

UNIVERSITÉ DE MONTRÉAL

INTÉGRATION ET OPTIMISATION MULTIDISCIPLINAIRE D'UNE
TURBOMACHINE AU STADE PRÉLIMINAIRE DE CONCEPTION

DENIS WALCH

DÉPARTEMENT DE GÉNIE MÉCANIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
NOVEMBRE 2007



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-25585-8

Our file Notre référence

ISBN: 978-0-494-25585-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

INTÉGRATION ET OPTIMISATION MULTIDISCIPLINAIRE D'UNE
TURBOMACHINE AU STADE PRÉLIMINAIRE DE CONCEPTION

présenté par: WALCH Denis

en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées

a été présentée au jury d'examen constitué de:

M. VO Huu Duc, Ph.D., président

M. TRÉPANIÉ Jean-Yves, Ph.D., membre et directeur de recherche

M. TRIBES Christophe, Doct. ès sciences, membre et codirecteur de recherche

M. VADEAN Aurelian, Doct., membre

“Check, Double-Check, and Check again”

Peu d’autres dictons n’auront mieux reflété mes travaux de recherche...

REMERCIEMENTS

Je tiens à remercier les membres du labo, et toutes les personnes qui ont contribué de près ou de loin à mes travaux de recherche.

Je voudrais remercier le CRIAQ pour son support financier qui m'a permis de réaliser mon projet de recherche en maintenant une bonne qualité de vie. En premier lieu, je tiens à remercier Pascal Prado pour son travail rigoureux, ses précieux conseils et ses explications au début de ma maîtrise qui m'ont permis de bien comprendre un système aussi complexe que PMDO. Je tiens aussi à remercier Harish Ganoory et Paul Tran, étudiants en génie informatique, pour leurs importantes contributions logicielles au projet PMDO. Daniel Tardif, étudiant en génie mécanique, a également contribué à la phase d'intégration.

Ensuite, malgré certains événements qui ont limité le support que je pouvais recevoir de Pratt & Whitney Canada, je tiens à remercier Yulia Panchenko et Jasmin Turcotte pour leur suivi, leurs commentaires et suggestions.

Christophe Tribes fut un guide patient, une aide précieuse et constante tout au long de ma maîtrise. Merci également à Jean-Yves Trépanier pour sa supervision discrète mais ferme et motivante.

Je tiens à remercier mes partenaires de laboratoire Jonathan Jean-Dallaire et Noémi Giammichele pour leurs contributions à la réussite de mes cours. Un merci particulier à Jocelyn Bourgeois pour les échanges constructifs qui m'ont permis de faire progresser mes recherches.

J'aimerais féliciter l'ensemble des membres du labo Luminos pour leur compréhension face à mes inquiétudes pour mes problèmes d'immigration.

Enfin, je tiens à remercier mon épouse Geneviève, sans qui je n'aurai probablement jamais ni entrepris ni achevé ma maîtrise. Elle m'a toujours soutenu dans "les bons coups et les échecs".

RÉSUMÉ

Ce mémoire présente l'intégration et l'optimisation d'une turbomachine au stade de conception préliminaire. Une stratégie est proposée pour mener des optimisations sur un tel problème multidisciplinaire complexe. La méthode présentée se base sur une analyse de sensibilité des variables en vue d'une sélection suivie d'un traitement spécifique sur les couplages des analyses multidisciplinaires. L'idée principale est de mener une optimisation en trois étapes, utilisant trois couches de variables distinctes.

L'analyse de sensibilité permet de sélectionner des variables d'entrée d'analyses multidisciplinaires. Le but est d'établir (i) un classement des variables d'entrée en fonction de leur influence relative sur les variables de sortie et (ii) une sélection pertinente de ces variables afin de les incorporer dans la stratégie globale. L'analyse de sensibilité proposée est basée sur un échantillonnage d'hypercube latin de l'espace de conception et sur la construction de modèles de surface de réponse. Ce modèle est une approximation au second ordre de données sur les variables d'entrée et de sortie, obtenu par une méthode de minimisation des moindres carrés. La sensibilité d'une variable d'entrée sur une variable de sortie est déterminée par les poids du modèle. Deux critères ont été établis pour le classement et la sélection des variables. Le premier critère est un seuil qui détermine si la variable est suffisamment influente. Le second critère définit un objectif de simplification : un nombre maximum de variables peut être sélectionné par sous-ensemble analysé. Des variables d'entrée sont ainsi sélectionnées pour chaque variable de sortie considérée. Dans une discipline, les listes sont alors comparées pour construire une liste finale par étape d'optimisation.

Ensuite, le traitement des couplages permet d'analyser des liens spécifiques pour les suspendre temporairement de l'analyse multidisciplinaire, sans comparaison relative aux autres liens. Les analyses multidisciplinaires nécessitent des ensembles de valeurs cohérentes pour les variables d'entrée, ce qui mène à des répartitions particulières de certaines valeurs. L'analyse de ces répartitions des valeurs des variables d'entrée, parallèlement

aux répartitions des variables de sortie, permet de déceler des incompatibilités potentielles pour les variables couplées (à la fois variables d'entrée et de sortie). Ensuite, des modèles de surface de réponses sont construits pour les variables impliquées dans un lien de couplage choisi, sur des sous-intervalles de l'espace de conception. Les termes linéaires des modèles sont interprétés comme des segments de fonction affines, et permettent de déterminer les sous-intervalles sur lesquels les variables varient le plus. Le croisement de ces informations avec la répartition des valeurs des variables d'entrée sur leurs intervalles permet de déterminer la fraction de valeurs d'une variable d'entrée qui ont le plus d'influence sur une variable de sortie.

L'étude montre que l'analyse de sensibilité est performante pour la sélection de variables influentes. En effet, une optimisation utilisant les seules variables les plus importantes permet de réduire le poids global en maintenant les performances de la turbomachine, dans une proportion comparable à une optimisation classique utilisant toutes les variables. La méthode proposée pour le traitement des couplages n'a pas pu permettre de retirer les liens les plus délicats à gérer. Enfin, la stratégie proposée est performante sur un problème qui possède des caractéristiques spécifiques. Or, le problème PMDO proposé par Pratt & Whitney Canada n'est pas parfaitement compatible avec ces caractéristiques. Toutefois, des optimisations exploitant l'analyse de sensibilité proposée sont capables de déterminer des géométries de turbomachine qui minimisent le poids dans un temps de calcul très compétitif.

ABSTRACT

This thesis presents the integration and optimization of an advanced turbomachinery design. A strategy proposes a three-step optimization procedure to cope with such a complex multi-disciplinary problem. This method is based on a sensitivity analysis on variables for selection and a specific process on the multi-disciplinary analyses couplings.

The sensitivity analyses allow to select input variables from multi-disciplinary analyses. The purpose is to establish (i) a ranked list of variables according to their relative influence on the outputs, and (ii) a relevant selection of these variables, to be incorporated into the global strategy. The proposed sensitivity analysis method is based on latin-hypercube sampling and a response surface model. This model is a second-order approximation of the sampled input and output variables obtained by a least-square fitting method. The sensitivity of an output variable to an input variable is obtained from the weights of the model. For the ranking and the selection of the variables, two criteria have been established. The first criterion is a threshold to decide if the variable has enough influence or not. The second defines the simplification objective : a maximum number of variables to be managed per subset. Then, some input variables are selected for each considered output variable. To take all output variables into account, lists are compared between the outputs to built a final list for each of the three optimization steps.

The coupling processing is made by conducting analyses on specific links to temporarily suspend them from the multi-disciplinary analysis (MDA), without any relative comparison to other links. The MDAs require consistent sets of values for the input variables. This leads to particular distributions of some variables. The analysis of the input variable values distributions, along with output variables values distributions allow to detect incompatibilities for coupled variables (that can be both an input and an output variable). Then, response surface models (RSM) are built for the variables that are involved in a chosen coupling link, on sub-intervals of the design space. Linear terms of the models

are interpreted as segments of linear functions and allow to determine the sub-intervals where the variables have the most important amplitude change. The variables values distribution on their intervals cross-referenced with these linear terms, makes possible to find the variables values fraction that give the most important influence on an output variable.

The study shows that the sensitivity analysis performs a relevant selection of influent variables. Indeed, an optimization that uses only the most important variables, enables a global weight reduction comparable with a classic optimization process that uses all the variables, while maintaining the engine performance. The proposed method for coupling processing did not help to suspend the most problematic links. Finally, the proposed strategy performs well on a problem that has specific characteristics. The PMDO problem from Pratt & Whitney Canada does not perfectly fit with these characteristics. However, some optimizations that uses the proposed sensitivity analysis, are able to find turbomachine geometries that reduce the overall weight within a very competitive timescale.

TABLE DES MATIÈRES

DÉDICACE	iv
REMERCIEMENTS	v
RÉSUMÉ	vi
ABSTRACT	viii
TABLE DES MATIÈRES	x
LISTE DES FIGURES	xiv
LISTE DES NOTATIONS ET DES SYMBOLES	xvi
LISTE DES TABLEAUX	xviii
LISTE DES ANNEXES	xix
INTRODUCTION	1
CHAPITRE 1 INTRODUCTION AU PROJET DE MAÎTRISE	3
1.1 Contexte	3
1.1.1 Pratt & Whitney Canada	3
1.1.2 Le prototype PMDO	3
1.2 Présentation du projet de maîtrise	5
1.2.1 Objectif général	5
1.2.2 Problématique	5
1.2.2.1 Outils fournis par P&WC	6
1.2.2.2 Exécution d'une discipline	7
1.2.2.3 Expérience des projets passés	9

1.2.3	Simplification du problème	9
1.2.4	Cadre théorique	10
1.2.4.1	Intégration	10
1.2.4.2	Optimisation	12
CHAPITRE 2	REVUE DE LITTÉRATURE	13
2.1	Stratégies générales d'optimisation	13
2.2	Importance des variables et leur classification	18
2.2.1	Modèle de surface de réponse	18
2.2.2	Filtrage par groupe ou "group screening"	20
2.2.3	Réseaux de neurones	21
2.2.4	Analyse par Composantes Principales	22
2.2.5	Algorithmes génétiques	23
2.2.6	Exploitation de la formulation analytique des fonctions coûts . .	24
2.2.7	Bilan sur les méthodes de sélection de variables	24
2.3	Étude des couplages interdisciplinaires	26
2.3.1	Gestion des liens interdisciplinaires	27
2.3.2	Mesures du couplage	28
2.3.2.1	Équations de sensibilités globales ou "Global Sensitivity Equations" (GSE)	28
2.3.2.2	Méthode des équations de sensibilité globale modifiée	31
2.3.2.3	Classification des couplages	34
2.4	Algorithmes d'optimisation	36
2.5	Conclusion de la revue et établissement de la démarche	40
CHAPITRE 3	CLASSIFICATION DES VARIABLES	43
3.1	Définition de la sensibilité	43
3.2	Influence et surface de réponse	45
3.2.1	Définition de l'influence	45

3.2.2	Influence relative et absolue	46
3.3	Modèle de surface de réponse	47
3.4	Classification	48
3.5	Stabilité de l'espace de départ et contraintes	54
CHAPITRE 4	ÉTUDE DES COUPLAGES	56
4.1	Notion de transport de couplage	56
4.2	Détermination de l'hypothèse de transport de couplage disjoint	57
4.2.1	Répartition des valeurs des variables couplées	58
4.2.2	Intervalle d'influence	60
4.2.3	Croisement des informations	64
4.3	Résultats de l'analyse	64
CHAPITRE 5	PROCESSUS GLOBAL D'OPTIMISATION : APPLICATION	66
5.1	Stratégie globale d'optimisation	66
5.2	Illustration par un cas test	67
5.2.1	Formulation	67
5.2.2	Historique et résultats d'optimisation	68
5.2.3	Conditions nécessaire d'efficacité	70
CHAPITRE 6	RÉSULTATS DES OPTIMISATIONS DU SYSTÈME PMDO	72
6.1	Caractéristiques du système PMDO	72
6.2	Optimisations avec Isight	74
6.2.1	Fonction coût	74
6.2.2	Historique d'optimisation	74
6.3	Implémentation d'un optimiseur à gradient	76
6.3.1	Fonction coût utilisée	76
6.3.2	Réglages de l'algorithme d'optimisation	77
6.3.3	Calcul du gradient	79

6.3.4	Optimisations avec les couches de variables	81
6.3.5	Bilan de l'optimisation par gradients	86
6.4	Discussion critique des résultats obtenus	86
6.4.1	Réduction de poids	86
6.4.2	Problème d'optimisation	87
6.5	Orientation des travaux futurs	88
6.5.1	Analyse de sensibilité	89
6.5.2	Étude des couplages	90
CONCLUSION		92
BIBLIOGRAPHIE		96
ANNEXES		102

LISTE DES FIGURES

FIG. 1.1	Architecture du prototype PMDO	4
FIG. 1.2	MDA initial	7
FIG. 1.3	MDA après simplifications	11
FIG. 2.1	Configurations possibles du MDA	27
FIG. 2.2	MDA considéré pour la méthode GSE	29
FIG. 2.3	Organisation des disciplines : problème maître, et deux sous-tâches	32
FIG. 3.1	Tracé des coefficients d'un modèle de surface de réponse . . .	47
FIG. 3.2	Analyse multidisciplinaire exploitant la couche de variables 1 .	53
FIG. 3.3	Analyse multidisciplinaire exploitant la couche de variables 1 et 2	53
FIG. 3.4	Analyse multidisciplinaire exploitant la couche de variables 1, 2 et 3	54
FIG. 4.1	Transport de couplage	57
FIG. 4.2	Illustration des intervalles joints et disjoints	58
FIG. 4.3	Répartition des valeurs sur les intervalles de référence	60
FIG. 4.4	Tendances de variables d'entrée par rapport à une variable de sortie de la discipline 1	62
FIG. 4.5	Découpages de l'espace pour déterminer les surfaces de réponses de X_1	62
FIG. 4.6	Analyse croisée des répartitions des valeurs des variables et des intervalles d'influence	63
FIG. 5.1	Processus général d'optimisation, avec 3 couches de variables et une discipline	67
FIG. 5.2	Surface de réponse pour la fonction cas test	68
FIG. 5.3	Historique d'optimisation : valeurs de la fonction coût à la fin de chaque itération	69
FIG. 6.1	Exemple d'un modèle de surface de réponse de PMDO	73

