



HAL
open science

Estimation d'incertitudes pour la propagation acoustique en milieu extérieur

Olivia Leroy

► **To cite this version:**

Olivia Leroy. Estimation d'incertitudes pour la propagation acoustique en milieu extérieur. Acoustique [physics.class-ph]. UNIVERSITE DU MAINE, 2010. Français. NNT : . tel-00906913

HAL Id: tel-00906913

<https://theses.hal.science/tel-00906913v1>

Submitted on 20 Nov 2013

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 DE DOCTORAT DE L'UNIVERSITÉ DU MAINE

spécialité ACOUSTIQUE

École doctorale Sciences pour l'Ingénieur, Géosciences, Architecture (ED 498)

présentée par

Olivia LEROY

pour obtenir le grade de
Docteur de l'Université du Maine

ESTIMATION D'INCERTITUDES POUR LA PROPAGATION ACOUSTIQUE EN MILIEU EXTÉRIEUR

soutenue le 7 décembre 2010

au Laboratoire Central des Ponts et Chaussées (Centre de Nantes)

devant le jury composé de :

Michel BÉRENGIER	Directeur de recherche HDR (LCPC)	Directeur de thèse
Laurent CARRARO	Professeur HDR (Telecom Saint-Etienne)	Rapporteur
Mathieu COUPLET	Ingénieur de recherche (EDF R&D)	Examineur
Gilles DAIGLE	Chercheur émérite (MG Acoustics)	Rapporteur
Benoit GAUVREAU	Chargé de recherche (LCPC)	Examineur
Fabrice JUNKER	Ingénieur de recherche (EDF R&D)	Examineur
Laurent SIMON	Professeur HDR (Université du Maine)	Examineur
Hans WACKERNAGEL	Maître de recherche HDR (ENSMP)	Président de jury

*A ma famille,
à mes amis,
à Franck*

REMERCIEMENTS

Ces quelques mots sont destinés à tous ceux qui ont contribué à ces travaux ainsi qu'à ceux qui ont partagé de près ou de loin cette expérience.

Je voudrais tout d'abord exprimer mes plus profonds remerciements à Benoit et Fabrice, pour leurs précieux conseils et leur présence au quotidien, ce qui m'a permis d'avancer et de persévérer dans ces travaux. A leurs côtés j'ai également appris les valeurs fondamentales du métier de chercheur et je leur en suis plus que reconnaissante.

Mes pensées se tournent vers Michel Bérengier qui malgré un emploi du temps de ministre était disponible pour me conseiller et me guider avec sagesse et justesse. Ce fut un réel plaisir d'évoluer à ces côtés pendant ces quelques années.

Je tiens également à remercier sincèrement Etienne de Rocquigny et Mathieu Couplet qui ont réussi à me faire appréhender et apprécier les statistiques avec des méthodes et des conseils toujours pertinents et simples. Merci à tous les deux pour votre implication dans mes travaux.

J'adresse de très sincères remerciements à Gilles Daigle et Laurent Carraro d'avoir accepté d'être les rapporteurs de mon travail et pour la grande attention portée à mon mémoire. Je remercie également Hans Wackernagel, Laurent Simon pour leur présence, leurs remarques constructives et encourageantes au cours de ma soutenance.

Ma reconnaissance s'adresse au groupe Acoustique du département AMA chez EDF où j'ai été accueillie à bras ouverts malgré de rares passages. La bonne humeur ambiante était d'ailleurs un bon leitmotiv pour venir m'égarer dans les contrées parisiennes.

Je me tourne maintenant vers les membres de l'équipe acoustique du LCPC pour l'ambiance chaleureuse qu'ils font régner au quotidien. Merci pour votre bonne humeur qui a rendu ces trois années un peu plus faciles. Une pensée également aux « copains de la navette », et notamment à Sandra avec qui les voyages furent toujours animés ! Un clin d'œil spécial à Dan, Guillaume, Gwenaël, Nicolas, Olivier avec qui j'ai passé d'excellents moments.

Une pensée à mes parents qui m'ont encouragé et soutenu tout au long de mes études, à ma sœur qui m'a toujours ouvert sa porte à toute heure du jour ou de la nuit. Enfin une pensée toute particulière à Franck qui a su me supporter, m'encourager et sans qui je n'en serais pas là aujourd'hui.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	vi
LISTE DES FIGURES	ix
LISTE DES TABLEAUX	xiii
NOTATIONS	xv
INTRODUCTION GÉNÉRALE	1
1 PROPAGATION ACOUSTIQUE EN MILIEU EXTÉRIEUR : INCERTITUDE ET VARIABILITÉ	5
1 PHYSIQUE DU SYSTÈME ÉTUDIÉ	7
1.1 Introduction	7
1.2 Caractérisation de la source sonore	7
1.3 Caractérisation du milieu de propagation	8
1.4 Bilan	29
2 MODÉLISATION DU SYSTÈME	30
2.1 Introduction	30
2.2 Incertitude de modèle	30
2.3 Modèles utilisés en propagation acoustique en milieu ouvert	31
2.4 Bilan	34
3 GRANDEURS D'INTÉRÊT	34
3.1 Introduction	34
3.2 Variabilité temporelle des niveaux sonores – De la turbulence au long terme . .	35
3.3 Variabilité spatiale des niveaux sonores - De l'échelle locale à l'échelle régionale	35
3.4 Incertitude sur les niveaux sonores	36
3.5 Bilan	37
4 CONCLUSION DU CHAPITRE	37
2 PROCESSUS DE QUANTIFICATION D'INCERTITUDES : MÉTHODOLOGIE	39
1 SPÉCIFICATION DU SYSTÈME	41
1.1 Introduction	41
1.2 Données mesurées	41
1.3 Choix du modèle	42
1.4 Bilan	48
2 DÉFINITION DES OBSERVABLES	48
2.1 Observables physiques	48
2.2 Observables définies par des outils statistiques	48
3 CALIBRATION DU SYSTÈME	49
3.1 Introduction	49
3.2 Approches statistiques	49
3.3 Définition du modèle statistique de Calibration sous Incertitude	51

3.4	Bilan	52
4	VALIDATION DU SYSTÈME	53
4.1	Identifiabilité des paramètres du modèle	53
4.2	Critères statistiques	53
4.3	Vérification des hypothèses initiales	54
4.4	Bilan	54
5	PRÉVISION	55
5.1	Variance d'estimation	55
5.2	Incertitude globale	56
6	CONCLUSION DU CHAPITRE	57
3	ESTIMATION D'INCERTITUDES ASSOCIÉES AUX CARACTÉRISTIQUES ACOUSTIQUES DE SOL	59
1	SPÉCIFICATION DU SYSTÈME	61
1.1	Données observées : mesures d'impédance de sol de la campagne expérimentale de Lannemezan 2005	61
1.2	Choix du modèle	62
2	DÉFINITION DES OBSERVABLES	64
2.1	Approche "physique"	64
2.2	Approche "statistique" basée sur la méthode d'Analyse en Composantes Principales fonctionnelles	68
2.3	Comparaison des méthodes de définition des observables	70
2.4	Bilan	76
3	CALIBRATION DES MODÈLES D'IMPÉDANCE DE SOL	77
3.1	Définition de la fonction coût	77
3.2	Calibration avec le modèle de Delany et Bazley	77
3.3	Calibration avec le modèle de Zwicker et Kosten	83
4	VALIDATION ET CHOIX DU MODÈLE STATISTIQUE OPTIMAL	88
4.1	Observables physiques	88
4.2	Observables définies par ACPf	89
4.3	Bilan	90
5	CONCLUSION DU CHAPITRE	90
4	ESTIMATION D'INCERTITUDE EN PRÉSENCE D'EFFETS MICROMÉTÉOROLOGIQUES	93
1	SPÉCIFICATION DU SYSTÈME	95
1.1	Données observées	95
1.2	Choix du modèle	99
1.3	Bilan	114
2	DÉFINITION DES OBSERVABLES	115
3	CALIBRATION DU SYSTÈME	116
3.1	Définition du modèle statistique de calibration	116
3.2	Estimation des paramètres incertains	116
4	VALIDATION DU SYSTÈME	118
5	PRÉVISION	122
5.1	Variance d'estimation des paramètres	122
5.2	Incertitude globale	123
5.3	Bilan	127
6	CONCLUSION DU CHAPITRE	127
5	DISCUSSION SUR LES INCERTITUDES EXPÉRIMENTALES ET LES MODÈLES DE PROPAGATION	129

1	INTRODUCTION	131
2	LIMITATIONS EXPÉRIMENTALES	131
2.1	Caractérisation de la source	131
2.2	Caractérisation du milieu de propagation	133
2.3	Validation de la base de données	135
2.4	Exemples de résultats	135
3	LIMITES DE MODÉLISATION	138
4	CONCLUSION ET PERSPECTIVES	140
	CONCLUSION ET PERSPECTIVES	143
	BIBLIOGRAPHIE	149
A	PRÉSENTATION DES BASES DE DONNÉES EXPÉRIMENTALES À DISPOSITION	157
1	STATION DE LONG TERME DE SAINT BERTHEVIN (53)	158
1.1	Cadre et objectifs	158
1.2	Dispositif expérimental	158
1.3	Exploitation	158
2	CAMPAGNE EXPÉRIMENTALE DE LANNEMEZAN 2005 (65)	159
2.1	Cadre et objectifs	159
2.2	Dispositif expérimental	159
2.3	Exploitation	172
3	CAMPAGNE MICROMÉTÉOROLOGIQUE DU LCPC NANTES 2007	175
3.1	Cadre et objectifs	175
3.2	Dispositif expérimental	175
3.3	Exploitation	175
	ANNEXES	157
B	NOTIONS DE GÉOSTATISTIQUES	177
1	VARIOGRAMME	177
2	KRIGEAGE	178
C	ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLES	179
D	IDENTIFIABILITÉ	181
1	GÉNÉRALITÉS	181
2	QUELQUES EXEMPLES	183
2.1	Cas linéaire	183
2.2	Prise en compte de biais	183
2.3	Prise en compte des variances de résidus	184
2.4	Prise en compte de deux paramètres physiques à estimer, de biais et de variances de résidus	185
2.5	Prise en compte de l'intercorrélacion entre les observables	186
E	PROPAGATION D'INCERTITUDES LIÉES AUX PARAMÈTRES MICROMÉTÉOROLOGIQUES	187

LISTE DES FIGURES

1.1	Parcours de rayons sonores issus d'une source ponctuelle au-dessus d'un sol présentant plusieurs types de comportement	11
1.2	Dispositif expérimental d'une mesure d'impédance de sol <i>in situ</i>	14
1.3	Incertitudes sur les niveaux d'atténuation sonores en dB(A) pour des sols très absorbants à réfléchissant en fonction de la distance à la source	15
1.4	Évolution schématique de la couche limite atmosphérique au cours d'une journée	16
1.5	Comparaison des profils de célérité de type logarithmique, linéaire et hybride	18
1.6	Profil de célérité de type hybride recalé sur des points de mesure	19
1.7	Effets d'un gradient vertical de température positif sur la propagation des rayons sonores en l'absence de vent	21
1.8	Effets d'un gradient vertical de température négatif sur la propagation des rayons sonores en l'absence de vent	21
1.9	Effets d'un gradient vertical de vent sur la propagation des rayons sonores en l'absence de gradient de température	22
1.10	Incertitudes sur les niveaux d'atténuation sonore à trois distances de propagation pour un sol herbeux	23
1.11	Incertitudes sur les niveaux d'atténuation sonore à trois distances de propagation pour un sol de type revêtement routier	24
1.12	Évolution du niveau d'atténuation sonore en fonction du gradient vertical de célérité du son	24
1.13	Spectre énergétique de la vitesse du vent	25
1.14	Effet de la turbulence atmosphérique en comparaison avec les mesures expérimentales issues de Wiener et Keast	27
1.15	Effet qualitatif d'une discontinuité d'impédance	28
1.16	Différentes sources d'incertitudes épistémiques liées à l'élaboration d'un modèle	31
1.17	Constitution du champ acoustique par la théorie des rayons	32
1.18	Variation des $L_{Aeq}(6h - 22h)$ au cours d'une année	36
1.19	Atténuation et dispersion des niveaux sonores en conditions favorables et en conditions défavorables	37
2.1	Représentation schématique du processus d'estimation d'incertitudes	40
2.2	Plan factoriel complet à trois paramètres et trois niveaux	45
2.3	Plan de Doehlert à trois facteurs	45
3.1	Dispositif expérimental	61
3.2	Évolution des valeurs de résistance spécifique au passage de l'air et de précipitations au cours de la campagne expérimentale de Lannemezan 2005	62
3.3	Évolution des valeurs d'épaisseur au cours de la campagne expérimentale de Lannemezan 2005	63
3.4	Groupes de spectres d'atténuation relative mesurés correspondant à la première et à la seconde période de la campagne expérimentale de Lannemezan 2005	63

3.5	Spectre calibré par calages manuels en aveugle (par 5 opérateurs) et histogrammes des valeurs de résistance spécifique au passage de l'air et d'épaisseur correspondant	65
3.6	Représentation des observables sur un spectre mesuré	66
3.7	Nuage de points représentant les observables définies par calage manuel des spectres .	67
3.8	Nuage de points représentant les observables 1 et 3 définies par calage manuel des spectres	67
3.9	Valeurs propres de l'ACPf	68
3.10	Deux premières fonctions propres définies par ACPf	69
3.11	Nuage de points représentant les observables définies par ACPf pour les deux groupes de données à l'étude	69
3.12	Nuage de corrélation des valeurs de résistance spécifique au passage de l'air estimées à partir des observables physiques et des coefficients d'ACPf pour le modèle de Delany et Bazley	71
3.13	Évolution des valeurs de résistance spécifique au passage de l'air estimées à partir des observables physiques et des coefficients d'ACPf pour le modèle de Delany et Bazley .	71
3.14	Nuage de corrélation des valeurs d'épaisseur estimées à partir des observables physiques et des coefficients d'ACPf pour le modèle de Delany et Bazley	72
3.15	Nuage de corrélation entre les résidus calculés après calibration spectre par spectre sur les observables physiques pour le modèle de Delany et Bazley	73
3.16	Nuage de corrélation entre les résidus calculés après calibration spectre par spectre sur les observables définies par ACPf pour le modèle de Delany et Bazley	74
3.17	Sensibilité du modèle de Zwicker et Kosten à la constante de structure	75
3.18	Évolution des valeurs de résistance spécifique au passage de l'air estimées en se basant sur les observables physiques et les coefficients d'ACPf pour le modèle de Zwicker et Kosten	75
3.19	Boîtes à moustaches des paramètres incertains du modèle de Delany et Bazley obtenus par le processus de bootstrap	79
3.20	Histogrammes et nuages des corrélation des résidus calculés sur l'ensemble des spectres	80
3.21	Histogrammes et nuages des corrélation des résidus calculés pour le groupe (1)	80
3.22	Histogrammes et nuages des corrélation des résidus calculés pour le groupe (2)	81
3.23	Boîtes à moustaches des paramètres incertains du modèle de Delany et Bazley pour les trois configurations étudiées obtenues par le processus de bootstrap	82
3.24	Boîtes à moustaches des paramètres incertains du modèle de Zwicker et Kosten pour les trois configurations étudiées (observables physiques)	84
3.25	Boîtes à moustaches des paramètres incertains du modèle de Zwicker et Kosten pour les trois configurations étudiées (coefficients de l'ACP) obtenues par la méthode de Bootstrap	86
4.1	Spectres d'atténuation expérimentaux mesurés à (50 m, 2 m) et (150 m, 2 m) de la source pour des conditions de propagation homogènes	96
4.2	Histogrammes du gradient vertical de célérité effective, de l'indice de turbulence normalisé et de la résistance spécifique au passage de l'air mesurés pour les spectres sélectionnés	97
4.3	Histogrammes de la température de l'air et de l'humidité relative mesurées pour les spectres sélectionnés	97
4.4	Spectres d'atténuation expérimentaux mesurés à 50 m de la source pour des conditions de propagation homogènes – mise en évidence de spectres pour des valeurs θ proches de la moyenne et pour des valeurs d_j distinctes	98
4.5	Spectres d'atténuation expérimentaux mesurés à 150 m de la source pour des conditions de propagation homogènes – mise en évidence de spectres pour des valeurs θ proches de la moyenne et pour des valeurs d_j distinctes	98
4.6	Nuages de corrélation entre les paramètres a_{log} et b_{lin} pour la campagne de Lannemezan 2005, les projets MEDD et Harmonoise	100

4.7	Profils verticaux de célérité du son effective calculés pour la campagne de Lannemezan 2005, les projets MEDD et Harmonoise	101
4.8	Spectres d'atténuation à (100 m, 2 m) modélisés par le code d'Équation Parabolique sans turbulence représentés par classes de propagation selon la valeur du paramètre de réfraction a_{log}	103
4.9	Spectres d'atténuation à (150 m, 2 m) en relatif à (10 m, 2 m) modélisés par le code d'Équation Parabolique sans turbulence représentés par classes de propagation selon la valeur du paramètre de réfraction a_{log}	104
4.10	Vues 2D du plan d'expériences numériques avec le paramètre de turbulence fixe	105
4.11	Plan factoriel complet pour une valeur de turbulence donnée	106
4.12	Surface de réponse à une distance à la source de 50 m (sans turbulence – $f = 500$ Hz)	107
4.13	Surface de réponse à une distance à la source de 50 m (turbulence intermédiaire – $f = 500$ Hz)	107
4.14	Surface de réponse à une distance à la source de 50 m (turbulence maximale – $f = 500$ Hz)	108
4.15	Surface de réponse à une distance à la source de 50 m (sans turbulence – $f = 2$ kHz)	108
4.16	Surface de réponse à une distance à la source de 50 m (turbulence intermédiaire – $f = 2$ kHz)	109
4.17	Surface de réponse à une distance à la source de 50 m (turbulence maximale – $f = 2$ kHz)	109
4.18	Surface de réponse à une distance à la source de 150 m (sans turbulence – $f = 500$ Hz)	110
4.19	Surface de réponse à une distance à la source de 150 m (turbulence intermédiaire – $f = 500$ Hz)	110
4.20	Surface de réponse à une distance à la source de 150 m (turbulence maximale – $f = 500$ Hz)	111
4.21	Surface de réponse à une distance à la source de 150 m (sans turbulence – $f = 2$ kHz)	111
4.22	Surface de réponse à une distance à la source de 150 m (turbulence intermédiaire – $f = 2$ kHz)	112
4.23	Surface de réponse à une distance à la source de 150 m (turbulence maximale – $f = 2$ kHz)	112
4.24	Comparaison des écarts-type estimés lors de la construction du métamodèle par krigage et des écarts-type expérimentaux en fonction de la fréquence	113
4.25	Valeurs propres de l'ACPf en appliquant la règle de Kaiser pour les spectres mesurés à (50 m, 2 m) et (150 m, 2 m) de la source	115
4.26	Six premières fonctions propres de l'ACPf pour les spectres mesurés à (150 m, 2 m) de la source	116
4.27	Spectres expérimentaux et spectres simulés par le métamodèle pour les paramètres moyens estimés θ pour des valeurs de d_j variables à 50 m de la source	117
4.28	Spectres expérimentaux et spectres simulés par le métamodèle pour les paramètres moyens estimés θ pour des valeurs de d_j variables à 150 m de la source	118
4.29	“p-p plots” des résidus relatifs à chaque observable calculés à une distance à la source de 50 m	120
4.30	“p-p plots” des résidus relatifs à chaque observable calculés à une distance à la source de 150 m	121
4.31	Spectres expérimentaux et spectres calculés par le métamodèle pour les paramètres estimés par bootstrap et pour des valeurs de d_j variables à 50 m de la source	124
4.33	Représentation des spectres expérimentaux, du spectre moyen modélisé et centré et des intervalles de confiance à 95 % relatifs à l'écart-type du modèle et à l'écart-type associé à l'incertitude globale à 50 m de la source. La différence entre les lignes pointillées de part et d'autre du spectre moyen correspond à l'écart-type résiduel.	126

4.34	Représentation des spectres expérimentaux, du spectre moyen modélisé et centré et des intervalles de confiance à 95 % relatifs à l'écart-type du modèle et à l'écart-type associé à l'incertitude globale à 150 m de la source. La différence entre les lignes pointillées de part et d'autre du spectre moyen correspond à l'écart-type résiduel.	126
5.1	Vue générale de la Station de Long Terme de Saint-Berthevin.	132
5.2	Représentation schématique de la source autoroutière de la SLT.	132
5.3	(a) Variogramme expérimental et modèle ajusté (courbe rouge) pour les mesures de résistance spécifique au passage de l'air R_s - (b) Carte de résistance spécifique au passage de l'air obtenue par krigeage - les points de mesure sont les points $I_j, j = 1, \dots, 12$ [22].	134
5.4	Résultats expérimentaux issus de la SLT (2002-2007) - niveaux sonores mesurés au mât A3 relativement au mât A1 à 2 et 5 m de hauteur (moyenne en trait plein et écarts-type en trait discontinu).	136
5.5	Niveaux sonores mesurés au mât A5 relativement au mât A1	136
5.6	Nuages de corrélation entre les paramètres de réfraction thermique et aérodynamique aux mâts météorologiques M1 et M4	137
5.7	Résultats expérimentaux issus de la SLT (2002-2007) - histogrammes des gradients verticaux de célérité aux mâts météorologiques M1 et M4.	138
5.8	Niveaux sonores mesurés au mât A5 relativement au mât A1	139
A.1	Échelles spatiales et temporelles explorées par les bases de données expérimentales et les modèles acoustiques à disposition	157
A.2	Présentation schématique de la Station de Long Terme sur le site de Saint-Berthevin .	159
A.3	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 125 Hz et 250 Hz . . .	160
A.4	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 500 Hz et 1 kHz . . .	161
A.5	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 2 kHz	162
A.6	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 125 Hz et 250 Hz . . .	163
A.7	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 500 Hz et 1 kHz . . .	164
A.8	Niveaux sonores mesurés au mât A3 relativement au mât A1 à 2 kHz	165
A.9	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 125 Hz et 250 Hz . . .	166
A.10	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 500 Hz et 1 kHz . . .	167
A.11	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 2 kHz	168
A.12	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 125 Hz et 250 Hz . . .	169
A.13	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 500 Hz et 1 kHz . . .	170
A.14	Niveaux sonores mesurés au mât A5 relativement au mât A1 à 2 kHz	171
A.15	Présentation schématique de la campagne expérimentale de Lannemezan 2005	172
A.16	Schéma du dispositif expérimental	175

LISTE DES TABLEAUX

1.1	Ordres de grandeur des valeurs de résistance spécifique au passage de l'air selon la nature du sol	12
1.2	Valeurs typiques du paramètre de rugosité z_0	17
2.1	Coordonnées du plan de Doehlert à trois facteurs	46
3.1	Moyenne des paramètres estimés spectre par spectre selon les deux méthodes de définition d'observables avec le modèle de Delany et Bazley	70
3.2	Moyenne des paramètres estimés spectre par spectre avec le modèle de Zwicker et Kosten et écarts-type associés	73
3.3	Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten	76
3.4	Paramètres estimés à partir des "observables physiques" avec le modèle de Delany et Bazley correspondant à la base de données complète et aux deux groupes de spectres	77
3.5	Paramètres estimés par la méthode de bootstrap à partir des observables physiques avec le modèle de Delany et Bazley et écarts-type associés	78
3.6	Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Delany et Bazley avec les observables physiques	79
3.7	Paramètres estimés à partir des coefficients d'ACPf avec le modèle de Delany et Bazley correspondant à la base de données complète et aux deux groupes de spectres	81
3.8	Paramètres estimés à partir des observables définies par les coefficients d'ACPf avec le modèle de Delany et Bazley et écarts-type d'estimation associés	82
3.9	Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Delany et Bazley avec les coefficients de l'ACPf	83
3.10	Paramètres estimés à partir des observables physiques avec le modèle de Zwicker et Kosten	84
3.11	Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Zwicker et Kosten avec les observables physiques	85
3.12	Paramètres estimés à partir des observables définies par ACPf avec le modèle de Zwicker et Kosten	86
3.13	Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Zwicker et Kosten avec les coefficients de l'ACPf	87
3.14	Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten basées sur les trois observables physiques	88
3.15	Critères statistiques calculés pour une valeur de fonction coût arbitraire constante pour un nombre d'observations $n = 479 \times 3$, pour un nombre de paramètres dépendant du système considéré	89
3.16	Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten basées sur les coefficients d'ACPf	89
4.1	Définition du plan d'expériences numériques avec le paramètre de turbulence fixe	104

4.2	Résultats statistiques de la validation croisée pour les métamodèles relatifs à chaque tiers d'octave	114
4.3	Paramètres physiques estimés à 50 m et 150 m de la source en considérant un groupe de spectres	116
4.4	Matrice de corrélation des résidus calculés à une distance à la source de 50 m	119
4.5	Matrice de corrélation des résidus calculés à une distance à la source de 150 m	119
4.6	Paramètres physiques et écarts-type d'estimation estimés par la méthode de bootstrap à 50 m et 150 m de la source en considérant un groupe de spectres représentatif de conditions de propagation homogènes	122
A.1	Résultats de calculs de variance normalisée entre les gradients de célérité effective pour toutes les directions de propagation et à toutes les distances à la source	173
A.2	Résultats de calculs du coefficient de corrélation entre les gradients de célérité effective pour toutes les directions de propagation et à toutes les distances à la source	174
A.3	Résultats vérifiant les conditions imposées sur le coefficient de corrélation et sur la variance normalisée entre les gradients de célérité effective pour toutes les directions de propagation et à toutes les distances à la source	174

NOTATIONS

GRANDEURS PHYSIQUES

r	distance entre la source et le récepteur en m
z	hauteur en m
f	fréquence de l'onde sonore en Hz
k	nombre d'onde acoustique en m^{-1}
T	température en K
$T_0 = 293,15$ K	température de référence (20 °C)
α_{atm}	coefficient d'absorption atmosphérique des ondes sonores en $dB.m^{-1}$
θ_v	température virtuelle en K
c	célérité du son en $m.s^{-1}$
c_{eff}	célérité du son en $m.s^{-1}$
$c_0 = 343$ $m.s^{-1}$	célérité du son de référence à $T = T_0$
Z	impédance du sol en $Pa.s.m^{-1}$
Z_0	impédance de l'air ($\rho_0 c_0$)
ρ_0	masse volumique de l'air
e	épaisseur en m
R_s	résistance spécifique au passage de l'air en $kN.s.m^{-4}$
Ω	porosité
C_s	constante de structure
z_0	longueur de rugosité de la couche limite de surface en m
$n = \frac{c_{eff}}{c_0}$	indice de réfraction acoustique du milieu
μ	partie fluctuante de l'indice de réfraction acoustique du milieu
$R = 286,69$ $J.kg^{-1}.K^{-1}$	constante de gaz parfaits
$\Gamma_d = 0,0098$ $K.m^{-1}$	gradient sec de température adiabatique
$\kappa = 0,4$	constante de von Kàrmàn
ϑ	direction de propagation en °
$\frac{\partial c}{\partial z}$	gradient vertical de célérité du son en s^{-1}
a_{T_p}	paramètre de réfraction thermique pondéré en s^{-1}
a_{v_p}	paramètre de réfraction aérodynamique projeté en s^{-1}
U^*	vitesse de frottement en $m.s^{-1}$
T^*	échelle de température en K
L_{MO}	longueur de Monin-Obukhov en m
ξ	paramètre adimensionnel de stabilité
P_t, P_d, P_r	champ transmis, champ direct, champ réfléchi
R_p	coefficient de réflexion en onde plane
Q	coefficient de réflexion en onde sphérique

GRANDEURS STATISTIQUES

ε_j	résidus
σ^2	variance des résidus
x	paramètres physiques incertains
$\theta = [x, \sigma^2]$	vecteur de paramètres incertains
$\sigma_{\hat{\theta}}^2$	variance d'estimation des paramètres
Γ	matrice de covariance des résidus
d_j	paramètres connus
Y_j	observables
$H(\mathbf{x}, d_j)$	modèle de prévision acoustique
$\mathcal{L} [(Y_j, d_j), \theta]$	fonction de log-vraisemblance
ρ	coefficient de corrélation
R^2	coefficient de détermination
AIC	critère d'information d'Akaike
BIC	critère d'information bayésien
λ	valeurs propres de l'ACP

INTRODUCTION GÉNÉRALE

« Si l'on s'en tient aux faits, les événements du monde peuvent être considérés tantôt comme déterminés, tantôt comme aléatoires : deux regards possibles sur un même monde. Un nouveau camouflet à notre image habituelle de la réalité. Pour apercevoir celle-ci avec une certaine précision, il nous faut, selon les cas et surtout selon les échelles, adopter des lunettes mentales différentes. »

Jean HAMBURGER (1999)

CONTEXTE

Le monde est régi par des phénomènes physiques et sociétaux complexes. Pour décrire ces phénomènes en évolution permanente, nous avons recours à des lois qui représentent l'état de notre connaissance du monde. Ces représentations approximatives de la réalité obligent à tenir compte de la notion d'incertitude. Dans la littérature, l'incertitude est généralement définie comme la combinaison d'un manque de connaissance et de la difficulté à prévoir l'évolution d'un système. Cette incertitude peut être liée à l'état d'un système et au caractère imprévisible de l'avenir comme l'a notamment défini Pierre Frédéric Gonod [61]. Dans un contexte sociétal par exemple, le mode de pensée ou les valeurs peuvent faire évoluer un système de façon inattendue. Il en est de même dans un contexte purement physique, où les vitesses de processus, une rupture du système peuvent modifier de façon irréversible leur évolution : les processus naturels peuvent présenter différents degrés de stabilité ainsi qu'un caractère fortement aléatoire voire chaotique. Cette variabilité engendre alors des phénomènes inattendus toujours plus difficiles à prévoir.

Si les notions d'*incertitude* et de *variabilité* sont souvent combinées, il est important de séparer ces deux notions pour parvenir à améliorer la connaissance des phénomènes. Le vocabulaire utilisé dans ce mémoire correspond à la terminologie largement utilisée en analyse des risques environnementaux [36, 51, 52]. Selon ces sources, la variabilité représente l'hétérogénéité inter-individuelle, ou dispersion vraie, d'une grandeur dans une population ou d'un paramètre du système étudié. Elle est inhérente au système observé et ne peut être réduite qu'en modifiant ce système, mais pas par des observations supplémentaires. L'incertitude est définie comme le manque de connaissance du modélisateur quant au système ou à un de ses paramètres. Elle n'est donc pas une propriété du système mais de l'état de connaissance du modélisateur. Elle peut être réduite par une acquisition de connaissances supplémentaires. Enfin, il est indispensable de mentionner l'expression « incertitude globale », qui englobe à la fois le concept de variabilité et le concept d'incertitude. D'autres termes peuvent également être utilisés pour désigner ce concept associant incertitude et variabilité, tel que l'« indéterminabilité » [125].

Le bruit et notamment celui inhérent aux infrastructures industrielles et routières est l'une des nuisances les moins bien perçues par le riverain : en 2010, deux-tiers des français se disent gênés par le bruit [117]. Une meilleure maîtrise des phénomènes sous-jacents et un contrôle des incertitudes associées sont donc requis. La problématique du bruit est notamment étudiée à travers le domaine physique de propagation acoustique en milieu extérieur. Elle implique en effet des phénomènes complexes tels que les effets liés au sol et les effets liés aux conditions micrométéorologiques qui contribuent au caractère fluctuant des niveaux sonores ainsi qu'à leur variabilité difficilement prévisible. Ces phénomènes étant complexes, leur interprétation et leur représentation physique est soumise à un certain manque de connaissance. Une approche classique pour décrire le bruit est basée sur la notion d'indicateurs, approche déjà clairement décrite à la fin des années soixante-dix dans un article issu d'une revue générale des routes et des aéroports [14] :

« Le bruit est un phénomène essentiellement fluctuant. Pour le caractériser, il est nécessaire d'utiliser des indicateurs constitués de moyennes sur des échantillons agrégés de niveaux sonores, ces échantillons étant situés dans une période de temps donnée. Ce type de représentation amène forcément des incertitudes sur les résultats. »

Un certain nombre de problématiques y sont abordées telles que la définition des indicateurs, leur représentativité, ainsi que la variabilité physique et les incertitudes liées aux niveaux sonores. Une démarche permettant à la fois d'utiliser les connaissances physiques actuelles en propagation et de déterminer les incertitudes associées permettrait donc une meilleure évaluation des situations d'exposition au bruit.

DÉMARCHE

Afin d'améliorer la qualité des prévisions de certains phénomènes environnementaux tels que les phénomènes météorologiques, océanographiques ou hydrogéologiques, des méthodes statistiques peuvent être mises en œuvre. Un certain nombre de méthodes statistiques telles que l'assimilation de données [47], la prévision d'ensemble [120], les méthodes géostatistiques [33] sont déjà communément utilisées dans divers domaines de la physique et notamment dans les problématiques impliquant des prises de décision ou des analyses de risques. Certaines ont déjà été appliquées pour répondre à notre problématique liée aux incertitudes et à la variabilité en propagation acoustique en milieu extérieur [18, 108], mais seuls des résultats qualitatifs avaient été obtenus.

L'objectif principal de ce travail de thèse consiste à définir une méthodologie statistique permettant la quantification de l'incertitude et de la variabilité liées aux phénomènes de propagation acoustique en milieu extérieur (complexe ou non), en intégrant à la fois des données expérimentales et des modèles de prévision acoustique. Les objectifs sous-jacents sont de maîtriser au mieux cette incertitude globale, d'associer à une prévision un niveau d'incertitude puis d'améliorer le contenu prédictif des modèles de prévision. Ce type de résultat permet notamment d'alimenter les travaux normatifs et d'apporter des éléments pour qu'un acousticien puisse optimiser la mise en place et le déroulement d'une étude d'impact.

La méthodologie statistique que nous proposons est basée sur la comparaison de données expérimentales et de données modélisées. Cependant cette simple démarche n'est possible que si l'ensemble des paramètres d'entrée du modèle de prévision est bien connu, ce qui est un cas idéalisé. Il est donc nécessaire dans la plupart des cas d'étude de procéder à la *calibration* des paramètres d'entrée. Cette phase de calibration permet de minimiser de façon considérable les écarts observés entre données modélisées et données expérimentales mais une part non-négligeable d'incertitude (globale) demeure malgré tout. La méthode de calibration des paramètres d'entrée considérés comme incertains permet

une estimation de la valeur de ces paramètres ainsi qu'une quantification de l'incertitude sur ces paramètres. A la calibration vient s'ajouter une phase de prévision permettant la quantification de la dispersion associée aux grandeurs d'intérêt.

ORGANISATION DU MÉMOIRE

Le chapitre 1 présente les phénomènes physiques liés à la propagation acoustique en milieu extérieur. Ce chapitre est organisé selon la vision du statisticien, en considérant les paramètres d'entrée, le modèle et les grandeurs d'intérêt. Des éléments mettant en évidence l'incertitude et la variabilité associées aux phénomènes ponctuent ce chapitre et visent à sensibiliser le lecteur à la problématique.

Le chapitre 2 détaille la méthodologie statistique proposée pour répondre à l'objectif premier de ce travail : l'estimation de l'incertitude et de la variabilité d'un système complexe. Le processus d'estimation d'incertitude décrit dans ce chapitre est basé sur un processus de calibration « évolué » et s'est construit tout au long des travaux de thèse. Chaque élément de la méthodologie a été ajouté au fil des applications, du fait des limites de la méthode initiale *i.e.* la calibration « simple ». La méthodologie statistique proposée s'organise selon cinq phases : la spécification du système, la définition des observables, la calibration, la validation et la prévision.

Cette méthodologie est ainsi utilisée pour étudier l'incertitude et la variabilité associées aux seuls paramètres acoustiques de sol (chapitre 3), en considérant de courtes distances de propagation (de l'ordre de quelques mètres). Cette application permet de tester la robustesse du processus de quantification d'incertitudes et d'accéder aux incertitudes liées aux caractéristiques acoustiques de sol. Cette application a également permis d'appréhender la robustesse de différents modèles acoustiques de sol.

Le chapitre 4 présente un cas d'application de la méthodologie d'estimation d'incertitude en considérant des distances de propagation de l'ordre de 150 mètres. Ces plus grandes distances de propagation nécessitent la prise en compte des effets micrométéorologiques en plus des effets de sol. Pour des raisons de ressources calculatoires, cette étude plus complexe a notamment nécessité le développement d'un modèle mathématique simplifié, appelé *métamodèle*, comme modèle de prévision de substitution.

Le chapitre 5 a pour objectif de mettre en évidence les limitations actuelles de l'application de telles méthodes statistiques au domaine de la propagation acoustique en milieu extérieur, aussi bien dans le domaine expérimental que dans le domaine de la modélisation.

1

PROPAGATION ACOUSTIQUE EN MILIEU EXTÉRIEUR : INCERTITUDE ET VARIABILITÉ

PLAN DU CHAPITRE

1	PHYSIQUE DU SYSTÈME ÉTUDIÉ	7
1.1	Introduction	7
1.2	Caractérisation de la source sonore	7
1.3	Caractérisation du milieu de propagation	8
1.4	Bilan	29
2	MODÉLISATION DU SYSTÈME	30
2.1	Introduction	30
2.2	Incertainité de modèle	30
2.3	Modèles utilisés en propagation acoustique en milieu ouvert	31
2.4	Bilan	34
3	GRANDEURS D'INTÉRÊT	34
3.1	Introduction	34
3.2	Variabilité temporelle des niveaux sonores – De la turbulence au long terme . . .	35
3.3	Variabilité spatiale des niveaux sonores - De l'échelle locale à l'échelle régionale	35
3.4	Incertainité sur les niveaux sonores	36
3.5	Bilan	37
4	CONCLUSION DU CHAPITRE	37

L'organisation de ce chapitre va refléter l'approche du statisticien lors de l'étude d'un système physique avec une partie relative à la description physique du problème permettant d'expliquer les paramètres d'entrée. La seconde partie concerne la modélisation du système, en mettant en avant les modèles de propagation acoustique que nous allons utiliser au cours de notre étude. Finalement, une partie est consacrée aux grandeurs d'intérêt à savoir les niveaux sonores. On s'attachera à chaque étape à rendre compte de la part d'incertainité et de la part de variabilité physique associées aux variables ou aux phénomènes considérés pour éclairer le lecteur sur les contributions en termes d'incertainité de chaque élément du système étudié.

1 PHYSIQUE DU SYSTÈME ÉTUDIÉ

1.1 INTRODUCTION

Dans une démarche statistique, on s'intéresse d'abord aux paramètres disponibles permettant de caractériser le processus étudié. Ces paramètres peuvent être mesurés, estimés à partir d'autres paramètres eux mêmes mesurés ou estimés, imposés de manière arbitraire à partir d'hypothèses ou issus de la littérature. On considère certains paramètres comme fixés et connus et d'autres paramètres comme incertains du fait de leur variabilité et de leur incertitude. Dans le cadre de la propagation acoustique en milieu extérieur, ces paramètres d'entrée permettent de caractériser la source sonore ainsi que le milieu de propagation.

1.2 CARACTÉRISATION DE LA SOURCE SONORE

La démarche proposée par la suite ne dépend pas du type de source de bruit. Il est cependant intéressant de garder à l'esprit que l'incertitude peut varier selon que la source soit liée au trafic routier, ferroviaire, aérien ou à une activité industrielle. Ces sources diffèrent entre elles par leur spectre, par leur géométrie, par leur localisation ou évolution spatio-temporelle. Il est donc essentiel de bien connaître les caractéristiques de la source sonore étudiée. A titre d'illustration, seules les problématiques routières et industrielles seront abordées dans ce qui suit.

1.2.1 Sources de bruit routier

Le niveau de bruit émis par chaque véhicule dépend de son fonctionnement et de son environnement. L'émission sonore globale d'un véhicule est donc sensible à un ensemble de paramètres qui représente la liste des variables à estimer pour déterminer l'émission de bruit [30, 80, 112].

1.2.1.1 Cas d'un véhicule isolé

Les paramètres caractérisant l'émission acoustique d'un véhicule individuel peuvent être regroupés en trois grandes familles :

- *les paramètres cinématiques* : il s'agit classiquement de la vitesse et de l'accélération. La vitesse est un paramètre primordial pour déterminer le bruit émis par un véhicule. Elle peut être utilisée comme unique variable descriptive lorsque le bruit de contact pneumatique-chaussée devient prédominant (pour des vitesses supérieures à 50 km/h). Cette source peut être modélisée par une source ponctuelle mobile avec une hauteur de source équivalente de quelques centimètres seulement ;
- *les paramètres mécaniques* qui décrivent le fonctionnement du moteur ;
- *les paramètres du véhicule et de l'infrastructure* donnant une indication sur l'environnement dans lequel évolue le véhicule (type de revêtement, de véhicule, de pneumatique, etc.). Dans les modèles de sources, les paramètres de l'infrastructure sont des paramètres descriptifs qui interviennent comme des correctifs par rapport à un environnement standard. Leur influence sur le bruit émis est constante quelque soit l'état cinématique et mécanique des véhicules.

1.2.1.2 Cas d'un flux de trafic

Dans le cas d'une circulation importante, il est impossible de connaître le comportement de chaque véhicule individuellement. Le comportement du flux peut être approché grâce à des variables plus globales telles que le débit (nombre de véhicules qui passent en un point par unité de temps) ou la densité (nombre de véhicules par unité de longueur à un instant donné). La vitesse, l'accélération et le rapport de boîte de vitesse peuvent être définis constant pour des portions de voies homogènes et appliqués pour chaque type de véhicule de cette section. Il est bien entendu nécessaire de prendre en compte la silhouette des véhicules : véhicule léger, poids lourd, deux roues, etc. Considérer la vitesse, l'accélération et le rapport de boîte de vitesse constants n'est justifié que si le comportement global

du trafic est stationnaire et si la dispersion des comportements individuels des véhicules n'est pas trop importante.

Si chaque véhicule peut être considéré comme une source ponctuelle mobile, à proximité de la route seul le véhicule le plus proche influe sur le niveau sonore. A une distance plus importante (au minimum à une demi-distance séparant les sources ponctuelles), les sources dites secondaires ne sont plus négligeables. Le bruit est alors constitué de la somme des contributions des sources individuelles en mouvement, ce qui peut être assimilé à une source linéique. La caractérisation expérimentale d'une telle source linéique peut être complexe du fait des contributions différentes selon la ligne de vue considérée.

Certains paramètres tels que la vitesse, le pourcentage de véhicules légers et de poids lourds, le revêtement (état de vieillissement de la chaussée) n'étant pas encore bien maîtrisés, des incertitudes peuvent être engendrées sur la caractérisation de la source. Par exemple, dans le cas d'un même revêtement routier, au passage d'un véhicule à une vitesse donnée (90 km/h), le niveau sonore peut varier de 2 dB.

1.2.2 Sources de bruit industriel

Les sources de bruit industriel sont extrêmement variées et peuvent être réparties selon les caractéristiques principales suivantes :

- leurs *tailles* : sources de petites dimensions (climatiseurs) et de grandes dimensions (salle des machines de centrale électrique) ;
- leurs *spectres d'émission* : puissance acoustique, directivité, aspects large bande (bruit d'écoulement) et aspects basses fréquences qui nécessitent notamment la connaissance de la répartition énergétique spectrale ;
- leur *caractère tonal* qui est défini par l'émergence d'un son pur dans un spectre de bruit large bande (transformateur) ;
- leur *caractère impulsif* qui permet également de caractériser la source en fonction du niveau des impulsions et de l'importance des impulsions par rapport au bruit de fond ;
- la *durée des cycles de fonctionnement* de certaines installations.

Une combinaison ou une interaction de ces différents paramètres complexifie la caractérisation d'une source de bruit industriel mais doit permettre de la décrire de manière optimale. Ces caractéristiques ne pouvant pas toutes être maîtrisées de façon optimale, elles constituent alors des causes d'incertitudes non négligeables. Selon les méthodes employées, l'estimation de la puissance acoustique varie de 1,5 à 4 dB(A)¹ selon les références normatives [5].

1.3 CARACTÉRISATION DU MILIEU DE PROPAGATION

Les sources sont situées dans un environnement complexe qui influe de façon conséquente sur la propagation acoustique et sur les niveaux sonores au(x) récepteur(s). Dans un cas idéalisé, le milieu de propagation peut être non borné (champ libre) et considéré comme homogène où seuls certains phénomènes dominent : la divergence géométrique ainsi que l'absorption atmosphérique. Dans le cas où le milieu est semi-ouvert (présence d'un sol), le champ acoustique dépend des propriétés des frontières du milieu et notamment des caractéristiques du sol. Le milieu de propagation étant en réalité hétérogène, à ces phénomènes viennent s'ajouter les effets liés au vent et à la température.

1. Le niveau pondéré A corrige un spectre en prenant en compte de façon approximative la sensibilité de l'oreille humaine (moins sensible aux basses fréquences qu'aux hautes fréquences).

1.3.1 Divergence géométrique

La divergence géométrique est le phénomène relatif à l'expansion du front d'onde dans le milieu en espace libre sans aucune frontière et en milieu non dissipatif. La progression de l'onde se réalise à énergie constante et demeure conservative à la surface du front d'onde. Le front d'onde s'étend lorsque l'on s'éloigne de la source tandis que la densité surfacique d'énergie va diminuer de manière inverse. Si r est la distance entre une source ponctuelle et le récepteur, l'amplitude des oscillations décroît alors proportionnellement à $1/r$. L'atténuation de cette onde sphérique est de 6 dB par doublement de distance. Une source linéique, dont les fronts d'onde sont de forme cylindrique et proportionnelle à r , voit son énergie décroître moins vite et l'atténuation n'est que de 3 dB par doublement de distance ($1/\sqrt{r}$). En s'éloignant d'une source (« champ lointain » *i.e.* r très supérieur à la longueur d'onde λ), le front d'onde va progressivement être assimilable à un plan.

1.3.2 Absorption atmosphérique

C'est en 1931 que Knudsen [76] a émis l'hypothèse selon laquelle les propriétés du milieu de propagation et notamment l'humidité avaient un effet d'absorption au cours de la propagation du son. De nombreux phénomènes thermodynamiques et échanges énergétiques ont lieu dans l'atmosphère du fait de la composition et des propriétés du milieu. Les différents phénomènes influents ont donc été mis en évidence ; la dissipation acoustique ou absorption dite classique est due à trois processus principaux que l'on considère des fluides simples (gaz par exemple) ou des fluides complexes (fluides à propriétés hétérogènes et polyphasiques) :

- la *viscosité* : elle est induite par le transport moléculaire de quantité de mouvement. Les forces (cisaillement et volume) entre les différentes couches à travers lesquelles se propage l'onde sont responsables de l'amortissement des ondes acoustiques ;
- la *conduction thermique* : l'onde subit un processus de diffusion du fait d'un transfert de chaleur entre les zones de surpression et de dépression ; zones créées au passage de l'onde acoustique dans le milieu de propagation par des processus thermodynamiques (loi de Lechâtelier) ;
- la *relaxation moléculaire* : du fait de sollicitation extérieure, un retard à l'établissement d'un équilibre peut survenir et ainsi induire un amortissement des ondes. Ce retard dépend directement de la concentration en molécules d'eau et donc de l'humidité relative. La compression au passage de l'onde acoustique induit une accélération du mouvement des particules (des processus de collisions transforment l'énergie de translation en énergie de vibration et de rotation). Lorsque la fréquence de l'onde correspond à la fréquence de relaxation du gaz, l'augmentation de l'énergie de vibration et l'apparition de la phase de détente induisent alors une perte d'énergie en pression.

Il a été démontré que les phénomènes de viscosité et la conduction thermique étaient négligeables si l'on considère la propagation sonore en milieu ouvert pour des gaz polyatomiques tels que l'air [28]. Seule la relaxation moléculaire est à prendre en considération dans le phénomène d'absorption atmosphérique. Un certain nombre d'auteurs ont tenté de modéliser le phénomène par des modèles semi-empiriques et des modèles théoriques [114]. L'absorption atmosphérique dépend essentiellement de l'humidité relative du milieu, de la température ainsi que de la fréquence considérée. Evans et al. [46] ont fourni des valeurs de coefficients d'absorption en fonction de ces différents paramètres. Des expressions analytiques permettant de calculer les valeurs du coefficient d'absorption noté α_{atm} (exprimé en dB.m^{-1}) pour chaque fréquence ont été établies [16]. L'absorption atmosphérique est définie par la relation suivante :

$$\alpha_{atm} = \frac{f}{p_s p_{s0}} \left[1,84 \cdot 10^{-11} \left(\frac{T}{T_0} \right)^{\frac{1}{2}} + \left(\frac{T}{T_0} \right)^{-\frac{5}{2}} \left[\frac{0,01278 \exp \frac{-2239,1}{T}}{f_{r,O} + \frac{f^2}{f_{r,O}}} + \frac{0,1068 \exp \frac{-3352}{T}}{f_{r,N} + \frac{f^2}{f_{r,N}}} \right] \right], \quad (1.1)$$

où f est la fréquence (Hz), p_s la pression atmosphérique (Pa), p_{s0} la pression atmosphérique de référence (1 atm = 101325 Pa), T la température atmosphérique (K), T_0 la température atmosphérique de référence (293,15 K), $f_{r,O}$ la fréquence de relaxation de la molécule d'oxygène (Hz), $f_{r,N}$ la fréquence de relaxation de la molécule d'azote (Hz). Il est donc nécessaire d'intégrer ce processus dans l'étude de la propagation acoustique à grande distance ; les documents normatifs [4] reprennent d'ailleurs ces propriétés.

L'absorption atmosphérique présente une certaine variabilité temporelle aux échelles journalière et saisonnière [77]. L'humidité relative atteint, par exemple, son maximum au lever du soleil et son minimum en milieu de journée lorsque les températures sont plus élevées. Les variations journalières sont d'ailleurs plus importantes au cours de la période estivale. L'utilisation de données globales moyennées est donc à éviter ; il est préférable d'utiliser des mesures locales suffisamment échantillonnées dans le domaine temporel.

1.3.3 Caractérisation acoustique du sol

1.3.3.1 Onde réfléchi et onde de sol

La source et le récepteur étant généralement près du sol, il est indispensable de prendre en compte ses effets. Selon la nature de la surface du sol, celle-ci peut plus ou moins absorber l'énergie de l'onde sonore. L'onde réfléchi sur le sol se combine avec l'onde directe et donne naissance à des interférences. Ce champ d'interférences dépend de la géométrie du problème et des propriétés acoustiques du sol qui déterminent l'amplitude et la phase de l'onde réfléchi [45].

La problématique de la propagation d'un point à un autre au-dessus d'une frontière a été étudiée par Sommerfeld dans le cadre de la propagation des ondes électromagnétiques [111]. La théorie de la propagation des ondes acoustiques au-dessus d'un milieu poreux a par la suite été établie sur cette base. Si l'on considère la propagation acoustique d'une onde issue d'une source ponctuelle à grande distance à proximité d'une interface plane entre deux milieux de propagation (une atmosphère homogène et le sol) dans le cadre d'une approximation en champ lointain (front d'onde sphérique assimilé à un front d'onde plan), le champ de pression acoustique est caractérisé par la formule de Weyl-Van der Pol [96] :

$$p_t = p_d + R_p p_r \text{ avec } R_p = \frac{\sin \phi - \frac{Z}{Z_0}}{\sin \phi + \frac{Z}{Z_0}}, \quad (1.2)$$

où p_t , p_d et p_r correspondent respectivement aux champs transmis, direct et réfléchi, le coefficient R_p correspond au coefficient de réflexion en onde plane dans le cas d'une réflexion spéculaire, Z_0 et Z correspondent respectivement à l'impédance de l'air et de sol, ϕ l'angle entre la surface et le rayon incident. Dans le cas des ondes sphériques, le coefficient de réflexion (en onde sphérique) est défini par la relation $Q = R_p + F(w)(1 - R_p)$. La contribution du second terme de Q sur le champ total agit comme une correction pour le passage d'une onde plane à une onde sphérique [13, 104]. Cette contribution est appelée *onde de sol* par analogie avec le terme utilisé en réception radio. La fonction $F(w)$ parfois appelée « facteur limite de perte » est une fonction complexe qui décrit l'interaction d'un front d'onde courbé avec le sol. Cette fonction est définie par :

$$F(w) = 1 + i\sqrt{\pi} \exp(-w^2) \operatorname{erfc}(-jw), \quad (1.3)$$

où erfc est la fonction d'erreur complémentaire [32] et w la distance numérique définie par l'équation :

$$w = \frac{2jk_0 r_r}{(1 - R_p)^2 \cos^2 \phi} \left(\frac{Z_0}{Z} \right)^2 \left(1 - \frac{k_0^2}{k^2} \cos^2 \phi \right). \quad (1.4)$$

Z_0 et Z correspondent respectivement à l'impédance de l'air et de sol, k_0 et k leur nombre d'onde associée et ϕ l'angle entre la surface et le rayon incident.

1.3.3.2 Effet d'épaisseur

Lorsqu'on suppose que le sol ne contient qu'une seule couche de matériau homogène semi-infini et que la réaction de la surface est indépendante de l'angle d'incidence de l'onde sonore, on dit qu'on se place dans le cadre d'une *réaction localisée*. C'est le cas par exemple d'un sol réfléchissant de type revêtement routier, ou d'un sol naturel (en première approche). L'onde transmise dans le sol se propage le long de la normale à la surface et les angles de l'onde incidente et de l'onde réfléchi par rapport au sol sont égaux (figure 1.1).

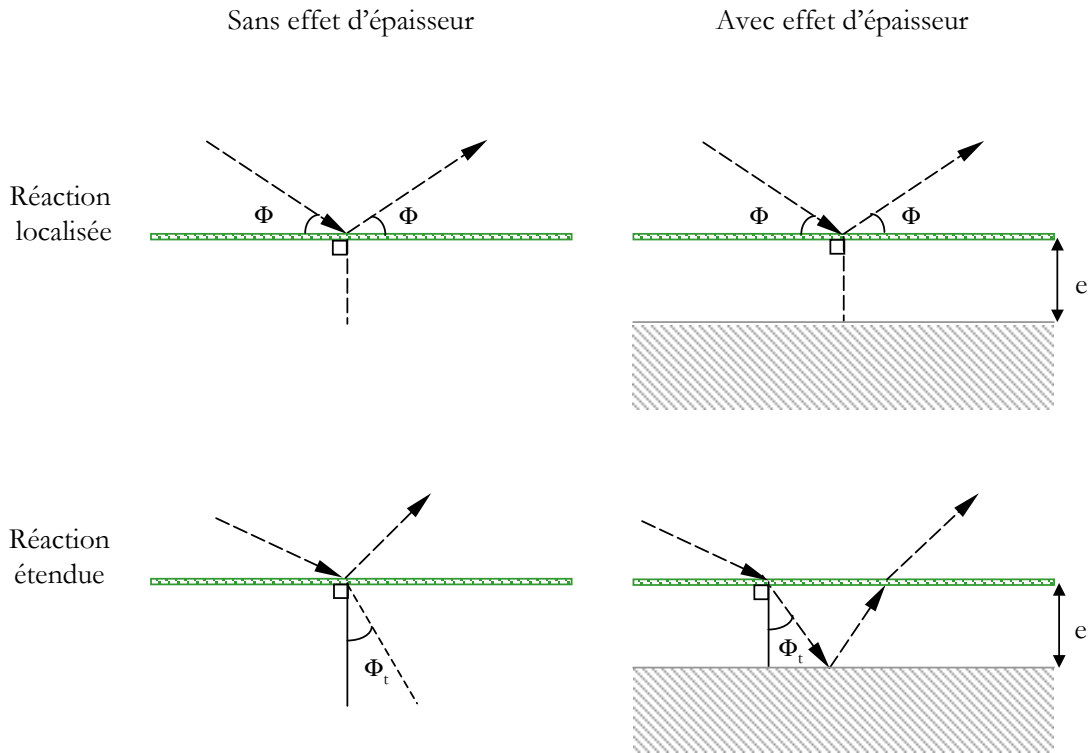


FIG. 1.1 – Parcours de rayons sonores issus d'une source ponctuelle au-dessus d'un sol présentant plusieurs types de comportement : réaction localisée ou réaction étendue, avec ou sans effet d'épaisseur.

Par ailleurs, lorsque le sol ne présente pas de propriétés vérifiant cette hypothèse, notamment les matériaux très absorbants d'épaisseur finie (revêtement routier poreux, laine minérale), on dit qu'on se place dans le cadre d'une *réaction étendue*. L'onde est alors transmise dans la couche de surface du sol (l'angle de transmission dépend de l'angle d'incidence et de la nature de l'interface). Si la seconde couche présente une impédance supérieure à la première, une onde confinée dans la première couche va transmettre de l'énergie au milieu de propagation, après perte d'énergie suite aux multiples réflexions sur les surfaces du dioptré ainsi formé. Ce phénomène est appelé *effet d'épaisseur* (figure 1.1). Dans les modèles à réaction étendue, le nombre de couches considéré est généralement égal à deux car il

est rare que les ondes sonores pénètrent à plus de quelques centimètres dans le sol ; le champ réfléchi des ondes issues des n -ièmes couches ($n > 2$) sont généralement négligeables vis-à-vis du champ total transmis. On peut par exemple supposer un nombre supérieur dans le cas de la modélisation de la propagation à travers des zones forestières. A noter que les deux modèles de réaction sont équivalents en incidence normale.

1.3.3.3 Paramètres caractéristiques et modèles d'impédance de sol

L'ensemble de ces phénomènes dépend finalement de la nature du sol et des paramètres qui les caractérisent. Un paramètre fondamental permet de caractériser les propriétés acoustiques du sol : la résistance spécifique au passage de l'air notée R_s (kN.s.m^{-4}). Des ordres de grandeur de ce paramètre en fonction de la nature du sol sont répertoriés dans le tableau 1.1. D'autres paramètres tels que la tortuosité ou la porosité peuvent être pris en compte ; tout dépend du modèle d'impédance utilisé dans laquelle sont décrits plus précisément ces différents modèles).

Nature du sol	Résistance spécifique au passage de l'air (kN.s.m^{-4})
Béton lisse	> 100 000
Revêtement routier	50000 à 100000
Terre compactée	2000 à 4000
Pré, champ de céréales	200 à 500
Sols engazonnés	150 à 300
Sous-bois	20 à 80
Neige fraîche, béton bitumineux poreux	5 à 20

TAB. 1.1 – Ordres de grandeur des valeurs de résistance spécifique au passage de l'air R_s (kN.s.m^{-4}) pour des types de sols spécifiques.

Les caractéristiques de sol sont souvent complexes à déterminer dans des situations réalistes. Il existe un certain nombre de modèles permettant d'estimer l'impédance du sol ; les modèles existant sont spécifiques au type de sol à caractériser (chaussée bitumineuse, surface herbeuse, etc.) et peuvent être plus ou moins complexes. On peut distinguer :

- des modèles phénoménologiques à un ou deux paramètres (*e.g.* modèle de Delany et Bazley) [41] ;
- des modèles permettant de caractériser la surface d'un point de vue microstructural (modèles de Stinson, d'Attenborough). Dans le cas par exemple du modèle d'Attenborough, quatre paramètres de sol doivent être bien connus (la résistance spécifique au passage de l'air, la forme des pores, la porosité ainsi que la tortuosité) [11, 12] ;
- des modèles adaptés pour les surfaces à résistance spécifique au passage de l'air élevée ;
- des modèles de relaxation [130] (la diffusion visqueuse et thermique est considérée comme un processus de relaxation) ;
- des modèles incluant le paramètre de rugosité de surface (se référer à [122] dans le cas de surfaces rugueuses dures et à [119] dans le cas de surfaces rugueuses à impédance finie).

On retiendra cependant deux modèles couramment utilisés pour le type de sol (sol herbeux) rencontré dans notre étude : le modèle de Delany et Bazley et le modèle de Zwicker et Kosten. Ces modèles sont moins évolués que des modèles microstructuraux, mais ils permettent toutefois une évaluation correcte des paramètres acoustiques de sol.

(a) *Modèle semi-empirique de Delany et Bazley*

Le modèle de Delany et Bazley était initialement un modèle de caractérisation de matériaux absorbants poreux [41]. Son domaine de validité était limité aux bornes $0,01 < \frac{f}{R_s} < 1 \text{ m}^3.\text{kg}^{-1}$ avec f la fréquence et R_s la résistance spécifique au passage de l'air. Chessel a montré que ce modèle était également adapté pour caractériser l'impédance des sols herbeux pour des valeurs de $\frac{f}{R_s}$ variant de 10^{-4} à $0,1 \text{ m}^3.\text{kg}^{-1}$. Il fournit une expression du nombre d'onde dans le sol et de l'impédance acoustique selon les relations suivantes :

$$\begin{cases} k = \frac{\omega}{c} \left(1 + 10,3 \left(\frac{f}{R_s} \right)^{-0,693} + 12,1j \left(\frac{f}{R_s} \right)^{-0,618} \right) \\ Z = Z_0 \left(1 + 9,08 \left(\frac{f}{R_s} \right)^{-0,754} + 11,9j \left(\frac{f}{R_s} \right)^{-0,732} \right) \end{cases}, \quad (1.5)$$

où c la célérité du son dans l'air, f la fréquence, R_s la résistance spécifique au passage de l'air du matériau considéré, k le nombre d'onde et $j^2 = -1$.

(b) *Modèle de Zwikker et Kosten*

Zwikker et Kosten ont également développé différents modèles caractérisant l'impédance de matériaux poreux. Dans ce type de modèles, le sol est considéré comme étant un milieu poreux semi-infini ou une couche poreuse à support rigide. Le nombre d'onde et l'impédance acoustique s'écrivent selon les expressions suivantes :

$$\begin{cases} k = \frac{\omega}{c} \sqrt{q_c \Omega} \\ Z = \sqrt{\frac{q_c}{\Omega}} \end{cases}, \quad (1.6)$$

où $q_c = \frac{C_s}{\Omega} + j \frac{R_s}{\rho \omega}$.

Ω correspond à la porosité, C_s à la constante de structure, R_s à la résistance spécifique au passage de l'air et ρ à la masse volumique de l'air. $\frac{1}{\Omega}$ est le volume disponible dans le matériau poreux ; la constante de structure C_s rend compte de la structure spécifique des pores et des propriétés du matériau.

(c) *Correction d'épaisseur*

Dans le cas d'un sol naturel ou herbeux, l'impédance Z peut être corrigée par le paramètre d'épaisseur e (m) :

$$Z_c = Z \coth(-jke). \quad (1.7)$$

Ces modèles restent cependant limités aux basses et moyennes fréquences car la rugosité du sol induit de la diffraction qui vient perturber le champ aux plus hautes fréquences. Ces modèles ne permettent malheureusement pas la prise en compte de tels phénomènes.

Les valeurs des paramètres sont propres à chacun des modèles : un même paramètre pour un même type de sol prendra des valeurs différentes selon le modèle utilisé. C'est pourquoi ces paramètres sont appelés paramètres effectifs. De plus les modèles de sol impliquent des paramètres qui ne sont généralement pas directement mesurables en milieu naturel.

1.3.3.4 Estimation des paramètres caractéristiques du sol

Les paramètres caractéristiques du sol (R_s par exemple) sont rarement mesurés de manière directe *in situ*. Cette caractéristique de sol R_s peut être accessible uniquement par méthode inverse en ayant recours à un modèle de propagation. La description du principe de la mesure est inspirée des travaux de O. Baume [21].

Le principe expérimental est basé sur le modèle de propagation de Rudnick pour lequel, dans le cas d'un sol homogène et sans effet de réfraction du milieu de propagation la pression acoustique p_t est donnée par l'expression générale $p_t(r) = p_d(r) + Q \cdot p_r(r)$, où p_d est la pression acoustique due à l'onde directe, p_r la pression acoustique issue de la source image et Q le coefficient de réflexion du sol en onde sphérique. L'objectif des mesures est d'extraire la valeur de l'atténuation due au champ réfléchi. Le dispositif expérimental est schématisé (figure 1.2) et rend compte de la disposition de la source et de deux microphones. Pour les deux microphones, il y a contribution du champ direct et du champ réfléchi. Le rapport des pressions mesurées aux deux microphones est indépendante de la puissance de la source et ne dépend que des propriétés du sol.

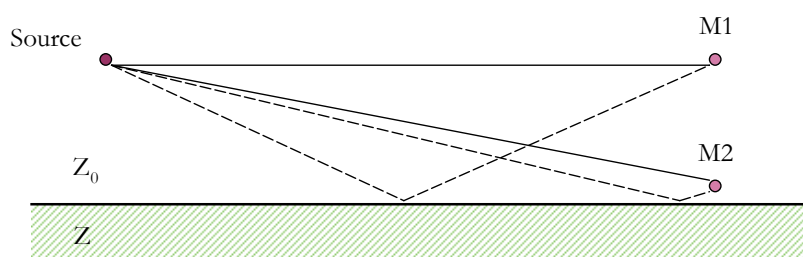


FIG. 1.2 – Dispositif expérimental d'une mesure d'impédance de sol *in situ*.

Le principe de la méthode consiste à ajuster un modèle d'atténuation aux données mesurées. Le modèle d'impédance acoustique généralement utilisé est le modèle de Delany et Bazley avec effet d'épaisseur. Deux paramètres phénoménologiques permettent ainsi d'ajuster le modèle aux données : la résistance spécifique au passage de l'air R_s et l'épaisseur de la couche poreuse e . Le calage est effectué en ajustant manuellement les paramètres du modèle selon des critères très dépendants de l'opérateur. C'est grâce à cette méthode inverse que sont accessibles les caractéristiques acoustiques de sol (pour un sol herbeux).

1.3.3.5 Incertitude et variabilité associées aux caractéristiques acoustiques de sol

L'incertitude expérimentale associée à ce paramètre peut atteindre environ 20 kN.s.m^{-4} et intègre l'incertitude de calage, l'incertitude de répétabilité du signal ainsi que l'incertitude métrologique. En effet, la géométrie et le choix du système expérimental permettant d'accéder à l'impédance acoustique a un impact non-négligeable sur la précision des données [116]. Selon la nature du sol considéré, une telle incertitude sur le paramètre R_s en tant que paramètre d'entrée des modèles, peut mener à des incertitudes considérables en termes de niveaux sonores comme l'illustre la figure 1.3.

Cette figure représente la dispersion que l'on peut observer sur des niveaux sonores en considérant des variations du paramètre de résistance spécifique au passage de l'air de 30 kN.s.m^{-4} (valeur arbitraire supérieure à l'incertitude expérimentale car des sources d'incertitudes complémentaires sont à prendre en compte) pour différents types de sol. Une telle variation de R_s se traduit par une sensibilité variable des niveaux sonores selon la nature du sol ainsi que la distance de propagation considérée. On s'aperçoit que plus le sol est réfléchissant (valeurs de R_s de l'ordre de 10^5 kN.s.m^{-4}) moins les niveaux sonores sont sensibles à une variation du paramètre R_s . Si l'on s'affranchit de

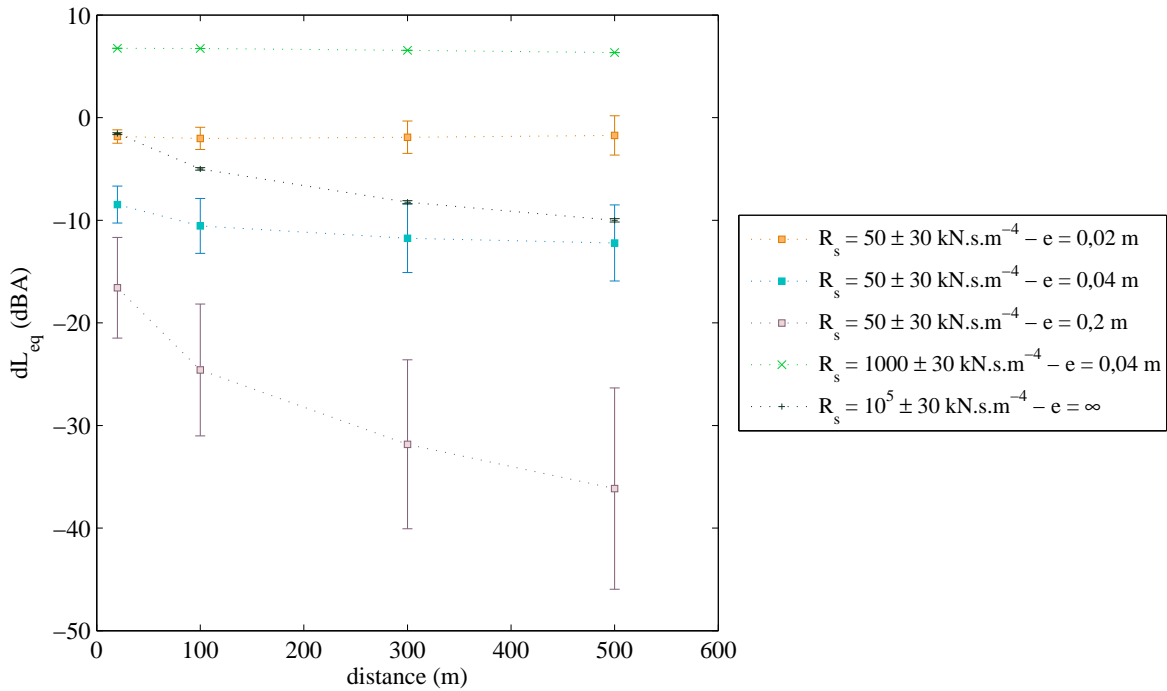


FIG. 1.3 – Incertitudes sur les niveaux d'atténuation sonores en dB(A) pour des sols très absorbants (avec ou sans effet d'épaisseur) à réfléchissant (sans effet d'épaisseur) en fonction de la distance à la source. Les écarts-type associés aux résistances spécifique au passage de l'air sont de 30 kN.s.m^{-4} . La source et le récepteur sont placés à $0,05 \text{ m}$ au-dessus du sol. Les calculs sont menés à l'aide d'un modèle de propagation analytique [104] et du modèle d'impédance de Delany et Bazley, en considérant un gradient de célérité positif (section 1.3.4.2) afin de renforcer les effets de sol.

l'effet d'épaisseur, les niveaux d'atténuation sonore sont d'autant plus sensibles et donc présentent une dispersion d'autant plus grande que la résistance spécifique au passage de l'air est faible (sol absorbant). Si l'on compare maintenant l'effet de l'incertitude en fonction de l'effet d'épaisseur (à savoir pour $R_s = 50 \text{ kN.s.m}^{-4}$), atténuation et dispersion augmentent pour une épaisseur e croissante. D'un point de vue spatial, l'atténuation des niveaux sonores augmente avec la distance, quel que soit le type de sol considéré. On observe bien que l'atténuation est d'autant plus grande lorsque le sol est absorbant. De plus, la dispersion associée aux niveaux d'atténuation augmente également pour des distances croissantes.

Ces valeurs évoluent donc aussi bien en fonction du type de sol mais aussi, pour des sols naturels, en fonction des conditions micrométéorologiques (rosée, humidité relative) voire météorologiques (présence de neige, gel, etc.). L'impédance de sol suit donc une variabilité journalière et saisonnière. Ces caractéristiques évoluent également spatialement ; la dispersion locale peut être considérable (ces éléments seront notamment étudiés au chapitre 5). Un site dont le sol est considéré comme homogène (sol naturel herbeux par exemple) et auquel on attribuerait une valeur unique, peut donc présenter une forte variabilité. Des études géostatistiques ont notamment été menées pour caractériser la structure spatio-temporelle de ce paramètre dans une telle situation [18, 22, 103]. Précipitations et ensoleillement ont une influence non-négligeable sur ce paramètre.

Notre problématique d'étude impliquant généralement la considération de sols herbeux (valeurs de R_s relativement faibles), la prise en compte des incertitudes associées à ces caractéristiques acoustiques de sol est importante.

1.3.4 Effets de vent et de température

1.3.4.1 Caractérisation de la couche limite de surface

La propagation des ondes sonores se produit principalement dans les basses couches de l'atmosphère et notamment dans une couche appelée Couche Limite Atmosphérique (C.L.A.). La C.L.A. s'étend sur des altitudes allant d'une centaine de mètres à environ 1 km ; son sommet est souvent caractérisé par la présence de nuages et notamment de cumulus ou stratocumulus. Les mécanismes *thermiques* et *aérodynamiques* sont plus ou moins actifs et jouent sur l'épaisseur de la C.L.A. (figure 1.4). D'un point de vue aérodynamique, la C.L.A. est définie comme étant la zone de l'atmosphère où l'écoulement du fluide est influencé par l'interaction avec la surface terrestre directement. Le temps de réponse à un forçage (rugosité, relief, couvert végétal, évaporation, transfert de chaleur, etc.) est court, de l'ordre de l'heure. D'un point de vue thermique, la C.L.A. est la zone de l'atmosphère au voisinage de la surface terrestre dans laquelle la variation diurne du rayonnement solaire est directement perceptible.

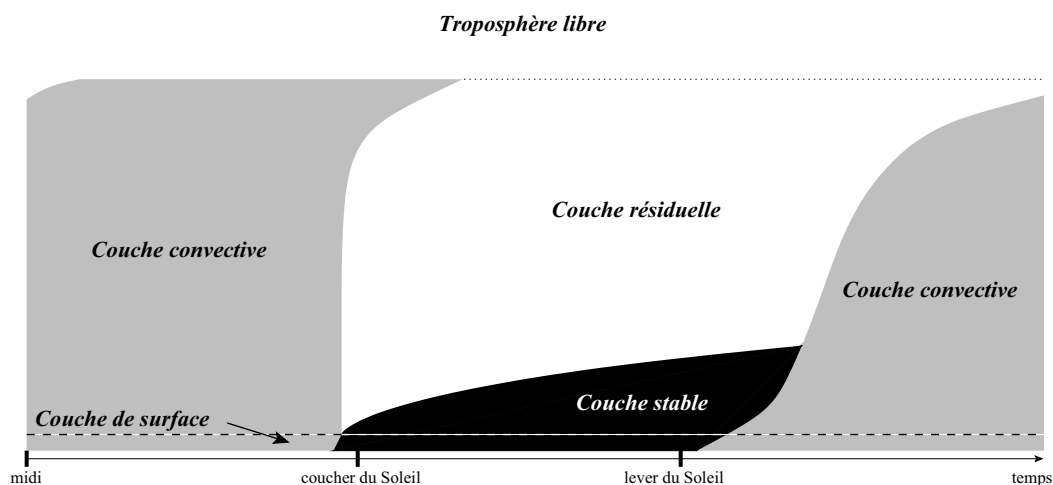


FIG. 1.4 – Évolution schématique de la couche limite atmosphérique au cours d'une journée [113].

En journée, le sol réchauffe la C.L.A. par transfert turbulent : le flux de chaleur turbulent est positif et le gradient de température négatif. Une couche de mélange ou couche convective caractérisée par une turbulence très forte apparaît alors. Cette turbulence contribue à homogénéiser l'ensemble des grandeurs associées au fluide (température, quantité de mouvement, etc.). L'atmosphère est en stratification *instable*. Au sommet de la C.L.A., une inversion thermique de stratification *stable* bloque les ascensions d'air qui replongent dans la couche de mélange provoquant l'entraînement des masses d'air de l'atmosphère libre dans la C.L.A..

Le soir, le sol se refroidit : la température du sol est inférieure à la température de l'air. Le flux de chaleur turbulent est négatif, dirigé vers le sol (gradient de température positif) et il y a destruction de la turbulence d'origine dynamique. Le mélange est peu turbulent et la couche limite nocturne est mince. L'atmosphère est en stratification *stable*. Au-dessus, la couche de mélange de la journée précédente devient une couche *résiduelle*, en général neutre (gradient de température nul).

Les phénomènes prépondérants pour la propagation acoustique se déroulent dans la Couche Limite de Surface (C.L.S.) dont l'épaisseur est de l'ordre de quelques dixièmes de l'épaisseur de la C.L.A. (une centaine de mètres les jours de vent fort et quelques mètres les nuits claires et par vent faible). Le profil vertical de la vitesse du vent est perturbé du fait du frottement de l'air sur le sol, où la présence d'obstacles ou de discontinuités modifie la trajectoire du vent. La C.L.S. est donc une zone à fort gradient de vitesse de vent. De plus, la partie inférieure de la C.L.S. (sur une hauteur qui varie de quelques centimètres à quelques dizaines de mètres) est fortement soumise à la rugosité du

sol ; l'écoulement de l'air est alors fortement turbulent, non homogène et non stationnaire. Cette zone est souvent caractérisée par une rugosité globale, homogène à une longueur reliée à la nature du recouvrement appelée longueur de rugosité que l'on note z_0 (m) (tableau 1.2).

Nature du sol	Paramètre de rugosité z_0 (m)
Etendues d'eau libre	0,001-0,01
Sol nu	0,005-0,02
Gazon ras/dense	0,001-0,02
Cultures agricoles - Habitats épars	0,1-0,6
Forêts - Zones périurbaines	0,5-1,8
Urbain dense	1,5-10

TAB. 1.2 – Valeurs typiques du paramètre de rugosité z_0 [92, 113].

De nombreux phénomènes physiques complexes se développent au sein de la C.L.S. et affectent la cohérence du champ acoustique étudié. Les facteurs thermiques et aérodynamiques rendent compte de l'hétérogénéité spatiale et de la variabilité au sein du milieu de propagation. Celui-ci peut alors être décrit par un indice de réfraction acoustique noté n , par analogie avec les lois de l'optique. Cet indice varie dans l'espace et principalement avec la hauteur. Son expression s'écrit comme suit [23, 132] :

$$n(\mathbf{r}) = \frac{c(\mathbf{r})}{c_0} = \langle n(\mathbf{r}) \rangle + \mu(\mathbf{r}), \quad (1.8)$$

où $c(\mathbf{r})$ est la célérité effective de l'onde sonore dans le milieu traversé et c_0 la célérité de référence ($c_0 \approx 343 \text{ m.s}^{-1}$ à la température $T_0 = 293,15 \text{ K}$).

On distingue donc une partie moyenne ou déterministe de l'indice de réfraction $\langle n(\mathbf{r}) \rangle$ et une partie fluctuante ou aléatoire $\mu(\mathbf{r})$. Deux phénomènes différents mais complémentaires agissent donc sur la propagation acoustique à grande distance : les phénomènes de réfraction moyenne et de turbulence atmosphérique. La combinaison et l'interaction de ces phénomènes complexifient manifestement les caractéristiques du milieu de propagation et par conséquent la description et la compréhension des phénomènes inhérents à la propagation acoustique.

1.3.4.2 Réfraction moyenne

(a) Généralités

La réfraction moyenne des ondes sonores est directement induite par la stratification verticale de l'atmosphère. Les profils verticaux moyens de célérité effective du son qui caractérisent cet effet sont calculés à partir de profils verticaux moyens de vent et de température : ces profils sont considérés comme stationnaires sur la période considérée. Les échelles de temps considérées sont de l'ordre de la minute à l'heure (généralement moyennés sur des périodes de quinze minutes) et intègrent la turbulence de grande échelle qui est notamment induite par l'intermittence du vent apparaissant par convection. D'un point de vue spatial, dans le cas où l'on considère un site plan et dégagé, une approximation supplémentaire est de négliger l'évolution des profils verticaux avec la distance. Les profils sont alors supposés indépendants de la distance à la source r . La célérité effective moyenne du son s'écrit alors selon l'expression suivante :

$$\langle c(z) \rangle = \sqrt{\gamma RT(z)} + \langle u(z) \rangle \cos \vartheta = c_0 \sqrt{\frac{\langle T(z) \rangle}{T_0}} + \langle u(z) \rangle \cos \vartheta, \quad (1.9)$$

où $\gamma \approx 1,4$ est le rapport des chaleurs spécifiques de l'air, $R = 286,69 \text{ J.kg}^{-1}.\text{K}^{-1}$ est la constante des gaz parfaits, ϑ est l'angle entre le vecteur vent et la direction de propagation du son entre source

et récepteur. $\cos \vartheta$ traduit la projection de la direction du vent sur la direction de propagation du son. $\langle T(z) \rangle$ et $\langle u(z) \rangle$ sont respectivement la température moyenne et la vitesse de vent moyenne à l'altitude z . L'équation 1.9 permet ainsi par dérivation et développement limité d'exprimer le gradient vertical moyen de célérité effective :

$$\frac{\partial \langle c(z) \rangle}{\partial z} \approx \frac{1}{2} \frac{\gamma R}{c_0} \frac{\partial \langle T(z) \rangle}{\partial z} + \frac{\partial \langle u(z) \rangle}{\partial z} \cos \vartheta. \quad (1.10)$$

Pour des raisons de commodité, on décompose cette expression comme la somme d'un paramètre de réfraction thermique pondéré a_{T_p} et d'un paramètre de réfraction aérodynamique projeté a_{V_p} , où $a_{T_p} = \frac{1}{2} \frac{\gamma R}{c_0} \frac{\partial \langle T(z) \rangle}{\partial z}$ et $a_{V_p} = \frac{\partial \langle u(z) \rangle}{\partial z} \cos \vartheta$. Ces notations peu communes permettent de simplifier l'expression du gradient vertical moyen de célérité du son et de quantifier les contributions relatives des effets thermiques et aérodynamiques [57].

Les gradients verticaux de célérité étant d'autant plus grands à proximité du sol qu'en altitude, une modélisation des profils de célérité par des fonctions linéaires n'est pas adéquate. Afin de prendre en compte cette tendance, l'introduction de profils logarithmiques [60], de profils multi-linéaires [90], ou de profils hybrides de type « log-lin » a été proposée. La figure 1.5 montre les différences qu'il peut exister entre de tels profils.

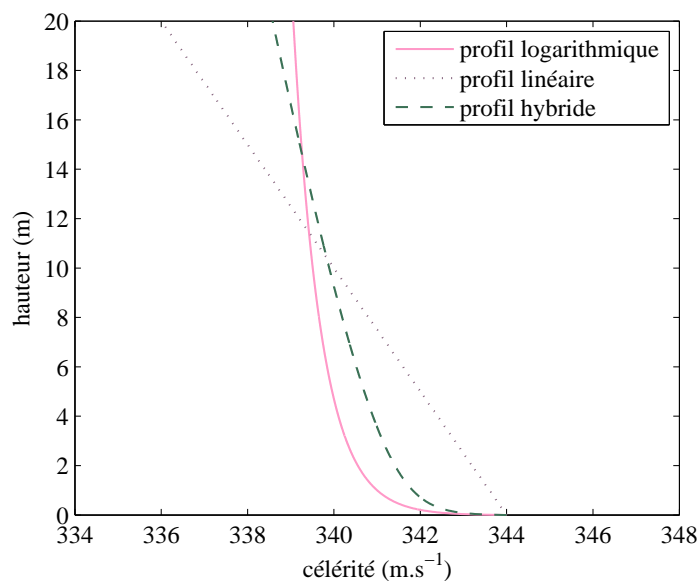


FIG. 1.5 – Comparaison des profils de célérité de type logarithmique, linéaire et hybride.

$$c(z) = \begin{cases} c_0 + a_{log} \log \left(1 + \frac{z}{z_0} \right) + b_{lin} (z - z_0) & \text{si } z > z_0 \\ c_0 & \text{si } 0 \leq z \leq z_0 \end{cases}, \quad (1.11)$$

où c_0 est la célérité de référence et a_{log} et b_{lin} sont les paramètres de réfraction. Ces paramètres peuvent être soit estimés soit calculés. L'estimation est basée sur l'ajustement du profil de célérité théorique (par minimisation au sens des moindres carrés) sur des points de mesures effectués à différentes hauteurs z , un minimum de trois points de mesures est requis. Ces mesures sont généralement menées à l'aide d'un mât micrométéorologique conventionnel présentant trois hauteurs de capteurs. Un exemple de profil de célérité recalé est donné figure 1.6.

