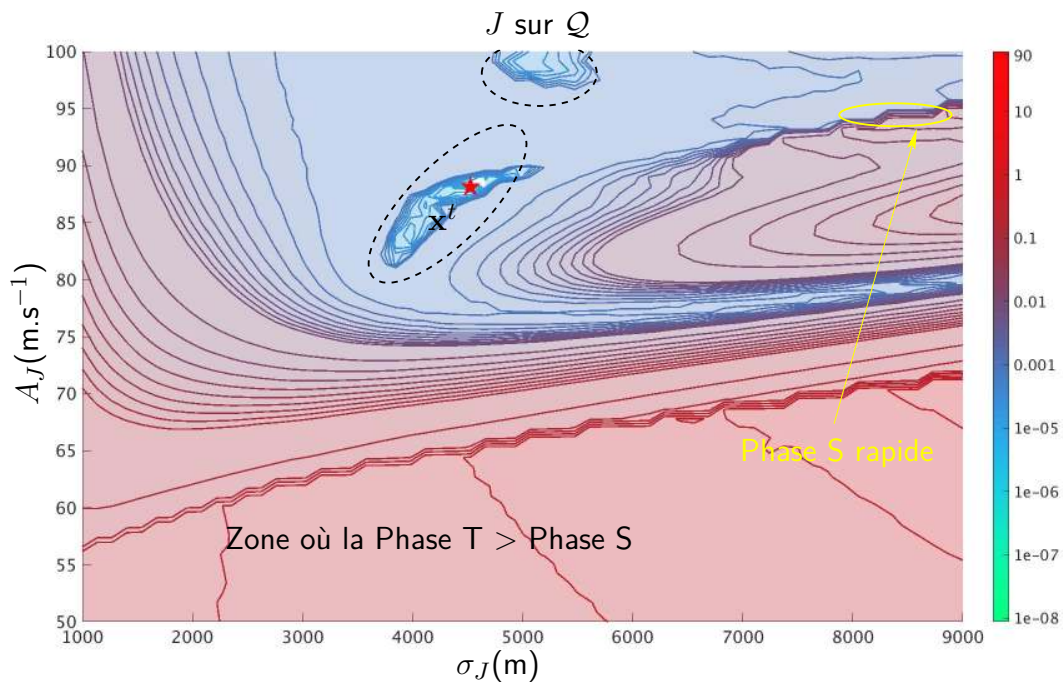


Figure 4.19 : Évolution de J sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge. Les iso-valeurs vont de 0 à 100 et + avec le même code couleur que la figure 4.11.

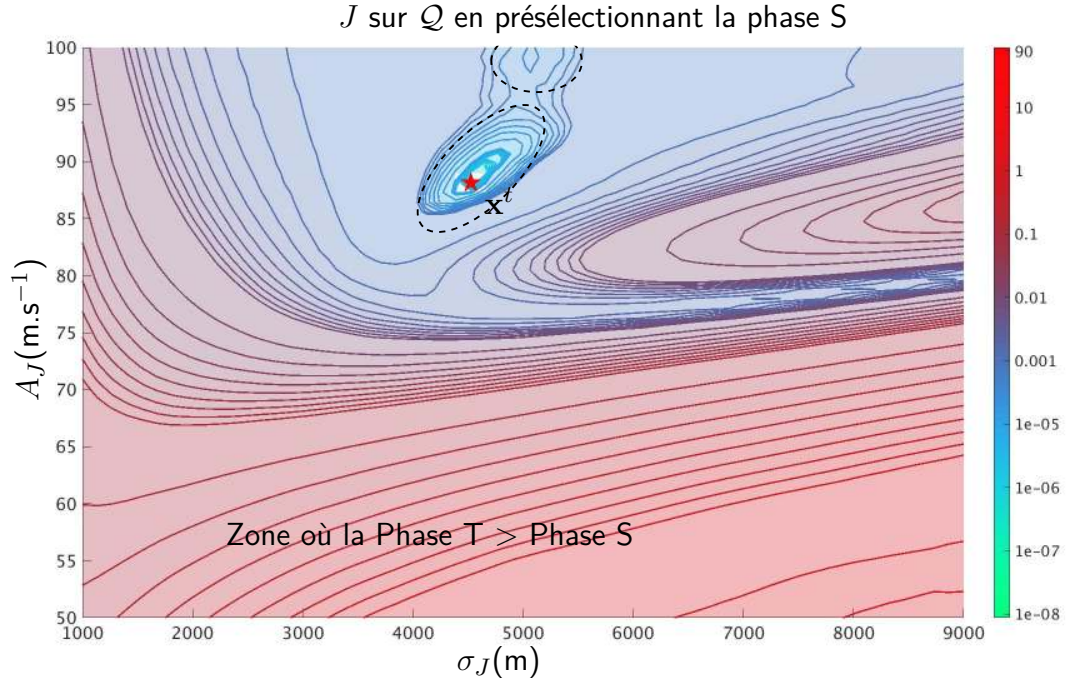


Figure 4.20 : Évolution de J en utilisant Δ_{ω_2} sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge. Les iso-valeurs vont de 0 à 100 et + avec le même code couleur que la figure 4.11.

iso-contours de la figure 4.19. On se propose alors d'utiliser la sélection de phase par troncature du signal, bien qu'elle implique un coût numérique supplémentaire (estimation a priori des temps de trajet et des durées des phases). Cette méthode d'extraction élimine la discontinuité de J associée au basculement de l'algorithme vers la *Phase T*. Pour lisser les non-régularités de J , on prendra pour Δ la définition basée sur la centroïde d'ordre 2. La figure 4.20 montre les valeurs de J en considérant cette nouvelle définition. Si on la compare à la figure 4.19, les non-régularités disparaissent avec cette définition de la durée et le problème possède moins de minima locaux. Cependant, le resserrement des iso-contours autour du minimum global indique une forte variation du gradient autour de celui-ci. Le minimum est piqué et donc particulièrement difficile d'accès avec un algorithme d'optimisation. Pour élargir le minimum, on applique une fonction f strictement croissante à J de manière à conserver les minima tout en les élargissant. On obtient un problème de minimisation anamorphe à celui de J pour $f(J)$. On applique la transformation suivante :

$$J \leftarrow K \log(1 + \beta J_{base}) \quad (4.13)$$

où J_{base} correspond à la définition (4.5) précédente de J . Le paramètre β permet de modifier l'élargissement de la courbure autour des minima. Après plusieurs essais, $\beta = 10^4$ et $K=20$ permettent un bon compromis entre l'élargissement du minimum global et celui des autres minima de J . Le comportement de J sur \mathcal{Q} est représenté sur la figure 4.21. J ne possède plus que deux minima visibles :

- le minimum global recherché en $A_J = 88 \text{ m.s}^{-1}$ et $\sigma_j = 4500 \text{ m}$ correspondant à l'état réel et qui est marqué par une convergence des directions du gradient présentées sur la figure 4.21 (b) ;
- un minimum local aux environs des $A_J \approx 98 \text{ m.s}^{-1}$ et $\sigma_J \approx 5000 \text{ m}$ où le gradient est très faible.

Dans ces conditions, le problème inverse possède une unique solution globale sur l'espace \mathcal{Q} .

4.3.3 Résultats du problème inverse

Les méthodes locales nécessitent généralement peu d'évaluations de J , mais peuvent converger vers des minima locaux. Malgré la présence d'un second minima local pour le problème qui nous intéresse, on se propose d'utiliser des algorithmes locaux d'optimisation pour minimiser J .

Algorithme de descente en gradient sans recherche en ligne

On utilise pour J la définition (4.13) et l'algorithme de descente en gradient en fixant le pas à $\|\alpha_k \mathbf{p}_k\| = \nabla_J \mathbf{X}$. Le pas utilisé pour estimer le gradient par différences finies est donné par ($\delta A_j = 1 \text{ m.s}^{-1}$, $\delta \sigma_j = 100 \text{ m}$). Afin de se prémunir d'une possible divergence de l'algorithme, on limite le nombre d'itérations à 100 et la progression minimale à 10^{-15} . Les critères de convergence correspondent à $\nabla J < 1.10^{-5}$ et une valeur de $J < 0.01$. On initialise la recherche au point $\mathbf{X} = (85, 2500)$ et on recherche la cible $\mathbf{y}^t = H(\mathbf{x}^t)$ pour ($A_j^t = 88 \text{ m.s}^{-1}$, $\sigma_j^t = 4500 \text{ m}$). On constate sur la figure 4.22 que :

- l'algorithme progresse rapidement vers la région de faible gradient, mais ne converge pas en moins de 100 itérations pour cette initialisation. Il oscille dès la 80^{ème} itération entre les points (87.84, 4277) et (87.79, 4983) ;
- En bloquant le pas égal au gradient, l'algorithme n'arrive pas à s'introduire dans le creux de l'espace où se trouve le minimum. Le minimum doit être encore trop piqué.

D'autres minimisations avec d'autres initialisations ont été faites mais elles ont conduit aux mêmes oscillations autour du minimum.

Algorithme de descente en gradient avec recherche en ligne

Nous avons alors cherché à limiter l'influence du mauvais dimensionnement du pas en ajoutant une recherche en ligne de type *back-tracking*. Le principe consiste à diminuer α_k depuis une valeur préalablement fixée, tant que cela n'amène pas à une diminution de J . On obtient alors au bout de 100 itérations en partant de la même initialisation le comportement présenté sur la figure 4.23. Il apparaît que :

- l'algorithme se rapproche du minimum global mais converge vers un point qui ne correspond pas au minimum de J , mais à $J = 0.14$;
- l'algorithme converge plus rapidement que précédemment.

Ici, l'algorithme s'arrête avant 100 itérations. Il n'a pas atteint les critères de convergence, mais son pas de progression est devenu trop petit dans la direction proposée par le gradient. Au voisinage du minimum, il devient compliqué d'estimer précisément le gradient et sa direction. Un moyen d'éviter ce problème, c'est de contraindre le problème inverse dans une direction donnée lorsque l'estimation du gradient devient délicate. Cela peut se faire en ajoutant une contrainte.

Optimisation locale sous contrainte

Dans la mesure où il est nécessaire d'estimer a priori le temps de trajet pour pouvoir utiliser le filtrage par troncature du signal et la définition de la durée basée sur la centroïde, on se propose d'utiliser le temps de trajet T_t de la *Phase S* en tant que contrainte. On cherche alors, dans ce cas précis, à estimer le lien entre T_t et (A_j, σ_j) en considérant pour

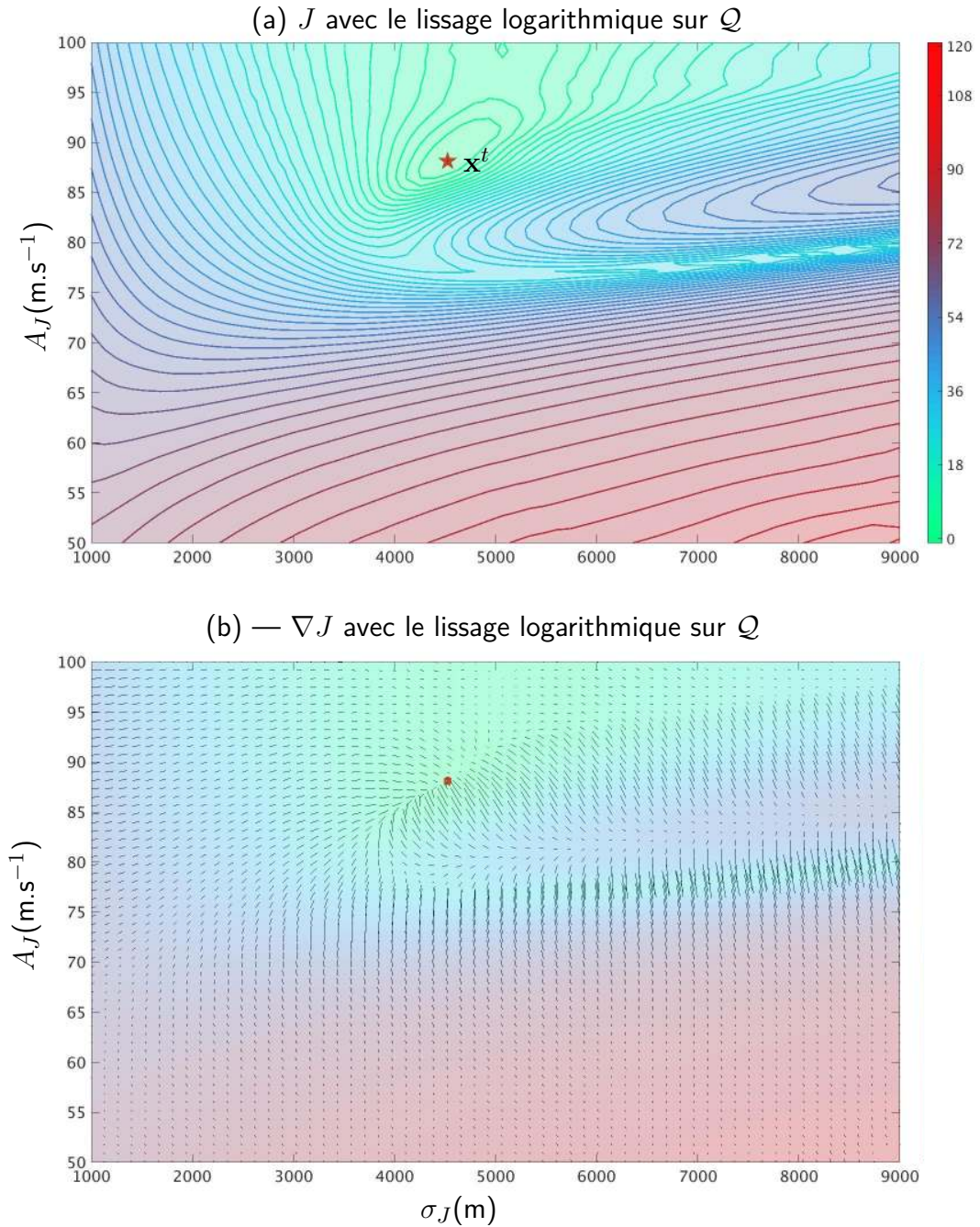


Figure 4.21 : Évolution de J et de son gradient sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge. (a) Valeur de J sur \mathcal{Q} . (b) La couleur indique la valeur de J de manière identique à (a). Les barres correspondent aux gradients ∇J . Leur longueur est proportionnelle à la valeur du gradient.

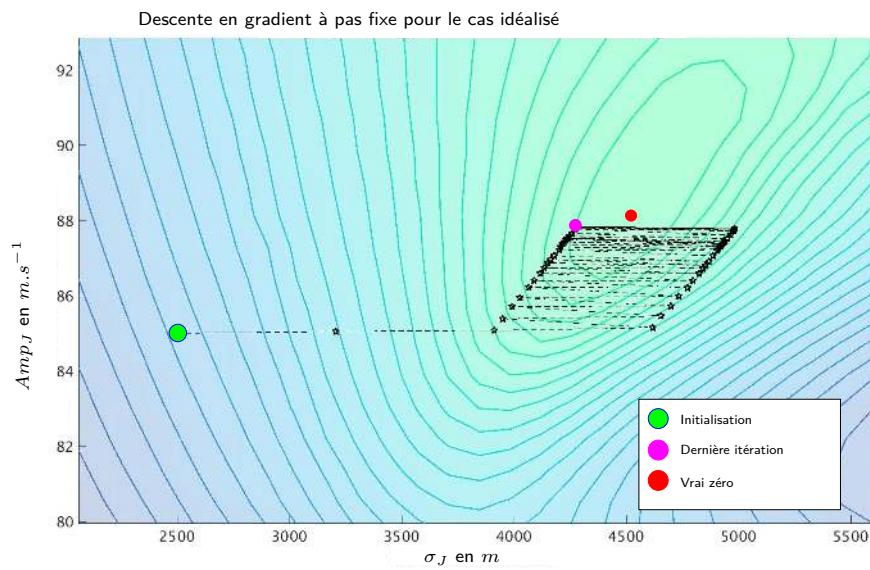


Figure 4.22 : Points donnés par la méthode de descente en gradient sans recherche en ligne pour une initialisation en $\mathbf{X} = (85, 2500)$.

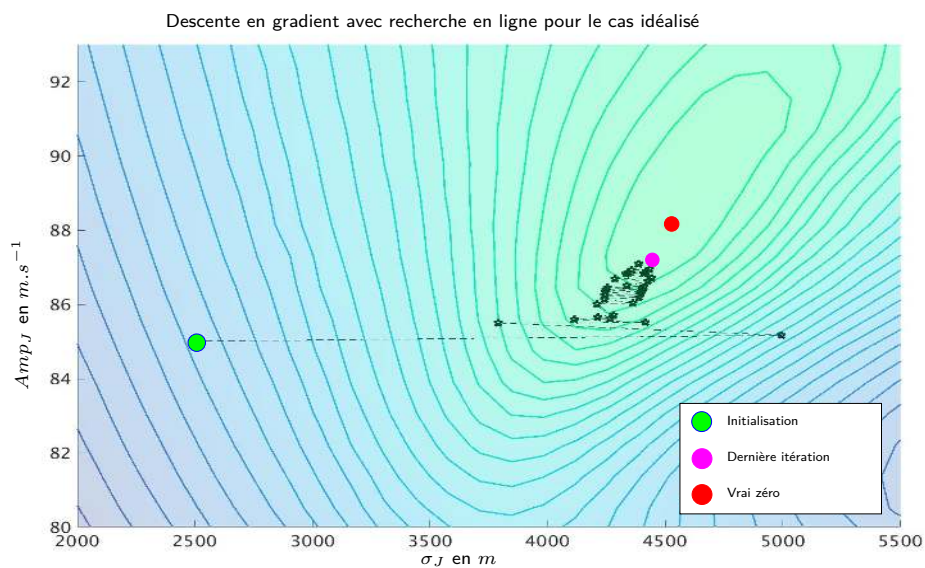


Figure 4.23 : Points donnés par la méthode « de descente de gradient avec recherche en ligne » pour une initialisation en $\mathbf{X} = (85, 2500]$.

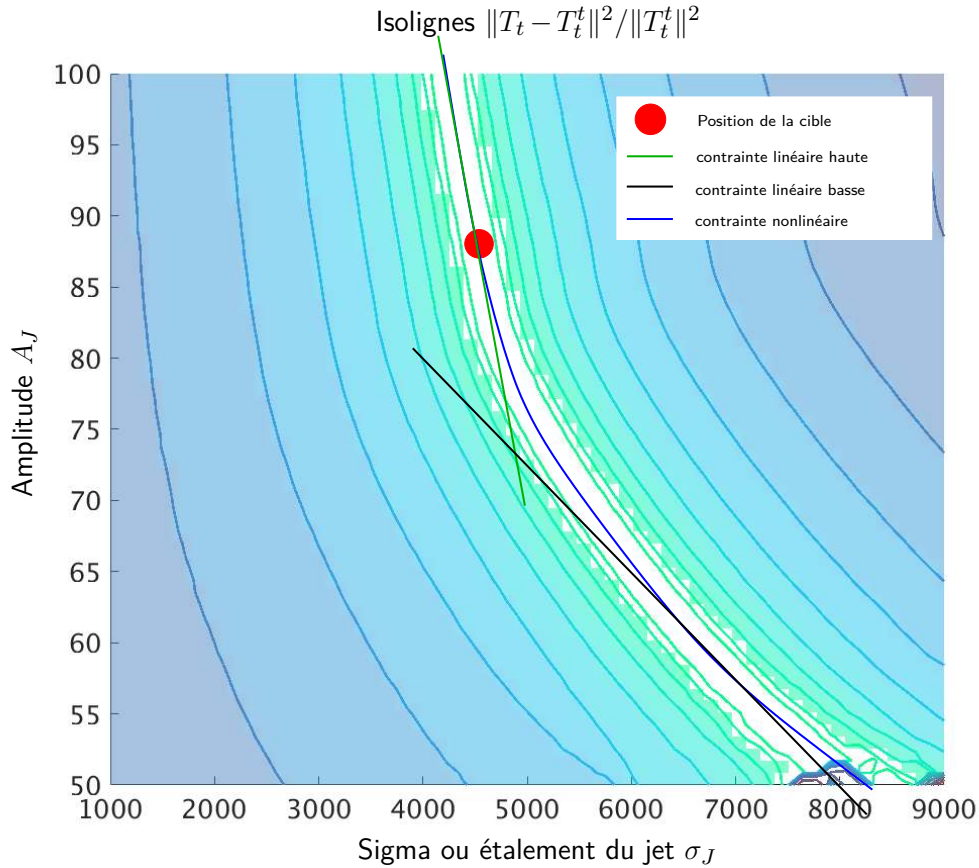


Figure 4.24 : Définition de la contrainte en T_t dans le cas de l'atmosphère idéalisée

cela une loi linéaire au voisinage du minimum global. On effectue une régression linéaire sur les points du quadrillage proches du minimum global vérifiant $\|T_t - T_t^t\|^2 = 0$, présentée dans la figure 4.24. On obtient une loi de contrainte linéaire de la forme $c_{T_P}(A_J, \sigma_J) = 0$. Cette estimation ad hoc de la contrainte n'est valable que pour ce vecteur d'observation \mathbf{Y}^t cible. On a donc à résoudre le problème non-régulier, non-convexe sous contrainte linéaire suivant :

$$\left\{ \begin{array}{l} \min_{(A_J, \sigma_J)} J(A_J, \sigma_J) = 20 \log(1 + 10^4 J_{base}) \\ c_{T_P}(\mathbf{x}) = 0 \\ 50 \text{ m.s}^{-1} \leq A_J \leq 250 \text{ m.s}^{-1} \\ 1000 \text{ m} \leq \sigma_J \leq 15000 \text{ m}. \end{array} \right. \quad (4.14)$$

On ajoute une contrainte sur le domaine de recherche pour que l'algorithme puisse s'arrêter rapidement si celui-ci s'éloigne fortement du minimum global. Pour résoudre les équations (4.14), on utilise la fonction *fmincon* de MATLAB en choisissant dans un premier temps une résolution à l'aide d'un algorithme de Quasi-Newton utilisant l'approximation BFGS (Broyden, 1970 ; Fletcher, 1970 ; Goldfarb, 1970 ; Shanno, 1970) pour la hessienne avec conditions de Wolfe sur le pas. Le choix s'est porté sur le BFGS, car il a montré un bon comportement sur des problèmes possédant de multiples minima locaux et non-réguliers (Curtis *et al.*, 2017). La minimisation de J avec comme initialisation $(A_J^0 = 85 \text{ m.s}^{-1}, \sigma_J^0 = 4500 \text{ m})$ conduit aux résultats suivants :

- l'algorithme converge en 11 itérations vers le minimum global ;
- l'algorithme n'a effectué que 73 évaluations de J .

Pour d'autres initialisations, l'algorithme n'a pas convergé vers le minimum global.

On se propose alors d'utiliser une *méthode de relaxation de la contrainte* pour mieux explorer l'espace. Ce nouvel algorithme montre qu'une autre initialisation parmi celles testées a permis à l'algorithme de converger vers le minimum global. L'algorithme converge également pour $(A_J^0=95 \text{ m.s}^{-1}, \sigma_J^0=9000 \text{ m})$. Les caractéristiques pour ces deux minimisations sont :

$ini = (85, 4500)$ Convergence en 15 itérations et 104 évaluations de J ;

$ini = (95, 9000)$ Convergence en 15 itérations et 157 évaluations de J .

Le coût calcul est néanmoins 1,5 fois plus élevé avec l'utilisation de la relaxation de contrainte.

4.4 Conclusions

Dans ce chapitre, on s'est intéressé aux moyens de mesurer l'écart entre des observations infrasons et des résultats de simulation pour résoudre le problème inverse par minimisation de cet écart. Au cours d'une étude de la propagation entre Rivesaltes et l'OHP, on a pu définir ce que représente une détection d'infrasons et quelles sont ces quantités caractéristiques. Les simulations par tracé de rayons n'ont pu expliquer le signal perçu en station, car celle-ci se trouvait en zone d'ombre géométrique. Le code de propagation FLOWS, lui, n'a permis d'expliquer qu'une partie des détections. La question alors s'est posée de savoir si le reste du signal détecté était associé à d'autres sources. On a pu voir cependant qu'une perturbation physiquement acceptable de l'atmosphère en un point précis permettait de justifier l'ensemble du signal détecté. Cette perturbation, dont le réalisme n'a pas été établi, illustre la sensibilité des résultats de simulation vis-à-vis de la précision de l'état atmosphérique utilisé en entrée.

Afin d'évaluer le comportement de l'inversion, nous avons défini une atmosphère idéalisée à deux paramètres. On a alors cherché une formulation de la fonction objective à minimiser qui traduise l'écart entre les résultats de simulation et les observations. Afin de limiter les coûts de calcul, nous avons employé des algorithmes d'optimisation locale pour minimiser la fonction objective. On en conclut que l'optimisation sous contrainte sur le temps T_t est la meilleure approche testée pour résoudre le problème inverse (4.14). On note que, même dans un cas idéalisé, la restitution de profils à partir de ces méthodes est compliquée. Il est donc important de considérer un a priori sur l'état atmosphérique afin de régulariser l'espace le plus possible, comme a pu le montrer Lalande (2012). La régularisation apportée par un a priori permet de résoudre le problème inverse sous réserve que son poids ne soit pas trop important dans la fonctionnelle et domine l'apport de la comparaison aux observations. Nous avons vu au chapitre 2, qu'il existe des ensembles d'états prévus ou analysés par les modèles de PNT. On se propose d'utiliser ces ensembles pour contraindre le problème inverse consistant à estimer des profils atmosphériques à partir des données infrasons et d'examiner s'ils permettent de réduire le nombre de minima. De plus, seulement trois quantités issues des détections ont été utilisées. On se propose d'utiliser une approche bayésienne pour extraire nos observations de l'atmosphère. Contrairement à ce chapitre, cette méthode sera évaluée sur la base de près de 6000 détections réelles.

Chapitre 5

Approche bayésienne pour l'inversion des données infrasons

Sommaire

5.1	Résumé et études préliminaires	78
5.1.1	Résumé	78
5.1.2	Points techniques révélés par l'étude	79
5.2	Bayesian selection	83
5.2.1	Abstract	83
5.2.2	Introduction	84
5.2.3	Infrasound propagation modeling	85
5.2.4	Infrasound-based atmospheric state selection	87
5.2.5	Application to the detections of Mount Etna	89
5.2.6	Discussion	95
5.2.7	Relevance of the best atmospheric states	97
5.2.8	Conclusions and Perspectives	101
5.3	Compléments à l'article	103
5.3.1	Annexe de l'article : <i>Numerical set-up of the Bayesian inversion over synthetic observations</i>	103
5.3.2	Utilisation des temps de trajet	107
5.3.3	Modèle de propagation <i>range-dependent</i>	107
5.3.4	Apport des prévisions à court terme	109
5.3.5	Cas d'une autre station	110

Les conclusions du chapitre précédent nous ont montré qu'il est important d'utiliser une information a priori sur la spécification du milieu pour mieux contraindre le problème d'inversion de données infrasons. Afin de se rapprocher de ce qui pourrait être le cadre de l'assimilation de données en prévision numérique du temps, cet a priori est celui qui est fourni par les ensembles météorologiques et nous utiliserons le formalisme bayésien régissant l'assimilation de données. L'objectif est ici de quantifier la probabilité qu'un profil atmosphérique prévu par le modèle de PNT a d'être proche de la réalité, en tirant parti des observations infrasonores et de sa probabilité a priori. Il faut alors définir les probabilités $\mathcal{P}(\mathbf{X})$ et $\mathcal{P}(\mathbf{Y}|\mathbf{X})$ de l'état a priori et de l'observation infrasonore conditionnée par cet état, respectivement.

La probabilité a priori d'un état est estimée en utilisant l'écart de l'état testé à l'ébauche du modèle (cf le chapitre 2). Cette ébauche est régie par la matrice de covariance d'erreurs \mathbf{B} et estimée par le moyen de l'assimilation d'ensemble. Classiquement, on suppose que

cette probabilité a priori suit une loi gaussienne dont l'espérance est égale à la moyenne de l'ensemble et de matrice de covariances \mathbf{B} . Cependant, cette définition fournit une probabilité a priori à chaque membre de l'ensemble à l'échelle du globe et non en un point précis et pour certaines de ses variables. C'est pourquoi, ignorant le comportement de cette loi à l'échelle locale, nous faisons le choix de considérer une loi uniforme pour la probabilité a priori.

Concernant la probabilité $\mathcal{P}(\mathbf{Y}|\mathbf{X})$, on considère qu'elle dépend de l'écart entre le signal mesuré en station et celui obtenu en le simulant avec un modèle de propagation considérant l'état de l'atmosphère \mathbf{X} en entrée. On se propose de réutiliser le formalisme du chapitre précédent pour quantifier cet écart. Il sera représenté comme une somme pondérée des écarts obtenus sur les différentes quantités d'intérêt des détections. La formulation de la probabilité de simuler fidèlement les observations infrasons et sa mise en application pour restituer des profils de température et de vent sur une base de cas réels associés à des éruptions volcaniques ont donné lieu à l'article *Bayesian selection of atmospheric profiles from an ensemble data assimilation system using infrasonic observations of a May 2016 Mount Etna eruptions* (Vanderbecken et al., 2020) publié au *Journal of Geophysical Research : Atmospheres* et présenté dans ce chapitre.

Ce chapitre est structuré en trois parties. La première partie présente un résumé en français de l'article ainsi que les résultats d'études préliminaires. La seconde partie contient l'article publié. Un ensemble d'études complémentaires est décrit et commenté dans la troisième partie.

5.1 Résumé et études préliminaires

5.1.1 Résumé

Dans cet article, la propagation des infrasons est considérée comme un moyen de fournir dans la moyenne atmosphère des informations utiles à un modèle de PNT. Pour exploiter cette information, une méthode bayésienne est proposée. Cette méthode est basée sur la comparaison de certaines observations et de leurs estimations par un modèle numérique de simulation de la propagation des infrasons. Cette méthode sélectionne l'état de l'atmosphère le plus probable parmi un ensemble d'analyses météorologiques. Cette sélection se fait pour des profils verticaux de température et de vitesse du vent zonal et du vent méridien. Les observations infrasonores, ici réduites à deux quantités d'intérêt que sont le back-azimut et la vitesse de passage, sont simulées à partir d'un code de tracé de rayons en coordonnées sphériques. Le milieu de propagation est spécifié à l'aide de profils atmosphériques issus du système d'assimilation d'ensemble de Météo-France. La méthode a été appliquée à une base de 6000 détections associées à une activité éruptive du Mont Etna en mai 2016. La méthode bayésienne permet d'identifier au sein de l'ensemble, les membres les plus pertinents vis-à-vis des observations infrasons, et de fournir ainsi une information utile dans la moyenne atmosphère. Lorsque les profils de vent et de température permettent d'utiliser l'approximation de célérité effective, il apparaît que les membres avec une forte vraisemblance possèdent des similarités en termes de profils de célérité effective mais présentent moins de ressemblances en termes de température et de vent. On peut en conclure que, pour bien caractériser l'information apportée sur la température et sur le vent par les infrasons, d'autres quantités d'intérêt de ces signaux apparaissent nécessaires (e.g. temps de trajet, amplitude, ...). On montre également que la probabilité des profils sélectionnés est sensible à l'adéquation temporelle entre les champs météorologiques et la date des détections. Elle est également sensible aux conditions de surface au niveau de la station réceptrice.

5.1.2 Points techniques révélés par l'étude

Les paragraphes suivants décrivent certaines difficultés rencontrées lors de l'étude qui ont conduit à certains choix techniques présentés dans l'article. Afin de ne pas alourdir le propos de détails techniques, ces choix n'ont pas été explicités dans l'article, mais ils méritent d'être présentés ici.

Le choix de la source et des stations

Durant la période éruptive de mai 2016, les conditions atmosphériques étaient favorables à la propagation au sein de guides d'ondes d'infrasons vers l'Ouest et le Nord-Ouest comme le montre la figure 5.1. Ces guides d'ondes sont caractérisés par un rapport entre la célérité effective à 50 km d'altitude et celle en surface supérieur à l'unité. Au début de cette étude, seules les stations de l'OHP et IS48 du SSI ont été identifiées comme stations ayant détecté les éruptions qui se sont produites au Mont Etna. C'est pourquoi seules les données enregistrées par ces stations ont été retenues pour l'étude. Par la suite, des études complémentaires ont montré que d'autres stations, plus lointaines, avaient également enregistré du signal en provenance du Mont Etna, dont certaines situées à plusieurs milliers de kilomètres. Cela se comprend assez bien, car les conditions météorologiques offraient des guides d'ondes stables dans la moyenne atmosphère dans de nombreuses directions comme le montre la figure 5.1. Les volcans sont donc des sources impactant de nombreuses stations en simultané et donc de grand intérêt car elles peuvent fournir des informations sur la circulation globale de la moyenne atmosphère aux modèles de PNT.

La propagation entre le Mont Etna et IS48 s'effectue principalement dans la direction zonale et la distance à parcourir est de 550 km. La propagation entre le Mont Etna et l'OHP est principalement dans la direction méridienne et d'une distance supérieure à 1000 km. Par la suite, on considérera que les profils atmosphériques sont invariants le long du trajet des ondes. Dans la mesure où cette hypothèse d'invariance permet de réduire le coût de calcul, il est possible de tester différentes stratégies pour sélectionner les « meilleurs » états atmosphériques, ce qui est intéressant dans le cadre d'une étude de faisabilité. Par contre, cette hypothèse exclut les stations qui seraient trop éloignées de la source, comme c'est le cas pour la station OHP. À terme, cette hypothèse pourra être relaxée.

Le choix du point d'extraction des profils

Dans l'article, on suppose que les caractéristiques de l'atmosphère n'évoluent pas entre la source des infrasons et la station réceptrice. Le milieu n'est donc défini que par les profils verticaux de température, de vent zonal, de vent méridien, de pression et de masse volumique extraits en un unique point entre la source et la station. Généralement, le point d'extraction est colocalisé soit avec la source, soit avec la station réceptrice. Le code de tracé de rayons GeoAc, utilisé pour cette étude, définit le milieu à partir d'une altitude zéro et considère qu'il n'y a pas d'orographie. Le choix de prendre l'altitude de la source ou de la station comme altitude zéro, tronque d'autant la troposphère, ce qui rapproche la moyenne atmosphère de la surface. Comme la moyenne atmosphère est à des altitudes plus basses, les ondes parcourent moins de distance avant d'être réfractées, ce qui modifie la distance entre les réflexions sur la surface. On définit ici l'altitude zéro comme étant celle du niveau de la mer pour ne pas tronquer la troposphère. Cela impose d'extrapoler les profils jusqu'au niveau de la mer.

Le choix du type d'extrapolation n'est pas trivial et elle peut avoir des conséquences sur la simulation de la propagation. Les profils de vitesse du son c , de vent zonal U et de

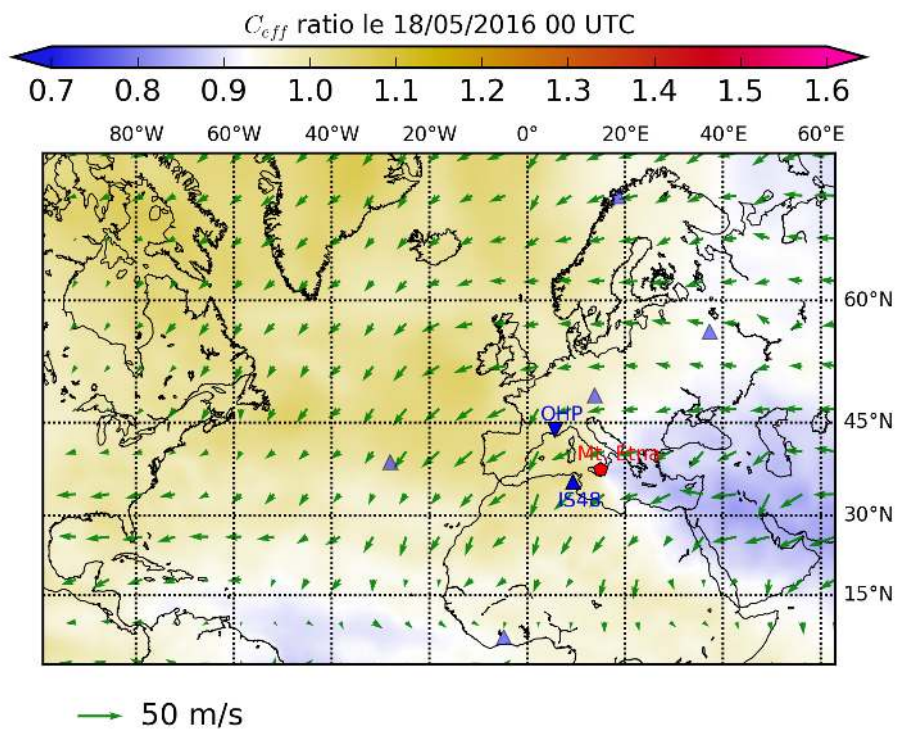


Figure 5.1 : Rapport de célérité effective entre la surface et 50 km d'altitude obtenu le 18 mai 2016. Le vent est projeté sur les grands cercles passant par le Mont Etna (pentagone rouge). Les champs météorologiques proviennent de l'analyse du réseau de 00 UTC du CEPMMT. Les flèches représentent le vent à 50 km. Les triangles bleus désignent des stations infrasons du SSI et de l'OHP.

vent méridional V représentés sur les figures 5.2 (a,b,c) sont extraits du membre #19 de l'AEARP du 18 mai 2016 à la verticale du Mont Etna. La résolution horizontale de l'AEARP fait que les profils considèrent la surface à une altitude de 163 m, il faudra donc les extrapoler jusqu'au niveau de la mer. Comme le modèle de propagation GeoAc ne fournit l'impact des rayons qu'au niveau de son altitude zéro, on a extrapolé l'atmosphère de manière homogène (i.e sans vent et isotherme), qui ne modifie pas la trajectoire des rayons la traversant. Ainsi, si l'altitude réelle de la station se trouve dans la couche extrapolée, il est facile à l'aide de règles de trigonométrie d'identifier les rayons propres et les quantités d'intérêt qu'ils simulent. La station IS48 est située à environ 1000 m d'altitude, soit largement au-dessus de la couche extrapolée initiale de 163 m. On a étendu la couche homogène pour inclure l'altitude de la station IS48. L'atmosphère devient donc homogène en dessous de l'altitude de la station (c'est pourquoi les profils des figures 5.2 (a,b,c) possèdent une couche homogène de 1000 m).