FIG. 6.2	Historique de la fonction coût	75
FIG. 6.3	Historique de la fonction coût : un calcul de gradient avec optimisation des disciplines et une recherche linéaire	80
FIG. 6.4	Historique de la fonction coût : un calcul de gradient avec exécution des disciplines et une recherche linéaire	81
FIG. 6.5	Coupe du moteur initial, Poids total 5870 livres	84
FIG. 6.6	Coupe du moteur : Optimisation de la couche 1, Poids total 5550 livres	85
FIG. 6.7	Coupe du moteur : Optimisation de la couche 2, Poids total 5566 livres	85
FIG. 6.8	Coupe du moteur : Optimisation de la couche 3, Poids total 5561 livres	85
FIG. I.1	Extrait de la feuille Excel d'analyse de l'hélice	105
FIG. I.2	Référence à une autre discipline	106
FIG. I.3	Référence à une feuille non intégrée	108
FIG. I.4	Calculs transférés dans la discipline	109
FIG. II.1	Définition du problème de lissage	111
FIG. II.2	Lissage avec 4 splines de degré 3	112
FIG. II.3	Lissage avec 2 splines de degré 3	113
FIG. II.4	Lissage avec 2 splines de degré 3, contraintes mixtes	114
FIG. II.5	Lissage avec une formulation spécifique	115
FIG. III.1	Décomposition DO	118

LISTE DES NOTATIONS ET DES SYMBOLES

Acronymes utilisés dans le texte

ANOVA	ANalysis Of VAriance : analyse de variance
API	Application Program Interface : routines permettant une extension d'un langage de programmation
DAO	Distributed Analysis Optimization : méthode de décomposition
DO	Distributed Optimization : méthode de décomposition
DOE	Design Of Experiment : plan d'expériences
FIO	Fully Integrated Optimization : méthode de décomposition
GA	Genetic Algorithm : algorithme génétique
GSE	General Sensitivity Equations : équations de sensibilité globales
GUI	Graphical User Interface : interface graphique
IFFD	Iterated Fractional Factorial Design : expériences par plan fractionnels-factoriels itératif
JDM	Java Data Manager : nom du serveur dans le prototype PMDO
MDO	Multidisciplinary Design Optimization : optimisation multidisciplinaire en conception
MDA	Multi-Disciplinary Analysis : analyse multidisciplinaire
PMDO	Preliminary Multi-disciplinary Design Optimization : optimisation multidisciplinaire en conception préliminaire
P&WC	Pratt & Whitney Canada
RSM	Response Surface Model : modèle de surface de réponse
SB	Sequential Bifurcation : bifurcation séquentielle
TCP/IP	Transmission Control Protocol/Internet Protocol : protocoles de communication réseau

Caractères usuels

C_n^k	Nombre de Combinaisons de k éléments parmi n
D_i	Discipline portant le numéro i
f	Fonction coût
R	Coefficient de détermination
$S_{i,j}$	Sensibilité de la variable i par rapport à la variable j
$\bar{S}$	Moyenne de S
$\sigma(S)$	Écart-type de S
X	Variable d'entrée
X'	Variable adimensionnée
X^*	Variable optimisée
$\hat{X}$	Variable partiellement optimisée
Y	Variable de sortie

LISTE DES TABLEAUX

TAB. 1.1	Liste des disciplines intégrées dans le système PMDO	6
TAB. 2.1	Classement des sensibilités, avec k_1 et k_2 des paramètres réglables	35
TAB. 2.2	Acronymes des optimiseurs utilisables dans Isight	37
TAB. 2.3	Fonctionnalités des optimiseurs de Isight	38
TAB. 6.1	Répartition du poids selon les composantes pour les 3 cas	83

LISTE DES ANNEXES

ANNEXE I	INTÉGRATION DES DISCIPLINES	102
I.1	Outil fourni par P&WC et outils nécessaires pour PMDO	102
I.2	Principe d'ajout d'une variable	103
I.3	Variables Locales	105
I.4	Variables couplées	106
I.5	Références à des feuilles non intégrées	107
I.6	Bilan	108
ANNEXE II	REMPLACEMENT DE FONCTION SPLINES	110
II.1	Problème posé	110
II.2	Formulations possibles	110
II.3	Formulation spécifique	114
II.4	Conséquence du remplacement	115
ANNEXE III	MÉTHODE DE DÉCOMPOSITION DO	117

INTRODUCTION

Le développement constant du transport aérien entraîne une forte compétition entre les fournisseurs de moteurs comme General Electric, Rolls Royce, SNECMA, et Pratt & Whitney. Face à cette concurrence internationale, des progrès de productivité, de performance et d'expertise réalisés plus rapidement que les concurrents peuvent résulter en des concepts de futurs produits qui soient plus compétitifs sur le marché. Pour stimuler ces progrès, les activités de recherche et développement (R&D) s'orientent suivant deux axes. Le premier vise une meilleure compréhension de la physique des phénomènes dans la chambre de combustion, de meilleures descriptions des écoulements, des modèles plus précis pour la prédiction des transferts de chaleur, etc. Bien qu'essentiels et très prometteuses, ces études sont souvent envisageables seulement sur des perspectives à long terme. Le deuxième axe vise en général une meilleure exploitation des outils existants. Dans cette optique, l'intégration des outils disponibles dans un code unique afin d'obtenir une meilleure compréhension des dépendances des composants les uns par rapport aux autres est une avenue intéressante. Ce type de code intégré permet alors d'explorer des régions de l'espace de conception qui aboutissent à des moteurs réalisables, et par la suite, d'identifier des solutions optimales suivant les critères choisis.

Ainsi, le projet MOSAIC (système d'Optimisation basé sur des Standards et des Analyses et permettant une Intégration Configurable) financé par le CRIAQ (Consortium de Recherche et d'Innovation en Aérospatiale au Québec), et les partenaires industriels Bombardier, Bell Helicopter Textron Canada et Pratt & Whitney Canada a pour objectifs :

- le développement de protocoles d'échanges de données permettant l'intégration harmonieuse des outils d'analyse et d'optimisation dans un contexte de conception multidisciplinaire (MDO),

- l'intégration des outils d'analyse et d'optimisation dans des systèmes de conception multidisciplinaires propres à chaque entreprise et permettant la conception optimale de composants aéronautiques.

Le projet PMDO (Preliminary Multi-disciplinary Design Optimization : optimisation multidisciplinaire pour la conception préliminaire) constitue une des tâches du projet MOSAIC. Il répond à ce besoin d'intégration d'outils pour l'optimisation multidisciplinaire pour la conception en aéronautique. Dans le cadre du projet PMDO et dans la suite du projet de M. Pascal Prado (2005), ce mémoire présente une stratégie de résolution du problème d'optimisation multidisciplinaire fourni par le partenaire industriel Pratt & Whitney Canada (P&WC), s'appuyant sur le prototype établi par M. Prado.

Une présentation du projet de maîtrise replace les travaux de recherche dans leur contexte, en expliquant notamment les outils déjà développés par M. Prado, lors de son projet de maîtrise, et pose les problématiques des travaux de recherche dans le premier chapitre. Le mémoire propose une démarche pour traiter le problème d'optimisation multidisciplinaire proposé par P&WC, présentée après une revue bibliographique dans le chapitre 2. Deux éléments déterminants de la stratégie proposée sont présentés dans les deux chapitres suivants : une méthode d'analyse de sensibilités des variables d'entrée est présentée au chapitre 3 et un traitement des couplages est exposé au chapitre 4. Le processus général d'optimisation et des résultats préliminaires sont présentés au chapitre 5. Enfin, les résultats d'optimisations et des recommandations pour l'application de cette stratégie figurent au chapitre 6.

CHAPITRE 1

INTRODUCTION AU PROJET DE MAÎTRISE

1.1 Contexte

1.1.1 Pratt & Whitney Canada

Pratt & Whitney Canada (P&WC) est le premier motoriste mondial¹ pour les turbines à gaz de petite et moyenne puissance. Cette compagnie fournit principalement les moteurs destinés aux avions d'affaires, hélicoptères et avions de transport régionaux. L'expertise dans la conception des turbomachines est divisée entre Mississauga et Longueuil. Un groupe d'ingénieurs du département d'ingénierie avancée (Advanced Engineering) de Longueuil est impliqué à titre de principaux consultants dans ce projet de maîtrise.

1.1.2 Le prototype PMDO

Ce projet de maîtrise s'inscrit dans la suite du projet de M. Prado. Le prototype PMDO (Prado, 2005) est un outil d'optimisation multidisciplinaire pour le développement de produits dans une phase de conception avancée. M. Prado a conçu une plateforme d'intégration portable et configurable afin d'utiliser divers outils d'analyse, dont le fonctionnement a été démontré sur des problèmes tests. Cette plateforme est l'élément qui relie les outils d'analyse à l'optimiseur et qui permet à l'utilisateur de spécifier, via une interface, le type de problème à résoudre. L'architecture modulaire, illustré sur la figure 1.1, autorise le remplacement aisé de l'un des composants (outils d'analyse, optimiseur, in-

¹D'après www.pwc.ca

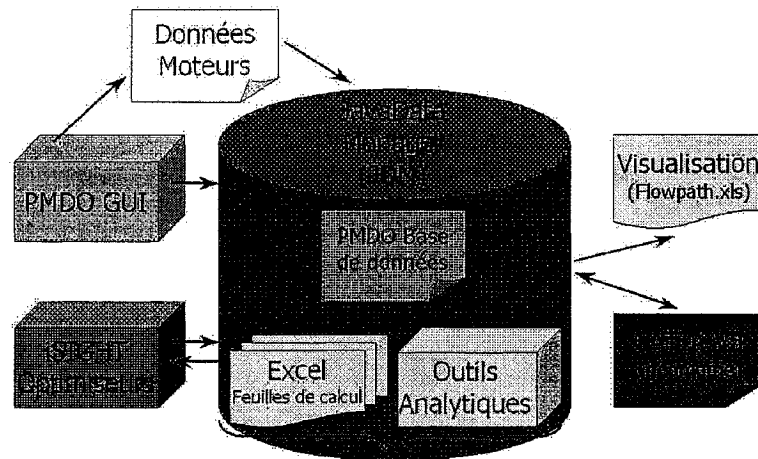


FIG. 1.1 Architecture du prototype PMDO

terface graphique) au besoin.

Les composantes principales du prototype PMDO, ainsi que leurs fonctions sont les suivantes :

- Le Serveur (Java Data Manager, JDM) gère les requêtes des clients suivant un protocole TCP/IP. C'est l'entité qui gère les exécutions disciplinaires et conserve à jour un fichier XML de base de données.
- Les Disciplines (formules analytiques ou feuilles de calcul Excel) sont les outils de dimensionnement des composants de la turbomachine.
- Le fichier PMDODatabase.xml est la base de données qui conserve les caractéristiques des variables.
- L'optimiseur (Isight) permet à l'utilisateur de définir le problème d'optimisation.

Cette plateforme d'intégration est codée en langage JAVA, excepté Isight, et peut être utilisée sur de nombreux systèmes d'exploitation. Une interface de programmation (API) en JAVA permet d'exploiter les feuilles de calcul du logiciel Microsoft Excel.

1.2 Présentation du projet de maîtrise

1.2.1 Objectif général

L'objectif général de ce travail consiste à formuler et tester une stratégie permettant de rendre l'optimisation réalisable avec la plateforme de Prado et dans un temps acceptable. Dans ce but, le projet de maîtrise comporte deux objectifs spécifiques. L'intégration d'outils d'analyse actuels dans le prototype, ainsi que l'établissement d'une stratégie d'optimisation permettant de diminuer le temps de résolution.

L'objectif principal est de rendre le prototype opérationnel pour l'analyse couplée d'une turbine à gaz durant la phase de conception préliminaire. Ce travail permettra de valider les choix de M. Prado et de tester les capacités d'intégration du prototype sur des outils d'analyses réels.

1.2.2 Problématique

Le but de l'optimisation est de déterminer une géométrie de la veine gazeuse qui minimise le poids. L'analyse multidisciplinaire (MDA) du problème est présentée sur la figure 1.2. Les disciplines correspondent aux analyses spécifiques des composants du moteur, et figurent dans le MDA sous la forme de boîtes labellées D_i , où $i \in \{1, \dots, 13\}$. Le tableau 1.1 donne une correspondance entre la numérotation des disciplines, le composant de la turbomachine, et un exemple de variable de sortie. Les lignes verticales représentent les variables d'entrée, et les chiffres caractérisent le nombre de variables. Les lignes horizontales indiquent les variables de sortie. Chaque discipline dépend ensuite de variables d'entrée locales (colonne de gauche), de variables partagées et de couplages figurant sur le MDA. Le nombre de variables de sortie pour chaque discipline est donné par les valeurs de la colonne de droite. Le problème d'optimisation comporte 557 va-

Dénomination	Nom complet de la discipline	Exemple de variable de sortie
D1	Hélice	<i>FanDiameter</i>
D2	Compresseur Basse Pression	<i>LPCLax</i>
D3	Conduit	<i>iCDuctLaxMin</i>
D4	Compresseur Haute Pression	<i>HPCBladeHeightOut</i>
D5	Compresseur Centrifuge - Aero	<i>Thpcinout</i>
D6	Compresseur Centrifuge - Structure	<i>HPTRotationalSpeedaero2</i>
D7	Chambre de combustion axiale	<i>CombDPperP</i>
D8	Chambre de combustion renversée	<i>RevCombDPperP</i>
D9	Turbine Haute Pression - Aero	<i>HPTDhperU2Stg1</i>
D10	Turbine Haute Pression - Structure	<i>BurstMarginHPTStress1</i>
D11	Turbine Basse Pression - Aero	<i>LPTMaxFlareAngle</i>
D12	Turbine Basse Pression - Structure	<i>BurstSpeedMargin1Err</i>
D13	Arbre de transmission	<i>MaxDeflectedRadius</i>

TAB. 1.1 Liste des disciplines intégrées dans le système PMDO

riables réparties en 13 disciplines. Les variables partagées sont des variables d'entrée d'au moins 2 disciplines simultanément, et sont au nombre de 28. 323 variables sont locales : gérées uniquement au niveau des disciplines. Les liens de couplages entre les disciplines sont caractérisés par le nombre de variables constituant le lien. La fonction coût est calculée à partir de certaines des 127 variables de sortie. Enfin, 79 variables décrivent des couplages entre les disciplines. Les variables couplées sont à la fois des variables de sortie d'une discipline et des variables d'entrée d'une ou de plusieurs disciplines.

