



HAL
open science

Estimation des indices de Sobol à l'aide d'un métamodèle multi-éléments : application à la dynamique du vol

Nicolas Vauchel

► **To cite this version:**

Nicolas Vauchel. Estimation des indices de Sobol à l'aide d'un métamodèle multi-éléments : application à la dynamique du vol. Milieux fluides et réactifs. Université de Lille, 2023. Français. NNT : 2023ULILN008 . tel-04324781

HAL Id: tel-04324781

<https://theses.hal.science/tel-04324781>

Submitted on 5 Dec 2023

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

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



Thèse

pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ DE LILLE

École Doctorale Sciences de l'Ingénierie et des Systèmes - ENGSYS 632

Univ. Lille, CNRS, ONERA, Arts et Métiers Institute of Technology, Centrale Lille,
UMR 9014 - LMFL - Laboratoire de Mécanique des Fluides de Lille - Kampé de Fériet,
F-59000 Lille, France
présentée par

Nicolas VAUCHEL

Estimation des indices de Sobol à l'aide d'un métamodèle multi-éléments : application à la dynamique du vol

soutenue le 14/03/2023 à Lille devant le jury composé de :

Olivier LE MAITRE	Directeur de Recherche - École polytechnique	Rapporteur
Franck CAZAURANG	Professeur des Universités - Université de Bordeaux	Rapporteur
Marcello MELDI	Professeur des Universités - LMFL	Président du jury
Sidonie LEFEBVRE	Ingénieure de recherche - Onera Palaiseau	Examineur
Éric GARNIER	Directeur de recherche - Onera Lille	Encadrant
Thomas GOMEZ	Enseignant Chercheur - ICAM	Directeur de thèse

Table des matières

Remerciements	7
Notations	9
Abréviations	9
Notations mathématiques	11
Général	11
Dynamique du vol	11
Analyse de sensibilité, quantification d’incertitudes et algorithme ME-ACD	15
Notations des algorithmes	18
Divers	18
Introduction	19
Contexte	19
Objectifs scientifiques de la thèse	20
Plan du rapport de thèse	23
Extraits d’articles	23
1 Dynamique du vol	25
1.1 Établissement du système dynamique	25
1.1.1 Généralités et vocabulaire sur les avions	25
1.1.2 Hypothèses de l’étude	26
1.1.3 Les trois trièdres de référence	27
1.1.4 Cinématique de l’aéronef	29
1.1.5 Système dynamique et réductions	31
1.2 Modèle de représentation des coefficients aérodynamiques globaux	38
1.2.1 Forme générale des coefficients aérodynamiques	38
1.2.2 Structure additive	39
1.2.3 Obtention des données expérimentales	42
1.2.4 Méthodes de blending	43
1.2.5 Modèle de représentation sélectionné	48
1.3 Théorie des bifurcations	49
1.3.1 Notions d’équilibre, de stabilité et de domaine d’attraction	49
1.3.2 Bifurcations	53
1.4 Premières simulations et méthodes testées	55
2 Analyse de sensibilité et quantification d’incertitudes : état de l’art	59
2.1 Introduction	59
2.2 QOI et paramètres d’entrée	59

2.2.1	Phénomène aléatoire	59
2.2.2	Variable aléatoire réelle	60
2.2.3	Paramètres d'entrée	62
2.2.4	QOI	64
2.2.5	Transformations dans le cas de variables dépendantes	65
2.3	Notation indicielle des interactions	65
2.3.1	Indices d'interaction	66
2.3.2	Ensemble fini de composantes	67
2.3.3	Vecteurs de paramètres d'entrée	68
2.3.4	Symboles de Kronecker étendus	68
2.3.5	Interactions d'ordre inférieur	69
2.4	Analyse de sensibilité	69
2.4.1	Analyse de sensibilité locale	70
2.4.2	Analyse de sensibilité globale (autre que les indices de Sobol)	70
2.4.3	Indices de Sobol	71
2.5	Métamodèle par chaos polynomial	74
2.5.1	Polynômes	74
2.5.2	Orthogonalité des polynômes	76
2.5.3	Polynômes du chaos	80
2.6	Améliorations possibles des polynômes du chaos	88
2.6.1	Polynômes du chaos arbitraires	88
2.6.2	Polynômes du chaos épars	89
2.6.3	Utilisation d'une autre base	89
2.6.4	Métamodèles multi-éléments	89
3	Développement de la méthode ME-ACD	93
3.1	Introduction	93
3.2	Algorithme ME-ACD	93
3.2.1	Espaces abstraits	93
3.2.2	Étapes de l'algorithme	94
3.2.3	Étape 1 - Échantillonnage de l'espace des paramètres	95
3.2.4	Étape 2 - Génération de la matrice d'information globale	96
3.2.5	Étape 3 - Estimation des dérivées	98
3.2.6	Étape 4 - Partition de l'espace des paramètres et entraînement des modèles locaux	99
3.2.7	Étape 5 - Entraînement du classifieur	106
3.2.8	Résumé	107
3.3	Applications de ME-ACD	108
3.3.1	Évaluation de ME-ACD sur des QOI discontinues	109
3.3.2	Évaluation de ME-ACD sur une QOI présentant une discontinuité partielle	114
3.3.3	Évaluation de ME-ACD sur une QOI modélisant une bifurcation	115
3.3.4	Évaluation de ME-ACD sur la fonction de Genz en dimension 5	120
3.4	Conclusion	121
3.5	Extension à une QOI dépendant du temps	122
4	Estimation des indices de Sobol à l'aide d'un métamodèle multi-éléments	125
4.1	Nouvelle approche concernant les indices de Sobol	125

4.1.1	Forme explicite des termes ANOVA	125
4.1.2	Matrice Ξ	127
4.1.3	Vecteur $\tilde{\Xi}$	129
4.1.4	Estimation des indices de Sobol avec un métamodèle PC avec cette nouvelle approche	131
4.1.5	Estimation des indices de Sobol avec un métamodèle ME	133
4.2	Méthode numérique	137
4.2.1	Partition en union d'hyperpavés	137
4.2.2	Changement de base	140
4.2.3	Implémentation du vecteur $\tilde{\Xi}$	143
4.3	Applications	143
4.3.1	Métrique d'erreur	143
4.3.2	Fonction d'Ishigami	144
4.3.3	Fonction de Genz discontinue : comparaison des trois procédures concernant les hyperpavés hybrides	145
4.3.4	Étude de sensibilité en présence d'une bifurcation	147
4.4	Conclusion	153
5	Applications à la Dynamique du vol	155
5.1	Méthodologie	155
5.2	Applications	157
5.2.1	Analyse de sensibilité dans le domaine de vol normal	157
5.2.2	Influence des paramètres d'amortissement	159
5.2.3	Sensibilité et robustesse du comportement des avions au modèle de représentation	170
5.3	Conclusion	173
	Conclusion	179
	Conclusion	179
	Discussions et perspectives	180
A	Outils d'apprentissage automatique	183
A.1	Classifieurs	183
A.1.1	SVM	183
A.1.2	Classifieur par réseau de neurone MLP	184
A.2	Gradient Boosted Tree	186
A.3	Ball-tree	188

Remerciements

En premier lieu, je remercie Docteur Olivier Le Maître et Professeur Franck Cazaurang d'avoir accepté d'être rapporteur pour ce présent manuscrit. Je remercie également Madame Sidonie Lefebvre et Professeur Marcello Meldi pour avoir accepté de faire parti du jury de la thèse.

Je remercie grandement l'Onera et la Région Haut-de-France pour m'avoir offert l'opportunité financière et matérielle de réaliser ces travaux de thèse.

Je ne remercierais jamais assez mes encadrants Docteur Thomas Gomez, mon directeur de thèse, Monsieur Olivier Rénier, mon encadrant lors de la première moitié de la thèse et Docteur Éric Garnier, mon encadrant lors de la seconde moitié, qui m'ont énormément apporté, tant au niveau scientifique qu'au niveau humain. Sans le temps qu'ils m'ont accordé à répondre à mes nombreuses questions et à me soutenir, la thèse n'aurait jamais pu aboutir.

Je remercie également Dominique Farcy, Laurent Planckaert, Patricia Coton, ainsi que tout le reste de l'équipe de dynamique du vol du centre Onera Lille pour leur accueil, pour leur disponibilité et bien sûr pour le partage de leur précieuses connaissance et expérience.

Je remercie tous les autres collaborateurs du centre Onera Lille que j'ai eu le plaisir de côtoyer et qui ont rendu l'espace de travail agréable à vivre au quotidien. Avoir appris et travaillé à leurs côtés a été un réel plaisir. Ainsi, je remercie Quentin Gallas pour son soutien, Brigitte Duda pour l'aide dans l'organisation de mon travail au quotidien et pour son soutien émotionnel, Anne Gilliot pour avoir pris le temps de répondre à mes nombreuses questions, Giovanni Carbone pour m'avoir supporté comme voisin de bureau pendant la première partie de la thèse, Charles Pivot et Pierre Dubois pour le partage de leurs connaissances en apprentissage automatique et Célestin Ott pour sa bonne humeur à toute épreuve. Je remercie également Thomas Huret pour sa bienveillance, Rollando Cruz Marquez pour ses questionnements philosophiques qui ont animé de nombreuses pauses, Lauriane Lefevre pour m'avoir supporté comme voisin de bureau cette fois-ci lors de la seconde moitié de la thèse, ainsi que Baptiste Isnard et Elvina Derhille qui ont grandement participé à la bonne humeur de l'Open space.

Enfin, j'adresse de chaleureux remerciements à Cécilia Di-Filippo, la femme de ma vie, qui a toujours fait preuve d'un soutien sans faille à mon égard, notamment lors de la thèse

et lors des rudes épreuves que j'ai dû endurer en parallèle de ces travaux. Mille mercis s'adressent alors également à Janine Barras, qui n'a jamais cessé d'être présente pour son petit-fils. Le soutien de ces deux êtres chers à mon cœur a été une source de motivation intarissable et indispensable pour traverser tous les obstacles qui se sont dressés en travers de mon chemin, tant au niveau personnel qu'au niveau professionnel.

Notations

Abréviations

Dans tout le manuscrit, des abréviations sont utilisées. Le choix s'est porté sur l'utilisation des abréviations anglaises correspondantes. La première fois que le terme qui sera abrégé apparaît dans un chapitre, sa traduction anglaise ainsi que l'abréviation correspondante sont indiquées entre parenthèses. La liste ci-dessous reprend l'ensemble des abréviations apparaissant dans le manuscrit, à l'exception de celles concernant des laboratoires, des écoles ou des entreprises.

- ODE : Ordinary Differential Equation - Équation différentielle ordinaire
- GSA : Global Sensitivity Analysis - Analyse de sensibilité globale
- UQ : Uncertainty Quantification - Quantification d'incertitudes
- QOI : Quantity Of Interest - Quantité d'intérêt
- MC : Monte-Carlo method - méthode Monte-Carlo
- ML : Machine learning - Apprentissage automatique
- PC : Polynomial Chaos - Chaos polynomial
- g-PC : generalised Polynomial Chaos - Polynômes du chaos généralisés (aussi désignés par PC)
- NI : Non-Intrusive - Non intrusif
- ME : Multi-Element - Multi-éléments
- pdf : probability density function - fonction densité de probabilité
- ANOVA : ANalyse Of Variance
- OLS : Ordinary Least Square - Méthode des moindres carrés ordinaires
- MSE : Mean Square Error - Erreur moyenne quadratique
- PRESS : Predicted Residual Sum of Squares
- SVD : Singular Value Decomposition - Décomposition en valeurs singulières
- MLP : Multi Layer Perceptron - Perceptron multicouche

- GBT : Gradient Boosted Trees

Notations mathématiques

Général

- \mathbf{U} : notation d'un vecteur.
- \mathbf{A} : notation d'une matrice.
- Ox : notation d'un axe passant par un point O et de vecteur directeur \mathbf{x} .
- $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots)$: notation d'une base de vecteurs.
- Oxy : notation d'un plan passant par un point O et généré par l'espace vectoriel de base (\mathbf{x}, \mathbf{y}) .
- $Oxyz$: notation d'un repère (aussi appelé trièdre par la suite) d'origine O et de base $(\mathbf{x}, \mathbf{y}, \mathbf{z})$.
- $\mathbf{u} \cdot \mathbf{v}$: produit scalaire entre les vecteurs \mathbf{u} et \mathbf{v} .
- $\langle \cdot, \cdot \rangle$: produit scalaire entre deux « vecteurs » au sens d'éléments d'un espace vectoriel. Dans le manuscrit, ces éléments sont des polynômes.
- $\|u\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$: norme associée au produit scalaire $\langle \cdot, \cdot \rangle$.
- $\mathbf{u} \wedge \mathbf{v}$: produit vectoriel entre les vecteurs \mathbf{u} et \mathbf{v} .
- $\frac{df}{dx}$: dérivée de la fonction f par rapport à x .
- $\frac{\partial f}{\partial x}$: dérivée partielle de la fonction f par rapport à x .
- $\int_a^b f dx$: intégrale d'une fonction f par rapport à x entre les bornes a et b .
- $\text{card}(E)$: cardinal de l'ensemble fini E .
- \bar{A} : complémentaire de l'ensemble A . Si A représente un événement aléatoire, \bar{A} est l'événement contraire de A .
- $\delta_{m,n}$: symbole de Kronecker standard.

Dynamique du vol

- Caractéristiques de l'environnement et de l'écoulement :
 - ρ : masse volumique de l'air ambiant, supposée constante.
 - g : intensité de la pesanteur.
 - Re : nombre de Reynolds de l'écoulement.
 - M : nombre de Mach de l'écoulement.
- Caractéristiques de l'avion :

- m : masse de l'avion supposée constante.
 - O : centre de gravité de l'avion supposé fixe par rapport à ce dernier et situé dans son plan de symétrie.
 - \mathbf{I}_O : matrice d'inertie de l'avion calculée en son centre de gravité.
 - A : moment d'inertie de l'avion par rapport à l'axe Ox .
 - B : moment d'inertie de l'avion par rapport à l'axe Oy .
 - C : moment d'inertie de l'avion par rapport à l'axe Oz .
 - E : produit d'inertie de l'avion par rapport au plan Oxz .
 - l : longueur de référence - corde moyenne d'une aile de l'avion dans l'étude.
 - S : surface de référence - surface d'une aile de l'avion dans l'étude.
- Paramètres de contrôle de l'avion :
 - δ_l : position angulaire des ailerons.
 - δ_m : position angulaire de la gouverne de profondeur.
 - δ_n : position angulaire de la gouverne de direction.
- Référentiels d'étude en dynamique du vol :
 - R_T : référentiel terrestre supposé galiléen. Son origine est O_T et sa base associée est $(\mathbf{x}_T, \mathbf{y}_T, \mathbf{z}_T)$.
 - R : référentiel lié à l'avion. Son origine est O et sa base associée est $(\mathbf{x}, \mathbf{y}, \mathbf{z})$.
 - R_A : référentiel aérodynamique lié à la vitesse de l'avion. Son origine est O et sa base associée est $(\mathbf{x}_a, \mathbf{y}_a, \mathbf{z}_a)$.
- Position de l'avion par rapport au référentiel terrestre :
 - $\mathbf{O}_T\mathbf{O}$: vecteur position du centre de gravité de l'aéronef par rapport au référentiel terrestre.
 - x_O : composante de $\mathbf{O}_T\mathbf{O}$ selon \mathbf{x}_T .
 - y_O : composante de $\mathbf{O}_T\mathbf{O}$ selon \mathbf{y}_T .
 - z_O : composante de $\mathbf{O}_T\mathbf{O}$ selon \mathbf{z}_T .
 - Φ : angle de gîte de l'avion.
 - Θ : assiette longitudinale de l'avion.
 - Ψ : azimuth de l'avion.

- Vitesse de l'aéronef :
 - $\mathbf{V}(\mathbf{R}/\mathbf{R}_T)$: vitesse de l'aéronef par rapport au référentiel terrestre.
 - V : norme de la vitesse de l'aéronef par rapport au référentiel terrestre.
 - α : angle d'incidence de l'avion.
 - β : angle de dérapage de l'avion.

- Vitesse de rotation de l'aéronef :
 - $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$: vitesse de rotation de l'aéronef par rapport au référentiel terrestre.
 - p : roulis ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{x} .
 - q : tangage ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{y} .
 - r : lacet ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{z} .
 - p_a : roulis aérodynamique ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{x}_a .
 - q_a : tangage aérodynamique ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{y}_a .
 - r_a : lacet aérodynamique ; la composante de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ selon \mathbf{z}_a .
 - w_{ss} : composante « état stationnaire » du vecteur rotation prise en compte pour l'utilisation des données de la balance rotative dans les méthodes de « blending ».
 - $\mathbf{\Omega}_{osc}$: vecteur contenant les composantes « oscillations forcées » du vecteur rotation, prises en compte pour l'utilisation des données en oscillations forcées dans les méthodes de « blending ».
 - p_{osc} : composante de $\mathbf{\Omega}_{osc}$ selon \mathbf{x} .
 - q_{osc} : composante de $\mathbf{\Omega}_{osc}$ selon \mathbf{y} .
 - r_{osc} : composante de $\mathbf{\Omega}_{osc}$ selon \mathbf{z} .

- Cinétique :
 - $\sigma_{\mathbf{O}}$: moment cinétique de l'aéronef par rapport à son centre de gravité.

- Actions mécaniques extérieures appliquées à l'aéronef :
 - \mathcal{P} : poids de l'avion.
 - \mathcal{F} : résultante des actions mécaniques exercées par le fluide sur l'avion.
 - X : composante de \mathbf{F} selon \mathbf{x} .
 - Y : composante de \mathbf{F} selon \mathbf{y} .

- Z : composante de \mathbf{F} selon \mathbf{z} .
 - X_a : composante de \mathbf{F} selon \mathbf{x}_a , appelée traînée.
 - Y_a : composante de \mathbf{F} selon \mathbf{y}_a .
 - Z_a : composante de \mathbf{F} selon \mathbf{z}_a , appelée portance.
 - \mathcal{T} : résultante des actions mécaniques de propulsion de l'aéronef. Cette force est introduite, mais non considérée dans l'étude.
 - \mathbf{M}_O : moment total des actions mécaniques exercées par le fluide sur l'avion calculé au centre de gravité O .
 - L_O : composante de \mathbf{M}_O selon \mathbf{x} .
 - M_O : composante de \mathbf{M}_O selon \mathbf{y} .
 - N_O : composante de \mathbf{M}_O selon \mathbf{z} .
 - \mathbf{M}_M : moment total des actions mécaniques exercées par le fluide sur l'avion calculé au point de mesure expérimentale M .
 - L_M : composante de \mathbf{M}_M selon \mathbf{x} .
 - M_M : composante de \mathbf{M}_M selon \mathbf{y} .
 - N_M : composante de \mathbf{M}_M selon \mathbf{z} .
- Coefficients aérodynamiques globaux :
 - C_x : forme adimensionnée de X .
 - C_y : forme adimensionnée de Y .
 - C_z : forme adimensionnée de Z .
 - C_l : forme adimensionnée de L_M .
 - C_m : forme adimensionnée de M_M .
 - C_n : forme adimensionnée de N_M .
 - $C_{l,O}$: forme adimensionnée de L_O .
 - $C_{m,O}$: forme adimensionnée de M_O .
 - $C_{n,O}$: forme adimensionnée de N_O .
 - C_{x_a} : forme adimensionnée de X_a .
 - C_{y_a} : forme adimensionnée de Y_a .
 - C_{z_a} : forme adimensionnée de Z_a .

Analyse de sensibilité, quantification d'incertitudes et algorithme ME-ACD

- $(\Omega, \mathcal{A}, Pr)$: espace de probabilité.
- Ω : espace contenant toutes les issues possibles de l'expérience aléatoire.
- $\mathcal{P}(\Omega)$: ensemble des parties de Ω , ensemble contenant tous les événements de l'expérience aléatoire.
- \mathcal{A} : tribu sur Ω .
- Pr : mesure de probabilité.
- \mathcal{C} : fonction de répartition.
- ρ : fonction densité de probabilité.
- $\mathcal{U}(a, b)$: loi uniforme sur $[a, b]$.
- $\mathcal{N}(\mu, \sigma)$: loi normale de moyenne μ et d'écart-type σ .
- $\Gamma(\alpha, \beta)$: loi Gamma de paramètres α et β .
- $\mathcal{B}(\alpha, \beta)$: loi Bêta de paramètres α et β .
- $E(Y)$: espérance d'une variable aléatoire réelle Y .
- $D(Y)$: variance d'une variable aléatoire réelle Y .
- $\mathbf{P} = (P_1, \dots, P_d)$: paramètres d'entrée, modélisés par des variables aléatoires réelles.
- $\mathbf{p} = (p_1, \dots, p_d)$: valeurs des paramètres d'entrée regroupées dans un vecteur.
- d : nombre des paramètres d'entrée, aussi appelée dimension.
- ρ_m : fonction densité de probabilité du paramètre d'entrée p_m .
- $\mathbf{x} = (x_1, \dots, x_m)$: paramètres d'entrée associés où chaque x_m suit une loi de distribution stochastique uniforme sur $[0, 1]$ et correspond à p_m par une transformation bijective prenant en compte ρ_m .
- f : quantité d'intérêt.
- \tilde{f} : quantité d'intérêt dépendant des paramètres d'entrée \mathbf{x} après changement de variable; cette notation est ensuite confondue avec celle du ou des métamodèles qui l'approchent.
- \hat{f} : métamodèle approchant \tilde{f} ; cette notation est utilisée lors de l'introduction des métamodèles, mais n'est ensuite plus utilisée quand le métamodèle est aussi désigné par \tilde{f} .
- $\Omega_d = [1, d]$: ensemble fini contenant les entiers allant de 1 à d .
- $\mathcal{I}_j^{(d)}$: ensemble contenant les indices multidimensionnels d'interaction d'ordre j entre d variables.
- $\mathcal{I}_l^{(d)}\{\mathbf{i}\}$: ensemble contenant les indices multidimensionnels d'interaction d'ordre $l < j$ entre l des j variables indexées par les composantes de \mathbf{i} .

- $\mathcal{S}(\mathbf{i})$: ensemble fini contenant les composantes du vecteur \mathbf{i} .
- $\mathbf{x}_{\mathbf{i}} = (x_{i_1}, \dots, x_{i_d})$
- $\mathbf{x}_{\sim \mathbf{i}}$: vecteur contenant toutes les composantes de \mathbf{x} qui ne sont pas présentes dans $\mathbf{x}_{\mathbf{i}}$.
- $\Delta_{\mathbf{a}, \mathbf{b}}^E$: premier symbole de Kronecker étendu.
- $\tilde{\Delta}_{\mathbf{a}, \mathbf{b}}^E$: second symbole de Kronecker étendu.
- \mathbb{H}_d : hypercube unité de dimension d ; c'est l'espace des paramètres, l'espace contenant tous les \mathbf{x} possibles.
- Δ_m : indice de sensibilité « One-at-time ».
- δ_m : effet élémentaire utilisé pour déterminer les indices de Morris.
- $\tilde{f}_{\mathbf{i}}^{(j)}$: terme de la décomposition ANOVA correspondant à l'interaction \mathbf{i} .
- $\mu_{\mathbf{i}}^{(j)}$: espérance de \tilde{f} sachant $\mathbf{x}_{\mathbf{i}}$.
- $D_{\mathbf{i}}^{(j)}$: variance partielle de la QOI correspondante à l'interaction \mathbf{i} .
- $S_{\mathbf{i}}^{(j)}$: indice de Sobol correspondant à l'interaction \mathbf{i} .
- $D_m^{(T)}$: quantité servant à déterminer l'indice de Sobol total associé au paramètre d'entrée p_m .
- S_m^T : indice de Sobol total correspondant au paramètre d'entrée p_m .
- P, Q : polynômes univariés pris en exemple.
- $\mathcal{R}_n[X]$: espace des polynômes univariés de degré inférieur ou égal à n , confondu avec l'espace des fonctions polynomiales associé.
- \mathcal{P}, \mathcal{Q} : polynômes à d variables pris en exemple.
- $\mathcal{R}_n[X_1, \dots, X_d]$: espace des polynômes à d variables de degré inférieur ou égal à n , confondu avec l'espace des fonctions polynomiales associé..
- $\mathcal{K}_n^{(d)}$: ensemble fini contenant les multi-indices indexant les termes d'un polynômes à d variables de degré n .
- $(\Psi_{\mathbf{k}})$: base canonique des polynômes à d variables.
- $\langle \cdot, \cdot \rangle_{\rho_m}$: produit scalaire associé à la densité de probabilité ρ_m .
- $\langle \cdot, \cdot \rangle_{a,b}$: produit scalaire correspondant à l'intervalle $[a, b]$ introduit pour les polynômes univariés.
- $H = [a_1, b_1] \times \dots \times [a_d, b_d]$: Hyperpavé de dimension d exemple.
- $\langle \cdot, \cdot \rangle_H$: produit scalaire correspondant à l'hyperpavé H introduit pour les polynômes à d variables .
- $(\phi_k^{a,b})_{k \in \llbracket 0, n \rrbracket}$: base des polynômes univariés orthonormale par rapport au produit scalaire $\langle \cdot, \cdot \rangle_{a,b}$

- $(\Phi_{\mathbf{k}}^H)_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$: base des polynômes à d variables orthonormale par rapport au produit scalaire $\langle \cdot, \cdot \rangle_H$
- L_n : polynômes de Legendre.
- $l_{n,k}$: coefficients du k^{eme} monôme du polynôme de Legendre L_n .
- $\lambda_{n,k}^{a,b}$: coefficients du k^{eme} monôme du polynôme $\phi_k^{a,b}$.
- v_H : hypervolume de l'hyperpavé H .
- $\mathcal{I}_{[a,b]}$: fonction indicatrice sur le segment $[a, b]$.
- $\mathcal{H}_{\mathbf{k}}$: polynôme à d variables étant un produit des polynômes d'Hermite univariés.
- \mathcal{T} : ensemble d'entraînement pour un métamodèle.
- \mathcal{P} : l'espace des paramètres originel de dimension d où les observations sont représentées par des points de coordonnées (p_1, \dots, p_d) .
- \mathcal{F} : l'espace de dimension $d + 1$ où les observations sont représentées par des points de coordonnées $(x_1, \dots, x_d, \tilde{f}(\mathbf{x}))$.
- \mathcal{D} : l'espace de dimension $d + 1$ où les observations sont représentées par des points de coordonnées $(\tilde{f}(\mathbf{x}), \frac{\partial \tilde{f}}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial \tilde{f}}{\partial x_d}(\mathbf{x}))$.
- N : le nombre d'observations d'entraînement générés par la séquence de Sobol pour la méthode ME-ACD.
- $N_{E_{min}}$: le nombre minimal d'éléments souhaité pour la méthode ME-ACD..
- $N_{E_{max}}$: le nombre maximal d'éléments souhaité pour la méthode ME-ACD..
- k_{min} : le nombre minimal de voisins dans \mathcal{F} utilisés dans l'estimation des dérivées partielles pour la méthode ME-ACD..
- k_{max} : le nombre maximal de voisins dans \mathcal{F} utilisés dans l'estimation des dérivées partielles pour la méthode ME-ACD..
- n_{min} : le degré minimal qu'un métamodèle polynomial local peut avoir pour la méthode ME-ACD..
- n_{max} : le degré maximal qu'un métamodèle polynomial local peut avoir pour la méthode ME-ACD..
- δ_{min} : un seuil optionnel pour effectuer un arrêt anticipé de la recherche de partition si une partition ayant une erreur inférieure à ce seuil a déjà été trouvée pour la méthode ME-ACD..
- N_l : nombre de taux d'apprentissage initiaux testés pour obtenir le classifieur optimal pour la méthode ME-ACD..
- $isConnect$: un booléen indiquant si un diagramme de connectivité est réalisé pour la méthode ME-ACD..
- $k_{connect}$ (optionnel) : nombre de voisins connectés qu'une observation a dans le diagramme de connectivité optionnel pour la méthode ME-ACD.

- Ξ : matrice contenant les $\xi_{\mathbf{u},\mathbf{v}}$ permettant les estimations des quantités de variances dans la nouvelle approche.
- $\tilde{\Xi}$: vecteur associé à la matrice Ξ .
- N_H : nombre d'hyperpavés dans la nouvelle partition de l'espace des paramètres.
- $\mathcal{S}_m = \{s_{m,1}, \dots, s_{m,\beta_{m+1}}\}$: m^{eme} composante des sommets des hyperpavés dans la partition de l'espace des paramètres.
- z : nombre d'hyperpavés par direction dans la partition cartésienne.
- $H_{\mathbf{h}} = [s_{1,h_1}, s_{1,h_1+1}] \times \dots \times [s_{d,h_d}, s_{d,h_d+1}]$: hyperpavé de la partition de l'espace des paramètres.
- $v_{\mathbf{h}}$: hypervolume de $H_{\mathbf{h}}$
- $v_{\mathbf{h}}^E$: volume de la projection de l'hypercube $H_{\mathbf{h}}$ sur les directions indexées par les entiers compris dans E .
- $\Phi_{\mathbf{h},\mathbf{k}}(\mathbf{x}) = \prod_{m=1}^d \phi_{m,k_m,h_m}(x_m)$
- $\mathcal{I}_{\mathbf{h}}$: fonction indicatrice sur l'hyperpavé $H_{\mathbf{h}}$.

Notations des algorithmes

- Dans tous les pseudo-codes présents dans le manuscrit, « $X \leftarrow$ » signifie que la variable X reçoit la valeur donnée.
- De plus, un « : » utilisé comme indice d'un vecteur, d'une matrice ou d'un tenseur désigne toutes les lignes, colonnes ou hypercolonnes de ce dernier.

Divers

- Dans le cas d'un vecteur de dimension 1, les notions de scalaire et de vecteur sont confondues.
- L'idée qu'une fonction dépende d'un vecteur à j composantes ou l'idée qu'elle dépende des j composantes directement sont également confondues.
- Les notions de polynôme et de fonction polynomiale sont confondues.

Introduction

Contexte

La thèse s'intéresse à une problématique concrète de sécurité des aéronefs. Le domaine de vol post-décroché est un domaine aérodynamiquement complexe où l'écoulement autour des surfaces portantes (ailes et gouvernes) peut présenter de fortes instabilités et peut être partiellement ou massivement décollé. Dans ce domaine de vol, atteignable de façon accidentelle ou volontaire (avions d'entraînement ou de voltige), les moyens de contrôle usuels sont moins efficaces, voire totalement inefficaces, ce qui peut mettre en danger le pilote et ses potentiels passagers. Un comportement particulièrement dangereux observable dans le domaine de vol post-décroché et qui est étudié dans les présents travaux est la vrille, menant à un mouvement en spirale de l'aéronef vers le sol. La perte de contrôle en plein vol, ou accident « LOC-I » pour « LOSS of control in flight », est l'une des causes principales d'accidents d'avion. D'après le rapport de l'Agence européenne de la sécurité aérienne [2], il s'agit de la première cause d'accident fatal de 2011 à 2015. Ces pertes de contrôle seraient en grande partie dues au décrochage qui survient lors du décollage de l'avion. Le projet européen SUPRA [4] pour « Simulation of Upset Recovery in Aviation », réalisé de 2009 à 2012, prouve le désir d'améliorer les simulateurs dans le but d'entraîner les pilotes à détecter les pertes de contrôle et à savoir récupérer le contrôle du véhicule. L'amélioration des modèles des forces et moments aérodynamiques afin de pouvoir traduire les effets complexes apparaissant dans le domaine décroché est donc un des besoins concrets afin de réduire le nombre d'accidents dus à la perte de contrôle.

Le travail de recherche s'intéresse alors à la détermination des prévisions de vol dans le domaine post-décroché, ainsi qu'à leurs dépendances aux structures de modèles utilisées pour les coefficients aérodynamiques et aux incertitudes des données expérimentales sur lesquelles ces modèles reposent. La dynamique du mouvement de l'avion est régie par un système dynamique d'équations différentielles ordinaires (Ordinary Differential Equation - ODE) autonomes non linéaires. Dans ces équations, les effets du fluide sur l'aéronef apparaissent par le biais des coefficients aérodynamiques globaux, qui sont les forces et les moments adimensionnés appliqués par le fluide sur l'aéronef. Ces coefficients dépendent de façon non-linéaire d'un grand nombre de variables, dont la géométrie de l'aéronef, sa vitesse et sa vitesse de rotation par rapport à la Terre, ainsi que des caractéristiques de l'écoulement qui l'entoure. Pour chaque coefficient, un modèle de représentation ayant une certaine structure est déterminé pour décrire ces dépendances complexes. Ce modèle s'appuie sur des données expérimentales recueillies sur des maquettes de taille réduite, les données de vol libre sur avion réel étant trop coûteuses et trop risquées à collecter

dans le domaine post-décroché. Une autre piste pour l'établissement de ces bases serait d'utiliser des données venant de calculs numériques. Néanmoins, le caractère instationnaire et complexe de l'écoulement autour de la géométrie 3D de l'aéronef semble rendre les simulations trop coûteuses en terme de temps de calcul pour le moment, même si des études récentes explorent cette direction de recherche [3]. Les modèles utilisés dans le cadre de notre étude sont bâtis exclusivement sur des données expérimentales. Dans le système dynamique, les coefficients aérodynamiques globaux sont évalués par interpolation dans ces tables de données d'après la structure du modèle choisie. De par la nécessité de sélectionner une structure simplificatrice du modèle de représentation des coefficients aérodynamiques globaux, ces modèles sont lacunaires. De plus, ils sont entachés d'incertitudes dues au caractère intrinsèque des expériences. Ces lacunes et ces incertitudes vont impacter les résultats des prévisions de vol. L'objectif de la thèse est d'étudier ces impacts.

Objectifs scientifiques de la thèse

L'objectif premier de la thèse est la recherche de corrélations entre les incertitudes et/ou les lacunes des modèles de représentation des coefficients aérodynamiques et les comportements de l'aéronef dans le domaine post-décroché. La sensibilité et la robustesse du comportement des aéronefs aux modèles de représentation peuvent alors être qualifiées. Ces études s'effectuent tant sur les impacts des incertitudes issues des expériences que sur les lacunes de la structure sélectionnée du modèle de représentation. La thèse vise alors à améliorer les techniques de prévisions de vol dans le domaine post-décroché en terme de fiabilité. Les méthodes envisagées sont évaluées autour d'un modèle s'appuyant sur les données obtenues expérimentalement sur balance rotative dans la soufflerie verticale SV4 du centre ONERA Lille sur une maquette d'avion léger générique à l'échelle 1/5.

De façon classique en dynamique du vol, le système dynamique est étudié de façon asymptotique. Le système contient de nombreux équilibres qui sont reliés à des comportements en vol. Dans le domaine de vol normal, les équilibres représentent des trajectoires de vol sûres. Dans le domaine de vol post-décroché, ces équilibres représentent des comportements potentiellement dangereux, comme le « deep stall » ou les vrilles. Pour un ensemble de valeurs des paramètres de contrôle, il existe plusieurs équilibres : certains sont stables et d'autres instables. Avec l'évolution des paramètres de contrôle, la valeur de ces équilibres est modifiée et leur stabilité peut être modifiée. Dans le cas de modification de la stabilité, il y a bifurcation ; un changement de paramètre rend un équilibre stable instable ou inversement, ce qui a pour conséquence un changement du comportement du système. Le système dynamique est alors étudié à l'aide de la théorie des bifurcations. Des diagrammes de bifurcation sont tracés pour illustrer la présence des équilibres et des bifurcations [13, 28, 24, 25]. Ces équilibres sont alors reliés à des comportements observables.

Dans le cadre des travaux de cette thèse, le système n'est pas étudié asymptotiquement mais l'évolution temporelle des comportements observés est étudiée. Le choix s'est porté sur les techniques d'analyse de sensibilité globale (Global Sensitivity Analysis - GSA) afin obtenir des informations sur l'impact des incertitudes et des lacunes sur le système

dynamique. Les indices de sensibilité de Sobol sont ceux qui sont retenus. Ces indices donnent une information sur l'impact de paramètres incertains sur une Quantité d'intérêt (Quantity Of Interest - QOI), la quantité étudiée. Ils quantifient l'impact de l'action des paramètres incertains et de leurs interactions lorsqu'ils varient sur l'ensemble de leur intervalle. Les indices de Sobol peuvent être obtenus à l'aide de la méthode Monte-Carlo (Monte-Carlo method - MC), mais un grand nombre de simulations est souvent nécessaire pour les obtenir avec précision. Dans le système dynamique du vol, les QOI sont les variables d'état étudiées à des instants donnés. Le système dynamique du vol est relativement rapide à intégrer, mais comme il possède de très nombreux comportements intéressants à étudier, le coût de l'étude pour un aéronef donné peut devenir trop élevé.

Une solution est alors d'utiliser un métamodèle par chaos polynomial (Polynomial Chaos - PC), une méthode issue du domaine de la quantification d'incertitudes (Uncertainty Quantification - UQ). En effet, une relation analytique lie les coefficients du modèle PC et les indices de Sobol et l'obtention des métamodèles PC est moins coûteuse que la méthode MC. Néanmoins, la précision de ces métamodèles se dégrade en présence d'irrégularités et de discontinuités dans la dépendance de la QOI aux paramètres incertains. Il faut alors un grand nombre de termes du métamodèle pour approcher une telle QOI, ce qui devient problématique avec la célèbre « malédiction de la dimension ». De plus, en présence d'une discontinuité, le phénomène de Gibbs altère la précision du modèle PC et perdure même si le degré du polynôme augmente. Dans les systèmes dynamiques, une bifurcation peut se traduire par une irrégularité ou une discontinuité par rapport aux paramètres incertains à un instant donné assez avancé pour que la bifurcation soit développée. De telles discontinuités sont alors attendues dans l'étude du système dynamique du vol.

Pour remédier à la perte de précision des métamodèles PC dans le cas des QOI discontinues, les métamodèles multi-éléments (Multi-element - ME) ont été développés. Ces métamodèles sont des modèles PC par morceaux, s'appuyant sur une partition de l'espace des paramètres afin d'entraîner des métamodèles polynomiaux locaux sur chaque sous-ensemble de l'espace des paramètres, où la QOI est plus régulière localement. Mais si un tel métamodèle est plus précis pour une telle QOI, aucune relation analytique ne permet de retrouver les indices de Sobol à partir de ce dernier.

La bibliographie sur les techniques ME a mené à l'émergence de deux objectifs scientifiques supplémentaires pour les travaux de thèse. Ces objectifs concernent le développement de deux méthodes qui transcendent le domaine de la dynamique du vol et tendent à répondre à des problématiques concernant les métamodèles ME :

- Un nouvel algorithme permettant d'obtenir un métamodèle ME basé sur les méthodes modernes d'apprentissage automatique (Machine Learning - ML) est développé. Il se nomme ME-ACD, pour « Multi-Element model based on Agglomerative Clustering on the Derivatives ». Cet acronyme anglais signifie « méthode multi-éléments basée sur le regroupement hiérarchique ascendant des dérivées ».
- La démonstration d'un lien analytique entre les métamodèles ME et les indices de Sobol est démontré. L'obtention d'un métamodèle ME permet alors d'obtenir

rapidement l'ensemble des variances partielles permettant d'obtenir les indices recherchés, même pour une QOI discontinue.

Le développement de ces méthodes transcendant la dynamique du vol a occupé une place importante lors des travaux, ce qui a engendré un changement du nom de la thèse pour celui adopté actuellement. Son nom originel était : « Sensibilité et robustesse du comportement des aéronefs au voisinage et au-delà du décrochage aux modèles de représentation », ce titre ne faisant alors référence qu'à l'objectif initial de la thèse. Ces méthodes développées ont pour but de pouvoir être utilisées dans des études ayant une QOI discontinue et ayant un nombre de paramètres incertains relativement bas.

Les méthodes développées sont ensuite appliquées au système dynamique du vol afin de quantifier la sensibilité et la robustesse du comportement des aéronefs aux lacunes et aux incertitudes du modèle des coefficients globaux sélectionnés. Une critique qui peut être faite au sujet de la méthodologie adoptée dans l'étude est que le développement de ces méthodes est intervenu tôt lors de l'étude du système. Ce développement concernant la méthode ME a été réalisé sans que les méthodes de GSA déjà existantes ne soient testées sur le système. Contrairement à ce que montraient les résultats des premières simulations, les variables d'état du système dynamique du vol ne sont pas discontinues, mais seulement légèrement irrégulières, et ce même en présence de bifurcations. L'hypothèse de présence de discontinuités des variables d'état par rapport aux paramètres incertains, était fondée sur une erreur d'analyse. Le gain en terme de précision avec l'utilisation de la nouvelle méthodologie par rapport à la méthodologie PC classique n'est alors pas conséquent.

L'utilisation de la méthode d'analyse de sensibilité globale couplée à l'utilisation d'un métamodèle PC a déjà été réalisée pour étudier l'influence des paramètres d'un modèle de coefficients aérodynamiques dans [52]. Dans cet article, l'auteur utilise les méthodes de quantification d'analyse de sensibilité globale et de quantification d'incertitude afin d'étudier l'impact de deux composantes du coefficient de portance d'un modèle de coefficients aérodynamique linéaire dans le domaine de vol normal. D'après les connaissances de l'auteur, l'utilisation de ces méthodes pour évaluer l'impact des modèles des coefficients aérodynamiques sur le comportement d'un aéronef dans le domaine post-décroché n'a jamais été réalisée. L'utilisation de ces méthodes a permis de valider et d'approfondir les résultats d'une étude précédente [17] réalisée sur le même modèle de coefficients aérodynamiques, modèle présenté à la section 1.2.5. De par l'ajout artificiel de termes d'amortissement faisant défaut au modèle, l'analyse de sensibilité permet d'identifier quel terme d'amortissement aurait le plus d'impact sur la stabilité d'une vrille, ainsi que les combinaisons de termes d'amortissement qui mèneraient à la stabilité du comportement. Pour un comportement identifié à l'aide d'un modèle de coefficients lacunaire (sans terme d'amortissement), cette étude permet d'apporter des informations supplémentaires sur la stabilité potentielle du comportement et ainsi sur son danger. La sensibilité et la robustesse des comportements à haute incidence par rapport aux incertitudes commises sur les coefficients aérodynamiques peuvent également être estimées grâce à ces méthodes.

Plan du rapport de thèse

Le chapitre 1 traite de la dynamique du vol. Les hypothèses de l'étude sont introduites et le système dynamique régissant le mouvement de l'aéronef est mis en place. La modélisation des coefficients aérodynamiques globaux est ensuite abordée, traitant du choix de la structure du modèle de représentation parmi ceux existants, ainsi que des expériences permettant d'établir les tables de données sur lesquelles s'appuient ces modèles. En dernier lieu, les notions d'équilibre d'un système dynamique, de bifurcations et de domaine d'attraction sont introduites. Leur application en dynamique du vol est mise en lumière.

Le chapitre 2 présente un état de l'art sur les techniques d'analyse de sensibilité et d'UQ, en introduisant notamment les notions d'indices de Sobol, de métamodèles PC et de métamodèles ME. Ce chapitre présente également l'approche considérée dans les travaux, utilisant ces concepts.

Le chapitre 3 traite du développement de la méthode ME-ACD basée sur les techniques modernes de ML afin d'obtenir un métamodèle ME.

Le chapitre 4 présente une méthode permettant d'estimer les indices de Sobol à l'aide d'un modèle ME. Une nouvelle approche concernant les indices de Sobol est proposée afin d'obtenir une formule permettant de les estimer avec un modèle ME. L'estimation des indices de Sobol avec un modèle PC est reconduite avec cette approche pour montrer la cohérence de la démarche avec la formule classique.

Le chapitre 5 utilise l'approche développée au chapitre 4 dans le cas du système dynamique du vol à 8 variables d'état introduit au chapitre 1. Un métamodèle ME est obtenu à l'aide de l'extension temporelle de la méthode ME-ACD développée au chapitre 3.

Extraits d'articles

La thèse comporte des sections qui sont des versions traduites et légèrement modifiées de parties d'articles scientifiques écrits dans le cadre de l'étude.

- Le chapitre 2 contient des sections qui sont des versions traduites et légèrement modifiées d'extraits de [66] et de [67].
- Le chapitre 3 contient des sections qui sont des versions traduites et légèrement modifiées d'extraits de [66].
- Le chapitre 4 contient des sections qui sont des versions traduites et légèrement modifiées d'extraits de [67].
- Le chapitre 5 contient des sections qui sont des versions traduites et légèrement modifiées d'extraits de [65].

- Le chapitre 5 présente une étude de sensibilité sur des comportements proche de celle présentée dans l'article de congrès [65].

Chapitre 1

Dynamique du vol

Ce chapitre présente un état de l'art de l'étude de la dynamique d'un avion en vol. La mise en place du système dynamique d'équations différentielles ordinaires (Ordinary Differential Equation - ODE) autonomes couplées est traitée en premier lieu. Ce système est obtenu en appliquant le principe fondamental de la dynamique à l'aéronef modélisé par un solide indéformable. Le développement menant au système dans ce manuscrit est grandement inspiré de [70] qui reprend les fondamentaux de la dynamique du vol. En second lieu, la modélisation des coefficients aérodynamiques globaux, représentant les forces et les moments adimensionnés, est abordée. Le choix de la structure du modèle parmi les structures existantes, ainsi que l'obtention des données sur lesquelles reposent ces modèles sont développés. Enfin, la théorie des bifurcations est abordée. Les notions d'équilibre, de stabilité et de domaine d'attraction sont introduites, et l'utilisation de la théorie des bifurcations en dynamique du vol est traitée.

1.1 Établissement du système dynamique

1.1.1 Généralités et vocabulaire sur les avions

Le monde de l'aéronautique a un vocabulaire bien particulier entrevu dans cette courte section. Les différentes parties d'un avion (aussi qualifié d'aéronef par la suite) sont :

- Le fuselage : la partie aménagée pour recevoir la charge utile de l'appareil.
- La voilure : l'ensemble des surfaces portantes.
- Les gouvernes : les parties mobiles agissant sur le fluide ambiant et qui permettent de diriger l'avion.
- L'empennage : l'ensemble de plans fixes situés sur la queue de l'avion qui assurent la stabilité longitudinale et latérale de l'avion.

Sur la figure 1, les gouvernes de l'avion, qui ont pour but de contrôler le mouvement du véhicule dans le fluide, sont mises en évidence :

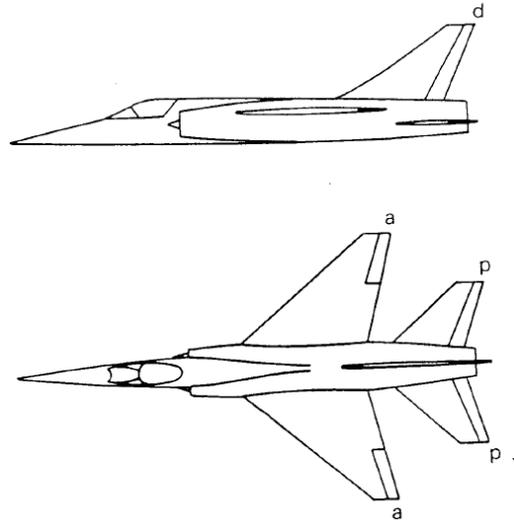


FIGURE 1 – Schéma d'un avion dont les gouvernes ont été mises en évidences (figure issue de [70]).

- Les gouvernes annotées « a » sont les ailerons (en anglais : « ailerons »). Ils répondent à la commande de gauchissement et permettent de réduire ou d'augmenter la portance sur chaque demi-voilure. Cela a pour but d'engendrer un couple de roulis, qui tend à incliner l'avion à droite ou à gauche.
- Les gouvernes annotées « p » sont les gouvernes de profondeur (en anglais : « elevator »). Elles répondent à la commande de profondeur et permettent de créer un moment de tangage qui tend à cabrer ou à faire piquer l'avion.
- La gouverne « d » est la gouverne de direction (en anglais : « rudder »). Elle répond à la commande de direction (ou palonnier) et a pour objectif d'engendrer un couple en lacet, qui tend à déplacer le nez de l'avion vers la gauche ou vers la droite.

Pour contrôler l'aéronef, ces gouvernes font un certain angle par rapport à leur position nominale aussi appelée « neutre » (position de référence lorsque tous les contrôles sont à zéro). Les positions angulaires des ailerons, de la gouverne de profondeur et de la gouverne de direction par rapport à leur position nominale sont respectivement notées δ_l , δ_m et δ_n .

Toutes ces gouvernes sont efficaces lorsque l'écoulement n'est pas décroché. Néanmoins, leur efficacité est réduite partiellement, voire totalement, au-delà du décrochage. Cela peut engendrer une perte de contrôle du véhicule.

1.1.2 Hypothèses de l'étude

- Le référentiel terrestre R_T est supposé galiléen. La rotondité de la Terre ainsi que les effets de sa rotation sur elle-même et de sa révolution autour du Soleil sont négligés.
- L'avion est modélisé par un solide indéformable.
- L'avion possède un plan de symétrie orthogonal à l'axe de son fuselage.

- La masse de l'avion est supposée constante. L'effet de la perte de carburant n'est alors pas pris en compte.
- Le centre de gravité de l'aéronef, supposé fixe, est situé dans son plan de symétrie.
- Les axes de l'avion sont supposés être ses axes d'inertie principaux. Le solide n'a donc aucun produit d'inertie, seuls ses 3 moments d'inertie sont considérés.
- La variation de la densité de l'air ρ avec la variation de l'altitude de l'aéronef est négligée.
- Aucune propulsion n'est considérée. Seul le vol plané ou la vrille sans propulsion sont étudiés.
- Les effets instationnaires de l'écoulement sur l'aéronef, dus au retard entre les actions du fluide entre les différentes surfaces portantes, sont négligés.

Ces hypothèses sont considérées (toutes ou en partie) de façon classique dans de nombreuses études de dynamique du vol [70, 28, 17, 23, 35, 59].

1.1.3 Les trois trièdres de référence

1.1.3.1 Définition des trièdres

Pour l'étude théorique du mouvement de l'avion, trois trièdres, associés à trois référentiels d'étude, sont considérés.

Le premier est un trièdre $O_T x_T y_T z_T$ associé au référentiel terrestre R_T . L'origine O_T est un point arbitraire fixe par rapport à la Terre et situé à la surface de celle-ci. L'axe $O_T z_T$ est orienté suivant la verticale descendante et les axes $O_T x_T$ et $O_T y_T$ sont choisis arbitrairement dans le plan horizontal. Ce référentiel est supposé galiléen d'après nos hypothèses. La rotondité de la Terre ainsi que les effets de sa rotation et de sa révolution autour du Soleil sont alors négligés.

L'avion est par hypothèse un solide indéformable. Le deuxième trièdre est alors un trièdre $Oxyz$ associé au référentiel R fixe par rapport à l'avion. L'origine O est le centre de gravité de l'aéronef (supposé fixe par rapport à ce dernier et situé dans son plan de symétrie d'après nos hypothèses). Les axes Ox et Oz sont deux axes orthogonaux du plan de symétrie de l'avion. L'axe Ox est parallèle à l'axe du fuselage et est orienté de l'arrière vers l'avant de l'avion et l'axe Oz est orienté du dos vers le ventre de l'avion. Oy vient compléter ce trièdre direct.

Le dernier trièdre, $Ox_a y_a z_a$, dit trièdre aérodynamique, est associé au référentiel aérodynamique R_A lié à la vitesse de l'aéronef. Son origine est le centre de gravité O de l'avion. L'axe Ox_a est le vecteur unitaire de même direction et de même sens que la vitesse de l'avion par rapport au référentiel terrestre. Oz_a est l'axe orthogonale à Ox_a situé dans le plan de symétrie de l'avion et orienté vers le ventre de celui-ci lorsque l'angle d'incidence

est nul. Oy_a complète ce trièdre direct.

Ces trièdres sont représentés sur les figures 2 et 3.

1.1.3.2 Positions relatives des trièdres

Les 3 trièdres de référence sont en mouvement les uns par rapport aux autres.

Le trièdre aérodynamique n'est qu'en rotation par rapport au trièdre avion, possédant la même origine que ce dernier. Les angles d'incidence, noté α , et de dérapage, noté β , définissent la position angulaire du trièdre aérodynamique par rapport au trièdre avion. L'angle d'incidence α est l'angle entre le plan $Ox_a y_a$ et l'axe Ox . C'est l'angle orienté de \mathbf{z}_a vers \mathbf{z} . L'angle de dérapage β est l'angle entre le plan Oxz et l'axe Ox_a . C'est l'angle orienté de \mathbf{y} vers \mathbf{y}_a . Ces angles sont représentés sur la figure 2.

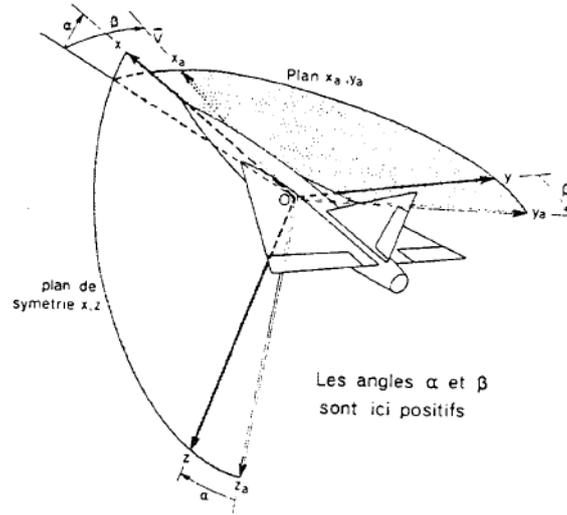


FIGURE 2 – Position du repère R_A par rapport au référentiel R , figure issue de [70]

La vitesse de rotation instantanée du repère R par rapport au repère R_A est donc

$$\boldsymbol{\Omega}(R/R_A) = \dot{\alpha}\mathbf{y} - \dot{\beta}\mathbf{z}_a \quad (1)$$

Le référentiel avion est quant à lui en translation et en rotation par rapport au référentiel terrestre. Son origine, le centre de gravité, est repéré par le vecteur position $\mathbf{O}_T\mathbf{O} = (x_0, y_0, -h)^T$ exprimé dans le repère R_T . h est l'altitude de l'avion, le signe « moins » étant présent car \mathbf{z}_T est la verticale descendante. Pour la rotation du trièdre avion par rapport au trièdre terrestre, les angles d'Euler illustrés sur la figure 3 sont utilisés. Ces angles Ψ , Θ et Φ portent respectivement les noms d'azimut, d'assiette longitudinale et d'angle de gîte.

Il existe également un positionnement avec des angles d'Euler définis pour le trièdre aérodynamique par rapport au trièdre terrestre mais il n'est pas introduit dans cette étude.

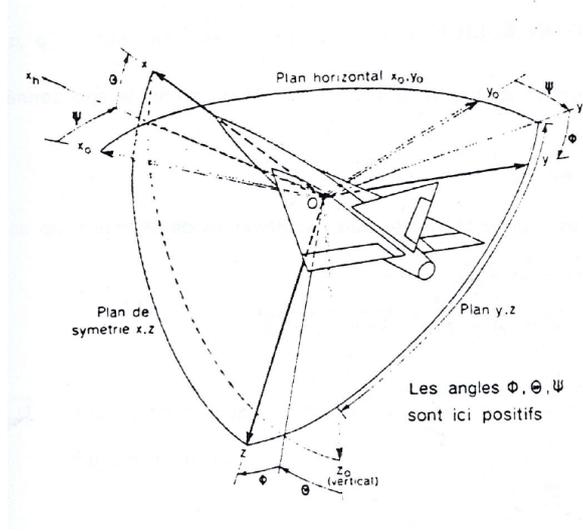


FIGURE 3 – Position du repère R par rapport au référentiel R_T , figure issue de [70]

1.1.3.3 Relations de passage

Soit un vecteur \mathbf{X} de composantes $(x_T, y_T, z_T)^T$ dans le trièdre terrestre, $(x, y, z)^T$ dans le trièdre avion et $(x_a, y_a, z_a)^T$ dans le trièdre aérodynamique. Les relations de passage entre ces composantes sont les suivantes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos \alpha \cos \beta & -\cos \alpha \sin \beta & -\sin \alpha \\ \sin \beta & \cos \beta & 0 \\ \sin \alpha \cos \beta & -\sin \alpha \sin \beta & \cos \alpha \end{bmatrix} \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} = \begin{bmatrix} \cos \alpha \cos \beta & \sin \beta & \sin \alpha \cos \beta \\ -\cos \alpha \sin \beta & \cos \beta & -\sin \alpha \sin \beta \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos \Psi \cos \Theta & \sin \Psi \cos \Theta & -\sin \Theta \\ \cos \Psi \sin \Theta \sin \Phi - \sin \Psi \cos \Phi & \sin \Psi \sin \Theta \sin \Phi + \cos \Psi \cos \Phi & \cos \Theta \sin \Phi \\ \cos \Psi \sin \Theta \cos \Phi + \sin \Psi \sin \Phi & \sin \Psi \sin \Theta \cos \Phi - \cos \Psi \sin \Phi & \cos \Theta \cos \Phi \end{bmatrix} \begin{bmatrix} x_T \\ y_T \\ z_T \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} x_T \\ y_T \\ z_T \end{bmatrix} = \begin{bmatrix} \cos \Psi \cos \Theta & \cos \Psi \sin \Theta \sin \Phi - \sin \Psi \cos \Phi & \cos \Psi \sin \Theta \cos \Phi + \sin \Psi \sin \Phi \\ \sin \Psi \cos \Theta & \sin \Psi \sin \Theta \sin \Phi + \cos \Psi \cos \Phi & \sin \Psi \sin \Theta \cos \Phi - \cos \Psi \sin \Phi \\ -\sin \Theta & \cos \Theta \sin \Phi & \cos \Theta \cos \Phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (5)$$

1.1.4 Cinématique de l'aéronef

1.1.4.1 Vitesse de l'avion par rapport au référentiel terrestre

La vitesse de l'avion par rapport à la Terre est $\mathbf{v}(R/R_T) = (V, 0, 0)^T$ exprimée dans le repère R_A , par définition de ce dernier.

1.1.4.2 Vitesse de rotation de l'avion par rapport au référentiel terrestre

Les composantes de la vitesse de rotation de l'avion par rapport au référentiel terrestre $\Omega(\mathbf{R}/\mathbf{R}_T)$ dans le trièdre avion portent des noms bien connus dans le monde de l'aéronautique :

- Le « roulis » est la composante sur \mathbf{x} , notée p .
- Le « tangage » est la composante sur \mathbf{y} , notée q .
- Le « lacet » est la composante sur \mathbf{z} , notée r .

Des définitions équivalentes existent pour les composantes autour des axes du trièdre aérodynamique :

- Le « roulis aérodynamique » est la composante sur \mathbf{x}_a . Cette composante sera notée p_a .
- Le « tangage aérodynamique » est la composante sur \mathbf{y}_a . Cette composante sera notée q_a .
- Le « lacet aérodynamique » est la composante sur \mathbf{z}_a . Cette composante sera notée r_a .

Ainsi, $\Omega(\mathbf{R}/\mathbf{R}_T) = (p, q, r)^T$ dans la base $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ et $\Omega(\mathbf{R}/\mathbf{R}_T) = (p_a, q_a, r_a)^T$ dans la base $(\mathbf{x}_a, \mathbf{y}_a, \mathbf{z}_a)$.

1.1.4.3 Relation entre les angles d'Euler et la vitesse de rotation

Par définition des angles d'Euler

$$p = -\dot{\Psi} \sin \Theta + \dot{\Phi} \quad (6)$$

$$q = \dot{\Psi} \cos \Theta \sin \Phi + \dot{\Theta} \cos \Phi \quad (7)$$

$$r = \dot{\Psi} \cos \Theta \cos \Phi - \dot{\Theta} \sin \Phi \quad (8)$$

et donc

$$\dot{\Psi} = \frac{1}{\cos \Theta} (q \sin \Phi + r \cos \Phi) \quad (9)$$

$$\dot{\Theta} = q \cos \Phi - r \sin \Phi \quad (10)$$

$$\dot{\Phi} = p + \tan \Theta (q \sin \Phi + r \cos \Phi) \quad (11)$$

$$(12)$$

1.1.4.4 Rappel de la formule de Bour

Soit \mathbf{A} un vecteur et R_1 un référentiel effectuant une rotation de vecteur $\boldsymbol{\Omega}(R_1/R)$ par rapport à un autre référentiel R . La formule de Bour permet de relier la dérivée du vecteur dans le trièdre lié à R_1 et celle dans le trièdre lié à R

$$\left[\frac{\partial \mathbf{A}}{\partial t}\right]_R = \left[\frac{\partial \mathbf{A}}{\partial t}\right]_{R_1} + \boldsymbol{\Omega}(R_1/R) \wedge \mathbf{A} \quad (13)$$

1.1.5 Système dynamique et réductions

1.1.5.1 Forme générale du système dynamique

La dynamique de l'avion peut être modélisée par un système de 12 ODE autonomes non-linéaires couplées mises sous la forme :

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \mathbf{U}) \quad (14)$$

où \mathbf{X} est le vecteur d'état dont les composantes décrivent l'état du système et \mathbf{U} est le vecteur de contrôle, vecteur contenant toutes les variables représentant les contrôles que peut avoir le pilote sur le système (positions angulaires des gouvernes dans l'étude). Les 12 variables d'état du système sont

$$\mathbf{X} = (V, \alpha, \beta, p, q, r, \Psi, \Theta, \Phi, x_0, y_0, h)^T \quad (15)$$

Les variables de contrôle sont

$$\mathbf{U} = (\delta_l, \delta_m, \delta_n)^T \quad (16)$$

Le pilote a bien sûr également un contrôle sur le moyen de propulsion de l'aéronef, mais aucune propulsion n'est considérée dans l'étude.

Les équations composant le système sont issues de l'application du principe fondamental de la dynamique à l'aéronef dans le référentiel terrestre R_T supposé galiléen, ainsi qu'à des relations cinématiques et des relations de dérivation.

1.1.5.2 Bilan des actions mécaniques extérieures s'appliquant à l'avion

Le poids de l'avion Dans le référentiel R_T , le poids s'exprime

$$\mathcal{P} = mg\mathbf{z}_T \quad (17)$$

D'après les relations de passage (4) et (3), il s'exprime dans le repère R_A

$$\begin{aligned}
\mathcal{P} = & mg[(-\sin \Theta \cos \alpha \cos \beta + \cos \Theta \sin \Phi \sin \beta + \cos \Theta \cos \Phi \sin \alpha \cos \beta)\mathbf{x}_a \\
& + (\sin \Theta \cos \alpha \sin \beta + \cos \Theta \sin \Phi \cos \beta - \cos \Theta \cos \Phi \sin \alpha \sin \beta)\mathbf{y}_a \\
& + (\sin \Theta \sin \alpha + \cos \Theta \cos \Phi \cos \alpha)\mathbf{z}_a]
\end{aligned} \tag{18}$$

D'après les hypothèses (Cf. section 1.1.2), la masse de l'aéronef m est supposée constante. La perte de masse due à la consommation de carburant n'est pas considérée dans l'étude.

Les actions mécaniques aérodynamiques Les actions mécaniques aérodynamiques sont celles exercées par le fluide sur l'aéronef au cours de son mouvement. Sur chaque surface portante (ailes et gouvernes), ces actions mécaniques se réduisent à une résultante et un couple en un point, le « foyer ». Si le calcul des moments est effectué au centre de gravité O , le moment total engendré par le fluide sur l'aéronef est la somme des couples ainsi que des moments issus des résultantes appliquées à chaque foyer. La résultante totale de ces forces est notée \mathcal{F} . Ses composantes sont notées $(X, Y, Z)^T$ dans le trièdre avion et $(X_a, Y_a, Z_a)^T$ dans le trièdre aérodynamique. Le moment total de ces actions calculé en O est noté \mathcal{M}_O . Ses composantes sont notées $(L_O, M_O, N_O)^T$ dans le trièdre avion.

Les coefficients aérodynamiques globaux sont les formes adimensionnées des forces et des moments aérodynamiques calculés en O

$$X = -\frac{1}{2}\rho SV^2 C_x \tag{19}$$

$$Y = \frac{1}{2}\rho SV^2 C_y \tag{20}$$

$$Z = -\frac{1}{2}\rho SV^2 C_z \tag{21}$$

$$L_O = \frac{1}{2}\rho SlV^2 C_{l,O} \tag{22}$$

$$M_O = \frac{1}{2}\rho SlV^2 C_{m,O} \tag{23}$$

$$N_O = \frac{1}{2}\rho SlV^2 C_{n,O} \tag{24}$$

où :

- S est la surface d'une aile de l'aéronef prise comme surface de référence.
- l est la corde moyenne d'une aile de l'aéronef prise comme longueur de référence.

Les coefficients aérodynamiques globaux $C_{x_a}, C_{y_a}, C_{z_a}$, formes adimensionnées des composantes de la force dans le trièdre aérodynamique (X_a, Y_a, Z_a) , sont définis de manière analogue. Un signe « moins » est également présent par convention dans l'adimensionnement des forces X_a et Z_a , pour que les coefficients, respectivement de « trainée » et de « portance », soient positifs en vol normal (la force de trainée étant dirigée selon $-\mathbf{x}_a$ et

la force de portance selon $-\mathbf{z}_a$).

Les coefficients adimensionnés dépendent de la géométrie de l'avion (et donc des positions angulaires des gouvernes), de son état, ainsi que des propriétés du fluide environnant. Après avoir choisi une structure de modèle de représentation de ces coefficients, leurs dépendances aux différentes variables sont obtenues à partir de données expérimentales. Cette modélisation est développée plus en détail à la section 1.2. Un des objectifs de la thèse est l'étude de l'influence de la structure de ces modèles de représentation ainsi que des incertitudes commises dans la détermination de ces derniers.

La prise de mesure des données expérimentales n'est pas réalisée au centre de gravité O , mais en un point de mesure M . La position du centre de gravité O est nommée le « centrage ». Ce dernier est exprimé dans cette étude par rapport au point de mesure M et est exprimé en ratio de corde OM/c . Plusieurs « centrages » peuvent être étudiés avec les mêmes données expérimentales, pour une même géométrie d'avion, les forces aérodynamiques ne dépendant pas de la répartition de masse de l'aéronef. Il est alors nécessaire de déplacer le moment obtenu à partir des tables de données expérimentales au centre de gravité de l'avion pour un centrage donné, car le centre de gravité est le point où est appliqué le principe fondamental de la dynamique pour mettre en place le système dynamique.

$$\mathcal{M}_O = \mathcal{M}_M + \mathbf{OM} \wedge \mathcal{F} \quad (25)$$

Les formes adimensionnées des composantes (L_M, M_M, N_M) de \mathcal{M}_M sont notées respectivement C_l, C_m et C_n .

Les forces et moments aérodynamiques dépendent de la masse volumique ρ du fluide environnant, qui dépend elle-même de l'altitude h à laquelle se trouve l'aéronef. Dans l'étude de ce présent manuscrit, la dépendance de la masse volumique à l'altitude est négligée par hypothèse (Cf. section 1.1.2).

Propulsion La propulsion est l'action mécanique permettant de pousser l'aéronef afin qu'il puisse se mouvoir. Le contrôle que possède le pilote sur la propulsion est simplifié ici par un paramètre de contrôle δ_T qu'il possède sur la norme $T(\delta_T)$ de la résultante \mathcal{T} de l'action mécanique de propulsion. Dans le cas de l'avion, elle est située dans le plan de symétrie de ce dernier, mais peut faire un angle non nul avec le vecteur \mathbf{x} . Cette angle est nommé « angle de calage » et est noté ω_{cal} . La direction de \mathbf{T} passe néanmoins par O , n'engendrant ainsi aucun moment autour de ce dernier

$$\mathcal{T} = T \cos \omega_{cal} \mathbf{x} - T \sin \omega_{cal} \mathbf{z} \quad (26)$$

La propulsion n'est pas considérée dans l'étude de ce présent manuscrit : $T = 0$ (Cf. section 1.1.2). Seuls le vol plané et la vrille sans propulsion sont considérés.

1.1.5.3 Équations du mouvement

Le principe fondamental de la dynamique est appliqué à l'avion dans le référentiel terrestre R_T supposé galiléen pour obtenir les 6 équations de mouvement qui sont les 6 premières équations du système dynamique. Ces équations peuvent être exprimées dans l'un des trièdres au choix. Pour faire apparaître directement les variables d'état sélectionnées, l'équation de la résultante est projetée dans le repère R_A et l'équation du moment est projetée dans le repère R .

Équation de la résultante

$$m \left[\frac{\partial \mathbf{v}}{\partial t} \right]_{R_T} = \mathcal{P} + \mathcal{F} \quad (27)$$

En utilisant la formule de Bour (13)

$$\left[\frac{\partial \mathbf{v}}{\partial t} \right]_{R_T} = \left[\frac{\partial \mathbf{v}}{\partial t} \right]_{R_A} + \boldsymbol{\Omega}(R_A/R_T) \wedge \mathbf{v} \quad (28)$$

or

$$\boldsymbol{\Omega}(R_A/R_T) = -\boldsymbol{\Omega}(R/R_A) + \boldsymbol{\Omega}(R/R_T) = -\dot{\alpha}(\sin \beta \mathbf{x}_a + \cos \beta \mathbf{y}_a) + \dot{\beta} \mathbf{z}_a + p_a \mathbf{x}_a + q_a \mathbf{y}_a + r_a \mathbf{z}_a \quad (29)$$

donc

$$m \left[\frac{\partial \mathbf{v}}{\partial t} \right]_{R_T} = m \dot{V} \mathbf{x}_a + mV(\dot{\alpha} \cos \beta \mathbf{z}_a + \dot{\beta} \mathbf{y}_a - q_a \mathbf{z}_a + r_a \mathbf{y}_a) \quad (30)$$

En projetant (27) dans le trièdre aérodynamique

$$m \dot{V} = mg[-\sin \Theta \cos \alpha \cos \beta + \cos \Theta \sin \Phi \sin \beta + \cos \Theta \cos \Phi \sin \alpha \cos \beta] - \frac{1}{2} \rho S V^2 C_{x_a} \quad (31)$$

$$mV(\dot{\beta} + r_a) = mg[\sin \Theta \cos \alpha \sin \beta + \cos \Theta \sin \Phi \cos \beta - \cos \Theta \cos \Phi \sin \alpha \sin \beta] + \frac{1}{2} \rho S V^2 C_{y_a} \quad (32)$$

$$mV(\dot{\alpha} \cos \beta - q_a) = mg[\sin \Theta \sin \alpha + \cos \Theta \cos \Phi \cos \alpha] - \frac{1}{2} \rho S V^2 C_{z_a} \quad (33)$$

Équations du moment Soit σ_O , le moment cinétique de l'avion par rapport à son centre de gravité. D'après l'équation du moment calculée en O

$$\left[\frac{\partial \sigma_O}{\partial t} \right]_{R_T} = \mathbf{M}_O \quad (34)$$

Soit \mathbf{I}_O , la matrice d'inertie de l'aéronef au point O . Comme le plan (Oxz) est un plan de symétrie pour l'avion, \mathbf{I}_O est de la forme :

$$\mathbf{I}_O = \begin{bmatrix} A & 0 & -E \\ 0 & B & 0 \\ -E & 0 & C \end{bmatrix} \quad (35)$$

A , B et C sont les moments d'inertie de l'avion et E est son produit d'inertie par rapport au plan Oxz . Le moment cinétique de l'avion autour de O est donc

$$\sigma_{\mathbf{O}} = \mathbf{I}_{\mathbf{O}} \cdot \boldsymbol{\Omega}(\mathbf{R}/\mathbf{R}_T) = (Ap - Er)\mathbf{x} + Bqy + (-Ep + Cr)\mathbf{z} \quad (36)$$

en utilisant la formule de Bour (13)

$$\begin{aligned} \left[\frac{\partial \sigma_{\mathbf{O}}}{\partial t} \right]_{\mathbf{R}_T} &= (A\dot{p} - E\dot{r} - Epq + Crq - Brq)\mathbf{x} \\ &+ (B\dot{q} + Apr - Er^2 + Ep^2 - Crp)\mathbf{y} \\ &+ (-E\dot{p} + C\dot{r} + Bpq - Apq + Erq)\mathbf{z} \end{aligned} \quad (37)$$

En injectant (37) dans (34) puis en projetant l'équation sur les axes du trièdre avion

$$A\dot{p} - E\dot{r} + (C - B)rq - Epq = \frac{1}{2}\rho SIV^2 C_{l,O} \quad (38)$$

$$B\dot{q} + (A - C)rp - Er^2 - Ep^2 = \frac{1}{2}\rho SIV^2 C_{m,O} \quad (39)$$

$$C\dot{r} - E\dot{p} + (B - A)qp + Erq = \frac{1}{2}\rho SIV^2 C_{n,O} \quad (40)$$

Une hypothèse supplémentaire est alors adoptée : E est négligé devant les moments d'inertie A , B et C (Cf. section 1.1.2). Cette hypothèse se vérifie avec les données inertielles de l'avion léger considéré. Ainsi, les axes du référentiel \mathbf{R} sont les axes principaux d'inertie de l'avion et les équations (38), (39) et (40) deviennent

$$A\dot{p} + (C - B)rq = \frac{1}{2}\rho SIV^2 C_{l,O} \quad (41)$$

$$B\dot{q} + (A - C)rp = \frac{1}{2}\rho SIV^2 C_{m,O} \quad (42)$$

$$C\dot{r} + (B - A)qp = \frac{1}{2}\rho SIV^2 C_{n,O} \quad (43)$$

1.1.5.4 Les 6 autres équations

Équations cinématiques Les équations du système dynamique relatives aux dérivées des angles d'Euler sont les relations (9), (10) et (11).