Les trajectoires des rayons lancés depuis le Mont Etna et présentées sur la figure 5.2 (e) montrent que cette couche isotherme conduit à une réfraction des rayons à environ 1000 m. Les rayons ayant une faible incidence n'atteignent jamais le niveau de la mer et donc ne contribuent pas à l'empreinte au sol. Ainsi, les rayons propres extraits (atteignant la station, cf chapitre 4) sont uniquement ceux qui correspondent aux fortes incidences et donc à des vitesses de passage élevées. Dans ces conditions, il est impossible d'expliquer par la simulation les phases avec de faibles vitesses de passage. Pour pallier ce problème, on a alors fait décroître la température selon le gradient vertical standard (i.e. en suivant la pente $-0,0065$ K/m) pour limiter le changement de gradient de température bloquant les rayons avec de faibles incidences. Toutefois les résultats ont montré que le problème n'était pas totalement résolu car il restait une discontinuité dans les profils de vent. Nous avons donc fait le choix d'extraire les champs en un point du grand cercle reliant le Mont Etna à la station, situé au niveau de la mer (≈ 10 m) et proche de la source pour limiter au maximum l'extrapolation verticale.

Données relatives à la source et choix de la période

L'information sur la source est constituée d'un ensemble de bulletins publiés toutes les minutes¹. Ces données ont été mises à la disposition par l'Université de Florence (UNIFI), dans le cadre du projet européen ARISE (Blanc *et al.*, 2018). Ces bulletins contiennent deux types d'informations :

- un nombre de détections cohérentes dont la durée est de 5 s ;
- une amplitude RMS² moyenne des détections en Pascal.

Ces informations permettent d'identifier les périodes d'intensification de l'activité du Mont Etna, c'est-à-dire les périodes où les amplitudes RMS sont supérieures à 1 Pa. La figure 5.3 montre la distribution des amplitudes des bulletins enregistrés entre novembre 2015 et janvier 2017.

On remarque trois périodes où la proportion de bulletins ayant une amplitude RMS moyennée supérieure à 1 Pa est plus forte : (1) fin novembre - début décembre 2015 ; (2) mi-mai à fin mai 2016 ; (3) mi-août 2016. D'après les résultats de Cannata *et al.* (2017), ces périodes correspondent toutes à des périodes d'éruptions stromboliennes³ dans un ou plusieurs des cratères du Mont Etna. Les deux premières périodes amènent de plus fortes détections. Il s'avère que durant les deux premières périodes, des jets de lave

1. <http://lgs.geo.unifi.it/index.php/monitoring/volcanoes/etna>

2. Root Mean Square

3. Les éruptions stromboliennes sont des poches de gaz présentes dans le magma qui explosent en surface.

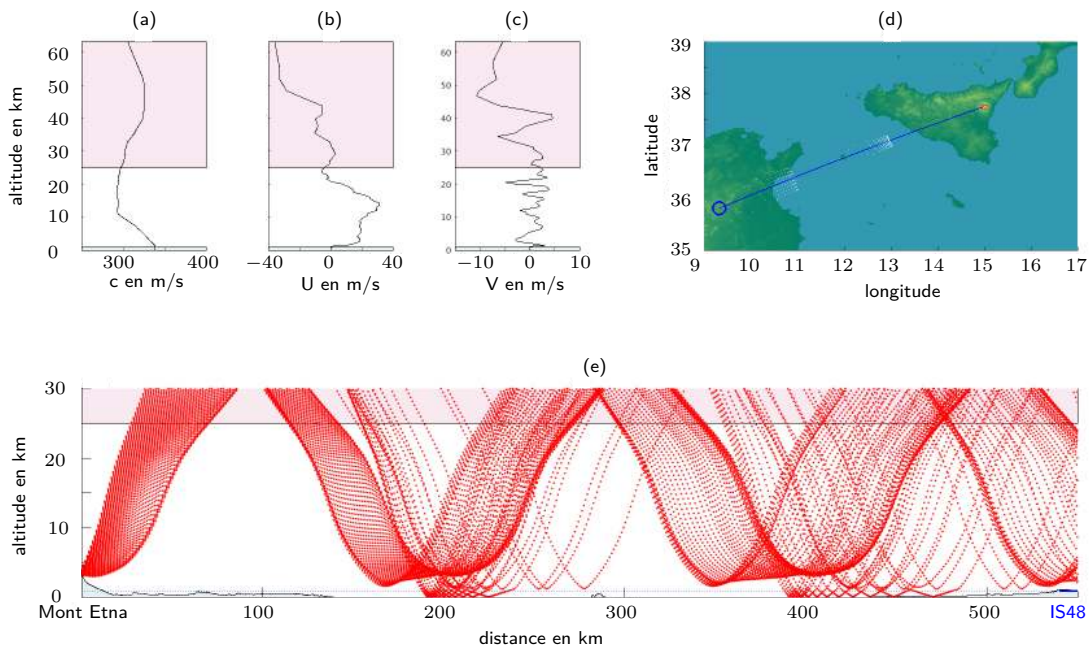


Figure 5.2 : Exemple de propagation entre le Mont Etna et IS48. (a,b,c) profils de vitesse du son, de vent zonal et de vent méridien du 18/05/2016 fournis par le membre #19 analysé de 00 UTC de l'AEARP. (d) Distribution des points d'impact des rayons (points blancs) issus du Mont Etna (étoile rouge) en direction de IS48 (rond bleu). Le grand cercle reliant le Mont Etna à IS48 est représenté par la ligne bleue. (e) trajectoires des rayons (rouge) dans le plan de propagation. L'orographie réelle est en gris. La ligne en pointillés indique la hauteur de la couche homogène.

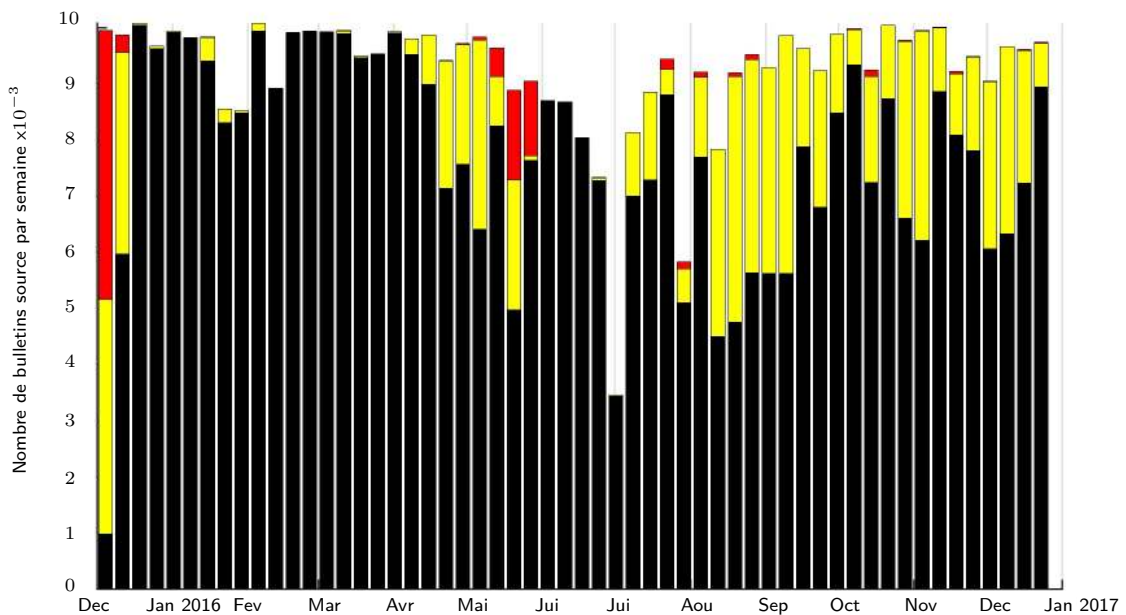


Figure 5.3 : Amplitudes RMS des bulletins de l'antenne infrason placée sur le Mont Etna au cours de l'année 2016. Les bulletins ayant vu aucune ou une seule détection de 5 s sont en noir. Les bulletins avec une amplitude moyenne RMS comprise entre 0 et 1 Pa sont en jaune. Les bulletins avec une amplitude RMS supérieure à 1 sont en rouge. La largeur de chacune des barres correspond à une semaine.

ont également eu lieu. Ces deux types d'éruptions sont effectivement à l'origine d'ondes infrasonores (Weill *et al.*, 1992 ; Vergnolle *et al.*, 1996), et devraient se traduire par des bulletins dont l'amplitude RMS est au-dessus de 1 Pa et des détections par IS48 (Tailpied *et al.*, 2013).

Les volcans peuvent être vus comme des sources ponctuelles et impulsionnelles du point de vue des infrasons. Un des problèmes liés à ce type de bulletin est qu'il n'est pas possible d'identifier précisément la séquence d'impulsions infrasonores générée par ces phénomènes éruptifs. En conséquence, lorsque le bulletin de l'antenne placée sur le Mont Etna présente une amplitude RMS supérieure à 1 Pa, nous pouvons dire qu'au moins une impulsion a lieu au sein de la minute de validité du bulletin. On ne peut que définir à une minute près l'instant t_0 de l'impulsion. Ce temps t_0 sera par la suite sera la date du bulletin. Comme un bulletin couvre une minute, il peut y avoir alors plusieurs éruptions qui auront la même valeur de t_0 .

Nous cherchons maintenant, parmi l'ensemble des bulletins de l'antenne située sur le Mont Etna, ceux caractérisant des éruptions ayant pu être détectées par la station infrasons Tunisienne IS48 du SSI. Pour simplifier le problème, seuls les bulletins associés à la fois à des conditions de propagation favorables et à des amplitudes RMS supérieures à 1 Pa, ont été retenus. En confrontant ces bulletins de l'antenne située sur le Mont Etna avec une première estimation grossière du signal cohérent enregistré à la station IS48, on peut associer 586 bulletins du Mont Etna à 586 détections à IS48. La majeure partie des associations ayant lieu pour la période éruptive de mai 2016.

À partir des 586 valeurs de t_0 approximatives données par les bulletins du Mont Etna retenus, une analyse PMCC plus fine a été faite sur les signaux enregistrés à IS48. Cette analyse produit un ensemble de 586 bulletins de détections à IS48 couvrant pour certains plus de 40 minutes et avec parfois des recouvrements. Une fois les détections présentes dans plusieurs bulletins éliminées, ces bulletins regroupent près de 6000 détections de phases cohérentes de la station IS48. Les temps de trajet des détections sont calculés alors vis-à-vis du temps t_0 que l'on a utilisé pour l'édition du bulletin à IS48. Les temps de trajet observés en ne considérant que ces valeurs de t_0 peuvent être supérieurs à l'heure, valeurs qui sont très supérieures aux temps de trajet simulés.

5.2 Article : *Bayesian selection of atmospheric profiles from an ensemble data assimilation system using infrasonic observations of May 2016 Mount Etna eruptions.* **publié en 2020 au Journal of Geophysical Research : Atmospheres**

5.2.1 Abstract

Observed infrasound arrivals are considered as a way to provide additional knowledge of the middle atmosphere for numerical weather prediction (NWP) models. To do so, a Bayesian approach, based on the discrepancies between predicted and observed infrasound arrival characteristics, is adopted for selecting the most likely atmospheric states, in terms of temperature and wind velocity profiles, among an ensemble of NWP model analyses. The predicted characteristics, in terms of trace velocity and the back azimuth angle, are computed using a three dimensional ray-tracing model and atmospheric profiles extracted once a day from ensemble of analyses produced by the Météo-France global NWP model. The performance of the method is demonstrated using a set of thousands of volcanic

eruptions that correspond to an upsurge in volcanic activity of Mount Etna (Sicily), during May, 2016. It is shown that the Bayesian approach allows the identification of the most likely members and, hence provides additional knowledge on the atmosphere truly probed by the infrasound. When the atmospheric temperature and wind profiles are reasonably meeting the effective sound speed assumption requirements, the members retrieved are similar in terms of effective sound speed but not in terms of wind and temperature. For improving both temperature and wind velocity profile selection, additional infrasound arrivals characteristics are then required (e.g. travel time, amplitude. . .). The confidence on the retrieved profiles is also sensitive to the time of validity of the meteorological analyses and on the definition of the surface conditions involved in the estimation of the predicted trace velocities.

5.2.2 Introduction

Infrasonic waves, which are usually defined as sound waves with frequencies ranging from 0.02 Hz to about 20 Hz, are emitted worldwide by various natural sources (Campus & Christie, 2010). These waves can propagate efficiently over long distances in the atmosphere, within multiple ducts that are produced by the wind and temperature vertical variations. While the most stable duct is potentially in the upper atmosphere (Gossard & Hooke, 1975), as a result of the steep increase in temperature above 90 km, the decrease in mean density produces substantial absorption coefficients (Sutherland & Bass, 2004). For this reason, the most efficient ducts, though not as stable, are often in the middle atmosphere, above the tropopause located between 10 and 17 km, and below the mesopause at 90 km. Although affected by atmospheric waves, the steadiest waveguide in the mid-latitudes appears to be located around the stratopause (De Groot-Hedlin & Hedlin, 2014 ; Hedlin & Drob, 2014 ; Drob *et al.*, 2013, among others) and characterized by a strong zonal stratospheric jet stream. At these altitudes the absorption coefficient is small and thus, it can be neglected in most practical applications.

The usual approach to calculate infrasound propagation consists of solving the wave equation in a given background atmospheric state that varies with altitude and horizontal distance. The atmospheric conditions are usually issued from atmospheric climate reanalyses, such as ERA-Interim (Dee *et al.*, 2011), upper atmospheric empirical models for winds (Drob *et al.*, 2008, 2015, among others) and temperature (Hedin, 1987, 1991 ; Picone *et al.*, 2002, among others), or reconstructions based on multiple data (Schwaiger *et al.*, 2019). This approach, however, fails in predicting important arrivals (Bertin *et al.*, 2014), essentially because the changing atmospheric conditions involve varying uncertainties (both in time and space) that cannot be captured using a single atmospheric state. On the other hand, in recent years, operational weather services, such as the European Centre for Medium-Range Weather Forecasts (ECMWF) or Météo-France (MF), extended their global Numerical Weather Prediction (NWP) up to the mesosphere (e.g., 0.1 hPa), often capturing atmospheric features that have a great impact on infrasound propagation. Ensemble analysis can be obtained from NWP, by slightly perturbing the model and the initial conditions, thus producing a set (or ensemble) of atmospheric states. In this case, instead of propagating infrasound in a single atmospheric state, multiple simulations can be conducted to account for the impact of atmospheric uncertainty in the infrasound characteristics, which is a form of Monte Carlo analysis.

A recent trend in the infrasound research community relies on perturbing the underlying atmospheric states, either by superimposing ad hoc fluctuations (Chunchuzov *et al.*, 2015, 2011 ; Assink *et al.*, 2014, 2013, among others), or by using an off-line gravity wave

model to generate the fluctuations from atmospheric states (Lalande & Waxler, 2016 ; Drob *et al.*, 2013). While these approaches are receiving increasing attention within the infrasound research community, the atmospheric relevance of the numerically obtained perturbations is unclear for two reasons. First, the disturbance is often described by superposition of a fairly small number of functions. These functions are obtained by performing a principal component analysis on the data (Assink *et al.*, 2013), by using ensemble meteorological data (Smets *et al.*, 2015), or simply by specifying a truncated set of basis functions (Bertin *et al.*, 2014). While these functions should cover the overall atmospheric variability, both in space and time, there remain the issue that the resulting profiles cannot exhibit structures which are not present in the functions, thereby reducing the dimensionality of the phase space. Second, in first attempts to incorporate physical laws, spectral schemes (Hines, 1997 ; Warner & McIntyre, 1996 ; Scinocca, 2003) have been used to describe the fluctuations from gravity waves energy spectra. These schemes consider that the entire spectrum of gravity waves is realized at each model grid point and at each time, somehow contradicting the observational fact that gravity waves often travel in narrow banded wave packets (Alexander, 2015 ; Plougonven *et al.*, 2013 ; Wright *et al.*, 2013 ; Hertzog *et al.*, 2012, 2008). Further, in this second strategy, it is difficult to separate epistemic uncertainty, due to lack of knowledge, from random variability. In NWP this second type of uncertainty can be reduced through adding observations during the data assimilation cycle.

The focus of this work is on selecting plausible atmospheric states, using discrepancies between predicted and observed infrasound measurements. Our approach differs from previous studies, as we focus specifically on obtaining a valid atmospheric state, i.e. with components (temperature, winds) that satisfy exactly the meteorological equations. The ensemble meteorological data issued from the MF ensemble data assimilation system (Berre *et al.*, 2015) is used as a set of realistic states that sample the current forecast and analysis uncertainties. A Bayesian approach already used at MF for the assimilation of radar reflectivities (Wattrelot *et al.*, 2014) and all-sky microwave radiance (Duruiseau *et al.*, 2019) has been chosen to measure the relevance of each member to predict infrasound observations. This study can be considered as a preliminary step towards the use of infrasound data for assimilation within NWP models in order to improve forecasts of the middle atmosphere.

The paper is organized as follows. Section 5.2.3 describes the infrasound propagation model and the atmospheric specifications. In section 5.2.4, the inverse problem is formulated using a Bayesian approach to select the best atmospheric specifications among the ensembles from the global NWP model of MF. The Bayesian approach is then applied in Section 5.2.5, using thousands of volcanic eruptions that correspond to an upsurge in volcanic activity of Mount Etna (Sicily), during May, 2016. It is found that these events can be used to improve our knowledge of the atmosphere in regions where NWP are currently lacking of observations, typically above 30 km (Polavarapu *et al.*, 2005). Section 5.2.6 discusses specific aspects of the results and their limitations.

5.2.3 Infrasound propagation modeling

A 3D geometric propagation technique

Direct Numerical Simulation (DNS) of infrasound propagation is computationally prohibitive (Marsden *et al.*, 2014 ; Sabatini *et al.*, 2016, 2019) and it is more a research tool than a brute-force solution to the Navier-Stokes equations. The wide range of scales in the atmosphere requires that care be taken in the numerical solution to avoid numerical instabilities (Millet *et al.*, 2007) and indeed, additional processing of the initial conditions

is generally necessary to obtain robust solutions (Sabatini *et al.*, 2016). In fact, the usual approach in studying infrasound propagation consists of solving the wave equation, using for this task numerical methods that are described in textbooks (Gossard & Hooke, 1975 ; Jensen *et al.*, 2011). In this paper, we rely on geometrical optics approximation and use the three dimensional linear ray-tracing software in spherical coordinate provided by GeoAc platform (Blom & Waxler, 2012 ; Blom, 2019) for computing the characteristics of the wavefront as it propagates away from a given source.

This method simulates the path of the normal to the wavefront, which is characterized by the eikonal vector, and the ray path length. The initial condition of the ray is specified by an elevation angle θ_0 and an azimuth angle ζ_0 . When rays are launched from a given source location with different initial conditions, the corresponding paths define the trajectories of the infrasonic waves, from which ground-based quantities can be extracted. These quantities are the travel time T_t , the trace velocity v , which is the horizontal component of the wavefront velocity, and the back azimuth ϕ . These quantities can be further compared to measurements when the rays reach an infrasound station. Three-dimensional linear ray tracing method requires atmospheric specifications in terms of temperature, wind vector components as functions of altitude, from ground level up to the top of the computational domain. These profiles can be extracted from NWP model outputs, which are generally confined to levels in the atmosphere where weather is most important, typically below 0.1 hPa. At mid-latitudes, the temperature gradient from winter pole to summer pole induces an easterly jet in the summer hemisphere and a westerly jet in the winter hemisphere, as a result of thermal wind balance, the maximum winds occurs near 55 km. This produces acoustic waveguides that are not well simulated by NWP models, mostly because of the lack of observations above 30 km. This uncertainty translates into varying travel times. Hence, to facilitate the separation of stratospheric and thermospheric phases, the profiles are extended for altitudes up to about 150 km, using the empirical models Mass Spectrometer-Incoherent Scatter (MSIS-00) model (Picone *et al.*, 2002) and Horizontal Wind Model (HWM07) (Drob *et al.*, 2008). The increase in temperature in the upper thermosphere provides a thermospheric arrival that can be clearly identified and compared to the observations.

Atmospheric profiles from NWP models

Operational NWP centers produce analyses and forecasts of meteorological fields. The most critical quantities are surface pressure, wind components, temperature, humidity, clouds and precipitation. Analyses are produced every 6 hours with global systems and forecasts are issued up to about a week (even though longer range forecasts are produced by atmosphere-ocean coupled systems). All atmospheric fields are provided on a grid associated with the horizontal and vertical discretization of prognostic dynamical equations solved by the NWP model. The vertical grid is generally a terrain following coordinate close to the ground that changes in a pressure coordinate at high levels (Simmons & Burridge, 1981). Atmospheric analyses are produced within data assimilation systems that combine optimally a short-range forecast with all available observations over a fixed time window (usually 6 hours around the analysis time). Global systems are strongly constrained by satellite observations, but conventional observations are still useful in particular to remove the biases in satellite measurements. As emphasized in the previous section, above 30 km, the number of observations is rather low to constrain the model state (channels from infra-red and micro-wave nadir sounders with broad weighting functions and bending angles from the radio-occultation of GNSS signals). The optimality of data assimilation systems strongly depends upon the specification of a priori and observation

errors (covariance matrixes). A priori errors are obtained from ensembles of short-range forecasts that are produced from ensembles of analyses where observations are randomly perturbed. For computational cost reasons, such ensembles are run at lower resolution than that of the NWP model producing the operational forecasts.

In the present study, the analyses from the ensemble data assimilation system of MF so-called AEARP (Assimilation d'Ensemble ARPEGE) will be considered. The data assimilation system is a 4 Dimensional-Variational data assimilation (4D-Var) run at a 40 km resolution ($T_L 479$ triangular linear spectral truncation) with a minimization performed at 134 km ($T_L 149$). The AEARP system has 25 members (Berre *et al.*, 2015). Surface conditions are defined from a soil vegetation model over land and by imposing the surface temperature from dedicated analysis over oceans. The vertical grid has 105 vertical levels between 10 m above the surface and pressure level 0.1 hPa (65 km in the mid-latitudes). The 105 hybrid-pressure levels of the ARPEGE models are not uniformly distributed (18 levels above 25 km and 28 levels under 2 km). The distribution of levels is done to optimize short-range tropospheric weather predictions (up to 4 days).

According to the assumptions made in the ARPEGE NWP model, the geopotential height Φ of the hybrid-pressure levels is obtained from the diagnostic hypsometric equation. The geopotential height at a half-level $i + 1/2$ of the hybrid coordinate can be expressed in terms of the value at the lower half-level $i - 1/2$, following the rule

$$\bar{\Phi}_{i+1/2} = \bar{\Phi}_{i-1/2} - r\tau_i \frac{\bar{p}_{i+1/2} - \bar{p}_{i-1/2}}{p_i}, \quad (5.1)$$

where $\tau_i = T_i(1 + 0.6078q_i)$ denotes the virtual temperature. In this equation q_i , T_i and p_i are respectively the specific humidity, pressure and dry air temperature, at level i . The air gas constant $r = 287.058 \text{ mol.kg}^{-1}$ is the one for dry air. The geopotential at level i , which is to be obtained from $\Phi_i = (\bar{\Phi}_{i-1/2} + \bar{\Phi}_{i+1/2})/2$, is initiated using the surface value and $\bar{\Phi}_{1/2} = z_g g$, where z_g is the ground altitude, and $g = 9.81 \text{ m.s}^{-2}$ is the Earth gravitational acceleration norm assumed constant. ARPEGE uses the ideal gas law for the moist air. As a consequence, the adiabatic sound speed is explicitly defined by $c = \sqrt{\gamma r \tau}$, where the isentropic expansion factor γ is set to 1.4. In the following, we will use c instead of the temperature.

5.2.4 Infrasound-based atmospheric state selection

The propagation model can be viewed as a function that maps an input \mathbf{X} to an output \mathbf{Y} , where \mathbf{X} is the atmospheric state. To determine \mathbf{Y} for a given atmospheric profile, rays are launched uniformly within a cone sampled with an elevation angle step $\delta\theta_0$ and an azimuth step $\delta\zeta_0$. Among all the rays launched, the M rays that link the infrasound source to a receiving station are named "eigenrays". Only the model outputs given by the M eigenrays are considered as suitable quantities for the Bayesian retrieval method as in Assink *et al.* (2013). The model output is as a random vector whose components are, respectively, the travel time $Y_1(\mathbf{X}) = t$, the trace velocity $Y_2(\mathbf{X}) = v$ and the back azimuth $Y_3(\mathbf{X}) = \phi$. The atmospheric conditions are described by four vertical profiles that specify the adiabatic sound speed, the zonal and meridional components of wind, and the pressure as functions of altitude. For the sake of simplicity, the profiles are extracted from the AEARP analyses at a given location along the source-receiver path.

If all atmospheric states are jointly uniformly distributed then the prior distribution of these states is constant over the entire parameter space and the Bayes theorem (Aitkin,

2010 ; Tarantola, 2005) can be written as

$$P(\mathbf{X}|\hat{\mathbf{Y}}) = \frac{P(\hat{\mathbf{Y}}|\mathbf{X})P(\mathbf{X})}{P(\hat{\mathbf{Y}})} \propto P(\hat{\mathbf{Y}}|\mathbf{X}), \quad (5.2)$$

where the righthand side $P(\hat{\mathbf{Y}}|\mathbf{X})$ is the likelihood $\mathcal{L}(\mathbf{X}|\hat{\mathbf{Y}})$. Using a Gaussian distribution for the model output, i.e.

$$\ln \mathcal{L}(\mathbf{X}|\hat{\mathbf{Y}}) = -\frac{1}{2}(\mathbf{Y}(\mathbf{X}) - \hat{\mathbf{Y}})^T \mathbf{R}^{-1}(\mathbf{Y}(\mathbf{X}) - \hat{\mathbf{Y}}) + A, \quad (5.3)$$

where \mathbf{R} is the covariance matrix of the observation error, and A is an undetermined normalization constant. The extra-diagonal elements of matrix \mathbf{R} are neglected, which is a common assumption for weakly correlated outputs (Assink et al., 2014). The diagonal of \mathbf{R} is composed of the variances σ_i^2 of the i -th random variable Y_i . In this study we will focus on the ability of the likelihood to rank the members of an atmospheric ensemble, according to their efficiency to explain observations at a given infrasound station. Since the likelihood ratio is the critical information for this task, we arbitrarily set the constant A to zero. Further, an optimal state can be obtained through selecting that which maximizes the likelihood function (or its logarithm), whatever the value of the constant. This maximum will be denoted as \mathcal{L}_{best} thereafter.

In this framework, the model output \mathbf{Y} and the observation vector $\hat{\mathbf{Y}}$ must have the same number of components. However, in many situations the number of eigenrays M connecting the source and the receiver for a given atmospheric state differs from an atmospheric state to another. If the atmospheric state provides no eigenrays (i.e. $M = 0$), then the likelihood is clipped to zero. Otherwise, we select an “optimal” eigenray among the computed eigenrays, for each atmospheric state, as being the closest on to the observations. Equation (3) can be then rewritten as :

$$\ln \mathcal{L}(\mathbf{X}|\hat{\mathbf{Y}}) = - \min_{j \in [1:M]} \left(\sum_{i=1}^3 \frac{(Y_{i,j} - \hat{Y}_i)^2}{2\sigma_i^2} \right), \quad (5.4)$$

where i is associated to a component of the observation vector $\hat{\mathbf{Y}}$ and j corresponds to a specific eigenray. The standard deviation of the travel time is fixed to $\sigma_1 = 10$ s. For the two other components of \mathbf{Y} (back azimuth angle, trace velocity), assuming that the modeling error is zero, the standard deviations can be deduced from the array geometry. For the infrasound station considered in this study, following Szuberla et Olson (2004) we set $\sigma_2 = 5 \text{ m}\cdot\text{s}^{-1}$ and $\sigma_3 = 0.5$ deg. To avoid misinterpretations, only the atmospheric states for which the likelihood is larger than $e^{-1/2} \simeq 0.6$ will be considered in the following. This criterion is used for testing whether the model output is sufficiently close to the observation or, in mathematical terms, if the model output lies in the confidence interval.

It is worthwhile to note that an eigenray is valid only if it reaches a small neighborhood of the receiver, which is often specified as a disk of given diameter ϵ . In fact the number of eigenrays that reach this region is not only driven by the atmospheric state, it also depends on the number of rays that are launched regarding the disk diameter, as well as other numerical parameters. In fact little is known about how small this diameter might be for given values of the discretization process. In this work, we use $\epsilon = 4$ km. The angular discretization for the azimuths and elevations is fixed to $\delta\theta = \delta\zeta_0 = 0.1$ deg and the azimuthal and elevation spread are fixed to 7 deg and 50 deg, respectively. These choices are justified in 5.3.1.

5.2.5 Application to the detections of Mount Etna

Description of raw data

In this section, the Bayesian method described in the previous section is applied using a sequence of eruptions emitted from Mount Etna, which is a volcano located along the eastern coast of Sicily (Italy), as shown in Figure 5.4. As reported by Cannata *et al.* (2017) and Edwards *et al.* (2018), an upsurge of its activity occurred during May 2016. In this study, we focus on the period between the 15th (00 UTC) and the 26th (00 UTC) of May 2016. This period is characterized by a first episode of strombolian activity at the North East Crater (NEC), on the 17th of May 2016, followed by a paroxysm at the Voragine crater (VOR) on the middle of the 18th of May. On early 19th and 21th, two new paroxysms occurred at VOR. This activity remained on the New South East Crater (NSEC) and NEC until the 26th of May.

Several infrasound signals were recorded by infrasound station IS48, which is one of the 51 currently certified infrasound arrays in the International Monitoring System (IMS) of the Comprehensive Nuclear Test-Ban Treaty Organization (CTBTO). This station is located in the northern part of Tunisia, as shown in Figure 5.4. The characteristics of this station, which is located 550 km away from the source, are given by Assink *et al.* (2014). For each infrasound event (hereafter referred to as a detection), an infrasound bulletin is published, using the Progressive Multi-Channel Correlation (PMCC) method (Cansi, 1995). These bulletins provide the back azimuth, the trace velocity, the starting time of the coherent signal, and the ending time of the coherent signal, leading to an estimate of the signal duration. The characteristics of the source are given by a near-field infrasound array, located at about 5 km away from the Mount Etna peak, and maintained by the University of Florence (UNIFI). Further in this study, it is assumed that the acoustic source is at the same location as that of the Mount Etna peak. The Mount Etna peak is identified in Figure 1 (a) by the red square. A near-field infrasound bulletin is published every minute and provides the averaged root mean square (RMS) amplitude of the coherent infrasound signal recorded within the minute. Since the acoustic source signal was not available, it was not possible to obtain an accurate value of the travel time, which leads us to withdraw this component in the observation vector.

Database of infrasound observations

This study focuses on infrasound waves that are refracted between the upper stratosphere and the lower mesosphere (e. g. in between 25 km and 65 km of altitude), for which the trace velocity at IS48 is smaller than $360 \text{ m}\cdot\text{s}^{-1}$. Thereafter those detections will be regarded as middle atmospheric arrivals. Since the atmosphere may skew the incoming wave from the source-receiver path, we restricted the analysis to waves that reach IS48 with a back azimuth of $64.9 \pm 15 \text{ deg N}$, where the center of the interval is the Mount Etna back azimuth. These two filters select only detections at IS48 that could be related to middle atmospheric arrivals from Mount Etna. However, other infrasound sources, such as the Stromboli volcano or the nearly constant Mediterranean surf noise, generate infrasound that meets also these two conditions. To avoid adding such events to the final database, and ensuring that remaining detections at IS48 are issued from the Mount Etna, two additional criteria have been used.

First, we have to ensure that the propagation between Mount Etna and IS48 station was supported by the meteorological conditions. Using ray theory and the effective sound speed approximation, the refraction of rays launched with a given elevation angle θ_0 occurs when the effective sound speed c_e becomes greater or equal to the value at ground level,

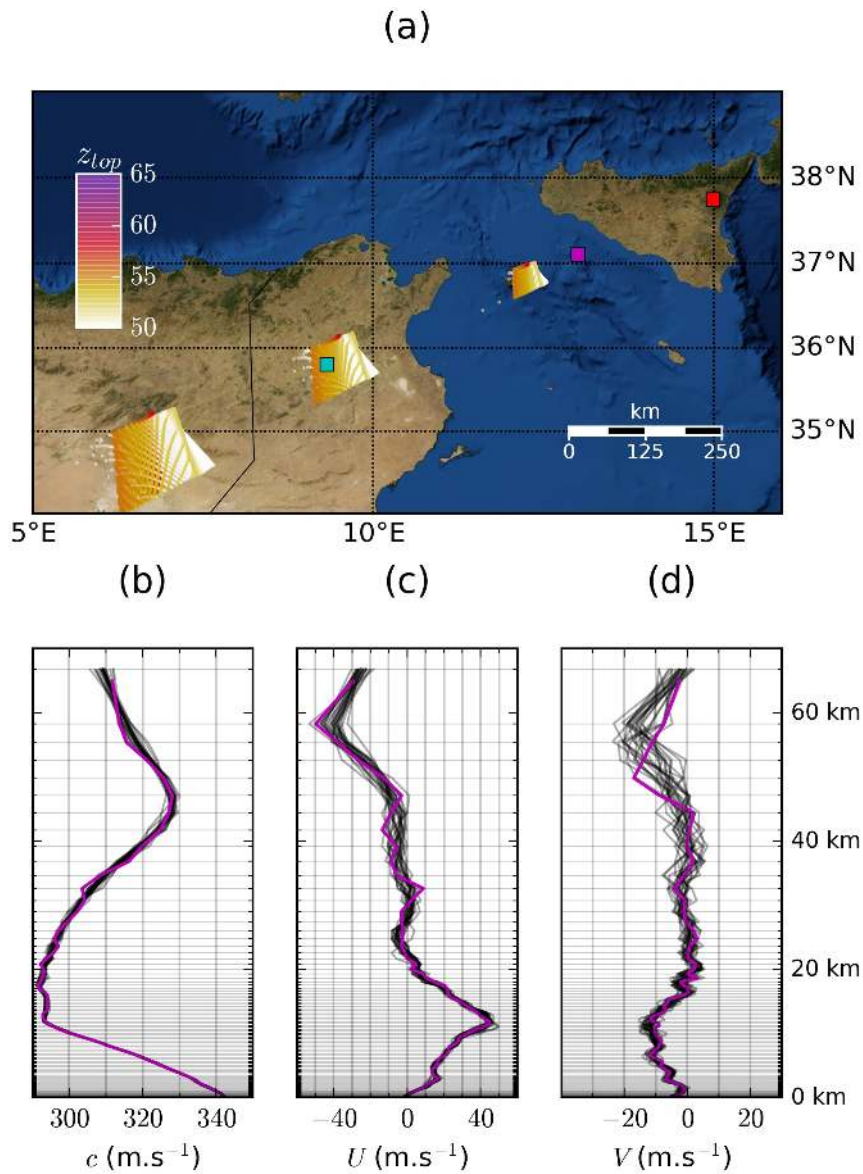


Figure 5.4 : (a) Satellite view of the domain of study. Mount Etna and IS48 infrasound station are represented by the red square and the cyan square, respectively. The magenta square represents the point of extraction of atmospheric profiles shown in (b), (c), and (d). The colored dots represent the simulated ground to ground infrasound propagation on May 18 2016 using member #19 from AEARP. The colorgrid displays the altitude of refraction z_{top} . The profiles in magenta are analysis from member #19 in terms of sound speed (b), zonal wind (c) and meridional wind (d). The profiles of the other members of the AEARP are displayed in light gray. The interval between vertical lines is 10 m.s⁻¹. The horizontal grid represents the 105 levels of member #19 altitudes.