1.2.2.1 Outils fournis par P&WC

Chaque discipline est modélisée dans le système sous la forme d'une feuille de calcul Excel spécifique. Initialement, ces feuilles indépendantes étaient reliées par des cellules partagées (liens de couplage entre les disciplines). La séparation du classeur fourni par le département d'ingénierie avancée de P&WC en ces disciplines indépendantes a consisté

Analyse Multidisciplinaire

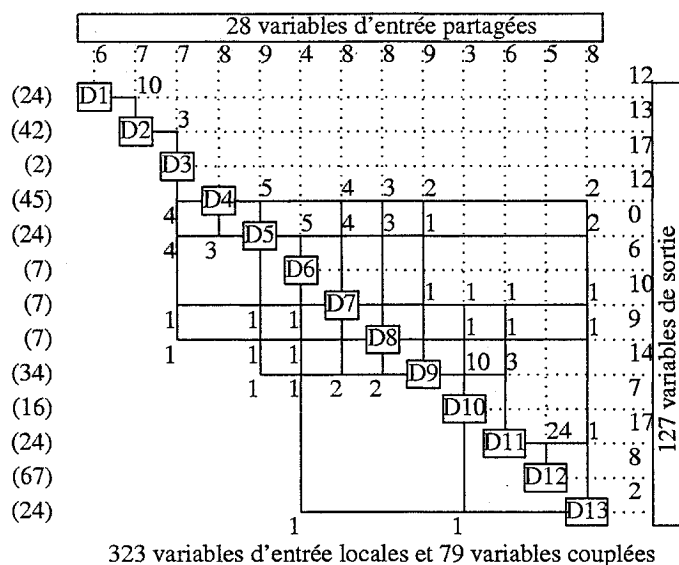


FIG. 1.2 MDA initial

en la suppression de ces cellules partagées et en l'introduction de variables de couplage pour conserver les liens entre les disciplines. Les résultats finaux (coupes de moteurs optimaux du chapitre 5) sont issues d'une mise à jour du classeur Excel par le JDM.

1.2.2.2 Exécution d'une discipline

Le serveur JAVA (JDM) gère la base de données ainsi que l'exécution des disciplines. La plateforme PMDO est conçue pour exploiter un protocole de communication TCP/IP au sein d'une architecture client-serveur, où le serveur reste le JDM et où le client est, en général, l'optimiseur intégré (Isight). Dans une telle architecture, le client est restreint à envoyer des messages et à attendre les réponses du serveur. Le résultat d'un calcul est obtenu à la fin de l'exécution des étapes suivantes :

- *Client* : mise à jour des valeurs de variables d'entrée de discipline
- *JDM* : récupère la base de données visée, le nom de la variable à mettre à jour, la

nouvelle valeur de la variable, et modifie la variable identifiée

- *Client* : exécution d'une discipline
 - *JDM* : récupère les valeurs des variables d'entrée appartenant à la discipline, exécute la discipline et met à jour les nouvelles valeurs des variables dans la base de données associée à la discipline
- *Client* : récupération de valeurs de variables de sortie
 - *JDM* : récupère la base de données visée, le nom de la variable dont la valeur est demandée, et renvoie la valeur de la variable inscrite

Les étapes de mise à jour de variables et de récupération de valeurs nécessitent autant de messages à envoyer qu'il y a de variables (respectivement en entrée et en sortie). Le calcul d'une fonction coût à partir des variables de sortie incombe ensuite au client (l'optimiseur Isight). Cette procédure de communication étant basée sur un protocole, tout client se conformant au même format d'envoi de messages est en mesure de recevoir les mêmes réponses. Une communication a ainsi pu être établie entre le JDM et plusieurs routines Matlab pour les analyses.

Prado (2005) avait identifié que ces échanges de messages constituaient le facteur limitant en termes de temps de calcul durant une optimisation dans Isight. Toutefois dans le cadre de ce projet de maîtrise, il a été décidé de ne pas se concentrer sur les aspects purement informatiques, le même protocole a donc été conservé, en raison de sa simplicité. L'intégration d'un nouveau composant d'optimisation dans le système fut ainsi facilitée. Des fonctionnalités ont été ajoutées avec l'aide d'étudiants en génie informatique. Le protocole de communication fait intervenir d'autres messages pour effectuer des sauvegardes de base de données, des mises à jour de fichier Excel, pour être en mesure de visualiser les résultats sous une forme graphique, et même de générer une esquisse au format CATIA V5. Ces données sont ensuite destinées à être utilisées comme canevas durant la phase de conception détaillée.

1.2.2.3 Expérience des projets passés

Considérant la taille du problème (13 disciplines et plus de 500 variables), il apparaît nécessaire d'établir une stratégie d'optimisation basée sur la connaissance des caractéristiques du problème pour tenter de réduire le temps de calcul. La méthodologie, bâtie sur l'expérience des projets passés, comportera certains éléments dont la pertinence a été démontrée :

- une méthode de décomposition à deux niveaux
- un algorithme à gradients pour l'optimisation

Partant de ces éléments éprouvés, le but sera de formuler et de tester une stratégie permettant de réaliser une optimisation, tout en limitant le temps de calcul.

1.2.3 Simplification du problème

Pour simplifier le problème et limiter l'espace de design à explorer, plusieurs paramètres ont été fixés. En effet, la consistance des valeurs des variables du cycle thermodynamique est assurée par un code déterminant des valeurs compatibles. Pour simplifier la tâche de l'optimiseur, ce cycle thermodynamique a été fixé. Ceci implique qu'un certain nombre de variables d'entrée sont fixées, incluant entre autres la pression et la température à l'entrée de la veine gazeuse.

De plus, pour unifier les types de variables à prendre en compte et avoir accès à un plus grand choix d'algorithmes d'optimisations, les nombres entiers ne sont pas pris en compte. En effet, ces nombres qualifient des caractéristiques très particulières d'une turbomachine. Certains nombres entiers qualifient la présence ou non de certains composants, par leur valeur unitaire ou nulle. D'autres traduisent le nombre d'étages du compresseur ou de la turbine. Ces valeurs déterminent, par leur modification un changement potentiellement radical dans le comportement de la turbomachine : présence ou absence d'un compresseur centrifuge, d'une soufflante qui transforme alors un turbo-

réacteur en turbosoufflante. De plus, ces paramètres ne sont pas, en général, considérés comme des variables d'optimisation : ces paramètres sont fixés au préalable par les ingénieurs, d'après leur expertise. En effet, l'interface graphique, validée par les ingénieurs de P&WC, requiert des données de configuration de base sur le moteur tels que le nombre d'étages, de la présence d'un étage centrifuge au compresseur, ou à la turbine. Ces choix permettent de simplifier la nature du problème. Ceci réduit le nombre de variables qui influencent les modifications majeures de comportement physique de turbomachine, et évite ainsi les plus fortes discontinuités qui pourraient advenir durant le processus d'optimisation. Cela n'entraîne pas de perte de généralisation du problème d'optimisation : si plusieurs optimisations sont menées pour chaque combinaison de nombres entiers, la comparaison est alors aisée entre les solutions convergées.

Les figures 1.2 et 1.3 montrent respectivement le MDA avant et après la simplification. Les simplifications sont mineures. En effet, même si 15% des variables locales et 75% des variables partagées sont éliminées, seuls deux liens de couplage sont retirés. Le nombre de variables simplifiées peut paraître élevé (42 variables sont fixées), mais est peu élevé comparativement au nombre total de variables du problème simplifié (10%), et représente une faible simplification du problème d'optimisation.

1.2.4 Cadre théorique

Cette section présente les éléments du problème pour les phases d'intégration et d'optimisation du projet de maîtrise.

1.2.4.1 Intégration

La phase d'intégration a permis d'exploiter les outils de conception de P&WC dans le prototype. Cette étape a consisté à définir au moyen de fichiers de configuration, les

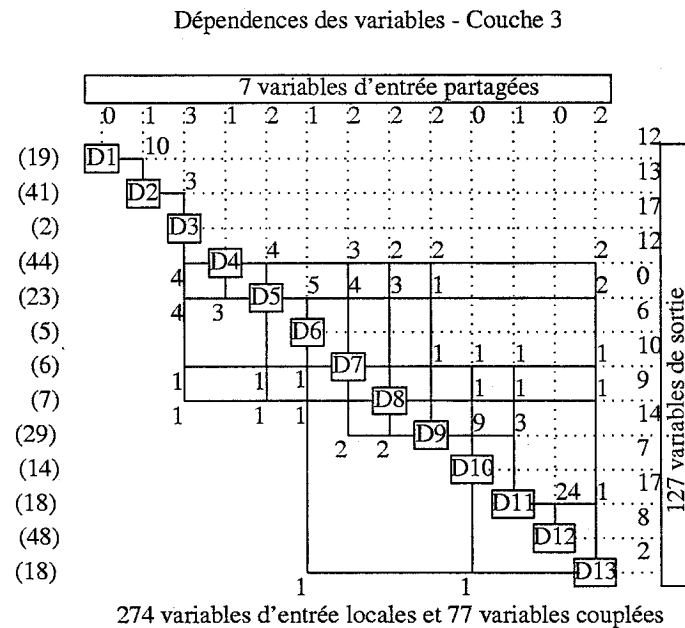


FIG. 1.3 MDA après simplifications

caractéristiques des variables utilisées dans les disciplines : nom, valeur, bornes de variation, type, localisation dans la feuille de calcul, etc.

L'intérêt théorique de cette intégration étant limité, les modalités de cette phase figurent dans la première annexe.

Les disciplines, bien que définies par des formules analytiques, sont perçues comme des boîtes noires dont les variables d'entrée (indépendantes) peuvent être modifiées pour observer les effets sur les variables de sortie (dépendantes). Les analyses se déroulent donc sans avoir accès au contenu des disciplines, mais seulement aux informations entrantes et sortantes. En effet, l'architecture a été développée pour utiliser des codes de calculs plus complexes tels que des analyses par éléments finis et non pour gérer les calculs de fonctions analytiques. Il est donc raisonnable de penser qu'une intégration de formules analytiques monobloc (une boîte noire unique), mènerait à un outil de conception plus performant, au détriment de la flexibilité de l'outil. En effet, une approche monobloc rend difficile l'ajout ou la suppression de discipline et l'intégration d'analyses variées.

Avec l'architecture proposée, les disciplines sont interreliées à travers les variables de couplage : ce sont les variables de sortie d'une discipline et d'entrée à une discipline tierce. Ces couplages reflètent les dépendances relatives des disciplines et rendent une optimisation du système complexe. En effet, l'objectif d'optimisation nécessite de modifier les valeurs des variables pour faire converger des itérations. Mais ces changements de valeurs de variables peuvent affecter la résolution du couplage, ou éloigner l'optimisation courante d'une région prometteuse.

Par ailleurs, des couplages existent entre les variables d'entrée d'une même discipline. Ces couplages internes sont résolus lorsque certaines des variables dépendantes, représentant une forme de résidus ont convergé vers zéro. La résolution de ces couplages ne peut être gérée autrement que par essai-erreur, et prise en compte sous la forme d'un terme de pénalité dans la fonction coût de la discipline. En effet, ces résidus sont implantés dans les feuilles de calcul, et ne peuvent pas être modifiés.

1.2.4.2 Optimisation

Partant de cette architecture fonctionnelle et une fois les disciplines disponibles, le projet consiste à exploiter les outils d'analyse actuels de P&WC pour parvenir à exécuter des optimisations sur des turbomachines. Prado (2005) a effectué la validation du prototype sur des problèmes simples comportant 24 variables, où le couplage interdisciplinaire est faible, et l'optimum connu. Le nouveau problème proposé par P&WC comporte plus de 500 variables, ainsi qu'un couplage plus complexe. Dans le cadre de ce projet de maîtrise, une attention particulière sera apportée pour valider le prototype dans le cadre d'une application pratique avec des contraintes de temps pour l'obtention du résultat.

Dans cette optique, deux axes de recherches ont un potentiel de réduction de la charge d'optimisation. Le premier consistera à classer les variables selon leur importance et d'utiliser cette classification lors de l'optimisation. Le deuxième axe de recherche consiste à étudier le MDA pour des simplifications potentielles et faciliter l'optimisation.

CHAPITRE 2

REVUE DE LITTÉRATURE

L'exploration de la littérature est organisée comme suit : la première section présentera des stratégies générales d'optimisation, et permettra d'en choisir une pour les besoins du projet. La section suivante présentera une revue détaillée de méthodes pour analyser les contributions des variables. La méthode choisie sera détaillée au chapitre 3. La troisième section exposera des outils pour traiter les liens, ou couplages, entre les disciplines et la technique choisie sera détaillée au chapitre 4. Enfin, la dernière section présentera le choix d'algorithme d'optimisation parmi les méthodes disponibles.

2.1 Stratégies générales d'optimisation

L'optimisation multidisciplinaire (MDO) fournit des méthodologies pour traiter les interactions entre plusieurs disciplines. Dans un tel cadre, le concepteur expérimenté dispose de plus de liberté pour influencer sur plus d'un domaine de spécialité à la fois. Le coût en terme de temps de calcul pour résoudre un problème de deux disciplines couplées n'étant pas la simple somme de la durée de résolution séquentielle des deux problèmes, la méthode utilisée va conditionner à la fois la durée du calcul et la qualité du résultat (le but étant d'aboutir au même résultat dans un temps plus court, ou à une meilleure optimisation dans le même temps). Ceci est dû au caractère fortement non linéaire des interactions entre les disciplines.

Sobieski & Kroo (1996) distinguent ainsi trois approches de recherche pour construire et développer les problèmes de MDO :

- Des problèmes comportant deux ou trois disciplines, où des analystes peuvent ac-

quérir une bonne connaissance dans ces différents domaines, permettent de développer une expertise sur leurs interactions. Une discipline englobant les autres est ainsi créée comme une nouvelle discipline à part entière. La prise en compte de ces interactions peut alors être intégrée dans des optimisations multidisciplinaires. En effet, dans la construction de systèmes complexes, des composants simples sont souvent intégrés peu à peu. L'aéroélasticité est un exemple de discipline créée pour prendre en compte l'interaction entre les disciplines aérodynamique et structurelle. Un système basé sur trois modèles simples d'aérodynamique, de structure et aéroélastique permet de prendre en compte des effets d'interaction, habituellement accessibles seulement au travers de deux disciplines aérodynamique et structure plus complexes et de temps de calcul beaucoup plus long. La reconfigurabilité et la flexibilité de ces systèmes est justement due à une approche très modulaire. Cela permet aussi de complexifier le processus d'optimisation par étapes.

- Une autre approche exploite des outils analytiques volontairement simples, car plus simples à intégrer dans une boucle d'analyse. Le but de cette approche est de parvenir à une modélisation simplifiée mais complète du produit ou du phénomène que l'on veut étudier. Les étapes suivantes du processus vont consister à raffiner progressivement les outils disponibles. Ceci simplifie les problèmes d'organisation de données et de circulation d'information. PMDO intègre ainsi des feuilles de calcul basées sur des formules analytiques simples, mais qui permettent de représenter les phénomènes physiques pour l'intégralité de la turbomachine. Dans le futur, ces disciplines simplifiées sont destinées à être remplacées par des calculs par éléments finis utilisant ANSYS ou FLUENT pour les calculs de structure et les calculs d'aérodynamique.
- La dernière approche consiste à gérer l'augmentation de complexité des systèmes MDO. Autant du point de vue matériel que du point de vue logiciel, les techniques proposent des moyens plus efficaces d'effectuer des optimisations sur des disciplines ayant des liens toujours plus complexes et de stocker des données en augmentation exponentielle. Des modèles et des techniques d'approximations sont de plus en plus

exploités pour réduire le coût lié aux analyses. En effet, un calcul par éléments finis reste long et coûteux, malgré les performances du matériel informatique récent.