Équations de dérivation Les équations du système dynamique relatives aux dérivées des coordonnées du centre de gravité dans le repère terrestre sont de simples relations de dérivation. Les trois dérivées correspondent aux composantes de la vitesse de l'aéronef dans le trièdre terrestre (avec un signe « - » pour l'altitude h). Ainsi, d'après les relations de passage (2) et (5)

$$\begin{aligned} \dot{x}_0 &= V \cos \alpha \cos \beta \cos \Psi \cos \Theta + V \sin \beta (\cos \Psi \sin \Theta \sin \Phi - \sin \Psi \cos \Phi) \\ &+ V \cos \beta \sin \alpha (\cos \Psi \sin \Theta \cos \Phi + \sin \Psi \sin \Phi) \end{aligned} \quad (44)$$

$$\begin{aligned} \dot{y}_0 &= V \cos \alpha \cos \beta \sin \Psi \cos \Theta + V \sin \beta (\sin \Psi \sin \Theta \sin \Phi + \cos \Psi \cos \Phi) \\ &+ V \cos \beta \sin \alpha (\sin \Psi \sin \Theta \cos \Phi - \cos \Psi \sin \Phi) \end{aligned} \quad (45)$$

$$\dot{h} = V \cos \alpha \cos \beta \sin \Theta - V \sin \beta \cos \Theta \sin \Phi - V \cos \beta \sin \alpha \cos \Theta \cos \Phi \quad (46)$$

1.1.5.5 Système dynamique à 12 variables

Le système dynamique est composé des équations (31), (32), (33), (41), (42), (43), (9), (10), (11), (44), (45) et (46) mises sous la forme (14)

$$\dot{V} = g(-\sin \Theta \cos \alpha \cos \beta + \cos \Theta \sin \Phi \sin \beta + \cos \Theta \cos \Phi \sin \alpha \cos \beta) - \frac{1}{2m} \rho S V^2 C_{x_a} \quad (47)$$

$$\dot{\alpha} = \frac{q_a}{\cos \beta} + \frac{g}{V \cos \beta} (\sin \Theta \sin \alpha + \cos \Theta \cos \Phi \cos \alpha) - \frac{1}{2mV \cos \beta} \rho S V^2 C_{z_a} \quad (48)$$

$$\dot{\beta} = -r_a + \frac{g}{V} (\sin \Theta \cos \alpha \sin \beta + \cos \Theta \sin \Phi \cos \beta - \cos \Theta \cos \Phi \sin \alpha \sin \beta) + \frac{1}{2mV} \rho S V^2 C_{y_a} \quad (49)$$

$$\dot{p} = \frac{B-C}{A} r q + \frac{1}{2A} \rho S I V^2 C_{l,O} \quad (50)$$

$$\dot{q} = \frac{C-A}{B} p r + \frac{1}{2B} \rho S I V^2 C_{m,O} \quad (51)$$

$$\dot{r} = \frac{A-B}{C} p q + \frac{1}{2C} \rho S I V^2 C_{n,O} \quad (52)$$

$$\dot{\Psi} = \frac{1}{\cos \Theta} (q \sin \Phi + r \cos \Phi) \quad (53)$$

$$\dot{\Theta} = q \cos \Phi - r \sin \Phi \quad (54)$$

$$\dot{\Phi} = p + \tan \Theta (q \sin \Phi + r \cos \Phi) \quad (55)$$

$$\dot{x}_0 = V \cos \alpha \cos \beta \cos \Psi \cos \Theta + V \sin \beta (\cos \Psi \sin \Theta \sin \Phi - \sin \Psi \cos \Phi) + V \cos \beta \sin \alpha (\cos \Psi \sin \Theta \cos \Phi + \sin \Psi \sin \Phi) \quad (56)$$

$$\dot{y}_0 = V \cos \alpha \cos \beta \sin \Psi \cos \Theta + V \sin \beta (\sin \Psi \sin \Theta \sin \Phi + \cos \Psi \cos \Phi) + V \cos \beta \sin \alpha (\sin \Psi \sin \Theta \cos \Phi - \cos \Psi \sin \Phi) \quad (57)$$

$$\dot{h} = V \cos \alpha \cos \beta \sin \Theta - V \sin \beta \cos \Theta \sin \Phi - V \cos \beta \sin \alpha \cos \Theta \cos \Phi \quad (58)$$

où, d'après la relation de passage (2) :

$$q_a = -p \cos \alpha \sin \beta + q \cos \beta - r \sin \alpha \sin \beta \quad (59)$$

$$r_a = -p \sin \alpha + r \cos \alpha \quad (60)$$

$$C_{x_a} = C_x \cos \alpha \cos \beta - C_y \sin \beta + C_z \sin \alpha \cos \beta \quad (61)$$

$$C_{y_a} = C_x \cos \alpha \sin \beta + C_y \cos \beta + C_z \sin \alpha \sin \beta \quad (62)$$

$$C_{z_a} = -C_x \sin \alpha + C_z \cos \alpha \quad (63)$$

1.1.5.6 Réduction à un système à 8 variables

Ce système à 12 variables peut être réduit à un système à 8 variables. En effet, pour être étudié avec les notions d'équilibre et de bifurcation (Cf. section 1.3), il est nécessaire de ne

pas considérer les variables Ψ , x_0 , y_0 et h comme variable d'état du système dynamique. Sinon, un équilibre correspondrait à un comportement où ces quantités sont constantes, soit un aéronef fixe par rapport à la Terre.

Retirer ces variables du système est possible sans hypothèse supplémentaire car aucun des membres de droite ne dépend de x_0 , de y_0 ou de h (ρ ne dépendant pas de h d'après nos hypothèses). De plus, seules les équations concernant la dérivée de ces 3 variables dépendent de Ψ . En effet, l'azimut et la position du centre de gravité n'influent pas sur la dynamique de l'aéronef (ρ ne dépendant pas de l'altitude). Les équations concernant la dérivée de ces 4 variables peuvent donc être retirées du système dynamique. L'évolution de ces variables pourra être déterminée a posteriori, une fois l'évolution temporelle de toutes les autres variables d'état obtenues à l'aide de l'intégration du système.

Ce système à 8 variables d'état est utilisé dans de nombreuses études de dynamique du vol utilisant la théorie des bifurcations [28, 35, 23, 17, 59]

1.1.5.7 Réduction à un système à 5 variables

D'après [24], dans les mouvements à haut taux de rotation, comme les vrilles, il est possible de considérer que le poids de l'aéronef est négligeable devant les effets d'inertie et les forces aérodynamiques. La vitesse est de plus considérée comme étant constante. De par ces hypothèses, l'assiette longitudinale Θ et l'angle de gîte Φ n'agissent plus dans la dynamique du mouvement de l'aéronef ; ils n'apparaissent plus dans le membre de droite des équations concernant les autres variables d'état. Ils peuvent donc, au même titre que les variables Ψ , x_0 , y_0 et h , être déterminés a posteriori. Les 5 équations sont donc

$$\dot{\alpha} = \frac{q_a}{\cos \beta} - \frac{1}{2mV \cos \beta} \rho S V^2 C_{z_a} \quad (64)$$

$$\dot{\beta} = -r_a + \frac{1}{2mV} \rho S V^2 C_{y_a} \quad (65)$$

$$\dot{p} = \frac{B - C}{A} r q + \frac{1}{2A} \rho S l V^2 C_{l,O} \quad (66)$$

$$\dot{q} = \frac{C - A}{B} p r + \frac{1}{2B} \rho S l V^2 C_{m,O} \quad (67)$$

$$\dot{r} = \frac{A - B}{C} p q + \frac{1}{2C} \rho S l V^2 C_{n,O} \quad (68)$$

Cette réduction est souvent utilisée pour simplifier le système pour les études de comportement à haut taux de rotation, comme le problème de couplage en roulis [24] ou l'étude de vrilles [59]. Dans [59], il est fait état de cette réduction pour étudier les régions d'attraction des équilibres du système simplifié (Cf. section 1.3.1.5). Il est alors mentionné que les équilibres du système réduit à 5 variables sont proches de ceux du système à 8 variables si l'on impose que la vitesse de l'aéronef, supposée constante, soit celle obtenue pour l'équilibre correspondant dans le système à 8 variables.

1.1.5.8 Réduction à un système à 3 variables

Un système encore plus réduit peut être obtenu en ne considérant que les effets de rotation et en négligeant q_a et r_a devant p_a , ce qui est vérifié pour la plupart des vrilles à très hauts angles d'incidence ($\alpha \approx \frac{\pi}{2}$). Avec ces hypothèses, d'après (2)

$$p = p_a \cos \alpha \cos \beta \quad (69)$$

$$q = p_a \sin \beta \quad (70)$$

$$r = p_a \sin \alpha \cos \beta \quad (71)$$

$$(72)$$

et les trois équations sont :

$$\dot{p} = \frac{B-C}{A}(p_a)^2 \sin \alpha \cos \beta \sin \beta + \frac{1}{2A} \rho S V^2 C_{l,O} \quad (73)$$

$$\dot{q} = \frac{C-A}{B}(p_a)^2 \cos \alpha \cos \beta \sin \alpha \cos \beta + \frac{1}{2B} \rho S V^2 C_{m,O} \quad (74)$$

$$\dot{r} = \frac{A-B}{C}(p_a)^2 \cos \alpha \cos \beta \sin \beta + \frac{1}{2C} \rho S V^2 C_{n,O} \quad (75)$$

Les points d'équilibre de ce système à 3 variables permettent donc d'estimer les valeurs de α , de β et p_a pour une vrille donnée. Ces paramètres peuvent constituer de bons points initiaux potentiels pour la recherche d'équilibre des systèmes possédant plus de variables (en déduisant de ces valeurs celles de p , q et r). C'est ce qui est réalisé dans [17].

1.2 Modèle de représentation des coefficients aérodynamiques globaux

1.2.1 Forme générale des coefficients aérodynamiques

Comme décrit à la section 1.1.5.2, les actions mécaniques du fluide sur l'aéronef sont représentées par une résultante et un moment exprimé en un point. Leurs composantes sont manipulées sous leur forme adimensionnée : les « coefficients aérodynamiques globaux ». Pour une géométrie d'aéronef donnée, les coefficients aérodynamiques globaux s'expriment :

$$\begin{aligned} C_i &= C_i(M, Re, \alpha, \beta, p, q, r, \delta_l, \delta_m, \delta_n, \dot{\alpha}, \dot{\beta}) \\ &= C_i(M, Re, \alpha, \beta, p_a, q_a, r_a, \delta_l, \delta_m, \delta_n, \dot{\alpha}, \dot{\beta}) \end{aligned}$$

où :

- $i \in \{x, y, z, l, m, n\}$
- M est le nombre de Mach
- Re est le nombre de Reynolds

Ces coefficients dépendent des caractéristiques de l'écoulement (traduites par M et Re), de la direction par laquelle il attaque la géométrie (traduite par α et β), de la vitesse de rotation de l'aéronef qui modifie la vitesse relative des molécules d'air attaquant la surface (traduite par p, q et r ou p_a, q_a et r_a), ainsi que des modifications de la géométrie induite par le braquage des gouvernes (traduites par δ_l, δ_m et δ_i). Les coefficients aérodynamiques peuvent également dépendre d'effets « instationnaires » (traduits par $\dot{\alpha}$ et $\dot{\beta}$), s'expliquant par des effets de retard entre les actions du fluide sur les différentes surfaces portantes, créant une dépendance à l'historique de variation des angles d'incidence et de dérapage. Ces effets instationnaires sont présents lorsque la dynamique du mouvement de l'aéronef est rapide comparée à celle de l'écoulement. Dans l'étude, ces effets instationnaires sont négligés (Cf. section 1.1.2).

Chaque coefficient aérodynamique global est donc une fonction non linéaire dépendant d'un grand nombre de variables. Pour un aéronef donné, l'évaluation de ces coefficients s'appuie sur des données, classiquement sur des données expérimentales. On peut néanmoins observer l'émergence de données venant de simulations CFD [3] pour étudier plus en détail l'effet du nombre de Reynolds dans le domaine post-décroché par exemple.

Actuellement, aucun essai expérimental ne permet d'étudier toutes ces dépendances à la fois et l'exploration d'un espace à 12 dimensions serait très coûteux en terme de temps et en terme financier. Une structure simplifiée est alors sélectionnée pour modéliser ces dépendances complexes. De façon classique, les données expérimentales sont obtenues par des essais réalisés en soufflerie. Dans les travaux, des données provenant d'essais sur une maquette d'avion léger montée sur une balance rotative située dans la soufflerie verticale du centre Onera Lille sont utilisées.

1.2.2 Structure additive

De façon classique en dynamique du vol, le choix se porte sur des modèles simplificateurs ayant une structure additive. Chaque coefficient aérodynamique global est modélisé par une somme de sous-fonctions non-linéaires dépendant d'un sous-ensemble des 12 variables identifiées ; chaque sous-fonction représentant la dépendance d'une variable ou de l'interaction entre plusieurs de ces variables. Sur l'ensemble de ces sous-fonctions, de nombreuses sont négligées, ce qui a un grand intérêt d'un point de vue théorique et expérimental. En effet, les sous-fonctions correspondant à des interactions négligeables ou qui sont inaccessibles avec les moyens expérimentaux sont ignorées. Au lieu de devoir balayer un espace à 12 dimensions afin d'ériger une unique base de données pour chaque coefficient, une base de données est obtenue pour les sous-fonctions non négligées de chaque coefficient. Comme les sous-fonctions dépendent de moins de variables, l'espace à balayer pour obtenir des points expérimentaux a une dimension inférieure, ce qui rend les expériences possibles avec les moyens actuels, au détriment de l'exhaustivité qu'aurait permis

une exploration de l'espace complet de tous les paramètres.

La logique de ces modèles additifs est illustrée sur une fonction théorique f . Soit une fonction théorique quelconque $f = f(x_1, x_2, x_3)$, dépendant de 3 variables x_1, x_2 et x_3 . f peut s'écrire, pour un triplet de constantes donné $(K_1, K_2, K_3) \in \mathbb{R}^3$:

$$\begin{aligned} f(x_1, x_2, x_3) &= f_0 + f_{x_1}(x_1) + f_{x_2}(x_2) + f_{x_3}(x_3) \\ &+ f_{x_1x_2}(x_1, x_2) + f_{x_1x_3}(x_1, x_3) + f_{x_2x_3}(x_2, x_3) \\ &+ f_{x_1x_2x_3}(x_1, x_2, x_3) \end{aligned} \quad (76)$$

Chaque sous-fonction s'annulant lorsque l'une de ses variables est égale à sa constante associée.

Exemple 1. $f_{x_1}(x_1 = K_1) = 0$, ou encore $f_{x_1, x_2}(x_1 = K_1, x_2) = f_{x_1, x_2}(x_1, x_2 = K_2) = f_{x_1, x_2}(x_1 = K_1, x_2 = K_2) = 0$.

Les sous-fonctions ne possédant qu'une variable en indice traduisent la dépendance au premier ordre de f par rapport à cette variable (l'effet de cette variable seule). Les sous-fonctions possédant plusieurs variables en indice traduisent la dépendance de f par rapport à l'interaction entre ce groupe de variables. Cette décomposition ressemble à la décomposition ANOVA (152) (présentée à la section 2.4.3.1), mais il s'agit d'une autre structure additive.

En adoptant ces propriétés, chaque sous-fonction a une forme bien précise dépendante de f et des sous-fonctions d'ordre inférieur.

Exemple 2.

$$f_0 = f(K_1, K_2, K_3) \quad (77)$$

$$f_{x_1}(x_1) = f(x_1, K_2, K_3) - f_0 \quad (78)$$

$$f_{x_1, x_2}(x_1, x_2) = f(x_1, x_2, K_3) - f_0 - f_{x_1}(x_1) - f_{x_2}(x_2) \quad (79)$$

Lorsque f dépend de 3 variables, elle se décompose en 8 sous-fonctions. Lorsque la fonction dépend de n variables, elle se décompose en 2^n sous-fonctions.

Chaque coefficient aérodynamique, qui dépend de 12 variables, se décompose alors en 4096 sous-fonctions non-linéaires. Mais de nombreuses de ces sous-fonctions sont négligées. En effet, comme il a été dit ci-dessus, l'accès à certaines dépendances est soit impossible soit trop coûteux expérimentalement ; ou certaines interactions sont négligeables. Cela se traduit alors par le fait d'ignorer des sous-fonctions dans le modèle additif. Dans les modèles de coefficients sélectionnés dans ces travaux, l'hypothèse de quasi-stationnarité ($\dot{\alpha} = 0$ et $\dot{\beta} = 0$) est considérée et la dépendance au nombre de Reynolds ainsi qu'au nombre de Mach n'est pas considérée. L'utilisation de données provenant d'une balance rotative sans inclinaison et l'utilisation d'aucune donnée en oscillations forcées (Cf. section 1.2.3), induit également le fait que les dépendances en q_a et en r_a ne soient pas accessibles. De plus, de nombreuses sous-fonctions traduisant des interactions d'ordre supérieur ou égale

à 3 (interactions entre 3 variables) sont négligées. Ces hypothèses réduisent très fortement le nombre de sous-fonctions composant le modèle de représentation. Ce modèle rend donc les expériences possibles, mais il devient lacunaire.

Une base de données est alors réalisée pour chaque sous-fonction non négligée de chaque coefficient aérodynamique. En reprenant l'exemple théorique, pour établir la base de donnée de f_{x_1, x_2} par exemple, il faut fixer x_3 à K_3 et faire varier x_1 et x_2 afin de récupérer les mesures expérimentales de f sur ces points. Les valeurs de f_0 , de f_{x_1} et de f_{x_2} mesurées au préalable en ces mêmes points (x_1, x_2) sont ensuite soustraites à ces valeurs afin d'obtenir la table souhaitée. Expérimentalement, seules les sous-fonctions auxquelles les expériences offrent l'accès et qui ont un effet non noyé dans le bruit sont mesurées.

Pour illustrer l'obtention des valeurs pour une table de données d'une sous-fonction d'un coefficient aérodynamique, voici l'exemple de $\Delta C_{i, p_a}(\alpha, \beta, p_a)$, la sous-fonction du coefficient C_i traduisant les effets des interactions d'un p_a non nul avec α et β :

$$\begin{aligned} \Delta C_{i, p_a}(\alpha, \beta, p_a) &= C_i(Ma = cst, Re = cst, \alpha, \beta, p_a, q_a = 0, r_a = 0, \delta_l = 0, \delta_m = 0, \delta_n = 0, \dot{\alpha} = 0, \dot{\beta} = 0) \\ &- C_{i_0}(\alpha, \beta) \end{aligned} \quad (80)$$

Les sous-fonctions d'un coefficient aérodynamique autre que C_{i_0} sont notées avec un Δ car elles sont interprétées comme des incréments à la sous-fonction traduisant les effets de α et de β seuls. L'obtention expérimentale des données relatives à $\Delta C_{i, p_a}$ suit celle de $C_{i_0}(\alpha, \beta)$ (correspondant à des mesures à $p_a = 0$). Les mesures sont réalisées aux mêmes valeurs de α et de β que cette dernière, en faisant cette fois-ci varier p_a . Les mesures de $C_{i_0}(\alpha, \beta)$ sont alors soustraites à celles obtenues avec un p_a non nul afin d'obtenir la table de données pour $\Delta C_{i, p_a}(\alpha, \beta, p_a)$. Pour être rigoureusement équivalent à l'exemple théorique de la fonction f pris ci-dessus, $\Delta C_{i, p_a}(\alpha, \beta, p_a)$ est une fonction à 3 variables ayant en indices α , β et p_a et $C_{i_0}(\alpha, \beta)$ est la somme de 4 sous fonctions : une constante (pour $\alpha = \beta = 0$), les effets d' α seul, les effets de β seul et les interactions entre α et β . De façon pratique, une unique table de données est réalisée pour la sous-fonction $C_{i_0}(\alpha, \beta)$.

Le modèle additif rend donc les campagnes expérimentales possibles concrètement, mais ce modèle est lacunaire. De plus, les tables de données souffrent de l'additivité des incertitudes de mesures, étant obtenues par soustraction de tables de données expérimentales entre elles. Par exemple, les incertitudes sur les données de la table $\Delta C_{i, p_a}(\alpha, \beta, p_a)$ sont supérieures à celles sur les données de $C_{i_0}(\alpha, \beta)$, les incertitudes de mesures de deux campagnes s'additionnant.

De façon classique, dans l'étude du mouvement de l'aéronef dans le domaine de vol linéaire ou lors de la linéarisation autour d'un équilibre, seules les dépendances au premier ordre sont considérées et sont de plus supposées linéaires autour de l'équilibre considéré. Les coefficients aérodynamiques se décomposent alors comme une somme de fonctions linéaires où les coefficients directeurs devant chaque variable sont parfois appelés « dérivées ».

1.2.3 Obtention des données expérimentales

Pour chaque sous-fonction ΔC_i , une table de données est érigée avec la méthode décrite à la section précédente, suite à des campagnes d'essais. Pour l'étude du mouvement de l'aéronef dans le domaine de vol post-décroché, des données issues de 3 types d'expériences sont utilisées de façon classique [33, 28, 44, 59, 3] : des données concernant les dépendances en α et β seuls obtenues à partir d'expériences en soufflerie (souvent appelées « static wind tunnel data »), des données obtenues à partir d'expériences en oscillations forcées (« dynamic forced oscillation data ») et des données obtenues à partir d'expériences sur balance rotative en soufflerie verticale (« dynamic rotary balance data »).

Ces expériences sont réalisées sur des maquettes à échelle réduite. On observe alors que les nombres de Reynolds des écoulements dans le cadre de l'expérience (de l'ordre de 10^5 - 10^6) sont inférieurs à ceux observés dans les écoulements autour des avions réels (de l'ordre de 10^7). Dans [3], des études numériques CFD sont réalisées sur une configuration partielle d'avion (le fuselage et les ailes, mais sans l'empennage), ce qui permet une première étude des coefficients aérodynamiques globaux à des nombres de Reynolds plus élevés et donc plus proches de ceux des écoulements autour de l'aéronef réel. Dans notre étude, seules des données expérimentales sont utilisées et les dépendances des coefficients aux nombres de Reynolds et au nombre de Mach sont alors négligées. De plus, l'hypothèse de quasi-stationarité est adoptée, comme il a été dit à la section 1.2.2.

1.2.3.1 Mesures

Les mesures des forces et des moments sont réalisées à l'aide d'une balance dynamométrique à 6 composantes. Ces mesures permettent, après traitement, d'obtenir la valeur des résultantes des forces aérodynamiques X , Y , Z et la valeur des moments aérodynamiques au point de mesure M : L_M , M_M et N_M .

1.2.3.2 Données statiques

Les forces et les moments sont mesurées alors que seuls l'angle d'incidence α et l'angle de dérapage β varient et que la maquette n'est pas en rotation. Les dépendances aux positions angulaires des gouvernes lorsque l'aéronef n'est pas en rotation sont également mesurées lors de ces expériences. Toutes ces dépendances sont souvent nommées dans la littérature comme étant les dépendances « statiques ». Ces mesures sont généralement réalisées en soufflerie horizontale, mais ont été réalisées dans le cas de la présente étude dans la soufflerie verticale sans rotation de la maquette.

1.2.3.3 Oscillations forcées

Cette expérience a pour but de déterminer les dépendances des forces et des moments vis à vis des composantes p , q et r de la vitesse de rotation de l'aéronef par rapport à la Terre. Pour ce faire, la maquette est fixée dans une soufflerie horizontale classique et est forcée d'osciller autour des axes Ox , Oy et Oz .

1.2.3.4 Balance rotative en soufflerie verticale

La balance rotative de la soufflerie verticale permet l'étude des vrilles car elle reproduit un mouvement qui présente des conditions proches de ces dernières. La vitesse de rota-



FIGURE 4 – Photographie de la maquette d’avion léger montée sur la balance rotative de l’Onera Lille. ©Onera

tion, qui correspond à p_a si l’axe de rotation de la balance est verticale (ce qui est le cas des campagnes d’essais ayant générées les données expérimentales utilisées), est fixée. Ce type d’expérience avec axe de rotation et axe vitesse colinéaires est nommé essais en rotation stationnaire, ou « coning motion » en Anglais.

Il est également possible d’avoir accès à d’autres dépendances (en q_a , en r_a , ou en $\dot{\alpha}$ et $\dot{\beta}$ par exemple) si l’axe de rotation de la balance est incliné par rapport à l’axe vent (qui est vertical dans les souffleries verticales). Ces essais sont alors appelés essais en rotation instationnaire ou « oscillatory coning motion ». Un dispositif présent sur la balance rotative du centre Onera Lille permet de réaliser l’inclinaison de l’axe mais dans le cadre des campagnes utilisées pour ces travaux de thèse, ces expériences ne sont pas réalisées.

1.2.4 Méthodes de blending

Comme il est mentionné dans [44], ni les données en oscillations forcées, ni les données issues de balance rotative ne permettent de représenter seules l’intégralité du mouvement complexe d’un départ en vrille. En effet, les pertes de contrôles engendrent des mouvements très complexes difficiles à représenter avec précision avec les montages expérimentaux actuels. Il est donc nécessaire d’utiliser les données venant des deux expériences. Les expériences en oscillations forcées permettant de mettre en évidence les dépendances des coefficients aux variables p , q et r et la balance rotative permettant de mettre en évidence la dépendance en p_a , il faut utiliser des techniques d’approximation afin d’utiliser les deux jeux de données en évitant de comptabiliser des effets de rotation deux fois. La thèse déjà citée [44] fait un état de l’art des méthodes de « blending » (qui signifie « mélange » en anglais). Cet état de l’art est rappelé dans cette section.

Ces techniques ont été utilisées avec succès pour créer des modèles de représentation pour les avions militaires [28, 24], et plus récemment pour les avions de transport (notamment avec le modèle GTM [44]).

1.2.4.1 Idée fondamentale des méthodes de blending

Les notations utilisées dans la thèse [44] sont reprises ici. Les méthodes de blending s'appuient sur la décomposition du vecteur rotation $\mathbf{\Omega}(\text{R}/\text{R}_T)$ en un vecteur $w_{ss}\mathbf{x}_a$, afin d'utiliser les données de la balance rotative et en un vecteur $\mathbf{\Omega}_{osc}$ afin d'utiliser les données en oscillations forcées

$$\mathbf{\Omega}(\text{R}/\text{R}_T) = w_{ss}\mathbf{x}_a + \mathbf{\Omega}_{osc} \quad (81)$$

avec

$$\mathbf{\Omega}_{osc} = p_{osc}\mathbf{x} + q_{osc}\mathbf{y} + r_{osc}\mathbf{z} \quad (82)$$

La structure additive permet d'utiliser les deux jeux de données. La sous-fonction dont les données sont issues de la balance rotative est $C_{i,w_{ss}}$ et les sous-fonctions dont les données sont issues des expériences en oscillations forcées sont $C_{i,p_{osc}}$, $C_{i,q_{osc}}$ et $C_{i,r_{osc}}$. Les méthodes diffèrent alors dans le choix de w_{ss} et des composantes de $\mathbf{\Omega}_{osc}$, qui serviront d'entrée pour évaluer les sous-fonctions $C_{i,w_{ss}}$, $C_{i,p_{osc}}$, $C_{i,q_{osc}}$ et $C_{i,r_{osc}}$.

1.2.4.2 Balance rotative seule

Cette méthode consiste à utiliser exclusivement les données issues de la balance rotative.

$$w_{ss} = p_a = p \cos \alpha \cos \beta + q \sin \beta + r \sin \alpha \cos \beta \quad (83)$$

$$\mathbf{\Omega}_{osc} = \mathbf{0} \quad (84)$$

$$p_{osc} = 0 \quad (85)$$

$$q_{osc} = 0 \quad (86)$$

$$r_{osc} = 0 \quad (87)$$

Il peut alors être conjecturé qu'avec cette méthode, les vrilles non oscillatoires établies sont retranscrites avec précision. Néanmoins, utiliser les données de la balance rotative seule présente des imperfections pour effectuer des simulations à 6 degrés de libertés, qui peuvent représenter des mouvements plus complexes comme les vrilles oscillantes.

1.2.4.3 Oscillations forcées seules

Cette méthode consiste à utiliser exclusivement les données issues d'expériences en oscillations forcées. Ainsi, tout effet d'interaction entre p , q et r sont ignorés

$$w_{ss} = 0 \quad (88)$$

$$\mathbf{\Omega}_{osc} = \mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T) \quad (89)$$

$$p_{osc} = p \quad (90)$$

$$q_{osc} = q \quad (91)$$

$$r_{osc} = r \quad (92)$$

Dans [44], l'auteur cite les deux articles [34, 37] déclarant que les oscillations forcées seules ne permettent pas de décrire les fortes dynamiques dans le domaine de vol post-décroché. La thèse [18] cite également des travaux allant dans ce sens. En effet, les mouvements complexes du domaine post-décroché s'éloignent des configurations des expériences en soufflerie. De plus, les taux de rotation rencontrés dans le domaine de vol post-décroché sont très importants et engendrent très souvent le fait d'être en extrapolation dans les tables de données issues d'oscillations forcées. La présence de données issues de balance rotative s'est alors avérée nécessaire pour décrire correctement les vrilles.

1.2.4.4 Méthode Direct Resolution Method (DIR)

La méthode DIR [7] consiste à prendre w_{ss} comme étant le roulis aérodynamique p_a et $\mathbf{\Omega}_{osc}$ comme étant la partie résiduelle du vecteur $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$

$$w_{ss} = p_a = p \cos \alpha \cos \beta + q \sin \beta + r \sin \alpha \cos \beta \quad (93)$$

$$\mathbf{\Omega}_{osc} = \mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T) - p_a \mathbf{x}_a = q_a \mathbf{y}_a + r_a \mathbf{z}_a \quad (94)$$

$$p_{osc} = p - p_a \cos \alpha \cos \beta = -q_a \cos \alpha \sin \beta - r_a \sin \alpha \quad (95)$$

$$q_{osc} = q - p_a \sin \beta = q_a \cos \beta \quad (96)$$

$$r_{osc} = r - p_a \sin \alpha \cos \beta = -q_a \sin \alpha \sin \beta + r_a \cos \alpha \quad (97)$$

Cette méthode peut présenter des problèmes car elle engendre souvent que p_{osc} et p n'aient pas le même signe, et de même pour r_{osc} et r . Les non-linéarités et l'asymétrie des données avec $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T)$ pour les manœuvres à hauts angles d'incidence peuvent alors engendrer des écarts conséquents.

1.2.4.5 Méthode DIR au premier ordre

Cette méthode est notamment utilisée dans [59] (sans néanmoins être explicitement dénommée telle quelle). Elle s'apparente à la méthode DIR mais fait appel à des hypothèses supplémentaires. Lors de vrilles, q_a et r_a sont généralement faibles par rapport à p_a . Cette méthode utilise alors les sous-fonctions $\Delta C_{i,q_a}$ et $\Delta C_{i,r_a}$ qui sont linéarisées autour de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T) = \mathbf{0}$. Ces dernières sont reliées à la linéarisation de $\Delta C_{i,p}$, $\Delta C_{i,q}$ et $\Delta C_{i,r}$ qui sont évaluées avec les données oscillations forcées.

Les coefficients directeurs des sous-fonctions linéarisés autour de $\mathbf{\Omega}(\mathbf{R}/\mathbf{R}_T) = \mathbf{0}$ (souvent dénommés « dérivées » dans la littérature, et dépendant de α et β) sont notés

$$\frac{\partial \Delta C_{i,X}}{\partial X}(\alpha, \beta, \boldsymbol{\Omega}(R/R_T) = \mathbf{0}) = C_i^X(\alpha, \beta) \quad (98)$$

Ainsi, les sous-fonctions linéarisées s'écrivent

$$\Delta C_{i,p}(\alpha, \beta, p) = C_i^p(\alpha, \beta)p \quad (99)$$

$$\Delta C_{i,q}(\alpha, \beta, q) = C_i^q(\alpha, \beta)q \quad (100)$$

$$\Delta C_{i,r}(\alpha, \beta, r) = C_i^r(\alpha, \beta)r \quad (101)$$

$$\Delta C_{i,q_a}(\alpha, \beta, q_a) = C_i^{q_a}(\alpha, \beta)q_a \quad (102)$$

$$\Delta C_{i,r_a}(\alpha, \beta, r_a) = C_i^{r_a}(\alpha, \beta)r_a \quad (103)$$

En appliquant la méthode DIR (en prenant $p = p_{osc}$, $q = q_{osc}$ et $r = r_{osc}$) et la linéarisation, nous obtenons

$$\Delta C_{i,p_{osc}}(\alpha, \beta, p_{osc}) = C_i^{p_{osc}}(\alpha, \beta)(-q_a \cos \alpha \sin \beta - r_a \sin \alpha) \quad (104)$$

$$\Delta C_{i,q_{osc}}(\alpha, \beta, q_{osc}) = C_i^{q_{osc}}(\alpha, \beta)q_a \cos \beta \quad (105)$$

$$\Delta C_{i,r_{osc}}(\alpha, \beta, r_{osc}) = C_i^{r_{osc}}(\alpha, \beta)(-q_a \sin \alpha \sin \beta + r_a \cos \alpha) \quad (106)$$

On identifie alors :

$$C_i^{q_a} = -C_i^{p_{osc}}(\alpha, \beta) \cos \alpha \sin \beta + C_i^{q_{osc}}(\alpha, \beta) \cos \beta - C_i^{r_{osc}}(\alpha, \beta) \sin \alpha \sin \beta \quad (107)$$

$$C_i^{r_a} = -C_i^{p_{osc}}(\alpha, \beta) \sin \alpha + C_i^{r_{osc}}(\alpha, \beta) \cos \alpha \quad (108)$$

1.2.4.6 Méthodes de Kalviste

Ces méthodes reposent sur la méthode proposée par Kalviste dans [33]. Elles existent sous trois formes : 3D-Kalviste (la méthode présentée par Kalviste dans son article), 2D-Kalviste et Hybride-Kalviste, la méthode 3D-Kalviste n'étant que rarement utilisée de nos jours. Ces méthodes consistent à réaliser une disjonction de cas pour déterminer les valeurs de w_{ss} et des composantes $\boldsymbol{\Omega}_{osc}$. Trois cas sont considérés. Soit $\boldsymbol{\Omega}_{xz} = p\mathbf{x} + r\mathbf{z}$, la projection de $\boldsymbol{\Omega}(R/R_T)$ dans le plan $(O, \mathbf{x}, \mathbf{z})$.

- *Cas 1* : le vecteur rotation a une composante w_{ss} selon \mathbf{x}_a et $\boldsymbol{\Omega}_{osc}$ a une composante en lacet et en tangage, si p et r ont le même signe et que :

$$- \theta_{xz} > \alpha \text{ si } p \text{ et } r \text{ sont positifs}$$

$$- 90^\circ < \theta_{xz} < 180^\circ + \alpha \text{ si } p \text{ et } r \text{ sont négatifs}$$

- *Cas 2* : le vecteur rotation a une composante w_{ss} selon \mathbf{x}_a et $\boldsymbol{\Omega}_{osc}$ a une composante en roulis et en tangage, si p et r ont le même signe et que :

TABLE 1 – Tableau résumant les valeurs des termes pris dans la méthode de blending Hybrid-Kalviste

	Cas	Cas 2	Cas 3
w_{ss}	$\frac{p}{\cos \alpha \cos \beta}$	$\frac{r}{\sin \alpha \cos \beta}$	0
p_{osc}	0	$p - w_{ss} \cos \alpha \cos \beta$	p
q_{osc}	$q - w_{ss} \sin \beta$	$q - w_{ss} \sin \beta$	q
r_{osc}	$r - w_{ss} \sin \alpha \cos \beta$	0	r

- $\theta_{xz} < \alpha$ si p et r sont positifs
- $180^\circ + \alpha < \theta_{xz} < 180^\circ$ si p et r sont négatifs

- *Cas 3* : Si le mouvement est non coordonné ($p \times r < 0$), la composante w_{ss} est prise comme étant égale à 0 et seules les données en oscillations forcées sont prises en compte.

La Table 1 résume les valeurs prises pour les composantes dans le cadre de la méthode Hybride-Kalviste et pour les 3 différents cas qu'elle introduit.

La méthode 2D-Kalviste reprend la même logique et les mêmes valeurs de la Table 1 mais avec $\beta = 0$.

1.2.4.7 Méthodes d'excès en roulis

Cette méthode [34] est en réalité le cas 2 de la méthode d'Hybride-Kalviste mais utilisée dans tous les cas de figure. La composante de lacet n'est jamais considérée.

1.2.4.8 Comparaison des méthodes de blending

Toutes ces méthodes de blending reposent sur des hypothèses simplificatrices. Dans la thèse [44], l'auteur compare les différentes méthodes de blending pour le modèle d'avion de ligne GTM. Selon ses conclusions, les méthodes Hybrid-Kalviste et DIR semblent être les plus proches des données de vol libre obtenues sur la maquette.

A la fin de la thèse [44], l'auteur propose une nouvelle méthode de blending prenant plus en compte les non-linéarités. Il propose de considérer la sous-fonction $\Delta C_{l,p_{osc}}$ par exemple comme étant :

$$\Delta C_{l,p_{osc}} = \Delta C_{l,p_{osc}}(\alpha, p) - \Delta C_{l,p_{osc}}(\alpha, w_{ss} \cos \alpha \cos \beta) \quad (109)$$

et non plus comme étant dans la méthode DIR :

$$\Delta C_{l,p_{osc}} = \Delta C_{l,p_{osc}}(\alpha, p - w_{ss} \cos \alpha \cos \beta) \quad (110)$$

Cette nouvelle méthode permet alors de prendre en compte les non-linéarités locales. Néanmoins, le problème d'extrapolation dans les bases de données obtenues en oscillations forcées est retrouvé.

1.2.5 Modèle de représentation sélectionné

1.2.5.1 Structure de modèle adoptée

Le modèle considéré est un modèle dont les données ont été obtenues à partir d'expériences en soufflerie réalisées sur une maquette d'avion léger disposée sur une balance rotative dans la soufflerie verticale SV4 du centre Onera Lille. Seules des données statiques et des données issues de la balance rotative sont considérées en premier lieu. Ainsi, la méthode « Balance rotative seule » est adoptée. Voici la structure adoptée :

$$C_x = C_{x0}(\alpha) \quad (111)$$

$$C_y = C_{y0}(\alpha, \beta) + \Delta C_{y,\delta_l}(\alpha, \delta_l) + \Delta C_{y,\delta_n}(\alpha, \delta_n) + \Delta C_{y,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (112)$$

$$C_z = C_{z0}(\alpha, \beta) + \Delta C_{z,\delta_l}(\alpha, \delta_l) + \Delta C_{z,\delta_m}(\alpha, \delta_m) + \Delta C_{z,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (113)$$

$$C_l = C_{l0}(\alpha, \beta) + \Delta C_{l,\delta_l}(\alpha, \delta_l) + \Delta C_{l,\delta_n}(\alpha, \delta_n) + \Delta C_{l,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (114)$$

$$C_m = C_{m0}(\alpha, \beta) + \Delta C_{m,\delta_l}(\alpha, \delta_l) + \Delta C_{m,\delta_m}(\alpha, \delta_m) + \Delta C_{m,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (115)$$

$$C_n = C_{n0}(\alpha, \beta) + \Delta C_{n,\delta_l}(\alpha, \delta_l) + \Delta C_{n,\delta_n}(\alpha, \delta_n) + \Delta C_{n,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (116)$$

où $\hat{p}_a = \frac{c}{V} p_a$ est le roulis aérodynamique adimensionné.

1.2.5.2 Données expérimentales concernant le modèle étudié

Les données sont issues d'expériences réalisées dans la soufflerie vertical SV4 du Centre Onera Lille sur une maquette d'avion léger générique à l'échelle 1/5. La vitesse d'écoulement pour ces campagnes expérimentales est de $V = 50m/s$ ($M = 0,15$) et le nombre de Reynolds est de $0,83 \times 10^6$.

En plus des caractéristiques dues à la structure de modèle adoptée, les données possèdent des spécificités supplémentaires. En effet, les campagnes d'essais ont été réalisées avec un « montage dorsal » (comme illustré par la figure 4) et un « montage ventral arrière », selon le mode de fixation de la maquette. Pour certaines sous-fonctions, 2 tables de données existent donc sur des pages de variables d'entrées qui s'entrecoupent.

La Table 2 résume l'ensemble des paramètres des tables de données obtenues à partir des expériences en soufflerie. Pour le modèle « ventrale arrière », une hypothèse de symétrie en β est réalisée. Seules les expériences à $\beta > 0$ sont réalisées et l'utilisation de l'hypothèse de symétrie est utilisée pour déterminer les valeurs des coefficients pour $\beta < 0$.

1.2.5.3 Des bases de données aux coefficients

Une fois les bases de données acquises, le choix de la méthode pour les utiliser lors de l'intégration numérique doit être fait. Pour le modèle sélectionné, en premier lieu, seules les données obtenues avec le montage dorsal sont utilisées pour les sous-fonctions statiques ($C_{i,0}$, C_{i,δ_l} , C_{i,δ_m} et C_{i,δ_n}). Pour la sous-fonction C_{i,\hat{p}_a} , les données du montage dorsal sont utilisées si $\alpha > 40^\circ$ et les données du montage ventral sont utilisées sinon. Pour chaque sous-fonction, une interpolation linéaire est réalisée dans la base de donnée correspondante. Une extrapolation linéaire pour les variables d'état est adoptée dans

	Montage dorsal				Montage ventral arrière			
	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\delta(^{\circ})$	\hat{p}_a	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\delta(^{\circ})$	\hat{p}_a
$C_{i,0}$	[5, 80]	[-10, 10]			[10, 40]	[0, 30]		
C_{i,δ_l}	[5, 80]		[-1, 1]		[10, 40]		[-1, 1]	
C_{i,δ_m}	[5, 80]		[-20, 5]		[10, 40]		[-20, 5]	
C_{i,δ_n}	[5, 80]		[-25, 25]		[10, 40]		[-25, 25]	
C_{i,\hat{p}_a}	[40, 80]	[-5, 5]		[-0.087, 0.087]	[0, 40]	[0, 15]		[-0.087, 0.087]

TABLE 2 – Intervalles de variations des variables d'état et de contrôles des tables de données concernant le modèle d'avion léger adopté obtenues à partir d'expériences en soufflerie dans la soufflerie verticale SV4 du centre Onera Lille.

une moindre mesure. Cette extrapolation prend en compte la pente définie par les deux points à l'extrémité (les deux premiers ou les deux derniers) dans ces bases. À partir d'une certaine valeur, la variable est mise à saturation. La Table 3 donne les intervalles d'interpolation/extrapolation. En dehors de ces plages, la valeurs est saturée au minimum ou au maximum de ces dernières.

	Montage dorsal			Montage ventral arrière		
	α	β	\hat{p}_a	α	β	\hat{p}_a
$C_{i,0}$	[0°, 90°]	[-30°, 30°]				
$C_{i,\delta}$	[0°, 90°]	[-30°, 30°]				
C_{i,\hat{p}_a}	[40°, 80°]	[-10°, 10°]	[-0.15, 0.15]	[0°, 40°]	[-15°, 15°]	[-0.15, 0.15]

TABLE 3 – Intervalles où l'interpolation/extrapolation a lieu. En dehors de ces plages, la valeur est saturée au minimum ou au maximum de ces plages.

1.3 Théorie des bifurcations

Dans cette section, les notions de point d'équilibre, de cycle limite, de stabilité, de bifurcation et de domaine d'attraction sont introduites pour un système dynamique théorique donné $\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \mathbf{U})$ avec \mathbf{U} fixé. \mathbf{f} est supposée à minima Lipschitz pour qu'à une condition initiale donnée ne correspond qu'une seule fonction temporelle $\mathbf{X}(t)$. La définition de certaines notions sont illustrées à l'aide d'un système dynamique simple et le lien avec le système de dynamique du vol est également établi. Pour l'étude des comportements de l'aéronef en tant qu'état d'équilibre d'un système dynamique, le système à 8 variables (Cf. section 1.1.5.6) est très souvent celui sélectionné. Les réductions à 5 et à 3 variables d'état peuvent également être utilisées.

1.3.1 Notions d'équilibre, de stabilité et de domaine d'attraction

Le livre [68] traite en profondeur de l'étude mathématique des systèmes dynamiques, des notions d'états d'équilibre, de stabilité et de domaines d'attraction.

1.3.1.1 Espace des phases

L'espace des phases est un espace abstrait où sont représentés tous les vecteurs d'état \mathbf{X} possibles. Une évolution possible des valeurs de ce vecteur $\mathbf{X}(t)$ au cours du temps est appelée une trajectoire de phase.

1.3.1.2 Système dynamique exemple : le pendule simple avec amortissement

L'ensemble de la section 1.3 est illustrée par le système dynamique correspondant à la dynamique du pendule simple avec un fil rigide où la masse du fil est négligée. Le pendule est constitué d'une « masse » de masse m et d'un fil rigide inextensible de longueur l de masse négligeable. Le fil fait un angle θ avec la verticale descendante, comme illustré à la figure 5. Le pendule est soumis à son poids \mathbf{P} , à la tension du fil \mathbf{T} ainsi qu'à une force de frottement fluide, simplifiée par $\mathbf{F} = -f\mathbf{v}$.

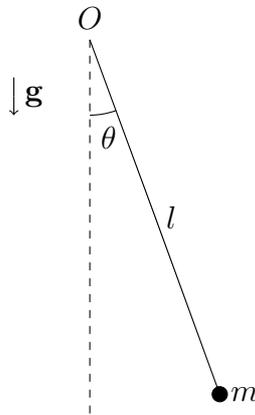


FIGURE 5 – Figure représentant le pendule simple avec amortissement étudié pour illustrer les définitions de cette sous-section.

L'équation du mouvement est alors :

$$\ddot{\theta} = -\frac{g}{l} \sin \theta - \frac{f}{m} \dot{\theta} \quad (117)$$

Les variables d'état sélectionnées sont θ et $\dot{\theta}$. L'espace des phases est alors de dimension 2. Ainsi, le système dynamique s'écrit

$$\dot{\mathbf{X}} = \begin{pmatrix} \dot{\theta} \\ \ddot{\theta} \end{pmatrix} = \mathbf{f}(\mathbf{X}) = \begin{pmatrix} \dot{\theta} \\ -\frac{g}{l} \sin \theta - \frac{f}{m} \dot{\theta} \end{pmatrix} \quad (118)$$

1.3.1.3 État d'équilibre

Pour un système dynamique $\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \mathbf{U})$ et pour un vecteur de contrôle \mathbf{U} donné, un état asymptotique (aussi appelé état d'équilibre) est un domaine de l'espace de phase tel que lorsqu'un état s'y trouve à un instant t , il y reste pour tout instant qui lui succède. Il existe deux types d'états asymptotiques habituellement rencontrés : les points d'équilibre et les orbites périodiques (aussi appelés cycles limites).

Un point d'équilibre est un point \mathbf{X}_e de l'espace de phase tel que $\mathbf{f}(\mathbf{X}_e, \mathbf{U}) = 0$.

Une orbite périodique est une trajectoire qui obéit à $\mathbf{X}_e(t) = \mathbf{X}_e(t + T)$ où T est la période de l'orbite (inconnue si le système est autonome). Pour étudier les orbites périodiques, il est possible d'utiliser la notion de section de Poincaré, un hyperplan bien choisi dans l'espace de phase. Le cycle limite vient alors couper cet hyperplan à intervalle de temps régulier de façon identique et le cycle a alors pour équivalent un point fixe dans le système dynamique discrétisé avec comme pas de temps T . Les notions concernant les cycles limites ne sont entrevues que brièvement dans le manuscrit.

Exemple 3. D'après (118), le système dynamique du pendule possède une infinité de points d'équilibres de la forme

$$\mathbf{X}_e = \begin{pmatrix} k\pi \\ 0 \end{pmatrix} \quad (119)$$

avec $k \in \mathbb{Z}$. Seuls les points d'équilibres $\mathbf{X}_{e1} = (0, 0)^T$ et $\mathbf{X}_{e2} = (\pi, 0)^T$ sont considérés dans l'exemple.

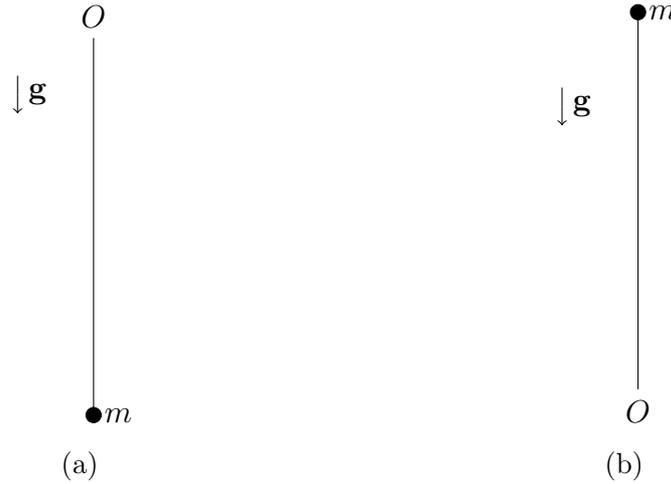


FIGURE 6 – Situations correspondant aux 2 points d'équilibre du pendule considérés. (a) $\mathbf{X}_{e1} = (0, 0)^T$ (b) $\mathbf{X}_{e2} = (\pi, 0)^T$

Pour le système dynamique du vol à 8 variables, un état d'équilibre correspond à un état où V , α , β , p , q , r , Θ et Φ sont constants. À l'inverse du pendule, un état d'équilibre n'est pas un état où l'aéronef n'est plus en mouvement. Dans le domaine de vol non décroché, un état d'équilibre correspond à un état de vol normal et non dangereux. Dans le domaine de vol décroché, un état d'équilibre est un comportement qui a de très grande chances d'être dangereux. Parmi ces comportements dangereux, il existe un comportement correspondant à une trajectoire hélicoïdale qui dirige l'avion vers le sol : une vrille.

1.3.1.4 Stabilité

Un point d'équilibre \mathbf{X}_e est stable au sens de Lyapunov (localement stable) si pour tout $\epsilon > 0$ et pour tout $t'_0 > t_0$, il existe $\delta > 0$ tel que

$$\|\mathbf{X}(t'_0) - \mathbf{X}_e(t'_0)\| < \delta \Rightarrow \|\mathbf{X}(t) - \mathbf{X}_e(t)\| < \epsilon \quad (120)$$

Un point d'équilibre est asymptotiquement stable s'il est stable au sens de Lyapunov et si

$$\lim_{t \rightarrow \infty} \|\mathbf{X}(t) - \mathbf{X}_e(t)\| = 0$$

Physiquement, un équilibre est asymptotiquement stable si, lorsque l'on écarte un peu le système de son équilibre, il y revient. Si l'équilibre est instable, lorsque le système est écarté de ce dernier, l'état évolue vers un équilibre stable ou ne trouve plus aucun équilibre. Pour étudier la stabilité d'un point d'équilibre d'un système dynamique, le système linéarisé autour de ce point est étudié. Soit $\mathbf{Z} = \mathbf{X} - \mathbf{X}_e$, le système linéarisé est

$$\dot{\mathbf{Z}} = \mathbf{J}\mathbf{Z} \quad (121)$$

où $\mathbf{J} = \left(\frac{\partial f_i}{\partial X_j}\right)_{\mathbf{X}=\mathbf{X}_e}$ est la jacobienne du système dynamique évaluée en $\mathbf{X} = \mathbf{X}_e$. Les valeurs propres de \mathbf{J} sont alors étudiées.

- Si la matrice est Hurwitz, c'est-à-dire si toutes ses valeurs propres ont une partie réelle strictement négative, alors le point d'équilibre est asymptotiquement stable.
- Si une des valeurs propres a une partie réelle nulle et que toutes les autres sont à partie réelle strictement négative, l'étude du système linéarisé ne suffit pas pour déterminer la stabilité du point d'équilibre.
- Si au moins l'une des valeurs propres de la jacobienne a une partie réelle positive, le point d'équilibre est instable.

Exemple 4. Au point d'équilibre $\mathbf{X}_{e1} = (0, 0)^T$ du pendule, la jacobienne vaut

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 1 \\ -\frac{g}{l} & -\frac{f}{m} \end{pmatrix} \quad (122)$$

qui possède des valeurs propres à partie réelle négative. L'équilibre est asymptotiquement stable.

Au point d'équilibre $\mathbf{X}_{e2} = (\pi, 0)^T$ du pendule, la jacobienne vaut

$$\mathbf{J}_2 = \begin{pmatrix} 0 & 1 \\ \frac{g}{l} & -\frac{f}{m} \end{pmatrix} \quad (123)$$

qui possède une valeur propre réelle positive. L'équilibre est instable.

En ce qui concerne le système dynamique du vol, des études similaires peuvent être réalisées pour étudier la stabilité de ses différents points d'équilibre. Un équilibre correspondant à une vrille stable représente alors un comportement dangereux, cette vrille tendant à perdurer.

Pour le cas des systèmes dynamiques discrets (et donc l'étude des cycles limites des systèmes continus dans leur section de Poincaré), une étude similaire existe. Les multiplicateurs de Floquet doivent se situer à l'intérieur du cercle unité pour que l'équilibre discret ou le cycle limite soit stable.

1.3.1.5 Domaines d'attraction

Un domaine d'attraction (aussi appelé région d'attraction) $A_{\mathbf{X}_e}$, correspondant à un point équilibre asymptotiquement stable \mathbf{X}_e est le sous-ensemble de l'espace des phases suivant :

$$A_{\mathbf{X}_e} = \{\mathbf{X}_0 \in \Omega : \text{si } \mathbf{X}(t = 0) = \mathbf{X}_0, \text{ alors } \lim_{t \rightarrow \infty} \mathbf{X}(t) = \mathbf{X}_e\} \quad (124)$$

L'obtention des domaines d'attractions exacts est un problème difficile en soi, qui reste non résolu pour des systèmes dynamiques non linéaires complexes. En dynamique du vol, la taille du domaine d'attraction associé à un état d'équilibre correspondant à une vrille semble représenter une bonne métrique du danger de cette dernière. Plus le domaine d'attraction associé à une vrille stable est grand, et plus il y a de chances que l'aéronef entre dans cet état de vrille pour un contrôle donné. C'est dans cette optique que les études sur les domaines d'attraction en dynamique du vol [59] ont été réalisées. Pour ces études, le système à 5 variables est utilisé pour des raisons de simplicité. De plus, pour l'utilisation de méthodes modernes permettant d'avoir des sous-ensembles de l'espace des phases sous la forme de deux ellipsoïdes encadrant $A_{\mathbf{X}_e}$, le système doit être approché sous la forme de polynômes. Autour des équilibres asymptotiquement stables étudiés, le modèle aérodynamique, ainsi que les termes trigonométriques des équations sont approchés par des polynômes.

1.3.2 Bifurcations

1.3.2.1 Définition de la bifurcation

Un point d'équilibre existe pour un vecteur de contrôle \mathbf{U} donné. Il est prouvé, d'après le théorème des fonctions implicites, qu'en faisant varier un paramètre de contrôle, un autre point d'équilibre existe au voisinage du premier dans l'espace des phases augmenté d'une dimension prenant en compte la valeur du paramètre de contrôle qui varie. Ce nouveau point d'équilibre est situé au voisinage du premier, mais peut avoir une autre stabilité. Il y a bifurcation si la variation d'un paramètre de contrôle engendre le changement de stabilité d'au moins un des états asymptotiques existants. Physiquement, une bifurcation peut se traduire par une légère modification d'un paramètre de contrôle engendrant une modification conséquente du comportement du système. En effet, si l'état actuel du système se trouve être un point d'équilibre stable et que le paramètre est modifié légèrement, mais que le nouvel équilibre proche du premier reste stable, l'état du système converge vers ce nouvel équilibre. Les modifications sur le système sont donc légers. Mais si le nouvel équilibre est instable, l'état du système va s'écarter de cet équilibre pour

potentiellement converger vers un autre équilibre, qui lui, est stable, ou va diverger. Une modification conséquente de l'état d'un système est alors observée pour une modification légère d'un paramètre de contrôle. Mathématiquement, cela se traduit par la traversée de l'axe imaginaire d'une ou plusieurs valeurs propres de la jacobienne évalué au point d'équilibre pour les systèmes continus linéarisés ou de la traversée du cercle unité d'un ou plusieurs multiplicateurs pour les systèmes discrets (et donc pour les cycles limites étudiés dans leur plan de Poincaré).

Il existe de nombreuses bifurcations, comme :

- La bifurcation « col-noeud » (ou « saddle-node » ou « fold » en anglais), qui traduit la co-existence d'un équilibre stable et d'un équilibre instable avant la bifurcation, puis plus d'aucun équilibre après. Dans le diagramme de bifurcation (Cf. section 1.3.2.2), un repliement de la surface d'équilibre est observé.
- La bifurcation de Hopf, où un point d'équilibre se déstabilise pour engendrer l'apparition d'un cycle limite.

La thèse [18] utilisant la théorie des bifurcations en dynamique du vol traite des bifurcations avec plus de détails.

Dans l'exemple du pendule amorti, si l'utilisateur a un contrôle sur $\frac{f}{m}$ et que cette valeur est mise à 0 (cas non amorti), le pendule oscille autour de \mathbf{X}_{e1} . La partie réelle des valeurs propres de la jacobienne J_1 croisent l'axe imaginaire (bifurcation de Hopf). Un cycle limite est observé autour du point $(0, 0)$, le pendule oscille à l'infini autour de l'équilibre à $\theta = 0$ et $\dot{\theta} = 0$.

1.3.2.2 Diagramme de bifurcation

Un diagramme de bifurcation est un diagramme présentant en abscisse la valeur d'un paramètre de contrôle et en ordonnée la valeur d'une variable d'état (un diagramme peut alors être tracé pour chaque variable d'état). Sur ce graphique est tracé l'ensemble des points d'équilibre du système, et ceux pour un ensemble de valeurs d'un paramètre de contrôle (les autres étant fixés). Une convention est adoptée pour représenter les équilibres stables et les équilibres instables. L'algorithme de continuation, basé sur le théorème implicite, est très utilisé parmi les méthodes numériques pour tracer les diagrammes de bifurcation.

La théorie de bifurcations a été utilisée la première fois dans le domaine de la dynamique du vol dans [13]. Cette théorie a ensuite été très souvent utilisée pour de nombreuses études en dynamique du vol, qu'ils s'agissent d'avions de combat ou d'avions de ligne [18, 24, 28, 25, 23, 17, 59]. Les diagrammes de bifurcation ont l'avantage de présenter une cartographie des états asymptotiques d'un aéronef pour un paramètre de contrôle qui varie.

La figure 7 montre un ensemble de diagrammes de bifurcation pour le modèle d'avion

léger adopté, obtenu à l'aide de l'algorithme de continuation Matcont [16]. Sur la figure 7, aucun équilibre stable n'apparaît dans le domaine post-décroché. Dans l'article [17], utilisant le même modèle aérodynamique en premier lieu, la cause de ce phénomène est attribuée à la non prise en compte de termes dits « d'amortissement », termes du modèle aérodynamique dépendant de q_a et de r_a que ne peut pas identifier la balance rotative seule. L'ajout de ces termes d'amortissement, provenant de données en oscillations forcées selon la méthode de blending 1.2.4.5, rend certains équilibres du domaine post-décroché stables. Ce phénomène est également étudié dans la présente thèse avec une autre méthode à la section 5.2.2.

1.4 Premières simulations et méthodes testées

Afin d'étudier la sensibilité et la robustesse des comportements dans le domaine post-décroché aux modèles de représentation, l'utilisation de la théorie des bifurcations a été considérée en premier lieu. L'article [17] propose une première exploration de la dépendance des comportements aux modèles en ajoutant des termes d'amortissement afin de rendre certains équilibres stables et en perturbant le modèle avec des incertitudes ΔC_i sur les coefficients aérodynamiques de moment. Les modifications induites sur les équilibres sont étudiées en comparant les différents diagrammes de bifurcations.

En s'inspirant de cet article, pour modéliser cette dépendance, en tout premier lieu dans les présents travaux, des coefficients de pondération ont été placés devant les sous-fonctions des modèles de coefficients aérodynamiques. La création de « diagrammes de sensibilité » a alors été testée. Ces diagrammes prennent pour base un diagramme de bifurcation selon un paramètre de gouverne réalisé de façon classique. Ce diagramme est ensuite discrétisé, et l'algorithme de continuation est appelé sur chaque point en étudiant la variation d'un autre paramètre de contrôle, l'un des coefficients de pondération ajouté. Une projection pour obtenir un diagramme 2D proche du diagramme original est alors réalisée. Mais ce type de figure ne donnant que des résultats qualitatifs et ne permettant que de faire varier un paramètre d'incertitude à la fois a vite été abandonné.

De façon générale, les travaux se sont éloignés de la théorie des bifurcations, pour éviter les limites suivantes :

- La théorie des bifurcation ne traite que des états asymptotiques lorsque les contrôles restent constants. Pour un paramètre de contrôle donné, les états asymptotiques sont les états atteints si le pilote ne modifie pas ces contrôles pendant un temps infini. Une étude lors des premiers instants pourrait s'avérer utile lors des études dans le domaine post-décroché.
- Les diagrammes de bifurcations ne permettent d'étudier que l'influence d'un paramètre à la fois. L'étude de l'influence de plusieurs variations de paramètres pourraient constituer un grand intérêt.

Les travaux se sont alors dirigés vers l'analyse de sensibilité et la quantification d'incer-

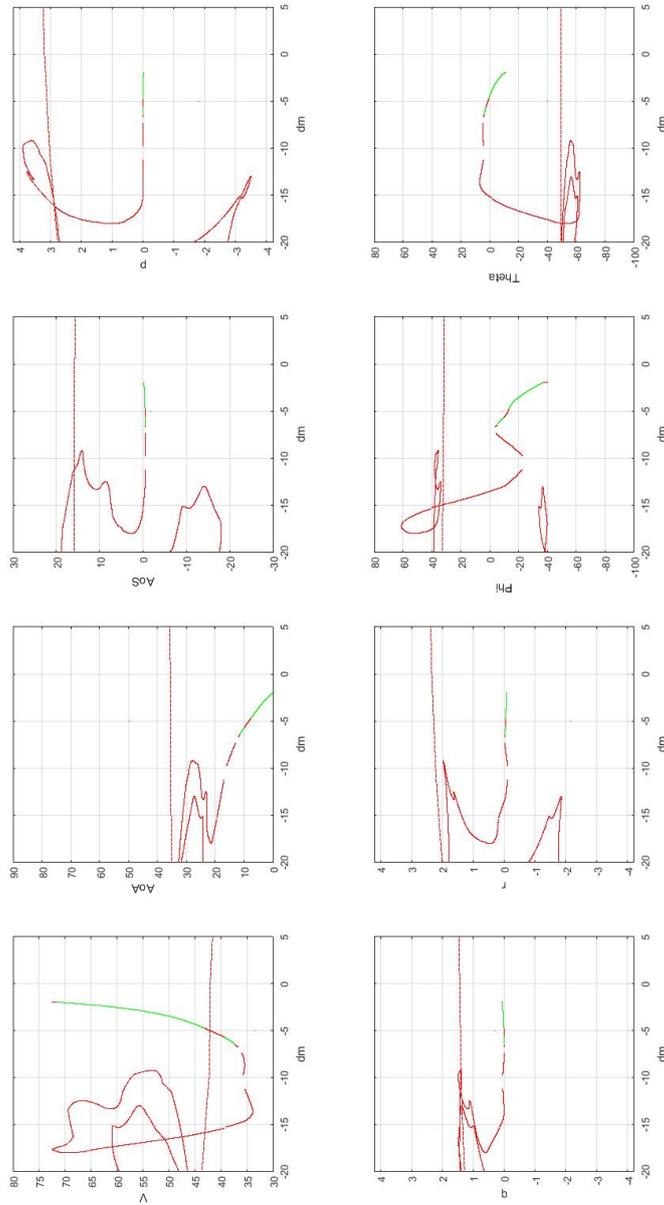


FIGURE 7 – Diagramme de bifurcation en gouverne de profondeur δ_m obtenu avec le système dynamique à 8 variables d'état et le modèle de représentation aérodynamique d'avion léger adopté. Les équilibres stables sont représentés en vert et les équilibres instables sont représentés en rouge.

titude. En premier lieu, l'impact des coefficients de pondération ajouté devant chaque sous-fonction sur les variables d'état a été étudié à l'aide de simulations de Monte-Carlo (Cf. section 2.4.3.5). En raison du nombre de comportements à étudier, l'utilisation d'un

métamodèle de polynômes du chaos (Polynomial Chaos - PC) a été envisagée (Cf. section 2.5). Néanmoins, lors de ces premières simulations, ces premiers métamodèles étaient très imprécis dus à la présence de discontinuités. Les recherches ont alors mené vers les métamodèles ME (Cf. section 2.6.4).

Néanmoins, afin d'utiliser moins de paramètres d'entrée, une autre modélisation des lacunes et des incertitudes est adoptée, plus proche de celle de l'article [17]. Comme indiqué au chapitre 5, les lacunes sont modélisées par des termes d'amortissement et les incertitudes par des termes ΔC_i ajoutés. Dans les nouvelles simulations, les irrégularités provoquées par les bifurcations des comportements étudiés n'altèrent pas la précision des métamodèles PC de façon significative. L'hypothèse de présence de discontinuités dans les premières simulations était le fruit d'une erreur d'analyse.

Chapitre 2

Analyse de sensibilité et quantification d'incertitudes : état de l'art

Ce chapitre traite du cadre d'étude de l'analyse de sensibilité et de la quantification d'incertitudes (Uncertainty Quantification - UQ). Il introduit des notations mathématiques, présente notamment les indices de Sobol, indices de l'analyse de sensibilité globale (Global Sensitivity Analysis - GSA), puis présente les métamodèles de polynômes du chaos (Polynomial Chaos - PC). Une fois les limites de ces métamodèles soulignées, un état de l'art sur les alternatives possibles est réalisé, avec un développement plus précis sur les métamodèles multi-éléments (Multi-Element - ME). Ce chapitre contient des extraits de [67] et [66] traduits et légèrement modifiés pour être intégrés aux travaux.

2.1 Introduction

Les simulations déterministes, où tous les paramètres sont connus et fixés, sont depuis longtemps au cœur des simulations numériques. Dans les simulations modernes, les études sur l'impact des incertitudes des systèmes gagnent de plus en plus d'importance. L'UQ a pour but d'identifier les paramètres incertains d'un système, de les modéliser comme des paramètres stochastiques, et d'étudier leur impact sur une quantité d'intérêt (Quantity of Interest - QOI) qui dépend de ces paramètres par une loi déterministe. D'un autre côté, l'analyse de sensibilité a pour but de relier l'incertitude commise sur la détermination d'une QOI avec les incertitudes commises sur les paramètres dont dépend cette QOI. Les deux domaines sont ainsi intimement liés et des méthodes faisant appel à des concepts provenant des deux champs d'étude coexistent.

2.2 QOI et paramètres d'entrée

2.2.1 Phénomène aléatoire

L'étude des incertitudes repose sur la notion de probabilités et de phénomène aléatoire. D'un point de vue mathématique, la modélisation des phénomènes aléatoires peut être

réalisée à l'aide d'un espace de probabilité, composé d'un triplet $(\Omega, \mathcal{A}, Pr)$.

2.2.1.1 L'Univers Ω

Ω est l'espace comportant toutes les issues possibles d'un phénomène aléatoire.

Exemple 5. *Pour un lancé de dé à 6 faces, $\Omega = \{1, 2, 3, 4, 5, 6\}$.*

2.2.1.2 L'espace des événements \mathcal{A}

Un sous-ensemble de Ω est appelé un « événement ».

Exemple 6. *Pour l'exemple du lancé de dé à 6 faces, l'événement : « Obtenir un nombre pair » est modélisé mathématiquement par un sous-ensemble A de Ω avec $A = \{2, 4, 6\}$*

Soit $\mathcal{P}(\Omega)$, l'ensemble des parties de Ω . $\mathcal{P}(\Omega)$ est l'ensemble contenant tous les sous-ensemble de Ω , soit tous les événements possibles. \mathcal{A} est une « tribu » (aussi appelé σ -algèbre) sur Ω , c'est-à-dire que c'est un sous-ensemble de $\mathcal{P}(\Omega)$ non vide, stable par complémentaire et stable par union dénombrable. \mathcal{A} est donc un ensemble contenant les événements que l'on souhaite étudier, ensemble possédant de plus les 3 propriétés décrites ci-dessus.

2.2.1.3 Mesure de probabilité Pr

Une mesure de probabilité est une application allant de \mathcal{A} dans $[0, 1]$ respectant

$$Pr(\Omega) = 1 \quad (125)$$

et soit (A_n) une suite d'événement de \mathcal{A} deux à deux incompatibles ($A_i \cap A_j = \emptyset, \forall i \neq j$),

$$Pr\left(\bigcup_n A_n\right) = \sum_n Pr(A_n) \quad (126)$$

Une telle fonction associe alors un événement de \mathcal{A} à une valeur entre $[0, 1]$: sa probabilité.