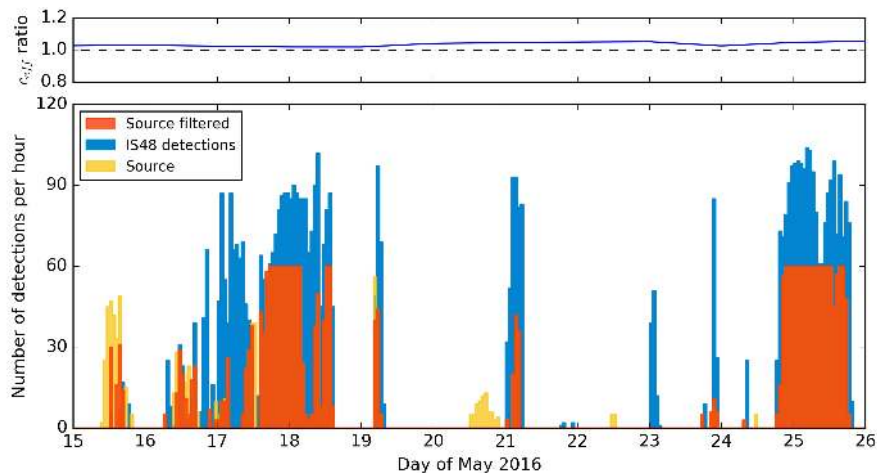


Figure 5.5 : Infrasound database. Selected events from May 15 2016 to May 26 2016 at Mount Etna (source) and IS48 station (detections). (Upper panel) The solid line depicts the effective sound speed ratio (first criterion, see Section 5.2.5) and the dotted line the value 1 against which the effective sound speed ratio is compared. (Lower panel) The yellow bars represent the hourly number of near field infrasound events with an averaged RMS amplitude over 1 Pa recorded by the Mount Etna infrasound array. The orange bars represent the fraction of the nearfield events obtained satisfying the second criterion. The blue bars show the hourly number of far field infrasound detections recorded by IS48 array.

c_{e0} , divided by $\cos\theta_0$. Then, whatever the initial θ_0 , a necessary condition for propagation in the middle atmosphere is that $c_e(z) \geq c_{e0}$ for $25 \leq z \leq 65$ km. This makes our first criterion. The maximum sound speed ratio between 25 and 65 km is shown in the upper panel of Figure 5.5. The atmospheric states having their maximal sound speed ratio greater than unity are referred to as ‘supportive’ atmospheric conditions in the following.

Second, we restrict the study to detections at IS48 that are preceded by a near-field event. We retain near field infrasound events associated to bulletins of the near field infrasound array with their averaged RMS amplitude over a minute larger than 1 Pa. These detections are represented in yellow in Figure 5.5. IS48 detections and near field events are associated when the estimated observed travel time is between 1528 and 2619 s. The near field events associated with IS48 detections are represented in orange in Figure 5.5. This second criterion ensures that the near-field events can be unambiguously associated with the IS48 detections. Finally according to these conditions, the final database is composed of about 6000 detections. The resulting associated IS48 detections are displayed in blue in Figure 5.5. We mention that the number of remaining detections is in excellent agreement with the Mount Etna events database described by Edwards *et al.* (2018).

For the ray-tracing simulations, the atmospheric profiles are extracted at the location 13.0 deg E and 37.1 deg N (magenta square in Figure 1 (a)), which lies along the source-receiver path, 190 km away from the Mount Etna. Typical vertical profiles of sound speed, zonal and meridional wind velocities are given in Figures 1 (b), (c) and (d), respectively. Each black profile represents a member of the AEARP analysis ensemble on the 18 of May 2016 at 00 UTC. One can notice that the differences between wind profiles increase with altitude, especially above 30 km. Profiles in magenta, that are extracted from member #19, are used to simulate infrasound propagation from Mount Etna in the azimuth direction of IS48 (Figure 1 (a)). The IS48 station will actually be affected by the second bounce of stratospheric arrivals according to Figure 5.4 (a). We mention

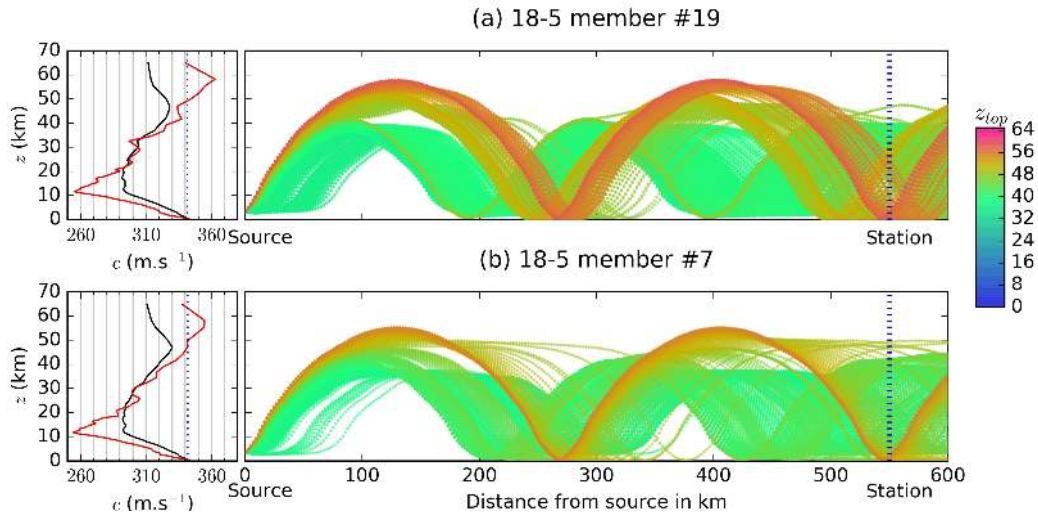


Figure 5.6 : Rays computed using atmospheric profiles from AEARP ensemble analysis on the 18th of May, 2016. Upper panels : member #19 adiabatic sound speed profile in black, effective sound speed in red and ray trajectories colored by turning height ; lower panel : member #7. adiabatic sound speed profile in black, effective sound speed in red and ray trajectories colored by altitude of refraction. The rays are launched for a fixed azimuth of 248.5 deg. The vertical dashed line represents the location of the infrasound station (right panels) or the effective sound speed at ground (left panels).

that this behavior is typical of the period considered in this study. The color of the dots displays the altitude of refraction z_{top} of the rays representing the altitude of refraction of infrasound. According to the color of the dots that reach the station, member #19 leads to three different stratospheric arrivals. One with $z_{top} \leq 50$ km represented by the white dots, one around with $z_{top} \approx 53$ km represented by the yellow ones and a third one with $z_{top} \geq 55$ km represented by dots that range from orange to red. This shows that a single member can lead to several arrivals according to the value of z_{top} . The advantage of using an ensemble is that one can test if different members share the same arrivals. Figure 5.6 shows raypaths obtained using two members (# 19 and #7) of the AEARP ensemble analysis, on the 18th of May, 2016. Their respective effective sound speed and adiabatic sound speed are also displayed. While the members #19 and #7 both lead to several stratospheric arrivals in the vicinity of the station mentioned previously, their corresponding raypaths are significantly different. Such differences are directly related to the increasing difference in member's atmospheric profiles above 30 km. This reflect the larger uncertainties we have at those altitudes, coming from reminiscent of the smaller amount of observations used in the analyses.

Figure 5.7 displays the day-to-day variability of the standard deviation profiles for sound speed, zonal and meridional winds. This dispersion accounts for uncertainties both in initial conditions and observations. Although the latter are larger in the stratosphere and the lower mesosphere, there may sometimes be situations for which the statistical dispersion is large below 10 km, belying the first impression of a glance. For the database considered in this study, this situation occurs in particular on the 19th of May.

Numerical results

In this section we apply the Bayesian method described in Section 5.2.4, using the infrasound database defined in Section 5.2.5 together with 25 atmospheric profiles. These profiles are extracted from the AEARP ensemble analysis at the location 13 deg E and 37.1 deg N and time 00 UTC. A ray tracing simulation is performed for each member of the

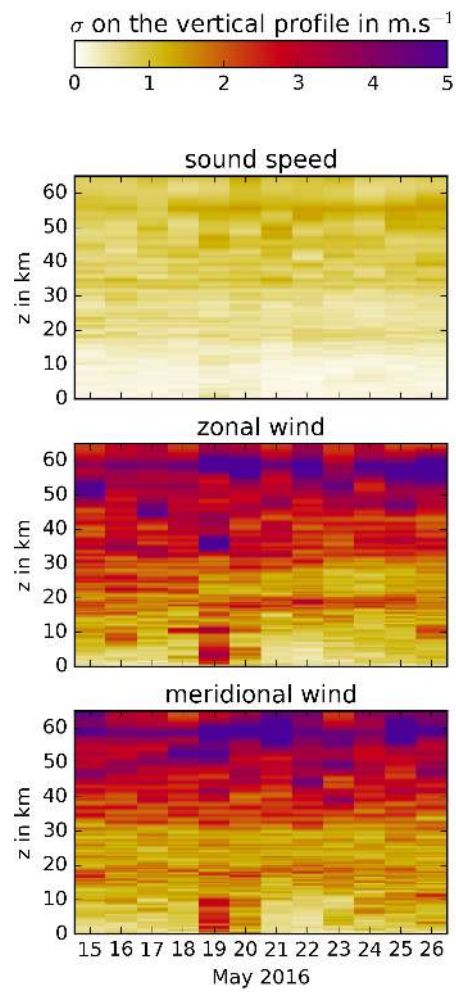


Figure 5.7 : Standard deviation of daily analyses from AEARP in terms of sound speed and horizontal wind components at location 13.0 deg E and 37.1 deg N (see Figure 5.4 (a)) between 15 May 2016 and 27 May 2016 (00 UTC).

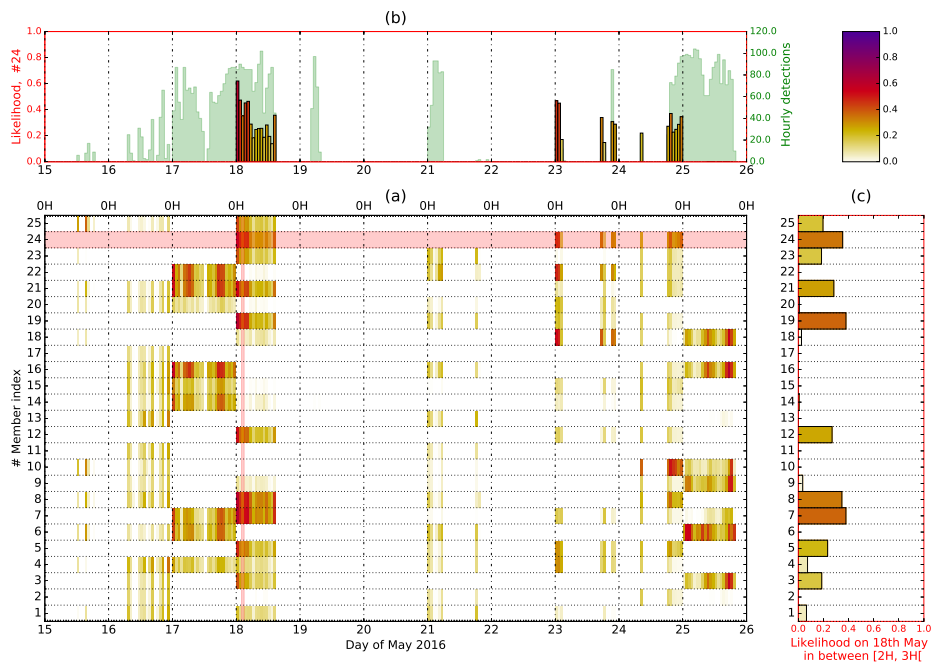


Figure 5.8 : (a) Hourly average likelihood for each AEARP member (vertical axis) over the eruptive period of May 2016. The pixel width represents an hour and the color represents the likelihood value. (b) The green bars represent the hourly number of detections at IS48 used to compute the likelihood. Other bars give the average likelihood for member #24 as a function of time. (c) Hourly averaged likelihood (between 02 and 03 UTC) for each member on the 18th May, 2016.

ensemble, each ray being launched from the Mount Etna peak location. The azimuth and elevation of each ray are chosen in the ranges 248.5 ± 3.5 deg and 20 ± 25 deg, respectively, with an angular resolution of 0.1 deg. As previously underlined, the components of the observation vector \hat{Y} are the trace velocity and the back azimuth as obtained using PMCC. The experimental methodology is fully justified in 5.3.1.

Figure 5.8 displays the evolution of the averaged likelihood value over IS48 detections that occurred within one hour reached by each AEARP member. The average likelihood distribution for a specific slot is shown in the right panel for the 18th of May, between 02 and 03 UTC. This figure shows that only few members explain the infrasound detections at IS48, thereby leading to a simple way to distinguish the atmosphere scenarii provided by the AEARP analyses. It can be seen that generally the hourly averaged likelihood is steadily decreasing over time except for some jumps at 00 UTC. These jumps correspond to updates of the AEARP ensemble analysis. Note also that an increasing number of detections do not necessarily imply a larger likelihood. For instance, and despite the occurrence of paroxysmal explosions the 19th of May, the average likelihood remains vanishingly small over the whole day, for any member of the ensemble. On the other hand, and in striking contrast with the 19th of May, the average likelihood values extend up to 0.5 one day earlier, depending on the member considered. In many cases, there is no unique dominant member and infrasound observations could be equally well explained by several atmospheric states. This behavior is coherent with the synthetic case study described in 5.3.1.

A final comment on the average likelihoods is that the maximum value over all members is almost always below 0.6, which basically suggests that the vast majority of

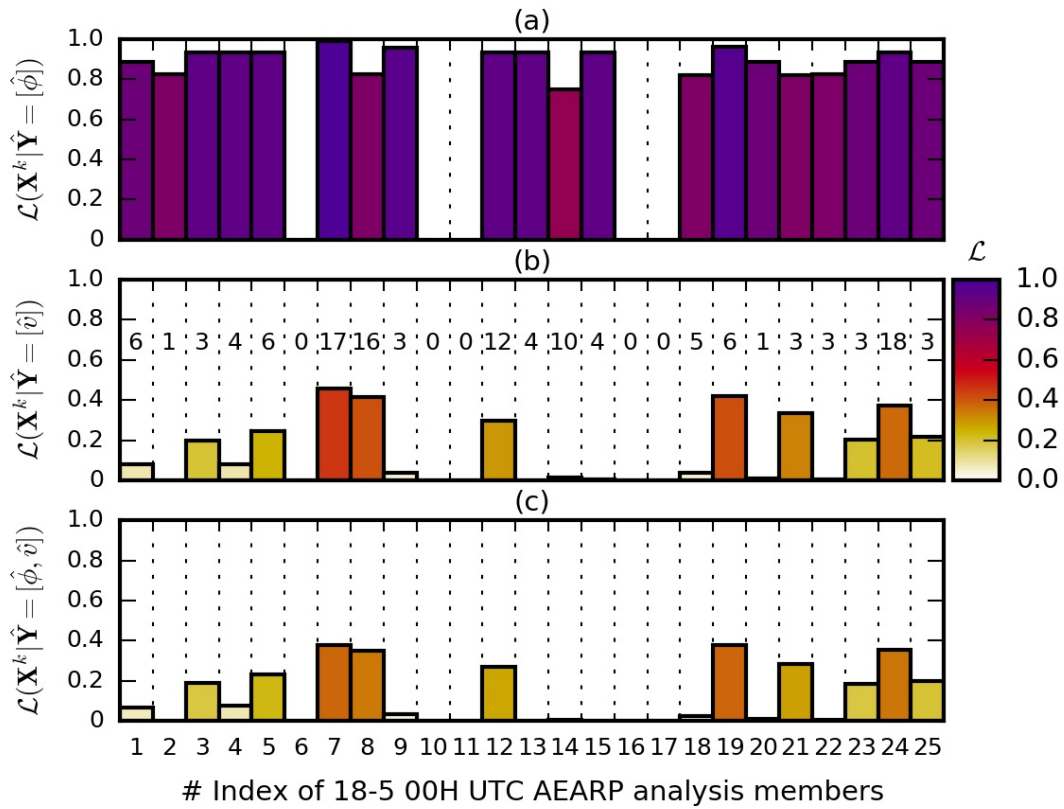


Figure 5.9 : Hourly averaged likelihood (02-03 UTC, 18 May 2016) for each AEARP member computed using the back azimuth (a), the trace velocity (b) and both observation components (c). The color scale is that of figure 5.8. The numbers shown on panel (b) are the number of eigenrays M available per member for the computation of the likelihood.

atmospheric states could not be used to explain the infrasound observations, at least with a good degree of confidence. This interpretation emphasizes the need to assess in quantitative terms the likelihood which is deemed meaningful. This is pursued in the next section.

5.2.6 Discussion

Influence of vector-valued observation

There are a few practical issues we face in setting-up a Bayesian approach. First, it is typical to use most useful information from PMCC for specifying the vector-valued observation. While this is a good starting point, the relative effect of each component on the likelihood is questionable. In Figure 5.9, the hourly averaged likelihood (02-03 UTC) is computed using the first component of the observation vector (i.e. $\hat{Y}_3 = \phi$), the second component ($\hat{Y}_2 = v$) or both components, for each AEARP member. The detailed comparison of the results obtained with these three choices suggests that each component does not contribute the same way to the likelihood. Taking the first component (Figure 5.9 (a)) yields to a relatively uniform distribution over the 25 members, unless no eigenrays can be computed (members 6, 10, 11, 16 and 17). Therefore a likelihood only based on the back azimuth is not a very selective criteria. Using the trace velocity leads to numerically large differences between the likelihood values, as Figure 5.9 (b) shows. According to the number of eigenrays provided by the different members, some members provide eigenrays like the #22 member without explaining the trace velocity

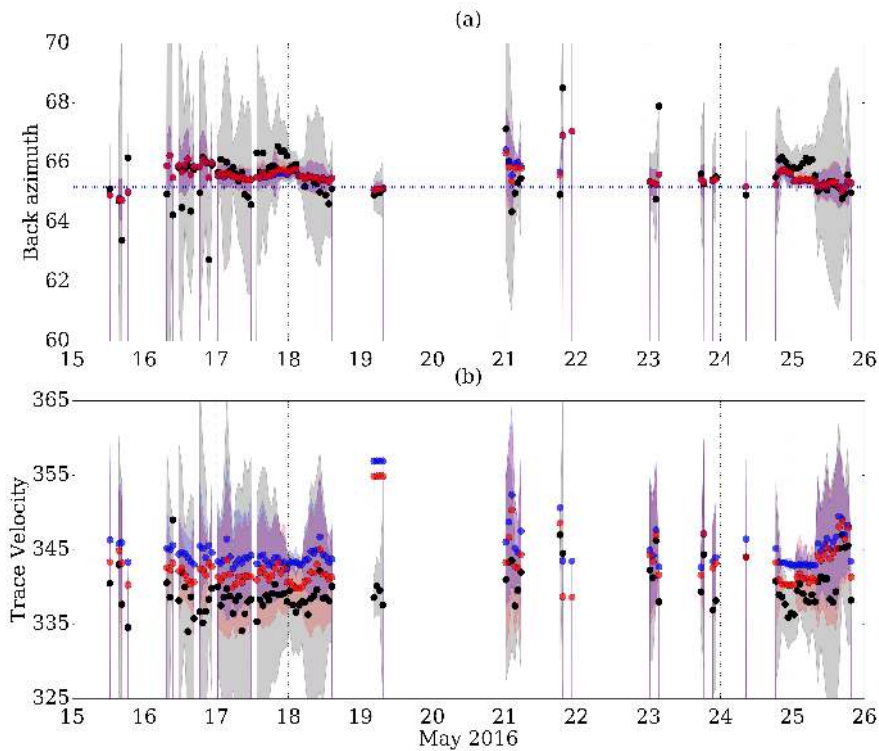


Figure 5.10 : Hourly averaged back azimuths (a) and hourly averaged trace velocities (b). Black circles correspond to the observations at IS48. The dotted blue line gives the theoretical back azimuth between the Mount Etna and the infrasound station IS48. Red and blue circles are (mean) predictions obtained using the sea surface temperature 190 km away from the source (figure 5.4) or the surface temperature at IS48, respectively. Shaded regions indicate two standard deviations of the mean.

observation since its likelihood is close to zero. This highlights that the trace velocity is a more selective criterion than the back azimuth according to our method. By comparing the likelihood value distribution using both the back azimuth and the trace velocity in Figure 5.9 (c) to the product of the distribution in Figure 5.9 (a) and (b) (not shown), it appears that the likelihood of few members are slightly lower than expected (e.g. member #7). This is due to our definition of the likelihood and more precisely to the way we choose the “optimal” eigenray. This highlights the fact that the eigenray that minimizes the difference between observed and simulated back azimuth, or that minimizes the difference between observed and simulated trace velocity, does not always minimize the sum of those differences. This result means that it is important to keep the two components of the observation vector in the definition (5.4).

Second, each observation is regarded as an input of the Bayesian method, despite the fact that it is often obtained through post-processing techniques. For example, the trace velocity is computed using $v = c_0 / \cos \theta$, where c_0 is the effective sound speed at ground level and θ is the elevation of the wavefront, to be computed for each eigenray. Since the winds at ground level were fixed to zero, the effective sound speed is reduced here to the adiabatic sound speed $c_0 = \sqrt{\gamma r T_0}$. It is common practice to overcome the lack of local measurement by using for T_0 the value at the lowest level of the extracted profile, so in our specific case at the sea level.

The trace velocities computed with this approach and observed at IS48 are displayed in Figure 5.10 (b). The computed velocities (blue, Figure 5.10 (b)) seem to be overestimated

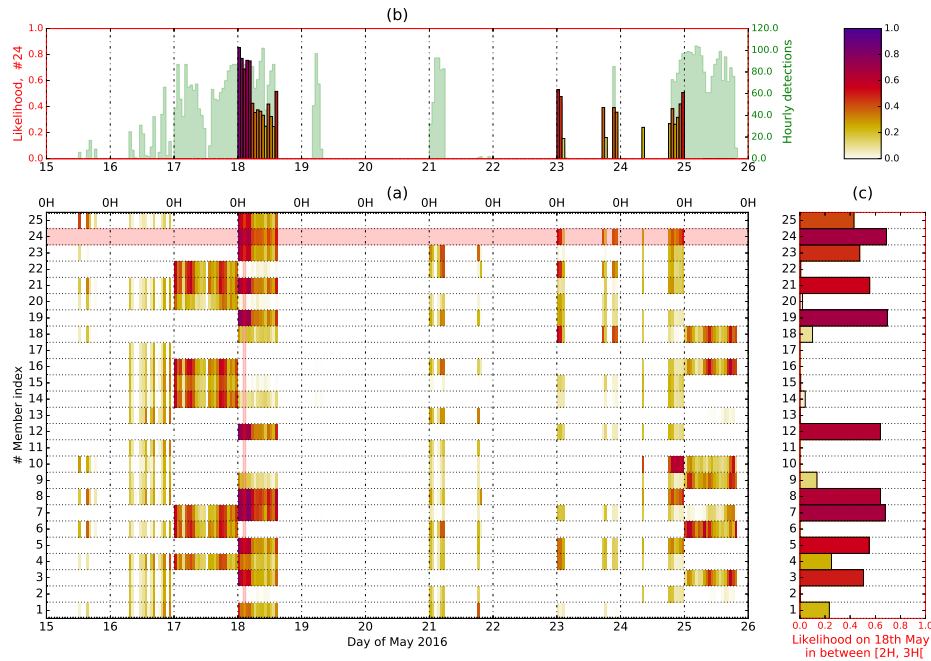


Figure 5.11 : Updated hourly averaged likelihood for AEARP members using the surface temperature at IS48 station for computing the trace velocity. See legend of figure 5.8.

with this definition of T_0 . Another choice for T_0 is provided by a linear interpolation of the surface model, for each AEARP member at IS48. Of course, since the horizontal grid spacing is fixed, neither the AEARP ensemble analysis nor any other atmospheric specification can capture the exact altitude of the infrasound station (900 m) and thus, there is inevitably some uncertainty associated with T_0 . Nevertheless, even with a rather coarse topography, the temperature is lower and the resulting trace velocities (red, Figure 5.10 (b)) are closer to the observations. Note that the computed back azimuth is not particularly affected by the change in definition of T_0 , as Figure 5.10 (a) shows.

In this case, Figure 5.11 shows that the average likelihood can be larger than 0.6 meaning a closer fit between the simulations and the observations. As we mention earlier in this article, the likelihood ratio provides the hard information. Since this ratio weakly depends on T_0 , the hierarchy of the members is not significantly affected by the choice of T_0 .

5.2.7 Relevance of the best atmospheric states

We now are studying if identical likelihoods highlight atmospheric states that share similar trends. Figure 5.12 (a-d) shows azimuth deviation $\phi_0 - \phi$ of infrasound at the surface predicted by the rays that reach the great circle linking the Mount Etna and IS48 station. Here ϕ_0 is the theoretical back azimuth, and the x-axis represents the distance from the source. The IS48 station, which is 550 km away from the Mount Etna, is represented by the solid blue line in Figure 5.12. Although the mean azimuth deviation is generally less than 1 deg for the members of maximum likelihood, the azimuth deviation spread can drastically change from one member to another.

Results are given for members #7, #19, #24, for which the average likelihood on the 18th of May (over the range 02 - 03 UTC) is roughly 0.7. The member #1, for which the likelihood is three times lower, is also depicted. As it was done in Figure 5.4 (a), the dots

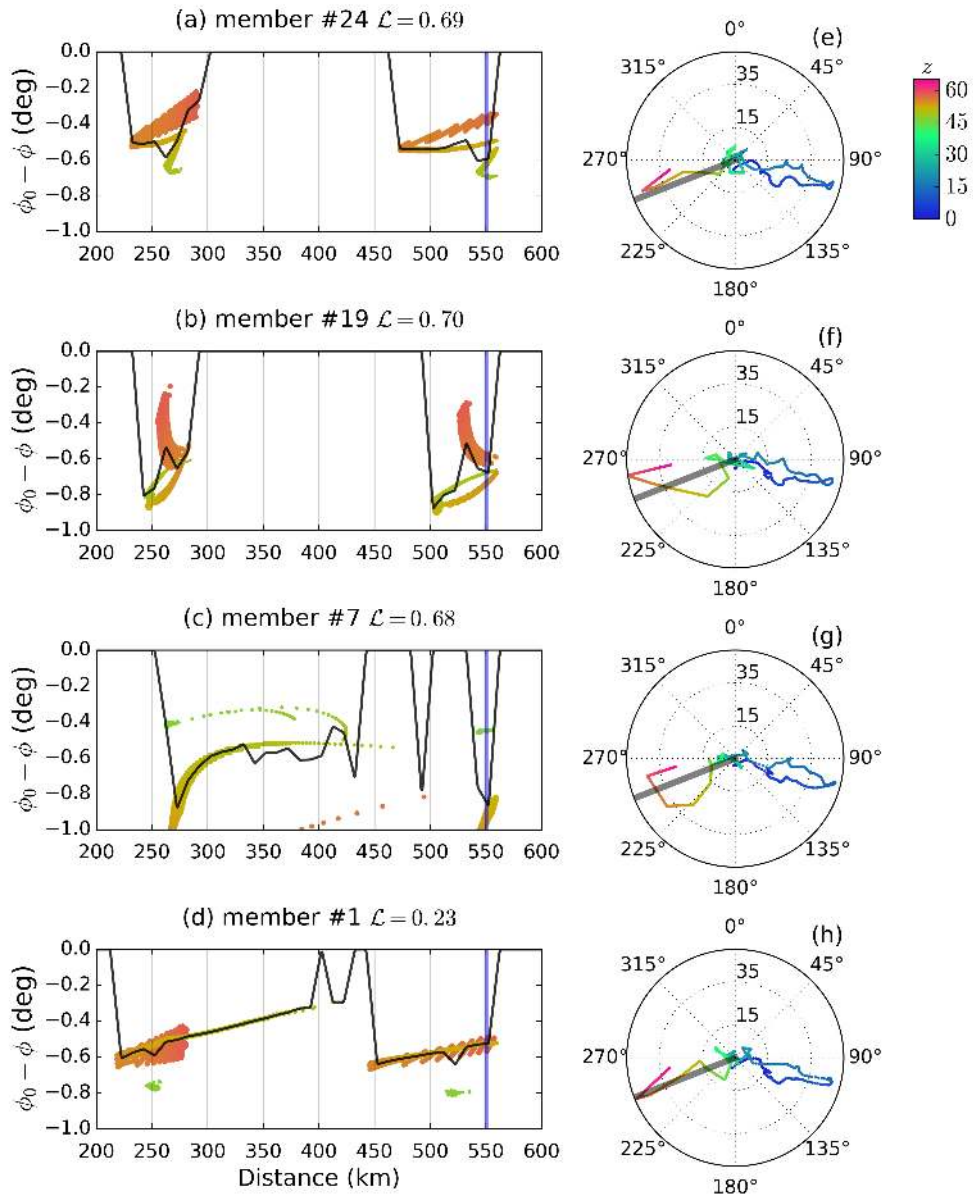


Figure 5.12 : (a-d) Deviation to the theoretical back azimuth ϕ_0 for members #24, #19, #7, #1; AERP analysis on 18 May 2016 (00 UTC). The black solid line is the mean azimuth deviation of impacting rays and the color of each point represents the altitude of refraction of the rays with a size proportional to v . The IS48 location is represented by the blue dashed line. (e-h) Hodographs with the same color scale for the altitude. The theoretical azimuth of IS48 from the view of Mount Etna is also represented by a thick gray line and isolines 15 m.s^{-1} and 35 m.s^{-1} are given by dashed circles.

that represent the impact of the rays are colored by their z_{top} to distinguish infrasound arrivals. Members #24 and #19 predict three distinct arrivals at IS48, and members #7 and #1 predict only two arrivals, as Figure 5.12 shows. Then one can expect similar features in the #19 and #24 wind profiles and in #7 and #1 wind profiles. Actually member #19 and #7 involve strong stratospheric winds perpendicular to the theoretical azimuth (see Figures 5.12 (f, g)) justifying the differences in azimuth deviations. Note that the wind profiles encountered by the rays are displayed by a hodograph, which is basically a line connecting the tips of wind vectors from one level to another. On the contrary, members #24 and #1, Figures 5.12 (e-h) show that the winds above 50 km are aligned with the source-receiver path, leading to less differences in terms of azimuth deviation between the arrivals. As mentioned earlier, all these four members share the same mean azimuth deviation. This can be due to the effect on the propagation of the strong tropospheric jet stream (e.g. eastward jet around 15 km of altitude) common to all displayed hodographs. Obviously, the comparison between wind profiles of the different members cannot justify by itself the differences in their likelihood values. A closer examination of the arrivals at IS48 shows that only the rays that are refracted around 49 km explain the observations, provided that the launch elevation is close to $\theta_0 = 20.5$ deg. This condition is satisfied for members #24, #19 and #7. In striking contrast with these members, the arrivals obtained with member #1 are associated both with higher refraction altitudes and elevation angles ($\theta_0 \simeq 24$ deg). In that case the trace velocities are larger than the one predicted by the other members and therefore its likelihood is lower. To explain this behavior we consider that since all the members have the same mean deviation, the propagation problem can be examined in a two-dimensional framework. This important property allows us to use the effective sound speed approximation (Godin, 2002).

A common approximation is to describe wave propagation through the effective sound speed $c_e = c + U$, where c is the adiabatic sound speed and U is the component in the direction of sound propagation. Figure 5.13 shows that the resulting profiles are similar for members considered in Figure 5.12, especially for altitudes in the range 48-55 km, although member #1 has a much lower likelihood than the other members. Using this approximation, the lower refraction altitude is located where the effective sound speed equals $c_{e0}/\cos\theta_0$, where $c_{e0} \simeq 320 \text{ m}\cdot\text{s}^{-1}$ is the effective sound speed at ground level and θ_0 is the launch angle (Jensen *et al.*, 2011). For $\theta_0 = 20.5$ deg, we get a threshold of $341.6 \text{ m}\cdot\text{s}^{-1}$, which is displayed in Figure 5.13. From this condition it is clear that the launch angle is larger for atmospheric specification provided by member #1. Hence, the value of θ_0 required for the ray to reach the infrasound station is not guaranteed, which leads to a reduced distance per bounce preventing them from reaching the station. This waveguide induces isolated arrivals just as shown in Figure 5.12 (d) (green arrivals). Note that member #7 is close to the threshold at 47 km, but remains slightly above the threshold which makes possible the ray to refract at 49 km. Members #19 and #24 are close to each other in terms of effective sound speed profiles, whereas one cannot see such similar trend in their respective hodograph displayed in Figure 5.12(e, f). This result highlights two critical aspects in computing the likelihood of a given member. First, the temperature profiles can be significantly different between two members to compensate the wind structures discrepancies. Second, the atmospheric state cannot be determined accurately by two components of the observation vector, even though the effective sound speed approximation is valid.

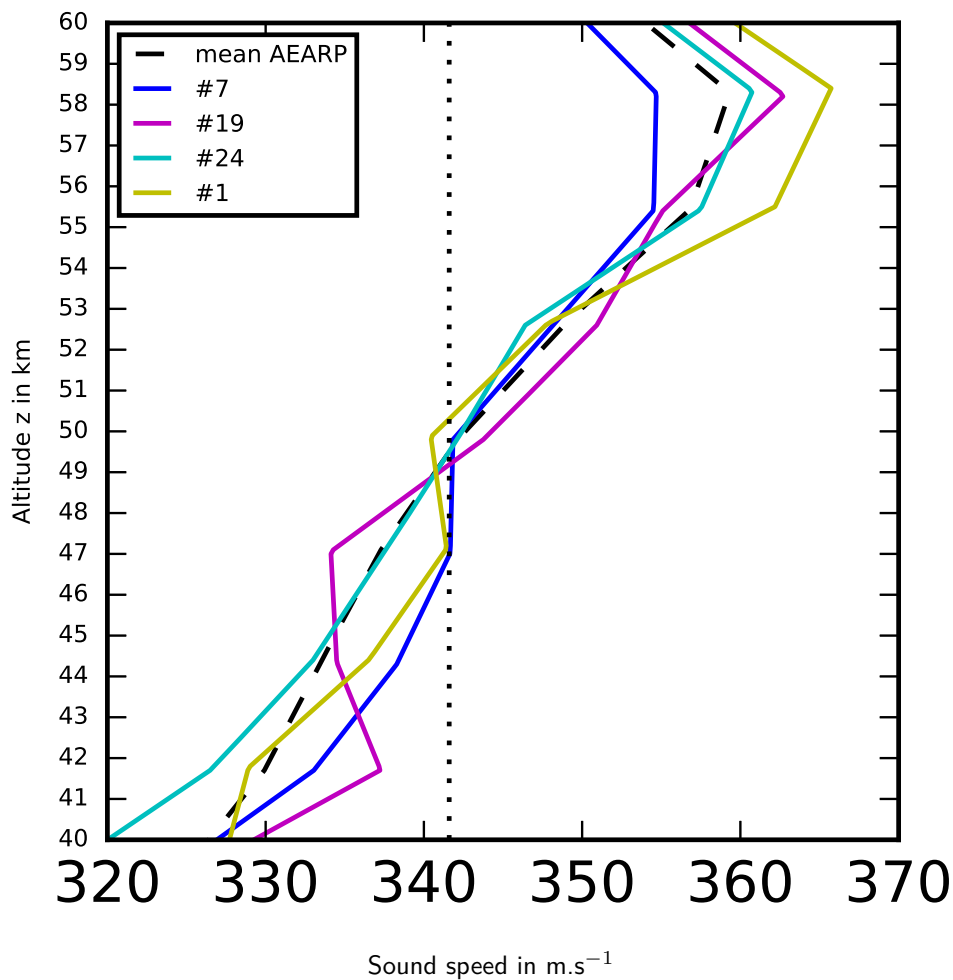


Figure 5.13 : Effective sound speed profiles for members #24, #19, #7 and #1 from the AEARP analysis of May 18 2016 (00 UTC) in the middle atmosphere. The effective sound speed profiles are extracted at location 13.0 deg E and 37.1 deg N. The mean AEARP profile is represented in dashed line and the gray region gives one standard deviation around the mean. The vertical dotted line is at 341.6 m.s⁻¹

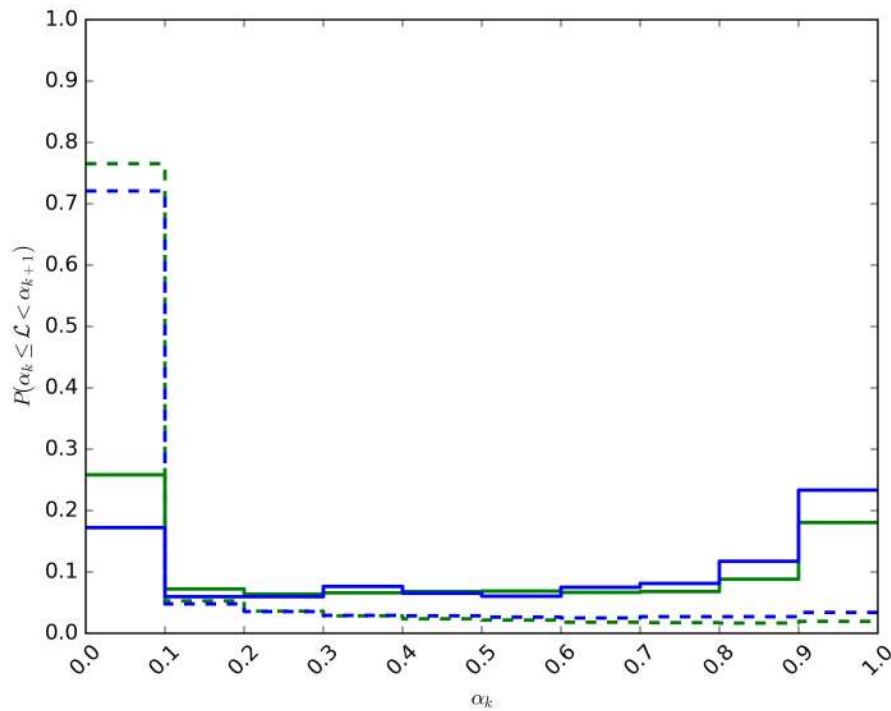


Figure 5.14 : Probability for the likelihood to lie within a given interval $\alpha_k \leq \mathcal{L} < \alpha_{k+1}$ over the period 15-26 May 2016. Solid lines correspond to \mathcal{L}_{best} and dashed lines to \mathcal{L} , for detections in the ranges 00 - 03 UTC (blue) and 00 - 24 UTC (green)

Synchronism between ensemble profiles and infrasound detections

From Figure 5.11 it can be seen that the likelihood value of member #24 generally decreases with time during a specific day. The AEARP analyses considered in this study are obtained through a data assimilation process (Talagrand, 1997 ; Courtier *et al.*, 1994) that uses observations available over a 6 hour window centered on 00 UTC. While these analyses are assumed to represent a valid, realistic atmospheric state over the 6h window, they have been considered to be valid during 24 hours starting at 00 UTC. To assess the impact of this assumption on the likelihood values, the probability of the likelihood to reach a specific value is shown in Figure 5.14. The probabilities of the likelihood values reached during this study are compared to the probability of the likelihood values computed using only IS48 detections between 00 and 03 UTC. One can notice that the likelihood values close to one are rare. The probability of likelihood values computed with observations between 00 and 03 UTC is higher close to unity and lower close to zero. This behavior is noticeable on the distribution of maximum likelihood values \mathcal{L}_{best} (Figure 5.14) as well. Thus the behavior noticed for member #24, as mentioned earlier (Figure 5.8, top panel), appears to be shared by all members. We conclude from this sensitivity study that the results from the Bayesian inversion are improved when the time difference between observations and model profiles is reduced. A better temporal consistency is therefore recommended in future studies.