Ces approches sont complémentaires pour mener à bien le développement de systèmes MDO. Le cas de PMDO illustre bien cette complémentarité. En effet, M. Prado a développé une plateforme permettant d'utiliser les outils de Pratt & Whitney Canada, pendant que les ingénieurs du département de Recherche et Développement ont développé plusieurs disciplines se basant sur des formules analytiques et des modèles simples, appelés à évoluer par la suite.

Pour la résolution des problèmes d'optimisation, Sobieski & Kroo (1996) recommandent l'utilisation de modèles simplifiés pour l'optimisation comparativement aux modèles exploités pour l'analyse. En effet, la recherche d'une solution optimale précise n'est pas forcément l'objectif premier. Des approximations suffisent pour traduire les interactions liant les disciplines, et prendre en compte leurs influences réciproques. Des modèles de surfaces de réponses ou des réseaux de neurones peuvent être interprétés comme des approximations globales. Des comportements généraux des variables découlent par déductions sur les pondérations des modèles. Suivant une idée très similaire, Egorov et al. (1999) proposent d'utiliser des outils d'analyse de précision variable pour résoudre des problèmes multidisciplinaires : l'objectif revient à représenter le phénomène physique avec un nombre de variables et de relations d'interactions réduites, sans nuire à la représentation. Une fois la convergence atteinte, le remplacement des outils simplifiés par les disciplines utilisant des formulations plus élaborées est effectué.

Par ailleurs, le coût en terme de calculs dépend principalement des couplages existants entre les disciplines, et à plus forte raison de l'intensité de ces couplages. Or les stratégies de décomposition de problèmes d'optimisation impliquent des techniques radicalement différentes pour gérer la résolution des couplages. Reyer & Papalambros (1999) distinguent deux principaux types de stratégies :

- Les stratégies séquentielles : Une méthode de type “passage simple” se contente d’optimiser les disciplines hiérarchiquement. Les disciplines subissent alors les choix des optimisations sur les analyses précédentes. Un mécanisme itératif permet de revenir à la première discipline optimisée pour boucler le système, et tenter d’améliorer les performances globales en faisant remonter de l’information. La résolution d’une équation non linéaire par une méthode itérative en est un bon exemple. Un résultat pertinent est rarement déterminé en une seule étape. En général, la résolution converge lorsque le résultat d’une itération est exploité pour entreprendre l’itération suivante.
- Les stratégies concurrentes permettent de gérer toutes les variables communes en même temps. Des méthodes (All-At-Once, FIO) s’affranchissent des problèmes de dépendance, en ne considérant pas la segmentation du problème. D’autres méthodes (DAO, CO, DO) brisent les couplages afin de créer des sous-tâches, et une tâche globale pour gérer la reconstruction des couplages via des contraintes de consistance. Un exemple de problème d’optimisation décomposé suivant la formulation DO est donné dans l’annexe III.

Reyer & Papalambros (2000) montrent numériquement et analytiquement la supériorité des techniques concurrentes en terme de précision et de temps de calcul. Sans surprise, la méthode séquentielle qui se contente d’un seul passage pour optimiser successivement toutes les disciplines est la pire. Toutefois, ce n’est pas la méthode qui segmente le plus les tâches d’optimisation ou d’analyse qui donne les meilleurs résultats. En effet, pour certains problèmes, une méthode de type FIO semble plus efficace qu’une méthode multi-niveaux. Papalambros & Michelena (2001) sont conscients de la limitation apparente des techniques bi-niveaux, et explique : *“les stratégies de décomposition entraînent presque toujours une augmentation du temps de calcul, en raison de la gestion plus complexe, ou du nombre global de variables plus élevé, en comparaison à une stratégie non décomposée”*. Toutefois, des améliorations de plusieurs types peuvent avantager certaines méthodes dépendamment du problème à résoudre. Ainsi, une décomposition qui suspend certains paramètres peut se montrer plus efficace si les variables fixées le sont à

une valeur proche de l'optimum. Turcotte (2004) a montré qu'une méthode de décomposition (DAO) peut se montrer supérieure à une méthode globale (FIO), si les calculs des sous-tâches (analyses, optimisation ou calculs de gradients) sont parallélisés. Il convient de considérer le temps d'exécution des analyses par rapport au temps consommé pour la gestion des données : les stratégies bi-niveaux (Prado, 2005) consomment, en effet, moins d'évaluations de fonctions que les méthodes globales (i.e., méthodes n'exploitant pas de décomposition en sous-tâches), notamment en raison du temps dépensé pour la résolution du couplage. En effet, les méthodes de décomposition gèrent les couplages comme des contraintes supplémentaires dans le problème d'optimisation, tandis que les stratégies globales ont recours à des méthodes itératives de résolution du couplage.

Cette revue des stratégies globales d'optimisation a fait ressortir trois aspects importants sur les boucles d'analyse multidisciplinaires :

- La décomposition du problème et sa représentation au travers de modèles plus ou moins simplifiée au fil de la convergence partagera la charge d'optimisation entre plusieurs optimiseurs,
- L'analyse des variables au sein des disciplines distinguera les variables importantes du bruit,
- L'étude des liens qui lient les disciplines par les variables de couplage simplifiera l'analyse multidisciplinaire.

Des stratégies de décomposition (FIO, DAO, DO) ont déjà été comparées dans des travaux précédents. En effet, Prado (2005) avait montré que les stratégies de décomposition étaient plus efficaces pour résoudre un problème d'optimisation. La décomposition DAO¹ avait été démontrée plus efficace que la décomposition DO. Le choix s'est toutefois orienté vers la décomposition DO, en raison du nombre beaucoup plus élevé de

¹La méthode de décomposition DAO est similaire à la décomposition DO, détaillée dans l'annexe III : DAO n'a pas de sous-problèmes d'optimisation, seulement des exécutions et toutes les variables sont gérées par le problème maître

variables à gérer que par rapport au cas test sur lequel ces conclusions sont fondées (plus de 500 contre 24). En effet, la décomposition DO segmente le problème d'optimisation en sous-problèmes afin d'alléger la tâche de l'optimiseur maître. Cette méthode autorise une recherche au niveau des disciplines potentiellement plus créative en raison de contraintes non imposées (Tribes & Trépanier, 2006). Le fonctionnement de la décomposition DO est présenté dans l'annexe III. Les recherches proposées porteront donc sur le traitement du problème en exploitant les deux autres aspects : l'analyse des variables et l'étude des couplages.

2.2 Importance des variables et leur classification

La littérature recèle de nombreuses techniques pour résoudre le problème dit de "Sélection de Variables". Ces techniques reviennent bien souvent à établir un classement basé sur une analyse de sensibilité. Le but de cette section n'est pas d'expliquer toutes ces techniques, mais de présenter sommairement les plus intéressantes du point de vue du problème actuel.

2.2.1 Modèle de surface de réponse

Une des méthodes les plus répandues est le modèle de surface de réponse (RSM) (Resit et al., 1998). Dans sa forme la plus simplifiée, un modèle RSM pondère linéairement les variables du système à étudier. Les coefficients sont alors déterminés par une méthode de moindres carrés. Pour prendre en compte des dépendances plus complexes entre les variables, le modèle devient une somme linéaire de fonctions dont les formulations découlent d'hypothèses. Cette méthode est simple, rapide et pourvue d'un estimateur d'erreur. Cette estimation est déterminée à partir des points connus d'une base de données et non à partir de la formulation de la fonction réelle. L'erreur est une mesure de la propor-

tion de variation des variables dépendantes non expliquée par le modèle ; elle permet de juger la qualité de l'approximation.

Pour illustrer cette méthode, le volume Y d'un parallépipède rectangle de dimensions X_1, X_2, X_3 , percé d'un trou borgne de dimensions X_4, X_5, X_6 d'un ordre de grandeur plus petits que les valeurs principales peut être optimisé en tenant compte d'un volume minimal. Le modèle RSM sera de la forme : $Y = a \cdot X_1 + b \cdot X_2 + c \cdot X_3 + d \cdot X_4 + e \cdot X_5 + f \cdot X_6$. Des coefficients élevés seront probablement associés aux variables X_1, X_2, X_3 , puisque ces variables caractérisent le corps principal, tandis que des coefficients plus modestes seront associés aux variables caractérisant les dimensions du trou.

Une technique de sélection de variables peut ainsi être obtenue en joignant ces méthodes statistiques à des méthodes d'échantillonnage (plans d'expériences "full-factorial", "fractional factorial" (Montgomery, 2000), de Taguchi (cité par Unal & Dean (1991)), ou d'autres techniques plus ou moins élaborées présentées par Buchsbaum et al. (2003)) et une analyse de variance (ANOVA, voir Montgomery (2000) et Trocine & Malone (2000)). Cette stratégie n'est pas toujours efficace sur des problèmes comportant de nombreuses variables. En effet, le nombre d'évaluations requises par les plans d'expériences classiques augmente de manière exponentielle avec le nombre de variables. De plus, lorsque le problème est sous-déterminé, les plans d'expérience classiques peuvent mener à de mauvaises interprétations des résultats : l'influence de plusieurs variables en interaction est perçue au travers de la variation d'une unique variable. Ces effets confondus ne permettent que difficilement de déterminer laquelle des variables du groupe prime sur les autres. Toutefois, des techniques d'échantillonnage comme l'hypercube latin (McKay et al., 2000) se révèlent plus efficaces en répartissant uniformément les points dans l'espace de conception.

2.2.2 Filtrage par groupe ou “group screening”

Les méthodes de filtrage par groupe (Trocine & Malone, 2001, 2000, Kleijnen et al., 2003, Wan et al., 2003, Chipman & Winker, 2000) (plus connues en anglais sous l'appellation “group screening methods”) aident à déterminer les variables les plus importantes d'une boucle d'analyse, en utilisant un nombre d'évaluations inférieur au nombre de variables. La procédure générale est très similaire à la construction d'un modèle statistique, à la différence près que ces méthodes de filtrage associent un groupe de variables à une pondération : l'influence d'un poids ne correspond plus à une seule variable. Ces méthodes de filtrage s'avèrent efficaces et peuvent être appliquées lorsque certaines hypothèses sont vérifiées. Ces hypothèses cherchent à garantir l'effet du groupe de variables sur la fonction coût :

- Les effets de variables indépendantes dans un groupe sur la fonction coût ou sur les variables dépendantes doivent être de même signe.
- Les effets d'interaction ne doivent pas se compenser à l'intérieur d'un groupe.
- La fonction coût doit montrer une variation monotone par rapport à toutes les variables d'entrée.

Sur l'exemple du parallélépipède troué, deux groupes pourraient être constitués : X_1, X_2, X_3 et X_4, X_5, X_6 , pour bâtir un modèle du type $Y = a \cdot x_1 + b \cdot x_2$. Moins de points d'échantillonnage sont alors requis pour déterminer le groupe de variables influentes. À titre d'illustration, Trocine & Malone (2001), utilisant une méthode de bifurcation séquentielle (“Sequential Bifurcation”), présentent un modèle écologique stochastique de 281 variables, où 15 effets principaux sont déterminés à partir de données de 144 points d'analyse. Campolongo, Kleijnen & Andres (1997), cités par Trocine & Malone (2000), sont parvenus, par une technique particulière (Iterated Fractional Factorial Design, IFFD), à déterminer 18 effets principaux sur un problème de 3000 variables. Cette dernière méthode requiert toutefois l'hypothèse supplémentaire que très peu de facteurs dominant.

2.2.3 Réseaux de neurones

Un réseau de neurones est composé d'entités (neurones) reliés ensemble. Chacun des neurones représente un type de fonction : transformation affine, incrémentation, parabolique, exponentielle, etc. Des valeurs pour les variables d'entrée sont fournies à ces neurones qui transforment la valeur suivant la fonction qu'ils représentent. En sortie, une somme pondérée des calculs effectués par les neurones est récupérée. Le but de la phase d'apprentissage est de dimensionner ces pondérations afin que les valeurs renvoyées par le réseau de neurones soit consistant avec les valeurs des variables de sortie calculées d'après le modèle réel. Une introduction des concepts des réseaux de neurones est présentée dans Diedrich (1990).

Une fois l'apprentissage achevé, une analyse de sensibilité peut être effectuée en exploitant le réseau. Souvent, les pondérations du réseau ne sont pas accessibles aisément. La mesure de la différence sur les variables de sortie lorsqu'une variable d'entrée est successivement soustraite, puis réintroduite dans le réseau permet de juger de son influence. Si son influence est forte, la soustraction de la variable dans le système aura pour effet de modifier les valeurs des variables de sortie, par rapport au cas où la variable est incluse. Un critère sur la modification acceptable des variables de sortie permet alors de classer les variables d'entrée testées les unes après les autres.

Pour reprendre l'exemple sur le parallélépipède, l'opération consiste à entraîner un réseau composé d'au moins 6 neurones (nombre de variables) qui pourraient contribuer linéairement pour déterminer les pondérations dans un modèle de la forme $Y = a \cdot g(X_1) + b \cdot g(X_2) + c \cdot g(X_3) + d \cdot g(X_4) + e \cdot g(X_5) + f \cdot g(X_6)$, où g est une fonction linéaire ($g : x \mapsto x$). L'analyse de sensibilité consisterait ensuite à tester la pertinence de modèles dérivés (i.e. $Y = a \cdot g(X_1) + b \cdot g(X_2) + c \cdot g(X_3) + d \cdot g(X_4) + e \cdot g(X_5)$, ou $Y = b \cdot g(X_2) + c \cdot g(X_3) + d \cdot g(X_4) + e \cdot g(X_5) + f \cdot g(X_6)$), pour déterminer quelles variables entraînent les plus fortes augmentations d'erreur entre le modèle dérivé et le modèle initial.