2.2.1.4 Espace de probabilité

Un espace de probabilité est alors un triplet $(\Omega, \mathcal{A}, Pr)$. Dans l'étude, il modélise tous les phénomènes aléatoires menant aux incertitudes sur les paramètres qui vont affecter la QOI. Néanmoins, cet espace est considéré de façon abstraite, aucun événement n'est mentionné explicitement. Seules sont explicitées les variables aléatoires et leur fonction densité de probabilité (probability density function - pdf).

2.2.2 Variable aléatoire réelle

2.2.2.1 Définition d'une variable aléatoire réelle

Une variable aléatoire réelle est une fonction allant de Ω dans \mathbb{R} . Cette fonction associe à chaque issue d'une expérience aléatoire une valeur réelle.

Exemple 7. Soit un jeu reposant sur un lancé de dé à 6 faces. Si le résultat est 1 ou 2, le joueur perd 2 points; si le résultat est 3 ou 4, le joueur ne remporte rien et si le résultat est 5 ou 6, le joueur remporte 2 points. La fonction qui associe aux issues le gain algébrique reçu est une variable aléatoire réelle.

La variable aléatoire prise en exemple est une variable aléatoire discrète. Dans les travaux, des variables aléatoires réelles continues sont considérées. La probabilité que la variable réelle soit située dans un intervalle correspond à la probabilité de l'événement impliquant que cette variable soit dans cet intervalle.

$$Pr(X \in [a, b]) = Pr\left(\{w \in \Omega : X(w) \in [a, b]\}\right) \quad (127)$$

Les paramètres incertains dont dépend la QOI, ainsi que la QOI elle-même, sont modélisés par des variables aléatoires réelles.

2.2.2.2 Fonction de répartition

La fonction de répartition \mathcal{C} d'une variable aléatoire réelle Y est la fonction qui associe à $y \in \mathbb{R}$

$$\mathcal{C}(y) = Pr(Y \leq y) \quad (128)$$

La fonction de répartition capture la distribution stochastique de Y . Cette fonction présente les propriétés suivantes :

- \mathcal{C} est croissante.
- $0 \leq \mathcal{C}(y) \leq 1$
- $\lim_{y \rightarrow -\infty} \mathcal{C}(y) = 0$
- $\lim_{y \rightarrow \infty} \mathcal{C}(y) = 1$

et respecte

$$Pr\left(Y \in [a, b]\right) = \mathcal{C}(b) - \mathcal{C}(a) \quad (129)$$

2.2.2.3 Fonction de densité de probabilité (pdf)

La pdf ρ d'une variable aléatoire réelle Y est la dérivée de sa fonction de répartition.

$$\rho = \frac{d\mathcal{C}}{dy} \quad (130)$$

Intuitivement, $\rho(y)$ est la probabilité que Y se trouve dans l'intervalle $[y, y+dy]$. ρ respecte alors

$$Pr\left(X \in [a, b]\right) = \int_a^b \rho(x)dx \quad (131)$$

Cette fonction est donc positive d'après les propriétés de la fonction de répartition et

$$\int_{\mathbb{R}} \rho(x)dx = 1 \quad (132)$$

La Table 4 explicite les pdf pour des distributions stochastiques communes.

Nom de la distribution	Notation	Support	pdf
Uniforme	$\mathcal{U}(a, b)$	$[a, b]$	$\rho(y) = \frac{1}{b-a}$
Normale	$\mathcal{N}(\mu, \sigma)$	$[-\infty, \infty]$	$\rho(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$
Gamma	$\Gamma(\alpha, \beta)$	$[0, \infty]$	$\rho(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \exp(-\beta y)$
Bêta	$\mathcal{B}(\alpha, \beta)$	$[0, 1]$	$\rho(y) = \frac{\Gamma(\alpha+\beta)y^{\alpha-1}(1-y)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}$

TABLE 4 – Exemple de distributions stochastiques continues communes et leur pdf

2.2.2.4 Espérance et variance

L'espérance d'une variable aléatoire Y est la quantité

$$E(Y) = \int_{-\infty}^{\infty} y\rho(y)dy \quad (133)$$

C'est la valeur moyenne prise par la variable aléatoire réelle en prenant en compte sa distribution stochastique.

La variance est une quantité traduisant la dispersion des valeurs que peut prendre la variable aléatoire par rapport à son espérance, toujours en prenant en compte sa distribution stochastique

$$D(Y) = \int_{-\infty}^{\infty} y^2\rho(y)dy - \left(E(Y)\right)^2 \quad (134)$$

Une variable aléatoire de second ordre est une variable aléatoire ayant une variance finie. Toutes les variables considérées dans l'étude sont supposées être de second ordre.

2.2.3 Paramètres d'entrée

Les paramètres incertains de l'étude dont dépend la QOI sont appelés les « paramètres d'entrée ». Ils sont modélisés par des variables aléatoires réelles de Ω dans \mathbb{R} , Ω représentant l'ensemble des issues possibles des phénomènes aléatoires menant à des variations des paramètres d'entrée. Ces événements sont considérés de façon abstraite, ils ne sont

pas explicités.

Les paramètres d'entrée peuvent être concrets (des variables physiques entachées d'incertitudes) ou abstraits (des paramètres mathématiques considérés pour introduire l'aléatoire). Chaque paramètre d'entrée concret est :

- Soit modélisé directement par une variable aléatoire ayant une certaine pdf.
- Soit modélisé comme une fonction d'un paramètre abstrait étant une variable aléatoire standard (suivant une pdf commune). Il peut être exprimé comme somme de polynômes [75] ou par une transformation isoprobabiliste [64]. Dans ce cas, les termes de paramètre d'entrée peuvent également faire référence à ces variables standards par abus de langage.

d paramètres d'entrée sont considérés. Ils sont notés (P_1, \dots, P_d) et sont regroupés dans un vecteur \mathbf{P} . Les valeurs prises par ces variables aléatoires pour une issue $\omega \in \Omega$ sont notées (p_1, \dots, p_d) et sont regroupées dans un vecteur \mathbf{p} .

$$\mathbf{p} = (p_1, \dots, p_d) = (P_1(\omega), \dots, P_d(\omega)) \quad (135)$$

Par abus de langage, les termes de « paramètres d'entrée » font également référence à \mathbf{p} et à ses composantes. Lorsque plusieurs variables aléatoires sont en jeu dans la même expérience, une fonction de répartition à plusieurs variables est définie

$$\mathcal{C}(p_1, \dots, p_d) = Pr(P_1 \leq p_1, \dots, P_d \leq p_d) \quad (136)$$

la densité de probabilité à plusieurs variables (joint pdf en anglais) ρ est alors définie comme étant la dérivée à l'ordre d de cette fonction de répartition par rapport à tous les paramètres d'entrée. À l'instar de la définition de la pdf à une variable

$$\rho(p_1, \dots, p_d) = Pr(P_1 \in [p_1 + dp_1], \dots, P_d \in [p_d + dp_d]) \quad (137)$$

La détermination de la pdf est une étape de l'UQ qui n'est pas considérée dans cette étude. Dans les travaux, chaque paramètre d'entrée P_m (concret ou abstrait) est supposé suivre une loi de distribution de pdf ρ_m connue et les P_m sont supposés être indépendants. Le cas de variables dépendantes est discuté à la section 2.2.5. Dans le cas de variables indépendantes

$$\rho(\mathbf{p}) = \prod_{m=1}^d \rho_m(p_d) \quad (138)$$

où ρ_m est la pdf de P_m .

Afin de rester au plus proche de l'article introduisant les indices de Sobol [62], étudié à la section 2.4.3, chaque P_m est associé à une variable aléatoire abstraite X_m suivant une loi

uniforme sur $[0, 1]$ par une transformation bijective prenant en compte la distribution de P_m . La valeur prise par la variable X_m pour une certaine issue est notée x_m . $\forall m \in [1, d]$

$$x_m = \mathcal{C}_m(p_m) \quad (139)$$

où \mathcal{C}_m est la fonction de répartition associée à la pdf ρ_m . La transformation est bijective si $\rho_m > 0$ et non seulement $\rho_m \geq 0$. Cette hypothèse est supposée vérifiée dans tous les travaux (les pdf sont seulement considérées sur leur support). Avec cette transformation, chaque X_m suit alors bien une loi uniforme sur $[0, 1]$. $\mathbf{x} = (x_1, \dots, x_d) \in [0, 1]^d$ est le vecteur contenant toutes les valeurs prises par ces paramètres. Les termes de « paramètres d'entrée » dénotent alors \mathbf{X} ou \mathbf{x} . L'hypercube unité de dimension d , noté $\mathbb{H}_d = [0, 1]^d$ représente l'espace des paramètres, l'espace contenant toutes les valeurs possibles de \mathbf{x} . Ainsi, le nombre de paramètres d'entrée d porte également le nom de « dimension », donnant la dimension de l'espace des paramètres. Avec cette transformation

$$\int \bullet \rho(\mathbf{p}) d\mathbf{p} = \int_{\mathbb{H}_d} \bullet d\mathbf{x} \quad (140)$$

Cette transformation n'est alors possible que si les variables sont indépendantes et qu'elles possèdent une fonction de répartition connue.

Dans d'autres études [64], cette transformation bijective n'est qu'une étape intermédiaire pour ensuite associer les paramètres d'entrée à d'autres variables suivant des pdfs voulues. Une fois la transformation (139) effectuée, la réciproque de la fonction de répartition de la variable finale que l'on souhaite manipuler est appliquée à x_m . Ces variables finales suivent alors une loi de distribution souhaitée et présentes certaines propriétés recherchées, comme être définies dans un certain support par exemple.

2.2.4 QOI

La QOI est notée f . f peut être une fonction des variables de temps et d'espace, mais en premier lieu, ces dépendances ne sont pas considérées. Seule la dépendance aux d paramètres d'entrée est considérée, sans perte de généralité. La QOI est une fonction déterministe des paramètres d'entrée, mais comme ces derniers sont stochastiques sur l'espace $(\Omega, \mathcal{A}, Pr)$, elle est elle-même stochastique sur cette espace. On note \tilde{f} la QOI dépendant des variables standards \mathbf{x} , après le changement de variable (139)

$$\tilde{f}(x_1, \dots, x_d) = f(p_1, \dots, p_d) \quad (141)$$

Dorénavant, le terme de « QOI » fait référence à \tilde{f} et non plus à f .

2.2.5 Transformations dans le cas de variables dépendantes

2.2.5.1 Variables dépendantes

Deux variables aléatoires sont dépendantes si la valeur de l'une influence la distribution de l'autre ou inversement. Dans le cas de variables dépendantes, la relation (138) n'est plus valable.

2.2.5.2 Transformée de Rosenblatt

La transformée de Rosenblatt, introduite dans [51], permet d'associer à un ensemble de variables aléatoires dépendantes \mathbf{P} , un ensemble de variables aléatoires indépendantes \mathbf{X} où chaque X_m suit une loi uniforme sur $[0, 1]$. La transformée de Rosenblatt est la suivante.

$$\begin{aligned}x_1 &= Pr(P_1 \leq p_1) \\x_2 &= Pr(P_2 \leq p_2 \mid P_1 = p_1) \\&\dots \\x_d &= Pr(P_d \leq p_d \mid P_1 = p_1, \dots, P_{d-1} = p_{d-1})\end{aligned}\tag{142}$$

Dans le cas de variables indépendantes, la bijection (139) est équivalente à cette transformée. Tout comme cette dernière, les réciproques de fonction de répartition peuvent être appliquées pour obtenir des variables aléatoires indépendantes suivant une distribution voulue. De façon classique, la transformée de Rosenblatt est utilisée pour obtenir des variables aléatoires suivant une loi normale centrée réduite.

2.2.5.3 Transformée de Nataf

Une autre transformée, utilisée de façon classique, est la transformée de Nataf [46]. Cette transformée en deux étapes est expliquée brièvement dans cette section. La première étape consiste à appliquer la réciproque de la fonction de répartition associée à la loi normale centrée réduite à chaque fonction de répartition marginale des variables dépendantes. La seconde étape consiste à décorréliser les variables gaussiennes obtenues en appliquant une transformation linéaire. Elle est utile lorsque seules les fonctions de répartition marginales et la matrice de corrélation sont connues.

2.3 Notation indicielle des interactions

En analyse de sensibilité et en UQ, l'impact des actions des paramètres d'entrée et des interactions entre ces derniers sur la QOI est étudié. Afin de pouvoir dénoter de façon compacte chaque action et chaque interaction dans les démonstrations, une logique de notation d'indexage multidimensionnel est introduite dans cette section. Un ensemble d'objets mathématiques qui sont utiles dans ces démonstrations sont également introduits. Les indices sont alors des multi-indices représentés sous la forme de vecteur. Les notions de vecteurs de dimension 1 et de scalaires sont confondus dans les travaux.

L'ensemble fini contenant tous les entiers allant de 1 à d , est noté $\Omega_d = [1, d] = \{1, \dots, d\}$.

2.3.1 Indices d'interaction

Chaque multi-indice désignant une action ou une interaction peut être le scalaire « 0 », un scalaire de Ω_d ou un vecteur de $(\Omega_d)^j$ où j représente l'ordre de l'interaction.

- L'indice scalaire « 0 » désigne la non-action, le fait de ne dépendre d'aucun paramètre d'entrée.
- La valeur d'un indice scalaire non nul i désigne l'action du paramètre x_i seul. Par exemple, $i = 2$ désigne l'action du paramètre x_2 seul.
- Les valeurs des composantes d'un multi-indice $\mathbf{i} \in (\Omega_d)^j$ déterminent les paramètres d'entrée en lien avec l'interaction à l'ordre j qu'il désigne. Par exemple, $\mathbf{i} = (1, 3)$ désigne l'interaction à l'ordre 2 entre x_1 et x_3 .

Par souci de simplicité des notations, la non-action et les actions à l'ordre 1 sont également désignées « interactions » (interactions à l'ordre 0 et interactions à l'ordre 1), et les indices scalaires i sont notés également \mathbf{i} dans les démonstrations.

Définition 1. $\forall d \in \mathbb{N}^*, \forall j \in \Omega_d, \mathcal{I}_j^{(d)}$ est l'espace fini des vecteurs dont les composantes sont les combinaisons sans remise entre j entiers de Ω_d (les composantes étant rangées par ordre croissant). Par convention, la définition de ces ensembles est étendue au cas $j = 0$ par $\mathcal{I}_0^{(d)} = \{0\}$.

Exemple 8. Pour $d = 3$

- $\mathcal{I}_0^{(3)} = \{0\}$
- $\mathcal{I}_1^{(3)} = \{1, 2, 3\}$
- $\mathcal{I}_2^{(3)} = \{(1, 2), (1, 3), (2, 3)\}$
- $\mathcal{I}_3^{(3)} = \{(1, 2, 3)\}$

$\mathcal{I}_j^{(d)}$ contient tous les multi-indices désignant les interactions à l'ordre j entre les d paramètres d'entrée.

Propriété 1.

$$\text{card}(\mathcal{I}_j^{(d)}) = \binom{d}{j} \quad (143)$$

Pour un indice $\mathbf{i} \in \mathcal{I}_j^{(d)}$, l'indice $\sim \mathbf{i}$ est quant à lui le vecteur de taille $d - j$ ayant pour composantes les entiers de Ω_d qui ne font pas partie des composantes de \mathbf{i} , rangés dans l'ordre croissant.

Si cet indexage est pratique pour les démonstrations mathématiques, un indexage scalaire est plus pratique pour les simulations numériques. Une transformation bijective pour associer à chaque multi-indice un indice scalaire pour les applications numériques est alors introduite.

$\forall d \in \mathbb{N}^*, \forall j \in [0, d]$, pour associer un indice scalaire à un indice de taille j , $\mathbf{i} \in \mathcal{I}_j^{(d)}$, voici les étapes proposées :

- un vecteur \mathbf{b} de taille d ayant 1 pour composantes indexées par les composantes de \mathbf{i} et 0 pour les autres composantes, est associé à \mathbf{i} .
- le vecteur \mathbf{b} peut être associé à un nombre binaire $(b)_2$ dont les chiffres sont les composantes de \mathbf{b} lues de droite à gauche. Le scalaire associé à \mathbf{i} , $M(\mathbf{i})$, est alors le nombre décimal tel que $M(\mathbf{i}) - 1$ soit égal à $(b)_2$ en base 2.

Avec cette procédure, l'indice scalaire de la non-action est toujours 1 et l'indice scalaire de l'interaction entre les d paramètres d'entrée est toujours le dernier indice 2^d .

Exemple. Avec $d = 4$

j	\mathbf{i}	\mathbf{b}	$(b)_2$	$M(\mathbf{i}) - 1$	$M(\mathbf{i})$
0	0	(0, 0, 0, 0)	(0000) ₂	0	1
1	1	(1, 0, 0, 0)	(0001) ₂	1	2
	2	(0, 1, 0, 0)	(0010) ₂	2	3
	3	(0, 0, 1, 0)	(0100) ₂	4	5
	4	(0, 0, 0, 1)	(1000) ₂	8	9
2	(1, 2)	(1, 1, 0, 0)	(0011) ₂	3	4
	(1, 3)	(1, 0, 1, 0)	(0101) ₂	5	6
	(1, 4)	(1, 0, 0, 1)	(1001) ₂	9	10
	(2, 3)	(0, 1, 1, 0)	(0110) ₂	6	7
	(2, 4)	(0, 1, 0, 1)	(1010) ₂	10	11
	(3, 4)	(0, 0, 1, 1)	(1100) ₂	12	13
3	(1, 2, 3)	(1, 1, 1, 0)	(0111) ₂	7	8
	(1, 2, 4)	(1, 1, 0, 1)	(1011) ₂	11	12
	(1, 3, 4)	(1, 0, 1, 1)	(1101) ₂	13	14
	(2, 3, 4)	(0, 1, 1, 1)	(1110) ₂	14	15
4	(1, 2, 3, 4)	(1, 1, 1, 1)	(1111) ₂	15	16

2.3.2 Ensemble fini de composantes

Pour un sous-ensemble $E \subset \Omega_d$, \overline{E} désigne son ensemble complémentaire dans Ω_d .

Définition 2. $\forall d \in \mathbb{N}^*, \forall j \in \Omega_d, \forall \mathbf{i} \in \mathcal{I}_j^{(d)}$, $\mathcal{S}(\mathbf{i})$ désigne l'ensemble fini d'entiers de cardinal j contenant les composantes de \mathbf{i} . Par convention, si $\mathbf{i} \in \mathcal{I}_0^{(d)}$ ($i = 0$), $\mathcal{S}(\mathbf{i}) = \emptyset$.

Exemple 9. Pour $d = 3, j = 2$ et $\mathbf{i} = (1, 3)$, $\mathcal{S}(\mathbf{i}) = \{1, 3\}$ et $\overline{\mathcal{S}(\mathbf{i})} = \mathcal{S}(\sim \mathbf{i}) = \{2\}$.

Lemme 1. $\forall d \in \mathbb{N}^*, \forall j_1 \in \Omega_d, \forall \mathbf{u} \in \mathcal{I}_{j_1}^{(d)}$ et $\forall j_2 \in \Omega_d, \forall \mathbf{v} \in \mathcal{I}_{j_2}^{(d)}$

$$\overline{\mathcal{S}(\mathbf{u})} \cup (\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v})) = \overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \quad (144)$$

Démonstration.

$$\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} = \overline{\mathcal{S}(\mathbf{u})} \cup \overline{\mathcal{S}(\mathbf{v})} = \overline{\mathcal{S}(\mathbf{u})} \cup \left((\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v})) \cup (\overline{\mathcal{S}(\mathbf{u})} \setminus \mathcal{S}(\mathbf{v})) \right) = \overline{\mathcal{S}(\mathbf{u})} \cup (\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v}))$$

□

Ce lemme est utilisé lors de démonstrations réalisées par la suite.

2.3.3 Vecteurs de paramètres d'entrée

Définition 3. $\forall d \in \mathbb{N}^*, \forall j \in \Omega_d, \forall \mathbf{i} \in \mathcal{I}_j^{(d)}, \mathbf{x}_{\mathbf{i}} = (x_{i_1}, \dots, x_{i_j}) \in [0, 1]^j$ désigne le vecteur contenant les j paramètres d'entrée indexés par les composantes de \mathbf{i} et $\mathbf{x}_{\sim \mathbf{i}} \in [0, 1]^{d-j}$ désigne le vecteur contenant les $d - j$ autres composantes de \mathbf{x} . Par convention, pour $i \in \mathcal{I}_0^{(d)}$ ($i = 0$), une fonction dépendante de $\mathbf{x}_{\mathbf{i}}$ ne dépend d'aucun paramètre d'entrée et $\mathbf{x}_{\sim \mathbf{i}} = \mathbf{x}$.

2.3.4 Symboles de Kronecker étendus

Définition 4. $\forall d \in \mathbb{N}^*, \forall \mathbf{a} \in \mathbb{N}^d, \forall \mathbf{b} \in \mathbb{N}^d, \forall E \subset \Omega_d$, soit $\Delta_{\mathbf{a}, \mathbf{b}}^E$

$$\Delta_{\mathbf{a}, \mathbf{b}}^E = \prod_{m \in E} \delta_{a_m, b_m} \quad (145)$$

avec $\delta_{m,n}$ le symbole de Kronecker classique

$$\delta_{a,b} = \begin{cases} 1 & \text{si } a = b \\ 0 & \text{sinon} \end{cases} \quad (146)$$

Pour deux vecteurs \mathbf{a} et \mathbf{b} appartenant à \mathbb{N}^d , $\Delta_{\mathbf{a}, \mathbf{b}}^{S(\mathbf{i})}$ est égal à 1 si toutes les composantes de \mathbf{a} et de \mathbf{b} indexées par les composantes \mathbf{i} sont égales. Par convention $\Delta_{\mathbf{a}, \mathbf{b}}^{S(\mathbf{0})} = \Delta_{\mathbf{a}, \mathbf{b}}^{\emptyset} = 1$.

Exemple 10. Pour $d = 4$, $\mathbf{i} = (1, 3)$ et $\mathbf{a} = (4, 5, 7, 9)$, si $\mathbf{b} = (1, 2, 7, 9)$, $\Delta_{\mathbf{a}, \mathbf{b}}^{S(\mathbf{i})} = 0$ car $a_1 \neq b_1$ mais si $\mathbf{b} = (4, 2, 7, 9)$, $\Delta_{\mathbf{a}, \mathbf{b}}^{S(\mathbf{i})} = 1$, car $a_1 = b_1$ et $a_3 = b_3$.

Propriété 2. $\forall E \subset \Omega_d, \forall F \subset \Omega_d$

$$\Delta_{\mathbf{a}, \mathbf{b}}^E \times \Delta_{\mathbf{a}, \mathbf{b}}^F = \Delta_{\mathbf{a}, \mathbf{b}}^{E \cup F} \quad (147)$$

Démonstration.

$$\Delta_{\mathbf{a}, \mathbf{b}}^E \times \Delta_{\mathbf{a}, \mathbf{b}}^F = \prod_{m \in E \setminus F} \delta_{a_m, b_m} \times \prod_{m \in F \setminus E} \delta_{a_m, b_m} \prod_{m \in E \cap F} (\delta_{a_m, b_m})^2 \quad (148)$$

Or, d'après la définition du symbole de Kronecker classique, $(\delta_{a,b})^2 = \delta_{a,b}$, donc (147) est prouvée. \square

Définition 5.

$$\tilde{\Delta}_{\mathbf{a},\mathbf{b}}^E = \prod_{m \in E} \delta_{a_m, b_m} \prod_{m \notin E} (1 - \delta_{a_m, b_m}) \quad (149)$$

Il s'agit d'un autre symbole de Kronecker étendu égal à 1 lorsque toutes les composantes des deux vecteurs sont indexées par les indices contenus dans E sont égales et quand toutes les autres composantes diffèrent. Si l'une de ces conditions n'est pas respectée, $\tilde{\Delta}_{\mathbf{a},\mathbf{b}}^E = 0$.

2.3.5 Interactions d'ordre inférieur

Définition 6. $\forall d \in \mathbb{N}^*, \forall j \in \Omega_d, \forall \mathbf{i} \in \mathcal{I}_j^{(d)}, \forall l \in \llbracket 0, j \rrbracket$

$$\mathcal{I}_l^{(j)}\{\mathbf{i}\} = \{\mathbf{u} \in \mathcal{I}_l^{(d)}, \mathcal{S}(\mathbf{u}) \subset \mathcal{S}(\mathbf{i})\}$$

$\mathcal{I}_l^{(d)}\{\mathbf{i}\}$ est l'ensemble fini des vecteurs de composantes étant les combinaisons sans remise entre l des j composantes de \mathbf{i} (les composantes étant rangées par ordre croissant). Par convention, $\mathcal{I}_0^{(d)}\{\mathbf{i}\} = \{0\}$. $\mathcal{I}_l^{(d)}\{\mathbf{i}\}$ correspond à l'ensemble des multi-indices indexant les interactions entre l des j paramètres d'entrée indexés par les composantes de \mathbf{i} .

Propriété 3.

$$\text{card}(\mathcal{I}_l^{(j)}\{\mathbf{i}\}) = \binom{j}{l}$$

Exemple 11. Pour $d = 4, j = 3$ et $\mathbf{i} = (1, 3, 4)$,

- $\mathcal{I}_0^{(j)}\{\mathbf{i}\} = \{0\}$
- $\mathcal{I}_1^{(j)}\{\mathbf{i}\} = \{1, 3, 4\}$
- $\mathcal{I}_2^{(j)}\{\mathbf{i}\} = \{(1, 3), (1, 4), (3, 4)\}$
- $\mathcal{I}_3^{(j)}\{\mathbf{i}\} = \{(1, 3, 4)\}$

2.4 Analyse de sensibilité

L'analyse de sensibilité a pour but d'expliquer l'incertitude commise sur la QOI en fonction des incertitudes commises sur les paramètres d'entrée, notamment en classant l'impact de chaque action et interaction entre ces paramètres d'entrée sur la QOI.

Les techniques de sensibilité peuvent être scindées en 2 groupes : les méthodes d'analyse de sensibilité locale et les méthodes d'analyse de sensibilité globale.

2.4.1 Analyse de sensibilité locale

L'analyse de sensibilité locale consiste à étudier les pentes $\frac{\partial f}{\partial x_m}$ autour d'un point nominal.

Dans [53], les auteurs mettent la lumière sur le fait que ce critère de sensibilité ne convient que localement autour du point nominal. Ils ne conviennent de façon globale que pour les systèmes linéaires, les d pentes obtenues n'étant constantes dans tout l'espace des paramètres que dans ce cas. Pour un modèle non-linéaire, le choix du point nominal, ainsi que les valeurs des autres paramètres d'entrée modifient la valeur de ces pentes.

2.4.2 Analyse de sensibilité globale (autre que les indices de Sobol)

L'analyse de sensibilité globale consiste à étudier l'influence des paramètres d'entrée sur la QOI lorsqu'ils varient sur l'ensemble de leur intervalle.

2.4.2.1 Méthode One At Time

La méthode One At time est une méthode consistant à choisir un point nominal au centre de l'espace des paramètres (au centre de l'hypercube unité en considérant les variables X_m). Pour chaque direction de l'espace $m \in \Omega_d$, la quantité $\Delta_m = \frac{f(x_1, \dots, 1, \dots, x_d) - f(x_1, \dots, 0, \dots, x_d)}{1-0}$ est considérée comme indice de sensibilité. Comme il est souligné dans [54], cette méthode ne permet pas d'explorer la totalité de l'espace des paramètres, ne pouvant atteindre les coins de cet espace. À mesure que le nombre de paramètres d'entrée augmente, qui correspond à la dimension de l'hypercube unité \mathbb{H}_d , l'espace inexploré représente une fraction de plus en plus importante de l'espace. Cette méthode ne semble alors elle aussi n'être adaptée qu'aux modèles linéaires.

2.4.2.2 Indices de Morris

Cette méthode, développée dans [43] et améliorée dans [12], est aussi appelée « méthode des effets élémentaires ». Elle repose sur un échantillonnage de l'espace des paramètres. Pour chaque point échantillonné \mathbf{x} et une valeur Δ sélectionnée, d effets élémentaires sont calculés

$$\delta_m(\mathbf{x}) = \frac{y(x_1, \dots, x_m + \Delta, \dots, x_d) - y(\mathbf{x})}{\Delta} \quad (150)$$

Classiquement, la moyenne et la variance de chaque δ_m sur l'ensemble des points échantillonnés sont calculées. La moyenne donne un ordre de grandeur de l'influence du paramètre x_m et la variance donne un ordre de grandeur des non-linéarités liées à ce paramètre. Dans [12], la moyenne de la valeur absolue de chaque effet élémentaire δ_m est considérée au lieu de la moyenne des effets élémentaires. Cela permet d'empêcher l'annulation d'effets élémentaires deux à deux pour une dépendance non monotone. Dans l'article original [43], une technique d'échantillonnage est proposée, qui est améliorée dans [12]. Cette méthode est alors vue comme une amélioration de la méthode OAT permettant une meilleure exploration de l'espace des paramètres.

2.4.2.3 Indices DGSM

Les indices DGSM, pour « derivative based global sensitivity measures » [63] ont une définition proche des indices de Morris concernant la moyenne des valeurs absolues des effets élémentaires

$$\int_{\mathbb{H}_d} \left(\frac{\partial f}{\partial x_m} \right)^2 d\mathbf{x} \quad (151)$$

Ces indices sont estimés à l'aide d'un échantillonnage Monte-Carlo ou Quasi Monte-Carlo (comme les échantillonnages de Sobol [61]). Il est alors montré dans [63] que ces indices sont supérieurs ou égaux aux indices de Sobol totaux (définis à la section 2.4.3.4). Il en résulte que si cet indice de sensibilité est faible, le paramètre d'entrée n'est pas influant. L'inverse n'est pas vrai, cet indice peut être élevé même si en réalité le paramètre n'est pas influant. Comme il nécessite moins d'échantillons pour obtenir une estimation précise de ces indices que pour obtenir les indices de Sobol totaux, ils peuvent être utilisés en première approche pour déterminer quels sont les paramètres non influents sur la QOI. Ainsi, ces paramètres non influents pourraient être fixés à une valeur nominale. Le nombre de paramètres d'entrée (la dimension du problème) peut alors être réduit.

2.4.3 Indices de Sobol

Par rapport aux autres indices de sensibilité décrits ci-dessus, les indices de Sobol se sont affirmés comme étant ceux offrant l'image de la sensibilité des interactions des paramètres d'entrée la plus fidèle.

2.4.3.1 Décomposition ANOVA (ANalysis Of VAriance)

La QOI \tilde{f} peut être décomposée en une unique somme de sous-fonctions, dénommées « summands » dans l'article [62]. Chaque sous-fonction correspond à une interaction à l'ordre $j = 0$ ou à l'ordre $j \in \Omega_d$.

$$\tilde{f}(\mathbf{x}) = \sum_{j=0}^d \sum_{\mathbf{i} \in \mathcal{I}_j^{(d)}} \tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) \quad (152)$$

$\tilde{f}_0^{(0)}$ est l'espérance du QOI

$$\tilde{f}_0^{(0)} = \int_{\mathbb{H}_d} \tilde{f}(\mathbf{x}) d\mathbf{x} \quad (153)$$

et, $\forall j \in \Omega_d, \forall \mathbf{i} \in \mathcal{I}_j^{(d)}, \forall m \in \mathcal{S}(\mathbf{i})$

$$\int_0^1 \tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) dx_m = 0 \quad (154)$$

Chaque $\tilde{f}_{\mathbf{i}}^{(j)}$ représente l'impact de l'interaction entre les j paramètres d'entrée indexés par les composantes de \mathbf{i} sur la QOI. Il est prouvé dans [62] que

$$\tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) - \sum_{l=0}^{j-1} \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) \quad (155)$$

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \int_{\mathbb{H}_{d-j}} \tilde{f}(\mathbf{x}_{\mathbf{i}}, \mathbf{x}_{\sim \mathbf{i}}) d\mathbf{x}_{\sim \mathbf{i}} \quad (156)$$

$\mu_{\mathbf{i}}^{(j)}$ est l'espérance de \tilde{f} sachant $\mathbf{x}_{\mathbf{i}}$. Chaque sous-fonction $\tilde{f}_{\mathbf{i}}^{(j)}$ est de ce fait égale à l'espérance de f sachant $\mathbf{x}_{\mathbf{i}}$ soustraite des « summands » correspondant aux interactions d'ordre $0 \leq l < j$ entre les composantes de \mathbf{i} .

Exemple 12. Pour $d \geq 3$, $j = 2$ et $\mathbf{i} = (1, 3)$:

$$\tilde{f}_{(1,3)}^{(2)}\left((x_1, x_3)\right) = \mu_{(1,3)}^{(2)}\left((x_1, x_3)\right) - \tilde{f}_1^{(1)}(x_1) - \tilde{f}_3^{(1)}(x_3) - \tilde{f}_0^{(0)}$$

2.4.3.2 Variances partielles

La variance de la QOI est notée D .

$$D = \int_{\mathbb{H}_d} \tilde{f}(\mathbf{x})^2 d\mathbf{x} - \left(\tilde{f}_0^{(0)}\right)^2 \quad (157)$$

Une conséquence directe de la décomposition ANOVA (152) est que la variance peut être décomposée en une somme de termes appelés « variances partielles » [62]

$$D = \sum_{j=1}^d \sum_{\mathbf{i} \in \mathcal{I}_j^{(d)}} D_{\mathbf{i}}^{(j)} \quad (158)$$

$$D_{\mathbf{i}}^{(j)} = \int_{\mathbb{H}_j} \left(\tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}})\right)^2 d\mathbf{x}_{\mathbf{i}} = \int_{\mathbb{H}_d} \left(\tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}})\right)^2 d\mathbf{x} \quad (159)$$

Chaque variance partielle $D_{\mathbf{i}}^{(j)}$ donne une valeur quantitative de l'impact de l'interaction \mathbf{i} sur la variance totale.

2.4.3.3 Indices de Sobol

Les indices de Sobol sont alors définis par

$$S_{\mathbf{i}}^{(j)} = \frac{D_{\mathbf{i}}^{(j)}}{D} \quad (160)$$

L'indice de Sobol $S_{\mathbf{i}}^{(j)}$ donne l'information sur l'impact de l'interaction \mathbf{i} sous la forme d'un ratio. Avec ces informations, les interactions ayant le plus d'impact sur la QOI peuvent

être déterminées et toutes les interactions peuvent être classées par importance d'impact sur la QOI. Comme la somme de toutes les variances partielles vaut D , la somme de tous les indices de Sobol vaut 1.

Les indices au premier ordre $S_m^{(1)}$ donne une image de l'influence du paramètre x_m seul sur la QOI. Les indices $S_{(m_1, m_2)}^{(2)}$ donne une image de l'impact de l'interaction entre x_{m_1} et x_{m_2} sur la QOI, et ainsi de suite.

2.4.3.4 Indices de Sobol totaux

Les indices de Sobol totaux sont introduits dans [27]. Il existe un indice de Sobol total par paramètre d'entrée. Pour un paramètre P_m donné, la quantité $D_m^{(T)}$ est définie comme étant la somme de toutes les variances partielles des interactions d'ordre 1 à d où m apparait. L'indice de Sobol total $S_m^{(T)}$ est défini comme étant cette quantité $D_m^{(T)}$ divisée par la variance D , à l'instar des indices de Sobol classiques. L'indice de Sobol total est donc la somme des indices de Sobol des interactions où m apparait. Si l'indice $S_m^{(T)}$ est proche de 0, cela signifie que toutes les interactions concernant le paramètre d'entrée x_m n'ont que peu d'influence sur la QOI par rapport aux autres.

Note : la somme des d indices totaux ne vaut pas 1.

2.4.3.5 Méthode Monte-Carlo

Les variances partielles $D_i^{(j)}$ et les quantités D_i^T peuvent être estimées à l'aide de la méthode Monte Carlo (Monte-Carlo method - MC) [62, 55]. Pour effectuer l'estimation d'une de ces quantités, une séquence de Sobol [61] (ou une autre séquence quasi-aléatoire) de N points de dimension $2d$ est générée. Chaque point de dimension $2d$, \mathbf{a} , peut voir ses composantes être séparées en deux groupes égaux (les d premières et les d dernières) pour donner 2 points de dimension d , \mathbf{y} et \mathbf{z}

$$\mathbf{a} = (y_1, \dots, y_d, z_1, \dots, z_d) \quad (161)$$

L'espérance et la variance peuvent être estimées de façon classique avec la séquence des N points \mathbf{y} ou des N points \mathbf{z} .

La variance partielle $D_i^{(j)}$ peut être estimée par :

$$D_i^{(j)} = \frac{1}{N} \sum_{n=1}^N f(\mathbf{y}_n) f(\mathbf{y}_{n, \mathbf{i}}, \mathbf{z}_{n, \sim \mathbf{i}}) - \sum_{l=1}^{j-1} \sum_{\mathbf{u} \in \mathcal{I}_l^{(d)} \{\mathbf{i}\}} D_{\mathbf{u}}^{(l)} - (f_0^{(0)})^2 \quad (162)$$

Pour estimer une variance partielle à l'ordre j , les variances partielles des ordres inférieurs doivent déjà avoir été estimées.

Pour la quantité D_i^T , l'estimation suivante est utilisée

$$D_i^{(T)} = \frac{1}{2N} \sum_{n=1}^N (f(\mathbf{y}_n) - f(\mathbf{y}_{n,\sim i}, \mathbf{z}_{n,i}))^2 \quad (163)$$

Cet estimateur est considéré comme étant le plus précis parmi d'autres estimateurs dans [55] (d'après les auteurs, cet estimateur est issu de [30]).

Obtenir l'intégralité des indices de Sobol nécessite alors de nombreux appels au système permettant d'obtenir la valeur de la QOI. Lorsqu'un appel est coûteux, ce qui est fait en général est que seul les indices au premier ordre et les indices totaux sont estimés. Avec les estimateurs ci-dessus, $N(d+2)$ appels sont nécessaires.

2.5 Métamodèle par chaos polynomial

2.5.1 Polynômes

Dans tout le manuscrit, les notions de polynômes et de fonctions polynomiales associées sont confondues.

2.5.1.1 Polynômes univariés

Un polynôme (fonction polynomiale) univarié(e) est une fonction de la forme

$$P(x) = \sum_{k=0}^n c_k x^k \quad (164)$$

où n est le degré du polynôme et les c_k sont ses coefficients, des nombres réels avec $c_n \neq 0$. Un terme de la somme se nomme un monôme. l'espace des polynômes de degré n , étant également l'espace des fonctions polynomiales dans notre simplification, est dénommé $\mathbb{R}_n[X]$.

2.5.1.2 Polynômes multivariés

À l'instar de la section 2.3, une logique d'indexage multidimensionnelle est introduite, ayant cette fois-ci pour objectif de désigner les monômes d'un polynôme multivarié ou les composantes d'un polynôme exprimées dans une certaine base (Cf. section 2.5.1.3).

Définition 7. $\forall d \in \mathbb{N}^*, \forall n \in \mathbb{N}, \mathcal{K}_n^{(d)}$ est l'ensemble des vecteurs de $[[0, d]]^d$ dont les composantes sont les arrangements avec remise entre d entiers de $[[0, d]]$ et dont la somme de ces composantes est inférieure ou égale à $n \in \mathbb{N}$.

Exemple 13. Pour $d = 3, n = 2$,

$$\mathcal{K}_2^{(3)} = \{ \begin{array}{l} (0, 0, 0), \\ (1, 0, 0), (0, 1, 0), (0, 1, 0), \\ (1, 1, 0), (1, 0, 1), (0, 1, 1), (2, 0, 0), (0, 2, 0), (0, 0, 2) \end{array} \}$$

Un polynôme à d variables de degré n est une fonction de la forme

$$\mathcal{P}(\mathbf{x}) = \mathcal{P}(x_1, \dots, x_d) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} c_{\mathbf{k}} \prod_{m=1}^d (x_m)^{k_m} \quad (165)$$

Les notions de monômes et de coefficients sont également valables pour les polynômes multivariés. Le degré d'un monôme indexé par \mathbf{k} est la somme des composantes de \mathbf{k} . Le degré n du polynôme est le maximum des degrés des monômes qui le composent. L'espace des polynômes à d variables de degré n est noté $\mathbb{R}_n[X_1, \dots, X_d]$. $\mathcal{K}_n^{(d)}$ contient alors l'ensemble des indices des monômes composant un polynôme à d variables de degré n .

Un polynôme multivarié peut également être vu comme étant le produit de d polynômes univariés. Le polynôme peut alors être indexé par $\mathbf{k} = (k_1, \dots, k_d)$, où chaque k_m est le degré de chaque polynôme univarié. Le polynôme est alors le produit de polynômes univariés de degré k_m dépendant chacun d'une variable différente x_m .

Exemple 14. Soit $\mathcal{P}_{\mathbf{k}}$ un polynôme à d variables indexé par \mathbf{k} .

$$\mathcal{P}_{\mathbf{k}}(\mathbf{x}) = \prod_{m=1}^d P_{m, k_m}(x_m) \quad (166)$$

où les P_{m, k_m} sont des polynômes univariés de degrés k_m .

Les polynômes faisant partie d'une base, notion introduite à la section 2.5.1.3, sont indexés par un indice multidimensionnel \mathbf{k} de cette façon.

2.5.1.3 Polynômes comme vecteurs

$\forall d \in \mathbb{R}, \forall n \in \mathbb{R}, \mathbb{R}_n[X_1, \dots, X_d]$ est un espace vectoriel. Ainsi, les polynômes peuvent être exprimés dans la base canonique ou dans d'autres bases de cet espace. La base canonique de $\mathbb{R}_n[X_1, \dots, X_d]$ est notée $(\Psi_{\mathbf{k}})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$

$$\Psi_{\mathbf{k}}(\mathbf{x}) = \prod_{m=1}^d x^{k_m} \quad (167)$$

Un polynôme à d variables de degré n peut donc être exprimé comme une combinaison linéaire de polynômes d'une base de $\mathbb{R}_n[X_1, \dots, X_d]$. La composante située devant le polynôme de la base indexé par \mathbf{k} (étant un produit de polynômes univariés de degré k_m dépendant de x_m) est elle-même indexée par \mathbf{k} .

Exemple 15. Soit $(\mathcal{P}_{\mathbf{k}})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ une base des polynômes à d variables de degré n . Un polynôme \mathcal{Q} à d variables de degré n peut s'écrire dans cette base

$$\mathcal{Q}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \mathcal{P}_{\mathbf{k}} \quad (168)$$

Le polynôme possède M composantes $\alpha_{\mathbf{k}}$

$$M = \binom{n+d}{n} \quad (169)$$

Il s'agit du cardinal de $\mathcal{K}_n^{(d)}$. L'évolution du nombre de composantes de l'expression du vecteur dans la base avec la dimension d et le degré n illustre la célèbre « malédiction de la dimension ». A haute dimension d , un degré n élevé peut devenir problématique en terme de coût tant au niveau de la mémoire qu'en terme de temps dans des applications numériques.

Il est plus pratique d'utiliser un indexage scalaire pour les simulations numériques. Pour cela, un indexage scalaire peut être adopté en rangeant les polynômes de la base par ordre croissant de degré, l'ordre entre les polynômes de même degré n'importe peu tant que la même logique est conservée tout au long de la simulation.

2.5.2 Orthogonalité des polynômes

2.5.2.1 Produits scalaires de l'espace des polynômes

En munissant un espace vectoriel réel d'un produit scalaire euclidien, on parle d'espace de Hilbert. Cela permet d'introduire la notion d'orthogonalité. Un produit scalaire associé à un espace vectoriel E est une opération qui à deux éléments de E associe un scalaire et qui possède les trois propriétés suivantes :

- Il est bilinéaire.
- Il est symétrique.
- Il est défini positif.

Dans cette section sont introduits des produits scalaires associés à $\mathbb{R}_n[X]$ ou $\mathbb{R}_n[X_1, \dots, X_d]$ qui sont utilisés dans les travaux.

Pour les polynômes univariés Pour les polynômes univariés, un produit scalaire associé à une pdf ρ peut être introduit. Pour deux polynômes P et Q de $\mathbb{R}_n[X]$

$$\langle P, Q \rangle_\rho = \int_{\mathcal{S}_\rho} P(x)Q(x)\rho(x)dx \quad (170)$$

où \mathcal{S}_ρ est le support de la pdf ρ . Comme $\rho > 0$ sur son support et par propriété de l'intégrale, l'application est bien un produit scalaire.

Soit $a \in [0, 1[$ et $b \in [a, 1]$, on définit de plus le produit scalaire suivant

$$\langle P, Q \rangle_{a,b} = \int_a^b P(x)Q(x)dx \quad (171)$$

Ce produit scalaire est semblable au produit scalaire associé à la loi uniforme sur $[a, b]$, étant égal à ce dernier à un facteur $\rho = 1/(b - a)$ près. Même s'il n'est pas associé à la loi uniforme, il s'agit tout de même d'un produit scalaire, étant une forme bilinéaire symétrique définie positive.

Pour les polynômes multivariés Pour les polynômes à d variables, le produit scalaire associé à la pdf ρ à plusieurs variables indépendantes est défini, pour les polynômes $\mathcal{P} = \prod_{m=1}^d P_m$ et $\mathcal{Q} = \prod_{m=1}^d Q_m$ de $\mathbb{R}_n[X_1, \dots, X_n]$

$$\langle \mathcal{P}, \mathcal{Q} \rangle_\rho = \prod_{m=1}^d \langle P_m, Q_m \rangle_{\rho_m} \quad (172)$$

De plus, soit un hyperpavé $H = [a_1, b_1] \times \dots \times [a_d, b_d]$, on définit le produit scalaire correspondant à H par

$$\langle \mathcal{P}, \mathcal{Q} \rangle_H = \int_H \mathcal{P}(\mathbf{x}) \mathcal{Q}(\mathbf{x}) d\mathbf{x} = \prod_{m=1}^d \langle P_{m,k_m}, Q_{m,k_m} \rangle_{a_m, b_m} \quad (173)$$

2.5.2.2 Orthogonalité

Deux polynômes \mathcal{P} et \mathcal{Q} de $\mathbb{R}_n[X_1, \dots, X_d]$ sont dits orthogonaux pour un produit scalaire $\langle \cdot, \cdot \rangle$ donné si

$$\langle \mathcal{P}, \mathcal{Q} \rangle = 0 \quad (174)$$

Une base de polynômes $(\mathcal{P}_{\mathbf{k}})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ est dite orthogonale si les polynômes qui la composent sont deux à deux orthogonaux. Elle est de plus dite orthonormée si, $\forall \mathbf{k} \in \mathcal{K}_n^{(d)}$

$$\|\mathcal{P}_{\mathbf{k}}\| = \sqrt{\langle \mathcal{P}_{\mathbf{k}}, \mathcal{P}_{\mathbf{k}} \rangle} = 1 \quad (175)$$

où $\|\cdot\|$ est la norme associée au produit scalaire.

Une base de polynômes multivariés orthonormaux par rapport à un produit scalaire est composée de polynômes respectant

$$\langle \mathcal{P}_{\mathbf{k}_1}, \mathcal{P}_{\mathbf{k}_2} \rangle = \Delta_{\mathbf{k}_1, \mathbf{k}_2}^{\Omega_d} \quad (176)$$

2.5.2.3 Définition de bases orthonormées spéciales

Pour chaque produit scalaire associé à une pdf classique, une base orthogonale de polynômes univariés existe. Le schéma de Wiener-Askey relie la loi de distribution à la base orthogonale associée.

Des lois classiques sont rappelées dans la Table 4. En ce qui concerne les polynômes multivariés associés à la distribution des paramètres d'entrée \mathbf{p} , il s'agit de polynômes $(\mathcal{P}_{\mathbf{k}})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$, où chaque $\mathcal{P}_{\mathbf{k}}$ est le produit de polynômes P_{m,k_m} suivant le schéma de Wiener

Loi de distribution	Famille de polynômes univariés associés	Support
Gaussian	Hermite	\mathbb{R}
Uniform	Legendre	$[-1, 1]$
Gamma	Laguerre	$[0, +\infty]$
Beta	Jacobi	$[-1, 1]$

TABLE 5 – Schéma de Wiener-Askey reliant loi de distribution stochastique à la famille de polynômes univariés orthogonaux

Askey (Cf. Table 5) pour la distribution de P_m . Chaque polynôme P_{m,k_m} est orthogonal d'après le produit scalaire associé à la loi de distribution du paramètre d'entrée $\langle \cdot, \cdot \rangle_{\rho_m}$ correspondant.

La base des polynômes univariés orthonormale par rapport à $\langle, \rangle_{a,b}$ ainsi que la base des polynômes à d variables orthonormale par rapport à \langle, \rangle_H sont également déterminées dans cette section. Cette dernière base est utilisée dans les métamodèles PC et multi-éléments (Multi-Element - ME) de différentes manières dans l'approche considérée.

Pour les polynômes univariés $\forall a \in [0, 1[, \forall b \in]a, 1]$, la base $(\phi_k^{a,b})_{k \in [0, n]}$ des polynômes univariés orthonormée par rapport à $\langle, \rangle_{a,b}$ est définie. Si $a = 0$ et $b = 1$, le produit scalaire est celui associé à la loi uniforme entre 0 et 1. $(\phi_k^{0,1})_{k \in [0, n]}$ est alors la base des polynômes de Legendre décalés étant normalisée pour être orthonormale (et non plus seulement orthogonale) à $\langle \cdot, \cdot \rangle_{0,1}$.

Dans le cas général, la base $(\phi_k^{a,b})$ est obtenue à partir de la base des polynômes de Legendre $(L_k)_{k \in [0, n]}$ standard, base orthogonale par rapport au produit scalaire associé à la loi uniforme sur $[-1, 1]$

$$\deg(L_k) = k \quad (177)$$

$$\frac{1}{2} \int_{-1}^1 L_m(x)L_n(x)dx = \frac{1}{2n+1} \delta_{m,n} \quad (178)$$

Avec le changement de variable $y = \frac{b-a}{2}x + \frac{a+b}{2}$

$$\frac{2n+1}{b-a} \int_a^b L_m\left(\frac{2y-(a+b)}{b-a}\right)L_n\left(\frac{2y-(a+b)}{b-a}\right)dy = \delta_{m,n} \quad (179)$$

Par identification

$$\phi_n^{a,b} = \sqrt{\frac{2n+1}{b-a}} L_n\left(\frac{2y-(a+b)}{b-a}\right) \quad (180)$$

En désignant les coefficients de ces polynômes exprimés dans la base canonique par

$$L_n(x) = \sum_{k=0}^n l_{n,k} x^k \quad (181)$$

$$\phi_n^{a,b} = \sum_{k=0}^n \lambda_{n,k}^{a,b} x^k \quad (182)$$

La relation entre les coefficients des polynômes de Legendre et ceux de la base orthonormée est obtenue en développant

$$L_n\left(\frac{2y - (a+b)}{b-a}\right) = \sum_{k=0}^n l_{n,k} \left(\frac{2y - (a+b)}{b-a}\right)^k \quad (183)$$

$$L_n\left(\frac{2y - (a+b)}{b-a}\right) = \sum_{k=0}^n \sum_{m=0}^k \frac{l_{n,k}}{(b-a)^k} \binom{k}{m} (2y)^m (-1)^{k-m} (a+b)^{k-m} \quad (184)$$

$$L_n\left(\frac{2y - (a+b)}{b-a}\right) = \sum_{m=0}^n \left[\sum_{k=m}^n \frac{l_{n,k}}{(b-a)^k} \binom{k}{m} (2)^m (-1)^{k-m} (a+b)^{k-m} \right] y^m \quad (185)$$

Par identification (après permutation des indices muets)

$$\lambda_{n,k}^{a,b} = \sqrt{\frac{2n+1}{b-a}} \sum_{m=k}^n \frac{l_{n,m}}{(b-a)^m} \binom{m}{k} (2)^k (-1)^{m-k} (a+b)^{m-k} \quad (186)$$

Avec, en particulier

$$\phi_0^{a,b} = \lambda_{0,0}^{a,b} = \frac{1}{\sqrt{b-a}} \quad (187)$$

Pour les polynômes multivariés En ce qui concerne les polynômes multivariés, pour un hyperpavé $H = [a_1, b_1] \times \cdots \times [a_d, b_d]$, la base $(\Phi_{\mathbf{k}}^H)_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ orthonormée par rapport à $\langle \cdot, \cdot \rangle_H$ est définie à l'aide des polynômes univariés définis au paragraphe précédent

$$\Phi_{\mathbf{k}}^H(\mathbf{x}) = \prod_{m=1}^d \phi_{k_m}^{a_m, b_m} \quad (188)$$

En particulier

$$\Phi_0^H = \prod_{m=1}^d \phi_0^{a_m, b_m} = \frac{1}{\sqrt{v_H}} \quad (189)$$

$$v_H = \prod_{m=1}^d (b_m - a_m) \quad (190)$$

où v_H est l'hypervolume de H .

2.5.3 Polynômes du chaos

2.5.3.1 Travaux de Wiener

Les polynômes du chaos reposent sur les travaux de Norbert Wiener [72], qui stipulent qu'une variable aléatoire de second ordre quelconque Y peut être exprimée

$$Y(\xi_1, \dots, \xi_\infty) = \sum \alpha_{\mathbf{k}} \mathcal{H}_{\mathbf{k}}(\xi_1, \dots, \xi_\infty) \quad (191)$$

Les $\alpha_{\mathbf{k}}$ sont appelés « les coefficients du chaos ». Les ξ_m sont une infinité de variables suivant une loi normale centrée réduite et les \mathcal{H} sont des polynômes multivariés étant des produits des polynômes de Hermite, polynômes étant orthogonaux par rapport au produit scalaire associé à la loi normale centrée réduite d'après le schéma de Wiener Askey (Cf. Table 5). La convergence L_2 de cette série a été prouvée dans [11]. Pour des applications physiques, un nombre fini de variables ainsi qu'un degré de troncature sont sélectionnés. Ce procédé peut alors être utilisé pour approcher la QOI, ou même pour approcher les paramètres d'entrée concrets [75], comme une série de polynômes des variables standards ξ_m . La première utilisation des polynômes du chaos en ingénierie a été réalisée dans [22], introduisant les méthodes de « Stochastic finite elements ».

2.5.3.2 g-PC

La méthode PC a ensuite été étendue dans [75] pour devenir la méthode « generalised Polynomial Chaos » (generalised Polynomial Chaos - g-PC). Les modèles g-PC sont des métamodèles polynomiaux pouvant utiliser des variables standards suivant une autre distribution que la loi centrale réduite et utilisant une autre base de polynômes orthogonaux. La nature de la base des polynômes multivariés dépend des pdf des variables standards sélectionnées. La base n'est alors plus $(\mathcal{H}_{\mathbf{k}})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ mais une base de polynômes multivariés, produit de polynômes suivant le schéma de Wiener Askey (Cf. Table 5). Dans le manuscrit, les abréviations PC et g-PC réfèrent toutes deux à un métamodèle g-PC.

Lorsqu'un paramètre d'entrée concret P_m , suivant une loi autre qu'une loi normale, est approché par une série g-PC, il est montré dans [75] que la convergence exponentielle de la série approchant la QOI peut être obtenue en utilisant des variables standards suivant une loi de même nature, car le paramètre d'entrée est mieux approché. En effet, si P_m suit une loi uniforme par exemple, l'approcher par une série de polynôme d'Hermite dépendant d'une variable gaussienne nécessite un haut degré de troncature et le modèle souffre du phénomène de Gibbs. Il peut néanmoins être approché directement par une série de polynôme de Legendre de degré 1 dépendant d'une variable suivant une loi uniforme sur $[-1, 1]$, ce qui est équivalent à effectuer une transformation linéaire entre le paramètre d'entrée concret P_m et la variable standard. L'extension g-PC a également permis les transformations introduites dans 2.2.3 entre les paramètres d'entrée, autre technique permettant d'utiliser des variables standards.

2.5.3.3 Dépendances en temps et en espace

Dans le cas d'une QOI dépendant du temps t et de l'espace \mathbf{r} , l'expansion en polynômes du chaos peut être étendue en considérant que les coefficients du chaos $\alpha_{\mathbf{k}}$ sont eux-mêmes des fonctions du temps et de l'espace. Pour une QOI solution d'un système dynamique, le modèle doit alors faire face à une complexité croissante avec le temps, appelé « long term integration problem » dans [69], la QOI ayant de grandes chances de devenir de plus en plus irrégulière dans son évolution avec les paramètres d'entrée.

Pour un paramètre d'entrée concret dépendant de l'espace, il est conseillé (notamment dans [38]) d'utiliser la transformée de Karhunen-Loeve [41]. Cette transformation n'est utilisable que sur les variables aléatoires dont la fonction de covariance est connue et ne peut donc pas être appliquée à la QOI. Cette transformation, aussi appelée « Principal Component Analysis » ou « Proper Orthogonal Decomposition » dans d'autres domaines, revient à décomposer la variables à l'aide des valeurs propres et des vecteurs propres de sa fonction de covariance. Grâce à cette expansion, un paramètre d'entrée dépendant de l'espace peut être étendu en une somme infinie où chaque terme est un produit entre un paramètre stochastique ne dépendant pas de l'espace, une constante (une valeur propre) et une fonction qui, elle, dépend de l'espace. Les paramètres stochastiques ne sont pas corrélés entre eux et cette somme peut être tronquée. En définitive, cette transformée permet d'étendre un paramètre d'entrée dépendant de l'espace en une fonction dépendant d'un nombre fini de variables aléatoires non corrélées entre elles. Au lieu d'utiliser le paramètre d'entrée dépendant de l'espace, les paramètres non corrélés peuvent être utilisés. Dans les travaux présentés dans ce manuscrit, aucune dépendance en espace n'est considérée.

2.5.3.4 Approche considérée

Dans les travaux, la transformation (139) est utilisée pour considérer les variables standards suivant une loi uniforme sur $[0, 1]$. Pour le modèle PC considéré pour la QOI, la base orthonormale de polynômes à une variable choisie est $(\phi_k^{0,1})$ pour chaque variable x_m . Le produit scalaire pour $a = 0$ et $b = 1$ correspond alors exactement au produit scalaire de la loi uniforme sur $[0, 1]$. Cette base est la base des polynômes de Legendre décalés normalisée. La base des polynômes multivariés orthonormale utilisée comme base des polynômes du chaos est alors $(\Phi_{\mathbf{k}}^{\mathbb{H}_d})$ et le modèle PC s'écrit, dans cette approche

$$\tilde{f}(\mathbf{x}) \approx \hat{f}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \Phi_{\mathbf{k}}^{\mathbb{H}_d}(\mathbf{x}) \quad (192)$$

Obtenir le modèle revient alors à obtenir les coefficients du modèle. Les méthodes d'obtention des coefficients se scindent en 2 grandes familles : les méthodes intrusives et les méthodes non-intrusives (Non Intrusive - NI). Un recueil de ces méthodes est disponible dans [74]. La connaissance des coefficients $\alpha_{\mathbf{k}}$ permet d'accéder analytiquement à des informations stochastiques sur \tilde{f} comme ses moments ou ses indices de Sobol, grâce aux propriétés d'orthogonalité de la base des polynômes.

2.5.3.5 Méthodes intrusives

Les méthodes de Galerkin sont les méthodes intrusives les plus couramment utilisées [75]. Le résidu de l'équation considérée pour obtenir la QOI est projeté sur la base des polynômes utilisée à l'aide du produit scalaire associé. Un système d'équations couplées sur les coefficients du chaos est alors obtenu, et sa résolution donne accès à ces derniers. L'utilisation des méthodes intrusives nécessitent la mise en place de ces nouvelles équations explicitement, mise en place pouvant s'avérer très complexe pour certaines QOI.

2.5.3.6 Méthodes non-intrusives

Les méthodes NI ont pour but d'obtenir les coefficients $\alpha_{\mathbf{k}}$ du métamodèle g-PC en lançant de nombreuses fois la simulation permettant d'obtenir la QOI avec des valeurs des paramètres d'entrée fixes différentes à chaque simulation. Un ensemble de valeur arrêté pour ces paramètres d'entrée avec la valeur associée du QOI est nommé dans le manuscrit une « observation », à l'instar des méthodes du domaine de l'apprentissage automatique (Machine Learning - ML). Dans la littérature, on retrouve également le terme de « nœud » (node). Les observations servant à obtenir les coefficients du modèle sont appelées les observations d'entraînement, et l'ensemble de ces dernières est appelé l'ensemble d'entraînement. Le nombre d'observations d'entraînement est noté N et l'ensemble d'entraînement est noté $\mathcal{T} = \{\mathbf{x}_i\}_{i=1}^N$. Trois méthodes NI sont rappelées ici, adaptées à l'approche considérée.

Projection pseudo-spectrale La projection pseudo-spectrale [73, 40] est une des méthodes NI se servant de l'orthogonalité de la base des polynômes. Elle utilise ces propriétés d'orthogonalité pour trouver une expression des coefficients. $\forall n \in [0, d], \forall \mathbf{k} \in \mathcal{K}_n^{(d)}$

$$\langle \tilde{f}, \Phi_{\mathbf{m}}^{\mathbb{H}_d} \rangle_{\mathbb{H}_d} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \langle \Phi_{\mathbf{k}}^{\mathbb{H}_d}, \Phi_{\mathbf{m}}^{\mathbb{H}_d} \rangle_{\mathbb{H}_d} = \alpha_{\mathbf{m}} \quad (193)$$

Des règles de quadrature sont ensuite utilisées pour évaluer l'intégrale présente dans la définition du produit scalaire. Cela implique que les observations d'entraînement aient une certaine structure, comme une grille de Gauss-Legendre ou une grille éparse de Smolyak [60] par exemple. Dans le contexte des présents travaux, les observations d'entraînement ne peuvent pas être structurées et l'utilisation de la méthode de pseudo-projection spectrale n'est pas possible.

Méthodes sur échantillonnage non structuré Deux méthodes NI sont capables de mener aux coefficients du métamodèle PC sur un échantillonnage d'observations d'entraînement non structuré : l'interpolation sur une grille non structurée et la régression linéaire. Contrairement à la méthode intrusive de Galerkin et à la méthode NI de pseudo-projection spectrale, ces méthodes n'utilisent pas les propriétés d'orthogonalité pour obtenir les coefficients. Ainsi, au lieu d'exprimer la QOI dans la base des $(\Phi_{\mathbf{k}}^{\mathbb{H}_d})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ (Cf. formule (192)), elle est exprimée tout d'abord dans la base canonique de $\mathbb{R}_n[X_1, \dots, X_d]$

$$\hat{f}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} c_{\mathbf{k}} \Psi_{\mathbf{k}}(\mathbf{x}) \quad (194)$$

Une fois les coefficients $c_{\mathbf{k}}$ obtenus, un changement de base vers la base des $(\Phi_{\mathbf{k}}^{\mathbb{H}^d})_{\mathbf{k} \in \mathcal{K}_n^{(d)}}$ peut permettre de trouver les $\alpha_{\mathbf{k}}$ du métamodèle PC (192) et les propriétés d'orthogonalité peuvent ainsi être utilisées pour déterminer des informations stochastiques sur la QOI.

Interpolation : L'interpolation sur une grille non structurée est développée dans [45]. Le métamodèle est forcé à être égal à la valeur de la QOI sur chaque observations d'entraînement. Cette méthode est notamment utilisée dans [29, 26]. Néanmoins, la régression linéaire est préférée dans ces travaux, afin de pouvoir bénéficier des avantages que cette dernière offre sur le calcul rapide d'une erreur de validation croisée.

Méthode des moindres carrés ordinaires : Dans les présents travaux, la méthode des moindres carrés ordinaires (Ordinary Least Square - OLS), une méthode de régression linéaire, est utilisée. Contrairement à l'interpolation, OLS n'implique pas que le métamodèle soit égal à la valeur de la QOI sur les observations d'entraînement. OLS vise à minimiser la distance euclidienne entre la valeur du QOI et la valeur du métamodèle sur ces observations.

$$\mathbf{C}_{opt} = \underset{\mathbf{C}}{\operatorname{argmin}} \|\mathbf{AC} - \mathbf{B}\|^2 \quad (195)$$

où $\mathbf{C} \in \mathbb{R}^M$ est le vecteur contenant les coefficients recherchés, $\mathbf{A} \in \mathcal{M}_{N,M}$ est la matrice d'information et $\mathbf{B} \in \mathbb{R}^N$ est le vecteur contenant les valeurs du QOI sur les échantillons d'entraînement. Dans l'objectif d'obtenir les coefficients du métamodèle exprimé dans la base canonique : $A_{ij} = \Psi_j^{(d)}(\mathbf{x}_i)$, $B_i = \tilde{f}(\mathbf{x}_i)$ et $C_j = c_j$

Le problème est alors équivalent à résoudre cette équation matricielle, souvent désignée par le terme d'« équation normale »

$$\mathbf{A}^T \mathbf{A} \mathbf{C} = \mathbf{A}^T \mathbf{B} \quad (196)$$

La solution de cette équation est

$$\mathbf{C}_{opt} = \mathbf{A}^+ \mathbf{B} \quad (197)$$

où $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$.

Le nombre M , nombre de composantes du polynôme, déterminé dans le cas des métamodèles du chaos par le degré de troncature choisi n et la dimension d d'après l'équation (169), est une variable cruciale pour la régression. Comparé au nombre d'observations d'entraînement N , si M est trop faible, la régression manque de précision (phénomène souvent appelé « sous-apprentissage » ou « biais » dans le domaine ML). Si M est trop élevé, le phénomène appelé « surapprentissage » ou « variance » dans le domaine ML peut apparaître. Ce dernier phénomène traduit le fait que le métamodèle soit précis sur les observations d'entraînement mais ne se généralise pas bien sur les autres observations, menant ainsi à un métamodèle de faible précision entre les observations d'entraînement.

Ces aspects peuvent être pris en compte en testant toutes les régressions avec un degré de troncature entre les degrés n_{min} et n_{max} choisis et en sélectionnant le modèle optimal comme celui qui minimise une erreur de validation croisée. L'erreur considérée doit être définie entre les valeurs du métamodèle et les vraies valeurs de la QOI sur les observations d'entraînement. Ainsi, cette erreur peut être utilisée lors de l'apprentissage pour sélectionner le degré de troncature menant au métamodèle optimal, ou plus généralement pour sélectionner les meilleurs hyperparamètres à utiliser parmi une série de valeurs proposées. Utiliser l'erreur moyenne quadratique (Mean Square Error - MSE) entre les valeurs du QOI et celles du métamodèle sur les observations d'entraînement donnerait une image faussée de la précision dans le cas du surapprentissage. L'utilisation de la validation croisée peut convenir pour éviter ce phénomène. L'erreur PRESS (pour Predicted Residual Sum of Squares), utilisée dans [8] dans le cadre des polynômes du chaos, est très efficace dans le contexte des métamodèles PC entraînés avec la méthode OLS. Elle est basée sur la validation croisée avec omission d'une valeur

$$\delta = \frac{1}{N} \sum_{i=1}^N \left(\hat{f}^{(i)}(\mathbf{x}_i) - \tilde{f}(\mathbf{x}_i) \right)^2 \quad (198)$$

où $\hat{f}^{(i)}$ est le métamodèle appris sur l'ensemble d'entraînement $\mathcal{T} \setminus \{\mathbf{x}_i\}$.

Un des avantages de l'utilisation de la méthode OLS est qu'il existe une formule analytique pour exprimer chaque terme de la somme sans avoir à ré-entraîner un nouveau modèle sur chaque ensemble d'entraînement $\mathcal{T} \setminus \{\mathbf{x}_i\}$ pour chaque observation \mathbf{x}_i . Cette formule [56], qui ne nécessite que l'apprentissage du modèle sur \mathcal{T} , s'écrit

$$\delta = \frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{f}(\mathbf{x}_i) - \tilde{f}(\mathbf{x}_i)}{1 - h_i} \right)^2 \quad (199)$$

où h_i est le i^{eme} terme diagonal de la matrice $\mathbf{H} = \mathbf{A}\mathbf{A}^+$ connue sous le nom de « matrice chapeau ». Cette erreur reflète correctement la précision du métamodèle, même en présence de surapprentissage.

Afin d'obtenir numériquement \mathbf{A}^+ , l'erreur PRESS et la matrice chapeau, la décomposition en valeur singulière (Singular Vector Decomposition - SVD) [32] de \mathbf{A} est utilisée. Si \mathbf{A} est une matrice de taille $N \times M$, trouver la SVD de \mathbf{A} consiste à trouver une matrice \mathbf{U} de taille $N \times N$, une matrice diagonale rectangulaire $\mathbf{\Sigma}$ de taille $N \times M$ ainsi qu'une matrice orthogonale \mathbf{V} de taille $M \times M$ tel que

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (200)$$

Les colonnes de \mathbf{U} sont appelés « vecteurs singuliers à gauche » de \mathbf{A} , ce sont les vecteurs propres de $\mathbf{A}\mathbf{A}^T$. Les colonnes de \mathbf{V} sont appelés « vecteurs singuliers à droite », ce sont les vecteurs propres de $\mathbf{A}^T\mathbf{A}$. Les termes diagonaux de la matrice $\mathbf{\Sigma}$ sont les « valeurs singulières » de \mathbf{A} , ce sont les racines carrées des valeurs propres de $\mathbf{A}\mathbf{A}^T$. Une convention prise est de classer les valeurs singulières sur la diagonale de $\mathbf{\Sigma}$ par ordre décroissant. Si

\mathbf{A} est de rang r , \mathbf{A} a r valeurs singulières non nulles.

La SVD est très pratique pour obtenir numériquement \mathbf{A}^+ . Il est prouvé dans [32] que

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T \quad (201)$$

où $\mathbf{\Sigma}^+$ est la matrice diagonale rectangulaire de taille $M \times N$ contenant les inverses des valeurs singulières non nulles sur sa diagonale et des zéros autre part.

La SVD permet également d'obtenir efficacement la matrice chapeau, et donc l'erreur PRESS. En effet,

$$\mathbf{H} = \mathbf{A}\mathbf{A}^+ = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^+\mathbf{U}^T = \mathbf{U}_r\mathbf{U}_r^T \quad (202)$$

où \mathbf{U}_r est la matrice carrée contenant les r premières lignes et colonnes de \mathbf{U} . La diagonale de \mathbf{H} , nécessaire pour obtenir l'erreur PRESS, est alors égale

$$h_{i,i} = \sum_{j=1}^r [(U_r)_{i,j}]^2 \quad (203)$$

Ainsi, l'entièreté des matrices \mathbf{U} et \mathbf{V} n'est pas nécessaire pour obtenir \mathbf{A}^+ et la diagonale de \mathbf{H} . Seules les r premières colonnes de \mathbf{U} , de $\mathbf{\Sigma}$ et les r premières lignes de \mathbf{V}^T (correspondant aux valeurs singulières non nulles) sont nécessaires. Obtenir ces matrices réduites est connu sous le nom de « SVD compacte ».

Algorithm 1 Pseudo-code appliquant l'OLS et obtenant l'erreur PRESS correspondante

$\mathbf{A} \leftarrow N \times M$ matrice d'information

$\mathbf{B} \leftarrow$ vecteur de taille N contenant les valeurs de la QOI sur les observations d'entraînement.