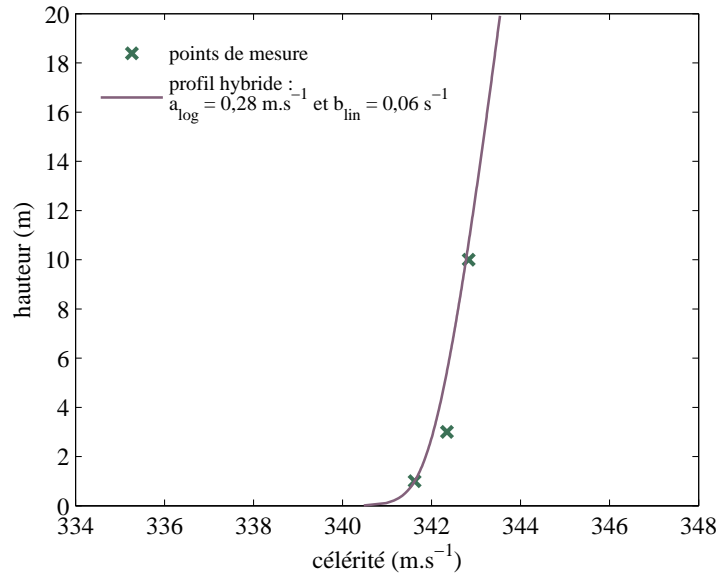


FIG. 1.6 – Profil de célérité de type hybride recalé sur des points de mesure.

La seconde méthode permettant d'accéder aux profils verticaux de célérité du son est basée sur la théorie de similitude de Monin-Obukhov [88, 113]. Cette théorie fait l'hypothèse que les flux turbulents sont stationnaires et constants selon l'altitude au sein de la C.L.S.. Elle permet de regrouper l'ensemble des grandeurs fondamentales sous la forme d'échelles caractéristiques. Sous les hypothèses de stratification verticale et de vent soutenu, la théorie de similitude de Monin-Obukhov permet d'exprimer les paramètres de réfraction a_{log} et b_{lin} à partir des trois échelles caractéristiques U^* , T^* et $\zeta = \frac{z}{L_{MO}}$:

En conditions neutres et stables ($0 \leq \zeta < 1$), le cisaillement joue un rôle prépondérant (plus particulièrement pour ζ positif) :

$$\begin{cases} a_{log} = \frac{1}{2} \frac{\gamma R}{c_0} \frac{0,74 T^*}{\kappa} + \frac{U^*}{k} \cos \vartheta \\ b_{lin} = \frac{1}{2} \frac{\gamma R}{c_0} \left(\frac{T^*}{\kappa} \frac{4,7}{L_{MO}} - \Gamma_d \right) + \frac{U^*}{\kappa} \frac{4,7}{L_{MO}} \cos \vartheta. \end{cases} \quad (1.12)$$

En conditions instables ($-0,8 < \xi < 0$), la poussée convective domine :

$$\begin{cases} a_{log} \approx \frac{1}{2} \frac{\gamma R}{c_0} \frac{0,74 T^*}{\kappa} + \frac{U^*}{\kappa} \cos \vartheta \\ b_{lin} \approx \frac{1}{2} \frac{\gamma R}{c_0} \left(\frac{T^*}{\kappa} \frac{0,74}{L_{MO}} - \Gamma_d \right) + \frac{U^*}{\kappa} \frac{1,0}{L_{MO}} \cos \vartheta, \end{cases} \quad (1.13)$$

où κ est la constante de von Kármán ($\kappa = 0,4$), Γ_d le gradient sec de température adiabatique ($\Gamma_d = 0,0098 \text{ K.m}^{-1}$), R la constante des gaz parfaits ($R = 286,69 \text{ J.kg}^{-1}.\text{K}^{-1}$), γ le rapport des capacités calorifiques massiques de l'air ($\gamma \approx 1,4$), ϑ la direction de propagation ($^\circ$). U^* est appelée vitesse de frottement (frottement engendré par le sol sur le fluide), T^* est l'échelle de température (et traduit les effets du flux de chaleur vertical) et L_{MO} est la longueur de Monin-Obukhov (*i.e.* la longueur à partir de laquelle les forces de cisaillement deviennent négligeables devant les forces convectives). Ces échelles caractéristiques sont définies par les expressions suivantes [113] :

$$\begin{cases} U^* = \left(\overline{u'w'^2} + \overline{v'w'^2} \right)^{1/4} \\ T^* = \frac{-\overline{w'\theta'_v}}{U^*} \\ L_{MO} = - \frac{\theta_v \left(\overline{u'w'} + \overline{v'w'} \right)^{3/4}}{\kappa g \overline{w'\theta'_v}} \end{cases}, \quad (1.14)$$

θ_v la température virtuelle définie par $\theta_v \approx T - 0,01z$ avec z la hauteur de mesure. Ces expressions montrent la dépendance des échelles caractéristiques avec les covariances de grandeurs mesurées, *i.e.* les composantes horizontales (u , v) et verticale (w) du vecteur vent et la température virtuelle. Le paramètre adimensionnel de stabilité $\xi = \frac{z}{L_{MO}}$ définit la stabilité de l'atmosphère et est ainsi directement reliés aux conditions de propagation. En pratique ces grandeurs sont accessibles expérimentalement à l'aide d'anémomètres ultrasoniques tridimensionnels.

Ces deux méthodes permettant l'estimation des paramètres de réfraction a_{log} et b_{lin} sont liées à des techniques de mesure distinctes, toutes deux soumises à des approximations. Les expressions de ces paramètres de réfraction ne sont à l'heure actuelle pas stabilisées et donnent toujours lieu à des travaux de recherche. De plus, le paramètre de rugosité z_0 intervenant dans l'expression du profil vertical de célérité (équation 1.11) est extrêmement influant mais délicat à estimer dans certains milieux. Méthodologie et approximations constituent ainsi une part d'incertitude conséquente sur les profils de célérité et sur les paramètres de réfraction. La présence de telles variables a un impact fondamental sur la propagation acoustique.

(b) Propagation en présence de gradients verticaux de célérité

En l'absence de vent, lorsque la température est plus élevée en altitude qu'au niveau du sol (gradient vertical de température $\frac{\partial T}{\partial z} > 0$), ce qui correspond à un phénomène d'inversion rencontré habituellement en période nocturne, les phénomènes de convection thermique sont quasi inexistantes et le milieu est dit *stable*. Les rayons de propagation acoustique se courbent alors vers le bas, le niveau sonore décroît plus lentement (figure 1.7). Ces conditions sont dites « favorables » à la propagation. L'énergie sonore étant concentrée proche du sol, les effets de sol (section 1.3.3) en sont alors renforcés.

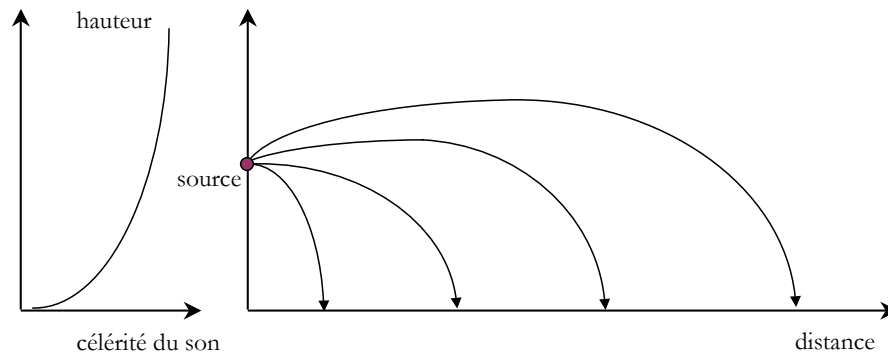


FIG. 1.7 – Effets d'un gradient vertical de température positif sur la propagation des rayons sonores en l'absence de vent.

A contrario lorsque la température est plus élevée au niveau du sol qu'en altitude (conditions micrométéorologiques dites *instables*), ce qui est généralement le cas en période diurne (fort ensoleillement, faible nébulosité), les rayons de propagation acoustique se courbent vers le haut : des zones d'ombre apparaissent de part et d'autre de la source et la vitesse du son décroît avec l'altitude (Figure 1.8). Ces conditions sont dites « défavorables » à la propagation. Au sein d'une zone d'ombre, les effets de sol sont quasiment négligeables, on notera cependant l'existence d'*ondes rampantes* fortement influencées par la nature du sol, notamment dans le cas d'un sol absorbant.

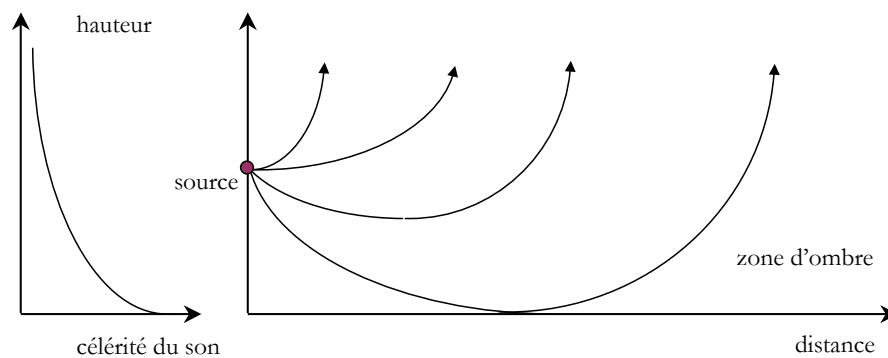


FIG. 1.8 – Effets d'un gradient vertical de température négatif sur la propagation des rayons sonores en l'absence de vent.

Dans le cas où le vent est portant, la célérité du son a tendance à augmenter avec l'altitude ; les rayons sonores se courbent vers le bas ce qui entraîne une augmentation des niveaux sonores (Figure 1.9). A contrario, en cas de vent contraire, la vitesse du vent se soustrait à la célérité du son ce qui entraîne une courbure des rayons acoustiques vers le haut et donc une diminution des niveaux sonores.

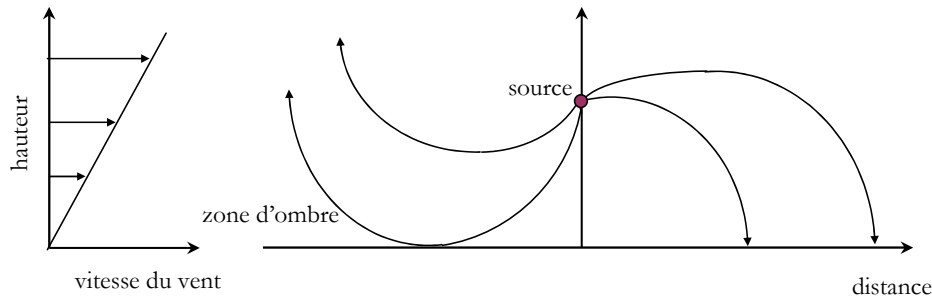


FIG. 1.9 – Effets d'un gradient vertical de vent sur la propagation des rayons sonores en l'absence de gradient de température.

Dans certaines conditions bien spécifiques et rares, le gradient vertical de célérité peut être nul : les conditions de propagation sont dites homogènes et conduisent à une propagation acoustique suivant des rayons rectilignes. Ce phénomène peut se produire dans le cas où la vitesse de vent est nulle et la température constante avec l'altitude. Une telle situation peut être rencontrée de façon éphémère au lever ou au coucher du soleil ou dans des conditions de couverture nuageuse totale. Ce phénomène peut également se produire lorsque les effets aérodynamiques et thermiques se compensent, comme par exemple en présence de vent contraire à la direction de propagation, de nuit, lorsque le ciel est dégagé. Ces deux phénomènes étant relativement rares, les conditions homogènes à la propagation constituent davantage un état passager entre deux modes de propagation qu'un mode de propagation à part entière.

(c) Effet de la réfraction sur les niveaux sonores

D'un point de vue théorique, l'évolution des niveaux sonores en fonction du gradient vertical de célérité présente une courbe en « S », avec une très forte sensibilité proche des conditions homogènes ($\frac{\partial c}{\partial z} = 0$) et une sensibilité moindre pour des conditions plus extrêmes (fort ou faible gradients) avec la présence de paliers. On se propose d'observer cette évolution d'un point de vue numérique puis expérimental.

D'un point de vue numérique, les calculs des niveaux d'atténuation sonores sont menés à trois distances de propagation (100, 300 et 500 m) relativement à un microphone situé à 10 m de la source pour un sol herbeux ($R_s = 100 \text{ kN.s.m}^{-4}$) et un sol réfléchissant ($R_s = 10^5 \text{ kN.s.m}^{-4}$). Afin de rendre compte de la sensibilité des niveaux sonores aux valeurs de gradients verticaux de célérité, on leur associe un écart-type de $0,06 \text{ s}^{-1}$. Cette valeur est choisie en regard d'une étude menée sur l'incertitude expérimentale associée aux gradients verticaux de célérité (Annexe E). Les calculs sont menés à l'aide d'une méthode de référence basée sur la résolution de l'Equation Parabolique (section 2.3). Les résultats sont représentés figures 1.10 et 1.11.

Dans les deux configurations (sol absorbant et sol réfléchissant), les résultats de calcul présentent une bonne cohérence de l'évolution des niveaux sonores et de leur étendue avec la distance : l'atténuation et l'étendue des niveaux sonores augmentent avec la distance. On vérifie également que plus la distance est grande plus les conditions micrométéorologiques sont influentes sur la propagation acoustique. De plus, la sensibilité des niveaux sonores aux gradients verticaux de célérité est plus marquée pour un sol réfléchissant. Les figures interférentielles dépendent évidemment de la valeur de la résistance spécifique au passage de l'air (section 1.3.3). À noter cependant que les calculs sont menés sans turbulence ce qui implique la présence d'une zone d'ombre notamment en conditions défavorables à la propagation, où les atténuations sont extrêmement marquées. Les résultats numériques présentés n'illustrent donc pas exactement la réalité physique mais plutôt l'influence de la variabilité du gradient de célérité seul.

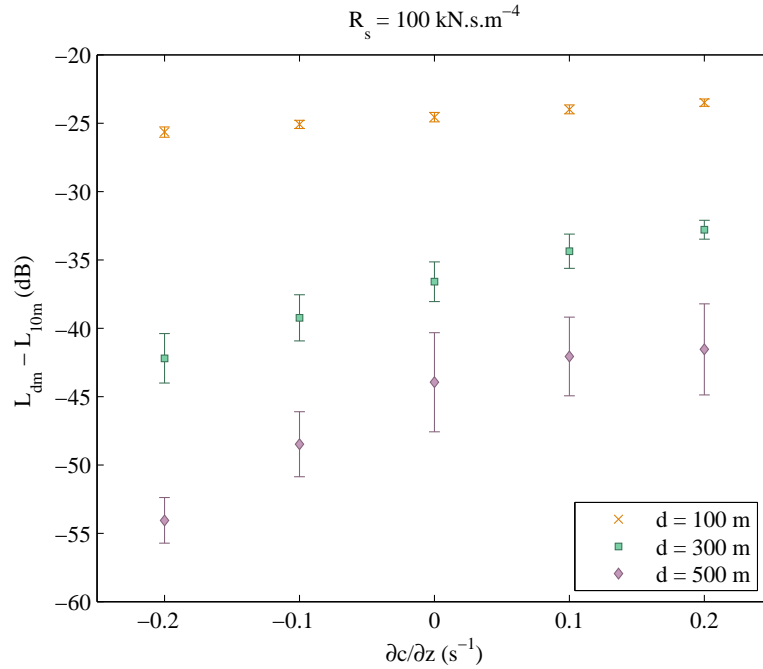


FIG. 1.10 – Incertitudes sur les niveaux d’atténuation sonore en dB(A) à trois distances de propagation d relativement à un microphone situé à 10 m de la source pour un sol herbeux ($R_s = 100 \text{ kN.s.m}^{-4}$). Les écarts-type associés à chaque gradient de célérité effective sont de $0,06 \text{ s}^{-1}$. La source et le récepteur sont placés respectivement à 0,05 m et 2 m de haut.

Pour examiner le *point de vue expérimental*, on utilise les données issues de la campagne expérimentale de Lannemezan 2005 (Annexe A). Ces mesures sont effectuées à une distance à la source ponctuelle de 150 m, au-dessus d’un sol herbeux plan. Elles rendent compte des effets de sol (la moyenne des valeurs de R_s est de 100 kN.s.m^{-4} avec un écart-type de 20 kN.s.m^{-4}) mais surtout des effets micrométéorologiques prépondérants à cette distance de propagation. Cette évolution est représentée par la figure 1.12 relativement à des mesures effectuées à 10 m de la source afin de s’affranchir de la variabilité de la source elle-même.

Ce graphique rend compte de l’incertitude et de la variabilité physique des conditions micrométéorologiques et des propriétés du sol sur une période de trois mois de mesure. Les niveaux d’atténuation sonores présentent un palier en conditions favorables à la propagation. De plus, ces atténuations et leur dispersion sont d’autant plus fortes que les conditions de propagation sont défavorables (soit $\frac{\partial c}{\partial z} < 0$). Il est intéressant de comparer la dispersion en conditions défavorables à la propagation où les effets de sol sont moins marqués. D’un point de vue numérique, à une distance de propagation de 100 m, la dispersion des niveaux d’atténuation est de l’ordre de 1 dB(A) pour une valeur de gradient négative (pas de variation de la résistance spécifique au passage de l’air). D’un point de vue expérimental, cette dispersion peut notamment atteindre 8 dB(A) pour un gradient de l’ordre de $-0,05 \text{ s}^{-1}$, ce qui est considérable en regard de la dispersion observée figure 1.10.

Les phénomènes de réfraction témoignent de l’effet micrométéorologique principal sur la propagation acoustique en milieu extérieur. Ce phénomène se déroule au sein de la C.L.S. dont l’état évolue selon un cycle journalier : profils verticaux de vent et de température sont variables dans le temps, ce qui permet globalement d’appréhender les effets en terme acoustique. Il est cependant nécessaire de ne pas se baser que sur cette structure générale et de considérer son évolution temporelle *i.e.* sa

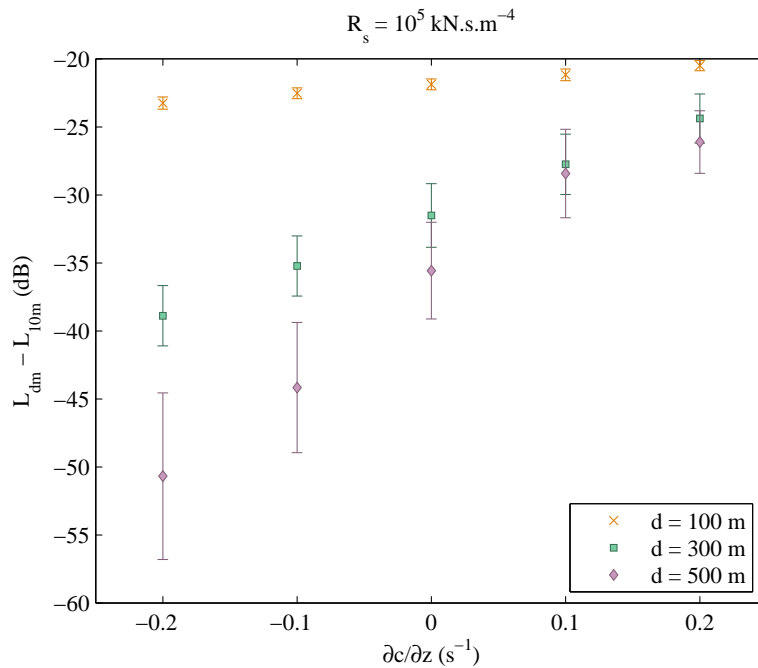


FIG. 1.11 – Incertitudes sur les niveaux d'atténuation sonore en dB(A) à trois distances de propagation d relativement à un microphone situé à 10 m de la source pour un sol de type revêtement routier ($R_s = 10^5 \text{ kN.s.m}^{-4}$). Les écarts-type associés aux gradients de célérité effective sont de $0,06 \text{ s}^{-1}$. La source et le récepteur sont placés respectivement à 0,05 m et 2 m de haut.

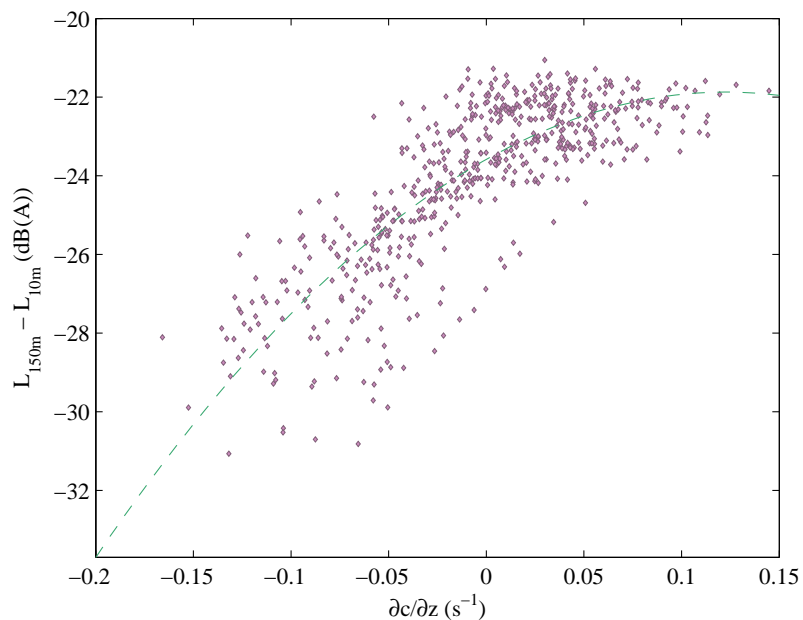


FIG. 1.12 – Évolution du niveau d'atténuation sonore en fonction du gradient vertical de célérité du son.

variabilité sur des périodes inférieures à l'heure, voire la minute, voire la seconde. Sur ces périodes, les phénomènes rencontrés sont liés à la turbulence atmosphérique.

1.3.4.3 Turbulence atmosphérique

(a) Additivité et continuité des champs turbulents

La partie aléatoire $\mu(\mathbf{r})$ de l'indice de réfraction peut être exprimée par une relation qui rend compte des variations spatio-temporelles de la vitesse du vent u' (composante longitudinale du vent alignée avec la direction de propagation) et de la température T' . Ces composantes aléatoires permettent d'écrire les fluctuations $\mu(\mathbf{r})$ [128] :

$$\mu(\mathbf{r}) = -\frac{T'}{2T_0} - \frac{u'}{c_0} \quad (1.15)$$

où T_0 est la température de référence ($T_0 = 293,15$ K) et c_0 est la célérité du son dans l'air à la température T_0 ($c_0 = 343$ m.s⁻¹).

L'analyse harmonique de la vitesse du vent suggère que la turbulence résulte de la superposition de mouvements oscillatoires plus ou moins irréguliers d'amplitude, de phase et de longueur d'onde différentes, en d'autres termes, de tourbillons de tailles différentes : l'énergie turbulente est liée à la turbulence générée par la viscosité et aux grandes circulations d'air convectives. Outre cette additivité, l'interaction entre ces échelles doit être prise en compte : il n'existe en effet pas de séparation complète entre les échelles micrométéorologiques ce qui est illustré par un spectre temporel de la vitesse du vent en surface (figure 1.13).

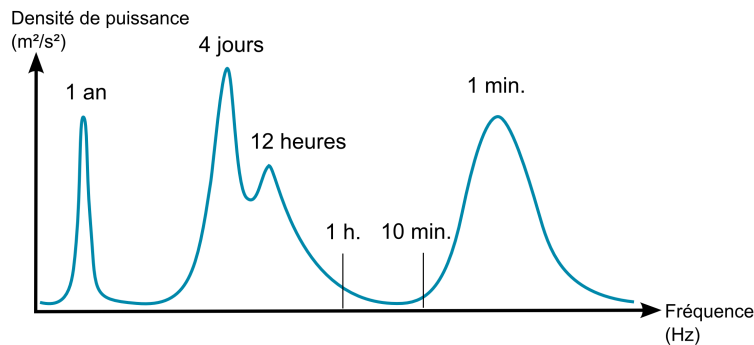


FIG. 1.13 – Spectre énergétique de la vitesse du vent [48].

Une analyse spectrale permet de retrouver l'échelle des fluctuations de l'écoulement de l'air. Les pics du spectre montrent dans quelle mesure les tourbillons participent effectivement à l'énergie turbulente :

- le pic relatif à une période d'environ 100 heures traduit les variations de vitesse du vent associées aux passages successifs de cyclones et d'anticyclones ;
- le pic relatif à une période d'environ 12 heures traduit l'augmentation de la vitesse du vent dans la journée et sa diminution durant la nuit. C'est la zone qui contient les structures de turbulence générées par convection thermique et dont les dimensions spatio-temporelles sont relatives à la partie moyenne de l'indice de réfraction ;
- le pic relatif à une période comprise entre 10 s et 10 min traduit la turbulence de petite échelle. Cette zone agit sur la diffusion de l'énergie sonore à travers la partie stochastique de l'indice de réfraction. L'énergie issue des grandes structures se transmet à des structures plus petites.

(b) *Fluctuations de l'indice de réfraction*

En tant que variable aléatoire, μ est caractérisée par sa fonction d'autocorrélation spatiale B :

$$B(\vec{r}_1, \vec{r}_2) = \overline{\mu(\vec{r}_1)\mu(\vec{r}_2)}, \quad (1.16)$$

où $\overline{\mu(\vec{r})}$ représente la moyenne de la fonction aléatoire $\mu(\vec{r})$. Si la turbulence est homogène, cette fonction reste inchangée par translation et ne dépend alors que de la séparation vectorielle entre les deux points de mesure. Le spectre en nombre d'onde correspond à la transformée de Fourier spatiale de la fonction de covariance qui dans un cas unidimensionnel s'écrit :

$$V(\kappa) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-j\kappa r) B(r) dr. \quad (1.17)$$

Les principaux spectres utilisés en micrométéorologie sont les spectres de Gauss, de Kolmogorov et de von Kármán. Pour un spectre gaussien (expression la plus simple) et pour une atmosphère présentant à la fois des fluctuations de température et de vent, $B(r)$ s'écrit :

$$B(r) = \langle \mu^2 \rangle \exp\left(\frac{-r^2}{l^2}\right), \quad (1.18)$$

avec l la longueur de corrélation et $\langle \mu^2 \rangle$ l'indice de turbulence. $\langle \mu^2 \rangle$ peut s'exprimer en fonction des composantes parallèle σ_u^2 et perpendiculaire σ_v^2 de la variance du vecteur vent, de la variance de la température σ_T^2 .

L'hypothèse de la « turbulence gelée » de Taylor permet de considérer que pendant la traversée de l'onde sonore, les paramètres régissant la turbulence sont invariants ce qui permet ainsi de caractériser le phénomène à un instant donné et d'exploiter les variances et covariances temporelles des grandeurs micrométéorologiques. Lorsque la propagation acoustique est considérée comme quasi-horizontale, la variance temporelle de $\mu(\mathbf{r})$ peut s'exprimer à partir des fluctuations horizontales du vent, soit les composantes parallèle σ_u et perpendiculaire σ_v de la variance du vecteur vent et l'angle entre le vent horizontal moyen et la direction de propagation [134] :

$$\langle \mu^2 \rangle \approx \frac{\sigma_u^2}{c_0^2} \cos^2 \vartheta + \frac{\sigma_v^2}{c_0^2} \sin^2 \vartheta + \frac{\sigma_u \sigma_T}{c_0 T_0} \cos \vartheta + \frac{\sigma_T^2}{4T_0^2}. \quad (1.19)$$

Les échelles caractéristiques de la turbulence U^* et T^* obtenues à partir de la théorie de similitude de Monin-Obukhov [88] permettent également d'obtenir une expression approchée de l'indice de turbulence [133] :

$$\langle \mu^2 \rangle \approx \frac{5U^{*2}}{c_0^2} + \frac{2,5U^*T^*}{c_0 T_0} \cos \vartheta + \frac{T^{*2}}{T_0}; \quad (1.20)$$

où T^* et U^* représentent respectivement les parties stochastiques de la température T et de la composante longitudinale du vent u , projetée selon la direction de propagation. T_0 et c_0 sont les valeurs de température et de célérité du son de référence.

(c) *Effets de la turbulence*

Les effets de la turbulence induisent des fluctuations aléatoires de la phase et de l'amplitude de l'onde sonore ainsi qu'une déformation des fronts d'onde. La turbulence présente deux effets majeurs sur les phénomènes de propagation acoustique :

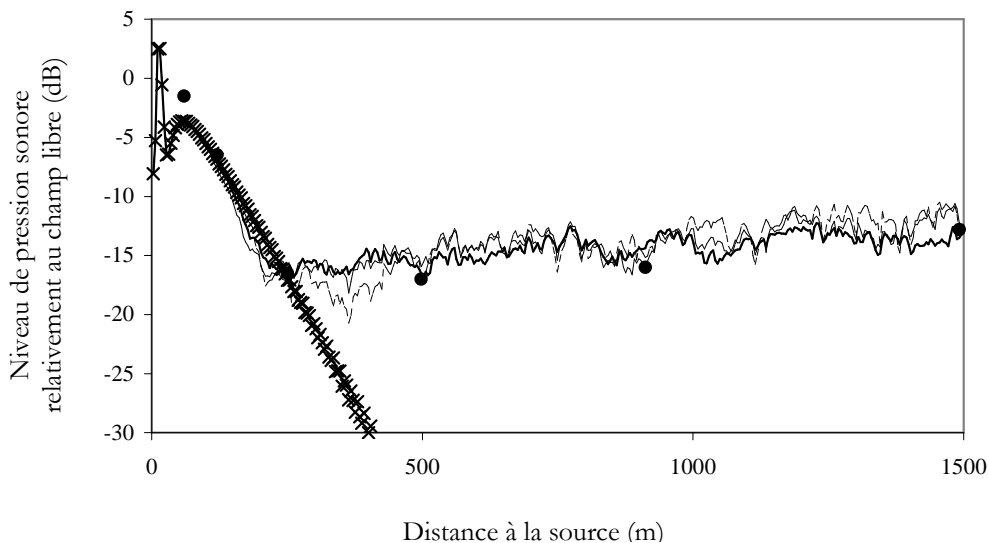


FIG. 1.14 – Effet de la turbulence atmosphérique - calculs déterministes sans turbulence (\times) et calculs avec turbulence (-), en comparaison avec les mesures expérimentales (\bullet) issues de Wiener et Keast [129]. $f = 424$ Hz, hauteur de source à 1,2 m, hauteur de récepteur 0,6 m, $R_s = 200$ kN.s.m $^{-4}$, $a_{log} = 0,5$ m.s $^{-1}$.

- la pénétration du son dans les zones d'ombre par diffraction des ondes acoustiques sur les structures turbulentes [129] ;
- la décorrélation des signaux acoustiques *i.e.* la perte de cohérence existant entre l'onde directe et l'onde réfléchi par le sol (réduction de l'importance des extrema des interférences) [37, 38, 134].

La figure 1.14 montre l'effet de pénétration du son dans une zone d'ombre par un calcul qui prend en compte la turbulence [27]. Dans des conditions défavorables à la propagation, le niveau de pression relatif au champ libre tend à se stabiliser entre -15 et -20 dB dans la zone d'ombre à environ 250 m de la source. Sans prise en compte de la turbulence, les niveaux de pression continuent à décroître avec la distance.

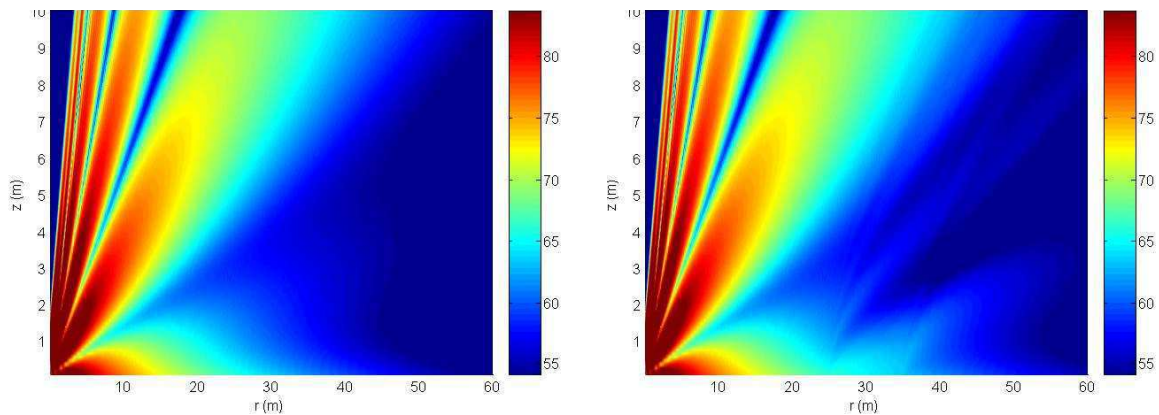
La turbulence rentre dans la classe des phénomènes chaotiques. Le caractère « imprévisible » du phénomène a d'ailleurs été montré par Lorenz dans les années soixante [86]. Les lois physiques régissant le phénomène sont bien connues mais la caractérisation du phénomène ne peut se faire qu'en se basant sur des approches statistiques. Un indicateur statistique peut être calculé mais il est difficile de prévoir avec précision les effets de la turbulence. La forte variabilité du phénomène dans le temps et dans l'espace induit par suite une part de variabilité (sur des échelles variant de quelques dixièmes de seconde à la minute) sur les niveaux sonores.

1.3.5 Effets mixtes dans des situations plus complexes

1.3.5.1 Diffraction par rupture d'impédance sur le sol

Lorsque des changements soudains de propriétés du sol interviennent sur le chemin de propagation, la discontinuité d'impédance induit un champ diffracté qui vient modifier la structure du champ acoustique. Cet effet est assimilable à la diffraction par un obstacle (une arête vive) [100, 101]. La diffraction consiste en une ré-émission de l'énergie sonore lorsque l'onde atteint un obstacle dont les dimensions sont de l'ordre de la longueur d'onde. Ce phénomène peut également apparaître lorsque le

sol est rugueux. Un exemple de diffraction par une discontinuité d'impédance est donné figure 1.15.



(a) 500 Hz - Sol Herbeux (300 kN.s.m^{-4}) sur tout le chemin de propagation

(b) 500 Hz - Sol Herbeux (300 kN.s.m^{-4}) - Présence d'un cours d'eau à une distance comprise entre 25 m et 35 m de la source (simulé par une résistance spécifique au passage de l'air de $30000 \text{ kN.s.m}^{-4}$)

FIG. 1.15 – Résultat d'une résolution numérique de l'équation parabolique avec et sans discontinuité d'impédance.

Au dessus d'un sol homogène, le front d'onde se propage sans être perturbé alors que dans le cas d'un sol hétérogène, on remarque un comportement similaire à la présence de sources secondaires au niveau des discontinuités qui ré-émettent l'énergie de l'onde incidente dans toutes les directions. Aux basses fréquences, la rugosité de la surface est invisible pour l'onde acoustique alors que pour des plus hautes fréquences celle-ci doit être prise en compte. Des études ont également été menée en prenant en compte l'effet d'une atmosphère réfringente [26, 58].

1.3.5.2 Effets des frontières de forme complexe

La forme de la surface au dessus de laquelle la propagation sonore est étudiée engendre aussi des phénomènes de diffraction ou de diffusion (propriété d'un obstacle à renvoyer l'énergie acoustique dans plusieurs directions). Le comportement des ondes sonores dépend donc fortement de la topographie. De plus, le type de terrain rencontré au cours de la propagation peut modifier localement de façon plus ou moins importante les profils micrométéorologiques (section 1.3.4.2). On trouve parfois dans la littérature une analogie entre les effets dus à la topographie et ceux dus à la réfraction par l'atmosphère [26]. On reprend ci-après le descriptif de l'influence des frontières « types » rencontrées en propagation extérieure.

(a) Influence des zones urbanisées

En zones urbanisées, le sol est souvent très réfléchissant. Réflexions multiples et diffuses, diffraction de l'onde par différents obstacles sont autant d'effets complexes ne permettant pas l'étude de la propagation acoustique avec la même approche qu'en espace ouvert ou semi-ouvert. La présence d'immeubles et de divers bâtiments réduit significativement la vitesse moyenne de vent et les comportements turbulents augmentent du fait de ces obstacles. La diffraction est relativement importante aux basses fréquences (100 Hz) et la diffusion domine aux plus hautes fréquences. Un effet thermique est également observé du fait de l'inertie des façades des bâtiments. En effet, dans la journée, les bâtiments vont emmagasiner de l'énergie radiative solaire ce qui aura pour effet de diminuer l'humidité

et d'entraîner de fortes circulations convectives. En période nocturne, cette énergie va être ré-émise ce qui va entraîner de fortes différences de température entre les zones urbaines et rurales.

(b) *Influence de la végétation*

Dans le cas d'une zone végétale, du fait de la porosité du sol, les moyennes fréquences sont fortement atténuées. En hautes fréquences, le phénomène de diffusion atténue sensiblement les niveaux sonores notamment lorsque les longueurs d'ondes sont de taille équivalente aux dimensions de certains « obstacles » comme par exemple les feuilles pour la propagation dans une zone boisée. Il s'avère que les premiers centimètres du sol sont beaucoup plus importants que les caractéristiques aériennes [39]. Dans les zones où la densité d'arbres est importante, une atténuation de l'amplitude des ondes sonores est observée. Cette atténuation dépend du diamètre des troncs, de la densité des arbres, de l'importance du feuillage, du type d'arbres ainsi que de la modification du type de sol [63]. Lorsque l'atmosphère est réfringente, les gradients verticaux de vent et de température, facteurs majeurs (section 1.3.4.2) dans la propagation au voisinage du sol [15] sont fortement influencés par la présence d'arbres.

Dans le cas de forêts dites fermées, les gradients verticaux de température sont en valeur absolue plus faibles qu'à l'extérieur (air plus frais en journée dans la zone boisée et plus chaud en période nocturne); l'humidité relative augmente du fait de la respiration des végétaux; la vitesse du vent est quasiment nulle sous les deux tiers de la hauteur de la canopée et augmente au dessus de la cime des arbres; l'effet turbulent diminue proportionnellement avec la densité du feuillage.

(c) *Influence du relief*

La surface du sol n'est jamais parfaitement plane et présente des discontinuités. Ces discontinuités influencent de manière notable le déplacement et le comportement des masses d'air. A l'encontre d'un obstacle, leur mouvement est désordonné, des tourbillons se forment et les profils verticaux de célérité sont modifiés [113]. Si l'on considère une colline, une zone d'ombre va se former car celle-ci va constituer une barrière naturelle pour les ondes sonores; s'ajoute à la formation d'une zone d'ombre le phénomène de réfraction. Les niveaux sonores dans la zone d'ombre vont alors dépendre de cet effet de réfraction des conditions micrométéorologiques *i.e.* de la réfringence de l'atmosphère (section 1.3.4.2).

Une autre façon de représenter le phénomène est de ne plus considérer un sol « courbé » et une atmosphère homogène mais un sol plan et une atmosphère réfringente. Une surface convexe sera modélisée par des conditions de propagation favorables [26] et une surface concave par des conditions de propagation défavorables.

1.4 BILAN

La propagation acoustique en milieu ouvert combine un certain nombre de phénomènes complexes. Les éléments majeurs à prendre en compte sont :

- la caractérisation de la source : une source réelle telle qu'une source de bruit routier fluctue considérablement dans le temps;
- les frontières du milieu et les caractéristiques acoustiques de sol : l'impédance de sol varie selon la nature du sol. L'impédance d'un sol herbeux évolue avec les événements météorologiques tels que les précipitations et l'ensoleillement. Une relation forte existe de plus entre les frontières du milieu et les phénomènes micrométéorologiques influents;
- l'indice de réfraction du milieu : la partie moyenne de l'indice de réfraction prend en compte les champs moyens de vent et de température sur des périodes d'une dizaine de minutes à une heure. La partie fluctuante de la réfraction se traduit grâce au paramètre de turbulence dont les effets sont traduits à l'aide d'indices statistiques. Ces effets sont de l'ordre de la seconde et sont significatifs sur les phénomènes de propagation acoustique en milieu ouvert.

Ces éléments fluctuent aussi bien dans le domaine spatial que temporel. Il est donc complexe d'accéder à une information précise par la mesure. A la variabilité spatio-temporelle viennent s'ajouter l'incertitude épistémique, l'incertitude de mesure, etc.

2 MODÉLISATION DU SYSTÈME

« Tous les modèles sont faux, certains sont utiles. »
Georges BOX

2.1 INTRODUCTION

Il existe différents types de modèles plus ou moins évolués qui peuvent être de nature déterministe ou de nature stochastique ; les approches stochastiques permettant d'accéder à une statistique des résultats. Tous restent cependant une représentation partielle de la « réalité » physique. Après une partie consacrée aux notions d'incertitude de modèle, on présentera différents types de modèles de propagation plus ou moins complexes (sans et avec prise en compte des conditions micrométéorologiques) en mettant en exergue leurs intérêts et leurs limites.

2.2 INCERTITUDE DE MODÈLE

Malgré des avancées toujours plus importantes dans le développement d'outils de modélisation, du fait notamment des moyens informatiques toujours plus performants, des incertitudes liés à la modélisation persistent. Différents aspects de ces incertitudes vont être abordés avec notamment l'identification des sources d'incertitudes. La construction ou l'élaboration de tout modèle est sujette à deux sources d'incertitudes : l'*incertitude épistémique*, due notamment à l'impossibilité de représenter parfaitement le problème étudié et l'*incertitude aléatoire*, due au fait que certaines variables ou paramètres supposés déterministes du modèle sont en fait des estimations de moments de variables aléatoires.

2.2.1 Incertitude épistémique

Le passage du phénomène réel au modèle engendre une erreur nommée *incertitude épistémique*. Chaque étape de construction d'un modèle donne lieu à de nouvelles incertitudes ; ces étapes sont résumées par la figure 1.16.

- *Du phénomène physique réel au modèle théorique.* Le phénomène considéré ne peut jamais être parfaitement décrit. Cette mauvaise compréhension génère une incertitude dans la définition du modèle théorique. Manque de connaissance, complexité du phénomène évoluant dans le temps, mauvaise adéquation entre le phénomène à modéliser et la physique utilisée sont autant d'éléments qui contribuent à l'incertitude épistémique ;
- *Du modèle théorique au modèle mathématique.* Le phénomène physique ne peut pas être considéré dans son intégralité et des approximations mathématiques (hypothèses trop fortes, conditions initiales) sont souvent requises ;
- *Du modèle mathématique au modèle numérique.* L'accès à une solution numérique requiert une discrétisation du modèle mathématique et est sujette à la précision de l'ordinateur utilisé. Ces approximations constituent une source d'incertitude supplémentaire.

Cette incertitude peut notamment être réduite à différents niveaux de l'élaboration du modèle par l'apport de données complémentaires. La compréhension humaine peut être améliorée grâce à de

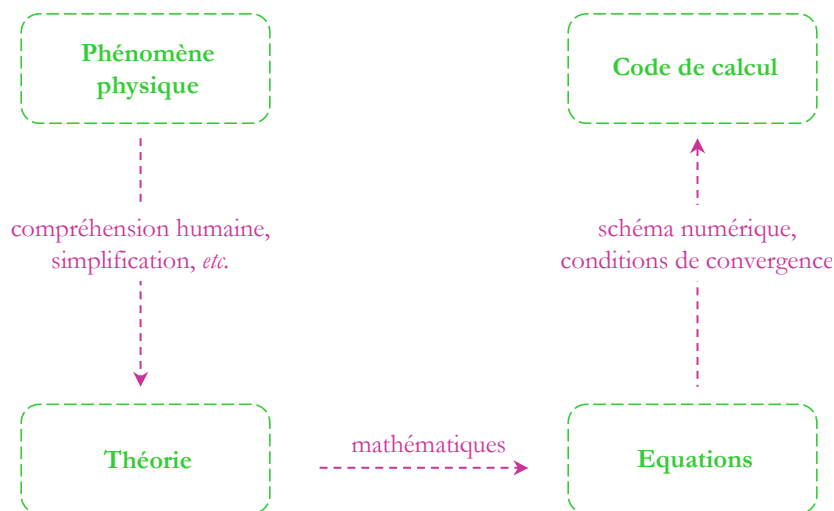


FIG. 1.16 – Différentes sources d’incertitudes épistémiques (flèches) liées à l’élaboration d’un modèle.

nouvelles données (expérimentales par exemple) et ces données peuvent également servir à valider le code de calcul développé.

2.2.2 Incertitude aléatoire

La seconde source d’incertitude est dite aléatoire. Elle est associée à l’estimation des paramètres du modèle. La variabilité des quantités physiques mesurées permet notamment d’expliquer pourquoi des résultats différents peuvent être obtenus en répétant une expérience plusieurs fois avec des conditions apparemment identiques. A titre d’illustration, d’un point de vue acoustique, on peut notamment relever le cas de la mesure de la puissance d’une source étalon qui présente un certain écart-type lié directement à la variabilité du comportement de la source (influence thermique, fatigue mécanique). Ce type d’incertitude est inhérent à tout phénomène ou processus physique et il est impossible de la supprimer.

Malgré l’existence de ces incertitudes relatives aux modèles, nous n’avons finalement pas d’information qualitative sur celles-ci, ce qui devra être intégré dans la méthodologie statistique proposée (chapitre 2), notamment grâce à un processus de calibration des paramètres et de leurs propriétés statistiques (variance).

2.3 MODÈLES UTILISÉS EN PROPAGATION ACOUSTIQUE EN MILIEU OUVERT

Différents modèles aux degrés de complexité plus ou moins élevés vont maintenant être présentés avec une mise en lumière sur la part d’incertitude qui leur est associée. Les modèles utilisés dans le domaine de l’acoustique sont nombreux et peuvent être classés en deux catégories selon leur niveau de sophistication : les modèles simplifiés (ou modèles d’ingénierie), généralement basés sur des méthodes analytiques et les modèles de référence basés sur des méthodes analytiques ou numériques.

La première catégorie de modèles donne lieu au développement de méthodes de calculs normalisées telles que la méthode européenne CNOSSOS [9, 72] et la méthode française NMPB08 [112]. Ces méthodes servent notamment à l’implémentation de modèles d’ingénierie communément utilisés par les bureaux d’étude.

La seconde catégorie de modèles (dits de référence), qu’ils soient numériques ou analytiques, peut

être classée en deux grandes classes : les modèles fréquentiels et les modèles temporels. Il existe un certain nombre de modèles de propagation temporels tels que la méthode des différences finies dans le domaine temporel (FDTD) appliquée aux équations d'Euler linéarisées [24], la méthode de lancer de particules (SPSS) [17], la méthode TLM (Transmission Line Matrix) [62]. Bien qu'ils constituent une nouvelle génération de modèles encourageants, la théorie de ces modèles ne sera pas développée dans ce mémoire.

Les modèles fréquentiels les plus utilisés sont :

- analytiques : modèle de Rudnick [45, 104] ;
- numériques : méthode des éléments de frontière (BEM - Boundary Element Method) [82], méthodes de résolution de l'équation parabolique (PE - Parabolic Equation).

Le modèle analytique de Rudnick et le modèle numérique basé sur la résolution de l'équation parabolique sont présentés ci-après.

2.3.1 Propagation acoustique au-dessus d'un sol plan homogène sans prise en compte des effets micrométéorologiques

Si l'on considère un milieu homogène et isotrope (soit $\frac{\partial c}{\partial z} = 0$), la formulation de Weyl-van der Pol permet de résoudre le problème de la propagation sur un terrain homogène [96]. Cette approche intègre trois ondes constitutives du champ sonore tel que :

$$p_t = p_d + R_p p_r + (1 - R_p) F(w) p_r, \quad (1.21)$$

où p_t , p_d et p_r correspondent respectivement aux champs transmis, direct et réfléchi, le coefficient R_p correspond au coefficient de réflexion en onde plane et $F(w)$ est une fonction appelée facteur limite de perte (cf. section 1.3.3). La figure 1.17 illustre de manière schématique la propagation au-dessus d'un sol plan à impédance constante.

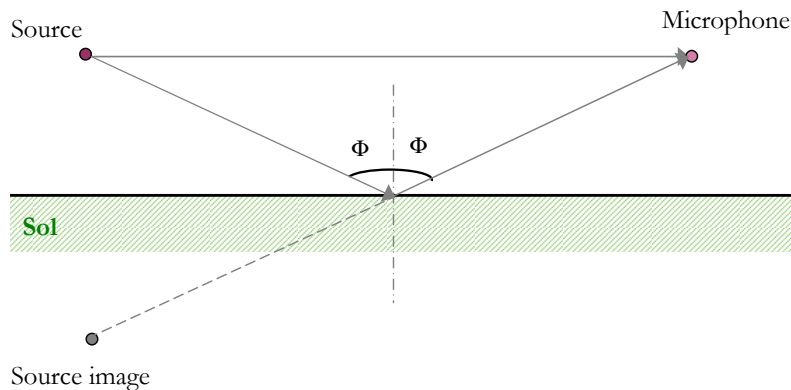


FIG. 1.17 – Constitution du champ acoustique par la théorie des rayons.

Ce modèle analytique simple peut être complété par des méthodes de prise en compte de ruptures d'impédance basées sur la formulation de Huyghens-Kirchoff [100], des effets météorologiques [26, 38] ainsi que des effets de diffusion [78].

2.3.2 Propagation acoustique avec prise en compte des effets micrométéorologiques

L'approximation parabolique a initialement été développée dans le domaine de l'électromagnétisme puis a été utilisée en acoustique sous-marine avant d'être adaptée à la propagation acoustique en milieu ouvert. Cette adaptation a notamment été initiée par Gilbert et White [60]. Depuis, la communauté

s'est intéressée au développement de cette méthode [55, 83]. L'équation parabolique permet de décrire la propagation du son en milieu hétérogène. L'impédance du sol et le profil de célérité peuvent varier dans l'espace sans augmenter sensiblement les temps de calcul. L'équation parabolique est fondée sur une approximation de l'équation d'onde elliptique de Helmholtz pour la pression acoustique p . Soit en coordonnées cylindriques dans un cas axisymétrique :

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2} + k(r, z)^2 \right) p(r, z) = 0, \quad (1.22)$$

avec r la distance horizontale, z la hauteur, $k(r, z) = \frac{\omega}{c(r, z)}$ le nombre d'onde effectif, $c(r, z)$ la célérité effective du son et ω la fréquence angulaire. La forme parabolique de l'équation est obtenue en introduisant une variable de la forme :

$$U(r, z) = \sqrt{r} p(r, z). \quad (1.23)$$

Ainsi à partir des équations 1.22 et 1.23, on obtient :

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{4r^2} + \frac{\partial^2}{\partial z^2} + k(r, z)^2 \right) U(r, z) = 0. \quad (1.24)$$

L'approximation de champ lointain se traduit par $r \gg k^{-1}$ ce qui permet de négliger le terme $\frac{1}{4r^2}$ et d'aboutir ainsi à l'équation suivante :

$$\left(\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial z^2} + k^2 \right) U(r, z) = 0. \quad (1.25)$$

L'indice de réfraction étant supposé faiblement variable avec la distance, les variations de k en fonction de r deviennent alors négligeables sur un pas de calcul. k est alors uniquement fonction de z (atmosphère stratifiée). Si l'on pose l'opérateur $S(z) = \frac{\partial^2}{\partial z^2} + k(z)^2$, l'équation 1.25 peut se réécrire sous la forme :

$$\left(\frac{\partial^2}{\partial r^2} + S(z) \right) U(r, z) = 0. \quad (1.26)$$

Par factorisation, cette équation fait apparaître deux termes : l'un correspondant à une onde se propageant dans le sens $r > 0$ (*onde propagative*) et l'autre à une onde se propageant en sens inverse (*onde rétrograde*). En négligeant l'onde rétropropagée, l'équation 1.26 devient :

$$\left(\frac{\partial}{\partial r} - j\sqrt{S(z)} \right) U(r, z) = 0. \quad (1.27)$$

L'implémentation numérique de cette équation nécessite une *approximation de l'opérateur de propagation* $\sqrt{S(z)}$. De plus, l'amélioration de la convergence numérique implique une réécriture du champ U ; celui-ci est remplacé par $\phi(r, z) = U(r, z) e^{-jk_0 r}$, où $e^{jk_0 r}$ est une porteuse variant selon r à la fréquence spatiale k_0 calculée à partir de la célérité effective $c(z=0)$. L'équation 1.27 devient :

$$\frac{\partial \phi(r, z)}{\partial r} = j(\sqrt{S} - k_0) \phi(r, z). \quad (1.28)$$

La résolution du problème s'effectue pas à pas de la source vers le récepteur, à partir d'un champ initial. La solution peut s'écrire sous la forme générale suivante [83] :

$$A\phi(r + dr) = B\phi(r) \quad (1.29)$$

où dr correspond à la résolution du maillage sur lequel le calcul est réalisé. Le champ au pas $r + dr$ s'obtient à partir du champ au pas r par inversion de la matrice A . Ainsi $\phi(r + dr) = A^{-1}B\phi(r)$. Le champ acoustique est ainsi calculé de façon itérative.

Il existe plusieurs techniques de résolution de l'équation parabolique standard. Les plus connues sont la méthode de Crank-Nicholson (CNPE), la méthode de Split-step Padé²[31], la méthode GT-PE (Generalised Terrain Parabolic Equation) [106], la méthode Split-step Fourier [31] et la méthode basée sur l'utilisation de la fonction de Green (GFPE). Elles se différencient par la façon d'exprimer l'opérateur $\sqrt{S(z)}$ et/ou par le schéma numérique de résolution de l'équation.

L'équation parabolique prend en compte les phénomènes micrométéorologiques se déroulant dans la C.L.A. (section 1.3.4.1) tels que les phénomènes de convection et les effets de réfraction (variations lentes de l'ordre de la dizaine de minutes) ainsi que les effets de diffusion dus aux variations plus rapides telles que les fluctuations turbulentes (variations de l'ordre de la seconde). La réfraction est prise en compte en considérant les profils de vent et de température. La turbulence atmosphérique, représentée par un champ aléatoire de fluctuations de l'indice de réfraction, est directement incluse dans le profil de célérité $c(z)$ grâce à la technique des modes de Fourier aléatoires [70].

Les modèles de propagation utilisent comme données d'entrée les paramètres effectifs des modèles d'impédance de sol (section 1.3.3.3) et d'autres paramètres non directement mesurables (paramètres micrométéorologiques). En plus de l'incertitude liée à l'élaboration du modèle, vont s'ajouter l'incertitude et la variabilité liées aux paramètres d'entrée.

2.4 BILAN

Les modèles vont ainsi être utilisés sans avoir accès à une valeur quantitative de leur précision ; seule la catégorie dans laquelle le modèle peut être placé donne accès à une précision relative et qualitative de celui-ci. Les modèles seront finalement considérés comme des boîtes noires et donc utilisés « en aveugle ». L'erreur du modèle sera ainsi répercutée sur les grandeurs d'intérêt en plus de la part d'incertitude et de variabilité associées aux paramètres d'entrée.

3 GRANDEURS D'INTÉRÊT

3.1 INTRODUCTION

L'indicateur le plus fréquemment utilisé dans les études d'impact acoustique est un niveau énergétique continu équivalent, calculé sur une période de temps représentative du phénomène sonore mis en évidence. Il est exprimé en décibels ou en décibels pondéré A ; ces indicateurs sont respectivement notés L_{eq} et L_{Aeq} . Cette approche permet d'obtenir un unique indicateur pour un phénomène par essence fluctuant. Les grandeurs observées sont imposées par le cadre réglementaire dont dépend chaque type de source. Dans le cas du bruit routier on parle de *valeur seuil* variant de 55 à 60 dB(A) selon la période de la journée et l'installation [1] alors que dans le cas du bruit industriel on raisonne en termes d'*émergence* par rapport à un niveau mesuré en l'absence de la source objet de l'étude. La limite d'émergence varie de 3 à 5 dB(A) selon la période de la journée [2]. Ces exigences réglementaires ne prennent cependant pas en compte l'incertitude et la variabilité associées à ces niveaux sonores ce qui est la principale difficulté dans l'appréciation des situations d'exposition au bruit.

Pour tenir compte de cette variabilité, les mesures et/ou prévisions sont généralement effectuées dans des conditions de propagation favorables à la propagation acoustique [131], ce qui ne témoigne pas

2. L'approximation de Padé autorise de grands angles de validité de l'équation parabolique (WAPE pour Wide Angle Parabolic Equation). C'est la méthode qui est implémentée dans le modèle que nous allons utiliser au chapitre 4.

de la variabilité réelle du phénomène mais permet seulement de fixer une valeur seuil de dépassement dans une situation d'exposition maximale. Pour expliquer la variabilité des niveaux sonores, ceux-ci doivent en effet être considérés à différentes échelles d'espace et de temps.

3.2 VARIABILITÉ TEMPORELLE DES NIVEAUX SONORES – DE LA TURBULENCE AU LONG TERME

3.2.1 Échelle de court terme

L'échelle de court terme est relative à des périodes de l'ordre de 10 minutes à 2 heures. Sur cette échelle, les profils verticaux de vitesse du son sont moyennés dans le temps et permettent alors de caractériser les conditions moyennes de propagation (section 1.3.4.2). On doit cependant garder à l'esprit que les phénomènes demeurent fluctuants à l'intérieur de cette échelle du fait de la turbulence atmosphérique (section 1.3.4.3). Notre étude s'intéressant spécifiquement à cette échelle temporelle, elle intégrera à la fois les effets de réfraction moyenne et les effets turbulents.

3.2.2 Échelle de moyen terme

Cette échelle correspond aux périodes de plusieurs heures. Les périodes jour/soir/nuit définies au niveau réglementaire [3] correspondent à cette échelle. Certains paramètres physiques peuvent être caractéristiques de phénomènes annuels ou saisonniers alors que d'autres paramètres sont sensibles à des micro-fluctuations temporelles. Une simple décomposition journalière n'est pas toujours représentative de l'ensemble des phénomènes car les profils verticaux de vent et de température varient largement dans le temps en raison des fluctuations aléatoires des conditions météorologiques.

3.2.3 Echelle de long terme

Un élément de référence est disponible pour caractériser une durée de « long terme » [136]. En effet, certaines études montrent que l'erreur maximale commise entre une durée de mesure de un an et une durée de mesure de 29 ans est inférieure au décibel pour une source ponctuelle. On peut donc considérer que l'année est bien représentative d'une situation de long terme. De plus pour une mesure de 6 mois entre le solstice d'été et celui d'hiver, l'erreur maximale sur la mesure est inférieure à 1,5 dB.

De récents travaux ont notamment permis de mettre en évidence l'influence de la saison sur la dispersion des niveaux $L_{Aeq}(6h-22h)$ [43]. Ce résultat est illustré par la figure 1.18.

On peut noter que cette dispersion est de l'ordre de 10 dB(A) en période estivale et de 15 dB(A) en période hivernale. A partir de données météorologiques acquises sur une période de 29 ans sur différents sites repérés en France, des calculs de propagation ont été réalisés. Une distribution des niveaux sonores (sur chaque situation météorologique) a ainsi été obtenue. Cette distribution rend compte également du pourcentage de présence d'un niveau sonore, du fait d'une situation météorologique à l'échelle de l'année. Ceci permet ensuite de calculer des niveaux sonores de long terme (sur 29 ans) notés L_{LT} [112]. Cette approche rend compte de la variabilité temporelle d'un niveau sonore (à l'échelle de la saison) à travers la variabilité temporelle des conditions micrométéorologiques et fournit des niveaux sonores de référence. Cette approche combine aussi bien des données de moyen terme et de court terme.

3.3 VARIABILITÉ SPATIALE DES NIVEAUX SONORES - DE L'ÉCHELLE LOCALE À L'ÉCHELLE RÉGIONALE

La variabilité spatiale des niveaux sonores est un phénomène connu. C'est d'ailleurs pour cette raison que l'on cherche à établir des cartographies du bruit. Les dépendances spatiales des niveaux sonores sont liées à la nature locale du sol, au couvert végétal, aux obstacles, etc., entre la source et

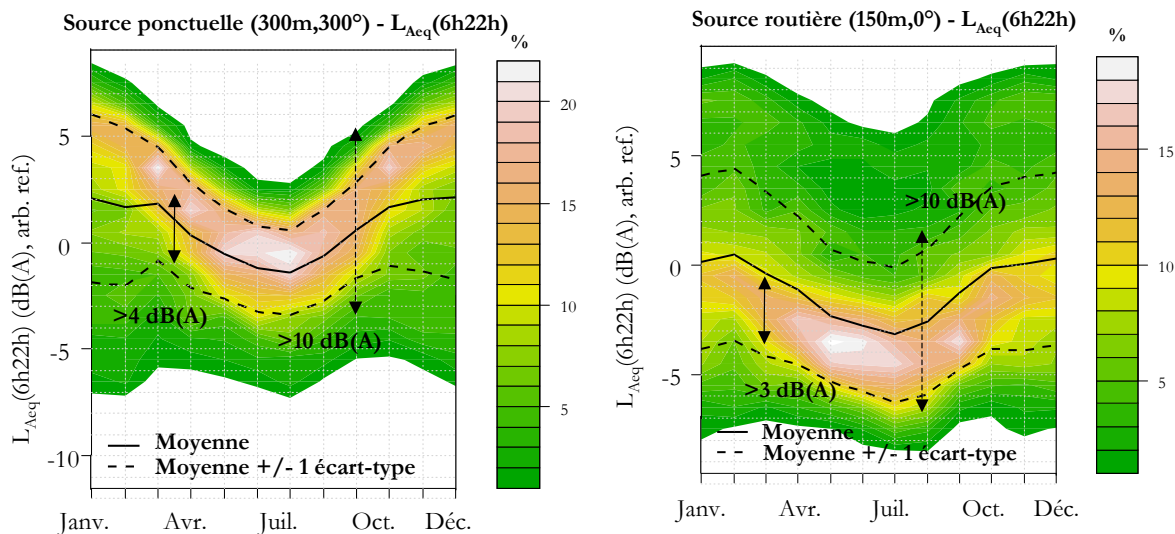


FIG. 1.18 – Variation des $L_{Aeq}(6h - 22h)$ au cours d'une année dans le cas d'une source ponctuelle et d'une source routière [43].

le récepteur [8]. Des méthodes géostatistiques ont été appliquées pour interpoler ce type de données expérimentales, en considérant une cohérence spatiale [20, 21]. Ce paramètre peut être très fluctuant sur des échelles spatiales de l'ordre du mètre. Cette variabilité est également liée aux échelles spatiales des phénomènes météorologiques influents :

- à l'échelle régionale : les phénomènes météorologiques synoptiques (déplacement des masses d'air sur des échelles de l'ordre de plusieurs centaines de kilomètres) ;
- à l'échelle locale : les phénomènes micrométéorologiques. Ces paramètres micrométéorologiques doivent être mesurés sur le site d'étude (entre la source et le récepteur) ; des mesures à quelques kilomètres du site ne permettent pas une caractérisation précise et optimale du site.

La figure 1.19 illustre la variabilité spatiale des phénomènes : l'augmentation de la dispersion des niveaux sonores avec la distance. Cette représentation rend également compte de différences majeures en termes de dispersion selon les conditions de propagation.

3.4 INCERTITUDE SUR LES NIVEAUX SONORES

3.4.1 Niveaux sonores mesurés

L'incertitude associée aux niveaux sonores mesurés est notamment liée aux erreurs de mesures : dérives de capteurs, une mauvaise calibration du microphone, le type de matériel utilisé, etc. Les conditions de mesures contribuent également à cette incertitude : différence inter-opérateurs, répétabilité, reproductibilité. Ces différents aspects sont cependant irréductibles et constituent une part d'incertitude aléatoire.

3.4.2 Niveaux sonores calculés

En ce qui concerne les niveaux sonores calculés, l'incertitude intègre l'ensemble des incertitudes abordées dans les sections précédentes : l'incertitude et la variabilité sur les paramètres d'entrée vont se propager dans le modèle de propagation. L'incertitude associée au(x) modèle(s) va alors se combiner aux incertitudes liées aux paramètres d'entrée. Cet ensemble d'incertitudes se reporte alors sur les paramètres d'intérêt : les niveaux sonores. Cette incertitude combinée est cependant difficilement

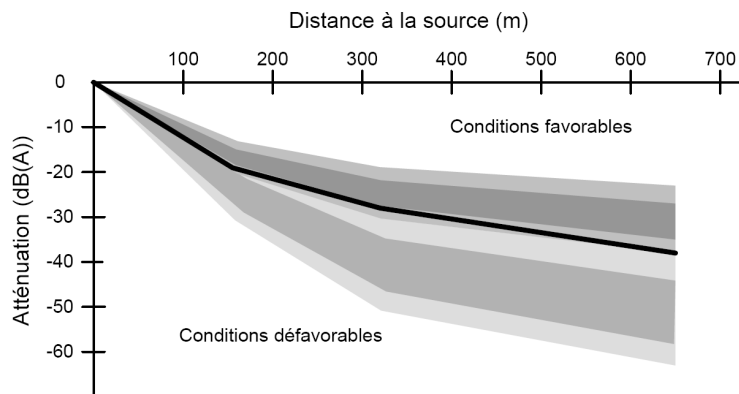


FIG. 1.19 – Atténuation et dispersion des niveaux sonores en conditions favorables et en conditions défavorables, comparés au niveau sonore théorique en condition homogène (trait plein) en fonction de la distance à la source (source située à 6 m au-dessus du sol et récepteurs à 1,5 m de hauteur) [137].

accessible, chaque source d'incertitudes n'étant pas quantifiée. Des méthodes statistiques basées sur la décomposition en modes propres de calculs de référence (POD - Proper Orthogonal Decomposition) [95], sur la théorie de la logique floue et l'analyse multidimensionnelle [135] ont été proposées. Ces méthodes permettent de caractériser qualitativement l'incertitude sur les niveaux sonores mais le problème de quantification de cette incertitude reste entier.

3.5 BILAN

A l'incertitude sur les niveaux sonores vient s'ajouter l'aspect variabilité des niveaux sonores. Cette combinaison supplémentaire rend d'autant plus complexe l'accès à cette incertitude combinée. A l'heure actuelle il existe très peu de travaux abordant aussi bien les problématiques de variabilité spatio-temporelle que d'incertitudes. Comment quantifier cette incertitude sur les niveaux sonores ne connaissant pas quantitativement les incertitudes sur les paramètres d'entrée et le modèle ? Quelle est la contribution de chacun des paramètres influents ? C'est à ces questions que vont tenter de répondre les prochains chapitres de ce mémoire.

4 CONCLUSION DU CHAPITRE

Cette revue bibliographique permet de prendre conscience de la problématique d'incertitude dans le domaine de la propagation acoustique en milieu ouvert. Incertitude et variabilité se retrouvent à chaque étape : de la description physique des phénomènes et des paramètres d'entrée à l'observation des grandeurs d'intérêt, en passant par le modèle décrivant le phénomène. L'incertitude finale *i.e.* sur les niveaux sonores est donc une combinaison complexe des différents phénomènes. Cette incertitude dépend de plus du milieu, des échelles spatiales et temporelles considérées. L'« opérateur » contribue de manière non-négligeable à cette incertitude par exemple en choisissant des conditions initiales de calcul, la mise en place d'un protocole expérimental, le modèle, etc. Tout l'enjeu réside alors dans la caractérisation et la quantification de cette incertitude combinée.

2

PROCESSUS DE QUANTIFICATION D'INCERTITUDES : MÉTHODOLOGIE

PLAN DU CHAPITRE

1	SPÉCIFICATION DU SYSTÈME	41
1.1	Introduction	41
1.2	Données mesurées	41
1.3	Choix du modèle	42
1.4	Bilan	48
2	DÉFINITION DES OBSERVABLES	48
2.1	Observables physiques	48
2.2	Observables définies par des outils statistiques	48
3	CALIBRATION DU SYSTÈME	49
3.1	Introduction	49
3.2	Approches statistiques	49
3.3	Définition du modèle statistique de Calibration sous Incertitude	51
3.4	Bilan	52
4	VALIDATION DU SYSTÈME	53
4.1	Identifiabilité des paramètres du modèle	53
4.2	Critères statistiques	53
4.3	Vérification des hypothèses initiales	54
4.4	Bilan	54
5	PRÉVISION	55
5.1	Variance d'estimation	55
5.2	Incertainude globale	56
6	CONCLUSION DU CHAPITRE	57

Le processus de quantification d'incertitudes mis en place est basé sur un processus de calibration entre un jeu de données expérimentales et un modèle de prévision acoustique. Une telle méthode permet de renforcer le contenu prédictif du modèle et d'accéder à une information quantitative relative aux données expérimentales et au modèle. La méthodologie développée pour cette problématique de quantification d'incertitudes s'organise selon les étapes suivantes (figure 2.1) :

1. la spécification du système ;
2. la définition des observables ;
3. la calibration du système ;
4. la validation du système ;
5. et la prévision.

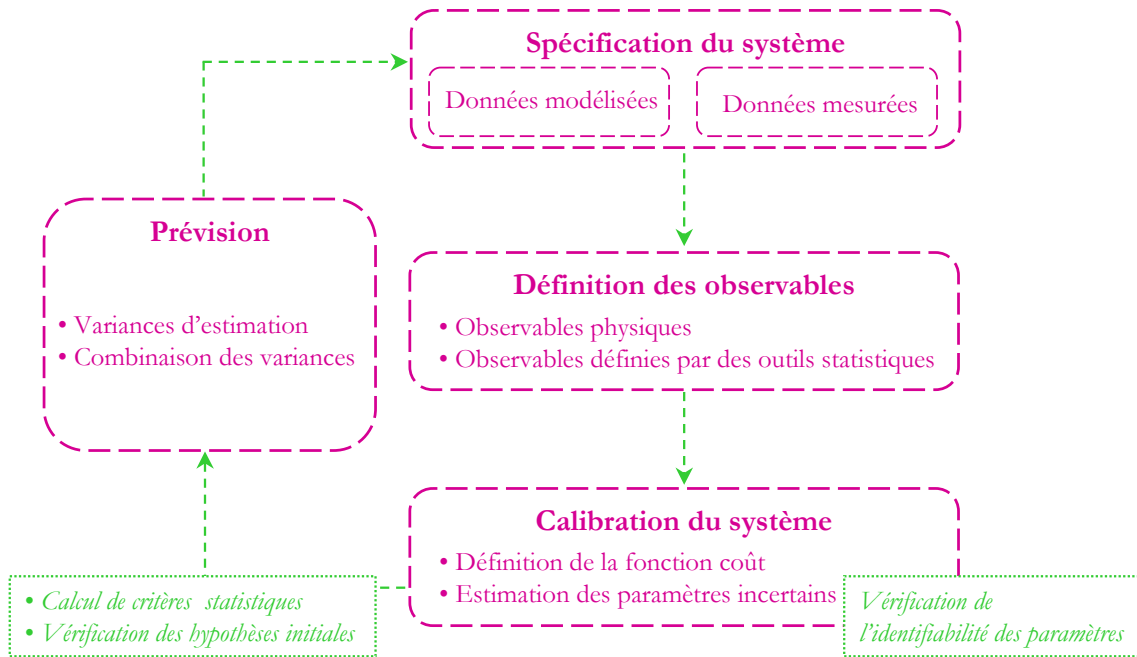


FIG. 2.1 – Représentation schématique du processus d'estimation d'incertitudes - les encadrés verts correspondent à la phase de validation du système.

Une telle méthodologie est mise en place pour répondre à différents objectifs : rendre compte de façon qualitative des sources d'incertitude et de variabilité prépondérantes, estimer l'incertitude associée aux paramètres d'entrée du système et quantifier l'incertitude globale associée aux grandeurs d'intérêt.

Cette démarche a un aspect potentiellement itératif : après avoir mené une étude complète (de la spécification du système à la prévision), nous pouvons être contraints de recommencer le processus en spécifiant un nouveau système plus robuste. Chaque étape du processus proposé est ainsi décrite dans ce chapitre.

1 SPÉCIFICATION DU SYSTÈME

1.1 INTRODUCTION

La méthodologie que nous proposons intègre à la fois des données expérimentales et des données numériques. Elles vont être confrontées et combinées, afin notamment de renforcer le contenu prédictif du système. La phase de spécification des données est basée sur la sélection de données expérimentales et sur la sélection d'un modèle de prévision adapté. Ce choix doit être en adéquation avec les objectifs de notre étude.

1.2 DONNÉES MESURÉES

Les données expérimentales que nous avons à disposition sont représentatives d'un panel de configurations (sol plan ou topographie complexe, source ponctuelle ou source linéique), chacune rendant compte d'un ensemble de situations atmosphériques, micrométéorologiques, etc. Ces bases de données sont présentées en Annexe A. Afin de répondre aux objectifs et enjeux de l'étude, une base de données adéquate peut être sélectionnée et des situations physiques spécifiques doivent être mises en évidence. Pour étudier une situation donnée, il est intéressant de disposer d'un nombre important de mesures caractéristiques de cette situation. Ceci nous conduit finalement à mener une réflexion sur la création de groupes de données caractéristiques.

1.2.1 Création de groupes de données expérimentales

Si l'on cherche à caractériser une campagne expérimentale dans son ensemble, le groupe de données avec lequel nous travaillerons sera constitué de l'intégralité de ces données. Cependant, dans la plupart des cas, les données expérimentales ne seront pas utilisées massivement et un travail plus approfondi de création de groupes de données doit être mené. La constitution de groupes de données présente un grand intérêt pour caractériser la variabilité des données sur une période considérée et l'incertitude associées à des conditions physiques spécifiques (par exemple pour des conditions de propagation données). Si l'on souhaite étudier la variabilité saisonnière, journalière ou horaire des niveaux sonores, le découpage dans le temps de la base de données est particulièrement aisé. Cependant, lorsque l'on essaie de dégager des situations plus spécifiques, il est nécessaire d'avoir recours à des techniques de classification évoluées.

Toutefois, la construction de classes d'évènements physiques particuliers n'est pas toujours aisée. En météorologie, il existe différentes méthodes de classification basées sur les états de stabilité de la couche limite de surface (chapitre 1 section 1.3.4.1) définies par Pasquill [94], Turner [121], ou Johnson *et al.* [67]. En propagation acoustique en milieu extérieur, il existe aussi des classifications basées sur les conditions de propagation avec la grille $U_i T_i$ de la normalisation française [108], ainsi qu'une méthode hybride, la grille $W_i S_i$ issue de travaux européens dans le cadre des projets Harmonoise et Imagine [64]. Mais de telles classes peuvent être discutables : la limite entre deux classes n'est pas toujours nette, particulièrement lorsqu'une réflexion sur l'incertitude et la variabilité est menée. La problématique de classification des données étant complexe, la création de groupes dépendra essentiellement du jeu de données à disposition. Une méthode basée sur un découpage *a priori* des paramètres du système (paramètres géométriques, paramètres micrométéorologiques ou paramètres de sol ou bien une combinaison de paramètres) a par exemple été développée dans l'une de nos études. Cette méthode propre à l'étude menée est détaillée à la section 1.1.1 du chapitre 4.

1.2.2 Paramètres d'entrée du système

Aux groupes de données constitués correspondent respectivement des paramètres influents à savoir, les paramètres géométriques, les paramètres micrométéorologiques, les paramètres de sol, etc. Ces

paramètres peuvent être mesurés ou non et bien connus (ils sont notés d_j) ou incertains (ils sont notés x). Ces paramètres incertains vont être recalés lors du processus de calibration (section 3).

1.3 CHOIX DU MODÈLE

Le choix du modèle de propagation acoustique est un élément essentiel. Échantillonnage en temps et étendue spatiale du domaine sont des éléments déterminants. Selon les cas, un modèle numérique de référence ou un modèle simplifié (section 2.3) pourra être utilisé ; temps de calculs et sensibilité du modèle aux paramètres d'entrée seront alors différents. La problématique des temps de calculs liée à l'utilisation de modèles numériques peut être résolue en ayant recours à une représentation mathématique simplifiée du modèle appelée *métamodèle*.

Kleijnen et al. [75] ont défini un métamodèle comme étant une approximation de la vraie fonction entrée/sortie implicitement définie par un modèle de simulation donné. De toute évidence, un métamodèle est un modèle simplifié se substituant à l'exécution de codes de calculs nécessitant trop de ressources. La construction d'un métamodèle a pour but d'obtenir un modèle mathématique représentatif du modèle étudié en termes de qualité prévisionnelle et dont le temps de calcul pour évaluer une réponse est négligeable en regard du modèle initial. À noter que les termes *surface de réponse* [25] et *émulateur* [91] sont couramment rencontrés dans la littérature pour désigner un métamodèle.

D'un point de vue pratique, un métamodèle est construit et ajusté à partir de calculs effectués avec le modèle de prévision initial (pour différents jeux de paramètres). Le nombre de calculs nécessaires pour constituer un plan d'expériences numériques robuste dépend essentiellement de la complexité du code. La nature du métamodèle (section 1.3.2) dépend quant à elle des objectifs d'utilisation du métamodèle (section 1.3.1). Ajustement et validation du métamodèle (section 1.3.4) constituent une étape clé dans la mise en place d'un métamodèle. L'ensemble de cette démarche est décrite pas à pas ci-après.

1.3.1 Objectifs d'application du métamodèle

Le recours à un métamodèle répond à différents objectifs : la compréhension du modèle, la résolution de problèmes d'optimisation de la réponse, la prévision rapide de nouvelles réponses et l'aide dans la vérification et la validation du modèle de prévision. Ces objectifs sont classés selon le niveau d'exigence qu'ils requièrent pour définir le métamodèle.

1.3.1.1 compréhension du modèle

Un objectif peut être de comprendre le comportement d'un modèle, comprendre le comportement des données de sortie du modèle, effectuer une analyse de sensibilité. Différents degrés de compréhension du système peuvent donc être étudiés :

1. la tendance/l'évolution qualitative des grandeurs d'intérêt : la variation d'une grandeur d'intérêt est-elle proportionnelle à la variation des paramètres d'entrée ?
2. le contrôle : quels facteurs ont des effets importants ?
3. les effets principaux, les interactions, les effets quadratiques et d'autres effets d'ordre plus élevé : quels sont les effets d'un facteur spécifique et éventuellement d'une combinaison avec d'autres facteurs ?

1.3.1.2 Vérification et validation du modèle de prévision

Le métamodèle peut être développé pour aider à la vérification et la validation du modèle de prévision (de référence ou d'ingénierie). Une première méthode consiste à déterminer si les grandeurs d'intérêt du métamodèle répondent aux changements des paramètres d'entrée dans le sens attendu du fait de la connaissance qualitative *a priori* du système réel. Une seconde peut être utilisée lorsque le

système réel est facilement observable. Dans ce cas, des comparaisons sont faites entre les résultats du métamodèle et les observations du système réel pour de nombreuses conditions expérimentales dans le domaine d'applicabilité du modèle de prévision.

1.3.1.3 Optimisation

Le métamodèle peut être utilisé pour déterminer le jeu de paramètres d'entrée qui optimise une fonction objectif spécifique définie à partir des grandeurs d'intérêt du modèle. Pour une valeur de sortie donnée, le but est de trouver les paramètres d'entrée correspondants. Ce type d'application nous intéressera par la suite.

1.3.1.4 Prévision

Le métamodèle peut remplacer le modèle de simulation pour fournir un jeu de grandeurs d'intérêt. Le métamodèle est alors utilisé à la place du modèle de simulation, étant donné sa rapidité et sa facilité d'utilisation.

1.3.2 Choix du métamodèle

Le choix du métamodèle est un problème à part entière qui dépend du contexte et du modèle de prévision initial. Une première stratégie est de privilégier la simplicité en adéquation avec les objectifs de l'étude [66]. Il existe différents types de métamodèles :

- les modèles les plus populaires sont les *modèles de régression polynômiale* d'ordre peu élevé [89]. Ce modèle peut suffire à expliquer un phénomène « simple » et ne peut en aucun cas être utilisé dans un cas d'étude de phénomènes complexes. Une relation de type polynômial entre les paramètres d'entrée et les grandeurs d'intérêt peut être utilisée dans le cas où l'on cherche à comprendre le système réel (des tendances globales cherchent à être dégagées sans grande précision) ;
- les *modèles polynômiaux par morceaux* de type MARS (Multivariate Additive Regression Spline) ou PolyMARS (Polychotomous regression based on MARS) [53] sont une généralisation de la régression linéaire où les fonctions de base sont des splines. Ces méthodes sont relativement souples et l'estimation repose sur la construction de splines dans chaque direction ;
- les *modèles additifs* constitués d'une somme de fonctions relatives à chaque paramètre d'entrée permettent de représenter leur effet sur les grandeurs d'intérêt mais il deviennent peu fiables lorsque des interactions entre les paramètres existent [29] ;
- les *modèles de krigeage* qui sont des modèles d'interpolation inspirés des méthodes géostatistiques et menant à la construction d'un modèle probabiliste. La fonction est choisie en passant par une fonction de corrélation et des paramètres déterminés par les observations [73]. La méthode de krigeage est décrite en Annexe B dans le cadre de la géostatistique.

Cette liste n'est cependant pas exhaustive ; il existe en effet beaucoup d'autres modèles pour ajuster les réponses de codes de calculs (les SVM, les réseaux de neurones, les arbres de régression, etc.). Le métamodèle et le plan d'expériences associé seront donc choisis en fonction du modèle de prévision (complexité du phénomène modélisé, interaction entre les paramètres) mais aussi en fonction des objectifs d'utilisation du métamodèle.

1.3.3 Construction du plan d'expériences numériques

Le métamodèle se construit à partir d'un plan d'expériences numériques dont la construction dépend du domaine d'applicabilité de l'étude (domaine de variation des paramètres). Dans le cas où un grand nombre de paramètres d'entrée intervient, un *criblage* [74] peut être mis en place pour

détecter les facteurs les plus influents et leur domaine de variation. L'expérimentateur peut aussi procéder lui-même à cette sélection.

Les qualités essentielles d'un plan d'expériences sont sa *robustesse* (capacité à analyser différents modèles), son *efficacité*, la *répartition des points* (remplissage homogène de l'espace échantillonné) ainsi qu'un *coût de construction minimum*. Le plan d'expériences numériques doit en effet satisfaire toutes les classes de métamodèles. Ainsi, un nombre approprié d'expériences et de niveaux pour chaque paramètre d'entrée doivent être choisis afin d'assurer en particulier le recouvrement de l'espace. Les méthodes de génération des plans doivent être peu coûteuses en temps d'exécution et doivent aussi permettre d'obtenir des plans efficaces avec un faible nombre d'expériences.