2.2.4 Analyse par Composantes Principales

Selon la définition proposée par Wikipedia ², “l’Analyse en Composantes Principales (ACP) est une méthode statistique [...] qui a pour but de comprendre et de visualiser comment les effets de phénomènes a priori isolés se combinent. [...] En dimension plus grande que deux, une ACP va toujours déterminer les axes [...], qui expliquent le mieux la dispersion du nuage des points disponibles”. C’est la transformation linéaire optimale qui permet de déterminer l’hyperplan, ou les sous-espaces de plus forte variance. L’axe déterminé par cette analyse donne la meilleure corrélation entre les variables indépendantes. Le vecteur colinéaire à la direction de l’axe est une combinaison linéaire des vecteurs de base. Or ces vecteurs de base représentent les axes de variation des variables indépendantes. À partir de cette représentation, un classement des variables peut être déterminé, selon la pondération dans les coordonnées du vecteur de la direction principale. Une description plus approfondie de cette technique est donnée dans Aluja-Banet (1998). Toutefois, certaines caractéristiques limitent l’utilisation de cette méthode. Plusieurs directions principales peuvent être déterminées par l’analyse. Or chaque direction étant unique, la combinaison linéaire des vecteurs de base sera donc distincte, et le classement des variables indépendantes sera alors modifié. Une moyenne peut être utilisée, mais une fonction de pondération doit être incluse afin que des directions associées à une variance plus faible n’aient pas une influence trop forte sur des directions plus importantes. L’analyse par composantes principales appliqué au cas du parallélépipède utiliserait un échantillonnage des variables d’entrée $X_1, X_2, X_3, X_4, X_5, X_6$ pour déterminer les valeurs du volume Y . L’analyse de ces résultats donnerait un vecteur probablement orienté suivant la direction $\vec{D}$ de la forme :

$$\vec{D} = a \cdot \vec{X}_1 + b \cdot \vec{X}_2 + c \cdot \vec{X}_3 + d \cdot \vec{X}_4 + e \cdot \vec{X}_5 + f \cdot \vec{X}_6$$

²http://fr.wikipedia.org/wiki/Analyse_en_composantes_principales

Les variables les plus influentes peuvent alors être identifiées par une analyse des coefficients de pondération a, b, c, d, e, f des directions de base $\vec{X}_1, \vec{X}_2, \vec{X}_3, \vec{X}_4, \vec{X}_5, \vec{X}_6$

2.2.5 Algorithmes génétiques

Dans une application bio-informatique (Jarvis & Goodacre, 2005), un algorithme génétique est exploité afin de rechercher, parmi des sous-ensembles de variables, la meilleure combinaison. Chacun de ces sous-ensembles traduit un arrangement des variables d'entrée en modèle mathématique. La fonction coût est l'erreur de représentation du modèle ainsi généré. Le meilleur modèle donne alors le meilleur sous-ensemble de variables et les pondérations associées qui garantissent une minimisation de l'erreur. Ces poids permettent d'établir un classement des variables.

Toutefois, au moins deux problèmes menacent la pertinence de cette méthode. Un objectif d'optimisation est donné, sans qu'aucune garantie ne soit possible sur l'existence d'une solution. En effet, si le nombre de variables réellement pertinentes est plus élevé que la dimension de l'espace sur lequel les sous-ensembles travaillent, alors l'erreur du modèle risque fortement d'être toujours plus élevée que l'objectif d'optimisation. Par ailleurs, dans le cas inverse (nombre de variables réellement pertinentes moins élevé que la dimension du sous-ensemble), une variable peut avoir été sélectionnée alors que son influence sur le modèle est minime (pondération très faible par rapport aux poids des autres variables).

L'application de cette méthode sur le cas du parallélépipède se concrétiserait par la construction de sous ensembles des variables initiales telles que $\{X_1, X_2, X_3, X_4\}$, ou $\{X_1, X_2, X_5, X_6\}$. La détermination de coefficients pour construire ensuite un modèle de surface de réponse permet ensuite de caractériser la pertinence d'approximation (avec respectivement des modèles de la forme $Y = a \cdot X_1 + b \cdot X_2 + c \cdot X_3 + d \cdot X_4$ ou $Y = a \cdot X_1 + b \cdot X_2 + e \cdot X_5 + f \cdot X_6$). Le sous-ensemble de variable donnant l'erreur la plus faible peut ensuite être considéré comme l'ensemble de variables influentes

recherché.

2.2.6 Exploitation de la formulation analytique des fonctions coûts

Azarm a présenté une méthodologie de réduction pour traiter les problèmes d'optimisation multi-objectifs. Ses exemples (voir Azarm & Eschenauer (1993) et Azarm (1994)) montrent une réduction du nombre de variables basée sur la formulation analytique de la fonction coût. Les variables locales impliquées dans des contraintes et non dans les fonctions coût, ont une valeur fixe pendant l'optimisation des variables globales. Ce principe ne pourrait pas être appliqué sur une boîte noire, puisqu'il est a priori impossible de déterminer si les variables locales influençant les contraintes influent aussi sur la fonction coût. Même pour les problèmes d'optimisation qui exploitent les formulations analytiques des disciplines, cette technique ne peut garantir aucune réduction du nombre de variables. Toutefois, cette réduction du nombre de variables n'était qu'un corollaire de ces articles. De plus, cette méthode est également fondée sur une classification préalable de variables primaires (influence globale), et secondaires (influence locale). Pour appliquer cette méthode, une telle information sur l'influence doit être supposée ou connue. Il est délicat d'appliquer cette méthode sur le problème du parallélépipède troué. En effet, en toute rigueur, cette méthode ne permet aucune simplification. Mais si une contrainte était ajoutée sur les variables déterminantes des dimensions du trou (X_4, X_5, X_6) l'optimisation pourrait être menée seulement sur les variables X_1, X_2, X_3 .

2.2.7 Bilan sur les méthodes de sélection de variables

Beaucoup de méthodes existent et certaines combinaisons peuvent servir à réaliser une analyse de sensibilité. Chaque technique a ses avantages et ses inconvénients. Ainsi, le choix d'une méthode dépend principalement de l'accessibilité des données (utilisation

des seules données existantes ou possibilités de calculs de nouveaux points), la taille du problème (les méthodes avec apprentissage nécessitent plus de données, qui doivent être générées, et ceci peut résulter en un temps de calcul rallongé) et l'information disponible sur le problème. Certaines techniques montrent des performances séduisantes, en contrepartie de certaines connaissances préalables sur le système. Cette expertise peut être supposée puis validée postérieurement au processus d'analyse. Toutefois, il est préférable de privilégier des techniques qui ne se basent pas sur des hypothèses trop fortes, surtout dans le cas des techniques de filtrage par groupe.

Par définition, peu ou pas d'information est disponible sur le problème initial. La plateforme configurable de PMDO en fait un système qui peut proposer, à partir de 13 disciplines intégrées, plus de 60000 variations du problème d'optimisation de turbomachine. Toute collecte de données afin d'entraîner un réseau de neurones peut s'avérer très longue pour obtenir une bonne corrélation entre la modélisation et le résultat de simulation. L'analyse multidisciplinaire comporte plus de 500 variables en considérant le système global, ce qui est beaucoup trop élevé pour utiliser les plans d'expériences classiques. Les méthodes candidates restantes peuvent réaliser une sélection non pertinente des variables (utilisation d'algorithmes génétiques et analyse par composantes principales). Pour corriger cette sélection, des traitements peuvent être appliqués, pour vérifier l'influence des variables en analysant les coefficients de pondération. Mais les spécifications sur de nouveaux seuils de sélection n'ont pas de signification particulière pour être réglés génériquement.

La méthode proposée dans le cadre de ce travail est basée sur une utilisation des statistiques, en exploitant des surfaces de réponses limitées à des portions du problème : les variables propres aux disciplines. Réduisant les modèles de surface de réponses aux variables d'entrée et de sortie d'une discipline à la fois plutôt que celles de toutes les disciplines simultanément, les nombres de variables sont réduits de façon significative. Ensuite, des techniques d'échantillonnage tels que l'hypercube latin permettent une distribution uniforme des valeurs des variables d'entrée dans l'ensemble de l'espace de

design. Par ailleurs, les paramètres à configurer ont une signification reliée aux données : leur valeur peut être calculée par rapport à la répartition des données en sortie. De plus, l'estimateur d'erreur de cette méthode permet de juger facilement de sa pertinence. Le chapitre 2 sur l'analyse de sensibilité des variables illustre comment cette technique est utilisée pour classer un ensemble de variables, dans le cas de PMDO. La finalité de cette opération est également replacée dans le contexte du projet de maîtrise et dans le processus global d'optimisation.

2.3 Étude des couplages interdisciplinaires

La gestion des couplages interdisciplinaires est reliée de manière plus ou moins directe au choix de la méthode de décomposition. Ces couplages sont parmi les éléments les plus délicats à gérer. En effet, une résolution séquentielle de deux disciplines couplées ne permet pas de résoudre le couplage et, en général, ne donne pas le même optimum si le problème A est résolu préalablement au problème B, comparativement au cas où A est résolu postérieurement à B. Fathy et al. (2001) ont également montré que les optimisations des sous-problèmes doivent être simultanées et non résolues séquentiellement.

Dans les cas de décomposition à plusieurs niveaux, un optimiseur maître gère les variables globales, et exécute des sous-optimisations. Ensuite, un mécanisme gère les couplages : des contraintes de consistance³, par exemple, relient les variables des sous-optimisations à l'optimiseur maître.

Rogers (Rogers, 1996a,b, Rogers et al., 1996) a établi une méthodologie pour gérer ces couplages interdisciplinaires.

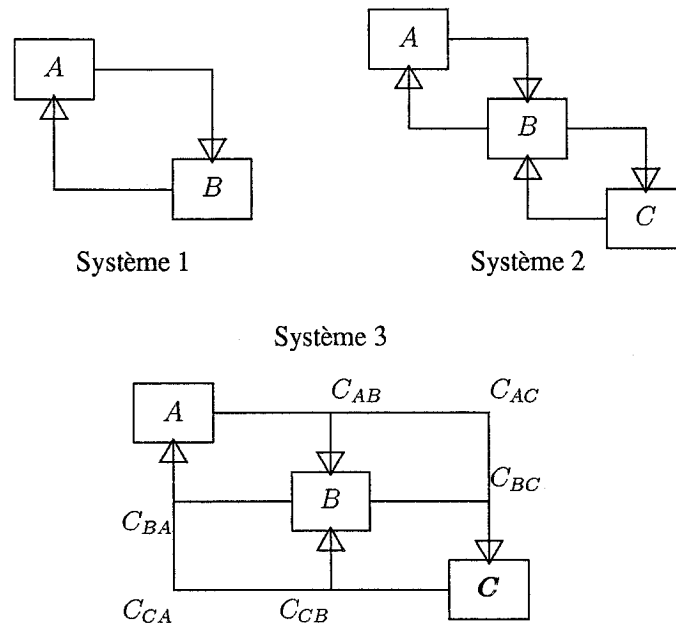


FIG. 2.1 Configurations possibles du MDA

2.3.1 Gestion des liens interdisciplinaires

Dans une résolution de problème d'optimisation, un des plus gros problèmes se présente lorsqu'un couplage lie deux disciplines dans les deux sens (Système 1 de la figure 2.1) : la discipline A donne une valeur à la discipline B, qui à son tour fournit une valeur à la discipline A, comme variable d'entrée. En effet, à chaque fois qu'un lien avant (de A vers B), se conjugue à un lien arrière (de B vers A), des itérations sont requises pour résoudre une inconsistance potentielle : A détermine des valeurs de sortie à partir de valeurs tests en entrée. Or B fournit une valeur à A en retour. Tant que cette valeur de retour n'est pas cohérente avec la valeur que A utilise, une nouvelle itération est nécessaire. Les boucles d'analyse présentent souvent plus de deux disciplines, et ainsi des liens comme ceux du système 2 apparaissent et la résolution du couplage A-B peut être perturbée par la résolution du couplage B-C. Dans une formulation générale, le système 3 montre une

³Ces contraintes assurent une résolution du couplage, et leur utilisation est présentée dans la méthode de décomposition DO, annexe III

notation qui permet de distinguer tous les liens de couplage et leur direction.

Dans une application pratique, Rogers & Bloebaum (1994) ont montré que les couplages les plus problématiques sont les liens de retour. Une mesure de l'intensité du couplage leur permet alors de suspendre certains liens faibles pour faciliter l'optimisation. La solution optimale alors obtenue est d'une précision variable, suivant les critères de suspension du couplage. De plus, la méthodologie que Rogers propose utilise une réorganisation des tâches afin que soient minimisées les boucles de retour, à travers une ou plusieurs disciplines.

2.3.2 Mesures du couplage

Les articles se rapportant aux problèmes de couplage exploitent souvent un paramètre variable permettant de modifier l'intensité du couplage et, dans certains cas, de changer la position de l'optimum, pour tester la robustesse de différents algorithmes. Toutefois, peu d'articles proposent une véritable mesure globale de l'intensité du couplage entre deux disciplines.

Les équations de sensibilités globales de Sobieszczanski-Sobieski (1990) proposent de déterminer la variation des variables de sortie Y (variables de couplage), par rapport à la variation de variables d'entrée X (non issues de couplages). L'approche de Rogers et al. (1996) se base sur le calcul d'un rapport de variations de deux variables couplées qui sont des variables de sortie de deux disciplines distinctes, mais couplées.

2.3.2.1 Équations de sensibilités globales ou “Global Sensitivity Equations” (GSE)

Cette méthode, présentée par Sobieszczanski-Sobieski (1990), permet de caractériser la sensibilité des couplages internes de systèmes complexes. En considérant l'analyse multidisciplinaire de la figure 2.2, trois disciplines (A , B , C) sont définies, et s'échangent des

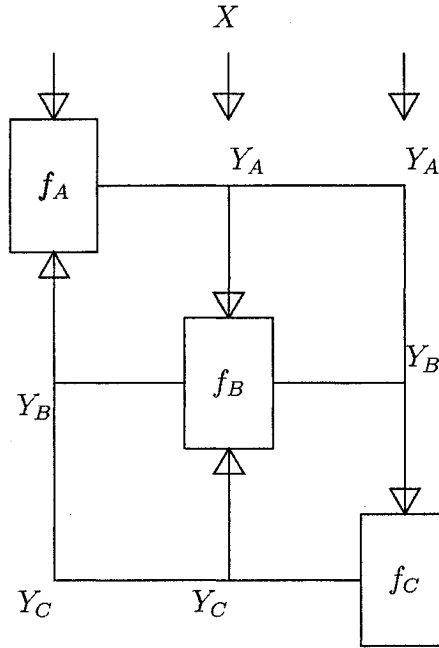


FIG. 2.2 MDA considéré pour la méthode GSE

informations sous la forme de variables de couplage (Y_A , Y_B , Y_C). Dans ce cas particulier, aucune distinction n'est faite pour différencier les variables de couplage de A vers B comparativement aux variables de A vers C . Une formulation plus générale existe, mais ce manque de différenciation ne gêne pas : cela permet de simplifier l'explication du principe.