$\mathbf{U}, \mathbf{\Sigma}, \mathbf{V} \leftarrow SVD(\mathbf{A})$ (forme compacte)

$\mathbf{\Sigma}^+ \leftarrow$ Matrice de taille $\min(N, M) \times \min(N, M)$ avec $\frac{1}{\sigma_{i,i}}$ sur sa diagonale et 0 autre part

$\mathbf{X} \leftarrow \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T\mathbf{B}$

$\mathbf{H} \leftarrow \sum_{j=1}^r [(U)_{i,j}]^2$ (Ici \mathbf{H} est le vecteur contenant les termes diagonaux de la matrice chapeau et non pas la matrice en elle-même)

$\mathbf{Y} = \mathbf{A}\mathbf{X}$

$\delta \leftarrow \frac{1}{N} \sum_{i=1}^N \left(\frac{Y_i - B_i}{1 - H_i} \right)^2$

Return \mathbf{X}, δ

L'Algorithme 1 présente le pseudo-code permettant d'effectuer numériquement l'OLS et d'obtenir l'erreur PRESS correspondante en utilisant les méthodes décrites précédemment.

Si \mathbf{A} est mal conditionnée, l'utilisateur peut ignorer un ensemble de valeurs singulières ayant une valeur inférieure à un seuil donné, ne prenant en compte que $\hat{r} \leq r$ valeurs singulières à la place de r .

On considère $N \geq M$, sans perte de généralité (si $M > N$, les opérations peuvent être conduites sur la transposée de la matrice). Dans le cas le plus coûteux, $r = M$. La complexité temporelle d'un produit matriciel entre une matrice \mathbf{A} de taille $N \times M$ et une matrice \mathbf{B} de taille $M \times P$ est en $\mathcal{O}(M^3 + M^2N + MN) = \mathcal{O}(M^2N)$. D'après les connaissances de l'auteur, la complexité de la SVD compacte est en $\mathcal{O}(M^2N)$. Ainsi, la complexité temporelle de l'algorithme OLS présenté ici est en $\mathcal{O}(M^2N)$.

2.5.3.7 Utilisation des propriétés d'orthogonalité

Une fois le métamodèle PC obtenu, l'utilisation de ses propriétés d'orthogonalité permet d'accéder à des informations sur la QOI.

En effet, son espérance est directement estimée par la valeur du premier coefficient $\alpha_{\mathbf{0}}$

$$E(\tilde{f}) \approx E(\hat{f}) = \langle \hat{f}, 1 \rangle_{\mathbb{H}_d} = \langle \hat{f}, \Phi_{\mathbf{0}}^{\mathbb{H}_d} \rangle_{\mathbb{H}_d} = \alpha_{\mathbf{0}} \quad (204)$$

La variance est quant à elle estimée par

$$D(\tilde{f}) \approx D(\hat{f}) = \langle \hat{f}, \hat{f} \rangle_{\mathbb{H}_d} - \alpha_{\mathbf{0}}^2 = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)} \setminus \{\mathbf{0}\}} \alpha_{\mathbf{k}}^2 \quad (205)$$

les propriétés d'orthogonalité peuvent également permettre d'accéder aux indices de Sobol et indices de Sobol totaux [64, 15]. En effet, le terme $\tilde{f}_{\mathbf{i}}^{(j)}$ de la décomposition ANOVA (152) pour le métamodèle g-PC \hat{f} est identifié comme étant

$$\tilde{f}_{\mathbf{i}}^{(j)} \approx \hat{f}_{\mathbf{i}}^{(j)} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \Phi_{\mathbf{k}}^{\mathbb{H}_d} \tilde{\Delta}_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \quad (206)$$

Cette fonction est la somme des termes de l'expansion PC traduisant les effets de l'interaction \mathbf{i} . Ces termes sont ceux qui dépendent de l'intégralité des j x_m indexés par l'interaction \mathbf{i} et seulement de ce groupe de j paramètres. Cela se traduit par le fait d'être un monôme indexé par un indice \mathbf{k} dont les composantes indexés par \mathbf{i} sont non nulles et dont toutes les autres sont nulles. Cette propriété est modélisée mathématiquement par $\tilde{\Delta}_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}}$. La somme de ces termes donne bien \hat{f} , et ces termes respectent la propriété (154). Par unicité, il s'agit bien des termes de la décomposition ANOVA.

Ainsi, d'après (159)

$$D_{\mathbf{i}}^{(j)} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} (\alpha_{\mathbf{k}})^2 \tilde{\Delta}_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \quad (207)$$

Comme la quantité $D_i^{(T)}$ est la somme de tous les $D_{\mathbf{i}}^{(j)}$ où i apparait, il s'agit de la somme de toutes les composantes ayant une dépendance en x_i .

$$D_i^{(T)} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} (\alpha_{\mathbf{k}})^2 (1 - \tilde{\Delta}_{\mathbf{k}, \mathbf{0}}^{S(\mathbf{i})}) \quad (208)$$

Les indices de Sobol peuvent alors être estimés une fois les coefficients du chaos obtenus. Contrairement aux méthodes MC (Cf. section 2.4.3.5), qui nécessitent de nouvelles observations pour estimer chaque $D_{\mathbf{i}}^{(j)}$, l'intégralité des observations peut être utilisée pour apprendre les coefficients du métamodèle PC, qui permet ensuite de déterminer l'intégralité des $D_{\mathbf{i}}^{(j)}$. Un meilleur balayage de l'espace est alors possible et le coût d'obtention des indices de Sobol est réduit.

2.5.3.8 Limites du chaos polynomial

Les métamodèles PC peuvent approcher une QOI présentant une dépendance assez régulière par rapport aux paramètres d'entrée avec précision. Néanmoins, ils se heurtent à deux limitations.

- Le nombre de composantes du polynôme est sujet à la malédiction de la dimension, évoluant d'après la formule (169). Pour un grand nombre de paramètres d'entrée (en haute dimension), utiliser un degré de troncature suffisamment élevé peut devenir trop coûteux.
- La précision du métamodèle est altérée pour une QOI possédant des irrégularités ou des discontinuités dans son évolution par rapport aux paramètres d'entrée. La Figure 8 illustre cette perte de précision pour une QOI constante par morceaux $\tilde{f}^{1d-discontinuous}$ ne dépendant que d'une variable. Le phénomène de Gibbs induit que, même en augmentant le degré de troncature du métamodèle, les oscillations proches de la discontinuités perdurent et altèrent la précision.

$$\tilde{f}^{1d-discontinuous}(x) = \begin{cases} -1 & \text{if } x < \frac{1}{2} \\ 1 & \text{otherwise} \end{cases} \quad (209)$$

Le système dynamique du vol qui est étudié dans les travaux présente de nombreuses bifurcations, bifurcations qui se traduisent par des irrégularités et des discontinuités dans la dépendance de la QOI par rapport aux paramètres d'entrée à un certain instant. Il est donc nécessaire d'utiliser une variante des métamodèles PC pour obtenir un métamodèle précis.

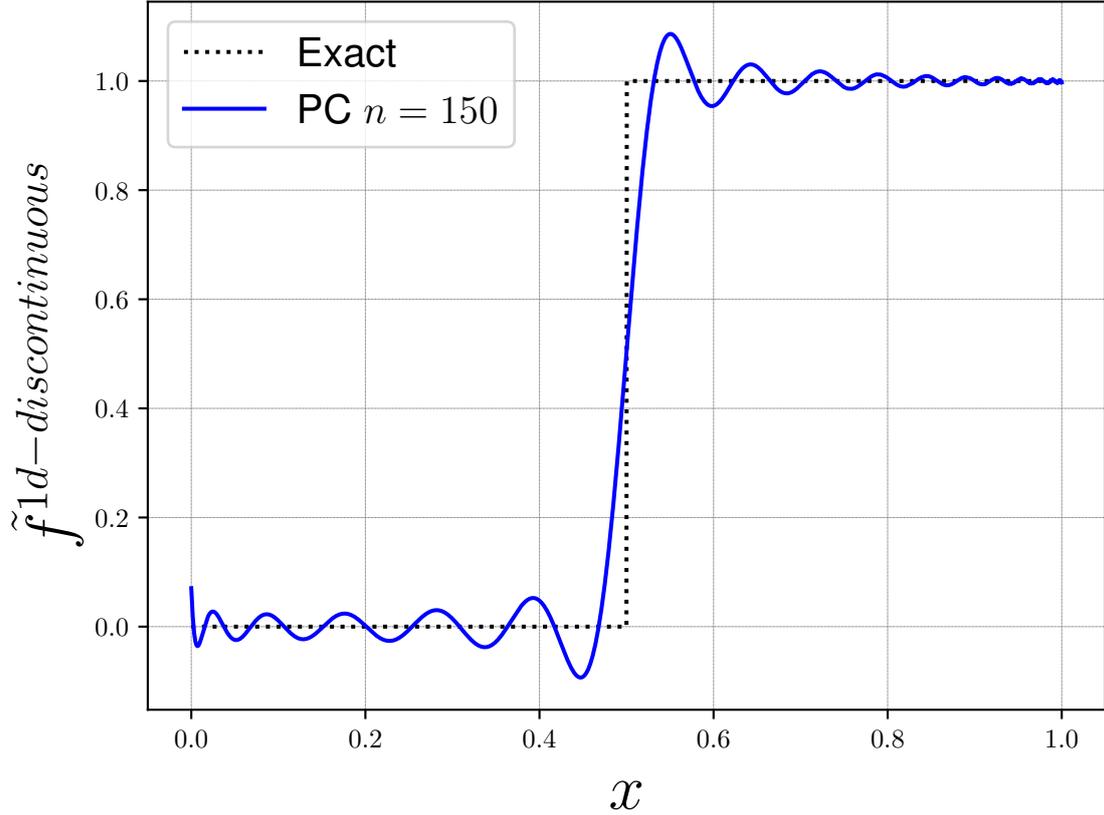


FIGURE 8 – Tracé de $\tilde{f}^{1d-discontinuous}(x)$ (en pointillés) et de son approximation PC $\hat{f}^{1d-discontinuous}(x)$ (en trait plein) avec un degré $n = 150$.

2.6 Améliorations possibles des polynômes du chaos

2.6.1 Polynômes du chaos arbitraires

Cette amélioration [47] suit la généralisation g-PC afin de proposer une base de polynômes orthogonaux pour des paramètres d'entrée suivant une distribution quelconque. Cette approximation porte le nom d'« arbitrary Polynomial Chaos » (aPC). Ce choix répond au fait que la connaissance sur les données des paramètres d'entrée et de la QOI n'est pas complète. La distribution, ainsi que de potentielles dépendances entre les paramètres d'entrée, peuvent ne pas être connues. La pdf de ces paramètres d'entrée n'est donc pas approchée par une loi commune et n'a pas besoin d'être connue ni explicitée. Seul un jeu de données est nécessaire. Dans [47], les moments stochastiques des données sont utilisés pour construire la base. L'article cite également d'autres méthodes pour obtenir cette base, comme l'algorithme d'orthogonalisation de Gram-Schmidt. Dans notre étude, les distributions des variables modélisant les incertitudes et les lacunes sont estimées et leur pdfs sont approchées par des lois communes. La méthode aPC n'est donc pas utilisée.

2.6.2 Polynômes du chaos épars

Pour contrer la malédiction de la dimension, il est proposé dans [8] de ne pas utiliser tous les termes de la série correspondant à un $\mathbf{k} \in \mathcal{K}_n^{(d)}$. Dans l'article, une méthode est suggérée pour ignorer certains termes de haut degré concernant les interactions d'ordre supérieur, ayant moins de chance d'avoir de l'impact sur la QOI. De plus, une méthode de sélection itérative des termes du modèle, basée sur la méthode de « Least Angle Regression », est proposée. Les polynômes du chaos épars ne sont pas utilisés dans l'étude. Comme des métamodèles ME sont privilégiés, le degré des polynômes locaux peut être relativement faible, ce qui contre déjà la malédiction de la dimension. L'utilisation des polynômes du chaos épars pourra représenter un axe d'amélioration des modèles développés par la suite.

2.6.3 Utilisation d'une autre base

Pour contrer la perte de précision des modèles g-PC pour une QOI irrégulière ou discontinue, des métamodèles ayant une structure non polynomiale mais proche des polynômes du chaos sont envisagés.

Dans [14], le modèle envisagé est un quotient de deux séries de polynômes. La forme de la fonction permet un apprentissage rapide du méta-modèle, mais contrairement aux modèles PC, aucun lien analytique ne relie les coefficients du modèles aux informations stochastiques. Ce métamodèle est utilisé pour obtenir des échantillons plus rapidement lors des méthodes MC. Dans l'étude du système dynamique du vol, il est souhaité que le métamodèle ait un lien analytique avec les indices de Sobol pour pouvoir être utilisé afin de réaliser une analyse de sensibilité. Cette base n'est donc pas utilisée dans le cas de l'étude.

Dans [39], une base orthogonale de fonctions sous forme des « ondelettes » de Haar est envisagée. À l'instar des polynômes du chaos, les propriétés orthogonales permettent d'accéder analytiquement aux propriétés stochastiques de la QOI. Néanmoins, d'après [39], l'approximation offre un métamodèle précis en présence de discontinuités, mais est surpassée par les polynômes du chaos standard dans le cas d'une QOI régulière. Dans l'étude du système dynamique du vol, comme l'évolution temporelle des variables d'état est approchée, ces variables sont régulières dans les premiers instants, puis peuvent devenir irrégulières et/ou discontinues à des instants plus avancés. L'utilisation des ondelettes n'est donc pas considérée pour l'étude.

Une solution pour approcher une QOI irrégulière ou discontinue est d'utiliser un métamodèle ME. Il s'agit d'un modèle polynomial par morceaux. Ils sont décrits plus en détail à la section 2.6.4.

2.6.4 Métamodèles multi-éléments

Les métamodèles ME consistent à partitionner l'espace des paramètres \mathbb{H}_d en $N_E \in \mathbb{N}^*$ sous-ensembles appelés éléments.

$$\mathbb{H}_d = \bigcup_{e=1}^{N_E} E_e \quad (210)$$

$$E_{e_1} \cap E_{e_2} = \emptyset \text{ if } e_1 \neq e_2 \quad (211)$$

Une fois la partition effectuée, un métamodèle PC local \hat{f}_e est entraîné sur chaque élément E_e sur les observations d'entraînement situés dans l'élément. Le métamodèle ME-PC global est donc un modèle par morceaux composé de tous les modèles locaux. Comme les propriétés d'orthogonalité ne sont pas utilisées pour apprendre le modèle, la base canonique de $\mathbb{R}[X_1, \dots, X_d]$ est utilisée pour les modèles locaux.

$$\hat{f}(\mathbf{x}) = \sum_{e=1}^{N_E} \hat{f}_e(\mathbf{x}) \mathcal{I}_{E_e}(\mathbf{x}) = \sum_{e=1}^{N_E} \left(\sum_{j=1}^{M_e} c_{ej} \Psi_j^{(d)}(\mathbf{x}) \right) \mathcal{I}_{E_e}(\mathbf{x}) \quad (212)$$

où M_e est le nombre de termes de la série polynomiale locale \hat{f}_e et \mathcal{I}_{E_e} est la fonction indicatrice de l'élément E_e satisfaisant

$$\mathcal{I}_{E_e}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in E_e \\ 0 & \text{otherwise} \end{cases} \quad (213)$$

Dans les applications numériques, le rôle de la fonction indicatrice peut être réalisé par une condition si les équations des frontières entre les éléments sont connues, ou ce rôle peut être joué par un classifieur. L'annexe A discute des différents classifieurs en détail. Une fois les métamodèles locaux et le potentiel classifieur entraînés, pour évaluer le modèle sur un nouveau point (représentant un nouveau jeu de paramètres d'entrée), l'élément E_e dans lequel le point se trouve est détecté et le métamodèle local correspondant est utilisé pour effectuer l'évaluation.

La figure 9 reprend l'exemple de la QOI à une variable $\tilde{f}^{1d-discontinuous}(x)$ définie à la section 2.5.3.7, avec cette fois-ci un métamodèle ME. Le phénomène de Gibbs n'est pas présent et les degrés des modèles locaux sont de 0.

La première méthode ME, ME-gPC, est présentée dans [69]. Elle repose sur une découpe de l'espace des paramètres en hyperpavés de plus en plus petits obtenus par dichotomies successives si un critère de précision n'est pas respecté. Les coefficients sont obtenus à l'aide d'une méthode intrusive. Son équivalent NI, ME-PCM, est développée dans [19]. Avec ces découpes successives, les irrégularités et les discontinuités sont confinées dans des hyperpavés de plus en plus petits, ce qui a moins d'impact sur la précision. Le nombre d'éléments évolue néanmoins de façon importante avec la dimension d et le nombre de dichotomie réalisée.

Plus récemment, des méthodes ME ont été développées pour éviter les découpes successives en hyperpavés de plus en plus petits, en partitionnant l'espace des paramètres en suivant cette fois-ci la forme des irrégularités ou des discontinuités. Les éléments ne sont donc plus forcés d'être des hyperpavés. Dans [57], les auteurs utilisent l'inférence

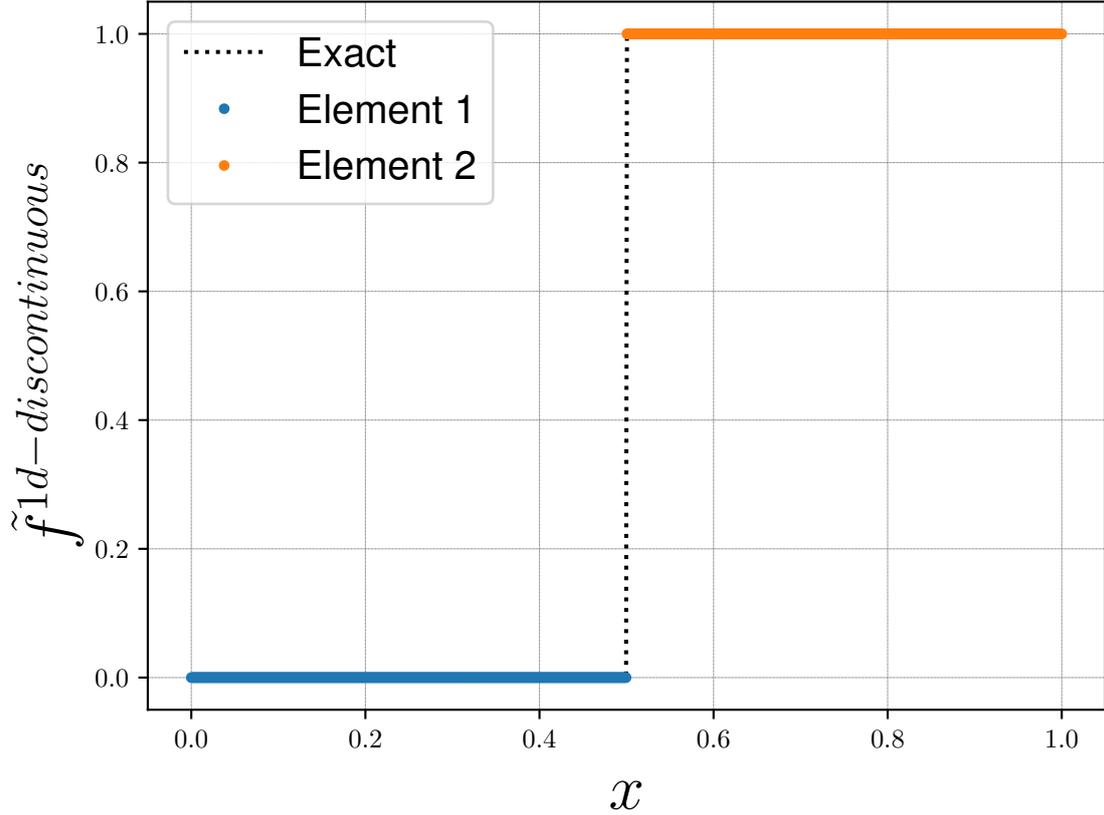


FIGURE 9 – Tracé de $\tilde{f}^{1d-discontinuous}(x)$ (en pointillés) et de son approximation ME (en trait plein) composé de deux modèles locaux de degré $n = 0$.

bayésienne pour paramétrer une équation de la discontinuité ou de l'irrégularité et utilisent ensuite la transformation de Rosenblatt comme transformation bijective de chaque élément en hyperpavé afin d'obtenir les coefficients des métamodèles locaux par projection pseudo-spectrale. Dans [29], les auteurs proposent une méthode basée sur un échantillonnage adaptatif ainsi que sur la détection de discontinuité qui permet également une partition de l'espace des paramètres sur des éléments quelconques épousant la forme des discontinuités. L'utilisation de l'interpolation sur une grille non structurée permet de ne pas avoir à utiliser de transformation bijective vers un hyperpavé. La seconde méthode est améliorée dans [26], proposant un échantillonnage adaptatif nécessitant moins d'observations et utilisant cette fois-ci une machine à vecteurs de support (Support Vector Machine -SVM, Cf Annexe A) comme classifieur. Néanmoins, la détection de discontinuité se focalise sur les discontinuités et non plus sur les irrégularités comme les premières méthodes ME. De plus, comme soulevé dans [26], la méthode ne permet pas de détecter des discontinuités ne parcourant que partiellement l'espace des paramètres.

Le système dynamique du vol est un système pouvant contenir de nombreuses bifurcations. L'objectif est d'étudier l'évolution temporelle des sensibilités de chaque variables d'état par rapport à des paramètres traduisant des lacunes et des incertitudes. Aux premiers instants, les variables d'état sont régulières, et avec le temps, des irrégularités risquent d'apparaître et risquent de se transformer petit à petit en discontinuités avec le temps. Afin d'éviter la multiplication des hyperpavés dans le cas d'une discontinuité ou

d'une irrégularité non alignée avec les axes cartésiens, l'utilisation d'une méthode moderne est privilégiée. Néanmoins, la détection des discontinuités ne se focalise que sur les discontinuités et non pas sur les irrégularités. Les méthodes modernes sont souvent utilisées sur des QOI indépendantes du temps possédant une ou plusieurs discontinuités. Dans [29], la méthode n'est testée que sur des QOI théoriques discontinues et indépendantes du temps. Dans [26], la méthode est utilisée pour étudier le « problème de rupture en barrage fluide » à un instant, qui est potentiellement discontinu. De plus, ces méthodes reposent sur des échantillonnage adaptatifs.

Les travaux de thèse ont alors donné naissance à deux objectifs dépassant le cadre de la dynamique du vol pour répondre à deux problématiques soulevées concernant les méthodes ME :

- Les méthodes actuelles possèdent des défauts qui ont été soulevés au paragraphe précédent. Les méthodes classiques ME confinent les irrégularités et les discontinuités, mais peuvent engendrer un nombre de dichotomies important. Les méthodes reposant sur la détection de discontinuité se focalise sur les discontinuités et non plus sur les irrégularités et reposent sur un échantillonnage adaptatif. Le Chapitre 3 présente alors le développement d'une nouvelle méthode permettant d'obtenir un métamodèle ME. La méthode est capable de partitionner l'espace des paramètres en épousant la forme des discontinuités, mais également des irrégularités, même si ces dernières ne parcourent que partiellement l'espace. La méthode a également pour but de fonctionner avec un échantillonnage quelconque. Dans cette méthode, chaque coefficient de chaque métamodèle local \hat{f}_e est obtenu avec l'algorithme OLS (Cf. section 2.5.3.6) sur les observations d'entraînement situées dans E_e . Un classifieur sous forme de réseau de neurones est utilisé pour jouer le rôle de la fonction indicatrice. Comme OLS fonctionne sur une grille d'observations d'entraînement non structurée, aucune transformation bijective vers des variables locales n'est utilisée. Cette méthode a pour but d'améliorer les algorithmes actuels d'obtention de métamodèles ME, afin qu'il puisse être utilisé dans de nombreux domaines. Il est de plus conçu pour être adapté à l'approximation des variables d'état du système dynamique du vol.
- Pour être utilisées, les propriétés d'orthogonalité nécessitent une intégration sur l'ensemble de l'espace des paramètres, d'après la définition du produit scalaire associé à ρ . Comme les métamodèles ME sont définis avec des modèles locaux, l'orthogonalité ne peut pas être utilisé pour trouver les indices de Sobol. Le Chapitre 4 développe une méthodologie pour palier à ce problème et approcher les indices de Sobol à partir des coefficients des polynômes locaux d'un métamodèle ME.

Chapitre 3

Développement de la méthode ME-ACD

3.1 Introduction

ME-ACD est un algorithme développé dans [66] dans le contexte de la thèse. Ce chapitre est une version traduite et légèrement modifiée de cet article afin d'être intégrée au manuscrit. L'algorithme a pour but d'obtenir un métamodèle multi-elements (Multi-Element - ME) en utilisant les méthodes modernes d'apprentissage automatique (Machine Learning - ML). Il suit alors les logiques de [57] et [29] en partitionnant l'espace des paramètres stochastiques suivant la forme des discontinuités et des irrégularités. La partition est réalisée en utilisant le regroupement hiérarchique ascendant sur les observations d'entraînement représentées dans un espace prenant en compte la valeur des dérivées partielles de la quantité d'intérêt (Quantity Of Interest - QOI) par rapport aux paramètres d'entrée en ces points. Le regroupement hiérarchique a déjà été utilisé dans le contexte des modèles ME dans [58] et [50], mais il est mis en œuvre sur les observations d'entraînement représentées par des points dans un espace prenant en compte la valeur des paramètres d'entrée et la valeur de la QOI, mais pas celle de ses dérivées partielles.

3.2 Algorithme ME-ACD

3.2.1 Espaces abstraits

Dans la méthode ME-ACD, les observations sont considérées dans différents espaces abstraits :

- \mathcal{P} : l'espace des paramètres originel de dimension d où les observations ont pour coordonnées (p_1, \dots, p_d) .
- \mathbb{H}_d : l'espace des paramètres après la transformation bijective introduite à la section 2.2 où les observations ont des coordonnées (x_1, \dots, x_d) . Cet espace, qui est

l'hypercube unité, est celui désigné par « espace des paramètres ».

- \mathcal{F} : l'espace de dimension $d + 1$ où les observations ont pour coordonnées $(x_1, \dots, x_d, \tilde{f}(\mathbf{x}))$. La valeur de la QOI est prise en compte dans cette représentation.
- \mathcal{D} : l'espace de dimension $d + 1$ où les observations ont pour coordonnées $(\tilde{f}(\mathbf{x}), \frac{\partial \tilde{f}}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial \tilde{f}}{\partial x_d}(\mathbf{x}))$. La valeur de la QOI et de ses dérivées partielles par rapport aux paramètres d'entrée sont prises en compte dans cette représentation.

De ce fait, une observation peut être représentée dans ces différents espaces. L'objectif des méthodes ME est de partitionner l'espace des paramètres \mathbb{H}_d en N_E sous-ensembles E_e pour ensuite entraîner des métamodèles locaux sur chacun de ces éléments. Pour effectuer cette partition, le regroupement hiérarchique ascendant est utilisé. Le résultat de ce regroupement est différent selon l'espace dans lequel les observations sont représentées. L'espace doit être sélectionné pour que le regroupement hiérarchique ascendant induise une partition efficace de l'espace des paramètres en présence de discontinuités et en présence d'irrégularités. L'espace sélectionné est alors \mathcal{D} , mais les autres espaces interviennent également dans l'algorithme.

3.2.2 Étapes de l'algorithme

ME-ACD est un algorithme en 5 étapes :

- Étape 1 - Échantillonnage de l'espace des paramètres : l'espace des paramètres est échantillonné avec un échantillonnage au choix de l'utilisateur. La valeur de la QOI est ensuite obtenue sur chaque échantillon. Les observations peuvent alors être représentées dans les espaces \mathcal{P} , \mathbb{H}_d et \mathcal{F} .
- Étape 2 - Génération de la matrice d'information globale : les métamodèles locaux sont obtenus à l'aide de la méthode des moindres carrés (Ordinary Least Square - OLS). Mais pour chaque élément et chaque degré de métamodèle polynomial testé, la création d'une nouvelle matrice d'information n'est pas nécessaire. Ces matrices sont les sous-matrices d'une matrice d'information prenant en compte toutes les observations d'entraînement et le degré maximal n_{max} considéré. Cette matrice, dite globale, est générée une fois pour toute lors de cette étape de prétraitement.
- Étape 3 - Estimation des dérivées partielles : afin de pouvoir représenter les observations dans l'espace \mathcal{D} pour utiliser le regroupement hiérarchique ascendant, il nous faut obtenir les valeurs des dérivées partielles de la QOI par rapport aux paramètres d'entrée sur les échantillons. Cette étape de prétraitement utilise alors la méthode OLS pour estimer la valeur de ces dérivées partielles.
- Étape 4 - Partitions de l'espace des paramètres et entraînement des modèles locaux : cette étape est le cœur de l'algorithme. En utilisant le regroupement hiérarchique ascendant sur les observations représentées dans l'espace \mathcal{D} , des groupes d'observations sont créés. Ces groupes représentent les éléments de la partition de l'espace des paramètres. Des partitions possédant entre $N_{E_{min}}$ et $N_{E_{max}}$ éléments sont testées, et pour chaque élément de chaque partition, des métamodèles polynomiaux de degré

n_{min} à n_{max} sont testés. Le degré est sélectionné comme étant celui correspondant au métamodèle locale menant à une erreur PRESS minimale. La partition minimisant le maximum des erreurs PRESS des métamodèles locaux de degré optimal est sélectionnée.

- Étape 5 - Entraînement du classifieur : une fois une partition obtenue à l'étape 4, un classifieur est entraîné afin de savoir quel métamodèle local utiliser lors d'un appel au métamodèle ME sur une combinaison de paramètres d'entrée ne faisant pas partie des observations d'entraînement.

Ces 5 étapes sont décrites plus en détail dans les sections suivantes.

3.2.3 Étape 1 - Échantillonnage de l'espace des paramètres

La première étape consiste à échantillonner l'espace des paramètres \mathbb{H}_d afin d'obtenir l'ensemble d'entraînement $\mathcal{T} = \{\mathbf{x}_i\}_{i=1}^N$ de N observations d'entraînement. Un des objectifs de ME-ACD est de fonctionner sur chaque échantillonnage, pour qu'il puisse être utilisé sur des données expérimentales par exemple. En ce sens, la méthode ne repose pas sur un échantillonnage adaptatif, contrairement aux méthodes [29, 26]. Dans ces travaux, à des fins théoriques, les séquences quasi-aléatoires de Sobol [61] sont utilisées. Tout comme l'échantillonnage par hypercube latin, cette séquence a pour but d'échantillonner l'espace de dimension d en laissant moins de zones éparses dans l'hypercube que les méthodes de Monte-Carlo. Plus d'informations concernant cet échantillonnage sont disponibles dans [31]. Les nombre initiaux utilisés pour générer ces séquences sont ceux suggérés par les auteurs et sont disponibles sur le site web : <https://web.maths.unsw.edu.au/~fkuo/sobol/> (dernier accès : 01/04/2022).

Avec cette stratégie, l'hypercube unité \mathbb{H}_d est échantillonné et les \mathbf{p}_i correspondant aux observations d'entraînement \mathbf{x}_i sont obtenus en inversant la formule (139). Dans certains cas, une modification légère est de rigueur pour les observations localisées sur les hyperfaces de l'hypercube. Pour éviter qu'un \mathbf{p}_i correspondant ait une coordonnée infinie pour une loi de distribution normale par exemple, le \mathbf{x}_i peut être translaté d'une distance de $\epsilon = 10^{-3}$ par rapport aux hyperfaces sur lesquelles il est normalement situé. Cela se traduit par l'ajout de ϵ à toutes les composantes nulles des \mathbf{x}_i et d'un retrait de ϵ à toutes les composantes égales à l'unité des \mathbf{x}_i .

La figure 10 montre un échantillonnage de Monte-Carlo ainsi qu'un échantillonnage de Sobol pour un espace des paramètres $2D$. 500 observations de \mathbf{p} avec p_1 suivant une loi uniforme sur $[1, 1.5]$ et p_2 suivant une loi normale de moyenne 10 et d'écart-type 5 échantillonnées avec les deux méthodes sont représentées dans \mathbb{H}_d et dans \mathcal{P} . Les échantillonnages sont très proches, mais l'échantillonnage de Sobol est élaboré pour garantir une meilleure homogénéité.

Une fois l'échantillonnage réalisé, la QOI est évaluée sur toutes les observations d'entraînement de \mathcal{T} pour obtenir un vecteur $\mathbf{B} \in \mathbb{R}^N$ avec :

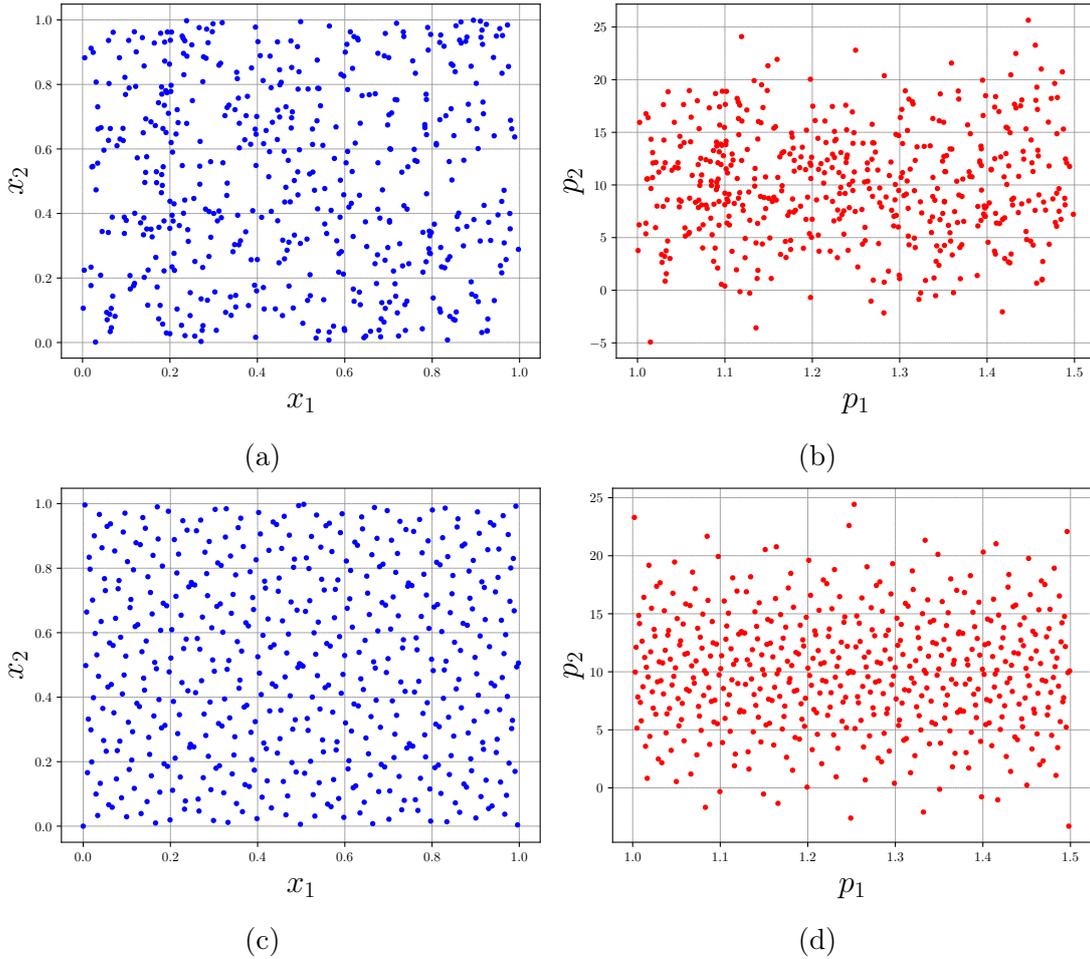


FIGURE 10 – $N = 500$ observations générées par les deux méthodes d'échantillonnage. (a) Échantillonnage de Monte-Carlo dans \mathbb{H}_d . (b) Transformation de l'échantillonnage de Monte-Carlo dans \mathcal{P} . (c) Échantillonnage de Sobol dans \mathbb{H}_d . (d) Transformation de l'échantillonnage de Sobol dans \mathcal{P} .

$$B_i = \tilde{f}(\mathbf{x}_i) = f(\mathbf{p}_i) \quad (214)$$

La complexité temporelle de la génération d'une séquence de Sobol est en $\mathcal{O}(N \times d)$. La complexité temporelle de l'évaluation de la QOI est indépendante de l'algorithme.

Cette étape est optionnelle si l'algorithme est utilisé sur des données déjà obtenues ; l'échantillonnage étant déjà réalisé.

3.2.4 Étape 2 - Génération de la matrice d'information globale

Lors de l'étape 4 de l'algorithme, décrite ci-après, l'espace des paramètres \mathbb{H}_d est partitionné et la méthode OLS est utilisée avec des métamodèles polynomiaux de degré allant de n_{min} à n_{max} pour trouver les coefficients des métamodèles locaux. Un avantage majeur de l'utilisation de la méthode OLS dans l'approche adoptée est qu'il n'est nécessaire de générer qu'une unique matrice d'information. Cette matrice est nommée matrice d'infor-

mation globale. Quand la méthode OLS est appelée pour apprendre les coefficients d'un métamodèle local, seule une sous-matrice de cette matrice globale sera utilisée. Les lignes de cette sous-matrice sont les lignes correspondantes aux observations d'entraînement situées dans l'élément. Ses colonnes sont les premières $M(n, d)$ colonnes selon le degré n du métamodèle considéré. Cette matrice d'information globale, noté \mathbf{A} , est une matrice de taille $N \times M(n_{max}, d)$. Les termes de cette matrice s'écrivent :

$$A_{i,j} = \Psi_j(\mathbf{x}_i) \quad (215)$$

où ici j est un indice scalaire correspondant à un certain multi-indice \mathbf{k} (Cf. section 2.5.1.2)

$$\Psi_j = \prod_{m=1}^d X_m^{k_m} \quad (216)$$

$$n = \sum_{m=1}^d k_m \quad (217)$$

avec $n \in [0, n_{max}]$ le degré de Ψ_j .

Utiliser une méthode naïve pour générer cette matrice d'information implique le fait de devoir re-générer chaque composante $(\mathbf{x}_i)_m^{k_m}$ pour le calcul de chaque $\Psi_j(\mathbf{x}_i)$ menant ainsi à une complexité temporelle en $\mathcal{O}(d^2)$ pour l'obtention de chaque $\Psi_j(\mathbf{x}_i)$. La complexité temporelle de la génération de la matrice globale est alors en $\mathcal{O}(N \times M(n_{max}, d) \times d^2)$.

Au lieu d'utiliser cette méthode naïve, pour chaque \mathbf{x}_i , les valeurs de chaque $(\mathbf{x}_i)_m^{k_m}$ peuvent être stockées dans une matrice $D^{(i)}$ de taille $d \times n_{max}$ avec $D_{m,l}^{(i)} = (\mathbf{x}_i)_m^l$. Chaque terme $A_{i,j}$ est ainsi le produit de termes de $D^{(i)}$

$$A_{i,j} = \prod_{m=1}^d D_{m,k_m}^{(i)} \quad (218)$$

Avec cette méthode, la génération de la matrice $D^{(i)}$ a une complexité temporelle en $\mathcal{O}(d \times n_{max})$. La génération du terme $\Psi_j(\mathbf{x}_i)$ est en $\mathcal{O}(d)$. Ainsi, la complexité temporelle de la génération de la matrice globale est alors en $\mathcal{O}(N \times d \times n_{max} + N \times M(n_{max}, d) \times d) = \mathcal{O}(N \times M(n_{max}, d) \times d)$.

L'Algorithme 2 est le pseudo-code de la génération de la matrice d'information globale avec cette méthode.

Algorithm 2 Pseudo-code de la génération de la matrice d'information globale

$d \leftarrow$ nombre de paramètres d'entrée (dimension de l'espace des paramètres)
 $\mathcal{T} \leftarrow$ paramètres d'entrée correspondant aux observations d'entraînement échantillonnées à l'étape 1
 $n_{max} \leftarrow$ degré maximal souhaité pour les métamodèles locaux
 $M \leftarrow \frac{(n_{max}+d)!}{n_{max}!d!}$
Initialisation d'une matrice \mathbf{A} de taille $N \times M$ avec des zéros
for $i \in [1, N]$ **do**
 $x \leftarrow \mathcal{T}_i$
 Initialisation d'une matrice \mathbf{D} de taille $d \times n_{max}$ avec des zéros
 for $l \in [1, d]$ **do**
 for $m \in [1, n_{max}]$ **do**
 $D_{l,m} \leftarrow x_l^m$
 end for
 end for
 for $j \in [1, M]$ **do**
 $\mathbf{k} \leftarrow$ multidégré du polynôme multivarié correspondant au degré scalaire j
 $a = 1$
 for $l \in [1, d]$ **do**
 $a \leftarrow a \times D_{l,k_l}$
 end for
 $A_{i,j} \leftarrow a$
 end for
end for

3.2.5 Étape 3 - Estimation des dérivées

À l'étape 4 de l'algorithme, le regroupement hiérarchique ascendant est utilisé sur les observations représentées comme des points de \mathcal{D} . Les valeurs des dérivées partielles du QOI sur la grille non structurée des points d'entraînement doivent alors être estimées. Pour ce faire, la formule de Taylor au premier ordre est utilisée. Pour deux observations différentes \mathbf{x} et \mathbf{x}' assez proches dans l'espace \mathbb{H}_d

$$\tilde{f}(\mathbf{x}) \approx \tilde{f}(\mathbf{x}') + \frac{\partial \tilde{f}}{\partial x_1}(\mathbf{x}') \times (x_1 - x'_1) + \dots + \frac{\partial \tilde{f}}{\partial x_d}(\mathbf{x}') \times (x_d - x'_d) \quad (219)$$

Ainsi, les valeurs des dérivées partielles du QOI pour un point donné \mathbf{x}' peuvent être obtenues en utilisant la méthode OLS, cette fois-ci pour résoudre la formule de Taylor avec l'observation et ses plus proches voisins. $\{\mathbf{x}^i\}_{i=1}^{N_{neigh}}$ désignant l'ensemble des N_{neigh} plus proches voisins de \mathbf{x}' , le système suivant doit être résolu

$$\begin{pmatrix} x_1^1 - x'_1 & \cdots & x_d^1 - x'_d \\ \vdots & & \vdots \\ x_1^{N_{neigh}} - x'_1 & \cdots & x_d^{N_{neigh}} - x'_d \end{pmatrix} \begin{pmatrix} \frac{\partial \tilde{f}}{\partial x_1}(\mathbf{x}') \\ \vdots \\ \frac{\partial \tilde{f}}{\partial x_d}(\mathbf{x}') \end{pmatrix} = \begin{pmatrix} \tilde{f}(\mathbf{x}^1) - \tilde{f}(\mathbf{x}') \\ \vdots \\ \tilde{f}(\mathbf{x}^{N_{neigh}}) - \tilde{f}(\mathbf{x}') \end{pmatrix} \quad (220)$$

Il ne reste plus qu'à déterminer quelles observations sont considérées comme voisines de \mathbf{x}' . Une première idée serait de considérer comme voisines les observations correspondant

aux points les plus proches dans \mathbb{H}_d . Néanmoins, si la QOI est discontinue, le calcul des dérivées avec des observations situées dans de part et d'autre de la discontinuité serait altéré comme la QOI n'est pas dérivable dans cette zone. De ce fait, les N_{neigh} plus proches voisins sont déterminés comme étant les plus proches voisins de l'observation correspondant à \mathbf{x}' dans l'espace \mathcal{F} . Dans le calcul des dérivées partielles, il y a donc de grandes chances que seuls les points venant du même côté de la discontinuité soient considérés comme voisins. L'algorithme « Ball-tree » [48] est utilisé pour la recherche de ces voisins. Pour déterminer le nombre de voisins optimal N_{neigh} entre k_{min} et k_{max} , l'algorithme OLS est utilisé avec toutes les valeurs entre ces deux bornes, et le nombre de voisins menant à l'erreur PRESS (Predicted Residual Sum of Squares) minimale est sélectionné.

Ainsi, les observations peuvent dorénavant être représentées par des points de l'espace \mathcal{D} et le regroupement hiérarchique ascendant peut être réalisé sur ces points. L'Algorithme 3 présente le pseudo-code de l'étape 3.

Cette étape contient la construction d'un Ball-tree sur les N observations d'entraînement représentées comme des points d'un espace de dimension $d + 1$. Le coût temporel de la construction de ce Ball-tree est en $\mathcal{O}(N \log(N))$ [48]. Selon les connaissances de l'auteur, le coût de parcours de ce Ball-tree est en $\mathcal{O}(d \log(N))$, et donc le coût de parcours pour obtenir les plus proches voisins est en $\mathcal{O}(dN \log(N))$. Cette étape contient également la création d'une matrice d'information pour chaque observation d'entraînement, pour une complexité temporelle totale en $\mathcal{O}(N \times dk_{max})$, et $k_{max} - k_{min} + 1$ régressions avec des sous-matrices de taille $N_{neigh} \times d$ de cette matrice sont effectuées. Si k_{min} et k_{max} sont proportionnels à d , la complexité de ces régressions pour une observation est en $\mathcal{O}(d^4)$. Ainsi, la complexité de cette étape a plus de chance d'être dominée par les régressions et est en $\mathcal{O}(N \times d^4)$.

3.2.6 Étape 4 - Partition de l'espace des paramètres et entraînement des modèles locaux

Cette étape est l'étape principale de la méthode ME-ACD. Elle consiste à tester différentes partitions de l'espace des paramètres \mathbb{H}_d en utilisant le regroupement hiérarchique ascendant et à sélectionner celle menant aux modèles locaux les plus précis. Des partitions avec $N_E \in [|N_{E_{min}}, N_{E_{max}}|]$ éléments sont considérées, menant ainsi à $N_p = N_{E_{max}} - N_{E_{min}} + 1$ partitions testées. Dans cette section, comme plusieurs partitions sont considérées, un élément est désigné par $E_{p,e}$, où $p \in [|1, N_p|]$ est l'indice de la partition à $N_E = N_{E_{min}} + p - 1$ éléments et $e \in [|1, N_E|]$ est l'indice de l'élément dans sa partition. $E_{1,e}$ correspond alors à l'élément e de la partition à $N_{E_{min}}$ éléments et $E_{N_p,e}$ correspond à l'élément e de celle à $N_{E_{max}}$ éléments.

Dans la méthode du regroupement hiérarchique ascendant, chaque point de l'espace considéré débute dans son propre groupe et, à chaque étape, les 2 groupes les plus proches sont fusionnés, jusqu'à ce que le nombre de groupe souhaité soit atteint. La méthode Ward [71] est utilisée. Avec cette méthode, les deux groupes ayant la la somme des carrés des distances entre paires de points (entre un point du premier groupe et un point du second) sont fusionnés à chaque étape. Pour profiter de l'approche ascendante, seule

Algorithm 3 Pseudo-code de l'étape 3

$d \leftarrow$ nombre de paramètres d'entrée (dimension de l'espace des paramètres)
 $\mathcal{T} \leftarrow$ paramètres d'entrée correspondant aux observations d'entraînement échantillonnées à l'étape 1
 $\mathbf{B} \leftarrow$ valeurs de la QOI sur les observations d'entraînement obtenues à l'étape 1
Initialisation d'une matrice \mathbf{D} de taille $N \times d$ avec des zéros.
Construction du Ball-tree avec les données d'entraînement considérées comme des points de \mathcal{F} (utilisant les informations de \mathcal{T} et de \mathbf{B})
 $\mathbf{J} \leftarrow$ une matrice de taille $N \times k_{max}$ où chaque ligne contient les indices des k_{max} plus proches voisins de l'observation d'indice celui de la ligne, rangés du plus proche au plus éloigné
for $i \in [1, N]$ **do**
 $\mathbf{x} \leftarrow \mathcal{T}_i$
 $\mathbf{I} \leftarrow J_{i,:}$
 Initialisation d'une matrice \mathbf{A} de taille $k_{max} \times d$ et d'un vecteur \mathbf{F} de taille k_{max} avec des zéros
 for $m \in [1, k_{max}]$ **do**
 $F_m \leftarrow B_i - B_{I_m}$
 for $l \in [1, d]$ **do**
 $A_{m,l} \leftarrow x_l - \mathcal{T}_{I_m,l}$
 end for
 end for
 Initialisation d'une matrice \mathbf{C} de taille $d \times (k_{max} - k_{min} + 1)$ et d'un vecteur \mathbf{E} de taille $k_{max} - k_{min} + 1$ avec des zéros
 for $m \in [k_{min}, k_{max}]$ **do**
 $C_{:,m-k_{min}+1}, E_m = OLS(A_{1:m,:}, F_{1:m})$
 end for
 $k_{opt} \leftarrow \underset{m \in [k_{min}, k_{max}]}{\operatorname{argmin}} E_m$
 $D_{i,:} \leftarrow C_{:,k_{opt}}$
end for
return(\mathbf{D})

la partition à $N_{E_{min}}$ éléments est effectuée. Les partitions ayant de $N_{E_{min}} + 1$ à $N_{E_{max}}$ éléments correspondent aux dernières étapes du regroupement hiérarchique ascendant à $N_{E_{min}}$ éléments, et sont alors également accessibles. Depuis la partition à $N_E = N_{E_{max}}$ éléments, deux des éléments sont fusionnés, menant ainsi à la partition à $N_{E_{max}} - 1$ éléments et ainsi de suite jusqu'à ce que la partition à $N_E = N_{E_{min}}$ soit atteinte. De ce fait, les éléments des partitions à $N_E < N_{E_{max}}$ sont des unions des éléments de la partition à $N_{E_{max}}$ éléments. Une façon compacte de stocker cette information est d'associer un entier à chaque observation d'entraînement correspondant à l'indice e de l'élément $E_{N_p,e}$ de la partition à $N_{E_{max}}$ éléments dans lequel le point associé à l'observation se trouve. Comme les éléments $E_{p,e}$ d'une certaine partition sont des unions des éléments $E_{N_p,e}$, les éléments peuvent être stockés comme une liste d'entiers correspondant aux indices des éléments $E_{N_p,e}$ qui le composent.

L'information concernant toutes ces partitions obtenus par le regroupement hiérarchique

peut alors être stockée sous la forme :

- D'un vecteur \mathcal{L} de taille N contenant les entiers associés à chaque observation d'entraînement, entier étant égal à l'indice e de l'élément $E_{N_p,e}$ (de la partition à $N_{E_{max}}$ éléments) dans lequel elle se trouve ;
- De N_p listes d'éléments \mathcal{E}_p , une par partition testée, contenant N_E éléments $E_{p,e}$. Comme $E_{p,e}$ est l'union de certains éléments $E_{N_p,e'}$, un élément est représenté comme une liste d'entiers, entiers correspondant aux indices e' des éléments $E_{N_p,e'}$ qui le composent.

La figure 11 illustre cette idée par un diagramme ayant une structure arborescente. Dans cet exemple, $N_{E_{min}} = 2$ et $N_{E_{max}} = 4$. Les éléments sont représentés par des cercles et chaque étape du regroupement hiérarchique est une profondeur de l'arbre. Dans l'exemple, \mathcal{L} est un vecteur de taille N contenant des entiers allant de 1 à 4 et il y a 3 listes :

- \mathcal{E}_1 est une liste contenant les éléments $E_{1,e}$ stockés sous la forme de $E_{1,1} = [1]$ et $E_{1,2} = [2, 3, 4]$;
- \mathcal{E}_2 est une liste contenant les éléments $E_{2,e}$ stockés sous la forme de $E_{2,1} = [1]$, $E_{2,2} = [2, 3]$ et $E_{2,3} = [4]$;
- \mathcal{E}_3 est une liste contenant les éléments $E_{3,e}$ stockés sous la forme de $E_{3,1} = [1]$, $E_{3,2} = [2]$, $E_{3,3} = [3]$ et $E_{3,4} = [4]$

Avec cette structure de données, pour une partition donnée, les observations situées dans un élément $E_{p,e}$ sont celles dont l'entier associé dans le vecteur \mathcal{L} fait parti des entiers de la liste représentant $E_{p,e}$. Par exemple, une observation située dans $E_{1,2}$ dans l'exemple précédent a un entier associé de 2, 3 ou 4 (correspondant aux indices des éléments de la partition à $N_{E_{max}}$ éléments dont l'union est $E_{1,2}$).

Appliquer le regroupement hiérarchique sur les observations d'entraînement représentées dans différents espaces mène à des résultats différents. Effectuer un regroupement hiérarchique dans \mathbb{H}_d n'aurait pas d'intérêt. L'effectuer avec les observations considérées comme des points de \mathcal{F} partitionnerait efficacement l'espace des paramètres pour une QOI discontinue, mais pas pour le cas d'une QOI continue irrégulière. Une irrégularité d'une QOI est due à une variation rapide de ses dérivées partielles. L'idée est alors d'utiliser le regroupement pour obtenir des éléments dans lesquels la QOI a des valeurs de dérivées partielles variant peu. De ce fait, le regroupement hiérarchique est effectué sur les observations représentées comme étant des points de \mathcal{D} . Ainsi, l'espace des paramètres \mathbb{H}_d est partitionné selon les discontinuités, comme la valeur de la QOI est présente dans les coordonnées d'un point de \mathcal{D} , et selon les irrégularités, comme les valeurs des dérivées partielles sont aussi présentes dans ces coordonnées. La figure 12 représente le tracé des observations de deux QOI dépendant d'une seule variable $f^{1d-discontinuous}$ (définie par (209)) et $\tilde{f}^{1d-irregular}$ (définie par (221) ci-dessous). Les observations sont représentés dans les espaces \mathcal{D} et \mathcal{F} . $\tilde{f}^{1d-irregular}$ est une QOI continue irrégulière définie par

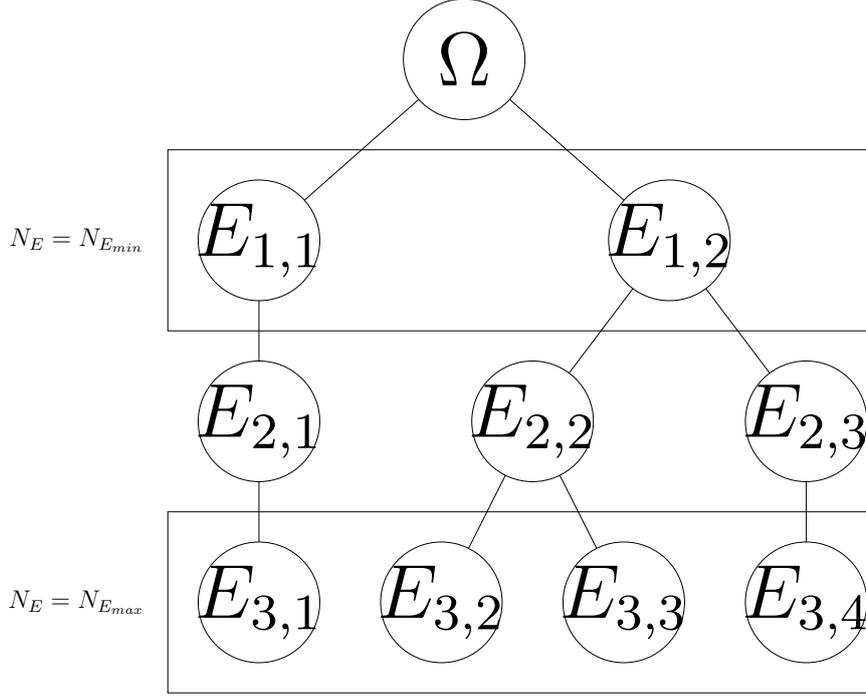


FIGURE 11 – Diagramme en structure arborescente donnant un exemple de regroupement hiérarchique ascendant. Les groupes (les éléments) sont représentés par des cercles et les étapes de $N_{E_{max}} = 4$ à $N_E = 1$, sont illustrées; $N_{E_{min}} = 2$ dans cet exemple; l’algorithme de regroupement n’effectue donc pas réellement l’étape $N_E = 1$.

$$\tilde{f}^{1d-irregular}(x) = 10 \tanh\left(10\left(x - \frac{1}{2}\right)\right) \quad (221)$$

Comme $d = 1$ pour les exemples de la figure, l’échantillonnage est une grille uniforme $1D$ avec un pas constant. La figure 12 illustre comment le regroupement est effectué dans l’espace \mathcal{D} (sous-figures (b) et (d)), et comment il permet de partitionner l’espace des paramètres $\mathbb{H}_1 = [0, 1]$ afin d’utiliser des modèles locaux de degré inférieur (sous-figures (a) et (c)) dans ces deux exemples.

Un diagramme de connectivité entre les observations d’entraînement représentés dans \mathbb{H}_d peut également être réalisé avant le regroupement (comme étape optionnelle entre l’étape 3 et l’étape 4). Avec un tel diagramme, seuls les groupes ayant au moins un couple de points connectés entre eux peuvent fusionner. Ainsi, cela réduit les chances de fusionner deux groupes étant situés loin de l’autre dans \mathbb{H}_d . Une façon de réaliser ce diagramme est de connecter chaque observations avec ses $k_{connect}$ plus proches voisins dans \mathbb{H}_d , $k_{connect}$ étant un hyperparamètre à choisir. Le choix d’utiliser ou non un tel diagramme de connectivité est modélisé par un booléen *isConnect*.

Une estimation imprécise des dérivées partielles calculées à l’étape 3 peut altérer l’étape de partition, mais si les coordonnées d’un point ne sont pas bien calculées, il a des chances d’être vu comme étant un « outlier » dans \mathcal{D} et a de grande chance d’être isolé par la partition, créant néanmoins un élément non nécessaire. Des actions peuvent être effec-

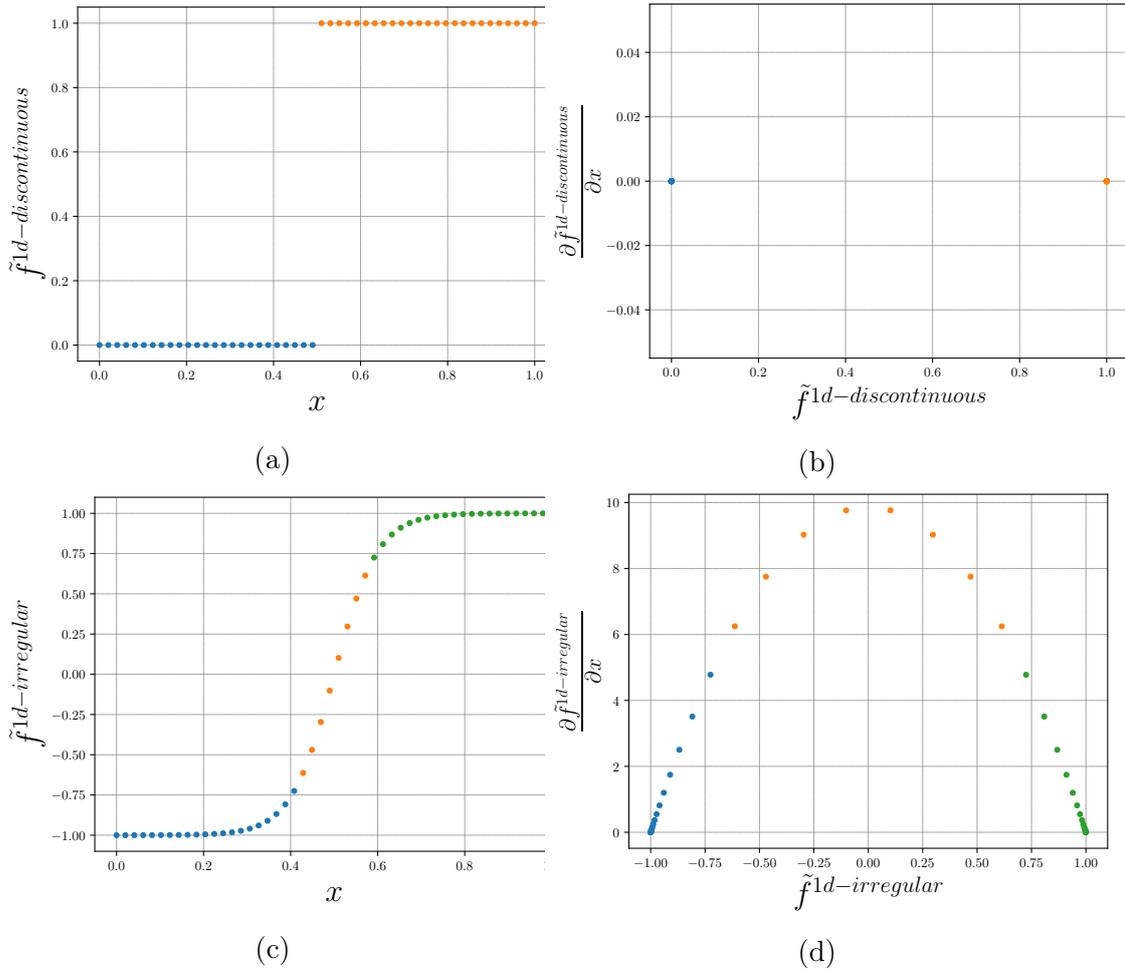


FIGURE 12 – Tracé de 50 observations pour 2 QOI $\tilde{f}^{1d-discontinuous}$ et $\tilde{f}^{1d-irregular}$ représentées par des points dans les espaces \mathcal{F} et \mathcal{D} . Les différentes couleurs des observations représentent les éléments dans lequel elles se situent. (a) Observations de $\tilde{f}^{1d-discontinuous}$ représentées dans \mathcal{F} . (b) Observations de $\tilde{f}^{1d-discontinuous}$ représentées dans \mathcal{D} . (c) Observations de $\tilde{f}^{1d-irregular}$ représentées dans \mathcal{F} . (d) Observations de $\tilde{f}^{1d-irregular}$ représentées dans \mathcal{D} .

tuées pour retirer ces observations de l'ensemble d'entraînement, ne comptant ainsi plus l'élément supplémentaire.

Pour chaque partition, des modèles locaux sont entraînés sur chacun des éléments. Une fois les N_p partitions obtenues, celle menant aux modèles locaux les plus précis est gardée. Le critère de sélection de la partition optimale est basée sur les erreurs PRESS des modèles locaux. La partition à $N_E \in [N_{E_{min}}, N_{E_{max}}]$ est considérée. Pour chacun de ses éléments $E_{p,e}$, un métamodèle local est entraîné en utilisant la méthode OLS sur les observations situées dans l'élément afin d'obtenir les coefficients du modèle par chaos polynomial (Polynomial Chaos - PC). Comme souligné plus haut à la section 2.5.3.6, le degré de troncature optimal du modèle polynomial est déterminé. Pour chaque métamodèle local, des régressions sont effectuées pour un degré de troncature entre n_{min} et n_{max} . Le degré de troncature choisi est celui minimisant l'erreur PRESS $\delta_{p,e,n}$ du métamodèle local \tilde{f}_e considéré sur les observations d'entraînement situées dans l'élément $E_{p,e}$ de la partition p

à N_E éléments et avec un degré de troncature n . En réalité, l'erreur PRESS $\delta_{p,e,n}$ utilisée est l'erreur PRESS normalisée par le maximum de la valeur absolue prise par la QOI sur ses observations d'entraînement.

$$\delta_{p,e} = \min_{n \in [n_{min}, n_{max}]} \delta_{p,e,n} \quad (222)$$

Cette erreur $\delta_{p,e}$ représente la précision du métamodèle local \tilde{f}_e entraîné sur les observations situées dans $E_{p,e}$ et ayant un degré optimal situé entre n_{min} et n_{max} . L'erreur δ_p , image de la précision de la partition p à N_E éléments, est définie comme étant la valeur maximale des erreurs $\delta_{p,e}$ de tous les métamodèles locaux entraînés sur leur élément avec un degré optimal :

$$\delta_p = \max_{e \in [1, N_E]} \delta_{p,e} \quad (223)$$

Pour obtenir la partition optimale parmi toutes celles testées, celle à $N_{E_{min}}$ éléments est tout d'abord considérée. Les métamodèles locaux de chacun des éléments de cette partition sont entraînés avec un degré optimal. D'une partition à la suivante, seul un élément est séparé en deux, redescendant la structure arborescente du regroupement hiérarchique. De ce fait, à chaque nouvelle partition testée, seuls les métamodèles locaux sur les deux nouveaux éléments doivent être entraînés, les autres ayant déjà été entraînés pour une partition précédente. Lors de ces tests de partitions, un test peut être ajouté : si l'erreur de sélection δ_p de la partition considérée est inférieure à un seuil donné δ_{min} , une partition acceptable est déjà trouvée et tester d'autres partitions n'est plus nécessaire. Ce test est désigné par la suite par arrêt anticipé de l'algorithme.

Comme il a été dit plus haut à la section 3.2.4, un des avantages majeurs d'utiliser la méthode OLS dans cette approche est que la matrice d'information globale \mathbf{A} , de taille $N \times M(n_{max}, d)$, n'est générée qu'une fois, à l'étape 2. Quand une régression pour un métamodèle local sur un élément est souhaitée, seule une sous-matrice composée des lignes correspondantes aux observations situées dans l'élément et aux colonnes correspondantes au degré de troncature (les $M(n, d)$ premières colonnes pour un degré n) est utilisée.

L'Algorithme 4 correspond au pseudo-code de l'étape 4.

La complexité temporelle de cette étape est donnée par celle du regroupement hiérarchique ascendant et par celles de toutes les régressions. La complexité temporelle de la méthode ward est en $\mathcal{O}(d \times N^3)$. Dans un cas sans diagramme de connectivité ni arrêt anticipé, $N_{E_{min}} + 2 \times (N_{E_{max}} - N_{E_{min}})$ métamodèles locaux sont entraînés. Chaque entraînement nécessite $n_{max} - n_{min} + 1$ appels à la méthode OLS pour obtenir le métamodèle de degré optimal, dominé par l'appel avec $n = n_{max}$. La complexité temporelle de toutes les régressions est alors en

$$\mathcal{O} \left(\left[N_{E_{min}} + 2 \times (N_{E_{max}} - N_{E_{min}}) \right] \times \left(\max(N, M(n_{max}, d)) \times \min(N, M(n_{max}, d))^2 \right) \right).$$

Algorithm 4 Pseudo-code de l'étape 4

$d \leftarrow$ nombre de paramètres d'entrée (dimension de l'espace des paramètres)
 $n_{min}, n_{max} \leftarrow$ degré minimum et degré maximum testés pour les métamodèles locaux
 $N_{E_{min}}, N_{E_{max}} \leftarrow$ nombre d'éléments minimum et nombre d'éléments maximum testés pour les partitions
 $N_p = N_{E_{max}} - N_{E_{min}} + 1$
 $\delta_{min} \leftarrow$ erreur de précision de partition minimale pour un arrêt anticipé potentiel
 $\mathbf{B} \leftarrow$ valeurs de la QOI sur les observations d'entraînement obtenues à l'étape 1
 $\mathbf{A} \leftarrow$ matrice d'information globale générée à l'étape 2.
 $\mathbf{D} \leftarrow$ valeurs des dérivées partielles de la QOI sur les observations d'entraînement générées à l'étape 3.
 $\mathcal{L}, \mathcal{E}_1, \dots, \mathcal{E}_{N_p} \leftarrow$ Résultats du regroupement hiérarchique ascendant (méthode ward) à $N_{E_{min}}$ groupes sur les observations d'entraînement considérées comme des points de \mathcal{D} (utilisant les informations de \mathbf{B} et \mathbf{D})
Initialisation de la variable δ_{opt} à une valeur très élevée
Initialisation de trois listes vides C , δ et $done_elements$
for $p \in [1, N_p]$ **do**
 $N_E = N_{E_{min}} + p - 1$
 for $e \in [1, N_E]$ **do**
 if la liste e de \mathcal{E}_p , $E_{p,e}$, n'est pas déjà dans $done_elements$ **then**
 $E_{p,e}$ est ajoutée dans la liste $done_elements$
 $I \leftarrow$ indices de \mathcal{L} dont les objets sont égaux aux entiers contenus dans $E_{p,e}$
 $\mathbf{b} \leftarrow$ sous-vecteur de \mathbf{B} composés des lignes d'indices dans I .
 Initialisation d'une liste vide c et initialisation d'un vecteur $\tilde{\delta}$ de taille $n_{max} - n_{min} + 1$ avec des zéros
 for $n \in [n_{min}, n_{max}]$ **do**
 $\mathbf{a} \leftarrow$ sous-matrice de \mathbf{A} composée de ses $M(n, d)$ premières colonnes et des lignes d'indices contenus dans I .
 le résultat de la méthode *OLS* avec les matrices \mathbf{a} et \mathbf{b} est ajouté à la liste c et l'erreur PRESS correspondante est stockée comme étant $\tilde{\delta}_{n-n_{min}+1}$
 end for
 $n_{opt} \leftarrow \underset{n \in [n_{min}, n_{max}]}{\operatorname{argmin}} \tilde{\delta}$
 $c_{n_{opt}-n_{min}+1}$ est ajouté à C
 $\delta_{n_{opt}-n_{min}+1}$ est ajouté à δ
 end if
 end for
 $\delta_p \leftarrow$ maximum entre les composantes de δ correspondant aux éléments de la partition p
 if $\delta_p \leq \delta_{opt}$ **then**
 $N_{E_{opt}} \leftarrow N_E$
 $\delta_{opt} = \delta_p$
 if $\delta_p \leq \delta_{min}$ **then**
 Arrêt anticipé de l'algorithme
 end if
 end if
end for
les coefficients des métamodèles locaux sont les coefficients dans C correspondant aux éléments de la partition à $N_{E_{opt}}$ éléments.

3.2.7 Étape 5 - Entraînement du classifieur

Une fois que la partition est obtenue, avec ses N_E éléments et donc ses N_E métamodèles locaux, un classifieur est nécessaire. En effet, pour les futures évaluations du modèle, la première étape est de déterminer dans quel élément E_e le point se trouve, pour savoir quel métamodèle local utiliser pour estimer la valeur de la QOI. L'utilisation d'un métamodèle correspondant au mauvais élément mène à des erreurs très élevées, les métamodèles polynomiaux pouvant être très divergents loin des points où ils ont été entraînés. Ce classifieur est l'équivalent numérique de la fonction indicatrice \mathcal{I}_{E_e} de la formule (212).

Le choix du classifieur est libre. Dans ces travaux, un perceptron multicouche (Multi layer Perceptron - MLP) est utilisé (Cf. annexe A). Le classifieur MLP est un réseau de neurone artificiel utilisé dans le but de classer des observations labellisées par un entier déterminant à quelle classe elles appartiennent. À l'étape 4, un entier correspondant à l'élément dans la partition à $N_E = N_{E_{max}}$ éléments est associé à chaque observation d'entraînement. Pour l'étape 5, les observations d'entraînement sont associées cette fois-ci à un entier égal à l'indice de l'élément dans lequel elle se trouve dans la partition à N_E éléments.

Pour toutes les applications ci-dessous, le classifieur MLP possède 2 couches cachées de 50 neurones. Pour ces couches cachées, la fonction d'activation est l'unité linéaire rectifiée (« ReLu »). Pour la couche de sortie, la fonction softmax est utilisée. En terme de fonction de coût, l'entropie croisée est minimisée. L'entraînement du classifieur est effectué par rétropropagation avec la méthode Adam [36] (avec les valeurs dégressives classiques $\beta_1 = 0.9$ et $\beta_2 = 0.999$, un mini-batch de 200 observations et $n_{epochs} = 200$ epochs sur l'ensemble des observations d'entraînements et un arrêt anticipé si les poids convergent). Plusieurs classifieurs sont entraînés avec des taux d'apprentissage initiaux différents. Celui qui maximise le nombre d'observations d'entraînement correctement labellisées, mesure de la précision du classifieur, est conservé. Dans les travaux, une méthode naïve testant les N_l taux d'apprentissage initiaux sélectionnés est adoptée. Aucun terme de régularisation n'est ajouté comme le surapprentissage n'est pas un problème dans cette approche de classification (toutes les observations d'entraînement doivent être bien classées pour éviter l'utilisation du mauvais modèle local).