5.2.8 Conclusions and Perspectives

The main goal of this study was to examine the interest of using infrasound propagation in order to improve wind and temperature profiles from a NWP model in the middle atmosphere (above 30 km). Indeed, only few observations are available in operational

atmospheric data assimilation systems to constrain the state of NWP models above 30 km. To reach that goal, an original methodology has been defined by selecting, through a Bayesian inversion, atmospheric profiles among an ensemble of analyses explaining best infrasound measurements in terms of back azimuth and trace velocity. More precisely, analyses from the Météo-France assimilation ensemble system AEARP with 25 members and the ray tracing model GeoAc have been used to simulate infrasound propagation between Mount Etna volcano and the IMS station IS48 in Tunisia over a 2-week period in May 2016.

The ray tracing method requires a large number of simulations with different initial incidence and azimuth angles together with a criterion defining rays reaching the station. Therefore, the experimental set-up has been considered with simulated observations in order to choose the best angular discretization and an optimal radius around the station for defining successful simulations (see 5.3.1). Similarly, the Bayesian inversion (based on a maximum likelihood estimate) has been adapted to the multiple solutions given by the ray tracing method for each atmospheric profile. This simulated observation framework has also been used to define optimal values of observation errors.

The capacity of the Bayesian inversion to select atmospheric profiles leading to simulations of back azimuth and trace velocity close to observations has been demonstrated for most events measured at IS48 from 16 May 2016 to 26 May 2016. The large spread of the AEARP profiles above 40 km allows a wide variety of infrasound simulations defining the a priori database from which the best ones are selected. It has been shown that when using only the back azimuth observation, only few members are discarded and that the selected ones have very similar likelihood values. The use of the trace velocity observation provides a better constraint for selecting a restricted number of profiles with high likelihood values. When examining in details profiles with such high values they exhibit significant differences in terms of azimuth deviation and altitude of refraction, despite some similarities in terms of vertical profiles of effective sound speed. Such result points out towards the need of additional observations to better constraint the inversion problem. This could come from the travel time of the front wave, but due to the lack of knowledge of the Mount Etna source, it has not been possible to consider such additional observable in the present study.

A number of hypotheses made in this study have been examined through sensitivity experiments. First the ray tracing model GeoAc ran in a homogeneous atmospheric profile along the propagation between the source and the station. We have shown that an accurate (unbiased) trace velocity can only be obtained when using surface temperature at the station. Atmospheric profiles from the AEARP are taken from analyses made at 00 UTC each day and assumed to be valid over a 24-hour period. By selecting only detections close to this analysis time, the likelihood values are somewhat increased showing the need to consider, in future studies, atmospheric profiles updated more frequently.

Future studies will address using GeoAc version that accounts for the full three-dimensional structure of the atmosphere along the propagation. This will allow to examine more precisely the validity of a single profile assumption that makes the inversion problem much simpler. Despite the demonstration made in this paper of the capacity of the Bayesian inversion to select a few profiles explaining infrasound observations it has not been possible to show that they are closer to an independent truth. It would be extremely valuable to consider infrasound stations collocated with dedicated instruments probing the middle atmosphere, such as Doppler LIDAR (Chanin *et al.*, 1989 ; Baumgarten, 2010) to assess the relevance of selected profiles by the Bayesian inversion. Improved knowledge about the source could be provided by continuous ones that could be more easily modelled than volcanic eruptions, such as ocean swells (Waxler & Gilbert, 2006 ; Ardhuin *et al.*,

2011 ; Le Pichon *et al.*, 2018). Similarly additional observables at detection stations could be simulated by other propagation models (e.g. full wave or reduced models can describe wave amplitude and duration of the signal (Bertin *et al.*, 2014)). Moreover, the use of several detection stations of the same infrasound source could help to improve the Bayesian selection (i.e. through the use of relative travel times between stations). This extension of the method is currently examined. Finally it would be interesting to consider ensemble short-range forecasts instead of analyses in the Bayesian inversion. This could allow selected profiles to be considered as "pseudo-observations" in atmospheric NWP data assimilation systems.

5.3 Compléments à l'article

5.3.1 Annexe de l'article : *Numerical set-up of the Bayesian inversion over synthetic observations*

The numerical set-up of Bayesian method presented in Section 5.2.4 has been defined using synthetic observations where the solution of the inverse problem is known. A punctual source at Mount Etna location is considered and the propagation domain remains the one between Mount Etna and Tunisia. The station measurements and the ensemble of atmospheric profiles are synthetic data.

Definition of the synthetic atmospheric profiles

The synthetic atmospheric profiles are designed to be consistent with the climatology of mid-latitudes summer (i.e. easterly zonal stratospheric circulation). The ensemble of profiles should have similar properties to those extracted from the AEARP analyses. However, for the sake of simplicity, the temperature profile of each member is taken from the US standard 1976 atmosphere model (US76). The temperature profile is built by considering a surface pressure value of 1013.25 hPa and a surface temperature of 288.15 K. The US76 model top is around 86 km, thus the temperature profile is extended by exponential increase to describe the thermospheric temperature. The baseline, denotes thereafter reference, wind profiles correspond to a single Gaussian stratospheric jet centered at 55 km of altitude blowing at $40\text{m}\cdot\text{s}^{-1}$ from the East. The reference atmospheric profiles (U , V and c) are shown in Figure 5.15 (a,b,c). To build our synthetic ensemble, the baseline wind profiles are perturbed by the addition of an vertically oscillating wind, which the amplitude, the direction and the wavelength are randomly drawn. This leads to $10\text{m}\cdot\text{s}^{-1}$ variability at 50 km of altitude on both winds components, which is consistent with the variability of the AEARP analyses at those levels. In order to keep this synthetic case similar to the real case study, 50 members are drawn. Then rays are computed with GeoAc for the fifty one members of this synthetic ensemble. In the case of the propagation for member #6, the station could detect multiple phases as shown on the projected trajectories of Figure 5.15 (e). The azimuth deviation induced by cross-winds is significant as shown in Figure 5.15 (d).

Bayesian inversion for the synthetic case

As an evaluation test, a ray computed with member #6 is used to generate synthetic observations, to be called thereafter the "test-ray". The synthetic station is defined as a disk centered on the impact point of the "test-ray". There is currently no straightforward method that provides the size of the region around the station that should be considered.

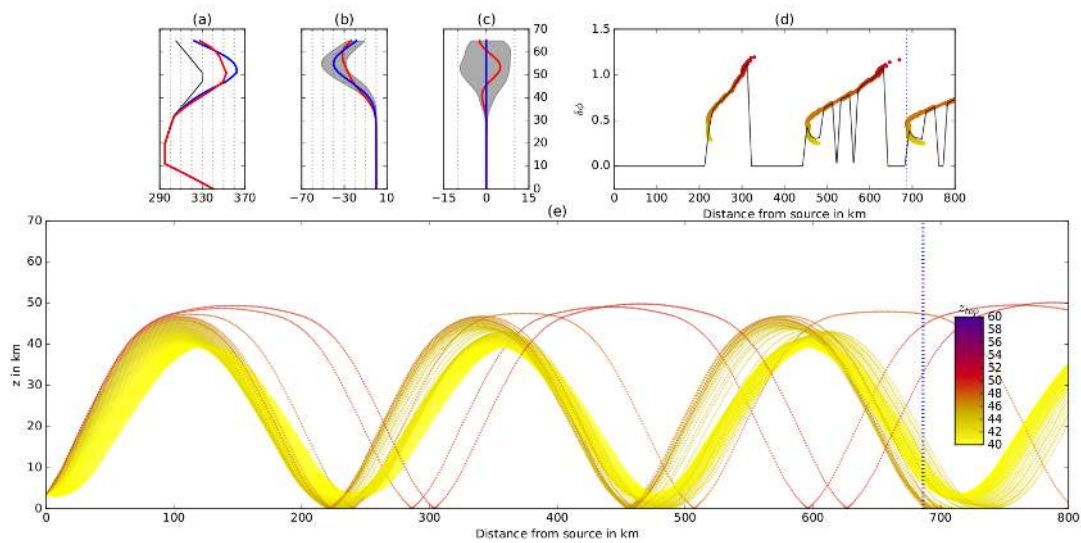


Figure 5.15 : Infrasound propagation for a set of atmospheric profiles from a synthetic ensemble of 50 members. (a) Profile of adiabatic sound speed (black), effective sound speed in the theoretical azimuth of members #0 (blue) and #6 (red). (b) Zonal wind profiles of members #0 (blue) and #6 (red) together with the ensemble spread (grey areas). (c) Meridional wind profiles of members #0 (blue) and #6 (red) together with ensemble spread (grey areas). (d) Deviation of azimuth of the rays along the geodesic. The solid line represents the mean deviation and the color scale represents the altitude of refraction. (e) Ray trajectories projected in the vertical plan along the great circle linking the station to the source. In both (d) and (e) panels, the blue dashed line represents the synthetic station location. The color scale shows the altitude of refraction.

Assink *et al.* (2014) proposed to define the disk diameter to consider in function of the array aperture. The diameter they took were equal to four times the aperture of the station they were studied. Thus, for our case study, the disk diameter should be set to 4 km which is four times the actual aperture of the Tunisian IMS station. In addition to this value of the diameter, the synthetic study was run with a diameter set to 20 km in order to evaluate the impact of such choice.

The test-ray characteristics (back azimuth angle, trace velocity, travel time) are considered as our synthetic observation vector. Their standard deviation errors are respectively set to 0.5° , 5 m.s^{-1} and 10 s. Then the rays are computed for the fifty other members of the synthetic ensemble and the Bayesian retrieval method is applied to the complete synthetic ensemble. Rays are computed every 0.1° between 245° and 252° of azimuth and -5° and 45° of elevation. This angle discretization is the most refine one we used, and it is denoted as the "highest". This leads to the likelihood values shown in Figure 5.16 (a).

As expected, member #6 gets a likelihood value of one which is significantly higher than the other members. Therefore the Bayesian retrieval is able to find and highlight the expected solution of this synthetic case. According to this distribution, member #17 does not provide any ray that reaches the station and thus according to the likelihood values, it is irrelevant.

By removing the travel time from the observation vector, the likelihood distribution exhibits several maxima, as shown in Figure 5.16 (b). The highest angle discretization is considered here. Two distributions are displayed : one where the station is a 20 km diameter disk and one where the diameter is reduced to 4 km. Regardless the station characteristics, member #2, #22 and #44 have a likelihood value close or equal to unity. Those members are different from member #6 highlighting the non-convexity of the inverse problem in such conditions. Then it is likely that for the real case study several likelihood maxima can be obtained. The use of the travel time makes the inverse problem better determined, however its use is non-trivial. A difficulty arises from the specification of the travel time error of an arrival especially when the source emission time is not well determined.

Sensitivity to the experimental set-up

By construction rays will sample the wavefront. Hence, a sparse distribution of rays can induce a misrepresentation of the ground to ground propagation of infrasound, or even the miss of an arrival. Thus the number of rays initially launched should be as high as possible but it has a computational cost. Thus a compromise needs to be found between the computational cost induced by the number of acoustic rays and the accuracy of the ground to ground propagation. A set of two angular discretizations has been considered : a first one where $\delta\zeta_0 = 1^\circ$ and $\delta\theta_0 = 0.5^\circ$, denotes thereafter as the coarsest discretization. And a second one where $\delta\zeta_0 = 0.1^\circ$ and $\delta\theta_0 = 0.1^\circ$, denotes thereafter as the most refined one. The impact of the two angular discretizations on the likelihood value reached by the fifty members of our synthetic ensemble is shown in Figure 5.16 (a) and (c). As shown in Figure 5.16 (c), with a low discretization and a 4 km diameter station, only 10 members have a non-zero likelihood and no other member than #6 has a significant likelihood value (i.e. over 0.6 corresponding to a simulation close enough to the observations). By comparing to the results shown in Figure 5.16 (a), the number of members with a non-zero likelihood increases significantly with a more refined discretization of rays when considering a 4 km diameter station. This means, that when not enough rays are launched, relevant members can see their likelihood value

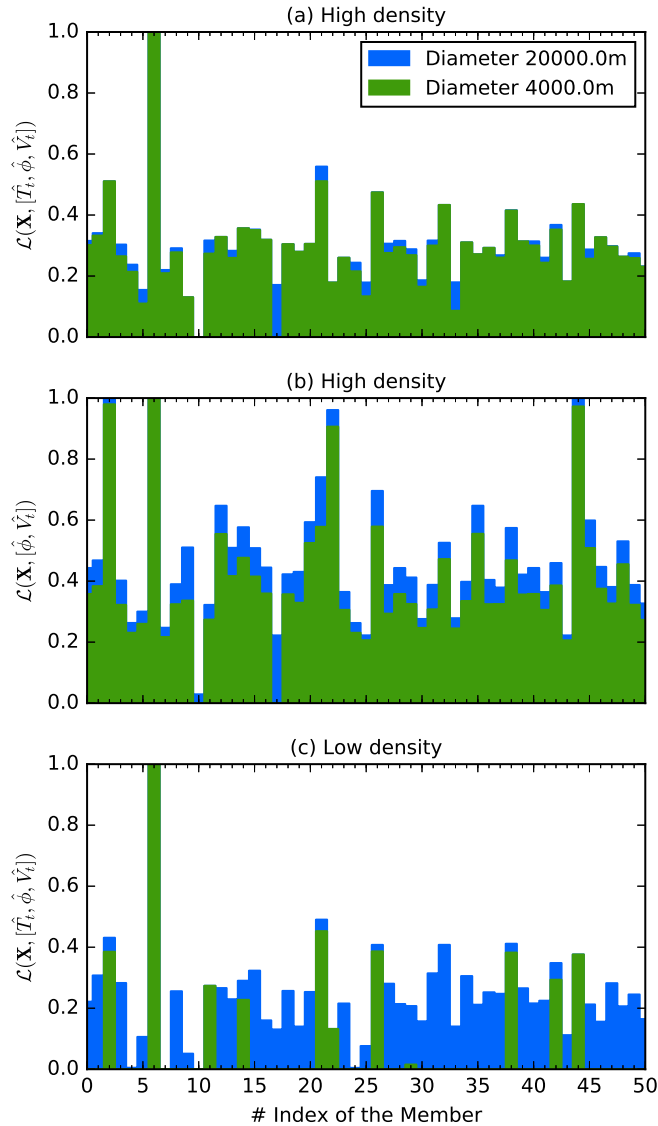


Figure 5.16 : Distribution of likelihood values for each member of the synthetic ensemble with different experimental specifications. (a) Values from a Bayesian inversion considering travel time, trace velocity and back azimuth measurements. The initial angular discretization is the most refine one (High density). (b) Values from a Bayesian inversion considering only back azimuth and trace velocity measurements. The initial angular discretization is the most refine one (High density). (c) Values from Bayesian inversion considering the three components of the observation vectors. The initial angular discretization is the coarsest one (Low density). Distributions in blue are obtained considering a 10 km radius around the station for sucessful ground to ground propagation, whereas for distributions in green this radius is reduced to 2 km.

drop to zero. Thus to avoid the removal of relevant members, the most refined angular discretization is chosen for the set-up. With this choice, equivalent likelihood distributions are obtained with a station of either 4 km or 20 km diameter. The choice of 4 km station has been retained in agreement with Assink *et al.* (2014).

Synthesis and expectations on the real case

The results obtained from the application of the Bayesian retrieval on synthetic observations reveal that the likelihood value will be equal to unity when the simulation results perfectly match the observations. Thus the likelihood defined in this study (5.4) is a good indicator of the fit between the simulation and the observations. The likelihood distribution can be non Gaussian leading to several members of the ensemble sharing similar likelihood values. Hence, interpretation of the temperature and wind retrieved profiles has to be done carefully. By examining the sensitivity of the likelihood distribution to various parameters, it appears that the highest density of angular discretization is preferable, whereas the 4 km diameter around the station allows us to be consistent with choices made in the literature.

5.3.2 Utilisation des temps de trajet

Dans le cadre de l'article, les temps de trajet obtenus à partir des valeurs t_0 des bulletins source (paragraphe 5.1.2) ont été évincés car trop incertains. Bien que les temps de trajet soient supérieurs à 1519 s, près de 700 d'entre eux ont une valeur supérieure à 2620 s, ce qui va à l'encontre du second critère de filtrage donné dans l'article. En réalité, ces détections ne sont pas aberrantes. Il est possible de trouver des bulletins source satisfaisant aux deux critères de sélection de l'article. En fait, ces détections sont le résultat d'une mauvaise association entre les bulletins à IS48 et les bulletins à la source. La figure 5.17 illustre l'impact de la prise en compte des temps de trajet, en plus des deux observations utilisées dans l'article (back-azimut et vitesse de passage), sur les vraisemblances.

Même si les temps de trajet pourraient être mieux pris en compte, par exemple en adaptant la fenêtre temporelle aux résultats de simulation, il restera toujours une part d'incertitude irréductible liée à la fréquence d'édition des bulletins de source.

5.3.3 Modèle de propagation *range-dependent*

L'une des hypothèses de l'article est le fait de supposer que l'atmosphère est homogène (i.e. ne varie pas) dans la direction horizontale. Cette hypothèse, qui est relativement classique, fait cependant débat au sein de la communauté des spécialistes des infrasons. C'est pourquoi, on réévalue ici la vraisemblance moyenne des membres, en suivant le même protocole que celui de l'article, mais en extrayant⁴ les profils au niveau de la source. On obtient alors la distribution présentée dans la figure 5.18. En utilisant ces profils, on constate que la vraisemblance permet toujours de hiérarchiser les membres au sein de l'ensemble, au sens où deux ou trois membres l'emportent sur les autres, même si les membres en question sont différents de ceux identifiés dans l'article. On remarque également qu'il est possible de hiérarchiser les membres pour la date du 19 mai, ce qui indique que les profils utilisés permettent, contrairement à ceux de l'article, d'assurer de

4. Notons que cette procédure ne permet pas, à proprement parler, de tester l'hypothèse dite *range-dependent*, mais d'évaluer l'impact d'un changement du point d'extraction du profil, en considérant toujours une propagation qui ne tient pas compte de l'évolution du profil entre la source et le récepteur.

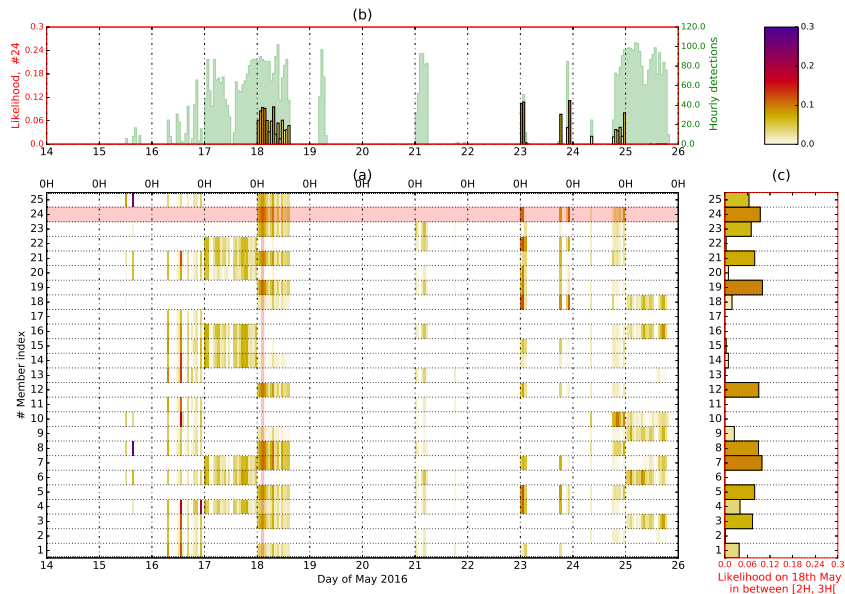


Figure 5.17 : Distribution des vraisemblances moyennes en tenant compte également des temps de trajet. Cette figure est à comparer avec la figure 5.11 de l'article. L'erreur d'observation considérée pour le temps de trajet est de 70 s (60 s pour le délai des bulletins source + 10 s pour la durée du signal à la station IS48)

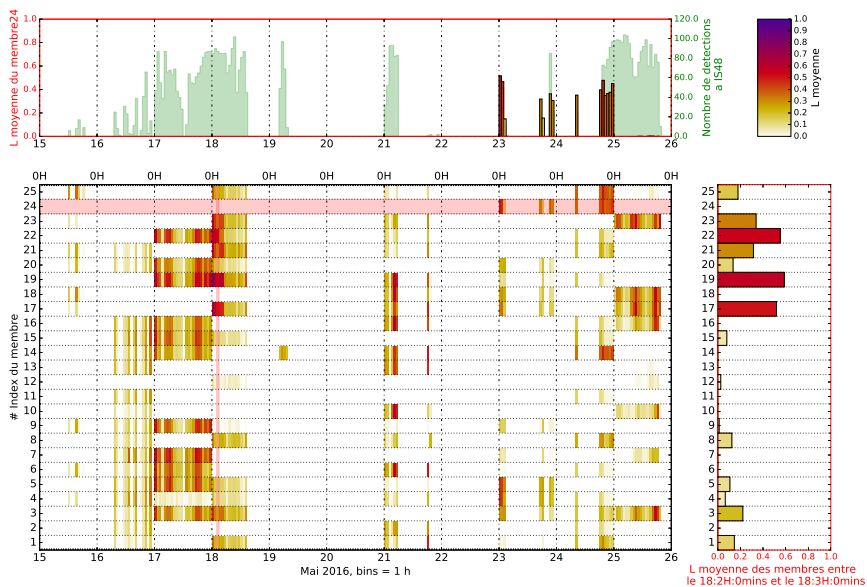


Figure 5.18 : Distribution des vraisemblances moyennes des membres sur la période éruptive du Mont Etna. Cette figure est à comparer avec la figure 5.11 de l'article. Les spécifications du milieu sont extraites au niveau de la source au lieu du point en mer.

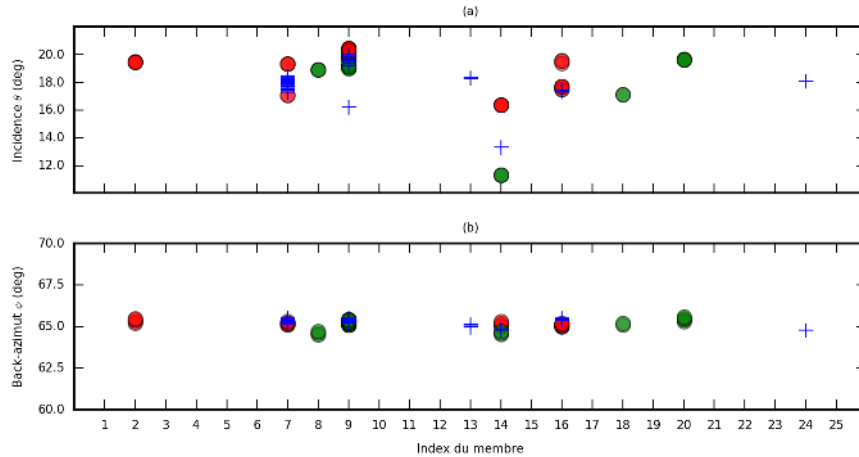


Figure 5.19 : Incidences et back-azimuts simulés en station le 19 mai. Croix bleues : résultats avec le code propagation *range-dependent*. Ronds rouges : résultats avec les profils extraits à 190 km de la source. Ronds verts : résultats avec les profils extraits au Mont Etna.

bonnes conditions de propagation entre la source et la station pour cette journée (i.e. produisent des rayons propres). Notons que les profils extraits au niveau du Mont Etna ont été extrapolés entre 0 et 163 m d'altitude en considérant une décroissance standard de la température et une atmosphère sans vent. Cette technique sera appliquée à tous les profils considérés dans cette partie.

Il est également possible de relaxer l'hypothèse *range-independent* considérée dans l'article. Pour cela nous utilisons une version du code GeoAc qui tient compte de l'évolution de l'état atmosphérique dans la direction horizontale (Blom, 2019) pour la date du 19 mai. Il apparaît que, là aussi, les membres se trouvent redistribués à la fois par rapport aux résultats de l'article et à ceux obtenus en changeant le point d'extraction des profils. D'autre part, on constate que le back-azimut des rayons propres reste proche du back-azimut théorique et que l'incidence des rayons reste élevée, comparée à celle attendue pour que la vitesse de passage simulée soit proche des observations. Le membre #14, quel que soit le protocole de simulation, sera celui qui a la moins mauvaise vraisemblance car il est celui conduisant à la simulation des incidences les plus faibles. Ainsi, l'ajout des effets *range-dependent* n'a pas corrigé l'écart entre les vitesses de passage observées et simulées. Toutefois, ce constat doit être nuancé dans la mesure où certains rayons sont réfractés vers le haut proche de la surface et non à réfléchis à la surface, comme le montre la figure 5.20. Cet effet numérique, probablement dû à la présence de la couche « extrapolée » peut néanmoins être corrigé en se plaçant au niveau du sol et non pas au niveau de la mer.

5.3.4 Apport des prévisions à court terme

Dans le cadre de l'article, seules des analyses ont été utilisées. On se propose dans ce paragraphe d'étendre la méthode aux ébauches de l'AEARP. Les ébauches de l'AEARP sont des prévisions perturbées à 6 h d'échéance et dont la perturbation a été amplifiée pour tenir compte de l'erreur du modèle de PNT (cf le chapitre 2). Ce sont ces ébauches qui servent d'a priori pour le cycle d'assimilation de données d'ensemble. Comme on peut le remarquer sur la figure 5.21, la dispersion et la moyenne des ensembles ne sont que peu impactées dans la moyenne atmosphère. La méthode d'inversion devrait aboutir au même type de résultats qu'avec les analyses. Si on examine l'incrément d'analyse, c'est-à-dire la différence entre les profils de l'ébauche et ceux de l'analyse, on remarque qu'un membre

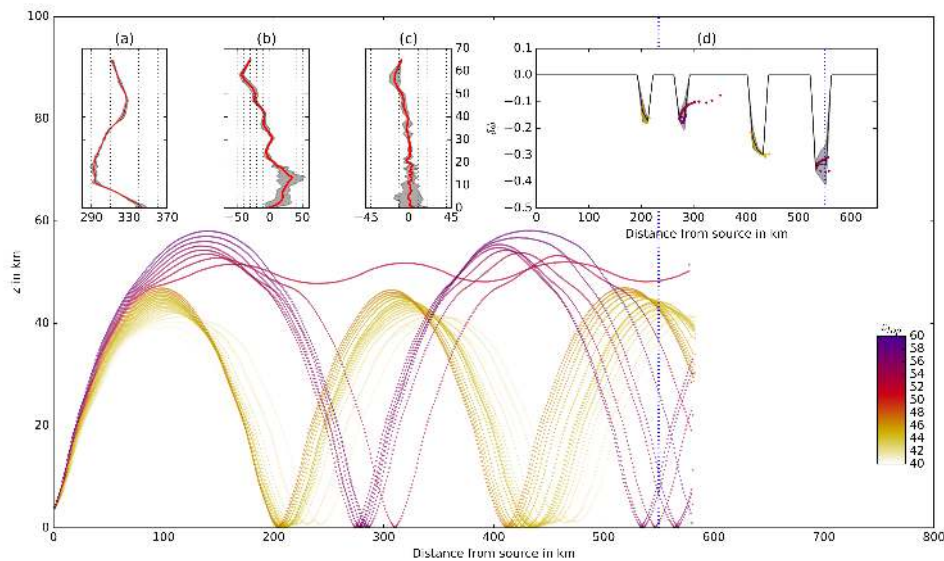


Figure 5.20 : Propagation utilisant le membre #16 en milieu hétérogène. (a,b,c) Profils respectifs de vitesses du son, de vent zonal, de vent méridien extraits dans la région de propagation. (d) Déviation d'azimut le long du grand cercle. (principal) Trajectoires des infrasons représentées par les lignes pointillées. Leur couleur est en fonction de l'altitude maximale de réfraction.

analysé peut voir son vent modifié de près de $10 \text{ m}\cdot\text{s}^{-1}$ au cours de l'analyse (figure 5.21). Cela implique qu'un membre de l'ébauche de forte vraisemblance peut avoir une faible vraisemblance vis-à-vis des infrasons après l'assimilation de données. On comprend alors l'intérêt d'appliquer la méthode de l'article pour proposer des pseudo observations à assimiler pour contraindre les analyses vers les ébauches de fortes vraisemblances.

Nous avons appliqué la méthode bayésienne aux ébauches des ensembles pour la base de détections de l'article. Les rayons sont lancés avec la discrétisation angulaire la plus lâche décrite dans le paragraphe 5.4.1. Pour compenser le manque de rayons lancés, la station IS48 est considérée comme étant un disque au sol de 10 km de rayon. Les profils utilisés sont localisés au niveau du Mont Etna. L'atmosphère sous la surface est supposée homogène (cf 5.1. la couche fait 163 m d'altitude). Les membres sont actualisés à 00 UTC et à 12 UTC. Les vraisemblances moyennes obtenues pour la période éruptive du Mont Etna en mai 2016 sont représentées en figure 5.22.

Comme attendu, les conclusions générales demeurent inchangées à celles de l'article. Certains membres, comme l'ébauche #19 du 18 mai, ont de très faibles vraisemblances et s'améliorent avec l'assimilation de données, alors que d'autres, comme le membre #13, voient leur vraisemblance fortement diminuer après analyse. L'actualisation toutes les 12 heures améliore un peu les résultats autour de midi.

5.3.5 Cas d'une autre station

Bien qu'une seule station ait été traitée dans le cadre de l'article, des études complémentaires ont permis d'associer des détections enregistrées à d'autres stations aux événements éruptifs considérés dans l'article. C'est notamment le cas de la station OHP dont l'analyse, appliquant le protocole décrit dans l'article, permet d'aboutir à la distribution des vraisemblances de la figure 5.23. Les profils ont été extraits en un point en mer le long du grand cercle liant le Mont Etna à l'OHP et les détections à l'OHP

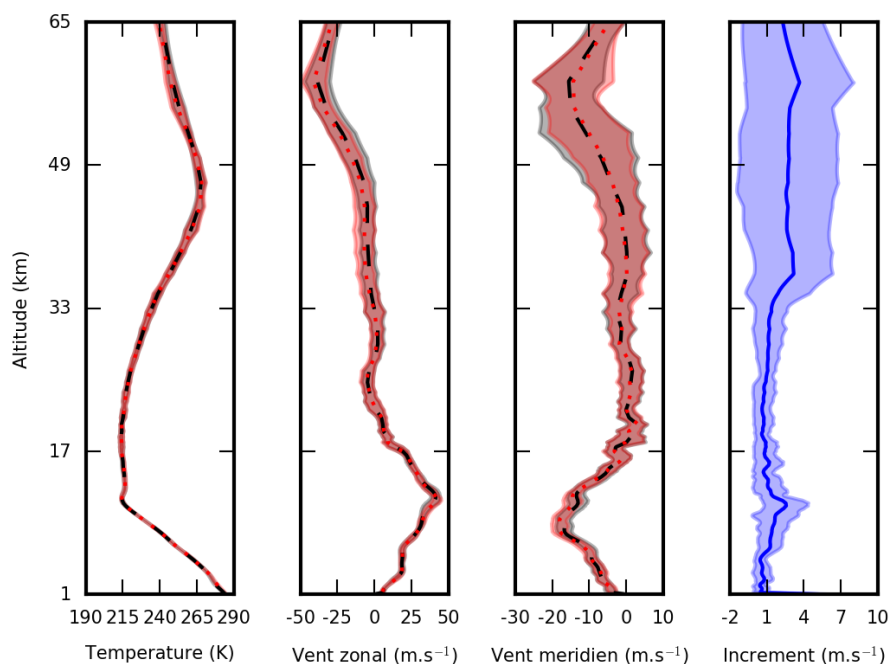


Figure 5.21 : Comparaison des ensembles de profils issus de l'ébauche et de l'analyse du 18 mai 00 UTC. En noir : données issues de l'ébauche de 00 UTC. En rouge : profils après assimilation de données. En bleu : incrément d'analyse sur la norme du vent. Les lignes sont les valeurs moyennes sur les ensembles et les surfaces couvrent deux écart-types autour de ces moyennes.

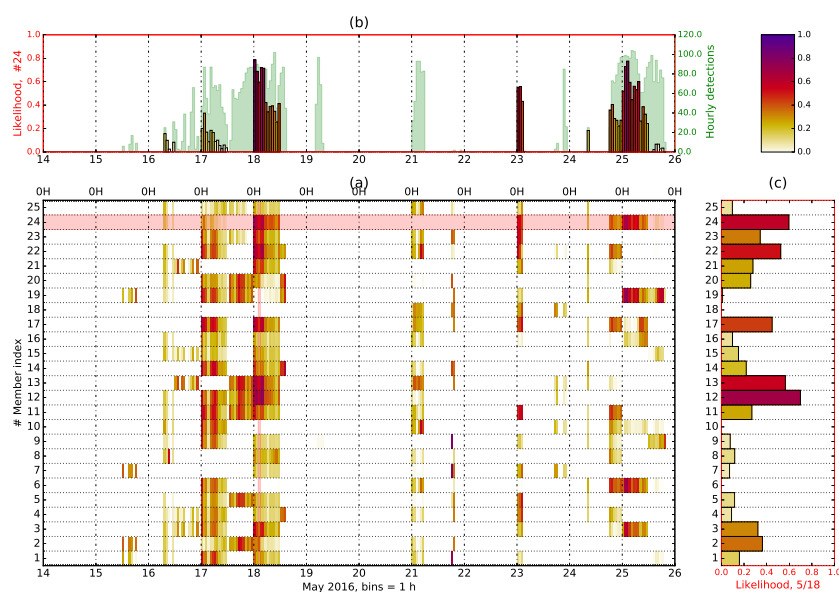


Figure 5.22 : Distribution des vraisemblances moyennes des membres de l'ébauche de l'AEARP sur la période éruptive du Mont Etna. Cette figure est à comparer avec la figure 5.11 de l'article. Les ébauches sont actualisées à 00 UTC et à 12 UTC.

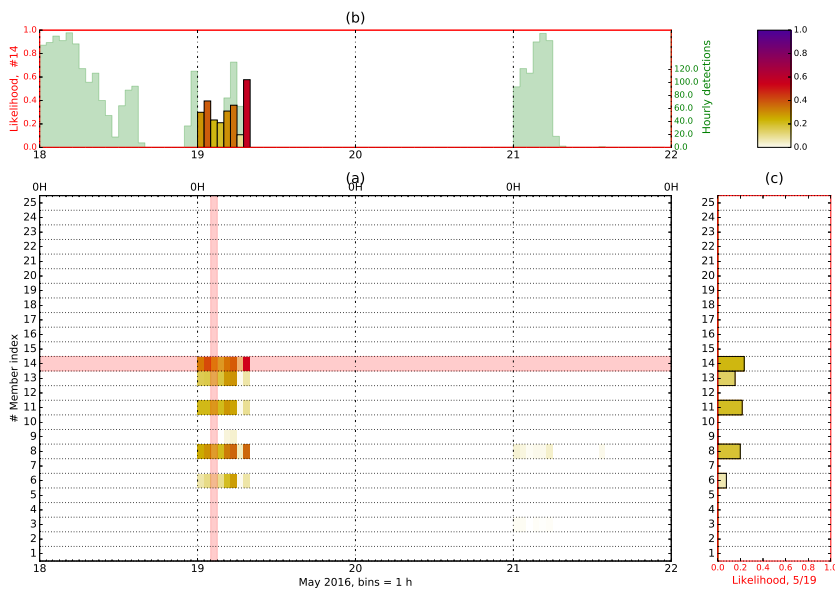


Figure 5.23 : Distribution des vraisemblances moyennes des membres pour les détections enregistrées à l'OHP entre le 18 mai et le 22 mai.

sont filtrées de la même manière que dans l'article. On remarque que seul l'événement paroxystique du 19 mai permet de hiérarchiser les membres, ce qui n'était pas le cas à partir des observations réalisées à IS48. Les valeurs de la vraisemblance sont faibles et s'expliquent par une sous-estimation des vitesses de passage simulées. Cela est dû au fait que l'altitude de l'OHP dans le modèle AEARP est plus élevée que dans la réalité et donc que la température de surface prise pour le calcul la vitesse du son au sol intervenant dans l'estimation de la vitesse de passage est surestimée. Pour conclure, une hiérarchisation en considérant les détections de plusieurs stations impliquerait le développement d'outils adaptés lorsqu'il n'est pas possible d'associer facilement un même événement à la source à des détections de différentes stations.