Définissant la fonction associée à chaque discipline par f_x , où x est la discipline, et Y_x les variables couplées des disciplines x , l'expression des fonctions est :

$$\text{Discipline } A \rightarrow f_A(X, Y_B, Y_C), \text{ Variable couplée en sortie } Y_A \quad (2.1)$$

$$\text{Discipline } B \rightarrow f_B(X, Y_A, Y_C), \text{ Variable couplée en sortie } Y_B \quad (2.2)$$

$$\text{Discipline } C \rightarrow f_C(X, Y_A, Y_B), \text{ Variable couplée en sortie } Y_C \quad (2.3)$$

où X représente les variables partagées et locales. La résolution du couplage est achevée lorsque les résidus F_x sont nuls (contraintes de consistance résolues) :

$$F_A(Y, X) = f_A(X, Y_B, Y_C) - Y_A = 0 \quad (2.4)$$

$$F_B(Y, X) = f_B(X, Y_A, Y_C) - Y_B = 0 \quad (2.5)$$

$$F_C(Y, X) = f_C(X, Y_A, Y_B) - Y_C = 0 \quad (2.6)$$

Dans une forme plus générale, le système peut s'écrire $F(Y, X) = 0$. La sensibilité de ce système vérifie $\frac{dF}{dX} = 0$, d'où :

$$\left[\frac{\partial F}{\partial Y} \right] \left(\frac{\partial Y}{\partial X_k} \right) = - \left(\frac{\partial F}{\partial X_k} \right) \quad (2.7)$$

En réimplantant une forme linéarisée des fonctions f_x , on obtient :

$$\begin{bmatrix} I & -\frac{\partial f_A}{\partial Y_B} & -\frac{\partial f_A}{\partial Y_C} \\ -\frac{\partial f_B}{\partial Y_A} & I & -\frac{\partial f_B}{\partial Y_C} \\ -\frac{\partial f_C}{\partial Y_A} & -\frac{\partial f_C}{\partial Y_B} & I \end{bmatrix} \begin{pmatrix} \frac{\partial Y_A}{\partial X_k} \\ \frac{\partial Y_B}{\partial X_k} \\ \frac{\partial Y_C}{\partial X_k} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_A}{\partial X_k} \\ \frac{\partial f_B}{\partial X_k} \\ \frac{\partial f_C}{\partial X_k} \end{pmatrix} \quad (2.8)$$

Et dans le cas où Y_x est un vecteur représentant plusieurs (n) variables, une partie du système précédent deviendrait :

$$\begin{bmatrix} \cdot & \begin{bmatrix} -\frac{\partial f_{A1}}{\partial Y_{B1}} & \dots & -\frac{\partial f_{A1}}{\partial Y_{Bm}} \\ \vdots & \ddots & \vdots \\ -\frac{\partial f_{An}}{\partial Y_{B1}} & \dots & -\frac{\partial f_{An}}{\partial Y_{Bm}} \end{bmatrix} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{pmatrix} \left(\frac{\partial Y_{A1}}{\partial X_k} \right) \\ \vdots \\ \left(\frac{\partial Y_{An}}{\partial X_k} \right) \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial f_{A1}}{\partial X_k} \right) \\ \vdots \\ \left(\frac{\partial f_{An}}{\partial X_k} \right) \\ \cdot \\ \cdot \end{pmatrix} \quad (2.9)$$

Par des calculs de différences finies, toutes les dérivées partielles de f_x sont simples à

déterminer, même si le processus peut être long. Pour déterminer les termes de sensibilité recherchés $\partial Y_x / \partial X_k$, il reste à résoudre un nouveau système linéaire pour chaque variable de sortie X_k . La charge de calcul devient lourde lorsqu'il s'agit de considérer les différentes variables partagées et locales qui, dans le cas de PMDO, sont à un nombre proche de 300. Afin d'éviter ce coût prohibitif, il semble plus approprié de tenter de déterminer de façon plus directe les quantités $\partial Y_x / \partial X_k$.

2.3.2.2 Méthode des équations de sensibilité globale modifiée

Alyaqout et al. (2005) présentent une fonction de mesure du couplage, dont certains des termes sont issus d'une forme modifiée de la méthode précédente. Cette modification permet de prendre en compte la condition d'optimalité de Karush-Kuhn-Tucker de la fonction coût sur certaines variables. Cette mesure d'intensité Γ_x permet de quantifier un couplage, et est exploitée pour suspendre les variables les moins influentes dans une boucle d'analyse.

Considérons le problème décomposé (voir figure 2.3) en deux sous-optimisations, pourvues de fonctions à minimiser f_1 et f_2 , et dont les variables sont respectivement X_1, X_3 et X_2, X_3 . Le problème maître est doté d'une fonction f_3 , composée des fonctions des sous tâches. Les couplages visés pour la suspension sont les liens des disciplines 1 et 2 avec l'optimiseur maître (3). La démarche suit le processus suivant :

- Étape 0 : Les sous-optimisations (1 et 2) sont lancées avec X_3 fixé. Les solutions optimales des sous-optimisations 1 et 2 correspondent à X_1^* et X_2^* respectivement. On complète l'étape 0 en optimisant le problème maître avec X_1^* et X_2^* , ce qui donne X_3^* .
- Étape 1 : Une mesure du couplage maître-discipline 1 à travers la variable X_1 peut être déterminée à partir de la fonction définie par Alyaqout et al. Cette fonction correspond à une estimation de la sensibilité de la fonction coût optimale f_3^* par rapport à une variable suspendue (X_1 fixé dans l'exemple qui suit) et est obtenue par une dérivation

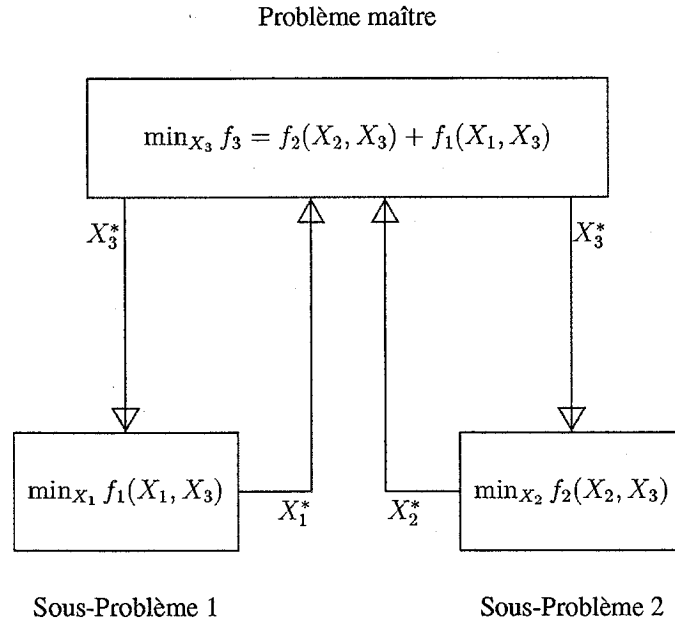


FIG. 2.3 Organisation des disciplines : problème maître, et deux sous-tâches

en chaîne de la fonction coût du problème maître :

$$\Gamma_1 = \frac{df_3^*(X_1)}{dX_1} = \frac{\partial f_3^*}{\partial X_1} + \frac{\partial f_3^*}{\partial X_2} \frac{d\hat{X}_2}{dX_1} + \frac{\partial f_3^*}{\partial X_3} \frac{d\hat{X}_3}{dX_1} \quad (2.10)$$

où $\hat{X}_2$, $\hat{X}_3$ sont les solutions de la sous-optimisation 2 et du problème maître pour X_1 fixé. $\hat{X}_2 = \hat{X}_2(X_1, \hat{X}_3)$ et $\hat{X}_3 = \hat{X}_3(X_1, \hat{X}_2)$ expriment la condition de stationnarité pour le problème maître et la sous-optimisation 2. Les dérivées partielles de f_3^* peuvent être évaluées localement par différences finies. Les dérivées de $\hat{X}_2$ et $\hat{X}_3$, par rapport à X_1 sont déterminées en résolvant un système modifié d'équations GSE :

$$\begin{bmatrix} I & -\frac{\partial \hat{X}_2}{\partial \hat{X}_3} \\ -\frac{\partial \hat{X}_3}{\partial \hat{X}_2} & I \end{bmatrix} \begin{bmatrix} \frac{d\hat{X}_2}{dX_1} \\ \frac{d\hat{X}_3}{dX_1} \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{X}_2}{\partial X_1} \\ \frac{\partial \hat{X}_3}{\partial X_1} \end{bmatrix} \quad (2.11)$$

où les dérivées partielles sont également évaluées localement. Un calcul similaire per-

met d'obtenir $\Gamma_2 = df_3^*(X_2)/dX_2$, lorsque la variable X_2 est fixée. Ces fonctions de couplage Γ_1 , Γ_2 représentent l'effet d'une perturbation δX_1 ou δX_2 sur la fonction coût optimale du problème maître. Selon que Γ_1 ou Γ_2 est faible, une variable locale peut être suspendue.

- Étape 2 : Un critère de comparaison des couplages est donc nécessaire pour transcrire sous forme mathématique le jugement de comparaison entre les Γ obtenus :

$$\|\Gamma_1\| < c\|\Gamma_2\| \rightarrow \text{Suspension de } X_1 \quad (2.12)$$

$$\|\Gamma_2\| < c\|\Gamma_1\| \rightarrow \text{Suspension de } X_2 \quad (2.13)$$

où c est un scalaire à définir empiriquement.

- Étape 3 : L'hypothèse que le calcul des dérivées de $\hat{X}_2$ et $\hat{X}_3$ par les GSE modifiées est réalisé proche de l'optimum de X_1 , doit être vérifiée.

En effet, si X_1 a été suspendu dans l'étape 2, une optimisation permet de déterminer de nouvelles valeurs pour X_2^* et X_3^* . À ces nouvelles valeurs correspondent une nouvelle valeur estimée pour X_1^* . En effet, une estimation de cet écart (δX_1) est disponible à partir des quantités calculées :

$$\frac{\delta X_2}{\delta X_1} = \frac{d\hat{X}_2}{dX_1} \text{ et } \frac{\delta X_3}{\delta X_1} = \frac{d\hat{X}_3}{dX_1} \quad (2.14)$$

Un critère sur l'écart de ce nouvel optimum par rapport à la valeur fixée précédente (δX_1 si X_1 fixé, δX_2 si X_2 fixé) permet alors de vérifier que les calculs de dérivées ont été réalisés dans un voisinage pertinent de l'optimum. Dans le cas contraire, la suspension de la variable est annulée et le processus reprend à l'étape 1 avec pour valeurs de départ X_2^* et X_3^* .

- Étape 4 : l'optimisation est terminée lorsque la variable X_3 a convergé : lorsque deux valeurs de X_3 obtenues successivement après avoir suivi le processus en entier sont suffisamment proches, suivant la précision désirée.

Cette méthode semble assez radicale : dans l'exemple fourni, une variable est fixée, et ceci suspend la discipline au complet, si le couplage ne remplit pas le critère de comparaison ($\|\Gamma_1\| < c\|\Gamma_2\|$). Par ailleurs, la stratégie se base sur une mesure du couplage dont les termes sont déterminés soit par différences finies soit par les équations de sensibilité globale modifiées. La mesure étant locale, ces calculs sont effectués plusieurs fois au cours de l'optimisation pour déterminer de nouvelles mesures du couplage à chaque itération. Dans un cas où le problème contient de nombreuses variables, le coût en calculs augmente rapidement. De plus, la proximité de l'optimum est une hypothèse forte. Toutefois, un mécanisme de vérification s'assure de la crédibilité de l'hypothèse, même si de nouveaux calculs sont requis.

Tout comme pour les GSE, une approche plus directe pour déterminer directement les dérivées associées aux variables de couplage semble plus appropriée, surtout dans le cas de PMDO.

2.3.2.3 Classification des couplages

L'approche proposée par Rogers & Bloebaum (1994) se focalise sur les termes de sensibilités et exploite une classification, basée sur des critères statistiques pour simplifier une analyse multidisciplinaire (MDA).

Soit Y_A la variable de sortie de la discipline A et Y_B la variable de sortie de la discipline B, une définition normalisée de la sensibilité du couplage de A vers B est :

$$S_{B,A} = \left(\frac{\partial Y_B}{\partial Y_A} \right)' = \frac{Y_A}{Y_B} \frac{\partial Y_B}{\partial Y_A} \quad (2.15)$$

où $S_{j,i}$ désigne la sensibilité d'une variable de sortie de la j ème discipline par rapport à une variable de sortie de la i ème discipline. Une boucle complète de couplage est caractérisée par un lien avant $S_{j,i}$, et un lien arrière $S_{i,j}$. Ce paramètre permet alors de caractériser un lien de couplage entre deux disciplines. Un classement est ensuite effec-

Catégorie	Bornes
Extrêmement Fort (ES)	$S_{i,j} \geq \bar{S} + k_1\sigma(S)$
Très Fort (VS)	$\bar{S} + \frac{3}{5}k_1\sigma(S) \leq S_{i,j} < \bar{S} + k_1\sigma(S)$
Fort (S)	$\bar{S} + \frac{1}{5}k_1\sigma(S) \leq S_{i,j} < \bar{S} + \frac{3}{5}k_1\sigma(S)$
Neutre (N)	$\bar{S} + \frac{1}{5}k_2\sigma(S) < S_{i,j} < \bar{S} + \frac{1}{5}k_1\sigma(S)$
Faible (W)	$\bar{S} - \frac{3}{5}k_2\sigma(S) < S_{i,j} \leq \bar{S} - \frac{1}{5}k_2\sigma(S)$
Très Faible (VW)	$\bar{S} - k_2\sigma(S) < S_{i,j} \leq \bar{S} - \frac{3}{5}k_2\sigma(S)$
Extrêmement Faible (EW)	$S_{i,j} \leq \bar{S} - k_2\sigma(S)$

TAB. 2.1 Classement des sensibilités, avec k_1 et k_2 des paramètres réglables

tué parmi toutes les sensibilités $S_{i,j}$ en fonction de leur moyenne ($\bar{S}$), et de leur écart-type ($\sigma(S)$) (voir tableau 2.1). Des règles accompagnent cette classification pour caractériser les liens de couplage au sein d'une boucle. Tous les liens classifiés "Extrêmement Faible" peuvent être suspendus sans comparaison. Ensuite, seules les liens de retour sont les plus problématiques. Si l'une des sensibilités est classifiée "Extrêmement Fort" à "Très Fort", en comparaison d'un lien classifié "Extrêmement Faible" à "Très Faible", le lien de couplage de plus faible sensibilité peut être suspendu.