D'après les connaissances de l'auteur, la complexité temporelle d'une étape d'optimisation des poids du classifieur avec propagation avant, rétropropagation et correction des biais des moments de la méthode Adam est en $\mathcal{O}\left((d \times n_{neurons} + n_{neurons}^2 + n_{neurons} \times N_E) \times N\right)$, avec $n_{neurons}$ étant le nombre de neurones dans les 2 couches cachées (50 dans les travaux). Dans ce cas, l'entraînement d'un classifieur est en $\mathcal{O}(n_{epochs} \times n_{neurons}^2 \times N_{train})$, où $n_{epochs} = 200$ dans le pire des cas. Ainsi, l'entraînement de tous les classifieurs est en $\mathcal{O}(N_l \times n_{epochs} \times n_{neurons}^2 \times N)$.

L'entrée du classifieur est un point de \mathbb{H}_d (un ensemble de valeur de paramètre d'entrée) qui est l'entrée du métamodèle par morceaux. Ce point est désigné par le point d'évaluation. La sortie du classifieur est un vecteur à N_E composantes contenant les probabilités que le point d'évaluation soit dans l'élément E_e . Ainsi, un test peut être ajouté pour vérifier si le point d'évaluation est situé dans une zone d'indécision : si la probabilité la

plus élevée est inférieure à θ , $\theta \in [0, 1]$ étant un seuil à choisir, le point d'évaluation est considéré comme situé dans une zone d'indécision. Dans ce cas, le point n'est pas évalué par le modèle et il est ignoré. θ est un paramètre de compromis. Plus θ est élevé, et plus de prudence est prise, mais plus de points d'évaluation proches de la frontière entre les éléments sont ignorés. Avec $\theta = 0$, aucun point d'évaluation n'est ignoré, mais ceux qui sont situés dans une zone d'indécision ont plus de chances d'altérer la précision du métamodèle global. En augmentant θ , un point d'évaluation a plus de chance d'être ignoré s'il affecte la précision. Après une certaine valeur, augmenter θ induit que certains points d'évaluation sont ignorés même s'ils n'affectent pas la précision. Dans [29], la méthode multi-éléments minimale présente une procédure similaire, utilisant la résolution de la stratégie d'échantillonnage pour ignorer des points situés à une distance d'une frontière inférieure à cette résolution. θ est un paramètre qui peut être affiné après l'entraînement des modèles locaux et du classifieur.

3.2.8 Résumé

Les paramètres et hyperparamètres numériques de l'algorithme sont :

- N : le nombre d'observations d'entraînement générés par la séquence de Sobol. Cela détermine la taille de l'ensemble d'entraînement.
- $N_{E_{min}}$: le nombre minimal d'éléments souhaité.
- $N_{E_{max}}$: le nombre maximal d'éléments souhaité.
- k_{min} : le nombre minimal de voisins dans \mathcal{F} utilisés dans l'estimation des dérivées partielles.
- k_{max} : le nombre maximal de voisins dans \mathcal{F} utilisés dans l'estimation des dérivées partielles.
- n_{min} : le degré minimal qu'un métamodèle PC local peut avoir.
- n_{max} : le degré maximal qu'un métamodèle PC local peut avoir.
- δ_{min} : un seuil optionnel pour effectuer un arrêt anticipé de la recherche de partition si une partition ayant une erreur inférieure à ce seuil a déjà été trouvée.
- Un ensemble de N_l taux d'apprentissage initiaux pour obtenir le classifieur optimal.
- *isConnect* : un booléen indiquant si un diagramme de connectivité est réalisé.
- $k_{connect}$ (optionnel) : nombre de voisins connectés qu'une observation a dans le diagramme de connectivité optionnel.

Pour le classifieur, les hyperparamètres sont donnés à la section 3.2.7. Le paramètre de précaution $\theta \in [0, 1]$ associé au classifieur est un paramètre de post-traitement. Une fois que le métamodèle est obtenu avec un classifieur associé, l'ajustage de θ est toujours possible sans lancer l'entraînement une nouvelle fois.

Pour le regroupement hiérarchique ascendant, la recherche des plus proches voisins et le classifieur par réseau de neurones, la bibliothèque Scikit-learn a été utilisée [49].

En terme de complexité temporelle :

- Étape 1 : $\mathcal{O}(N \times d) + N \times$ la complexité temporelle pour obtenir une valeur de f (indépendante de l'algorithme).
- Étape 2 : $\mathcal{O}(N \times M(n_{max}, d) \times d)$
- Étape 3 : $\mathcal{O}(N \times d^4)$ (si k_{min} et k_{max} sont proportionnels à d)
- Étape 4 : $\mathcal{O}\left(\left[N_{E_{min}} + 2 \times (N_{E_{max}} - N_{E_{min}})\right] \times \left(\max(N, M(n_{max}, d)) \times \min(N, M(n_{max}, d))\right)^2\right)$
- Étape 5 : $\mathcal{O}(N_l \times n_{epochs} \times n_{neurons}^2 \times N)$

L'étape dominante dépend du choix des hyperparamètres et de la dimension de la QOI étudiée.

3.3 Applications de ME-ACD

Dans cette section, les performances de ME-ACD sont étudiées sur un ensemble de QOI test. Pour évaluer ces performances, les possibles causes d'erreur entre le métamodèle par morceaux et la QOI qu'il approche sont identifiées. Ces erreurs proviennent de 3 phénomènes :

- **Sous-apprentissage (ou biais)** : les degrés (minimisant l'erreur PRESS) des métamodèles locaux sont trop faibles pour approcher la QOI sur les éléments. Augmenter N , n_{max} et/ou $N_{E_{max}}$ peuvent possiblement résoudre le problème.
- **Surapprentissage (ou variance)** : comme les partitions de l'espace des paramètres et le degré optimal des métamodèles locaux sont sélectionnés avec l'erreur PRESS, le surapprentissage n'a que très peu de chance de se produire. Dans le cas contraire, diminuer n_{min} et/ou augmenter $N_{E_{max}}$ peuvent possiblement résoudre le problème.
- **Mauvaise classification** : comme les métamodèles polynomiaux locaux divergent de façon assez prononcée loin de leurs observations d'entraînement, si un point d'évaluation est classé comme étant dans un mauvais élément, le mauvais métamodèle local est utilisé, ce qui affecte la précision. En rappel, le paramètre de précaution θ est présent pour éviter ces erreurs, ignorant le point quand la probabilité donnée par le classifieur d'être dans un élément est strictement inférieure à ce dernier. Augmenter N ou ajouter des observations d'entraînement proches des frontières des éléments et/ou modifier les paramètres du classifieur peuvent possiblement résoudre le problème. Augmenter θ rend l'algorithme plus strict et le conduit à ignorer

plus de points d'évaluation ayant plus de chances d'être responsable des erreurs de mauvaise classification.

Pour le classifieur, le taux d'apprentissage initial est celui dans $[10^{-m}, 2 \times 10^{-m}, 5 \times 10^{-m}]$, $\forall m \in [1, 7]$ menant au classifieur maximisant la précision sur l'ensemble d'entraînement lui-même (le classifieur optimal entre les $N_l = 21$ entraînés est ainsi sélectionné). Le paramètre de précaution est fixé à $\theta = 0.85$, ce qui signifie qu'un point d'évaluation n'est pas ignoré si le classifieur est sûr à plus de 85% qu'il soit situé dans un élément.

Les performances de l'algorithme sont comparées à celles d'un métamodèle PC standard de degré optimal entre n_{min-PC} et n_{max-PC} et de celles d'un algorithme de Gradient Boosted Trees (GBT) [20] (Cf. annexe A) tout deux entraînés sur le même ensemble d'entraînement. En réalité, lors des applications, le métamodèle standard PC est obtenu avec la méthode ME-ACD avec $N_{E_{min}} = N_{E_{max}} = 1$. Pour régler les 3 hyperparamètres de GBT : le taux d'apprentissage, le nombre d'arbres et la profondeur maximale des arbres ; un ensemble de validation de $N_{val} = 10^5$ observations, pour chaque ensemble d'entraînement, est échantillonné. Une recherche de type brute-force est alors réalisée pour chaque ensemble d'entraînement : chaque combinaison d'hyperparamètres est testée pour entraîner un modèle GBT et celui minimisant la MSE entre la vraie valeur de la QOI et la valeur du métamodèle GBT sur l'ensemble de validation est sélectionné. Dans les applications, les taux d'apprentissage sont situés dans $[10^{-m}, 2 \times 10^{-m}, 5 \times 10^{-m}]$, $\forall m \in [1, 7]$, le nombre d'arbres dans $[100, 200, 300, 400, 500]$ et la profondeur maximale des arbres dans $[2, 3, 4, 5]$.

Pour évaluer la précision de ces algorithmes, un ensemble de test de $N_{test} = 10^5$ observations est échantillonné aléatoirement. Cet ensemble est différent de l'ensemble de validation utilisé pour régler les hyperparamètres du modèle GBT, mais reste le même pour une étude de convergence (et donc pour plusieurs ensembles d'entraînement). À l'inverse, les ensembles de validation utilisés pour régler les hyperparamètres des modèles GBT sont différents pour chaque ensemble d'entraînement différent (pour chaque N d'une étude de convergence).

La précision de chaque métamodèle est donnée par E , la MSE entre la vraie valeur de la QOI et la valeur estimée par le métamodèle sur l'ensemble de test. Afin de mettre en valeur l'impact du paramètre de précaution, E_θ est également considérée. Il s'agit de la même erreur, mais évaluée sur les N_θ observations tests gardées par la procédure de précaution du métamodèle ME-ACD avec $\theta = 0.85$. La valeur $V_{\theta=0.85} = \frac{N_{\theta=0.85}}{N_{test}}$, fraction des observations tests gardée par la procédure de précaution, est donnée pour chaque N de l'étude de convergence.

3.3.1 Évaluation de ME-ACD sur des QOI discontinues

Dans cette section, 3 QOI discontinues prenant $d = 2$ paramètres d'entrée sont étudiées :

$$\tilde{f}_{hyperplane} = \begin{cases} 1 & \text{si } x_1 + 0.3x_2 - 0.5 > 0 \\ -1 & \text{sinon} \end{cases} \quad (224)$$

Cette fonction présente une discontinuité qui est une droite d'équation $x_1 + 0.3x_2 - 0.5 = 0$.

$$\tilde{f}^{hypersphere} = \begin{cases} 1 & \text{si } (x_1 - 0.5)^2 + (x_2 - 0.5)^2 < (0.25)^2 \\ -1 & \text{sinon} \end{cases} \quad (225)$$

Cette fonction présente une discontinuité qui est un cercle de centre $(0.5, 0.5)$ et de rayon 0.25 .

$$\tilde{f}^{multiple} = \begin{cases} 3 & \text{si } (x_1)^2 + (x_2)^2 < (0.25)^2 \\ 2 & \text{sinon si } x_1 < 0.5 - 0.3x_2 \\ 1 & \text{sinon si } (x_1 - 1)^2 + (x_2 - 1)^2 > 0.3^2 \\ 0 & \text{sinon} \end{cases} \quad (226)$$

Cette fonction présente 3 discontinuités, 2 fractions de cercles dans les coins de l'hypercube unité et la même droite que $\tilde{f}^{hyperplane}$.

Pour ces fonctions test, les éléments optimaux sont clairement définis. De ce fait, à chaque observation d'entraînement peut être associé un « vrai » entier indiquant dans quel élément il se trouve. Avec ces entiers, la précision du classifieur, désignée par F et représentant la fraction d'observations tests bien classées, peut être déterminée. Pour les futures applications, comme la procédure Multi-éléments est également utilisée pour étudier les QOI continues irrégulières avec un modèle polynomial de degré inférieur, la partition optimale n'est pas clairement définie et l'obtention de cette information n'est pas possible.

La Figure 13 montre 1000 observations d'entraînement classées par le regroupement de l'étape 4 de ME-ACD et représentées dans l'espace des paramètres pour les 3 fonctions discontinues. Cette première visualisation souligne que la méthode ME-ACD est capable de détecter des discontinuités et de partitionner l'espace des paramètres selon ces discontinuités.

	PC	ME-ACD
N	[100, 2000]	
$N_{E_{min}}$	1	1
$N_{E_{max}}$	1	10
k_{min}	non applicable	6
k_{max}	non applicable	10
n_{min}	0	0
n_{max}	20	0
δ_{min}	10^{-5}	10^{-5}
Taux d'apprentissage initiaux testés	Cf section 3.3	
$isConnect$	Faux	Faux
θ	non applicable	0.85

TABLE 6 – Hyperparamètres et paramètres sélectionnés pour ME-ACD et PC pour l'étude de convergence de $\tilde{f}^{hyperplane}$, $\tilde{f}^{hypersphere}$ et $\tilde{f}^{multiple}$.

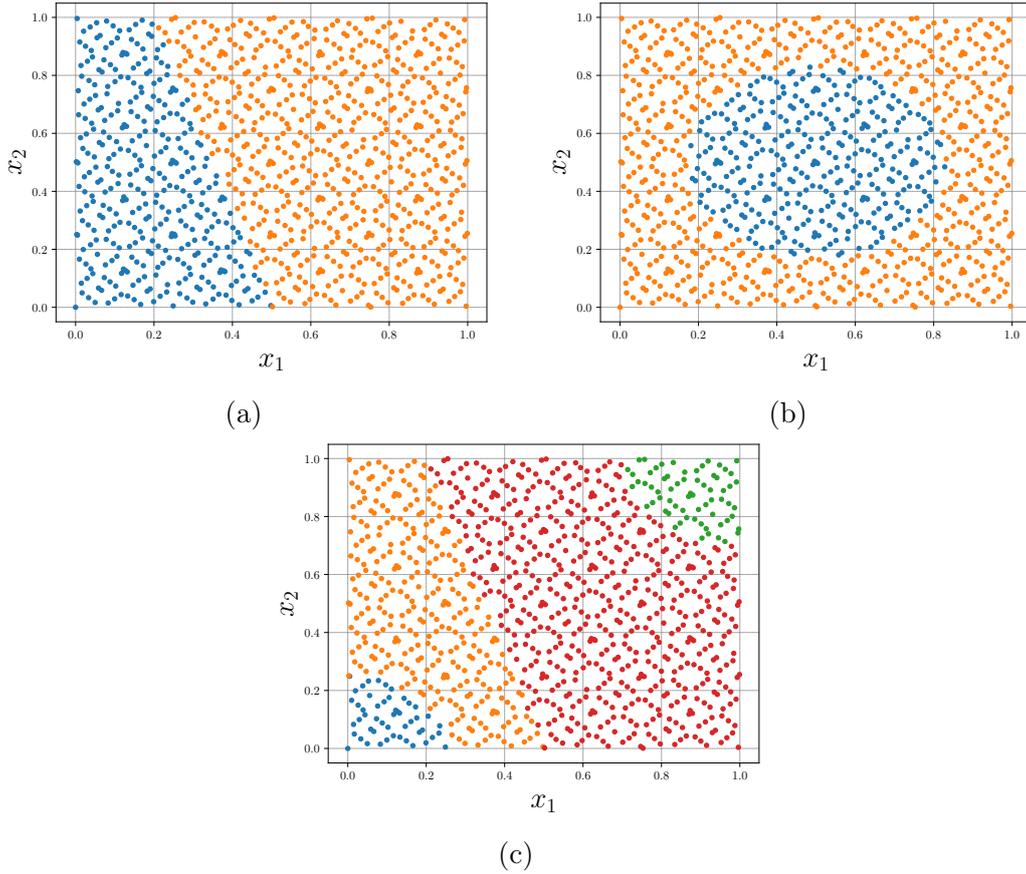


FIGURE 13 – Visualisation de 1000 observations d’entraînement classées et représentées dans l’espace des paramètres $2D \mathcal{H}_2$ partitionné pour les 3 fonctions tests. Les différentes couleurs représentent les différents éléments obtenus par le regroupement. (a) Visualisation de $\tilde{f}^{hyperplane}$. (b) Visualisation de $\tilde{f}^{hypersphere}$. (c) Visualisation de $\tilde{f}^{multiple}$.

La Figure 14 rassemble tous les diagrammes de convergence de E , de la précision F du classifieur de ME-ACD et de la fraction d’observations tests gardée par la procédure de précaution de ME-ACD sur les 3 fonctions tests sur les observations tests. La Table 6 contient les hyperparamètres et les paramètres sélectionnés pour PC et ME-ACD pour ces études de convergence. Les hyperparamètres de GBT sont les meilleurs parmi ceux testés par la recherche de type brute-force décrite à la section 3.3 pour chaque N des études de convergence.

Les diagrammes de convergence de E (Figures 14a, Figure 14c et Figure 14e) soulignent que la méthode ME-ACD est capable de mener à un modèle plus précis que les 2 autres méthodes, mais l’erreur est toujours relativement élevée. Comme les fonctions sont constantes par morceaux ($n_{min} = n_{max} = 0$), la cause de ces erreurs pour le modèle ME est la mauvaise classification des observations. Comme F est proche de 99% pour chaque N , seul 1% des observations sont responsables de toute l’erreur. Ce sont les observations situées proches des discontinuités.

La Figure 15 montre cette fois-ci un diagramme de convergence de la MSE sur les N_θ observations tests gardées par la procédure de précaution pour chaque N . Avec $\theta = 0.85$,

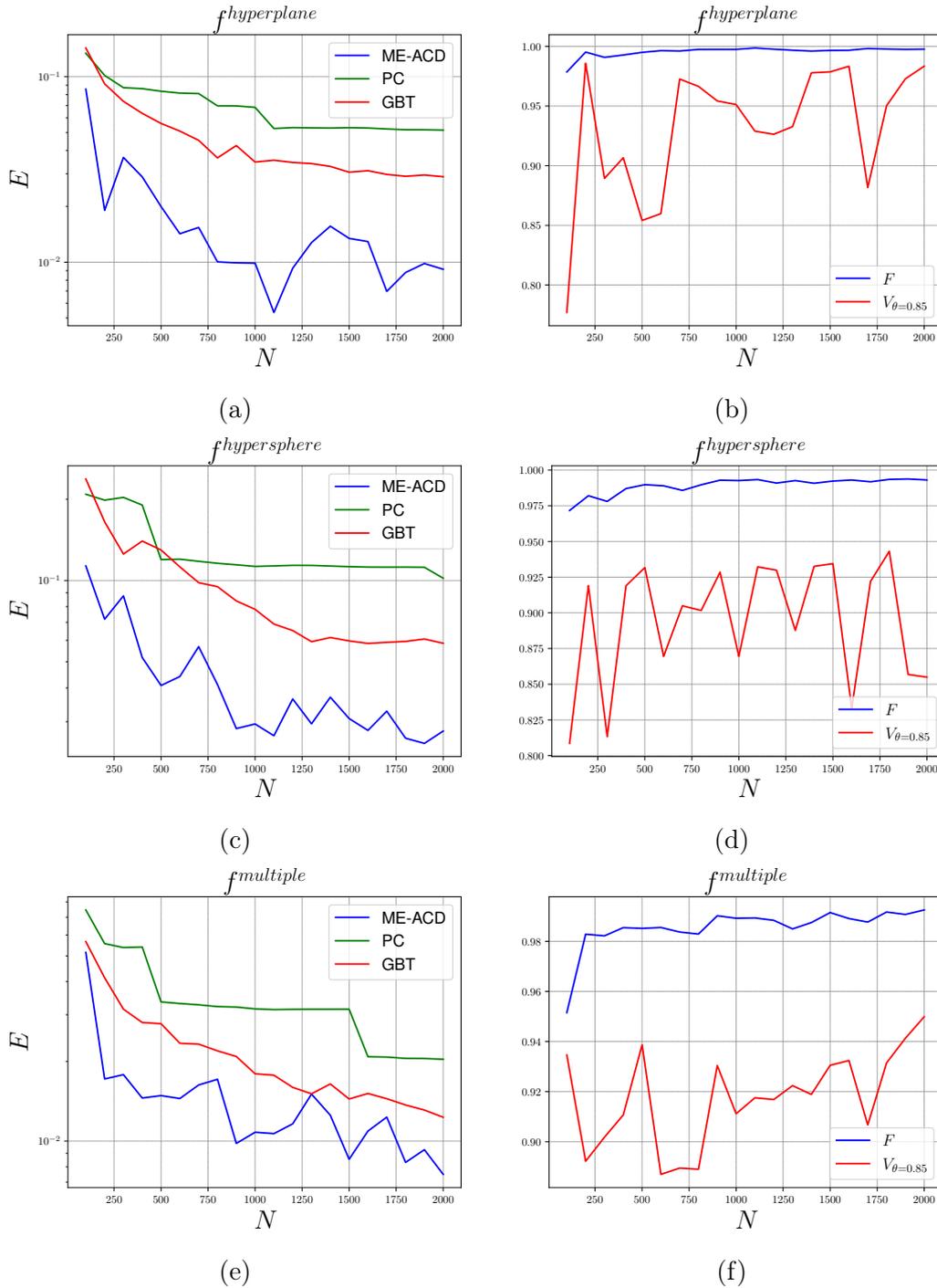


FIGURE 14 – Sur la gauche, les diagrammes de convergence de la MSE entre les valeurs des métamodèles et la valeur vraie des 3 fonctions discontinues sur l'ensemble test. (a) $\tilde{f}^{\text{hyperplane}}$. (c) $\tilde{f}^{\text{hypersphere}}$. (e) $\tilde{f}^{\text{multiple}}$. Sur la droite, diagramme de convergence de la précision du classifieur de ME-ACD et de la fraction d'observations gardées par la procédure de précaution de ME-ACD, sur l'ensemble test. (b) $\tilde{f}^{\text{hyperplane}}$. (d) $\tilde{f}^{\text{hypersphere}}$. (f) $\tilde{f}^{\text{multiple}}$.

V_{θ} est d'environ 90% pour chaque N (comme montré sur la Figure 14b, la Figure 14d et la Figure 14f). Dans la plupart des cas, toutes les observations problématiques sont ignorées par la procédure de précaution, menant à des erreurs de l'ordre de 10^{-30} , ce qui

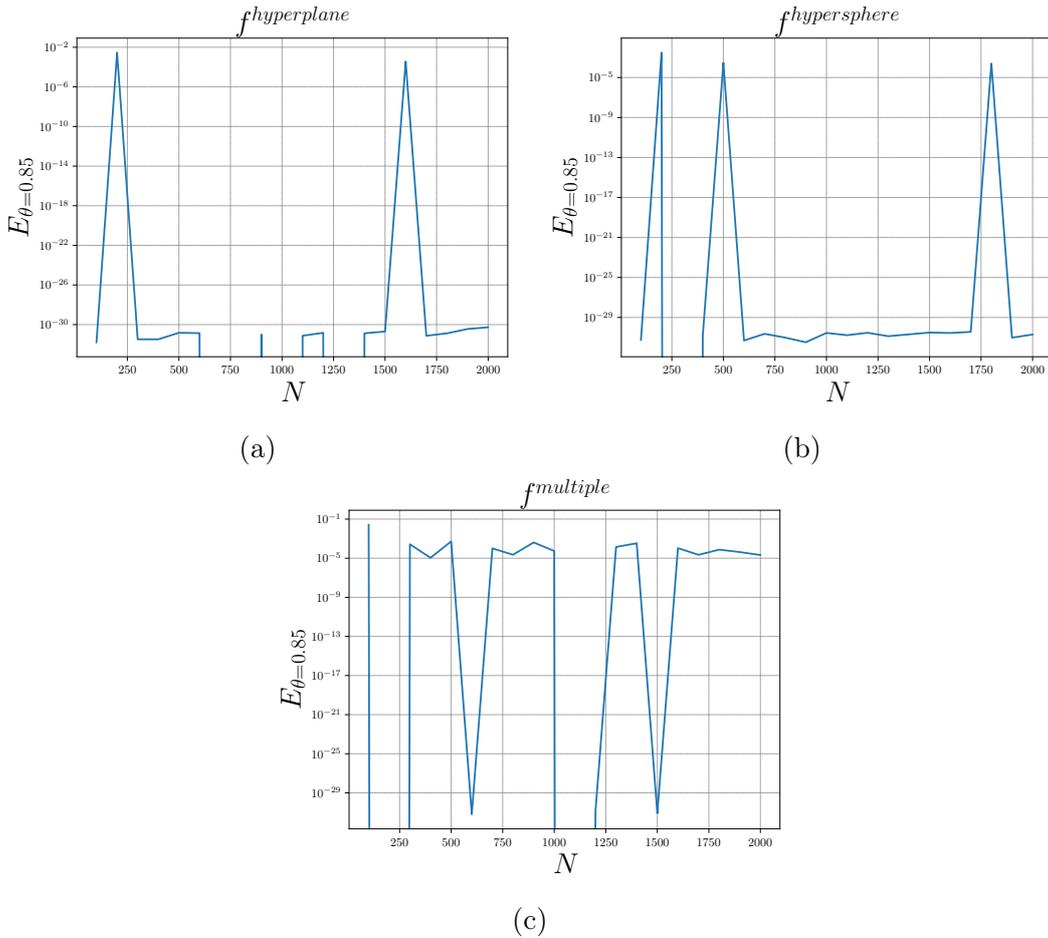


FIGURE 15 – Diagramme de convergence de la MSE entre les valeurs estimées par les métamodèles et la vraie valeur, sur les $N_{\theta=0.85}$ observations tests gardées par la procédure de précaution pour chaque N . (a) $\tilde{f}^{\text{hyperplane}}$. (b) $\tilde{f}^{\text{hypersphere}}$. (c) $\tilde{f}^{\text{multiple}}$.

démontre que toutes les observations restantes sont bien classées. Dans d'autres cas, les erreurs sont d'ordre 10^{-5} , ce qui atteste que seule une dizaine des $N_{\theta} \approx 10^4$ observations tests gardées ne sont pas bien classées. Avec $\theta = 85\%$, environ 9% des observations tests sont ignorées malgré le fait qu'elles ne mènent à aucune erreur de mauvaise classification. La procédure de précaution est alors capable de détecter les observations problématiques et des actions peuvent être menées en conséquence pour éviter des erreurs de mauvaise classification.

Des résultats encore plus précis pour le classifieur (au niveau de F et de V_{θ}) auraient pu être obtenus avec un autre échantillonnage. Ici, un échantillonnage de Sobol est utilisé pour modéliser un échantillonnage de Monte-Carlo plus homogène, menant au même échantillonnage indépendamment des potentielles discontinuités. Dans [26], un échantillonnage adaptatif est développé. Cet échantillonnage est capable de générer plus d'observations proches des discontinuités. Si le choix de l'échantillonnage est possible, plus d'observations proches des discontinuités devraient mener à une meilleure précision du classifieur et à une procédure de précaution plus efficace.

La Figure 16 montre le degré optimal des métamodèles PC pour les études de convergence des 3 fonctions tests. Le nombre de termes du métamodèle polynomial PC à 2 variables correspondant est indiqué entre parenthèses. Avec un nombre croissant d'observations d'entraînement, PC est capable de sélectionner un degré optimal plus élevé sans souffrir de surapprentissage. Comparés à ces métamodèles PC, les métamodèles polynomiaux locaux de ME-ACD n'ont qu'un terme chacun et le métamodèle global ME est plus précis. L'économie en nombre de termes dans les séries polynomiales pourrait être encore plus marquée à plus haute dimension.

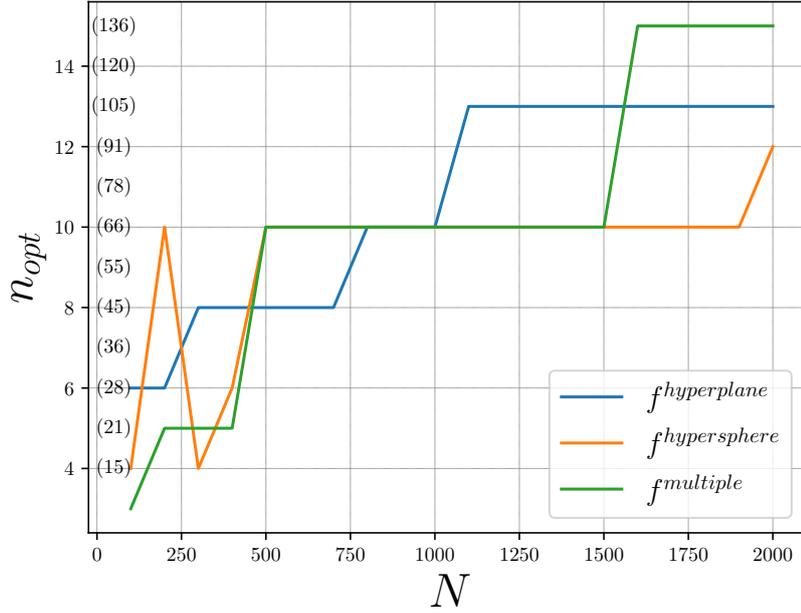


FIGURE 16 – Évolution du degré optimal des métamodèles PC pour les études des 3 fonctions discontinues $\tilde{f}_{hyperplane}$, $\tilde{f}_{hypersphere}$ and $\tilde{f}_{multiple}$. Le nombre de termes du métamodèle $M(n, d = 2)$ est également indiqué entre parenthèses sur l'axe des ordonnées.

3.3.2 Évaluation de ME-ACD sur une QOI présentant une discontinuité partielle

Dans cette sous-section, la QOI suivante, dépendant de $d = 2$ paramètres d'entrée, est étudiée

$$\tilde{f}^{partial} = \begin{cases} 0.5x_2 - 0.25 & \text{si } x_1 \geq 0.5 \\ -0.5x_2 + 0.25 & \text{sinon} \end{cases} \quad (227)$$

Cette fonction définie par morceaux présente une discontinuité traversant « partiellement » l'espace des paramètres. La fonction est linéaire de part et d'autre de l'hyperplan $x_1 = 0.5$. Néanmoins, quand x_2 devient proche de 0.5, l'écart entre les deux surfaces est tellement petit que numériquement, cette région est régulière (même si la discontinuité est toujours présente mathématiquement). D'après les connaissances de l'auteur, aucune des méthodes ME précédentes ne permet de partitionner l'espace des paramètres en présence

d'une discontinuité partielle. D'après [26], c'est une problématique pour les méthodes ME modernes.

	PC	ME-ACD
N	[100, 2000]	
$N_{E_{min}}$	1	1
$N_{E_{max}}$	1	10
k_{min}	non applicable	6
k_{max}	non applicable	10
n_{min}	0	0
n_{max}	20	10
δ_{min}	10^{-5}	10^{-5}
Taux d'application initiaux testés	Cf section 3.3	
$isConnect$	non applicable	Faux
θ	non applicable	0.85

TABLE 7 – Hyperparamètres et paramètres sélectionnés pour PC et ME-ACD pour l'étude de convergence de $\tilde{f}^{partial}$.

La Figure 17, contenant les résultats obtenus avec les hyperparamètres et les paramètres rassemblés dans la Table 7, montre des résultats similaires avec les études de convergence des fonctions discontinues. ME-ACD est alors capable de gérer les discontinuités ne passant que partiellement à travers l'espace des paramètres de la même manière que les discontinuités traversant intégralement l'espace. À l'inverse de l'annihilation polynomiale utilisée dans [5, 29], le regroupement hiérarchique des observations dans \mathcal{D} n'est pas affecté par un écart plus petit aux frontières de deux éléments.

3.3.3 Évaluation de ME-ACD sur une QOI modélisant une bifurcation

Un des avantages majeurs de ME-ACD est qu'il est capable d'obtenir un métamodèle précis dans le cas de fonctions continues régulières, de fonctions continues irrégulières et de fonctions discontinues. L'adaptation de ME-ACD pour obtenir des modèles de solutions d'équations différentielles ordinaires (Ordinary Differential Equations - ODE) est présentée à la section 3.5. Cette adaptation semble alors convenable pour approcher les solutions d'ODE lors de l'apparition d'une bifurcation (menant alors à des irrégularités et des discontinuités par rapport aux paramètres d'entrée), comme le système dynamique du vol présenté au chapitre 1.

La QOI étudiée est la suivante

$$\tilde{f}^{bif-t}(x_1, x_2) = (1 + x_2) \tanh\left(\frac{1}{2}(3x_1 - 1)(1 + x_2)t\right) \quad (228)$$

Cette QOI est la solution à l'instant t de l'ODE :

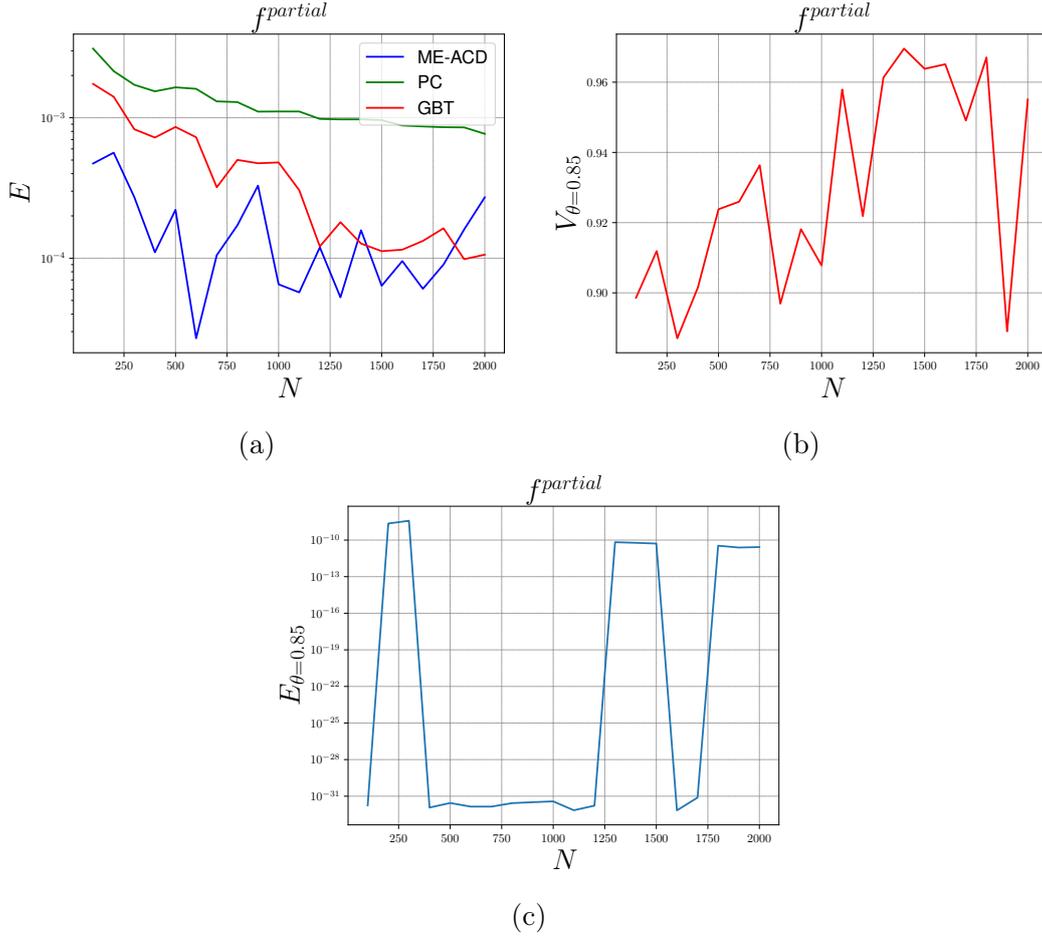


FIGURE 17 – (a) Diagramme de convergence de la MSE pour $\tilde{f}^{partial}$ sur son ensemble test. (b) Diagramme de convergence de la précision du classifieur de ME-ACD et de la fraction d’observations gardée par la procédure de précaution de ME-ACD pour $\tilde{f}^{partial}$ sur son ensemble test. (c) Diagramme de convergence de la MSE pour $\tilde{f}^{partial}$ sur les $N_{\theta=0.85}$ observations tests gardées par la procédure de précaution de ME-ACD.

$$\dot{y} = -p_1(y^2 - p_2^2) \quad (229)$$

avec $y(t = 0) = 0$, p_1 suivant une loi de distribution uniforme sur $[-\frac{1}{2}, 1]$ et p_2 suivant une loi de distribution uniforme sur $[1, 2]$. La transformation bijective pour obtenir le \mathbf{x} correspondant apparait dans la forme de la fonction \tilde{f}^{bif-t} . Cette ODE modélise l’apparition d’une bifurcation. Les deux équilibres de ce système (en ignorant le cas ponctuel $p_1 = 0$) sont $-p_2$, stable si $p_1 < 0$, et p_2 , stable si $p_1 > 0$.

Aux premiers instants, la solution est régulière en fonction de \mathbf{x} . Quand t augmente, la solution est de plus en plus irrégulière par rapport à \mathbf{x} . Asymptotiquement ($t \rightarrow \infty$), la solution présente une discontinuité, qui est une droite d’équation $p_1 = 0$ ($x_1 = \frac{1}{3}$), symbole d’une bifurcation par rapport à p_1 (ou x_1).

Les 3 instants qui sont considérés sont $t = 1$ (régulier), $t = 10$ (irrégulier) et $t = 1000$ (discontinu). La Figure 18 montre les variations de la fonction avec x_1 pour $x_2 = 1$ (la

valeur de x_2 menant au plus grand écart).

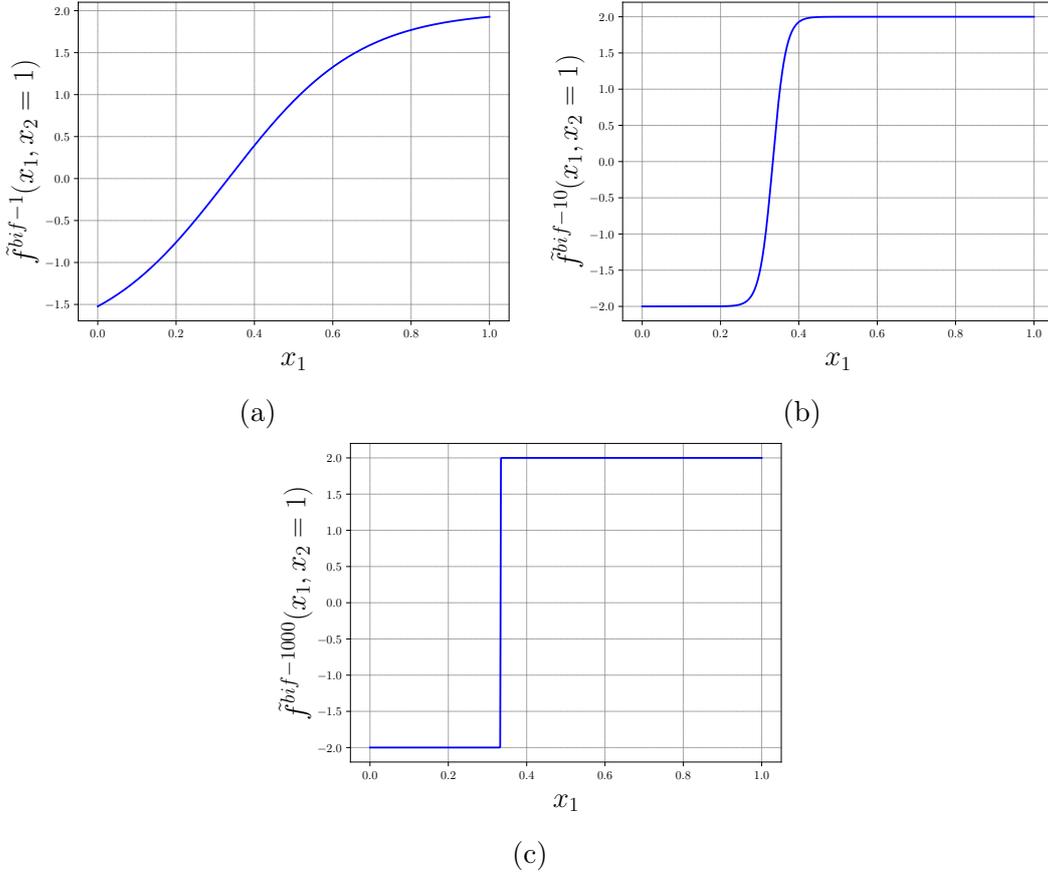


FIGURE 18 – Graphique de $\tilde{f}^{bif-t}(x_1, x_2 = 1)$ en fonction de x_1 à 3 instants : $t = 1$ (régulier), $t = 10$ (irrégulier) et $t = 1000$ (discontinu).

De façon similaire aux études précédentes, la Figure 19 montre les résultats des études de convergence de la MSE sur l’intégralité de l’ensemble test ainsi l’évolution de la fraction des observations tests gardées par la procédure de précaution de ME-ACD. La Figure 20 montre la convergence de la MSE $E_{\theta=0.85}$ sur les observations gardées. La Table 8 rassemble les hyperparamètres et les paramètres utilisés pour ces études.

À $t = 1$, l’arrêt anticipé de ME-ACD induit que la partition sélectionnée soit à 1 élément. Le modèle est ainsi strictement équivalent à PC, mais avec un degré maximal n_{max} moins important. PC montre donc naturellement de meilleures performances que ME-ACD à cet instant. Avec un n_{max} similaire, les deux méthodes auraient donné des résultats identiques. Comme à cet instant, $N_{E_{opt}} = 1$ pour tous les N , 100% des observations tests sont gardées par la procédure de précaution.

À $t = 10$, ME-ACD partitionne l’espace des paramètres en 3 éléments ou plus pour utiliser des métamodèles locaux de degré inférieur par rapport à PC. ME-ACD est capable de donner une approximation précise d’une QOI continue irrégulière en utilisant des métamodèles locaux de faible degré. Comme montré à la Figure 20b, la procédure de précaution permet d’avoir un modèle encore plus précis, mais comme la QOI n’est pas

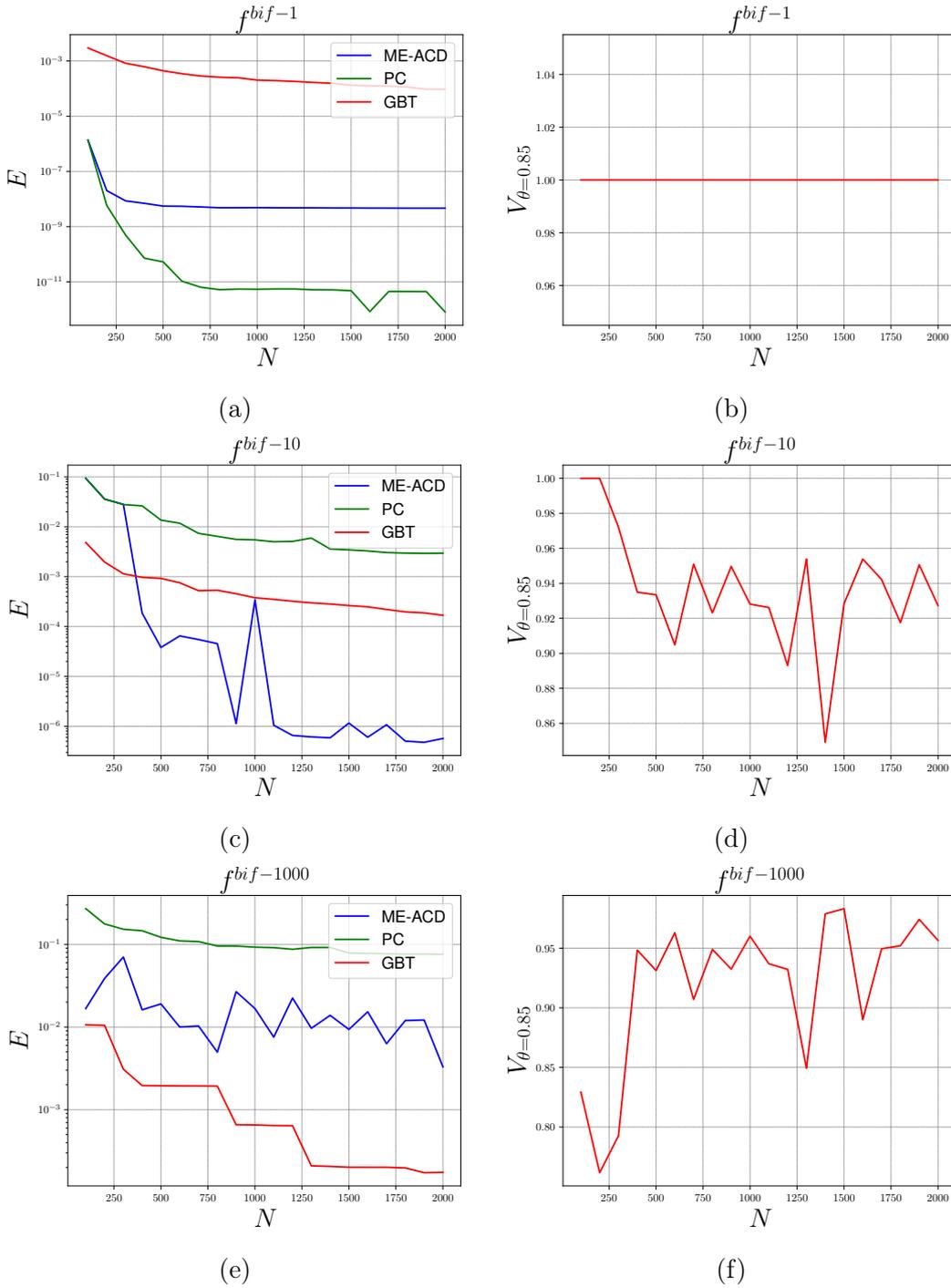


FIGURE 19 – Sur la gauche, les diagrammes de convergence de la MSE entre les valeurs des métamodèles et la valeur vraie de la fonction considérée aux 3 instants sur l’ensemble test. (a) \tilde{f}^{bif-1} . (c) \tilde{f}^{bif-10} . (e) $\tilde{f}^{bif-1000}$. Sur la droite, diagramme de convergence de la précision du classifieur de ME-ACD et de la fraction d’observations gardées par la procédure de précaution de ME-ACD, sur l’ensemble test. (b) \tilde{f}^{bif-1} . (d) \tilde{f}^{bif-10} . (f) $\tilde{f}^{bif-1000}$.

discontinue, les erreurs de mauvaise classification ont moins d’impact sur la précision.

À $t = 1000$, ME-ACD se comporte de manière quasi-identique qu’avec la QOI $\tilde{f}^{hyperplane}$.

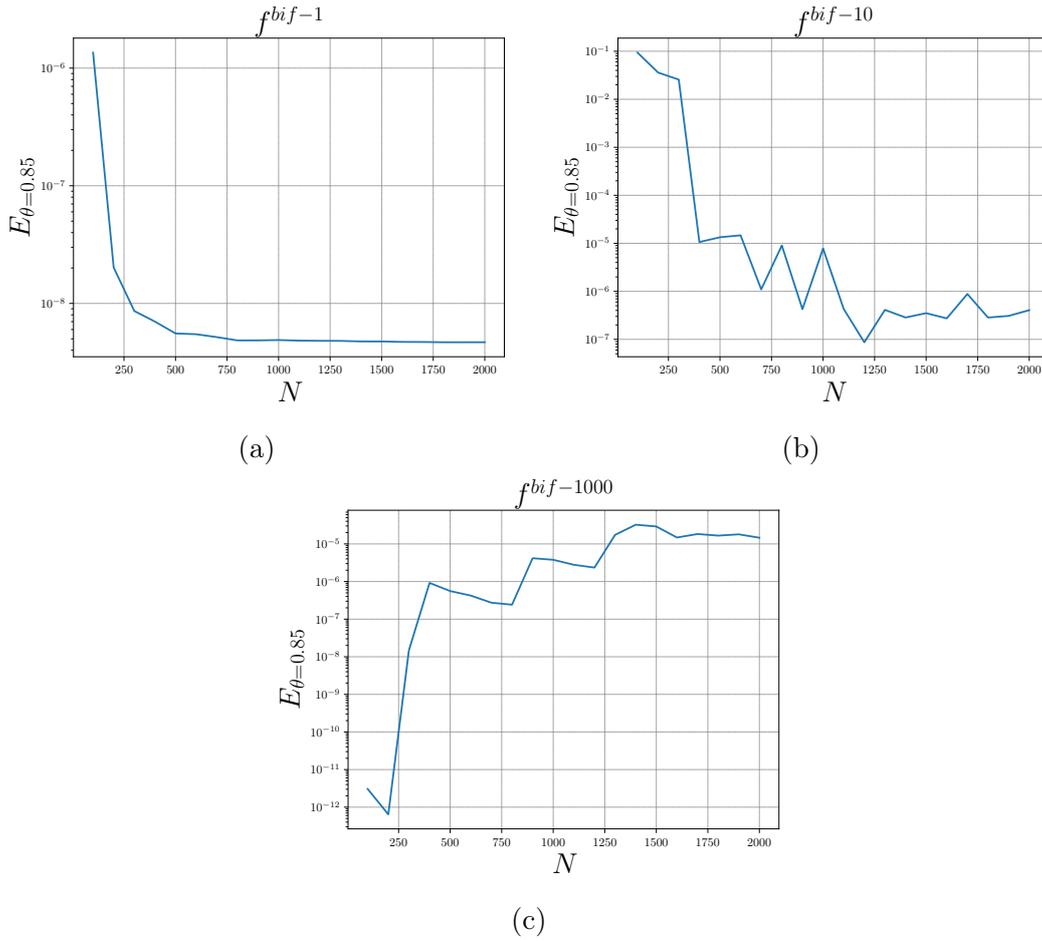


FIGURE 20 – Diagramme de convergence de la MSE entre les valeurs des méta-modèles et la vraie valeur, sur les $N_{\theta=0.85}$ observations tests gardées par la procédure de précaution pour chaque N . (a) \tilde{f}^{bif-1} . (b) \tilde{f}^{bif-10} . (c) $\tilde{f}^{bif-1000}$.

	PC	ME-ACD
N	[100, 2000]	
$N_{E_{min}}$	1	1
$N_{E_{max}}$	1	10
k_{min}	non applicable	6
k_{max}	non applicable	10
n_{min}	0	0
n_{max}	20	10
δ_{min}	10^{-5}	10^{-5}
taux d'apprentissage initiaux testés	Cf section 3.3	
$isConnect$	Faux	Vrai
$k_{connect}$	non applicable	10
θ	non applicable	0.85

TABLE 8 – Hyperparamètres et paramètres sélectionnés pour PC et ME-ACD pour les étude de convergence de \tilde{f}^{bif-1} , \tilde{f}^{bif-10} et $\tilde{f}^{bif-1000}$.

Seules plusieurs observations d'entraînement proches de la quasi-discontinuité ont une valeur du QOI différente de $-p_2$ ou p_2 . Ces observations d'entraînement altèrent la pré-

cision des métamodèles locaux, comme elles ne sont pas assez pour former à elles-seules un troisième élément. La procédure de précaution peut ignorer les observations menant à une erreur de mauvaise classification, mais n’ignore pas ces observations problématiques. Cela explique pourquoi l’erreur décroît très lentement, comme l’échantillonnage de Sobol n’ajoute des observations d’entraînement que très rarement dans la zone problématique qui est très fine. Avec un faible N , aucune observation n’est située dans cette zone, et un modèle à 2 éléments avec 2 hyperplans est obtenu par ME-ACD, ce qui explique que la précision soit meilleure avec un faible N . Générer plus d’observations proches de la quasi-discontinuité avec un échantillonnage adaptatif pourrait résoudre le problème, menant à assez d’observations pour faire un troisième élément. Une autre solution serait de retirer les quelques points problématiques de l’ensemble d’entraînement en les considérant comme des « outliers ».

ME-ACD est alors capable de donner un modèle précis dans les cas d’une QOI régulière, d’une QOI continue mais irrégulière et d’une QOI discontinue, ce qui en fait un métamodèle capable d’approcher une grande variété de QOI et un bon candidat pour approcher les solutions d’une ODE avec de potentielles bifurcations.

3.3.4 Évaluation de ME-ACD sur la fonction de Genz en dimension 5

Dans cette section, une QOI de la famille des fonctions de Genz [21], dépendant de 5 paramètres d’entrée, est étudiée :

$$\tilde{f}^{genz}(\mathbf{x}) = \begin{cases} 0 & \text{si l'un des } x_i > 0.5 \\ \exp(\sum_{i=1}^5 x_i) & \text{sinon} \end{cases} \quad (230)$$

	PC	ME-ACD
N	50000	
$N_{E_{min}}$	1	1
$N_{E_{max}}$	1	5
k_{min}	non applicable	15
k_{max}	non applicable	25
n_{min}	0	0
n_{max}	10	10
δ_{min}	10^{-5}	10^{-5}
Taux d’apprentissage initiaux testés	Cf section 3.3	
$isConnect$	non applicable	Vrai
$k_{connect}$	non applicable	30
θ	non applicable	0.85

TABLE 9 – Hyperparamètres et paramètres sélectionnés pour PC et ME-ACD pour l’étude de convergence de \tilde{f}^{genz}

La Table 9 rassemble les hyperparamètres et les paramètres pour l’étude de convergence, et la Table 10 montre quant à elle les performances des 3 méthodes avec $N = 50000$

	PC	GBT	ME-ACD
E	0.15415751	0.01737417	0.02062906
$E_{\theta=0.85}$	non applicable		0.00182077
$V_{\theta=0.85}$	non applicable		99.651%

TABLE 10 – Performances des 3 méthodes de la QOI $5d f^{genz}$ entraînée sur $N = 50000$ observations. Les performances sont évaluées sur un ensemble d’entraînement de $N_{test} = 10^5$ observations tests.

observations d’entraînement. Encore une fois, ME-ACD est capable de donner un modèle relativement précis et est capable de réduire les erreurs de mauvaise classification avec la procédure de précaution. Pour cette application, ME-ACD et PC utilisent la même matrice d’information globale ayant $M(n_{max} = 10, d = 5) = 3003$ colonnes, mais ME-ACD mène à une bien meilleure précision avec la procédure ME. Le degré optimal de PC est égal à 9.

3.4 Conclusion

Dans ce chapitre, une nouvelle méthode, ME-ACD, a été développée. Cette méthode est capable de partitionner efficacement l’espace des paramètres pour obtenir un méta-modèle PC par morceaux dans le cas d’une QOI discontinue, tout comme les méthodes reposant l’inférence bayésienne [57] ou sur la détection par annihilation polynomiale [29]. ME-ACD est aussi capable de partitionner l’espace en présence d’une discontinuité ne parcourant pas l’intégralité de l’espace des paramètres d’un côté de l’hypercube à l’autre, ce qui n’était pas possible avec les méthodes précédentes, d’après les connaissances de l’auteur. De plus, la méthode permet d’étudier une QOI continue mais irrégulière, qui nécessite un grand nombre de termes dans un modèle PC standard. Le modèle ME est composé de métamodèles PC locaux de degré inférieur, prenant ainsi la même logique que la méthode ME-PCM [19] mais avec des partitions suivant les irrégularités. La méthode a été testée avec succès sur des QOI discontinues et des QOI continues ayant des pentes raides. D’après ces premiers résultats, ME-ACD semble générale et est prometteuse pour approcher de nombreuses QOI.

ME-ACD a été développée dans le contexte de l’UQ, mais, comme les autres méthodes non-intrusives, elle peut être utilisée dans un autre contexte comme une nouvelle méthode de régression. ME-ACD donne des modèles précis dans le cas de QOI discontinues et de QOI continues irrégulières dépendant d’un nombre modéré de paramètres d’entrée. ME-ACD peut mener à une interprétation physique, partitionnant l’espace des paramètres en éléments pouvant être interprétés. Avec l’utilisation d’un réseau de neurones comme classifieur, le métamodèle est capable de déterminer si un point d’évaluation se trouve dans une zone d’indécision entre les éléments, et est alors capable de faire preuve de précaution avec ces points, en les ignorant à la place de donner une valeur risquant d’être imprécise du fait des erreurs de mauvaise classification.

ME-ACD est en développement et présente des limites qui seront le sujet de futures re-

cherches. L’approche par regroupement hiérarchique est sensible à l’échantillonnage et la partition peut ainsi manquer de robustesse dans certains cas. Si le choix de l’échantillonnage n’est pas possible (à cause de contraintes expérimentales par exemple), la partition obtenue peut ne pas être convenable avec l’échantillonnage donné (des observations étant trop loin les unes des autres pour donner une estimation correcte des gradients par exemple). De plus, pour le moment, l’utilisation de la méthode est limitée à un nombre modéré de paramètres d’entrée ($d \leq 9$). Pour une dimension plus élevée, la notion de plus proches voisins utilisée notamment dans le cas de l’estimation des dérivées perd son sens, comme souligné dans [6]. Comme un des objectifs de la méthode est de contrer la malédiction de la dimension en utilisant des métamodèles locaux de plus faibles degré, utiliser cette méthode en plus haute dimension pourrait s’avérer utile. Des travaux visant à adapter cette méthode pour des dimensions supérieures seront alors menés.

3.5 Extension à une QOI dépendant du temps

Cet algorithme peut être étendu pour approcher une QOI étant la solution d’une ODE, et dépendant donc du temps. Cette section est une version traduite et légèrement modifiée d’un extrait de [65]. Comme précisé à la section 2.5.3.3, les coefficients des métamodèles locaux sont également des fonctions du temps.

$$\hat{f}(t, \mathbf{x}) = \sum_{e=1}^{N_E} \hat{f}_e(t, \mathbf{x}) \mathcal{I}_{E_e}(\mathbf{x}) = \sum_{e=1}^{N_E} \left(\sum_{j=1}^{M_e} c_{ej}(t) \Psi_j^{(d)}(\mathbf{x}) \right) \mathcal{I}_{E_e}(\mathbf{x}) \quad (231)$$

Un ensemble fini de N_t instants $\{t_1, \dots, t_n\}$ (indexés par ordre croissant) est considéré. En prenant avantage des potentielles similarités de la QOI d’un instant au suivant, pour chaque instant t_τ , avec $\tau \geq 2$, la partition obtenue à l’instant précédent $t_{\tau-1}$ est testée en premier lieu et les métamodèles locaux correspondant aux éléments de cette partition sont entraînés avec les informations de la QOI à l’instant τ . L’erreur de partition à l’instant τ , δ_τ est calculée (par le biais de (223)). Cette erreur est alors une métrique de la précision du métamodèle global à l’instant t_τ avec la partition correspondant à l’instant précédent. Si $\delta_\tau \leq \delta_{max}$, δ_{max} étant un seuil à sélectionner, la partition et les métamodèles locaux sont conservés pour cet instant. Dans ce cas, la partition est la même pour $t_{\tau-1}$ et t_τ , mais les coefficients des métamodèles locaux diffèrent potentiellement. Pour le premier instant $\tau = 1$ ou si $\delta_\tau > \delta_{max}$ pour $\tau > 1$, une nouvelle partition convenable est obtenue à l’aide de l’algorithme ME-ACD. Les métamodèles locaux sont alors entraînés avec les valeurs de la QOI à l’instant t_τ . Un nouveau classifieur est alors entraîné pour cette nouvelle partition. L’algorithme ME-ACD temporel mène alors à $N_p \leq N_t$ partitions avec leur classifieur correspondant. Une liste \mathbb{T} de taille N_t relie les instants avec leur partition. Pour chaque instant, les coefficients des métamodèles locaux sont disponibles, pour chaque élément de la partition correspondante à cet instant.

Pour un système dynamique, composé de plusieurs ODE couplées, cet algorithme peut être répété pour chaque variable d’état du système, chaque variable étant traitée de façon indépendante. Comme ces variables ont différentes échelles, une erreur de partition normalisée δ_τ peut être utilisée pour être comparée aux paramètres choisis δ_{min} et δ_{max} ,

ces paramètres étant choisis pour toutes les variables. Dans la présente étude, δ_τ est normalisée par le maximum de la valeur absolue que peut prendre la variable d'état à l'instant t_τ dans tous les scénari correspondant aux observations d'entraînement.

Chapitre 4

Estimation des indices de Sobol à l'aide d'un métamodèle multi-éléments

Ce chapitre a pour but de présenter les développements d'une méthodologie pour estimer les indices de Sobol à l'aide d'un métamodèle multi-éléments (Multi-Element - ME). Il s'agit d'une version traduite et modifiée de l'article [67], en cours d'écriture, qui introduit cette méthodologie.

4.1 Nouvelle approche concernant les indices de Sobol

Pour obtenir la formule analytique entre les coefficients d'un métamodèle par chaos polynomial (Polynomial Chaos - PC) et les indices de Sobol, ce qui est réalisé de façon classique dans la littérature [64, 15] est l'identification des termes de la décomposition ANOVA (Cf. section 2.5.3.7). Les variances partielles sont ainsi déterminées à l'aide des propriétés d'orthogonalité pour obtenir la formule (207). Dans le cas ME, l'identification des termes n'est pas aussi directe. Une autre approche est alors considérée, nécessitant cette fois-ci une forme analytique des termes d'espérance conditionnelle $\mu_{\mathbf{i}}^{(j)}$ définis par (156).

4.1.1 Forme explicite des termes ANOVA

Théorème 1. *Dans le cas d'une quantité d'intérêt (Quantity Of Interest - QOI) de forme quelconque, $\forall j \in \Omega_d$, $\forall \mathbf{i} \in \mathcal{I}_j^{(d)}$*

$$\tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{l=0}^j \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} (-1)^{j-l} \mu_{\mathbf{u}}^{(l)}(\mathbf{x}_{\mathbf{u}}) \quad (232)$$

Démonstration. $\forall d \in \mathbb{N}^*$, la proposition P(j) (la formule (232)) est prouvée par récurrence $\forall j \in \Omega_d$.

Initialisation $P(0)$ est directement vérifiée. Pour $\mathbf{i} \in \mathcal{I}_{j=0}^{(d)}$, $\mathbf{i} = \mathbf{0}$ et

$$\tilde{f}_0^{(0)} = E(\tilde{f}) = (-1)^{0-0} \mu_0^{(0)}$$

Hérédité Soit $j \in \llbracket 0, d-1 \rrbracket$, $\forall l \in \llbracket 0, j \rrbracket$, $P(l)$ est supposée vérifiée. $\mathbf{i} \in \mathcal{I}_{j+1}^{(d)}$ est dorénavant considéré. D'après (155)

$$\tilde{f}_{\mathbf{i}}^{(j+1)} = \mu_{\mathbf{i}}^{(j+1)} - \sum_{l=0}^j \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}^{(l)} \quad (233)$$

D'après $P(0)$, $P(1)$, et ainsi de suite jusqu'à $P(j)$, chaque $\tilde{f}_{\mathbf{u}}^{(l)}$ ($l \in \llbracket 0, j \rrbracket$) de la somme est lui-même une somme de termes $(-1)^{l-m} \mu_{\mathbf{v}}^{(m)}$, $m \in \llbracket 0, l \rrbracket$, $\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{u}\}$.

Ainsi, pour un $l \in \llbracket 0, j \rrbracket$ donné, la somme des $\tilde{f}_{\mathbf{u}}^{(l)}$ sur tous les $\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}$ est une somme de somme de termes $(-1)^{l-m} \mu_{\mathbf{v}}^{(m)}$. Par énumération de ces termes, regroupés par valeur de m :

- $m = 0$: chaque $\tilde{f}_{\mathbf{u}}^{(l)}$ de la somme apporte un terme $(-1)^l \mu_0^{(0)}$. Ainsi, on énumère $\binom{j+1}{l}$ termes $(-1)^l \mu_0^{(0)}$ dans la somme des $\tilde{f}_{\mathbf{u}}^{(l)}$.
- $m = 1$: pour un $\mathbf{v} \in \mathcal{I}_{m=1}^{(j)}\{\mathbf{i}\}$, ($\mathbf{v} = v$ est un scalaire) donné, tous les $\tilde{f}_{\mathbf{u}}^{(l)}$ de la somme n'apporte pas de terme $(-1)^{l-1} \mu_{\mathbf{v}}^{(1)}$. Les $\tilde{f}_{\mathbf{u}}^{(l)}$ correspondant aux interactions d'ordre l ne contenant pas v parmi leurs composantes n'apportent pas de termes $(-1)^{l-1} \mu_{\mathbf{v}}^{(1)}$. On dénombre $\binom{j+1-1}{l}$ $\tilde{f}_{\mathbf{u}}^{(l)}$ n'apportant pas ce terme. D'après la relation de Pascal qui stipule que $\binom{j+1}{l} = \binom{j+1-1}{l} + \binom{j+1-1}{l-1}$, $\binom{j+1-1}{l-1}$ $\tilde{f}_{\mathbf{u}}^{(l)}$ apportent un terme $(-1)^{l-1} \mu_{\mathbf{v}}^{(1)}$.
- $m = 2$: pour un couple $\mathbf{v} = (v_1, v_2) \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}$ donné, $\binom{j-2}{l-2}$ termes apportent $(-1)^{l-2} \mu_{(\mathbf{v}_1, \mathbf{v}_2)}^{(2)}$. En effet, d'après le point précédent, $\binom{j+1-1}{l-1}$ termes de la somme dépendent du paramètre indexé par v_1 . Parmi ces termes, $\binom{j+1-2}{l-1}$ ne dépend pas du paramètre d'entrée indexé par v_2 . D'après la relation de Pascal, $\binom{j+1-2}{l-2}$ termes $(-1)^{l-2} \mu_{\mathbf{v}}^{(2)}$ sont énumérés dans la somme.
- $m \in \llbracket 3, l \rrbracket$: en réitérant la même logique, il y a $\binom{j+1-m}{l-m}$ termes $(-1)^{l-m} \mu_{\mathbf{v}}^{(m)}$ dans la somme.

Ainsi, pour un $l \in \llbracket 0, j \rrbracket$ donné

$$\sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}^{(l)} = \sum_{m=0}^l \binom{j+1-m}{l-m} \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} (-1)^{l-m} \mu_{\mathbf{v}}^{(m)} \quad (234)$$

ce qui implique

$$\sum_{l=0}^j \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}^{(l)} = \sum_{l=0}^j \sum_{m=0}^l (-1)^{l-m} \binom{j+1-m}{l-m} \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \mu_{\mathbf{v}}^{(m)} \quad (235)$$

En réordonnant les termes de la somme

$$\sum_{l=0}^j \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}^{(l)} = \sum_{m=0}^j \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \mu_{\mathbf{v}}^{(m)} \sum_{l=m}^j (-1)^{l-m} \binom{j+1-m}{l-m}$$

Or, $\forall m \in [0, j]$

$$\sum_{l=m}^j (-1)^{l-m} \binom{j+1-m}{l-m} = \sum_{a=0}^{j+1-m} (-1)^a \binom{j+1-m}{a} - (-1)^{j+1-m} = (1-1)^{j+1-m} - (-1)^{j+1-m} = (-1)^{j-m} \quad (236)$$

Donc

$$\sum_{l=0}^j \sum_{\mathbf{u} \in \mathcal{I}_l^{(j)}\{\mathbf{i}\}} \tilde{f}_{\mathbf{u}}^{(l)} = \sum_{m=0}^j \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} (-1)^{j-m} \mu_{\mathbf{v}}^{(m)}$$

et, d'après (155)

$$\tilde{f}_{\mathbf{i}}^{(j+1)} = \mu_{\mathbf{i}}^{(j+1)} + \sum_{m=0}^j \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} (-1)^{j+1-m} \mu_{\mathbf{v}}^{(m)} = \sum_{m=0}^{j+1} \sum_{\mathbf{v} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} (-1)^{j+1-m} \mu_{\mathbf{v}}^{(m)}$$

$P(j+1)$ est vérifié.

Conclusion La formule (232) est prouvée $\forall j \in \Omega_d$. □

Exemple 16. Pour $d = 5$, $j = 3$, $\mathbf{i} = (1, 3, 5)$,

$$\begin{aligned} f_{(1,3,5)}^{(3)} &= \mu_{(1,3,5)}^{(3)} - f_{(1,3)}^{(2)} - f_{(1,5)}^{(2)} - f_{(3,5)}^{(2)} - f_1^{(1)} - f_3^{(1)} - f_5^{(1)} - f_0^{(0)} \\ &= \mu_{(1,3,5)}^{(3)} - \mu_{(1,3)}^{(2)} - \mu_{(1,5)}^{(2)} - \mu_{(3,5)}^{(2)} + \mu_1^{(1)} + \mu_3^{(1)} + \mu_5^{(1)} - \mu_0^{(0)} \end{aligned} \quad (237)$$

4.1.2 Matrice Ξ

Dans cette section, une matrice Ξ de taille $2^d \times 2^d$ est définie. Il est ensuite montré que les variances partielles $D_{\mathbf{i}}^{(j)}$ peuvent être retrouvées à l'aide d'une combinaison linéaire des termes de cette matrice.

Définition 8. $\forall a_1 \in \Omega_d$, $\forall a_2 \in \Omega_d$, $\forall \mathbf{u} \in \mathcal{I}_{a_1}^{(d)}$, $\forall \mathbf{v} \in \mathcal{I}_{a_2}^{(d)}$

$$\xi_{\mathbf{u}, \mathbf{v}} = \int_{\mathcal{H}_d} \mu_{\mathbf{u}}^{(a_1)}(\mathbf{x}_{\mathbf{u}}) \mu_{\mathbf{v}}^{(a_2)}(\mathbf{x}_{\mathbf{v}}) d\mathbf{x} \quad (238)$$

Proposition 1.

$$D_{\mathbf{i}}^{(j)} = \sum_{m_1=0}^j \sum_{m_2=0}^j \sum_{\mathbf{u} \in \mathcal{I}_{m_1}^{(j)}\{\mathbf{i}\}} \sum_{\mathbf{v} \in \mathcal{I}_{m_2}^{(j)}\{\mathbf{i}\}} (-1)^{-m_1-m_2} \xi_{\mathbf{u},\mathbf{v}} \quad (239)$$

Démonstration. D'après (232)

$$\left(\tilde{f}_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) \right)^2 = \sum_{m_1=0}^j \sum_{m_2=0}^j \sum_{\mathbf{u} \in \mathcal{I}_{m_1}^{(j)}\{\mathbf{i}\}} \sum_{\mathbf{v} \in \mathcal{I}_{m_2}^{(j)}\{\mathbf{i}\}} (-1)^{-m_1-m_2} \mu_{\mathbf{u}}^{(m_1)}(\mathbf{x}_{\mathbf{u}}) \mu_{\mathbf{v}}^{(m_2)}(\mathbf{x}_{\mathbf{v}}) \quad (240)$$

(239) est alors obtenue en injectant (240) dans (159). \square

Trouver les termes $\xi_{\mathbf{u},\mathbf{v}}$ permet alors de trouver les variances partielles. Les termes $\xi_{\mathbf{u},\mathbf{v}}$ peuvent être ordonnés dans une matrice Ξ de taille $2^d \times 2^d$ en appliquant la transformation bijective suggérée à la section 2.3.1 pour chacun des deux multi-indices. Les indices des termes de la matrice sont gardés sous forme multidimensionnelle dans les démonstrations analytiques.