Chapitre 6

Conclusions et perspectives

Sommaire

6.1 Conclusions générales	113
6.2 Perspectives de la thèse	116

6.1 Conclusions générales

Dans le cadre de cette thèse, nous avons étudié comment et pourquoi des ondes infrasonores peuvent permettre de sonder la moyenne atmosphère. Pour ce faire, nous avons d'abord vu comment, à grande échelle, ces ondes acoustiques que sont les infrasons, sont affectées par la circulation générale de l'atmosphère terrestre. Il en ressort que la structure générale de l'atmosphère tend à confiner les infrasons au sein de guides d'ondes permettant ainsi leur propagation sur de grandes distances pouvant atteindre plusieurs milliers de kilomètres. Aujourd'hui, l'ensemble de l'activité infrasonore détectée par le réseau global du SSI ne permet pas de caractériser leurs sources, c'est pourquoi on a recours à des outils de simulation de la propagation. Cependant les résultats de simulations sont souvent significativement différents des observations. Ces écarts résultent, entre autres, de la difficulté à simuler finement, dans l'espace et le temps, la moyenne atmosphère (zone située entre 15 et 80 km d'altitude). En effet, cette couche de l'atmosphère n'est pas historiquement la région d'intérêt de la météorologie et donc des modèles de PNT. Ceux-ci sont principalement dédiés à la modélisation fine de la circulation troposphérique impactant directement la sécurité des biens et des personnes au travers de phénomènes météorologiques à enjeux (tempêtes, vents violents, pluies intenses, brouillard, ...). La moyenne atmosphère est une région difficile d'accès par les moyens de mesures in situ, et de ce fait, elle est moins bien documentée que la troposphère. Cependant, dès les années 70, des pionniers ont utilisé les infrasons pour corriger le vent de la moyenne atmosphère au moyen d'un code de tracé de rayons (Donn & Rind, 1971 ; Rind *et al.*, 1973 ; Rind & Donn, 1975). Ces premiers résultats, bien que très encourageants, étaient limités par les modèles atmosphériques et les modèles de propagation de l'époque. C'est pourquoi de nombreuses études sur l'évaluation et la quantification des corrections à apporter aux modèles atmosphériques ont été effectuées depuis cette époque, grâce à l'amélioration continue des moyens de simulation, des moyens de mesure des infrasons et de notre connaissance de l'atmosphère. De nos jours les modèles globaux de PNT décrivent partiellement ou totalement la moyenne atmosphère, mais les écarts entre les résultats de simulations et les observations infrasonores montrent qu'il existe encore des erreurs significatives sur l'état de la moyenne atmosphère qu'ils simulent.

L'intérêt grandissant des services météorologiques pour la circulation de la moyenne atmosphère, lié à l'étude de l'évolution de la couche d'ozone et à l'extension de la limite de prévisibilité des modèles, fait qu'aujourd'hui les modèles de PNT ont plus que jamais besoin d'informations dans ces régions. C'est pourquoi cette thèse a exploré les moyens d'extraire de l'information fournie par la propagation des infrasons en vue d'être assimilée dans les modèles de PNT. Il a été nécessaire dans un premier temps d'examiner quelles informations fournissent les modèles d'atmosphère sur l'état de la moyenne atmosphère. Les modèles de PNT sont aujourd'hui ceux qui permettent d'accéder à l'information la plus précise sur l'état de la moyenne atmosphère, et ce, malgré leur faible résolution verticale à ces altitudes (plusieurs kilomètres). La moyenne atmosphère n'est évidemment pas la seule région où les modèles de PNT présentent des incertitudes. C'est pourquoi les services météorologiques de prévision numérique du temps ont mis en place deux outils pour à la fois réduire et documenter les erreurs de modélisation. D'une part, ils utilisent un processus d'assimilation de données permettant d'améliorer l'estimation de l'état global de l'atmosphère grâce aux observations disponibles. D'autre part, ils utilisent des approches statistiques, usuellement appelées ensemblistes, afin d'estimer les erreurs respectives sur l'état atmosphérique analysé et celui simulé. En raison des contraintes actuelles de coût de calcul, les ensembles sont de petite taille et définis sur des grilles de plus basse résolution spatiale. Ainsi, on distingue une estimation déterministe de l'état de l'atmosphère (à haute résolution) et des estimations ensemblistes (à basse résolution). Il est toutefois important de noter que ces configurations déterministe et ensembliste restent intimement liées.

En vue d'assimiler des données infrasons, il est nécessaire de définir un opérateur d'observation. Si l'on souhaite assimiler directement des signaux infrasonores, l'opérateur d'observation doit alors correspondre à un modèle de propagation des ondes. Or, les ondes acoustiques ne sont pas résolues par les modèles globaux de PNT, essentiellement parce que la résolution numérique complète des équations régissant la propagation des infrasons sur le globe nécessiterait des moyens de calculs incompatibles avec les contraintes opérationnelles. C'est pourquoi nous avons retenu des méthodes de modélisation dites rapides que sont la décomposition modale et le tracé de rayons. Ces deux modèles se basent sur des hypothèses simplificatrices et donnent, de façon complémentaire, accès à différentes quantités d'intérêt des signaux infrasonores comme le back-azimut, la vitesse de passage, l'amplitude, ...

Mais, malgré l'utilisation de différents modèles, il apparaît difficile de simuler précisément les valeurs observées en raison des diverses sources d'erreurs. En effet, la source est rarement connue conduisant à des incertitudes de modélisation et les modèles peuvent ne pas simuler certaines quantités d'intérêt car limités par certaines de leurs hypothèses simplificatrices respectives. Lors de l'étude du cas Rivesaltes, la modélisation par tracé de rayons n'a pas pu fournir une estimation des back-azimuts et des vitesses de passage des ondes en accord avec les mesures faites à la station de l'OHP¹. La station de l'OHP se trouvait en effet dans une zone d'ombre géométrique. Au contraire, la modélisation par décomposition modale a permis de justifier une partie du signal observé. Comme les pionniers Donn et Rind en avaient l'intuition, il a fallu ajouter une perturbation aux profils de vent dans la moyenne atmosphère pour simuler l'ensemble du signal. Cette perturbation montre que l'adéquation entre le signal enregistré en station et les résultats de simulations dépend de la pertinence de l'état atmosphérique, vis-à-vis des infrasons, utilisé pour la simulation. Cependant, les complications liées à la modélisation de la propagation des signaux mises en évidence indiquent qu'il est trop compliqué de construire un opérateur d'observation efficace pour assimiler les infrasons directement.

1. Observatoire Haute Provence

Une stratégie alternative est alors examinée, qui consiste à proposer d'assimiler des profils verticaux de température et de vitesse du vent résultant de l'inversion des signaux plutôt que l'assimilation directe des signaux infrasons. L'opérateur d'observation se réduit alors à une interpolation spatiale des profils de vent et de température, ce qui est beaucoup plus simple à développer au sein d'un système d'assimilation de données pour un modèle de PNT. Nous avons alors résolu le problème inverse dans un cadre idéalisé à deux degrés de liberté, afin d'étudier sa capacité à restituer deux paramètres définissant la forme d'un profil vertical de célérité effective. Ce problème inverse, bien qu'idéalisé, a montré qu'il était non-convexe et difficile - mais pas impossible - à résoudre par le moyen de méthodes d'optimisation locale. Il requiert l'utilisation d'un terme a priori pour limiter l'espace des possibles et les minima locaux.

Nous avons ensuite modifié notre méthodologie d'inversion pour inclure une information a priori sur l'état de l'atmosphère. Il a été fait le choix de s'inspirer d'approches bayésiennes mises en place pour l'assimilation opérationnelle de certaines observations dans les modèles de PNT. Nous avons fait le choix d'utiliser des ensembles météorologiques comme information a priori. Afin de tester cette méthodologie, nous avons construit une base d'événements infrasons liés à une source connue : le Mont Etna. Lorsque les conditions météorologiques sont favorables (c'est-à-dire lorsque qu'il existe un guide d'ondes), les éruptions volcaniques du Mont Etna peuvent être enregistrées à la station du SSI IS48 située à 550 km du volcan. Sur l'ensemble des détections de IS48, nous avons isolé une période de 10 jours consécutifs où la station a détecté le volcan. Nous avons alors développé et évalué notre nouvelle approche sur cette base de détections. La méthode considère les ensembles d'ébauches ou d'analyses issus de l'assimilation d'ensemble du modèle global ARPEGE de Météo-France (AEARP) comme a priori, en supposant les différents membres équiprobables. La vraisemblance des membres est basée sur les écarts entre les back-azimuts et les vitesses de passage simulés par tracé de rayons et ceux observés. Nous avons pu mettre en évidence la capacité de la vraisemblance à hiérarchiser les membres selon leurs adéquations avec les mesures infrasons. Sur la période d'étude, les valeurs prises par la vraisemblance ont montré que deux ou trois membres parmi les vingt-cinq de l'ensemble expliquent mieux les observations. En comparant alors les profils de vent et de température, nous avons noté qu'ils étaient significativement différents, alors qu'en utilisant l'approximation de célérité effective, ces différences sont réduites. Ainsi, un même profil de célérité effective peut correspondre à différents états « vraisemblables ». Nous avons pu noter également que le profil de température joue un rôle non négligeable. L'étude du comportement de la vraisemblance a mis en évidence la nécessité d'avoir des informations précises sur les conditions de surface en station et l'importance d'actualiser fréquemment les profils atmosphériques a priori.

Pour conclure, la méthode développée dans le cadre de cette thèse a permis de sélectionner des profils de température et de vent conduisant à un bon accord avec des mesures infrasons au moyen d'un modèle de propagation. Il faut néanmoins retenir que l'information infrasonore n'a pas été pleinement exploitée, et donc qu'il est sans doute possible de rendre la sélection plus discriminante. En effet, d'autres stations ont entendu les éruptions étudiées. Il existe certaines informations plus précises sur la source, permettant d'utiliser d'autres quantités d'intérêt comme les temps de trajet, dont l'importance a été évoquée au cours de cette thèse. La méthode d'inversion est rapide, applicable aux ensembles d'ébauches et, moyennant des optimisations des coûts de calcul, pourrait répondre aux contraintes opérationnelles de l'assimilation de données. Le caractère discriminant de cette vraisemblance nous indique que l'assimilation d'une information liée aux infrasons aurait un apport, mais la quantification de cet apport nécessite d'autres études.

6.2 Perspectives de la thèse

À partir de la méthode proposée ici, des expériences d'assimilation pourraient être réalisées, en considérant des profils de vent et de température issus de la sélection bayésienne, ou des combinaisons de ces profils. Toutefois on peut s'attendre à ce que l'assimilation d'un seul profil dans un modèle de PNT n'ait qu'un impact faible sur la qualité de l'analyse et de la prévision au regard du nombre d'observations utilisées actuellement dans ces modèles. Pour espérer obtenir un impact significatif il serait nécessaire de disposer d'une base couvrant une période de temps assez longue, sur une large région du globe. La présence d'anomalies saisonnières telles que les réchauffements stratosphériques soudains serait de nature à renforcer l'impact attendu en apportant une plus-value par rapport aux systèmes d'observations opérationnels. Des expériences pourraient être faites d'abord avec des modèles de circulation générale simplifiés, comme cela a été réalisé dans le cadre de la thèse d'Angot (2013).

La méthode bayésienne développée ici a été appliquée aux données issues d'une seule station alors que plusieurs stations ont détecté les éruptions du Mont Etna. Le développement d'une méthode utilisant plusieurs stations simultanément se complique pour deux raisons : d'une part, la disponibilité d'un modèle de propagation "range-dependent" en coordonnées sphériques qui tient compte de l'orographie ; d'autre part la gestion de l'association des détections entre les stations à une source donnée. Des améliorations des codes de propagation sont attendues notamment de tracé de rayons sans zones d'ombre (thèse d'Annie Zelias au CEA, communication interne). L'association des détections progresse, elle aussi, au travers de méthodes bayésiennes (Millet *et al.*, 2018). Elle pourrait ainsi permettre d'étendre la sélection bayésienne à des stations multiples.

Le problème majeur associé à l'assimilation de données infrasons est celui des sources. En effet, peu de sources naturelles sont bien connues et instrumentées en permanence. Les volcans, lorsqu'ils sont bien instrumentés, sont de bons candidats, mais ils ne sont pas en permanence en éruption. Les observations sur l'état de la moyenne atmosphère obtenues seraient discontinues dans le temps et l'espace et seraient assimilées de manière irrégulière. Une autre source pourrait être les orages (Lacroix *et al.*, 2018, 2019), qui sont aussi discontinus dans l'espace et le temps mais qui peuvent être prévus par les modèles météorologiques notamment à fine échelle. Une autre source continue d'infrasons est les microbaromes. Ces sources sont étendues et résultent d'interactions entre la surface des océans et le vent. Des modèles pour ces sources sont en cours de développement (De Carlo *et al.*, 2020). Aujourd'hui, certains modèles de PNT sont couplés à des modèles de vagues et d'océan. Il serait alors envisageable d'intégrer les modèles de sources microbaromes au sein de modèles de PNT. L'existence d'un champ de sources au sein des modèles de PNT modifie l'accessibilité de l'assimilation directe de données infrasons. La définition d'un opérateur d'observation, correspondant à un modèle de propagation, devient envisageable sans toutefois résoudre les équations de propagation. En effet, des travaux récents montrent qu'il est possible de définir un méta-modèle pour la propagation des ondes par décomposition en polynômes de chaos (Thèse de Goupy A.). Une autre option serait de disposer de sources anthropiques contrôlées et dans ces conditions l'assimilation de données infrasons pour corriger des profils de vents devient exploitable (Amezcuca *et al.*, 2020).

Comme l'a montré Smets (2018), les infrasons peuvent être utilisés pour évaluer la circulation dans la moyenne atmosphère prévue par les modèles de PNT, que ce soit d'un point de vue déterministe ou ensembliste. Les infrasons pourraient alors devenir, au même titre que les observations de précipitations, un moyen d'évaluer les performances

des modèles météorologiques. L'approche bayésienne utilisant les infrasons, pourrait ainsi permettre d'évaluer certaines paramétrisations physiques (telle que celle des ondes de gravité) comme cela a été fait par Ribstein, B. (communications personnelles).

Toutes ces perspectives nécessitent de disposer de mesures directes de l'état de la moyenne atmosphère (lidar, radiomètres) pour valider les résultats d'inversion ou de bases de données infrasons fiables (source, bulletins de détection).

Annexe A

Extraction des données météorologiques

Sommaire

A.1 Les sorties d'un modèle de PNT	119
A.1.1 Extraction des champs d'intérêt	119
A.1.2 Construction de la grille verticale en pression	121
A.2 Construction des profils pour les modèles de propagation	125
A.2.1 Conversion altitude-pression	125
A.2.2 Données lacunaires et conditions opérationnelles	126
A.3 Calcul de la vitesse du son	127
A.3.1 Extrapolation dans la haute atmosphère et conditions de surface	131

Les fichiers natifs, ou historiques, de sortie des modèles de PNT sont de larges fichiers binaires incluant l'ensemble des champs d'intérêt de la météorologie. De nombreux champs sont inutiles pour la modélisation de la propagation des ondes infrasonores, c'est pourquoi cette annexe décrit les principales étapes nécessaires pour passer des fichiers natifs à des spécifications atmosphériques pertinentes pour la modélisation de la propagation des ondes infrasonores.

A.1 Les sorties d'un modèle de PNT

A.1.1 Extraction des champs d'intérêt

Dans ce paragraphe, nous parlons des fichiers natifs et de leur manipulation. Ces fichiers binaires sont organisés selon des normes précises. Dans le cadre de la thèse, ces fichiers étaient au format interne de Météo-France 'FA'¹ et au format standard « GRIB »², imposé par l'OMM. Bien qu'au vu des périodes regardées dans cette thèse, ce soit les fichiers 'FA' qui aient été majoritairement manipulés, nous décrivons ici la manipulation des fichiers 'GRIB', car ils correspondent au format standard de fourniture des données des modèles de PNT à la communauté des utilisateurs.

1. Ces fichiers sont constitués d'une couche logicielle dite « FA »(conteneur) pour gérer l'enveloppe dans laquelle les données sont enregistrées et d'un encodeur/décodeur permettant de transformer les données natives au format GRIB.

2. Le format GRIB est un format de fichier avec un conteneur et un encodeur/décodeur créés par l'OMM. L'encodeur/décodeur existe en 3 versions : GRIB-0, GRIB-1 et GRIB-2

Variable	<i>discipline</i>	<i>parameterCategory</i>	<i>parameterNumber</i>	<i>level</i>	<i>typeOfFirstFixedSurface</i>
Température air sec	0	0	0	1..105	119
Humidité spécifique (A.9)	0	1	0	1..105	119
Vent zonal	0	2	2	1..105	119
Vent méridien	0	2	3	1..105	119
Pression de surface	0	3	0	0	1
Géopotential de surface	0	3	4	0	1
Température de surface	2	3	18	0	1

Tableau A.1 : Description des champs selon les tables GRIB2 pour un fichier AEARP

Les fichiers GRIB, de l'anglais *GRidded Binary*, sont des fichiers séquentiels dans lesquels chaque enregistrement est auto-documenté. Chaque champ est précédé d'un en-tête de plusieurs octets, sans signe, contenant les informations tabulées suivantes :

- méta-données sur l'en-tête lui même (la longueur, usage des octets, présence de sous-en-têtes optionnels, le type de calendrier, de produits : analyse ou prévision, ...);
- la description standardisée du champ ;
- les transformations à effectuer pour obtenir les valeurs du champ ;
- les propriétés et les caractéristiques géométriques de la grille dans laquelle les données ont été produites.

Pour extraire un champ d'un fichier GRIB (dont on a vérifié préalablement l'origine), il faut en premier lieu connaître sa description standardisée (table 4 disponible sur le site https://www.nco.ncep.noaa.gov/pmb/docs/grib2/grib2_doc/). Par la suite, on supposera que l'on dispose d'une librairie permettant de lire le contenu de fichiers de type GRIB-2 dans un langage de programmation évolué comme Fortran, C ou Python, qui associe les octets d'en-têtes aux différentes clés tabulées. La première clé est la *discipline*. Pour un produit météorologique, la clé *discipline* du champ prend la valeur 0 alors que, par exemple, pour un produit océanographique, elle prendra la valeur 10. Une fois la discipline connue, les variables sont rangées dans différentes catégories indexées par la clé *parameterCategory*. Au sein de ces grandes catégories, les divers champs sont indexés selon la clé *parameterNumber*. Par exemple, si l'on souhaite extraire un champ de température d'un fichier Grib, alors il faut récupérer le champ répondant à la description [*discipline* :0, *parameterCategory* :0, *parameterNumber* :0]. Si le fichier GRIB contient le champ de température à différents niveaux (10-m, 850 hPa, ...), la description précédente n'est pas suffisante. Il faut connaître le niveau et le type de niveau utilisé dans le fichier GRIB (clés *level*, *typeOfFirstFixedSurface* et *typeOfSecondFixedSurface* si besoin).

Dans le cadre de la modélisation de la propagation des infrasons, les champs à extraire des fichiers natifs sont : la température, la vitesse du vent zonal, la vitesse du vent méridien, l'humidité spécifique³, la température de surface, la pression de surface et le géopotential de surface. La description des champs est fournie au tableau A.1 Les champs extraits peuvent être répartis sur une grille de Gauss (pour les sorties de AEARP, PEARP et ARPEGE). Les modèles de propagation utilisés ne considèrent que des champs sur des grilles uniformes. C'est pourquoi on interpole linéairement les champs sur des grilles uniformes dans la direction horizontale afin de satisfaire les conditions d'entrée des codes de propagation. La valeur 119 prise par la clé *typeOfFirstFixedSurface* au tableau A.1, indique que ces champs sont calculés sur des niveaux pressions hybrides (cf chapitre 2). Il faut donc disposer des paramètres de ces niveaux pour reconstruire les champs de pression.

3. l'humidité spécifique est la fraction massique en vapeur d'eau de l'air humide.

A.1.2 Construction de la grille verticale en pression

Dans le cadre des fichiers natifs de sortie d'ARPEGE ou d'AEARP, les champs sont disponibles sur des niveaux pression dits pleins. Dans le modèle numérique, lors du calcul, l'atmosphère est découpée verticalement en L couches délimitées par $L + 1$ demi-niveaux. Les valeurs de pression aux niveaux pleins, que l'on notera p_i , sont reliées à celles des demi-niveaux, que l'on notera $p_{i\pm 1/2}$ par l'équation :

$$p_i = \frac{1}{2}(p_{i+1/2} + p_{i-1/2}), \forall i \in \llbracket 1; L \rrbracket. \quad (\text{A.1})$$

Les $L + 1$ demi-niveaux, quant à eux, suivent la définition suivante :

$$p_{i+1/2} = A(i) + B(i)p_{gnd}, \forall i \in \llbracket 0; L \rrbracket, \quad (\text{A.2})$$

où A et B sont des tableaux de taille $(L + 1, 1)$ et p_{gnd} la pression de surface. Les $A(i)$ et $B(i)$ sont des constantes qui définissent la coordonnée verticale. Historiquement, on cherche à avoir des couches plus fines dans la troposphère et la couche limite planétaire, c'est-à-dire là où les gradients verticaux des différentes variables météorologiques sont les plus importants et les plus variables. On souhaite également que les niveaux les plus hauts soient des isobares pour éviter d'avoir l'influence numérique du relief dans la stratosphère et au-dessus. On souhaite également que la coordonnée verticale suive le terrain près du sol pour un traitement plus aisé de la condition à la limite inférieure. On parle ainsi de coordonnée verticale hybride.

Supposons que parmi les L couches, on souhaite avoir l niveaux isobares pour des pressions inférieures à p_0 . Le basculement se fait à la coordonnée $\eta_0 = l/L$. On se dote d'un paramètre ajustable σ_0 pour contruire B . On pose :

$$B(\eta) = \begin{cases} 0 & \eta \leq \eta_0 \\ \frac{(\eta - \eta_0)^2(\eta - \eta_2)}{(1 - \eta_0)^2(1 - \eta_2)} & \eta \geq \eta_0, \end{cases} \quad (\text{A.3})$$

où la coordonnée η varie entre 0 et 1 (l'indice du demi-niveau divisé par le nombre total de demi-niveaux) et η_2 est obtenue à partir de la relation $B(1 - 1/L) = \sigma_0$. On se dote d'un autre paramètre η_1 où l'on souhaite voir s'annuler A . Les valeurs de A suivent alors la définition :

$$A(\eta) = \frac{\eta(1 - \eta)(\eta_1 - \eta)}{\eta_0(1 - \eta_0)(\eta_1 - \eta_0)} p_0. \quad (\text{A.4})$$

A s'annule en η_1 et aux extrémités du segment $[0, 1]$. B s'annule en 0 et est égale à 1 en 1. Cela veut dire que le premier demi-niveau est l'isobare 0 Pa et que le dernier demi-niveau est la pression de surface. Les valeurs prises par A et B dépendent des paramètres ajustables $(\eta_0, p_0, \sigma_0, \eta_1)$. La seule contrainte est de s'assurer que la pression reste une fonction croissante de η pour toutes les valeurs réalistes de p_{gnd} .

Les valeurs prises pour A et B dans le cycle 42 d'ARPEGE et d'IFS sont données au tableau A.2. Le choix des valeurs permet de considérer suffisamment de niveaux dans la couche limite planétaire. La figure 2.8 de la thèse représente la distribution des 105 (137) niveaux d'ARPEGE (d'IFS) à la verticale d'un relief synthétique jusqu'à l'isobare 0.1 hPa (0.01 hPa). Ainsi, en connaissant les valeurs de A et B , on reconstruit, à partir du champ de pression de surface, l'ensemble du champ de pression.

Indice	A(ARPEGE)	B(ARPEGE)	A(IFS)	B(IFS)
0	0.00000000e+00	0.00000000e+00	0.00000000e+00	0.00000000e+00
1	2.00000000e+01	0.00000000e+00	2.00036500e+00	0.00000000e+00
2	3.00900250e+01	0.00000000e+00	3.10224100e+00	0.00000000e+00
3	4.45725840e+01	0.00000000e+00	4.66608400e+00	0.00000000e+00
4	6.48628530e+01	0.00000000e+00	6.82797700e+00	0.00000000e+00
5	9.31103600e+01	0.00000000e+00	9.74696600e+00	0.00000000e+00
6	1.32443415e+02	0.00000000e+00	1.36054240e+01	0.00000000e+00
7	1.87540847e+02	0.00000000e+00	1.86089310e+01	0.00000000e+00
8	2.63803409e+02	0.00000000e+00	2.49857180e+01	0.00000000e+00
9	3.67681628e+02	0.00000000e+00	3.29857100e+01	0.00000000e+00
10	5.06562941e+02	0.00000000e+00	4.28792420e+01	0.00000000e+00
11	6.88379537e+02	0.00000000e+00	5.49554630e+01	0.00000000e+00
12	9.20841495e+02	0.00000000e+00	6.95205760e+01	0.00000000e+00
13	1.20750433e+03	0.00000000e+00	8.68958820e+01	0.00000000e+00
14	1.54981194e+03	0.00000000e+00	1.07415741e+02	0.00000000e+00
15	1.94724209e+03	0.00000000e+00	1.31425507e+02	0.00000000e+00
16	2.39652359e+03	0.00000000e+00	1.59279404e+02	0.00000000e+00
17	2.89231163e+03	0.00000000e+00	1.91338562e+02	0.00000000e+00
18	3.42850265e+03	0.00000000e+00	2.27968948e+02	0.00000000e+00
19	4.00008754e+03	0.00000000e+00	2.69539581e+02	0.00000000e+00
20	4.60537749e+03	0.00000000e+00	3.16420746e+02	0.00000000e+00
21	5.24846284e+03	0.00000000e+00	3.68982361e+02	0.00000000e+00
22	5.92002418e+03	0.00000000e+00	4.27592499e+02	0.00000000e+00
23	6.60829462e+03	0.00000000e+00	4.92616028e+02	0.00000000e+00
24	7.23606118e+03	6.58361400e-04	5.64413452e+02	0.00000000e+00
25	7.79244141e+03	1.98661550e-03	6.43339905e+02	0.00000000e+00
26	8.29217133e+03	3.75863290e-03	7.29744141e+02	0.00000000e+00
27	8.74111759e+03	5.85926300e-03	8.23967834e+02	0.00000000e+00
28	9.14469299e+03	8.20855070e-03	9.26344910e+02	0.00000000e+00
29	9.50906444e+03	1.07558012e-02	1.03720117e+03	0.00000000e+00
30	9.84133832e+03	1.34825104e-02	1.15685364e+03	0.00000000e+00
31	1.01494365e+04	1.64075393e-02	1.28561035e+03	0.00000000e+00
32	1.04418470e+04	1.95935371e-02	1.42377014e+03	0.00000000e+00
33	1.07272843e+04	2.31554825e-02	1.57162292e+03	0.00000000e+00
34	1.10034669e+04	2.71081824e-02	1.72944898e+03	0.00000000e+00
35	1.12682828e+04	3.14653107e-02	1.89751929e+03	0.00000000e+00
36	1.15198069e+04	3.62398827e-02	2.07609595e+03	0.00000000e+00
37	1.17563104e+04	4.14449575e-02	2.26543164e+03	0.00000000e+00
38	1.19762581e+04	4.70945886e-02	2.46577051e+03	0.00000000e+00
39	1.21782921e+04	5.32050399e-02	2.67734814e+03	0.00000000e+00
40	1.23612000e+04	5.97962823e-02	2.90039135e+03	0.00000000e+00
41	1.25238679e+04	6.68937902e-02	3.13511938e+03	0.00000000e+00
42	1.26652150e+04	7.45306536e-02	3.38174365e+03	0.00000000e+00
43	1.27841101e+04	8.27500364e-02	3.64046826e+03	0.00000000e+00
44	1.28792685e+04	9.16080249e-02	3.91149048e+03	0.00000000e+00
45	1.29490555e+04	1.01164293e-01	4.19493066e+03	0.00000000e+00
46	1.29914074e+04	1.11438526e-01	4.49081738e+03	0.00000000e+00

47	1.30044792e+04	1.22432417e-01	4.79914941e+03	0.00000000e+00
48	1.29867987e+04	1.34154290e-01	5.11989502e+03	0.00000000e+00
49	1.29371816e+04	1.46611202e-01	5.45299072e+03	0.00000000e+00
50	1.28547155e+04	1.59809664e-01	5.79834473e+03	0.00000000e+00
51	1.27387333e+04	1.73756529e-01	6.15607422e+03	0.00000000e+00
52	1.25887811e+04	1.88460004e-01	6.52694678e+03	0.00000000e+00
53	1.24045784e+04	2.03930778e-01	6.91187061e+03	0.00000000e+00
54	1.21859749e+04	2.20183254e-01	7.31186914e+03	0.00000000e+00
55	1.19329054e+04	2.37236866e-01	7.72741211e+03	7.00000000e-06
56	1.16453445e+04	2.55117439e-01	8.15935400e+03	2.40000000e-05
57	1.13232648e+04	2.73858604e-01	8.60852539e+03	5.90000000e-05
58	1.09666013e+04	2.93503223e-01	9.07640039e+03	1.12000000e-04
59	1.05752255e+04	3.14104805e-01	9.56268262e+03	1.99000000e-04
60	1.01489334e+04	3.35728904e-01	1.00659785e+04	3.40000000e-04
61	9.69068810e+03	3.58297060e-01	1.05846318e+04	5.62000000e-04
62	9.20426496e+03	3.81717071e-01	1.11166621e+04	8.90000000e-04
63	8.69416286e+03	4.05883804e-01	1.16600674e+04	1.35300000e-03
64	8.16548514e+03	4.30680413e-01	1.22115479e+04	1.99200000e-03
65	7.62379567e+03	4.55979918e-01	1.27668730e+04	2.85700000e-03
66	7.07495626e+03	4.81647077e-01	1.33246689e+04	3.97100000e-03
67	6.52495243e+03	5.07540487e-01	1.38813311e+04	5.37800000e-03
68	5.97971616e+03	5.33514819e-01	1.44321396e+04	7.13300000e-03
69	5.44495367e+03	5.59423111e-01	1.49756152e+04	9.26100000e-03
70	4.92598602e+03	5.85119022e-01	1.55082568e+04	1.18060000e-02
71	4.42760955e+03	6.10458985e-01	1.60261152e+04	1.48160000e-02
72	3.95398177e+03	6.35304182e-01	1.65273223e+04	1.83180000e-02
73	3.50853667e+03	6.59522285e-01	1.70087891e+04	2.23550000e-02
74	3.09393148e+03	6.82988956e-01	1.74676133e+04	2.69640000e-02
75	2.71153576e+03	7.05618732e-01	1.79016211e+04	3.21760000e-02
76	2.36194653e+03	7.27341465e-01	1.83084336e+04	3.80260000e-02
77	2.04505614e+03	7.48102778e-01	1.86857188e+04	4.45480000e-02
78	1.76013834e+03	7.67864256e-01	1.90312891e+04	5.17730000e-02
79	1.50594655e+03	7.86603411e-01	1.93435117e+04	5.97280000e-02
80	1.28085569e+03	8.04310413e-01	1.96200430e+04	6.84480000e-02
81	1.08295320e+03	8.20987846e-01	1.98593906e+04	7.79580000e-02
82	9.10128593e+02	8.36650263e-01	2.00599316e+04	8.82860000e-02
83	7.60156637e+02	8.51323688e-01	2.02196641e+04	9.94620000e-02
84	6.30868213e+02	8.65034441e-01	2.03378633e+04	1.11505000e-01
85	5.20135633e+02	8.77813340e-01	2.04123086e+04	1.24448000e-01
86	4.25907991e+02	8.89695232e-01	2.04420781e+04	1.38313000e-01
87	3.46238826e+02	9.00718573e-01	2.04257188e+04	1.53125000e-01
88	2.79306705e+02	9.10925064e-01	2.03618164e+04	1.68910000e-01
89	2.23429418e+02	9.20359334e-01	2.02495117e+04	1.85689000e-01
90	1.77072555e+02	9.29068668e-01	2.00870859e+04	2.03491000e-01
91	1.38853271e+02	9.37102783e-01	1.98740254e+04	2.22333000e-01
92	1.07540044e+02	9.44513630e-01	1.96085723e+04	2.42244000e-01
93	8.20492750e+01	9.51355235e-01	1.92902266e+04	2.63242000e-01
94	6.15200200e+01	9.57656976e-01	1.89174609e+04	2.85354000e-01
95	4.51846750e+01	9.63448592e-01	1.84897070e+04	3.08598000e-01

96	3.23647680e+01	9.68760110e-01	1.80069258e+04	3.32939000e-01
97	2.24659420e+01	9.73621788e-01	1.74718398e+04	3.58254000e-01
98	1.49723630e+01	9.78064067e-01	1.68886875e+04	3.84363000e-01
99	9.44074800e+00	9.82117536e-01	1.62620469e+04	4.11125000e-01
100	5.49419100e+00	9.85812896e-01	1.55966953e+04	4.38391000e-01
101	2.81596200e+00	9.89180938e-01	1.48984531e+04	4.66003000e-01
102	1.14339600e+00	9.92252519e-01	1.41733242e+04	4.93800000e-01
103	2.62019000e-01	9.95058550e-01	1.34277695e+04	5.21619000e-01
104	0.00000000e+00	9.97629984e-01	1.26682578e+04	5.49301000e-01
105	0.00000000e+00	1.00000000e+00	1.19013398e+04	5.76692000e-01
106	nan	nan	1.11333047e+04	6.03648000e-01
107	nan	nan	1.03701758e+04	6.30036000e-01
108	nan	nan	9.61751562e+03	6.55736000e-01
109	nan	nan	8.88045312e+03	6.80643000e-01
110	nan	nan	8.16337500e+03	7.04669000e-01
111	nan	nan	7.47034375e+03	7.27739000e-01
112	nan	nan	6.80442188e+03	7.49797000e-01
113	nan	nan	6.16853125e+03	7.70798000e-01
114	nan	nan	5.56438281e+03	7.90717000e-01
115	nan	nan	4.99379688e+03	8.09536000e-01
116	nan	nan	4.45737500e+03	8.27256000e-01
117	nan	nan	3.95596094e+03	8.43881000e-01
118	nan	nan	3.48923438e+03	8.59432000e-01
119	nan	nan	3.05726562e+03	8.73929000e-01
120	nan	nan	2.65914062e+03	8.87408000e-01
121	nan	nan	2.29424219e+03	8.99900000e-01
122	nan	nan	1.96150000e+03	9.11448000e-01
123	nan	nan	1.65947656e+03	9.22096000e-01
124	nan	nan	1.38754688e+03	9.31881000e-01
125	nan	nan	1.14325000e+03	9.40860000e-01
126	nan	nan	9.26507813e+02	9.49064000e-01
127	nan	nan	7.34992188e+02	9.56550000e-01
128	nan	nan	5.68062500e+02	9.63352000e-01
129	nan	nan	4.24414063e+02	9.69513000e-01
130	nan	nan	3.02476563e+02	9.75078000e-01
131	nan	nan	2.02484375e+02	9.80072000e-01
132	nan	nan	1.22101563e+02	9.84542000e-01
133	nan	nan	6.27812500e+01	9.88500000e-01
134	nan	nan	2.28359380e+01	9.91984000e-01
135	nan	nan	3.75781300e+00	9.95003000e-01
136	nan	nan	0.00000000e+00	9.97630000e-01
137	nan	nan	0.00000000e+00	1.00000000e+00

Tableau A.2 : Valeurs de A et B pour les grilles verticales de ARPEGE et IFS.

A.2 Construction des profils pour les modèles de propagation

La propagation des ondes infrasonores est dépendante des grandeurs météorologiques suivantes :

- la température T .
- la composante zonale du vent, que l'on notera U .
- la composante méridionale du vent, que l'on notera V .
- la masse volumique, que l'on notera ρ .

Ces grandeurs doivent être exprimées en fonction de leur localisation dans les trois dimensions (latitude, longitude, altitude sol). Comme on a pu le voir dans les paragraphes précédents, les champs de ces grandeurs sont estimés sur une grille verticale en pression. Il faut à présent établir les correspondances altitude-pression.

A.2.1 Conversion altitude-pression

On cherche alors à passer d'une grille en pression hybride à une grille en altitude. Les modèles ARPEGE comme IFS considèrent l'équilibre hydrostatique (2.1), dès lors nous faisons l'hypothèse que les demi-niveaux pressions sont suffisamment proches les uns des autres pour que l'équilibre hydrostatique puisse être approximé par :

$$\Delta\Phi = -\frac{\Delta p}{\rho}, \quad (\text{A.5})$$

où Φ représente le géopotential, ρ est la masse volumique moyenne d'air humide (i.e. contenant de la vapeur d'eau) sur l'intervalle de pression Δp , où Δp représente l'écart entre deux demi-niveaux $p_{i-1/2}$ et $p_{i+1/2}$.

Le géopotential est défini par $\Phi = \int_0^h g(\text{lat}, z) dz$ où g représente ici la pesanteur terrestre. Par définition, la pesanteur g est dépendante de l'altitude z de la manière suivante $g = G \frac{m_{\text{terre}}}{(R)^2}$, où $R (= a + z)$ est la distance par rapport au centre de la terre, $G (= 6,674 \times 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2})$ la constante universelle de gravitation et m_{terre} la masse de la Terre. On considère que l'on peut approximer la gravité terrestre par \bar{g} la moyenne sur la sphère de g pour une altitude nulle. Alors pour toute altitude z , on approxime la gravité par $g \approx \bar{g}(1 - 2\frac{z}{a})$. Les altitudes maximales considérées sont de l'ordre de 100 km, donc la diminution de g par rapport à g_0 est de l'ordre de 1,5% à 100 km d'altitude. Pour simplifier, on supposera donc que $g \approx \bar{g} = 9,81 \text{ m} \cdot \text{s}^{-2}$ et que cette valeur est constante pour toutes altitudes⁴. Le géopotential Φ devient alors le produit de la gravité avec l'altitude. Cette approximation conduit à une erreur sur l'altitude des niveaux les plus hauts de l'ordre du kilomètre lors de l'intégration de l'équation (A.5).