Certains plans d'expériences *numériques* ont été initialement développés dans le cadre de la planification d'expériences *classique* ; ces plans mènent à une bonne qualité concernant le remplissage de l'espace. On peut distinguer les plans factoriels complets (choix de points sur une grille régulière dans le domaine de variation des paramètres), les plans factoriels fractionnaires (sélection de sous-ensembles de plans factoriels complets afin de réduire le nombre de simulations), les plans composites centrés (constitués de trois catégories d'essais telles que les essais du plan factoriel, des essais « en étoile », des essais au centre du domaine), les plans de Box-Behnken [84] (structures symétriques dans lesquelles chaque facteur prend trois niveaux : combinaison de plans factoriels à deux niveaux, avec des blocs incomplets équilibrés suivant un arrangement particulier avec plusieurs essais au centre du domaine expérimental), ou bien même les plans de Doehlert (ou réseaux uniformes) [84]. Il existe également des plans combinant deux approches par exemple les plans LHS maximin.

D'autres plans dits « marginaux » présentent de bonnes propriétés en termes de non-redondance et de non-alignement sur certains sous-espaces. Hypercubes latins [87] et tableaux orthogonaux [93] font notamment partie de cette catégorie de plans. Ces plans n'assurent cependant pas un remplissage optimal de l'espace. À l'heure actuelle, des études sont en cours sur l'utilisation de nouveaux plans d'expériences plus modulables basés sur des méthodes de Monte-Carlo par Chaînes de Markov. Cette technique consiste à remplir l'espace de façon aléatoire, en contrôlant la proximité entre les points [49, 50]. Des outils graphiques tels que le « radar d'uniformité » permettent notamment d'étudier l'uniformité des points sur les sous-espaces de dimension 2 et 3.

Le but n'étant pas de procéder à une description détaillée de l'ensemble des plans d'expériences existant, le lecteur pourra se référer à de nombreux ouvrages pour en savoir davantage sur le sujet [84]. Plans factoriels complets et plans de Doehlert seront utilisés dans notre étude (chapitre 4) et leurs propriétés sont davantage développées ci-après.

1.3.3.1 Plans factoriels complets

Les plans factoriels complets (à d niveaux) sont les plus simples et sont aussi très utiles car ils forment la base de tous les débuts d'étude. Les premiers résultats obtenus grâce à ces plans peuvent toujours être complétés par de nouvelles expériences permettant d'atteindre le degré de précision et d'information recherché. Un plan factoriel complet permet de couvrir de façon régulière l'espace des paramètres d'entrée (figure 2.2).

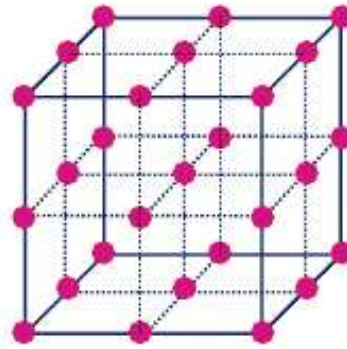


FIG. 2.2 – Plan factoriel complet à trois paramètres et trois niveaux.

Les plans factoriels complets deviennent difficiles à utiliser dès que le nombre de paramètres q (appelés facteurs dans le domaine de la planification d'expériences) dépasse la valeur de trois ou quatre. En effet, le nombre d'expériences n à réaliser est égal à d^q , ce qui conduit rapidement à un nombre important de calculs avec le modèle de prévision.

1.3.3.2 Plans de Doehlert

Les plans de Doehlert appartiennent à la famille des réseaux. Ils sont formés d'un essai au centre du domaine et de $q^2 + q$ essais répartis le plus uniformément possibles sur une sphère de rayon 1 (axes des paramètres normés). La méthode permettant de générer des points est itérative et considère un simplexe initial¹ dans le domaine d'applicabilité de l'étude. Des transformations géométriques (translations, rotations) sont appliquées à l'un des sommets de ce simplexe ; l'opération est renouvelée à plusieurs reprises afin d'obtenir une distribution de points particulière (figure 2.3).

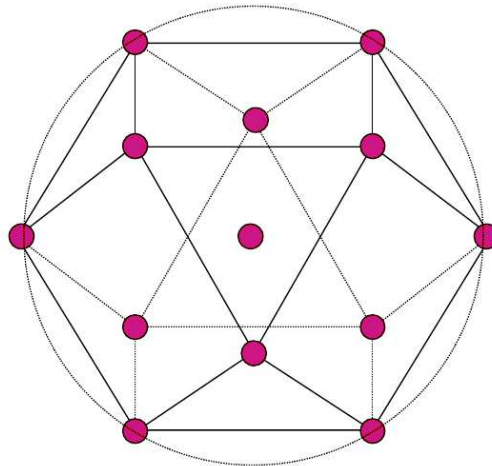


FIG. 2.3 – Plan de Doehlert à trois facteurs.

1. La notion de simplexe est la généralisation à une dimension quelconque de celle d'un triangle en dimension 2, ou celle d'un tétraèdre en dimension 3. Un simplexe de dimension n est un polytope de dimension n à $n + 1$ sommets

Par exemple pour $q = 3$, les coordonnées du plan de Doehlert sont synthétisées par le tableau 2.1.

indice de l'essai	X_1	X_2	X_3
1	0	0	0
2	-1	0	0
3	1	0	0
4	-1/2	$-\sqrt{3}/2$	0
5	1/2	$-\sqrt{3}/2$	0
6	-1/2	$\sqrt{3}/2$	0
7	1/2	$\sqrt{3}/2$	0
8	-1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
9	0	$1/\sqrt{3}$	$-\sqrt{2/3}$
10	1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
11	-1/2	$1/2\sqrt{3}$	$\sqrt{2/3}$
12	0	$-1/\sqrt{3}$	$\sqrt{2/3}$
13	1/2	$1/2\sqrt{3}$	$\sqrt{2/3}$

TAB. 2.1 – Coordonnées du plan de Doehlert à trois facteurs soit $n = 13$.

Les plans de Doehlert présentent un certain nombre de propriétés intéressantes :

- le faible nombre d'expériences à mener en grande dimension. Pour un nombre de facteurs donné q , le nombre minimal de points différents à calculer est $n = q^2 + q + 1$, ce qui est un avantage conséquent en grande dimension (vis-à-vis d'un plan factoriel complet) ;
- ces plans permettent de plus une démarche séquentielle dans la construction d'une surface de réponse. Si la zone expérimentale la plus intéressante n'est pas dans le domaine défini préalablement, mais située à son voisinage, il est alors possible de construire autour de l'un des points du domaine initial un nouvel ensemble d'expériences qui intègre une partie des simulations déjà effectuées ;
- chaque simplexe étant déductible du simplexe de dimension immédiatement inférieur, il est alors facile d'augmenter le nombre de facteurs progressivement sans pour autant recommencer toutes les expériences ;
- il est possible de simplifier l'étude à une zone plus restreinte en changeant uniquement les dimensions du simplexe initial.

Le principal inconvénient de tels plans est la perte d'information due aux alignements dans l'espace d'autant qu'il existe des techniques de génération aléatoire de plans uniformes permettant de limiter les alignements [49].

1.3.4 Ajustement et validation du métamodèle

Le plan d'expériences numériques permet d'obtenir un jeu de données entrée/sortie pour l'ajustement du métamodèle. Le métamodèle est utilisé pour calculer les données de sortie correspondant au jeu de données d'entrée. Ces estimations sont ensuite évaluées en utilisant des critères statistiques. Le processus de validation du métamodèle est similaire au processus de validation d'un modèle de simulation. Cependant, cette procédure peut varier selon la nature du métamodèle. Dans le cadre de la validation on peut en effet distinguer deux types de métamodèles :

1. les métamodèles basés sur une *régression* pour lesquels le prédicteur n'est pas interpolé ;
2. les métamodèles d'*interpolation* tels les métamodèles MARS ou le krigeage (modèle probabiliste dont le prédicteur est un interpolateur).

Cette validation peut être effectuée selon les méthodes définies ci-après.

1.3.4.1 Ajustement dans le cas de métamodèles de régression

Un critère d'ajustement classique est le *coefficient de détermination* R^2 défini comme suit. Si l'on note \hat{y}_i les variables de sortie du métamodèle, y_i les résultats de référence, $\bar{y} = \frac{1}{n} \sum_i y_i$ une estimation moyenne de la réponse et n le nombre de points du plan d'expériences, le coefficient de détermination traduit l'adéquation entre \hat{y}_i et y_i . Il indique le taux de variation expliquée par la variation entre les deux variables. Le coefficient R^2 s'écrit :

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (2.1)$$

R^2 est un paramètre compris entre 0 et 1. R^2 est égal à 0 si les sorties du métamodèle sont constantes pour les n combinaisons de facteurs avec $i = 1, \dots, n$. La mesure R^2 est égale à 1 si et seulement si les sorties du métamodèle sont égales aux sorties du modèle de simulation soit $\hat{y}_i = y_i$. La mesure R^2 présente cependant un inconvénient important : cette mesure augmente toujours lorsque des variables de régression sont ajoutées. Un coefficient de détermination ajusté peut alors être calculé [107]. Il n'existe malheureusement pas de seuil inférieur pour ce critère.

1.3.4.2 Validation préalable dans le cas de métamodèles de régression

Dans ce cas, le modèle est alors $y = \mathbb{E}(Y|X = x) = f(x)$, où la loi de probabilité jointe (X, Y) est le fruit d'une régression sur $(x_i, y_i)_{i=1, \dots, n}$. On a alors en général des résidus² $\varepsilon_i = y_i - f(x_i) = y_i - \hat{y}_i$ non nuls. De plus, le modèle est généralement sous la forme $Y = f(X) + \varepsilon$ où $\varepsilon \sim \mathcal{N}(0, \Sigma)$. Il convient d'analyser les résidus ε_i vis-à-vis de cette hypothèse de normalité et pour s'assurer que l'écart entre réalisations y_i et prédictions \hat{y}_i est faible. Pour ce faire, un certain nombre de tests de normalité existent [107] et l'observation des histogrammes des résidus fournit également ces informations. Dans le cas où les propriétés des résidus sont vérifiées, le coefficient de détermination R^2 peut être calculé.

1.3.4.3 Validation du métamodèle

Ici, on voit le métamodèle comme une fonction déterministe, même si le métamodèle a une dimension probabiliste (régression ou krigeage par exemple) : $y = f(x)$. Deux approches sont possibles :

1. sur les données $(x_i, y_i)_{i=1, \dots, n}$ disponibles, $n - p$ sont utilisées pour construire le métamodèle et les restants pour mener la validation ;
2. la validation croisée³ (*k-fold cross-validation*, *leave-one-out cross-validation*, *leave-v-out cross-validation*, *split-sample method*). C'est la méthode *leave-one-out* qui est utilisée dans ce mémoire ; celle-ci est détaillée en section 1.2.3 du chapitre 4.

Quelque soit l'approche, plusieurs indicateurs statistiques peuvent être calculés :

- l'erreur quadratique moyenne = $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$, qui correspond à la moyenne des erreurs de prévision au carré ;
- l'erreur absolue moyenne = $\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ (moins sensible aux valeurs extrêmes que l'erreur quadratique moyenne) ;
- ou le résidu maximal = $\max |y_i - \hat{y}_i|$ pour $1 \leq i \leq n$ (vérification que l'ensemble des prévisions est correct).

2. Les résidus correspondent ici à la différence entre les résultats numériques de référence et les résultats métamodélisés.

3. le mot *validation* a une signification différente ; l'*ajustement de sous-ensembles* serait un meilleur terme pour la *validation croisée*

Dans ce cas un coefficient « assimilé » au coefficient de détermination (adapté pour la régression) peut être calculé : le coefficient de prédictivité. Ce coefficient dont la terminologie est peu employée en statistique est comparable au coefficient plus connu nommé PRESS (PRedictive Error Sum of Squares) [127].

1.4 BILAN

La constitution du jeu de données expérimentales, le choix du modèle et la construction éventuelle d'un métamodèle nécessitent une grande rigueur. Tout le reste du processus de quantification d'incertitude découle de cette première étape majeure qu'est la spécification du système. Données expérimentales non validées, choix de modèle de propagation inapproprié, oubli d'un paramètre influent majeur (comme la non prise en compte des caractéristiques de sol pour une étude de propagation en considérant des distances de l'ordre de la centaine de mètres) sont autant d'éléments pouvant induire une erreur considérable sur l'étude de quantification d'incertitude. Une fois que tous les éléments sont vérifiés, la base de données expérimentales et le (méta)modèle peuvent être utilisés dans le cadre du processus de calibration sous incertitude.

2 DÉFINITION DES OBSERVABLES

Une *observable* est par définition une quantité physique représentative des données mesurées et des données modélisées. Le processus de quantification d'incertitudes que nous proposons étudie les résidus ε (différence entre les données modélisées et les données mesurées) correspondant à chaque observable comme on le verra à la section 3.3 étant donné le modèle statistique retenu. Les résidus pour chacune des observables doivent vérifier les hypothèses suivantes : être *aléatoires*, *gaussiens* et *indépendants*. Un choix judicieux des observables est donc nécessaire afin que ces hypothèses soient vérifiées. Les observables doivent être définies en regard de la sensibilité du modèle et des phénomènes physiques étudiés. Il est également important de noter qu'un nombre trop important d'observables mènent à une indétermination du système d'étude. Il est donc essentiel de sélectionner des observables appropriées, représentatives des données à disposition et permettant de vérifier les hypothèses d'applicabilité du processus définies à la section 3.

2.1 OBSERVABLES PHYSIQUES

Une approche physique de définition des observables est utilisée dans le chapitre 3 en se basant sur une connaissance des données et des éléments de plus forte influence. Il s'agit alors de déterminer un nombre optimal d'observables faisant un compromis entre qualité d'information recherchée et la restriction du nombre d'observables pour que le problème reste identifiable (section 4.1).

2.2 OBSERVABLES DÉFINIES PAR DES OUTILS STATISTIQUES

Une seconde méthode de définition d'observables est proposée. Elle est basée sur une méthode d'Analyse en Composantes Principales fonctionnelles. L'analyse en composantes principales, ou ACP [107], est souvent la première méthode utilisée pour mener une analyse statistique descriptive. Cette méthode permet de dégager les premiers modes ou « composantes principales » de variation d'un jeu de données et combien d'entre eux paraissent substantiels. Les sommes cumulées des valeurs propres associées aux vecteurs propres de la base « des modes » sélectionnés sont les indices usuels de l'importance des composantes principales. Le nombre de modes de variation est requis pour obtenir une approximation satisfaisante des données originales. Seuls les modes dominants améliorent le rapport signal sur bruit des données conservées.

La sortie des modèles de propagation acoustique étant de nature fonctionnelle⁴, l'ACP est donc réalisée dans un cadre fonctionnel. Ces fonctions de base sont orthogonales et représentent avec précision les fonctions réelles du système sur une base à faibles dimensions. Elle est basée sur les travaux de Ramsay [99] et est décrite en détails en Annexe C.

Les observables correspondent ainsi aux composantes de la matrice de projection des variables sur les fonctions propres ; ces observables sont notées Y_j . Le nombre d'observables est sélectionné en prenant en compte les parts d'inertie des valeurs propres (la sélection sera faite selon la règle de Kaiser [107]). Le recours à de telles observables permet de diminuer la dimension du système d'étude. Les avantages principaux de la définition des observables par ACPf est la décorrélation des observables et la possibilité de revenir aisément au système physique initial. Pour mener une analyse en composantes principales fonctionnelles sur les données à disposition, les fonctions développées sous Matlab par J. Ramsay, G. Hooker et S. Graves [98] seront utilisées.

3 CALIBRATION DU SYSTÈME

3.1 INTRODUCTION

La phase de calibration⁵ consiste à déterminer les valeurs des paramètres d'entrée du modèle et leur variance d'estimation en utilisant une base de données expérimentales. Un aspect fondamental du traitement des incertitudes et variabilité a trait à l'utilisation de la notion de probabilité. Développée à partir du 17ème siècle avec une référence toute particulière aux jeux de hasard, la théorie des probabilités permet d'appréhender les processus dont la réalisation présente une certaine variabilité qui est le fruit du hasard ou considéré comme tel. Deux visions complémentaires de la notion de probabilité co-existent : il existe des différences philosophiques profondes entre approches fréquentiste (section 3.2.1) et bayésienne (section 3.2.2). Il est donc essentiel de faire un bilan général sur ces approches afin de comprendre les fondements de la méthode de calibration sous incertitude (section 3.2.3).

3.2 APPROCHES STATISTIQUES

3.2.1 Approche fréquentiste

L'approche fréquentiste, ou objectiviste, définit la probabilité d'un événement comme la valeur (unique) vers laquelle converge la fréquence d'observation de cet événement. De cette approche découlent les statistiques inférentielles classiques, dans lesquelles toute estimation se fonde exclusivement sur l'observation de données. En l'absence de données, ou si les données disponibles ne sont pas pertinentes, cette approche ne permet pas de construire d'estimation.

En termes de calibration, l'approche fréquentiste utilise un échantillon de données pour estimer une valeur fixe du vecteur de paramètres d'entrée incertains \mathbf{x} . L'approche fréquentiste utilise des fonctions appelées *estimateurs* qui permettent de relier les valeurs estimées des paramètres aux données observées. La méthode des moindres carrés ordinaires peut être appliquée pour calculer l'estimateur du vecteur des paramètres d'entrée. Celui-ci minimise la somme des carrés des écarts entre les données observées et les données modélisées. Cette méthode est mise en application pour des modèles non-linéaires avec des algorithmes itératifs de type Gauss-Newton. Le principal défaut de ces méthodes locales est que le résultat est très dépendant de la valeur de départ du vecteur de paramètres

4. Une variable aléatoire est dite fonctionnelle si ses valeurs sont dans un espace de dimension infinie. Une observation d'une variable fonctionnelle est appelée donnée fonctionnelle. Des données fonctionnelles sont des données issues de la mesure répétée d'un paramètre évolutif issu d'un processus.

5. Le terme *calibration* sera utilisé tout au long de ce mémoire malgré le fait qu'il constitue un abus de langage. Il est en effet important de noter que cette formulation est un anglicisme.

(l'algorithme convergeant vers un autre minimum local). Des algorithmes itératifs d'optimisation globale permettent de s'assurer d'une convergence vers le minimum global, *e.g.*, l'algorithme du « recuit simulé » ou des algorithmes génétiques [10].

3.2.2 Approche bayésienne

L'approche bayésienne, ou subjectiviste, définit la probabilité d'un événement comme le degré de croyance d'une personne quant à la survenue de cet événement. C'est donc une propriété de l'événement mais aussi du niveau de connaissance du modélisateur (dit « expert ») au moment de la modélisation. La probabilité d'un événement n'est unique que si tout le monde, à tout moment, a la même connaissance. Dans cette approche, le processus d'estimation repose à la fois sur la connaissance *a priori* du modélisateur (concernant par exemple le « support » du paramètre, *i.e.* l'intervalle à l'intérieur duquel il est certain que le paramètre se trouve) et sur les données observées s'il y a lieu.

En termes de calibration, les méthodes bayésiennes utilisent la connaissance *a priori* sur les paramètres θ . Cette méthode consiste d'abord à définir la distribution *a priori* $P(\theta)$; les bornes de cette distribution définissent une gamme de variation des paramètres et reflètent le niveau d'incertitude sur les paramètres avant de les confronter avec les données observées. Ensuite, la distribution *a posteriori* $P(\theta|Y)$ est calculée en multipliant la distribution *a priori* avec une fonction de vraisemblance $P(Y|\theta)$:

$$P(\theta|Y) \propto P(\theta) \times P(Y|\theta), \quad (2.2)$$

où \propto traduit une relation de proportionnalité. Cette formule correspond à un calcul classique de probabilité conditionnelle (formule de Bayes).

La fonction de vraisemblance correspond à la probabilité des sorties connaissant les paramètres d'entrée du modèle. Elle est déterminée par la distribution de probabilité des erreurs sur les observations en considérant les erreurs comme indépendantes, normalement distribuées et centrées en zéro. Cette estimation d'une nouvelle distribution des paramètres à partir des données expérimentales est basée sur le théorème de Bayes. Ces méthodes sont clairement expliquées par Van Oijen et *al.* [124], Gallagher [54] et Svenssona et *al.* [115]. Une méthode de maximum de vraisemblance pénalisé appelée Maximum *a posteriori* ou Mode *a posteriori* (MAP) consiste à maximiser le produit entre la fonction de vraisemblance et une distribution *a priori* sur les paramètres θ . Dans le cas où la distribution est uniforme, cette approche s'apparente à un cadre fréquentiste.

L'approche bayésienne est plus robuste que l'approche fréquentiste car elle permet d'intégrer des informations bien connues (informations *a priori*) associées au système ce qui permet d'éviter les problèmes d'identifiabilité (section 4.1). De plus, cette méthode fournit des informations aussi bien qualitatives que quantitatives afin de mieux comprendre la nature de la calibration et/ou de l'aspect prédictif du problème. Cependant cette méthode a un coût d'implémentation numérique important et la définition de la loi *a priori* peut s'avérer être une étape complexe.

3.2.3 Calibration sous Incertitude

La méthode de calibration choisie est basée sur l'approche fréquentiste présentée à la section 3.2.1. Les méthodes fréquentistes classiques sont utilisées pour calibrer les paramètres de modèles en prenant en compte l'incertitude sur les données (appelé bruit de mesures) et l'incertitude associée au modèle lui-même (chapitre 1 section 2.2).

La calibration sous incertitude est un processus de calibration amélioré qui permet initialement de prendre en compte les incertitudes d'entrée et finalement de quantifier ces incertitudes d'un point de vue *global*. Les différences avec d'autres méthodes de calibration sont que :

- l'on considère les incertitudes globales *i.e.* la combinaison des incertitudes épistémiques et de la variabilité ;
- les propriétés structurelles des modèles physiques et statistiques sont validées à chaque étape du processus sous des hypothèses initiales spécifiques.

Une telle méthode probabiliste est appelée méthode de Calibration sous Incertitude ou de façon plus synthétique CUU (Calibration Under Uncertainty). C'est cette méthode qui a été retenue et qui est décrite infra. Cette approche est efficace dans l'exploration et l'étude de l'incertitude de prévision. Elle présente de plus l'avantage d'une implémentation mathématique et numérique relativement simple. Cependant, la différentiabilité de la relation entre les paramètres du modèle et les sorties du modèle est requise, ce qui n'est pas toujours le cas du fait par exemple de valeurs seuils dans certains modèles. Lorsque les conditions de différentiabilité sont vérifiées, le processus de calibration est robuste et fournit un moyen « rapide » d'estimer l'incertitude de prévision. C'est donc une méthode puissante pour étudier l'incertitude d'un point de vue qualitatif et quantitatif.

3.3 DÉFINITION DU MODÈLE STATISTIQUE DE CALIBRATION SOUS INCERTITUDE

Le modèle statistique doit permettre de réduire l'incertitude sur les paramètres incertains et d'améliorer le contenu prédictif du processus de calibration. Celui-ci doit également permettre de minimiser les résidus et leur variance associée. Une méthode de minimisation communément utilisée est la méthode de minimisation par moindres carrés. Cependant on propose ici d'avoir recours à une méthode plus générale basée sur un principe différent de la minimisation de la fonction coût classique : la *maximisation de la fonction de vraisemblance*.

On considère un n -échantillon $\mathbf{y} = (y_1, y_2, \dots, y_n)$ de réalisations indépendantes, de n variables stochastiques Y_j tel que $Y_j = H(\mathbf{x}, d_j) + \varepsilon_j$ avec $\varepsilon_j \sim \mathcal{N}(0, \sigma^2)$. H correspond au modèle considéré, ε_j sont les résidus entre la mesure et le modèle, \mathbf{x} est le vecteur de paramètres (physiques) incertains et σ^2 est la variance des résidus. La fonction de log-vraisemblance est alors définie comme la densité de probabilité d'obtenir (Y_j, d_j) connaissant θ , par la relation $f[(Y_j, d_j) | \theta]$ avec $\theta = (\mathbf{x}, \sigma^2)$ pour des valeurs connues \mathbf{x} et σ^2 . Sous l'hypothèse selon laquelle les résidus sont normalement distribués, les observables mesurées Y_j peuvent aussi s'écrire comme une distribution gaussienne $Y_j \sim \mathcal{N}(H(\mathbf{x}, d_j), \sigma^2)$, centrée sur les données modélisées. La fonction de vraisemblance correspondante est alors définie par l'équation 2.3 :

$$\begin{aligned} L[(Y_j, d_j), \theta] &= \prod_{j=1}^n f[(Y_j, d_j) | \theta] \\ &= \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(Y_j - H(\mathbf{x}, d_j))^2}{2\sigma^2}\right]. \end{aligned} \quad (2.3)$$

Une écriture équivalente du problème de maximisation, mais cependant plus robuste numériquement, consiste à considérer le logarithme de cette fonction et à écrire la fonction coût sous la forme d'une fonction de log-vraisemblance (équation 2.4) :

$$\mathcal{L}[(Y_j, d_j), \theta] = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (Y_j - H(\mathbf{x}, d_j))^2. \quad (2.4)$$

Une généralisation de cette équation sous forme vectorielle s'écrit :

$$\mathcal{L} \left[(Y_j^l, d_j), \theta \right] = -\frac{n}{2} \log(\det[2\pi\Gamma]) - \frac{1}{2} \sum_{j=1}^n \left[(Y_j^l - H(\mathbf{x}, d_j))' \Gamma^{-1} (Y_j^l - H(\mathbf{x}, d_j)) \right], \quad (2.5)$$

où l est l'indice des observables, j l'indice des données d'entrée et Γ la matrice de covariance des résidus. Par analogie avec la définition de l'équation de log-vraisemblance précédente (équation (2.4)), la variance devient la matrice de covariance Γ calculée sur tous les résidus. La matrice de covariance constitue une source d'information majeure sur le processus de calibration : elle apporte des connaissances sur la structure des résidus et sur leur interaction *i.e.* sur le comportement des données mesurées et modélisées. Il est également essentiel de vérifier les propriétés de la matrice de covariance Γ car elle doit être inversible. La forme générique de la matrice de covariance s'écrit :

$$\Gamma = \begin{pmatrix} \sigma_1^2 & \cdots & \gamma_{1l}\sigma_1\sigma_l & \cdots & \gamma_{1n_{obs}}\sigma_1\sigma_{n_{obs}} \\ \vdots & \ddots & & & \vdots \\ \gamma_{l1}\sigma_l\sigma_1 & & \sigma_l^2 & & \gamma_{ln_{obs}}\sigma_l\sigma_{n_{obs}} \\ \vdots & & & \ddots & \vdots \\ \gamma_{n_{obs}l}\sigma_{n_{obs}}\sigma_l & \cdots & \gamma_{n_{obs}l}\sigma_{n_{obs}}\sigma_l & \cdots & \sigma_{n_{obs}}^2 \end{pmatrix}, \quad (2.6)$$

où $l = 1, \dots, n_{obs}$ (n_{obs} étant le nombre d'observables). σ_l sont les écarts-type des résidus, γ les coefficients d'intercorrélation entre chaque observable. Le produit entre les écarts-type σ_l et le coefficient d'intercorrélation correspond à la covariance entre les observables.

La maximisation de la fonction de log-vraisemblance (vectorielle ou non) est équivalente à une minimisation par moindres carrés. La formulation de la fonction de log-vraisemblance dépend du système considéré et peut être écrite de façon plus ou moins complexe selon ses propriétés et plus particulièrement celles des résidus ε . Le maximum de vraisemblance $\hat{\theta}$ est la solution des équations de vraisemblance $\frac{\partial \mathcal{L} [(Y_j, d_j), \theta]}{\partial \theta} = 0$.

La fonction coût choisie dans le processus de calibration sous incertitude est une fonction de log-vraisemblance. La nature linéaire ou non-linéaire du modèle considéré, la dimension des variables de sortie, l'apport d'un avis *a priori* sur les paramètres d'entrée sont autant d'éléments influents sur la forme finale de cette fonction de log-vraisemblance. Différents cas peuvent être abordés selon les propriétés du système et par suite les propriétés de la fonction coût. Les résidus modèle-mesure peuvent être *homoscédastiques* (la variance du résidu se situe autour d'une tendance centrale constante, *i.e.* σ fixe pour toutes les observables) ou *hétéroscédastiques* (la variance des résidus dépend des observables). L'hypothèse d'hétéroscédasticité permet de pondérer les paramètres selon leur importance. Il est notamment nécessaire de prendre en compte une précision variable du modèle lorsque les observables dépendent de différents capteurs et/ou de conditions environnementales variables. L'hypothèse de normalité est également faite.

3.4 BILAN

L'approche de calibration sous incertitude est une méthode fréquentiste permettant éventuellement d'intégrer de l'information *a priori* au système d'étude. Cette phase du processus de quantification d'incertitudes consiste à construire le modèle statistique en écrivant la fonction coût (la fonction de log-vraisemblance) à minimiser. L'écriture de la fonction de log-vraisemblance dépend des propriétés

des données expérimentales à disposition mais aussi des objectifs de l'étude ; les paramètres incertains sont ainsi estimés en maximisant cette fonction de log-vraisemblance. L'étape suivante du processus consiste à valider le système d'étude en se basant sur un certain nombre de propriétés et critères.

4 VALIDATION DU SYSTÈME

La validation du modèle statistique est une phase indispensable au processus de quantification d'incertitudes. Cette phase permet de vérifier l'identifiabilité du système, de vérifier les hypothèses initiales sur les propriétés des résidus ε (indépendants, distribution gaussienne) et de sélectionner le modèle statistique optimal.

4.1 IDENTIFIABILITÉ DES PARAMÈTRES DU MODÈLE

L'estimation de paramètres (ou identification) est une phase importante dans la modélisation d'un système [42, 126]. Si le système n'est pas identifiable, la convergence des paramètres vers la vraie valeur lors de la phase d'optimisation n'est pas garantie. L'analyse d'identifiabilité du système permet de détecter des modèles inadéquats ou mal définis et peut apporter des avantages tels que l'apprentissage des liens entre les paramètres d'entrée et les grandeurs d'intérêt de sortie. Cette connaissance peut être très utile dans la conception de la structure du modèle, dans la sélection du modèle ainsi que dans l'estimation numérique. L'approche d'identifiabilité des paramètres du modèle est développée en Annexe D. Cependant dans la pratique, la vérification de l'identifiabilité est complexe. En effet, on maîtrise l'identifiabilité uniquement dans un cas linéaire alors que nos systèmes étudiés sont généralement non-linéaires.

4.2 CRITÈRES STATISTIQUES

Dans la théorie de l'information, il n'existe pas de vrais modèles. La réalité peut être seulement approchée par des modèles dépendant de plusieurs paramètres physiques (*e.g.* ceux liés au sol, à la micrométéorologique). Le meilleur modèle basé sur le jeu de données considéré devrait être la meilleure approximation de la réalité. Certains critères statistiques permettent d'estimer la fiabilité des modèles et de les comparer entre eux. Ces critères ne sont efficaces et comparables que pour des modèles calculés à partir de variables similaires.

4.2.1 Critère d'information d'Akaike AIC

La perte d'information du fait de l'approximation de la « vérité » est mesurée par le nombre d'information de Kullback-Leibler (KL). Il n'est pas possible de calculer le nombre d'information KL directement car celui-ci dépend de la vérité qui est inconnue. Akaike a établi une approximation de la quantité KL appelée Critère d'Information d'Akaike, notée AIC [7]. Ce critère est défini par l'équation (2.7) :

$$AIC = 2q - 2\mathcal{L} [(Y_j, d_j), \hat{\theta}], \quad (2.7)$$

où q est le nombre de paramètres du modèle ajustés et \mathcal{L} est la valeur maximisée de la fonction de log-vraisemblance pour le modèle estimé. Pour qu'un modèle soit optimal, il faut que la valeur de l'AIC soit minimale. La méthodologie AIC tend à trouver le modèle qui explique au mieux des données avec un minimum de degrés de liberté et aussi un modèle qui fait le meilleur compromis biais-variance pour le nombre de données disponibles.

4.2.2 Critère d'information Bayésien BIC

Dans un cadre Bayésien, on cherche à sélectionner le modèle le plus vraisemblable au vu des données. Schwarz [109] a proposé une quantité appelée Critère d'Information Bayésien, noté BIC et

défini par l'équation suivante :

$$BIC = 2q \ln n - 2\mathcal{L} [(Y_j, d_j), \hat{\theta}], \quad (2.8)$$

où n est la taille de l'échantillon. Le critère BIC pénalise plus fortement le nombre de paramètres q que le critère AIC. Pour qu'un modèle soit optimal, la valeur de BIC doit être minimale. Le terme principal est simplement l'estimateur du terme de vraisemblance. Le second terme est affecté par une approche Bayésienne du fait des singularités de la distribution *a priori*. Le critère BIC est une façon d'estimer le meilleur modèle en utilisant seulement les estimations attendues et finalement de trouver le modèle qui prédit au mieux la réalité. Une présentation plus complète sur la construction de ce critère est donnée par Raftery [97].

4.2.3 Discussion concernant les critères AIC et BIC

Le critère AIC a tendance à favoriser un modèle trop riche en paramètres et d'autant plus compliqué qu'il y a plus de données. Il est cependant important de noter qu'il s'applique davantage dans le cas de *modèles explicatifs* alors que le critère BIC s'applique dans le cas de *modèles prédictifs*. L'objectif de notre étude d'incertitudes est en premier lieu de comprendre le comportement du modèle et par suite de procéder à une étude prédictive ; ces deux critères seront donc utilisés [79].

4.2.4 Coefficient de détermination R^2

Ce coefficient indique la part de variance expliquée par la relation entre les deux variables (section 1.3.4.1). Lorsque la variance associée au modèle est équivalente à la variance associée aux données expérimentales (soit R^2 proche de l'unité), cela signifie que le modèle prédit correctement toutes les valeurs observées Y_j à partir des données d'entrée. Dans le cas contraire, le modèle acoustique et les estimations ne permettent pas d'expliquer la variation observée et la valeur du coefficient R^2 est proche de zéro.

4.3 VÉRIFICATION DES HYPOTHÈSES INITIALES

Les hypothèses initiales sont relatives aux propriétés des résidus. Afin de vérifier l'indépendance des résidus, les coefficients de corrélation sont calculés (la valeur doit être au plus proche de zéro pour vérifier l'absence de corrélation des résidus), les nuages de corrélation entre résidus et les nuages de corrélation entre les résidus et les paramètres d'intérêt peuvent être tracés. Ces derniers permettent de mettre en évidence d'éventuelles propriétés non retranscrites par le coefficient de corrélation.

Il est de plus nécessaire de vérifier le caractère gaussien et non biaisé des distributions des résidus. Pour ce faire, on utilise les histogrammes, les représentations « p-p plot » (représentation d'une distribution de probabilité cumulative empirique en regard d'une distribution cumulative gaussienne ; cette représentation est particulièrement adaptée dans le cas où des petits échantillons de données sont considérés) ou des tests de Kolmogorov-Smirnov [107]. Le test de normalité de Kolmogorov-Smirnov fournit un indicateur, la *p-value*, qui rend compte de la probabilité qu'une distribution ne soit pas gaussienne. La *p-value* doit donc être typiquement $< 0,1$ pour rendre compte d'une distribution gaussienne.

4.4 BILAN

Ces outils statistiques doivent être utilisés de façon combinée afin de valider et choisir le modèle statistique adéquat. L'identifiabilité du système est une étape fondamentale dans le processus de calibration mais dans la pratique elle est très complexe à mettre œuvre dans le cas de systèmes non-linéaires à variables corrélées. De plus, les critères statistiques proposés doivent être interprétés en

regard des résultats physiques obtenus et de notre connaissance *a priori* des phénomènes physiques. Si un critère favorise un modèle préférentiellement à un autre alors que l'on sait que ce modèle n'est pas adapté, ou que le résultat obtenu n'a pas de signification physique, notre connaissance prendra alors le dessus sur les critères. Il est intéressant de se baser sur de tels indicateurs qu'il faut cependant interpréter avec un regard critique.

5 PRÉVISION

Une fois le système statistique validé, il est possible de procéder à la phase de prévision. La calibration du système permet en premier lieu d'estimer les paramètres d'entrée du modèle de propagation acoustique ainsi que les paramètres statistiques inclus dans la fonction coût (la variance des résidus et éventuellement leur corrélation). Lorsque θ est supposé fixé, la connaissance d'un intervalle de confiance sur θ ou sur une fonction d'intérêt de θ est une aide à la prise de décision. Lorsqu'une estimation ponctuelle $\hat{\theta}$ de θ est nécessaire, obtenir une mesure de l'incertitude de cette estimation est indispensable.

5.1 VARIANCE D'ESTIMATION

La variance d'estimation peut être estimée de deux façons différentes : en inversant la hessienne de la fonction de log-vraisemblance (ce qui est souvent complexe en pratique) et en ayant recours à des méthodes de bootstrap.

5.1.1 Par inversion de la hessienne de la fonction de log-vraisemblance

Les paramètres incertains θ sont estimés en maximisant la fonction de log-vraisemblance (respectivement en minimisant l'opposé de la fonction de log-vraisemblance) (section 3.3). Les paramètres physiques incertains \mathbf{x} et les paramètres statistiques incertains σ^2 sont estimés en résolvant les équations $\frac{\partial \mathcal{L}}{\partial \theta} = 0$. Sous certaines hypothèses, les estimateurs sont asymptotiquement gaussiens et leurs variances peuvent être (asymptotiquement) calculées pour estimer leur précision, en inversant la hessienne⁶ de la fonction de log-vraisemblance (équation 2.9) :

$$\sigma_{\hat{\theta}}^2 = \text{var} \hat{\theta} = \text{var} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\sigma}^2 \end{bmatrix} \cong \left[-\frac{\partial^2 \mathcal{L}}{\partial \theta \partial \theta'} \right]^{-1}. \quad (2.9)$$

Les variances associées aux résidus peuvent aussi être considérées comme paramètres inconnus et estimés. Cependant, cette approche étant souvent complexe à mettre en œuvre, on utilisera préférentiellement des procédures de bootstrap décrites ci-après pour calculer cette variance d'estimation.

5.1.2 Par procédures de bootstrap

Les méthodes de bootstrap ou méthodes de Cyrano⁷, introduites par Efron [44] et depuis couramment utilisées [118], permettent de pallier l'absence d'intervalles de confiance ou de proposer une alternative à leur construction. Contrairement à une construction théorique, ces procédures sont liées à un estimateur $\hat{\theta}$ dont elles estiment la variance notée $\sigma_{\hat{\theta}}^2$.

Il existe deux types de bootstrap :

6. La hessienne d'une fonction est la matrice carrée de ses dérivées partielles secondes.

7. Le nom de la méthode a été choisi par Bradley Efron, en référence aux Aventures du Baron de Münchhausen par Rudolph Erich Raspe. Parmi ses nombreuses histoires de voyages et exploits, le Baron de Münchhausen raconte qu'il a réussi à sortir d'un marécage en tirant sur ses bottes et en se propulsant dans les airs. Il se serait aidé des anneaux cousus sur le bord de ses bottes, appelés bootstraps en anglais. La méthode statistique étant *autocentrée* et *itérative*, sa dénomination est tout à fait justifiée. La méthode est également appelée méthode de Cyrano, car cette scène avait été reprise par Edmond Rostand dans Cyrano de Bergerac.

- le *bootstrap non paramétrique* (adapté aux échantillons de grande taille) simule M nouveaux échantillons en piochant aléatoirement des données Y_j avec remise ;
- le *bootstrap paramétrique* (préférable dans le cas des échantillons identiquement distribués) recrée intégralement M échantillons par des méthodes de Monte-Carlo à partir des moyennes et écarts-type des échantillons.

Pour chaque échantillon i (où $i = 1, \dots, M$), on procède à une estimation des paramètres $\hat{\theta}_i$ par maximisation de la fonction de log-vraisemblance (équation 2.4). A partir de cet échantillonnage de paramètres ainsi estimés, il est possible d'obtenir des intervalles de confiances à partir des quantiles observés et à partir de l'estimateur empirique $\sigma_{\hat{\theta}}^2$ (hypothèse gaussienne). Cet estimateur s'écrit selon l'équation 2.10,

$$\sigma_{\hat{\theta}}^2 = \frac{1}{M-1} \sum_{i=1}^M \left(\hat{\theta}_i - \frac{1}{M} \sum_{i=1}^M \hat{\theta}_i \right)^2, \quad (2.10)$$

et offre ainsi une estimation de la variance de l'estimateur $\hat{\theta}$. Les justifications de ce schéma (en particulier la consistance asymptotique) sont précisées par exemple dans Van der Waart [123]. Une fois l'estimateur $\sigma_{\hat{\theta}}^2$ produit, le calcul d'un intervalle de confiance est rendu possible. Plusieurs méthodes permettent de construire des intervalles de confiance pour θ (et pour toute fonction de θ). Le choix de la méthode dépend essentiellement du nombre d'échantillons à disposition (échantillons « bootstrappés »).

Il est possible que la méthode d'estimation par bootstrap produise un biais estimé par l'équation 2.11 :

$$b(\hat{\theta}) = \frac{1}{M} \sum_{i=1}^M [\hat{\theta}_i - \hat{\theta}(Y_j)]. \quad (2.11)$$

Les fractiles empiriques doivent être débiaisés en regard de l'estimation initiale $\hat{\theta}(Y_j)$, en leur retirant la quantité $b(\hat{\theta})$.

Ces méthodes de bootstrap ont pour vocation de répéter la procédure d'estimation sur des échantillons dont l'origine est supposée reproduire les conditions réelles de l'expérience. Par ailleurs, l'écart entre l'estimation et les valeurs centrales d'une distribution bootstrap permettent d'estimer le biais de la méthode, voire de le corriger. Cette méthode sera ainsi appliquée pour rendre compte de l'incertitude associée aux estimateurs du modèle. Le type de méthode employée (paramétrique ou non) dépendra des données sur lesquelles l'étude est menée.

5.2 INCERTITUDE GLOBALE

Dans un processus de quantification d'incertitudes complet, l'incertitude combinée d'un système, *i.e.* les variances associées aux résidus et au modèle, peut être calculée. Un processus de Monte-Carlo ou des méthodes de simulation avancées sont utilisées. Des distributions de probabilité gaussiennes sont générées, avec des valeurs moyennes égales aux paramètres estimés $\hat{\theta}$ (paramètres physiques et biais) et des variances associées. Ces variances correspondent aux variances d'estimation notées $\sigma_{\hat{\theta}}^2$. Les distributions de paramètres générées sont des distributions normales de la forme $\theta_p \sim \mathcal{N}(\hat{\theta}, \sigma_{\hat{\theta}}^2)$.

Ces paramètres constituent un nouveau jeu de paramètres d'entrée du modèle ; de nouvelles données modélisées et de nouveaux résidus sont ainsi calculés pour chaque observable spécifique. L'incertitude combinée est alors calculée en associant variances de résidus et variances du modèle :

$$\sigma_{Y_j}^2 = \sigma_{H_j}^2 + \sigma_{\varepsilon_j}^2, \quad (2.12)$$

où $\sigma_{Y_j}^2$ correspond à la variance des observables expérimentales, $\sigma_{H_j}^2 = \text{var}[H_j(\theta_p)]$ à une variance de modèle du fait de la variance d'estimation des paramètres et $\sigma_{\varepsilon_j}^2$ à la variance des résidus. Cette dernière rend compte de toutes les incertitudes et de la variabilité qui ne peuvent être expliquées grâce au processus de quantification d'incertitudes. Les résidus intègrent donc tous les éléments non-expliqués par le modèle acoustique, l'incertitude expérimentale ainsi qu'une part de variabilité non-expliquée par les paramètres. Des intervalles de confiance associés à chaque observable peuvent être calculés à partir de ces variances ; la procédure est la même que dans le cas du bootstrap (section 5.1.2).

Le processus de quantification étant mené pour des groupes de données rendant compte de classes de propagation acoustique particulières, ces variances pourront ainsi être associées à chacun de ces groupes. Le contenu prédictif du modèle (ici acoustique) en sera ainsi renforcé ; à une prévision pourra être associé un écart-type, représentatif de la variabilité et de l'incertitude relatives aux phénomènes.

6 CONCLUSION DU CHAPITRE

Nous avons explicité la méthodologie statistique envisagée pour répondre à la problématique d'incertitudes en propagation acoustique en milieu extérieur. Le recours à des processus de calibration n'est pas nouveau dans le domaine de la statistique et a notamment été développée dans le cadre d'analyses de risques (économétrie, problématiques industrielles, etc.), ou d'études environnementales soumises elles aussi à de fortes incertitudes. La méthodologie complète que nous proposons a été construite autour du processus de calibration envisagé initialement. A ce jour, cette méthodologie n'a jamais été appliquée à notre domaine d'étude et s'avère d'une grande utilité pour répondre à l'objectif principal des travaux : la quantification de l'incertitude globale sur les niveaux sonores du fait de l'incertitude et de la variabilité des paramètres associés aux phénomènes influents.

Ces travaux n'ayant pas pour ambition de recouvrir l'ensemble des phénomènes et situations rencontrés en réalité, seuls quelques cas d'études seront présentés. Le premier cas d'étude (chapitre 3) considère uniquement l'influence des effets de sol sur la propagation acoustique (courtes distances de propagation) pour deux modèles de sol différents. Le second cas d'étude (chapitre 4) prend en compte des distances de propagation plus grandes, ce qui induit de prendre également en compte les phénomènes micrométéorologiques. Ces applications concrètes permettent de présenter la mise en œuvre pratique, l'apport de la méthode ainsi que les limites auxquelles nous avons pu être confrontés.

3

ESTIMATION D'INCERTITUDES ASSOCIÉES AUX CARACTÉRISTIQUES ACOUSTIQUES DE SOL

PLAN DU CHAPITRE

1	SPÉCIFICATION DU SYSTÈME	61
1.1	Données observées : mesures d'impédance de sol de la campagne expérimentale de Lannemezan 2005	61
1.2	Choix du modèle	62
2	DÉFINITION DES OBSERVABLES	64
2.1	Approche "physique"	64
2.2	Approche "statistique" basée sur la méthode d'Analyse en Composantes Principales fonctionnelles	68
2.3	Comparaison des méthodes de définition des observables	70
2.4	Bilan	76
3	CALIBRATION DES MODÈLES D'IMPÉDANCE DE SOL	77
3.1	Définition de la fonction coût	77
3.2	Calibration avec le modèle de Delany et Bazley	77
3.3	Calibration avec le modèle de Zwicker et Kosten	83
4	VALIDATION ET CHOIX DU MODÈLE STATISTIQUE OPTIMAL	88
4.1	Observables physiques	88
4.2	Observables définies par ACPf	89
4.3	Bilan	90
5	CONCLUSION DU CHAPITRE	90

La quantification des incertitudes associées aux caractéristiques de sol est une première application simple de l'approche de calibration sous incertitude. Comme nous avons pu le voir au chapitre 1, ces paramètres sont soumis à de fortes incertitude et variabilité qu'il est donc essentiel d'appréhender. Cette étude de quantification d'incertitudes est effectuée afin de répondre à un certain nombre d'objectifs :

1. tester l'approche de calibration sous incertitude ;
2. faire évoluer la méthodologie initiale en fonction des "obstacles" rencontrés et appliquer la méthodologie complète présentée au chapitre 2 ;
3. avoir une meilleure connaissance des incertitudes liées aux paramètres acoustiques de sol ;
4. comparer la robustesse de différents modèles acoustiques de sol vis-à-vis du processus de calibration.

1 SPÉCIFICATION DU SYSTÈME

1.1 DONNÉES OBSERVÉES : MESURES D'IMPÉDANCE DE SOL DE LA CAMPAGNE EXPÉRIMENTALE DE LANNEMEZAN 2005

Les données expérimentales sont issues de la campagne de Lannemezan 2005 (Annexe A). Les paramètres de sol sont mesurés selon le dispositif expérimental fixe représenté par la figure 3.1 du chapitre 1. Les données à disposition sont des niveaux de pression sonore mesurés au microphone 1 (M1) à 2 m de hauteur par rapport au sol relativement aux niveaux de pression sonore mesurés au microphone 2 (M2) à 30 cm du sol (bandes fines). La base de données est constituée d'un total de 479 spectres d'atténuation représentatifs de l'ensemble de la campagne expérimentale (3 mois) après validation et nettoyage des données initiales [68]. Les caractéristiques de sol étant difficilement mesurables *in situ*, on détermine la valeur de ces paramètres par méthode inverse : on recale un modèle de sol sur une mesure (spectre acoustique).

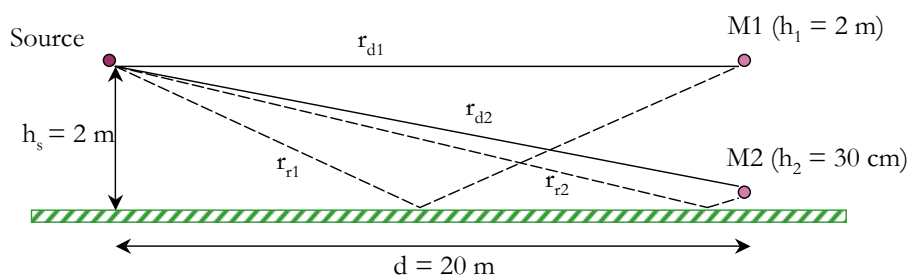


FIG. 3.1 – Dispositif expérimental.

1.1.1 Création de groupes de données expérimentales

La constitution de groupes de spectres, représentatifs de périodes temporelles données, présente un grand intérêt pour étudier la variation de la valeur de l'écart-type associé aux résidus (différence entre les données expérimentales et les données calculées) en fonction du groupe. L'application de l'approche de calibration sous incertitudes à ces groupes est une manière d'étudier la variabilité intrinsèque, la sensibilité du modèle associée à chaque groupe et finalement l'incertitude combinée du système physique.

Une première étude basée sur la calibration manuelle (recalage des paramètres par des acousticiens expérimentés) de chaque spectre mesuré montre la tendance globale des valeurs de résistance spécifique au passage de l'air sur l'intégralité de la campagne expérimentale (figure 3.2). La présence d'eau dans le sol peut induire une diminution de la porosité : l'absorption de l'eau implique une diminution des vides d'air dans le sol et le lavage des particules induit le colmatage des pores. Par conséquent, l'impédance de sol augmente ; cette augmentation est d'autant plus importante lorsque la saturation est proche [34]. Ainsi, en cas d'événements de précipitations, un retard est observé pour le retour à la valeur initiale de la résistance spécifique au passage de l'air. Ce retard dépend de la durée des événements de précipitations et de l'ensoleillement. La variabilité observée sur des échelles temporelles plus courtes s'explique par le cycle jour/nuit avec la formation de rosée en fin de période nocturne et le réchauffement de la surface en début de période diurne.

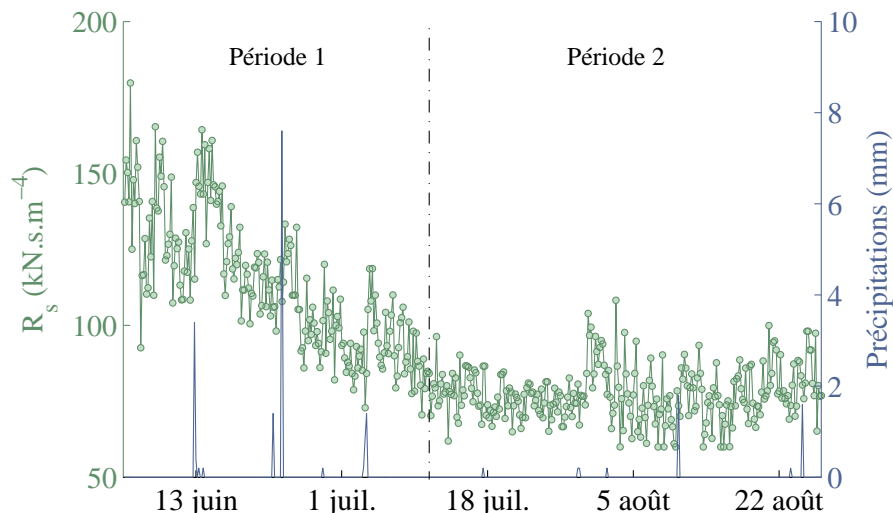


FIG. 3.2 – Évolution des valeurs de résistance spécifique au passage de l'air R_s et des précipitations au cours de la campagne expérimentale de Lannemezan 2005. Deux périodes sont considérées.

Cette tendance met en évidence deux périodes temporelles spécifiques au cours de la campagne expérimentale, qui sont fortement corrélées avec les conditions micrométéorologiques et plus particulièrement avec les événements de précipitations :

1. la première période de la campagne (du 6 juin 2005 au 10 juillet 2005) montre une diminution globale des valeurs de résistance spécifique au passage de l'air et une amplitude de variation globalement importante (d'environ 100 kN.s.m^{-4}). L'augmentation locale des valeurs de résistance spécifique au passage de l'air est due aux événements de précipitations prédominants au cours de cette période ;
2. la seconde période (du 10 juillet 2005 à la fin de la campagne expérimentale) présente une phase stable avec des variations en amplitude plus faibles de la résistance spécifique au passage de l'air (de l'ordre de 50 kN.s.m^{-4}).

A noter qu'une tendance inverse est classiquement observée pour le paramètre d'épaisseur e (figure 3.3). La sensibilité du paramètre d'épaisseur aux événements de précipitations est moindre.

Deux groupes de spectres correspondant à ces périodes temporelles sont donc constitués (figure 3.4) avec 200 échantillons pour le premiers groupe et 279 pour le second. Ces deux groupes rendent compte d'une position du creux d'effet de sol légèrement différente en fréquence : on repère le creux d'effet de sol plus près de 200 Hz dans la deuxième période que dans la première (figure 3.4). De plus, la dispersion des spectres relatifs à la première partie de la campagne expérimentale est plus forte, ce qui rend bien compte d'une variabilité plus importante des données au cours de cette période.

1.1.2 Paramètres d'entrée du système

Les paramètres d'entrée du système correspondent aux paramètres acoustiques de sol qui dépendent directement du modèle d'impédance de sol utilisé.

1.2 CHOIX DU MODÈLE

Le modèle que nous considérons dans cette étude est constitué d'un modèle d'impédance de sol (cf. chapitre 1 section 1.3.3.3) et d'un modèle de propagation analytique simple (Rudnick) (cf. chapitre 1 section 2.3). L'étude que nous menons ici est fondée sur l'estimation de l'incertitude associée aux

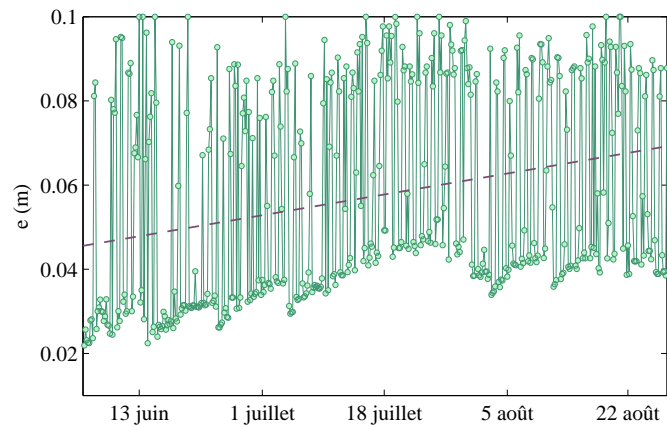


FIG. 3.3 – Évolution des valeurs d'épaisseur e au cours de la campagne expérimentale de Lannemezan 2005 - la ligne pointillée rend compte de la tendance "globalement croissante" de l'évolution du paramètre e .

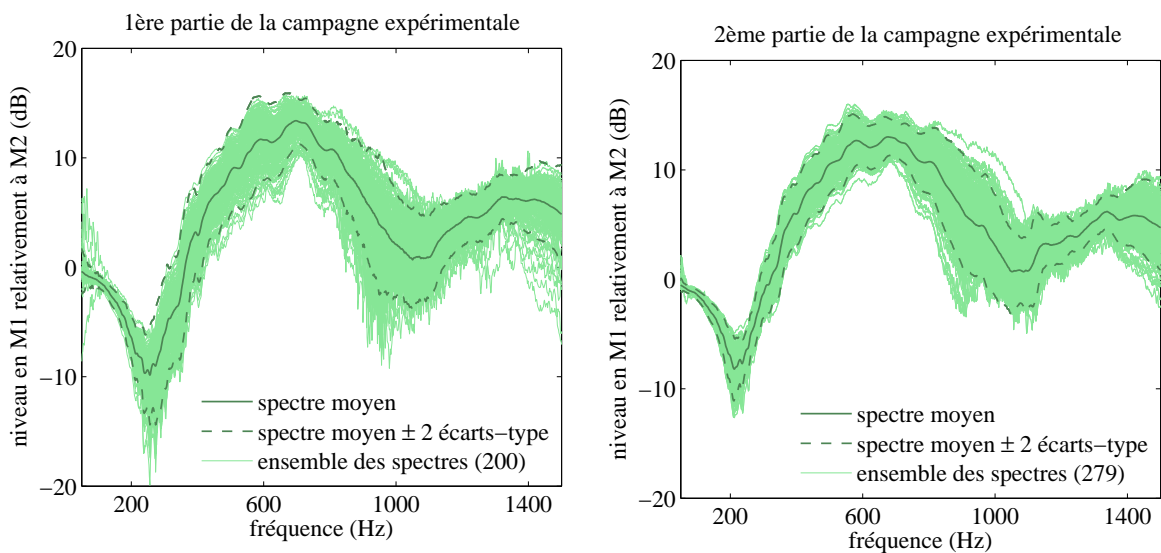


FIG. 3.4 – Deux groupes de spectres d'atténuation relative mesurés correspondant à la première et à la seconde période de la campagne expérimentale de Lannemezan 2005.

paramètres du modèle et non à l'estimation de l'incertitude associée aux spectres de niveaux de pression mesurés aux microphones 1 et 2.

Deux modèles d'impédance de sol vont être considérés dans cette étude en prenant en compte l'effet d'épaisseur. Le modèle semi-empirique à un paramètre de Delany et Bazley, ainsi que le modèle à trois paramètres de Zwicker et Kosten. Etant donnés les modèles utilisés, les paramètres incertains θ *i.e.* les paramètres à calibrer sont :

- la résistance spécifique au passage de l'air notée R_s et l'épaisseur e , dans le cas du modèle de Delany et Bazley ;
- la résistance spécifique au passage de l'air R_s , la porosité Ω , la constante de structure C_s (caractérisant la tortuosité du sol) et l'épaisseur e dans le cas du modèle de Zwicker et Kosten.

L'ensemble de ces paramètres est défini au Chapitre 1. Ces modèles dépendent également de paramètres connus notés d_j : position des capteurs, distance source-récepteur et température de l'air. L'application du processus de quantification d'incertitudes à deux modèles différents permet de tester la robustesse du processus selon la complexité du modèle considéré, liée ici au nombre de paramètres d'entrée.

2 DÉFINITION DES OBSERVABLES

Le choix des observables consiste à mettre en évidence les éléments caractéristiques du spectre d'atténuation relative. Dans une étude préliminaire de l'application de l'approche de calibration sous incertitude, le spectre d'atténuation en bandes fines est utilisé comme observables. Les observables d'étude correspondraient dans ce cas à l'ensemble des composantes fréquentielles. Ceci a conduit à une sur-paramétrisation de l'information et à la singularité du modèle statistique (annulation du déterminant de la matrice de covariance Γ) du fait de la forte corrélation entre chaque observable (cf. chapitre 2). Ce choix n'apporte donc pas d'information majeure quant aux phénomènes étudiés et ne permet pas de mettre en évidence les éléments caractéristiques du spectre. Par conséquent il a été nécessaire de définir de nouvelles observables plus représentatives de l'information donnée par le modèle physique. Deux approches ont ainsi été envisagées pour définir des observables robustes représentatives des données étudiées :

1. une approche "physique" basée sur le comportement d'un opérateur lors de la calibration manuelle de spectres ;
2. une approche "statistique" basée sur une Analyse en Composantes principales fonctionnelles (notée ACPf) des spectres d'atténuation, qui consiste à ne considérer que les coefficients d'ACPF expliquant un maximum de variance des données (chapitre 2).

Ces deux approches sont détaillées dans ce qui suit.

2.1 APPROCHE "PHYSIQUE"

Le comportement d'un expert/opérateur lors de calages de spectres d'atténuation à partir d'un modèle d'impédance de sol a été observé et analysé. Le modèle considéré ici est le modèle d'impédance de sol de Delany et Bazley [41] corrigé par l'effet d'épaisseur et combiné au modèle de propagation analytique de Rudnick [104]. Le calage manuel d'un spectre consiste donc à ajuster deux paramètres, R_s et e . L'étude du comportement d'un opérateur lors d'un calage de spectre renseigne sur les zones du spectre auxquelles l'opérateur accorde davantage d'importance. Ces zones sont les plus caractéristiques des propriétés du sol.

2.1.1 Étude du comportement d'un expert lors du calage manuel des données

Afin de mener à bien cette étude, cinq spectres d'atténuation relative sont générés aléatoirement une centaine de fois. Pour chaque spectre généré, l'opérateur procède "en aveugle" (*i.e.* sans connaître quel paramètre il fait varier ni les valeurs de ces paramètres) au calage des paramètres de résistance spécifique au passage de l'air R_s ($kN.s.m^{-4}$) et d'épaisseur e (m). Les résultats de calage pour l'ensemble des opérateurs et pour un spectre sont représentés figure 3.5.

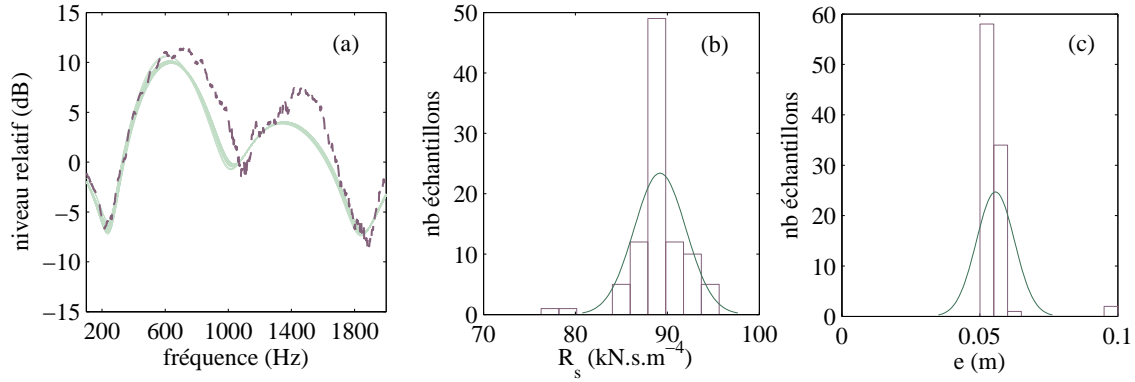


FIG. 3.5 – (a) : spectre calibré par calages manuels en aveugle (par 5 opérateurs) - la courbe en pointillés correspond au spectre mesuré et l'ensemble de courbes continues constitue le fuseau de spectres calibrés - (b) et (c) : histogrammes des valeurs de résistance spécifique au passage de l'air R_s et d'épaisseur e correspondant à ces calages.

Les conclusions suivantes sont faites par observation des résultats de calage sur l'ensemble des spectres et sont validées grâce aux témoignages et remarques des opérateurs se prêtant à l'exercice.

1. les opérateurs “calent” globalement un spectre mesuré de façon similaire : les calages d'un même spectre n fois par différents opérateurs présentent une dispersion relativement faible. L'intérêt principal d'un calage en aveugle est que l'opérateur n'est pas influencé par les valeurs des paramètres acoustiques de sol ;
2. les opérateurs se concentrent essentiellement sur la zone “ basses fréquences ” du spectre. En effet, le modèle de Delany et Bazley ne modélise pas de manière optimale le comportement aux fréquences plus élevées (diffraction expérimentale par la rugosité de surface) ;
3. les opérateurs accordent une grande importance à la position en fréquence du creux d'effet de sol (très dépendante de la valeur de la résistance spécifique au passage de l'air R_s) ;
4. les opérateurs essaient (lors d'un calage) de minimiser les écarts entre le spectre modélisé et le spectre mesuré de part et d'autre du creux d'effet de sol.

2.1.2 Choix des observables

Suite à l'étude présentée en section 2.1.1, trois observables ont été retenues. Elles sont représentées par la figure 3.6 et correspondent à :

1. la valeur de la fréquence correspondant au premier creux d'effet de sol notée Y_1 , caractéristique de la valeur du paramètre R_s ;
2. l'intégrale sur le spectre d'atténuation noté L_{att} sur la partie située à gauche du premier creux d'effet de sol :

$$Y_2 = \int_{f_{min}}^{Y_1+dx} L_{att}, \quad (3.1)$$

où $f_{min} = 50$ Hz et $dx = 100$ Hz, caractéristique de la partie basses fréquences du spectre. A noter que les valeurs choisies pour f_{min} , f_{max} et dx dépendent de l'allure générale du spectre d'atténuation et doivent être adaptées de manière empirique au cas d'étude considéré, notamment si le type de sol ou la géométrie du problème sont très différents ;

3. l'intégrale sur le spectre d'atténuation entre la position du creux d'effet de sol et la fréquence maximale :

$$Y_3 = \int_{Y_1+dx}^{f_{max}} L_{att}, \quad (3.2)$$

où $f_{max} = 500$ Hz et $dx = 100$ Hz, caractéristique de la partie moyenne fréquence du spectre.

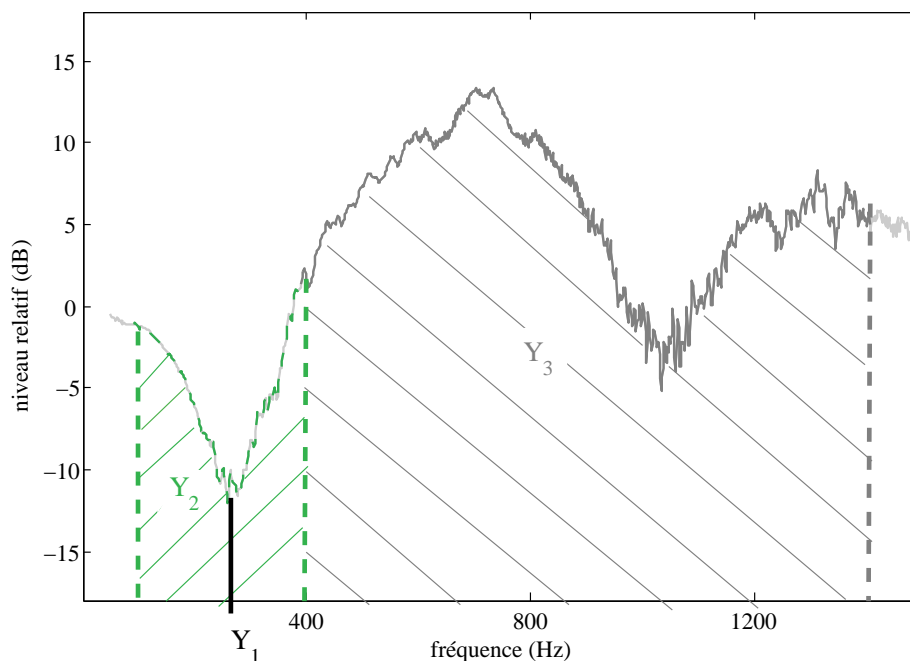


FIG. 3.6 – Représentation des observables sur un spectre mesuré.

2.1.3 Bilan

Les observables calculées sur l'ensemble de la campagne expérimentale sont représentées figure 3.7. Cette représentation permet d'avoir une vision globale du comportement des observables entre elles et fait apparaître une corrélation entre les observables 1 et 3 (figure 3.8). Etant donné que le processus de calibration requiert l'indépendance des résidus ε (chapitre 2 section 4.3), la corrélation des observables - qui n'implique pas nécessairement la corrélation des résidus - pourrait poser des problèmes conséquents en termes de mise en œuvre de la méthodologie ou en termes de résultats.

La représentation (figure 3.7) met également en évidence le découpage de la base de données expérimentales en regard des observables. Deux groupes de données bien distincts se dégagent du nuage de points comme pressenti lors du découpage initial (section 1.1.1).

L'avantage majeur de la définition physique des observables est la prise en compte du comportement d'un opérateur vis-à-vis du modèle en regard des données. Il reflète les éléments importants du spectre ainsi que leur sensibilité aux paramètres à recalculer. Cette définition est relativement flexible selon la forme des spectres étudiés : les bornes du domaine d'étude sont modulables et peuvent être adaptées en fonction des données à disposition.

Les inconvénients majeurs de ce type de description sont que les observables ne sont pas homogènes *i.e.* de grandeurs physiques semblables ; les observables sont alors difficilement comparables¹. En terme prédictif, le retour à une grandeur purement acoustique (*e.g.* à un spectre acoustique) n'est pas aisé. En effet, l'étude d'estimation d'incertitude étant menée relativement aux trois observables, l'extrapolation des résultats à un spectre complet n'est pas une démarche rigoureuse. Il est en effet complexe voire

1. Une possibilité pourrait être de centrer et réduire ces observables mais leur interprétation physique n'en serait pas plus aisée.

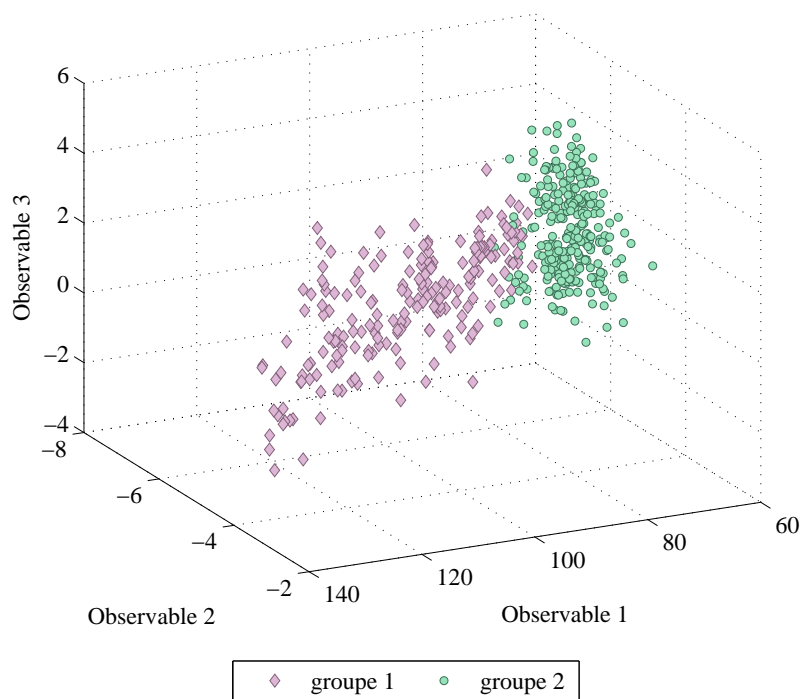


FIG. 3.7 – Nuage de points représentant les observables définies par calage manuel des spectres pour les deux groupes de données à l'étude. On remarque bien deux groupes de données distincts.

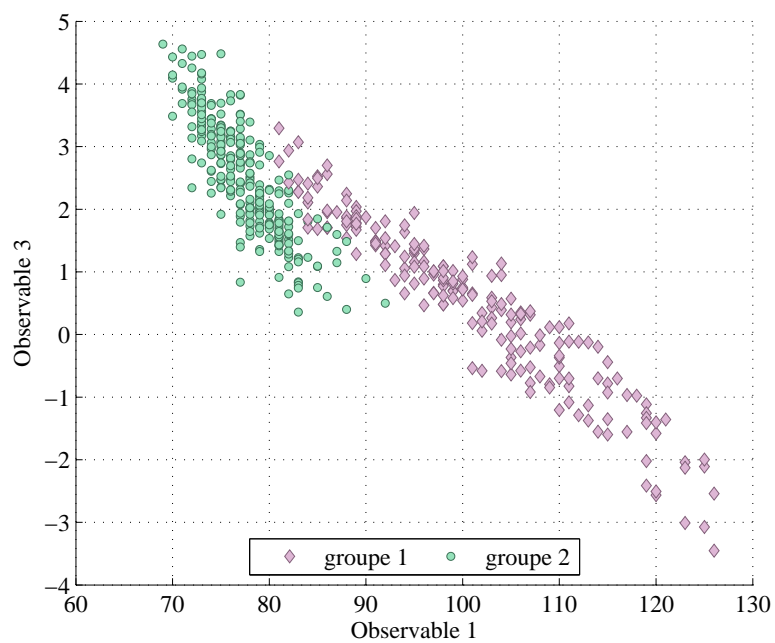


FIG. 3.8 – Nuage de points représentant les observables 1 et 3 définies par calage manuel des spectres pour les deux groupes de données à l'étude. Il existe une corrélation entre les deux observables.

impossible de reconstituer la forme d'un spectre sur une bande de fréquences donnée uniquement à partir de trois niveaux d'information. Cette méthode de définition des observables ne peut alors être appliquée que dans le cadre d'un processus d'estimation des paramètres et d'erreurs associées à ces paramètres qui ne nécessite pas de s'intéresser à la prévision des spectres. Cette méthode étant "subjective", d'autres observables auraient également pu être définies en s'assurant qu'elles soient pleinement représentatives d'un spectre d'atténuation.

2.2 APPROCHE "STATISTIQUE" BASÉE SUR LA MÉTHODE D'ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLES

La seconde méthode de définition des observables proposée est basée sur la méthode d'Analyse en Composantes Principales fonctionnelles (ACPF) non réduite². Cette méthode est expliquée en détails au chapitre 2 section (2) et est appliquée à la même base de données d'impédance de sol de la campagne expérimentale de Lannemezan 2005 en conservant le même domaine d'étude que dans le cas de la définition des observables d'un point de vue physique. Il est essentiel que les observables soient calculées sur un même domaine afin que les approches soient comparables.

2.2.1 Choix des observables

Ces observables correspondent aux coefficients de l'ACPF. Le choix du nombre de coefficients à sélectionner est basé sur la valeur des valeurs propres qui traduit le pourcentage de variance expliquée par les fonctions propres de l'ACPF. La sélection des fonctions propres a été menée selon la règle de Kaiser [107] (figure 3.9).

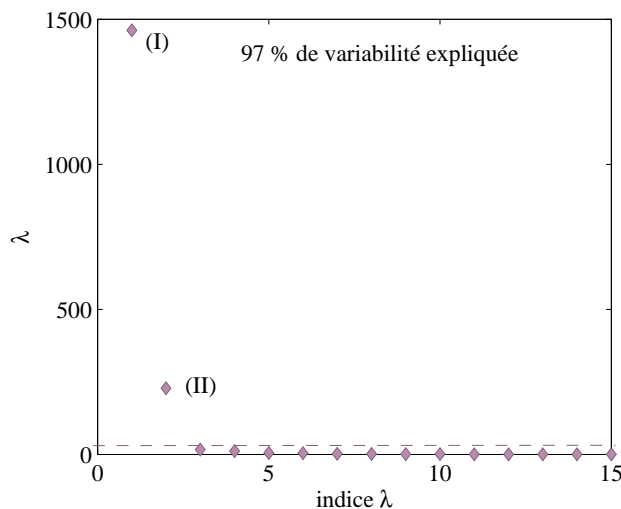


FIG. 3.9 – Valeurs propres de l'ACPF en appliquant la règle de Kaiser, les deux premières valeurs propres permettent d'expliquer 97 % de la variance.

La règle de Kaiser stipule que seules les fonctions dont les valeurs propres sont supérieures à l'unité doivent être retenues. Dans notre cas, les deux premières valeurs propres sont supérieures à l'unité et permettent au total d'expliquer 97 % de la variance des données. Deux fonctions propres seront alors conservées et utilisées pour projeter les observables *i.e.* les coefficients de l'ACPF. Les fonctions propres sélectionnées sont représentées figure 3.10.

². Ce choix de métrique privilégie les variables les plus dispersées pour lesquelles les différences entre individus sont les plus fortes. A noter qu'un élément intéressant pourrait être d'appliquer une pondération non uniforme en fonction de la fréquence.

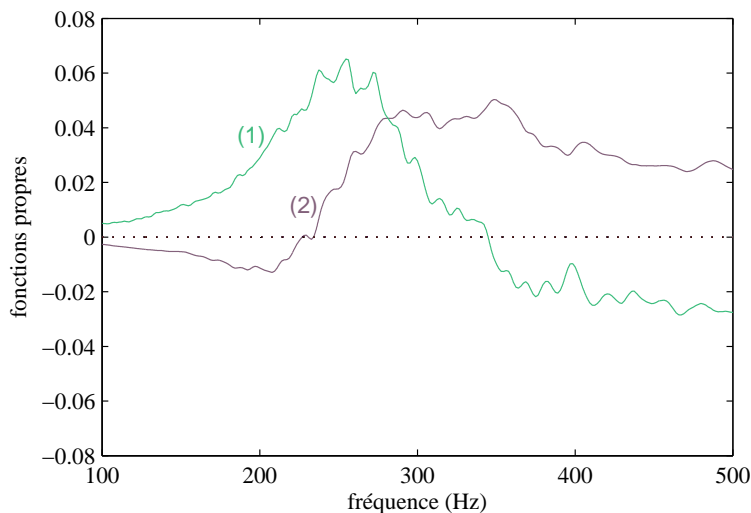


FIG. 3.10 – Deux premières fonctions propres définies par ACPf.

2.2.2 Bilan

Les observables ainsi définies sont représentées par un nuage de points bidimensionnel (figure 3.11).

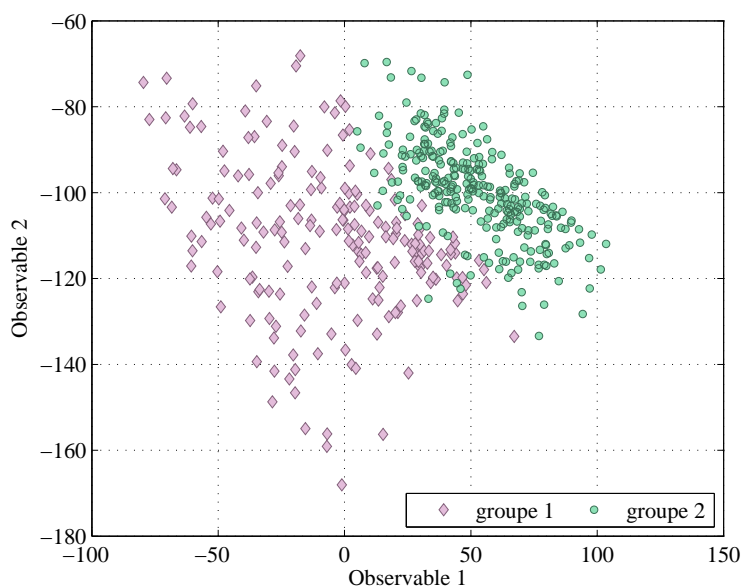


FIG. 3.11 – Nuage de points représentant les observables définies par ACPf pour les deux groupes de données à l'étude.

Cette représentation permet de montrer l'absence de corrélation entre les observables, par comparaison avec les observables physiques (figure 3.7), ce qui est l'un des avantages majeurs de la définition des observables par ACPf, qui sont par définition non corrélées. En effet, la projection des données sur la base induit l'indépendance entre les coefficients de l'ACPf. La définition des observables par la méthode d'ACPf présente donc un avantage non-négligeable en comparaison des observables physiques : l'indépendance des observables – qui favorise l'indépendance des résidus – devrait permettre de vérifier les hypothèses initiales.

Les deux groupes de spectres qui vont être étudiés sont également bien distincts dans le domaine des observables, ce qui permet de juger une fois encore de la cohérence de cette constitution de groupes. De plus, les observables définies par ACPf sont homogènes. Le nombre d'observables varie en fonction du taux de variabilité des données à expliquer ; un critère peut être aisément fixé pour choisir une valeur seuil de variabilité. Il est cependant difficile de savoir quelle signification physique ont les coefficients de l'ACPf. En contrepartie, le retour au domaine physique est simple : on peut en effet re-synthétiser des spectres approchés par des fonctions de base.

2.3 COMPARAISON DES MÉTHODES DE DÉFINITION DES OBSERVABLES

Cette étude comparative a pour objectif de définir des observables optimales afin de garantir la robustesse du modèle statistique pour l'estimation des paramètres incertains. Cette étude est basée sur la comparaison des paramètres du modèle et des résidus modèle-mesures obtenus par calibration spectre par spectre de l'ensemble des données à disposition, pour chacune des définitions d'observables et pour chaque modèle d'impédance de sol. Cette approche spectre par spectre est plus simple que la calibration par groupe des spectres ce qui permet ainsi une étude préliminaire de notre système d'étude peu coûteuse en temps de calcul.

2.3.1 Modèle de Delany et Bazley

2.3.1.1 Estimation et interprétation des paramètres

Les résultats de la calibration spectre par spectre sur l'ensemble des spectres sont synthétisés dans le tableau 3.1 ; ce tableau rend compte des moyennes $\langle \theta \rangle$ et écarts-type σ_θ des paramètres estimés pour les deux types d'observables étudiés.

Paramètres incertains	Obs physiques		Coefficients ACPf	
	$\langle \theta \rangle$	σ_θ	$\langle \theta \rangle$	σ_θ
R_s	93	24	101	26
e	0,06	0,03	0,05	0,01

TAB. 3.1 – *Moyenne des paramètres estimés spectre par spectre selon les deux méthodes de définition d'observables, avec le modèle de Delany et Bazley et écarts-type associés sur la base de données complète.*

Les paramètres estimés sont relativement cohérents d'une définition d'observables à l'autre en regard des valeurs des écarts-type. Ces écarts-type rendent compte de la variabilité physique des spectres dans le temps. En se basant sur les coefficients de l'ACPf la moyenne des valeurs de résistance spécifique au passage de l'air est légèrement plus élevée que dans de cas de la définition physique des observables. Cet écart n'est cependant pas significatif en regard de la précision expérimentale sur R_s qui est de l'ordre de $\pm 20 \text{ kN.s.m}^{-4}$ pour les types de sol considérés [18]. Les valeurs moyennes d'épaisseur sont également proches et les ordres de grandeur de ces paramètres sont cohérents en regard des valeurs de résistance spécifique au passage de l'air associées (épaisseur plus élevée pour une résistance spécifique au passage de l'air plus faible).

Dans le cas d'un modèle d'impédance de sol à seulement deux paramètres, deux méthodes de définition d'observables fournissent donc des résultats presque similaires. De plus, ces estimations faites spectre par spectre présentent une bonne corrélation. Celle-ci peut être observée grâce au graphe de corrélation des valeurs de résistance spécifique au passage de l'air estimées à partir des observables physiques et des observables par ACPf (figure 3.12).

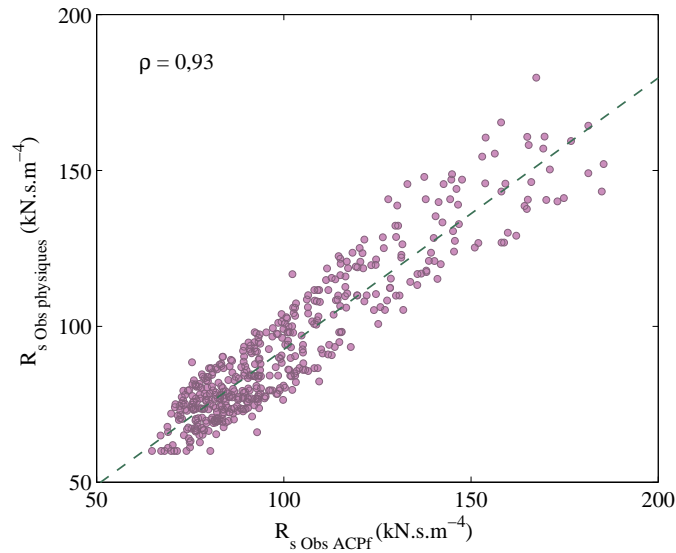


FIG. 3.12 – Nuage de corrélation des valeurs de résistance spécifique au passage de l’air R_s estimées à partir des observables physiques et des coefficients d’ACPF pour le modèle de Delany et Bazley.