Proposition 2. *L'espérance de la QOI vaut, dans cette nouvelle approche*

$$E = \mu_{\mathbf{0}}^{(0)} = \sqrt{\xi_{\mathbf{0},\mathbf{0}}} \quad (241)$$

Démonstration. La relation est directement obtenue à partir de la définition du terme $\xi_{\mathbf{0},\mathbf{0}}$. \square

Proposition 3. *Soit $\mathbf{z} \in \mathcal{I}_d^{(d)}$, l'indice dénommant l'interaction à l'ordre d entre tous les paramètres d'entrée. La variance vaut, dans cette nouvelle approche*

$$D = \xi_{\mathbf{z},\mathbf{z}} - \xi_{\mathbf{0},\mathbf{0}} \quad (242)$$

Démonstration. Soit $\mathbf{z} \in \mathcal{I}_d^{(d)}$

$$\mu_{\mathbf{z}}^{(d)}(\mathbf{x}) = f(\mathbf{x}) \quad (243)$$

il en résulte alors

$$\xi_{\mathbf{z},\mathbf{z}} = \int_{\mathcal{H}_d} \left(f(\mathbf{x}) \right)^2 d\mathbf{x} \quad (244)$$

(242) est alors trouvée en injectant (244) dans (157). \square

Une fois que la matrice Ξ est construite, les variances partielles et la variance de la QOI sont accessibles. Ainsi, tous les indices de Sobol peuvent être déterminés à l'aide des composantes de cette matrice.

La matrice Ξ présente des propriétés intéressantes, qui peuvent être utilisées dans le cas d'une QOI générale. Pour les cas particuliers de métamodèles PC ou ME, les propriétés de la section 4.1.3, qui sont plus strictes, sont utilisées.

Propriété 4. Ξ est symétrique, d'après la commutativité de la multiplication

$$\xi_{\mathbf{u},\mathbf{v}} = \xi_{\mathbf{v},\mathbf{u}} \quad (245)$$

Propriété 5. Si $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v}) = \emptyset$

$$\xi_{\mathbf{u},\mathbf{v}} = \int_{\mathcal{H}_d} \mu_{\mathbf{u}}^{(m_1)}(\mathbf{x}_{\mathbf{u}}) \mu_{\mathbf{v}}^{(m_2)}(\mathbf{x}_{\mathbf{v}}) d\mathbf{x} = \int_{\mathcal{H}_{m_1}} \mu_{\mathbf{u}}^{(m_1)}(\mathbf{x}_{\mathbf{u}}) d\mathbf{x}_{\mathbf{u}} \times \int_{\mathcal{H}_{m_2}} \mu_{\mathbf{v}}^{(m_2)}(\mathbf{x}_{\mathbf{v}}) d\mathbf{x}_{\mathbf{v}} = \left(f_0^{(0)}\right)^2 \quad (246)$$

Cette formule est valable en particulier quand $\mathbf{u} \in \mathcal{I}_0^{(d)}$ ou $\mathbf{v} \in \mathcal{I}_0^{(d)}$.

4.1.3 Vecteur $\tilde{\Xi}$

Dans les cas particuliers d'un métamodèle PC ou ME, la matrice Ξ présente une propriété supplémentaire

Propriété 6. $\forall d \in \mathbb{N}^*, \forall j \in [1, d], \forall m_1, m_2, m_3, m_4 \in [1, j], \forall \mathbf{u} \in \mathcal{I}_{m_1}^{(d)}, \forall \mathbf{v} \in \mathcal{I}_{m_2}^{(d)}, \mathbf{w} \in \mathcal{I}_{m_3}^{(d)}, \forall \mathbf{y} \in \mathcal{I}_{m_4}^{(d)}$

$$\xi_{\mathbf{u},\mathbf{v}} = \xi_{\mathbf{w},\mathbf{y}} \text{ si } \mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v}) = \mathcal{S}(\mathbf{w}) \cap \mathcal{S}(\mathbf{y}) \quad (247)$$

Le fait que la matrice Ξ possède cette propriété dans les cas PC et ME est démontré aux sections 4.1.4 et 4.1.5.3. Grâce à cette propriété, de nombreuses composantes de la matrice Ξ sont égales et la matrice peut être réduite à un vecteur $\tilde{\Xi}$ avec

$$\tilde{\xi}_{\mathbf{u}} = \xi_{\mathbf{v},\mathbf{w}} \text{ avec } \mathcal{S}(\mathbf{u}) = \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w}) \quad (248)$$

L'espérance devient alors

$$E = \sqrt{\tilde{\xi}_{\mathbf{0}}} \quad (249)$$

et la variance devient

$$D = \tilde{\xi}_{\mathbf{z}} - \tilde{\xi}_{\mathbf{0}} \quad (250)$$

Théorème 2. En ce qui concerne les variances partielles, la formule (239) devient

$$D_{\mathbf{i}}^{(j)} = \sum_{m=0}^j \sum_{\mathbf{u} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} (-1)^{j-m} \tilde{\xi}_{\mathbf{u}} \quad (251)$$

Démonstration. En premier lieu, les intersections $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})$ apparaissant dans (239) sont énumérées. Soit $m \in \Omega_d$, une interaction $\mathbf{w} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}$ est considérée. Pour un couple donné (m_1, m_2) , le nombre de couples $\mathbf{u} \in \mathcal{I}_{m_1}^{(j)}\{\mathbf{i}\}$ et $\mathbf{v} \in \mathcal{I}_{m_2}^{(j)}\{\mathbf{i}\}$ satisfaisant $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v}) = \mathcal{S}(\mathbf{w})$ sont énumérés.

Tout d'abord, tous les \mathbf{u} avec $\mathcal{S}(\mathbf{w}) \subset \mathcal{S}(\mathbf{u})$ sont énumérés. Comme tous les éléments de $\mathcal{S}(\mathbf{w})$ doivent être dans $\mathcal{S}(\mathbf{u})$, les \mathbf{u} possibles sont ceux possédant les composantes de \mathbf{w} et dont les autres $m_1 - m$ composantes sont les possibles combinaisons sans remise entre les $j - m$ entiers de Ω_d n'étant pas dans $\mathcal{S}(\mathbf{w})$. Ainsi, $\binom{j-m}{m_1-m}$ possibles \mathbf{u} sont énumérés.

Ensuite, tous les \mathbf{v} avec $\mathcal{S}(\mathbf{w}) \subset \mathcal{S}(\mathbf{v})$ et $(\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{w})) \not\subset \mathcal{S}(\mathbf{v})$ sont énumérés (afin que $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v}) = \mathcal{S}(\mathbf{w})$ et que l'intersection ne contienne pas d'entier supplémentaire n'étant pas dans $\mathcal{S}(\mathbf{w})$). Pour respecter ces conditions, il faut que $m_2 \geq m$ pour que $\mathcal{S}(\mathbf{v})$ contienne tous les entiers de $\mathcal{S}(\mathbf{w})$ et $m_2 \leq j - m_1 + m$ sinon $\mathcal{S}(\mathbf{v})$ contiendrait au moins un élément de $\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{w})$. Les possibles \mathbf{v} sont alors les vecteurs possédant les m entiers de $\mathcal{S}(\mathbf{w})$ parmi ses composantes et dont les autres composantes sont les combinaisons sans remise entre les $j - m_1$ entiers de Ω_d qui ne sont ni dans $\mathcal{S}(\mathbf{u})$ ni dans $\mathcal{S}(\mathbf{w})$. Ainsi, $\binom{j-m}{m_1-m}$ possibles \mathbf{v} respectent les propriétés souhaitées.

En conclusion, avec m_1 et m_2 fixés, $\binom{j-m}{m_1-m} \times \binom{j-m_1}{m_2-m}$ couples (\mathbf{u}, \mathbf{v}) avec $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v}) = \mathcal{S}(\mathbf{w})$ sont énumérés.

D'après (248) et l'énumération réalisée, (239) peut se réécrire

$$D_{\mathbf{i}}^{(j)} = \sum_{m=0}^j \sum_{\mathbf{w} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \left[\sum_{m_1=m}^j \sum_{m_2=m}^{j-m_1+m} \binom{j-m}{m_1-m} \binom{j-m_1}{m_2-m} (-1)^{m_1+m_2} \right] \tilde{\xi}_{\mathbf{w}} \quad (252)$$

Or

$$\sum_{m_1=m}^j \sum_{m_2=m}^{j-m_1+m} \binom{j-m}{m_1-m} \binom{j-m_1}{m_2-m} (-1)^{m_1+m_2} = \sum_{a_1=0}^{j-m} \binom{j-m}{a_1} (-1)^{a_1} \sum_{a_2=0}^{j-a_1-m} \binom{j-a_1-m}{a_2} (-1)^{a_2} \quad (253)$$

De plus

$$\sum_{a_2=0}^{j-a_1-m} \binom{j-a_1-m}{a_2} (-1)^{a_2} = (1-1)^{j-a_1-m} = 0^{j-a_1-m} \quad (254)$$

Donc

$$\sum_{a_1=0}^{j-m} \binom{j-m}{a_1} (-1)^{a_1} \sum_{a_2=0}^{j-a_1-m} \binom{j-a_1-m}{a_2} (-1)^{a_2} = \sum_{a_1=0}^{j-m} \binom{j-m}{a_1} (-1)^{a_1} 0^{j-m-a_1} = (-1)^{j-m} \quad (255)$$

En conclusion, (251) est trouvée en injectant (255) dans (252). \square

Concernant les quantités permettant de calculer les indices de Sobol totaux,

$$D_i^{(T)} = \tilde{\xi}_{\mathbf{z}} - \tilde{\xi}_{\sim i} \quad (256)$$

Démonstration. Il est montré dans [55] que $D_i^{(T)} = D - \left(\int_{\mathcal{H}_{d-j}} \left(\mu_{\sim i}^{(d-j)} \right)^2 d\mathbf{x}_{\sim i} - (f_0)^2 \right)$

De ce fait, le résultat est direct. \square

4.1.4 Estimation des indices de Sobol avec un métamodèle PC avec cette nouvelle approche

Dans cette section, une formule analytique entre les coefficients du chaos d'un métamodèle PC et les composantes du vecteur $\tilde{\Xi}$ est démontrée. Pour cela, une forme analytique des termes $\mu_{\mathbf{i}}^{(j)}$ est explicitée, ce qui mène à l'expression des termes de la matrice Ξ . Il est ensuite montré que, pour un métamodèle PC, la matrice Ξ respecte (247) et les termes du vecteur $\tilde{\Xi}$ correspondant sont ainsi explicités en fonction des coefficients du chaos.

La première étape est alors de trouver une expression de $\mu_{\mathbf{i}}^{(j)}$ en fonction des coefficients du chaos.

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \int_{\mathcal{H}_{d-j}} \Phi_{\mathbf{k}}^{\mathcal{H}_d}(\mathbf{x}_{\mathbf{i}}, \mathbf{x}_{\sim i}) d\mathbf{x}_{\sim i} \quad (257)$$

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \prod_{m \in \overline{\mathcal{S}(\mathbf{i})}} \int_0^1 \phi_{k_m}^{0,1}(x_m) dx_m \prod_{m \in \mathcal{S}(\mathbf{i})} \phi_{k_m}^{0,1}(x_m) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \Delta_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \prod_{m \in \mathcal{S}(\mathbf{i})} \phi_{k_m}^{0,1}(x_m) \quad (258)$$

En ce qui concerne les termes de la matrice Ξ

$$\mu_{\mathbf{v}}^{(m_1)}(\mathbf{x}_{\mathbf{v}}) \mu_{\mathbf{w}}^{(m_2)}(\mathbf{x}_{\mathbf{w}}) = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \sum_{\mathbf{l} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \alpha_{\mathbf{l}} \Delta_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v})}} \Delta_{\mathbf{l}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{w})}} \prod_{q \in \mathcal{S}(\mathbf{v})} \phi_{k_q}^{0,1}(x_q) \prod_{r \in \mathcal{S}(\mathbf{w})} \phi_{l_r}^{0,1}(x_r) \quad (259)$$

Or

$$\prod_{q \in \mathcal{S}(\mathbf{v})} \phi_{k_q}^{0,1}(x_q) \prod_{r \in \mathcal{S}(\mathbf{w})} \phi_{l_r}^{0,1}(x_r) = \prod_{q \in \mathcal{S}(\mathbf{v}) \setminus \mathcal{S}(\mathbf{w})} \phi_{k_q}^{0,1}(x_q) \prod_{r \in \mathcal{S}(\mathbf{w}) \setminus \mathcal{S}(\mathbf{v})} \phi_{l_r}^{0,1}(x_r) \prod_{t \in \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \phi_{k_t}^{0,1}(x_t) \phi_{l_t}^{0,1}(x_t) \quad (260)$$

Donc, par orthogonalité

$$\xi_{\mathbf{v}, \mathbf{w}} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \sum_{\mathbf{l} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \alpha_{\mathbf{l}} \Delta_{\mathbf{k}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v})}} \Delta_{\mathbf{l}, \mathbf{0}}^{\overline{\mathcal{S}(\mathbf{w})}} \Delta_{\mathbf{k}, \mathbf{0}}^{\mathcal{S}(\mathbf{v}) \setminus \mathcal{S}(\mathbf{w})} \Delta_{\mathbf{l}, \mathbf{0}}^{\mathcal{S}(\mathbf{w}) \setminus \mathcal{S}(\mathbf{v})} \Delta_{\mathbf{k}, \mathbf{l}}^{\mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \quad (261)$$

D'après (147) et d'après le lemme 1

$$\xi_{\mathbf{v},\mathbf{w}} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} \sum_{\mathbf{l} \in \mathcal{K}_n^{(d)}} \alpha_{\mathbf{k}} \alpha_{\mathbf{l}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})}} \Delta_{\mathbf{l},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})}} \Delta_{\mathbf{k},\mathbf{l}}^{\mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \quad (262)$$

soit

$$\xi_{\mathbf{v},\mathbf{w}} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} (\alpha_{\mathbf{k}})^2 \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})}} \quad (263)$$

D'après (263), chaque terme $\xi_{\mathbf{u},\mathbf{v}}$ avec une intersection $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})$ similaire sont égaux. Ainsi, la proposition (247) est respectée et les composantes du vecteur $\tilde{\Xi}$ correspondant sont

$$\tilde{\xi}_{\mathbf{u}} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} (\alpha_{\mathbf{k}})^2 \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \quad (264)$$

La variance peut alors être estimée par (250), les variances partielles par (251) et les quantités $D_i^{(T)}$ par (256) en utilisant les composantes du vecteur $\tilde{\Xi}$.

Il peut alors être montré que cette nouvelle approche mène au même résultat que le raisonnement développé à la section 2.5.3.7. En injectant (264) dans (251)

$$D_{\mathbf{i}}^{(j)} = \sum_{\mathbf{k} \in \mathcal{K}_n^{(d)}} (\alpha_{\mathbf{k}})^2 \sum_{m=0}^j (-1)^{j-m} \sum_{\mathbf{u} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \quad (265)$$

Néanmoins

$$\tilde{\Delta}_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} = \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \times (1 - \delta_{i_1}) \times \dots \times (1 - \delta_{i_j}) = \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \times (1 - \delta_{i_1} - \delta_{i_2} + \delta_{i_1} \delta_{i_2}) \times \dots \times (1 - \delta_{i_j}) \quad (266)$$

et ainsi de suite jusqu'à obtenir

$$\tilde{\Delta}_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} = \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \sum_{m=0}^j (-1)^m \sum_{\mathbf{u} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \Delta_{\mathbf{k},\mathbf{0}}^{\mathcal{S}(\mathbf{u})} \quad (267)$$

Comme toutes les interactions d'ordre inférieur impliquant les composantes de \mathbf{i} sont présentes dans la somme, chaque sous-ensemble $\mathcal{S}(\mathbf{u})$ peut être vu comme un sous ensemble $\mathcal{S}(\mathbf{i}) \setminus \mathcal{S}(\mathbf{v})$, avec $\mathbf{v} \in \mathcal{I}_{j-m}^{(j)}\{\mathbf{i}\}$. La double somme peut donc être réordonnée

$$\tilde{\Delta}_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} = \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \sum_{m=0}^j (-1)^{j-m} \sum_{\mathbf{u} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \Delta_{\mathbf{k},\mathbf{0}}^{\mathcal{S}(\mathbf{i}) \setminus \mathcal{S}(\mathbf{u})} = \sum_{m=0}^j (-1)^{j-m} \sum_{\mathbf{u} \in \mathcal{I}_m^{(j)}\{\mathbf{i}\}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \quad (268)$$

et (207) est retrouvée en injectant (268) dans (265). La nouvelle approche est donc équivalente à l'approche classique reliant les coefficients du chaos aux variances partielles. Il peut néanmoins être conjecturé que, pour un métamodèle PC classique, utiliser cette approche de façon numérique mène à des résultats moins précis car l'approche est moins directe. Ce sont les termes $\tilde{\xi}_{\mathbf{u}}$ qui sont exprimés en fonction des coefficients du chaos et non pas les variances partielles directement. Les erreurs commises sur les $\tilde{\xi}_{\mathbf{u}}$ s'ajoutent alors pour la détermination de la variance partielle.

4.1.5 Estimation des indices de Sobol avec un métamodèle ME

Le raisonnement est reconduit avec un métamodèle ME de la forme (212). Ce métamodèle peut être obtenu à l'aide de l'algorithme ME-ACD développé au chapitre 3 ou par une autre méthode ME. Comme la base canonique a été utilisée pour les métamodèles locaux dans notre approche, un changement de base vers une base orthogonale est réalisé. De plus, les avantages des modèles ME modernes étant que les éléments ont des formes arbitraires et non pas des hyperpavés pour épouser les formes des discontinuités et des irrégularités va représenter un désavantage dans notre cas. En effet, les propriétés d'orthogonalité utilisées reposent sur le produit scalaire défini à la section 2.5.2.1 correspondant à un hyperpavé, et non pas à un élément de forme quelconque. Le choix à tout de même était fait de privilégier un métamodèle ME appris sur des éléments de forme quelconque afin d'éviter toute imprécision due au phénomène de Gibbs. Une approximation des éléments sous forme d'union d'hyperpavés est alors réalisée une fois le métamodèle ME entraîné, avec l'aide du classifieur de la méthode. Ce choix est discuté à la section 4.4. La méthodologie développée ci-dessous peut ainsi convenir à tout type de métamodèle ME, qu'ils soient classiques utilisant des éléments directement sous forme d'hyperpavés, ou qu'ils appliquent la logique de création de nombre minimal d'éléments.

4.1.5.1 Partition en union d'hyperpavés

L'hypercube unité \mathbb{H}_d , partitionné sous forme d'éléments de forme quelconque $\{E_e\}_{e=1}^{N_E}$, va ici être repartitionné en une union de N_H hyperpavés $\{H_h\}_{h=1}^{N_H}$ ne se chevauchant pas.

$$\bigcup_{h=1}^{N_h} H_h = \mathcal{H}_d \quad (269)$$

$$H_{h_1} \cap H_{h_2} = \emptyset \text{ if } h_1 \neq h_2 \quad (270)$$

De plus, afin de respecter des propriétés utilisées dans les démonstrations de la section 4.1.5.3, les sommets des hyperpavés doivent former une grille de dimension d . Le pas de cette grille peut être irrégulier pour une direction donnée et entre les différentes directions. Pour chaque direction $m \in \Omega_d$, le segment $[0, 1]$ composant l'hypercube unité $\mathbb{H}_d = [0, 1]^d$ dans la direction m est divisé en β_m segments dont les $\beta_m + 1$ extrémités sont rangées dans un ensemble fini

$$\mathcal{S}_m = \{s_{m,1}, \dots, s_{m,\beta_m+1}\} \quad (271)$$

où $s_{m,1} = 0$ et $s_{m,\beta_m+1} = 1$. Ces valeurs sont alors les m^{eme} coordonnées possibles des sommets des hyperpavés qui sont ainsi disposés sur une grille à pas irréguliers. Comme les sommets sont sur une grille, une autre logique d'indexage multidimensionnel peut être introduite.

Définition 9. *Un hyperpavé de la décomposition peut être indexé par un multi-indice \mathbf{h} où*

$$H_{\mathbf{h}} = [s_{1,h_1}, s_{1,h_1+1}] \times \cdots \times [s_{d,h_d}, s_{d,h_d+1}] \quad (272)$$

\mathcal{D}_h est alors l'ensemble des indices multidimensionnels possibles de la décomposition en hyperpavés. C'est l'ensemble des vecteurs de taille d qui ont pour m^{ieme} composante un entier de l'intervalle $[[1, \beta_m]]$.

Avec cette nouvelle partition de l'espace, le métamodèle ME va être lui-même approché par

$$\tilde{f}(\mathbf{x}) \approx \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} c_{\mathbf{h},\mathbf{k}} \Psi_{\mathbf{k}}(\mathbf{x}) \mathcal{I}_{H_{\mathbf{h}}}(\mathbf{x}) \quad (273)$$

Pour un hyperpavé $H_{\mathbf{h}}$, si $H_{\mathbf{h}} \subset E_e$ $n_{\mathbf{h}} = n_e$ et $\forall \mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}$, $c_{\mathbf{h},\mathbf{k}} = c_{e,\mathbf{k}}$. Dans ce cas, l'hyperpavé est appelé « complet ». Si une frontière entre au moins deux éléments traverse un hyperpavé, ce dernier est appelé hybride. Dans ce cas, $n_{\mathbf{h}} \leq \max_{e \in [[1, N_E]]} n_e$ et $\forall \mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}$

$$c_{\mathbf{h},\mathbf{k}} = \sum_{e=1}^{N_E} \omega_{\mathbf{h},e,\mathbf{k}} c_{e,\mathbf{k}} \quad (274)$$

où $c_{e,\mathbf{k}} = 0$ si $\mathbf{k} \in \mathcal{K}_n^{(d)}$ avec $n > n_e$.

Trois procédures sont alors envisagées pour déterminer les coefficients $c_{\mathbf{h},\mathbf{k}}$ en fonction des $c_{e,\mathbf{k}}$:

- **Ignorance** : cette procédure simple ignore tous les hyperpavés hybrides en stipulant que $\omega_{\mathbf{h},e,\mathbf{k}} = 0$ pour ces derniers.
- **Vote** : cette procédure ne prend en compte que l'élément occupant la fraction d'hypervolume de l'hyperpavé la plus importante, en stipulant que $\omega_{\mathbf{h},e,\mathbf{k}} = 1$ pour e , l'indice de l'élément dont l'intersection avec l'hyperpavé occupe l'hypervolume le plus important et $\omega_{\mathbf{h},e,\mathbf{k}} = 0$ pour les autres.
- **Raccordement** : les coefficients $w_{e,\mathbf{k}}$ sont pris comme étant la fraction de l'hypervolume de l'hyperpavé $H_{\mathbf{h}}$ occupée par l'élément E_e .

4.1.5.2 Changement de base

Pour chaque métamodèle local correspondant à un hyperpavé $H_{\mathbf{h}}$, un changement de base est réalisé afin d'utiliser la base orthonormale par rapport au produit scalaire $\langle \cdot, \cdot \rangle_{H_{\mathbf{h}}}$

définie par (188). Le métamodèle ME adapté sur les hyperpavés peut donc s'écrire

$$\tilde{f}(\mathbf{x}) \approx \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \Phi_{\mathbf{h},\mathbf{k}}(\mathbf{x}) \mathcal{I}_{H_{\mathbf{h}}}(\mathbf{x}) \quad (275)$$

avec, comme nouvelles notations

$$\Phi_{\mathbf{h},\mathbf{k}}(\mathbf{x}) = \prod_{m=1}^d \phi_{m,k_m,h_m}(x_m) = \prod_{m=1}^d \phi_{k_m}^{h_m,h_m+1}(x_m) \quad (276)$$

$$\mathcal{I}_{H_{\mathbf{h}}}(\mathbf{x}) = \prod_{m=1}^d \mathcal{I}_{m,h_m}(x_m) = \prod_{m=1}^d \mathcal{I}_{[s_m,h_m,s_m,h_m+1]}(x_m) \quad (277)$$

4.1.5.3 Estimation du vecteur $\tilde{\Xi}$

Définition 10. $\forall E \subset \Omega_d$,

$$v_{\mathbf{h}}^E = \prod_{m \in E} (s_{m,h_m+1} - s_{m,h_m}) \quad (278)$$

Il s'agit de l'hypervolume de la projection de $H_{\mathbf{h}}$ sur les directions indexées par les entiers compris dans E . Cette définition est étendue à $v_{\mathbf{h}}^{\emptyset} = 1$

Le raisonnement appliqué à la section 4.1.4 est alors reconduit ici dans le cas d'un métamodèle ME de forme (275).

En premier lieu, le terme $\mu_{\mathbf{i}}^{(j)}$ est exprimé en fonction des coefficients des modèles locaux

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \int_{\mathcal{H}_{d-j}} \Phi_{\mathbf{h},\mathbf{k}}(\mathbf{x}) \mathcal{I}_{H_{\mathbf{h}}}(\mathbf{x}) d\mathbf{x}_{\sim \mathbf{i}} \quad (279)$$

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \prod_{m \in \overline{\mathcal{S}(\mathbf{i})}} \int_0^1 \phi_{m,h_m,k_m}(x_m) \mathcal{I}_{m,h_m}(x_m) dx_m \prod_{m \in \mathcal{S}(\mathbf{i})} \phi_{m,h_m,k_m}(x_m) \mathcal{I}_{m,h_m}(x_m) \quad (280)$$

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{i})}}} \prod_{m \in \overline{\mathcal{S}(\mathbf{i})}} \langle \phi_{m,h_m,k_m}, \phi_{m,h_m,0} \rangle_{s_{m,h_m}, s_{m,h_m+1}} \prod_{m \in \mathcal{S}(\mathbf{i})} \phi_{m,h_m,k_m}(x_m) \mathcal{I}_{m,h_m}(x_m) \quad (281)$$

$$\mu_{\mathbf{i}}^{(j)}(\mathbf{x}_{\mathbf{i}}) = \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{i})}}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{i})}} \prod_{m \in \mathcal{S}(\mathbf{i})} \phi_{m,h_m,k_m}(x_m) \mathcal{I}_{m,h_m}(x_m) \quad (282)$$

Les termes de la matrice Ξ peuvent ensuite être déduits

$$\mu_{\mathbf{u}}^{(a_1)} \mu_{\mathbf{v}}^{(a_2)} = \sum_{\mathbf{g} \in \mathcal{D}_h} \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{m} \in \mathcal{K}_{n_{\mathbf{g}}}^{(d)}} \sum_{\mathbf{n} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{g},\mathbf{m}} \alpha_{\mathbf{h},\mathbf{n}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{u})}}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{v})}}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v})}} \prod_{q \in \mathcal{S}(\mathbf{u})} \phi_{q,g_q,m_q} \mathcal{I}_{q,g_q} \prod_{r \in \mathcal{S}(\mathbf{v})} \phi_{r,h_r,n_r} \mathcal{I}_{r,h_r} \quad (283)$$

Or

$$\prod_{q \in \mathcal{S}(\mathbf{u})} \phi_{q,g_q,a_q} \mathcal{I}_{q,g_q} \prod_{r \in \mathcal{S}(\mathbf{v})} \phi_{r,h_r,b_r} \mathcal{I}_{r,h_r} = \prod_{q \in \mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v})} \phi_{q,g_q,a_q} \mathcal{I}_{q,g_q} \prod_{r \in \mathcal{S}(\mathbf{v}) \setminus \mathcal{S}(\mathbf{u})} \phi_{r,h_r,b_r} \mathcal{I}_{r,h_r} \prod_{t \in \mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \phi_{t,g_t,a_t} \mathcal{I}_{t,g_t} \phi_{t,h_t,b_t} \mathcal{I}_{t,h_t} \quad (284)$$

et

$$\prod_{t \in \mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \phi_{t,g_t,a_t} \mathcal{I}_{t,g_t} \phi_{t,h_t,b_t} \mathcal{I}_{t,h_t} = \Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \prod_{t \in \mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \phi_{t,h_t,a_t} \phi_{t,h_t,b_t} \mathcal{I}_{t,h_t} \quad (285)$$

donc

$$\xi_{\mathbf{u},\mathbf{v}} = \sum_{\mathbf{g} \in \mathcal{D}_h} \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{a} \in \mathcal{K}_{n_{\mathbf{g}}}^{(d)}} \sum_{\mathbf{b} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{g},\mathbf{a}} \alpha_{\mathbf{h},\mathbf{b}} \sqrt{v_{\mathbf{g}}^{\overline{\mathcal{S}(\mathbf{u})}}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{v})}}} \sqrt{v_{\mathbf{g}}^{\overline{\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v})}}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{v}) \setminus \mathcal{S}(\mathbf{u})}}} \Delta_{\mathbf{a},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \Delta_{\mathbf{b},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{v})}} \Delta_{\mathbf{a},\mathbf{0}}^{\mathcal{S}(\mathbf{u}) \setminus \mathcal{S}(\mathbf{v})} \Delta_{\mathbf{b},\mathbf{0}}^{\mathcal{S}(\mathbf{v}) \setminus \mathcal{S}(\mathbf{u})} \Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \Delta_{\mathbf{a},\mathbf{b}}^{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \quad (286)$$

Appliquer la formule (144) souligne le fait que toutes les composantes de \mathbf{a} et de \mathbf{b} avec des indices dans $\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})}$ doivent être égales à 0. De plus, toutes les composantes de \mathbf{a} et de \mathbf{b} d'indices dans $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})$ doivent être égales. Ainsi

$$\xi_{\mathbf{u},\mathbf{v}} = \sum_{\mathbf{g} \in \mathcal{D}_h} \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h},\mathbf{g}}}^{(d)}} \alpha_{\mathbf{g},\mathbf{k}} \alpha_{\mathbf{h},\mathbf{k}} \sqrt{v_{\mathbf{g}}^{\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})}}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})}}} \Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})}} \quad (287)$$

où $n_{\mathbf{h},\mathbf{g}} = \min(n_{\mathbf{h}}, n_{\mathbf{g}})$.

D'après (287), tous les termes $\xi_{\mathbf{u},\mathbf{v}}$ ayant le même $\mathcal{S}(\mathbf{u}) \cap \mathcal{S}(\mathbf{v})$ sont égaux. La proposition (247) est alors respectée et le vecteur $\tilde{\Xi}$ correspondant est

$$\tilde{\xi}_{\mathbf{u}} = \sum_{\mathbf{g} \in \mathcal{D}_h} \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h},\mathbf{g}}}^{(d)}} \alpha_{\mathbf{g},\mathbf{k}} \alpha_{\mathbf{h},\mathbf{k}} \sqrt{v_{\mathbf{g}}^{\overline{\mathcal{S}(\mathbf{u})}}} \sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{u})}}} \Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u})} \Delta_{\mathbf{k},\mathbf{0}}^{\overline{\mathcal{S}(\mathbf{u})}} \quad (288)$$

La variance peut alors être estimée par (250), les variances partielles par (251) et les quantités $D_i^{(T)}$ par (256) en utilisant les composantes du vecteur $\tilde{\Xi}$.

Pour montrer la cohérence de la formule, le cas PC peut être retrouvé. En effet, un métamodèle PC peut être vu comme un métamodèle ME particulier de forme (275) avec seulement un hyperpavé dans sa décomposition, l'hypercube unité. Dans ce cas particulier, $\mathcal{D}_h = \{(0, \dots, 0)\}$, $\sqrt{v_{\mathbf{h}}^{\overline{\mathcal{S}(\mathbf{u})}}} = 1$, et $\Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u})} = 1$, ce qui permet d'obtenir directement (264) à partir de (288).

4.2 Méthode numérique

Dans cette section, une méthode numérique est proposée pour obtenir les termes du vecteur $\tilde{\mathbf{E}}$ dans le cas d'un métamodèle ME. Un algorithme vectorisé est alors disponible à chaque étape, si l'utilisateur dispose d'assez de ressource mémoire pour les utiliser. Dans le cas contraire, une méthode itérative « on-the-run » peut être envisagée en alternative à chaque étape.

Pour illustrer chacune des étapes de la méthode, la QOI suivante est étudiée comme exemple

$$g^{\text{discontinuous-2}}(\mathbf{x}) = \begin{cases} \exp(x_1 + x_2) = e^{x_1} \times e^{x_2} & \text{if } x_1 < \frac{\pi}{6} \text{ and } x_2 < \frac{\pi}{6} \\ 0 & \text{sinon} \end{cases} \quad (289)$$

Il s'agit de la fonction de Genz discontinue [21] de dimension 2. La discontinuité est réglée pour être le carré $[0, \frac{\pi}{6}] \times [0, \frac{\pi}{6}]$, afin d'éviter que l'extrémité de ses côtés coïncident avec un nombre de la forme $1/n$ avec $n \in \mathbb{N}$, ce qui présente un avantage pour la méthode présentée dans cette section.

4.2.1 Partition en union d'hyperpavés

Partition cartésienne En premier lieu, l'hypercube unité est partitionné en z^d hypercubes dont les sommets forment une grille de pas régulier, avec z hypercubes par direction (le pas est donc de $1/z$ dans chaque direction). Les $(z + 1)^d$ sommets des hypercubes composent donc cette grille de dimension d . Numériquement, ces hypercubes peuvent être représentés dans un tenseur d'entiers \mathbf{Q} de taille (z, \dots, z) . $H_{\mathbf{h}}$ est alors représenté par l'entier Q_{h_1, \dots, h_d} . Pour un hypercube complet, l'entier correspondant est e , l'indice de l'élément qui contient l'entièreté de l'hypercube ($H_{\mathbf{h}} \subset E_e$). Pour un hypercube hybride, l'entier correspondant est arbitrairement $N_E + 1$. Pour obtenir \mathbf{Q} , les $(z + 1)^d$ sommets des hypercubes sont générés et classifiés à l'aide de la procédure de classification associée au modèle ME considéré (un réseau de neurone dans le cas de ME-ACD). Pour chaque hypercube, si tous ses sommets sont déterminés comme étant dans le même élément par la méthode de classification, l'hypercube est estimé être complet et son entier associé est l'indice de l'élément. Dans le cas contraire, si au moins deux sommets sont déterminés comme étant dans deux éléments différents, l'hypercube est estimé être hybride et l'entier correspondant est $N_E + 1$.

L'algorithme 5 présente un algorithme vectorisé permettant d'obtenir le tenseur \mathbf{Q} modélisant les hypercubes de la première étape de la partition. Si chaque entier du tenseur est codé sur un octet, la taille maximum du tenseur dans l'algorithme est de $(z + 1)^d$ octets. Le vecteur \mathbf{X} contenant $d \times (z + 1)^d$ flottants est le plus problématique en terme de coût mémoire. Afin d'économiser de la mémoire, la première implémentation du tenseur \mathbf{Q} contenant les labels des points de \mathbf{X} peut être effectué en considérant des sous ensembles de ces points, et une parallélisation peut également être envisagée.

Algorithm 5 Pseudo-code de l'étape de partition cartésienne (Vectorisé)

$z \leftarrow$ nombre d'hypercubes par direction

$\mathbf{x} \leftarrow$ vecteur de taille $z + 1$ où $x_i = \frac{i-1}{z}$

$\mathbf{X} \leftarrow$ points de la grille de dimension d de pas $1/z$, points dont les composantes sont les combinaisons avec remplacement des d composantes de \mathbf{x} .

$\mathbf{Q} \leftarrow$ indices des éléments contenant les points \mathbf{X} donnés par la procédure de classification de la méthode ME utilisée mis sous la forme d'un tenseur de forme $(z + 1, \dots, z + 1)$

for $m \in \llbracket 1, d \rrbracket$ **do**

$\mathbf{D} \leftarrow$ tenseur de forme (z, \dots, z) contenant les différences entre les composantes de \mathbf{Q} avec le précédent voisin dans la direction m

 Toutes les composantes de \mathbf{Q} ayant une composante de même indice dans \mathbf{D} différente de 0 est mis à $N_E + 1$ (indiquant un hypercube hybride). Les composantes de \mathbf{Q} ayant comme m^{eme} indice $z + 1$ (les points situés sur les hyperfaces supérieures) restent inchangées (ils seront retirés à la dernière étape).

end for

$\mathbf{Q} \leftarrow \mathbf{Q}$ sans ses composantes ayant au moins un de ses indices égal à $z + 1$.

Une fois ces hypercubes obtenus, une étape de fusion peut être considérée. En effet, dans les démonstrations de la section 4.1.5.3, les hyperpavés n'ont pas la contrainte d'être des hypercubes. Le nombre d'hyperpavés $N_h = \beta_1 \times \dots \times \beta_d$ est une variable cruciale pour le coût de l'algorithme, que ce soit en espace mémoire ou en temps de calcul. Cette étape a donc pour but de réduire le nombre d'hyperpavés, fusionnant les hypercubes avec leur voisin de même label pour obtenir $N_H \leq z^d$ hyperpavés. Néanmoins, les sommets des hyperpavés doivent toujours former une grille de dimension d . Pour respecter cette contrainte, pour chaque direction m , un hyperpavé ayant h_m comme m^{eme} composante d'indice est fusionné avec un hyperpavé ayant $h_m + 1$ comme m^{eme} composante d'indice (nommé son voisin) si seulement tous les hyperpavés ayant h_m comme m^{eme} composante d'indice ont le même label que leur voisin dans la direction m .

Algorithm 6 Pseudo-code de l'étape de fusion (Vectorisée)

$\mathbf{Q} \leftarrow (z, \dots, z)$ tenseur d'entier représentant les hypercubes obtenu par partition cartésienne

$\mathbf{x} \leftarrow$ vecteur de taille $z + 1$ où $x_i = \frac{i-1}{z}$

$S \leftarrow$ liste contenant d listes identiques avec les éléments de \mathbf{x} ; la m^{eme} liste de S , est appelée S_m

for $m \in \llbracket 1, d \rrbracket$ **do**

$\mathbf{D} \leftarrow$ tenseur de forme (z, \dots, z) contenant les différences entre les composantes de \mathbf{Q} et leur précédent voisin dans la direction m

$I \leftarrow$ liste des entiers de $\llbracket 1, z \rrbracket$ où toutes les composantes de \mathbf{D} ayant comme m^{eme} indice cet entier sont égaux à 0.

$\mathbf{Q} \leftarrow \mathbf{Q}$ sans les composantes ayant comme m^{eme} indice un indice dans I .

$S_m \leftarrow S_m$ sans les composantes ayant un indice dans I

end for

L'algorithme 6 présente le pseudo-code de l'algorithme vectorisé permettant d'obtenir les hyperpavés labellisés à partir des hypercubes obtenus à l'étape précédente. Les hyperpavés sont représentés par le tenseur \mathbf{Q} d'ordre d et de taille $(\beta_1, \dots, \beta_d)$ contenant les labels de

ces hyperpavés et par S une liste de d listes de taille β_m représentant numériquement les d ensembles finis \mathcal{S}_m .

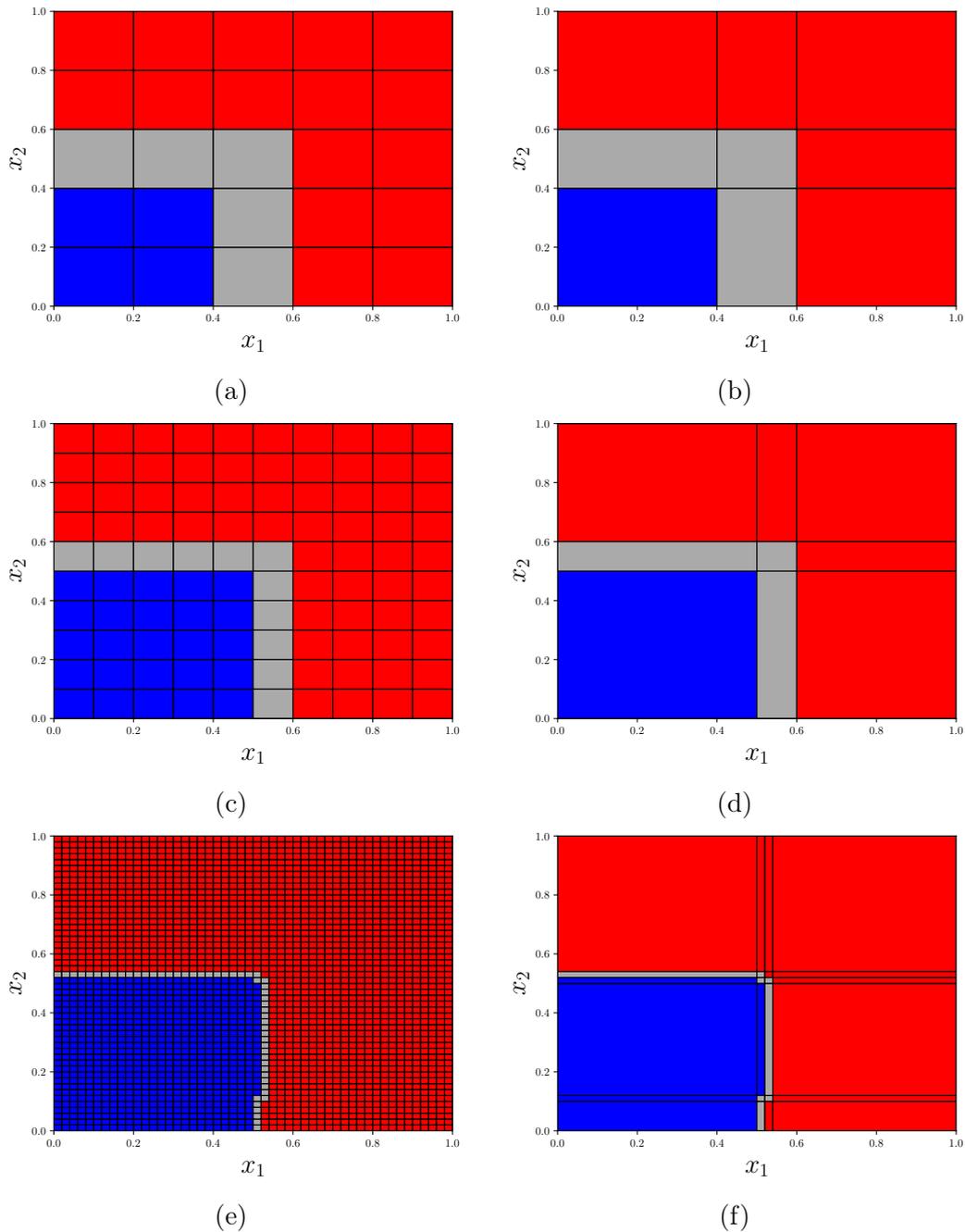


FIGURE 21 – Sur la gauche, la partition de l’espace des paramètres \mathbb{H}_2 en z^d hypercubes dans le cas de la QOI exemple $g^{discontinue-2}$. Les couleurs rouge et bleue représentent les deux éléments déterminés par le classifieur de l’algorithme ME-ACD. Les hypercubes hybrides sont représentés en gris. (a) $z = 5$. (c) $z = 10$. (e) $z = 50$. Sur la droite, la partition en hyperpavés finale obtenue après l’étape de fusion sur les z^d hypercubes dans le cas de la QOI exemple $g^{discontinue-2}$. (b) $z = 5$ menant à $N_h = 9$ hyperpavés. (d) $z = 10$ menant à $N_h = 9$ hyperpavés. (f) $z = 50$ menant à $N_h = 24$ hyperpavés.

La figure 21 présente les résultats de la partition cartésienne (Figures 21a, 21c et 21e) et de l’étape de fusion (Figures 21b, 21d et 21f) pour la QOI exemple $g^{discontinue-2}$. Comme

il est illustré sur la figure, si z est relativement faible, le nombre d'hyperpavés est relativement faible, mais les hyperpavés hybrides (en gris sur la figure) occupent un hypervolume relativement important, ce qui altère la précision de l'estimation. Si z est trop élevé, la partition en hyperpavés devient sensible aux frontières données par la procédure de classification (si cette procédure est autre que l'utilisation d'une formule analytique de l'équation des frontières), comme les hyperpavés sont petits et situés entre les observations d'entraînement. Cela peut ainsi mener à une augmentation significative du nombre d'hyperpavés. De plus, comme l'hyperpavé est situé entre les observations d'entraînement, la classification n'est pas effectuée avec une grande confiance et l'hyperpavé peut être mal labellisé. z est en ce sens un paramètre de compromis.

4.2.2 Changement de base

Dans le but d'avoir accès aux coefficients $\alpha_{\mathbf{h},\mathbf{k}}$ des métamodèles locaux exprimés dans les bases orthonormées par rapport au produit scalaire $\langle \cdot, \cdot \rangle_{H_{\mathbf{h}}}$, un changement de base est effectué pour chaque métamodèle local (et donc pour chaque hyperpavé). Pour chaque hyperpavé, les coefficients des polynômes de la base $(\Phi_{\mathbf{k}}^{H_{\mathbf{h}}})_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}}$ doivent être implémentés, où n_h est le degré du métamodèle sur l'hypercube $H_{\mathbf{h}}$. Dans le but d'obtenir ces coefficients, une matrice de passage $\Lambda_{\mathbf{h}}$ de taille $\binom{n_{\mathbf{h}}+d}{n_{\mathbf{h}}}, \binom{n_{\mathbf{h}}+d}{n_{\mathbf{h}}}$ doit être construite.

Toutes ces matrices peuvent être stockées dans un tenseur d'ordre 3, Λ en associant le multi-index \mathbf{h} à un scalaire. Ce tenseur est de taille (N_H, n, n) , avec n le maximum des degrés des métamodèles locaux. Pour chaque \mathbf{h} , la matrice $\Lambda_{\mathbf{h}}$ est complétée de 0 pour les termes d'indices strictement supérieur à $\binom{n_{\mathbf{h}}+d}{n_{\mathbf{h}}}$ en ligne ou en colonne.

Les termes de la matrice de passage $\Lambda_{\mathbf{h}}$ sont les composantes des polynômes multivariés de la base orthonormale $(\Phi_{\mathbf{H}_{\mathbf{h}}}^k)_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}}$ exprimés dans la base canonique de $\mathbb{R}[X_1, \dots, X_d]$. Pour un métamodèle local

$$\sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} c_{\mathbf{h},\mathbf{k}} \Psi_{\mathbf{k}}(\mathbf{x}) = \sum_{\mathbf{q} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \left(\sum_{\mathbf{k} \in \mathcal{K}_{n_{\mathbf{h}}}^{(d)}} \alpha_{\mathbf{h},\mathbf{k}} \Lambda_{\mathbf{h},\mathbf{k},\mathbf{q}} \right) \Psi_{\mathbf{q}}(\mathbf{x}) \quad (290)$$

En notant $\mathbf{C}_{\mathbf{h}}$ le vecteur contenant les coefficients $c_{\mathbf{h},\mathbf{k}}$ et $A_{\mathbf{h}}$ le vecteur contenant les $\alpha_{\mathbf{h},\mathbf{k}}$

$$\mathbf{C}_{\mathbf{h}}^T = \Lambda_{\mathbf{h}}^T A_{\mathbf{h}} \quad (291)$$

Une résolution triangulaire supérieure permet d'obtenir efficacement $A_{\mathbf{h}}$.

Comme démontré à la section 2.5.2.3, les composantes des polynômes orthogonaux sont

$$\Lambda_{\mathbf{h},\mathbf{k},\mathbf{q}} = \prod_{m=1}^d \lambda_{k_m, q_m}^{s_{h_m}, s_{h_m+1}} \quad (292)$$

De nombreux couples (s_{h_m}, s_{h_m+1}) peuvent être similaires suite à la partition de l'espace. Pour éviter de répéter des calculs, une base de données de $\lambda_{k,q}^{a,b}$ peut être construite. Pour chaque couple (a, b) , un indice w est attribué. Le nombre de couples (a_w, b_w) différents est noté N_w .

$$\lambda_{w,k,q} = \sqrt{\frac{2k+1}{b_w - a_w}} \sum_{y=q}^k \frac{l_{k,y}}{(b_w - a_w)^y} \binom{y}{q} (2)^q (-1)^{y-q} (a_w + b_w)^{y-q} \quad (293)$$

Deux matrices sont alors introduites pour implémenter les coefficients et le tenseur :

- I est une matrice de taille (N_h, d) où $I_{\mathbf{h},m}$ vaut un entier w qui signifie que les extrémités du segment $1D$ de direction m de l'hypercube $H_{\mathbf{h}}$ est le couple (a_w, b_w) indexé par w .
- J est une matrice de taille $(N_w, 2)$ où la ligne w relie l'indice w à un couple (a_w, b_w)

$$\Lambda_{\mathbf{h},\mathbf{k},\mathbf{q}} = \prod_{m=1}^d \lambda_{I_{\mathbf{h},m},k_m,q_m} \quad (294)$$

Exemple 17. Avec la QOI exemple $g^{disccontinue-2}$ et $z = 5$, dont la partition en hyperpavés est illustrée sur la Figure 21b, $S_1 = S_2 = [0, 0.4, 0.6, 1]$, $N_h = 9$, $N_w = 3$ et

$$I = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{pmatrix} \quad J = \begin{pmatrix} 0 & 0.4 \\ 0.4 & 0.6 \\ 0.6 & 1 \end{pmatrix}$$

Exemple 18. Avec la QOI exemple $g^{disccontinue-2}$ et $z = 50$, dont la partition en hyperpavés est illustrée sur la Figure 21f, $S_1 = [0, 0.5, 0.52, 0.54, 1]$, $S_2 = [0, 0.1, 0.12, 0.5, 0.52, 0.54, 1]$, $N_h = 24$, $N_w = 7$ et

$$I = \begin{pmatrix} 1 & 5 \\ 2 & 5 \\ 3 & 5 \\ 4 & 5 \\ 1 & 6 \\ 2 & 6 \\ 3 & 6 \\ 4 & 6 \\ 1 & 7 \\ 2 & 7 \\ 3 & 7 \\ 4 & 7 \\ 1 & 2 \\ 2 & 2 \\ 3 & 2 \\ 4 & 2 \\ 1 & 3 \\ 2 & 3 \\ 3 & 3 \\ 4 & 3 \\ 1 & 4 \\ 2 & 4 \\ 3 & 4 \\ 4 & 4 \end{pmatrix} \quad J = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0.52 \\ 0.52 & 0.54 \\ 0.54 & 1 \\ 0 & 0.1 \\ 0.1 & 0.12 \\ 0.12 & 0.5 \end{pmatrix}$$

Chaque coefficient $\lambda_{w,k,q}$ de la base de données peut être exprimé

$$\lambda_{w,k,q} = \sum_{y=q}^k B_{w,q,y} L_{k,q,y} \quad (295)$$

$$B_{w,q,y} = \frac{1}{\sqrt{b_w - a_w}} \left(-\frac{1}{a_w + b_w} \right)^q \left(-\frac{a_w + b_w}{b_w - a_w} \right)^y \quad (296)$$

$$L_{k,q,y} = 2^q \sqrt{2k+1} \binom{y}{q} l_{k,y} \quad (297)$$

où le tenseur \mathbf{L} est indépendant de la partition et le tenseur \mathbf{B} est dépendant de la partition. Ces deux tenseurs peuvent être implémentés avec un algorithme vectorisé. De ce fait, la base de données des coefficients $1D \lambda_{w,k,q}$ peut être implémentée efficacement et le tenseur de passage Λ peut également être obtenu par vectorisation.

Maintenant que toutes les matrices $\Lambda_{\mathbf{h}}$ sont implémentées, la seule étape restante est celle consistant à implémenter $\mathbf{C}_{\mathbf{h}}$ pour obtenir ensuite les coefficients $\mathbf{A}_{\mathbf{h}}$ pour chaque hyperpavé. Pour un hyperpavé complet, les composantes de $\mathbf{C}_{\mathbf{h}}$ sont les $c_{e,\mathbf{k}}$. Pour un hyperpavé hybride, la formule (274) est utilisée. Si la procédure sélectionnée est « Ignorance », $\mathbf{C}_{\mathbf{h}}$ est le vecteur nul. Sinon, la fraction de l'hypermolume de $H_{\mathbf{h}}$ occupée par l'élément E_e doit être estimée. Pour cela, N_{frac} observations sont générées par une séquence de Sobol dans l'hyperpavé $H_{\mathbf{h}}$. Chaque point est ensuite classé par la procédure de classification

de la méthode ME et la fraction de l'hypervolume occupée par l'élément e est estimée comme étant le ratio entre le nombre de points ayant comme label e sur les N_{frac} points échantillonnés. Les coefficients $\omega_{\mathbf{h},e,\mathbf{k}}$ sont ensuite déterminés en fonction de la procédure sélectionnée et de cette estimation de l'hypervolume occupé.

4.2.3 Implémentation du vecteur $\tilde{\Xi}$

(288) peut être réécrite

$$\tilde{\xi}_{\mathbf{u}} = \sum_{\mathbf{g} \in \mathcal{D}_h} \sum_{\mathbf{h} \in \mathcal{D}_h} \sum_{\mathbf{k} \in \mathcal{K}_{\mathbf{h},\mathbf{g}}^{(d)}} \mathcal{V}_{\mathbf{g},\mathbf{k}}^{\mathcal{S}(\mathbf{u})} \mathcal{V}_{\mathbf{h},\mathbf{k}}^{\mathcal{S}(\mathbf{u})} \Delta_{\mathbf{g},\mathbf{h}}^{\mathcal{S}(\mathbf{u})} \quad (298)$$

$$\mathcal{V}_{\mathbf{h},\mathbf{k}}^{\mathcal{S}(\mathbf{u})} = \alpha_{\mathbf{h},\mathbf{k}} \sqrt{v_{\mathbf{h}}^{\mathcal{S}(\mathbf{u})}} \Delta_{\mathbf{k},\mathbf{0}}^{\mathcal{S}(\mathbf{u})} \quad (299)$$

De même qu'avec les étapes précédentes, $\tilde{\Xi}$ peut être implémenté de façon vectorisée ou de façon « on-the-run » avec une potentielle parallélisation.

4.3 Applications

4.3.1 Métrique d'erreur

Afin de comparer la précision de l'estimation des variances partielles avec un métamodèle PC et un métamodèle ME, la métrique suivante est utilisée

$$\left(E_{\mathbf{i}}^{(j)}\right)_{model} = \frac{\left| \left(D_{\mathbf{i}}^{(j)}\right)_{model} - \left(D_{\mathbf{i}}^{(j)}\right)_{ref} \right|}{\left(D\right)_{ref}} \quad (300)$$

où :

- $\left(D\right)_{ref}$ est la valeur de la variance obtenue par la méthode Monte-Carlo avec un nombre élevé d'observations N_{MC} .
- $\left(D_{\mathbf{i}}^{(j)}\right)_{ref}$ est la valeur de la variance partielle obtenue par la méthode Monte-Carlo (Cf. section 2.4.3.5) avec un nombre élevé d'observations N_{MC} (N_{MC} observations sont générées par quantité).
- $\left(D_{\mathbf{i}}^{(j)}\right)_{model}$ est la valeur de la même variance partielle obtenue par estimation à l'aide du métamodèle considéré et de sa méthode associée ; la méthode utilisée pour l'estimation à l'aide d'un métamodèle PC étant la méthode classique.

Pour les simulations Monte-Carlo, le nombre d'observations N_{MC} est spécifié, la convergence étant assurée après comparaison des résultats en augmentant successivement le nombre d'observations. Le métamodèle ME est obtenu en utilisant la méthode ME-ACD développée au chapitre 3, les hyperparamètres utilisés pour chaque simulation étant spécifiés. Pour la reproductibilité des résultats, l'ensemencement aléatoire du réseau de neurones classifieur a pour valeur 0. Pour entraîner le métamodèle PC, la méthode OLS est utilisée sur les mêmes observations d'entraînement utilisées pour entraîner le métamodèle ME, mais avec des métamodèles PC de degrés allant de n_{min-PC} à n_{max-PC} . Le métamodèle sélectionné est celui minimisant l'erreur PRESS. Comme l'erreur PRESS est une erreur de validation croisée, le surapprentissage a de grandes chances d'être évité.

L'échantillonnage des observations d'entraînement (pour toutes les méthodes) est réalisé à l'aide d'une séquence de Sobol dans l'hypercube unité. Les méthodes sont comparées avec deux échantillonnages de nombre d'observations d'entraînement N différents, un avec un nombre d'observations relativement faible et un second avec un nombre d'observations plus élevé.

4.3.2 Fonction d'Ishigami

La fonction d'Ishigami est la QOI suivante

$$f^{Ishigami}(p_1, p_2, p_3) = \sin(p_1) + a(\sin(p_2))^2 + bp_3^4 \sin(p_1) \quad (301)$$

les paramètres d'entrée p_1 , p_2 et p_3 suivant tous une loi uniforme sur $[-\pi, \pi]$.

C'est une QOI étudiée de façon classique en analyse de sensibilité [15, 64]. De façon classique, le choix des paramètres a et b est de $a = 7$ et $b = 0.1$. La QOI est relativement régulière et ses variances partielles sont connues. Les valeurs théoriques de ces dernières ainsi que les valeurs théoriques des quantités $D_i^{(T)}$ sont disponibles à la Table 11. Pour cette application seulement, les valeurs de référence $\left(D_i^{(j)}\right)_{ref}$ et $\left(D\right)_{ref}$ utilisées pour le calcul de l'erreur définie à la section 4.3.1 sont ces valeurs et non pas des valeurs obtenues à l'aide de la méthode MC.

Les hyperparamètres utilisés sont indiqués dans la Table 12. La Table 13 rassemble quant à elle les résultats des estimations obtenues à l'aide des méthodes PC et ME.

Pour cette application, comme la QOI est assez régulière, ME-ACD offre un modèle identique à celui obtenu par la méthode PC classique, sans partition de l'espace des paramètres. La comparaison est donc essentiellement entre l'approche classique et l'approche développée au chapitre 4 permettant de relier les coefficients d'un métamodèles PC aux quantités de variance. Il a été conjecturé à la section 4.1.4 que la nouvelle approche avec le vecteur $\tilde{\Xi}$ dans le cas PC serait moins précise que l'approche standard. Cela est vérifié, mais les différences numériques sont très faibles, si bien qu'elles n'apparaissent pas dans la Table 13. Il est alors vérifié que la nouvelle approche avec le vecteur $\tilde{\Xi}$ est quasiment équivalente à la méthode classique pour cette application. Les erreurs entre les valeurs

Variances	Valeur théorique	Valeur arrondie avec $a = 7$ et $b = 0.1$
D	$\frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{1}{2} + \frac{a^2}{8}$	13.845
$D_1^{(1)}$	$\frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} + \frac{1}{2}$	4.346
$D_2^{(1)}$	$\frac{a^2}{8}$	6.125
$D_3^{(1)}$	0	0
$D_{(1,2)}^{(2)}$	0	0
$D_{(1,3)}^{(2)}$	$\frac{8b^2\pi^8}{225}$	3.374
$D_{(2,3)}^{(2)}$	0	0
$D_{(1,2,3)}^{(3)}$	0	0
$D_1^{(T)}$	$\frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{1}{2}$	7.720
$D_2^{(T)}$	$\frac{a^2}{8}$	6.125
$D_3^{(T)}$	$\frac{8b^2\pi^8}{225}$	3.374

TABLE 11 – Valeurs théoriques des variances partielles $D_i^{(j)}$ et des quantités $D_i^{(T)}$ et D pour la QOI fonction d’Ishigami.

	PC	ME-ACD
N		{70, 150}
$N_{E_{min}}$		1
$N_{E_{max}}$		5
k_{min}		5
k_{max}		10
n_{min}	0	0
n_{max}	10	10
δ_{min}		10^{-5}
Taux d’apprentissage initiaux testés		Cf section 3.3
$isConnect$		Vrai
$k_{connect}$		10
z		50
N_{frac}		1000

TABLE 12 – Hyperparamètres sélectionnés pour les méthodes PC et ME-ACD pour l’étude de la fonction d’Ishigami.

théoriques et les valeurs numériques proviennent des erreurs d’approximation du méta-modèle PC/ME. L’erreur PRESS normalisée vaut $\delta = 0.83$ pour $N = 70$ (le degré optimal étant $n = 1$, ce qui explique que les variances partielles d’ordre supérieure ou égale à 2 ne peuvent pas être estimées) et $\delta = 0.26$ pour $N = 150$ ($n = 6$).

4.3.3 Fonction de Genz discontinue : comparaison des trois procédures concernant les hyperpavés hybrides

Dans cette section, la QOI suivante est étudiée

	$N = 70$					$N = 150$			
	$D_{i,ref}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$
D	13.8446	2.1187	2.1187	$8.47 \cdot 10^{-1}$	$8.47 \cdot 10^{-1}$	14.3965	14.3965	$3.99 \cdot 10^{-2}$	$3.99 \cdot 10^{-2}$
$D_1^{(1)}$	4.3459	2.1067	2.1067	$1.62 \cdot 10^{-1}$	$1.62 \cdot 10^{-1}$	4.6200	4.6200	$1.98 \cdot 10^{-2}$	$1.98 \cdot 10^{-2}$
$D_2^{(1)}$	6.1250	0.0104	0.0104	$4.42 \cdot 10^{-1}$	$4.42 \cdot 10^{-1}$	5.8360	5.8360	$2.09 \cdot 10^{-2}$	$2.09 \cdot 10^{-2}$
$D_3^{(1)}$	0.0000	0.0016	0.0016	$1.17 \cdot 10^{-4}$	$1.17 \cdot 10^{-4}$	0.0140	0.0140	$1.01 \cdot 10^{-3}$	$1.01 \cdot 10^{-3}$
$D_{(1,2)}^{(2)}$	0.0000	0.0000	0.0000	$1.00 \cdot 10^{-14}$	$1.00 \cdot 10^{-14}$	0.0955	0.0955	$6.90 \cdot 10^{-3}$	$6.90 \cdot 10^{-3}$
$D_{(1,3)}^{(2)}$	3.3737	0.0000	0.0000	$2.44 \cdot 10^{-1}$	$2.44 \cdot 10^{-1}$	3.4697	3.4697	$6.94 \cdot 10^{-3}$	$6.94 \cdot 10^{-3}$
$D_{(2,3)}^{(2)}$	0.0000	0.0000	0.0000	$1.00 \cdot 10^{-14}$	$1.00 \cdot 10^{-14}$	0.2084	0.2084	$1.51 \cdot 10^{-2}$	$1.51 \cdot 10^{-2}$
$D_{(1,2,3)}^{(3)}$	0.0000	0.0000	0.0000	$1.28 \cdot 10^{-16}$	$1.00 \cdot 10^{-14}$	0.1528	0.1528	$1.10 \cdot 10^{-2}$	$1.10 \cdot 10^{-2}$
$D_1^{(T)}$	7.7196	2.1067	2.1067	$4.05 \cdot 10^{-1}$	$4.05 \cdot 10^{-1}$	8.3380	8.3380	$4.47 \cdot 10^{-2}$	$4.47 \cdot 10^{-2}$
$D_2^{(T)}$	6.1250	0.0104	0.0104	$4.42 \cdot 10^{-1}$	$4.42 \cdot 10^{-1}$	6.2928	6.2928	$1.21 \cdot 10^{-2}$	$1.21 \cdot 10^{-2}$
$D_3^{(T)}$	3.3737	0.0016	0.0016	$2.44 \cdot 10^{-1}$	$2.44 \cdot 10^{-1}$	3.8449	3.8449	$3.40 \cdot 10^{-2}$	$3.40 \cdot 10^{-2}$

TABLE 13 – Estimation des variances partielles pour la QOI $f^{Ishigami}$ à l’aide d’un métamodèle PC et d’un métamodèle ME.

$$g^{discontinue-4}(\mathbf{x}) = \begin{cases} \exp(x_1 + x_2 + x_3 + x_4) & \text{si } \forall i \in [1, 4], x_i < \frac{\pi}{6} \\ 0 & \text{sinon} \end{cases} \quad (302)$$

où les x_i suivent une loi uniforme sur $[0, 1]$.

Pour cette QOI, toutes les interactions entre les paramètres d’entrée ont un impact. Dans la nature, il est très rare d’avoir ce cas de figure. Cette QOI théorique permet néanmoins de mettre à l’épreuve la méthodologie et d’étudier précisément l’impact de certains paramètres sur l’estimation des quantités d’ordre différents. Dans cette application, l’impact du choix de la procédure de gestion des hyperpavés hybrides (« Ignorance », « Vote » ou « Raccordement ») est discuté à l’aide de la QOI.

Les hyperparamètres utilisés sont indiqués dans la Table 14. Les Tables 15, 16 et 17 contiennent les estimations obtenues respectivement avec les procédures « Ignorance », « Vote » et « Raccordement ». Les résultats avec le métamodèle PC sont identiques dans ces trois tables, cette procédure ne concernant que le métamodèle ME. Le nombre d’observations d’entraînement pour estimer les valeurs de référence est de $N_{MC} = 10^6$ observations.

Pour $N = 100$, le degré du métamodèle PC optimal est de 2 (erreur PRESS normalisée de $\delta = 8.1 \times 10^{-2}$); ceux des métamodèles locaux sont de 0 et de 1 (pour une erreur de partition normalisée de $\delta = 7,4 \times 10^{-4}$). Avec un degré $n = 1$, le métamodèle PC ne peut pas estimer les variances partielles d’ordre $j > 2$ avec $N = 100$. Le métamodèle ME permet quant à lui d’estimer ces valeurs même si les degrés sont inférieurs à 2. En augmentant le nombre d’observations d’entraînement à $N = 1000$, le métamodèle PC peut avoir un degré optimal plus élevé, $n = 5$ (erreur PRESS normalisée de $\delta = 3.1 \times 10^{-2}$), ce qui permet de mieux estimer les différentes quantités. Pour une procédure de gestion des hyperpavés hybrides autre qu’« Ignorance », le métamodèle ME permet d’avoir des estimations globalement plus précises que celles du métamodèle PC pour cette QOI dis-

	PC	ME-ACD
N		{100, 1000}
$N_{E_{min}}$		1
$N_{E_{max}}$		5
k_{min}		5
k_{max}		10
n_{min}	0	0
n_{max}	10	10
δ_{min}		10^{-5}
Taux d'apprentissage initiaux testés		Cf section 3.3
$isConnect$		Vrai
$k_{connect}$		10
z		25
N_{frac}		1000

TABLE 14 – Hyperparamètres sélectionnés pour les méthodes PC et ME-ACD pour l'étude de $g^{discontinue-4}$.

continue. Si le métamodèle PC est capable d'estimer les quantités au premier ordre avec une précision comparable à celle du métamodèle ME, les effets des interactions d'ordre supérieur sont quant à eux estimés avec moins de précision, le métamodèle PC approchant la discontinuité par une pente. Avec $N = 1000$, les degrés des métamodèles locaux pour la méthode ME sont de 0 et de 3 (erreur de partition normalisée $\delta = 1.2 \times 10^{-11}$). En augmentant le nombre d'observations, on observe des estimations globalement plus précises, même si pour certaines quantités, la précision est dégradée. Le métamodèle ME permet alors d'estimer les quantités avec plus de précision.

En ce qui concerne la comparaison des procédures de gestion des hyperpavés hybrides, la procédure « Ignorance » ne mène pas à une estimation précise, ignorant l'information située à l'interface entre les éléments. L'ajout d'observations d'entraînement n'améliore que très peu la précision des estimations. Les procédures « Vote » et « Raccordement » permettent quant à elle des estimations plus précises qu'avec le métamodèle PC en utilisant des métamodèles locaux ayant des degrés de troncature moins importants. La précision des estimations pour les deux procédures est assez similaire. On observe néanmoins que les estimations sont très sensibles à l'échantillonnage de l'espace (et donc au classifieur). Par exemple, la précision de la variance se dégrade dans cette application pour la procédure « Raccordement » alors que le nombre d'observations augmente. Ce manque de robustesse par rapport à l'échantillonnage a été observé également dans d'autres applications non rapportées dans le manuscrit. Il sera l'objet de futures recherches afin d'améliorer la méthodologie. La procédure utilisée par la suite est la procédure « Raccordement ».

4.3.4 Étude de sensibilité en présence d'une bifurcation

L'objectif final de l'étude est d'utiliser la méthodologie afin de conduire une étude de sensibilité efficace pour le système dynamique du vol (Cf. section 1.1.5.6), système présentant de nombreuses bifurcations. Un premier test sur un système théorique modélisant

	$N = 100$					$N = 1000$			
	$D_{i,ref}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$
D	0.1186	0.0959	0.0519	$1.92 \cdot 10^{-1}$	$5.62 \cdot 10^{-1}$	0.0905	0.0849	$2.37 \cdot 10^{-1}$	$2.85 \cdot 10^{-1}$
$D_1^{(1)}$	0.0088	0.0058	0.0087	$2.53 \cdot 10^{-2}$	$5.87 \cdot 10^{-4}$	0.0054	0.0080	$2.89 \cdot 10^{-2}$	$7.04 \cdot 10^{-3}$
$D_2^{(1)}$	0.0088	0.0059	0.0073	$2.46 \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$	0.0054	0.0081	$2.82 \cdot 10^{-2}$	$6.00 \cdot 10^{-3}$
$D_3^{(1)}$	0.0089	0.0041	0.0055	$4.02 \cdot 10^{-2}$	$2.82 \cdot 10^{-2}$	0.0054	0.0076	$2.91 \cdot 10^{-2}$	$1.05 \cdot 10^{-2}$
$D_4^{(1)}$	0.0087	0.0072	0.0070	$1.30 \cdot 10^{-2}$	$1.44 \cdot 10^{-2}$	0.0057	0.0080	$2.56 \cdot 10^{-2}$	$6.36 \cdot 10^{-3}$
$D_{(1,2)}^{(2)}$	0.0081	0.0056	0.0061	$2.15 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$	0.0053	0.0059	$2.39 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$
$D_{(1,3)}^{(2)}$	0.0079	0.0039	0.0030	$3.40 \cdot 10^{-2}$	$4.14 \cdot 10^{-2}$	0.0055	0.0056	$2.00 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$
$D_{(2,3)}^{(2)}$	0.0081	0.0040	0.0031	$3.46 \cdot 10^{-2}$	$4.15 \cdot 10^{-2}$	0.0054	0.0053	$2.27 \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$
$D_{(1,4)}^{(2)}$	0.0079	0.0071	0.0039	$7.20 \cdot 10^{-3}$	$3.40 \cdot 10^{-2}$	0.0059	0.0058	$1.75 \cdot 10^{-2}$	$1.82 \cdot 10^{-2}$
$D_{(2,4)}^{(2)}$	0.0080	0.0075	0.0052	$3.87 \cdot 10^{-3}$	$2.39 \cdot 10^{-2}$	0.0057	0.0061	$1.90 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$
$D_{(3,4)}^{(2)}$	0.0081	0.0049	0.0021	$2.69 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	0.0056	0.0058	$2.07 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$
$D_{(1,2,3)}^{(3)}$	0.0073	0.0045	0.0000	$2.38 \cdot 10^{-2}$	$6.17 \cdot 10^{-2}$	0.0061	0.0038	$1.06 \cdot 10^{-2}$	$2.95 \cdot 10^{-2}$
$D_{(1,2,4)}^{(3)}$	0.0072	0.0088	0.0000	$1.28 \cdot 10^{-2}$	$6.10 \cdot 10^{-2}$	0.0062	0.0044	$8.50 \cdot 10^{-3}$	$2.37 \cdot 10^{-2}$
$D_{(1,3,4)}^{(3)}$	0.0072	0.0057	0.0000	$1.25 \cdot 10^{-2}$	$6.10 \cdot 10^{-2}$	0.0067	0.0042	$4.71 \cdot 10^{-3}$	$2.60 \cdot 10^{-2}$
$D_{(2,3,4)}^{(3)}$	0.0072	0.0062	0.0000	$8.65 \cdot 10^{-3}$	$6.07 \cdot 10^{-2}$	0.0060	0.0041	$1.02 \cdot 10^{-2}$	$2.59 \cdot 10^{-2}$
$D_{(1,2,3,4)}^{(4)}$	0.0068	0.0148	0.0000	$6.82 \cdot 10^{-2}$	$5.69 \cdot 10^{-2}$	0.0102	0.0022	$2.93 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$
$D_1^{(T)}$	0.0612	0.0562	0.0217	$4.29 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0512	0.0399	$8.44 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$
$D_2^{(T)}$	0.0613	0.0572	0.0216	$3.46 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0503	0.0400	$9.25 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$
$D_3^{(T)}$	0.0611	0.0481	0.0138	$1.09 \cdot 10^{-1}$	$3.99 \cdot 10^{-1}$	0.0509	0.0387	$8.56 \cdot 10^{-2}$	$1.88 \cdot 10^{-1}$
$D_4^{(T)}$	0.0612	0.0622	0.0182	$8.99 \cdot 10^{-3}$	$3.62 \cdot 10^{-1}$	0.0521	0.0406	$7.69 \cdot 10^{-2}$	$1.74 \cdot 10^{-1}$

TABLE 15 – Estimation des variances pour la QOI $g^{discontinue-4}$ avec la procédure « Ignorance ».