La masse volumique n'est pas directement disponible dans les champs archivés par les modèles, mais comme l'atmosphère est considérée comme un mélange de gaz parfaits, sa valeur peut être déduite de l'équation suivante :

$$p_i = \rho_i \frac{R}{M_{\text{air}}} T_i, \quad (\text{A.6})$$

où $R (= 8,314 \text{ J} \cdot \text{K}^{-1} \cdot \text{mol}^{-1})$ est la constante des gaz parfaits, M_{air} représente la masse molaire moyenne de l'air sur l'intervalle Δp , et T_i la température de l'air au niveau plein. L'air est supposé être un mélange d'air sec, de masse molaire M_{dry} et à la pression

4. Ce sont des hypothèses faites au niveau du modèle de PNT.

partielle p_{dry} , et de vapeur d'eau, de masse molaire M_{vp} et à la pression partielle p_{vp} . Dès lors M_{air} s'exprime de la façon suivante :

$$M_{air} = \frac{p_{dry}}{p_i} M_{dry} + \frac{p_{vp}}{p_i} M_{vp} = \frac{p_i - p_{vp}}{p_i} M_{dry} + \frac{p_{vp}}{p_i} M_{vp}, \quad (A.7)$$

On se rappellera que la loi de Dalton exige que la somme de ces pressions partielles (p_{vp} , p_{dry}) soit égale à p_i . On cherche à faire intervenir uniquement des variables primitives du modèle. On note $\epsilon = \frac{M_{vp}}{M_{dry}}$ le rapport entre les masses molaires et on pose $r_{dry} = \frac{R}{M_{dry}}$ que l'on suppose constant et indépendant de l'altitude. En introduisant la définition de M_{air} (A.7) et les paramètres précédents dans l'équation (A.6), on obtient :

$$p_i = \rho_i \frac{R}{M_{dry} \left(\frac{p_i - p_{vp}}{p_i} + \frac{p_{vp}}{p_i} \frac{M_{vp}}{M_{dry}} \right)} T_i = \rho_i r_{dry} \frac{1}{1 - \frac{p_{vp}}{p_i} (1 - \epsilon)} T_i. \quad (A.8)$$

On retrouve alors une grandeur utilisée en météorologie qui est la température virtuelle ($Tv = \frac{1}{1 - \frac{p_{vp}}{p_i} (1 - \epsilon)} T_i$). Cette température correspond à la température qu'aurait un air sec avec la même pression et la même masse volumique.

En introduisant la fraction massique en vapeur d'eau

$$q = \frac{\frac{p_{vp}}{p} M_{vp}}{M_{air}} = \frac{\frac{p_{vp}}{p} M_{vp}}{\frac{p_{dry}}{p} M_{dry} + \frac{p_{vp}}{p} M_{vp}}, \quad (A.9)$$

aussi appelée humidité spécifique, dans la définition de Tv , on met en évidence alors la dépendance linéaire de la température virtuelle avec l'humidité spécifique $Tv = (1 + \frac{1-\epsilon}{\epsilon} q)T$. La correspondance géopotential-pression (A.5) s'effectue alors de la manière suivante⁵ :

$$\begin{cases} \Phi_{1/2} = \Phi_{gnd} \\ \Phi_{i+1/2} = \Phi_{i-1/2} - r_{dry} \left(1 + \frac{1-\epsilon}{\epsilon} q_i\right) T_i \frac{p_{i+1/2} - p_{i-1/2}}{p_i}, \forall i \in \llbracket 1; L \rrbracket \\ \Phi_i = \frac{1}{2} (\Phi_{i+1/2} + \Phi_{i-1/2}), \forall i \in \llbracket 1; L \rrbracket, \end{cases} \quad (A.10)$$

où Φ_{gnd} est le géopotential de surface. Par identification, on a également :

$$\rho_i = \frac{p_i}{r_{dry} \left(1 + \frac{1-\epsilon}{\epsilon} q_i\right) T_i}. \quad (A.11)$$

Ainsi les grandeurs caractéristiques du milieu de propagation (U , V , Tv , ρ) sont maintenant définies en fonction de l'altitude sol.

A.2.2 Données lacunaires et conditions opérationnelles

Les données du CEPMMT disponibles au CEA de manière opérationnelle ne possèdent pas les champs d'humidité spécifique, ni de géopotential de surface. L'intégration de l'équation (A.10) n'est pas possible telle qu'elle est décrite précédemment. Les fichiers du CEPMMT disponibles possèdent par contre le champ de pression réduite au niveau de la mer p_{rmsl} . On se propose alors de calculer le géopotential de surface en intégrant

5. Ici le niveau avec l'indice le plus élevé est celui le plus haut. D'où $\Phi_{1/2} = \Phi_{gnd}$

l'hypsométrie (cf chapitre 2) entre la pression p_{rmsl} et la pression de surface. On supposera que l'atmosphère est isotherme entre les deux niveaux. Cela conduit à la relation suivante :

$$\Phi_{gnd} \approx -r_{dry} T \ln(p_{gnd}/p_{rmsl}) \quad (\text{A.12})$$

On a alors les conditions de surface nécessaires à l'intégration de l'équation (A.10).

Historiquement le CEPMMT fournit au CEA des fichiers qui contiennent les champs météorologiques sur 90 niveaux verticaux. Comme la grille verticale des versions récentes du modèle a changé pour une grille avec 137 niveaux, les fichiers reçus actuellement comportent les champs des grandeurs météorologiques sur 90 niveaux pleins parmi ceux de la grille 137 niveaux⁶. Pour intégrer alors l'équation (A.10), il faut substituer les champs de température des niveaux manquants. Pour ces niveaux, les températures qui ont été utilisées sont les moyennes entre les valeurs des deux niveaux les plus proches. Ces approximations se traduisent par des écarts avec la correspondance altitude/pression trouvée avec les fichiers comportant les données sur l'ensemble des 137 niveaux. La figure A.1 montre les écarts en kilomètres associés à chaque niveau vertical. L'erreur générée est de l'ordre du kilomètre pour le point de la grille horizontale représentée. Pour estimer précisément l'impact de ces approximations, il faudrait répéter cette comparaison sur l'ensemble de la grille horizontale.

A.3 Calcul de la vitesse du son

Le modèle de décomposition modale a besoin en entrée d'une estimation de la vitesse adiabatique du son (en réalité sa célérité effective). Comme l'air est considéré comme un gaz parfait, la vitesse du son peut s'écrire en fonction de la température du milieu ambiant. On obtient la relation suivante :

$$c_0 = \sqrt{\gamma r T}. \quad (\text{A.13})$$

Ainsi, pour les latitudes tempérées et au niveau du sol, la vitesse du son est de l'ordre de 350 m.s^{-1} . Comme précédemment, $r (= R/M)$ est le rapport entre la constante des gaz parfaits R et la masse molaire de l'air humide. On décompose alors l'équation (A.13) de manière à faire apparaître les termes en air sec et en vapeur d'eau.

$$\left\{ \begin{array}{l} c_0^2 = \gamma r_{dry} \left(1 + \frac{1-\epsilon}{\epsilon} q\right) T = \gamma r_{dry} T v, \text{ avec} \\ \gamma = \frac{p_{dry} M_{dry} C_{p_{dry}} + p_{vp} M_{vp} C_{p_{vp}}}{p_{dry} M_{dry} C_{v_{dry}} + p_{vp} M_{vp} C_{v_{vp}}}, \end{array} \right. \quad (\text{A.14})$$

où $C_{p_{dry}}$ et $C_{v_{dry}}$ sont respectivement les capacités massiques isobare et isochore de l'air sec, et $C_{p_{vp}}$ et $C_{v_{vp}}$ sont respectivement les capacités thermiques massiques isobare et isochore de la vapeur d'eau. Cette nouvelle formulation fait intervenir la température virtuelle Tv et γ qui dépendent tous les deux de l'humidité de l'air. On cherche alors à estimer γ en fonction d'un $\gamma_{dry} = \frac{C_{p_{dry}}}{C_{v_{dry}}}$ et de l'humidité spécifique q . On met donc en facteur γ_{dry} dans l'équation (A.14) et on élimine les termes en p_{dry} avec la loi de Dalton. On obtient alors :

$$\gamma = \gamma_{dry} \frac{p - p_{vp}(1 - \epsilon\alpha)}{p - p_{vp}(1 - \epsilon\beta)}, \quad (\text{A.15})$$

6. correspondance disponible via le lien : <https://www.ecmwf.int/en/forecasts/documentation-and-support/correspondence-between-l91-and-l137-model-levels>.

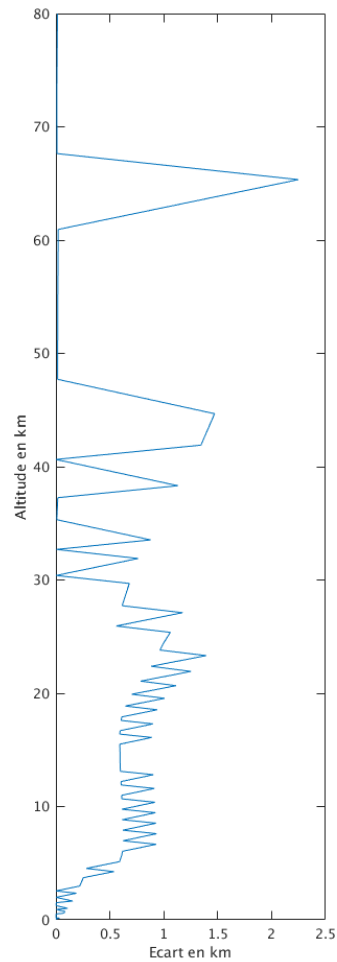


Figure A.1 : Comparaison entre les altitudes des 90 niveaux à partir des fichiers du CEPMMT et celles obtenues à partir des fichiers disponibles opérationnellement au CEA . Données issues de l'analyse du réseau de 00 h du 10/08/2016. La comparaison est effectuée en un point de grille particulier (0 °N,0 °E).

Constantes	Valeurs (Unité)	Constantes	Valeurs (Unité)
γ_{dry}	1.4 (SU)	R	8.314 (J.K ⁻¹ .mol ⁻¹)
M_{dry}	28.96 (g.mol ⁻¹)	M_{vp}	18.01 (g.mol ⁻¹)
Cv_{dry}	0.717 (kJ.kg ⁻¹)	Cv_{vp}	1.39 (kJ.kg ⁻¹)
Cp_{dry}	1.00 (kJ.kg ⁻¹)	Cp_{vp}	1.85 (kJ.kg ⁻¹)
ϵ	0.622 (SU)	a	6371 (km)
α	1.839 (SU)	β	1.932 (SU)

Tableau A.3 : Valeur des constantes physiques utilisées dans la thèse

T=303K, q = 0.02647	Air sec	Air sec avec Tv	Air Humide	Approximation (A.20)
c en m.s ⁻¹	349.01	351.80	351.38	351.37
erreur en m.s ⁻¹	- 2.37	+ 0.42	0.	10 ⁻²

Tableau A.4 : Exemple de correction induite par la prise en compte de l'humidité dans le calcul de la vitesse du son pour une température de 293 K, une humidité spécifique saturante et une pression de 1013 hPa. L'erreur est calculée par rapport à la correction en γ et Tv sans approximation (colonne 3).

avec $\alpha = \frac{Cp_{vp}}{Cp_{dry}}$ et $\beta = \frac{Cv_{vp}}{Cv_{dry}}$ dont les valeurs sont données au tableau A.3. D'après la définition de q , la pression partielle de vapeur d'eau p_{vp} est reliée à l'humidité spécifique par la relation $p_{vp} = \frac{qp}{q + \epsilon(1-q)}$. On introduit cette relation dans l'équation (A.15) et on obtient :

$$\gamma = \gamma_{dry} \frac{1 + (\alpha - 1)q}{1 + (\beta - 1)q}. \quad (\text{A.16})$$

Le rapport γ diminue avec l'humidité spécifique, et dans le cas où $q = 0.02647$ kg.kg⁻¹ (saturation en vapeur à 303 K et 1013 hPa), on obtient $\gamma = 0.9976\gamma_{dry}$ en prenant les valeurs présentes au tableau A.3.

Si l'on note $c_{dry} = \sqrt{\gamma_{dry} r_{dry} T}$ la vitesse du son dans un air sec, on trouve que :

$$c_0^2 = c_{dry}^2 \left(\frac{1 + (\alpha - 1)q}{1 + (\beta - 1)q} \right) \left(1 + \frac{1 - \epsilon}{\epsilon} q \right). \quad (\text{A.17})$$

L'apport de la prise en compte de l'humidité sur la valeur de c_0 est montré au tableau A.4.

D'après les résultats au tableau A.4, la considération d'un air sec au lieu d'un air humide, plus réaliste, engendre une sous-estimation de la vitesse du son. La seule prise en compte de la température virtuelle améliore de manière significative l'estimation de la vitesse du son, mais elle se trouve alors légèrement surestimée. La correction en γ est faible et sa dépendance en q compliquée, on souhaite alors l'approximer. On propose de multiplier numérateur et dénominateur par $(1 - (\beta - 1)q)$ dans la définition (A.16) de γ . On obtient alors :

$$\gamma = \gamma_{dry} \frac{(1 + (\alpha - 1)q)(1 - (\beta - 1)q)}{1 - (\beta q)^2}. \quad (\text{A.18})$$

Comme q est petit devant 1, on considère que le terme en βq^2 est donc très petit. On effectue un développement limité à l'ordre 0 de $\frac{1}{1 - (\beta q)^2}$, on trouve alors que $\gamma \approx \gamma_{dry}(1 + (\alpha - 1)q)(1 - (\beta - 1)q)(1 + o(\beta^2 q^2))$. En négligeant les termes d'ordre 2 en q on obtient l'approximation linéaire suivante :

$$\gamma \approx \gamma_{dry}(1 + (\alpha - \beta)q). \quad (\text{A.19})$$

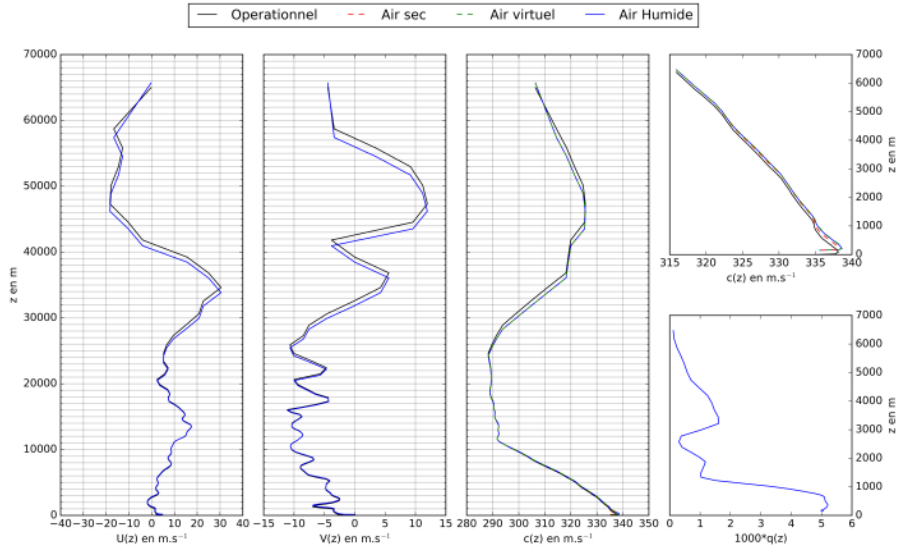


Figure A.2 : Impact des différentes hypothèses dans la définition des profils de vent et de vitesse du son. Les profils sont extraits du membre 1 de l'AEARP du 28-12-2018 du réseau de 00 H à la verticale du Mont Etna. Extraction des profils de vent par les méthodes opérationnelles du CEA (courbe noire). Extraction des profils de vent avec la conversion altitude-pression (A.10)(courbe bleue). La vitesse du son est calculée pour un air sec (courbe rouge). La vitesse du son est calculée avec la température virtuelle (courbe verte). La vitesse du son est calculée pour un air humide (courbe bleue).

La vitesse du son du milieu en altitude s'exprime donc de la manière suivante :

$$c_0 \approx \sqrt{\gamma_{dry} r_{dry} (1 - (\beta - \alpha)q) \left(1 + \frac{1 - \epsilon}{\epsilon} q\right) T} \quad (\text{A.20})$$

L'approximation amène à une sous-estimation de la vitesse de moins de $1 \times 10^{-2} \text{ m.s}^{-1}$ dans le pire des cas (i.e. humidité spécifique saturante pour une forte température) comme le montrent les résultats de la 4ème colonne du tableau A.4, ce qui est négligeable.

La figure A.2 présente les profils de température et de vent en un point de grille spécifique en considérant différentes conditions :

1. conversion altitude/pression considérant que les champs sont situés sur des niveaux hybrides en pression en faisant l'hypothèse d'une pression de surface homogène et égale à 1013.25 hPa sur le globe (opérationnel au CEA) et d'un air sec pour le calcul de la vitesse du son ;
2. conversion altitude/pression via le système (A.10) et considération d'un air sec pour le calcul de la vitesse du son ; conversion altitude/pression via le système (A.10) et prise en compte de la température virtuelle dans c ;
3. conversion altitude/pression via le système (A.10) et prise en compte de l'humidité dans le calcul de c (via l'équation (A.20)).

On remarque que les profils dits « opérationnels » sont décalés de quelques kilomètres sur la verticale. La différence entre le profil de vitesse du son considérant un air humide (en bleu) et celui considérant seulement la température virtuelle (en vert) est négligeable. On utilisera donc par la suite la définition basée sur la température virtuelle.

A.3.1 Extrapolation dans la haute atmosphère et conditions de surface

Les modèles de propagation des infrasons ont besoin de profils atmosphériques depuis la surface jusqu'à la thermosphère. Quel que soit le modèle PNT global utilisé (tableau 2.3), il faut combiner ses profils avec ceux d'un modèle de prévision de l'état de la thermosphère. On a vu dans le chapitre 2 que la meilleure estimation disponible dans la thermosphère correspond à celle proposée par les modèles empiriques tels que MSIS, HWM et plus récemment MENTAT. Le raccordement effectué actuellement au CEA entre les profils des modèles de PNT et ceux des modèles empiriques utilise des splines cubiques. Il commence à 80 km d'altitude et peut donc modifier les données fournies par le modèle IFS du CEPMMT. Nous avons plutôt opté pour un raccordement utilisant une fonction tangente hyperbolique placée de manière à ne pas modifier les données météorologiques.

Les modèles sont également sensibles au type d'interpolation verticale effectuée. En effet, une interpolation spline peut générer des structures supplémentaires non simulées par le modèle. De la même manière, une interpolation linéaire lisse toutes les structures et génère des ruptures de pentes qui peuvent poser des problèmes lors du calcul des champs de pression acoustique ou des rayons. Cet effet est accentué dans la stratosphère et la mésosphère, où les modèles ont une faible résolution. Une solution serait alors de modéliser certaines structures de petites échelles verticales présentes en réalité, en utilisant une paramétrisation des ondes de gravité (Gardner *et al.*, 1993).

Annexe B

Intérêt des prévisions d'ensemble

Sommaire

B.1 Les paramétrisations physiques pour le modèle AR-PEGE et la PEARP	133
B.1.1 Les paramétrisations physiques	134
B.1.2 PEARP et perturbations par approche multi-physiques .	135

Dans le cadre de cette thèse, ce sont les ensembles du cycle d'assimilation d'ARPEGE (AEARP) qui ont été utilisés pour extraire des informations sur la moyenne atmosphère à partir des données infrasons. Comme il est indiqué dans le chapitre 2, ces ensembles d'analyses et d'ébauches perturbées sont respectivement représentatifs de l'incertitude de l'analyse et de la prévision à courte échéance (6 h). Pour estimer l'incertitude sur les prévisions à plus longue échéance (jusqu'à 4 jours à Météo-France), on dispose d'un autre système : la prévision d'ensemble. La prévision d'ensemble du modèle ARPEGE, aussi appelée PEARP, a été envisagée dans nos travaux mais, pour des raisons décrites ci-après, les produits de l'assimilation d'ensemble lui ont été préférés. Cette annexe synthétise les études menées avec la PEARP.

B.1 Les paramétrisations physiques pour le modèle ARPEGE et la PEARP

Les prévisions d'ensemble utilisent entre autres les incertitudes sur les paramétrisations physiques pour générer ses différents membres. En premier lieu, on définit un ensemble d'états initiaux à partir desquels sont effectuées les prévisions. Pour la PEARP, ces états initiaux s'appuient sur des combinaisons linéaires entre les membres de l'ensemble d'ébauches perturbées de l'AEARP et des vecteurs singuliers¹ de l'atmosphère calculés dans plusieurs régions d'intérêt. Les perturbations des membres de l'ensemble d'ébauches utilisé ici, sont amplifiées par un coefficient d'inflation dépendant de l'altitude. Cette méthode permet de définir 34 des 35 états initiaux de la PEARP. Le 35^{ème} état est celui de référence et part de la prévision issue du modèle déterministe ARPEGE. On obtient un ensemble de 35 membres, pour initialiser les prévisions d'ensemble. Les champs météorologiques prévus seront calculés et stockés sur une grille de Gauss basculée et étirée de troncature T_L798 et de coefficient d'étirement 2.4. La résolution était de 10 km sur la France et de 60 km aux antipodes. Ils sont calculés pour 90 niveaux pression hybrides

1. perturbations initiales les plus instables sur des périodes de temps variant entre 18 et 24 h (Buizza & Palmer, 1995).

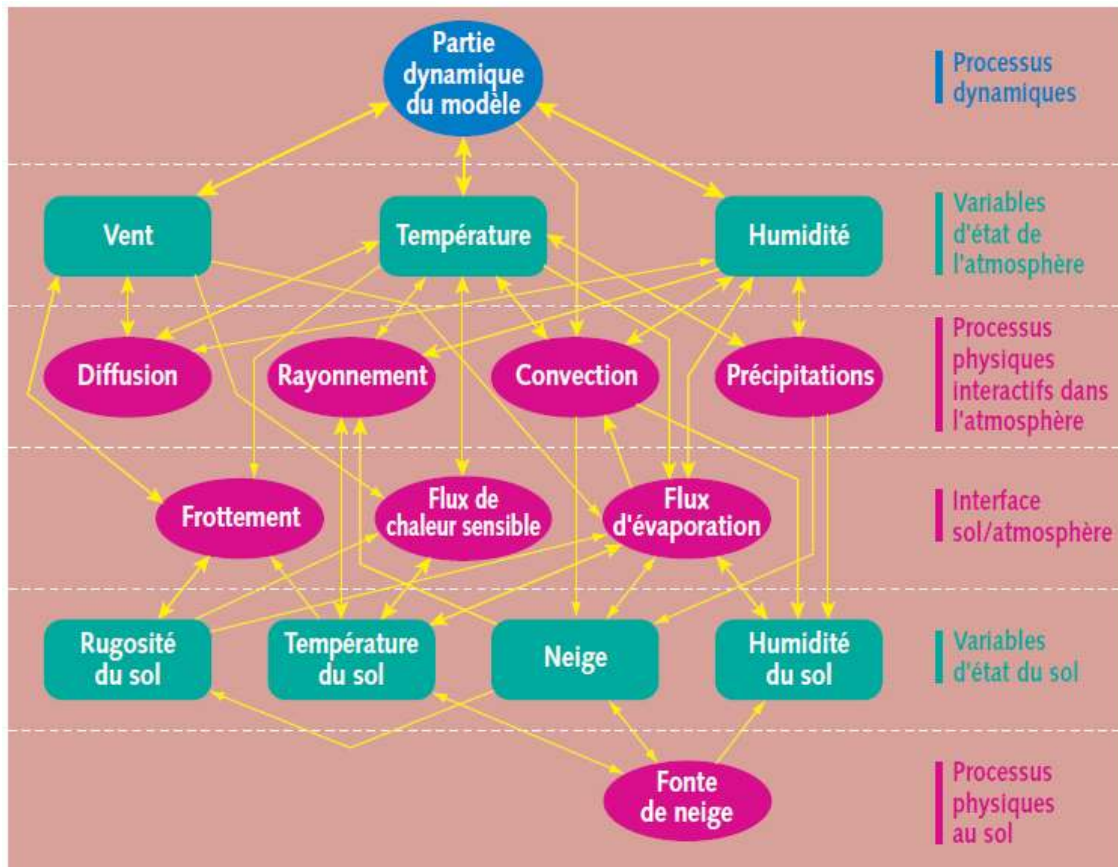


Figure B.1 : Schéma conceptuel des processus physiques qui doivent être introduits dans les modèles PNT et des interactions qui les lient. Tirée de Coiffier (2000).

dont le plus haut se trouve à 1 hPa (soit environ 50 km d'altitude). On perturbe ensuite les paramétrisations physiques différemment pour chacun des membres de la prévision d'ensemble afin de décrire les erreurs modèles.

B.1.1 Les paramétrisations physiques

La partie physique des modèles de PNT rend compte des phénomènes suivants :

- la diffusion par la turbulence de petite échelle dans la couche limite planétaire au moyen d'un schéma en énergie cinétique turbulent (TKE²) ;
- la convection peu profonde (nuages de la couche limite : cumulus, strato-cumulus) ;
- la convection profonde (nuages très développés sur la verticale comme les cumulonimbus ;
- les transferts radiatifs dans les domaines solaire et thermique incluant les interactions avec les nuages ;
- le cycle de l'eau atmosphérique : formation des nuages et des précipitations ;
- les effets de l'orographie comme les ondes de gravité et les blocages aérodynamiques des flux.

Les interactions entre la partie dynamique et les paramétrisations peuvent être schématisées de la manière proposée par la figure B.1.

Les paramétrisations influent sur les différentes variables d'état du modèle que sont la température, la vitesse du vent et la proportion des différentes phases de l'eau. Elles

2. Turbulent Kinetic Energy.

Numéro Physique	Diffusion turbulente	Convection peu profonde	Convection profonde	Rayonnement	Recouvrement Nuageux	Forçage orographique
0	oper	oper	oper	oper	oper	oper
1	oper	PCMT (Piriou <i>et al.</i> , 2007)	PCMT (Piriou <i>et al.</i> , 2007)	oper	oper	oper
2	Sans advection verticale	oper	oper	oper	oper	oper
3	oper	K-gradient formulation (Lancz <i>et al.</i> , 2018)	oper	oper	oper	oper
4	oper	oper	oper	oper	oper	déferlement onde de gravité
5	oper	oper	PCMT (Piriou <i>et al.</i> , 2007)	oper	oper	oper
6	oper	K-gradient formulation (Lancz <i>et al.</i> , 2018)	PCMT (Piriou <i>et al.</i> , 2007)	oper	oper	oper
7	oper	PCMT (Piriou <i>et al.</i> , 2007)	PCMT (Piriou <i>et al.</i> , 2007)	CEPMMT	MclCa jeu 1	oper
8	oper	PCMT (Piriou <i>et al.</i> , 2007)	PCMT (Piriou <i>et al.</i> , 2007)	CEPMMT	MclCa2 jeu 2	oper
9	oper	oper	oper	CEPMMT	MclCa2 jeu 2	oper

Tableau B.1 : Modifications apportées aux différentes paramétrisations de la PEARP vis-à-vis de celles utilisées pour la production déterministe du modèle ARPEGE, notée oper.

interagissent pour certaines avec le modèle de surface couplé au modèle de PNT dont les processus sont aussi paramétrés. Les interactions entre la surface et l'atmosphère sont générées par le modèle de surface SURFEX (Masson *et al.*, 2013). SURFEX décompose toute maille du globe en quatre grands types de surfaces : nature, lac, ville, mer. Au cours d'un pas de temps du modèle, SURFEX reçoit d'ARPEGE la température de l'air, l'humidité spécifique, la vitesse horizontale du vent, la pression, les précipitations totales et le forçage radiatif. Le modèle de surface SURFEX calcule et transmet alors à ARPEGE, les flux de surface moyens de quantité de mouvement, d'évaporation et de chaleur sensible, la température de surface, l'albédo de surface et l'émissivité de surface.

Dans la PEARP, les paramètres du modèle de surface ne sont actuellement pas perturbés.

B.1.2 PEARP et perturbations par approche multi-physiques

Alors que le CEPMMT perturbe de manière stochastique les tendances des paramétrisations physiques, Météo-France représente l'incertitude de ses prévisions par une approche multi-physiques (Descamps *et al.*, 2015). Pour chacun des 35 membres, une paramétrisation de la physique parmi un jeu de dix est tirée aléatoirement. Les différentes paramétrisations³ sont présentées au tableau B.1. On notera que les paramétrisations 0 sont celles du modèle ARPEGE déterministe. Le membre 0, ou membre de référence, utilise les mêmes paramètres et conditions initiales que la prévision déterministe d'ARPEGE.

Le schéma de la figure B.2 synthétise le fonctionnement de la PEARP. On part d'un ensemble d'ébauches perturbées issu de l'AEARP (analyse de référence + perturbation AEARP dans le schéma), puis on les combine avec les vecteurs singuliers afin de créer les états initiaux et enfin on leur associe une paramétrisation de la physique. L'écart au membre de référence augmente avec l'échéance de prévision à cause du comportement chaotique de l'atmosphère.

Si l'on compare la prévision à courte échéance que l'on a en un point spécifique avec celle à longue échéance comme cela est fait sur la figure B.3, on remarque que les perturbations s'amplifient et se propagent verticalement. Cependant la dispersion autour du membre de référence au delà de la tropopause (12 km) reste faible. Au bout de 102 h,

3. MclCa = Monte Carlo Independant Column approximation.(Morcrette *et al.*, 2008)

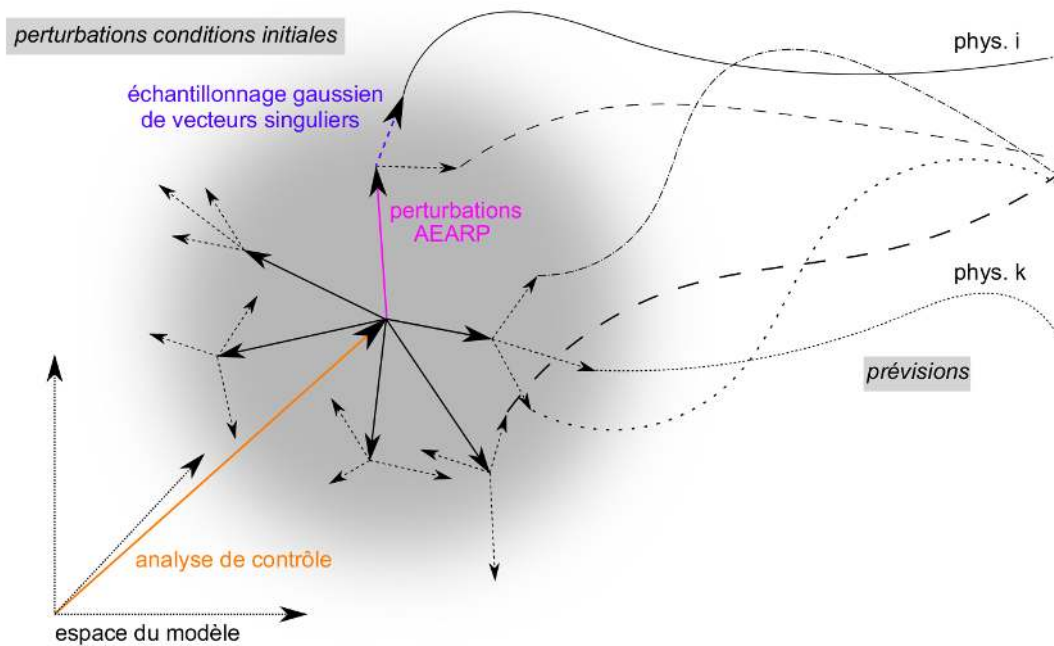


Figure B.2 : Schéma conceptuel de la construction de la PEARP. La figure est extraite de la description de la PEARP4 datant de Novembre 2017 et signée par C. Labadie, P. Cébron et L. Descamps.

la variabilité sur les profils dans la troposphère est de l'ordre de 50 m.s^{-1} et la valeur médiane s'éloigne de plus en plus du membre de référence.

Si l'on examine les champs de température vers 50 hPa à l'échelle globale comme sur la figure B.4, on constate que la moyenne des prévisions PEARP diffère peu de la prévision d'analyse ARPEGE déterministe. Dans ce cas particulier du réchauffement stratosphérique soudain de février 2018, même à 102 h, le champ de température prévu est en bon accord avec celui de l'analyse. Cependant, plus l'échéance de prévision est lointaine, plus les petites structures de l'atmosphère sont lissées comme le montre la figure B.4. Ainsi la moyenne des prévisions PEARP à 102 h, ne donne qu'une estimation fidèle des grandes échelles.

Le fait que le toit du modèle pour les prévisions d'ensemble PEARP se situe à 1 hPa (50 km d'altitude) et que la dispersion dans la stratosphère soit faible étaient problématiques pour les applications de la thèse orientées vers la moyenne atmosphère. C'est pourquoi, nous avons préféré utiliser les ensembles produits par l'AEARP, car disponible plus haut dans l'atmosphère et avec une plus grande variabilité dans la moyenne atmosphère. Néanmoins, pour d'autres types d'applications, mettant en jeu des réfractions dans la troposphère et dans la basse stratosphère, il serait intéressant d'appliquer la méthode bayésienne décrite dans le chapitre 5 avec les membres de la PEARP et, par exemple, d'examiner sa capacité à distinguer certains jeux de paramétrisations.

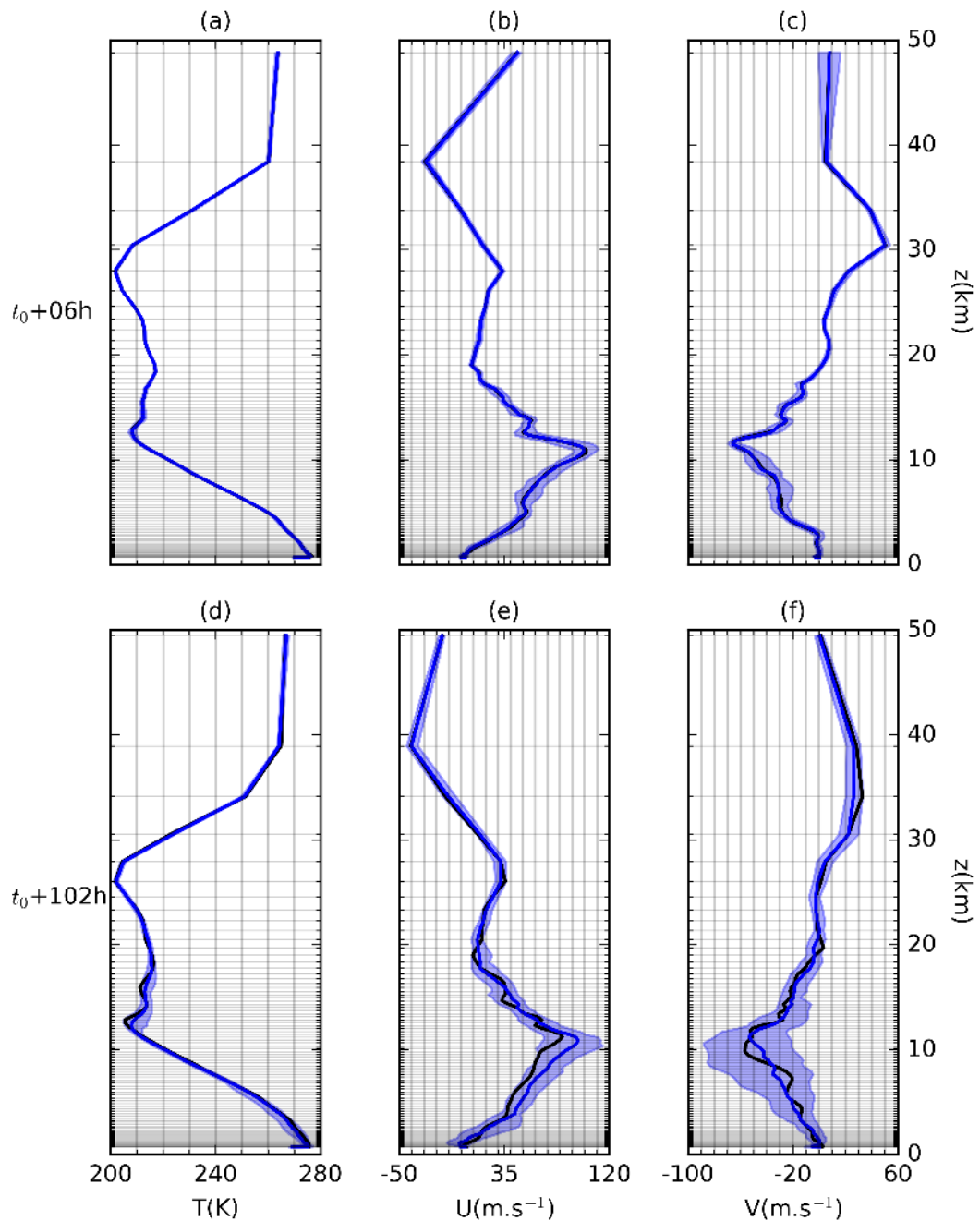


Figure B.3 : Comparaison des profils de température et de vitesse du vent prévus en un point spécifique. (a-b-c) prévision d'ensemble effectuée à partir d'un état initial daté de 6 heures auparavant. (d-e-f) prévision d'ensemble effectuée à partir d'un état initial daté de 102 heures auparavant. Les profils de l'état de référence sont en noir. Les profils du médian de l'ensemble sont en bleu. Les aires bleues en transparence représentent l'écart entre le 5ème percentile et le 95ème percentile. La grille horizontale correspond à la distribution des niveaux. L'intervalle entre les lignes verticales représente 10 unités.