Rogers et Bloebaum (Rogers, 1996a, Rogers et al., 1996, Rogers, 1996b) montrent que ces simplifications permettent une résolution du MDA avec une très bonne précision. Par ailleurs, une modification de l'ordre dans lequel les disciplines sont résolues, permet de réduire le nombre de liens arrière. Rogers et al. ont également développé un code (DeMaid) pour optimiser la réorganisation des tâches afin de minimiser les liens de couplage retour.

Cette dernière méthode semble la plus intéressante à appliquer dans le cas de PMDO, puisqu'elle se base sur un traitement direct des termes de sensibilité. Une version modifiée de cette méthode est utilisée pour contribuer au traitement des variables. Rogers exploite des rapports de variations des variables de sortie de disciplines distinctes l'une par rapport à l'autre, tandis que la sélection de variables exploite des rapports de varia-

tion d'une variable de sortie sur une variable d'entrée au sein d'une discipline à la fois. Toutefois, l'objectif de la sélection de variables n'est pas de caractériser les couplages, mais l'influence des variables. Enfin, c'est une méthode inspirée de cette caractérisation de la sensibilité des variables qui est exploitée dans le traitement des couplages.

2.4 Algorithmes d'optimisation

À la base de toute méthode, stratégie ou approche d'un système MDO, un algorithme d'optimisation est la pièce maîtresse qui va déterminer comment la recherche de points est menée. Le prototype PMDO intègre, par l'optimiseur Isight, un choix de nombreux algorithmes d'optimisation, dont les caractéristiques permettent de déterminer celui qui convient le mieux au problème que l'on souhaite résoudre. Le tableau 2.2 détaille les noms reliés aux algorithmes, et le tableau 2.3 présente les méthodes disponibles et leurs caractéristiques dans l'optimiseur intégré au prototype. L'algorithme d'optimisation choisi parmi la liste proposée par Isight est NLPQL, pour plusieurs raisons :

- C'est un algorithme à gradient. En effet, la stratégie a pour but d'exploiter l'exploration au voisinage de points supposés être proches du point optimal.
- Cet algorithme supporte les non-linéarités, probablement fortes dans le problème de PMDO, en raison des fonctions *if* des feuilles de calcul Excel.
- La faible quantité de mémoire requise par l'algorithme permet de longs calculs.
- Enfin, le principe de la méthode est bien connue au laboratoire, pour être en mesure d'analyser l'historique des optimisations.

La revue bibliographique fut portée à l'origine sur d'autres types de méthodes, dont les caractéristiques auraient pu être intéressantes pour PMDO. Toutefois leur efficacité n'a pas été vérifiée, même si leur intégration dans Isight est possible. Le projet étant financé sur deux ans par Pratt & Whitney Canada, le cadre de recherche devait tenir compte de cette limitation. Ainsi, les méthodes explorées pour le projet furent :

Acronyme	Nom Complet
DHS	Directed Heuristic Search
DSO	Downhill Simplex Optimizer
EP	Exterior Penalty
LSGRG2	Generalized Reduced Gradient
HJ	Hook-Jeeves Direct Search Method
CONMIM	Method of Feasible Directions
MOST	Mixed Integer Optimization
MFD	Modified Method of Feasible Directions
MIGA	Multi-Island Genetic Algorithm
NCGA	Neighborhood Cultivation Genetic Algorithm
NSGA-II	Non-dominated Sorting Genetic Algorithm
PAO	Pointer Automatic Optimizer
PRTO	Pointer Run Trained Optimizer
PTO	Pointer Trained Optimizer
STOA	Satisficing Trade off Analysis
SLP	Sequential Linear Programming
DONLP	Linear Sequential Quadratic Programming
NLPQL	Non-Linear Sequential Quadratic Programming
SAM	Successive Approximation Method

TAB. 2.2 Acronymes des optimiseurs utilisables dans Isight

Optimiseurs de Isight	Gradient	Connaissance	Nombres entiers	Non-linéaire	Longs calculs	Espace discontinu	Utilisable dans PMDO
DHS	Non	Oui	Oui	Oui	Oui	Non	Non
DSO	Non	Non	Non	?	Oui	Oui	Oui
EP	Oui	Non	Non	Oui	Oui	Oui	Oui
LSGRG2	Oui	Non	Non	Oui	Oui	Non	Non
HJ	Non	Non	Oui	Non	Non	Non	Non
CONMIM	Oui	?	Non	Oui	Oui	Non	Oui
MOST	Oui	Oui	Oui	?	Oui	?	Non
MFD	Oui	Non	Non	Oui	Oui	Non	Oui
MIGA	Non	Non	Oui	Oui	Non	Oui	Non
NCGA	Non	Non	Oui	Oui	Non	Oui	Non
NSGA-II	Non	Non	Oui	Oui	Non	Oui	Non
PAO	Partiel	Non	Oui	Oui	Oui	Oui	Non
PRTO	Partiel	Non	Oui	Oui	Oui	Oui	Non
PTO	Partiel	Non	Oui	Oui	Oui	Oui	Non
STOA	Oui	Oui	Non	?	Non	Non	Non
SLP	Oui	Non	Non	Non	Oui	Non	Non
DONLP	Oui	Non	Non	Oui	Non	Non	Non
NLPQL	Oui	Non	Non	Oui	Oui	Non	Oui
SAM	Oui	Non	Oui	Non	Oui	Non	Non

TAB. 2.3 Fonctionnalités des optimiseurs de Isight

- les algorithmes génétiques (Martin & Quinn, 1995, Nye, 2004, Wagner et al., 2004, Buche et al., 2004, Spinellis & Papadopoulos, 1999), qui supportent bien les non-linéarités, et sont souvent utilisés pour des optimisations sur des systèmes mécaniques complexes,
- le recuit simulé, offrant une convergence théorique vers l'optimum global (Martin & Quinn, 1995, Spinellis & Papadopoulos, 1999),
- la recherche Tabu, qui capitalise les régions de l'espace déjà explorées (Martin & Quinn, 1995, Dengiz & Alabas, 2000),
- les systèmes experts qui exploitent une base de données pour prendre des décisions sur des directions de recherche, suivant la localisation dans l'espace de conception (Mookerjee & Mannino, 1997),
- les méthodes d'exploration avec plans d'expériences, utilisant les tables de Taguchi (Unal & Dean, 1991), ou une version améliorée de ces tables (Buchsbaum et al., 2003).

Ces méthodes peuvent alors utiliser des modèles remplaçant les appels à d'éventuelles fonction coût complexes comme :

- des réseaux de neurones (Moerland & Fiesler, 1996, Kozma et al., 2003),
- des chaînes de Markov (Abbeel & Ng, 2004),
- des surfaces de réponses (Resit et al., 1998).

Par ailleurs, d'autres méthodes moins connues, et dont l'efficacité même théorique n'a pas été vérifiée se trouvent dans la littérature :

- la méthode "DIRECT" (Siah et al., 1999), qui subdivise méthodiquement l'espace de conception pour réaliser des approximations locales,
- la méthode "MDD" (Kott & Gabriele, 1999), qui tente de parcourir l'espace en maximisant la distance par rapport au dernier point exploré,
- une méthode qui propose d'exploiter une mesure de la longueur d'onde de l'espace de conception pour l'exploration en s'affranchissant du bruit ou des minima locaux (Morelle, 1999).

D'autres méthodes pourraient être développées pour introduire des nouvelles techniques de recherche. Le but initial du projet de maîtrise était d'exploiter l'optimiseur intégré, sans avoir recours à d'autres composants.

Les résultats d'optimisation figurent au chapitre 6. La modularité du prototype PMDO a été exploitée pour intégrer un optimiseur exécutable BFGS (méthode à gradient), dans un but de comparaison avec l'optimisation effectuée avec NLPQL par Isight. La stratégie établie étant théoriquement plus adaptée pour améliorer une optimisation à gradient, seuls ces deux types d'algorithmes d'optimisation ont été testés.

2.5 Conclusion de la revue et établissement de la démarche

La revue de la littérature a montré l'importance du choix de la décomposition et de l'algorithme d'optimisation pour résoudre un problème multidisciplinaire. Le prototype PMDO existant permet à l'utilisateur de choisir une méthode de décomposition parmi une liste et de choisir l'algorithme d'optimisation parmi ceux proposés par Isight. Le projet propose de traiter deux aspects fondamentaux de l'analyse et de l'optimisation multidisciplinaire : l'analyse des variables et l'étude des couplages.

Sobieszcanski-Sobieski (1990) et Egorov et al. (1999) suggèrent l'emploi de modèles simplifiés pour l'optimisation puisque le but est de représenter les interactions et non d'obtenir des résultats de simulation précis. Par ailleurs, les outils de sélections de variables cherchent à exploiter la règle du 80-20 énoncée par Montgomery (2000) : 20% des efforts produisent 80% des effets. En extrapolant ce principe en optimisation, un nombre réduit de variables pourrait expliquer une part majeure des variations des variables de sortie. Toutes les disciplines ainsi exploitées par des nombres réduits de va-

riables permettent de simplifier l'analyse multidisciplinaire sans affecter les interactions. Le problème d'optimisation est alors allégé de variables. Si des variables de couplage sont suspendues suite à cette sélection, alors certains couplages sont également suspendus. L'analyse des liens interdisciplinaires permet ensuite de tenter de continuer à simplifier l'analyse multidisciplinaire. Toutefois, pour conserver une certaine précision et éviter une simplification à outrance, il est préférable de revenir au système non simplifié, une fois la convergence atteinte.

Cette tactique est prometteuse car en effet, une optimisation à gradient est, en général, plus efficace si le point de départ est proche de l'optimum. Si l'optimum du problème simplifié est proche de l'optimum du problème non simplifié, alors l'optimisation du problème complet est potentiellement plus rapide par comparaison à une optimisation partant d'un point de départ plus éloigné de l'optimum. Cette proximité des points optimaux est plus que probable. En effet, il est raisonnable de penser que la position de l'optimum dépend plus fortement des variables les plus importantes dans le système, puisque la fonction coût est calculée à partir des mêmes variables de sortie que le problème soit simplifié ou non. Et ce sont les variations de ces variables de sortie qui auront permis de déterminer quelles variables d'entrée les influencent le plus. L'optimum déterminé par la formulation simplifiée sera donc probablement proche de l'optimum de la formulation complexe, puisque la valeur de la fonction coût associée ne subira pas de forte évolution, si l'analyse de sensibilité est fiable.

Cette hypothèse a été testée sur une fonction polynomiale simple, afin de déterminer sa pertinence. Les résultats de cette courte étude montrent que cette stratégie peut s'avérer très intéressante, si plusieurs caractéristiques sont réunies. Ces critères figurent au chapitre 5 et les résultats sont présentés au chapitre 6.

Pour poursuivre la simplification de l'analyse multidisciplinaire suite à la sélection de variables, la stratégie proposée inclut une étude des couplages. En effet, la méthode proposée pour caractériser les couplages est ciblée sur des liens spécifiques, plutôt qu'une classification et une comparaison basée sur des critères statistiques, qui nécessiteraient

un calcul de toutes les sensibilités. C'est un outil complémentaire à l'analyse de sensibilité pour l'étude de certains couplages qui complexifient un problème d'optimisation.

Un test sur une fonction analytique a permis d'établir qu'une méthode utilisant trois étapes d'optimisation, au lieu de deux comme évoqué dans le début de cette section, peut donner des résultats exploitables avant même la fin du processus d'optimisation complet. La même démarche est appliquée sur le problème PMDO.

CHAPITRE 3

CLASSIFICATION DES VARIABLES

Cette section explique la méthodologie de réalisation de l'analyse de sensibilité et du classement de variables. Certaines vérifications sont menées afin de s'assurer de la pertinence de l'analyse.

3.1 Définition de la sensibilité

La sensibilité d'une variable de sortie (Y) par rapport à une variable d'entrée, au premier ordre, peut s'écrire :

$$S_{Y,X} = \frac{\Delta Y}{\Delta X} \quad (3.1)$$

Montano & Palmer (2003) ont proposé une définition différente pour être en mesure de prendre en compte les nombres entiers dans la définition précédente :

$$S_{Y,X} = \frac{\bar{Y}(X_{\max}) - \bar{Y}(X_{\min})}{X_{\max} - X_{\min}} \quad (3.2)$$

où $\bar{Y}(X_{\max})$ est la moyenne de Y quand la variable d'entrée X est fixée à sa valeur maximale, les autres variables d'entrée étant non fixées, et $\bar{Y}(X_{\min})$ quand X est fixée à sa valeur minimale. Ces définitions de premier ordre ne permettent pas de connaître la sensibilité des interactions de deux variables ou plus. L'approche proposée consiste à construire une surface de réponse (RSM), représentée par une fonction polynomiale, composée de fonctions linéaires, d'interactions du second ordre et quadratiques des variables d'entrée. Une minimisation des moindres carrés des résidus permet de déterminer

les pondérations du modèle :

$$\begin{aligned}
 Y = & \underbrace{I}_{\text{moyenne}} + \underbrace{a_1 X_1 + a_2 X_2 + \dots}_{\text{linéaires}} \\
 & + \underbrace{b_1 X_1 X_2 + b_2 X_1 X_3 + \dots + c_1 X_2 X_3 + c_2 X_2 X_4 + \dots}_{\text{interactions}} \\
 & + \underbrace{\alpha_1 X_1^2 + \alpha_2 X_2^2 + \dots}_{\text{quadratiques}}
 \end{aligned} \tag{3.3}$$

Les variables d'entrée (X_i) doivent être adimensionnées grâce à l'information des intervalles définis pour chaque variable. Une adimensionnalisation doit aussi être appliquée sur les variables de sortie, en considérant des intervalles de référence, sensés représenter leurs variations typiques. La section suivante démontre l'importance de cette adimensionnalisation pour la comparaison des influences des variables.