	$N = 100$					$N = 1000$			
	$D_{i,ref}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$
D	0.1186	0.1235	0.0519	$4.08 \cdot 10^{-2}$	$5.62 \cdot 10^{-1}$	0.1117	0.0849	$5.84 \cdot 10^{-2}$	$2.85 \cdot 10^{-1}$
$D_1^{(1)}$	0.0088	0.0088	0.0087	$3.57 \cdot 10^{-4}$	$5.87 \cdot 10^{-4}$	0.0073	0.0080	$1.22 \cdot 10^{-2}$	$7.04 \cdot 10^{-3}$
$D_2^{(1)}$	0.0088	0.0085	0.0073	$2.02 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$	0.0076	0.0081	$9.84 \cdot 10^{-3}$	$6.00 \cdot 10^{-3}$
$D_3^{(1)}$	0.0089	0.0059	0.0055	$2.47 \cdot 10^{-2}$	$2.82 \cdot 10^{-2}$	0.0074	0.0076	$1.23 \cdot 10^{-2}$	$1.05 \cdot 10^{-2}$
$D_4^{(1)}$	0.0087	0.0107	0.0070	$1.63 \cdot 10^{-2}$	$1.44 \cdot 10^{-2}$	0.0080	0.0080	$6.43 \cdot 10^{-3}$	$6.36 \cdot 10^{-3}$
$D_{(1,2)}^{(2)}$	0.0081	0.0074	0.0061	$6.24 \cdot 10^{-3}$	$1.73 \cdot 10^{-2}$	0.0067	0.0059	$1.22 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$
$D_{(1,3)}^{(2)}$	0.0079	0.0052	0.0030	$2.33 \cdot 10^{-2}$	$4.14 \cdot 10^{-2}$	0.0068	0.0056	$9.19 \cdot 10^{-3}$	$1.94 \cdot 10^{-2}$
$D_{(2,3)}^{(2)}$	0.0081	0.0050	0.0031	$2.55 \cdot 10^{-2}$	$4.15 \cdot 10^{-2}$	0.0067	0.0053	$1.13 \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$
$D_{(1,4)}^{(2)}$	0.0079	0.0095	0.0039	$1.32 \cdot 10^{-2}$	$3.40 \cdot 10^{-2}$	0.0073	0.0058	$5.13 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$
$D_{(2,4)}^{(2)}$	0.0080	0.0098	0.0052	$1.49 \cdot 10^{-2}$	$2.39 \cdot 10^{-2}$	0.0073	0.0061	$6.16 \cdot 10^{-3}$	$1.63 \cdot 10^{-2}$
$D_{(3,4)}^{(2)}$	0.0081	0.0064	0.0021	$1.42 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	0.0071	0.0058	$8.30 \cdot 10^{-3}$	$1.94 \cdot 10^{-2}$
$D_{(1,2,3)}^{(3)}$	0.0073	0.0053	0.0000	$1.67 \cdot 10^{-2}$	$6.17 \cdot 10^{-2}$	0.0071	0.0038	$2.18 \cdot 10^{-3}$	$2.95 \cdot 10^{-2}$
$D_{(1,2,4)}^{(3)}$	0.0072	0.0104	0.0000	$2.67 \cdot 10^{-2}$	$6.10 \cdot 10^{-2}$	0.0070	0.0044	$1.84 \cdot 10^{-3}$	$2.37 \cdot 10^{-2}$
$D_{(1,3,4)}^{(3)}$	0.0072	0.0069	0.0000	$2.68 \cdot 10^{-3}$	$6.10 \cdot 10^{-2}$	0.0075	0.0042	$2.38 \cdot 10^{-3}$	$2.60 \cdot 10^{-2}$
$D_{(2,3,4)}^{(3)}$	0.0072	0.0073	0.0000	$4.87 \cdot 10^{-4}$	$6.07 \cdot 10^{-2}$	0.0068	0.0041	$3.13 \cdot 10^{-3}$	$2.59 \cdot 10^{-2}$
$D_{(1,2,3,4)}^{(4)}$	0.0068	0.0164	0.0000	$8.15 \cdot 10^{-2}$	$5.69 \cdot 10^{-2}$	0.0110	0.0022	$3.59 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$
$D_1^{(T)}$	0.0612	0.0699	0.0217	$7.26 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0608	0.0399	$4.09 \cdot 10^{-3}$	$1.80 \cdot 10^{-1}$
$D_2^{(T)}$	0.0613	0.0701	0.0216	$7.44 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0602	0.0400	$9.43 \cdot 10^{-3}$	$1.80 \cdot 10^{-1}$
$D_3^{(T)}$	0.0611	0.0585	0.0138	$2.20 \cdot 10^{-2}$	$3.99 \cdot 10^{-1}$	0.0605	0.0387	$4.97 \cdot 10^{-3}$	$1.88 \cdot 10^{-1}$
$D_4^{(T)}$	0.0612	0.0773	0.0182	$1.36 \cdot 10^{-1}$	$3.62 \cdot 10^{-1}$	0.0620	0.0406	$7.33 \cdot 10^{-3}$	$1.74 \cdot 10^{-1}$

TABLE 16 – Estimation des variances pour la QOI $g^{discontinue-4}$ avec la procédure « Vote ».

	$N = 100$					$N = 1000$			
	$D_{i,ref}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$
D	0.1186	0.1168	0.0519	$1.53 \cdot 10^{-2}$	$5.62 \cdot 10^{-1}$	0.1056	0.0849	$1.09 \cdot 10^{-1}$	$2.85 \cdot 10^{-1}$
$D_1^{(1)}$	0.0088	0.0087	0.0087	$7.74 \cdot 10^{-4}$	$5.87 \cdot 10^{-4}$	0.0074	0.0080	$1.20 \cdot 10^{-2}$	$7.04 \cdot 10^{-3}$
$D_2^{(1)}$	0.0088	0.0086	0.0073	$1.82 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$	0.0075	0.0081	$1.05 \cdot 10^{-2}$	$6.00 \cdot 10^{-3}$
$D_3^{(1)}$	0.0089	0.0059	0.0055	$2.46 \cdot 10^{-2}$	$2.82 \cdot 10^{-2}$	0.0074	0.0076	$1.21 \cdot 10^{-2}$	$1.05 \cdot 10^{-2}$
$D_4^{(1)}$	0.0087	0.0107	0.0070	$1.63 \cdot 10^{-2}$	$1.44 \cdot 10^{-2}$	0.0079	0.0080	$7.16 \cdot 10^{-3}$	$6.36 \cdot 10^{-3}$
$D_{(1,2)}^{(2)}$	0.0081	0.0073	0.0061	$6.99 \cdot 10^{-3}$	$1.73 \cdot 10^{-2}$	0.0065	0.0059	$1.32 \cdot 10^{-2}$	$1.89 \cdot 10^{-2}$
$D_{(1,3)}^{(2)}$	0.0079	0.0051	0.0030	$2.39 \cdot 10^{-2}$	$4.14 \cdot 10^{-2}$	0.0068	0.0056	$9.43 \cdot 10^{-3}$	$1.94 \cdot 10^{-2}$
$D_{(2,3)}^{(2)}$	0.0081	0.0050	0.0031	$2.57 \cdot 10^{-2}$	$4.15 \cdot 10^{-2}$	0.0066	0.0053	$1.21 \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$
$D_{(1,4)}^{(2)}$	0.0079	0.0094	0.0039	$1.21 \cdot 10^{-2}$	$3.40 \cdot 10^{-2}$	0.0072	0.0058	$6.41 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$
$D_{(2,4)}^{(2)}$	0.0080	0.0097	0.0052	$1.44 \cdot 10^{-2}$	$2.39 \cdot 10^{-2}$	0.0070	0.0061	$8.06 \cdot 10^{-3}$	$1.63 \cdot 10^{-2}$
$D_{(3,4)}^{(2)}$	0.0081	0.0063	0.0021	$1.48 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	0.0070	0.0058	$9.59 \cdot 10^{-3}$	$1.94 \cdot 10^{-2}$
$D_{(1,2,3)}^{(3)}$	0.0073	0.0050	0.0000	$1.91 \cdot 10^{-2}$	$6.17 \cdot 10^{-2}$	0.0065	0.0038	$6.60 \cdot 10^{-3}$	$2.95 \cdot 10^{-2}$
$D_{(1,2,4)}^{(3)}$	0.0072	0.0098	0.0000	$2.19 \cdot 10^{-2}$	$6.10 \cdot 10^{-2}$	0.0066	0.0044	$5.60 \cdot 10^{-3}$	$2.37 \cdot 10^{-2}$
$D_{(1,3,4)}^{(3)}$	0.0072	0.0065	0.0000	$5.81 \cdot 10^{-3}$	$6.10 \cdot 10^{-2}$	0.0070	0.0042	$1.76 \cdot 10^{-3}$	$2.60 \cdot 10^{-2}$
$D_{(2,3,4)}^{(3)}$	0.0072	0.0069	0.0000	$2.94 \cdot 10^{-3}$	$6.07 \cdot 10^{-2}$	0.0064	0.0041	$6.80 \cdot 10^{-3}$	$2.59 \cdot 10^{-2}$
$D_{(1,2,3,4)}^{(4)}$	0.0068	0.0118	0.0000	$4.29 \cdot 10^{-2}$	$5.69 \cdot 10^{-2}$	0.0077	0.0022	$8.32 \cdot 10^{-3}$	$3.80 \cdot 10^{-2}$
$D_1^{(T)}$	0.0612	0.0637	0.0217	$2.07 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0558	0.0399	$4.63 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$
$D_2^{(T)}$	0.0613	0.0641	0.0216	$2.41 \cdot 10^{-2}$	$3.34 \cdot 10^{-1}$	0.0550	0.0400	$5.32 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$
$D_3^{(T)}$	0.0611	0.0527	0.0138	$7.08 \cdot 10^{-2}$	$3.99 \cdot 10^{-1}$	0.0555	0.0387	$4.69 \cdot 10^{-2}$	$1.88 \cdot 10^{-1}$
$D_4^{(T)}$	0.0612	0.0712	0.0182	$8.41 \cdot 10^{-2}$	$3.62 \cdot 10^{-1}$	0.0568	0.0406	$3.70 \cdot 10^{-2}$	$1.74 \cdot 10^{-1}$

TABLE 17 – Estimation des variances pour la QOI $g^{discontinue-4}$ avec la procédure « Raccordement ».

l'apparition d'une bifurcation est réalisé. La QOI de cette section est la solution à l'instant t du système suivant

$$\dot{X} = -(p_1 + p_2)(X - p_3)(X - p_4) \quad (303)$$

Ce système est semblable à celui de l'application de la section 3.3.3, mais fait intervenir plus de paramètres pour mener une analyse de sensibilité plus intéressante. Le système possède deux équilibres stables (ignorant le cas $p_1 + p_2 = 0$) : $X_e = p_3$ stable si $p_1 + p_2 < 0$, et $X_e = p_4$ stable si $p_1 + p_2 > 0$. De ce fait, le système modélise l'apparition d'une bifurcation. La Table 18 indique la loi de distribution des paramètres d'entrée.

Paramètres d'entrée	Loi de distribution
p_1	$\mathcal{U}[-2, 1]$
p_2	$\mathcal{U}[0, 4]$
p_3	$\mathcal{U}[1, 5]$
p_4	$\mathcal{U}[5, 9]$

TABLE 18 – Loi de distribution des paramètres d'entrée pour la QOI modélisant l'apparition d'une bifurcation.

À un instant t assez élevé, la surface de réponse de la QOI $f^{(bif4-t)}(\mathbf{p}) = X(t, \mathbf{p})$ est irrégulière voire discontinue. L'instant considéré est $t = 20$. La Table 19 présente les hyperparamètres utilisés et la Table 20 présente les estimations des quantités de variances à l'instar des précédentes applications. La procédure de gestion des hyperpavés hybrides sélectionnée est « Raccordement ». Le nombre d'observations d'entraînement utilisé pour calculer la référence est de $N_{MC} = 2000$.

	PC	ME-ACD
N		{100, 500}
$N_{E_{min}}$		1
$N_{E_{max}}$		5
k_{min}		5
k_{max}		10
n_{min}	0	0
n_{max}	10	10
δ_{min}		10^{-3}
Taux d'apprentissage initiaux testés		Cf section 3.3
<i>isConnect</i>		Faux
z		20
N_{frac}		1000

TABLE 19 – Hyperparamètres sélectionnés pour les méthodes PC et ME-ACD pour l'étude de $f^{bif4-20}$.

Pour $N = 100$, le degré optimal du métamodèle PC est de 2 (erreur PRESS normalisée $\delta = 0.75$). Il est de $n = 4$ pour $N = 500$ (erreur PRESS normalisée $\delta = 0.45$). Les degrés

	$N = 100$					$N = 500$			
	$D_{i,ref}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$	$D_{i,ME}^{(j)}$	$D_{i,PC}^{(j)}$	$E_{i,ME}^{(j)}$	$E_{i,PC}^{(j)}$
D	15.27	14.32	12.59	$6.23 \cdot 10^{-2}$	$1.75 \cdot 10^{-1}$	14.79	12.89	$3.17 \cdot 10^{-2}$	$1.56 \cdot 10^{-1}$
$D_1^{(1)}$	2.99	2.72	3.01	$1.80 \cdot 10^{-2}$	$1.01 \cdot 10^{-3}$	2.88	3.00	$7.59 \cdot 10^{-3}$	$7.39 \cdot 10^{-4}$
$D_2^{(1)}$	5.05	4.49	5.04	$3.67 \cdot 10^{-2}$	$5.45 \cdot 10^{-4}$	4.73	4.75	$2.10 \cdot 10^{-2}$	$1.95 \cdot 10^{-2}$
$D_3^{(1)}$	0.04	0.03	0.08	$2.48 \cdot 10^{-4}$	$2.58 \cdot 10^{-3}$	0.05	0.04	$9.84 \cdot 10^{-4}$	$3.08 \cdot 10^{-4}$
$D_4^{(1)}$	0.97	0.95	1.03	$1.29 \cdot 10^{-3}$	$4.33 \cdot 10^{-3}$	0.85	0.86	$7.49 \cdot 10^{-3}$	$7.11 \cdot 10^{-3}$
$D_{(1,2)}^{(2)}$	5.87	5.74	2.86	$8.88 \cdot 10^{-3}$	$1.97 \cdot 10^{-1}$	5.75	3.75	$8.18 \cdot 10^{-3}$	$1.39 \cdot 10^{-1}$
$D_{(1,3)}^{(2)}$	-0.02	0.04	0.31	$4.07 \cdot 10^{-3}$	$2.21 \cdot 10^{-2}$	0.03	0.03	$3.73 \cdot 10^{-3}$	$3.64 \cdot 10^{-3}$
$D_{(2,3)}^{(2)}$	0.06	0.06	0.09	$5.36 \cdot 10^{-4}$	$1.58 \cdot 10^{-3}$	0.09	0.06	$1.87 \cdot 10^{-3}$	$2.85 \cdot 10^{-4}$
$D_{(1,4)}^{(2)}$	0.07	0.04	0.03	$2.02 \cdot 10^{-3}$	$2.26 \cdot 10^{-3}$	0.04	0.08	$1.98 \cdot 10^{-3}$	$8.64 \cdot 10^{-4}$
$D_{(2,4)}^{(2)}$	0.04	0.06	0.11	$1.33 \cdot 10^{-3}$	$4.78 \cdot 10^{-3}$	0.09	0.14	$3.85 \cdot 10^{-3}$	$6.66 \cdot 10^{-3}$
$D_{(3,4)}^{(2)}$	0.00	0.00	0.03	$9.22 \cdot 10^{-6}$	$1.69 \cdot 10^{-3}$	0.00	0.03	$1.98 \cdot 10^{-5}$	$1.85 \cdot 10^{-3}$
$D_{(1,2,3)}^{(3)}$	0.13	0.09	0.00	$2.77 \cdot 10^{-3}$	$8.57 \cdot 10^{-3}$	0.13	0.04	$1.05 \cdot 10^{-4}$	$5.89 \cdot 10^{-3}$
$D_{(1,2,4)}^{(3)}$	0.06	0.09	0.00	$2.20 \cdot 10^{-3}$	$3.73 \cdot 10^{-3}$	0.12	0.04	$4.44 \cdot 10^{-3}$	$1.13 \cdot 10^{-3}$
$D_{(1,3,4)}^{(3)}$	-0.00	0.00	0.00	$3.95 \cdot 10^{-4}$	$3.26 \cdot 10^{-4}$	0.00	0.03	$3.50 \cdot 10^{-4}$	$2.36 \cdot 10^{-3}$
$D_{(2,3,4)}^{(3)}$	0.00	0.00	0.00	$8.22 \cdot 10^{-5}$	$3.45 \cdot 10^{-5}$	0.00	0.03	$6.55 \cdot 10^{-6}$	$2.01 \cdot 10^{-3}$
$D_{(1,2,3,4)}^{(4)}$	0.00	0.02	0.00	$1.34 \cdot 10^{-3}$	$3.82 \cdot 10^{-5}$	0.01	0.00	$7.58 \cdot 10^{-4}$	$4.84 \cdot 10^{-6}$
Total 1	9.44	8.73	6.22	$4.64 \cdot 10^{-2}$	$2.11 \cdot 10^{-1}$	8.96	6.98	$3.14 \cdot 10^{-2}$	$1.61 \cdot 10^{-1}$
Total 2	11.23	10.54	8.10	$4.52 \cdot 10^{-2}$	$2.05 \cdot 10^{-1}$	10.93	8.81	$1.96 \cdot 10^{-2}$	$1.59 \cdot 10^{-1}$
Total 3	0.23	0.24	0.51	$1.06 \cdot 10^{-3}$	$1.83 \cdot 10^{-2}$	0.32	0.27	$6.27 \cdot 10^{-3}$	$2.70 \cdot 10^{-3}$
Total 4	1.11	1.16	1.20	$3.20 \cdot 10^{-3}$	$6.21 \cdot 10^{-3}$	1.12	1.21	$1.05 \cdot 10^{-3}$	$6.65 \cdot 10^{-3}$

TABLE 20 – Estimation des variances pour la QOI f^{bif-4} avec la procédure « Raccordement ».

optimaux des métamodèles locaux sont de 1 dans tous les éléments pour les deux échantillonnages (erreur de partition normalisée $\delta = 3.6 \times 10^{-6}$ pour $N = 100$ et $\delta = 1.0 \times 10^{-3}$ pour $N = 500$)

Note : L'erreur de partition normalisée augmente de $N = 100$ à $N = 500$ à cause d'observations situées sur la quasi-discontinuité, observations non présentes dans le premier échantillonnage à $N = 100$.

Avec un nombre relativement faible d'observations d'entraînement, l'impact de l'interaction (1, 2) est bien estimé à l'aide du métamodèle ME. C'est l'interaction responsable de la bifurcation et donc responsable d'une grande partie de la variance. Le métamodèle PC sous-estime grandement cet impact. Selon le métamodèle PC (avec $N = 500$ et donc $n = 4$), les 3 interactions ayant le plus d'impact sont dans l'ordre l'interaction 2 ($S_2^{(1)} = 37\%$), l'interaction (1, 2) ($S_{(1,2)}^{(2)} = 29\%$) puis l'interaction 1 ($S_1^{(1)} = 23\%$). Il ne réussit pas à classer correctement les interactions par ordre d'impact, ne parvenant pas à approcher correctement la discontinuité. À l'inverse, le métamodèle ME estime correctement l'ordre des interactions ayant le plus d'impact, avec des degrés plus faibles. Néanmoins, l'erreur commise lors des estimations des quantités d'ordre inférieur est plus élevée par rapport aux estimations du métamodèle PC.

4.4 Conclusion

En conclusion, ce chapitre présente une nouvelle approche concernant les indices de Sobol, reposant sur la construction d'une matrice dénotée Ξ et de son vecteur correspondant $\tilde{\Xi}$. Avec cette nouvelle approche, une nouvelle formule analytique reliant les indices de Sobol aux coefficients d'un métamodèle PC a été démontrée et sa cohérence avec la formule classique est soulignée. La démonstration est alors reconduite dans le cas d'un métamodèle ME afin d'obtenir une formule analytique reliant les coefficients des métamodèles locaux aux indices de Sobol. Pour utiliser cette formule avec un métamodèle ME quelconque, une méthode numérique est proposée.

La méthodologie a alors été testée sur des applications présentant des discontinuités. À la lueur de ces applications théoriques, la nouvelle méthodologie est prometteuse. Elle surpasse la méthode classique utilisant un métamodèle PC, méthode étant incapable de bien retranscrire certaines interactions provoquant des discontinuités.

La méthodologie souffre néanmoins de défauts. En premier lieu, le nombre d'hyperpavés peut augmenter de façon significative avec la dimension d et le pas de la grille initiale z . Ce défaut va de pair avec le défaut de la méthode ME-ACD (Cf. chapitre 3), étant moins effective à haute dimension. Il est alors conseillé à l'utilisateur de réduire la dimension du problème avant d'utiliser la méthodologie sur le système réduit. L'évolution du nombre significatif d'hyperpavés est d'autant plus vrai lorsque les frontières entre les éléments ne sont pas alignés avec les axes cartésiens. De ce fait, la méthodologie souffre des mêmes défauts que les méthodes ME classiques ME-gPC [69] et ME-PCM [19].

En second lieu, la méthodologie semble être assez sensible à l'échantillonnage et peut parfois manquer de robustesse pour certaines configurations. L'utilisation d'un échantillonnage adaptatif, comme par exemple celui développé dans [26], pourrait mener à des résultats plus précis. Ces imprécisions sont très probablement issues des erreurs commises lors de la gestion des hyperpavés hybrides. L'utilisation des méthodes ME-gPC [69] et ME-PCM [19] à la place de méthodes ME visant à utiliser un nombre minimal d'éléments serait une piste à explorer afin d'améliorer la précision méthode, ne nécessitant pas de faire appel aux méthodes numériques d'approximations en hyperpavés développées dans ce chapitre. Avec la méthode ME actuelle, un métamodèle très précis sans phénomène de Gibbs peut être obtenu. Néanmoins, une imprécision est commise due au fait que l'espace des paramètres doit être repartitionné en union d'hyperpavés, et qu'une procédure de gestion spéciale pour déterminer les coefficients des métamodèles locaux sur les hyperpavés hybrides doit être réalisée. Avec un métamodèle ME-gPC ou ME-PCM, cette procédure n'aurait pas lieu, les éléments étant déjà des hyperpavés. Les imprécisions de l'approximation du métamodèle dues aux phénomènes de Gibbs confinés dans certains éléments, mènerait alors peut-être finalement à des estimations plus précises.

Dans le chapitre 5, la nouvelle méthodologie est appliquée au système de la dynamique du vol afin d'étudier la sensibilité et la robustesse du comportement des avions aux modèles de représentation dans le domaine post-décroché.

Chapitre 5

Applications à la Dynamique du vol

5.1 Méthodologie

Ce chapitre comporte des applications semblables à celles de l'article de congrès [65], écrit dans le cadre de la thèse.

Dans ce chapitre, une analyse de sensibilité est conduite sur le système dynamique du vol introduit à la section 1.1.5.6. Le but est alors d'étudier la sensibilité et la robustesse du comportement des avions aux modèles de représentation dans le domaine post-décroché afin de répondre à l'objectif initial de la thèse. Le modèle des coefficients aérodynamiques est le modèle sélectionné à la section 1.2.5 reposant sur les données des expériences réalisées sur une maquette d'avion léger dans la soufflerie verticale SV4 du centre Onera Lille. Dans [17], une étude a été menée sur le même modèle en comparant des diagrammes de bifurcation avec et sans ajout de termes enrichissant ce modèle. Dans ce chapitre, une étude similaire est conduite avec une analyse de sensibilité sur l'évolution temporelle des variables d'état, et non pas sur les états asymptotiques.

Pour effectuer l'analyse de sensibilité, un métamodèle multi-éléments (Multi-Element - ME) temporel est entraîné pour chaque variable d'état pour un ensemble discret de N_t instants. Pour cela, l'algorithme ME-ACD temporel (Cf. section 3.5) est utilisé. La méthodologie développée au chapitre 4 est alors appliquée pour chaque variable d'état à chaque instant considéré. Une partition en union d'hyperpavés n'est réalisée qu'à chaque fois qu'une nouvelle partition est nécessaire pour une variable d'état à un instant donné. L'évolution temporelle de l'ensemble des variances partielles peut alors être estimée avec un nombre assez faible d'observations d'entraînement.

Afin de démontrer la potentielle efficacité de la méthode ME par rapport aux méthodes déjà existantes, la méthode Monte-carlo (Monte-Carlo - MC), ainsi que la méthode par chaos polynomiale (Polynomial Chaos - PC) sont également utilisées sur les comportements étudiés. Une critique qui pourrait être faite sur la démarche qui a été adoptée pour ces travaux est que les méthodes déjà existantes n'ont pas été testées sur le système avant le développement de ME-ACD et l'utilisation de la méthode ME sur ce dernier. Selon les

premières simulations réalisées avec le système dynamique du vol, certains comportements présentaient de fortes irrégularités ou des discontinuités, ce qui a motivé l'utilisation de métamodèles ME directement.

Pour effectuer l'analyse de sensibilité, 3 méthodes sont donc utilisées sur des comportements de dynamique du vol afin de comparer leurs performances. Les 3 méthodes sont les méthodes rencontrées aux chapitres précédents : la méthode MC, ainsi que les méthodes PC et ME pour obtenir les variances partielles sur un ensemble fini de N_t instants.

Pour la méthode MC, qui nécessite de nouvelles intégrations du système pour chaque estimation de variance partielle, seules les variances partielles d'ordre 1 et les quantités $D_i^{(T)}$ utilisées pour le calcul des indices de Sobol totaux sont estimées, comme il est réalisé de façon classique dans les applications concrètes. Afin de comparer les 3 méthodes, un nombre d'observations doit être sélectionné. Les métamodèles PC et ME sont entraînés sur les mêmes N observations d'entraînement obtenues par une séquence de Sobol. Pour un nombre d'observations N utilisées pour les méthodes PC et ME, $N_{MC} = N/(d + 2)$ observations sont utilisées pour chaque échantillonnage (séquences de Sobol) de la méthode MC, afin d'utiliser un nombre égal d'observations pour estimer les quantités (les méthodes PC et ME permettant d'obtenir l'ensemble des variances partielles, à l'inverse de la méthode MC). Pour comparer la précision des 3 méthodes, une méthode MC, nommée « référence », est réalisée à l'aide de $N_{ref} = 10000$ observations obtenues également par une séquence de Sobol pour chaque quantité estimée. La convergence des estimations est vérifiée pour toutes les applications. Cette référence est utilisée pour calculer une erreur semblable à celle définie à la section 4.3.1 pour la détermination des quantités pour chaque variable d'état à chaque instant. Les extrema des degrés testés des métamodèles PC sont les mêmes que les extrema des degrés testés pour les métamodèles ME.

L'erreur commise par les métamodèles PC et ME pour estimer les variables d'état elles-mêmes est représentée par l'évolution temporelle de l'erreur PRESS pour le premier et de l'erreur de partition pour le second, toutes deux normalisées par le maximum de la valeur absolue de la variable d'état à l'instant considéré. Le choix se portait sur ces erreurs ou sur des erreurs semblables à celles définie à la section 3.3, représentant l'écart moyen au carré sur des observations tests. L'utilisation d'une telle erreur aurait permis d'utiliser la procédure de précaution de ME-ACD. Comme la procédure de précaution n'entre pas en compte lors de l'utilisation de la formule analytique liant les métamodèles ME aux variances, mais qu'il y a une procédure de raccordement, le choix s'est porté sur ces erreurs de validation croisée internes.

Pour des raisons de clarté des figures, seuls les résultats concernant la vitesse V , l'angle d'incidence α et l'angle de dérapage β sont présentés.

5.2 Applications

5.2.1 Analyse de sensibilité dans le domaine de vol normal

Afin d'effectuer une première application sur le système, une analyse de sensibilité est réalisée dans le domaine de vol normal. L'analyse de sensibilité porte sur l'impact des positions angulaires des surfaces de contrôle. Les paramètres d'entrée sont donc δ_l , δ_m et δ_n . Ces variables de contrôle sont artificiellement modélisées comme des paramètres d'entrée suivant une loi uniforme. Leurs lois de distribution sont indiquées à la Table 21.

Paramètre d'entrée	Loi de distribution
δ_l	$\mathcal{U}(-0.1, 0.1)$
δ_m	$\mathcal{U}(-7, -4)$
δ_n	$\mathcal{U}(-4, 4)$

TABLE 21 – Lois de distribution des positions angulaires des gouvernes dans l'étude de leur impact sur l'évolution des variables d'état dans le domaine normal.

Pour un triplet de valeurs de ces paramètres d'entrée, cela modélise un échelon à $t = 0$ s des contrôles au neutre à ce triplet de valeur. L'avion se dirige alors vers un équilibre stable. Dans ce domaine, les équilibres sont des trajectoires sans danger. Les conditions initiales sélectionnées à $t = 0$ s sont :

- $V_0 = 45 \text{ m.s}^{-1}$
- $\alpha_0 = 0^\circ$
- $\beta_0 = 0^\circ$
- $p_0 = 0 \text{ rad.s}^{-1}$
- $q_0 = 0 \text{ rad.s}^{-1}$
- $r_0 = 0 \text{ rad.s}^{-1}$
- $\Theta_0 = 0^\circ$
- $\Phi_0 = 0^\circ$

La Table 22 indique les hyperparamètres utilisés pour ME-ACD et l'estimation des variances partielles. Les instants considérés sont les $N_t = 101$ instants de $t = 0$ s à $t = 25$ s espacés de façon régulière. Les extrema des degrés testés sont les mêmes pour le méta-modèle PC que pour le métamodèle ME.

La figure 22 représente l'ensemble des évolutions temporelles correspondant aux observations d'entraînement (une observation correspondant à un triplet de paramètres d'entrée). L'évolution temporelle moyenne est tracée en gras. Comme attendu, les comportements convergent vers un équilibre dont la valeur dépend des paramètres. Les différentes évolutions sont assez proches les unes des autres, ce qui présume une évolution régulière des

Hyperparameters	ME-ACD
N	150
N_{Emin}	1
N_{Emax}	5
k_{min}	5
k_{max}	15
n_{min}	0
n_{max}	5
δ_{min}	2.5×10^{-3}
δ_{max}	2.5×10^{-3}
$k_{connect}$	10
$N_{outliers}$	0
taux d'apprentissage initiaux testés	Cf section 3.3
z	25
N_{frac}	100

TABLE 22 – Hyperparamètres sélectionnés pour l'étude de l'impact des positions angulaires des gouvernes dans le domaine de vol normal.

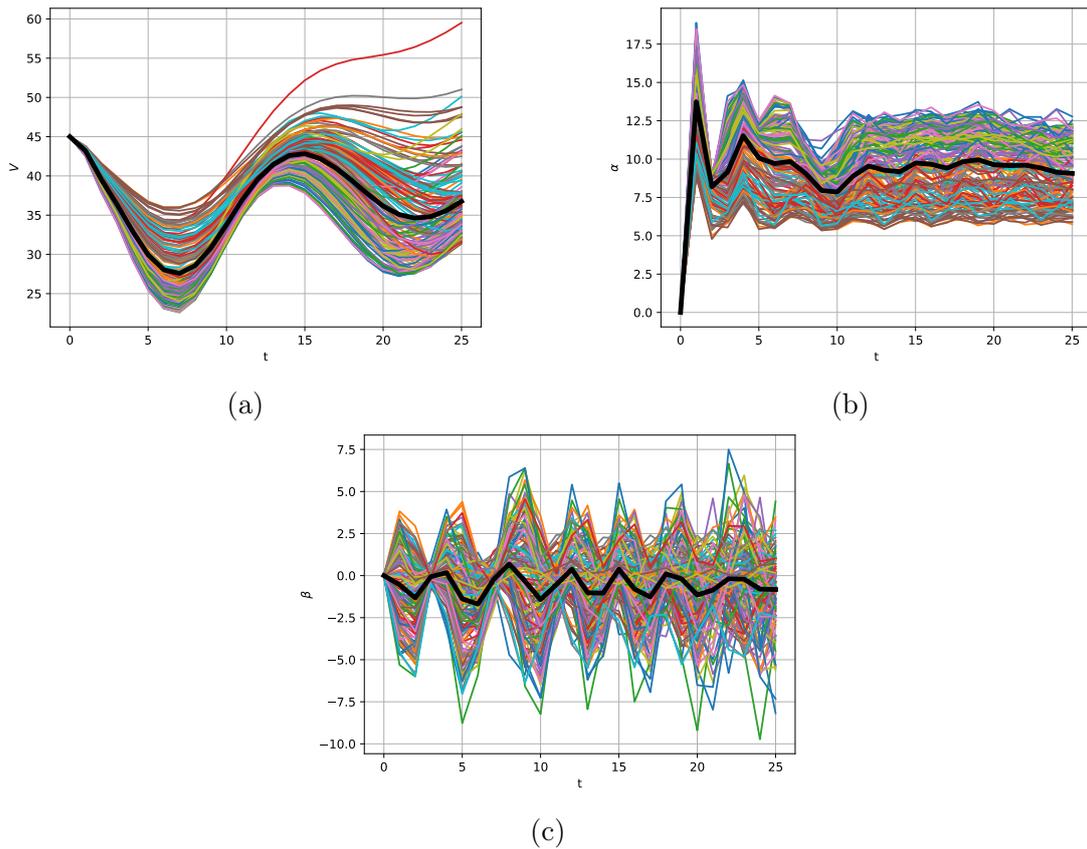


FIGURE 22 – Évolutions temporelles des variables d'état dans l'étude dans le domaine de vol linéaire pour tous les scénari correspondant aux observations d'entraînement et évolution moyenne (tracée en gras) pour (a) V , (b) α et (c) β .

variables d'état par rapport aux paramètres d'entrée pour chaque instant.

La figure 23 présente les résultats de l'analyse de sensibilité réalisée avec le métamodèle ME. Les résultats sont comparés à ceux de références. Les résultats de cette analyse de sensibilité correspondent aux attentes. Tout au long de son évolution temporelle, α dépend en très grande partie de la position angulaire de la gouverne de profondeur δ_m et β dépend en grande majorité de δ_n . La valeur de la vitesse dépend de l'ensemble des contrôles. Les résultats valident l'effectivité des contrôles dans le domaine de vol linéaire. Ils montrent de plus le découplage des effets latéraux (δ_l et δ_n) et longitudinaux (δ_m). Les oscillations observées proviennent du régime transitoire en réponse à l'échelon. Le métamodèle ME donne des résultats proches de la référence. L'évolution temporelle de l'erreur des estimations des quantités de sensibilité par rapport à la référence est représentée sur les figures 24 et 26.

En ce qui concerne la comparaison entre les différentes méthodes, la figure 24 montre l'évolution temporelle de l'erreur de l'estimation des quantités de sensibilité avec les méthodes ME et MC (avec $N_{MC} = 150/5 = 30$). Pour la méthode MC, seules les erreurs concernant les variances partielles d'ordre 1 sont disponibles. Avec un même nombre d'observations d'entraînement, la précision de la méthode ME surpasse celle de la méthode MC dans la grande majorité des cas ou est à minima du même ordre, en plus de pouvoir estimer l'intégralité des quantités.

Le métamodèle PC est lui aussi très performant dans le domaine de vol linéaire. La figure 25 compare les erreurs de validation croisée des métamodèles PC et ME dans l'estimation des variables d'état. Pour certains instants, les deux erreurs sont confondues car, le métamodèle ME ne propose aucune partition de l'espace des paramètres, offrant ainsi un métamodèle identique au métamodèle PC. Pour les instants où le métamodèle ME propose une partition de l'espace, la précision des deux métamodèles sont du même ordre. Ce résultat était prévisible, l'évolution des variables d'état par rapport aux paramètres d'entrée étant assez régulière. La figure 26 représente l'évolution temporelle des erreurs lors de l'estimation des variances, variances partielles et quantités pour calculer les indices de Sobol totaux. Pour ces estimations, les erreurs sont du même ordre. On remarque néanmoins que, même si le métamodèle ME propose à chaque instant une estimation des variables d'état ayant des métamodèles locaux avec des erreurs de validation croisée inférieures ou égales à l'erreur PRESS du métamodèle PC, le gain en précision sur les quantités estimées n'est pas significatif. Dans certains cas, les erreurs dues à l'estimation de la formule (288) à l'aide de la découpe en hyperpavés induisent même une estimation moins précise que l'estimation PC. Ces erreurs sont dues en grande partie à la divergence des métamodèles locaux en dehors de leur élément couplée à la méthode de raccordement.

5.2.2 Influence des paramètres d'amortissement

Dans cette section, une étude de sensibilité est réalisée dans le but d'évaluer l'impact de lacunes du modèle de représentation des coefficients aérodynamiques. Comme précisé à la section 1.3.2.2, avec des modèles de coefficients aérodynamiques ne prenant en compte que des données venant de la balance rotative, des équilibres du domaine post-décroché apparaissent comme instables [17]. Il a été montré que le manque de connaissance de la dépendance des coefficients aérodynamiques en termes de q_a et de r_a , composantes

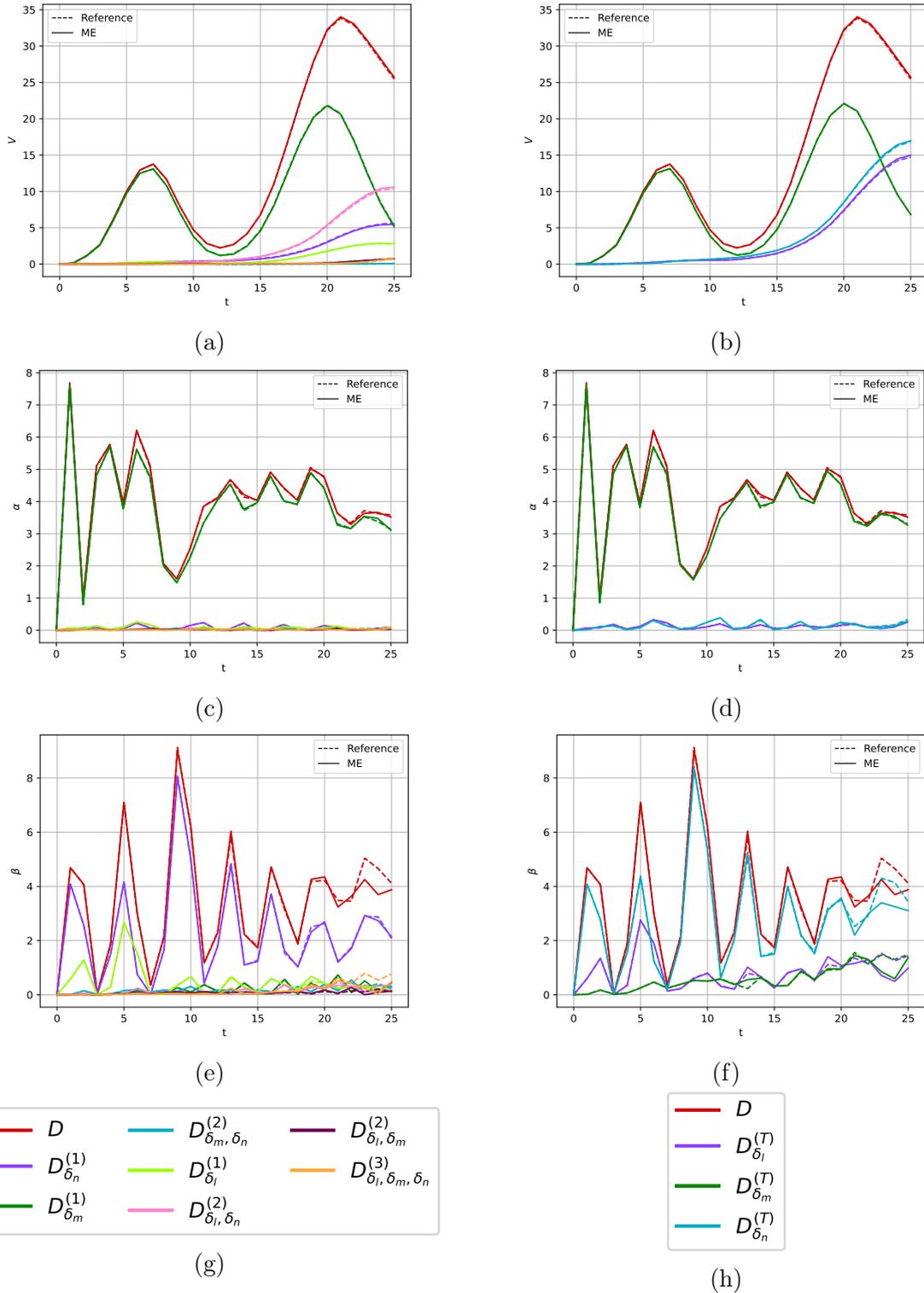


FIGURE 23 – Comparaison de l'évolution temporelle des quantités estimées par la méthode ME et de la référence pour l'étude dans le domaine de vol linéaire. Sur la gauche, évolution des estimations des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.

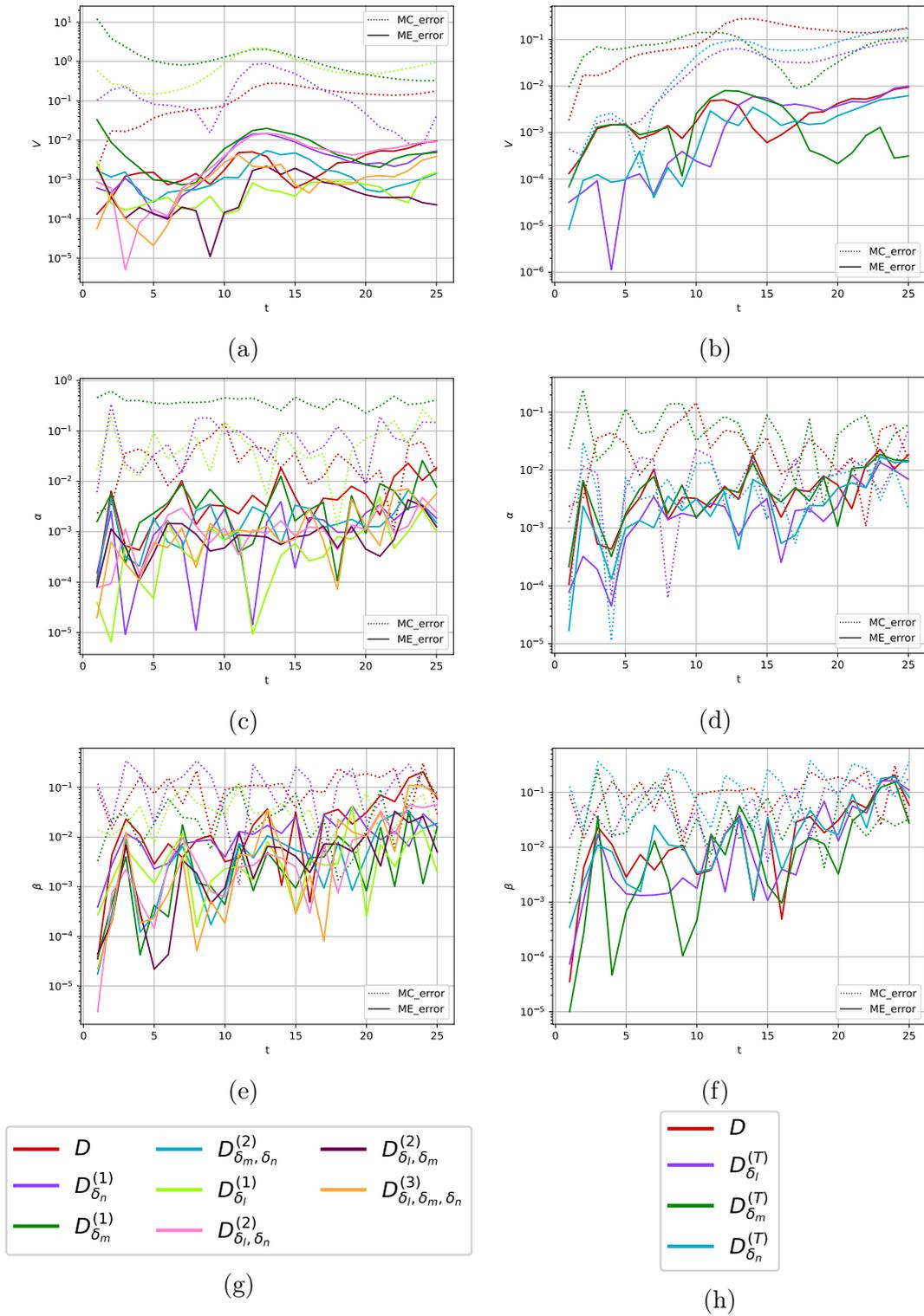


FIGURE 24 – Évolution temporelle des erreurs commises par les méthodes MC et ME dans l'estimation des quantités pour l'étude dans le domaine de vol linéaire. Sur la gauche, évolution des erreurs des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des erreurs des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.

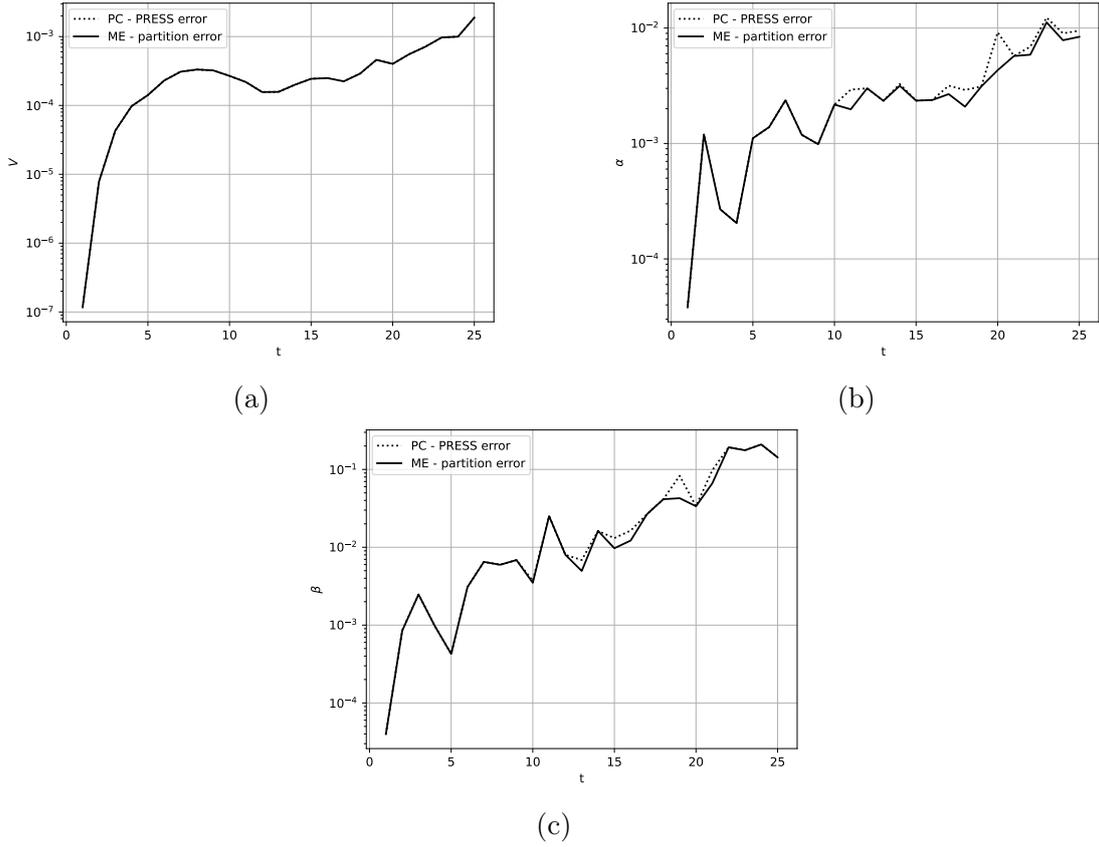


FIGURE 25 – Évolution temporelle des erreurs de validation croisée des métamodèle PC (erreur PRESS) et ME (erreur de partition) pour l'étude du domaine linéaire pour (a) V , (b) α et (c) β .

traduisant le non-alignement entre le vecteur rotation et le vecteur vitesse, peut en être la cause. Ces termes ont un effet d'amortissement qui stabilise certaines vrilles. Sans ces termes, le comportement de l'aéronef devient instable par manque de modélisation de la résistance de l'air avec ces composantes de rotation.

Une méthode proche de la méthode de blending « DIR au premier ordre » (Cf section 1.2.4.5) est alors utilisée sur le modèle des coefficients adopté.

$$C_l^{amort} = C_l + C_l^{\hat{q}_a} \hat{q}_a + C_l^{\hat{r}_a} \hat{r}_a \quad (304)$$

$$C_q^{amort} = C_q + C_m^{\hat{q}_a} \hat{q}_a \quad (305)$$

$$C_n^{amort} = C_n + C_n^{\hat{q}_a} \hat{q}_a + C_n^{\hat{r}_a} \hat{r}_a \quad (306)$$

où C_l , C_m et C_n sont les coefficients de moments adimensionnés obtenus à l'aide du modèle introduit à la section 1.2.5 et rappelé ci-dessous

$$C_l = C_{l0}(\alpha, \beta) + \Delta C_{l,\delta_l}(\alpha, \delta_l) + \Delta C_{l,\delta_n}(\alpha, \delta_n) + \Delta C_{l,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (307)$$

$$C_m = C_{m0}(\alpha, \beta) + \Delta C_{m,\delta_l}(\alpha, \delta_l) + \Delta C_{m,\delta_m}(\alpha, \delta_m) + \Delta C_{m,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (308)$$

$$C_n = C_{n0}(\alpha, \beta) + \Delta C_{n,\delta_l}(\alpha, \delta_l) + \Delta C_{n,\delta_n}(\alpha, \delta_n) + \Delta C_{n,\hat{p}_a}(\alpha, \beta, \hat{p}_a) \quad (309)$$

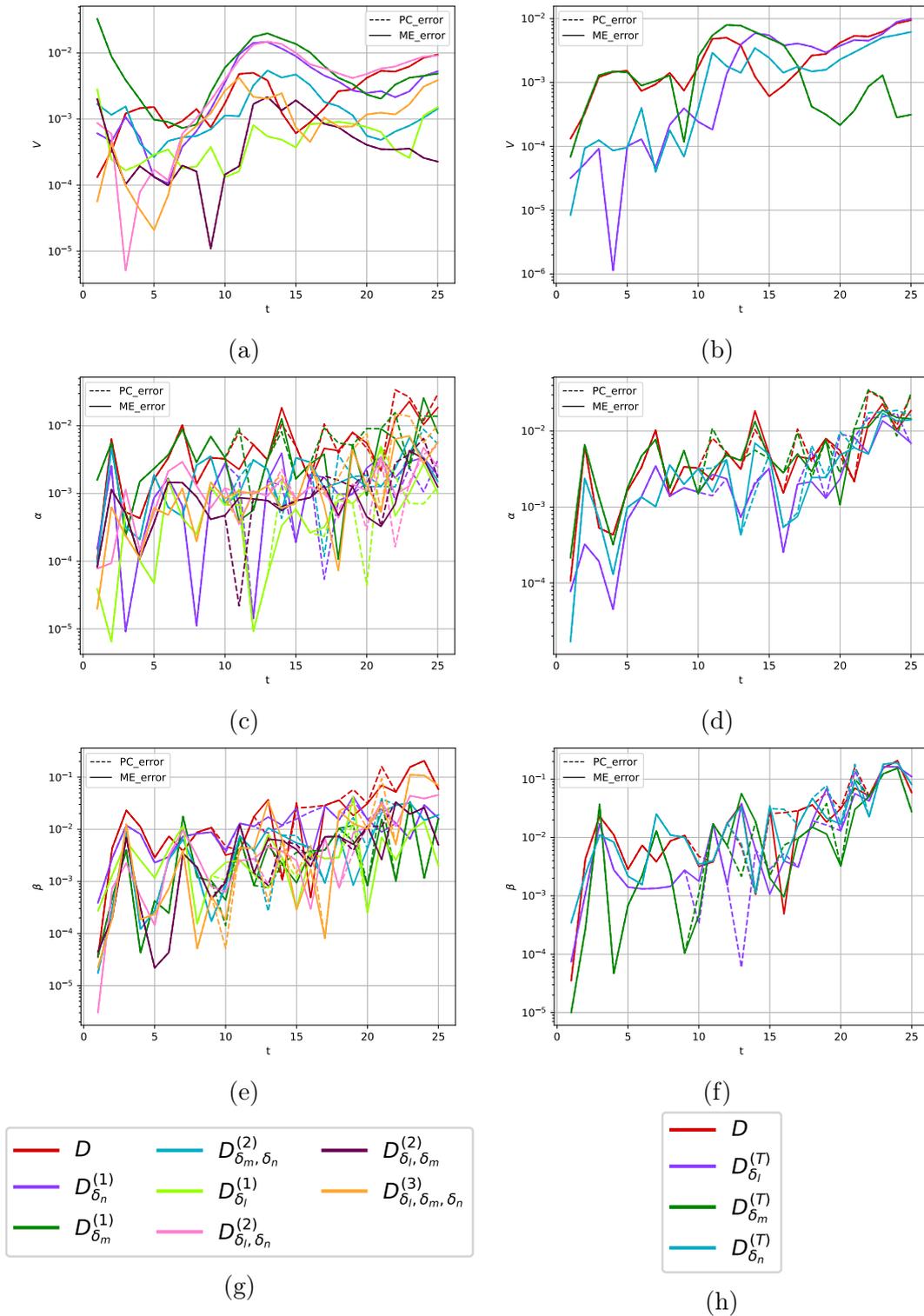


FIGURE 26 – Évolution temporelle des erreurs commises par les méthodes PC et ME dans l'estimation des quantités pour l'étude dans le domaine de vol linéaire. Sur la gauche, évolution des erreurs des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des des erreurs des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.

Dans [17], les coefficients directeurs $C_l^{\hat{q}_a}$, $C_l^{\hat{r}_a}$, $C_m^{\hat{q}_a}$, $C_n^{\hat{q}_a}$ et $C_n^{\hat{r}_a}$ sont obtenus à partir de données en oscillations forcées provenant d'expériences sur une maquette d'avion léger similaire à celui étudié. Un changement de base est ensuite réalisé avec ces données en oscillations forcées (Cf. section 1.2.4.5) pour obtenir les termes d'amortissement. Avec cette technique, les termes sont dépendants de α . Dans l'approche considérée ici, ils sont supposés être indépendants de α et sont modélisés par des variables aléatoires suivant une loi de distribution uniforme. Ce sont alors les paramètres d'entrée de l'étude. La Table 23 indique la loi uniforme que suit chaque paramètre d'entrée. Les intervalles des lois de distribution sont estimés à l'aide des extrema des valeurs des oscillations forcées disponibles dans [17] avec un intervalle plus étendu pour mener une étude plus approfondie de l'impact des termes d'amortissement.

Paramètre d'entrée	Loi de distribution
$C_l^{\hat{q}_a}$	$\mathcal{U}(-0.1, 0.1)$
$C_l^{\hat{r}_a}$	$\mathcal{U}(-0.1, 0.1)$
$C_m^{\hat{q}_a}$	$\mathcal{U}(-10, 0)$
$C_n^{\hat{q}_a}$	$\mathcal{U}(-0.1, 0.1)$
$C_n^{\hat{r}_a}$	$\mathcal{U}(-0.1, 0.1)$

TABLE 23 – Lois de distribution des coefficients directeur des termes d'amortissement ajoutés artificiellement dans le modèle des coefficients aérodynamiques.

De par la différence d'échelle d'évolution entre ces termes, il peut être conjecturé que le paramètre $C_m^{\hat{q}_a}$ a un impact plus important, mais au vu des nombreuses non-linéarités présentes dans le système, cette conjecture doit être vérifiée ou infirmée.

Les conditions initiales sélectionnées sont celles correspondant à une vrille, également étudiée dans [17].

- $V_0 = 27 \text{ m.s}^{-1}$
- $\alpha_0 = 59^\circ$
- $\beta_0 = 5.9^\circ$
- $p_0 = 1.8 \text{ rad.s}^{-1}$, $q_0 = 0.58 \text{ rad.s}^{-1}$, $r_0 = 3.03 \text{ rad.s}^{-1}$, menant à un taux de rotation de $\Omega = 3.6 \text{ rad.s}^{-1} = 0.57 \text{ tr.s}^{-1}$
- $\Theta_0 = -32^\circ$
- $\Phi_0 = 13^\circ$

et les positions angulaires des surfaces de contrôle sont fixées à :

- $\delta_l = -100\%$
- $\delta_m = -20^\circ$
- $\delta_n = -20^\circ$

Hyperparamètres	ME-ACD
N	1001
$N_{E_{min}}$	1
$N_{E_{max}}$	10
k_{min}	30
k_{max}	60
n_{min}	0
n_{max}	5
δ_{min}	2.5×10^{-3}
δ_{max}	2.5×10^{-3}
$k_{connect}$	25
$N_{outliers}$	0
taux d'apprentissage initiaux testés pour le classifieur	Cf section 3.3
z	5
N_{frac}	100

TABLE 24 – Hyperparamètres sélectionnés pour l'étude de l'impact des termes d'amortissement sur une vrille instable.

Cette vrille est instable sans terme d'amortissement ajouté et mène à une divergence après 10 s de simulation. Pour un tel comportement comme la vrille, cette durée où le comportement reste relativement stable est assez importante sachant que les contrôles restent fixes. Avec le modèle lacunaire sans terme d'amortissement, le comportement est prédit comme étant un équilibre instable, et donc difficilement atteignable en vol. Néanmoins, il est possible que le comportement réel soit stable, mais que l'absence de dépendance en q_a et r_a des coefficients aérodynamiques rendent l'équilibre prédit correspondant instable. Le fait que l'équilibre prédit ait ce caractère instable (même si l'instabilité n'apparaît qu'à des instants assez lointains) mène au fait que ce comportement n'a aucun domaine d'attraction. En ajoutant artificiellement des données d'amortissement, le comportement prédit devient stable. Il a alors un domaine d'attraction potentiellement grand. La compréhension de l'impact des termes d'amortissement est alors importante afin d'améliorer les modèles et prédire le danger de ces comportements.

L'application a pour but d'étudier l'influence des termes d'amortissement sur la stabilité de l'équilibre prédit. Si pour un jeu de termes d'amortissement, l'équilibre est stable et pour un autre jeu, l'équilibre est instable, cela va générer de la variance pour la variable d'état à un instant assez avancé. En effet, dans le premier cas, la variable d'état n'évolue pas au cours du temps et dans le second, la variable d'état évolue vers un équilibre stable ou diverge. Les termes d'amortissement générant le plus de variance sont ainsi identifiés comme étant les termes clés pour la stabilité.

Les instants considérés sont les $N_t = 21$ instants entre $t = 0$ s et $t = 20$ s.

La Table 24 indique les hyperparamètres sélectionnés. Sur la figure 27, représentant les évolutions temporelles des variables d'état dans les scénari correspondant aux combinaisons entre paramètres d'entrée, le développement de l'instabilité à partir de 10 s est

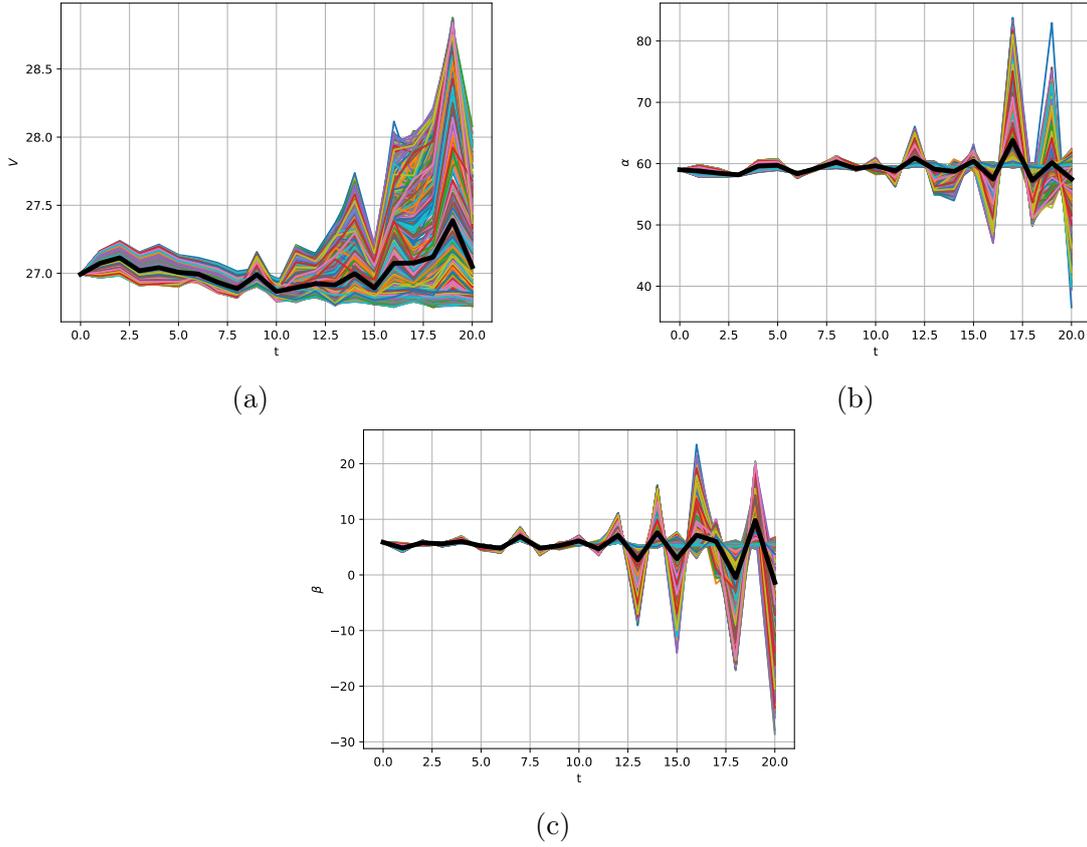


FIGURE 27 – Évolutions temporelles des variables d'état dans l'étude de l'influence des termes d'amortissement pour tous les scénari correspondant aux observations d'entraînement et évolution moyenne (tracée en gras) pour (a) V , (b) α et (c) β .

visible. La variance augmente dramatiquement à partir de cet instant.

La figure 28 présente l'évolution temporelle des quantités estimées pour V , α et β . Les variables d'état ont une haute variance aux derniers instants, due à l'instabilité de l'équilibre avec certaines combinaisons de paramètres d'entrée et à sa stabilité avec d'autres combinaisons de paramètres. Comme il peut être vu sur la figure, $C_m^{\hat{q}_a}$ est le paramètre clé pour la stabilité de cette ville. La conjecture est alors vérifiée. Dans les derniers instants, le paramètre $C_n^{\hat{q}_a}$ impacte également légèrement les variables d'état.

La figure 29a représente l'évolution de la valeur du métamodèle ME approchant $\alpha(t = 17 \text{ s})$ dans sa dépendance aux paramètres d'entrée $C_{m\hat{q}_a}$ et $C_{n\hat{q}_a}$. Sur la figure, les paramètres sont les paramètres transformés par (139) pour appartenir à l'intervalle $[0, 1]$. Sur la figure, les observations d'entraînement, projetées dans l'espace réduit à 3 dimensions, apparaissent en noir. Un échantillonnage test provenant d'une séquence de Sobol de 2000 observations dans l'espace des paramètres de dimension 5 est également effectué. Ces points sont évalués par le métamodèle ME et sont représentés sur la figure par un point d'une certaine couleur indiquant l'élément dans lequel il se trouve. Les différents éléments du métamodèle peuvent être interprétés. L'élément vert représente l'ensemble des valeurs des paramètres menant à une stabilité de l'équilibre. Une faible valeur de $C_{m\hat{q}_a}$ et $C_{n\hat{q}_a}$ mènent donc à la stabilité de la ville. Une irrégularité correspondant à cette frontière

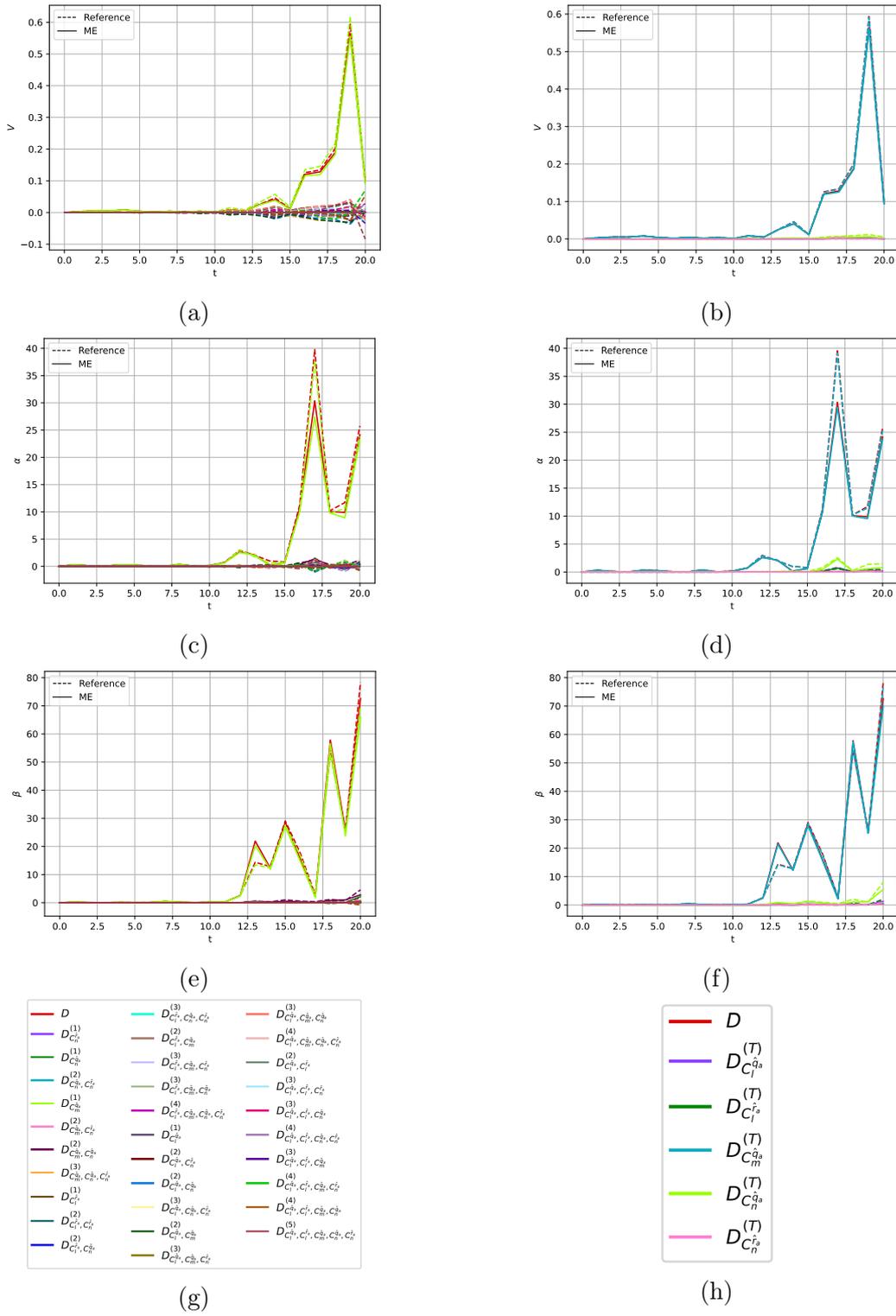


FIGURE 28 – Comparaison de l'évolution temporelle des quantités estimées par la méthode ME et la référence pour l'étude de l'ajout des termes d'amortissement. Sur la gauche, évolution des estimations des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.

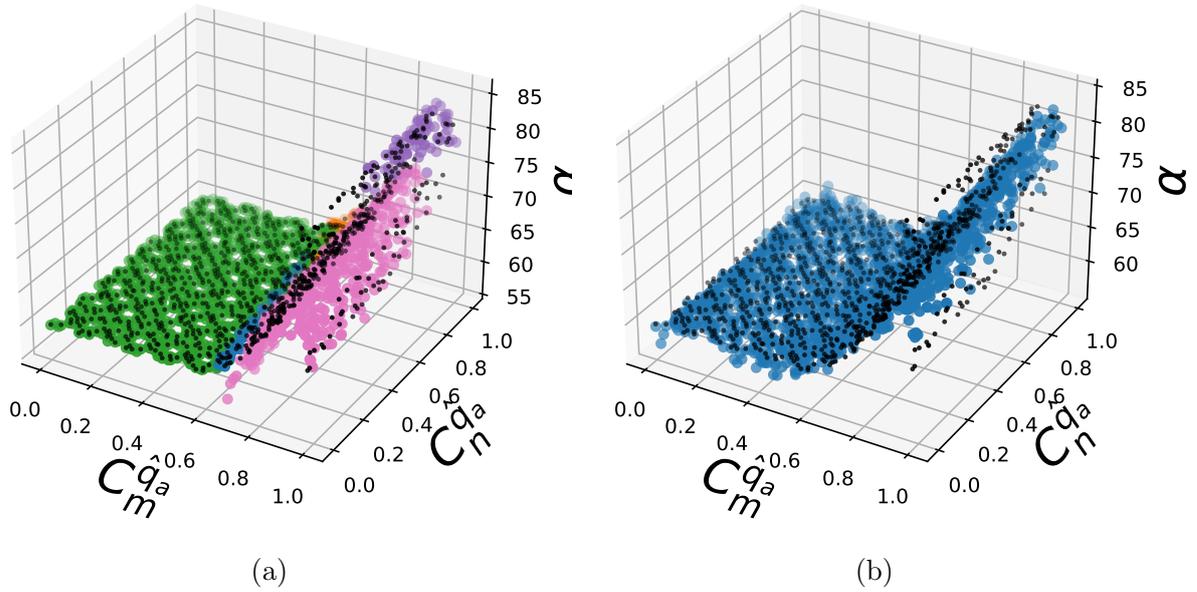


FIGURE 29 – Tracé de la valeur des métamodèles ME et PC approchant $\alpha(t = 17s)$ dans sa dépendance aux paramètres d’entrée $C_{m\hat{q}_a}$ et $C_{n\hat{q}_a}$, les deux paramètres ayant le plus d’impact. Les paramètres sont les paramètres transformés par (139). Les points noirs sont les observations d’entraînement. (a) Approximation du métamodèle ME. Les couleurs autres que noir représentent les différents éléments du métamodèle. (b) Approximation du métamodèle PC. Les évaluations tests apparaissent en bleues.

entre stabilité et instabilité de l’équilibre est observée, dont l’équation dépend des valeurs de $C_{m\hat{q}_a}$ et de $C_{n\hat{q}_a}$. Pour une valeur de $C_{m\hat{q}_a}$ inférieure ou égale à environ -4 (inférieure à $0,6$ pour le paramètre transformé illustré sur la figure), α garde une valeur proche de sa valeur initiale de 59° : la vrille est stable. L’utilisation d’un métamodèle ME permet alors une interprétation des différents éléments. Les différents classifieurs, propres à chaque variables d’état sur un intervalle de temps, permettent ensuite de déterminer rapidement dans quel élément se trouve le point. Dans le cas de la présente étude, cela permet d’avoir une estimation des combinaisons de paramètres rendant la vrille stable. On peut alors remarquer que la partition de ME-ACD crée, à cet instant, un élément rose avec des observations situées sur la pente ascendante et sur la pente descendante de l’évolution de α . Une partition séparant ces observations en deux éléments distincts aurait pu mener à un métamodèle possiblement plus précis. Comme indiqué à la section 3.4, ME-ACD manque de robustesse dans sa partition.