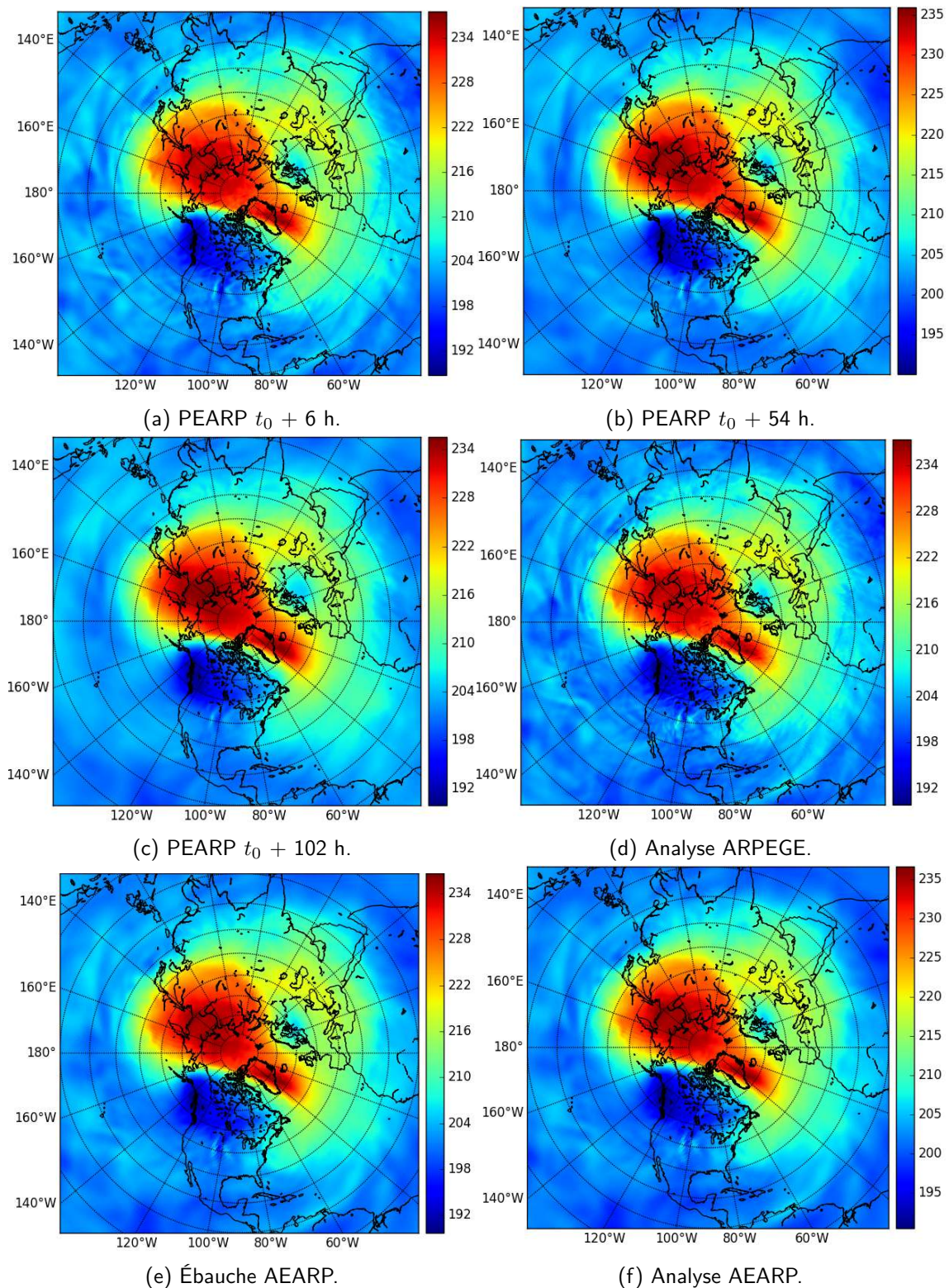


Figure B.4 : Comparatif de la température prévue par ARPEGE, PEARP et AEARP pour l'échéance du 15/02/2018 00 UTC à l'isobare 50 hPa. Les champs de température sont des moyennes d'ensemble (sauf pour l'analyse ARPEGE). L'échelle de couleur correspond à des températures exprimées en Kelvin.

Références

- Aitkin, M. (2010). *Statistical inference: an integrated Bayesian/likelihood approach*. Chapman and Hall/CRC, London. (ISBN : 978-1-4200-9343-8) doi : 10.1201/EBK1420093438
- Alexander, M. J. (2015). Global and seasonal variations in three-dimensional gravity wave momentum flux from satellite limb-sounding temperatures. *Geophysical Research Letters*, 42(16), 6860-6867. Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL065234> doi : 10.1002/2015GL065234
- Amezcuca, J., Näsholm, S. P., Blixt, E. M., & Charlton-Perez, A. J. (2020). Assimilation of atmospheric infrasound data to constrain tropospheric and stratospheric winds. *Quarterly Journal of the Royal Meteorological Society*, 146(731), 2634-2653. Consulté sur <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3809> doi : 10.1002/qj.3809
- Andrews, D. G., Leovy, C. B., & Holton, J. R. (1987). *Middle atmosphere dynamics* (Vol. 40). Academic press, Cambridge. (ISBN : 978-0-1205-8576-2)
- Angot, G. (2013). *Contribution des échauffements stratosphériques à la variabilité et à l'évolution à long terme de la moyenne atmosphère : observations et modélisations numériques* (Thèse de doctorat, Université Paris VI). Consulté sur <http://www.theses.fr/2013PA066655/document> (2013PA066655)
- Ardhuin, F., Stutzmann, E., Schimmel, M., & Mangeney, A. (2011). Ocean wave sources of seismic noise. *Journal of Geophysical Research: Oceans*, 116(C9). Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JC006952> doi : 10.1029/2011JC006952
- Assink, J., Le Pichon, A., Blanc, E., Kallel, M., & Khemiri, L. (2014). Evaluation of wind and temperature profiles from ECMWF analysis on two hemispheres using volcanic infrasound. *Journal of Geophysical Research: Atmospheres*, 119, 8659-8683. doi : 10.1002/2014JD021632
- Assink, J., Waxler, R., Frazier, W., & Lonzaga, J. (2013). The estimation of upper atmospheric wind model updates from infrasound data. *Journal of Geophysical Research: Atmospheres*, 118, 10707-10724.
- Assink, J., Waxler, R., & Velea, D. (2017). A wide-angle high Mach number modal expansion for infrasound propagation. *Journal of the Acoustical Society of America*, 141, 1781-1792.
- Attenborough, K. (1988). Review of ground effect on outdoor sound propagation from continuous broadband sources. *Applied acoustics*, 24, 289-319.
- Baer, F. (1972). An alternate scale representation of atmospheric energy spectra. *Journal of the Atmospheric Sciences*, 29(4), 649-664.
- Baldwin, M., Gray, L., Dunkerton, T., Hamilton, K., Haynes, P., Randel, W., Holton, J., Alexander, M., Hirota, I., Horinouchi, T., Jones, D., Kinnnersley, J., Marquardt, C., Sato, K., & Takahashi, M. (2001). The quasi-biennial oscillation. *Reviews of Geophysics*, 39(2), 179-229.
- Baldwin, M., Stephenson, D. B., Thompson, D., Dunkerton, T. J., Charlton, A., & O'Neill,

- A. (2003). Stratospheric memory and skill of extended-range weather forecasts. *Science*, 301(5633), 636–640.
- Barnett, J., & Corney, M. (1985). Middle atmosphere reference model derived from satellite data. In *International council of scientific unions middle atmosphere program. handbook for map, vol. 16 p 47-85* (Vol. 16).
- Baumgarten, G. (2010). Doppler Rayleigh/Mie/Raman lidar for wind and temperature measurements in the middle atmosphere up to 80 km. *Atmospheric Measurement Techniques*, 3(6), 1509–1518. Consulté sur <https://www.atmos-meas-tech.net/3/1509/2010/> doi : 10.5194/amt-3-1509-2010
- Bergmann, P. (1946). The wave equation in a medium with a variable index of refraction. *Journal of the Acoustical Society of America*, 17(4), 329-333.
- Berre, L., Varella, H., & Desroziers, G. (2015). Modelling of flow-dependent ensemble based background-error correlations using a wavelet formulation in 4D-Var at Météo-France. *Quarterly Journal of the Royal Meteorological Society*, 141, 2803-2812.
- Bertin, M. (2014). *Propagations des incertitudes dans un modèle réduit de propagation d'infrasons* (Thèse de doctorat, ENS Cachan). Consulté sur <http://www.theses.fr/2014DENSO020/document> (Thèse de doctorat dirigée par Bouché, Daniel Sciences Cachan, École normale supérieure 2014)
- Bertin, M., Millet, C., & Bouche, D. (2014). A low-order reduced model for the long range propagation of infrasounds in the atmosphere. *Journal of the Acoustical Society of America*, 136, 37-52.
- Blanc, E., Pol, K., Le Pichon, A., Hauchecorne, A., Keckhut, P., Baumgarten, G., Hildebrand, J., J.and Höffner, Stober, G., Hibbins, R., Espy, P., Rapp, M., Kaifler, B., Ceranna, L., Hupe, P., Hagen, J., Rüfenact, R., Kämpfer, N., & Smets., P. (2018). Toward an improved representation of middle atmospheric dynamics thanks to the ARISE project. *Surveys in geophysics*, 39(2), 171–225.
- Blokhintzev, D. (1946). The propagation of sound in an inhomogeneous and moving medium I. *Journal of the Acoustical Society of America*, 18(2), 322.
- Blom, P. (2019). Modeling infrasonic propagation through a spherical atmospheric layer-Analysis of the stratospheric pair. *Journal of the Acoustical Society of America*, 145(4), 2198-2208. doi : 10.1121/1.5096855
- Blom, P., & Waxler, R. (2012). Impulse propagation in the nocturnal boundary layer: Analysis of the geometric component. *Journal of the Acoustical Society of America*, 131, 3680-3690. doi : 10.1121/1.3699174
- Blom, P., & Waxler, R. (2017). Modeling and observations of an elevated, moving infrasonic source: Eigenray methods. *Journal of the Acoustical Society of America*, 141, 2681-2692.
- Brachet, N., Brown, D., Le Bras, R., Cansi, Y., Mialle, P., & Coyne, J. (2010). *Monitoring the Earth's atmosphere with the global IMS infrasound network*. Springer, Dordrecht : Le Pichon, A. and Blanc, E. and Hauchecorne, A. (ISBN : 978-1-4020-9507-8)
- Broyden, C. (1970). The convergence of a class of double-rank minimization algorithms 1. General considerations. *IMA Journal of Applied Mathematics*, 6(1), 76-90. doi : 10.1093/imamat/6.1.76
- Buizza, R., & Palmer, T. N. (1995). The singular-vector structure of the atmospheric global circulation. *Journal of the Atmospheric Sciences*, 52(9), 1434–1456.
- Campus, P., & Christie, D. (2010). *Worldwide observations of infrasonic waves*. Springer, Dordrecht : Le Pichon, Alexis and Blanc, Elisabeth and Hauchecorne, Alain. (ISBN : 978-1-4020-9507-8)
- Cannata, A., Di Grazia, G., Giuffrida, M., Gresta, S., Palano, M., Sciotto, M., Viccaro, M., & Zuccarello, F. (2017). Space-time evolution of magma storage and transfer at

- Mt. Etna volcano (Italy): The 2015-2016 reawakening of Voragine crater. *Journal of Geophysical Research: Geochemistry, Geophysics, Geosystems*, 19, 471-495. doi : 10.1002/2017GC007296
- Cansi, Y. (1995). An automatic seismic event processing for detection and location : The P.M.C.C. method. *Geophysical Research Letters*, 22, 1021-1024. doi : 10.1029/95GL00468
- Canuto, C., Hussaini, M. Y., Quarteroni, A., & Zang, T. A. (2006). *Spectral methods in fluid dynamics*. Springer-Verlag, Berlin. (ISBN : 978-3-540-52205-08)
- Chabot, V., Berre, L., & Desroziers, G. (2017). Diagnosis and normalization of gridpoint background-error variances induced by a block-diagonal wavelet covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, 143, 1268-1279.
- Chanin, M. L., Garnier, A., Hauchecorne, A., & Porteneuve, J. (1989). A Doppler lidar for measuring winds in the middle atmosphere. *Geophysical Research Letters*, 16(11), 1273-1276. Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL016i011p01273> doi : 10.1029/GL016i011p01273
- Chapnik, B., Desroziers, G., Rabier, F., & Talagrand, O. (2004). Properties and first application of an error-statistics tuning method in variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 130(601), 2253-2275.
- Charlton, A. J., & Polvani, L. M. (2007). A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *Journal of Climate*, 20(3), 449-469. doi : 10.1175/JCLI3996.1
- Chatillon, J. (2011). Perception des infrasons. *Acoustique & Techniques*, 67, 4-10.
- Chunchuzov, I. P., Kulichkov, S., Popov, O., Perepelkin, V., Vasil'ev, A., Glushkov, A., & Firstov, P. (2015). Characteristics of a fine vertical wind-field structure in the stratosphere and lower thermosphere according to infrasonic signals in the zone of acoustic shadow. *Izvestiya, Atmospheric and Oceanic Physics*, 51(1), 57-74.
- Chunchuzov, I. P., Kulichkov, S. N., Popov, O. E., Waxler, R., & Assink, J. (2011, 22). Infrasound scattering from atmospheric anisotropic inhomogeneities. *Izvestiya, Atmospheric and Oceanic Physics*, 47(5), 540. Consulté sur <https://doi.org/10.1134/S0001433811050045> doi : 10.1134/S0001433811050045
- Coiffier, J. (2000). Un demi-siècle de prévision numérique du temps. *La Météorologie*, 8(30), 11-31.
- Courtier, P., & Geleyn, J.-F. (1988). A global numerical weather prediction model with variable resolution: Application to the shallow-water equations. *Quarterly Journal of the Royal Meteorological Society*, 114(483), 1321-1346.
- Courtier, P., Thépaut, J.-N., & Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120, 1367-1387.
- Curtis, F., Mitchell, T., & Overton, M. (2017). A BFGS-SQP method for nonsmooth, nonconvex, constrained optimization and its evaluation using relative minimization profiles. *Optimization Methods and Software*, 32, 148-181.
- Daley, R. (1993). *Atmospheric data analysis* (N° 2). Cambridge university press, Cambridge. (ISBN : 978-0-5214-5825-2)
- Damiens, F., Millet, C., & Lott, F. (2018). An investigation of infrasound propagation over mountain ranges. *Journal of the Acoustical Society of America*, 143, 563-574. doi : 10.1121/1.5020783
- De Carlo, M., Arduin, F., & Le Pichon, A. (2020). Atmospheric infrasound radiation from ocean wave in finite depth: A unified generation theory and application to radiation patterns. *Geophysical Journal International*, 221, 568-585. doi : 10.1093/gji/ggaa015

- Dee, D., Uppala, S., Simmons, A., Berrisford, P., Polo, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Belsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553-597. doi : 10.1002/qj.828
- De Groot, S., & Mazur, P. (1963). Non-equilibrium thermodynamics. *American Journal of Physics*, *31*(7), 558-559.
- De Groot-Hedlin, C. D., & Hedlin, M. A. H. (2014). Infrasound detection of the Chelyabinsk meteor at the USArray. *Earth and Planetary Science Letters*, *402*, 337-345. doi : 10.1016/j.epsl.2014.01.031
- Descamps, L., Labadie, C., Joly, A., Bazile, E., Arbogast, P., & Cébron, P. (2015). PEARP, the Météo-France short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *141*, 1671-1685.
- Desroziers, G., & Ivanov, S. (2001). Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, *127*(574), 1433-1452.
- Donn, W. L., & Rind, D. (1971). Natural infrasound as an atmospheric probe. *Geophysical Journal International*, *26*(1-4), 111-133.
- Drob, D. P., Broutman, D., Hedlin, M., Winslow, N. W., & Gibson, R. G. (2013). A method for specifying atmospheric gravity wavefields for long-range infrasound propagation calculations. *Journal of Geophysical Research: Atmospheres*, *118*, 3933-3943. doi : 10.1029/2012JD018077
- Drob, D. P., Emmert, J. T., Crowley, G., Picone, J. M., Shepherd, T. G., Skinner, W., Hays, P., Niciejewski, R. J., Larsen, M., She, C. Y., Meriwether, J. W., Hernandez, G., Jarvis, M. J., Sipler, D. P., Tepley, C. A., O'Brien, M. S., Bowman, J. R., Wu, Q., Murayama, Y., Kawamura, S., Reid, I. M., & Vincent, R. A. (2008). An empirical model of the earth's horizontal wind fields: HWM07. *Journal of Geophysical Research: Space Physics*, *113*(A12). Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JA013668> doi : 10.1029/2008JA013668
- Drob, D. P., Emmert, J. T., Meriwether, J., Makela, J., Doornbos, E., Conde, M., Hernandez, G., Noto, J., Zawdie, K. A., McDonald, S., Huba, J., & Klenzing, J. (2015). An update to the horizontal wind model (HWM): The quiet time thermosphere. *Earth and Space Science*, *2*. doi : 10.1002/2014EA000089
- Drob, D. P., Meier, R., Picone, J., & Garcès, M. (2010). *Inversion of infrasound signals for passive atmospheric remote sensing*. Springer, Dordrecht : A. Le Pichon, E. Blanc and A. Hauchecorne. (ISBN : 978-1-4020-9507-8)
- Drob, D. P., Picone, J., & Garcès, M. (2003). Global morphology of infrasound propagation. *Journal of Geophysical Research: Atmospheres*, *108*, 4680-4692.
- Duruiseau, F., Chambon, P., Wattrelot, E., Barreyat, M., & Mahfouf, J.-F. (2019). Assimilating cloudy and rainy microwave observations from SAPHIR on board Megha Tropiques within the ARPEGE global model. *Quarterly Journal of the Royal Meteorological Society*, *145*(719), 620-641. Consulté sur <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3456> doi : 10.1002/qj.3456
- Edwards, M., Pioli, L., Andronico, D., Scollo, S., Ferrari, F., & Cristaldi, A. (2018). Shallow factors controlling the basaltic magmas : The 17-25 May 2016 eruption of Etna Volcano (Italy). *Journal of Volcanology and Geothermal Research*, *357*, 425-436. doi : 10.1016/j.jvolgeores.2018.05.015

- Evers, L., & Haak, H. (2010). *The characteristics of infrasound, its propagation and some early history*. Springer, Dordrecht : A. Le Pichon, E. Blanc and A. Hauchecorne. (ISBN : 978-1-4020-9507-8)
- Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal*, 13(3), 317-322. doi : 10.1093/comjnl/13.3.317
- Gainville, O. (2008). *Modélisation de la propagation atmosphérique des ondes infrasonores par une méthode de tracé de rayons non linéaires* (Thèse de doctorat, Ecole centrale Lyon). Consulté sur <http://www.theses.fr/2008ECDL0007/document> (Thèse de doctorat dirigée par Blanc-Benon, Philippe Acoustique Ecully, École centrale de Lyon 2008)
- Gao, H., Xu, J., Ward, W., & Smith, A. K. (2011). Temporal evolution of nightglow emission responses to SSW events observed by TIMED/SABER. *Journal of Geophysical Research: Atmospheres*, 116(D19).
- Gardner, C., Hostetler, C., & Franke, S. (1993). Gravity wave models for the horizontal wave number spectra of atmospheric velocity and density fluctuations. *Journal Geophysical Research: Atmospheres*, 98, 1035-1049. doi : 10.1029/92JD02051
- Georges, T. M., & Beasley, W. H. (1977). Refraction of infrasound by upper atmospheric winds. *Journal of the Acoustical Society of America*, 61(1), 28-34. Consulté sur <https://doi.org/10.1121/1.381263> doi : 10.1121/1.381263
- Godin, O. (2002). An effective quiescent medium for sound propagating through inhomogeneous, moving fluid. *Journal of the Acoustical Society of America*, 112, 1269-1275.
- Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109), 23-26.
- Gossard, E., & Hooke, W. (1975). *Waves in the atmosphere: Atmospheric infrasound and gravity waves, their generation and propagation (developments in atmospheric science)*. Elsevier Science Ltd, Amsterdam. (ISBN : 978-0444411969)
- Gottlieb, D., Hussaini, M. Y., & Orszag, S. A. (1984). Theory and applications of spectral methods. In *Proc. of the symposium of spectral methods for partial differential equations* (p. 1-55).
- Hagen, J., Murk, A., Rüfenacht, R., Khaykin, S., Hauchecorne, A., & Kämpfer, N. (2018). WIRA-C : a compact 142-GHz-radiometer for continuous middle-atmospheric wind measurements. *Atmospheric Measurement Techniques*, 11(9), 5007-5024. Consulté sur <https://www.atmos-meas-tech.net/11/5007/2018/> doi : 10.5194/amt-11-5007-2018
- Hanique-Cockenpot, G. (2011). *Étude numérique de la propagation non linéaire des infrasons dans l'atmosphère* (Thèse de doctorat, École centrale Lyon). Consulté sur <http://www.theses.fr/2011ECDL0032> (Thèse de doctorat en acoustique dirigée par Bailly, Christophe et Marsden, Olivier à Ecully, École centrale de Lyon 2011)
- Hedin, A. E. (1983). A revised thermospheric model based on mass spectrometer and incoherent scatter data: MSIS-83. *Journal of Geophysical Research: Space Physics*, 88(A12), 10170-10188.
- Hedin, A. E. (1987). MSIS-86 thermospheric model. *Journal of Geophysical Research: Space Physics*, 92(A5), 4649-4662. doi : 10.1029/JA092iA05p04649
- Hedin, A. E. (1991). Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research: Space Physics*, 96(A2), 1159-1172. Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/90JA02125> doi : 10.1029/90JA02125
- Hedin, A. E., Spencer, N., & Killeen, T. (1988). Empirical global model of upper

- thermosphere winds based on Atmosphere and Dynamics Explorer satellite data. *Journal of Geophysical Research: Space Physics*, 93(A9), 9959–9978.
- Hedlin, M. A. H., & Drob, D. P. (2014). Statistical characterization of atmospheric gravity waves by seismoacoustic observations. *Journal of Geophysical Research: Atmospheres*, 119(9), 5345–5363. doi : 10.1002/2013jd021304
- Hertzog, A., Alexander, M. J., & Plougonven, R. (2012). On the intermittency of gravity wave momentum flux in the stratosphere. *Journal of the Atmospheric Sciences*, 69(11), 3433–3448. Consulté sur <https://doi.org/10.1175/JAS-D-12-09.1> doi : 10.1175/JAS-D-12-09.1
- Hertzog, A., Boccara, G., Vincent, R. A., Vial, F., & Cocquerez, P. (2008). Estimation of gravity wave momentum flux and phase speeds from quasi-Lagrangian stratospheric balloon flights. Part II: Results from the Vorcore campaign in Antarctica. *Journal of the Atmospheric Sciences*, 65(10), 3056–3070. Consulté sur <https://doi.org/10.1175/2008JAS2710.1> doi : 10.1175/2008JAS2710.1
- Hines, C. O. (1997). Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part II: Broad and quasi monochromatic spectra, and implementation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 59(4), 387–400. Consulté sur <http://www.sciencedirect.com/science/article/pii/S1364682696000806> doi : [https://doi.org/10.1016/S1364-6826\(96\)00080-6](https://doi.org/10.1016/S1364-6826(96)00080-6)
- Holton, J. (2004). *An introduction to dynamic meteorology, 4th edition* (Vol. 88). Academic Press, Cambridge. (ISBN : 978-0-1235-4015-7)
- Hortal, M., & Simmons, A. (1991). Use of reduced Gaussian grids in spectral models. *Monthly Weather Review*, 119(4), 1057–1074.
- Jensen, F., Kuperman, W., Porter, M., & H, S. (2011). *Computational ocean acoustics (Modern acoustics and signal processing) 2nd Edition*. Springer, New-York. (ISBN : 978-1-4939-3704-2)
- Khaykin, S., Hauchecorne, A., Mzé, N., & Keckhut, P. (2015). Seasonal variation of gravity wave activity at midlatitudes from 7 years of COSMIC GPS and Rayleigh lidar temperature observations. *Geophysical Research Letters*, 42(4), 1251–1258.
- Killeen, T. L. (1987). Energetics and dynamics of the Earth's thermosphere. *Reviews of Geophysics*, 25(3), 433–454.
- Kulichkov, S. (2010). *On the prospect for acoustic sounding of the fine structure of the middle-atmosphere, in infrasound monitoring for atmospheric studies*. Springer, Dordrecht : Le Pichon, A. and Blanc, E. and Hauchecorne, A. (ISBN : 978-1-4020-9507-8)
- Kulichkov, S., Popov, O. Y., Mishenin, A., Chunchuzov, I. P., Chkhetiani, O., & Tsybul'skaya, N. (2017). Using the acoustic-pulse conservation law in estimating the energy of surface acoustic sources by remote sounding. *Izvestiya, Atmospheric and Oceanic Physics*, 53(6), 603–612.
- Lacroix, A., Coulouvrat, F., Marchiano, R., Farges, T., & Ripoll, J.-F. (2019). Acoustical energy of return strokes: A comparison between a statistical model and measurements. *Geophysical Research Letters*, 46, 11479–11489. doi : 10.1029/2019GL085369
- Lacroix, A., Farges, T., Marchiano, R., & Coulouvrat, F. (2018). Acoustical measurement of natural lightning flashes: Reconstructions and statistical analysis of energy spectra. *Journal of Geophysical Research: Atmospheres*, 123(21), 12–040.
- Lalande, J.-M. (2012). *Caractérisation des vents dans la moyenne atmosphère et basse thermosphère à partir d'observations d'ondes infrasonores* (Thèse de doctorat, École centrale Lyon). Consulté sur <http://www.theses.fr/2012ECDL0015/document> (Thèse de doctorat dirigée par Blanc-Benon, Philippe et Blanc, Elisabeth Acoustique

- Ecully, École centrale de Lyon 2012)
- Lalande, J.-M., & Waxler, R. (2016). The interaction between infrasonic waves and gravity wave perturbations: Application to observations using UTTR rocket motor fuel elimination events. *Journal of Geophysical Research: Atmospheres*, *121*, 5585-5600.
- Lancz, D., Szintai, B., & Honnert, R. (2018). Modification of a parametrization of shallow convection in the grey zone using a mesoscale model. *Boundary-Layer Meteorology*, *169*(3), 483-503.
- Le Pichon, A., Blanc, E., Drob, D. P., Lambotte, S., Dessa, J. X., Lardy, M., Bani, P., & Vergnolle, S. (2005). Infrasound monitoring of volcanoes to probe high-altitude winds. *Journal of Geophysical Research: Atmospheres*, *110*, D13106. doi : 10.1029/2004JD005587
- Le Pichon, A., De Carlo, M., Ardhuin, F., Arnal, T., & Cerana, L. (2018). Characterizing ocean ambient noise using infrasound network. *Journal of the Acoustical Society of America*, *143*(3), 1808-1808. Consulté sur <https://doi.org/10.1121/1.5035923> doi : 10.1121/1.5035923
- Leuenberger, D., Koller, M., Fuhrer, O., & Schär, C. (2010). A generalization of the SLEEVE vertical coordinate. *Monthly Weather Review*, *138*(9), 3683-3689.
- Lighthill, J. (1978). Acoustic streaming. *Journal of Sound and Vibration*, *61*(3), 391-418.
- Lions, J. L. (1971). *Optimal control of systems governed by partial differential equations*. Springer-Verlag, Berlin. (ISBN : 978-3-642-65026-0)
- Machenhauer, B., & Rasmussen, E. (1972). *On the integration of the spectral hydrodynamical equations by a transform method*. Institut de météorologie de l'Université de Copenhague. Consulté sur https://www.researchgate.net/publication/24362044_On_the_integration_of_the_spectral_hydrodynamical_equations_by_a_transform_method
- Marsden, O., Bogey, C., & Bailly, C. (2014). A study of infrasound propagation based on high-order finite difference solutions of the Navier-Stokes equations. *Journal of the Acoustical Society of America*, *135*(3), 1083-1095. Consulté sur <https://doi.org/10.1121/1.4864793> doi : 10.1121/1.4864793
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., & Voldoire, A. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of Earth surface variables and fluxes. *Geoscientific Model Development*, *6*, 929-960. Consulté sur <https://hal.archives-ouvertes.fr/hal-00968042> doi : 10.5194/gmd-6-929-2013
- Matsuno, T. (1971). A dynamical model of the stratospheric sudden warming. *Journal of the Atmospheric Sciences*, *28*(8), 1479-1494.
- McKisic, J. M. (1997). *Infrasound and the infrasonic monitoring of atmospheric nuclear explosions : A literature review* (Rapport technique). 1601 Research Boulevard Rockville, Maryland, 20850 : Tracor applied sciences Inc Rockville MD.
- Millet, C. (2015). Infrasound propagation and model reduction in randomly layered media. *Journal of the Acoustical Society of America*, *137*, 2372. doi : <https://doi.org/10.1121/1.492060>
- Millet, C., Bertin, M., & Mialle, P. (2018). Bayesian association of multiple infrasound events using long-range propagation models. *Journal of the Acoustical Society of America*, *144*(3), 1728-1728.

- Millet, C., Robinet, J.-C., & Roblin, C. (2007). On using computational aeroacoustics for long-range propagation of infrasounds in realistic atmospheres. *Geophysical Research Letters*, *34*(14). doi : 10.1029/2007GL029449
- Morcrette, J., Barker, H. W., Cole, J., Iacono, M. J., & Pincus, R. (2008). Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. *Monthly Weather Review*, *136*(12), 4773-4798.
- Nocedal, J., & Wright, S. (2006). *Numerical optimization second edition*. Springer, New-York. (ISBN : 978-1-4939-3711-0)
- Ostashev, V. E., Wilson, D. K., Liu, L., Aldridge, D. F., Symons, N. P., & Marlin, D. (2005). Equations for finite-difference, time-domain simulation of sound propagation in moving inhomogeneous media and numerical implementation. *Journal of the Acoustical Society of America*, *117*(2), 503-517. doi : 10.1121/1.1841531
- Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. (2002). NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research: Space Physics*, *107*(A12). Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JA009430> doi : 10.1029/2002JA009430
- Pierce, A. (1965). Extension of the method of normal modes to sound propagation in an almost-stratified medium. *Journal of the Acoustical Society of America*, *37*, 218-227.
- Piriou, J.-M., Redelsperger, J.-L., Geleyn, J.-F., Lafore, J.-P., & Guichard, F. (2007). An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations. *Journal of the Atmospheric Sciences*, *64*(11), 4127-4139. Consulté sur <https://doi.org/10.1175/2007JAS2144.1> doi : 10.1175/2007JAS2144.1
- Plougonven, R., Hertzog, A., & Guez, L. (2013). Gravity waves over Antarctica and the Southern ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. *Quarterly Journal of the Royal Meteorological Society*, *139*(670), 101-118. Consulté sur <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1965> doi : 10.1002/qj.1965
- Polavarapu, S., Shepherd, T. G., Rochon, Y., & Ren, S. (2005). Some challenges of middle atmosphere data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *131*, 3513-3527. doi : 10.1256/qj.05.87
- Qaddouri, A., & Lee, V. (2011). The Canadian Global Environmental Multiscale model on the Yin-Yang grid system. *Quarterly Journal of the Royal Meteorological Society*, *137*(660), 1913-1926.
- Randel, W., Udelhofen, P., Fleming, E., Geller, M., Gelman, M., Hamilton, K., Karoly, D., Ortland, D., Pawson, S., Swinbank, R., Wu, F., Baldwin, M., Chanin, M.-L., Keckhut, P., Labitzke, K., Remsberg, E., Simmons, A., & Wu, D. (2004). The SPARC intercomparison of middle-atmosphere climatologies. *Journal of Climate*, *17*(5), 986-1003.
- Raynaud, L., Berre, L., & Desroziers, G. (2012). Accounting for model error in the Météo-France ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *138*, 249-262.
- Reed, R. J., Campbell, W. J., Rasmussen, L. A., & Rogers, D. G. (1961). Evidence of a downward-propagating, annual wind reversal in the equatorial stratosphere. *Journal of Geophysical Research*, *66*(3), 813-818.
- Rind, D., & Donn, W. (1975). Further use of natural infrasound as a continuous monitor of the upper atmosphere. *Journal of the Atmospheric Sciences*, *32*, 1694-1704. doi : 10.1175/1520-0469(1975)032<1694:FUONIA>2.0.CO;2

- Rind, D., Donn, W., & Dede, E. (1973). Upper air wind speeds calculated from observations of natural infrasound. *Journal of the Atmospheric Sciences*, 30(8), 1726–1729.
- Rind, D., & Donn, W. L. (1978). Infrasound observations of variability during stratospheric warmings. *Journal of the Atmospheric Sciences*, 35(3), 546–553.
- Rishbeth, H. (1972). Thermospheric winds and the F-region: A review. *Journal of Atmospheric and Terrestrial Physics*, 34(1), 1–47.
- Romanowicz, B. (2003). Global mantle tomography: progress status in the past 10 years. *Annual Review of Earth and Planetary Sciences*, 31(1), 303–328.
- Sabatini, R., Marsden, O., Bailly, C., & Bogey, C. (2016). A numerical study of nonlinear infrasound propagation in a windy atmosphere. *Journal of the Acoustical Society of America*, 140, 641–656. doi : 10.1121/1.4958998
- Sabatini, R., Marsden, O., Bailly, C., & Gainville, O. (2019). Three-dimensional direct numerical simulation of infrasound propagation in the Earth's atmosphere. *Journal of Fluid Mechanics*, 859, 754–789. doi : 10.1017/jfm.2018.816
- Saint-Martin, D. (2010). *Étude comparative du rôle de la dynamique et la chimie dans la modélisation de l'atmosphère moyenne* (Thèse de doctorat, Université de Toulouse III Paul Sabatier). Consulté sur <http://www.theses.fr/2010T0U30253/document> (Thèse de doctorat dirigée par Cariolle, Daniel Physicochimie de l'atmosphère Toulouse 3 2010)
- Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., & Lupu, C. (2018). An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7), 2717–2737. Consulté sur <https://www.geosci-model-dev.net/11/2717/2018/> doi : 10.5194/gmd-11-2717-2018
- Schwaiger, H. F., Iezzi, A. M., & Fee, D. (2019). AVO-G2S : A modified, open-source Ground-to-Space atmospheric specification for infrasound modeling. *Computers & Geosciences*, 125, 90–97. Consulté sur <http://www.sciencedirect.com/science/article/pii/S0098300418301560> doi : <https://doi.org/10.1016/j.cageo.2018.12.013>
- Scinocca, J. F. (2003). An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *Journal of the Atmospheric Sciences*, 60(4), 667–682. Consulté sur [https://doi.org/10.1175/1520-0469\(2003\)060<0667:AASNGW>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0667:AASNGW>2.0.CO;2) doi : 10.1175/1520-0469(2003)060<0667:AASNGW>2.0.CO;2
- Shanno, D. F. (1970). Conditioning of Quasi-Newton methods for function minimization. *Mathematics of Computation*, 24(111), 647–656.
- Simmons, A., & Burridge, D. (1981). An energy and angular momentum conserving vertical finite difference scheme and hybrid vertical coordinates. *Monthly Weather Review*, 109, 758–766.
- Smets, P. (2018). *Infrasound and the Dynamical Stratosphere : A new application for operational weather and climate prediction* (Thèse de doctorat, TU Delft Applied Geophysics and Petrophysics). (Thèse de doctorat délivrée par l'Université Technique de Delft) doi : 10.4233/uuid:517f8597-9c24-4d01-83ed-0f430353e905
- Smets, P., Evers, L., Nasholm, S., & Gibbons, S. (2015). Probabilistic infrasound propagation using realistic atmospheric perturbations. *Geophysical Research Letters*, 42, 6510–6517. doi : 10.1002/2015GL064992
- Sutherland, L., & Bass, H. (2004). Atmospheric absorption in the atmosphere up to 160 km. *Journal of the Acoustical Society of America*, 115, 1012–1032.
- Symons, G., Judd, J., Strachey, R., Wharton, W., Evans, F., Russell, F., Archibald, D., & Whipple, G. (1888). The eruption of Krakatoa, and subsequent phenomena.