La corrélation entre les paramètres R_s atteint une valeur égale à 0,93. Ce nuage de corrélation rend compte d’un léger “offset” entre les valeurs ; les valeurs de R_s sont globalement plus grandes dans le cas des observables définies par ACPf. Cet offset s’élève à environ 10 kN.s.m^{-4} , ce qui reste de l’ordre de grandeur de la précision expérimentale. De plus, il est intéressant d’observer l’évolution dans le temps de la résistance spécifique au passage de l’air estimée par les deux types de définition d’observables (figure 3.13).

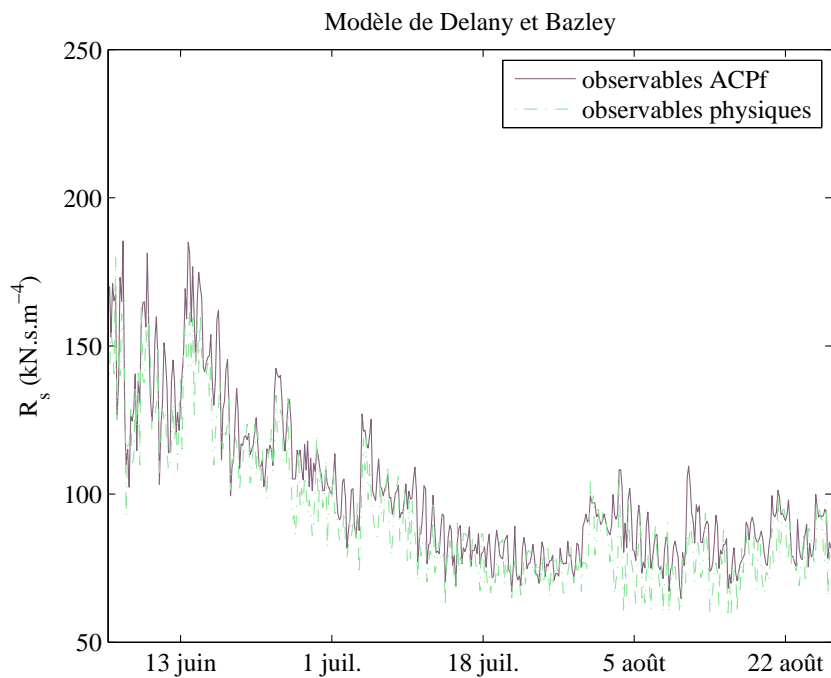


FIG. 3.13 – Évolution des valeurs de résistance spécifique au passage de l’air R_s estimées à partir des observables physiques et des coefficients d’ACPF pour le modèle de Delany et Bazley.

L'offset entre les valeurs estimées de R_s est bien présent en général mais pas systématique. La variation temporelle du paramètre est la même dans les deux cas considérés ; la variation jour/nuit est bien marquée ainsi que la tendance générale de l'évolution de R_s sur l'ensemble de la campagne expérimentale (les deux grandes périodes de la campagne se démarquent clairement).

L'évolution des valeurs du paramètre d'épaisseur e est difficilement interprétable ; celle-ci présente un aspect aléatoire qui rend illisible les tendances de ce paramètre dans le temps. Il est cependant intéressant d'observer le nuage de corrélation entre l'estimation à partir des coefficients d'ACPF et les observables physiques (figure 3.14).

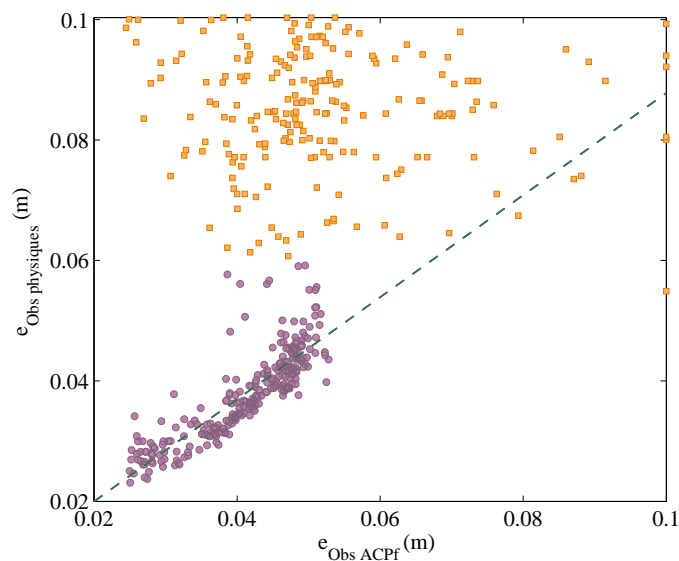


FIG. 3.14 – Nuage de corrélation des valeurs d'épaisseur e estimées à partir des observables physiques et des coefficients d'ACPF pour le modèle de Delany et Bazley. Le coefficient de corrélation sur l'ensemble des données s'élève à $\rho = 0,52$ et sur les valeurs inférieures à $0,06$ à $\rho = 0,84$.

Ce nuage de corrélation montre une bonne cohérence entre les valeurs d'épaisseur de $0,01$ à $0,05$ m. Or, au-delà de $0,05$ m, on sait qu'une augmentation de l'épaisseur e n'a plus d'influence sur le résultat ce qui explique l'étalement du nuage de corrélation. Cette observation permet de vérifier que les estimations à partir du modèle semi-empirique de Delany et Bazley sont robustes et cohérentes avec la physique des phénomènes.

2.3.1.2 Effet sur les résidus

A titre indicatif, on peut également observer les nuages de corrélation entre les résidus (différence entre observables expérimentales et observables recalées) pour chaque définition d'observables (figures 3.15 et 3.16). On vérifie ainsi qu'après calibration des spectres, les résidus relatifs aux observables sont bien non-corrélés. Cette observation est également valable pour les résidus relatifs aux observables physiques qui sont fortement corrélées notamment les observables 1 et 3 (figure 3.8). La corrélation entre les observables n'implique donc pas nécessairement une corrélation entre les résidus.

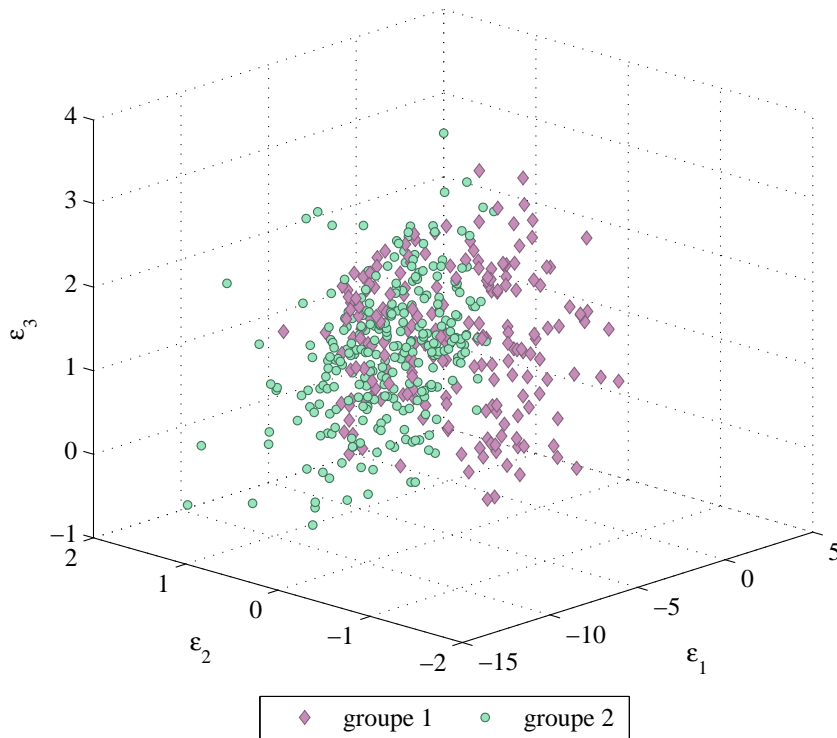


FIG. 3.15 – Nuage de corrélation entre les résidus calculés après calibration spectre par spectre sur les observables physiques pour le modèle de Delany et Bazley. Les résidus entre les observables sont bien indépendants ($\rho_{\varepsilon_1\varepsilon_2} = -0,3$, $\rho_{\varepsilon_1\varepsilon_3} = 0,14$ et $\rho_{\varepsilon_2\varepsilon_3} = -0,14$).

2.3.2 Modèle de Zwikker et Kosten

2.3.2.1 Estimation et interprétation des paramètres

Dans le cas du modèle de Zwikker et Kosten, le recours aux observables physiques est plus délicat car ces observables avaient été initialement définies en se basant sur le modèle de Delany et Bazley. Les zones d'importance sont cependant bien mises en exergue et cette définition d'observables devrait pouvoir s'adapter à d'autres modèles de sol. Les résultats présentés dans le tableau 3.2 le confirment.

Paramètres incertains	Obs physiques		Coefficients ACPf	
	$\langle\theta\rangle$	σ_θ	$\langle\theta\rangle$	σ_θ
R_s	87	32	94	35
e	0,06	0,02	0,07	0,02
Ω	0,95	0,06	0,95	0,07
C_s	1,0	1,8	0,2	0,8

TAB. 3.2 – Moyenne des paramètres estimés spectre par spectre avec le modèle de Zwikker et Kosten et écarts-type associés sur la base de données complète.

En effet, les valeurs moyennes des paramètres estimés par les deux méthodes de définition d'observables sont comparables compte-tenu des écarts-type. Seule la constante de structure présente des résultats relativement distincts; les écarts-type associés à ce paramètre étant dans les deux cas élevés, il est peu aisé de comparer ces paramètres moyens. Certaines dépendances entre les paramètres ont également une influence non négligeable sur l'estimation de l'ensemble des paramètres d'entrée du

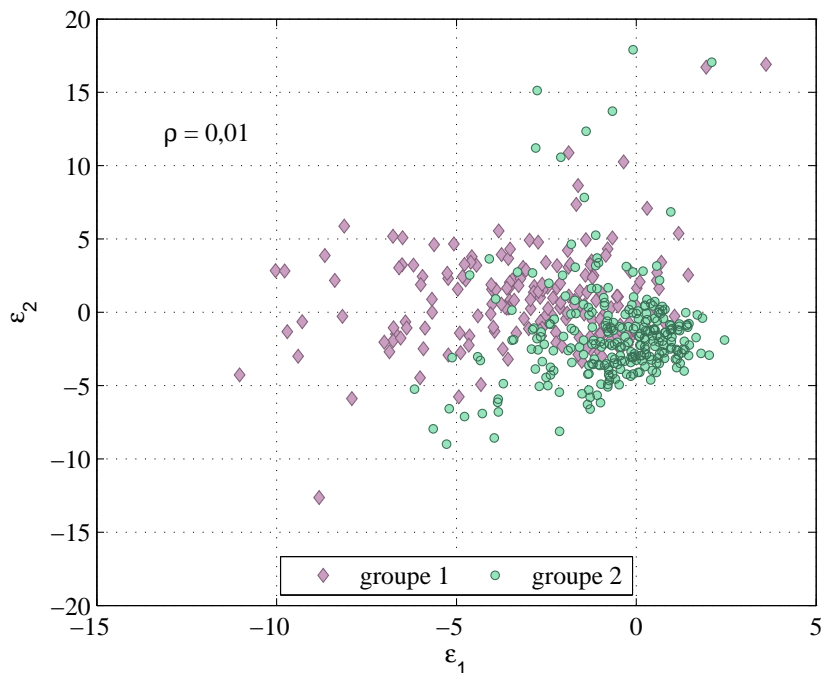


FIG. 3.16 – Nuage de corrélation entre les résidus calculés après calibration spectre par spectre sur les observables définies par ACPf pour le modèle de Delany et Bazley. Les résidus entre les observables sont bien indépendants.

modèle et sur l'estimation de la constante de structure C_s . A noter que ce paramètre est de moindre importance dans le modèle en regard de la résistance spécifique au passage de l'air, de l'épaisseur ou de la porosité. Ce modèle microstructural étant initialement adapté pour décrire des matériaux poreux, il n'est pas surprenant d'observer que ce paramètre soit difficile à interpréter d'un point de vue physique. En effet, on s'aperçoit figure 3.17, que pour la configuration définie en figure 3.1, pour des valeurs fixées de résistance spécifique au passage de l'air, d'épaisseur et de porosité, le modèle est très peu sensible à la constante de structure C_s . Cette faible sensibilité se traduit également sur les observables, ce qui explique pourquoi les estimations associées à C_s sont difficilement interprétables.

L'observation de l'évolution de la résistance spécifique au passage de l'air dans le temps (figure 3.18) permet de mettre en évidence des évolutions temporelles du paramètre R_s semblables, avec cependant de légères différences en amplitude entre les estimations. Une fois de plus, le modèle statistique défini à partir des coefficients de l'ACPf estime des valeurs légèrement supérieures à celui qui est défini à partir des observables physiques. La variation en amplitude globale du paramètre est bien supérieure dans le cas du modèle de Zwikker et Kosten que dans le cas du modèle de Delany et Bazley : ces amplitudes s'élèvent respectivement à 200 et 120 kN.s.m^{-4} . L'utilisation du modèle de Zwikker et Kosten rend compte d'une plus grande sensibilité aux conditions météorologiques et aux précipitations à travers le paramètre R_s . Cependant, il est difficile d'interpréter les valeurs des trois autres paramètres du modèle. En effet, la relation de l'impédance définie par l'équation 1.6 au chapitre 1 montre que des effets compensatoires peuvent apparaître entre la constante de structure C_s et la porosité Ω par exemple. Ces paramètres n'agissent donc pas de manière indépendante sur la valeur d'impédance acoustique du sol Z . De telles interactions se retrouvent également dans le cas de matériaux granulaires où le paramètre d'épaisseur et la tortuosité sont fortement liés. Une augmentation de la tortuosité tend à réduire l'épaisseur et inversement.

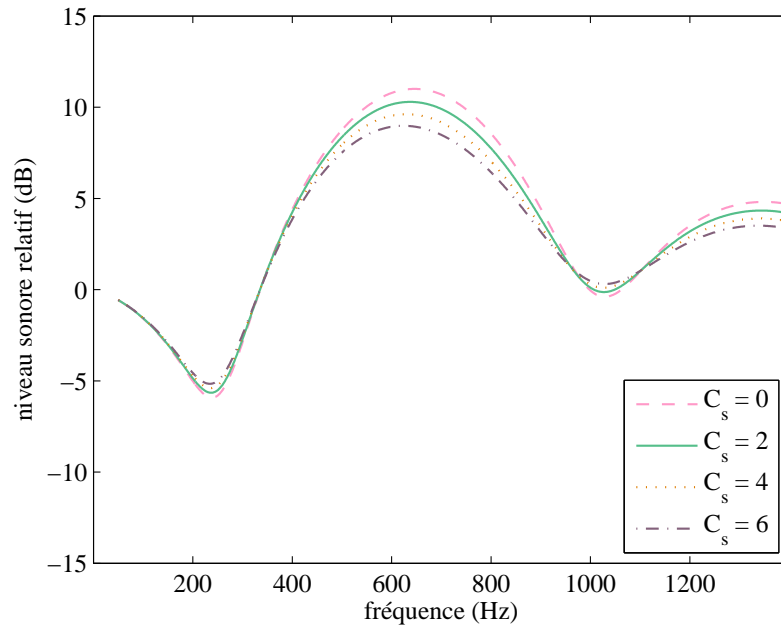


FIG. 3.17 – Sensibilité du modèle de Zwicker et Kosten à la constante de structure C_s avec $R_s = 90 \text{ kN.s.m}^{-4}$, $e = 0,06 \text{ m}$ et $\Omega = 0,95$ (correspondant aux valeurs moyennes estimées des paramètres).

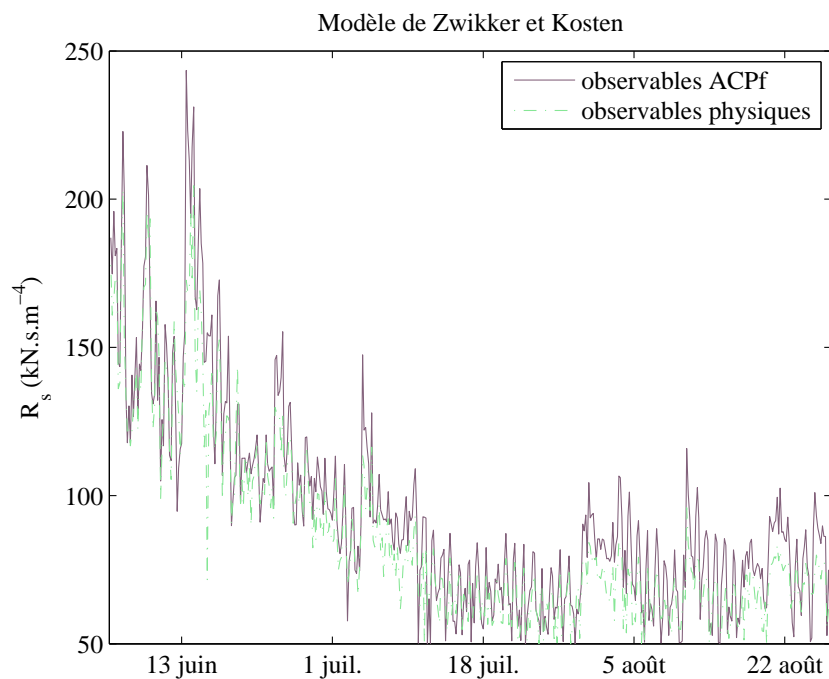


FIG. 3.18 – Évolution des valeurs de résistance spécifique au passage de l'air R_s estimées en se basant sur les observables physiques et les coefficients d'ACPF pour le modèle de Zwicker et Kosten.

2.3.2.2 Effet sur les résidus

Les résidus calculés après le calage des spectres à l'aide du modèle de Zwicker et Kosten sont également non-corrélés dans le cas des observables définis par ACPf ($\rho_{\varepsilon_1\varepsilon_2} = 0,34$) mais une corrélation "forte" est observée dans le cas des observables physiques. En effet, une corrélation de l'ordre de 0,7 entre les résidus relatifs aux observables 2 et 3. Les observables physiques doivent être utilisées avec précaution lorsqu'on considère un autre modèle que le modèle de Delany et Bazley.

2.3.3 Critères statistiques

Afin de compléter cette étude comparative et de valider les précédentes observations, un certain nombre de critères statistiques (R^2 , AIC, BIC), définis en section 2.3 du chapitre 2, sont calculés pour chaque modèle et chaque définition d'observables (tableau 3.3).

	Modèle de Delany et Bazley		Modèle de Zwicker et Kosten	
	Obs physiques	Coefficients ACPf	Obs physiques	Coefficients ACPf
R^2	0,95	0,99	0,92	0,93
AIC	9	11	12	15
BIC	13	16	22	24

TAB. 3.3 – Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten, avec des choix d'observables distincts.

Ces critères sont à considérer tout en ayant conscience que le nombre d'observables et/ou que le nombre de paramètres peuvent être différents selon le modèle considéré. Le coefficient de détermination R^2 , qui traduit l'adéquation entre les données modélisées et mesurées, est plus grand pour les observables définies par ACPf que pour les observables définies physiquement, que ce soit avec le modèle de Delany et Bazley ou le modèle de Zwicker et Kosten. L'estimation basée sur les observables définies par ACPf permet donc d'expliquer davantage de variance que celle basée sur les observables physiques. Les critères AIC et BIC, qui permettent de comparer la qualité de différents modèles (ici au sens du modèle statistique) avec une plus forte pénalisation pour le sur-paramétrage pour BIC, suivent la tendance inverse. Les écarts entre les critères sont cependant assez faibles et peuvent être aisément expliqués par le nombre d'observables en jeu : trois dans le cas physique et seulement deux dans le cas de l'ACPf. Malgré le faible nombre d'observables, la méthode par ACPf permet d'estimer les paramètres avec robustesse, comme en témoignent les valeurs des coefficients de détermination. Ces tendances se retrouvent aussi bien pour le modèle de Delany et Bazley que pour le modèle de Zwicker et Kosten.

2.4 BILAN

Les méthodes de définition d'observables proposées semblent donc présenter un intérêt d'un point de vue statistique et seront exploitées dans un cadre plus complexe en considérant les groupes de données définis à la section 1.1.1. Cette première étude comparative a également permis de montrer que la définition des observables basée sur la méthode d'analyse en composantes principales fonctionnelles était aussi valable qu'une définition d'observables basée sur des critères physiques et qu'elle assure de plus l'indépendance des observables. Cette méthode est facilement généralisable et pourrait permettre à l'avenir de faciliter la définition des observables.

3 CALIBRATION DES MODÈLES D'IMPÉDANCE DE SOL

A cette étape de l'étude, les spectres ne sont plus considérés indépendamment mais par groupes de spectres. Les paramètres estimés sont alors représentatifs d'une situation observée pendant une période donnée et non plus d'un événement ponctuel. L'étude en est alors statistiquement plus robuste car elle permet la prise en compte d'une grande quantité d'informations pour une situation donnée. Procéder à une calibration spectre par spectre revient à considérer que chaque spectre représente parfaitement une situation donnée sur une période temporelle donnée. Or nous savons pertinemment qu'une telle hypothèse est erronée du fait de la variabilité non-négligeable des caractéristiques de sol et donc de ces spectres d'atténuation relative (chapitre 1 section 1.3.3). Il est donc plus pertinent de prendre en compte un groupe de spectres rendant compte de la variabilité d'une situation caractéristique. Nous étudions ici les mêmes modèles de sol que pour l'étude précédente. L'écriture de la fonction coût dépend du nombre de paramètres à estimer et du nombre d'observables définissant le système d'étude (chapitre 2 section 3.3). Les deux méthodes de définition d'observables seront étudiées.

3.1 DÉFINITION DE LA FONCTION COÛT

Le problème à résoudre présente différents *niveaux de complexité* :

- la *non-linéarité* ;
- la *multidimensionnalité* : deux ou quatre paramètres physiques incertains et leurs écarts-type associés doivent être estimés à partir d'un jeu de données de référence conséquent ;
- l'*hétéroscédasticité* : une étude préliminaire a permis d'observer que les variances de résidus étaient variables sur la dimension des données de sortie.

La fonction coût est une fonction de log-vraisemblance définie au chapitre 2 par l'équation 2.5. Dans le cas des observables physiques, $n_{obs} = 3$ et dans le cas des observables définies par ACPf, $n_{obs} = 2$.

3.2 CALIBRATION AVEC LE MODÈLE DE DELANY ET BAZLEY

Les paramètres incertains sont estimés en considérant un groupe composé de tous les spectres et de deux groupes définis en section 1.1.1. Les paramètres physiques incertains R_s (résistance spécifique au passage de l'air) et e (épaisseur) sont estimés par maximisation de la fonction de log-vraisemblance. La variance d'estimation des paramètres est obtenue par la méthode de bootstrap détaillée au chapitre 2 section 5.1.2.

3.2.1 Observables physiques

Les résultats sont synthétisés dans le tableau 3.4.

	<i>tous les spectres</i>	<i>groupe (1)</i>	<i>groupe (2)</i>
R_s	92,1	109	82,5
e	0,044	0,033	0,052

TAB. 3.4 – Paramètres estimés à partir des "observables physiques" avec le modèle de Delany et Bazley correspondant à la base de données complète et aux deux groupes de spectres.

Ces résultats montrent une différence importante entre les valeurs moyennes (notées $\langle \hat{\theta} \rangle$) de résistance spécifique au passage de l'air entre le groupe (1) et le groupe (2). Cet écart atteint presque 30 kN.s.m^{-4} , ce qui se situe au-delà des ordres de grandeurs liés à l'erreur de mesures ($\pm 10 \text{ kN.s.m}^{-4}$). Cette différence significative sur R_s rend donc bien compte des phénomènes micrométéorologiques rencontrés au cours de la campagne expérimentale et par suite de la variabilité temporelle liée à ce paramètre. Ces résultats rendent compte d'une première période de campagne pluvieuse et d'une

deuxième période relativement sèche, ce qui implique des valeurs plus fortes de R_s en début de campagne. Une telle différence sur de faibles valeurs de résistance spécifique au passage de l'air a un effet non-négligeable en termes de niveau sonore (cf. figure 1.3). A noter que le paramètre R_s estimé sur l'ensemble des spectres correspond à la moyenne des paramètres pour les groupes (1) et (2). Il est de plus comparable à la valeur moyenne de R_s estimée spectre par spectre (tableau 3.1).

Les tendances sont moins marquées pour l'effet d'épaisseur. Le paramètre e estimé pour l'ensemble des spectres correspond à la moyenne des paramètres pour les groupes (1) et (2). On vérifie que pour une augmentation de R_s , une diminution de e est observée et inversement.

La mise en place du processus de bootstrap (chapitre 2 section 5.1.2) fournit les écarts-type d'estimation associés aux paramètres incertains. La méthode de bootstrap est basée sur la reconstitution de la base de données initiale en générant une nouvelle base de données à partir de la moyenne et de l'écart-type de chacune des observables. Des jeux de données vont ainsi être générés 50 fois et le processus d'estimation de paramètres sera ainsi renouvelé pour chaque nouveau jeu de données. Afin que le processus de bootstrap ne soit pas trop chronophage, seuls des petits groupes de spectres seront générés (20 nouveaux spectres seront générés à chaque itération). Les résultats sont répertoriés dans le tableau 3.5 et sont représentés figure 3.19 à l'aide de "boîtes à moustaches"³.

Paramètres incertains	<i>tous les spectres</i>		<i>groupe (1)</i>		<i>groupe (2)</i>	
	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \theta \rangle}$
R_s	94	7	117	10	81	3
e	0,05	0,01	0,04	0,01	0,05	0,01

TAB. 3.5 – Paramètres estimés par la méthode de bootstrap à partir des "observables physiques" avec le modèle de Delany et Bazley et écarts-type associés correspondant à la base de données complète et aux deux groupes de spectres.

Les écarts-type témoignent de la variabilité propre à chaque groupe de spectres : l'écart-type sur R_s pour le groupe (1) est trois fois supérieur à l'écart-type pour le groupe (2). La résistance au passage de l'air moyenne associée à "tous les spectres" et représentative de l'intégralité de la campagne expérimentale constitue bien la moyenne des paramètres pour chacun des groupes. De plus, les valeurs moyennes des paramètres estimées par la méthode de bootstrap sont comparables aux valeurs estimées préalablement (tableau 3.4). L'écart-type sur le paramètre d'épaisseur est constant ce qui traduit un bon recalage quelque soit le groupe de données considéré. Les valeurs moyennes du paramètre d'épaisseur témoignent de l'évolution de e au cours de la campagne expérimentale ; cette tendance est illustrée grâce à la représentation des résultats par les boîtes à moustaches (figure 3.19). La méthode de bootstrap permet donc d'accéder à des écarts-type d'estimation sur R_s et e dont les ordres de grandeur sont réalistes. On constate cependant que ces valeurs sont plus faibles que celles calculées dans l'étude spectre par spectre (tableau 3.1). Cette valeur ne prétend pas traduire l'intégralité de la dispersion des spectres et elle est davantage associée à l'estimation des paramètres θ . Cette valeur rend compte de la variation qu'il est possible d'expliquer avec le modèle considéré.

3. La boîte contient 50 % des données. Le bord inférieur de la boîte indique le 75ème percentile de l'ensemble des données et le bord supérieur indique le 25ème percentile. La ligne dans cette zone indique la valeur médiane des données. Si cette ligne n'est pas à égale distance des bords de la boîte, les résultats sont biaisés. Les extrémités des lignes verticales ou "moustaches" indiquent les valeurs maximales ou minimales des données, à moins que des valeurs aberrantes ne soient présentes. Les points notés + se situant à l'extérieur des moustaches correspondent à ces valeurs aberrantes.

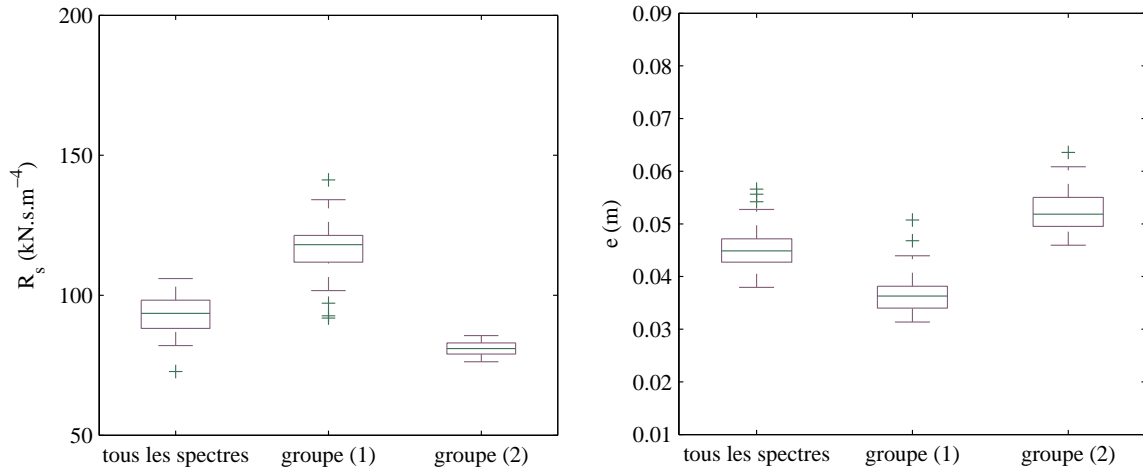


FIG. 3.19 – Boîtes à moustaches des paramètres incertains du modèle de Delany et Bazley pour les trois configurations étudiées (observables physiques) obtenus par le processus de bootstrap. Les + correspondent aux valeurs extrêmes de la distribution.

Une autre étape importante est la vérification des hypothèses faites pour mener l'étude. Cette étape consiste à vérifier les propriétés des résidus (ε_l où l est l'indice de l'observable) à savoir leur *normalité* et leur *indépendance*. Les résidus correspondent à la différence entre les observables mesurées et les observables calculées. Les propriétés des résidus sont synthétisées par le tableau 3.6 indiquant :

- la *p-value* qui permet de vérifier la normalité des résidus (test de Kolmogorov-Smirnov présenté au chapitre 2 section 4.3) Une *p-value* inférieure à 0,1 indique que la distribution est gaussienne ;
- les coefficients de corrélation entre les résidus.

	<i>p-value</i>		<i>coefficient de corrélation</i>			
	ε_1	ε_2	ε_3	$\rho_{\varepsilon_1\varepsilon_2}$	$\rho_{\varepsilon_1\varepsilon_3}$	$\rho_{\varepsilon_2\varepsilon_3}$
<i>tous les spectres</i>	0,2	0,1	0,02	-0,7	-0,7	0,2
<i>groupe (1)</i>	0,1	0,1	0,06	-0,5	-0,8	-0,02
<i>groupe (2)</i>	0,1	0,05	0,05	0,01	-0,7	-0,5

TAB. 3.6 – Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Delany et Bazley avec les observables physiques.

Ces propriétés sont également visibles sur les histogrammes et les nuages de corrélation sur l'ensemble des spectres (figure 3.20), ainsi que pour le groupe (1) (figure 3.21) et le groupe (2) (figure 3.22).

Lorsqu'on calibre le modèle sur l'ensemble des spectres (figure 3.20), les résidus présentent globalement un caractère gaussien à part pour les résidus (ε_1) associés à la première observable (*i.e.* position du premier creux d'interférence) ; la distribution de ε_1 présente en effet une tendance bimodale correspondant aux deux périodes. La *p-value* associée montre une valeur plus élevée que dans les autres cas ($> 0,1$), ce qui traduit cette particularité. Ce caractère bimodal disparaît lorsque la base de données est découpée en deux groupes⁴. De fortes corrélations sont également observées entre les résidus ε_1 et ε_2 , ainsi qu'entre les résidus ε_1 et ε_3 . Cette corrélation est réduite entre ε_1 et ε_2

4. Ces remarques nous laissent à penser qu'on pourrait également envisager d'utiliser l'observation des distributions des résidus pour procéder au découpage de la base de données et ainsi former des groupes dans le cadre d'une étude préliminaire.

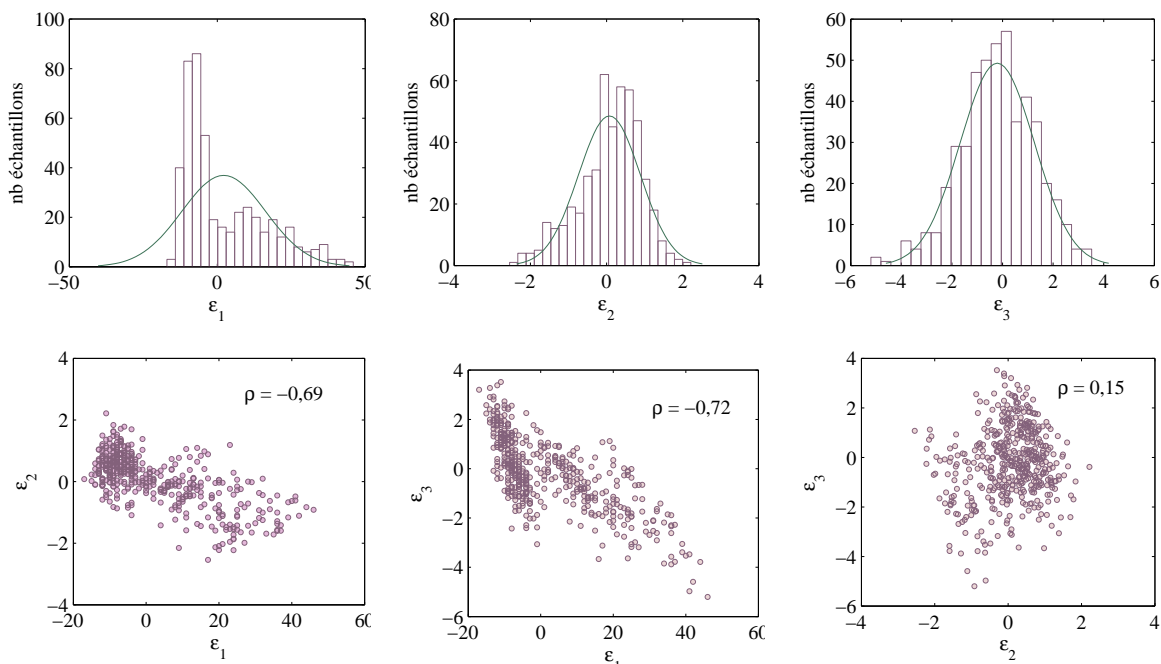


FIG. 3.20 – Histogrammes et nuages des corrélations des résidus calculés sur l'ensemble des spectres.

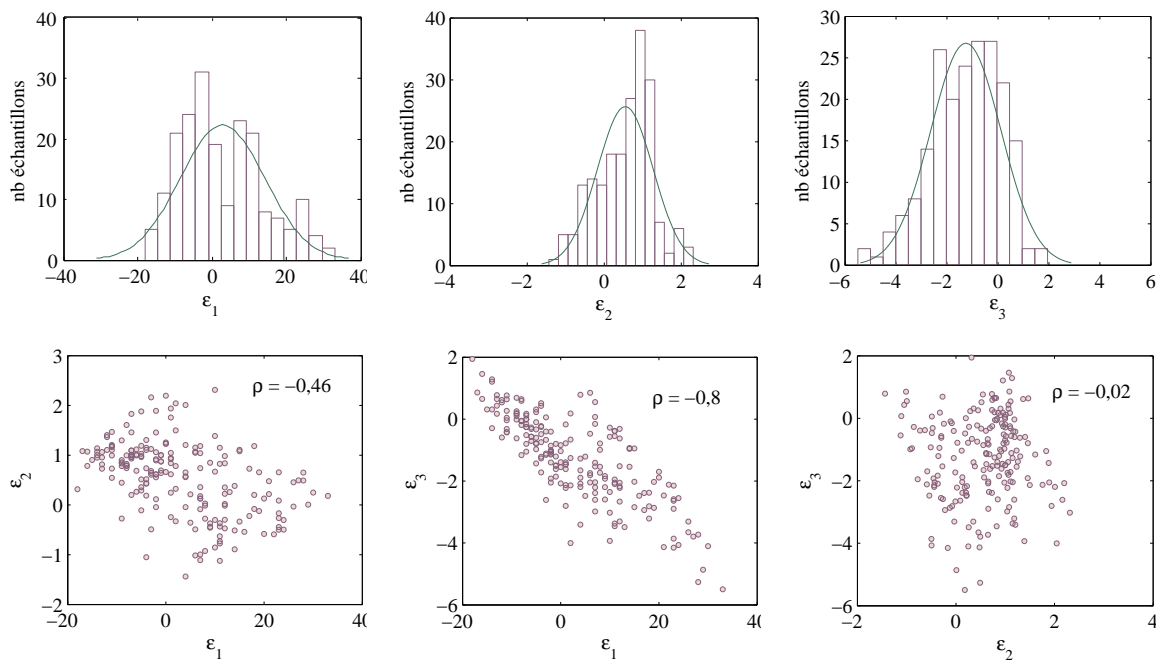


FIG. 3.21 – Histogrammes et nuages des corrélations des résidus calculés pour le groupe (1).

lorsque la base de données est scindée en deux groupes mais elle persiste entre ε_1 et ε_3 . L'observable Y_3 semble donc plus sensible que l'observable Y_2 aux variations de Y_1 *i.e.* à la position fréquentielle du creux d'effet de sol. Ces différents critères illustrent l'intérêt certain du découpage de la base de données expérimentale en deux parties avec une amélioration de la forme des distributions des résidus (caractère gaussien renforcé) et une diminution globale mais pas systématique de la corrélation entre les résidus. Ce découpage permet de plus de vérifier que l'estimation des paramètres est cohérente avec la physique des phénomènes. Le contenu prédictif des modèles en est alors renforcé. La persistance

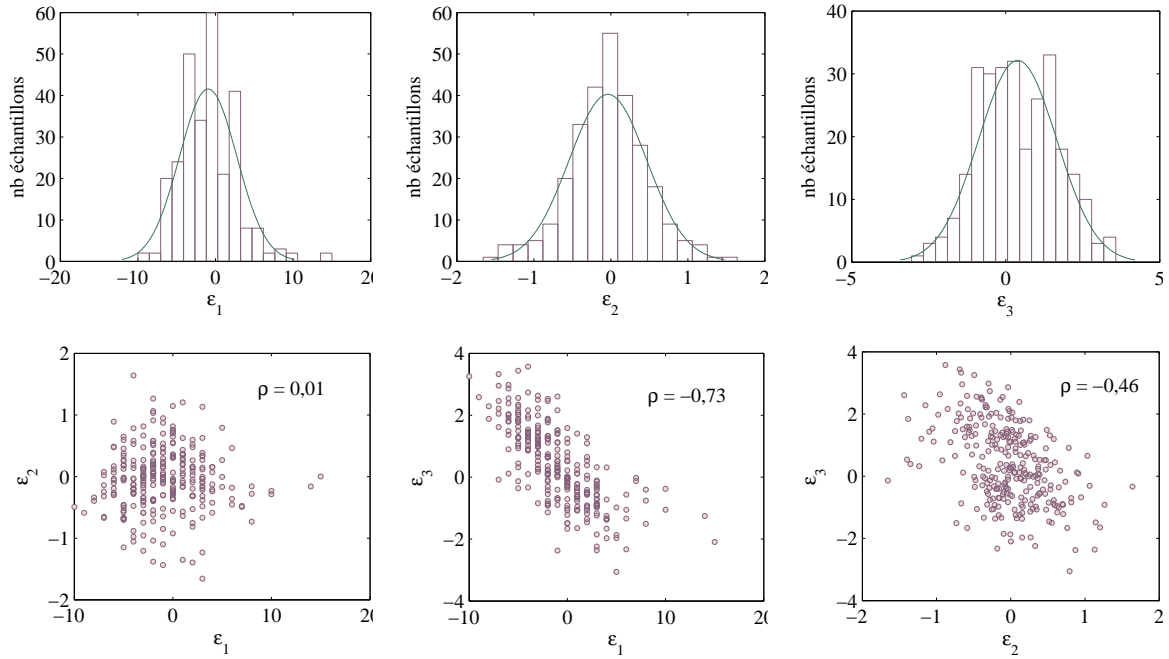


FIG. 3.22 – Histogrammes et nuages des corrélations des résidus calculés pour le groupe (2).

des corrélations entre les résidus montre cependant que le modèle statistique est parfait, par un autre choix d'observables physiques par exemple.

3.2.2 Observables définies par ACPf

La même étude est menée dans le cas d'observables définies par la méthode d'ACPf (section 2.2). Les tendances générales sur les paramètres estimés (tableau 3.7) sont les mêmes que dans le cas des observables dites physiques. Les valeurs moyennes de R_s sont légèrement supérieures à celles données au tableau 3.4 et notamment pour le groupe (1) pour lequel les spectres présentent la plus forte variabilité. Les valeurs du paramètre d'épaisseur sont identiques aux valeurs estimées à partir des observables physiques.

Paramètres	tous les spectres	groupe (1)	groupe (2)
R_s	99	124	84
e	0,0432	0,035	0,05

TAB. 3.7 – Paramètres estimés à partir des coefficients d'ACPf avec le modèle de Delany et Bazley correspondant à la base de données complète et aux deux groupes de spectres.

La méthode de bootstrap permet une fois encore d'avoir accès aux valeurs des écarts-type d'estimation réalistes d'un point de vue physique. Les résultats sont référencés dans le tableau 3.8 et représentés par les boîtes à moustaches (figure 3.23).

Paramètres incertains	tous les spectres		groupe (1)		groupe (2)	
	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$
R_s	97	13	120	16	85	2
e	0,04	0,01	0,03	0,01	0,05	0,01

TAB. 3.8 – Paramètres estimés à partir des observables définies par les coefficients d'ACPF avec le modèle de Delany et Bazley et écarts-type d'estimation associés obtenus par le processus de bootstrap correspondant à la base de données complète et aux deux groupes de spectres.

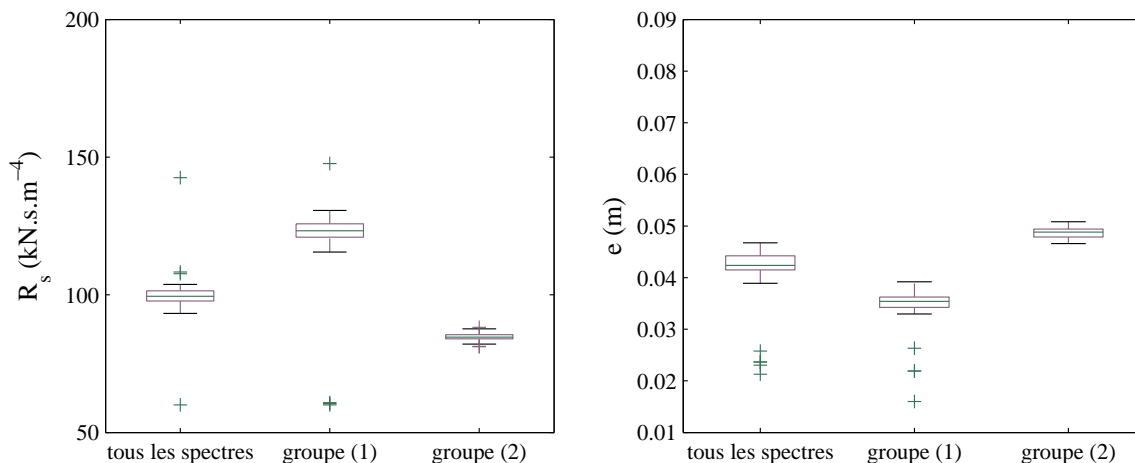


FIG. 3.23 – Boîtes à moustaches des paramètres incertains du modèle de Delany et Bazley pour les trois configurations étudiées (fonctions propres de l'ACPF) obtenues par le processus de bootstrap. Les + correspondent aux valeurs extrêmes de la distribution.

Les paramètres d'application du bootstrap sont similaires au cas précédent : 50 jeux de 20 spectres sont générés à partir des moyennes et écarts-type de la base de données initiale. Les résultats montrent une convergence des valeurs moyennes des paramètres vers les paramètres estimés précédemment (tableau 3.7). Les écarts-type associés à e sont constants en fonction des groupes et sont similaires à l'estimation faite dans le cas d'observables physiques. Cependant, les écarts-type liés à la résistance spécifique au passage de l'air augmentent considérablement dans le cas où tous les spectres sont considérés et dans le cas du groupe (1) : l'augmentation de $\sigma_{\langle \hat{\theta} \rangle}$ est de 6 kN.s.m^{-4} pour chacun. La valeur de ces écarts-type est notamment dû à quelques événements extrêmes intervenus lors de la procédure de bootstrap comme observé sur les boîtes à moustaches (figure 3.23). Ces événements ne sont pas pour autant à marginaliser et à mettre de côté : ils représentent la variabilité qui peut être rencontrée au cours d'une telle campagne expérimentale.

Les propriétés des résidus sont vérifiées afin de valider l'intérêt du découpage de la base de données, ainsi que la démarche globale : le tableau 3.9 rend compte du caractère gaussien des résidus et de leur éventuelle corrélation.

	<i>p-value</i>		<i>coefficient de corrélation</i>
	ϵ_1	ϵ_2	$\rho_{\epsilon_1\epsilon_2}$
<i>tous les spectres</i>	0,09	0,06	0,01
<i>groupe (1)</i>	0,06	0,04	0,02
<i>groupe (2)</i>	0,07	0,07	0,01

TAB. 3.9 – *Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Delany et Bazley avec les coefficients de l'ACPF.*

D'après les valeurs obtenues, chaque distribution de résidus est bien gaussienne (nous avons également vérifié par observation des histogrammes que ces résidus sont bien centrés en 0). Le découpage de la base de données initiale en deux groupes permet donc d'améliorer les propriétés gaussiennes des résidus. La corrélation entre les résidus pour chaque groupe de spectres considéré est presque inexistante ; les coefficients de corrélation s'élèvent en effet entre 0,01 et 0,02. Les résidus sont indépendants parce qu'ils sont directement liés aux coefficients de l'ACPF, ce qui est cohérent avec leurs propriétés.

3.2.3 Bilan

Les estimations effectuées par le biais des deux définitions d'observables sont équivalentes d'un point de vue quantitatif. La définition des observables basée sur les coefficients d'ACPF permet de vérifier les hypothèses de départ faites sur les résidus, ces hypothèses n'étant pas systématiquement vérifiées dans le cas des observables physiques. Une amélioration est apportée en scindant la base de données expérimentale en deux groupes. Grâce à ce découpage, il est ainsi possible d'identifier différentes caractéristiques de sol. Les écarts-type calculés par bootstrap traduisent la variabilité physique des données. Ce type d'écart-type peut être utilisé pour prévoir un jeu d'observables (par des méthodes de Monte Carlo par exemple).

Dans une démarche prédictive, on pourra donc prévoir la position du creux Y_1 , les observables Y_2 et Y_3 définies respectivement par les équations 3.1 et 3.2 ne présentent pas un intérêt majeur car le lien avec le spectre complet est difficile. Cependant, cette limite est sans conséquence car cette étude a pour but d'obtenir des informations statistiques sur les paramètres relatifs à la caractérisation acoustique du sol qui sont dans ce cas les paramètres d'entrée du modèle. Dans le cas des observables définies par ACPF, on peut facilement prévoir un jeu de spectres en calculant les observables (coefficients d'ACPF) et en re-synthétisant des spectres approchés par des fonctions de base.

3.3 CALIBRATION AVEC LE MODÈLE DE ZWIKKER ET KOSTEN

Le processus de calibration est appliqué ici au modèle de Zwicker et Kosten. L'étude est à nouveau menée à partir des deux définitions d'observables. Les écarts-type d'estimation sont calculés par la méthode de bootstrap.

3.3.1 Observables physiques

Les résultats de la calibration sont synthétisés dans le tableau 3.10. Celui-ci regroupe les résultats de la maximisation de la fonction de log-vraisemblance sur les groupes complets ainsi que les résultats issus de la mise en œuvre de la méthode de bootstrap. Ces résultats sont également représentés par les boîtes à moustaches (figure 3.24).

Paramètres incertains	tous les spectres			groupe (1)			groupe (2)		
	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$
R_s	99	101	14	127	125	29	76	75	6
e	0,06	0,06	0,02	0,06	0,06	0,02	0,06	0,06	0,02
Ω	0,95	0,97	0,03	0,94	0,94	0,08	0,96	0,96	0,05
C_s	1,6	1,6	1,6	1,5	1,4	1,6	1,7	1,7	1,9

TAB. 3.10 – Paramètres estimés à partir des observables physiques avec le modèle de Zwikker et Kosten. Les paramètres moyens et écarts-type associés sont obtenus par le processus de bootstrap correspondant à la base de données complète et aux deux groupes de spectres.

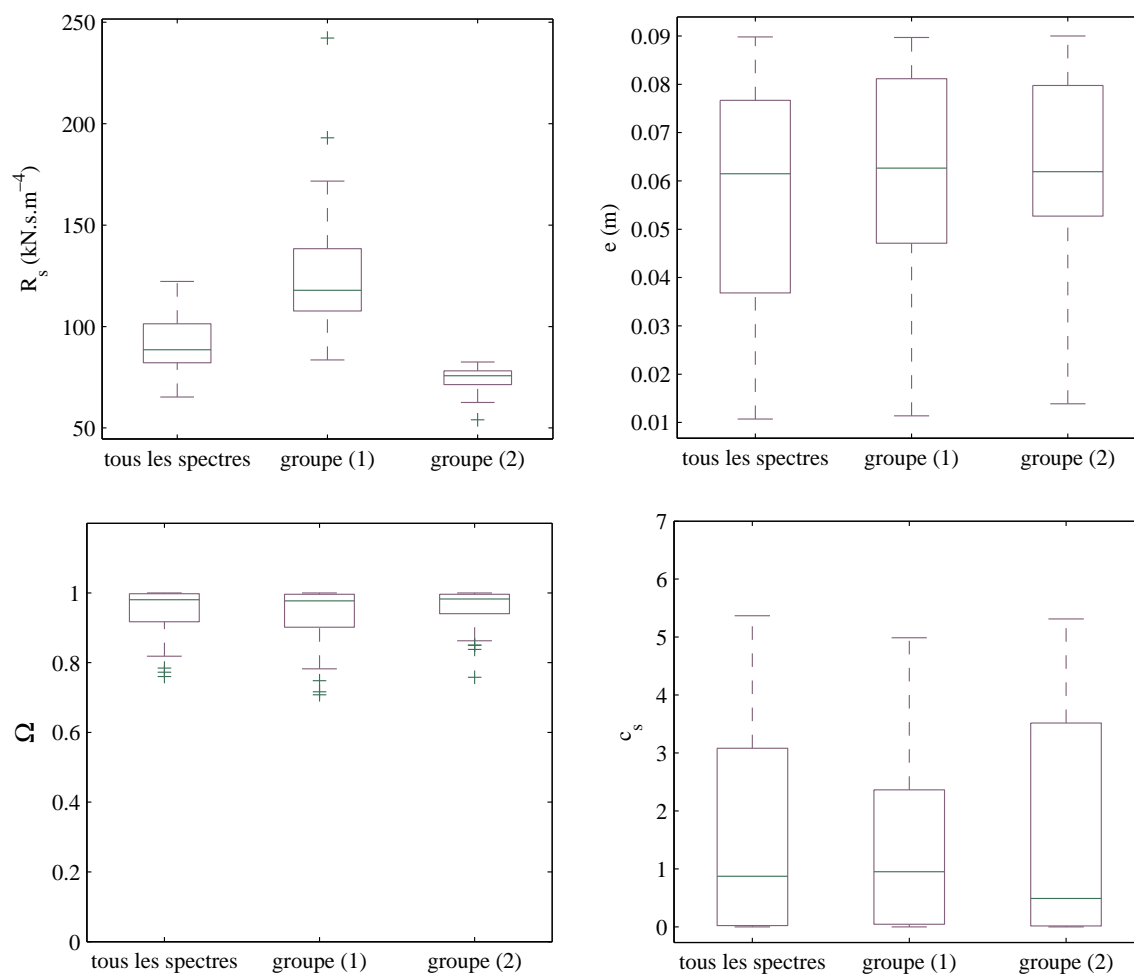


FIG. 3.24 – Boîtes à moustaches des paramètres incertains du modèle de Zwikker et Kosten pour les trois configurations étudiées (observables physiques).

Les valeurs de résistance spécifique au passage de l'air montrent une tendance déjà rencontrée dans le cas où l'étude est menée à l'aide du modèle de Delany et Bazley : valeur plus élevée pour le groupe (1) que pour le groupe (2) et valeur intermédiaire dans le cas où tous les spectres sont pris en compte. Aucune tendance marquée ne se dégage à partir des autres paramètres d'étude, à savoir l'épaisseur e , la porosité Ω et la constante de structure C_s . Le modèle et les observables définies sont moins sensibles à ces paramètres que le paramètre R_s et ne sont vraisemblablement pas indépendants. Cette forte dépendance ne permet donc pas de dégager d'évolution temporelle : pas de différences

entre les estimations du groupe (1) et du groupe (2). Ces paramètres semblent souvent atteindre les bornes fixées lors du processus de maximisation de la fonction de log-vraisemblance⁵, ce qui annule toute interprétation physique. En termes d'incertitudes sur les paramètres :

- les écarts-type de R_s présentent une tendance bien connue, liée à la variabilité physique observée au cours de la campagne (section 1.1.1), avec une erreur plus grande pour le groupe (1) que pour le groupe (2) ;
- les écarts-type de e sont constants pour chacun des groupes considérés ;
- les valeurs moyennes de porosité sont très semblables mais leurs écarts-type permettent de dégager une légère tendance avec une augmentation de l'erreur sur Ω pour le groupe (1) à forte variabilité ;
- la constante de structure est très peu variable et présente des écarts-type considérables en regard de l'ordre de grandeur du paramètre moyen.

En ce qui concerne la vérification des hypothèses initiales (tableau 3.11), les résultats sont très proches des résultats obtenus avec le modèle de Delany et Bazley : amélioration du caractère gaussien des distributions de résidus en scindant la base de données en deux groupes et diminution de la corrélation entre les résidus ε_1 et ε_2 . Les résultats ne sont pas significatifs d'un point de vue purement physique (à part pour le paramètre R_s) mais ils vérifient en partie les hypothèses initiales. Malgré des estimations de paramètres difficilement exploitables, ce modèle se comporte comme le modèle de Delany et Bazley d'un point de vue de la statistique des résidus. Une grande partie de la variation du modèle est reportée sur R_s .

	<i>p-value</i>			<i>coefficient de corrélation</i>		
	ε_1	ε_2	ε_3	$\rho_{\varepsilon_1\varepsilon_2}$	$\rho_{\varepsilon_1\varepsilon_3}$	$\rho_{\varepsilon_2\varepsilon_3}$
<i>tous les spectres</i>	0,2	0,1	0,02	-0,7	-0,7	0,2
<i>groupe (1)</i>	0,1	0,1	0,1	-0,4	-0,8	-0,02
<i>groupe (2)</i>	0,1	0,05	0,05	0,07	-0,7	-0,5

TAB. 3.11 – *Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Zwicker et Kosten avec les observables physiques.*

Cette application du processus de calibration ne paraît pas concluante non pas en termes statistiques mais du fait du faible apport des paramètres estimés autres que R_s et de leur erreur. Ces résultats peuvent également révéler que le modèle de Zwicker et Kosten n'est pas le plus adapté pour l'étude de ce type de sol (sol herbeux). De plus, certains paramètres estimés atteignent les bornes définies lors du processus d'optimisation ce qui peut témoigner de la présence de minima locaux. Ces résultats peu satisfaisants peuvent aussi être liés au fait que les observables physiques sont définies à partir du modèle de Delany et Bazley et non du modèle de Zwicker et Kosten ; la définition des observables à partir d'un modèle donné peut ne pas être transposable à d'autres modèles. Une solution alternative serait donc de définir les observables par une méthode plus générale, ne se rapportant pas à un modèle particulier, telle que la méthode de définition des observables par ACPf.

3.3.2 Observables définies par ACPf

L'étude est maintenant menée avec les observables définies par les coefficients de l'ACPF ; la définition des observables ne dépend désormais plus d'un modèle spécifique. Les estimations par

5. Les paramètres à estimer sont bornés au cours du processus d'optimisation de la fonction de vraisemblance, afin que les valeurs estimées ne soient pas aberrantes. Cependant, lorsque plusieurs paramètres estimés atteignent ces bornes, il devient difficile de leur donner une réelle interprétation physique. Ces résultats témoignent de l'impossibilité du système à converger ou bien que certains paramètres ont une influence moindre sur les données.

maximisation de la fonction de log-vraisemblance et les résultats obtenus pas le processus de bootstrap sont donnés au tableau 3.12 et à la figure 3.25. Les estimations $\langle \hat{\theta} \rangle$ issues du processus de bootstrap convergent bien vers les estimations $\hat{\theta}$. De faibles biais sont observés mais ceux-ci sont négligeables en regard des écarts-type d'estimation.

Paramètres incertains	tous les spectres			groupe (1)			groupe (2)		
	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\hat{\theta}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$
R_s	85	83	4	100	102	6	73	73	3
e	0,03	0,02	0,01	0,02	0,02	0,02	0,04	0,05	0,01
Ω	0,76	0,79	0,21	0,6	0,55	0,26	0,8	0,75	0,21
C_s	3,6	3,5	1,6	3	3,1	1,7	4,2	4,1	0,9

TAB. 3.12 – Paramètres estimés à partir des observables définies par ACPf avec le modèle de Zwikker et Kosten. Les paramètres moyens et écarts-type associés sont obtenus par le processus de bootstrap correspondant à la base de données complète et aux deux groupes de spectres.

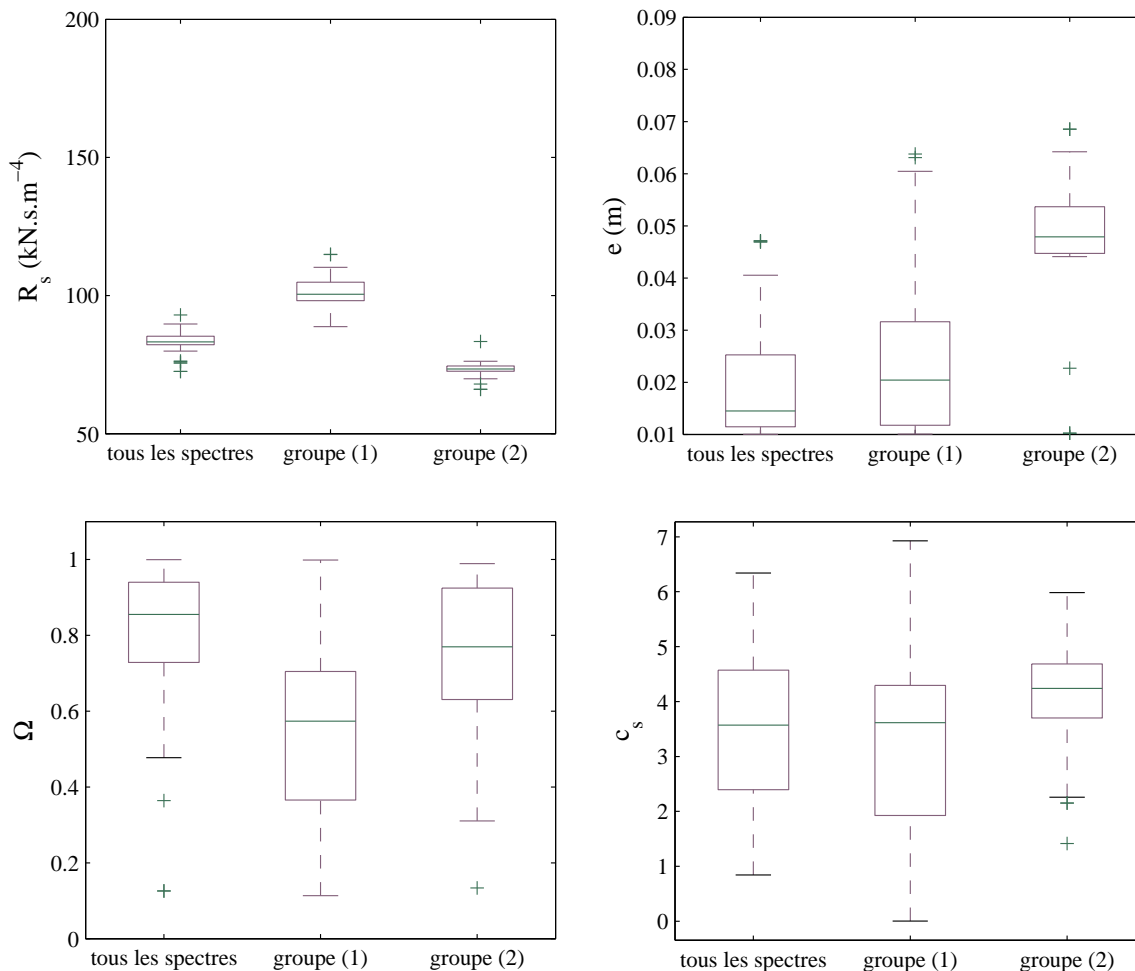


FIG. 3.25 – Boîtes à moustaches des paramètres incertains du modèle de Zwikker et Kosten pour les trois configurations étudiées (coefficients de l'ACP) obtenues par la méthode de Bootstrap.

Les estimations montrent que, dans le cas des coefficients d'ACPF pris comme observables, l'ordre de grandeur du paramètre de résistance spécifique au passage de l'air est plus faible que dans le cas des observables physiques (tableau 3.10). Les écarts-type associés à R_s ont très fortement diminué : $\sigma_{\langle \hat{\theta} \rangle}$ passe par exemple de 29 à 6 kN.s.m⁻⁴ dans le cas du groupe (1). Cette diminution des écarts-type rend compte d'une diminution de l'influence de R_s et conduit alors à une influence plus marquée des autres paramètres dont la porosité. Celle-ci a en effet tendance à augmenter pour la seconde partie de la campagne expérimentale pendant laquelle le sol est sec, ce qui est tout à fait cohérent d'un point de vue physique. Cet effet n'est pas très marqué (passage de 0,6 pour le groupe (1) à 0,8 pour le groupe (2)), mais il est révélateur de réels changements des propriétés structurales et acoustiques du sol. Le paramètre de porosité est intimement lié à la constante de structure ainsi qu'au paramètre d'épaisseur, qui présentent tous deux des tendances assez similaires au passage du groupe (1) au groupe (2). Concernant les estimations sur l'intégralité des spectres, les paramètres e et Ω ne correspondent pas à des valeurs intermédiaires de paramètres estimés pour les groupes (1) et (2). La dépendance entre les paramètres permettrait d'expliquer ce phénomène.

En termes d'erreur, seuls les écarts-type associés à R_s permettent de mettre en évidence la variabilité temporelle observée au cours de la campagne expérimentale (cf. figure 3.2) avec une dispersion forte des spectres d'atténuation relative en début de campagne et une dispersion plus réduite au cours de la seconde partie de la campagne.

La vérification des hypothèses initiales (tableau 3.13) montre que les distributions des résidus sont gaussiennes pour l'ensemble des groupes de spectres considérés, avec une valeur légèrement supérieure à 0,1 pour les résidus ε_2 du groupe (1), mais ces résultats sont tout à fait acceptables. L'indépendance entre les résidus est à nouveau vérifiée, ce qui valide une fois de plus la méthode d'ACPF quant à l'obtention de résidus indépendants.

	<i>p-value</i>		<i>coefficient de corrélation</i>
	ε_1	ε_2	$\rho_{\varepsilon_1\varepsilon_2}$
<i>tous les spectres</i>	0,09	0,07	0,01
<i>groupe (1)</i>	0,06	0,12	0,02
<i>groupe (2)</i>	0,05	0,04	0,01

TAB. 3.13 – Récapitulatif des propriétés des résidus dans le cas des paramètres estimés avec le modèle de Zwicker et Kosten avec les coefficients de l'ACPF.

3.3.3 Bilan

Malgré la forte dépendance entre les paramètres du modèle de Zwicker et Kosten, le processus de bootstrap (qui masque d'éventuelles erreurs d'estimations) permet d'estimer des paramètres expliquant les phénomènes physiques et de décrire "l'évolution" de l'état structural du sol au cours de la campagne expérimentale. La définition d'observables basées sur les coefficients d'ACPF renforce le processus de calibration. La définition d'observables physiques à partir du modèle de Delany et Bazley ne peut être transposée pour l'étude d'autres modèles d'impédance de sol. Il serait donc nécessaire de définir des observables physiques relatives à chaque modèle d'impédance de sol utilisé. Le recours au modèle de Zwicker et Kosten révèle certaines limites non négligeables dans l'estimation des paramètres. La méthode de bootstrap permet de contourner le problème en masquant d'éventuelles erreurs d'estimation.

De plus, afin d'éviter les minima locaux rencontrés au cours du processus d'optimisation de la fonction de log-vraisemblance, une approche complémentaire serait de mettre en place un processus

d'optimisation plus évolué comme par exemple un processus basé sur des algorithmes génétiques [10]. L'optimisation constitue une problématique à part entière et ne sera pas davantage développée dans ces travaux. Afin de compléter et de finaliser cette étude comparative, les résultats de validation et le choix du modèle statistique optimal sont présentés ci-après.

4 VALIDATION ET CHOIX DU MODÈLE STATISTIQUE OPTIMAL

La phase de validation est menée dans une première étape par la vérification des hypothèses initiales sur les distributions des résidus, phase déjà menée lors des études pour chaque modèle physique et chaque méthode de définition des observables. Une seconde phase consiste à calculer des critères statistiques relatifs à chaque modèle statistique considéré. Ces critères sont calculés à titre indicatif car il est délicat de comparer de tels critères relatifs à des modèles statistiques différents. Les critères tiennent compte du nombre d'observables et traduisent un compromis entre le nombre de paramètres et la qualité du recalage.

4.1 OBSERVABLES PHYSIQUES

Les critères statistiques (R^2 , AIC, BIC) sont calculés dans le cas des modèles définis par des observables physiques (tableau 3.14). Ces critères sont calculés dans deux configurations : pour le groupe de spectres complet (*tous les spectres*) et pour la réunion des deux groupes de spectres (*groupes (1) + (2)*) considérés précédemment. Cette approche permet ainsi de comparer des systèmes dont la quantité d'information est identique (même nombre de spectres dans les deux systèmes considérés).

	Modèle de Delany et Bazley		Modèle de Zwicker et Kosten	
	<i>tous spectres</i>	<i>groupes (1)+(2)</i>	<i>tous spectres</i>	<i>groupes (1)+(2)</i>
R^2	$0,01 \pm 0,01$	$0,45 \pm 0,24$	0	$0,45 \pm 0,24$
AIC	1	16	3	19
BIC	100	216	129	269

TAB. 3.14 – Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten basées sur les trois observables physiques.

Le point essentiel est le passage de l'ensemble des spectres au découpage de la base de données en deux groupes. Que ce soit pour le modèle de Delany et Bazley ou le modèle de Zwicker et Kosten, le coefficient de détermination tend donc à augmenter de façon significative ce qui rend compte d'une meilleure adéquation entre données mesurées et modélisées lorsque la base de données est découpée. Le coefficient de détermination associé à l'ensemble des spectres est quasiment nul, que ce soit pour le modèle de Delany et Bazley ou le modèle de Zwicker et Kosten ; ce résultat témoigne de la faible sensibilité du modèle aux paramètres d_j (la température de l'air). Le découpage de la base de données en deux groupes tend à augmenter le contenu prédictif des modèles.

Cependant, les coefficients AIC et BIC (chapitre 2 section 2.3) ne vont pas en ce sens. En effet, pour qu'un modèle soit considéré comme optimal, les valeurs de AIC et BIC doivent être minimales. Or leurs valeurs augmentent de façon significative en découpant la base de données en deux groupes, que ce soit pour le modèle de Delany et Bazley ou le modèle de Zwicker et Kosten.

Le passage d'un système considérant un groupe de spectres et d'un système intégrant deux groupes de spectres permet de minimiser la valeur de la fonction coût (soit de maximiser la fonction de log-vraisemblance \mathcal{L}) ; les critères AIC et BIC devraient donc diminuer. Or le passage d'un système à un autre modifie également le nombre de paramètres à estimer ; le passage d'un groupe à deux

sous-groupes (groupes (1) + (2)) multiplie alors par deux le nombre de paramètres à estimer ce qui tend alors à augmenter la valeur des critères de façon considérable. A titre d'illustration, pour une fonction coût constante fixée de façon arbitraire ($\mathcal{L} = 10$), on se propose de calculer les critères AIC et BIC en fonction du nombre de paramètres, en l'occurrence le nombre de paramètres estimés pour les modèles d'impédance de sol à l'étude (tableau 3.15).

	Modèle de Delany et Bazley		Modèle de Zwicker et Kosten	
	<i>tous spectres</i>	<i>groupes (1)+(2)</i>	<i>tous spectres</i>	<i>groupes (1)+(2)</i>
q	8	8×2	10	10×2
AIC	-4	12	0	20
BIC	96	213	125	271

TAB. 3.15 – Critères statistiques calculés pour une valeur de fonction coût arbitraire constante ($\mathcal{L} = 10$) pour un nombre d'observations $n = 479 \times 3$, pour un nombre de paramètres q dépendant du système considéré.

A travers ces résultats, on s'aperçoit que la variation du critère AIC au passage d'un groupe à deux groupes est du même ordre de grandeur que la variation observée tableau 3.15. Les légères différences sont notamment dues à l'augmentation de la fonction de log-vraisemblance au passage à deux groupes de spectres. Pour un nombre de paramètres similaires, les critères AIC et BIC seraient alors *très légèrement* inférieurs dans le cas où deux groupes de spectres sont considérés. En regard des critères AIC, BIC et surtout du coefficient de détermination R^2 , on s'aperçoit donc que le découpage de la base de données en deux groupes de spectres permet de renforcer le contenu prédictif du système.

Un élément important à relever est que ces paramètres ne rendent pas compte des limitations rencontrées au cours du processus de calibration. Si ces seuls critères sont utilisés pour valider le processus, le modèle de Zwicker et Kosten pourrait être utilisé au même titre que le modèle de Delany et Bazley, alors qu'il n'a pas été développé pour ce type de sol (section 3.3.1). Il est donc important de suivre avec attention le déroulement du processus de calibration et de relever d'éventuels "incidents" (problème d'optimisation, biais) qui pourraient aller à l'encontre des critères statistiques calculés. L'apport d'une expertise physique est également essentielle.

4.2 OBSERVABLES DÉFINIES PAR ACP_F

Les observations faites dans le cas des observables physiques sont transposables au cas des observables par coefficients d'ACP_F (tableau 3.16). L'augmentation du coefficient de détermination R^2 est prépondérante et rend davantage compte d'une amélioration du modèle statistique.

	Modèle de Delany et Bazley		Modèle de Zwicker et Kosten	
	<i>tous spectres</i>	<i>groupes (1)+(2)</i>	<i>tous spectres</i>	<i>groupes (1)+(2)</i>
R^2	$0,01 \pm 0,01$	$0,33 \pm 0,33$	0	$0,33 \pm 0,32$
AIC	-7	3	-3	6
BIC	51	120	79	163

TAB. 3.16 – Critères statistiques calculés à partir des estimations de paramètres avec le modèle de Delany et Bazley et le modèle de Zwicker et Kosten basées sur les coefficients d'ACP_F.

On peut cependant relever qu'en comparaison avec les résultats obtenus avec des observables physiques (tableau 3.14), le coefficient de détermination R^2 est plus faible et l'écart-type associé à ce critère est de l'ordre de grandeur de ce critère. En regard de ce critère, on pourrait alors choisir préférentiellement une définition physique des observables. Cependant, l'analyse des propriétés des résidus joue en faveur d'une définition d'observables à l'aide de l'ACPF, qui par définition implique leur indépendance et par conséquent l'indépendance des résidus.

4.3 BILAN

L'analyse d'un point de vue physique et statistique permet de renforcer les conclusions sur les résultats obtenus, pour les différents systèmes statistiques et physiques étudiés. Le découpage de la base de données permet d'améliorer le contenu prédictif d'un système d'étude comme en témoignent les critères statistiques AIC, BIC, R^2 , ainsi que la statistique associée aux résidus. La variabilité physique des données dans le temps est ainsi accessible.

Les observables doivent être adaptées au modèle physique considéré; les observables physiques définies à partir du modèle de Delany et Bazley ne doivent en aucun cas être utilisées pour l'étude d'un autre modèle. En effet, les modèles ne présentent pas la même sensibilité aux paramètres d'entrée et ne reproduisent pas de façon similaire la dynamique des sorties du modèle. La définition des observables par ACPF est donc à utiliser préférentiellement pour des études comparatives entre deux modèles d'impédance de sol.

5 CONCLUSION DU CHAPITRE

Ce chapitre a permis de mener une étude de calibration presque complète (la phase prévisionnelle n'a pas été abordée car l'estimation des paramètres de sol est un processus de calibration de paramètres d'entrée - problème inverse) et de détailler les différentes étapes du processus de quantification d'incertitudes. Le choix et la constitution de la base de données expérimentales constituent une étape importante dont dépendra le choix du (ou des) modèle(s) physique(s). La définition des observables est également un élément majeur dans un tel processus de quantification d'incertitudes. Ces observables dépendent directement de la base de données expérimentales ainsi que du modèle physique utilisé pour mener l'étude d'estimation de paramètres (voire même l'étude de prévision). De plus, le découpage de la base de données expérimentale initiale met en évidence des populations distinctes d'un point de vue spectral et aussi du point de vue des paramètres physiques les caractérisant. L'erreur associée à ces paramètres rend bien compte de la dispersion associée à chaque groupe. Cette méthode de calibration permet également de mettre en évidence des "classes" de caractéristiques acoustiques de sol - dans notre cas d'étude selon les conditions météorologiques - par simple observation des histogrammes résiduels sans *a priori* préalable.

L'estimation par maximisation de la fonction de log-vraisemblance permet l'estimation d'un paramètre moyen. L'inversion de la Hessienne de la fonction de log-vraisemblance étant complexe à mettre en œuvre, le processus de bootstrap devient alors une étape obligatoire pour calculer les variances d'estimation des paramètres. Ce processus ne rend cependant pas compte d'éventuels problèmes liés au processus d'optimisation et à l'estimation, masqués par la répétition du processus d'estimation. La méthode de bootstrap ne renforce en aucun cas la fiabilité du modèle mais elle met en évidence certaines propriétés des modèles en regard des données expérimentales. C'est une méthode simple pour quantifier la variabilité et l'incertitude associées aux paramètres d'un système (écart-types des paramètres) et par suite pour quantifier la variabilité et l'incertitude sur les prévisions du modèle. Le processus de bootstrap peut malheureusement s'avérer très chronophage du fait d'itérations répétées du processus de calibration.

Finale­ment, la méthodologie de quantification d'incertitudes est efficace à condition que le modèle d'étude et les observables soient choisies de façon optimale. Le recours à un ensemble de critères statistiques "classiques" de choix ou de validation de modèles et notamment le coefficient de détermination R^2 , permettent de caractériser la qualité d'un modèle statistique. L'observation des paramètres physiques et des propriétés des résidus est une étape essentielle pour l'interprétation des résultats. L'interprétation des critères AIC et BIC n'est pas toujours simple.

Une telle approche pourrait aisément être mise en pratique dans le cadre de campagnes expérimentales afin de d'estimer en "temps réel" les caractéristiques acoustiques de sol et leur associer un écart-type. Une application inspirée de cette approche a déjà été mise en place sur les mesures de la Station de Long Terme de Saint-Berthevin (cf. Annexe A) afin d'estimer seulement les paramètres du modèle [103].

4

ESTIMATION D'INCERTITUDE EN PRÉSENCE D'EFFETS MICROMÉTÉOROLOGIQUES

PLAN DU CHAPITRE

1	SPÉCIFICATION DU SYSTÈME	95
1.1	Données observées	95
1.2	Choix du modèle	99
1.3	Bilan	114
2	DÉFINITION DES OBSERVABLES	115
3	CALIBRATION DU SYSTÈME	116
3.1	Définition du modèle statistique de calibration	116
3.2	Estimation des paramètres incertains	116
4	VALIDATION DU SYSTÈME	118
5	PRÉVISION	122
5.1	Variance d'estimation des paramètres	122
5.2	Incertitude globale	123
5.3	Bilan	127
6	CONCLUSION DU CHAPITRE	127

La quantification des incertitudes en propagation acoustique à grande distance (de l'ordre de la centaine de mètres) implique de tenir compte des paramètres micrométéorologiques en plus des paramètres acoustiques de sol. Cette étude constitue donc un cas d'application plus complexe que celui présenté au chapitre précédent du fait que l'ensemble des phénomènes influents en propagation acoustique en milieu extérieur doit être pris en compte. Afin d'appliquer la méthode de quantification d'incertitude (chapitre 2), il est nécessaire d'identifier les données expérimentales (mesures) à disposition, les paramètres incertains et influents à estimer ainsi que le(s) modèle(s) à disposition présentant une bonne précision et une certaine ergonomie. On s'appuiera dans ce chapitre sur la base de données expérimentales issue de la campagne de Lannemezan 2005 qui rend compte de phénomènes sur un site plan en présence d'une source considérée comme ponctuelle (Annexe A).

Cette nouvelle étude de quantification d'incertitudes est menée afin de répondre à un certain nombre d'objectifs :

1. présenter des cas d'applications de quantification d'incertitude complets, impliquant notamment le développement d'un métamodèle ;
2. accéder à une information quant à l'incertitude et à la variabilité associées aux paramètres influents dans des conditions de propagation données incluant les effets micrométéorologiques ;

3. traduire cette information en termes de niveaux sonores ;
4. rendre compte de la variabilité spatiale ;
5. aboutir à la phase de prévision intégrant des incertitudes globales.

1 SPÉCIFICATION DU SYSTÈME

1.1 DONNÉES OBSERVÉES

La campagne expérimentale de Lannemezan 2005 est décrite en détails Annexe A. Cette campagne présente des données échantillonnées dans le temps (campagne d'une durée de 3 mois avec des spectres en tiers d'octave en L_{eq} sur quinze minutes) et dans l'espace (selon trois directions de propagation, à des distances à la source allant de 25 à 200 mètres à différentes hauteurs, à 2 et 4 m). Cette base de données est constituée d'informations sur les caractéristiques micrométéorologiques, sur les caractéristiques acoustiques de sol et sur les niveaux sonores. La base de données à disposition avait déjà subi un travail de validation et filtrage [68], mais un post-traitement conséquent a été mené en complément (afin de prendre en compte la variabilité liée à la source, direction du vent, etc.)¹.

Pour simplifier l'étude de quantification d'incertitude, seul un point de mesure micrométéorologique pour chaque direction de propagation sera pris en compte. La sélection de ce point de mesure est basée sur la comparaison des gradients verticaux de célérité du son mesurés en différents points sur l'ensemble de la campagne expérimentale : cette étude est présentée section 2.3 de l'Annexe A. Elle montre notamment que les mesures effectuées à 75 m de la source peuvent être considérées comme représentatives du site.

1.1.1 Création de groupes de données expérimentales

Dans l'étude précédente menée sur les caractéristiques acoustiques de sol (chapitre 3), nous avons montré qu'il était plus intéressant de mener l'étude de calibration sur des groupes de spectres et non spectre par spectre (sur l'ensemble de la base de données à disposition) : cette démarche permet en effet de renforcer le contenu prédictif du modèle.

Des groupes de données sont donc sélectionnés à partir de découpages *a priori* des paramètres influents (les paramètres micrométéorologiques avec le gradient vertical de célérité du son $\frac{\partial c}{\partial z}$ et le paramètre de turbulence $\langle \mu^2 \rangle$, ainsi que les caractéristiques de sol avec la résistance spécifique au passage de l'air R_s) pour lesquels les spectres présenteront une certaine variabilité. L'approche consiste dans une première étape à distinguer des classes de propagation "grossières" (favorable, défavorable et homogène), puis pour chaque groupe sont constitués des sous-groupes en fonction de valeurs de R_s et $\langle \mu^2 \rangle$. Une seconde étape consiste à comparer les groupes ainsi créés et à distinguer s'ils sont significativement différents ou non : dans le premier cas le groupe restera tel quel et dans le second cas, les groupes non-significativement différents pourront être agrégés.

Les spectres expérimentaux sélectionnés pour des conditions de propagation dites "homogènes" (chapitre 1 section 1.3.4.2) et les histogrammes des paramètres correspondants sont représentées figures 4.1 et 4.2.

La figure 4.1 montre logiquement une dispersion des niveaux sonores plus importante à 150 m qu'à 50 m de la source. Cette différence est due à la fois à la variabilité et à l'incertitude liées aux caractéristiques acoustiques de sol (chapitre 3) et aux conditions de propagation bien que celles-ci aient été sélectionnées. En effet, les conditions de propagation homogènes (chapitre 1 section 1.3.4.2) sont très rarement observées et il se peut qu'un échantillon de 15 minutes (en *moyenne* homogène) présente en fait une grande variabilité pendant cette période. En outre, des conditions de propagation parfaitement homogènes ne se rencontrent qu'en l'absence de vent, qui peut être vu comme une source de dispersion à travers la notion de célérité effective (chapitre 1 section 1.3.4.2). Pour la constitution des groupes nous avons supprimé les conditions pour lesquelles le vent de travers est trop marqué soit pour $\cos \theta > 0,8$.

1. Ce travail ne sera pas présenté dans le présent document.

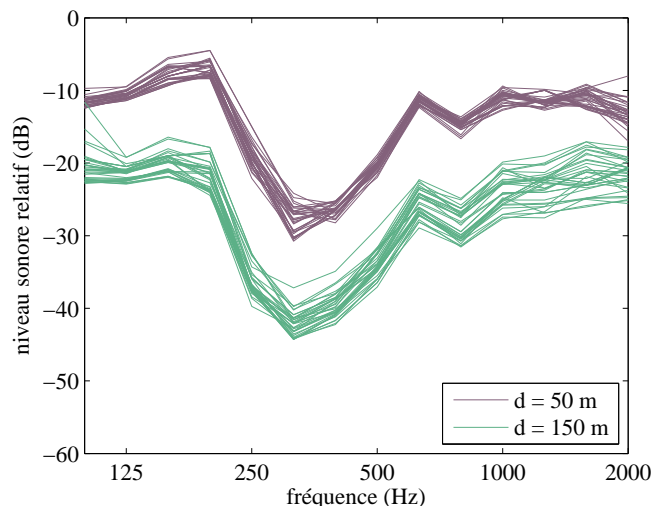


FIG. 4.1 – Campagne de Lannemezan 2005 - spectres d'atténuation expérimentaux mesurés à (50 m, 2 m) et (150 m, 2 m) relativement à un microphone situé à (10 m, 2 m) de la source pour des conditions de propagation homogènes.