La figure 30 montre l’évolution temporelle des 8 variables d’état et de l’azimut Ψ dans le cas du modèle sans ajout de termes d’amortissement et dans le cas où les termes $C_{m\hat{q}_a} = -10$ et $C_{n\hat{q}_a} = -0,1$ ont été ajoutés. Le changement de stabilité de l’équilibre est souligné par la figure. Le fait que les courbes oranges, correspondant au cas amorti, ne représentent pas des fonctions constantes mais soient des courbes convergeant vers un équilibre est dû au fait que les valeurs de l’équilibre sont légèrement modifiées avec l’ajout de termes d’amortissement.

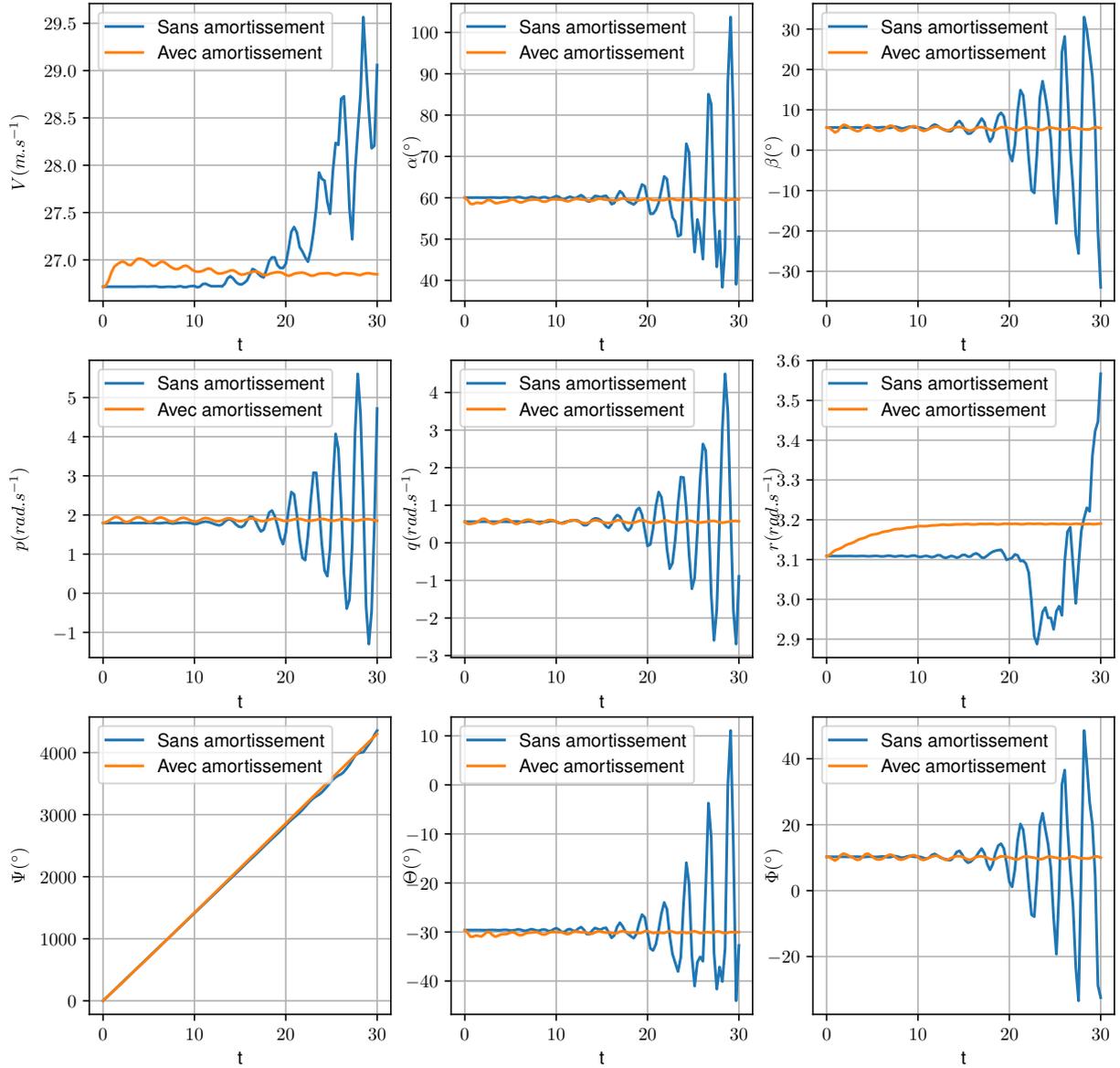


FIGURE 30 – Évolution des 8 variables d'état et de l'azimut Ψ dans le cas du modèle sans ajout de termes d'amortissement et dans le cas où les termes $C_{m\hat{q}_a} = -10$ et $C_{n\hat{q}_a} = -0.1$ ont été ajoutés.

En ce qui concerne la comparaison entre les différentes méthodes, à l'instar de la section précédente, la figure 31 compare les précisions des méthodes MC (avec $N_{MC} = N/7 = 143$) et ME pour estimer les quantités de sensibilité. La figure 32 compare les précisions des métamodèles ME et PC dans l'estimation des variables d'état et la figure 33 compare les précisions des méthodes PC et ME pour estimer les quantités de sensibilité. Contrairement aux premières conclusions prises à partir des toutes premières simulations des travaux, la bifurcation ne provoque pas de discontinuité. Si les irrégularités engendrées affectent la précision du métamodèle PC aux derniers instants, cette perte de précision n'est pas aussi importante qu'il avait été conjecturé. La figure 29b illustre la présence de légères oscillations, ainsi que de certains écarts entre le métamodèle et les observations d'entraînement proche de l'irrégularité. Comme le montre la figure 32, l'utilisation du métamodèle ME ne permet pas d'obtenir des métamodèles beaucoup plus précis au cours

du temps ; les erreurs sont de même ordre.

D'après la figure 31, les estimations menées avec les métamodèles ME sont dans l'ensemble plus précises que les estimations MC classiques pour les variances partielle. On dénombre néanmoins des instants où la méthode ME fournit des estimations particulièrement mauvaises, comme à $t = 17s$ en α ou $t = 13s$ en β , où les quantités sont sous-évaluées. Ces pics d'erreur ne sont pas présents dans les estimations de la méthodes PC, comme le montre la figure 33. Néanmoins, à ces instants, le métamodèle ME en lui-même plus précis que le métamodèle PC sur les observations d'entraînement. Comme déjà observé à l'application précédente, la méthode ME d'estimation des quantités de sensibilité induit des imprécisions. Ces imprécisions, couplées au manque de robustesse de l'algorithme ME-ACD, rendent la méthode d'estimation ME utilisée peu robuste.

En conclusion, l'analyse de sensibilité permet d'identifier quels sont les termes d'amortissement ayant le plus d'impact sur la vrille. Une réduction du nombre de paramètres peut alors être envisagée pour une étude plus poussée de l'impact des termes d'amortissement sur cette vrille. Il s'agit alors d'un pas supplémentaire vers la compréhension des phénomènes complexes agissant lors d'une vrille à haut angle d'incidence comme celle-ci. Une piste de recherche serait alors de remonter aux phénomènes physiques que représentent les termes d'amortissement impactant. Une fois ces termes d'amortissement clés ajoutées, plus d'équilibres stables, prédisant un comportement potentiellement dangereux, pourront être détectés et étudiés.

Comme discuté avec des collègues chercheurs, l'impact des termes d'amortissement sur cet équilibre pourrait également être étudié à l'aide de la notion de stabilité à temps court. D'une autre conversation a émergé l'idée d'utiliser l'analyse de fiabilité pour traiter ce genre de problèmes, au lieu d'utiliser la variance des variables d'état et les éléments du métamodèle comme métrique de robustesse. Ces deux pistes sont des axes de recherche à envisager pour la suite des études sur le sujet.

5.2.3 Sensibilité et robustesse du comportement des aéronefs au modèle de représentation

Une analyse de sensibilité et de robustesse de comportements stables par rapport aux incertitudes du modèle aérodynamique peut être effectuée. Comme le modèle aérodynamique repose sur des données expérimentales, des incertitudes affectent la précision du modèle. Ces incertitudes peuvent être modélisées par des termes ΔC_l , ΔC_m et ΔC_n , ajoutés aux trois coefficients de moments. Ces termes sont les paramètres d'entrée de l'étude, suivant une certaine loi de distribution. Les variances des variables d'état donnent alors des informations sur la robustesse du comportement par rapport à ces incertitudes, une faible variance correspondant à un comportement robuste. Les variances partielles donnent quant à elles une information sur la sensibilité de chaque variable d'état par rapport à chaque interaction entre les incertitudes.

La vrille étudiée est la même que celle de la section précédente (section 5.2.2). Ainsi, les conditions initiales de l'étude sont identiques. Les instants considérés sont également les

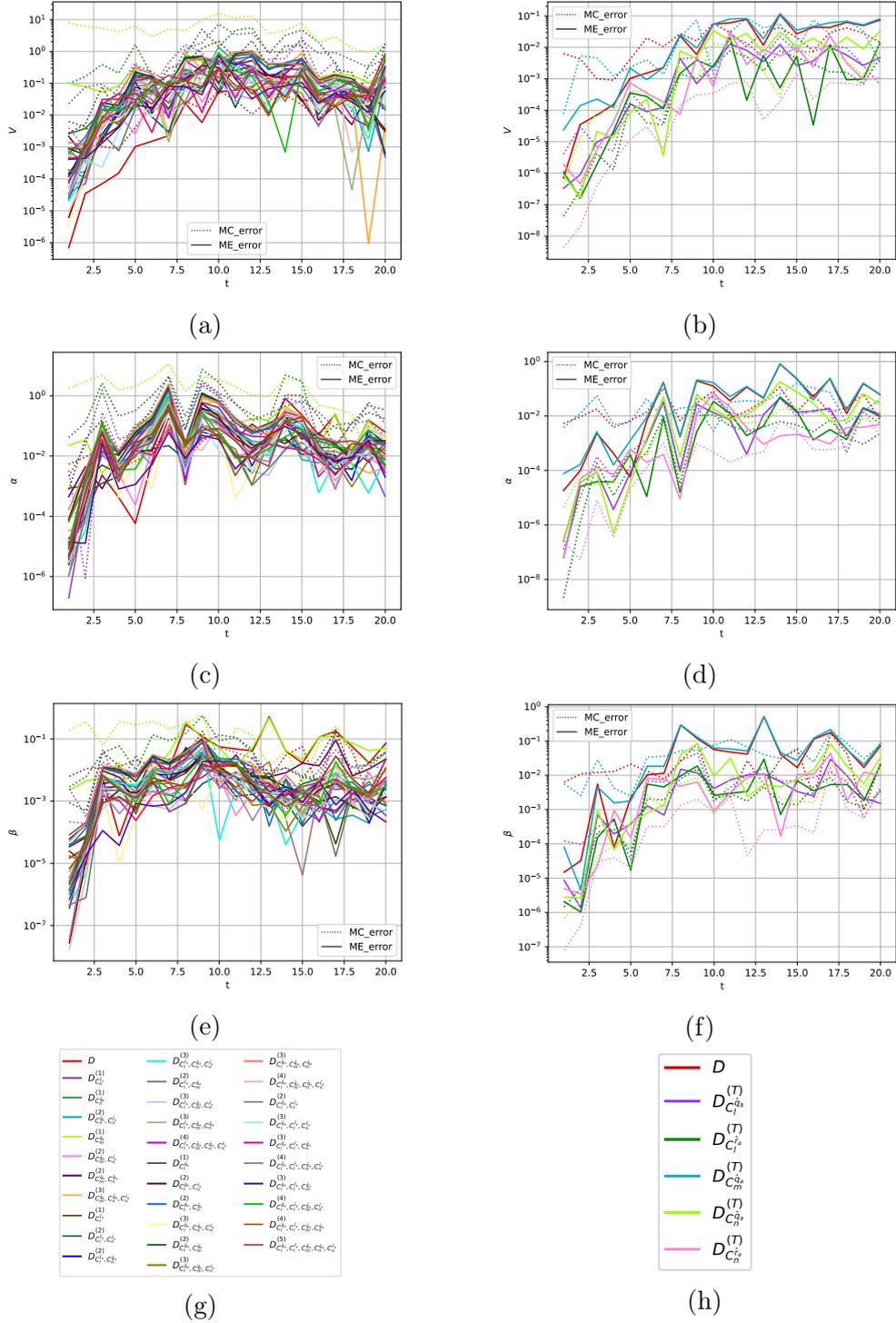


FIGURE 31 – Évolution temporelle des erreurs commises par les méthodes MC et ME dans l'estimation des quantités pour l'étude de l'ajout des termes d'amortissement. Sur la gauche, évolution des erreurs des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des erreurs des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.

mêmes. Des termes d'amortissement $C_{m\hat{q}_a} = -10$ et $C_{n\hat{q}_a} = -0.1$ sont ajoutés afin de rendre la ville stable, d'après les résultats précédents.

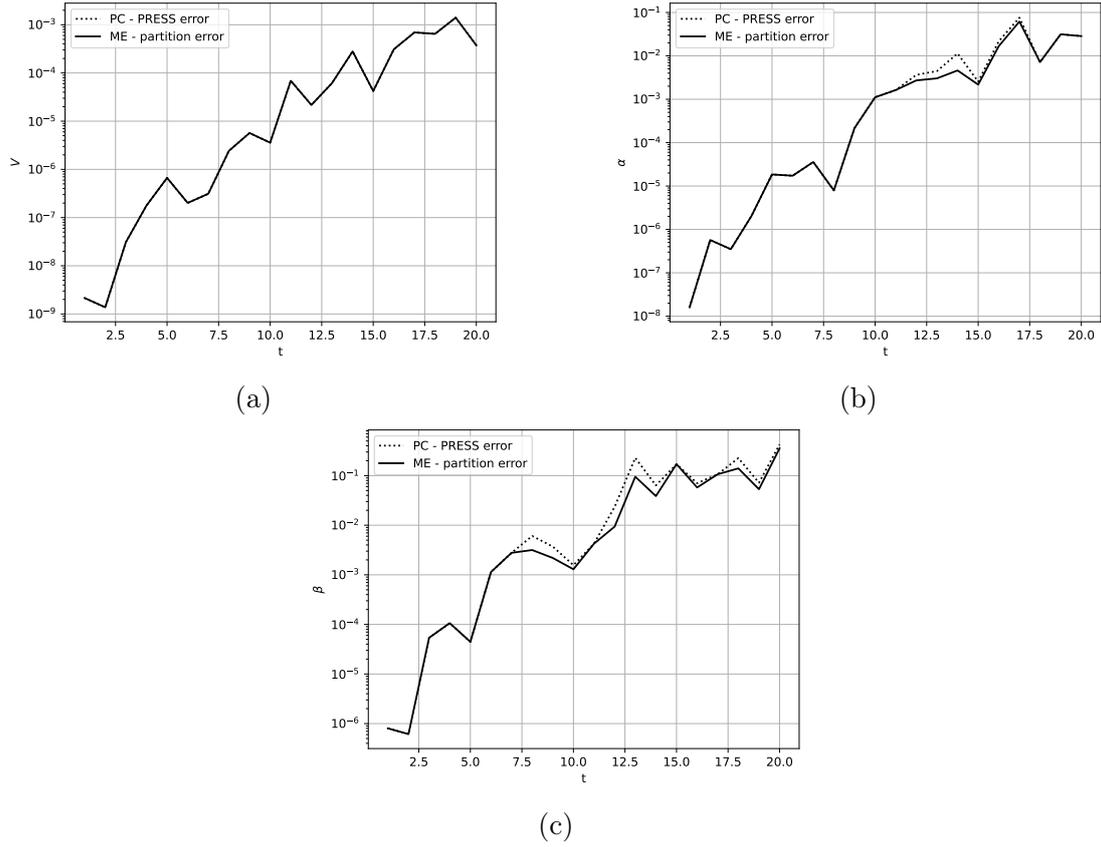


FIGURE 32 – Évolution temporelle des erreurs de validation croisée des métamodèle PC (erreur PRESS) et ME (erreur de partition) pour l'étude de l'ajout des termes d'amortissement pour (a) V , (b) α et (c) β .

Les paramètres d'entrée sélectionnés sont donc ΔC_l , ΔC_m et ΔC_n . Dans cette première étude, ces paramètres sont ajoutés de façon artificielle avec une loi de distribution choisie arbitrairement. Ils suivent tous trois une loi normale dont les caractéristiques sont données à la Table 25. Les écart-types des lois normales sont estimés comme étant égaux au tiers de 10% de la valeur absolue du coefficient de moment adimensionné à $t = 0$ s. Une étude de quantification d'incertitude plus poussée mériterait d'être menée afin de déterminer une loi de distribution proche de celle rencontrée lors des expériences.

Paramètre d'entrée	Loi de distribution
ΔC_l	$\mathcal{N}(0, 0.01)$
ΔC_m	$\mathcal{N}(0, 0.0333)$
ΔC_n	$\mathcal{N}(0, 0.00166)$

TABLE 25 – Lois de distribution des termes modélisant les incertitudes expérimentales ajoutés artificiellement dans le modèle des coefficients aérodynamiques.

Les hyperparamètres sélectionnés sont ceux indiqués dans la Table 26.

Les figures 34 et 35 soulignent que le comportement est robuste aux incertitudes, au vu des faibles variances. En effet, comme le montre la figure 35, la variance de la vitesse

Hyperparamètres	ME-ACD
N	1001
N_{Emin}	1
N_{Emax}	5
k_{min}	5
k_{max}	10
n_{min}	0
n_{max}	5
δ_{min}	2.5×10^{-3}
δ_{max}	2.5×10^{-3}
$k_{connect}$	10
$N_{outliers}$	0
taux d'apprentissage initiaux testés pour le classifieur	Cf section 3.3
z	25
N_{frac}	100

TABLE 26 – Hyperparamètres sélectionnés pour l'étude de l'impact des incertitudes des coefficients de moment.

ne dépasse pas $0,01 (m/s)^2$, et celle des angle d'incidence et de dérapage ne dépasse pas $0,35deg^2$. L'évolution des variables d'état par rapport aux paramètres d'entrée est régulière et le métamodèle ME est en réalité un métamodèle PC pour toutes les variables d'état à tous les instants. La précision de ces métamodèles dans l'approche des variables d'état est représentée à la figure 36. Au vu des faibles variances rencontrées, une analyse de sensibilité n'est pas de rigueur.

5.3 Conclusion

Dans ce chapitre, des analyses de sensibilité sont appliquées au système de dynamique du vol, dans le but de répondre à la problématique initiale de la thèse portant sur la sensibilité et la robustesse des comportements aux modèles de représentation des coefficients aérodynamiques. En ce sens, les études menées dans [17] sont reconduites sous l'angle de l'analyse de sensibilité globale. D'après les connaissances de l'auteur, aucune étude n'a effectué d'analyse de sensibilité dans ce but par le passé.

La méthodologie développée aux chapitres 3 et 4 est appliquée pour effectuer l'analyse de sensibilité. A priori, cette méthodologie devait être plus efficace que la méthode PC, en raison des bifurcations induisant des irrégularités et/ou des discontinuités dans les évolutions des variables d'état par rapport aux paramètres d'entrée. Néanmoins, comme l'a montré les différentes applications, les métamodèles PC suffisent pour conduire une analyse intéressante. En effet, les bifurcations n'engendrent pas de discontinuités, ni d'irrégularités assez pentues pour entacher la précision des métamodèles PC de façon significative. Si la méthode ME produit des métamodèles plus précis, les erreurs d'estimation soulevées à la section 4.4 induisent que les estimations des quantités utiles aux calculs des indices de Sobol et indices de Sobol totaux sont moins précises que l'estimation PC. L'utilisation

d'un métamodèle permet d'avoir accès à plus de quantités et de façon plus précise que les estimations MC classiques pour un même nombre d'observations d'entraînement. Un avantage propre aux métamodèles ME est que les éléments de sa partition peuvent être interprétés, ce qui est réalisé dans le cas de l'étude des termes d'amortissement.

Néanmoins, même si l'intérêt de l'utilisation de la méthode ME par rapport la méthode PC n'est pas validée sur les comportements étudiés, l'analyse de sensibilité réalisée dans le chapitre confirme les résultats des travaux présentés dans [17] et offre de nouvelles informations. La présente étude a permis de mettre en évidence les termes d'amortissement ayant le plus d'impact pour un exemple de vrille. La sensibilité et la robustesse de comportements face à des incertitudes de mesures ont également pu être étudiées à l'aide des estimations des variances partielles et de la variance totale.

La présente étude a pour but d'évaluer les capacités d'une méthode innovante pour étudier la sensibilité et la robustesse d'un comportement de l'aéronef par rapport à la modélisation des coefficients aérodynamiques. Des premières applications ont été réalisées, et si l'utilisation des métamodèles ME n'a pas été concluante pour ces premiers comportements en vol, l'utilisation de la méthode sur d'autres comportements prédits par le modèle, ainsi que par d'autres modèles de coefficient pourrait mener à d'autres conclusions. Ces tests pourront être le sujet de futures recherches.

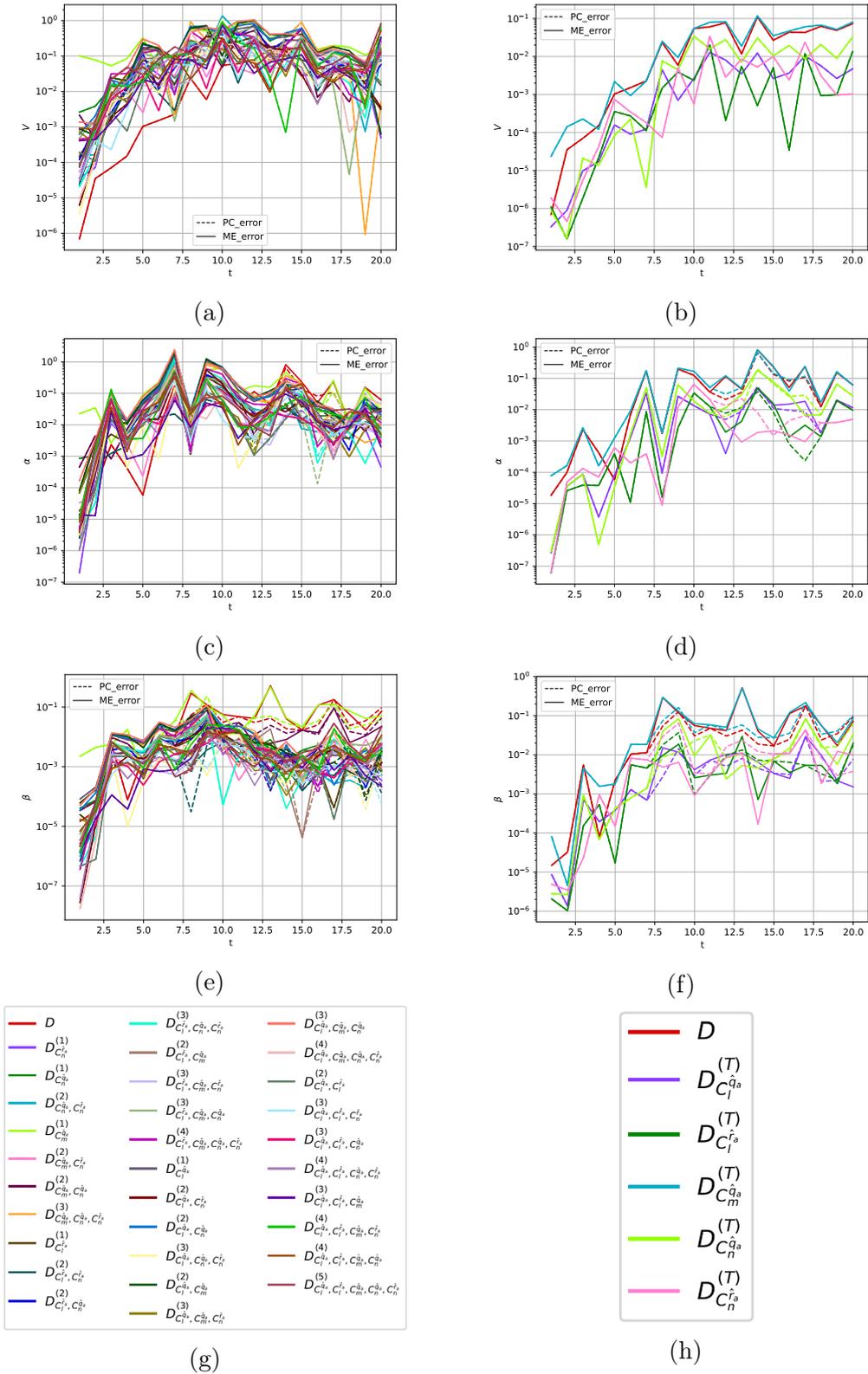
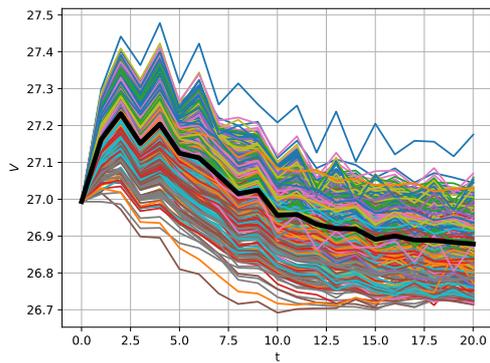
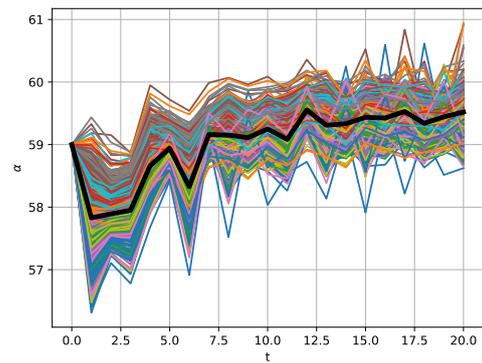


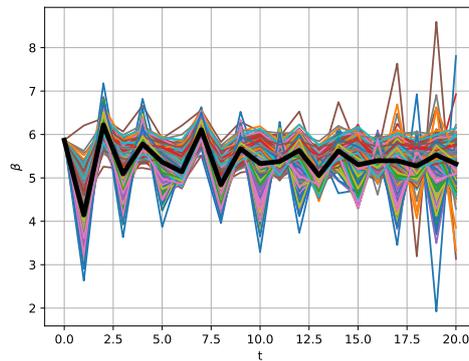
FIGURE 33 – Évolution temporelle des erreurs commises par les méthodes PC et ME dans l'estimation des quantités pour l'étude de l'ajout des termes d'amortissement. Sur la gauche, évolution des erreurs des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des erreurs des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.



(a)

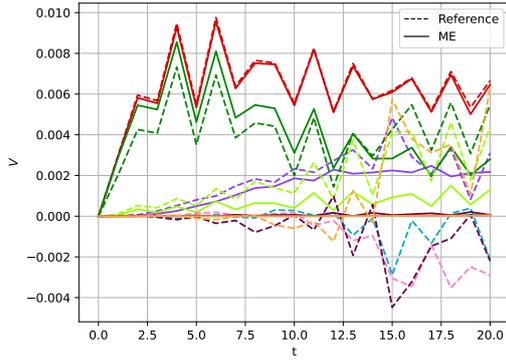


(b)

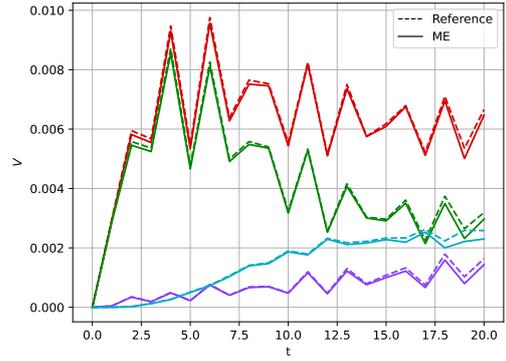


(c)

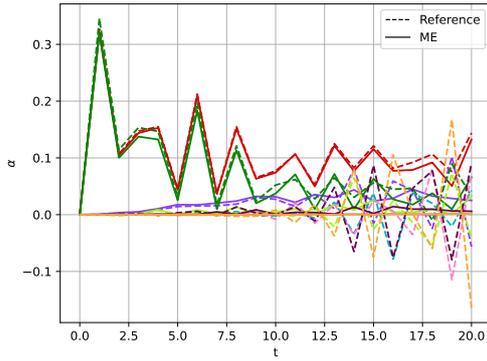
FIGURE 34 – Évolutions temporelles des variables d'état dans l'étude de l'influence des incertitudes des coefficients de moments pour tous les scénari correspondant aux observations d'entraînement et évolution moyenne (tracée en gras) pour (a) V , (b) α et (c) β .



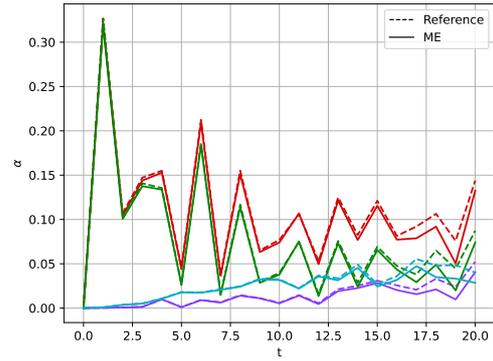
(a)



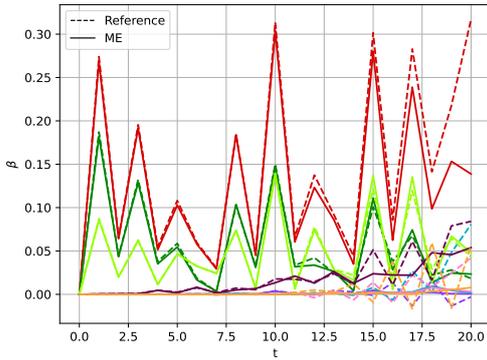
(b)



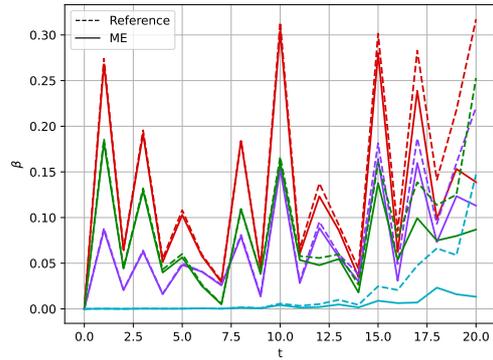
(c)



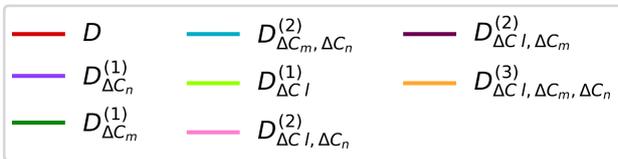
(d)



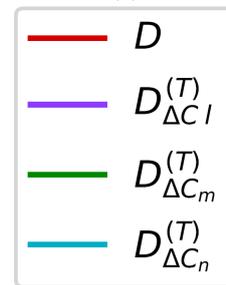
(e)



(f)

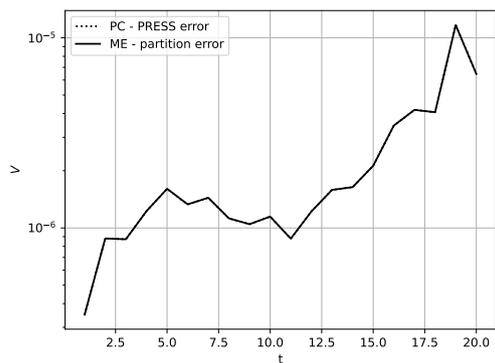


(g)



(h)

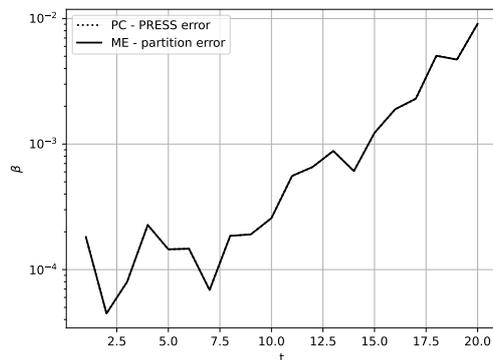
FIGURE 35 – Comparaison de l'évolution temporelle des quantités estimées par la méthode ME et la référence pour l'étude de l'influence des incertitudes des coefficients de moment. Sur la gauche, évolution des estimations des variances partielles d'ordre 1 à 3 pour (a) V , (c) α et (e) β . (g) Légende pour les variances partielles. Sur la droite, évolution des quantités utilisées dans le calcul des indices de Sobol totaux pour (b) V , (d) α et (f) β . (h) Légende pour ces quantités.



(a)



(b)



(c)

FIGURE 36 – Évolution temporelle des erreurs de validation croisée des métamodèle PC (erreur PRESS) et ME (erreur de partition) pour l'étude de l'influence des incertitudes de coefficients de moment pour (a) V , (b) α et (c) β .

Conclusion

Conclusion

Pour répondre à l'objectif initial de la thèse, portant sur l'étude de la dépendance des prévisions de vol à la modélisation des coefficients aérodynamiques, l'angle de vue de l'analyse de sensibilité globale (Global Sensitivity Analysis - GSA) est envisagé. L'évolution de la variance des variables d'état est considérée comme image de la robustesse et l'évolution des variances partielles (ou des indices de Sobol) donne l'évolution temporelle de la sensibilité de chaque variables d'état par rapport aux paramètres étudiés. Pour éviter d'utiliser la méthode de Monte-Carlo (Monte-Carlo - MC) classique pour obtenir ces valeurs, nécessitant de très nombreuses simulations, la construction d'un métamodèle est préférée. Les métamodèles par chaos polynomial (Polynomial Chaos - PC) permettent d'obtenir la variance et les variances partielles à l'aide de formules analytiques. Néanmoins, comme le système dynamique du vol contient de nombreuses bifurcations, il a été conjecturé qu'il y aurait de nombreuses irrégularités et/ou discontinuités, altérant la précision des métamodèles PC. Les premières simulations au début de l'étude semblaient aller dans ce sens. La solution envisagée a alors été d'utiliser un métamodèle multi-éléments (Multi-Element - ME).

La thèse s'est alors orientée vers l'élaboration de deux méthodes dans le but d'améliorer les techniques concernant les métamodèles ME :

En premier lieu, un algorithme permettant d'obtenir un métamodèle ME, ME-ACD, est développé. Son développement a donné lieu à l'écriture de [66]. Cette méthode repose sur l'utilisation de l'algorithme des plus proches voisins, de la régression, du regroupement hiérarchique ascendant et de la classification par réseau de neurones. Cette méthode tend alors à créer un nombre minimal d'éléments, épousant la forme des discontinuités et des irrégularités. La méthode a été testée sur des quantités d'intérêt (Quantity of Interest - QOI) de faible dimension ou de dimension modérée présentant une ou plusieurs discontinuités ou irrégularités. La méthode s'est avérée prometteuse, étant capable d'approcher des QOI possédant une grande variété d'irrégularités et de discontinuités. La méthode surpasse la méthode PC classique, souffrant de pertes de précision pour les QOI irrégulières. Comparée aux méthodes ME déjà existantes, ME-ACD est capable de gérer les irrégularités et/ou discontinuités ne parcourant que partiellement l'espace des paramètres et n'a pas besoin d'échantillonnage spécifique. Néanmoins, la méthode n'est valable que pour un nombre de paramètres modéré et souffre d'un manque de robustesse par rapport

à l'échantillonnage.

Ensuite, la formule liant les coefficients d'un métamodèle PC aux variances partielles est étendue au cas ME. Un article [67] reprenant les résultats du Chapitre 4 est en cours d'écriture à ce sujet. La formule s'appuie sur une nouvelle approche concernant le calcul des variances partielles. La validité de la nouvelle approche lorsque la QOI est un métamodèle PC est mise en lumière. Une estimation des variances partielles lorsque la QOI est approchée par un métamodèle ME est alors proposée. La formule s'accompagne d'une procédure numérique pour réaliser cette estimation. La formule développée fonctionne alors avec un métamodèle ME quelconque. Pour l'utilisation de cette formule, il est néanmoins nécessaire que l'espace des paramètres soit partitionné en une union d'hyperpavés, afin de pouvoir utiliser les propriétés orthogonales des polynômes des métamodèles locaux. Le choix a alors été fait, dans ce manuscrit, d'utiliser un métamodèle ME avec des éléments de forme quelconque (obtenu à l'aide de la méthode ME-ACD), afin d'obtenir un métamodèle précis avec un nombre minimal d'éléments, puis d'utiliser la procédure de classification du métamodèle pour effectuer une nouvelle partition de l'espace en union d'hyperpavés. La précision de l'estimation des variances partielles est évaluée sur un ensemble de QOI théoriques. La nouvelle procédure s'est avérée donner des estimations plus précises que les métamodèles PC standards, incapables de retranscrire certaines interactions menant à des discontinuités dans l'espace des paramètres. Néanmoins, la nouvelle procédure souffre de défauts. En effet, la méthode souffre d'un manque de robustesse par rapport à l'échantillonnage, défaut allant de pair avec la méthode ME-ACD sur laquelle elle repose. De plus, le nombre d'hyperpavés augmentent de façon importante avec la dimension, à l'instar des méthodes ME-gPC et ME-PCM classiques. La procédure de découpe en hyperpavés à partir du classifieur par réseau de neurones mène également à des erreurs de précision.

Le nouveau procédé reposant sur les deux méthodes développées a été ensuite appliqué au système dynamique du vol dans le but d'étudier la sensibilité et la robustesse des comportements aux modèles de représentation dans le domaine de vol post-décroché. Le nouvel angle de vue de la GSA a permis d'identifier les termes d'amortissement (ajoutés artificiellement) ayant le plus d'impact sur la stabilité d'une vrille prise en exemple, déjà étudiée dans [17] dans le cadre de la théorie des bifurcations. Cela a également permis d'identifier les combinaisons entre ces termes d'amortissement menant à la stabilité de la vrille. De plus, la sensibilité et la robustesse du comportement par rapport aux incertitudes commises sur les coefficients aérodynamiques peuvent être étudiées.

Discussions et perspectives

Comme il a été mentionné lors de la description des objectifs scientifiques, la méthodologie appliquée lors de l'étude peut être critiquée. En effet, une mauvaise interprétation sur les premières simulations a mené à l'hypothèse que l'évolution des variables d'état du système dynamique du vol étaient discontinues par rapport aux paramètres d'entrée. La recherche d'une solution efficace s'est alors orientée vers les métamodèles ME. De par l'identification de problématiques dans le domaine des métamodèles ME, et par un désir

de développement qui bénéficierait à de nombreux domaines scientifiques, la thèse s'est tournée vers l'établissement d'une nouvelle méthodologie reposant sur deux nouvelles méthodes. Ce développement s'éloigne alors de la problématique initiale. Les bifurcations des différents comportements de vol étudiés n'ont pas mené à des discontinuités, comme il a été conjecturé dans un premier temps. De ce fait, pour les comportements étudiés, l'apport de la nouvelle méthode n'a pas été conséquent. Le métamodèle PC suffit. En effet, les bifurcations ne provoquent pas d'irrégularité ni de discontinuité altérant la précision des métamodèles PC de façon significative. La thèse propose alors l'étude de sensibilité réalisée avec les métamodèles PC, qui permettent d'estimer les indices de Sobol de façon plus efficace qu'avec la méthode MC.

En ce qui concerne les perspectives :

- La méthode ME-ACD peut être améliorée. En effet, l'amélioration de la robustesse de la méthode par rapport à l'échantillonnage sera le sujet de futures recherches. Le développement d'un échantillonnage adaptatif, pour les utilisateurs ayant le choix de cet échantillonnage, pourra également améliorer la précision des métamodèles obtenus. Enfin, améliorer la méthode pour qu'elle puisse approcher des QOI dépendant d'un nombre plus important de paramètres d'entrée sera également un axe de recherche.
- Pour la formule analytique liant les indices de Sobol aux coefficients des métamodèles ME, l'utilisation d'une méthode ME classique reposant sur une partition de l'espace des paramètres en hyperpavés pourrait améliorer la précision et la robustesse de la méthode. Dans le manuscrit, l'utilisation d'une méthode partitionnant l'espace en fonction des discontinuités a été sélectionnée, afin d'éviter les imprécisions dues aux phénomènes de Gibbs ou à l'augmentation importante du nombre d'hyperpavés. Néanmoins, la procédure de partition en hyperpavés avec le classifieur mène également à des imprécisions.
- Pour l'analyse de sensibilité en dynamique du vol, l'utilisation d'un métamodèle PC semble suffisante pour les comportements étudiés. Le test de la méthode sur d'autres comportements et/ou avec d'autres modèles pourrait mener à d'autres conclusions. De plus, une étude de fiabilité ou une étude de stabilité à temps court pourraient être plus adaptées pour l'étude de l'influence des termes d'amortissement sur la stabilité d'un comportement.

Annexe A

Outils d'apprentissage automatique

Dans les travaux, des outils du domaine de l'apprentissage automatique (Machine learning - ML) sont utilisés. Cette annexe a pour objectif de présenter ces outils. De façon concrète dans les codes numériques développés, la bibliothèque Python scikit-learn [49] est utilisée.

A.1 Classifieurs

Dans le cadre des métamodèles multi-éléments (Multi-Element - ME) (Cf. section 2.6.4), des outils de classifications sont utilisés. Lorsqu'une évaluation sur un point représentant un jeu de paramètres d'entrée est réalisée, la première étape consiste à déterminer dans quel élément se trouve le point, pour ensuite utiliser le bon modèle local.

Un classifieur est un algorithme de classification statistique supervisé. Son objectif est, à l'aide d'un ensemble d'observations d'entraînement représentées par des points et possédant une certaine classe (modélisée par un entier), de pouvoir déterminer les probabilités qu'un point ne faisant pas parti de l'ensemble d'entraînement appartienne aux différentes classes.

A.1.1 SVM

Dans l'article [26], un classifieur SVM est utilisé pour incarner le rôle de la fonction indicatrice. SVM veut dire Support Vector Machine et est traduit en français par « machine à vecteurs de support » ou encore « séparateur à vaste marge ».

L'idée des SVM s'appuient sur la notion de marge maximale et la notion de fonction noyau. Ils ont été développés dans [9]. l'objectif est d'obtenir une frontière entre les observations d'entraînement à l'aide de l'optimisation quadratique. Les observations les plus proches de la frontière sont les « vecteurs de support ».

Dans le cas d'une SVM linéaire, une frontière sous forme d'hyperplan est recherchée. La marge maximale est la distance entre la frontière et les vecteurs de supports. En maxi-

misant cette quantité, l'hyperplan séparateur est optimal.

Néanmoins, les données ne sont pas toujours linéairement séparables (séparables par un hyperplan). L'astuce du noyau est alors utilisée. Cela permet de trouver un hyperplan séparateur des données, mais dans un espace transformé, si bien que dans l'espace non transformé, la frontière est non-linéaire.

Afin de trouver les coefficients définissant la frontière optimale, l'optimisation quadratique dans un espace dual peut être utilisée.

A.1.2 Classifieur par réseau de neurone MLP

Un réseau de neurones est un outil d'apprentissage supervisé pouvant être utilisé pour la régression ou la classification. Sa structure est inspirée de celle du cerveau animal. Leur développement résulte d'une longue suite de travaux ayant débuté en 1943 [42].

Dans cette section, les MLP, ou perceptrons multi-couches, sont présentés. Les MLP sont composés d'une succession de couches de cellules de base. Chaque cellule de base est appelée un « neurone ». Un neurone est représenté à la figure 37.

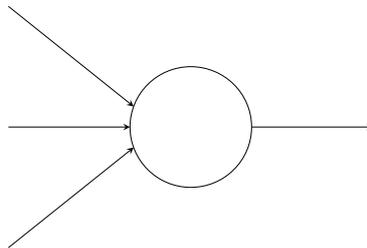


FIGURE 37 – Diagramme d'un neurone.

Dans les MLP, tous les neurones d'une couche sont connectés à ceux de la couche suivante. La figure 38 illustre un MLP. La première couche se nomme la couche d'entrée, la dernière couche se nomme la couche de sortie, et les couches intermédiaires se nomment les couches cachées.

Un poids est attaché à chaque couple de neurones connectés. Un scalaire supplémentaire, le « seuil », est associé à chaque neurone de la couche cachée. Ces poids et ces seuils, souvent initialisés à des valeurs aléatoires, sont les paramètres du réseau qui vont être ajustés au cours de l'apprentissage pour que ce dernier agisse comme un classifieur.

La couche d'entrée reçoit les différentes composantes d'une donnée vectorielle, une composante par neurone de la couche d'entrée. Dans le cas où les données d'entrée du classifieur sont des points d'un espace de dimension d , il y a d neurones d'entrée. Sur la figure 38, les données d'entrée sont de dimension 6. Ces neurones d'entrée transmettent les données d'entrée aux neurones de la première couche cachée.

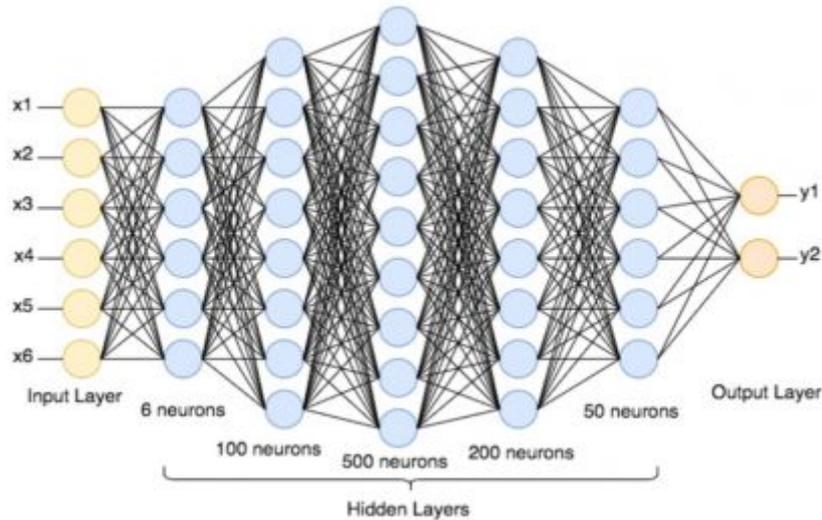


FIGURE 38 – Diagramme d’un réseau de neurones (source :<https://metalblog.ctif.com/2021/04/12/les-reseaux-de-neurones-et-la-data-augmentation/>).

Un neurone d’une couche cachée prend en entrée les données de sortie des neurones de la couche précédente. Il réalise ensuite une somme pondérée de ces entrées avec les poids correspondant au couple de neurones connectés, ajoute son seuil, puis compose cette somme par une fonction d’activation, qui est une fonction propre à la couche. Le résultat, la sortie du neurone, est ensuite envoyé aux neurones de la couche suivante et ainsi de suite.

La fonction d’activation modélise, comme son nom l’indique, l’activation ou non d’un neurone. Un neurone est dit activé s’il transmet son information aux neurones de la couche suivante. Les fonctions d’activation sont alors des fonctions valant 0 si le neurone n’est pas activé ou rendant compte de l’intensité de la somme pondérée sinon. Dans l’étude, la fonction ReLu, ou « rectified linear unit function » est utilisée. Cette fonction vaut

$$a(x) = \max(0, x) \tag{310}$$

Comme l’entrée de la fonction d’activation est une somme pondérée des sorties des neurones de la couche précédente ajoutée au seuil dépendant du neurone lui-même, le terme de seuil prend son sens. Si l’information reçue, sous la forme de la somme pondérée des sorties des neurones de la couche précédentes, dépasse le seuil, le neurone est activé, sinon, il ne l’est pas.

Dans le cas de la classification, il y a autant de neurones de sortie que de classes différentes. Sur la figure 38, les données sont classifiées en 2 classes. Chaque neurone de sortie prend les données de sortie des neurones de la dernière couche cachée, réalise une somme pondérée, mais n’applique aucune fonction d’activation. Les valeurs des neurones de sortie, vues comme les composantes d’un vecteur \mathbf{z} , sont ensuite mis en entrée des fonctions softmax_{*i*} associées à chaque classe *i* afin d’obtenir une valeur entre 0 et 1 pour chaque sortie.

$$\text{softmax}_i(\mathbf{z}) = \frac{\exp(z_i)}{\sum_{k=1}^K \exp(z_k)} \quad (311)$$

La valeur $\text{softmax}_i(\mathbf{z})$ est alors interprétée comme étant la probabilité que la donnée d'entrée soit de la classe i .

Le réseau de neurones est de ce fait une suite de compositions de fonctions non-linéaires capable d'approcher un comportement complexe une fois les poids et les seuils correctement réglés. L'apprentissage de ces poids et de ces seuils s'appuie sur la minimisation d'une fonction de coût à l'aide d'un jeu d'observations d'entraînement de classes connues. Pour une observation d'entraînement de classe connue, la fonction de coût prend en entrée les sorties du réseau de neurones obtenues en entrant les données d'entrée de l'observation, ainsi que la vraie classe de cette dernière. La fonction de coût évalue alors l'erreur commise par la classification du réseau de neurones pour évaluer dans quelle classe se trouve l'observation. Minimiser la fonction de coût sur cette observation d'entraînement permet de régler les poids et les seuils afin que le réseau de neurones donne une probabilité forte que les observations proches de cette dernière soient de sa classe. La minimisation de la moyenne des valeurs de la fonction de coût sur le jeu d'observations d'entraînement est alors réalisée dans le but d'obtenir un réseau de neurones donnant une bonne estimation proche de l'ensemble des données d'entraînement. Dans le cas de la classification multi-classes, la fonction de coût est l'entropie croisée moyenne. Afin d'obtenir les valeurs des poids qui minimisent cette fonction à l'aide des poids, un algorithme de type descente en gradient est utilisé. Pour obtenir le gradient de la fonction de coût par rapport aux poids, nécessaire pour l'algorithme, la méthode de rétropropagation est appliquée.

A.2 Gradient Boosted Tree

Dans l'étude, le Gradient Boosted Tree est utilisé comme métamodèle de régression afin de comparer les performances du métamodèle ME-ACD avec les siennes (Cf. section 3.3). Ce métamodèle, développé dans [20], repose sur l'idée de boosting. Le boosting est une technique de ML ayant pour but d'obtenir un métamodèle capable d'approcher une relation complexe, ce métamodèle étant une somme pondérée de métamodèles plus simples, moins précis et moins généralisables nommés « prédicteurs faibles ». Dans le cas du Gradient Boosted Tree, le prédicteur faible sélectionné est l'arbre de décision.

Un arbre de décision est une méthode de régression ou de classification supervisée qui a pour avantage d'être simple à comprendre et à représenter. Dans l'étude, le Gradient Boosted Tree est utilisé comme outil de régression, et ainsi sont également utilisés les arbres de décision sur lesquels il repose. Un arbre de décision utilisé pour la régression peut être vu comme un métamodèle constant par morceaux. Les arbres de décisions considérés sont des arbres obtenus à l'aide de l'algorithme CART [10], pour « Classification and Regression Trees ». Un tel arbre de décision repose sur une structure binaire. Il est composé de nœuds, et à chaque étape, un nœud peut se séparer en deux nœuds enfants. Chaque nœud « non-feuille » (un nœud qui se sépare) correspond à une condition. Cette

condition, ou ce test, correspond à vérifier si une des composantes de la donnée est inférieure ou égale à un certain seuil. À chaque nœud, les données d'entraînement sont ainsi séparées en deux groupes selon la condition. Les nœuds « feuilles », les derniers nœuds de la structure, renvoie la moyenne des valeurs de la quantité d'intérêt (Quantity Of Interest - QOI) sur les observations d'entraînement qui composent le nœud. Une fois l'arbre entraîné, lorsqu'une évaluation est nécessaire sur une donnée ne faisant pas partie de l'ensemble d'entraînement, les tests associés à chaque nœud sont successivement effectués depuis le nœud « racine » jusqu'au nœud « feuille » atteint avec la donnée mise en entrée. La valeur du QOI estimée par le métamodèle est alors la valeur renvoyée par le nœud « feuille » atteint en parcourant l'arbre.

Entraîner un arbre de décision revient alors à déterminer les conditions à associer à chaque nœud (la composante de la donnée qui sera soumise au test, le seuil du test...) et ainsi quelles dichotomies successives des observations d'entraînement sont à réaliser. Chaque séparation vise à diminuer le plus possible la fonction de coût sur les observations d'entraînement composant le nœud. Pour la régression, la fonction de coût est l'erreur quadratique moyenne. Dans chaque nœud, une recherche de la meilleure séparation est réalisée par le biais d'une méthode gloutonne. Pour chaque direction de l'espace des paramètres d'entrée, le meilleur seuil de séparation est déterminé, puis parmi ces points, la direction (avec son seuil associé déterminé à l'étape précédente) menant à la plus grande diminution de la fonction de coût est sélectionnée afin de séparer le nœud. Plusieurs critères d'arrêt peuvent être utilisés, comme le fait d'atteindre un nombre minimal d'observations d'entraînement constituant un nœud « feuille » ou le fait d'atteindre une certaine profondeur de l'arbre.

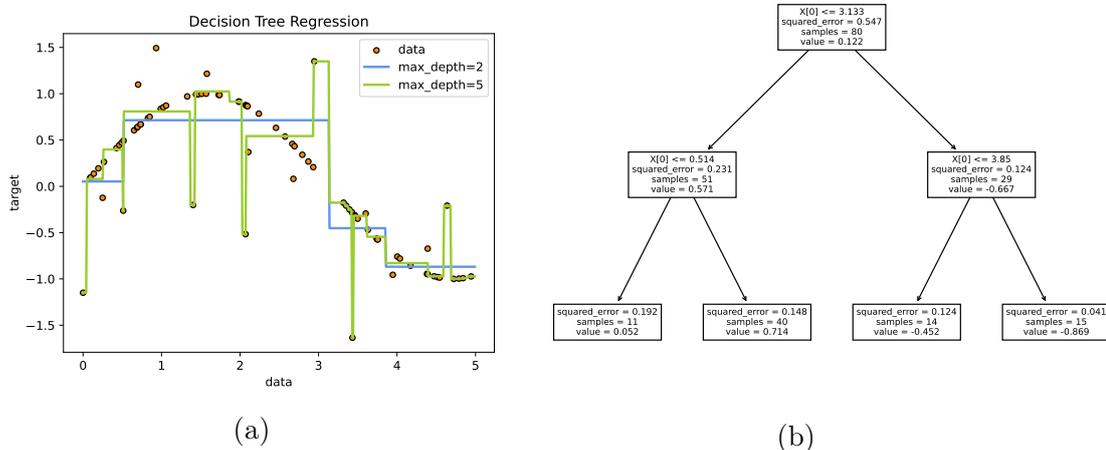


FIGURE 39 – (a) Régression de la fonction sinus (avec des observations bruitées) par deux arbres de régressions, un ayant une profondeur de 2 et l'autre ayant une profondeur de 5. (b) Représentation de l'arbre ayant une profondeur de 2.

La figure 39 illustre l'exemple d'une régression à l'aide d'un arbre de décision. Cet exemple est un exemple 1D proposé par scikit-learn [1]. La QOI dépendant d'une unique variable est un sinus additionné à un bruit. Comme le montre la figure 39a, si la profondeur maximale de l'arbre augmente, le métamodèle devient sensible à l'overfitting. La figure 39b illustre le fait que l'arbre de décision est facilement compréhensible et représentable.

Le métamodèle Gradient Boosted Tree est alors une somme pondérée d'arbres de décision [20] qui sont ajoutés progressivement de sorte à améliorer la précision du métamodèle, et à augmenter sa capacité de généralisation (éviter l'overfitting). À la première étape, le métamodèle Gradient Boosted Tree est une constante additionnée à un arbre de décision (multiplié par un coefficient de pondération). Sa précision est alors relativement mauvaise et il est sujet à l'overfitting. À chaque étape de boosting, un nouvel arbre de décision est appris et ajouté au modèle composé de la somme pondérée des arbres déterminés aux étapes précédente. Les paramètres du nouvel arbre ajouté et son coefficient de pondération sont ajustés de sorte que l'erreur entre la vraie valeur de la QOI et la valeur prédite par le métamodèle sur les observations d'entraînement soit réduite. Pour ajuster ces paramètres, une méthode de descente en gradient peut être utilisée. Un algorithme est proposé dans [20]. Le nouvel arbre s'additionne alors avec le métamodèle précédent pour donner le nouveau métamodèle boosté, tendant à corriger son erreur, et ainsi de suite.

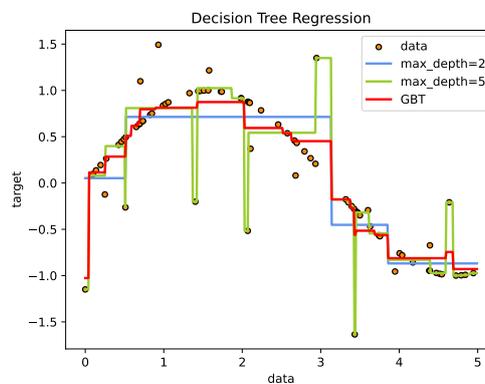


FIGURE 40 – Régression de la fonction sinus (avec des observations bruitées) par les deux arbres de décision du précédent exemple, ainsi que par un métamodèle GBT.

La figure 40 reprend l'exemple précédent, mais en illustrant également l'approximation par le métamodèle GBT composé d'une somme pondérée de 200 arbres de décisions de profondeur 1. L'augmentation en terme de précision et de généralisation est notable sur la figure.

A.3 Ball-tree

Dans l'étude, le Ball-tree, introduit dans [48], est utilisé pour la recherche des plus proches voisins pour ensuite estimer le gradient de la QOI par rapport aux paramètres d'entrée. Le Ball-tree est une structure en arbre binaire où chaque nœud correspond à l'hypersphère la plus petite contenant un sous-ensemble des observations. Les hypersphères peuvent avoir une intersection entre elles. D'un passage d'un nœud au suivant, les données sont séparées dans deux hypersphères minimales différentes. Le fait d'utiliser ces hypersphères minimales induit des propriétés idéales pour déterminer les plus proches voisins. L'article [48] compare empiriquement les performances de 5 algorithmes de construction de Ball-tree.

Bibliographie

- [1] Exemple de régression à l'aide d'arbres de décision, proposé par scikit-learn. URL : https://scikit-learn.org/stable/auto_examples/tree/plot_tree_regression.html#sphx-glr-auto-examples-tree-plot-tree-regression-py.
- [2] Loss of control in general aviation. URL : <https://www.easa.europa.eu/downloads/18980/en>.
- [3] NB Abramov, MG Goman, Alexander N Khrabrov, and BI Soemarwoto. Aerodynamic modeling for poststall flight simulation of a transport airplane. *Journal of Aircraft*, 56(4) :1427–1440, 2019.
- [4] Nikolay Abramov, M Goman, A Khrabrov, E Kolesnikov, M Sidoruk, B Soemarwoto, and H Smaili. Aerodynamic model of transport airplane in extended envelope for simulation of upset recovery. In *28th International Congress of the Aeronautical Sciences*, 2012.
- [5] Rick Archibald, Anne Gelb, and Jungho Yoon. Polynomial fitting for edge detection in irregularly sampled signals and images. *SIAM journal on numerical analysis*, 43(1) :259–279, 2005.
- [6] Kevin Beyer, Jonathan Goldstein, Raghu Ramakrishnan, and Uri Shaft. When is “nearest neighbor” meaningful? In *International conference on database theory*, pages 217–235. Springer, 1999.
- [7] William Bihrlé and Billy Barnhart. Spin prediction techniques. *Journal of Aircraft*, 20(2) :97–101, 1983.
- [8] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6) :2345–2367, 2011.
- [9] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.
- [10] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. *Classification and regression trees*. Routledge, 2017.
- [11] Robert H Cameron and William T Martin. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, pages 385–392, 1947.

- [12] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10) :1509–1518, 2007.
- [13] James V Carroll and Raman K Mehra. Bifurcation analysis of nonlinear aircraft dynamics. *Journal of Guidance, Control, and Dynamics*, 5(5) :529–536, 1982.
- [14] Tonkid Chantrasmi, Alireza Doostan, and Gianluca Iaccarino. Padé–legendre approximants for uncertainty analysis with discontinuous response surfaces. *Journal of Computational Physics*, 228(19) :7159–7180, 2009.
- [15] Thierry Crestaux, Olivier Le Maître, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7) :1161–1172, 2009.
- [16] Annick Dhooge, Willy Govaerts, Yu A Kuznetsov, Hil Gaétan Ellart Meijer, and Bart Sautois. New features of the software matcont for bifurcation analysis of dynamical systems. *Mathematical and Computer Modelling of Dynamical Systems*, 14(2) :147–175, 2008.
- [17] Dominique Farcy, Alexander Khrabrov, and Maria Sidoryuk. Sensitivity of spin parameters to uncertainties of the aerodynamic model of a light aircraft. In *Proceedings of the 8th European Conference for Aerospace Sciences (EUCASS)*, 2019.
- [18] Ralph D Fero. Investigation of the influence of rotary aerodynamics on the study of high angle of attack dynamics of the f-15b using bifurcation analysis. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING, 1991.
- [19] Jasmine Foo, Xiaoliang Wan, and George Em Karniadakis. The multi-element probabilistic collocation method (me-pcm) : Error analysis and applications. *Journal of Computational Physics*, 227(22) :9572–9595, 2008.
- [20] Jerome H Friedman. Greedy function approximation : a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.
- [21] Alan Genz. Testing Multidimensional Integration Routines. In *Proc. Of International Conference on Tools, Methods and Languages for Scientific and Engineering Computation*, pages 81–94, New York, NY, USA, 1984. Elsevier North-Holland, Inc.
- [22] Roger G Ghanem and Pol D Spanos. *Stochastic finite elements : a spectral approach*. Courier Corporation, 2003.
- [23] Stephen J Gill, Mark H Lowenberg, Simon A Neild, Bernd Krauskopf, Guilhem Puyou, and Etienne Coetzee. Upset dynamics of an airliner model : a nonlinear bifurcation analysis. *Journal of Aircraft*, 50(6) :1832–1842, 2013.
- [24] MG Goman, GI Zagainov, and AV Khramtsovsky. Application of bifurcation methods to nonlinear flight dynamics problems. *Progress in Aerospace Sciences*, 33(9-10) :539–586, 1997.
- [25] Philippe Guicheteau. Bifurcation theory : a tool for nonlinear flight dynamics. *Philosophical Transactions of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 356(1745) :2181–2201, 1998.

- [26] Yous V Halder, Benjamin Sanderson, and Barry Koren. An adaptive minimum spanning tree multielement method for uncertainty quantification of smooth and discontinuous responses. *SIAM Journal on Scientific Computing*, 41(6) :A3624–A3648, 2019.
- [27] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1) :1–17, 1996.
- [28] C Jahnke and FEC Culick. Application of bifurcation theory to the high-angle-of-attack dynamics of the f-14. *Journal of Aircraft*, 31(1) :26–34, 1994.
- [29] John D Jakeman, Akil Narayan, and Dongbin Xiu. Minimal multi-element stochastic collocation for uncertainty quantification of discontinuous functions. *Journal of Computational Physics*, 242 :790–808, 2013.
- [30] Michiel JW Jansen. Analysis of variance designs for model output. *Computer Physics Communications*, 117(1-2) :35–43, 1999.
- [31] Stephen Joe and Frances Y Kuo. Notes on generating sobol sequences. *ACM Transactions on Mathematical Software (TOMS)*, 29(1) :49–57, 2008.
- [32] Dan Kalman. A singularly valuable decomposition : the svd of a matrix. *The college mathematics journal*, 27(1) :2–23, 1996.
- [33] Juri Kalviste. Use of rotary balance and forced oscillation test data in six degrees of freedom simulation. In *9th Atmospheric Flight Mechanics Conference*, page 1364, 1982.
- [34] Jacob Kay. Acquiring and modeling unsteady aerodynamic characteristics. In *Atmospheric Flight Mechanics Conference*, page 3907, 2000.
- [35] A Khrabrov, M Sidoryuk, and M Goman. Aerodynamic model development and simulation of airliner spin for upset recovery. *Progress in flight physics*, 5 :621–636, 2013.
- [36] Diederik P Kingma and Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*, 2014.
- [37] Brianr Kramer. Experimental evaluation of superposition techniques applied to dynamic aerodynamics. In *40th AIAA aerospace sciences meeting & exhibit*, page 700, 2002.
- [38] Olivier Le Maître and Omar M Knio. *Spectral methods for uncertainty quantification : with applications to computational fluid dynamics*. Springer Science & Business Media, 2010.
- [39] Olivier Le Maître, Omar M Knio, Habib N Najm, and Roger G Ghanem. Uncertainty propagation using wiener–haar expansions. *Journal of computational Physics*, 197(1) :28–57, 2004.
- [40] Guang Lin, Xiaoliang Wan, Chau-Hsing Su, and George Em Karniadakis. Stochastic computational fluid mechanics. *Computing in Science & Engineering*, 9(2) :21, 2007.
- [41] Michel Loeve. Fonctions aleatoires du second ordre. *Processus stochastique et mouvement Brownien*, pages 366–420, 1948.

- [42] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4) :115–133, 1943.
- [43] Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2) :161–174, 1991.
- [44] Austin Murch. *Aerodynamic modeling of post-stall and spin dynamics of large transport airplanes*. PhD thesis, Georgia Institute of Technology, 2007.
- [45] Akil Narayan and Dongbin Xiu. Stochastic collocation methods on unstructured grids in high dimensions via interpolation. *SIAM Journal on Scientific Computing*, 34(3) :A1729–A1752, 2012.
- [46] Andre Nataf. Determination des distribution don’t les marges sont donnees. *Comptes rendus de l’Académie des Sciences*, 225 :42–43, 1962.
- [47] Sergey Oladyshkin and Wolfgang Nowak. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, 106 :179–190, 2012.
- [48] Stephen M Omohundro. *Five balltree construction algorithms*. International Computer Science Institute Berkeley, 1989.
- [49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn : Machine learning in Python. *Journal of Machine Learning Research*, 12 :2825–2830, 2011.
- [50] Prashant Rai. *Sparse low rank approximation of multivariate functions—Applications in uncertainty quantification*. PhD thesis, Ecole Centrale de Nantes (ECN), 2014.
- [51] Murray Rosenblatt. Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3) :470–472, 1952.
- [52] Bojana V Rosić and Jobst H Diekmann. Methods for the uncertainty quantification of aircraft simulation models. *Journal of Aircraft*, 52(4) :1247–1255, 2015.
- [53] Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. Why so many published sensitivity analyses are false : A systematic review of sensitivity analysis practices. *Environmental modelling & software*, 114 :29–39, 2019.
- [54] Andrea Saltelli and Paola Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12) :1508–1517, 2010.
- [55] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer physics communications*, 181(2) :259–270, 2010.
- [56] Gilbert Saporta. *Probabilités, analyse des données et statistique*. Editions technip, 2006.
- [57] Khachik Sargsyan, Cosmin Safta, Bert Debusschere, and Habib Najm. Uncertainty quantification given discontinuous model response and a limited number of model runs. *SIAM Journal on Scientific Computing*, 34(1) :B44–B64, 2012.

- [58] Khachik Sargsyan, Cosmin Safta, Habib N Najm, Bert J Debusschere, Daniel Ricciuto, and Peter Thornton. Dimensionality reduction for complex models via bayesian compressive sensing. *International Journal for Uncertainty Quantification*, 4(1), 2014.
- [59] Maria E Sidoryuk and Alexander N Khrabrov. Estimation of regions of attraction of aircraft spin modes. *Journal of Aircraft*, 56(1) :205–216, 2019.
- [60] Sergei Abramovich Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Doklady Akademii Nauk*, volume 148, pages 1042–1045. Russian Academy of Sciences, 1963.
- [61] Ilya M Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4) :784–802, 1967.
- [62] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1-3) :271–280, 2001.
- [63] Ilya M Sobol and Sergei Kucherenko. Derivative based global sensitivity measures. *Procedia-Social and Behavioral Sciences*, 2(6) :7745–7746, 2010.
- [64] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7) :964–979, 2008.
- [65] Nicolas Vauchel, Eric Garnier, and Thomas Gomez. Estimation of sobol indices with a multi-element polynomial chaos model : application to flight dynamics. In *25ème Congrès Français de Mécanique*, 2022.
- [66] Nicolas Vauchel, Éric Garnier, and Thomas Gomez. A multi-element non-intrusive polynomial chaos method using agglomerative clustering based on the derivatives to study irregular and discontinuous quantities of interest. *Journal of Computational Physics*, 473 :111763, 2023.
- [67] Nicolas Vauchel, Éric Garnier, and Thomas Gomez. Estimation of sobol indices using a multi-element polynomial chaos model. *en cours d'écriture*.
- [68] Mathukumalli Vidyasagar. *Nonlinear systems analysis*. SIAM, 2002.
- [69] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2) :617–642, 2005.
- [70] Jean-Claude Wanner. *Dynamique du vol et pilotage des avions*. ONERA/P-83/1. ONERA, 1983. URL : <https://books.google.fr/books?id=b-ESmgEACAAJ>.
- [71] Joe H Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301) :236–244, 1963.
- [72] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4) :897–936, 1938.
- [73] Dongbin Xiu. Efficient collocational approach for parametric uncertainty analysis. *Commun. Comput. Phys*, pages 293–309, 2007.

- [74] Dongbin Xiu. Fast numerical methods for stochastic computations : a review. *Communications in computational physics*, 5(2-4) :242–272, 2009.
- [75] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2) :619–644, 2002.