- Report of the Krakatoa Committee of the Royal Society.*
- Szuberla, C., & Olson, J. V. (2004). Uncertainties associated with parameter estimation in atmospheric infrasound arrays. *Journal of the Acoustical Society of America*, 115, 253-258. doi : 10.1121/1.1635407
- Tailpied, D., Le Pichon, A., Marchetti, E., Ripepe, M., Kallel, M., Ceranna, L., & Brachet, N. (2013). Remote infrasound monitoring of Mount Etna: Observed and predicted network detection capability. *InfraMatics*, 2(01), 1-11.
- Talagrand, O. (1997). Assimilation of observations, an introduction. *Journal of Meteorological Society of Japan*, 75, 191-209.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society of Industrial and Applied Mathematics, Philadelphia. (ISBN : 0-89871-572-5)
- Vanderbecken, P., Mahfouf, J.-F., & Millet, C. (2020). Bayesian selection of atmospheric profiles from an ensemble data assimilation system using infrasonic observations of May 2016 Mount Etna eruptions. *Journal of Geophysical Research: Atmospheres*(2), e2019JD31168. doi : 10.129/2019JD31168
- Vargin, P., & Kiryushov, B. (2019). Major sudden stratospheric warming in the Arctic in February 2018 and its impacts on the troposphere, mesosphere, and ozone layer. *Russian Meteorology and Hydrology*, 44(2), 112-123.
- Vergnolle, S., Brandeis, G., & Mareschal, J.-C. (1996). Strombolian explosions: 2. eruption dynamics determined from acoustic measurements. *Journal of Geophysical Research: Solid Earth*, 101(B9), 20449-20466. doi : 10.1029/96JB01925
- Veryard, R. G., & Ebdon, R. A. (1961). Fluctuations in tropical stratospheric winds. *Meteorology Magazine*, 90, 125-143.
- Warner, C. D., & McIntyre, M. E. (1996). On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere. *Journal of the Atmospheric Sciences*, 53(22), 3213-3235. Consulté sur [https://doi.org/10.1175/1520-0469\(1996\)053<3213:OTPADO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<3213:OTPADO>2.0.CO;2) doi : 10.1175/1520-0469(1996)053<3213:OTPADO>2.0.CO;2
- Wattrelot, E., Caumont, O., & Mahfouf, J.-F. (2014). Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Monthly Weather Review*, 142(5), 1852-1873. doi : 10.1175/MWR-D-13-00230.1
- Waxler, R. (2002). A vertical eigenfunction expansion for the propagation of sound in a downward-refracting atmosphere over a complex impedance plane. *Journal of the Acoustical Society of America*, 112(6), 2540-2552. doi : 10.1121/1.1514930
- Waxler, R., Evers, L. G., Assink, J., & Blom, P. (2015). The stratospheric arrival pair in infrasound propagation. *Journal of the Acoustical Society of America*, 137(4), 1846-1856.
- Waxler, R., & Gilbert, K. E. (2006). The radiation of atmospheric microbaroms by ocean waves. *Journal of the Acoustical Society of America*, 119(5), 2651-2664. Consulté sur <https://doi.org/10.1121/1.2191607> doi : 10.1121/1.2191607
- Weill, A., Brandeis, G., Vergnolle, S., Baudin, F., Bilbille, J., Fèvre, J.-F., Piron, B., & Hill, X. (1992). Acoustic sounder measurements of the vertical velocity of volcanic jets at Stromboli volcano. *Geophysical Research Letters*, 19(23), 2357-2360.
- Wright, C. J., Osprey, S. M., & Gille, J. C. (2013). Global observations of gravity wave intermittency and its impact on the observed momentum flux morphology. *Journal of Geophysical Research: Atmospheres*, 118(19), 10,980-10,993. Consulté sur <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50869> doi : 10.1002/jgrd.50869

Table des figures

1.1	Profils empiriques de température (a) et de norme du vent (b) à 38 degrés de latitude Nord.	3
1.2	Influence du vent sur la propagation des ondes infrasonores depuis le Mont Etna (14.99 ° E, 37.75 ° N) le 25/12/2018.(a) Profils empiriques de température (ligne bleue), de vent zonal (tirets noirs) et de vent méridien (pointillés noirs).(b) Empreinte au sol du champ d'ondes infrasonores prédit depuis le Mont Etna (pentagone vert). Chaque point est coloré en fonction de l'altitude maximale atteinte par les ondes qu'il représente.(d)-(g) Trajectoires des ondes infrasonores dans les plans verticaux correspondant aux sections A et B de la figure (b). (c)-(e)-(f)-(h) Profils de célérité effective du milieu respectivement dans la direction Ouest, Est, Sud et Nord, la célérité effective c_e étant l'addition de la vitesse du son et du vent dans la direction choisie. Les lignes grisées indiquent la valeur de c_e au sol. Lorsque la valeur de c_e devient supérieure à celle au sol, il y a un guide d'ondes.	4
1.3	Répartition des 51 stations infrasons du SSI certifiées par l'OTICE au 9 octobre 2019.	5
1.4	Exemple illustratif des observations faites à la station IS48 (Tunisie)(a) Carte de navigation autour de la station IS48 (rond magenta).(b) Image satellite de la tempête Miguel au large de la France le 6/06/2019. <i>Image de Meteosat/FVALK</i> (c) Phases détectées par IS48 du SSI entre le 03/06/2019 et le 22/07/2019 en fonction de leurs azimuts de provenance (back-azimut échelle de gauche). Chaque point est une phase détectée. La couleur indique la fréquence moyenne du signal de la phase détectée. La courbe noire indique l'amplitude du bruit ambiant incohérent pouvant occulter les détections (échelle de droite).	6
1.5	Représentation schématique du système d'observation global de l'atmosphère.	8
1.6	Exemple de distribution des observations directes et indirectes de l'atmosphère disponible au cours d'une période de 6 heures.	9
1.7	Cumuls mensuels des observations in-situ et satellitaires utilisées par Météo-France pour initialiser le modèle de PNT globale ARPEGE. <i>Tirée du rapport CNRM DirOP/COMPAS du 3/10/2019.</i>	10
1.8	Gammes d'altitudes sondées par les observations de température et de vent actuellement utilisées dans les modèles de PNT (en noir) et les nouvelles observations identifiées dans le cadre du projet européen d'infrastructure ARISE (en rouge). Les zones plus claires correspondent à des régions où l'observation est occasionnelle. <i>Adaptée du rapport du projet européen ARISE 2 (Grant 653980).</i>	12

2.1	Schéma illustratif de la circulation générale dans la troposphère.(a) Distribution des cellules (b) Coupe longitudinale. <i>Ces figures ont été extraites du site de la NOAA https://www.weather.gov/jetstream/ et complétées pour cette thèse.</i>	17
2.2	Coupes latitude-pression de la moyenne mensuelle de la température (K) et moyennées en longitude d'après la climatologie SPARC (Randel <i>et al.</i> , 2004) pour le mois de janvier (à gauche) et de juillet (à droite). L'intervalle des contours est de 10 K. <i>Cette figure est tirée de la thèse Saint-Martin (2010).</i>	18
2.3	Coupes latitude-pression de la moyenne mensuelle de la température (K) pour le mois de janvier et moyennées en longitude : à gauche celle d'après la climatologie SPARC (Randel <i>et al.</i> , 2004) ; à droite celle résultant de l'équilibre radiatif de la Terre d'après Holton (2004). L'intervalle des contours est de 10 K. <i>Cette figure est tirée de la thèse Saint-Martin (2010).</i>	19
2.4	Distribution moyenne du vent méridien au cours d'une année en trois points du globe.(a) 60 °N, 0 °W. (b) 0 °N, 0 °W. (c) 60 °S, 0 °W. Les données sont issues du modèle semi-empirique HWM (Drob <i>et al.</i> , 2008).	20
2.5	Distribution moyenne du vent zonal au cours d'une année en trois points du globe. Même points et données que sur la figure 2.4.	21
2.6	Distribution moyenne de la température au cours d'une année en trois points du globe. Les points sont les mêmes que ceux de la figure 2.4. Les données sont issues du modèle semi-empirique MSIS (Picone <i>et al.</i> , 2002).	21
2.7	Projection stéréographique polaire du champ de température (K) à 50 hPa (≈ 21 km) 15/02/2018 00 UTC. (a) Niveau 21(≈ 50 hPa) de l'analyse haute-résolution d'ARPEGE.(b) Niveau 48 (≈ 50 hPa) de l'analyse de IFS du CEPMMT.(c) Climatologie fournie par le modèle MSIS-90 à 21 km d'altitude (≈ 50 hPa) pour le 15 février.	25
2.8	Distribution des niveaux hybrides en pression au-dessus d'un même relief pour les 105 niveaux d'ARPEGE (bleu) et pour les 137 niveaux d'IFS (rouge).	28
2.9	Exemples de grilles horizontales dans une projection ortho centrée sur la France. (a) Une grille uniforme avec un point tout les 1.5 degrés (b) Une grille de Gauss linéaire réduite avec T_L479 . (c) Une grille basculée et étirée avec T_L798 et un coefficient d'étirement $C = 2.4$.(Pour la lisibilité de la figure, tous les points de grilles ne sont pas représentés).	30
2.10	Distribution de la température analysée par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018. . .	31
2.11	Distribution du vent zonal analysé par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018. . . .	31
2.12	Distribution du vent méridien analysé par le modèle IFS au cours de l'année 2018 en trois points du globe. Les points sont les mêmes que pour la figure 2.4. Les valeurs ne sont pas disponibles pour le 09/01/2018. . .	32
2.13	Représentation schématique des liens entre le cycle d'assimilation « déterministe » et celui « ensembliste »	36

3.1	Profils de c_{eff} (a) et spectre de l'opérateur $\hat{\mathcal{L}}$ pour une fréquence de 0.1 Hz (b) obtenu à partir des méthodes de discrétisation : par collocation spectrale (SC) et par différences finies (FD). Le nombre de points N_d de grille ou de collocation est donné en légende. $\mathcal{R}(x)$ et $\mathcal{I}(x)$ sont respectivement la partie réelle et la partie imaginaire de x . <i>Cette figure est adaptée du manuscrit de thèse de Goupy A.</i>	45
3.2	Rapport de célérité effective dans le plan de propagation. R est la distance à la source et le rond bleu symbolise la station réceptrice. <i>Conditions météorologiques déduites de l'analyse du CEPMMT du 6/11/2011 à minuit entre Rivesaltes et l'Observatoire de Haute Provence.</i>	48
3.3	Champs de pertes par transmission pour la fréquence 0.2 Hz (en haut) et la fréquence 0.8 Hz (en bas) dans le plan de propagation (colonne de gauche) et au sol (colonne de droite). La station réceptrice est représentée par un rond bleu et la source par le rond rouge.	49
4.1	Résultats d'analyse avec l'algorithme PMCC pour les signaux infrasons enregistrés à l'OHP le 06/11/2015 entre 1150 UTC et 1215 UTC.(haut) Vitesse de passage du signal cohérent dans l'espace temps-fréquence.(milieu) Back-azimut du signal cohérent arrivant dans l'espace temps-fréquence.(bas) Signal cohérent moyen.	52
4.2	Propagation d'infrasons à partir de Rivesaltes simulée par un modèle de tracé de rayons.(a) Profil de célérité effective dans la direction de l'OHP (axe horizontal inférieur) et de vent zonal (tiret bleu) et de vent méridien (pointillé bleu) (axe horizontal supérieur).(b) Trajectoires des rayons depuis Rivesaltes (pentagone vert) en direction de l'OHP (rond rouge). Les rayons sont colorés en fonction de leurs altitudes de réfraction z_{top} .(c) empreinte au sol donnée par les rayons.	54
4.3	Schéma illustratif de la relation entre l'incidence du rayon et la vitesse de passage.	55
4.4	Carte des pertes par transmission (dB) pour une fréquence de 0.8 Hz dans le cas d'une propagation d'ondes infrasonores entre Rivesaltes et l'OHP (rond rouge) pour le 06/11/2015 00 UTC. En gris, les trajectoires de la figure 4.2.	56
4.5	Résultats de la modélisation par FLOWS de la propagation depuis Rivesaltes vers l'OHP pour le 06/11/2015.(a) Signaux cohérents moyens enregistrés par l'OHP dans la gamme de fréquence 0.5-7 Hz.(b) Signaux simulés à l'OHP dans une gamme de fréquence entre 0-0.8 Hz pour une source de fréquence caractéristique 0.4 Hz.(c) Distributions temporelle et spatiale de l'enveloppe du signal en fonction de la distance en km par rapport à la source r et du temps retardé ($t-r/c_0$, avec $c_0 = 0,35 \text{ km.s}^{-1}$). Les points sont les point d'impacts identifiés par un code de tracé de rayons. La couleur des points est fonction de l'altitude de réfraction (bleu pour les phases troposphériques W , violet pour les phases stratosphériques S et gris pour les phases thermosphériques T). (d) Profil de célérité effective du son utilisé.	57
4.6	Résultats de la modélisation de la propagation depuis Rivesaltes vers l'OHP dans le cas d'un membre spécifique de l'ensemble issu de la perturbation de l'état atmosphérique par une structure de grande échelle.	58

4.7	Résultats de la modélisation de la propagation depuis Rivesaltes vers l'OHP dans le cas du membre perturbé de la figure 4.6 auquel l'on a ajouté une perturbation de petite échelle.	59
4.8	Profils de célérité effective pour plusieurs atmosphères idéalisées (c_2 : jet de 50 m.s ⁻¹ , c_1 : jet de 75 m.s ⁻¹ .)	60
4.9	Signaux normalisés obtenus avec les profils de célérité effective c_1 et c_2 à 285 km, présentés sur la figure 4.8. Les signaux sont normalisés par le maximum de leurs enveloppes respectives.	61
4.10	Filtrage visant à éliminer les Phases T appliqué sur le signal simulé pour différentes distances source-récepteur (R). Les signaux sont obtenus pour une même atmosphère idéalisée.	62
4.11	Amplitude maximale P_k des signaux enregistrés à 285 km de la source en fonction des paramètres de jet stratosphérique A_J et σ_J . Les iso-valeurs vont de 5.0×10^{-3} Pa à 2.0×10^{-2} Pa.	64
4.12	Temps de trajet Tt des signaux enregistrés à 285 km en fonction de A_J et σ_J . Les iso-valeurs vont de 900 s à 1000 s et +.	64
4.13	Extraction des quantités d'intérêt simulées. L'amplitude des signaux est en Pascal.	65
4.14	Durée Δ_{RSB} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 15 s à 30 s et +.	66
4.15	Durée Δ_G (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 30 s à 40 s et +.	66
4.16	Durée Δ_{ω_1} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 10 s à 20 s et +.	67
4.17	Durée Δ_{ω_2} (voir définition dans le texte) des signaux enregistrés à 285 km de la source en fonction de A_J et σ_J . Les iso-valeurs vont de 10 s à 25 s et +.	67
4.18	Schéma de la « recherche en ligne » avec conditions de Wolfe pour la fonction J . En violet les zones où α_k satisfait les conditions (4.11) et (4.12). En rouge les zones où α_k vérifie uniquement la condition de courbure (4.12). En bleu les zones où α_k vérifie uniquement la condition de réduction (4.11). En cyan le gradient au point $J(\mathbf{X}_k)$. Figure inspirée de celle proposée dans Nocedal et Wright (2006).	70
4.19	Évolution de J sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge. Les iso-valeurs vont de 0 à 100 et + avec le même code couleur que la figure 4.11.	70
4.20	Évolution de J en utilisant Δ_{ω_2} sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge. Les iso-valeurs vont de 0 à 100 et + avec le même code couleur que la figure 4.11.	71
4.21	Évolution de J et de son gradient sur le quadrillage \mathcal{Q} en (A_J, σ_J) pour une atmosphère cible \mathbf{X}^t en rouge.(a) Valeur de J sur \mathcal{Q} .(b) La couleur indique la valeur de J de manière identique à (a). Les barres correspondent aux gradients ∇J . Leur longueur est proportionnelle à la valeur du gradient.	73
4.22	Points donnés par la méthode de descente en gradient sans recherche en ligne pour une initialisation en $\mathbf{X} = (85, 2500)$	74
4.23	Points donnés par la méthode « de descente de gradient avec recherche en ligne » pour une initialisation en $\mathbf{X} = (85, 2500)$	74

4.24	Définition de la contrainte en T_t dans le cas de l'atmosphère idéalisée . . .	75
5.1	Rapport de célérité effective entre la surface et 50 km d'altitude obtenu le 18 mai 2016. Le vent est projeté sur les grands cercles passant par le Mont Etna (pentagone rouge). Les champs météorologiques proviennent de l'analyse du réseau de 00 UTC du CEPMMT. Les flèches représentent le vent à 50 km. Les triangles bleus désignent des stations infrasons du SSI et de l'OHP.	80
5.2	Exemple de propagation entre le Mont Etna et IS48.(a,b,c) profils de vitesse du son, de vent zonal et de vent méridien du 18/05/2016 fournis par le membre #19 analysé de 00 UTC de l'AEARP.(d) Distribution des points d'impact des rayons (points blancs) issus du Mont Etna (étoile rouge) en direction de IS48 (rond bleu). Le grand cercle reliant le Mont Etna à IS48 est représenté par la ligne bleue. (e) trajectoires des rayons (rouge) dans le plan de propagation. L'orographie réelle est en gris. La ligne en pointillés indique la hauteur de la couche homogène.	82
5.3	Amplitudes RMS des bulletins de l'antenne infrason placée sur le Mont Etna au cours de l'année 2016. Les bulletins ayant vu aucune ou une seule détection de 5 s sont en noir. Les bulletins avec une amplitude moyenne RMS comprise entre 0 et 1 Pa sont en jaune. Les bulletins avec une amplitude RMS supérieure à 1 sont en rouge. La largeur de chacune des barres correspond à une semaine.	82
5.4	(a) Satellite view of the domain of study. Mount Etna and IS48 infrasound station are represented by the red square and the cyan square, respectively. The magenta square represents the point of extraction of atmospheric profiles shown in (b), (c), and (d). The colored dots represent the simulated ground to ground infrasound propagation on May 18 2016 using member #19 from AEARP. The colorgrid displays the altitude of refraction z_{top} . The profiles in magenta are analysis from member #19 in terms of sound speed (b), zonal wind (c) and meridional wind (d). The profiles of the other members of the AEARP are displayed in light gray. The interval between vertical lines is 10 m.s^{-1} . The horizontal grid represents the 105 levels of member #19 altitudes.	90
5.5	Infrasound database. Selected events from May 15 2016 to May 26 2016 at Mount Etna (source) and IS48 station (detections). (Upper panel) The solid line depicts the effective sound speed ratio (first criterion, see Section 5.2.5) and the dotted line the value 1 against which the effective sound speed ratio is compared. (Lower panel) The yellow bars represent the hourly number of near field infrasound events with an averaged RMS amplitude over 1 Pa recorded by the Mount Etna infrasound array. The orange bars represent the fraction of the nearfield events obtained satisfying the second criterion. The blue bars show the hourly number of far field infrasound detections recorded by IS48 array.	91

5.6	Rays computed using atmospheric profiles from AEARP ensemble analysis on the 18th of May, 2016. Upper panels : member #19 adiabatic sound speed profile in black, effective sound speed in red and ray trajectories colored by turning height ; lower panel : member #7. adiabatic sound speed profile in black, effective sound speed in red and ray trajectories colored by altitude of refraction. The rays are launched for a fixed azimuth of 248.5 deg. The vertical dashed line represents the location of the infrasound station (right panels) or the effective sound speed at ground (left panels).	92
5.7	Standard deviation of daily analyses from AEARP in terms of sound speed and horizontal wind components at location 13.0 deg E and 37.1 deg N (see Figure 5.4 (a)) between 15 May 2016 and 27 May 2016 (00 UTC).	93
5.8	(a) Hourly average likelihood for each AEARP member (vertical axis) over the eruptive period of May 2016. The pixel width represents an hour and the color represents the likelihood value. (b) The green bars represent the hourly number of detections at IS48 used to compute the likelihood. Other bars give the average likelihood for member #24 as a function of time. (c) Hourly averaged likelihood (between 02 and 03 UTC) for each member on the 18th May, 2016.	94
5.9	Hourly averaged likelihood (02-03 UTC, 18 May 2016) for each AEARP member computed using the back azimuth (a), the trace velocity (b) and both observation components (c). The color scale is that of figure 5.8. The numbers shown on panel (b) are the number of eigenrays M available per member for the computation of the likelihood.	95
5.10	Hourly averaged back azimuths (a) and hourly averaged trace velocities (b). Black circles correspond to the observations at IS48. The dotted blue line gives the theoretical back azimuth between the Mount Etna and the infrasound station IS48. Red and blue circles are (mean) predictions obtained using the sea surface temperature 190 km away from the source (figure 5.4) or the surface temperature at IS48, respectively. Shaded regions indicate two standard deviations of the mean.	96
5.11	Updated hourly averaged likelihood for AEARP members using the surface temperature at IS48 station for computing the trace velocity. See legend of figure 5.8.	97
5.12	(a-d) Deviation to the theoretical back azimuth ϕ_0 for members #24, #19, #7, #1; AEARP analysis on 18 May 2016 (00 UTC). The black solid line is the mean azimuth deviation of impacting rays and the color of each point represents the altitude of refraction of the rays with a size proportional to v . The IS48 location is represented by the blue dashed line. (e-h) Hodographs with the same color scale for the altitude. The theoretical azimuth of IS48 from the view of Mount Etna is also represented by a thick gray line and isolines 15 m.s ⁻¹ and 35 m.s ⁻¹ are given by dashed circles.	98
5.13	Effective sound speed profiles for members #24, #19, #7 and #1 from the AEARP analysis of May 18 2016 (00 UTC) in the middle atmosphere. The effective sound speed profiles are extracted at location 13.0 deg E and 37.1 deg N. The mean AEARP profile is represented in dashed line and the gray region gives one standard deviation around the mean. The vertical dotted line is at 341.6 m.s ⁻¹	100

- 5.14 Probability for the likelihood to lie within a given interval $\alpha_k \leq \mathcal{L} < \alpha_{k+1}$ over the period 15-26 May 2016. Solid lines correspond to \mathcal{L}_{best} and dashed lines to \mathcal{L} , for detections in the ranges 00 - 03 UTC (blue) and 00 - 24 UTC (green) 101
- 5.15 Infrasound propagation for a set of atmospheric profiles from a synthetic ensemble of 50 members. (a) Profile of adiabatic sound speed (black), effective sound speed in the theoretical azimuth of members #0 (blue) and #6 (red). (b) Zonal wind profiles of members #0 (blue) and #6 (red) together with the ensemble spread (grey areas). (c) Meridional wind profiles of members #0 (blue) and #6 (red) together with ensemble spread (grey areas). (d) Deviation of azimuth of the rays along the geodesic. The solid line represents the mean deviation and the color scale represents the altitude of refraction. (e) Ray trajectories projected in the vertical plan along the great circle linking the station to the source. In both (d) and (e) panels, the blue dashed line represents the synthetic station location. The color scale shows the altitude of refraction. 104
- 5.16 Distribution of likelihood values for each member of the synthetic ensemble with different experimental specifications. (a) Values from a Bayesian inversion considering travel time, trace velocity and back azimuth measurements. The initial angular discretization is the most refine one (High density).(b) Values from a Bayesian inversion considering only back azimuth and trace velocity measurements. The initial angular discretization is the most refine one (High density).(c) Values from Bayesian inversion considering the three components of the observation vectors. The initial angular discretization is the coarsest one (Low density). Distributions in blue are obtained considering a 10 km radius around the station for sucessful ground to ground propagation, whereas for distributions in green this radius is reduced to 2 km. 106
- 5.17 Distribution des vraisemblances moyennes en tenant compte également des temps de trajet. Cette figure est à comparer avec la figure 5.11 de l'article. L'erreur d'observation considérée pour le temps de trajet est de 70 s (60 s pour le délai des bulletins source + 10 s pour la durée du signal à la station IS48) 108
- 5.18 Distribution des vraisemblances moyennes des membres sur la période éruptive du Mont Etna. Cette figure est à comparer avec la figure 5.11 de l'article. Les spécifications du milieux sont extraites au niveau de la source au lieu du point en mer. 108
- 5.19 Incidences et back-azimuts simulés en station le 19 mai. Croix bleues : résultats avec le code propagation *range-dependent*. Ronds rouges : résultats avec les profils extraits à 190 km de la source. Ronds verts : résultats avec les profils extraits au Mont Etna. 109
- 5.20 Propagation utilisant le membre #16 en milieu hétérogène.(a,b,c) Profils respectifs de vitesses du son, de vent zonal, de vent méridien extraits dans la région de propagation.(d) Déviation d'azimut le long du grand cercle. (principal) Trajectoires des infrasons représentées par les lignes pointillées. Leur couleur est en fonction de l'altitude maximale de réfraction. 110

5.21	Comparaison des ensembles de profils issus de l'ébauche et de l'analyse du 18 mai 00 UTC. En noir : données issues de l'ébauche de 00 UTC. En rouge : profils après assimilation de données. En bleu : incrément d'analyse sur la norme du vent. Les lignes sont les valeurs moyennes sur les ensembles et les surfaces couvrent deux écart-types autour de ces moyennes.	111
5.22	Distribution des vraisemblances moyennes des membres de l'ébauche de l'AEARP sur la période éruptive du Mont Etna. Cette figure est à comparer avec la figure 5.11 de l'article. Les ébauches sont actualisées à 00 UTC et à 12 UTC.	111
5.23	Distribution des vraisemblances moyennes des membres pour les détections enregistrées à l'OHP entre le 18 mai et le 22 mai.	112
A.1	Comparaison entre les altitudes des 90 niveaux à partir des fichiers du CEPMMT et celles obtenues à partir des fichiers disponibles opérationnellement au CEA . Données issues de l'analyse du réseau de 00 h du 10/08/2016. La comparaison est effectuée en un point de grille particulier (0 °N,0 °E).	128
A.2	Impact des différentes hypothèses dans la définition des profils de vent et de vitesse du son. Les profils sont extraits du membre 1 de l'AEARP du 28-12-2018 du réseau de 00 H à la verticale du Mont Etna. Extraction des profils de vent par les méthodes opérationnelles du CEA (courbe noire). Extraction des profils de vent avec la conversion altitude-pression (A.10)(courbe bleue). La vitesse du son est calculée pour un air sec (courbe rouge). La vitesse du son est calculée avec la température virtuelle (courbe verte). La vitesse du son est calculée pour un air humide (courbe bleue).	130
B.1	Schéma conceptuel des processus physiques qui doivent être introduits dans les modèles PNT et des interactions qui les lient. <i>Tirée de Coiffier (2000)</i>	134
B.2	Schéma conceptuel de la construction de la PEARP. <i>La figure est extraite de la description de la PEARP4 datant de Novembre 2017 et signée par C. Labadie, P. Cébron et L. Descamps</i>	136
B.3	Comparaison des profils de température et de vitesse du vent prévus en un point spécifique.(a-b-c) prévision d'ensemble effectuée à partir d'un état initial daté de 6 heures auparavant.(d-e-f) prévision d'ensemble effectuée à partir d'un état initial daté de 102 heures auparavant. Les profils de l'état de référence sont en noir. Les profils du médian de l'ensemble sont en bleu. Les aires bleues en transparence représentent l'écart entre le 5ème percentile et le 95ème percentile. La grille horizontale correspond à la distribution des niveaux. L'intervalle entre les lignes verticales représente 10 unités.	137
B.4	Comparatif de la température prévue par ARPEGE, PEARP et AEARP pour l'échéance du 15/02/2018 00 UTC à l'isobare 50 hPa. Les champs de température sont des moyennes d'ensemble (sauf pour l'analyse ARPEGE). L'échelle de couleur correspond à des températures exprimées en Kelvin.	138

Liste des tableaux

1.1	Quelques propriétés des principales sources d'infrasons. En orange : les sources impulsives. En bleu : les sources dites entretenues. <i>Ce tableau est adapté des travaux de McKisic (1997) et Campus et Christie (2010).</i>	2
1.2	Nomenclature des phases	3
2.1	Exemple de phénomènes de différents échelles présents dans l'atmosphère.	22
2.2	Base de données utilisée par HWM. En bleu celles modifiées par la version de 2014 (Drob <i>et al.</i> , 2015).Ce tableau provient de <i>Drob et al. (2008)</i> .	24
2.3	Modèles PNT globaux disponibles. Valeurs au 12 juillet 2019.	26
3.1	Comparaison entre les codes de propagation acoustique basés sur le tracé de rayons et ceux utilisant la méthode de décomposition modale.	50
4.1	Quelques quantités d'intérêt des détections de l'OHP associées à la source Rivesaltes de 1150 UTC. L'amplitude maximale est notée P_k . Les fréquences Freq., les back-azimuts ϕ et les vitesses de passage V_t des détections sont des valeurs moyennes sur la durée de la détection.	53
A.1	Description des champs selon les tables GRIB2 pour un fichier AEARP	120
A.2	Valeurs de A et B pour les grilles verticales de ARPEGE et IFS.	124
A.3	Valeur des constantes physiques utilisées dans la thèse	129
A.4	Exemple de correction induite par la prise en compte de l'humidité dans le calcul de la vitesse du son pour une température de 293 K, une humidité spécifique saturante et une pression de 1013 hPa. L'erreur est calculée par rapport à la correction en γ et Tv sans approximation (colonne 3).	129
B.1	Modifications apportées aux différentes paramétrisations de la PEARP vis-à-vis de celles utilisées pour la production déterministe du modèle ARPEGE, notée oper.	135

Glossaire

- ARISE** *Atmospheric dynamics Research InfraStructure in Europe* projet Européen financé par la Commission européenne. 11, 12, 147
- ARPEGE** Action de Recherche Petite Échelle Grandes Échelle. 10, 24, 25, 27, 28, 147, 148
- CEA** Commissariat à l'Énergie Atomique et aux énergies alternatives. 7, 11
- CEPMMT** Centre Européen pour les Prévisions Météorologiques à Moyen Termes. 24, 25, 30, 148
- CID** Centre International de Données de l'OTICE situé à Vienne, Autriche. 7
- CNRM** Centre National de Recherches Météorologiques. 10, 147
- GNSS-RO** *Global Navigation Satellite System - Radio Occultation*. 10
- HWM** *Horizontal Wind Model*. 22, 24, 25
- IFS** Integrated Forecasting System. 24, 25, 27, 28, 30, 31, 148
- IS48** 48^{ème} antenne acoustique du SSI située en Tunisie. 5–7, 147
- MF** Service météorologique national français Météo-France. 24
- MSIS** Modèle empirique de température et de composition de l'atmosphère. 25
- OMM** Organisation Météorologique Mondiale. 7
- OTICE** Organisation du Traité d'Interdiction Complète des Essais nucléaires. 5, 7
- PNT** Prévision Numérique du Temps. 8, 10–13, 24, 25, 29, 30, 147
- SPARC** *Stratospheric Processes And their Role in Climate*. 18, 19, 148
- SSI** Système de Surveillance International. 5–7, 14, 37, 147
- SSW** Réchauffement stratosphérique soudain (Sudden Stratospheric Warming).. 30
- TICE** Traité d'Interdiction Complète des Essais nucléaires. 5

Annexe C

Résumé étendu

Les ondes infrasonores, aussi appelées infrasons, sont des ondes acoustiques comprises entre 0 et 20 Hz émises par une large variété de sources. Leurs basses fréquences font que les gradients verticaux de vent et de température conduisent à la formation de guides d'ondes au sein desquels ces ondes peuvent se propager sur des milliers de kilomètres en étant très peu atténuées. Comme ce type de propagation est intrinsèquement lié à la circulation atmosphérique, ces ondes peuvent être utilisées comme traceurs. La circulation générale dans la stratosphère et la mésosphère, que l'on regroupe sous le terme générique de moyenne atmosphère, constitue pour les infrasons les guides d'ondes les plus efficaces. On se propose ici d'exploiter les signaux infrasons reçus au sol pour apporter de l'information nouvelle sur la circulation de la moyenne atmosphère. Elle pourra être utilisée afin d'améliorer l'initialisation des modèles de prévisions numériques du temps (PNT). Pour qu'une observation puisse être assimilée, il est nécessaire de définir un opérateur d'observation, capable de la simuler à partir des prévisions du modèle. L'objectif de cette thèse est d'examiner comment assimiler les infrasons en construisant un opérateur d'observation dédié.

Après avoir présenté l'état de connaissance sur la circulation atmosphérique et le principe de l'assimilation de données, nous examinons les données infrasons du système de surveillance international. Ce système est maintenu par l'Organisation du Traité d'Interdiction Complète des Essais nucléaires (OTICE). Un opérateur d'observation compatible avec les besoins de l'assimilation peut être construit à partir d'un modèle de propagation en tirant parti de méthodes de propagation rapide et des sources naturelles connues comme les volcans. Compte tenu des limitations associées à ces méthodes et ces sources, il est difficile de construire un opérateur d'observation permettant l'assimilation directe des données infrasons. Nous proposons alors de résoudre le problème inverse pour obtenir des profils verticaux de vent et de température cohérents avec les signaux infrasons mesurés.

Dans un premier temps nous avons mis en évidence la nécessité d'ajouter une perturbation aux profils atmosphériques pour restituer des signaux acoustiques réels par tracés de rayons ou décomposition modale. Ensuite, dans le cadre d'une atmosphère simplifiée à deux degrés de liberté et des signaux synthétiques, nous avons examiné la capacité des méthodes locales d'optimisation à estimer la perturbation atmosphérique minimisant l'écart entre les signaux modélisés et observés. Les résultats de cette étude montrent la difficulté d'obtenir une perturbation optimale sans information a priori sur celle-ci. Ce constat nous a amené à développer une approche bayésienne, en prenant pour a priori un ensemble d'analyses et de prévisions à courte échéance du modèle global de PNT de Météo-France. On utilise l'écart entre les résultats de simulations par un code de tracés de rayons et les détections infrasons pour évaluer la vraisemblance de chacun des membres de

l'ensemble. Cette méthode a été évaluée sur une base de détections de signaux infrasons issus des éruptions du Mont Etna en mai 2016. Nous montrons que cette méthode permet de sélectionner les membres de l'ensemble les plus vraisemblables, et ainsi fournir des profils atmosphériques, en accord avec les infrasons observés. Ils pourront ensuite être considérés comme des pseudo-observations dans un système d'assimilation de données.

Mots clés : propagation infrasonore, modélisation de l'atmosphère et assimilation de données.

Apport des infrasons pour l'assimilation de données dans un modèle global de prévision numérique du temps.

Abstract : Infrasons are acoustic waves between 0 and 20 Hz emitted by a wide variety of sources. Due to their low frequencies, vertical gradients of winds and temperature lead to the formation of waveguides in which they can propagate over thousands of kilometers with little attenuation. Since this type of propagation is intrinsically linked to the atmospheric circulation, these waves can be used as tracers. The general circulation in the stratosphere and the mesosphere, which forms the middle atmosphere, provides the most efficient acoustic waveguides. It is proposed here to use ground-based infrasound signals to provide new information on the circulation of the middle atmosphere. It can be used to improve the initialization of numerical weather prediction (NWP) models. For an observation to be assimilated, it is necessary to define an observation operator, able to simulate it from model predictions. The aim of this thesis is to examine how to assimilate infrasons by designing a dedicated observation operator.

After presenting the state of knowledge on atmospheric circulation and the principle of data assimilation, we examine the infrasound data of the international monitoring system. This system is maintained by the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). An observation operator compatible with the needs of assimilation can be designed from a propagation model by taking advantage of fast propagation methods and well-known natural sources such as volcanoes. Given the limitations associated with these methods and sources, it is difficult to design an observation operator allowing the direct assimilation of infrasound data. We propose then to solve the inverse problem by retrieving vertical profiles of winds and temperature consistent with the recorded infrasound signals.

Firstly, we have highlighted the need to superimpose a disturbance onto the atmospheric profiles in order to simulate real acoustic signals using ray tracing or modal decomposition methods. Then, within the frame of a simplified atmosphere with two degrees of freedom and simulated signals, we examined the ability of local optimization methods to estimate the atmospheric disturbance minimizing the discrepancies between modelled and observed signals. The results of this study show the difficulty in obtaining an optimal disturbance without prior information on it. This led us to develop a Bayesian approach, taking as prior information an ensemble of analyzes and short-term forecasts of the Météo-France global NWP model. The differences between the simulation results from a ray tracing code and the infrasound measurements are used to compute the likelihood of each member of the ensemble. This method was assessed using infrasound signals emitted by Mount Etna eruptions during May 2016. We show that this method allows to select the most likely members of the ensemble, and thus to provide atmospheric profiles, in agreement with observed infrasons. These profiles could then be considered as pseudo-observations in a data assimilation system.

Key words : infrasound propagation modelling, numerical weather prediction and data assimilation.

Mémoire de thèse de doctorat de l'Université de Toulouse III Paul Sabatier

Spécialité : Océan, Atmosphère et Climat.

Apport des infrasons pour l'assimilation de données dans un modèle global de prévision numérique du temps.

de P.J. Vanderbecken soutenue le 22 juin 2020 à Météo-France, Toulouse France.

Résumé : Cette thèse a pour but d'améliorer la prévision numérique du temps (PNT) dans la moyenne atmosphère (entre 17 km et 100 km) via l'apport d'informations déduites de la propagation d'ondes infrasonores. Ces ondes peuvent être vues comme des traceurs de l'état de la moyenne atmosphère. On se propose d'exploiter les bulletins de détections du réseau global d'antennes acoustiques de l'Organisation du Traité d'Interdiction Complète des Essais nucléaires pour fournir des informations sur la température et les vents. Pour résoudre ce problème inverse, des méthodes d'optimisation locale ont d'abord été testées. Puis, il a été développé une approche bayésienne pour sélectionner les états de l'atmosphère les plus vraisemblables prévus par un modèle de PNT. Cette approche a été appliquée sur une base de détections issues de l'activité volcanique du Mont Etna en mai 2016. La thèse montre qu'à partir des détections infrasonores, il est possible d'identifier, au sein d'un ensemble de prévisions, des états atmosphériques réalistes qui soient en accord avec les détections. Il peut alors être déduit des profils de températures et de vents dans la moyenne atmosphère à assimiler dans les modèles de PNT.

Mots clefs : propagation infrasonore, modélisation de l'atmosphère et assimilation de données.

Nota : thèse effectuée sous la supervision de J.-F. Mahfouf et de C. Millet au sein du CEA-DAM-DIF, F-91297 CEDEX Arpajon et du laboratoire CNRM-UMR 3589, F-31057 CEDEX Toulouse.



UNIVERSITÉ
TOULOUSE III
PAUL SABATIER



Université
de Toulouse