La figure 4.2 rend compte de la dispersion des paramètres influents relatifs aux groupes de spectres sélectionnés. La sélection de groupes a permis ici d'isoler des conditions de propagation en moyenne proches de l'homogène pour des caractéristiques de sol "peu" variables expérimentalement.

1.1.2 Paramètres d'entrée du système

En regard des paramètres qui ont servi à la constitution de groupes de spectres caractéristiques (ici de conditions de propagation homogènes), les paramètres influents constitueront le vecteur de paramètres d'entrée incertains θ du système :

- le gradient vertical de célérité effective du son calculé à partir des coefficients a_{log} ($m.s^{-1}$) et b_{lin} (s^{-1}) suivant la relation $\frac{\partial c}{\partial z} = \frac{a_{log}}{z} + b_{lin}$, où z est la hauteur considérée (chapitre 1 section 1.3.4.2) ;
- la résistance spécifique au passage de l'air R_s ($kN.s.m^{-4}$) caractéristique de l'impédance acoustique du sol (chapitre 1 section 1.3.3) ;
- le paramètre de turbulence $\langle \mu^2 \rangle$ (chapitre 1 section 1.3.4.3).

D'autres paramètres considérés comme maîtrisés (notés d_j) sont à prendre en considération :

- les paramètres relatifs à la configuration géométrique de la situation physique étudiée tels que la hauteur de la source, la hauteur du récepteur, la distance source-récepteur ;
- l'absorption atmosphérique (section 1.3.2) qui va dépendre des paramètres géométriques bien connus (distance de propagation), de l'humidité relative ainsi que de la température de l'air (figure 4.3). Aux distances source-récepteur considérées, l'influence de l'absorption atmosphérique sur les niveaux sonores est peu marquée (notamment aux basses fréquences) en comparaison des paramètres précédemment cités.

Les paramètres physiques incertains notés θ vont constituer les paramètres du modèle à calibrer durant le processus de quantification d'incertitudes. Les paramètres bien connus notés d_j vont permettre d'expliquer une part de la variabilité des données à l'étude (cf. figures 4.4 et 4.5).

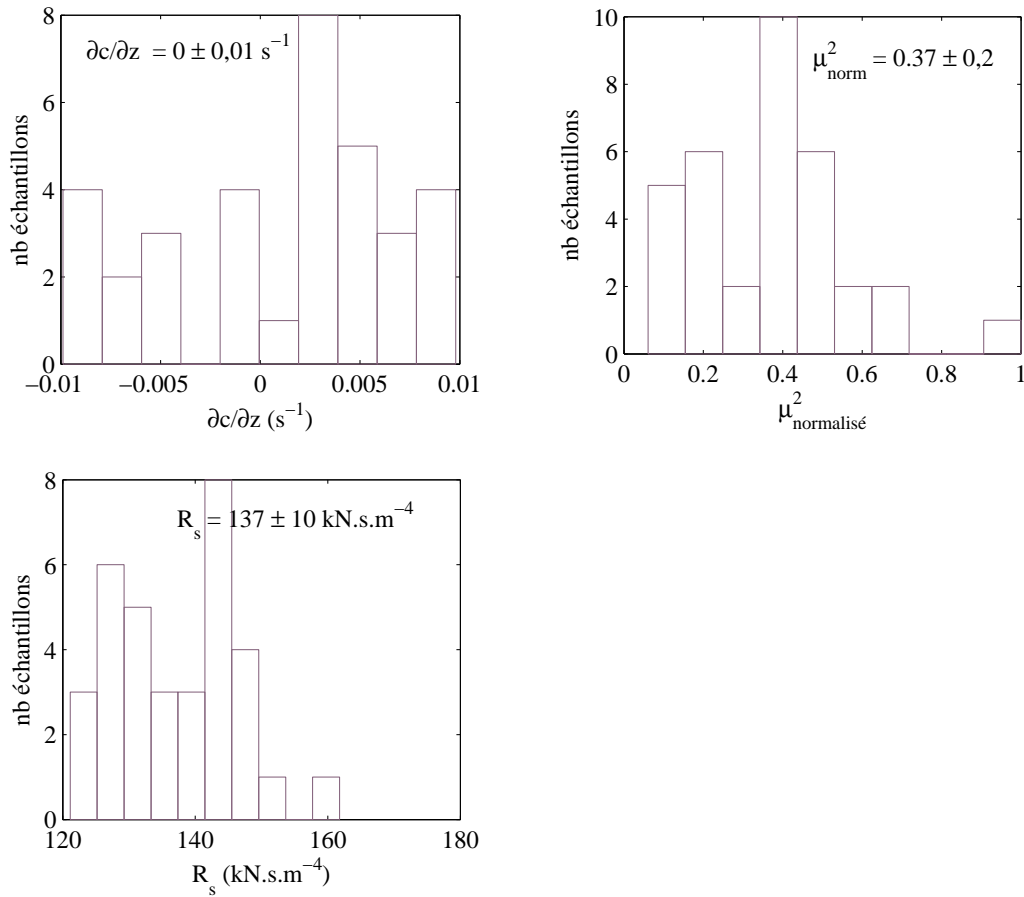


FIG. 4.2 – Campagne de Lannemezan 2005 - histogrammes du gradient vertical de célérité effective $\frac{\partial c}{\partial z}$, de l'indice de turbulence normalisé $\langle \mu_{norm}^2 \rangle$ et de la résistance spécifique au passage de l'air R_s mesurés pour les spectres sélectionnés. La moyenne et l'écart-type associés à chaque paramètre sont précisés.

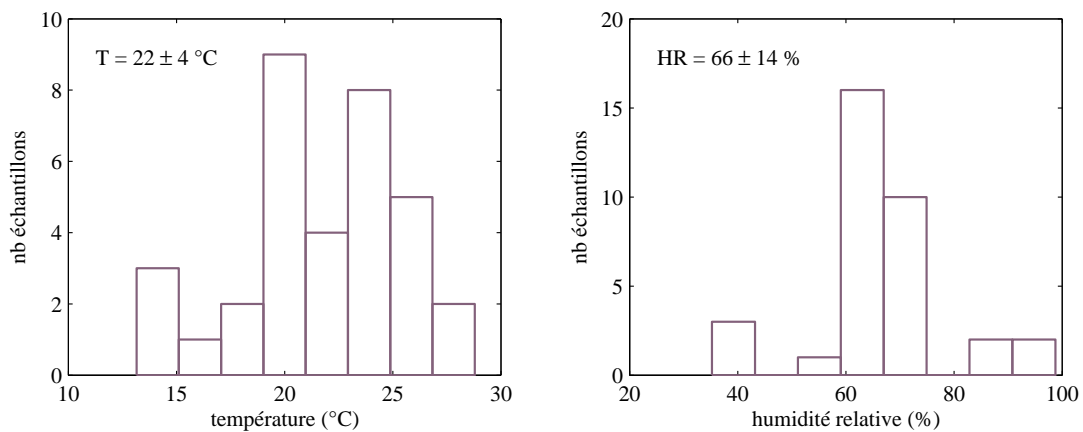


FIG. 4.3 – Campagne de Lannemezan 2005 - histogrammes de la température de l'air et de l'humidité relative mesurées pour les spectres sélectionnés. La moyenne et l'écart-type associés à chaque paramètre sont précisés.

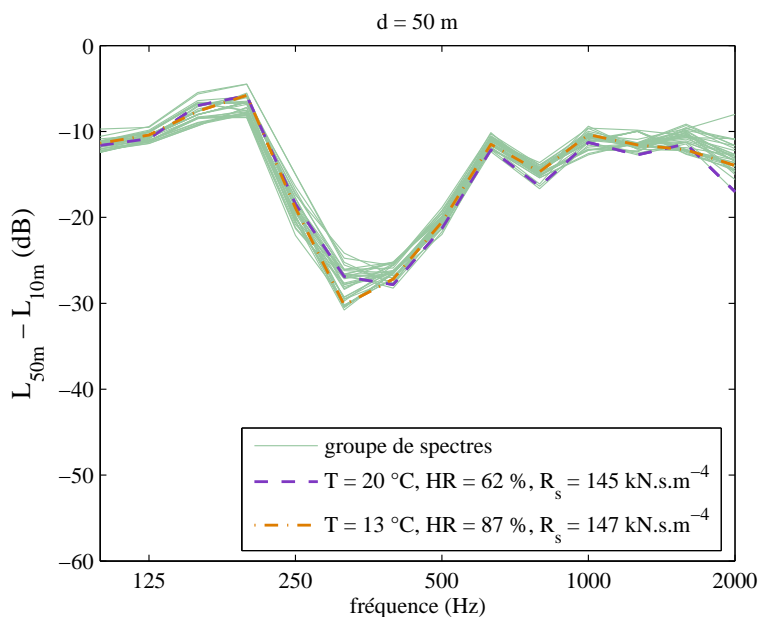


FIG. 4.4 – Campagne de Lannemezan 2005 - spectres d'atténuation expérimentaux mesurés à 50 m relativement à un microphone situé à 10 m de la source pour des conditions de propagation homogènes – mise en évidence de spectres pour des valeurs θ proches de la moyenne et pour des valeurs d_j distinctes.

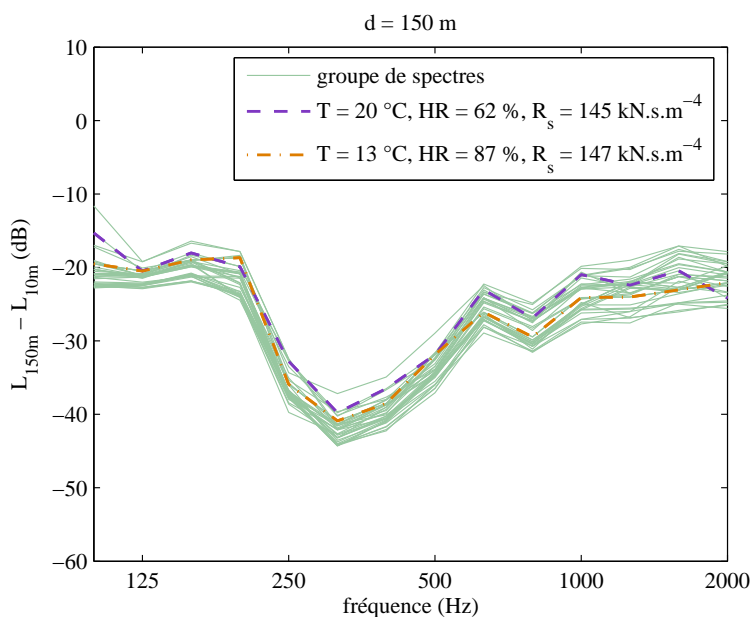


FIG. 4.5 – Campagne de Lannemezan 2005 - spectres d'atténuation expérimentaux mesurés à 150 m relativement à un microphone situé à 10 m de la source pour des conditions de propagation homogènes - mise en évidence de spectres pour des valeurs θ proches de la moyenne et pour des valeurs d_j distinctes.

Les figures 4.4 et 4.5 illustrent l'influence des valeurs des paramètres bien connus d_j pour des valeurs θ très proches de la moyenne pour les groupes de spectres sélectionnés à 50 et 150 m. On

s'aperçoit qu'une part non négligeable de la dispersion peut être expliquée par les paramètres d_j . Cet ensemble de paramètres constitue les paramètres d'entrée du modèle acoustique de simulation utilisé dans le processus de quantification d'incertitudes par calibration.

1.2 CHOIX DU MODÈLE

Le modèle de propagation acoustique utilisé est un modèle numérique de référence basé sur la résolution de l'Équation Parabolique (chapitre 1 section 2.3). Ce modèle coûteux en temps de calcul sera remplacé par un *métamodèle* (chapitre 2 section 1.3). Il va permettre de calculer l'atténuation des niveaux sonores à une distance donnée relativement à 10 m de la source, en tenant compte de l'absorption atmosphérique (chapitre 1 section 1.3.2).

1.2.1 Objectifs et nature du métamodèle

Le processus de quantification d'incertitudes étant basé sur une méthode d'optimisation requérant un grand nombre d'appels au modèle, un modèle simplifié appelé *métamodèle* (chapitre 2 section 1.3) a dû être construit afin d'optimiser le coût numérique. L'optimisation du modèle de simulation constitue de plus une étape indispensable pour s'assurer de la qualité du processus de calibration. Un élément intéressant est de lui associer une variance de prévision. Cette variance de prévision peut être mise en regard de la variance finale estimée. Un métamodèle de krigeage² (cf. Annexe B) est ainsi construit pour répondre à ce besoin.

1.2.2 Construction du plan d'expériences numériques

1.2.2.1 Identification du domaine d'applicabilité

Le métamodèle est utilisé à terme pour être comparé à des données expérimentales (section 1.1). Le domaine d'applicabilité doit donc recouvrir les situations rencontrées au cours de la campagne de mesure. De plus, le domaine d'applicabilité du plan d'expériences numériques utilisé pour construire le métamodèle est défini à partir des paramètres du modèle de propagation par les éléments suivants.

(a) *Domaine fréquentiel*

Il est défini par les tiers d'octave centrés sur les fréquences comprises entre 100 et 2000 Hz, soit un total de 14 tiers d'octave \times nombre de fréquences par tiers d'octave = 49. Les calculs sont menés pour un certain nombre de fréquences puis sont recomposés pour constituer les bandes de tiers d'octaves. Cette recombinaison permet notamment de lisser d'éventuels phénomènes interférentiels surestimés par un calcul en bandes fines ou fréquences pures (ces phénomènes sont d'autant plus marqués aux hautes fréquences). Le nombre de fréquences par tiers d'octave choisi sera donc plus élevé aux hautes fréquences qu'aux basses fréquences (conformément à la référence [69]).

(b) *Configuration géométrique*

La hauteur de la source est fixée à 2 m, la hauteur de récepteur à 2 et 4 m et la distance source-récepteur s'étend jusqu'à 500 m.

(c) *Résistance spécifique au passage de l'air*

Le domaine d'étude de R_s est défini en se basant sur l'étude des paramètres de sol menée au chapitre 3 : R_s varie de 50 à 290 kN.s.m⁻⁴ ce qui rend compte de l'ensemble des conditions de sol rencontrées au cours de la campagne.

2. Ce métamodèle a été construit sous Matlab grâce à l'outil DACE (Design and Analysis of Computer Experiments) [85]

(d) Paramètres de réfraction

Concernant les paramètres micrométéorologiques c'est-à-dire les coefficients a_{log} et b_{lin} définissant le gradient vertical de célérité du son (chapitre 1 section 1.3.4.2), le choix du jeu de paramètres nécessite une étude plus approfondie. Ces paramètres étant dépendants, il est nécessaire de les définir corrélativement. Afin de déterminer un jeu de paramètres, il est intéressant de voir comment ont été définis ces mêmes paramètres dans le cadre d'autres projets et études. Ainsi, la figure 4.6 présente les valeurs des coefficients a_{log} et b_{lin} issues de la campagne expérimentale de Lannemezan 2005, des plans de calculs définis dans le cadre d'un projet européen (Harmonoise) [40] et dans le cadre d'un projet national (MEDD) [68].

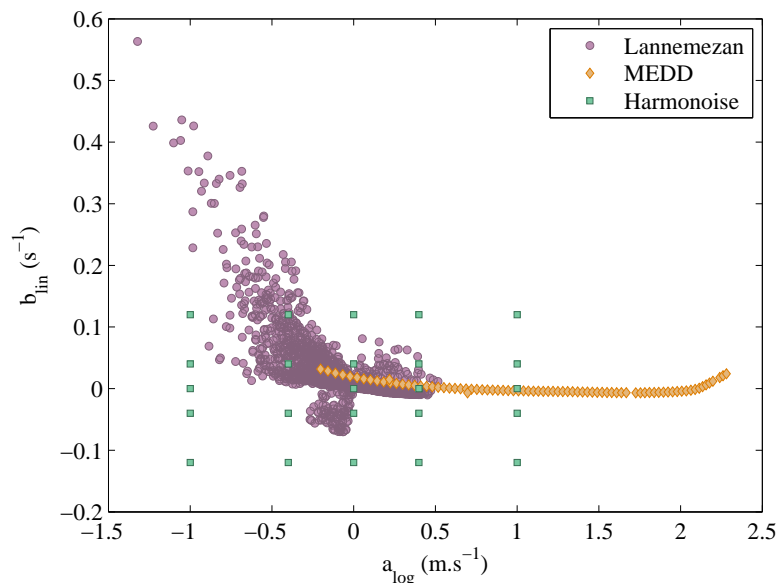


FIG. 4.6 – Nuages de corrélation entre les paramètres a_{log} et b_{lin} pour la campagne de Lannemezan 2005, les projets MEDD et Harmonoise.

Les données définies dans le cadre du projet Harmonoise permettent de recouvrir un grand domaine de données $[a_{log} - b_{lin}]$ alors que les données issues du projet MEDD contraignent les valeurs de b_{lin} autour de 0 s^{-1} (entre $-0,05 \text{ s}^{-1}$ et $0,03 \text{ s}^{-1}$) : les profils modélisés dans cette dernière étude étaient essentiellement influencés par le vent et donc de forme logarithmique marquée. Cette différence provient également des objectifs et des échelles spatiales des projets respectifs. Les profils de célérité correspondants sont alors calculés et représentés figure 4.7.

Dans le cadre de ces travaux de thèse, afin de trouver des valeurs de gradients équivalentes à la campagne de Lannemezan 2005 tout en minimisant le nombre de calculs, il nous a semblé plus judicieux de s'appuyer sur la base de données du projet Harmonoise, en affinant les valeurs entre chaque point. Ce choix permet globalement de recouvrir l'ensemble des profils expérimentaux de la campagne de Lannemezan 2005. Finalement le domaine d'étude définissant les coefficients a_{log} et b_{lin} retenu est le suivant : $a_{log} = [-0,80 : 0,80] \text{ m.s}^{-1}$ et $b_{lin} = [-0,12 : 0,12] \text{ s}^{-1}$. Cet échantillonnage ne prend cependant pas en compte certaines conditions de propagation, à savoir les valeurs de $b_{lin} > 0,12 \text{ s}^{-1}$ représentatives des inversions de température et les valeurs $a_{log} < -0,8 \text{ m.s}^{-1}$ représentatives de conditions aérodynamiques très défavorables. Cette perte d'information n'est pas problématique car notre étude est basée sur des conditions homogènes et n'intègre pas des conditions de propagation "extrêmes".

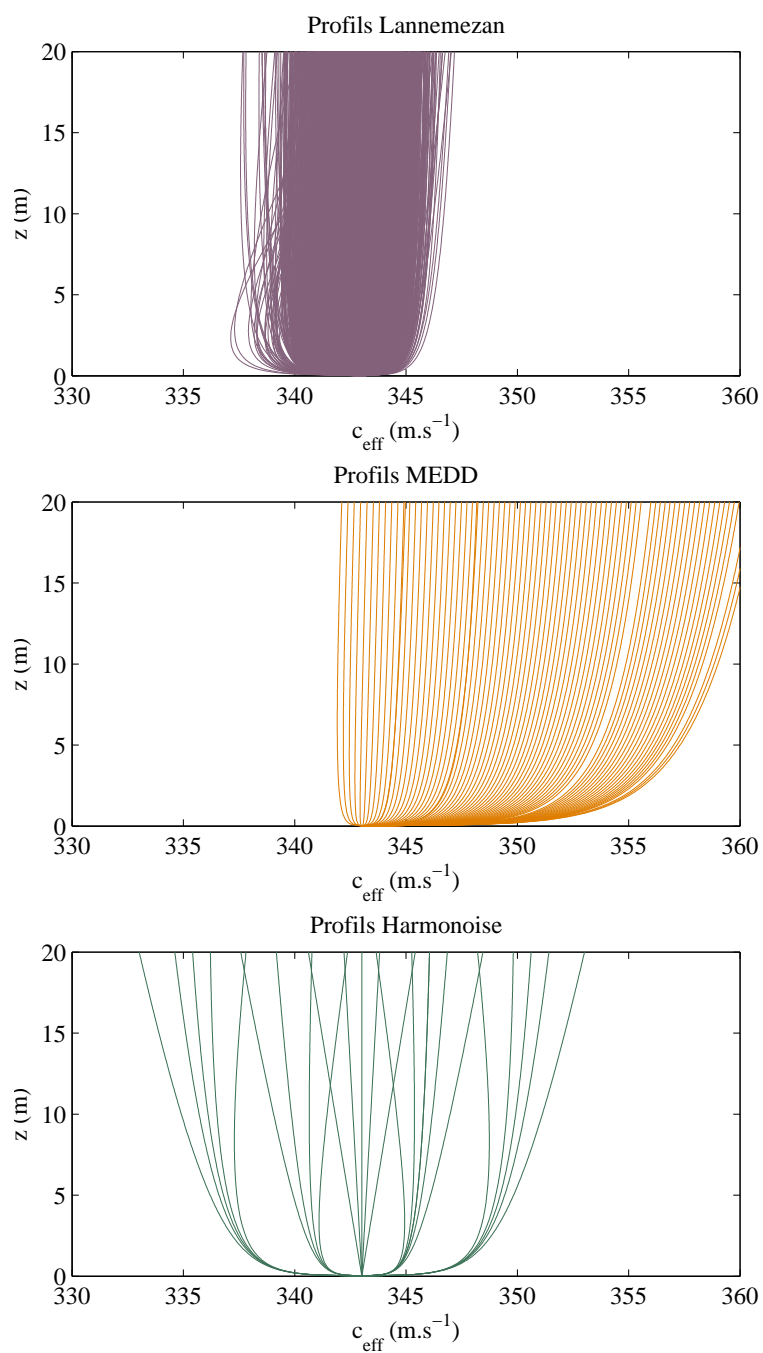


FIG. 4.7 – Profils verticaux de célérité du son effective calculés pour la campagne de Lannemezan 2005, les projets MEDD et Harmonoise.

(e) Paramètre de turbulence

Inclure le paramètre de turbulence $\langle \mu^2 \rangle$ dans le modèle d'Équation Parabolique est extrêmement coûteux en temps de calcul : la durée d'un calcul est multipliée par 50 à 100 par rapport au cas sans turbulence. La construction d'un plan de calculs numériques prenant en compte des valeurs variables du paramètre de turbulence devient alors rédhibitoire. Il est cependant indispensable de l'intégrer dans les calculs car la turbulence a notamment pour effet d'apporter de l'énergie en zone d'ombre et de lisser les figures interférentielles (perte de cohérence) aux fréquences les plus élevées (chapitre 1 section 1.3.4.3). Afin de rendre compte au mieux de ces phénomènes physiques et d'intégrer la turbulence dans notre plan de calcul tout en optimisant les temps de calcul, plusieurs approches ont été envisagées :

1. *ne pas prendre en compte la turbulence* dans le calcul induit de trop fortes atténuations en conditions défavorables, *i.e.* pour $a_{log} < -0,2 \text{ m.s}^{-1}$ (figures 4.8 et 4.9) surtout en hautes fréquences. En effet, à une distance source-récepteur de 150 m et à 2000 Hz, l'atténuation peut atteindre -120 dB ce qui n'a aucun sens physique. Une proposition est de *ne travailler qu'en conditions de propagation homogènes* ($\frac{\partial c}{\partial z} \sim 0$) et *favorables* ($\frac{\partial c}{\partial z} > 0$) pour lesquelles l'effet de la turbulence peut être considéré comme négligeable (à de "faibles" distances source-récepteur, inférieures à 100 m et pour des fréquences basses à moyennes) en regard de l'effet de turbulence en conditions de propagation défavorables. Cette approche restreint malheureusement de façon notable le domaine d'étude ;
2. *corriger les spectres modélisés* à l'aide d'une méthode définie par Salomons [40]. Selon lui, les effets de la turbulence atmosphérique peuvent être pris en compte après avoir procédé aux calculs de propagation en fixant une valeur "seuil" limite de 15 dB par rapport au champ direct. Les calculs de propagation sont effectués dans une atmosphère sans turbulence avec une absorption atmosphérique. Cependant cette approche ne permet pas d'avoir un paramètre de turbulence variable et dépendant de la fréquence (ce qui est le cas en réalité) ;
3. *construire un premier plan de calcul sans turbulence* ($\langle \mu^2 \rangle = 0$) et *un second plan de calcul avec une valeur de turbulence considérée comme maximale* ($\langle \mu_{max}^2 \rangle = 1.10^{-5}$) [38]. Afin de considérer des valeurs intermédiaires de turbulence, les plans construits pour ces valeurs intermédiaires sont calculés par interpolation. Cette démarche ne rend pas compte d'une grandeur physique du paramètre de turbulence mais d'une valeur normalisée de ce paramètre variant de 0 pour $\langle \mu^2 \rangle = 0$ à 1 pour $\langle \mu^2 \rangle = 1.10^{-5}$. Elle est calculée comme le rapport entre une valeur donnée de turbulence et la valeur de turbulence maximale considérée soit $\langle \mu_{norm}^2 \rangle = \frac{\langle \mu^2 \rangle}{\langle \mu_{max}^2 \rangle}$.

Considérant ces différentes approches et pour ne pas restreindre le domaine d'applicabilité du méta-modèle aux conditions homogènes et favorables, la troisième approche sera développée.

L'ensemble de ces paramètres et leur domaine d'applicabilité permettent ainsi de construire un plan d'expériences numériques.

1.2.2.2 Plan d'expériences numériques

(a) Pour une valeur de turbulence maximale

Le plan d'expériences pour une valeur de turbulence maximale doit permettre de minimiser le nombre de calculs effectués (très coûteux en temps dès lors qu'on prend en compte la turbulence) tout en permettant une répartition homogène des données dans l'espace des paramètres (bornes et intérieur du domaine). Doehlert a proposé une répartition uniforme des points expérimentaux dans l'espace expérimental (chapitre 2 section 1.3.3). Tous les points sont à la même distance du centre

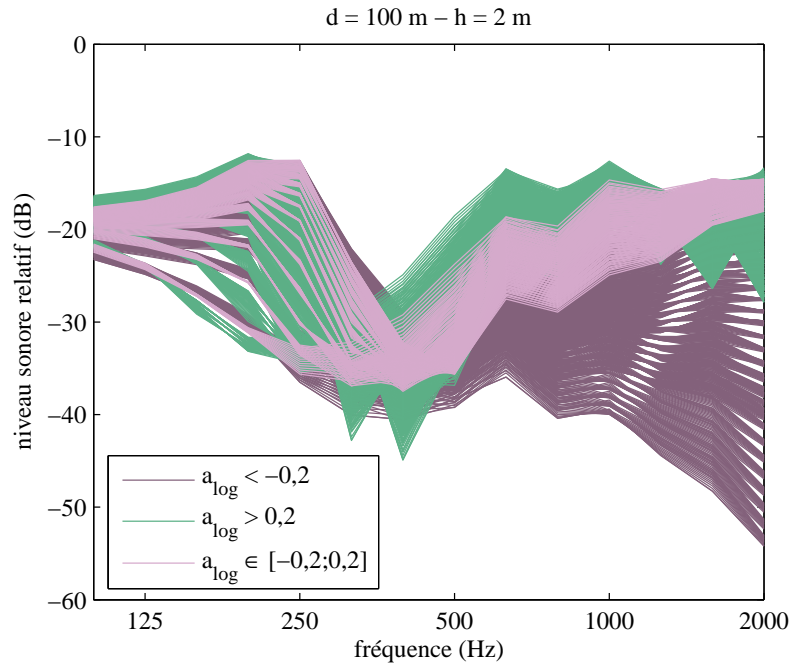


FIG. 4.8 – Spectres d'atténuation à (100 m, 2 m) en relatif à (10 m, 2 m) modélisés par le code d'Équation Parabolique sans turbulence représentés par classes de propagation selon la valeur du paramètre de réfraction a_{log} . On remarque une (trop) grande dispersion des résultats en hautes fréquences.

du domaine expérimental et sont régulièrement disposés sur un cercle trigonométrique dans le cas où le nombre de paramètres s'élève à deux ; ils forment ainsi un hexagone régulier. Dans le cas où le nombre de paramètres s'élève à trois, aux six points de l'hexagone régulier s'ajoutent six points supplémentaires formant ainsi un tétraèdre régulier. En plus d'une répartition homogène, ce type de plan présente l'avantage d'être complété de façon séquentielle. Le nombre d'essais est très réduit si on connaît les facteurs influents et le domaine d'étude. Un plan de Doehlert minimise également le nombre de calculs pour un nombre de paramètres donné ce qui est un avantage non négligeable dans le cas où les calculs sont longs. Un tel plan présente donc toutes les propriétés que nous recherchons.

Le plan de Doehlert est construit à partir d'une matrice d'expériences bien définie correspondant aux coordonnées de l'hexagone (deux paramètres), du tétraèdre (trois paramètres), etc. ; cette matrice de variables centrées réduites est donnée par le tableau 2 du chapitre 2.1. Pour une valeur fixe de turbulence, les autres paramètres influents sont la résistance spécifique au passage de l'air R_s , ainsi que les paramètres de réfraction a_{log} et b_{lin} . Les paramètres a_{log} et b_{lin} permettent tous les deux de caractériser les conditions atmosphériques de propagation, on choisit donc dans une première étape de construire le plan à partir des paramètres R_s et a_{log} (ce dernier ayant une influence prépondérante sur le gradient vertical de célérité du son). Dans une deuxième phase, il est possible de compléter le plan avec le paramètre b_{lin} . En regard de la matrice de variables centrées réduites (chapitre 2 tableau 2.1), les vecteurs de paramètres choisis sont les suivants :

- $X_1 = R_s = [50 \quad 110 \quad 170 \quad 230 \quad 290] \text{ kN.s.m}^{-4}$;
- $X_2 = a_{log} = [-0,8 \quad -0,4 \quad -0,2 \quad 0 \quad 0,2 \quad 0,4 \quad 0,8] \text{ m.s}^{-1}$;
- $X_3 = b_{lin} = [-0,12 \quad 0 \quad 0,12] \text{ s}^{-1}$.

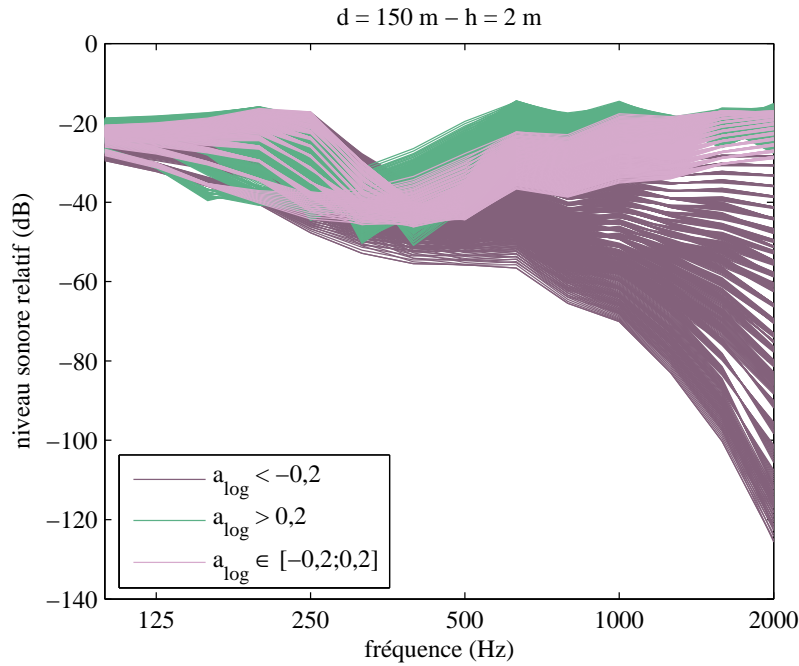


FIG. 4.9 – Spectres d'atténuation à (150 m, 2 m) en relatif à 10 m modélisés par le code d'Équation Parabolique sans turbulence représentés par classes de propagation selon la valeur du paramètre de réfraction a_{log} . On remarque une (trop) grande dispersion des résultats en hautes fréquences.

La répartition des paramètres est synthétisée dans le tableau 4.1.

n	R_s (kN.s.m ⁻⁴)	a_{log} (m.s ⁻¹)	b_{lin} (s ⁻¹)
1	170	0	0
2	50	0	0
3	290	0	0
4	110	-0,8	0
5	230	-0,8	0
6	110	0,8	0
7	230	0,8	0
8	110	-0,2	-0,12
9	170	0,4	-0,12
10	230	-0,2	-0,12
11	110	0,2	0,12
12	170	-0,4	0,12
13	230	0,2	0,12

TAB. 4.1 – Définition du plan d'expériences numériques avec le paramètre de turbulence fixé à 1.10^{-5} .

Le plan d'expériences ainsi construit est représenté par la figure 4.10 qui rend compte de la répartition homogène des paramètres.

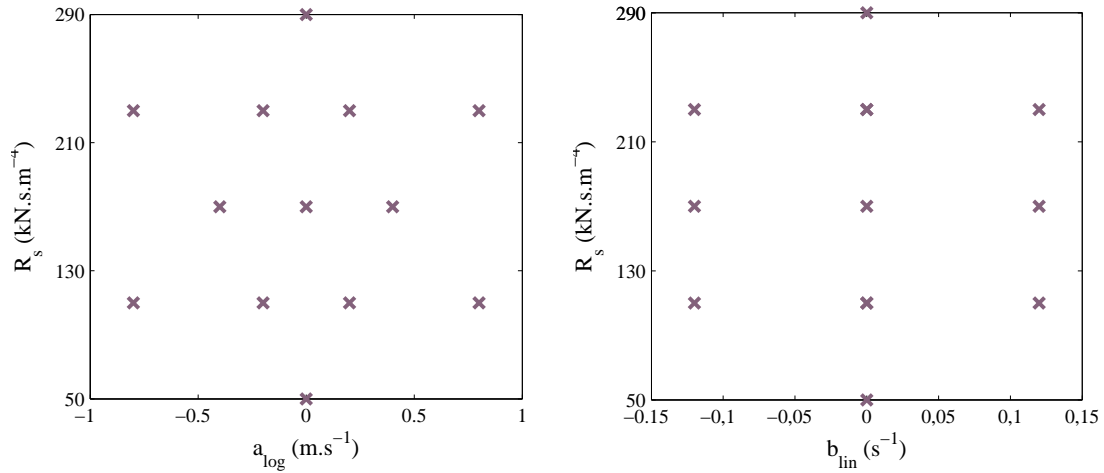


FIG. 4.10 – Vues 2D du plan d’expériences numériques avec le paramètre de turbulence fixé à 1.10^{-5} .

Le plan d’expériences avec une valeur maximale de turbulence étant défini ($\langle \mu_{norm}^2 \rangle = 1$), le second plan de calcul sans turbulence va alors être construit.

(b) *Pour une valeur de turbulence nulle*

Les calculs menés avec une turbulence nulle étant moins coûteux en temps de calcul, il est possible de construire un plan d’expériences numériques plus “robuste” : un plan factoriel complet (chapitre 2 section 1.3.3.1). Les paramètres sont ainsi discrétisés comme suit :

- $R_s = [50 : 20 : 290] \text{ kN.s.m}^{-4}$;
- $a_{log} = [-0,8 : 0,2 : 0,8] \text{ m.s}^{-1}$;
- $b_{lin} = [-0,12 : 0,02 : 0,12] \text{ s}^{-1}$.

Les calculs des niveaux sonores avec et sans turbulence présentant des tendances globales similaires (hormis en conditions défavorables à la propagation où l’atténuation est très forte lorsqu’on ne tient pas compte de la turbulence) dans l’espace des paramètres, le plan de Doehlert avec turbulence est interpolé pour reproduire la même répartition que le plan factoriel complet sans turbulence. Cette démarche est discutable mais une procédure de validation croisée (section 1.2.3) a permis de la conforter.

(c) *Combinaison des plans avec et sans turbulence*

Désormais un troisième plan intermédiaire (représentatif de $\langle \mu_{norm}^2 \rangle = 0,5$) peut être construit par interpolation des deux plans factoriels complets avec et sans turbulence. C’est sur la base de ce plan d’expériences numériques, constitué de trois plans factoriels complets, qu’est construit le métamodèle de krigeage.

(d) *Plan d’expériences “final”*

Une mise en œuvre préliminaire du processus de quantification d’incertitudes a montré une sur-paramétrisation du métamodèle du fait que les paramètres micrométéorologiques a_{log} et b_{lin} ne sont pas indépendants. Un seul paramètre peut suffire pour caractériser les phénomènes de réfraction acoustique : le gradient vertical de célérité effective. Finalement le plan de calcul numériques est représenté figure 4.11 ; les paramètres sont répartis comme suit :

- $R_s = [50 : 20 : 290] \text{ kN.s.m}^{-4}$;
- $\frac{\partial c}{\partial z} = [-0,25 : 0,05 : 0,25] \text{ s}^{-1}$;
- $\langle \mu_{norm}^2 \rangle = [0 \quad 0,5 \quad 1] \text{ s}^{-1}$.

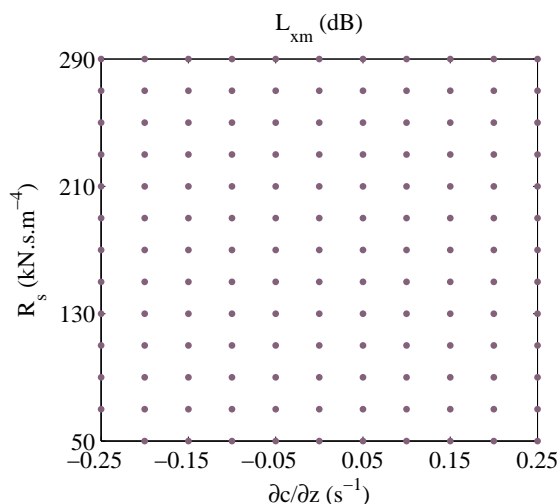


FIG. 4.11 – Plan factoriel complet pour une valeur de turbulence donnée.

La création du métamodèle est ainsi basée sur ce dernier plan d'expériences numériques calculé par interpolation d'un plan de Doehlert à deux paramètres (R_s et $\frac{\partial c}{\partial z} = \frac{a_{log}}{z}$, où z est la hauteur de calcul du gradient, soit $z = 3$ m) représentatif de la turbulence et d'un plan factoriel complet sans prise en compte de la turbulence. Cette approche permet alors de prendre en compte les effets de la turbulence de manière relativement robuste tout en minimisant les coûts de calculs. Le métamodèle sera ainsi construit à partir de ce plan d'expériences à trois paramètres.

1.2.3 Ajustement et validation du métamodèle

1.2.3.1 Ajustement du métamodèle

Le métamodèle étant un métamodèle de krigeage (universel), l'ajustement et la validation passe notamment par la minimisation de l'erreur quadratique moyenne entre le métamodèle et le modèle de référence. Les phénomènes à modéliser étant relativement continus, un métamodèle robuste peut être construit. Ce métamodèle est construit pour simuler plusieurs configurations géométriques pour un récepteur placé à des distances de : 50 et 150 m de la source et à une hauteur de 2 m. La construction du métamodèle est menée grâce à la *Toolbox Matlab DACE* [85]. Les modèles de variogrammes choisis sont des modèles exponentiels. A noter que dans certains cas, un filtrage peut être effectué afin d'obtenir une surface plus régulière au risque de ne pas passer par tous les points de données.

Des exemples de surfaces de réponse (représentant l'atténuation relative au microphone de référence (10 m, 2 m)) créées à partir du métamodèle, à 50 m et 150 m de la source pour deux valeurs de turbulence, à 500 Hz et 2 kHz sont représentés sur les figures 4.12 à 4.23. L'erreur quadratique moyenne associée au métamodèle est précisée pour chacune des surfaces. Sur ces surfaces sont indiqués les points de calculs effectués pour la construction du plan d'expériences numériques. Les points de calculs pour $\langle \mu_{norm}^2 \rangle = 0,5$ n'ont pas tous été calculés. Ces points sont issus de l'interpolation entre les calculs avec et sans turbulence. L'erreur quadratique moyenne relative à chacune des surfaces ne sera donc donnée que pour les points réellement calculés.

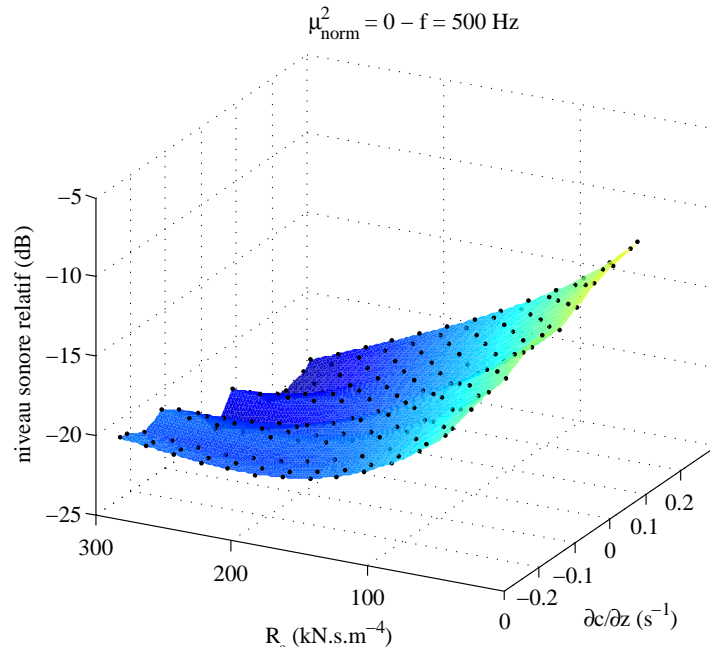


FIG. 4.12 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 0 - f = 500 \text{ Hz}$).
L'erreur quadratique moyenne sur la surface est de $0,03 \text{ dB}^2$.

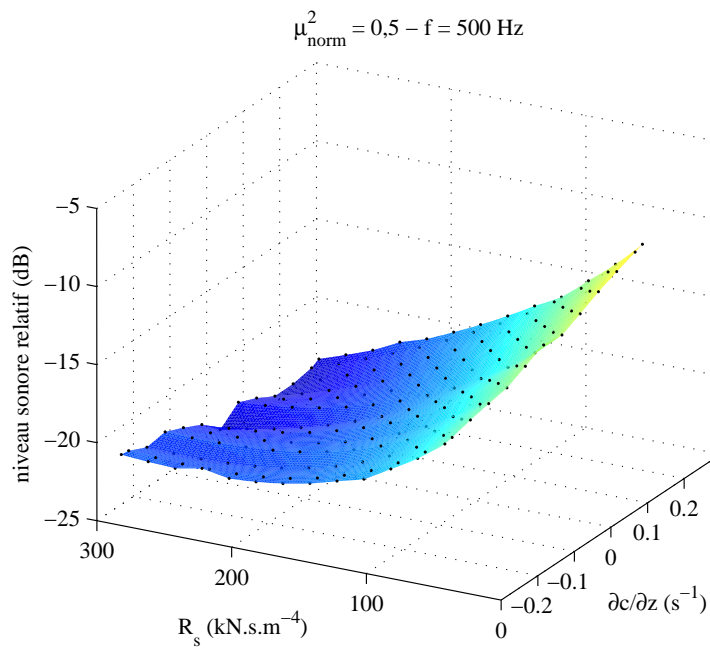


FIG. 4.13 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 0,5 - f = 500 \text{ Hz}$).

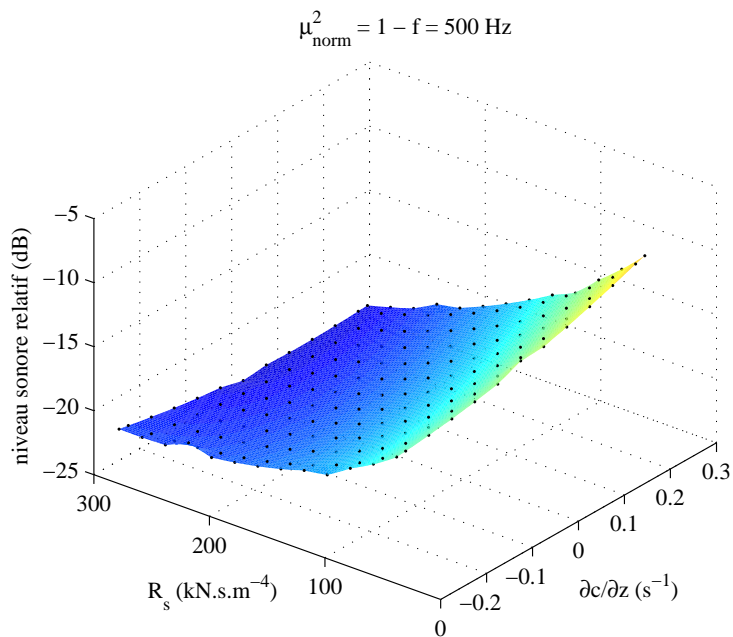


FIG. 4.14 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 1 - f = 500 \text{ Hz}$).
L'erreur quadratique moyenne sur la surface est de 0,03 dB².

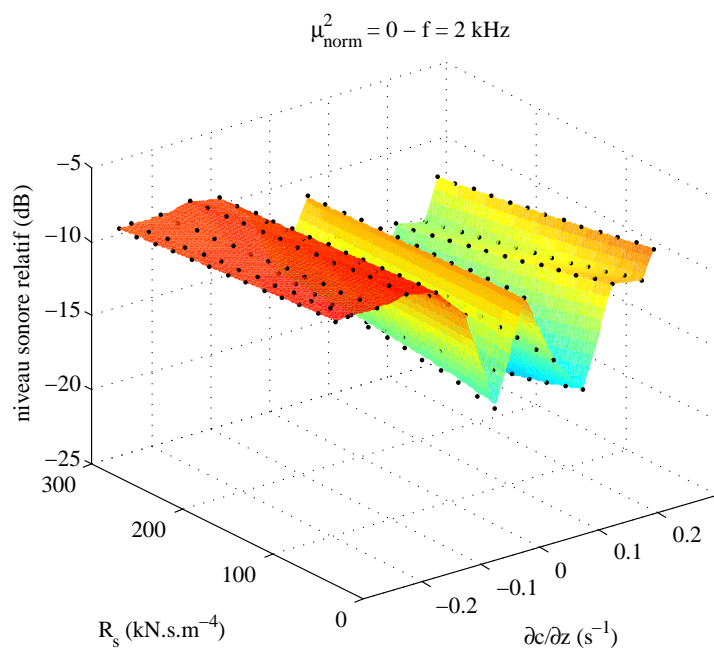


FIG. 4.15 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 0 - f = 2 \text{ kHz}$).
L'erreur quadratique moyenne sur la surface est de 0,3 dB².

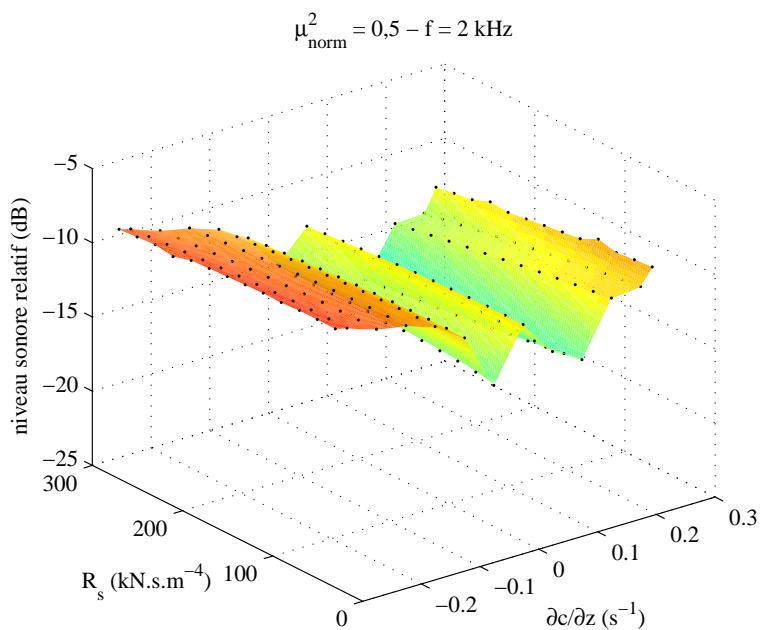


FIG. 4.16 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 0,5 - f = 2 \text{ kHz}$).

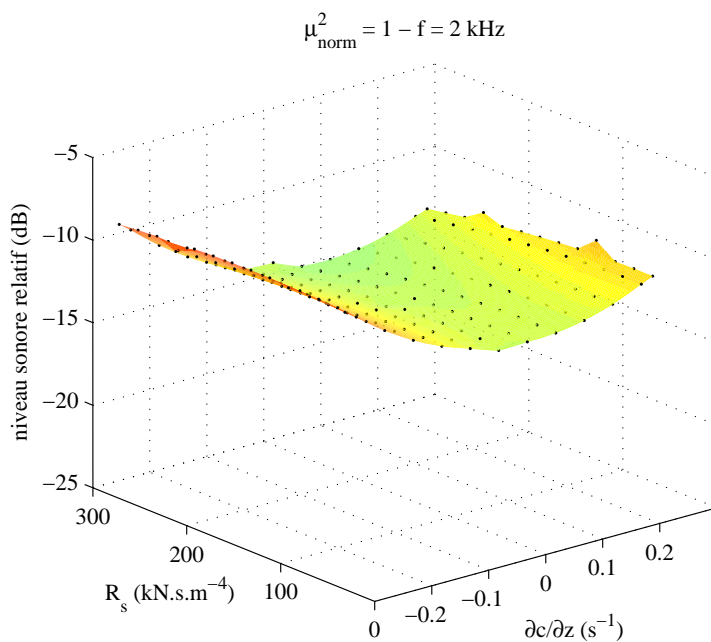


FIG. 4.17 – Surface de réponse à une distance à la source de 50 m ($\langle \mu_{\text{norm}}^2 \rangle = 1 - f = 2 \text{ kHz}$).
L'erreur quadratique moyenne sur la surface est de $0,3 \text{ dB}^2$.

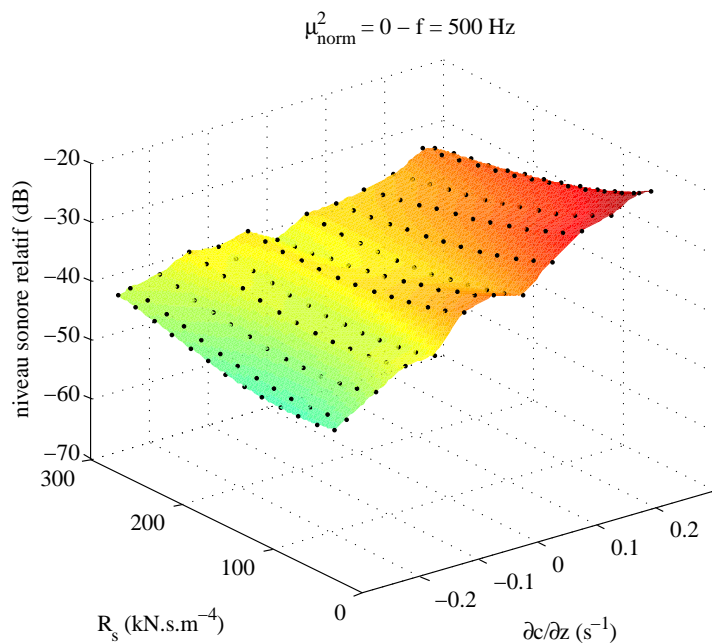


FIG. 4.18 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 0 - f = 500 \text{ Hz}$).
L'erreur quadratique moyenne sur la surface est de 0,1 dB².

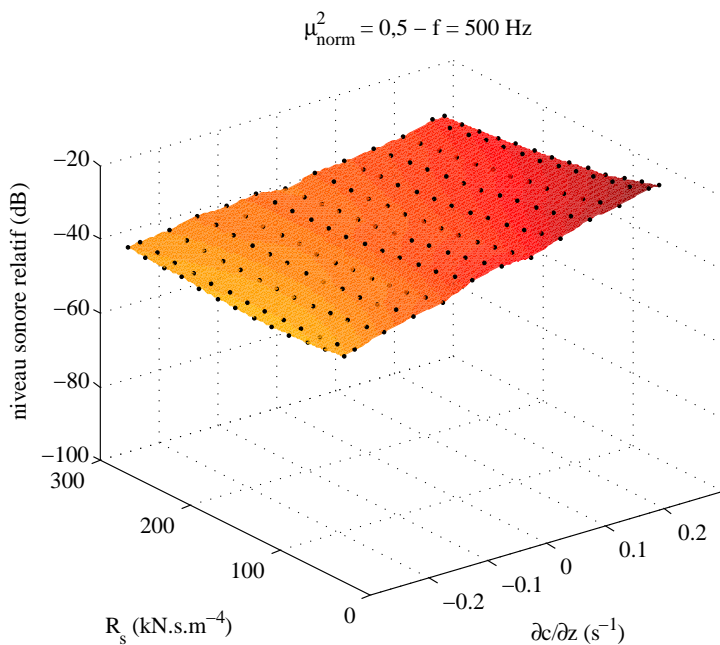


FIG. 4.19 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 0,5 - f = 500 \text{ Hz}$).

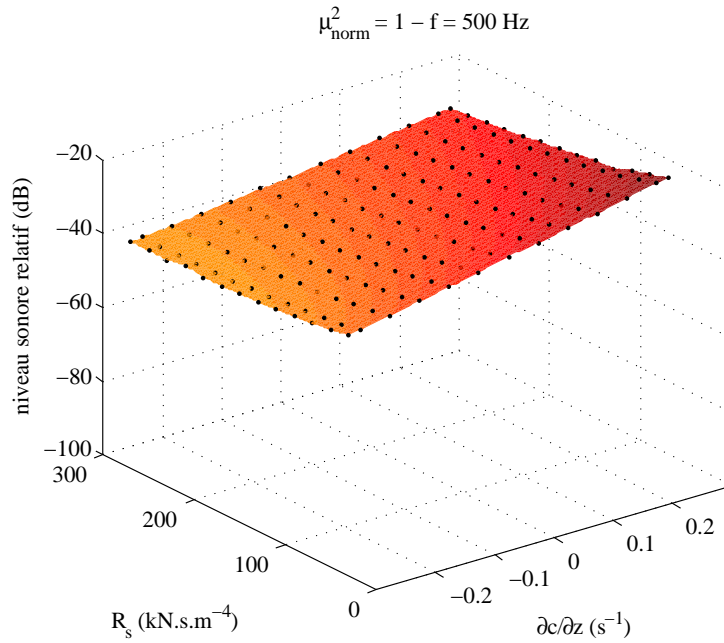


FIG. 4.20 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 1 - f = 500 \text{ Hz}$).
L'erreur quadratique moyenne sur la surface est de 0,1 dB².

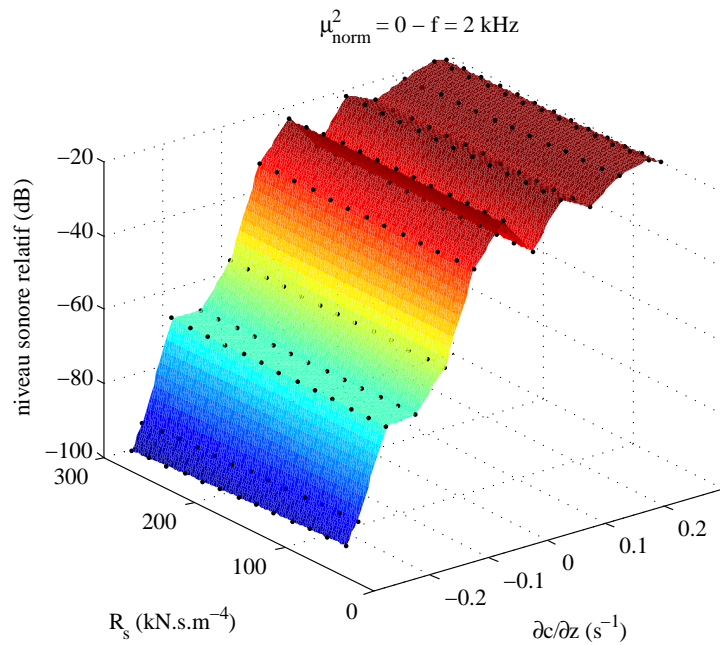


FIG. 4.21 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 0 - f = 2 \text{ kHz}$).
L'erreur quadratique moyenne sur la surface est de 1,8 dB².

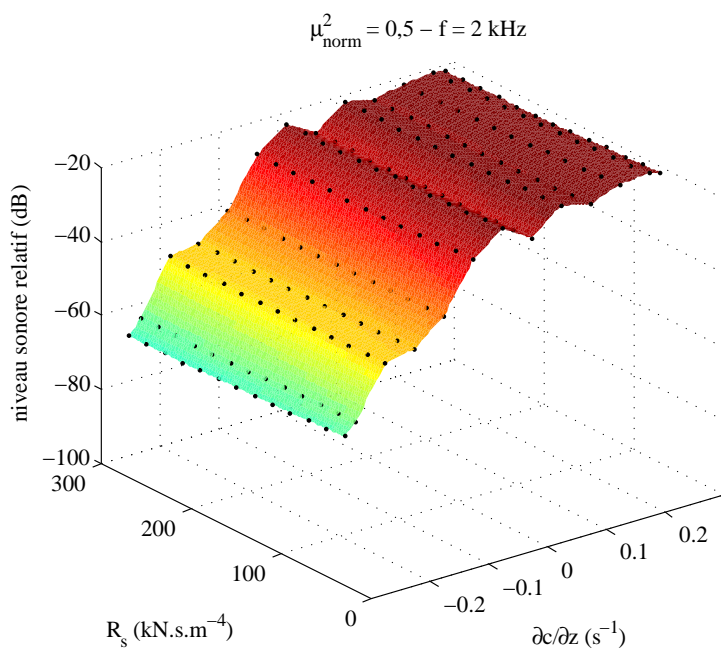


FIG. 4.22 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 0,5 - f = 2 \text{ kHz}$).

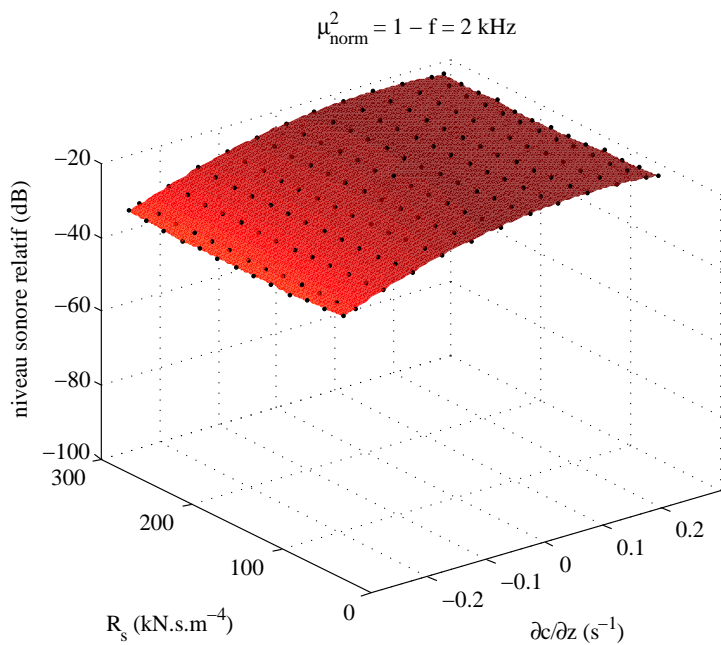


FIG. 4.23 – Surface de réponse à une distance à la source de 150 m ($\langle \mu_{\text{norm}}^2 \rangle = 1 - f = 2 \text{ kHz}$).
L'erreur quadratique moyenne sur la surface est de 1,8 dB².

L'observation de ces surfaces permet de mettre en évidence un certain nombre d'éléments et de vérifier les propriétés physiques bien connues liées aux phénomènes de propagation acoustique en milieu extérieur :

- on vérifie bien que l'effet de turbulence est plus marqué aux hautes fréquences (2 kHz) qu'à 500 Hz. Une valeur plus forte de turbulence apporte une quantité d'énergie non négligeable en conditions défavorables à la propagation (figures 4.22 et 4.23) mais permet également d'apporter de l'énergie dans les zones interférentielles pour des valeurs plus élevées de gradients verticaux de célérité effective. Ce phénomène est bien visible sur les figures 4.16 et 4.17. Cet effet existe à 500 Hz mais il est beaucoup moins marqué ;
- dans la configuration géométrique considérée, les niveaux d'atténuation sonore sont plus sensibles aux effets de sol à 500 Hz qu'à 2 kHz (cf. figures 4.13 et 4.16) ;
- l'erreur quadratique moyenne associée à la surface de réponse augmente avec :
 - la distance source-récepteur,
 - la fréquence.

Cette erreur est très faible en regard de la dispersion associée aux spectres expérimentaux (figure 4.24).

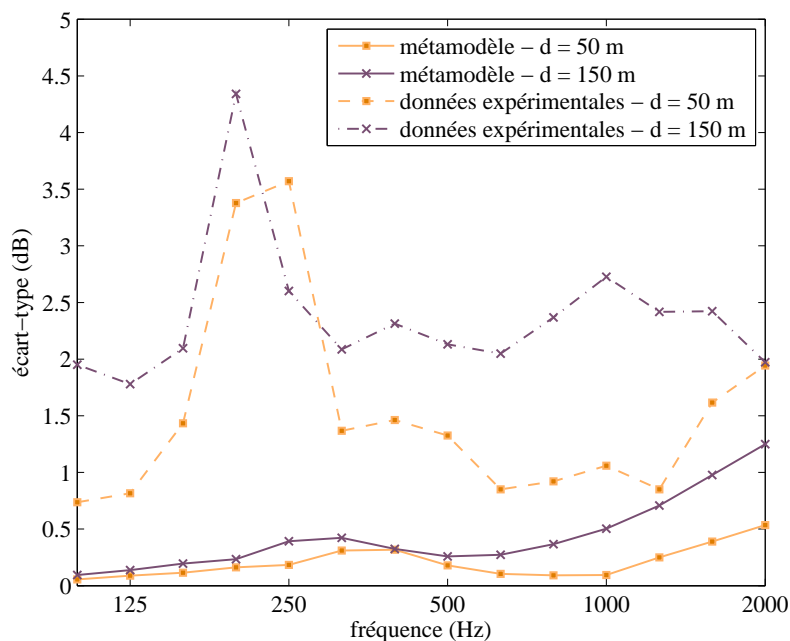


FIG. 4.24 – Comparaison des écarts-type estimés lors de la construction du métamodèle par krigeage et des écarts-type expérimentaux en fonction de la fréquence à (50 m, 2 m) et (150 m, 2 m) de la source.

On vérifie bien que l'écart-type associé au métamodèle a globalement tendance à augmenter avec des fréquences croissantes mais aussi avec la distance. Ces tendances sont cohérentes avec la variabilité physique des spectres d'atténuation relative qui est plus importante à grande distance de la source. De plus, les écarts-type relatifs au métamodèle de krigeage étant dans l'ensemble bien inférieurs aux écarts-type associés aux spectres expérimentaux, on considèrera que le modèle est valide et on négligera dans cette approche l'erreur liée au métamodèle.

1.2.3.2 Validation croisée

Le principe de cette technique de validation de modèles est le suivant : une donnée simulée par le modèle de prévision acoustique $L(x_j)$ est temporairement soustraite au plan d'expériences numériques complet ($L(x_i)_{i=1,\dots,p}$) avec p l'ensemble des points du plan d'expériences numériques). Un métamodèle est alors re-calculé sans ce point et est utilisé pour ré-estimer la valeur de ce point. Cette estimation est notée $L^*(x_j)$. Cette opération est menée pour tous les points $L(x_i)_{i=1,\dots,p}$ pris un à un et permet de calculer la moyenne de l'erreur (différence entre la valeur ré-estimée par le métamodèle $L^*(x_j)$ et la valeur calculée par le modèle de prévision acoustique $L(x_j)$) ainsi que la variance de l'erreur. Moyenne et variance de l'erreur sont respectivement définies par les équations 4.1 et 4.2 :

$$m_e = \frac{1}{p} \sum_{i=1}^p (L^*(x_j) - L(x_j)) \quad (4.1)$$

$$\sigma_e^2 = \frac{1}{p} \sum_{i=1}^p (L^*(x_j) - L(x_j))^2 \quad (4.2)$$

La moyenne de l'erreur doit tendre vers zéro pour vérifier que le métamodèle de krigeage est non-biaisé. Les résultats obtenus sont synthétisés dans le tableau 4.2.

fréquence (Hz)	distance = 50 m		distance = 150 m	
	m_e	σ_e^2	m_e	σ_e^2
100	0	0	0	0
125	0	0	0	0
160	0	0	0	0
200	0	0	0	0,05
250	0	0	0	0,07
320	0	0	-0,01	0,07
400	0	0	0	0,07
500	0	0	0,01	0,07
630	0	0	0,02	0,09
800	0	0	0,01	0,12
1000	0	0	0	0,17
1250	0	0	-0,02	0,25
1600	-0,02	0	-0,06	0,50
2000	-0,04	0,01	-0,16	1,17

TAB. 4.2 – Résultats statistiques de la validation croisée pour les métamodèles relatifs à chaque tiers d'octave à des distances source-récepteur de (50 m, 2 m) et (150 m, 2 m).

Ces résultats statistiques rendent compte d'une erreur très faible que l'on considère une distance source-récepteur de 50 ou 150 m. Cette erreur est nulle pour les métamodèles développés en basses fréquences et elle tend à augmenter progressivement pour des fréquences croissantes. Ce constat n'est pas surprenant et est expliqué par la plus forte variabilité des données et une influence plus marquée de la turbulence aux fréquences élevées. Ces ordres de grandeur restent cependant tout à fait acceptables en regard des écarts-type expérimentaux (figure 4.24). Cette procédure permet ainsi de valider le métamodèle.

1.3 BILAN

Cette phase de l'étude a permis de choisir différents groupes de données et également de développer un métamodèle de simulation opérationnel. Le métamodèle que nous avons développé est

robuste pour simuler rapidement (la durée d'un calcul dans une configuration donnée est inférieure à la seconde alors qu'avec le modèle initial, cette durée est de l'ordre d'une dizaine d'heures) et avec une précision acceptable les configurations voulues. Il est cependant perfectible. En effet, un nombre plus grand de calculs intégrant la turbulence permettrait de renforcer le métamodèle, notamment en conditions défavorables à la propagation. Les temps de calcul étant un obstacle majeur, la parallélisation du code de calculs permettrait de procéder à plus de calculs pour des temps de calcul réduits.

Malgré ces limitations, le métamodèle est complet et permet de reproduire de façon optimale les données calculées par le modèle de simulation acoustique. De plus, aux prévisions du métamodèle peuvent être associées des erreurs liées au métamodèle lui-même. Le métamodèle sera ainsi utilisé en toute confiance comme modèle alternatif au modèle de simulation acoustique basé sur la résolution de l'Équation Parabolique.

Avant de procéder à toute étude comparative entre données expérimentales et données (méta)modélisées sur les groupes de spectres sélectionnés, il est nécessaire de définir les observables d'étude.

2 DÉFINITION DES OBSERVABLES

L'étude menée sur les caractéristiques acoustiques de sol a permis de valider la robustesse de la méthode de définition d'observables par ACPf (chapitre 3 section 2.2) et c'est pourquoi nous allons de nouveau appliquer cette méthode. Les observables définies correspondent aux coefficients de l'ACPf dont la sélection est basée sur la valeur des valeurs propres et/ou le pourcentage de variation expliquée par les fonctions propres de l'ACPf. La sélection des coefficients d'ACPf a été menée selon la règle de Kaiser [107] (figure 4.25).

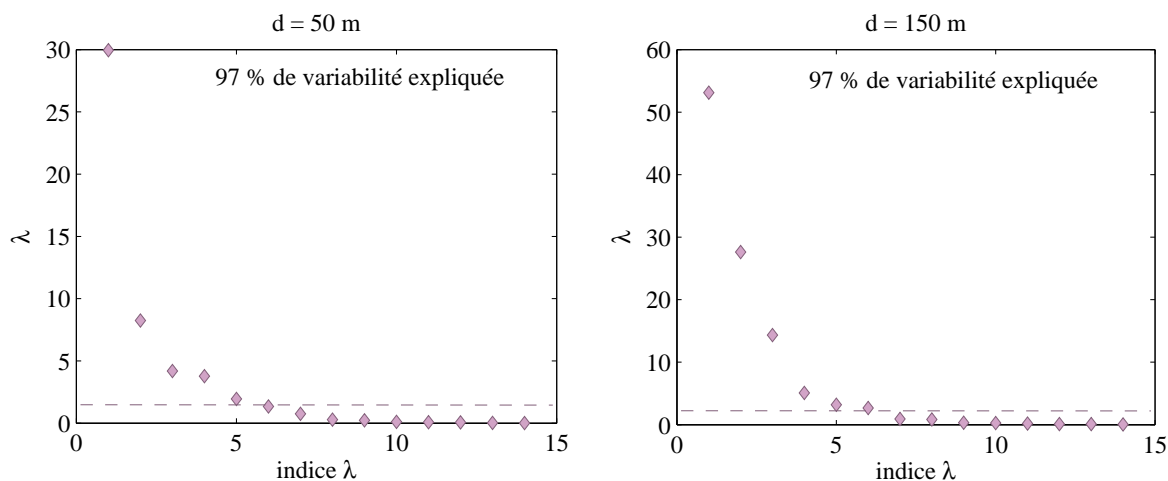


FIG. 4.25 – Valeurs propres de l'ACPf en appliquant la règle de Kaiser pour les spectres mesurés à (50 m, 2 m) et (150 m, 2 m) de la source. Dans les deux cas, les six premières valeurs propres permettent d'expliquer 97 % de la variance.

La règle de Kaiser stipule que seules les fonctions dont les valeurs propres sont supérieures à l'unité doivent être retenues. Dans notre cas, les six premières valeurs propres sont supérieures à l'unité et permettent au total d'expliquer 97 % de la variabilité des données. Les fonctions propres correspondantes sont représentées par la figure 4.26. Six coefficients seront alors conservés et utilisés

en tant qu'observables du processus de quantification d'incertitudes. On réduit donc de 14 à 6 la dimension du problème à étudier.

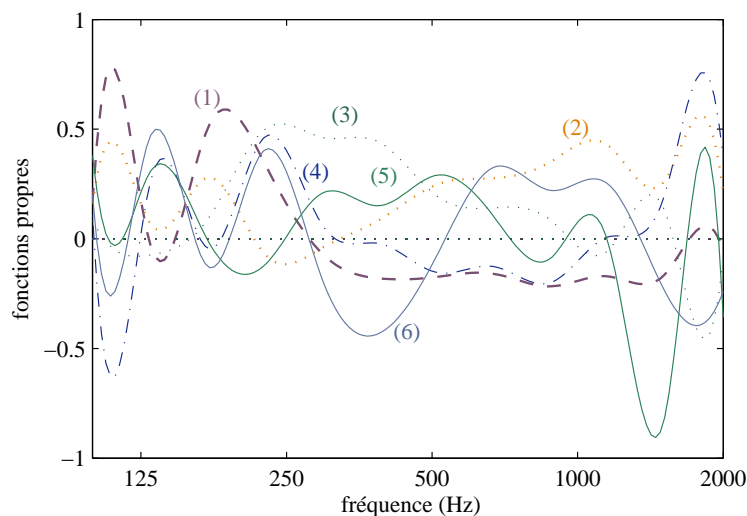


FIG. 4.26 – Six premières fonctions propres de l'ACPF pour les spectres mesurés à (150 m, 2 m) de la source.

3 CALIBRATION DU SYSTÈME

3.1 DÉFINITION DU MODÈLE STATISTIQUE DE CALIBRATION

Le système à l'étude est non-linéaire (modèle complexe), multidimensionnel (des groupes de spectres sont considérés et trois paramètres incertains ainsi que leurs écarts-type associés doivent être estimés) et hétéroscédastique (les variances des résidus dépendent des observables considérées). La formulation de la fonction coût, *i.e.* la fonction de log-vraisemblance, est la fonction définie au chapitre 2 par l'équation 2.5, en considérant un nombre d'observables $n_{obs} = 6$ (observables définies par ACPf).

3.2 ESTIMATION DES PARAMÈTRES INCERTAINS

Le processus de calibration est appliqué aux groupes de spectres définis en section 1.1.1 à 50 et 150 m de la source. Les paramètres estimés associés à ces groupes sont représentatifs de conditions de propagation moyennes ; ces paramètres sont synthétisés dans le tableau 4.3.

	distance = 50 m	distance = 150 m
$\partial c / \partial z$	$7,1 \cdot 10^{-3}$	0,05
R_s	103	111
$\langle \mu_{norm}^2 \rangle$	0,31	0,34
$\langle \mu^2 \rangle$	$3,1 \cdot 10^{-6}$	$3,4 \cdot 10^{-6}$

TAB. 4.3 – Paramètres physiques estimés à 50 m et 150 m de la source en considérant un groupe de spectres.

Ces paramètres rendent compte d'une légère différence à 50 et 150 m, notamment lorsque l'on considère le gradient vertical de célérité du son. En effet, à 50 m de la source ce paramètre rend compte de conditions très proches des conditions homogènes ($\frac{\partial c}{\partial z} \sim 0$) alors qu'à 150 m le paramètre estimé est égal à $0,05 \text{ s}^{-1}$ et rend compte de conditions plutôt favorables à la propagation. Cette observation pourrait être expliquée par une pente de profil linéaire équivalent variable avec la distance. Davantage d'explications sont données en section 5.1 en regard des écarts-type d'estimation. La valeur estimée pour le groupe de spectres mesurés à 50 m est proche de la valeur moyenne du gradient vertical de célérité mesuré (figure 4.2).

Concernant les valeurs de résistance spécifique au passage de l'air R_s , l'écart entre les valeurs (8 kN.s.m^{-4}) est en deçà de la précision expérimentale associée à ce paramètre qui est de l'ordre de 10 kN.s.m^{-4} ce qui ne témoigne pas d'une quelconque variabilité spatiale du paramètre. Ces valeurs estimées sont inférieures aux valeurs expérimentales d'environ 30 kN.s.m^{-4} , ce qui traduit l'intérêt de calibrer les paramètres du (méta)modèle sur les données expérimentales.

Les valeurs du paramètre de turbulence (normalisé ou non) sont très proches et ne rendent pas compte d'une variabilité spatiale du paramètre. Par contre elles témoignent de la nécessité de prendre en compte ce phénomène même dans des conditions homogènes ou favorables.

Le recours à un paramètre moyen estimé θ et à des paramètres variables bien connus (d_j), en tant que paramètres d'entrée du (méta)modèle, donne accès à une partie de l'incertitude épistémique associée au système physique étudié, *i.e.* la part d'incertitude expliquée par la variation des paramètres d_j . Ces paramètres connus sont la température et l'humidité relative variables dans le temps, qui ont notamment une incidence sur la valeur de l'absorption atmosphérique. Les figures 4.27 et 4.28 traduisent cette part d'incertitude dans le domaine spectral.

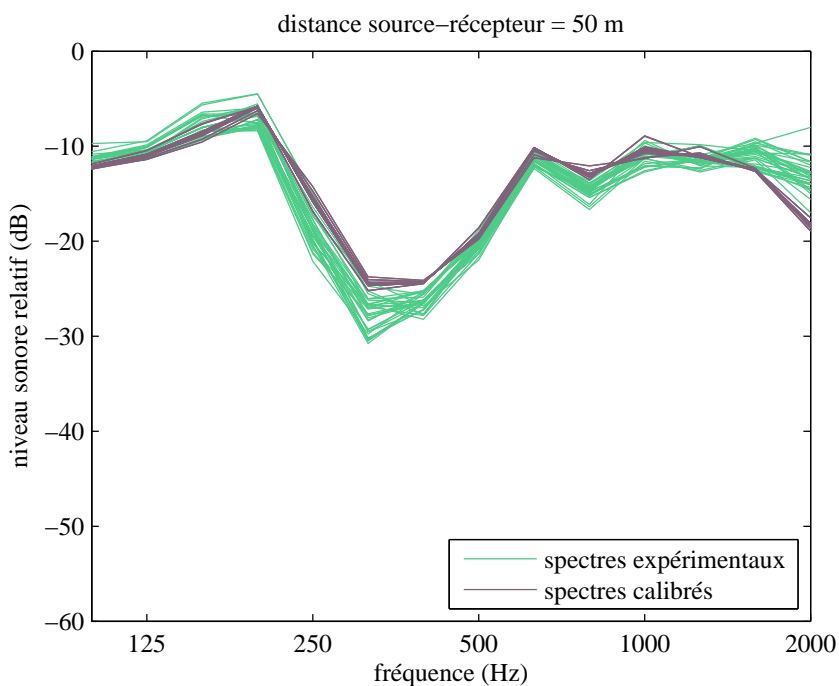


FIG. 4.27 – Spectres expérimentaux et spectres simulés par le métamodèle pour les paramètres moyens estimés θ pour des valeurs de d_j variables à 50 m de la source.

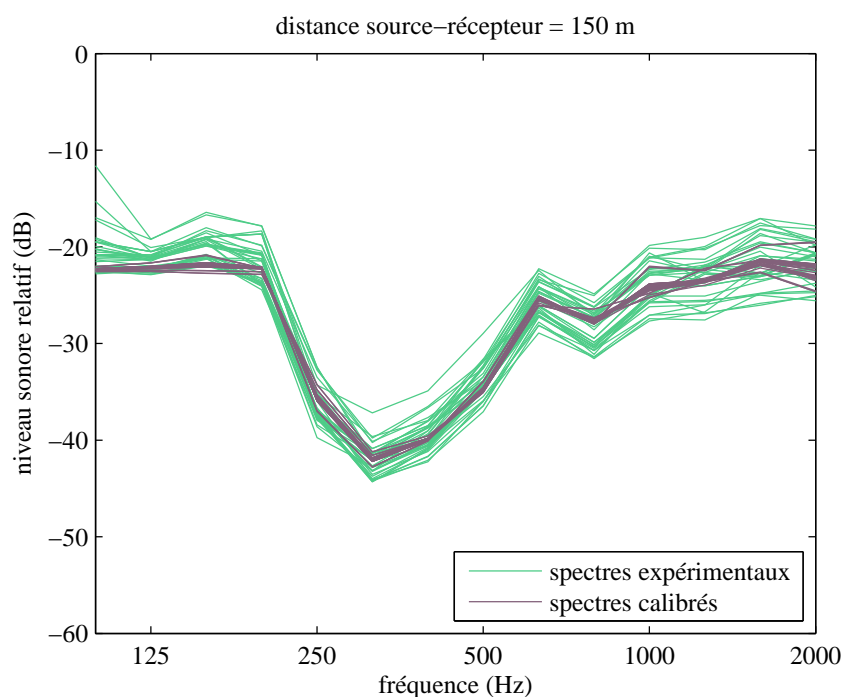


FIG. 4.28 – Spectres expérimentaux et spectres simulés par le métamodèle pour les paramètres moyens estimés θ pour des valeurs de d_j variables à 150 m de la source.

Ces représentations spectrales rendent compte de la présence de biais non négligeables à 50 m de la source, notamment entre 250 et 500 Hz (de l'ordre de 3 dB) et à 150 m de la source en basses fréquences (d'environ 3 dB également). Un élément peut potentiellement expliquer ces divergences : le (méta)modèle peut ne pas simuler de telles situations réelles, ces limites n'étant pas directement liées au métamodèle mais au modèle de prévision acoustique lui-même. Cette remarque est justifiée par l'observation des grandeurs d'intérêt en regard des paramètres d'entrée du système car le domaine des spectres expérimentaux n'est pas recouvert par les spectres modélisés.

Les paramètres d_j ne permettent d'expliquer qu'une part "minime" de la dispersion totale des spectres expérimentaux. Afin de rendre compte du taux de variabilité expliquée par ces paramètres, on se propose de calculer le coefficient de détermination R^2 entre les spectres mesurés et les spectres (méta)modélisés. En moyenne sur l'ensemble des bandes de tiers d'octave, $R^2 = 0,12$ avec un écart-type $\sigma_{R^2} = 0,11$ à 50 m de la source et $R^2 = 0,22$ avec un écart-type $\sigma_{R^2} = 0,13$ à 150 m de la source. Ces proportions traduisent une influence plus forte des paramètres d_j à 150 m, ce qui s'explique aisément par l'effet plus marqué de l'absorption atmosphérique à grande distance. Ce résultat est donc tout à fait cohérent avec la physique des phénomènes. Cependant, ces paramètres ne suffisent pas à expliquer l'incertitude et la variabilité totales associées aux spectres expérimentaux et c'est pourquoi des procédures de bootstrap sont mises en œuvre suite à la validation du système.

4 VALIDATION DU SYSTÈME

Dans cette phase de validation, on s'assure que les hypothèses sur l'indépendance des résidus et le caractère gaussien de leur distribution sont bien vérifiées. Les outils de vérification sont définies en section 4.3 du chapitre 2. Les matrices de corrélation des résidus sont ainsi calculées à 50 et 150 m de

la source (tableaux 4.4 et 4.5).

$\rho_{\varepsilon_i, \varepsilon_j}$	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
ε_1	1	0,31	-0,17	-0,1	0,5	0,17
ε_2		1	0,39	-0,17	-0,08	0,09
ε_3			1	-0,52	-0,7	-0,44
ε_4				1	-0,07	0,06
ε_5					1	0,49
ε_6						1

TAB. 4.4 – Matrice de corrélation des résidus calculés à une distance à la source de 50 m.

A 50 m de la source, les coefficients de corrélation entre les résidus sont faibles dans l'ensemble, ce qui traduit leur indépendance. On relèvera tout de même des valeurs un peu plus élevées pour les résidus relatifs aux observables 3 et 4, aux observables 3 et 6 ainsi qu'aux observables 5 et 6. Les valeurs de ces coefficients ne sont pas significatives mais il est cependant important de noter leur existence.

$\rho_{\varepsilon_i, \varepsilon_j}$	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
ε_1	1	-0,08	-0,14	-0,4	0,33	-0,5
ε_2		1	0,52	0,37	0,16	-0,09
ε_3			1	0,37	0,17	0,33
ε_4				1	0,13	0,47
ε_5					1	-0,49
ε_6						1

TAB. 4.5 – Matrice de corrélation des résidus calculés à une distance à la source de 150 m.

Comme pour les résidus à 50 m de la source, les coefficients de corrélation relatifs aux résidus à 150 m de la source sont globalement indépendants. On note cependant trois valeurs se démarquant légèrement des autres mais elles sont peu significatives.

On vérifie le caractère gaussien des distributions des résidus par des représentations graphiques appelées “p-p plots” (ou “probability-probability plots”) particulièrement bien adaptées dans le cas de petits échantillons (ici le nombre de spectres pour chaque groupe est de 34). Ces représentations sont données par les figures 4.29 et 4.30. Les points, correspondant à la distribution empirique, doivent être les plus proches de la ligne en pointillés, correspondant à la distribution gaussienne théorique. Ces graphes montrent dans l'ensemble un bon alignement des courbes empiriques sur les courbes théoriques.

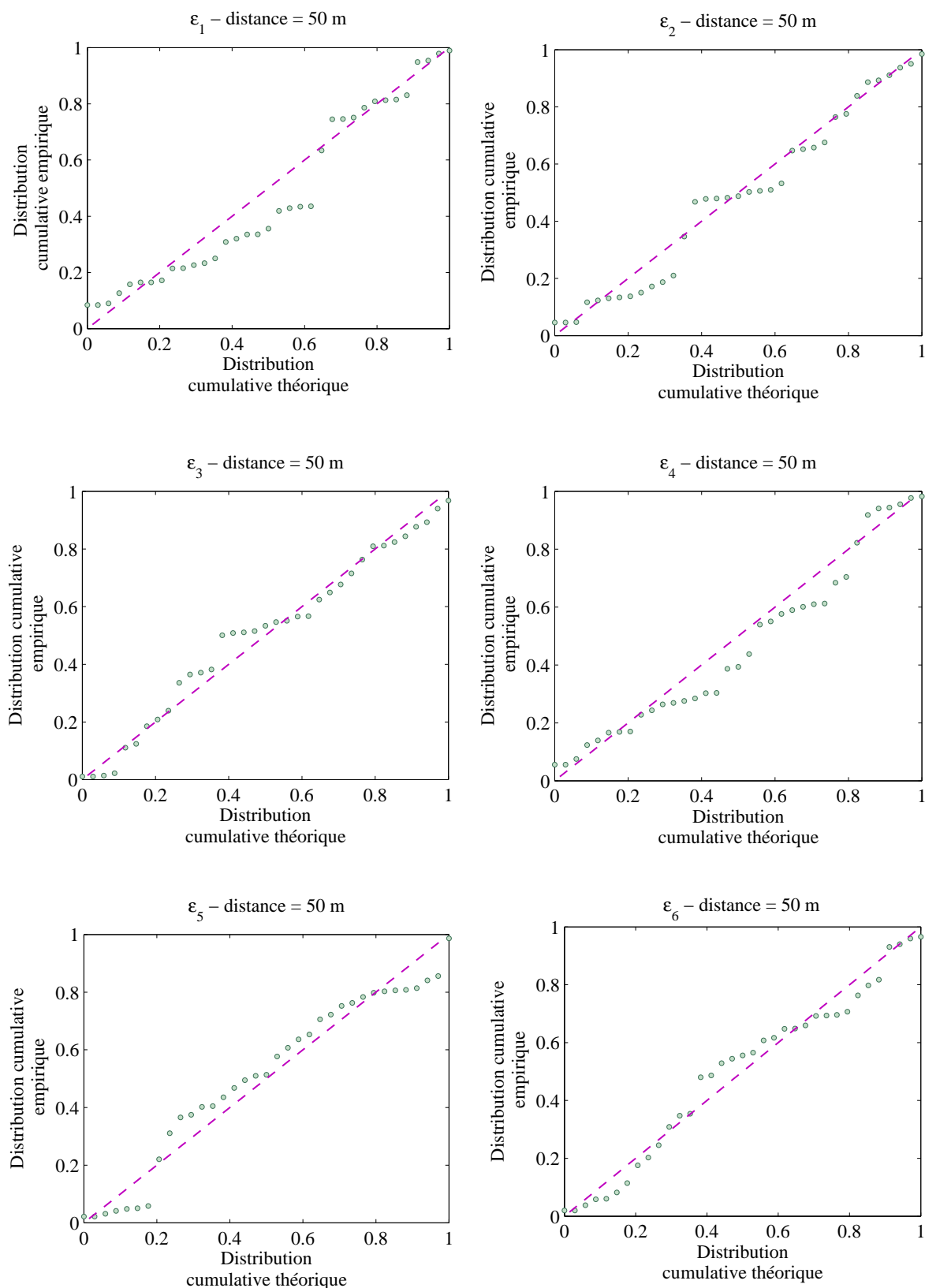


FIG. 4.29 – “p-p plots” des résidus relatifs à chaque observable calculés à une distance à la source de 50 m – la courbe en pointillés correspond à une distribution gaussienne théorique et les points à la distribution des résidus.

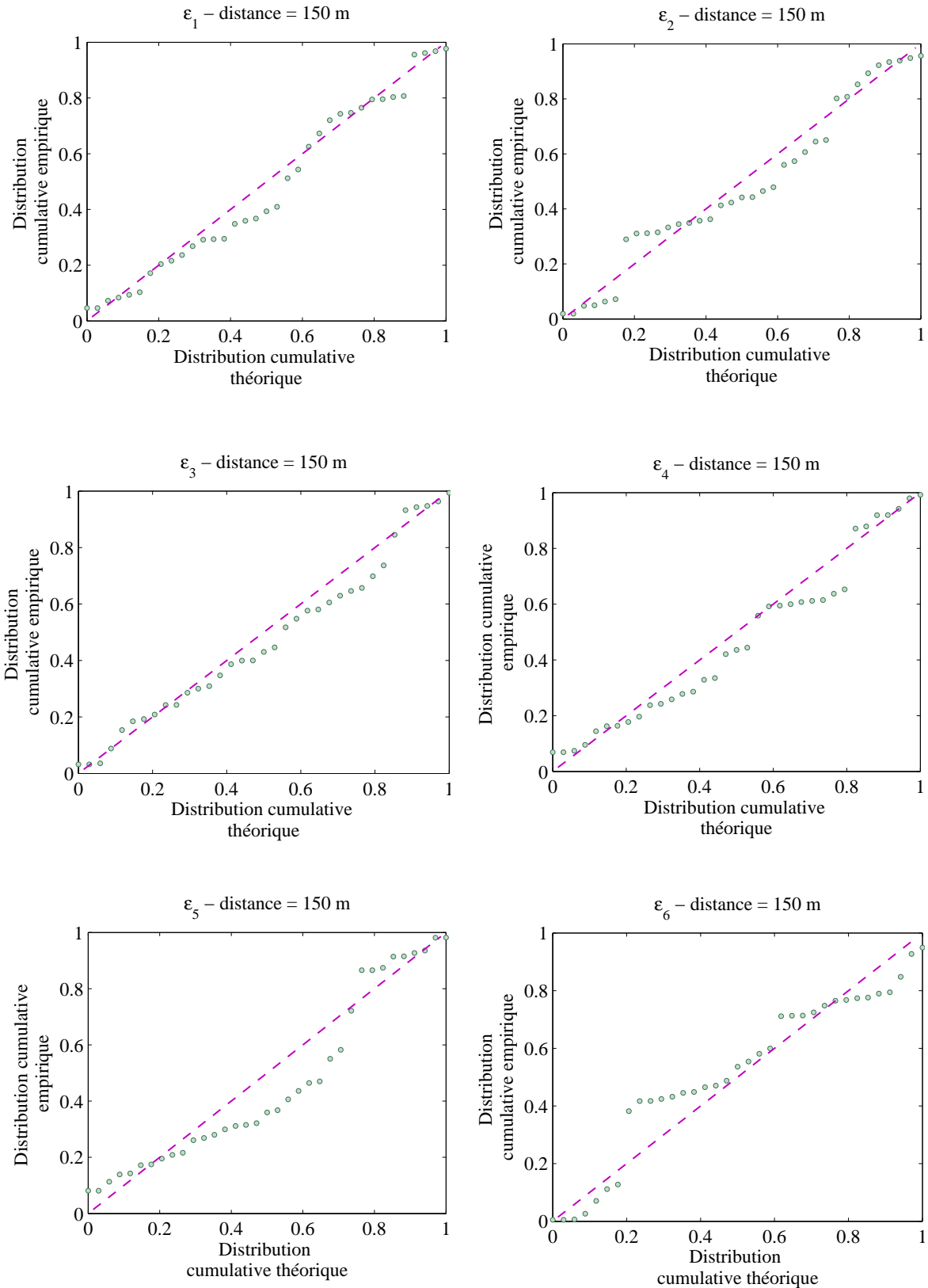


FIG. 4.30 – “p-p plots” des résidus relatifs à chaque observable calculés à une distance à la source de 150 m – la courbe en pointillés correspond à une distribution gaussienne théorique et les points à la distribution des résidus.

Des divergences sont cependant observées pour les résidus relatifs à l'observable 1 à 50 m de la source. Etant donné que l'observable 1 correspond au premier coefficient de l'ACPF, qui rend compte de la plus forte variabilité des données, ces divergences doivent se traduire par des écarts marqués entre les spectres estimés et les spectres expérimentaux. Cette remarque est d'ailleurs en adéquation avec les observations faites sur la figure 4.27 en section 3.2 où un biais important existe. L'écartement de la distribution des résidus d'une distribution gaussienne se traduit bien par des écarts non négligeables entre spectres estimés et spectres expérimentaux dans des zones de forte variabilité. Des écarts sont également observés pour les résidus relatifs aux observables 5 et 6 à 150 m de la source. Ces écarts ont une incidence sur les spectres ce qui a d'ailleurs été observé en section 3.2 où un biais apparaît.

En regard des résultats obtenus en termes de corrélation et de normalité des résidus, on considère que les hypothèses initiales sont bien vérifiées. Dans cette phase de validation, on se contente de vérifier ces hypothèses. En effet, ici un seul modèle statistique est considéré et le calcul de critères de validation tels que les critères AIC, BIC et le coefficient de détermination R^2 ne présente pas d'intérêt particulier (pas de prise de décision sur le choix d'un modèle statistique). De plus, le système statistique étant non linéaire à variables corrélées, la vérification de l'identifiabilité est complexe à mettre en place en pratique.

5 PRÉVISION

Le modèle statistique développé étant validé, il est maintenant possible de procéder à la phase de prévision. Celle-ci est basée sur le calcul des variances d'estimation puis des variances de prévision.

5.1 VARIANCE D'ESTIMATION DES PARAMÈTRES

Afin d'accéder à la variance d'estimation qui rend compte de la variabilité physique des données, la méthode de *bootstrap paramétrique* (chapitre 2 section 5.1.2) est mise en œuvre. Un certain nombre de spectres sont recréés intégralement par des méthodes de Monte-Carlo à partir des moyennes et des écarts-type des groupes des spectres expérimentaux à l'étude pour constituer un nouveau groupe de spectres. Les paramètres associés à ce nouveau groupe sont estimés par mise en œuvre du processus de calibration sous incertitude. Cette opération est répétée cinquante fois afin d'obtenir un échantillonnage de paramètres représentatifs du groupe de spectres expérimentaux. Cette démarche a permis d'obtenir les estimations synthétisées dans le tableau 4.6.

	distance = 50 m		distance = 150m	
	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$	$\langle \hat{\theta} \rangle$	$\sigma_{\langle \hat{\theta} \rangle}$
$\partial c / \partial z$	8.10^{-3}	0,05	0,05	0,03
R_s	103	8	109	8
$\langle \mu_{norm}^2 \rangle$	0,34	0,17	0,39	0,17
$\langle \mu^2 \rangle$	$3,4.10^{-6}$	$1,7.10^{-6}$	$3,9.10^{-6}$	$1,7.10^{-6}$

TAB. 4.6 – Paramètres physiques et écarts-type d'estimation estimés par la méthode de bootstrap à 50 m et 150 m de la source en considérant un groupe de spectres représentatif de conditions de propagation homogènes.

Ces résultats montrent la convergence des paramètres estimés par procédure de bootstrap $\langle \hat{\theta} \rangle$ vers les valeurs estimées à partir du groupe de spectres initial $\hat{\theta}$ (tableau 4.3). Les variances relatives à l'indice de turbulence normalisé $\langle \mu_{norm}^2 \rangle$ et à la résistance spécifique au passage de l'air R_s sont similaires quelque soit la distance à la source considérée, ce qui ne rend pas compte d'une éventuelle

variabilité spatiale sur ces paramètres. De plus, la variance d'estimation associée à la résistance spécifique au passage de l'air R_s est tout à fait cohérente avec les variances obtenues avec le modèle de Delany et Bazley au chapitre 3.

On retrouve ici encore une différence entre les valeurs de gradients verticaux de célérité en fonction de la distance ainsi qu'une différence entre les écarts-type de bootstrap associés à ce paramètre. Cette remarque traduit potentiellement qu'à un groupe de spectres donné - initialement considéré comme représentatif de conditions de propagation homogènes - peut être associé un paramètre présentant une forte variabilité. En effet, comparativement aux valeurs expérimentales du gradient vertical de célérité du son (figure 4.3), l'écart-type estimé par bootstrap est 3 à 5 fois supérieur. Ce paramètre rend compte également d'une variabilité spatiale : la valeur à 50 m est proche des conditions homogènes, ce qui coïncide avec les valeurs expérimentales de gradient mesurées à 75 m de la source, alors que la valeur à 150 m s'apparente plutôt à des conditions favorables à la propagation : les mesures effectuées en un point du site ne sont donc pas représentatives de l'ensemble du site et ne devraient pas être utilisées comme telles. Ces résultats remettent en cause l'hypothèse selon laquelle les gradients verticaux de célérité du son mesurés à 75 m sont représentatives de l'ensemble du site. La calibration des paramètres a donc permis de mettre en évidence cette propriété et il serait donc préférable pour de futures études de ne pas faire cette hypothèse concernant le gradient vertical de célérité du son. La variance d'estimation permet donc une interprétation plus robuste des paramètres physiques estimés.

5.2 INCERTITUDE GLOBALE

La phase finale du processus de quantification d'incertitude consiste maintenant à combiner l'ensemble des résultats précédents. En effet, à partir des différentes estimations de paramètres, il est possible d'accéder à l'incertitude liée à l'estimation des paramètres (qui rend compte de la variabilité physique des données expérimentales), à l'incertitude épistémique (qui peut être expliquée en partie par les paramètres d_j) et finalement à l'incertitude et à la variabilité inexpliquées (résidus).

5.2.1 Incertitude et variabilité liées à l'estimation des paramètres

A partir des paramètres moyens et des écarts-type d'estimation obtenus par la méthode de bootstrap notés $\sigma_{\langle \hat{\theta} \rangle}$, on génère des distributions de paramètres par une méthode de Monte-Carlo. Ces distributions auxquelles on associe les distributions de paramètres d_j - générées à partir de leurs moyennes et écarts-type - constitueront ainsi un nouveau jeu de paramètres à appliquer en entrée du (méta)modèle afin de propager l'incertitude sur les paramètres dans le modèle et ainsi rendre compte des effets de la variabilité des paramètres dans le domaine spectral (figures 4.31 et 4.32).

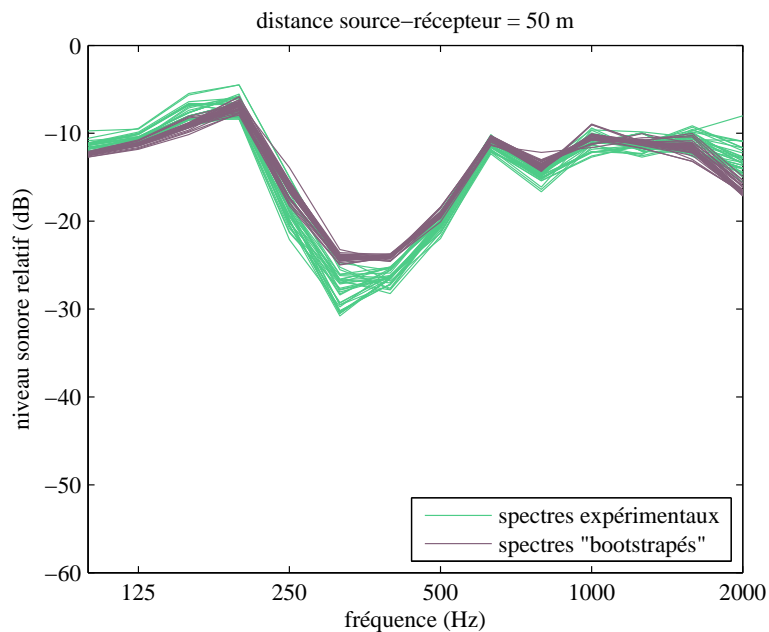


FIG. 4.31 – Spectres expérimentaux et spectres calculés par le métamodèle pour les paramètres estimés par bootstrap et pour des valeurs de d_j variables à 50 m de la source.

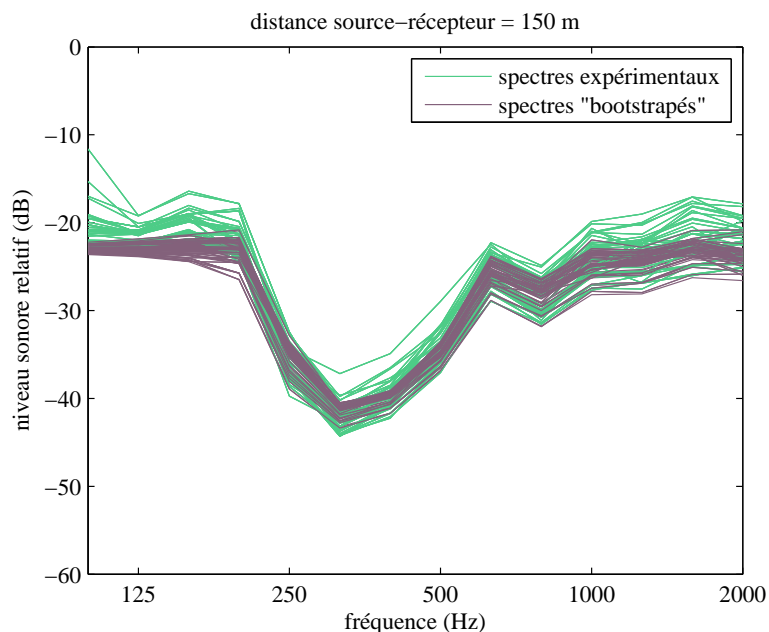


FIG. 4.32 – Spectres expérimentaux et spectres calculés par le métamodèle pour les paramètres estimés par bootstrap et pour des valeurs de d_j variables à 150 m de la source.

La propagation de la variance d'estimation à travers le système rend compte désormais d'une variabilité un peu plus prononcée des données modélisées, notamment à 150 m de la source où la dispersion associée aux spectres est bien plus grande que celle observée précédemment en tenant compte uniquement de la variabilité des paramètres connus d_j (figure 4.27). Cet effet est logiquement moins prononcé à 50 m, distance à laquelle les effets propagatifs sont moins marqués. Afin de rendre

compte de l'apport de la procédure de bootstrap pour l'estimation de l'incertitude et de la variabilité, on calcule le coefficient de détermination entre données (méta)modélisées et données expérimentales. Les résultats sont les suivants :

- à 50 m de la source, $R^2 = 0,44 \pm 0,15$;
- à 150 m de la source, $R^2 = 0,50 \pm 0,16$.

En regard des résultats obtenus précédemment à la section 3.2, l'augmentation des coefficients de détermination aux deux distances source-récepteur considérées rend compte d'une nette amélioration du contenu prédictif du modèle grâce à la prise en compte des variances des paramètres calculées par bootstrap. Une part de l'incertitude et de la variabilité physique associées aux spectres est ainsi expliquée à la fois par les écarts-type $\sigma_{\hat{\theta}}$ et par la variation des paramètres d_j . La part (non négligeable) d'incertitude et de variabilité inexpliquée est rejetée dans les résidus.

5.2.2 Incertitude et variabilité inexpliquées

Les résidus réunissent toute l'information qui n'a pas pu être expliquée par le processus de calibration sous incertitudes ni par les paramètres physiques bien connus mais moins influents (d_j). Ils intègrent notamment les éléments induits par le modèle de prévision acoustique (qui ne rend pas parfaitement compte de la réalité), ainsi qu'une part d'incertitude épistémique. L'incertitude résiduelle est représentée par la variance des résidus $\sigma_{\varepsilon_j}^2$ calculée comme suit :

$$\sigma_{\varepsilon_j}^2 = \sigma_{Y_j}^2 - \sigma_{H_j}^2. \quad (4.3)$$

$\sigma_{Y_j}^2$ et $\sigma_{H_j}^2$ correspondent respectivement à la variance des observables expérimentales et à la variance du modèle. Cette dernière est issue de la propagation de la variance d'estimation des paramètres θ mais aussi de la variance des paramètres d_j à travers le modèle.

5.2.3 Incertitude globale

Les étapes précédentes de l'étude ont montré que les estimations de paramètres rendaient compte de la présence de biais entre les données expérimentales et les données modélisées. Ces biais, dus essentiellement aux limites du modèle de prévision acoustique, se traduisent à travers des moyennes résiduelles très différentes de 0 pour certaines fréquences (biais moyens de l'ordre de 3 dB pour certaines fréquences).

On se propose de réduire ces écarts, *i.e.* de recentrer les résidus sur 0, en corrigeant les spectres modélisés par une dérive. Cette approche a été choisie alternativement à une approche de recalage des spectres au cours du processus de calibration (*i.e.* en intégrant des paramètres à estimer associés directement aux observables et permettant de minimiser les résidus). Elle favorise en effet le recalage des biais au détriment de l'estimation des paramètres influents ; un recalage de biais est plus direct et plus simple d'un point de vue de l'optimisation. L'approche choisie consiste donc à recentrer les distributions des résidus sur 0 par un correctif, calculé après recalage, apporté à chaque fréquence. Ce correctif est ensuite associé au modèle dans un cadre prédictif.

A partir des variances et des corrections de biais sont calculés des intervalles de confiance (IC) à 95 % représentés par les figures 4.33 et 4.34. Les intervalles de confiance représentés traduisent l'écart-type du modèle σ_{H_j} (écart-type calculé à partir de l'écart-type d'estimation de θ et de la variation des d_j) et l'écart-type des données expérimentales σ_{Y_j} qui traduit l'incertitude globale sur l'estimation. L'écart entre ces intervalles de confiance rend compte de l'écart-type résiduel σ_{ε_j} relatif à l'information inexpliquée.

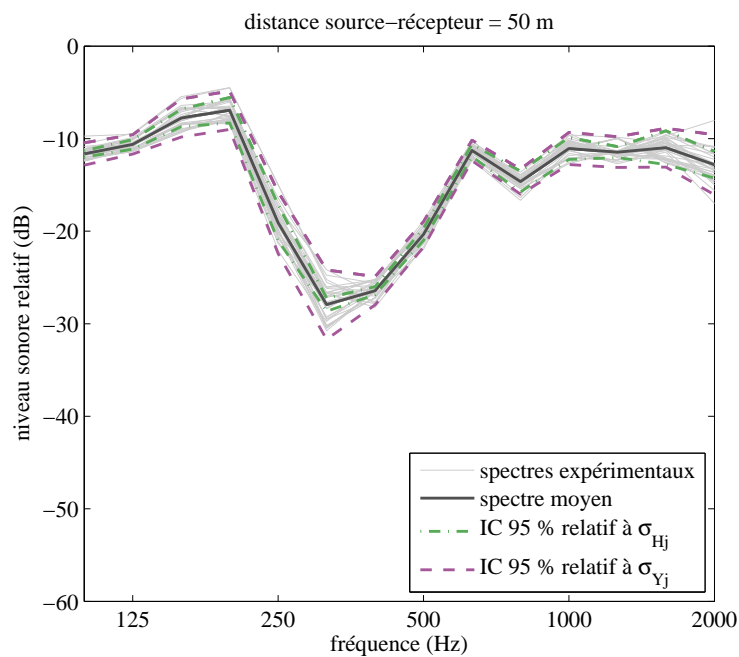


FIG. 4.33 – Représentation des spectres expérimentaux, du spectre moyen modélisé et centré et des intervalles de confiance à 95 % relatifs à l'écart-type du modèle et à l'écart-type associé à l'incertitude globale à 50 m de la source. La différence entre les lignes pointillées de part et d'autre du spectre moyen correspond à l'écart-type résiduel.

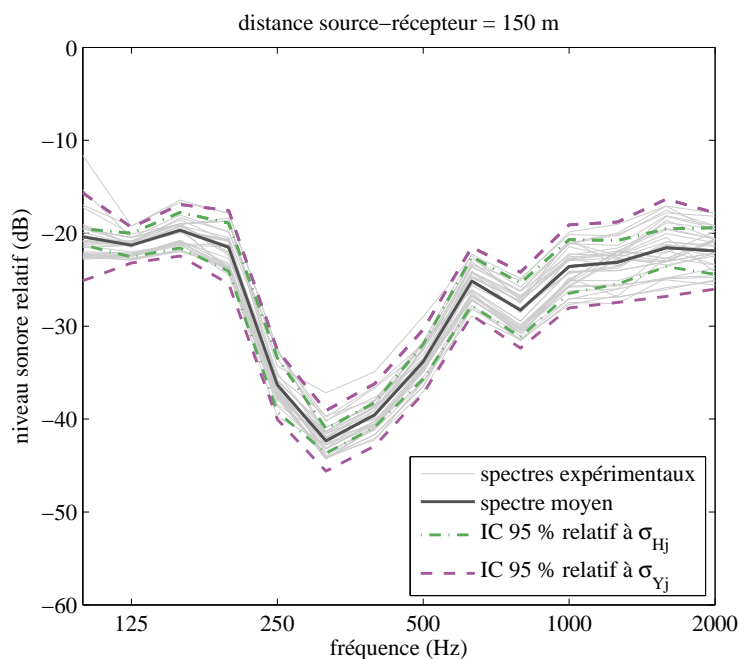


FIG. 4.34 – Représentation des spectres expérimentaux, du spectre moyen modélisé et centré et des intervalles de confiance à 95 % relatifs à l'écart-type du modèle et à l'écart-type associé à l'incertitude globale à 150 m de la source. La différence entre les lignes pointillées de part et d'autre du spectre moyen correspond à l'écart-type résiduel.

Des intervalles de confiance peuvent désormais être associés au modèle de prévision acoustique traduisant une réelle incertitude et variabilité physique. Ces intervalles de confiance dépendent bien entendu de la distance à la source considérée ainsi que de la fréquence.

5.3 BILAN

Pour ces conditions de propagation spécifiques (dans l'étude actuelle des conditions de propagation homogènes), le processus de calibration permet de déterminer une *incertitude globale*. Une grande partie de cette *incertitude globale* est liée à la variabilité des phénomènes qui peut être en partie expliquée par la variance d'estimation des paramètres influents (calculée par bootstrap). Chaque élément apporte une information complémentaire permettant d'expliquer une part de cette incertitude globale, comme par exemple les paramètres de température et d'humidité relative qui viennent corriger les spectres (absorption atmosphérique). Un intervalle de confiance est désormais associé aux calculs effectués dans une configuration donnée et pour des conditions de propagation bien spécifiques.

Le modèle de prévision acoustique semble présenter certaines limites qui ne sont pas directement liées au métamodèle mais au modèle de prévision acoustique lui-même : certaines situations observées physiquement dans la nature ne sont pas accessibles par la modélisation. De ce fait, des biais importants sont observés entre les spectres modélisés et les spectres expérimentaux. Afin de renforcer le contenu prédictif du modèle une correction des biais par une dérive est apportée aux spectres modélisés (après calibration) en regard des données expérimentales.

6 CONCLUSION DU CHAPITRE

Ce chapitre avait pour objectif de présenter la mise en place d'un processus *complet* de quantification d'incertitude avec notamment toute la problématique liée à la construction d'un métamodèle et la phase de prévision qui n'avait pas été abordée dans la partie précédente (chapitre 3).

Ce cas d'application permet de dégager les différentes étapes clés d'un tel processus de quantification d'incertitude :

- la première phase de l'étude consiste à créer des groupes de spectres. Plusieurs approches peuvent être envisagées en découpant la base de données dans le temps pour mettre en évidence la variabilité temporelle ou bien en sélectionnant des conditions de propagation bien particulières (*a priori* sur les paramètres). Les paramètres influents relatifs aux groupes présentent alors une très faible dispersion ;
- la calibration est menée à l'aide d'un modèle numérique de propagation acoustique parfois coûteux en temps de calcul. Afin de minimiser les coûts de calculs, l'utilisation d'un *métamodèle* s'est donc avéré indispensable. La construction d'un tel métamodèle requiert de suivre une méthodologie rigoureuse. Outre cette méthodologie, des décisions purement physiques viennent se greffer. Une des questions majeures a été la façon de prendre en compte différents niveaux de turbulence tout en conservant des coûts de calculs réalistes. Les choix effectués ont notamment été guidés par la recherche d'un compromis entre simplicité de mise en œuvre et robustesse finale du métamodèle. Cette construction du métamodèle est vraisemblablement perfectible mais ce dernier est tout à fait adéquat pour l'étude menée. Cet outil est d'autant plus intéressant qu'il pourrait également être utilisé dans le cadre d'autres études (sensibilité ou calibration) ;
- après avoir défini les *observables* d'étude (coefficients de l'ACPF), la *calibration du système* a été menée afin d'estimer en premier lieu les paramètres incertains θ . Cette estimation montre qu'une unique valeur de gradient vertical de célérité peut ne pas être suffisante et représentative

d'un site (différence marquée entre 50 et 150 m). Résistance spécifique au passage de l'air et paramètre de turbulence sont plus "cohérents" avec les valeurs expérimentales et ne présentent pas d'écarts conséquents entre les deux distances source-récepteur considérées. Ces estimations traduisent de plus la présence de biais importants entre spectres modélisés et spectres expérimentaux ;

- la *validation du système* a permis de confirmer la présence de biais, avec des distributions de résidus qui s'écartent sensiblement d'une distribution gaussienne théorique autour de la valeur centrale pour certaines observables. A noter que ces biais sont différents à 50 et 150 m de la source ;
- la phase de *prévision* finale rend compte des variances associées aux paramètres d'entrée du modèle (variances d'estimation) qui sont propagées dans le (méta)modèle. Cette part d'incertitude et de variabilité est traduite par un intervalle de confiance dont la moyenne est calculée en recentrant les résidus (correction des biais). La dispersion expérimentale non-expliquée par les paramètres d'entrée est rejetée dans le terme résiduel. Ecart-type du modèle (du fait de l'écart-type d'estimation de θ et de d_j) et écart-type résiduel constituent finalement l'incertitude globale. A un calcul de spectres ou de niveaux sonores peut ainsi être associés des intervalles de confiance : le modèle peut ainsi rendre compte de la variabilité physique et de l'incertitude. Cette estimation est cependant propre au site d'étude et à la géométrie considérés.

5

DISCUSSION SUR LES INCERTITUDES EXPÉRIMENTALES ET LES MODÈLES DE PROPAGATION

PLAN DU CHAPITRE

1	INTRODUCTION	131
2	LIMITATIONS EXPÉRIMENTALES	131
2.1	Caractérisation de la source	131
2.2	Caractérisation du milieu de propagation	133
2.3	Validation de la base de données	135
2.4	Exemples de résultats	135
3	LIMITES DE MODÉLISATION	138
4	CONCLUSION ET PERSPECTIVES	140

Dans les chapitres précédents, la méthodologie d'estimation d'incertitudes a été appliquée à des configurations relativement simples (site plan, source ponctuelle) et a permis de quantifier les incertitudes et la variabilité. Ce chapitre tente de dégager les limites rencontrées aussi bien d'un point de vue expérimental que numérique, lorsqu'on envisage d'appliquer une telle méthodologie à une configuration plus « réaliste » et donc plus complexe. Quelques perspectives applicatives sont aussi présentées.

1 INTRODUCTION

La base de données expérimentales à l'étude est la base de données issue de la Station de Long Terme (SLT) de Saint-Berthevin. Cette base de données contient six ans de mesures (période 2002-2007) acoustiques et météorologiques sur un site à topographie complexe (présence de plans d'eau, fort dénivelé). Cette base de données est davantage décrite en Annexe A. A noter de plus que chaque échantillon (acoustique et micrométéorologique) est représentatif d'une période de 15 minutes.

La mise en œuvre de la méthode de quantification d'incertitudes pour le cas d'un site à topographie complexe et d'une source routière réelle s'avère particulièrement difficile. En effet, caractérisation et modélisation de la source, gestion de la base de données, modélisation du site, etc. constituent les limites rencontrées au cours de notre étude. Elles sont liées aux aspects expérimentaux et numériques développés ci-après.

2 LIMITATIONS EXPÉRIMENTALES

2.1 CARACTÉRISATION DE LA SOURCE

Dans le cas de la SLT, on considère une source de bruit routier située sur un viaduc autoroutier surplombant le site expérimental et présentant un double tablier. Cette configuration présente des caractéristiques complexes, avec notamment des phénomènes de « fuite acoustique » et de diffraction entre les tabliers : l'ensemble formé de la source de bruit routier et du viaduc constitue la source à l'étude. La géométrie de cette source étant particulièrement complexe, la *modélisation* de cette source en est alors d'autant plus difficile (la modélisation « classique » par une source linéique n'est en effet pas suffisante pour caractériser l'ensemble des phénomènes observés).

La source que nous étudions est relative à un flux de trafic routier. Or, le trafic sur cette voie autoroutière n'est pas très soutenu et le *rapport signal sur bruit* est alors particulièrement faible et difficilement détectable à grande distance de la source. En outre, l'amplitude de la source de bruit routier n'est pas toujours suffisante pour obtenir un rapport signal sur bruit acceptable mesuré à chacun des microphones : la qualité de la mesure en est alors affectée.

Afin de s'affranchir des fluctuations des caractéristiques de la source routière (spectre et amplitude), un microphone de référence (placé sur le plateau à proximité de la voie autoroutière, figure A.2) permet d'exprimer les niveaux sonores en termes d'atténuation relative à ce microphone de référence. Cependant celui-ci ne permet pas d'obtenir des niveaux sonores de référence valables pour l'ensemble des points de mesures du site. Ce microphone de référence étant situé sur le plateau, sa *ligne de vue* est alors similaire à la ligne de vue d'un autre microphone situé sur le plateau (A5). Ce microphone de référence est également utilisé relativement à un second microphone situé en fond de vallée (A3) dont la ligne de vue est très différente de la ligne de vue sur le plateau (figure 5.1).



FIG. 5.1 – Vue générale de la Station de Long Terme de Saint-Berthevin.

En effet, ce microphone de référence ne caractérise la source que localement et ne prend donc pas en compte l'ensemble des contributions de la source du fait de son étendue spatiale (figure 5.2).

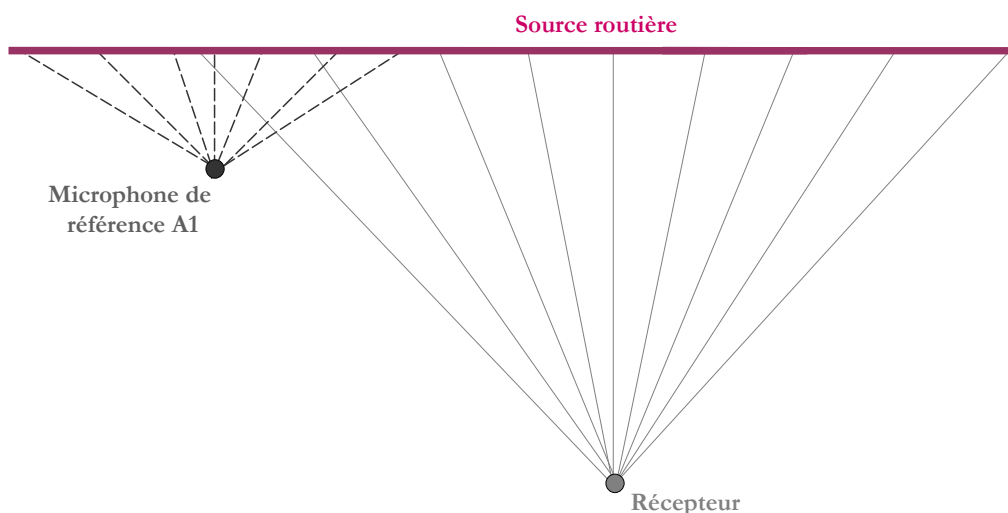


FIG. 5.2 – Représentation schématique de la source autoroutière de la SLT.

De plus, ce microphone étant très proche de la bordure autoroutière, les mesures enregistrées par celui-ci sont alors affectées par des effets de *diffraction* sur cette bordure. Ces contributions complémentaires non mesurées à plus grande distance de la source affectent alors les mesures effectuées par ce microphone de référence. Une solution pour s'affranchir du problème de diffraction et de ligne de vue serait de placer le microphone de référence un peu plus loin de la voie autoroutière et également d'ajouter un second microphone de référence permettant de rendre compte de la même ligne de vue que le microphone situé en fond de vallée.

Par ailleurs, la source de bruit routier étudiée est extrêmement *fluctuante* dans le temps. Un travail de sélection de périodes représentatives de flux de trafic est également nécessaire ; ce travail est mené par le Laboratoire Régional des Ponts et Chaussées de Blois selon les conditions inspirées des travaux normatifs [6], en utilisant les informations fournies par la station de comptage de trafic, ainsi que les

informations fournies par le microphone en bordure de voie (A1) lors du passage des véhicules. Cette sélection de conditions de trafic permet ainsi d'isoler les régimes de trafic qui nous intéressent (flux de trafic routier essentiellement et non le passage de véhicules isolés) et s'applique simultanément sur le reste des bases de données acoustique et météorologique, ce qui réduit considérablement la taille de la base de données initiale (les périodes nocturnes ne sont alors que très peu représentées).

En regard de cet ensemble d'éléments, la caractérisation de la source est donc complexe pour ce site d'étude et il devient alors difficile d'étudier les phénomènes propagatifs d'atténuation sonore relative à ce microphone de référence notamment en fond de vallée.

2.2 CARACTÉRISATION DU MILIEU DE PROPAGATION

La SLT présente des caractéristiques spatiales hétérogènes. L'altitude (présence d'un dénivelé d'environ 30 m), la présence d'obstacles (buttes de terres, arbres), la composition du sol (sol herbeux, présence de cailloux, plans d'eau, ruisseau) constituent un ensemble d'éléments qui rendent difficile la caractérisation expérimentale des paramètres influents sur la propagation acoustique. Parmi ces paramètres, on distingue ceux liés aux frontières du milieu (caractéristiques de sol) et ceux liés aux propriétés du milieu de propagation (paramètres micrométéorologiques).

2.2.1 Caractéristiques acoustiques de sol

Les caractéristiques acoustiques de sol de la SLT sont à l'heure actuelle mal connues. Ces propriétés sont variables dans le temps (chapitre 3) et dans l'espace [21]. Une valeur unique de résistance spécifique au passage de l'air mesurée en un point donné du site ne suffit pas à le caractériser correctement en termes d'impédance acoustique. Des mesures d'impédance de sol ont donc été effectuées mensuellement au cours de l'année 2006 en différents points du site. Les premières études statistiques menées à partir de ces données ne permettent pas de mettre en évidence clairement une évolution saisonnière des caractéristiques de sol. Afin de mettre en évidence la variabilité spatiale, des méthodes géostatistiques ont été utilisées [18, 22].

Ces méthodes d'interpolation des données utilisent la corrélation spatiale. Le degré de corrélation entre deux points de l'espace est accessible grâce à une fonction ajustée sur un variogramme expérimental (cette notion est abordée en annexe B). La méthode d'interpolation utilisée en géostatistique est le krigeage; il est basé sur la combinaison linéaire des données dont le critère de qualité est la minimisation de la *variance de l'erreur d'estimation* [33].

La variabilité observée entre plusieurs mesures autour d'un même point (sur une surface de l'ordre de 10 m²) est du même ordre de grandeur que celle observée à l'échelle du site, ce qui rend difficile toute analyse. Cependant, en moyenne pour chaque point de mesure, on peut montrer qu'il existe une variabilité spatiale non négligeable des caractéristiques acoustiques de sol (résistance spécifique au passage de l'air R_s) pour des zones du site *a priori* équivalentes *i.e.* dont le type de sol et le relief semblent identiques (figure 5.3) : de l'ordre de 100 kN.s.m⁻⁴ en fond de vallée (entre I2 et I5) et de l'ordre 300 kN.s.m⁻⁴ sur deux plateaux différents (entre I7 et I9).

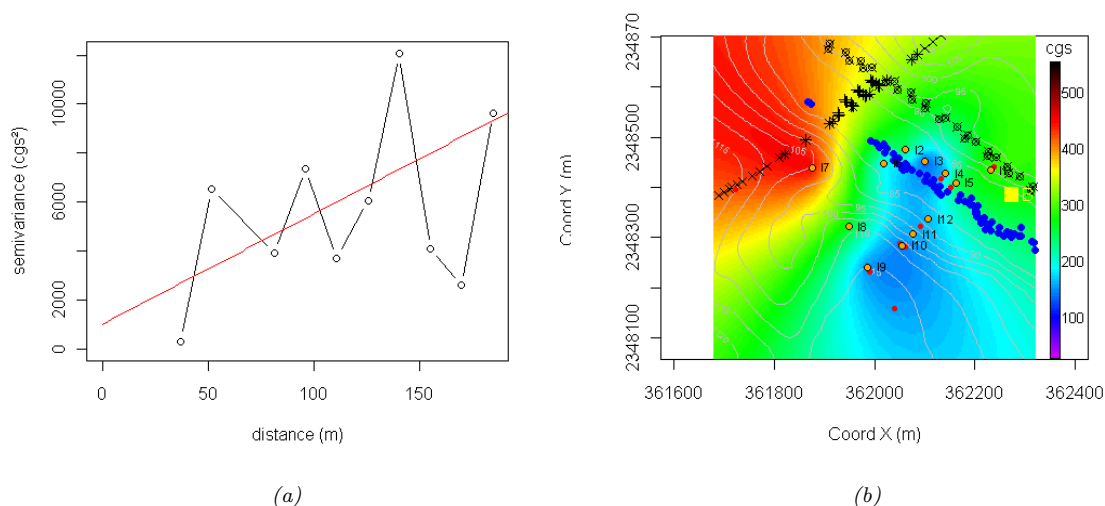


FIG. 5.3 – (a) Variogramme expérimental et modèle ajusté (courbe rouge) pour les mesures de résistance spécifique au passage de l'air R_s - (b) Carte de résistance spécifique au passage de l'air obtenue par krigeage - les points de mesure sont les points I_j , $j = 1, \dots, 12$ [22].

Plus récemment, un système de suivi continu des caractéristiques acoustiques de sol a été installé en un point du site (à proximité du mât A4) où des mesures sont effectuées toutes les quatre heures. Des mesures complémentaires (température et humidité de l'air et du sol, humectation, etc.) également effectuées en continu ont permis de dégager des lois d'évolution de l'impédance acoustique des sols avec des caractéristiques plus facilement mesurables [103].

Ces travaux sur la variabilité spatiale [22] et temporelle [103] des propriétés acoustiques du sol ont fait l'objet de stages co-encadrés en raison des interactions fortes avec ces travaux de thèse.

2.2.2 Paramètres micrométéorologiques

Les gradients verticaux de célérité du son ne sont accessibles qu'à partir des mesures par des mâts micrométéorologiques sur lesquels les capteurs de vitesse de vent et de température sont placés à 3, 10 et 25 m de hauteur. Cette configuration ne permet pas d'accéder rigoureusement à la forme logarithmique des profils de célérité car cette tendance logarithmique n'est marquée qu'à proximité du sol (en deçà de 3 m de hauteur). Seuls des profils linéaires sont donc accessibles, ce qui est une forte approximation au voisinage du sol et une source d'erreur non négligeable pour le calcul. Ces aspects météorologiques et méthodologiques alimentent encore les activités de recherche actuelles (chapitre 1 section 1.3.4.2). De plus, les paramètres météorologiques à disposition sur les six années de données ne fournissent pas d'information quant à la turbulence atmosphérique (chapitre 1 section 1.3.4.3). La prise en compte de celle-ci est pourtant un élément essentiel du point de vue de la modélisation ; le chapitre 3 a en effet montré qu'il était indispensable d'intégrer le paramètre de turbulence aux modèles afin d'obtenir des niveaux sonores réalistes (on retiendra notamment la dispersion très forte observée dans des conditions de propagation défavorables). Par contre ces données rendent bien compte de la variabilité spatiale des phénomènes micrométéorologiques, notamment l'effet « canyon » observé en comparant les mesures effectuées sur le plateau et en fond de vallée [57]. Celui-ci se traduit par la modification des conditions micrométéorologiques en fond de vallée du fait de la présence de dénivelés « forts » créant ainsi une circulation dite cellulaire qui ralentit les échanges avec les couches supérieures de l'atmosphère.

2.3 VALIDATION DE LA BASE DE DONNÉES

Les informations sur le trafic, issues de la station de comptage du trafic routier, font l'objet d'un post-traitement à part entière permettant de fournir pour chaque échantillon temporel et pour chaque voie le nombre de véhicules classés selon leur silhouette et leurs vitesses moyennes respectives.

Les données micrométéorologiques brutes nécessitent plusieurs traitements de validation basés sur les aspects métrologiques (dérives de capteurs, dysfonctionnements, pannes, etc.) mais également sur des aspects purement physiques : exclusion des régimes de vent extrêmes¹ *i.e.* en deçà de 1 m.s^{-1} et au delà de 5 m.s^{-1} .

En ce qui concerne les données acoustiques brutes, un travail important de validation a été mené. Ces outils de validation ont été développés en 2007 [81] et sont basés sur :

- la *comparaison inter-capteurs* : comparaison des deux microphones sur un même mât. Si des différences majeures entre les deux hauteurs de mesures surviennent les données sont alors invalidées pour les deux microphones ;
- la *détection du passage de trains* : lorsqu'une valeur seuil est atteinte au mât se trouvant en bordure de voie ferroviaire (A2), les mesures acoustiques sur l'ensemble des mâts sont alors invalidées simultanément ;
- la création d'un *gabarit de long terme* calculé à partir d'un an de mesures : ce gabarit permet d'isoler les événements sonores « marginaux » ;
- la vérification de la *continuité temporelle* des données ;
- etc.

La base de données finale est issue de la validation puis de la sélection croisée des données micrométéorologiques, des informations sur la source et des données acoustiques mesurées aux points récepteurs. Ces phases de validation et de sélection, indispensables pour tous travaux d'analyse, entraînent une réduction importante du nombre de données disponibles et par conséquent une perte d'informations sur la variabilité des conditions de propagation rencontrées sur ce site.

2.4 EXEMPLES DE RÉSULTATS

2.4.1 Mesures acoustiques

La comparaison des mesures acoustiques effectuées sur le plateau et en fond de vallée ne met pas en exergue de façon marquée une dépendance spatiale de la variabilité des données : les dispersions des spectres en A3 et en A5 sont en effet très semblables (figures 5.4 et 5.5).

1. A noter que deux régimes de vent principaux se distinguent : un régime de vent fort dont la vitesse moyenne mesurée à 6 m de hauteur s'étend de 3 à 5 m.s^{-1} et un régime de vent faible de 1 à 3 m.s^{-1} . Cette distinction est valable au sens de l'acoustique, en regard de l'influence du vent sur la propagation acoustique. Cette terminologie issue de travaux normatifs [6] ne correspond pas à celle couramment utilisée en météorologie.

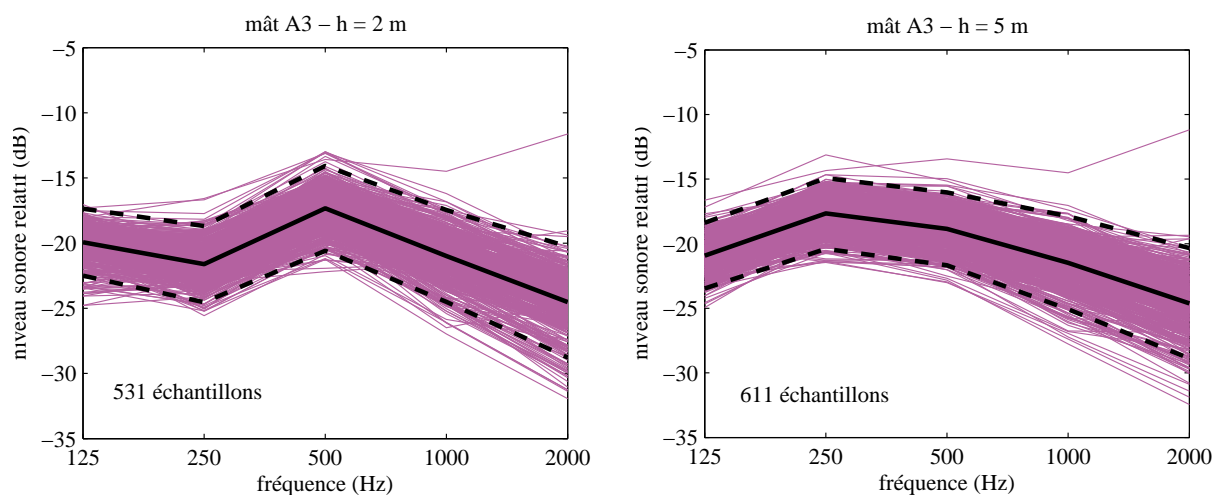


FIG. 5.4 – Résultats expérimentaux issus de la SLT (2002-2007) - niveaux sonores mesurés au mât A3 relativement au mât A1 à 2 et 5 m de hauteur (moyenne en trait plein et écarts-typé en trait discontinu).

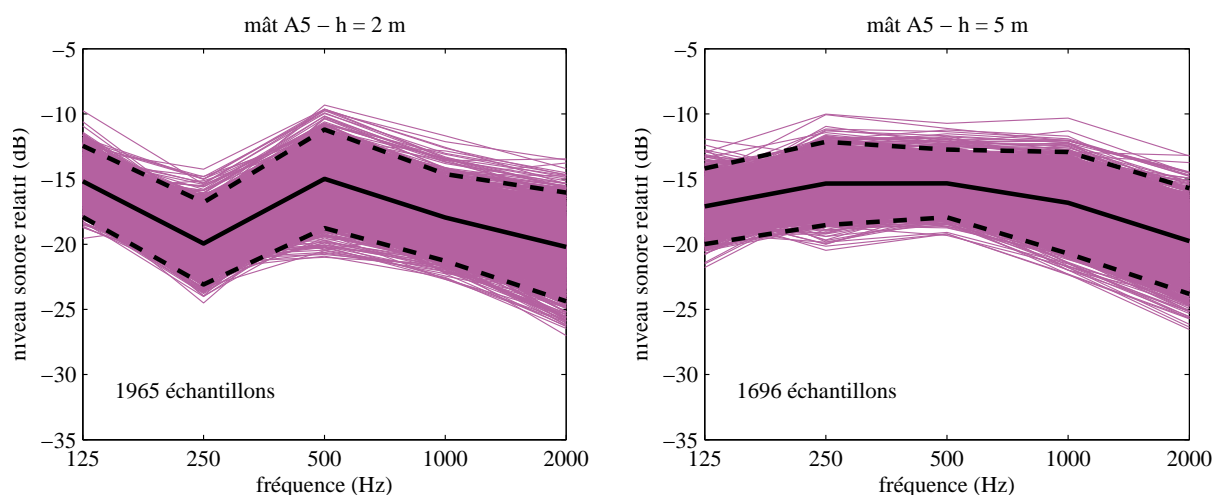


FIG. 5.5 – Résultats expérimentaux issus de la SLT (2002-2007) - niveaux sonores mesurés au mât A5 relativement au mât A1 à 2 et 5 m de hauteur (moyenne en trait plein et écarts-typé en trait discontinu).

On observe que l'atténuation des niveaux sonores relativement à A1 est plus élevée en A3 qu'en A5 avec une différence d'environ 5 dB pour l'ensemble des bandes d'octaves considérées. Cette observation peut être expliquée par les différences qui existent entre les deux microphones tant du point de vue de l'émission (ligne de vue de la source linéique, diffraction par le tablier) que de la propagation (effets de sol et effets micrométéorologiques). On peut notamment se référer aux figures 5.1 et 5.2 à titre d'illustration. De plus, que ce soit en A3 ou en A5, la forme des spectres est bien différente entre les mesures faites à 2 ou 5 m de hauteur, notamment à 250 Hz où on observe une figure interférentielle pour la géométrie considérée.

La variabilité temporelle est bien présente : malgré une validation rigoureuse des données (section 2.3), les dispersions totales des spectres au mât A3 et au mât A5 peuvent atteindre 5 à 10 dB sur l'ensemble de la gamme de fréquences toutes conditions de propagation confondues. Cette variabilité

temporelle associée aux niveaux sonores intègre la variabilité des phénomènes influents (effets de sol et effets micrométéorologiques) qui évoluent sur leurs propres échelles temporelles.

2.4.2 Mesures micrométéorologiques

Les gradients verticaux de célérité peuvent être reliés aux paramètres de réfraction thermique a_{T_p} et aérodynamique a_{V_p} (chapitre 1 section 1.3.4.2). La figure 5.6 montre que les conditions défavorables thermiques (avec des gradients thermiques forts tels que $a_{T_p} \ll 0$, représentatives de périodes diurnes ensoleillées) ne sont pas ou peu représentées dans la base de données. En effet, dans ces conditions de propagation, les niveaux sonores présentent un faible rapport signal sur bruit, ce qui mène à une élimination de ces échantillons. Le paramètre aérodynamique présente une distribution plus symétrique en raison de l'aspect vectoriel des effets de vent. En effet, la projection du vecteur vent sur la direction de propagation (chapitre 1 section 1.3.4.2) intègre la variabilité du vent sur la période considérée.

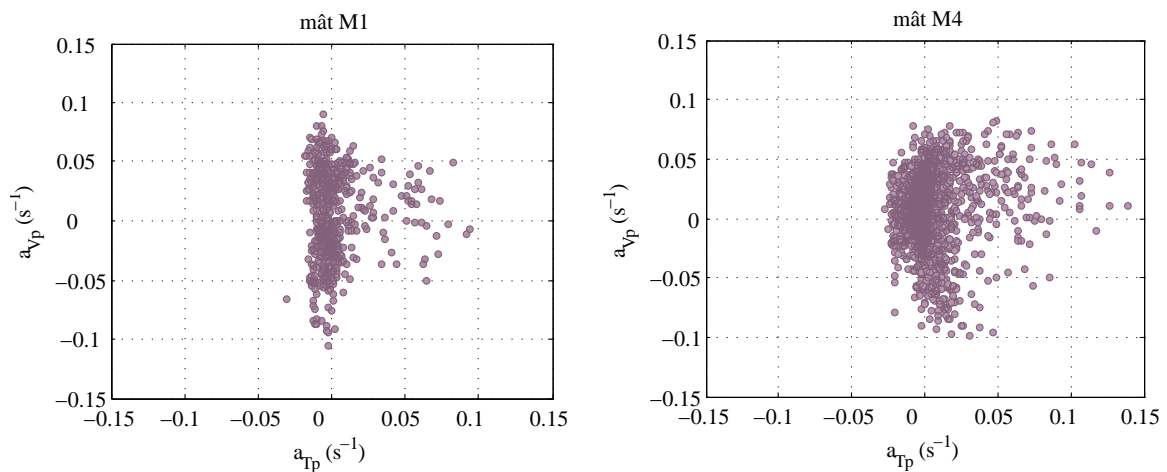


FIG. 5.6 – Résultats expérimentaux issus de la SLT (2002-2007) - nuages de corrélation entre le paramètre de réfraction thermique a_{T_p} et le paramètre de réfraction aérodynamique a_{V_p} aux mâts météorologiques M1 et M4.

Grâce à ces dernières propriétés, la dissymétrie observée sur a_{T_p} n'a alors pas d'incidence en termes de distribution des gradients verticaux de célérité du son (figure 5.7) où l'ensemble des grandes classes de conditions de propagation est représenté (défavorables, homogènes, favorables). A noter que les conditions de propagation proches des conditions homogènes sont les plus représentées alors que les conditions extrêmes sont peu présentes en raison des critères de validité définis.

La discrétisation arbitraire des histogrammes des gradients de célérité (figure 5.7) est plus fine que la précision expérimentale à laquelle nous avons accès en réalité (Annexe E). Cette discrétisation différente suivant que l'on considère le mât M1 ou le mât M4 (liée au nombre d'échantillons à disposition après la phase de validation), rend compte de caractéristiques gaussiennes et centrées de la distribution des gradients verticaux de la célérité du son. Les différences entre M1 et M4 (effet de site) sont davantage visibles sur la figure 5.6, en particulier pour les conditions $a_{T_p} \gg 0$.

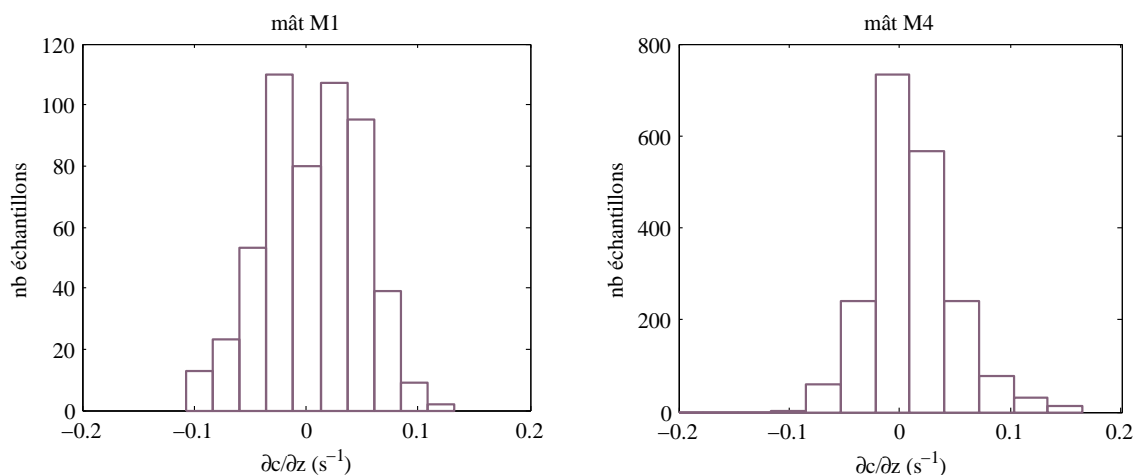


FIG. 5.7 – Résultats expérimentaux issus de la SLT (2002-2007) - histogrammes des gradients verticaux de célérité aux mâts météorologiques M1 et M4.

Les figures 5.8 présentent les niveaux sonores relatifs au microphone de référence en fonction des paramètres de réfraction thermique a_{T_p} , aérodynamique a_{v_p} et du gradient vertical de célérité $\frac{\partial c}{\partial z}$ pour les octaves centrés sur les fréquences 500 Hz et 2 kHz, mesurés à une hauteur de 5 m au microphone A5 (plateau). D'autres résultats sont fournis en section 1.3 de l'Annexe A pour d'autres bandes d'octaves mesurés aux mâts A3 (fond de vallée) et A5 à des hauteurs de 2 et 5 m. Ces figures présentent dans l'ensemble une forte dispersion. Elles montrent cependant des tendances « grossières », avec une augmentation du niveau sonore relatif pour des valeurs croissantes de gradients et de paramètres de réfraction, en particulier au microphone A5 et à 2 kHz. Dans ce cas, la légère tendance d'évolution croissante des niveaux sonores avec le gradient vertical de célérité est principalement expliquée par celle observée en fonction du paramètre thermique a_{T_p} . Cette observation n'est pas vérifiée en A3 (Annexe A section 1.3) en raison à la fois du moindre nombre d'échantillons et du caractère plus diffus de la source vue depuis le fond de la vallée (section 2.1).

Comme la théorie le laissait prévoir, ces tendances expérimentales obtenues dans le cas d'une source routière de type linéique sont moins marquées que dans le cas d'une source ponctuelle (chapitre 1 figure 1.12). On vérifie également que la dispersion des niveaux sonores tend à augmenter avec la fréquence, comme on le constate au passage de 500 Hz à 2 kHz, ainsi que pour les autres bandes d'octaves (Annexe A).

3 LIMITES DE MODÉLISATION

Une étude de calibration dans une configuration complexe telle que celle de la SLT nécessite l'utilisation de modèles de propagation évolués aux coûts de calculs prohibitifs. A ce jour, une telle configuration ne peut pas être parfaitement modélisée par les modèles de référence (ou de laboratoire) disponibles (chapitre 1 section 2.3).

Une alternative a donc été envisagée, en ayant recours à des modèles de propagation simplifiés du type ingénierie. Malheureusement les modèles actuellement à disposition tels que les modèles Harmonoise (méthode issue des travaux du projet européen Harmonoise [40] et pressentie par la commission européenne pour être la méthode harmonisée [9, 71, 72, 110]) et NMPB08 [112] ne permettent pas de modéliser de façon rigoureuse un site dont la configuration est aussi complexe, notamment pour les raisons suivantes :

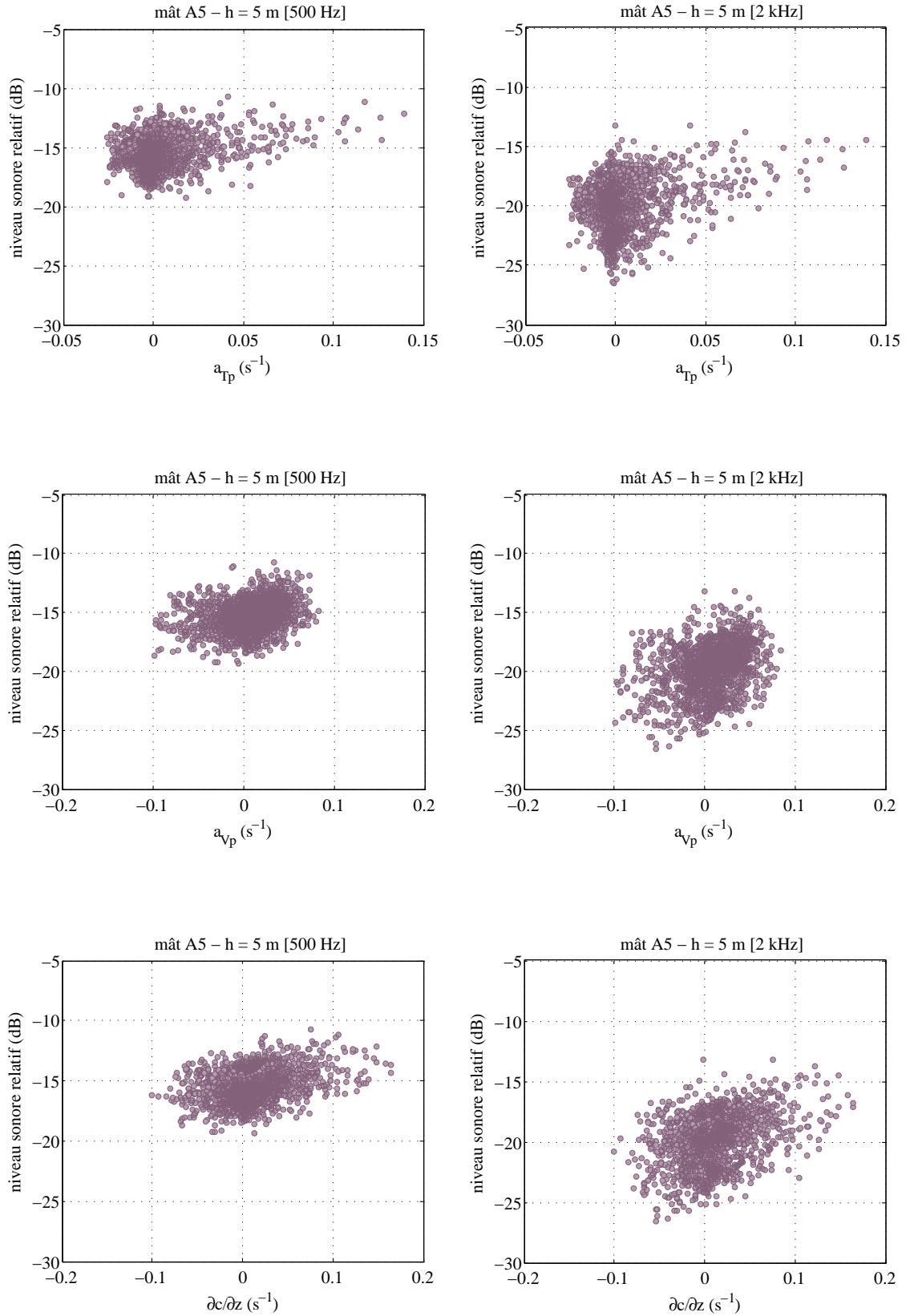


FIG. 5.8 – Résultats expérimentaux issus de la SLT (2002-2007) - niveaux sonores mesurés au mât A5 à 5 m de hauteur relativement au mât A1 à 500 Hz et 2 kHz.

- la prise en compte de la topographie est très approximative. La topographie entre source et récepteur est en effet simplifiée par un « plan moyen » qui lisse considérablement la topographie. Dans le cas d'un site à topographie complexe, une telle simplification ne rend alors plus compte d'éventuels effets majeurs sur la propagation (chapitre 1 section 1.3.5). La topographie devrait en outre être modélisée avec davantage de précision ;
- la prise en compte des caractéristiques acoustiques de sol passe, par exemple, par un coefficient variant entre 0 et 1. Le(s) modèle(s) n'étant pas sensible(s) à de faibles variations de ce paramètre, une étude de quantification ou d'estimation d'incertitudes basée notamment sur celui-ci n'est donc pas envisageable. De plus, notre connaissance de ce paramètre sur le site n'est à l'heure actuelle que partielle (section 2.2.1) ;
- les modèles d'ingénierie ne permettent pas de prendre en compte toute la complexité de la source de bruit routier sur la SLT du fait de la présence du tablier autoroutier (section 2.1) ;
- les paramètres d'entrée relatifs aux conditions micrométéorologiques ne rendent pas compte de toute l'étendue des conditions de propagation rencontrées en réalité (voire même parfois pas de conditions défavorables). De plus, ces modèles présentent une moindre sensibilité à ces paramètres micrométéorologiques d'entrée que les modèles de référence tels que les modèles basés sur la résolution de l'Équation Parabolique (chapitre 1 section 2.3).

4 CONCLUSION ET PERSPECTIVES

L'analyse des résultats expérimentaux de la Station de Long Terme rend compte de la problématique générale des incertitudes associées aux niveaux sonores et à leurs paramètres influents. Nous avons vu que ces incertitudes sont liées aux aspects expérimentaux et à la modélisation, pour lesquels nous dressons ci-dessous un bilan et des perspectives.

D'un point de vue expérimental, les sources d'incertitudes sont multiples et sont liées à l'acquisition des données (méthodologie et métrologie), aux caractéristiques intrinsèques du site (source routière, sol hétérogène, micrométéorologie, topographie complexe) ainsi qu'aux procédures de validation et de sélection des données. Les six années de données actuellement disponibles et validées rendent compte de grandes tendances qu'il serait nécessaire d'affiner grâce à des informations expérimentales plus précises. La classification actuelle des conditions de propagation est rendue complexe du fait de la dispersion importante des données acoustiques en regard des conditions de propagation. L'amélioration du protocole expérimental pourrait aider à définir une classification des conditions de propagation plus judicieuse en intégrant à la fois les caractéristiques de la source, l'aspect micrométéorologique (en prenant en compte les paramètres de réfraction thermiques et aérodynamiques ainsi que le coefficient de turbulence) et les caractéristiques acoustiques de sol.

- D'une manière plus générale, la base de données pourrait être améliorée selon les axes suivants :
- par l'ajout de microphones de référence équidistants de la source ou le long du chemin de propagation *e.g.* plusieurs microphones entre le mât de référence (A1) et le mât se situant sur le plateau (A5) ;
 - par l'acquisition par bande de tiers d'octaves plus que par bande d'octaves afin de raffiner la caractérisation des phénomènes ;
 - par l'étude plus approfondie du champ sonore au voisinage du viaduc (par antennerie par exemple) ;
 - en ayant davantage de connaissances sur l'évolution temporelle de l'impédance de sol avec la mise en place d'un système de suivi continu (mis en place en 2009) ;
 - en complétant les données micrométéorologiques par des informations sur la turbulence et en accédant à des profils verticaux de célérité du son *log-lin* plutôt que linéaires ou logarithmiques seulement (mise en place d'anémomètres ultrasoniques tridimensionnels en 2008).

D'un point de vue de la modélisation, ni les modèles de référence ni les modèles simplifiés ne permettent aujourd'hui de décrire rigoureusement et rapidement les phénomènes physiques rencontrés sur le site de la SLT. Ainsi, dans le domaine de la modélisation numérique, des développements sont encore nécessaires pour utiliser des modèles de référence (fréquentiels ou temporels, chapitre 1 section 2.3) et un certain nombre d'éléments doivent être améliorés en ce qui concerne les modèles simplifiés :

- la modélisation de la source linéique complexe : la configuration géométrique du viaduc doit être prise en compte avec soin (diffraction par le double tablier) ;
- la prise en compte de la topographie ne doit pas être réduite à un plan moyen entre la source et le récepteur ;
- les modèles devraient être plus sensibles aux variations des paramètres relatifs aux effets de sol et aux effets micrométéorologiques ;
- les modèles devraient de plus rendre compte d'une plus grande étendue des conditions de propagation et notamment des conditions de propagation défavorables qui sont parfois oubliées.

Dans des perspectives d'application de la méthode de quantification d'incertitude, il serait préférable d'utiliser des modèles de référence malgré leur coût numérique. Les temps de calculs pour l'ensemble du processus de calibration pourraient alors être réduits en ayant recours à des métamodèles (chapitre 2 section 1.3).

Ce cas d'étude rend compte des limites auxquelles nous sommes encore confrontées malgré l'existence de modèles relativement robustes et d'une base de données expérimentales très riche. En regard de ces remarques, il n'est pas possible à l'heure actuelle de mettre rigoureusement en place un processus de quantification d'incertitudes à partir de ces données expérimentales et des modèles que nous avons à notre disposition. L'amélioration des techniques expérimentales et des méthodes de modélisation pour des sources et sites complexes devraient permettre de mettre en place le processus de quantification d'incertitudes proposé.

CONCLUSION ET PERSPECTIVES

La problématique de la propagation acoustique en milieu extérieur est complexe et implique de nombreux phénomènes physiques liés aux caractéristiques acoustiques de sol ainsi qu'aux conditions micrométéorologiques (chapitre 1). Ces phénomènes fluctuent dans le temps et dans l'espace ce qui se traduit d'un point de vue acoustique par une forte *variabilité* des niveaux sonores. A cette variabilité physique vient s'ajouter une part d'*incertitude épistémique* liée à un manque de connaissance. Caractérisation de la source, interdépendance des phénomènes influents, méthodologie (configuration géométrique, métrologique, influence de l'opérateur et du matériel, etc.) sont autant d'éléments qui contribuent à cette incertitude. L'estimation d'un niveau sonore représentatif d'un site et d'une situation donnée à un instant donné est donc délicate en pratique. Ainsi ces travaux de thèse avaient pour objectifs d'apporter des connaissances et des outils pour quantifier incertitude et variabilité.

BILAN DU TRAVAIL DE THÈSE

MÉTHODOLOGIE STATISTIQUE DE QUANTIFICATION D'INCERTITUDE

L'élément central de ce travail de thèse est le développement d'une méthode statistique de quantification d'incertitude basée sur la « calibration » des paramètres influents du système (chapitre 2). Cette approche avait été appliquée dans des domaines scientifiques pour lesquels la notion de « risque » (de dépassement de seuil par exemple) est déterminante (risque de crue, risque nucléaire, etc.). Elle n'avait jamais été utilisée dans le domaine de la propagation acoustique en milieu ouvert et nous avons montré son intérêt pour estimer les paramètres d'entrée, l'incertitude qui leur est associée et l'incertitude sur les grandeurs d'intérêt du système. Cette méthode, fondée sur un processus de *calibration fréquentiste*, utilise des modèles de prévision acoustique ainsi que des groupes de données expérimentales caractéristiques de conditions de propagation particulières (période temporelle, type de conditions météorologiques, etc.), ce qui permet ainsi de renforcer le contenu prédictif des modèles. Ce processus statistique appelé *Calibration sous Incertitude*, consiste alors à *calibrer* les paramètres du modèle par minimisation d'une fonction coût qui contient notamment les *résidus* calculés entre les données expérimentales et les données modélisées. Nous avons fait évoluer cette méthode statistique au fil des travaux du fait des limitations et besoins rencontrés, ce qui a conduit à mettre en place un certain nombre d'outils.

Ce processus de quantification d'incertitude se décompose en cinq étapes principales. (1) La *spécification du système* qui consiste à sélectionner les données expérimentales et le modèle de prévision acoustique voire de construire un modèle de substitution appelé *métamodèle* dans le cas où le modèle choisi est trop chronophage. En outre, ce métamodèle présente l'intérêt de réduire les temps de calcul et de visualiser facilement les phénomènes dans l'espace des paramètres influents. La phase (2) de *définition des observables* a été développée suite à une indétermination du système statistique observée lors d'une étude préliminaire. Cette étape consiste en quelque sorte à sélectionner l'information utile contenue dans les données expérimentales. Ceci a conduit à une redéfinition des observables d'étude

selon deux méthodes : l'une basée sur des critères physiques dépendant fortement du modélisateur et une autre plus générale basée sur une analyse en composantes principales fonctionnelles (ACPF). Ces méthodes permettent de lever l'indétermination du problème et de réduire la dimension du problème. (3) A partir de ces éléments, le *système de calibration* peut être défini (en particulier la fonction de log-vraisemblance dont la forme dépend du système d'étude) puis *validé*. (4) Cette phase de validation consiste à vérifier les hypothèses initiales, calculer des critères statistiques et vérifier l'identifiabilité du système. L'identifiabilité est indispensable mais très difficilement vérifiable dans la pratique notamment dans le cas où des systèmes non linéaires à variables corrélées sont étudiés. Cet aspect nécessite donc des développements complémentaires qui dépassent le cadre de ces travaux. Cependant, on se rend aisément compte de la cohérence du système lors de la mise en œuvre du processus de calibration si la valeur des paramètres estimés est physiquement réaliste. Il peut également être intéressant d'observer la fonction coût dans l'espace des paramètres à estimer pour rendre compte de la continuité ou non de celle-ci. Dans le cas où le système n'est pas validé, celui-ci doit être redéfini ce qui implique de reprendre le processus depuis la première étape de spécification du système. Dans le cas où le système est validé, il est possible de passer à la phase *prévisionnelle* (5). Les paramètres influents sont estimés et les variances d'estimation associées sont calculées puis propagées à travers le modèle de prévision afin de traduire une part d'incertitude et de variabilité des données expérimentales. La part de variation inexpliquée par les paramètres influents et les variances d'estimation est reportée dans la partie *résiduelle*. La combinaison additive de l'écart-type lié à l'estimation et de l'écart-type résiduel constitue ainsi l'écart-type global caractérisant l'*incertitude globale* associée au système.

La méthodologie que nous proposons est adaptée à toute problématique de quantification d'incertitude : elle permet d'associer un écart-type aux paramètres d'entrée du système ainsi qu'aux prévisions du modèle dans une configuration donnée. Cette approche a ainsi été appliquée à différents cas d'études de niveaux de complexité croissants, dans le domaine de l'acoustique environnementale.

INCERTITUDE ASSOCIÉE AUX CARACTÉRISTIQUES ACOUSTIQUES DE SOL

Le premier cas d'application considéré est relatif à la prise en compte des propriétés des frontières du milieu de propagation, *i.e.* aux caractéristiques acoustiques de sol (chapitre 3). Celles-ci sont estimées par méthode inverse et sont donc très dépendantes des modèles d'impédance acoustique utilisés. Nous avons donc choisi de comparer deux modèles d'impédance de sol pour rendre compte de la sensibilité de chacun de ces modèles aux paramètres d'entrée. Nous avons également comparé différentes définitions du système statistique. Cette première application a ainsi permis de mettre en évidence l'avantage de la définition des observables par ACPF, du découpage de la base de données en plusieurs groupes représentatifs de situations spécifiques (dans notre cas d'étude période pluvieuse ou non), ainsi que l'avantage de travailler sur des groupes de spectres plutôt que sur un ensemble de spectres. Il est en effet statistiquement plus intéressant de raisonner sur un ensemble de spectres auxquels peut être attribué un jeu de paramètres moyens (complété par un écart-type) plutôt que sur des spectres indépendants.

Ainsi le découpage de la base de données en deux groupes distincts a permis d'isoler des conditions de propagation bien différentes en termes de propriétés acoustiques et de rendre compte de leur variabilité dans le temps. Des tendances journalières sont observées ainsi que des tendances liées à des paramètres atmosphériques tels que les précipitations. Les écarts-type d'estimation associés notamment à la résistance spécifique au passage de l'air traduisent clairement ces tendances. Cette approche montre qu'il est rarement concevable de considérer une valeur unique de résistance spécifique au passage de l'air pour caractériser un site. Pour notre base de données expérimentales acquises sur une durée de trois mois (période estivale), les caractéristiques de sol présentent une variabilité temporelle non négligeable que nous avons ainsi quantifiée à l'aide des outils développés. Une approche similaire pourrait être menée sur des échelles plus fines pour quantifier la variabilité des caractéristiques de sol

sur une journée.

Une telle approche par méthode inverse pose cependant le problème de la dépendance des résultats au modèle que l'on choisit : les tendances observées sur les paramètres estimés avec le modèle de Delany et Bazley ou le modèle de Zwikker et Kosten sont similaires, mais la corrélation entre les paramètres du modèle de Zwikker et Kosten rend parfois difficile l'estimation des paramètres. En effet, cette estimation de paramètres peut être problématique lorsqu'un nombre trop important de paramètres (parfois inaccessibles *in situ* par la mesure) doit être estimé. Dans ce type d'application, le recours à des modèles empiriques simples tel que le modèle de Delany et Bazley peut donc être préférable à l'utilisation de modèles microstructurels plus complexes.

Plusieurs éléments pourraient être envisagés pour améliorer les estimations des paramètres relatifs aux caractéristiques acoustiques de sol : la mise en œuvre d'une approche bayésienne intégrant des informations *a priori* sur les paramètres influents (et renforçant ainsi le processus), ou bien une « amélioration » des modèles d'impédance de sol. Il serait en effet intéressant de pouvoir intégrer des paramètres influents facilement mesurables sur site dans les modèles d'impédance acoustique de sol utilisés, dans le cas de sols naturels. L'intégration d'informations telles que le taux de précipitations, la température du sol ou l'hygrométrie du sol permettrait de raffiner les modèles d'impédance de sol.

INCERTITUDE ASSOCIÉE AUX CARACTÉRISTIQUES ACOUSTIQUES DE SOL ET AUX CONDITIONS MICROMÉTÉOROLOGIQUES

Ce second cas d'application de la méthode de quantification d'incertitude est à la fois plus complexe et plus complet (chapitre 4). On considère dans ce cas la propagation acoustique sur des distances plus importantes que précédemment (de l'ordre de 150 m). Ceci implique de prendre en compte les paramètres micrométéorologiques en plus des caractéristiques acoustiques de sol dans le processus de calibration. Cette nouvelle application constitue également une étude plus complète du processus de quantification d'incertitude, qui a notamment impliqué la construction d'un *métamodèle de krigeage* pour substituer le modèle de propagation acoustique. Ce métamodèle a été construit à partir de résultats d'un plan de calculs issus d'un modèle numérique de propagation de référence basé sur la résolution de l'équation parabolique. Cette étude a été menée sur des groupes de spectres, mesurés à 50 et 150 m de la source, représentatifs de conditions de propagation dites « homogènes ». Elle a permis de calibrer les paramètres incertains du modèle et de rendre compte de divergences certaines avec les paramètres expérimentaux et plus particulièrement avec la résistance spécifique au passage de l'air ce qui traduit l'intérêt de procéder à une telle calibration des paramètres. Cependant des biais non-négligeables sont observés à 50 m et 150 m entre spectres calibrés et spectres expérimentaux. C'est pourquoi une correction de biais (lié aux limites du modèle de prévision acoustique) peut être apportée en aval sur les spectres modélisés. Nous avons de plus montré que les paramètres influents connus et les écarts-type d'estimation des paramètres influents incertains permettaient d'expliquer une part conséquente de la variabilité et de l'incertitude associées aux spectres expérimentaux. L'incertitude inexpliquée est alors rejetée dans la partie résiduelle qui peut aussi être utilisée pour exprimer un intervalle de confiance dans une phase prévisionnelle.

Une telle application présente des apports pratiques – pour les études d'impact par exemple – car elle permet de quantifier l'incertitude et la variabilité pour un modèle donné et pour des données mesurées sur un site donné. Il est ainsi possible de caractériser un « site » pour un ensemble de situations acoustiques. Cette approche pourrait notamment être mise en place pour caractériser l'incertitude et la variabilité pour des périodes données (jour / soir / nuit par exemple) ou des situations micrométéorologiques particulières. Elle peut également être utilisée en ayant recours à des bases de données de référence caractérisant des configurations différentes (source, topographie du site, nature

du sol, etc.), notamment pour alimenter les travaux normatifs dans le domaine.

Cette méthode de quantification d'incertitude peut également servir dans le cadre de la modélisation numérique. En effet, l'estimation de l'incertitude peut permettre de corriger et améliorer des modèles présentant (par exemple) d'éventuels biais. En outre, la constitution de groupes de données de référence permet de renforcer le contenu prédictif du modèle de prévision acoustique considéré et de plus, d'associer un intervalle de confiance à un calcul effectué par modèle.

DISCUSSION SUR LES INCERTITUDES EXPÉRIMENTALES ET LIMITATIONS DES MODÈLES DE PRÉVISION ACOUSTIQUE

Nous avons également appliqué la méthode de quantification d'incertitudes à un cas bien plus complexe avec une source routière réelle et un site à topographie complexe (chapitre 5). Cependant des limites majeures ont été rencontrées aussi bien d'un point de vue des données expérimentales que d'un point de vue des capacités à modéliser cette situation complexe. En regard de ces observations, des propositions sont faites pour améliorer la qualité de la base de données expérimentales : ajout de microphones de référence, mesures continues des caractéristiques acoustiques de sol, etc.

Une amélioration en termes de modélisation acoustique est également requise. Les modèles de prévision actuels ne sont en effet pas capables de décrire rigoureusement et rapidement la physique de l'ensemble des phénomènes rencontrés en réalité : source linéique complexe, effets de diffusion et de diffraction tridimensionnels, topographie, etc. Si l'on souhaite à l'avenir mettre en place cette méthode de quantification d'incertitude pour des situations complexes, il serait donc souhaitable de poursuivre l'effort de développement de modèles numériques de référence et de construire ensuite un métamodèle associé afin de minimiser les temps de calcul. Cette application permettrait de rendre compte de l'incertitude associée à des niveaux sonores dans de tels cas et sur des échelles allant du court terme au long terme, ce qui n'a jamais été fait à l'heure actuelle.

PERSPECTIVES GÉNÉRALES

Outre l'efficacité avérée de la méthode de Calibration sous Incertitude basée sur des méthodes de calibration *fréquentiste*, il serait intéressant pour des études futures de mettre en œuvre des méthodes de *calibration bayésienne* reposant sur la connaissance *a priori* que possède le modélisateur sur les paramètres d'entrée du système à l'étude. Ce type d'étude requiert cependant la connaissance préalable des paramètres d'entrée et permettrait d'aborder le problème sous une approche nouvelle.

En ce qui concerne les sources d'incertitudes, la méthode de quantification d'incertitude et de variabilité nous a permis d'accéder à une information quant à l'*incertitude globale* sur les niveaux sonores. Une méthode de propagation d'incertitude a été proposée (cf. Annexe E) pour quantifier les incertitudes liées aux seuls paramètres micrométéorologiques (avec une comparaison des méthodologies de calculs de paramètres), mais cette approche ne combine malheureusement pas d'autres paramètres influents dans le système. Il serait donc intéressant de poursuivre les développements de la méthode de quantification d'incertitude dans l'objectif de dissocier les différentes sources d'incertitude (effets de sol, effets micrométéorologiques, approximation du modèle de prévision, incertitudes expérimentales, etc.) et de les quantifier indépendamment.

Par ailleurs, l'utilisation d'un métamodèle s'est avérée particulièrement utile comme modèle de substitution au modèle de propagation acoustique. Or, l'utilisation d'un tel outil pourrait être généralisée pour mener des études de sensibilité, ou bien pour modéliser des phénomènes non ou mal maîtrisés et parfois uniquement accessibles expérimentalement. Un modèle mathématique permettant de modéliser des effets météorologiques réalistes, mal maîtrisés ou non modélisés à l'heure actuelle

dans des outils de type ingénierie, pourrait ainsi être développé.

Enfin, une évolution de la méthode de calibration sous incertitude ambitieuse et plus complexe à mettre en œuvre serait une application de la méthode dans un contexte dynamique (assimilation de données) pour prévoir des niveaux sonores et une incertitude associée « en temps réel », ce qui est particulièrement intéressant dans le cas de systèmes de mesure permanents (Station de Long Terme du LCPC, réseaux de mesures des observatoires du bruit mis en place par les villes et aéroports).

BIBLIOGRAPHIE

- [1] Article 2 de l'arrêté du 5 mai 1995 relatif au bruit des infrastructures routières.
- [2] Article 3 de l'arrêté du 23 janvier 1997 relatif à la limitation des bruits émis dans l'environnement par les installations classées pour la protection de l'environnement.
- [3] Loi "bruit" 31 décembre 1992 + textes d'applications.
- [4] AFNOR : Norme française NF ENV 13005, guide pour l'expression de l'incertitude de mesure.
- [5] AFNOR : Norme internationale ISO 9614-2, détermination par intensimétrie des niveaux de puissance acoustique émis par les sources de bruit - partie 2 : Mesurage par balayage, 1996.
- [6] AFNOR : Acoustique - caractérisation et mesurage du bruit dû au trafic routier, 2002.
- [7] H. AKAIKE : Information theory and an extension of the maximum likelihood principle. *In 2nd International Symposium on Information Theory*, 1973.
- [8] J. ALBEROLA, I.H. FLINDELL et A.J. BULLMORE : Variability in road traffic noise levels. *Applied Acoustics*, 66:1180–1195, 2005.
- [9] F. ANFOSSO-LÉDÉE, M. PAVIOTTI, S. KEPHALOPOULOS et B. GERGELY : Common noise assessment methods in eu : Cnossos-eu (part ii : Application for strategic noise mapping and action planning). *In Internoise*, Lisbon (Portugal), 2010.
- [10] P. ARGOUL, N. POINT et G. DUTILLEUX : *Problèmes inverses en génie civil*. Laboratoire Central des Ponts et Chaussées, 2008.
- [11] K. ATTENBOROUGH : Ground parameter information for propagation modeling. *Journal of the Acoustical Society of America*, 92(1):418–427, 1981.
- [12] K. ATTENBOROUGH : Acoustical impedance models for outdoor ground surfaces. *Journal of Sound and Vibration*, 99(4):521–544, 1985.
- [13] K. ATTENBOROUGH et *al.* : Propagation of sound above a porous half-space. *Journal of the Acoustical Society of America*, 68(5):1493–1501, 1980.
- [14] P. BAR : Le bruit des transports terrestres - prévision, prévention, protection. *Revue générale des Routes et des Aéroports*, 534, 1977.
- [15] N. BARRIÈRE : *Etude théorique et expérimentale de la propagation du bruit de trafic en forêt*. Thèse de doctorat, Ecole Centrale de Lyon, 1999.
- [16] H. E. BASS, L. C. SUTHERLAND et A. J. ZUCKERWAR : Atmospheric absorption of sound : Update. *Journal of the Acoustical Society of America*, 88(4):2019–2021, 1990.
- [17] D.M. BATES et D.G. WATTS : *Application numérique du concept de particules sonores à la modélisation des champs sonores en acoustique architecturale*, volume 258-259. Thématique Méthodes numériques en Génie Civil, 2005.
- [18] O. BAUME : *Approche géostatistique de l'influence des paramètres physiques sur la propagation acoustique à grande distance*. Thèse de doctorat, Université du Maine, 2006.
- [19] O. BAUME, B. GAUVREAU, M. BÉRENGIER, F. JUNKER et F. LAUZIN : Long term monitoring site at saint-berthevin (france-53) : a tool for traffic noise characterization using space and time statistical variability of acoustical and micrometeorological events. *In CFA / DAGA 04*, Strasbourg (F), 2004.
- [20] O. BAUME, B. GAUVREAU, M. BÉRENGIER, F. JUNKER, H. WACKERNAGEL et J.P. CHILÈS : Geostatistical modeling of sound propagation : principles and a field application experiment. *Journal of the Acoustical Society of America*, 126(6):2894–2904, 2009.

- [21] O. BAUME, H. WACKERNAGEL, B. GAUVREAU, F. JUNKER, M. BÉRENGIER et J.-P. CHILÈS : Space and time exploration of acoustical series under the influence of various micrometeorological stability conditions. *In GEOstatistics for ENVironmental applications*, pages 45–57, Rhodes (G), 2006.
- [22] M. BELLANGER : Application d'outils géostatistiques et de méthodes de classification aux données de la station de long terme. Rapport de stage de Master ingénierie mathématique spécialité statistiques, LCPC - Université de Nantes, 2009.
- [23] P. BLANC-BENON, B. LIPKENS, L. DALLOIS, M.F. HAMILTON et D.T. BLACK-STOCK : Propagation of finite amplitude sound through turbulence : Modelling with geometrical acoustics and the parabolic approximation. *Journal of the Acoustical Society of America*, 111(1):487–498, 2002.
- [24] B. BOGEY et C. BAILLY : A family of low dispersive and low dissipative explicit scheme for flow and noise computations. *Journal of Computational Physics*, 194(1):194–214, 2004.
- [25] G.E. BOX et N.R. DRAPER : *Empirical model building and response surfaces*. Wiley Series in Probability and Mathematical Statistics, 2005.
- [26] M. BÉRENGIER et G.A. DAIGLE : Diffraction of sound above a curved surface having an impedance discontinuity. *Journal of the Acoustical Society of America*, 84(3):1055–1060, 1988.
- [27] M. BÉRENGIER, B. GAUVREAU et D. JUVÉ : Outdoor sound propagation : A short review on analytical and numerical approaches. *Acta Acustica united with Acustica*, 89(6):980–991, 2003.
- [28] M. BRUNEAU : *Manuel d'acoustique fondamentale*. Hermès, 1998.
- [29] A. BUJA, T. HASTIE et R. TIBSHIRANI : Linear smoothers and additive models. *The Annals of Statistics*, 17(2):453–555, 1989.
- [30] J. CESBRON : *Influence de la texture de chaussée sur le bruit de contact pneumatique / chaussée*. Thèse de doctorat, Ecole Centrale de Nantes, 2007.
- [31] P. CHEVRET : *Simulation numérique des effets de la turbulence sur la propagation du son dans l'atmosphère*. Thèse de doctorat, Ecole Centrale de Lyon, 1994.
- [32] C.F. CHIEN et W.W. SOROKA : Sound propagation along an impedance plane. *Journal of Sound and Vibration*, 43:9–20, 1975.
- [33] J.P. CHILÈS et P. DELFINER : *Geostatistics - modelling spatial uncertainty*. Wiley Series in Probability and Statistics, 1999.
- [34] A.J. CRAMOND. et C.G. DON : Effects of moisture content on soil impedance. *Journal of the Acoustical Society of America*, 82(1):293–301, 1987.
- [35] N. CRESSIE : *Statistics for Spatial Data*. Wiley Series in Probability and Statistics, 1993.
- [36] A.C. CULLEN et H.C. FREY : *Probabilistic techniques in exposure assessment. A handbook for dealing with variability and uncertainty in models and inputs*. Plenum Publishing Corporation, New York, 1999.
- [37] G.A. DAIGLE : Effects of atmospheric turbulence on the interference of sound waves above a finite impedance boundary. *Journal of the Acoustical Society of America*, 65(1):45–49, 1979.
- [38] G.A. DAIGLE, J.E. PIERCY et T.F.W. EMBLETON : Effects of atmospheric turbulence on the interference on the interference of sound waves near a hard boundary. *Journal of the Acoustical Society of America*, 64(2):622–630, 1978.
- [39] B.A. de JONG, A. MOERKERKEN et J.D. van der TOORN : Propagation of sound over grassland and over an earth barrier. *Journal of Sound and Vibration*, 86(1):23–46, 1983.
- [40] J. DEFRAANCE, E. SALOMONS, I. NOORDHOEK, D. HEIMANN, B. PLOVSING, G. WATTS, H. JONASSON, X. ZHANG, E. PREMAM, I. SCHMICH, F.-E. ABALLÉA, M. BAULAC et F. de ROO : Outdoor sound propagation reference model developed in the european harmonoise project. *Acta Acustica united with Acustica*, 93(2):213–227, 2007.

- [41] M.E. DELANY et E.N. BAZLEY : Acoustical properties of fibrous absorbent materials. *Applied Acoustics*, 3:105–116, 1970.
- [42] L. DENIS-VIDAL, G. JOLY-BLANCHARD et C. NOIRET : System identifiability (symbolic computation) and parameter estimation (numerical computation). *Numerical Algorithms*, 34:283–292, 2003.
- [43] D. ECOTIÈRE : Road noise : characterization and estimation of uncertainty due to meteorological effects. In *ASA/EEA/SFA joint congress, Acoustics'08*, Paris (F), 2008.
- [44] B. EFRON : The jackknife, the bootstrap and other resampling plans. *CBMS-NSF Regional Conference Series in Applied Mathematics*, 38, 1982.
- [45] T.F.W. EMBLETON : Tutorial on sound propagation outdoors. *Journal of the Acoustical Society of America*, 100(1):31–48, 1996.
- [46] L.B. EVANS, H.E. BASS et L.C. SUTHERLAND : Atmospheric absorption of sound : Theoretical predictions. *Journal of the Acoustical Society of America*, 5(2):1565–1575, 1972.
- [47] G. EVENSEN : *Data assimilation, The Ensemble Kalman Filter*. 2009.
- [48] T. FOKEN : *Micrometeorology*. 2006.
- [49] J. FRANCO : *Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*. Thèse de doctorat, Ecole Nationale Supérieure des Mines de Saint-Etienne, 2008.
- [50] J. FRANCO, X. BAY, B. CORRE et D. DUPUY : Strauss processes : A new space-filling design for computer experiments. In *Congrès conjoint de la Société Statistique du Canada et de la SFdS*, Ottawa (Canada), 2008.
- [51] H.C. FREY et D.E. BURMASTER : Methods for characterizing variability and uncertainty : comparison of bootstrap simulation and likelihood-based approaches. *Risk Analysis*, 19:109–130, 1999.
- [52] H.C. FREY et S.R. PATIL : Identification and review of sensitivity analysis methods. *Risk Analysis*, 22:553–578, 2002.
- [53] J. FRIEDMAN, T. HASTIE et R. TIBSHIRANI : Multivariate adaptative regression splines. *Annals of Statistics*, 10:337–407, 1991.
- [54] M. GALLAGHER et J. DOHERTY : Parameter estimation and uncertainty analysis for a watershed model. *Environmental Modelling & Software*, 22:1000–1020, 2007.
- [55] B. GAUVREAU : *Influence des conditions micrométéorologiques sur l'efficacité des écrans acoustiques*. Thèse de doctorat, Université du Maine, 1999.
- [56] B. GAUVREAU : Methodological and metrological benchmarks for determining experimental values of meteorological parameters. In *Invited paper to ASA/EEA/SFA joint congress, Acoustics'08*, Paris (F), 2008.
- [57] B. GAUVREAU, M. BELLANGER, B. BONHOMME et H. POIRIER : Exploitation de la base de données expérimentale de la “station de long terme” pour la classification des conditions de propagation acoustique. In *10ème Congrès Français d'Acoustique*, Lyon (F), 2010.
- [58] B. GAUVREAU, M. BÉRENGIER, P. BLANC-BENON et C. DEPOLLIÉ : Traffic noise prediction with the parabolic equation method : validation of split-step padé approach in complex environments. *Journal of the Acoustical Society of America*, 112(6):2680–2687, 2002.
- [59] B. GAUVREAU, D. ECOTIÈRE, H. LEFÈVRE et B. BONHOMME : *Propagation acoustique en milieu extérieur complexe. Caractérisation expérimentale in-situ des conditions micrométéorologiques. Eléments méthodologiques et métrologiques*. 2009.
- [60] K.E. GILBERT, M.J. WHITE et L.C. SUTHERLAND : Application of the parabolic equation to sound propagation in refracting atmosphere. *Journal of the Acoustical Society of America*, 85(2):630–637, 1989.

- [61] P.F. GONOD : Prospective anthropologique - concepts, 2007.
- [62] G. GUILLAUME, J. PICAUT, G. DUTILLEUX et G. GAUVREAU : Use of the transmission line matrix method for the sound propagation modelling in open-spaces. *In 13th Long Range Sound Propagation Symposium*, Lyon (F), 2008.
- [63] D. HEIMANN : Numerical simulation of wind and sound propagation through an idealised stand of trees. *Acta Acustica*, 89:779–788, 2003.
- [64] D. HEIMANN, M. BAKERMANS, J. DEFRANCE et D. KUNHER : Vertical sound speed profiles determined from meteorological measurements near the ground. *Acta Acustica*, 93:228–240, 2007.
- [65] B. HOCHWALD et A. NEHORAI : On identifiability and information regularity in parametrized normal distributions. *Circuits Systems and Signal Processing*, 16(1):83–89, 1997.
- [66] B. IOOSS : *Contributions au traitement des incertitudes en modélisation numérique : propagation d’ondes en milieu aléatoire et analyse statistique d’expériences simulées*. Habilitation à diriger des recherches, Université Toulouse III Paul Sabatier, 2009.
- [67] M.A. JOHNSON, R. RASPET et M.T. BOBAK : A turbulence model for sound propagation from an elevated source above level ground. *Journal of the Acoustical Society of America*, 81(3):638–646, 1987.
- [68] F. JUNKER et B. GAUVREAU : Classification de l’influence relative des paramètres physiques affectant les conditions de propagation à grande distance. *Rapport avancement projet MEDD*, 2005.
- [69] F. JUNKER, B. GAUVREAU, C. CREMEZI-CHARLET, Ph. BLANC-BENON, B. COTTÉ et D. ECOTIÈRE : Classification of relative influence of physical parameters for long range acoustic propagation. *In Internoise*, Honolulu (EUA), 2006.
- [70] D. JUVÉ, Ph. BLANC-BENON et P. CHEVRET : Sound propagation through a turbulent atmosphere : Influence of the turbulence model. *In 6th Long Range Sound Propagation Symposium*, Ottawa (Canada), June 1994.
- [71] S. KEPHALOPOULOS, B. GERGEL, F. ANFOSSO-LÉDÉE, M. PAVIOTTI et P. de VOS : Common noise assessment methods in eu : Cnossos-eu (part iv : Aircraft noise prediction for the purpose of strategic noise mapping). *In Internoise*, Lisbon (Portugal), 2010.
- [72] S. KEPHALOPOULOS, B. GERGELY, F. ANFOSSO-LÉDÉE et M. PAVIOTTI : Common noise assessment methods in eu : Cnossos-eu (part I : The roadmap). *In Internoise*, Lisbon (Portugal), 2010.
- [73] J.P.C. KLEIJNEN : Kriging metamodeling in simulation : A review. Discussion Paper 2007-13, Tilburg University, Center for Economic Research, 2007.
- [74] J.P.C. KLEIJNEN : Screening experiments for simulation : A review. Discussion Paper 2007-21, Tilburg University, Center for Economic Research, 2007.
- [75] J.P.C. KLEIJNEN, S. M. SANCHEZ, T. W. LUCAS et T.M. CIOPPA : A user’s guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, 17(3):263–289, 2005.
- [76] V.O. KNUDSEN : The effect of humidity upon the absorption of sound in a room and a determination of the coefficients of absorption of sound in air. *Journal of the Acoustical Society of America*, 3(1A):126–138, 1931.
- [77] C. LARSSON : Seasonal and diurnal variations of atmospheric absorption for horizontal sound propagation. pages 17–26, Lyon (F), 1996.
- [78] T. LE POLLÈS, M. PICAUT, J. Bérengier et C. BARDOS : Sound field modeling in a street canyon with partially diffusely reflecting boundaries by the transport theory. *Journal of the Acoustical Society of America*, 116(5):2969–2983, 2004.

- [79] E. LEBARBIER et T. MARY-HUARD : Le critère BIC : fondements théoriques et interprétation. Rapport technique, INRIA, 2004.
- [80] L. LECLERCQ : *Modélisation dynamique du trafic et applications à l'estimation du bruit routier*. Thèse de doctorat, INSA Lyon, 2002.
- [81] O. LEROY : Développement théorique et pratique d'un logiciel de validation automatique des données acoustiques de la station de long terme. Rapport de stage de Master acoustique, LCPC - Université Pierre et Marie Curie, Paris VI, 2007.
- [82] K.M. LI et Q. WANG : A bem approach to assess acoustic performance of noise barriers in a refracting atmosphere. *Journal of Sound and Vibration*, 211(4):663–681, 1998.
- [83] B. LIHOREAU, B. GAUVREAU, P. BLANC-BENON et M. BÉRENGIER : Outdoor sound propagation modeling in realistic environments : a coupling method using parabolic equation and atmospheric model. *Journal of the Acoustical Society of America*, 120(1):110–119, 2006.
- [84] R. LINDER : *Les plans d'expériences - Un outil indispensable à l'expérimentateur*. Presses de l'ENPC, 2005.
- [85] S.N. LOPHAVEN, H.B. NIELSEN et J. SONDERGAARD : Aspects of the matlab toolbox dace. *DTU, Report IMM-REP-2002-13, Informatics and Mathematical Modelling*, 2002.
- [86] E.N. LORENZ : *The Essence of Chaos*. Jessie and John Danz Lecture Series, 1993.
- [87] M.D. MC KAY, W.J. CONOVER et R.J. BECKMAN : A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245, 1979.
- [88] A. MONIN et A. OBUKHOV : Basic laws of turbulent mixing in the atmosphere near the ground. *Tr. Akad. Nauk. SSR Geofiz. Inst.*, 24:163–187, 1954.
- [89] R.H. MYERS et D.C. MONTGOMERY : *Response surface methodology*. 1995.
- [90] L. NIJS et C.P.A. WAPENAAR : The influence of wind and temperature gradients on sound propagation, calculated with the two way wave equation. *Journal of the Acoustical Society of America*, 87(5):1987–1998, 1990.
- [91] J.E. OAKLEY et A. O'HAGAN : Probabilistic sensitivity analysis of complex models : A bayesian approach. *Journal of the Royal Statistical Society*, 66:751–769, 2004.
- [92] T.R. OKE : *Boundary Layer Climates*. Routledge Publishers, 1992.
- [93] A.B. OWEN : Orthogonal arrays for computer experiments, integration and visualiazation. *Statistica Sinica*, 2:439–452, 1992.
- [94] F. PASQUILL : The estimation of the dispersion of windborne material. *Meteorology Magazine*, 90:33–49, 1961.
- [95] Chris L. PETTIT et D. Keith WILSON : Proper orthogonal decomposition and cluster weighted modeling for sensitivity analysis of sound propagation in the atmospheric surface layer. *Journal of the Acoustical Society of America*, 122(3):1374–1390, 2007.
- [96] J.E. PIERCY, T.F.W. EMBLETON et L.C. SUTHERLAND : Review of noise propagation in the atmosphere. *Journal of the Acoustical Society of America*, 61(6):1403–1418, 1977.
- [97] A.E. RAFTERY : A bayesian model selection in social research. Rapport technique 94-12, University of Washington - Demography Center Working, 1995.
- [98] J. O. RAMSAY, G. HOOKER et S. GRAVES : *Functional Data Analysis with R and Matlab*. Use R. Springer, 2009. ISBN : 978-0-387-98184-0.
- [99] J. O. RAMSAY et B.W. SILVERMAN : *Functional data analysis*. Springer Series in Statistics, 1997.
- [100] K.B. RASMUSSEN : A note on the calculation of sound propagation over jumps and screens. *Journal of Sound and Vibration*, 84(4):598–602, 1982.

- [101] J.S. ROBERTSON, P.J. SCHLATTER et W.L. SIEGMANN : Sound propagation over impedance discontinuities with the parabolic approximation. *Journal of the Acoustical Society of America*, 99(2):761–767, 1996.
- [102] T.J. ROTHENBERG : Identification in parametric models. *Econometrica*, 39(3):577–591, 1971.
- [103] R. ROUFFAUD : Traitement et analyse d'une base de données expérimentale des caractéristiques d'impédance de sol (station de long terme). Rapport de stage de Master acoustique, LCPC - Université du Maine, 2010.
- [104] I. RUDNICK : The propagation of an acoustic wave along a boundary. *Journal of the Acoustical Society of America*, 19(3):348–356, 1947.
- [105] J. SACKS, W.J. WELCH, T.J. MITCHELL et H.P. WYNN : Design and analysis of computer experiments. *Statistical Science*, 4:409–423, 1989.
- [106] E.M. SALOMONS : *Computational atmospheric acoustics*. Kluwer Academic Publisher, 2001.
- [107] G. SAPORTA : *Probabilités, analyse des données et statistique*. Editions Technip, 2006.
- [108] E. SÉCHET : *Modélisation d'une connaissance imprécise sur les influences des conditions météorologiques dans la propagation du son à partir de données expérimentales*. Thèse de doctorat, Université de Paris IX Dauphine, 1996.
- [109] G. SCHWARZ : Estimating the dimension of a model. *Annals of statistics*, 6:461–464, 1978.
- [110] S. SHILTON, N. JONES, T. WERST, D. van MAERCKE, P. de VOS, F. ANFOSSO, M. PAVIOTTI et S. KEPHALOPOULOS : Common noise assessment methods in eu : Cnossos-eu (part iii : Guidance for the competent use for strategic noise mapping purposes). In *Internoise*, Lisbon (Portugal), 2010.
- [111] A. SOMMERFELD : Uber die ausbreitung der wellen in der drahtlosen telegraphie. *Annalen der Physik*, 333(4):665–736, 1909.
- [112] SÉTRA : *Prévision du bruit routier 1- Calcul des émissions sonores dues au trafic routier*. juin 2009.
- [113] R.B. STULL : *An introduction to boundary layer meteorology*. Kluwer Academic, Dordrecht, 1988.
- [114] L.C. SUTHERLAND et H.E. BASS : Atmospheric absorption in the atmosphere at high altitude. *7th Long Range Sound Propagation*, pages 1–15, 1996.
- [115] M. SVENSSONA, P.E. JANSSONA, D. GUSTAFSSONA, D.B. KLEJAA, O. LANGVALL et A. LINDROTH : Bayesian calibration of a model describing carbon, water and heat fluxes for a swedish boreal forest stand. *Ecological Modelling*, 213:331–344, 2008.
- [116] S. TAHERZADEH : Sources of uncertainty in ground impedance measurements and estimations. In *Internoise*, Ottawa (Canada), 2009.
- [117] TNS SOFRES : *Les français et les nuisances sonores*, 2010.
- [118] M. TINGLEY et C. FIELD : Small-sample confidence intervals. *Journal of the American Statistical Association*, 85:427–434, 1990.
- [119] I. TOLSTOY : Coherent sound scatter from a rough interface between arbitrary fluids with particular reference to roughness element shapes and corrugated surfaces. *Journal of the Acoustical Society of America*, 72(3):960–972, 1982.
- [120] Z. TOTH et E. KALNAY : Ensemble forecasting at nmc : the generation of perturbations. *Bulletin of the American Meteorological Society*, 74(12):2317–2330, 1993.
- [121] D.B. TURNER : *Workbook of atmospheric dispersion estimates*. Rapport technique, 1994.
- [122] V. TWERSKY : Reflection and scattering of sound by correlated rough surfaces. *Journal of the Acoustical Society of America*, 73:68–84, 1983.
- [123] A.W. van der VAART : *Asymptotic Statistics*. Cambridge University Press, 1998.

- [124] M. VAN OIJEN, J. ROUGIER et R. SMITH : Bayesian calibration of process-based forest models : bridging the gap between models and data. *Tree Physiol.*, 25:915–927, 2005.
- [125] D. VOSE : *Risk Analysis, a quantitative guide*. John Wiley & Sons, 2000.
- [126] E. WALTER et L. PRONZATO : *Identification of Parametric Models from Experimental Data*. Springer-Verlag, Heidelberg, 1997.
- [127] S. WEISBERG : *Applied Linear Regression*. John Wiley & Sons, 1985.
- [128] K. WERT, P. BLANC-BENON et D. JUVÉ : Effect of turbulence scale resolution on numerical simulation of atmospheric sound propagation. In *4th AIAA/CEAS Aeroacoustics Conference*, pages 246–256, Toulouse (F), 1998.
- [129] F.M. WIENER et D.N. KEAST : Experimental study of the propagation of sound over ground. *Journal of the Acoustical Society of America*, 31(6):724–733, 1959.
- [130] D.K. WILSON : Simple relaxation models for the acoustical properties of porous media. *Applied Acoustics*, 50:171–188, 1997.
- [131] D.K. WILSON : Uncertainty in outdoor noise measurement and prediction. *Journal of the Acoustical Society of America*, 118(3):1890, 2005.
- [132] D.K. WILSON, J.G. BRASSEUR et K.E. GILBERT : Acoustic scattering and the spectrum of atmospheric turbulence. *Journal of the Acoustical Society of America*, 105(1):30–34, 1999.
- [133] D.K. WILSON et D.W. THOMSON : Acoustic propagation through anisotropic, surface layer turbulence. *Journal of the Acoustical Society of America*, 96(2):1080–1095, 1994.
- [134] D.K. WILSON, J.C. WYNGAARD et D.I. HAVELOCK : The effect of turbulence intermittency on scattering into an acoustic shadow zone. *Journal of the Acoustical Society of America*, 99(6), 1996.
- [135] V. ZOUBOFF, Y. BRUNET, M. BÉRENGIER et E. SÉCHET : A qualitative approach of atmospheric effects on long range sound propagation. In *6th Long Range Sound Propagation Symposium*, Ottawa (Canada), June 1994.
- [136] V. ZOUBOFF, B. GAUVREAU, M. BÉRENGIER et Y. BRUNET : *Méthode de reconstitution d'un niveau sonore de long terme - Principe et applications*. Collection Etudes et Recherches des Laboratoires des Ponts et Chaussées, 2005.
- [137] V. ZOUBOFF, J.C. LAPORTE et Y. BRUNET : *Effets des conditions météorologiques sur la propagation du bruit - Prise en compte pratique*, volume Méthode d'essai n° 51. 1998.

A

PRÉSENTATION DES BASES DE DONNÉES EXPÉRIMENTALES À DISPOSITION

Les données expérimentales à disposition sont issues de différentes campagnes de mesures qui ont été mises en place afin de développer la recherche méthodologique, d'avoir des bases de données de référence pour la validation des modèles numériques et d'étudier la variabilité spatio-temporelle des phénomènes étudiés ; la distinction de ces échelles est essentielle (Figure A.1).

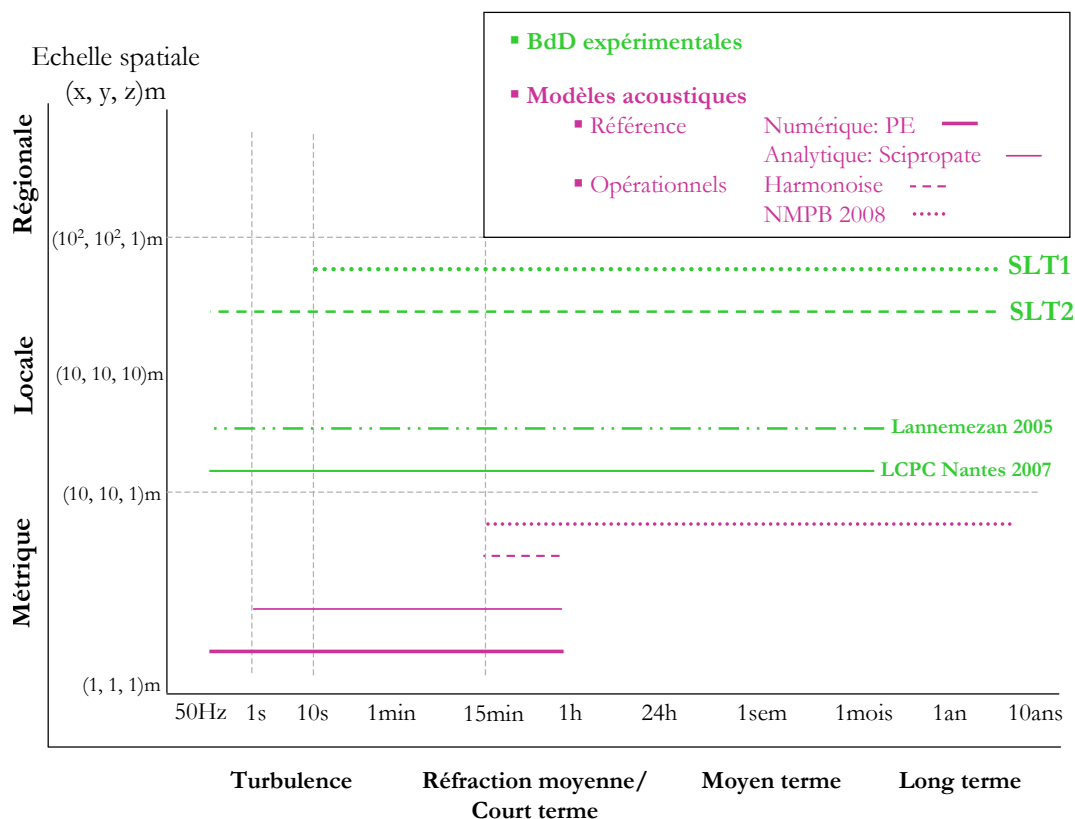


FIG. A.1 – Échelles spatiales et temporelles explorées par les bases de données expérimentales et les modèles acoustiques à disposition.

La campagne de Lannemezan 2005 permet de travailler à une échelle locale (maillage fin) et sur des échelles temporelles allant de la turbulence au moyen terme. La station de Long Terme fournit des données à une échelle locale (voire régionale) et sur des échelles temporelles allant de la turbulence au long terme. La campagne micrométéorologique menée sur le site du LCPC Nantes en 2007 permet d'étudier la variabilité temporelle des données ainsi que les incertitudes de mesures (comparaisons

inter-capteurs) sur une échelle spatiale locale et sur des échelles temporelles allant de la turbulence au moyen terme.

1 STATION DE LONG TERME DE SAINT BERTHEVIN (53)

1.1 CADRE ET OBJECTIFS

Dans le but de mieux maîtriser et comprendre les phénomènes physiques relatifs à la propagation acoustique en milieu extérieur hétérogène, un site expérimental permanent a été mis en place par le LCPC : la Station de Long Terme (SLT) de Saint-Berthevin [19]. Cette « opération de développement » du LCPC, lancée en 1999, est opérationnelle depuis 2001 et est prévue pour couvrir une période d'au moins 10 ans. Un tel site d'observation en continu sur une longue période présente plusieurs intérêts majeurs et notamment la possibilité de quantifier l'influence des conditions micrométéorologiques sur la variabilité d'un champ acoustique à l'échelle locale (celle du site) relativement à l'échelle régionale (celle des stations Météo-France). Les travaux de recherche relatifs à la SLT portent de plus sur l'aspect temporel (court/long terme), dans le but d'accéder à l'estimation de la variabilité temporelle de l'ensemble des grandeurs d'intérêt sur un site quelconque à court, moyen et long termes. Les données issues de la SLT doivent également permettre de valider les modèles numériques acoustiques et micrométéorologiques développés par le LCPC et par des organismes de recherche associés [83].

1.2 DISPOSITIF EXPÉRIMENTAL

Le site sur lequel devait être implantée la Station de Long Terme devait nécessairement respecter un cahier des charges spécifique. Dans ce cahier des charges, les principales caractéristiques d'intérêt étaient ainsi définies : topographie complexe, sources de bruit d'origine routière et ferroviaire, sol de nature relativement homogène, proximité d'une station Météo-France. Le site de Saint-Berthevin (Mayenne) respectait tous ces critères et a donc été choisi pour installer cette station expérimentale. Une acquisition en continu (échantillonnage temporel : 10 s) des grandeurs acoustiques (niveau global, 1/1 octave puis 1/3 octave à partir de l'année 2008, etc.), micrométéorologiques (vitesse et direction du vent, température de l'air et du sol, pluviométrie, solarimétrie, etc.) et de caractéristiques de sol (impédance) en plusieurs points du site : 5 mâts acoustiques de 5 m (2 hauteurs de capteurs), 2 mâts météo de 25 m (3 hauteurs de capteurs) et 2 mâts météo de 10 m (2 hauteurs de capteurs) équipés de sondes micrométéorologiques (figure A.2). Un mât acoustique de référence est placé à proximité de la voie autoroutière et également d'une station de comptage de trafic routier (qui est également en fonctionnement 24 h/ 24 et 365 j/an). Les récents développements et équipements de la SLT en collaboration avec le CCEP¹ d'Angers (SLT2) permettent de travailler à des échelles temporelles plus fines permettant d'identifier et de caractériser plus en détails les phénomènes de turbulence et de Long Terme.

1.3 EXPLOITATION

Un travail de validation de la base de données acoustique et météorologique de la Station de Long Terme a également été mené en collaboration avec le LRPC² Blois. Cette validation est indépendante de la validation quotidienne des données qui permet la détection ponctuelle de dérives de capteurs et/ou de problèmes météorologiques. Cette base de données ainsi validée permet notamment de procéder à la validation de modèles numériques de propagation acoustique et de mener des analyses statistiques approfondies. Des exemples de résultats sont données après validation de la base de données par les figures A.3 à A.14.

1. Centre d'Etudes et de Conception de Prototypes

2. Laboratoire Régional des Ponts et Chaussées

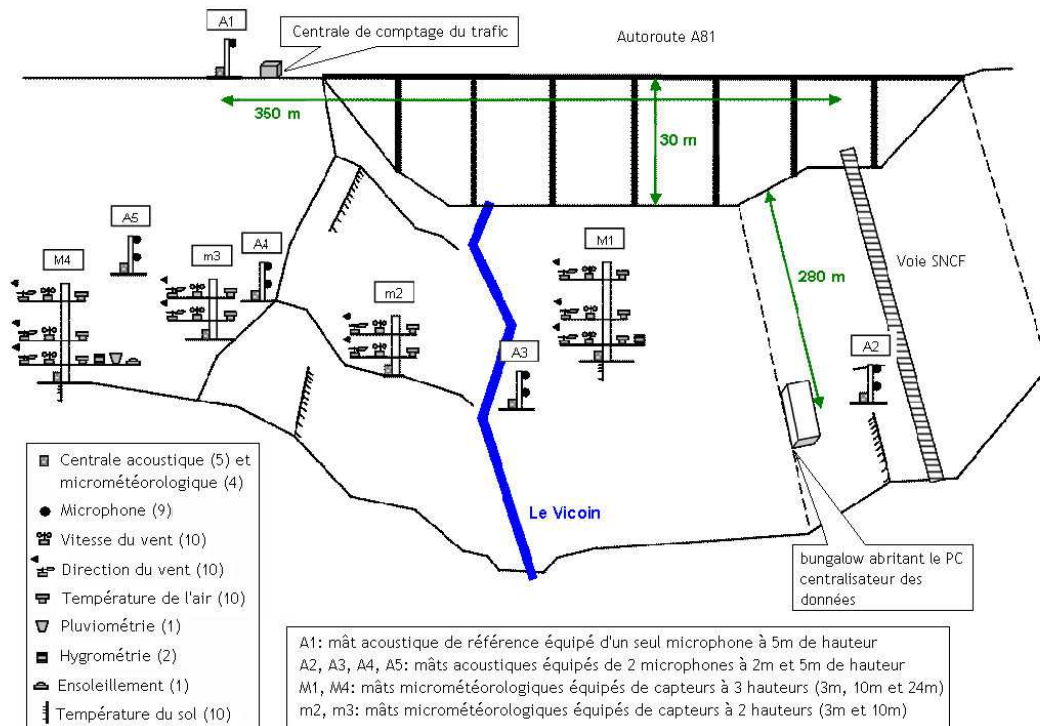


FIG. A.2 – Présentation schématique de la Station de Long Terme sur le site de Saint-Berthevin.

2 CAMPAGNE EXPÉRIMENTALE DE LANNEMEZAN 2005 (65)

2.1 CADRE ET OBJECTIFS

L'objectif de cette campagne expérimentale est de constituer un important jeu de données expérimentales afin de quantifier l'influence relative de certains paramètres relatifs à la propagation acoustique à grande distance en espace semi ouvert (dégagé de tout obstacle) et plan [68]. Ainsi le processus expérimental qui a été retenu vise à quantifier les effets de sol et les effets micrométéorologiques sur la propagation acoustique, en fonction de la géométrie source-récepteur considérée (distance, hauteurs). Ces mesures permettent également la validation de modèles numériques et analytiques de prévision de niveaux sonores, modèles qui pourront servir à compléter le jeu de données. Ils permettent en effet de simuler des situations et conditions (micrométéorologiques, caractéristiques de sol, distances, etc.) qui n'ont pas été rencontrées ou qui n'ont pas pu être mesurées au cours de la campagne de mesures. Cette campagne s'est déroulée en deux phases : une campagne courte (de 3 semaines d'acquisition permanente) avec un nombre maximum de capteurs et un taux d'échantillonnage élevé, puis une campagne longue (de 2 mois d'acquisition permanente) avec des échantillonnages en temps (fréquence d'acquisition) et en espace (position et nombre de capteurs) moins conséquents. Cette campagne expérimentale permet donc de récolter des jeux de données à une échelle spatiale locale uniquement mais avec un échantillonnage et un maillage relativement fin.

2.2 DISPOSITIF EXPÉRIMENTAL

La figure A.15 représente le dispositif expérimental sur le site de Lannemezan. Ce dispositif se compose des éléments suivants :

- *source sonore* : elle est placée relativement loin des arbres, afin de limiter l'effet de ces derniers (perturbation de l'écoulement d'air) à son voisinage. La source est omnidirectionnelle et station-

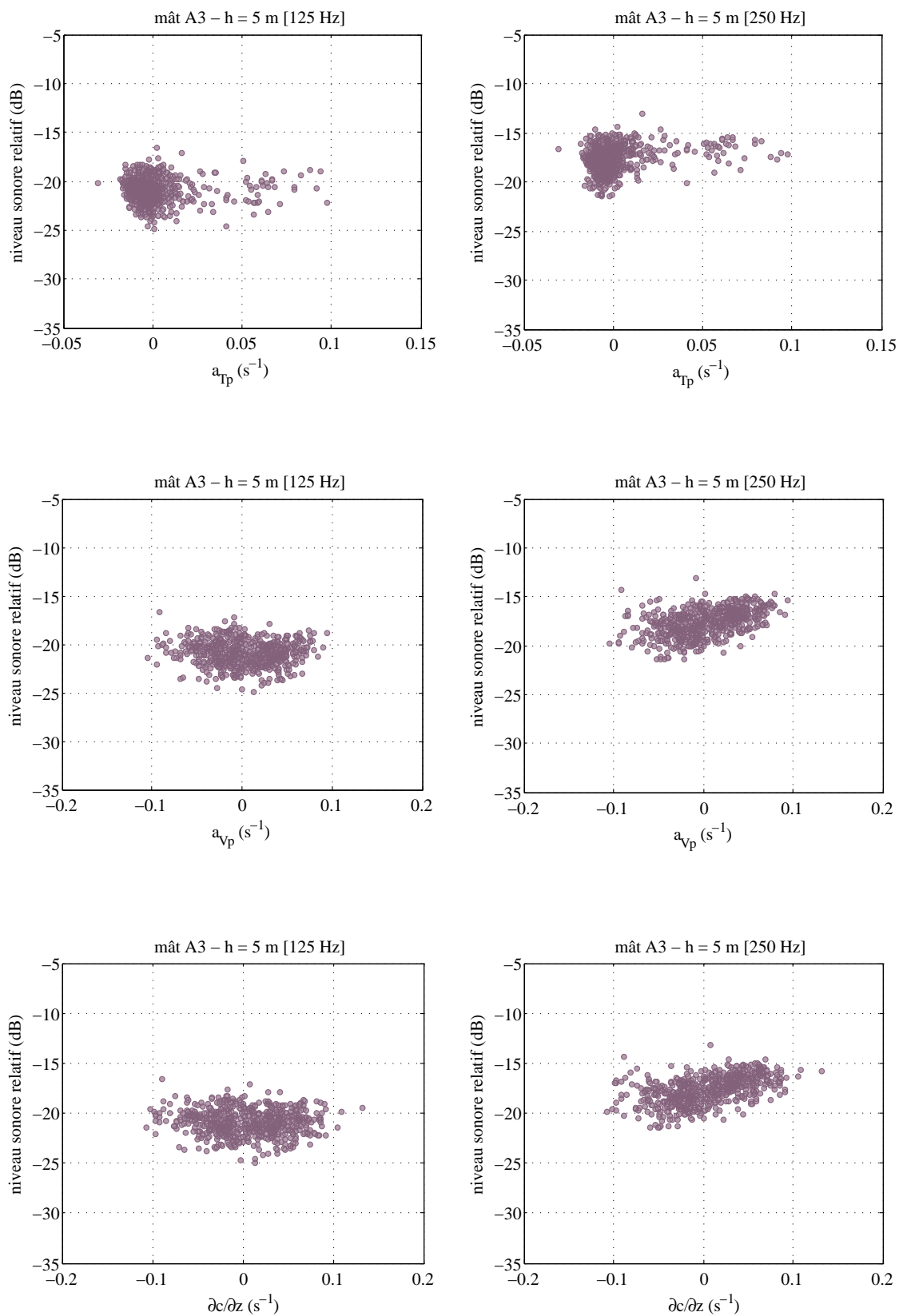


FIG. A.3 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 5 m de hauteur relativement au mât A1 à 125 Hz et 250 Hz.

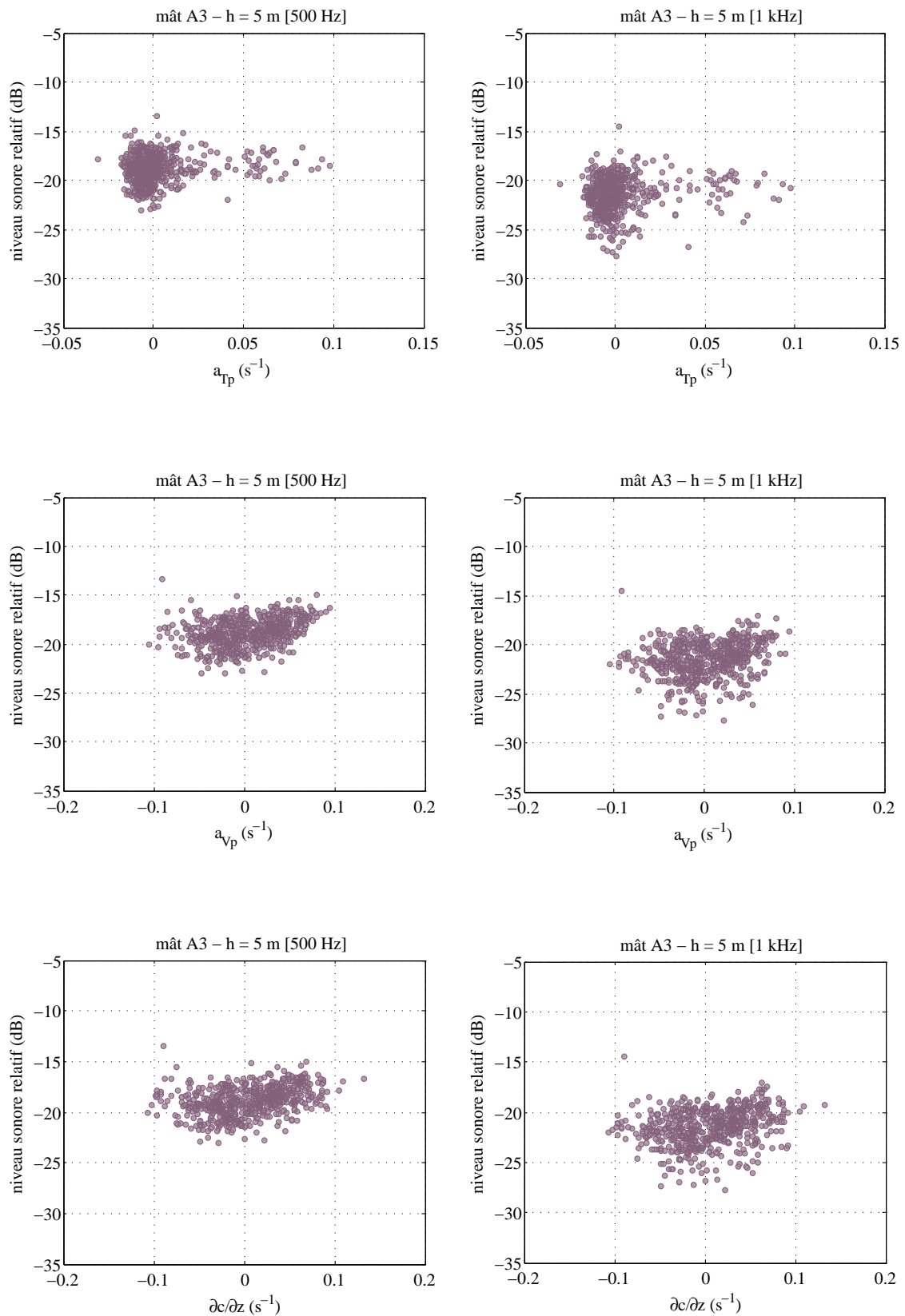


FIG. A.4 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 5 m de hauteur relativement au mât A1 à 500 Hz et 1 kHz.

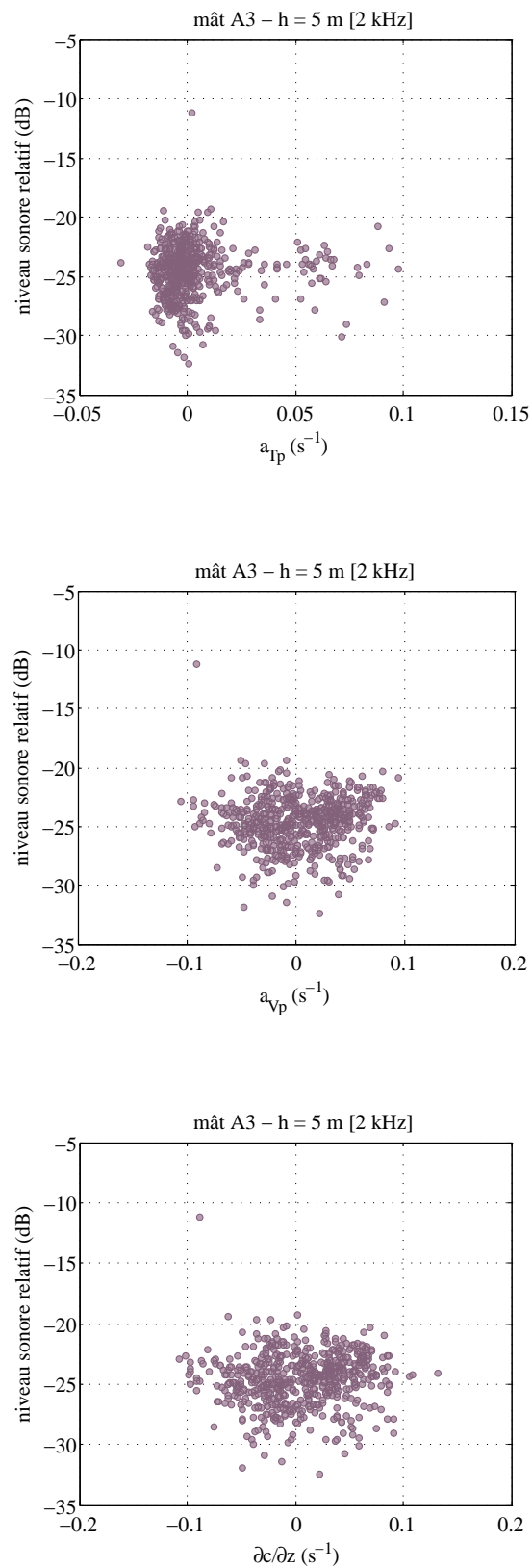


FIG. A.5 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 5 m de hauteur relativement au mât A1 à 2 kHz.

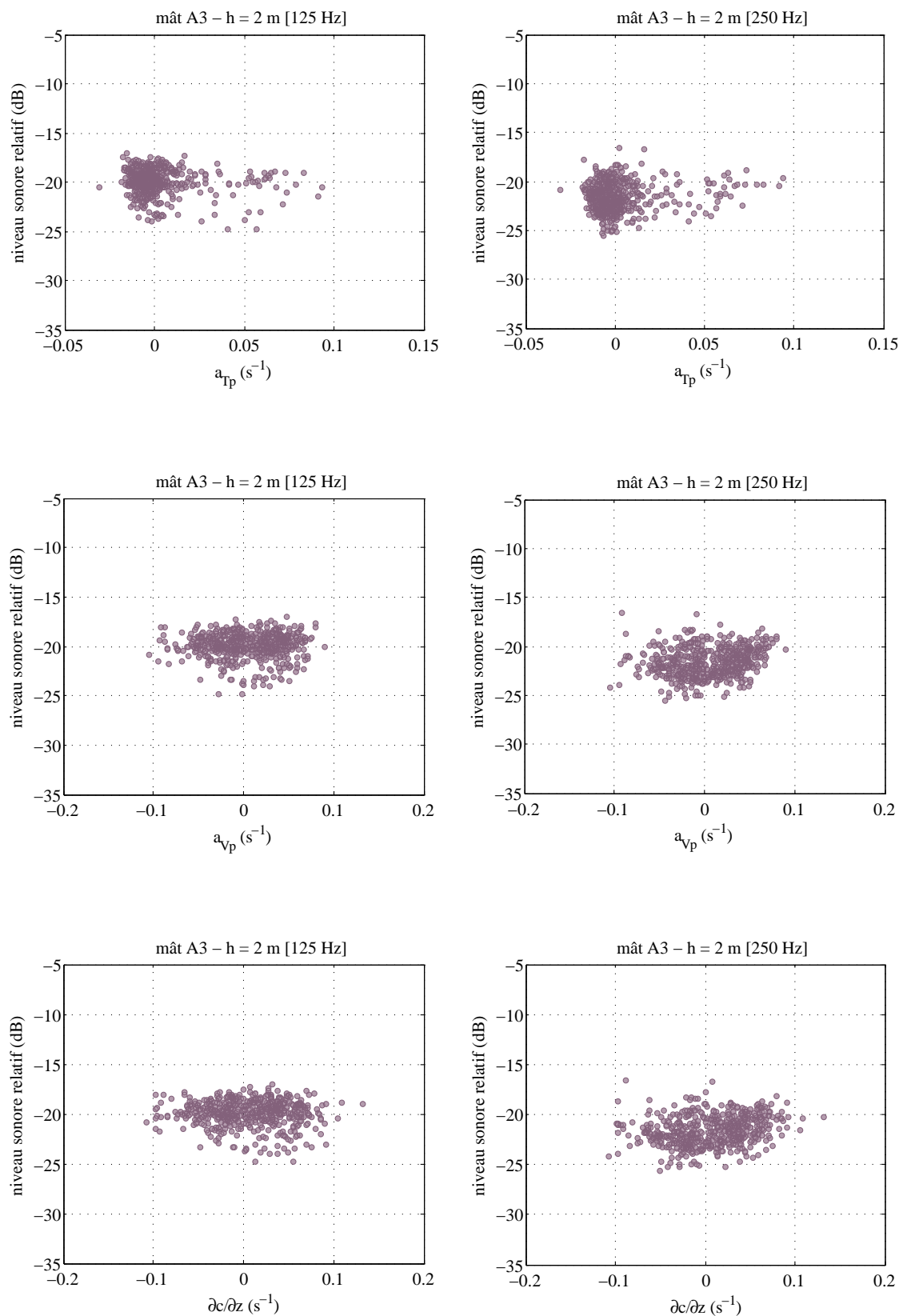


FIG. A.6 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 2 m de hauteur relativement au mât A1 à 125 Hz et 250 Hz.

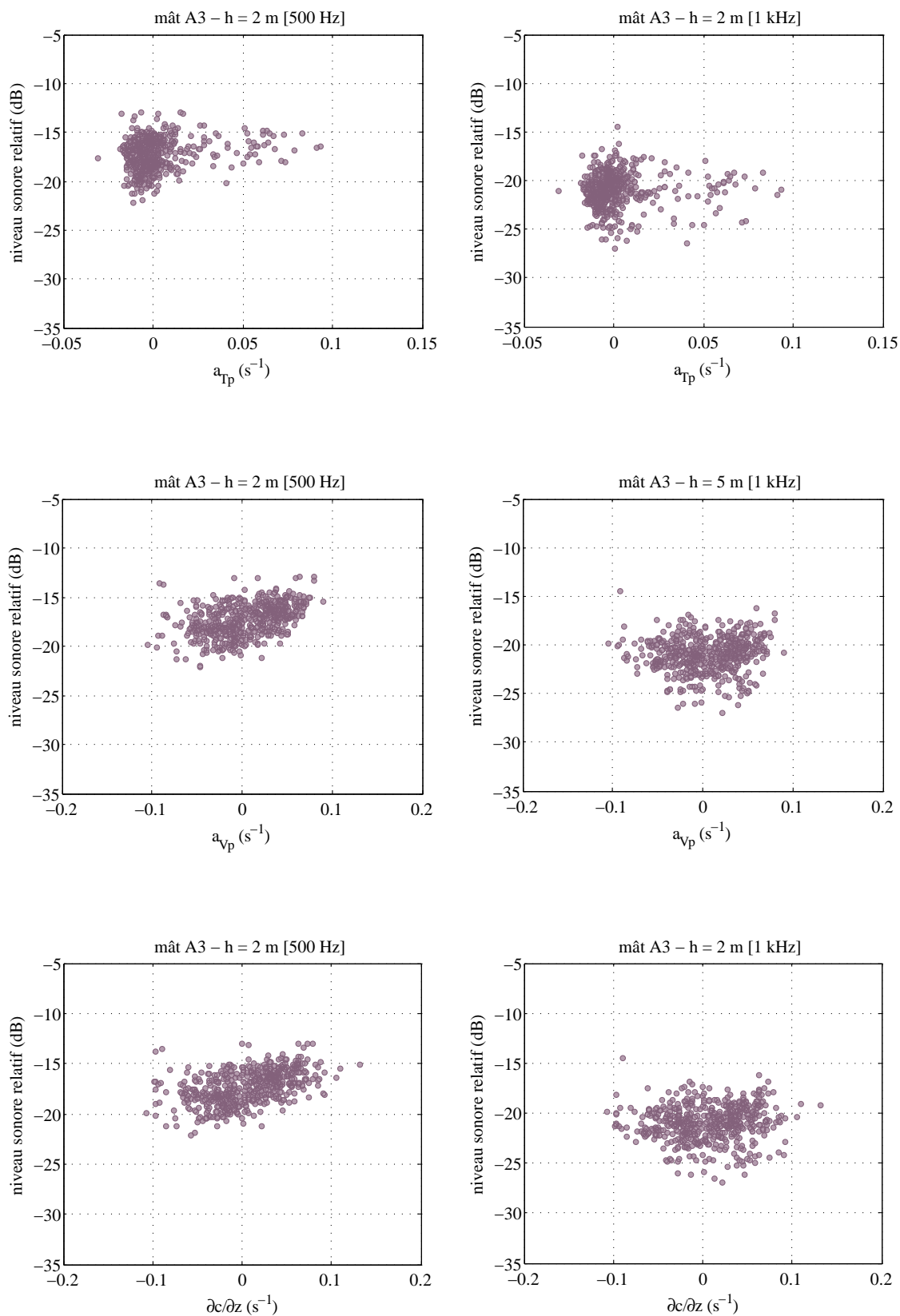


FIG. A.7 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 2 m de hauteur relativement au mât A1 à 500 Hz et 1 kHz.

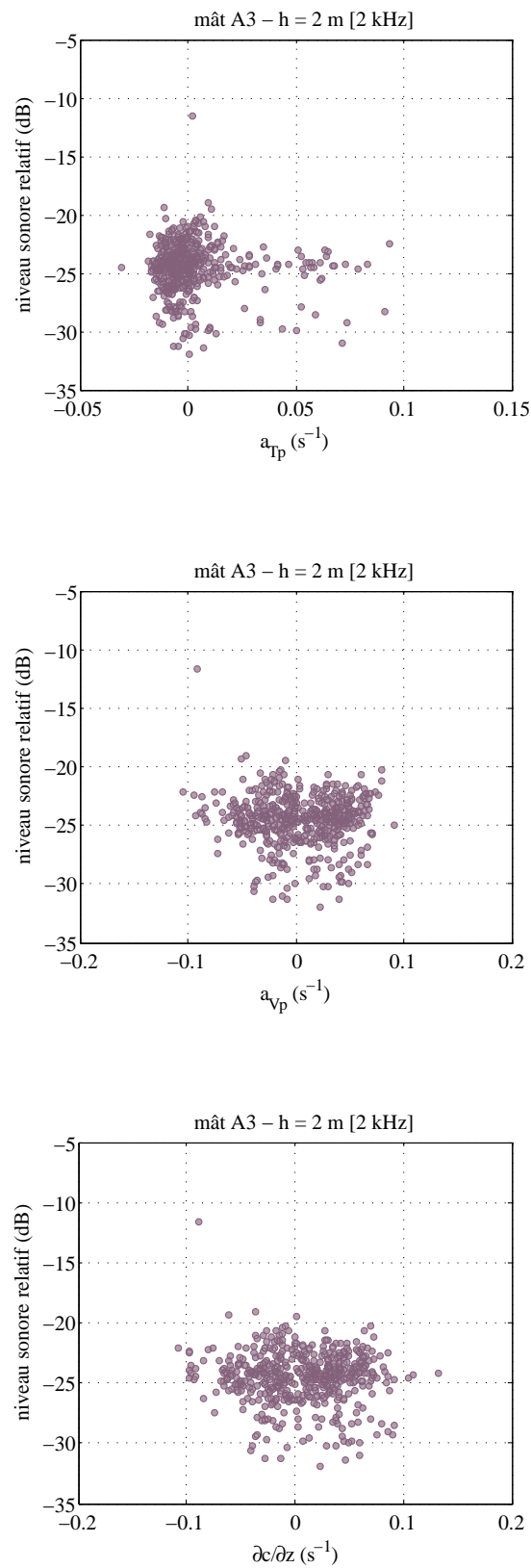


FIG. A.8 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A3 à 2 m de hauteur relativement au mât A1 à 2 kHz.

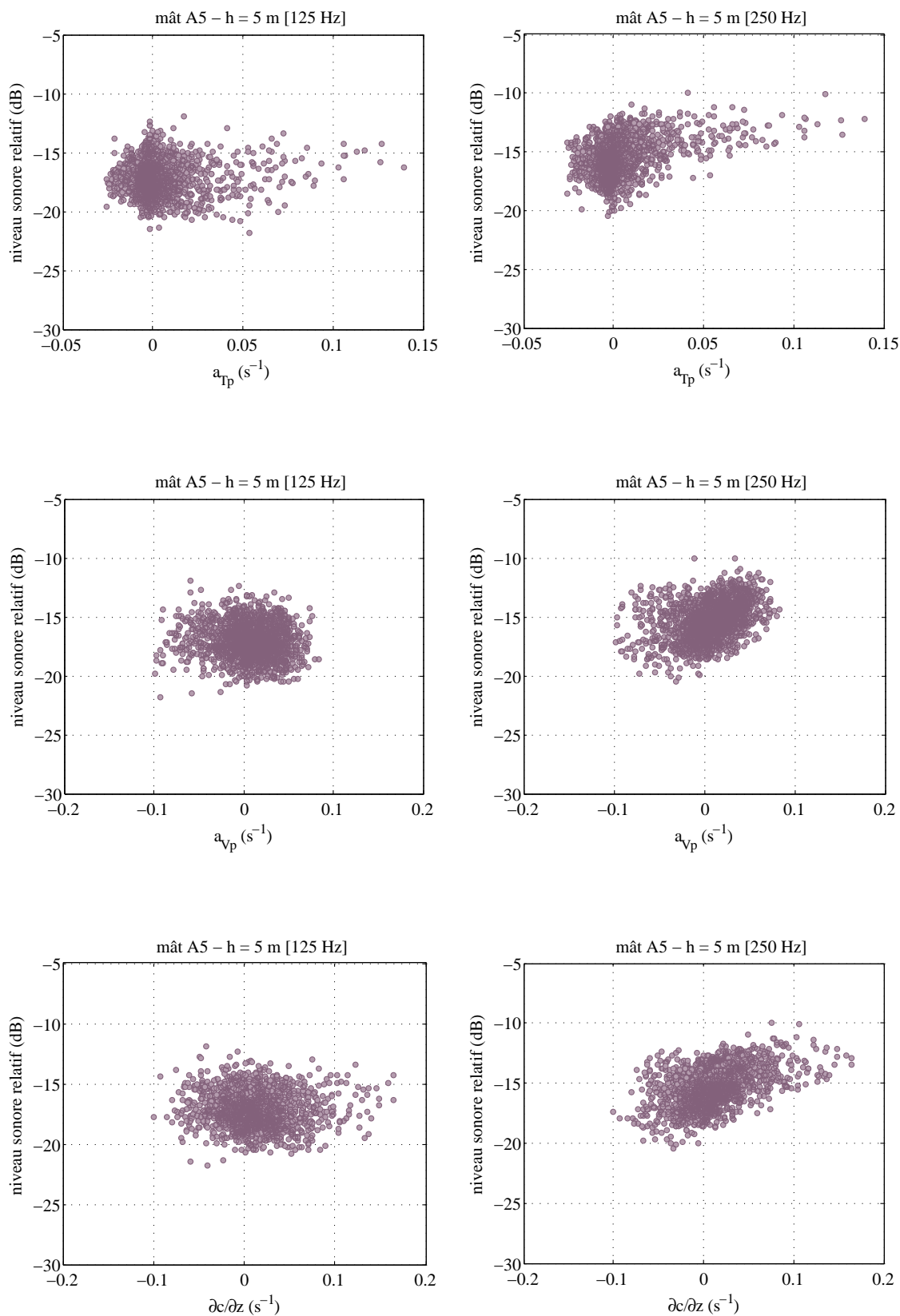


FIG. A.9 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 5 m de hauteur relativement au mât A1 à 125 Hz et 250 Hz.

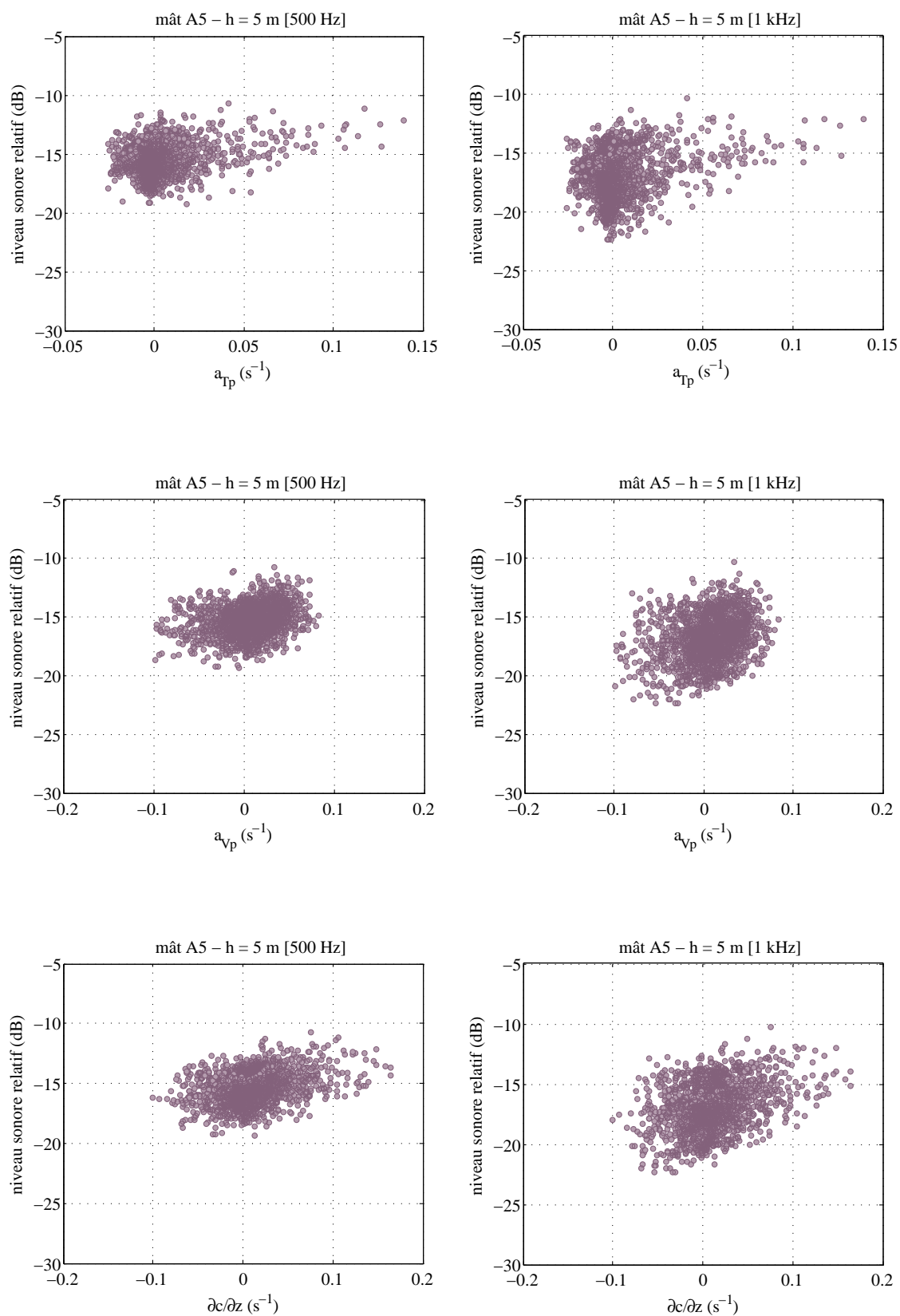


FIG. A.10 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 5 m de hauteur relativement au mât A1 à 500 Hz et 1 kHz.

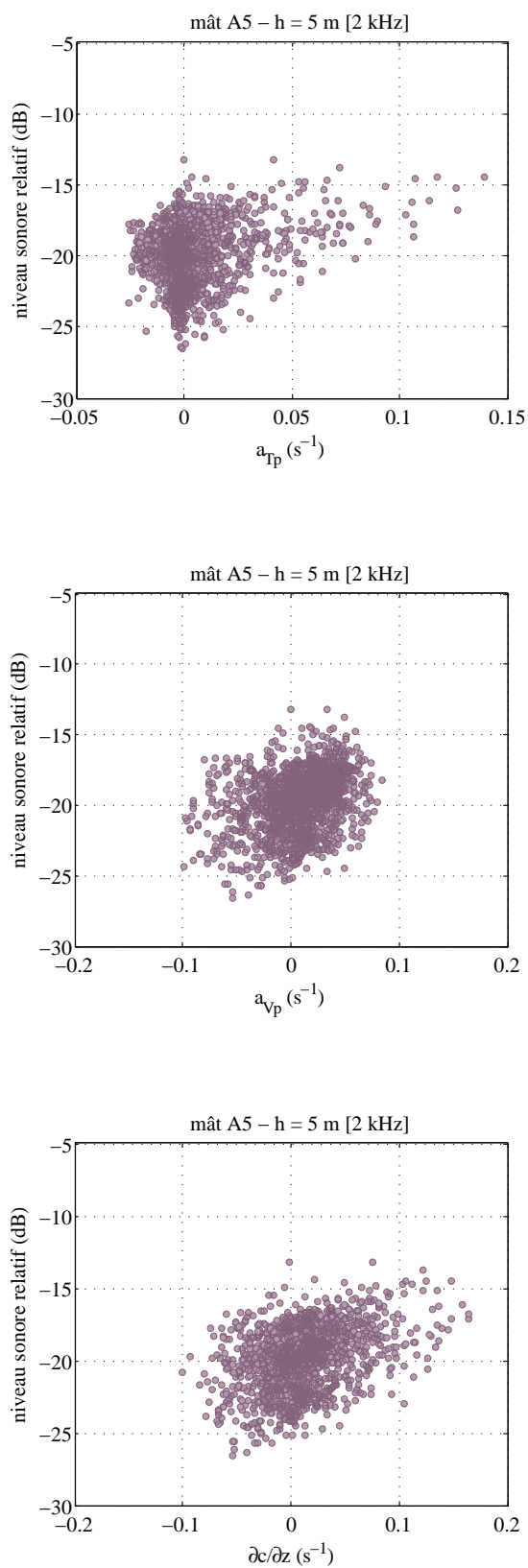


FIG. A.11 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 5 m de hauteur relativement au mât A1 à 2 kHz.

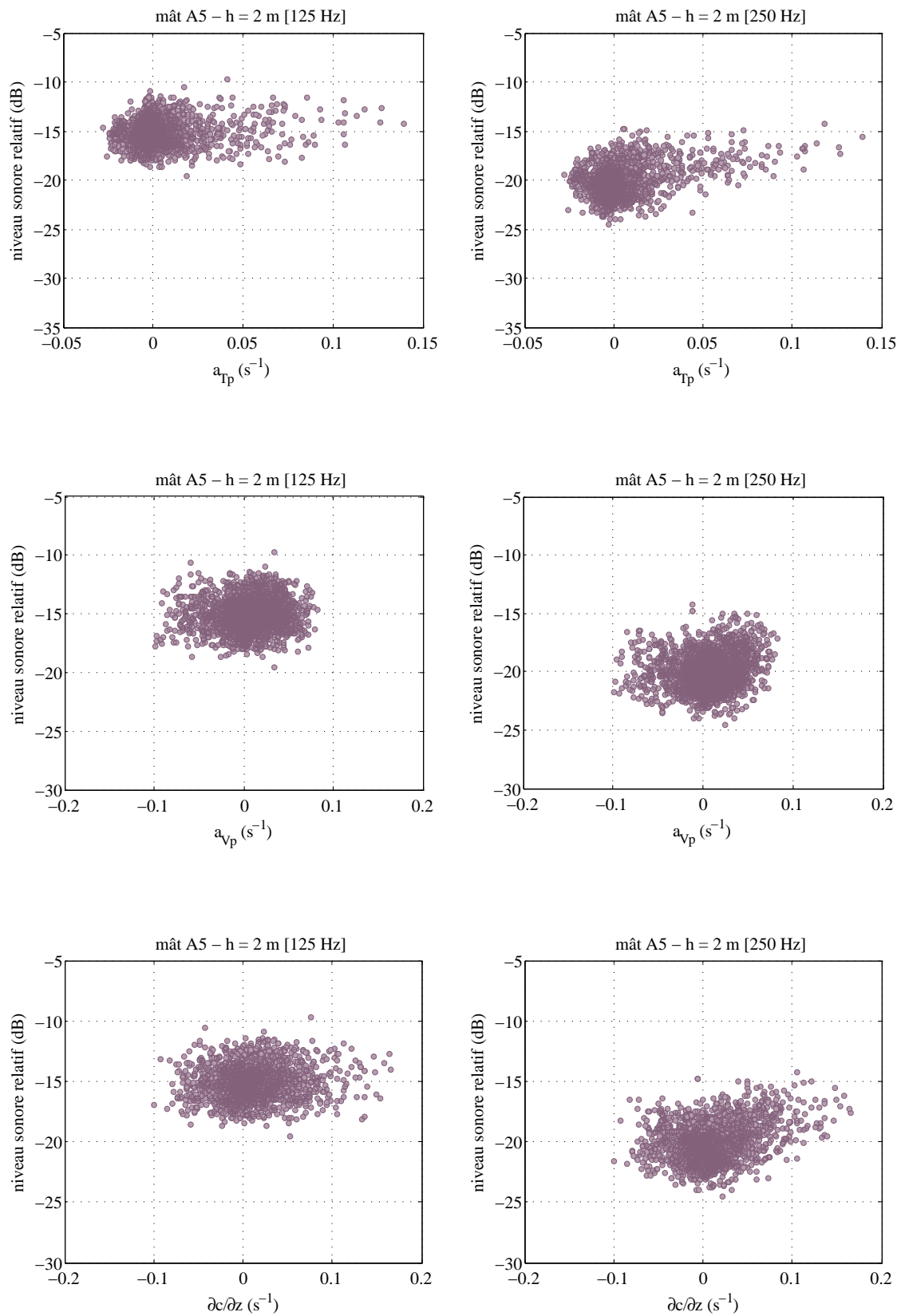


FIG. A.12 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 2 m de hauteur relativement au mât A1 à 125 Hz et 250 Hz.

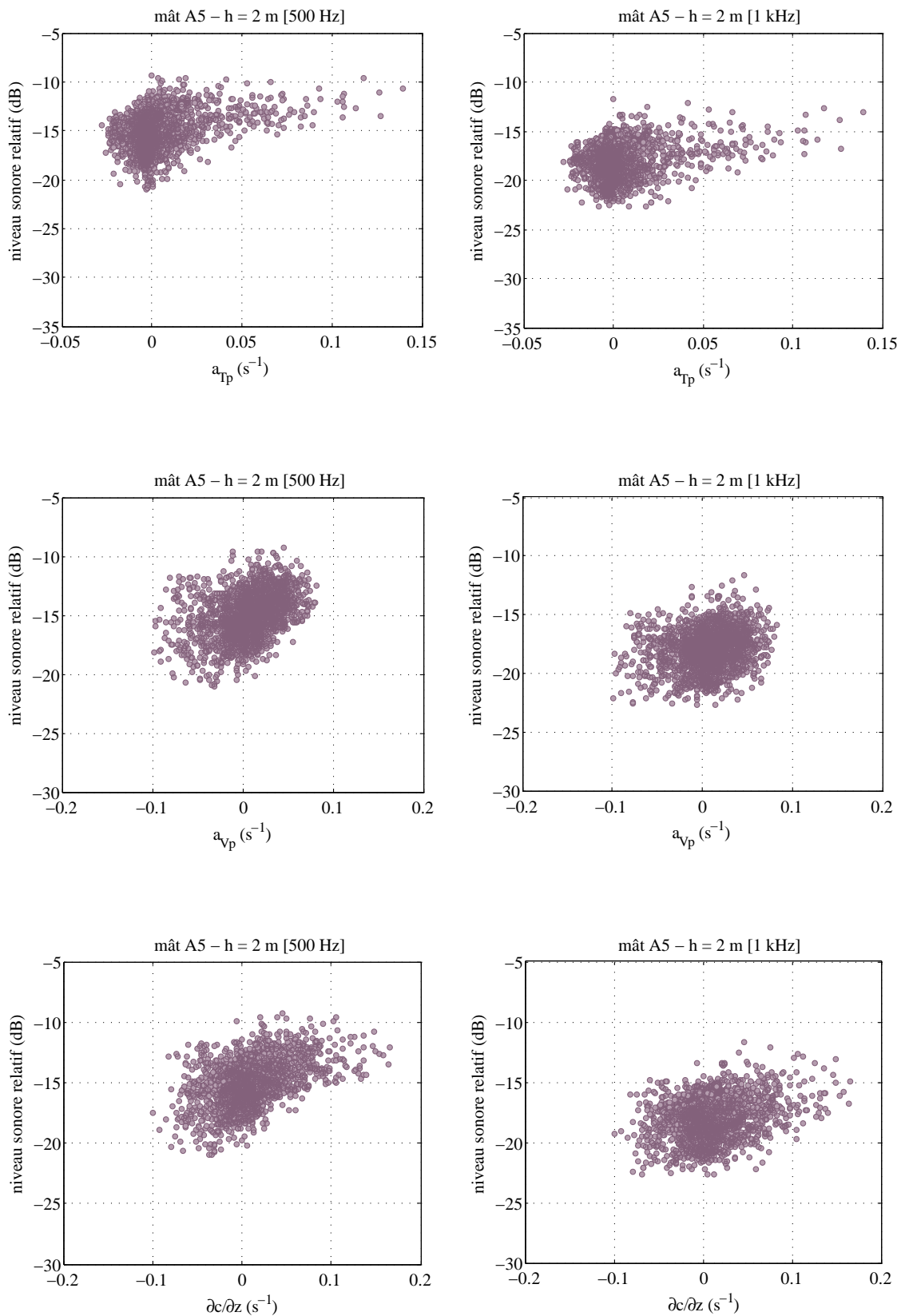


FIG. A.13 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 2 m de hauteur relativement au mât A1 à 500 Hz et 1 kHz.

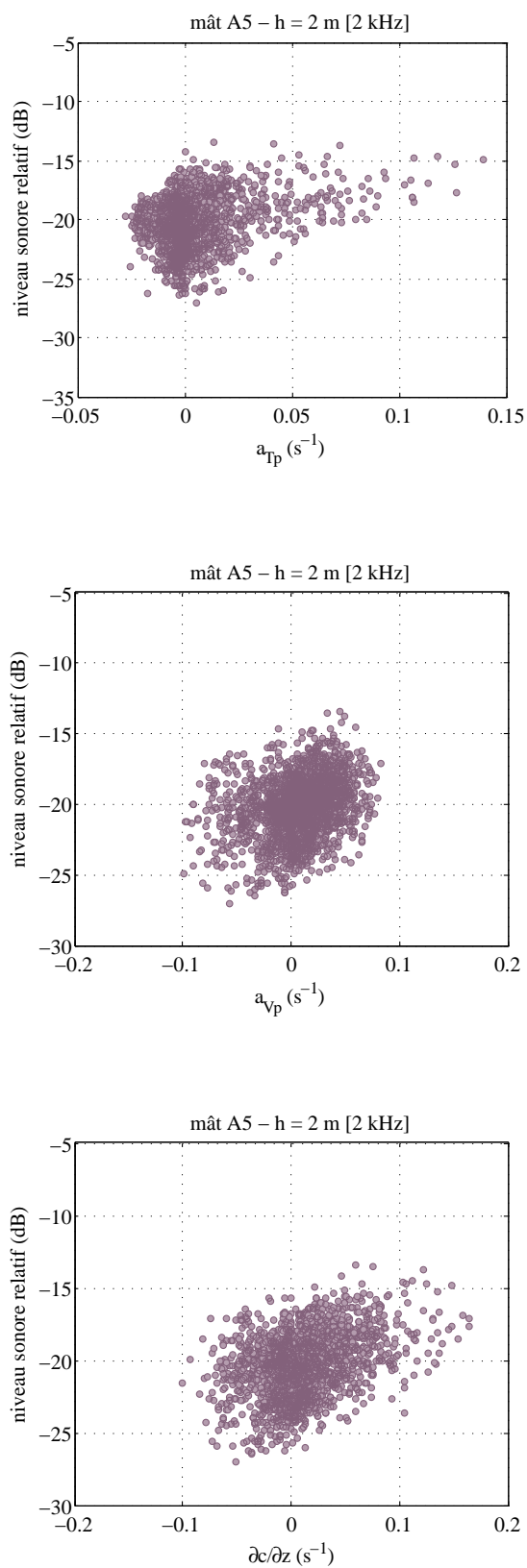


FIG. A.14 – Résultats expérimentaux issus de la SLT (2002-2007) – niveaux sonores mesurés au mât A5 à 2 m de hauteur relativement au mât A1 à 2 kHz.

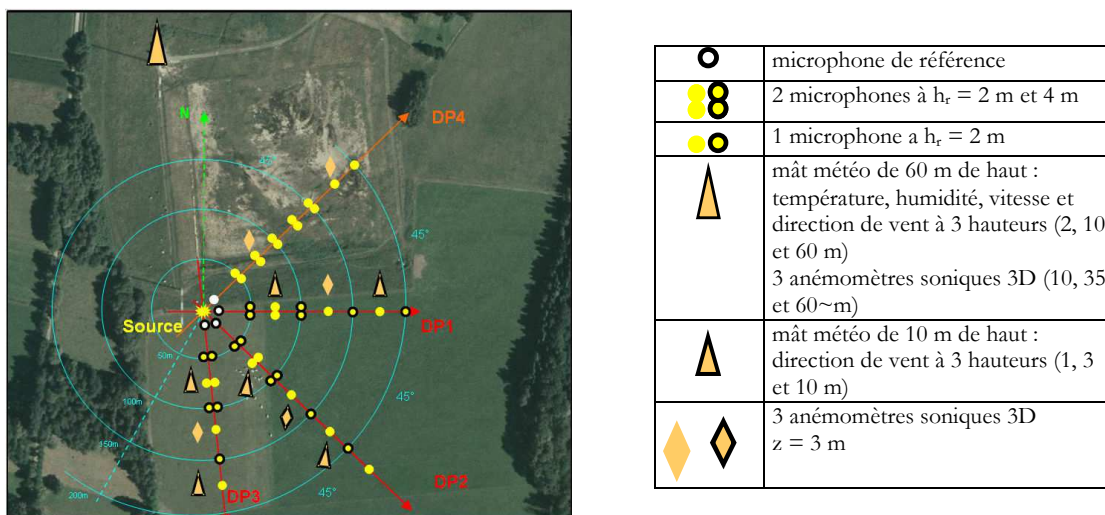


FIG. A.15 – Présentation schématique de la campagne expérimentale de Lannemezan 2005.

naire (en spectre et en amplitude). Elle a été placée à 2 m de hauteur et peut être considérée comme ponctuelle en regard des distances considérées ;

- *les capteurs acoustiques* : ils ont été disposés suivant 4 directions de propagation privilégiées, par pas de 25 m à partir de 50 m de la source. Sur chacune des directions, certaines voies sont doublées, en particulier celles de référence (proches de la source) et celles très éloignées, afin de vérifier la précision du matériel et d'assurer un secours en cas de dysfonctionnement du matériel. Pour chaque distance à la source, l'acquisition se fait systématiquement à deux hauteurs par rapport au sol (2 m et 4 m), sauf pour les récepteurs les plus éloignés (à partir de 125 m). L'acquisition se fait par *Leq1s* pour l'ensemble des 1/3 d'octave centrés sur les fréquences comprises entre 50 Hz et 5 kHz ;
- *les capteurs micrométéorologiques* : plusieurs systèmes de caractérisation de profils verticaux sont installés sur chacune des directions de propagation afin notamment de vérifier l'hypothèse de stationnarité des profils avec la distance à la source. Dans le cas contraire (non stationnarité), les systèmes permettent d'observer l'évolution spatiale des profils verticaux de célérité du son en plus de l'évolution temporelle. Ces informations sont obtenues grâce à des capteurs météorologiques du type anémomètre, girouette, sonde de température ventilée ou bien grâce à un anémomètre sonique tridimensionnel ;
- *l'impédance acoustique du sol* : les deux méthodes permettant de mesurer l'impédance acoustique de sol sont la mesure en continu en un seul point au voisinage de la source (méthode EDF R&D) et la mesure journalière en plusieurs points du site (méthode LCPC). Afin de perturber un minimum les acquisitions acoustiques, les points de mesure d'impédance sont spatialement distribués entre les mâts de mesures acoustiques et micrométéorologiques. Le suivi temporel continu de l'impédance est assuré en un seul point au voisinage de la source (sur DP3, à 10 m de la source).

2.3 EXPLOITATION

Afin de conserver un maximum de données, seules les données micrométéorologiques mesurées à 75 m de la source seront utilisées pour les différentes directions de propagation considérées (DP1,

DP2 et DP3). La cohérence de ce choix peut être vérifiée en calculant le rapport des variances noté NSD (Normalised Standard Deviation) (tableau A.1) et le coefficient de corrélation ρ (tableau A.2) sur les mesures du gradient vertical de célérité, entre les différentes directions de propagation et aux différentes distances à la source considérées.

Le calcul de la corrélation permet de vérifier la cohérence temporelle entre les données et le calcul de la NSD les variations en amplitude du gradient de célérité au cours d'une période de mesures donnée. Ces deux paramètres combinés constituent de bons outils de comparaison. Ces calculs se font à partir de jeux de données complets. Si par exemple à un instant t une donnée est inexistante en $DP1_{75}$, les données sur toutes autres directions de propagation et à toutes les distances seront « supprimées » à cet instant t .

NSD	$DP1_{75}$	$DP2_{75}$	$DP3_{75}$	$DP1_{125}$	$DP2_{125}$	$DP3_{125}$	$DP1_{175}$	$DP2_{175}$	$DP3_{175}$
$DP1_{75}$	1	1,34	1,34	0,39	0,44	0,49	1,26	1,30	1,44
$DP2_{75}$		1	1	0,29	0,33	0,36	0,94	0,98	1,08
$DP3_{75}$			1	0,29	0,33	0,36	0,94	0,98	1,08
$DP1_{125}$				1	1,11	1,22	3,17	3,28	3,64
$DP2_{125}$					1	1,10	2,85	2,96	3,27
$DP3_{125}$						1	2,59	2,68	2,97
$DP1_{175}$							1	1,04	1,15
$DP2_{175}$								1	1,11
$DP3_{175}$									1

TAB. A.1 – Résultats de calculs de variance normalisée (NSD) entre les gradients de célérité effective pour toutes les directions de propagation (DP) et à toutes les distances à la source (75, 125, 175 m).

Pour une même distance le long d'une ligne de propagation et pour les différentes directions de propagation, les valeurs de NSD sont proches de l'unité; l'amplitude des données présente donc une bonne cohérence spatiale. Cependant, à matériel équivalent, si l'on compare maintenant les résultats selon les différentes distances à la source pour une même direction de propagation, une correspondance est observée entre les données à 75 et 175 m; les valeurs de NSD ne sont alors pas cohérentes lorsque l'on considère les valeurs à 125 m. Ces dernières observations sont vérifiées pour chacune des directions de propagation. En croisant à la fois les directions de propagation et les distances à la source, les observations sont similaires au cas précédent : il existe une divergence entre les données à 125 m et les données aux deux autres distances considérées. Cette différence s'explique par le fait que les valeurs de gradient à 75 et 175 m sont issues d'un ajustement après mesure par mâât météorologique et que les valeurs à 125 m ont été reconstituées à partir de mesures effectuées par anémomètre ultrasonique tridimensionnel. Il existe donc des disparités de résultats selon la méthode utilisée. Une étude comparative et quantitative a d'ailleurs été menée en ce sens et rend compte d'un écart conséquent entre les méthodes (Annexe E).

Pour une direction de propagation donnée, les gradients verticaux de célérité présentent tous une bonne cohérence temporelle quelles que soient les distances considérées (tableau A.2) : les valeurs de gradients sont donc cohérentes aussi bien d'un point de vue temporel que spatial. A une même distance à la source mais pour différentes directions de propagation, les valeurs du coefficient de corrélation entre DP3 et tous les autres points de mesure, présentent systématiquement une divergence. Cette différence peut être notamment expliquée par la présence d'une barrière d'arbres le long de DP3; l'effet de cette barrière d'arbres est donc à prendre en compte. Les calculs entre DP1 et DP2 montrent cependant une bonne cohérence temporelle.

ρ	DP175	DP275	DP375	DP1125	DP2125	DP3125	DP1175	DP2175	DP3175
DP175	1	0,73	0,23	0,91	0,71	0,05	0,96	0,65	0,16
DP275		1	0,77	0,74	0,91	0,61	0,74	0,90	0,70
DP375			1	0,29	0,70	0,89	0,29	0,80	0,95
DP1125				1	0,74	0,14	0,90	0,68	0,22
DP2125					1	0,58	0,75	0,87	0,63
DP3125						1	0,12	0,66	0,87
DP1175							1	0,71	0,24
DP2175								1	0,77
DP3175									1

TAB. A.2 – Résultats de calculs du coefficient de corrélation (ρ) entre les gradients de célérité effective pour toutes les directions de propagation et à toutes les distances à la source (75, 125, 175 m).

D'après ces observations, il va être nécessaire de traiter chacune des directions de propagation de manière indépendante et de ne pas considérer les mesures en un point de l'espace comme représentatives de l'ensemble du site. Des conditions vont ainsi être imposées sur les valeurs de coefficient de corrélation et de variance normalisée à « respecter » pour que les gradients de célérité estimés en un point soient considérés comme similaires aux gradients mesurés en un autre point du site. On imposera des conditions plus sévères sur le coefficient de corrélation que sur la variance normalisée car en effet, la variance normalisée correspond aux variances calculées sur toute la période considérée alors que le calcul du coefficient de corrélation prend en compte la variation des données pas à pas. Les conditions de validation sont les suivantes : $\rho > 0,6$ et $0,4 < NSD < 2$. On indexera les résultats par la valeur 1 si les conditions sont respectées ou 0 dans le cas contraire (Tableau A.3).

	DP175	DP275	DP375	DP1125	DP2125	DP3125	DP1175	DP2175	DP3175
DP175	1	1	0	0	1	0	1	1	0
DP275		1	1	0	0	0	1	1	1
DP375			1	0	0	0	0	1	1
DP1125				1	1	0	0	0	0
DP2125					1	0	0	0	0
DP3125						1	0	0	0
DP1175							1	1	0
DP2175								1	1
DP3175									1

TAB. A.3 – Résultats vérifiant les conditions imposées sur le coefficient de corrélation et sur la variance normalisée entre les gradients de célérité effective pour toutes les directions de propagation et à toutes les distances à la source (75, 125, 175 m).

Le tableau A.3 permet de faire la synthèse des tableaux précédents. On retrouve bien une divergence entre les données à 125 m et les données mesurées à 75 m et 175 m, divergence mise en évidence par le calcul de la variance normalisée. De plus on retrouve cette différence observée en DP3 par comparaison avec les autres directions de propagation et ce à toutes les distances. Il est donc nécessaire de traiter de manière indépendante les différentes directions de propagation pour éviter tout biais et préférable de travailler avec les mesures de gradient à 75 m ou 175 m. Conclusion qui est de plus justifiée par le fait que les mesures de gradients verticaux de célérité effective à 125 m n'ont été effectuées qu'au cours de la première partie de la campagne et cet échantillonnage n'est donc pas représentatif de toute la campagne de mesure. Le nombre d'échantillons à 75 m étant le plus important et donc le

plus représentatif statistiquement, les mesures à 75 m sont donc utilisées pour caractériser le site et ce pour chacune des directions de propagation.

3 CAMPAGNE MICROMÉTÉOROLOGIQUE DU LCPC NANTES 2007

3.1 CADRE ET OBJECTIFS

Cette campagne expérimentale a été mise en place dans le but d'étudier les origines de la dispersion que l'on observe sur les mesures du profil vertical de célérité du son. En effet, il existe différentes approches permettant la caractérisation expérimentale du milieu, présentant des incertitudes qui leur sont propres : influence du type de matériel (mât équipé ou anémomètre ultrasonique tridimensionnel, marque des capteurs), de la méthodologie (hauteur des capteurs, fréquence d'échantillonnage, intervalle d'étude, post-traitement, etc.).

3.2 DISPOSITIF EXPÉRIMENTAL

Les mesures sont faites en un même point du site choisi (à quelques mètres près), à des hauteurs différentes afin de comparer l'influence de la position du matériel, à une même hauteur mais avec des marques de matériel différentes et avec un dispositif différent (mât micrométéorologiques *vs* anémomètre ultrasonique tridimensionnel).

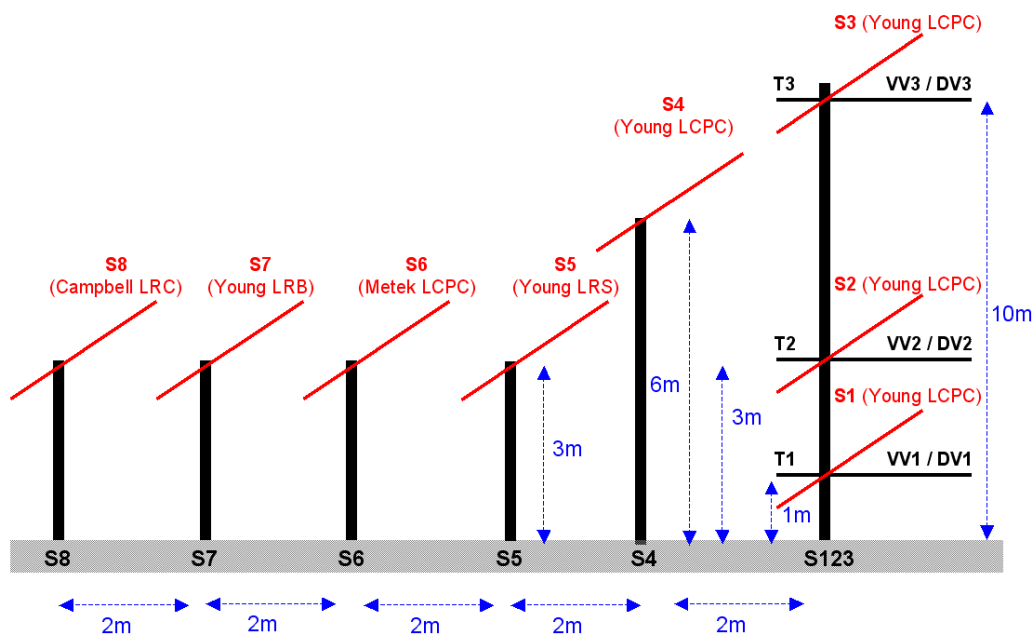


FIG. A.16 – Schéma du dispositif expérimental (coupe).

T : température ventilée

VV : vitesse de vent (anémomètre à hélice)

DV : direction du vent (girouette)

S : anémomètre ultrasonique 3D

3.3 EXPLOITATION

Cette campagne expérimentale est utilisée notamment pour différencier les différentes sources d'incertitudes, discerner laquelle est à l'origine de la dispersion généralement observée, pour finalement

qualifier et quantifier l'influence de chaque source sur le résultat final *i.e.* le gradient vertical de célérité effective du son [56, 59].

B

NOTIONS DE GÉOSTATISTIQUES

Cette annexe a pour objectif de présenter les outils géostatistiques utilisés ou abordés dans nos travaux. Ces outils ont été mentionnés dans le cadre de la construction d'un métamodèle de krigeage utilisé dans l'approche de quantification d'incertitudes et également dans le cadre de travaux menés pour mettre en évidence la variabilité spatiale des caractéristiques acoustiques de sol sur le site de la Station de Long Terme de Saint-Berthevin.

1 VARIOGRAMME

Le variogramme est un outils de la géostatistique développé pour l'analyse de la structure d'auto-corrélation de données notées Y (variables aléatoires). Il est calculé à partir de la différence quadratique entre deux échantillons (x_1 et x_2) et est défini par la relation suivante :

$$\gamma^*(x_2 - x_1) = \frac{1}{2} \text{var} [Y(x_2) - Y(x_1)]. \quad (\text{B.1})$$

La valeur moyenne est alors calculée pour toutes les paires d'échantillons (n_h) distants de $h = x_2 - x_1$. La fonction correspondante est appelée variogramme expérimental :

$$\gamma(h) = \frac{1}{2n_h} \sum_{i=1}^{n_h} [Y(x_i + h) - Y(x_i)]^2. \quad (\text{B.2})$$

Cette fonction variographique est croissante en fonction de h et elle synthétise des informations essentielles concernant le comportement des variables aléatoires. Un variogramme est caractérisé par :

- la portée *i.e.* la distance à laquelle deux observations ne sont plus corrélées ;
- le palier correspondant à la variance de la variable aléatoire ;
- l'effet de pépite traduisant la discontinuité du processus à courtes distances.

Une fonction mathématique théorique est ajustée sur le variogramme expérimental. En pratique cet ajustement se fait généralement par visualisation sans critère formel. Des méthodes de moindres carrés ordinaires ou pondérés peuvent également être utilisées pour ajuster une fonction théorique. Il existe un certain nombre de modèles variographiques : le modèle gaussien, exponentiel, sphérique, etc. [33]. Ces fonctions doivent rendre compte du comportement du variogramme à l'origine dont les propriétés sont déterminantes et sont liées à la continuité des variables. Les comportements typiques à l'origine sont :

- un comportement parabolique caractérisant une variable régulière généralement différentiable ;
- un comportement linéaire signifiant que la variable est continue mais non différentiable ;
- une discontinuité à l'origine caractérisant un effet de pépite : la variable n'est pas continue et très irrégulière ;
- une discontinuité du processus à très courte distance traduisant un effet de pépite pur ou un bruit blanc (absence totale de structure spatiale).

2 KRIGEAGE

Sacks [105] a été le premier à proposer l'utilisation de la méthode de krigeage en planification d'expériences numériques. Le krigeage doit son nom à l'ingénieur minier sud africain D.G. Krige qui a développé cette technique dans le cadre de la recherche d'or. Le krigeage est une méthode d'interpolation qui prédit les valeurs inconnues d'une fonction ou d'un processus aléatoire régionalisé, *i.e.* qui peut être décrit par sa corrélation spatiale [35]. Une prévision par krigeage est une combinaison linéaire pondérée de toutes les valeurs de sortie observées. Ces pondérateurs dépendent des distances entre les variables d'entrée pour lesquelles les variables de sortie doivent être prédites et les variables d'entrée déjà simulées (prise en compte de la structure spatiale des données). Le krigeage est fréquemment utilisé en simulation déterministe ; dans ce cas, la méthode de krigeage est une interpolation exacte. En simulation aléatoire, cette propriété disparaît ; chaque scénario est simulé plusieurs fois et le krigeage permet d'interpoler la moyenne des données résultantes. Le krigeage peut s'écrire selon la formulation suivante :

$$Y(\mathbf{x}) = m(\mathbf{x}) + \eta(\mathbf{x}) \text{ avec } \eta(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2(\mathbf{x})), \quad (\text{B.3})$$

où $m(\mathbf{x})$ est la moyenne du processus stochastique $Y(\mathbf{x})$ et $\eta(\mathbf{x})$ le bruit supposé indépendant, identiquement distribué de moyenne nulle et de variance $\sigma^2(\mathbf{x})$. L'application du krigeage ordinaire n'est possible que sous l'hypothèse de stationnarité de $Y(\mathbf{x})$, *i.e.* $m(\mathbf{x})$ est constante et les covariances de $Y(\mathbf{x} + \mathbf{h})$ et $Y(\mathbf{x})$ dépendent seulement de la distance euclidienne $\|\mathbf{h}\| = \|(\mathbf{x} + \mathbf{h}) - \mathbf{x}\|$. Le prédicteur de krigeage pour des paramètres d'entrée non observés x_0 (non simulés) est une combinaison linéaire $\hat{Y}(x_0)$ pondérée de l'ensemble des données résultantes observées :

$$\hat{Y}(\mathbf{x}_0) = \sum_{i=1}^n \lambda_i Y(x_i) = \lambda' \mathbf{Y}, \quad (\text{B.4})$$

avec $\sum_{i=1}^n \lambda_i = 1$, $\lambda = (\lambda_1, \dots, \lambda_n)'$ et $\mathbf{Y} = (Y(x_1), \dots, Y(x_n))'$. Les pondérateurs λ_i sont calculés en minimisant l'erreur quadratique moyenne de prévision. La technique utilise le variogramme défini ci-dessus.

Les pondérateurs optimaux s'écrivent alors selon l'équation B.5 :

$$\lambda' = \left(\gamma + \mathbf{1} \frac{1 - \mathbf{1}\Gamma^{-1}\gamma}{\mathbf{1}\Gamma^{-1}\mathbf{1}} \right) \Gamma^{-1}, \quad (\text{B.5})$$

où γ est le vecteur $(\gamma(x_0 - x_1), \dots, \gamma(x_0 - x_n))'$, Γ est une matrice $n \times n$ dont le (i, j) ème élément est $\gamma(x_i - x_j)$ et $\mathbf{1}$ est un vecteur unité. A noter que les pondérateurs λ_i de l'équation B.5 varient en fonction des points de prévision.

La régularité de la fonction est choisie au travers d'une fonction de corrélation et de paramètres déterminés par les observations. Le krigeage permet de plus d'associer une variance de prévision au métamodèle. Ainsi, il est intéressant de tenir compte de cette information en jouant différemment sur les zones connues avec précision et celles où très peu d'observations sont disponibles.

Les principaux inconvénients du krigeage sont essentiellement numériques. En effet, l'inversion de la matrice de corrélation peut être difficile du fait d'un éventuel mauvais conditionnement. Ceci est observé en grande dimension du fait d'un nombre d'observations souvent insuffisant pour permettre l'identification des paramètres. De plus, la fonction coût permettant l'estimation des paramètres est souvent plate au voisinage de l'optimum. Le profil de cette fonction coût peut également présenter des optima locaux. Il convient alors de choisir des algorithmes d'optimisation adaptés.

C

ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLES

Les données fonctionnelles sur lesquelles l'analyse en composantes principales est effectuée se modélisent par p variables aléatoires pour lesquelles chaque réalisation est une courbe et non plus un nombre réel comme pour l'analyse en composantes principales discrète. Dans le cas présent, les variables aléatoires sont des fonctions continues. L'approche classique de l'analyse en composantes principales fonctionnelles formule une discrétisation du problème par l'intermédiaire de la décomposition dans une base de fonctions $\langle \phi_1, \dots, \phi_m \rangle$. Soient :

- $\mathbf{X} = (X_1, \dots, X_p)$ le vecteur aléatoire de p variables aléatoires centrées ;
- $\phi = (\phi_1, \dots, \phi_m)$ le vecteur colonne des m fonctions de base ;
- \mathcal{W}_ϕ la matrice carrée d'ordre m des produits scalaires des éléments de ϕ ;
- \mathbf{C} la matrice $p \times m$ des coefficients des décompositions des X_i dans ϕ .

L'objectif de la méthode consiste à trouver un ensemble de p fonctions notées f_j ($1 \leq j \leq p$) et appelées fonctions propres (notées FP), qui maximisent la somme des carrés des covariances entre les fonctions propres et les X_i :

$$\sum_{i=1}^p \langle f_j, X_i \rangle^2 = \sum_{i=1}^p \left(\int_{\Omega} f_j(\omega) \cdot X_i(\omega) dP(\omega) \right)^2. \quad (\text{C.1})$$

Les FP forment une base orthonormale du sous-espace vectoriel engendré par les X_i :

$$\langle f_j, f_{j'} \rangle = \int_{\Omega} f_j(\omega) \cdot f_{j'}(\omega) dP(\omega) = \delta_{jj'}. \quad (\text{C.2})$$

En tant qu'éléments de l'espace des fonctions, les fonctions propres se décomposent également en somme pondérée de fonctions de base. Soit $u' = (u'_1, \dots, u'_m)^t$ les coefficients de la décomposition d'une fonction propre f_j dans ϕ :

$$\forall \omega \in \Omega, f_j(\omega) = \sum_{k=1}^m u'_k \phi_k(\omega) = \phi(\omega) u', \quad (\text{C.3})$$

où $\phi(\omega)$ dénote le vecteur des images de ω par les fonctions de base. Ces fonctions sont définies telles que :

$$\phi(\omega) = (\phi_1(\omega), \dots, \phi_m(\omega)). \quad (\text{C.4})$$

La contrainte de normalisation sur les fonctions propres impose :

$$\int_{\Omega} |f_j(\omega)|^2 dP(\omega) = \int_{\Omega} (\phi(\omega) u')^t (\phi(\omega) u') dP(\omega) = \int_{\Omega} u'^t \phi^t(\omega) \phi(\omega) u' dP(\omega) = u'^t \mathcal{W}_\phi u' = 1. \quad (\text{C.5})$$

On considère la fonction définie sur $\Omega \times \Omega$ qui à (ω, ω') associe le produit scalaire usuel entre les individus associés :

$$w(\omega, \omega') = \mathbf{X}(\omega) \mathbf{X}^t(\omega') = \sum_{i=1}^p X_i(\omega) X_i(\omega'). \quad (\text{C.6})$$

Soit l'opérateur associé :

$$W(f) = \int_{\Omega} w(\cdot, \omega') f(\omega') dP(\omega'). \quad (\text{C.7})$$

Par définition des fonctions f_j , il est possible d'écrire :

$$W(f_j)(\omega) = \int_{\Omega} w(\omega, \omega') f_j(\omega') dP(\omega') = \lambda f_j(\omega), \quad (\text{C.8})$$

où λ est la valeur propre associée à f_j . Si chaque X_i s'exprime comme une somme de m fonctions de base ϕ_k :

$$\mathbf{X} = \phi \mathbf{C}. \quad (\text{C.9})$$

$W(f_j)$ peut donc s'écrire de la façon suivante :

$$W(f_j)(\omega) = \int_{\Omega} w(\omega, \omega') f_j(\omega') dP(\omega') \quad (\text{C.10})$$

$$= \int_{\Omega} X^t(\omega) X(\omega') f_j(\omega') dP(\omega') \quad (\text{C.11})$$

$$= \int_{\Omega} (\phi(\omega) \mathbf{C})^t (\phi(\omega') \mathbf{C}) (\phi(\omega') u') dP(\omega') \quad (\text{C.12})$$

$$= \phi(\omega) \mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi} u' = \lambda f_j(\omega) = \lambda \phi(\omega) u'. \quad (\text{C.13})$$

Ce résultat est vrai pour ω et la formule (C.8) est donc équivalente à $\phi \mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi} u' = \lambda \phi u'$. ϕ étant de rang maximal, l'équation peut être simplifiée par :

$$\mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi} u' = \lambda u'. \quad (\text{C.14})$$

Le principe de l'ACP fonctionnelles se résume donc à trouver les coefficients de la décomposition des fonctions propres dans ϕ , c'est-à-dire à diagonaliser la matrice $\mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi}$. D'après l'équation (C.5), en posant $u = \mathcal{W}_{\phi}^{1/2} u'$, alors $u^t u = 1$ et l'équation (C.14) devient :

$$\mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi} u' = \mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi}^{1/2} u = \lambda u' = \lambda \mathcal{W}_{\phi}^{-1/2} u, \quad (\text{C.15})$$

ce qui est équivalent à $\mathcal{W}_{\phi}^{1/2} \mathbf{C}^t \mathbf{C} \mathcal{W}_{\phi}^{1/2} u = \lambda u$.

On calcule la matrice de projection \mathbf{F} de taille $p \times p$ des X_i sur les fonctions propres. Si $\mathbf{B} = \mathcal{W}_{\phi}^{-1/2} \mathbf{U}$ est la matrice $m \times m$ des projections des fonctions propres dans la base ϕ :

$$\mathbf{F}_{ij} = \int_{\Omega} X_i(\omega) f_j(\omega) d\omega = [(\phi \mathbf{C})^t (\phi \mathbf{B})]_{ij} = (\mathbf{C} \mathcal{W}_{\phi} \mathbf{B})_{ij} \quad (\text{C.16})$$

$$= (\mathbf{C} \mathcal{W}_{\phi}^{1/2} \mathcal{W}_{\phi}^{1/2} \mathbf{B})_{ij} = (\mathbf{C} \mathcal{W}_{\phi}^{1/2} \mathbf{U})_{ij}. \quad (\text{C.17})$$

\mathbf{F}_{ij} est la valeur de la projection de X_i sur la j ème fonction propre.

D

IDENTIFIABILITÉ

1 GÉNÉRALITÉS

Dans la littérature, la plupart des études d'identifiabilité peuvent être scindées en deux thématiques selon le type de modèle *i.e.* le cadre stochastique pour les modèles probabilistes et un cadre adapté aux modèles dynamiques. Etant donné le domaine d'application de notre étude, seul l'aspect stochastique sera abordé dans cette section. Le cadre stochastique concerne l'identifiabilité des paramètres de la fonction de distribution de probabilité d'un modèle. On considère que x est un vecteur aléatoire de \mathbb{R}^n dont la fonction de distribution est contrôlée par un jeu de paramètres $\theta = (\theta_1, \theta_2, \dots, \theta_p)$, où p est le nombre total de paramètres et où $\theta \in \mathcal{S}$, l'espace des paramètres.

Définition 1.0.1. *On dit qu'un modèle $H(\theta)$ d'une structure H (resp. qu'une fonction $f(\theta)$ des paramètres d'un modèle) est localement identifiable s'il existe un voisinage ouvert \mathcal{O} de θ , dans le domaine des paramètres admissibles D , tel que*

$$\begin{aligned} & \forall \theta' \in \mathcal{O} \quad \mathcal{C}(\theta) = \mathcal{C}(\theta') \Rightarrow \theta = \theta' \\ (\text{resp.}) \quad & \forall \theta' \in \mathcal{O} \quad \mathcal{C}(\theta) = \mathcal{C}(\theta') \Rightarrow f(\theta) = f(\theta'), \end{aligned}$$

où $\mathcal{C}(\theta)$ désigne le comportement entrée-sortie de $H(\theta)$.

Définition 1.0.2. *Avec les mêmes notations que dans la définition précédente, on dit qu'un modèle $H(\theta)$ d'une structure H (resp. qu'une fonction $f(\theta)$ des paramètres d'un modèle) est globalement identifiable si*

$$\begin{aligned} & \forall \theta' \in D \quad \mathcal{C}(\theta) = \mathcal{C}(\theta') \Rightarrow \theta = \theta' \\ (\text{resp.}) \quad & \forall \theta' \in D \quad \mathcal{C}(\theta) = \mathcal{C}(\theta') \Rightarrow f(\theta) = f(\theta'). \end{aligned}$$

L'identifiabilité structurelle est étroitement liée à l'unicité des paramètres déterminés à partir du comportement entrée-sortie du modèle.

Définition 1.0.3. *On dit qu'une propriété définie pour les modèles d'une structure H est une propriété structurelle de H si elle est vérifiée par presque tous les modèles de la structure, au sens de la mesure uniforme sur \mathbb{R}^p .*

La propriété d'identifiabilité structurelle est une condition nécessaire non suffisante pour procéder à l'estimation des paramètres. Un modèle structurellement identifiable est en fait seulement potentiellement identifiable dans le sens où le succès de l'estimation n'est pas garanti [126]. En effet, les tentatives pour identifier des modèles structurellement identifiables peuvent échouer pour différentes raisons : données manquantes, bruit, estimateurs imprécis, etc. Cependant ces facteurs ne devraient

en aucun cas nuire aux propriétés d'identifiabilité structurelle du système. La principale raison de non-identifiabilité est inhérente au modèle lui-même : même avec de solides plans d'expériences, une grande quantité de données et des estimateurs précis, il est parfois impossible d'obtenir des estimations raisonnables dans le cas de modèles structurellement non-identifiables. Un outil important permettant de tester l'identifiabilité d'un système est la Matrice d'Information de Fisher (*MIF*) définie par :

$$MIF : \mathcal{F} = \left(E \left[\frac{\partial \log f}{\partial \theta_i} \cdot \frac{\partial \log f}{\partial \theta_j} \right] \right)_{p \times p} . \quad (D.1)$$

Rothenberg [102] a étudié l'identifiabilité de modèles paramétriques basée sur l'examen du rang des *MIF*. Sous certaines conditions, la propriété d'identifiabilité locale est équivalente à la non-singularité de la Matrice d'Information de Fisher. Des critères d'identifiabilité globale ont également été définis pour des modèles exponentiels de distribution de probabilité. Hochwald et Nehorai [65] ont étudié le lien entre identifiabilité et régularité de la *MIF* pour des modèles de distribution de probabilité gaussienne. L'identifiabilité structurelle est une propriété intrinsèque du modèle. Les causes les plus évidentes de non-identifiabilité structurelle sont la *surdétermination du problème* ou la *redondance de paramètres*.

Dans notre étude, la fonction que l'on cherche à identifier est une fonction de log-vraisemblance. Cette fonction est plus ou moins complexe selon la problématique considérée (section 3.3). Le choix des observables, l'ajout de paramètres de correction du modèle (biais par exemple), la transformation du modèle, l'estimation de variances ou de coefficients d'inter-corrélation peuvent rendre le problème plus complexe et mener à une non-identifiabilité du système. De plus le système considéré présente généralement une surdétermination.

La mise sous forme matricielle d'un problème surdéterminé linéaire donne : $Hx = Y$, où $H \in \mathcal{M}_{n,p}$, $Y \in \mathfrak{R}^n$ et $x \in \mathfrak{R}^p$. L'ensemble des matrices de \mathfrak{R} de dimensions n, p est noté $\mathcal{M}_{n,p}$ et forme un espace vectoriel sur \mathfrak{R} . Une condition d'existence d'une solution x est que Y appartienne à $\mathfrak{S} H$ (*i.e.* un sous ensemble de l'espace vectoriel de H). Cependant cette condition n'est pas toujours vérifiée et dans ce cas, l'objectif est de trouver une solution x qui vérifie au mieux le système $Hx = Y$. Ainsi le vecteur ε appelé résidu est introduit par la relation $\varepsilon = Hx - Y$. Il s'agit donc de trouver une solution \hat{x} qui minimise le résidu. Comme $x \in \mathfrak{R}^p$, la notion de « plus petit possible » doit être associée à une norme $\|\cdot\|_n$. Dans le cas qui nous intéresse, la norme utilisée est la norme euclidienne $\|x\| = \sqrt{\sum_i x_i^2}$ et par conséquent on définit le problème des moindres carrés comme suit :

Définition 1.0.4. Soit $H \in \mathcal{M}_{n,p}$ et $Y \in \mathfrak{R}^n$ données avec $n \geq p$. On appelle problème des moindres carrés le problème :

$$\min \|Hx - Y\|^2. \quad (D.2)$$

La norme du résidu peut s'écrire selon l'équation (D.3).

$$\|u\|^2 = \|Hx - Y\|^2 = (Hx - Y)^t (Hx - Y) = x^t H^t Hx - 2x^t H^t Y + Y^t Y \quad (D.3)$$

Le problème des moindres carrés peut donc se formuler en $\min_{x \in \mathfrak{R}^p} (x^t H^t Hx - 2(H^t Y)^t x)$.

Définition 1.0.5. On appelle fonction quadratique une fonction $f : \mathfrak{R}^p \rightarrow \mathfrak{R}$ de la forme :

$$f(x) = (x^t H^t Hx - 2(H^t Y)^t x), \quad (D.4)$$

où $H^t H$ est une matrice de $\mathcal{M}_{p,p}$ symétrique et Y un vecteur donné de \mathfrak{R}^p . La résolution d'un problème de moindres carrés nécessite donc la minimisation d'une fonction quadratique. La théorie

de la minimisation montre que la condition nécessaire pour que \hat{x} soit le vecteur minimisant f i.e. $f(\hat{x}) \leq f(x) \forall x \in \mathbb{R}^p$ est que $\nabla_x f(\hat{x}) = 0$ avec $\nabla_x f(\hat{x}) = 2H^t(Hx - Y)$.

Théorème 1.0.1. Une condition nécessaire pour qu'un vecteur x soit solution du problème de minimisation des moindres carrés est que : $H^t Hx = H^t Y$. On appelle cette équation l'équation normale du problème des moindres carrés.

Théorème 1.0.2. Si H une matrice de $\mathcal{M}_{n,p}$ est une matrice telle que $\text{Rg } H = p$ alors le problème des moindres carrés admet une solution unique \hat{x} vérifiant : $H^t H\hat{x} = H^t Y$.

Finalement, si $\text{Rg } H = p$ i.e. H est de rang plein (une matrice de rang plein est une matrice dont les colonnes sont indépendantes), alors $H^t H$ est définie positive donc inversible. Le système $H^t Hx = H^t Y$ admet donc une unique solution \hat{x} qui est un minimum de $x^t H^t Hx - (H^t Y)x$. La résolution d'un problème de moindres carrés peut alors être effectuée à partir de l'équation normale. Différents cas sont développés en fonction de la forme de la fonction de log-vraisemblance \mathcal{L} .

2 QUELQUES EXEMPLES

2.1 CAS LINÉAIRE

L'observable Y étudiée s'écrit sous la forme $Y = H(x) + \varepsilon$ où $H(x)$ est le modèle et ε le résidu. Le modèle $H(x)$ étant considéré comme linéaire, le problème peut alors se réécrire selon l'équation $\mathbf{Y} = \mathbf{H}\mathbf{x} + \varepsilon$, où $\mathbf{H} = \frac{\partial H}{\partial x}$, $\mathbf{H} \in \mathcal{M}_{n,p}$ et $\mathbf{Y} \in \mathbb{R}^n$ données avec $n \geq p$ (n correspond au nombre d'observations et p au nombre de paramètres à estimer). Afin qu'il y ait unicité de la solution, la matrice \mathbf{H} doit être de rang égal à p i.e. la matrice $\mathbf{H}^t \mathbf{H}$ doit être définie positive. Détaillons alors plus spécifiquement ce système à notre cas d'étude en considérant un jeu de données expérimentales (p observations), k observables et un nombre de paramètres à estimer $n = 2$. Le système est défini selon l'équation (D.5) :

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_p \end{pmatrix} = \mathbf{H}\mathbf{x} + \varepsilon. \quad (\text{D.5})$$

La matrice \mathbf{H} s'écrit sous la forme :

$$\mathbf{H} = \begin{bmatrix} \frac{\partial H_1^1}{\partial x_1} & \frac{\partial H_1^1}{\partial x_2} \\ \vdots & \vdots \\ \frac{\partial H_1^k}{\partial x_1} & \frac{\partial H_1^k}{\partial x_2} \\ \vdots & \vdots \\ \frac{\partial H_p^1}{\partial x_1} & \frac{\partial H_p^1}{\partial x_2} \\ \vdots & \vdots \\ \frac{\partial H_p^k}{\partial x_1} & \frac{\partial H_p^k}{\partial x_2} \end{bmatrix}. \quad (\text{D.6})$$

Afin qu'il y ait existence et unicité de la solution, la matrice $\mathbf{H}^t \mathbf{H}$ doit être définie positive.

2.2 PRISE EN COMPTE DE BIAIS

Considérons maintenant le même système auquel on ajoute un paramètre à estimer ; ce paramètre est un biais noté μ associé à l'observable Y^k . Le système peut alors s'écrire selon l'équation (D.7) :

$$\mathbf{Y} = \mathbf{H} \begin{pmatrix} x_1 \\ x_2 \\ \mu \end{pmatrix} + \varepsilon. \quad (\text{D.7})$$

L'observable Y^k à laquelle s'applique le biais s'écrit sous la forme suivante : $Y_j^k = H_j^k \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mu + \varepsilon_j^k$.

Si on applique un développement de Taylor, on obtient l'expression (D.8) :

$$Y_j^k = \left(\frac{\partial H_j^k}{\partial x_1}, \frac{\partial H_j^k}{\partial x_2}, 1 \right) \cdot \begin{pmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \\ \mu \end{pmatrix} + H_j^k(x_1^0, x_2^0) + \varepsilon_j^k. \quad (\text{D.8})$$

La matrice à laquelle s'applique la condition d'unicité est alors modifiée par rapport à la précédente et s'écrit selon le système matriciel \mathbf{H} :

$$\mathbf{H} = \begin{bmatrix} \frac{\partial H_1^1}{\partial x_1} & \frac{\partial H_1^1}{\partial x_2} & 0 \\ \vdots & \vdots & \\ \frac{\partial H_1^k}{\partial x_1} & \frac{\partial H_1^k}{\partial x_2} & 1 \\ \vdots & \vdots & \\ \frac{\partial H_p^1}{\partial x_1} & \frac{\partial H_p^1}{\partial x_2} & 0 \\ \vdots & \vdots & \\ \frac{\partial H_p^k}{\partial x_1} & \frac{\partial H_p^k}{\partial x_2} & 1 \end{bmatrix}. \quad (\text{D.9})$$

La matrice \mathbf{H} doit être définie positive pour que le problème ait une solution unique : chaque colonne doit alors être linéairement indépendante.

2.3 PRISE EN COMPTE DES VARIANCES DE RÉSIDUS

2.3.1 Cas homoscedastique où un seul paramètre physique doit être estimé

Si l'on considère un système simplifié du précédent (bidimensionnel) mais que l'on y ajoute des paramètres à estimer (les variances associées aux résidus), les observables suivent alors la relation :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \cdot x + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \quad (\text{D.10})$$

La variance des résidus intervient dans la définition de la fonction de log-vraisemblance. Dans un cas homoscedastique (variance des résidus constante sur les observables), la fonction de log-vraisemblance s'écrit selon l'équation :

$$\mathcal{L}(x, \sigma^2) = -\frac{p}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{j=1}^p \frac{1}{\sigma^2} (y_j - h_j x)^2. \quad (\text{D.11})$$

Une condition au premier ordre pour qu'il y ait unicité de la solution est que la dérivée partielle de la fonction de log-vraisemblance en x soit nulle, ce qui implique la relation $x = \frac{h_1 y_1 + h_2 y_2}{h_1^2 + h_2^2}$. Si on généralise cette expression à p dimensions, on obtient alors la solution $\mathbf{x} = (\mathbf{H}^t \mathbf{H})^{-1} \mathbf{H} \mathbf{Y}$. Le fait que le terme $\mathbf{H}^t \mathbf{H}$ soit inversible signifie que la matrice \mathbf{H} doit être de rang plein. De même si l'on cherche à estimer la (ou les) valeur(s) de variances de résidus, la dérivée partielle de la fonction de log-vraisemblance en

σ^2 doit être nulle. Dans ce cas, $\sigma^2 = \frac{1}{2} \sum_{j=1}^2 (y_j - h_j x)^2$. Une condition pour qu'il y ait unicité de la solution est que la matrice (\mathbf{Y}, \mathbf{H}) soit de rang plein.

2.3.2 Cas homoscédastique où deux paramètres doivent être estimés

Dans le cas où deux paramètres sont à estimer, le système est défini par l'équation :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \quad (\text{D.12})$$

Si \mathbf{H} est une matrice de rang plein, les paramètres \mathbf{x} sont alors identifiables. L'estimation de la variance peut cependant s'avérer impossible si cette condition n'est pas vérifiée. Dans ce cas, afin de rendre les colonnes de \mathbf{H} indépendantes, il pourrait être nécessaire que le nombre d'observations soit supérieur au nombre de paramètres.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{31} & h_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}. \quad (\text{D.13})$$

Cependant, le fait que la matrice \mathbf{H} soit de rang plein *i.e.* que la matrice $\mathbf{H}^t \mathbf{H}$ soit inversible, ne permet pas nécessairement d'identifier les variances des résidus ; il est en effet possible que ces variances soient nulles. Les observations doivent donc être bruitées pour accéder à cette condition. Afin d'estimer les variances des résidus, la matrice (\mathbf{Y}, \mathbf{H}) doit être de rang plein. C'est une condition suffisante pour que le résidu soit non nul. De plus, si la matrice (\mathbf{Y}, \mathbf{H}) n'est pas de rang plein, il est alors possible de trouver des solutions \tilde{x}_1 et \tilde{x}_2 permettant d'annuler le résidu. La condition (\mathbf{Y}, \mathbf{H}) de rang plein est donc bien une condition nécessaire à l'identifiabilité des paramètres \mathbf{x} et des variances.

2.3.3 Cas hétéroscédastique où deux paramètres doivent être estimés

Si on considère désormais un cas hétéroscédastique à deux observables et donc deux variances et un nombre d'observations p , le système s'écrit selon la relation :

$$\begin{pmatrix} y_1^1 & y_2^1 \\ \vdots & \vdots \\ y_1^p & y_2^p \end{pmatrix} = \mathbf{H} \mathbf{x} + \begin{pmatrix} \sigma_1^2 \varepsilon_1^1 & \sigma_2^2 \varepsilon_1^1 \\ \vdots & \vdots \\ \sigma_1^2 \varepsilon_1^p & \sigma_2^2 \varepsilon_1^p \end{pmatrix} \quad (\text{D.14})$$

Si on considère que les observables sont indépendantes, la fonction de log-vraisemblance s'écrit alors sous la forme :

$$\mathcal{L}(\mathbf{x}, \sigma_1, \sigma_2) = -\frac{n}{2} \left[\log(2\pi\sigma_1^2) + \log(2\pi\sigma_2^2) \right] - \frac{1}{2} \sum_{j=1}^p \left[\frac{1}{\sigma_1^2} (y_1^j - h_1^j x)^2 + \frac{1}{\sigma_2^2} (y_2^j - h_2^j x)^2 \right] \quad (\text{D.15})$$

Une condition d'identifiabilité du problème est que la matrice $(\mathbf{H}^t \mathbf{H})^{-1}$ soit de rang plein. Les matrices $(\mathbf{Y}_1, \mathbf{H}_1)$ et $(\mathbf{Y}_2, \mathbf{H}_2)$ doivent également être de rang plein (CNS) et ce pour les mêmes raisons que dans le cas précédent. Une autre condition du second ordre est que la Hessienne de la fonction de log-vraisemblance soit définie positive. Si c'est le cas, le problème est convexe et il y a existence et unicité de la solution.

2.4 PRISE EN COMPTE DE DEUX PARAMÈTRES PHYSIQUES À ESTIMER, DE BIAIS ET DE VARIANCES DE RÉSIDUS

Posons alors le problème dans le cas où l'on considère 2 observables, 2 valeurs de variances (cas hétéroscédastique), la présence d'un biais (associé par exemple à l'observable y_1) et un nombre d'observations p . Le système est défini comme suit :

$$\begin{pmatrix} y_1^1 & y_2^1 \\ \vdots & \vdots \\ y_1^p & y_2^p \end{pmatrix} = \mathbf{H}\mathbf{x} + \boldsymbol{\mu} + \begin{pmatrix} \sigma_1^2 \varepsilon_1^1 & \sigma_2^2 \varepsilon_2^1 \\ \vdots & \vdots \\ \sigma_1^2 \varepsilon_1^p & \sigma_2^2 \varepsilon_2^p \end{pmatrix} \quad (\text{D.16})$$

Si on considère que les observables sont indépendantes, la fonction de log-vraisemblance s'écrit selon l'expression :

$$\mathcal{L}(\mathbf{x}, \sigma_1, \sigma_2, \boldsymbol{\mu}) = -\frac{p}{2} \left[\log(2\pi\sigma_1^2) + \log(2\pi\sigma_2^2) \right] - \frac{1}{2} \sum_{j=1}^p \left[\frac{1}{\sigma_1^2} (y_1^j - h_1^j x - \mu)^2 + \frac{1}{\sigma_2^2} (y_2^j - h_2^j x)^2 \right] \quad (\text{D.17})$$

La Hessienne de cette fonction doit être définie positive pour qu'il y ait existence et unicité de la solution. Pour être définie positive, la matrice doit vérifier certaines propriétés :

- les valeurs propres de la matrice sont strictement positives ;
- $\mathbf{x}^t \mathbf{H} \mathbf{0} \mathbf{x} > 0$.

Cependant, le problème peut être traité de façon différente en intégrant le biais au modèle \mathbf{H} . Le modèle considéré deviendra le modèle $\tilde{\mathbf{H}}$ où $\tilde{\mathbf{H}} = \mathbf{H}(\boldsymbol{\mu})$. Le système peut ainsi se réécrire comme suit :

$$\begin{pmatrix} y_1^1 & y_2^1 \\ \vdots & \vdots \\ y_1^p & y_2^p \end{pmatrix} = \tilde{\mathbf{H}}\mathbf{x} + \begin{pmatrix} \sigma_1^2 \varepsilon_1^1 & \sigma_2^2 \varepsilon_2^1 \\ \vdots & \vdots \\ \sigma_1^2 \varepsilon_1^p & \sigma_2^2 \varepsilon_2^p \end{pmatrix} \quad (\text{D.18})$$

Dans ce cas, afin de pouvoir estimer les paramètres x , le biais et les variances, une condition nécessaire et suffisante est que les matrices $(\mathbf{Y}_1, \tilde{\mathbf{H}}_1)$ et $(\mathbf{Y}_2, \tilde{\mathbf{H}}_2)$ soient de rang plein.

2.5 PRISE EN COMPTE DE L'INTERCORRÉLATION ENTRE LES OBSERVABLES

On peut redéfinir le système en considérant un coefficient d'intercorrélacion γ inconnu :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \end{pmatrix} \mathbf{x} + \boldsymbol{\varepsilon}. \quad (\text{D.19})$$

Soit $\boldsymbol{\varepsilon} = \begin{pmatrix} y_1 - H_1 x \\ y_2 - H_2 x \end{pmatrix} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$, où $\boldsymbol{\Sigma}$ est la matrice de covariance des résidus $\boldsymbol{\varepsilon}$ et se définit

$$\text{comme } \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \gamma\sigma_1\sigma_2 \\ \gamma\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

La fonction de log-vraisemblance associée à ce système s'écrit :

$$\mathcal{L}(\mathbf{x}, \sigma_1, \sigma_2, \boldsymbol{\mu}) = -\frac{p}{2} \log(\det[2\pi\boldsymbol{\Sigma}]) - \frac{1}{2} \sum_{j=1}^p [\mathbf{y} - \mathbf{H}\mathbf{x}]^t \boldsymbol{\Sigma}^{-1} [\mathbf{y} - \mathbf{H}\mathbf{x}]. \quad (\text{D.20})$$

L'identifiabilité de systèmes à variables corrélées est une problématique complexe qui nécessiterait une recherche bibliographique approfondie. Cependant cette problématique dépasse le cadre de cette thèse et ne sera donc pas abordée dans ces travaux.

E

PROPAGATION D'INCERTITUDES LIÉES AUX PARAMÈTRES MICROMÉTÉOROLOGIQUES

Une méthode classique dans l'étude des incertitudes, la méthode de propagation d'incertitudes, a été mise en place afin de rendre compte de l'effet des erreurs de mesures des paramètres micrométéorologiques sur les niveaux sonores. Pour des raisons de cohérence, cette approche n'a pas été développée dans le corps du document. Seule la communication au Congrès Français d'Acoustique 2010 retracera ces travaux.

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

Propagation d'incertitudes pour l'estimation des paramètres micrométéorologiques influents en acoustique environnementale

Olivia Leroy¹, Benoit Gauvreau¹, Fabrice Junker², Michel Bérengier¹

¹ Laboratoire Central des Ponts et Chaussées, Route de Bouaye, F-44341 Bouguenais Cedex

² EDF R&D, 1 avenue du Général de Gaulle, F-92141 Clamart Cedex

Les phénomènes de propagation acoustique à grande distance présentent une forte variabilité et sont soumis à de nombreuses sources d'incertitudes. Variabilité et incertitudes induisent une dispersion importante des niveaux sonores. Cette dispersion est notamment associée aux phénomènes micrométéorologiques influents, et en particulier aux profils verticaux de célérité. Ces sources d'incertitudes et leurs effets sur les niveaux sonores sont difficilement quantifiables. Aux incertitudes météorologiques et à la variabilité "physique" des phénomènes vient s'ajouter la diversité des méthodes de mesures (mât équipé de capteurs lents, anémomètres ultrasoniques 3D) et l'influence de la période d'observation (de la minute à l'heure). L'ensemble de ces paramètres conditionne l'estimation des champs de vent et de température et par suite le calcul des paramètres d'intérêt (profils verticaux de vent et de température).

Une méthode statistique de propagation d'incertitudes est proposée. Elle permet d'étudier le comportement des incertitudes expérimentales en propageant cette information dans le système, et finalement de quantifier cet effet sur les paramètres d'intérêt puis sur les observables d'intérêt. Cet article présente les bases théoriques de la méthode de propagation d'incertitudes et une application de cette méthode au calcul des profils verticaux de célérité du son. Une campagne expérimentale menée en 2007 sur le site du LCPC Nantes nous fournit une base de données conséquente. Cette étude permet d'appréhender la sensibilité de ces diverses approches par la propagation et la quantification des incertitudes.

1 Introduction

Les normes et réglementations actuelles dans le domaine de la propagation acoustique en milieu extérieur ne rendent pas suffisamment compte des incertitudes associées aux niveaux sonores. Les incertitudes expérimentales liées aux paramètres influents sur la propagation et notamment les paramètres micrométéorologiques, sont liées aussi bien aux aspects purement météorologiques (disposition des capteurs, précision de la mesure, échantillonnage) qu'aux aspects méthodologiques (types de capteurs utilisés, méthodes de calcul choisies). Des études concernant la caractérisation des incertitudes et de la variabilité des phénomènes ont déjà été menées en propagation acoustique en milieu extérieur en considérant les incertitudes dans leur globalité [1] [2] [3]. On propose dans ce papier de ne considérer que les incertitudes expérimentales et de quantifier ces dernières par une méthode statistique communément utilisée, la méthode de propagation d'incertitudes. Cette méthode va permettre d'étudier l'effet des différentes sources d'incertitudes expérimentales à travers les systèmes physiques considérés. Après une présentation théorique de la méthode, une application à la propagation acoustique en milieu extérieur est présentée. L'application de la propagation d'incertitude à deux méthodes de mesures des paramètres micrométéorologiques, combinés à un modèle de propagation acoustique de référence, va permettre d'apprécier quantitativement la sensibilité des niveaux sonores aux

incertitudes expérimentales et météorologiques associées aux paramètres micrométéorologiques.

2 Propagation d'incertitudes

2.1 Choix de la méthode

La propagation d'incertitudes dans un système consiste à estimer la dispersion des sorties du modèle due aux incertitudes associées aux entrées du système (erreurs expérimentales, précision de mesures, *etc.*). Cette dispersion peut être caractérisée par différentes quantités d'intérêt telles que des intervalles de prévision, des moments voire même une fonction de probabilité liée aux paramètres (on parle alors de mesure "complète de l'incertitude"). Différentes méthodes de propagation d'incertitudes peuvent être utilisées selon la nature statistique des quantités d'intérêt (statistiques) recherchées. Certaines contraintes méthodologiques sont également liées aux caractéristiques des variables d'intérêt étudiées. Les principales méthodes de propagation d'incertitudes (approximation de Taylor [4], simulation de Monte-Carlo [5], *etc.*) dépendent d'éléments clés [6] :

- la formulation du critère de décision (probabilité de dépassement du paramètre estimé) ;
- les caractéristiques du modèle d'incertitude : variables discrètes ou continues, les dépendances entre les variables, le champ spatial, la dépendance temporelle ou la multidimensionnalité des va-

riables de sortie ;

- les caractéristiques du système et notamment les contraintes en temps de calcul.

Les systèmes étudiés ici correspondent à des méthodes analytiques permettant de calculer les profils et gradients verticaux de célérité du son à partir de mesures micrométéorologiques (cf section 3). Par suite, les fonctions de densité de probabilité des profils et gradients seront utilisées comme données d’entrée d’un modèle numérique de propagation acoustique (Équation Parabolique) [7]. Ce système présente donc un caractère multidimensionnel des variables de sortie (obtention finale de spectres en tiers d’octave). De plus, le système étudié n’étant pas une simple fonction mathématique, il est difficile de définir un modèle de propagation d’incertitudes de manière purement analytique (hypothèses de linéarité, vérification du théorème Central-Limite). Afin de passer outre ces contraintes et estimer des fonctions de densité de probabilité, la méthode de Monte-Carlo est la méthode la plus adaptée à notre problématique.

2.2 Méthode de Monte-Carlo

Par convention, le modèle mathématique sera défini par la relation $Y = f(X)$, où Y est une variable de sortie scalaire et X représente les N paramètres d’entrée (X_1, \dots, X_N) . Chaque valeur d’entrée X_i est considérée comme une variable aléatoire, Y est donc également une variable aléatoire. Les variables X_i peuvent être représentées par une fonction de densité de probabilité $p_{x_i}(x_i)$ (x_i sont les valeurs possibles prises par X_i). L’*espérance* ou la meilleure estimation de la valeur X_i est notée $E(X_i)$ et l’*incertitude* associée à cette valeur $u(x_i) = \sigma(X_i)$. Ces quantités sont calculées à partir de la fonction de densité de probabilité de X_i . A partir de ces données et connaissant leurs propriétés, notre problème est l’estimation de l’incertitude associée aux variables de sortie notée $u(Y)$. La méthode de Monte-Carlo permet de résoudre des problèmes numériquement en simulant M fois les paramètres d’entrée X (cf section 3.4).

Les valeurs x_{ik} des variables d’entrée aléatoires sont simulées à partir de valeurs aléatoires distribuées selon la fonction de densité de probabilité définie préalablement (cf section 3.3). Il est donc possible de créer des “échantillons réels” de Y pour chaque tirage k en appliquant l’équation suivante :

$$y_k = f(x_{1k}, x_{2k}, \dots, x_{nk}), \quad k = 1, \dots, M. \quad (1)$$

L’évaluation de l’incertitude d’une “mesure” indirecte par la méthode de Monte-Carlo peut être synthétisée suivant les étapes suivantes :

1. définir le système étudié ;
2. sélectionner les variables d’entrée X_i qui contribuent à l’incertitude de Y ;
3. identifier la fonction de densité de probabilité g_{X_i} correspondant aux variables d’entrée préalablement sélectionnées ;
4. définir le nombre de tirages de Monte-Carlo ;
5. simuler M séries de nombres aléatoires $(x_{i1}, x_{i2}, \dots, x_{ik}, \dots, x_{iM})$ pour chaque variable d’entrée sélectionnée x_i où $i = 1, \dots, N$;

6. calculer les M échantillons de Y à partir de l’équation $y_k = f(x_{1k}, x_{2k}, \dots, x_{nk})$ où $k = 1, \dots, M$.

L’incertitude $u(Y)$ est assimilée à l’écart-type de l’échantillon :

$$u(Y) = \sqrt{\frac{\sum_{k=1}^M y_k^2 - \frac{1}{M} \left(\sum_{k=1}^M y_k \right)^2}{M - 1}}. \quad (2)$$

2.3 Discussion

La méthode de Monte-Carlo est une méthode numérique très efficace pour la propagation des distributions. Cette méthode peut être implémentée pour évaluer l’incertitude de processus analytiques. Les principaux avantages d’une simulation de Monte-Carlo sont que le caractère non-linéaire d’une fonction n’est pas limitant, aucune hypothèse n’est faite sur la distribution des variables, le calcul du nombre de degrés de liberté effectif n’est pas nécessaire, il n’est pas non plus nécessaire de calculer les dérivées partielles des fonctions étudiées. La méthode présente cependant certaines limitations : les temps de calculs peuvent être longs dans certains cas complexes, la sélection de la fonction de densité de probabilité des variables d’entrée peut être difficile du fait de l’imprécision sur les données ou d’un manque de compréhension du processus, la précision de la méthode de Monte-Carlo dépend aussi de la qualité du générateur de nombres aléatoires.

3 Applications

3.1 Systèmes d’étude

Les systèmes étudiés sont des méthodes de mesure indirecte des profils verticaux de célérité du son via les profils verticaux de température et de vitesse du vent. Ces profils peuvent être mesurés grâce à des capteurs “lents” (ou “conventionnels”) placés sur un mât à trois hauteurs (système (a)) et des capteurs “rapides” (anémomètre ultrasonique tridimensionnel) (système (b)). Selon la technique de mesure utilisée, les méthodes de calculs des profils sont différentes.

3.1.1 Paramètres de réfraction

Système (a) : les paramètres mesurés par les capteurs “lents” sont la vitesse de vent horizontal $u(m.s^{-1})$, la direction du vent $\theta(^{\circ})$ et la température $T(^{\circ}C)$ à trois hauteurs (par exemple 1, 3, 10 m). Ces paramètres permettent alors de calculer la célérité du son aux différentes hauteurs : $c(z) = \sqrt{\gamma RT(z)} + u(z) \cos \theta$. La méthode de calcul des paramètres de réfraction consiste alors à ajuster par minimisation au sens des moindres carrés un profil vertical de célérité défini par l’équation (6) sur les trois valeurs de célérité.

Système (b) : les anémomètres ultrasoniques mesurent avec une certaine fréquence (typiquement 20 Hz) les composantes de vent dans le plan horizontal et le plan vertical (respectivement u , v et w ($m.s^{-1}$)) ainsi que la température sonique. La température sonique

est la température mesurée par l'anémomètre ultrasonique ; elle est généralement inférieure à la température de l'air (d'environ 1 °C) mais présente la même structure temporelle que les composantes de vent. A partir de ces mesures instantanées, complétées par des mesures de température de l'air, sont calculées les échelles caractéristiques de Monin-Obukhov U^* (la vitesse de frottement), T^* (l'échelle de température) et L_{MO} (la longueur de Monin-Obukhov *i.e.* l'altitude à partir de laquelle les forces de cisaillement deviennent négligeables devant les forces convectives). Ces échelles basées sur la théorie de Monin-Obukhov sont définies par les expressions suivantes [8] :

$$\left\{ \begin{array}{l} U^* = \left(\overline{u'w'^2} + \overline{v'w'^2} \right)^{1/4} \\ T^* = \frac{-\overline{w'\theta'_v}}{U^*} \\ L_{MO} = -\frac{\theta_v (\overline{u'w'} + \overline{v'w'})^{3/4}}{kg\overline{w'\theta'_v}} \end{array} \right. , \quad (3)$$

où k est la constante de von Karman ($k = 0,4$), θ_v la température virtuelle définie par $\theta_v \approx T - 0,01z_{mes}$ avec z_{mes} la hauteur de mesure. Ces expressions montrent la dépendance des échelles caractéristiques avec les covariances de grandeurs mesurées. Le paramètre de stabilité $\xi = \frac{z_{mes}}{L_{MO}}$ définit la stabilité de l'atmosphère. Des paramètres de réfraction a_{log} et b_{lin} peuvent être définis en fonction de ces conditions de stabilité [9].

En conditions neutres et stables ($0 \leq \xi < 1$) :

$$\left\{ \begin{array}{l} a_{log} = \frac{1}{2} \frac{\gamma R}{c_0} \frac{0,74T^*}{k} + \frac{U^*}{k} \cos \theta \\ b_{lin} = \frac{1}{2} \frac{\gamma R}{c_0} \left(\frac{T^*}{k} \frac{4,7}{L_{MO}} - \Gamma_d \right) + \frac{U^*}{k} \frac{4,7}{L_{MO}} \cos \theta. \end{array} \right. \quad (4)$$

En conditions instables ($-0,8 < \xi < 0$) :

$$\left\{ \begin{array}{l} a_{log} \approx \frac{1}{2} \frac{\gamma R}{c_0} \frac{0,74T^*}{k} + \frac{U^*}{k} \cos \theta \\ b_{lin} \approx \frac{1}{2} \frac{\gamma R}{c_0} \left(\frac{T^*}{k} \frac{0,74}{L_{MO}} - \Gamma_d \right) + \frac{U^*}{k} \frac{1,0}{L_{MO}} \cos \theta, \end{array} \right. \quad (5)$$

avec le gradient sec de température adiabatique Γ_d ($= 0,0098 \text{ K m}^{-1}$), la constante des gaz parfaits R ($= 287 \text{ J.kg}^{-1}.\text{K}^{-1}$), le rapport des capacités calorifiques massiques de l'air ($\gamma = C_p/C_v$), la direction de propagation $\theta(^{\circ})$.

3.1.2 Profils et gradients verticaux de célérité du son

Les paramètres de réfraction ainsi obtenus par les systèmes (a) et (b) permettent de calculer profils et gradients verticaux de célérité du son. Les profils verticaux de célérité peuvent être approchés par l'équation (6) :

$$c(z) = c_0 + a_{log} \log \left(1 + \frac{z}{z_0} \right) + b_{lin} z, \quad (6)$$

où a_{log} et b_{lin} sont les paramètres de réfraction, c_0 est la célérité effective de référence ($= 343 \text{ m.s}^{-1}$ à 20 °C),

$z(m)$ le paramètre de hauteur et $z_0(m)$ la longueur de rugosité (sa valeur dépend de la nature du sol et s'étend de 10^{-2} à plusieurs mètres). Les gradients de célérité sont ensuite calculés par dérivation de la formulation du profil vertical de célérité :

$$\frac{\partial c(z)}{\partial z} = \frac{a_{log}}{z} + b_{lin}. \quad (7)$$

Ces paramètres calculés (gradients et/ou paramètres de réfraction) sont alors appliqués comme paramètres d'entrée d'un modèle de propagation acoustique. Le modèle considéré dans cette étude est le modèle de propagation de référence basé sur la résolution de l'équation parabolique [7].

3.2 Sélection des paramètres d'entrée

Les données utilisées dans cette étude sont issues d'une campagne expérimentale de trois mois menée en 2007 sur le site du LCPC Nantes [10] [11]. Les mesures issues d'un anémomètre ultrasonique (échantillonnage 20 Hz, hauteur 3 m) et les mesures effectuées à l'aide d'un mât équipé de capteurs conventionnels (1, 3, 10 m) seront utilisées. Les incertitudes expérimentales des deux techniques métrologiques seront étudiées sur des intervalles de temps de 10 et 60 minutes.

- **Système (a)** : dans le cas de mesures par capteurs "lents", les paramètres expérimentaux incertains sont la vitesse horizontale de vent u_a (m.s^{-1}), la température de l'air T_a ($^{\circ}\text{C}$), ainsi que la direction du vent θ_a ($^{\circ}$) aux trois hauteurs du mât ;
- **Système (b)** : dans le cas de mesures par anémomètre ultrasonique, les paramètres expérimentaux entachés d'incertitude sont les trois composantes de la vitesse de vent (u_b, v_b, w_b) (m.s^{-1}), la température sonique T_{sb} ainsi que la température de l'air T_b ($^{\circ}\text{C}$).

Dans le cadre de la propagation d'incertitudes, ces paramètres constituent les paramètres d'entrée et sont considérés comme aléatoires. Dans une approche probabiliste, ces paramètres sont notés X .

3.3 Identification de la fonction de densité de probabilité

Dans l'implémentation d'une méthode de Monte-Carlo, M vecteurs sont construits à partir des fonctions de densité de probabilité $g_{X_i}(\xi_i)$ pour les N quantités d'entrée X_i . Les incertitudes expérimentales sur les paramètres d'entrée servent alors d'incertitude standard $u(x)$ associée à l'estimation x . Estimation et incertitude standard étant les seules informations disponibles en regard d'une quantité X , d'après le principe du maximum d'entropie, une distribution de probabilité gaussienne $\mathcal{N}(x, u^2(x))$ peut être associée à X . La fonction de densité de probabilité de X s'écrit :

$$g_X(\xi) = \frac{1}{\sqrt{2\pi}u(x)} \exp \left(-\frac{(\xi - x)^2}{2u^2(x)} \right). \quad (8)$$

L'espérance et la variance de X s'écrivent alors :

$$E(X) = x \quad V(X) = u^2(x). \quad (9)$$

Pour effectuer un échantillonnage à partir d'une distribution normale $\mathcal{N}(x, u^2(x))$, on procède à un tirage aléatoire κ à partir de la distribution normale $\mathcal{N}(0, 1)$ et on génère ainsi les données $\xi = x + u(x)\kappa$.

3.4 Nombre de tirages de Monte-Carlo

Il est nécessaire de déterminer le nombre de tirages de Monte-Carlo M , *i.e.* le nombre d'évaluations du modèle. Etant donné le grand nombre de données expérimentales à disposition (trois mois de mesures), le nombre de calculs à effectuer ne facilite pas l'application d'une valeur M élevée pour obtenir une connaissance suffisante des distributions des quantités de sortie. Dans ce cas, une alternative est de considérer que la fonction de densité de probabilité des variables de sortie suit une loi normale et procéder comme suit : pour une faible valeur M (50 ou 100), la moyenne et l'écart-type des M valeurs résultantes du modèle sont considérés respectivement y et $u(y)$. Une fonction de densité de probabilité gaussienne $g_Y = \mathcal{N}(y, u^2(y))$ est alors attribuée pour caractériser Y .

4 Résultats et discussion

4.1 Effets sur les paramètres micrométéorologiques

L'étude est menée sur l'intégralité de la campagne expérimentale en sélectionnant cependant des conditions de propagation caractéristiques. Les gradients verticaux de célérité ont initialement été calculés à une hauteur de trois mètres par les deux méthodes sur des intervalles de temps de 10 et 60 minutes sur l'ensemble de la campagne. La figure 1 montre une corrélation forte entre les gradients, malgré la présence d'écarts entre les méthodes, essentiellement en conditions favorables et défavorables (point d'inflexion en conditions homogènes). Ces écarts peuvent être dus à la hauteur de calcul des gradients. En effet, à trois mètres, la tendance logarithmique des profils est forte pour les mesures par capteurs rapides et n'est pas aussi marquée avec les mesures par capteurs lents. Une hauteur de mesures plus élevée (où les profils ont une tendance linéaire) pourrait pallier ces écarts. De plus, la dispersion des données peut s'expliquer par les différences entre les techniques de mesure. Ces remarques sont cependant propres aux conditions micrométéorologiques considérées ; les gradients verticaux de célérité peuvent en effet être scindés en deux populations bien distinctes selon leur signe.

Dans cet article, seuls les résultats obtenus dans des conditions nettement favorables à la propagation (valeur du gradient vertical de célérité déterministe $> 0,1 \text{ s}^{-1}$) seront présentés. Un échantillon caractéristique et représentatif de ces conditions est sélectionné dans la base de données sur des intervalles de 10 et 60 minutes.

4.1.1 Intervalle de 10 minutes

Les profils verticaux de célérité estimés après propagation d'incertitudes sur un intervalle de temps de 10 minutes présentent des disparités entre les mesures par

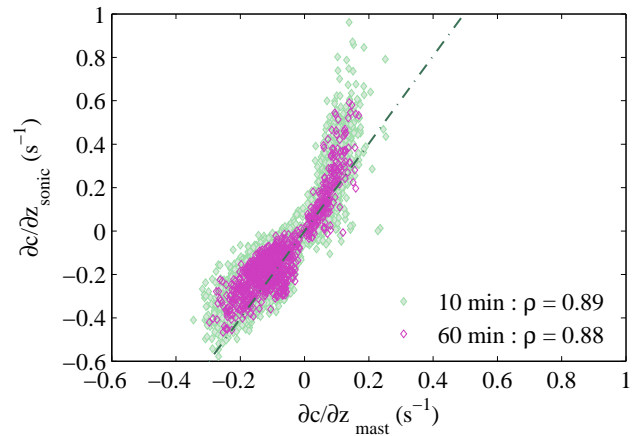


FIGURE 1 – Corrélation entre les gradients de célérité calculés par capteurs lents et capteurs rapides sur des intervalles de 10 et 60 min.

capteurs lents ou rapides. La figure 2 montre une certaine cohérence, une faible dispersion et une tendance logarithmique bien marquée des profils mesurés à l'aide du système (b). Les profils mesurés par le système (a) sont davantage dispersés et ne présentent pas toujours une tendance logarithmique marquée. Dans les deux cas, la dispersion des profils augmente avec la hauteur z . Cette disparité sur les profils verticaux de célérité a un effet sur les valeurs de gradients verticaux de célérité à trois mètres. Les histogrammes des gradients (Figure 2) rendent compte de l'étendue des données avec des gradients verticaux moyens proches, aux écarts près observés figure 1 (0,15 pour le système (a) et 0,12 pour le système (b)). Les écarts-types sur les gradients mesurés par les systèmes (a) et (b) sont respectivement de $0,06 \text{ s}^{-1}$ et $0,01 \text{ s}^{-1}$. Ces résultats témoignent de la sensibilité des deux approches et invitent à utiliser le système (b). Le système (a) est en effet soumis à des erreurs météorologiques majeures dues à la multiplication des capteurs, capteurs qui eux-mêmes présentent une imprécision plus grande qu'un anémomètre ultrasonique. De plus, la méthode d'estimation des paramètres de réfraction par ajustement des profils sur les points de mesure est soumise à des remises en question (manque d'informations en $z = z_0$, les éventuelles inversions peuvent avoir un sens physique ou peuvent être dues à des erreurs de mesures, la valeur z_0 est généralement une constante et pourrait être une variable ajustable, *etc.*).

4.1.2 Intervalle de 60 minutes

Le passage à un intervalle de temps de 60 minutes montre une augmentation de la dispersion globale des profils verticaux de célérité (Figure 3), menant à des valeurs de gradients verticaux de célérité négatives pour le système (a). L'augmentation de la dispersion des profils verticaux de célérité pour le système (b) est plus nette. Comme dans le cas d'un intervalle de 10 minutes, la dispersion augmente avec l'altitude mais peut atteindre 6 m.s^{-1} à 20 m (contre 1 m.s^{-1} sur 10 min). Les écarts-types sur les gradients mesurés par le système (a) et par le système (b) sont respectivement de $0,12 \text{ s}^{-1}$ et $0,08 \text{ s}^{-1}$.

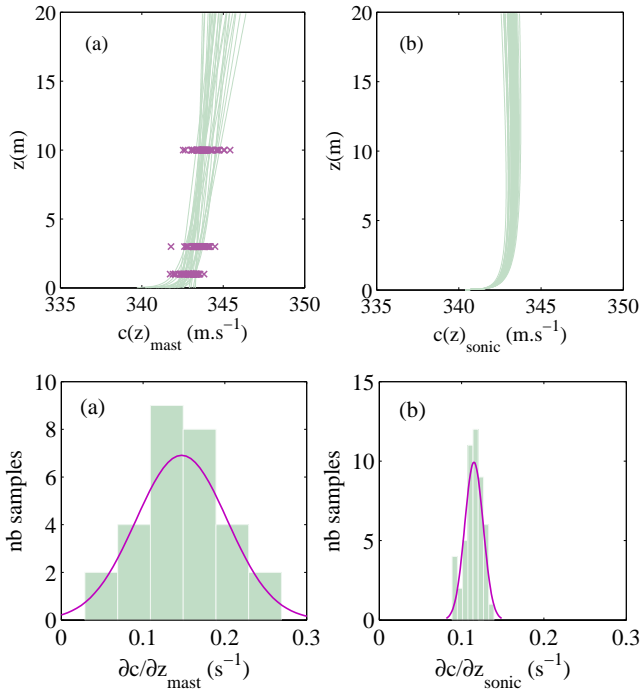


FIGURE 2 – Profils verticaux de célérité et histogrammes des gradients verticaux de célérité par capteurs lents (a) et capteurs rapides (b) sur un intervalle de 10 minutes.

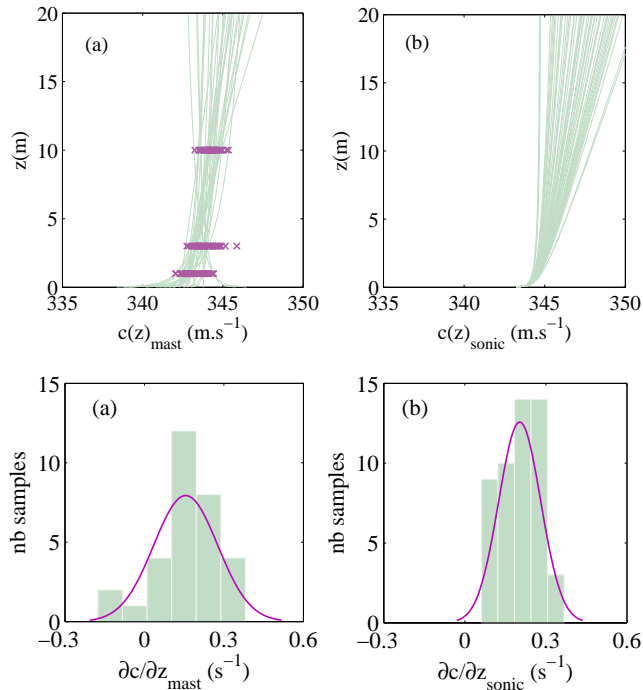


FIGURE 3 – Profils verticaux de célérité et histogrammes des gradients verticaux de célérité par capteurs lents (a) et capteurs rapides (b) sur un intervalle de 60 minutes.

Ces écarts-type rendent à nouveau compte de la disparité entre les méthodologies.

4.1.3 Bilan

Pour ces échantillons spécifiques, le passage d'un intervalle de 10 à 60 minutes a multiplié l'incertitude

sur les gradients verticaux de célérité par deux pour le système (a) et par huit pour le système (b). Cette étude rend compte de la variabilité temporelle des phénomènes et permet de quantifier cet effet sur l'estimation des paramètres micrométéorologiques. Cette étude donne des pistes intéressantes sur les approches métrologiques à envisager pour estimer des paramètres de réfraction en minimisant les erreurs sur ces estimations. L'approche par capteur rapide permet de minimiser l'étendue des valeurs des paramètres de réfraction et par suite de gradients verticaux de célérité acoustique. Ces résultats peuvent être liés à l'utilisation de la formule approchée pour a_{log} et b_{lin} (Equations (4) et (5)) qui contraint le domaine de variation des profils et réduit de ce fait la dispersion. Le choix de l'intervalle de temps a une incidence non-négligeable sur l'estimation des paramètres de réfraction et des profils verticaux de célérité.

4.2 Effet sur les niveaux sonores

Ces erreurs métrologiques et méthodologiques sur les paramètres micrométéorologiques ont des effets sur la prévision de la propagation acoustique et par conséquent des niveaux sonores. Les distributions des paramètres de réfraction sont appliquées en entrée du modèle d'équation parabolique afin d'appréhender l'effet des erreurs sur les spectres (en relatif à 10 m de la source). La configuration choisie est la suivante : source et récepteur à une hauteur de 2 m, site plan, sol herbeux (résistance spécifique au passage de l'air $\sigma = 150 \text{ kN.s.m}^{-4}$ et épaisseur $e = 0,04 \text{ m}$), distance de propagation jusqu'à 500 m, température fixée à $20 \text{ }^\circ\text{C}$, calcul en tiers d'octave puis calcul du niveau global (en dB), non prise en compte de la turbulence (le gradient de célérité étant positif, l'influence de la turbulence est moindre). La moyenne et l'étendue (écart entre le minimum et le maximum) des niveaux sonores en relatif à dix mètres de la source à trois distances de propagation (50, 250 et 500 m) sont présentées dans le Tableau (1) pour les systèmes considérés.

		10 min		60 min	
Système		(a)	(b)	(a)	(b)
50 m	moyenne	-13,4	-13,6	-13,3	-13,4
	étendue	0,3	0,1	0,8	0,4
250 m	moyenne	-25	-27,3	-25,1	-24,2
	étendue	6,8	0,9	13,7	3,7
500 m	moyenne	-30,0	-33,5	-30,9	-30,4
	étendue	5,8	0,4	23,3	2,1

TABLE 1 – Moyenne et étendue des niveaux sonores relatifs aux niveaux à 10 m (dB) pour les systèmes (a) et (b) à trois distances de propagation.

L'étendue des niveaux sonores est toujours plus grande dans le cas du système (a) que dans le cas du système (b) ce qui s'explique aisément par une plus forte dispersion des paramètres de réfraction et correspond aux résultats attendus. Le passage à des distances de propagation plus grandes, de 50 m à 250 m ou de 50 m à 500 m, montre une augmentation de l'étendue des niveaux sonores. En effet, plus les distances de propagation sont grandes, plus les effets de réfraction

sont marqués. Cette observation n'est cependant pas systématiquement vérifiée en comparant l'étendue des niveaux à 250 et 500 m. En effet la figure 4 montre une augmentation de l'étendue des niveaux sonores jusqu'à une distance d'environ 350 m. L'étendue tend alors à diminuer de 1 dB pour le système (a) et de 2 dB pour le système (b) jusqu'à la distance d'étude maximale. Ces observations sont expliquées par l'existence de zones interférentielles notamment à 500 m.

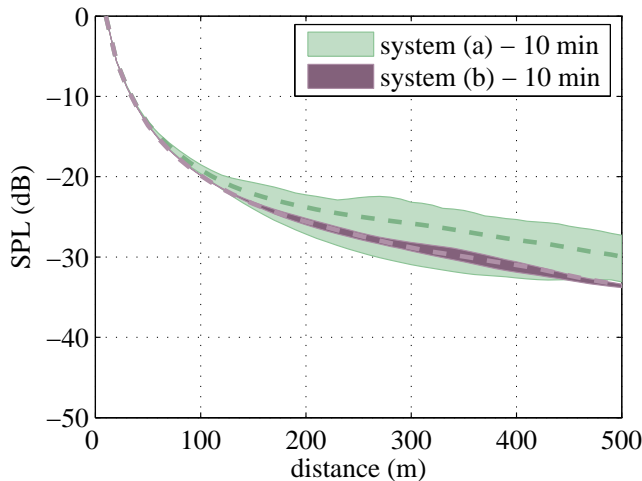


FIGURE 4 – Moyenne et étendue des niveaux sonores calculés sur 10 min en fonction de la distance à la source pour les systèmes (a) et (b).

Cette étude permet de montrer la forte sensibilité des niveaux sonores aux paramètres micrométéorologiques et notamment aux erreurs expérimentales et aux approches méthodologiques utilisées. A une distance donnée, l'étendue des niveaux sonores peut atteindre 24 dB (système (a) sur 60 minutes) pour ce cas d'étude. Le choix de l'intervalle de temps et de la méthode de mesure a donc un impact considérable sur l'estimation des niveaux sonores.

5 Conclusion

L'application de la méthode de propagation d'incertitude aux paramètres micrométéorologiques influents en propagation acoustique en milieu extérieur permet d'appréhender la dispersion des niveaux sonores due aux erreurs expérimentales, à la variabilité temporelle et aux approches méthodologiques utilisées. Cette étude tend à remettre en question la méthode conventionnelle de mesure indirecte des profils et gradients verticaux de célérité du son par capteurs lents. La prise en compte du paramètre de rugosité z_0 dans la procédure d'ajustement des profils pourrait remédier aux fortes dispersion observées dans cette étude. Le choix de l'intervalle de temps d'étude est également discutable du fait d'une grande dispersion des paramètres micrométéorologiques impliquant une dispersion importante sur les niveaux sonores globaux. Dans ce papier, l'étude est menée pour des conditions favorables à la propagation ; les cas de conditions homogènes et défavorables seront présentés dans des études ultérieures. Cette méthode statistique pourrait également être appliquée afin d'estimer quan-

titativement le paramètre le plus influent sur les profils et gradients verticaux de célérité et par suite sur les niveaux sonores, en propageant l'incertitude associée à chaque paramètre individuellement ou en procédant à des études de sensibilité. Cette étude ne rend compte que de la variabilité des paramètres micrométéorologiques et pourrait être combinée avec une étude de sensibilité sur les paramètres acoustiques de sol, sur les paramètres géométriques, et en prenant en compte les incertitudes de modélisation.

Références

- [1] Ecotière D., "Influence des fluctuations des conditions météorologiques sur la dispersion des niveaux sonores", *Acoustique et Technique* 51, 12-17 (2008)
- [2] Baume O., Wackernagel H., Gauvreau B., Junker F., Bérengier M., Chilès J.-P., "Space and time exploration of acoustical series under the influence of various micrometeorological stability conditions", *GEOstatistics for ENVironmental applications*, Rhodes (G), 45-57 (2006)
- [3] Pettit Chris L., Wilson D. Keith, "Proper orthogonal decomposition and cluster weighted modeling for sensitivity analysis of sound propagation in the atmospheric surface layer", *J. Acoust. Soc. Am.* 122(3), 1374-1390 (2007)
- [4] "Guide to the Expression of Uncertainty in Measurements" International Organisation for Standardisation, Geneva, Switzerland (1993) corrected and reprinted in 1995
- [5] Robert C., Casella G., "Monte-Carlo Statistical Methods", New York : Springer-Verlog (1999)
- [6] de Rocquigny E., Devictor N., Tarantola S., "Uncertainty in industrial practice - A guide to Quantitative Uncertainty Management", *John Wiley & Sons* (2008)
- [7] Lihoreau B., Gauvreau B., Blanc-Benon P., Bérengier M., "Outdoor sound propagation modeling in realistic environments : a coupling method using parabolic equation and atmospheric model", *J. Acoust. Soc. Am.* 120(1), 110-119 (2006)
- [8] Stull R.B., "An introduction to boundary layer meteorology", Kluwer Academic, Dordrecht, The Netherlands (1988)
- [9] Heimann D., Bakermans M., Defrance J., Kunher D., "Vertical sound speed profiles determined from meteorological measurements near the ground", *Acta Acustica* 93, 228-240 (2007)
- [10] Gauvreau B., "Methodological and metrological benchmarks for determining experimental values of meteorological parameters", In Invited paper to ASA/EEA/SFA joint congress, Acoustics'08, Paris (2008)
- [11] Gauvreau B., Ecotière D., Lefèvre H., Bonhomme B., "Propagation acoustique en milieu extérieur complexe. Caractérisation expérimentale in-situ des conditions micrométéorologiques. Eléments méthodologiques et métrologiques", *Etudes et recherches des laboratoires des ponts et chaussées* (2009)

Résumé

Les phénomènes de propagation acoustique en milieu extérieur sont soumis à de fortes incertitudes et à une variabilité spatio-temporelle non-négligeable. Effets de sol, phénomènes micrométéorologiques, variabilité de la source sont autant d'éléments qui agissent de façon combinée et qui rendent ainsi les phénomènes complexes. Les niveaux sonores sont alors difficilement mesurables et/ou estimés avec la précision requise d'un point de vue réglementaire et normatif. Une méthode statistique a donc été mise en place afin de pouvoir estimer ces incertitudes pour tous types de situations rencontrées : la méthode de calibration sous incertitude. Cette méthode intègre aussi bien des données modélisées que des données expérimentales ; des bases de données représentatives de sites différents plus ou moins complexes permettent d'appliquer le processus développé et de l'améliorer pas à pas. Cette méthode rend compte de la spécification du système (réflexions sur la constitution de jeux de données), la définition des observables, la calibration du système (l'écriture de la fonction coût et l'estimation des paramètres d'intérêt), la validation du système et la prévision. La méthode désormais aboutie peut être appliquée à tout type de cas d'étude dans la limite des possibilités des modèles de prévision et des données expérimentales à disposition en regard de la complexité des phénomènes ; ces limites sont également mises en évidence.

Mots-clés : propagation acoustique en milieu extérieur, site complexe, incertitude, variabilité, calibration, métamodèle, prévision.

ASSESSMENT OF UNCERTAINTY FOR OUTDOOR SOUND PROPAGATION

Abstract

Outdoor sound propagation phenomena are subject to substantial uncertainties and non-negligible variability in space and time. The combination of ground effects, micrometeorological phenomena, variability of the source gives rise to complex phenomena. It is so difficult to measure and/or assess sound pressure levels with the required accuracy from a regulatory and normative point of view. A statistical method has been defined in order to assess these uncertainties for any configurations : the Calibration Under Uncertainty (CUU) methodology. This method integrates both modeled data and experimental data ; available databases which are representative of more or less complex sites allow to apply the statistical process and to improve it step by step. System specification (datasets constitution), observables definition, system calibration (cost function formulation and parameters estimation), system validation and prediction are the main steps of the statistical process. This finalized process can be applied to any case study within the limitations of available prediction models and experimental data regarding the complexity of phenomena. These limitations are also highlighted.

Keywords: outdoor sound propagation, complex site, uncertainty, variability, calibration, metamodeling, forecasting.