Le nombre de coefficients de pondération est une fonction simple du nombre de variables. Ce nombre détermine le nombre minimum d'évaluations de fonctions requis pour la construction de la surface de réponse. Le nombre total de coefficients à déterminer est donné par l'expression suivante, où $N_{\text{entrée}}$ est le nombre de variables d'entrée :

$$\begin{aligned}
 N_{\text{termes}} = & \underbrace{1}_{\text{moyenne}} + \underbrace{N_{\text{entrée}}}_{\text{linéaires}} + \underbrace{C_{N_{\text{entrée}}}^2}_{\text{interactions}} + \underbrace{N_{\text{entrée}}}_{\text{quadratiques}} \\
 = & \sum_{k=0}^2 (C_{N_{\text{entrée}}}^k) + N_{\text{entrée}} \\
 = & \frac{1}{2} N_{\text{entrée}} (N_{\text{entrée}} + 1)
 \end{aligned} \tag{3.4}$$

où $C_{N_{\text{entrée}}}^2$ est le nombre de combinaisons possibles de 2 éléments parmi $N_{\text{entrée}}$. Resit et al. (1998) suggèrent des modèles sur-déterminés de 20 à 50% pour la construction de surfaces de réponses. Mais, des tests préliminaires sur le problème PMDO ont montré qu'un échantillonnage de l'espace de conception pourvu d'un nombre de points deux à trois fois plus élevé que le nombre de coefficients du modèle ($N_{\text{évaluations}} \simeq 2N_{\text{termes}}$)

est requis pour permettre une bonne reproductibilité de résultats. La répartition des points dans l'espace de conception a été réalisée par la technique de l'hypercube latin qui permet une répartition la plus uniforme possible des points d'échantillonnage sur l'espace de conception comme recommandé par McKay et al. (2000). Cette méthode performante est simple à mettre en place. Dans le cas de PMDO, les termes quadratiques furent ajoutés pour des raisons de stabilité du modèle : la surdétermination du modèle nécessaire à la reproductibilité est moindre lorsque les termes quadratiques sont présents. Ces termes représentent un surplus de coefficients à déterminer, donc requérant plus d'évaluations de la fonction coût, mais permettent des échantillons de données plus modestes.

3.2 Influence et surface de réponse

3.2.1 Définition de l'influence

L'influence d'une variable est alors la pondération associée au terme auquel elle appartient. Un terme linéaire ($a_i, i \in \{1, \dots, N_{\text{entrée}}\}$), comme un terme quadratique ($\alpha_i, i \in \{1, \dots, N_{\text{entrée}}\}$), donne son influence à une variable. Un terme d'interaction ($b_i, i \in \{1, \dots, N_{\text{entrée}} - 1\}$, $c_i, i \in \{1, \dots, N_{\text{entrée}} - 2\}$, etc.) valorise deux variables, chacune avec la même influence. Dans le cas de PMDO, ces influences reliées aux variables sont analysées sur une base de comparaison commune.

Ainsi, un modèle de surface de réponse est bâti pour approximer la valeur d'une variable de sortie en fonction des n variables d'entrée. m surfaces sont construites pour modéliser les variations des m variables de sortie. L'amplitude des coefficients déterminent alors l'influence de variables d'entrée relativement à une variable de sortie.

3.2.2 Influence relative et absolue

L'adimensionnalisation est recommandée par la plupart des auteurs dans le but de faciliter les comparaisons. Cette tâche est simple pour les variables d'entrée car les intervalles de définitions, devant être définis pour tout problème d'optimisation, sont déjà disponibles et peuvent servir à l'adimensionnement suivant une formulation classique :

$$X' = \frac{X - X_{inférieur}}{X_{supérieur} - X_{inférieur}} \quad (3.5)$$

où $X' \in [0, 1]$ est la forme adimensionnelle de $X \in [X_{inférieur}, X_{supérieur}]$. Le comportement des variables de sortie devrait être connu, pour détecter un comportement anormal. À partir de recommandations d'ingénieurs expérimentés de P&WC, la définition des bornes d'intervalles des variables de sortie pour le problème PMDO a été vérifiée, en effectuant des plans d'expériences sur les variables d'entrée.

Les variables d'entrée et de sortie ainsi adimensionnées, varient en principe dans un intervalle normalisé $([0, 1])$. Les résultats montrent que l'amplitude des coefficients du modèle de surface de réponse ne reste pas dans le même ordre de grandeur suivant la variable de sortie considérée. À titre d'exemple, la figure 3.1 montre des coefficients dont l'amplitude varie de 0 à 14. La comparaison simultanée de plusieurs surfaces de réponses est alors impossible. La méthodologie utilise alors la notion d'influence relative. Dans le cas de PMDO, le poids du compresseur basse pression (LPC, Low Pressure Compressor) peut recevoir des contributions majeures de plusieurs variables d'entrée, jugées influentes par le coefficient auquel elles sont reliées. Mais, ces variables ne peuvent pas être considérées influentes pour toutes les variables de sortie de la discipline du compresseur basse pression. En effet, puisque les coefficients des RSM ne sont pas du même ordre de grandeur, la comparaison de coefficients spécifiques n'est pas cohérente d'une RSM à une autre. Si l'amplitude de variation des coefficients était du même ordre de grandeur, tous les coefficients pourraient être comparés de façon absolue à travers les

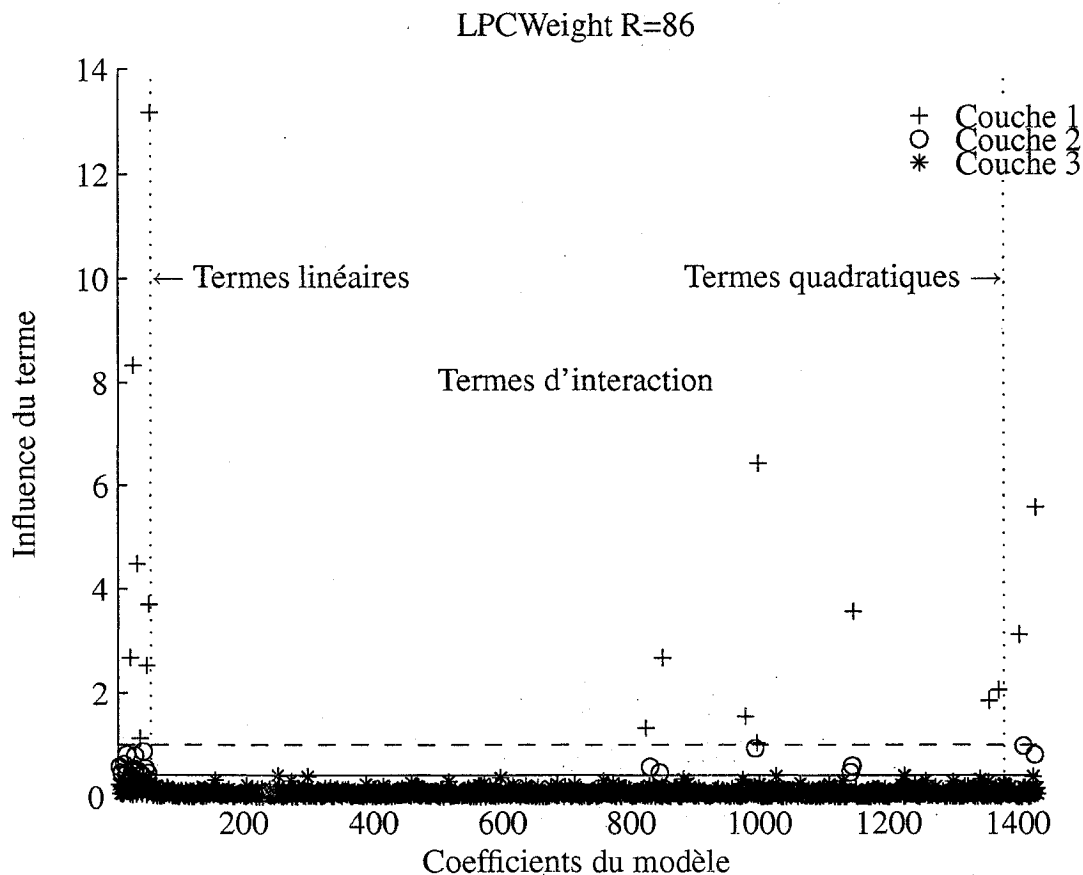


FIG. 3.1 Tracé des coefficients d'un modèle de surface de réponse

surfaces de réponses, sans comparaison relative à une variable de sortie en particulier. Cette distinction est importante et sera exploitée lors de la classification des variables.

3.3 Modèle de surface de réponse

La figure 3.1 montre le tracé des coefficients de la surface de réponse associée au poids du compresseur basse pression. Dans cet exemple, les 52 premiers points représentent les pondérations des contributions linéaires, les 52 derniers points indiquent les poids des contributions quadratiques. Les autres points montrent les contributions des interactions. Le coefficient R, dans le titre du graphe, est le coefficient de détermination, estimateur

d'erreur du modèle. Ce coefficient permet d'exprimer la capacité du modèle, à partir des fonctions de base, d'expliquer la variation d'une variable de sortie, en fonction des variables d'entrée, sous leur différentes formes de contributions. Dans l'exemple montré, $R = 86$, ce qui signifie que le modèle permet d'expliquer 86% de la variation de la variable de sortie par les contributions linéaires, les interactions du second ordre et les fonctions quadratiques des variables d'entrée. Les 14% restants correspondent à une variation non expliquée par ces fonctions. Cette erreur pourrait être réduite en utilisant des fonctions de base plus complexes, ou des fonctions polynomiales d'ordre plus élevé. Dans le cas de PMDO, l'analyse devant rester la plus générique possible pour toutes les variables, aucune contribution plus complexe n'est étudiée.

3.4 Classification

Exploitant une méthode que Rogers & Bloebaum (1994) ont employé pour réduire le nombre de liens de couplage interdisciplinaires, une technique simplifiée est utilisée pour effectuer un classement de variables. Pour classer l'intensité des couplages, Rogers et Bloebaum utilisaient 7 catégories, dont les seuils étaient basés sur les moyenne et écart type des sensibilités de couplage étudiées. Pour le classement des variables, la moyenne des coefficients ($\overline{RSM}$) et leur écart-type ($\sigma(RSM)$) sont exploités pour définir deux bornes :

$$Seuil_1 = \overline{RSM} + 1,5 \cdot \sigma(RSM) \quad (3.6)$$

$$Seuil_2 = \overline{RSM} + 0,5 \cdot \sigma(RSM) \quad (3.7)$$

Les seuils déterminent une limite inférieure aux coefficients. La pondération de l'écart type a été déterminée en exploitant une loi normale décentrée, supposément proche de la répartition réelle. Ainsi, de manière arbitraire, les 10% plus importants coefficients de-

vraient être supérieurs au $Seuil_1$, et les 50% plus importants coefficients devraient être supérieurs au $Seuil_2$. Ces pourcentages arbitraires proviennent d'un classement effectué selon l'expérience des ingénieurs de Pratt & Whitney Canada qui déterminent environ 10% du nombre total de variables comme étant importants d'un point de vue physique. Le but de ces limites est de détecter les poids les plus importants dans le modèle de surface de réponse, qui restent supérieurs à un certain seuil, afin d'être capable d'établir un classement de variables d'influence décroissante. L'algorithme débute avec les plus importants coefficients, et analyse leur statut. Si le coefficient est linéaire ou quadratique, un nom de variable est associé à sa numérotation dans le modèle. Ce nom est stocké en première place dans le fichier de classement. Si le coefficient est une interaction, deux noms de variables sont alors reliés à la numérotation dans le modèle. Ces deux noms sont alors stockés dans le fichier de classement. Ici, la même variable peut être sélectionnée deux ou trois fois, en raison de la nature des coefficients linéaire, d'interaction et quadratique. Évidemment, la liste finale ne verra apparaître qu'une seule occurrence de la même variable. Cette méthode basée sur des seuils permet à l'algorithme de rechercher de nouvelles variables jusqu'à atteindre la limite d'amplitude minimale sur les coefficients du modèle.

La figure 3.1 montre trois couches de données, parce que le processus d'optimisation comprend trois étapes. Les variables associées aux coefficients de la première couche (points +) sont les variables utilisées durant la première étape du processus. La seconde étape du processus exploite les variables reliées aux coefficients de la première et de la deuxième couche (points o). Enfin, la dernière étape du processus utilise toutes les variables initiales, sans réduction ni traitement particulier.

La méthodologie de classement est complète en considérant le nombre de variables influentes lors de la recherche dans les coefficients du modèle. En effet, la liste bâtie peut être très longue, représentant jusqu'à plus de 50% des variables initiales, alors que le but est de ne conserver que les plus importantes. Les $Seuil_x$ ($x \in \{1, 2\}$) agissent comme sélecteur sur les coefficients du modèle, et non sur les variables. La sélection de deux

coefficients peut pointer sur une variable (une contribution linéaire et une contribution quadratique, référant à la même variable) autant que sur quatre variables (deux contributions d'interactions, associées à deux variables différentes). Pour éviter qu'un trop grand nombre de variables ne soit sélectionné, un nombre maximum de variables a été défini pour chaque couche de variable. Ainsi, pour chaque variable de sortie, une valeur arbitraire de 30% des variables d'entrée seront sélectionnées pour la première couche. Et 60% d'entre elles seront sélectionnées pour la deuxième couche. En effet, la méthodologie utilisant trois couches de variables, un objectif de simplification pour ramener le nombre de variables sélectionnées au tiers du nombre total semble pertinent pour la première couche. Ce critère limite le nombre de variables sélectionnées pour les couches dans l'analyse d'une surface de réponse pour une variable de sortie. Ceci ne garantit en rien un nombre de variables maximum lorsque les listes sont comparées d'une variable de sortie à une autre. En principe, les variables de sortie sont indépendantes, et dépendent de variables d'entrée spécifiques. Mais il est légitime de penser qu'un groupe limité de variables d'entrée puisse être identifié comme le principal contributeur à plusieurs variables de sortie. Ces limites de nombre de variables sont une tentative pour fixer un objectif de simplification de l'analyse multidisciplinaire. En limitant le nombre de variables d'entrée à 30% pour chaque variable de sortie, l'analyse disciplinaire pourrait donc compter environ 30% des variables du système initial.

Deux critères conditionnent donc l'établissement des listes de variables. Pour une variable de sortie, les coefficients de la surface de réponse sont sélectionnés par rapport à deux seuils basés sur leurs moyenne et écart type. Ensuite, les variables associées aux coefficients sont identifiées et ajoutées sur une liste jusqu'à ce qu'un certain nombre d'entre elles soient sélectionnées. Ce nombre maximum est fonction du nombre total de variables d'entrée. Ceci permet de cerner le groupe de variables qui contribue le plus à la variable de sortie considérée. Une fois ces listes générées pour chaque variable de sortie, leur comparaison permet d'établir une liste finale de variables contributrices aux variables de sorties d'une discipline. Cette même sélection est ensuite appliquée aux

autres disciplines du MDA.

Afin de fixer les idées, un algorithme simplifié est présenté pour montrer comment passer des coefficients de la surface de réponse de la figure 3.1, à des listes de variables d'influence décroissante. Avant d'entrer dans cet algorithme, les coefficients de la surface de réponse ($RSM(i)$) sont déterminés et le nombre total de variables d'entrée de la discipline est connu ($NbVarEntrée$). L'algorithme se décompose selon les étapes suivantes :

- Calcul des nombre de variables à sélectionner par couche ($NbMaxVarCche_1 = 0, 3 \cdot NbVarEntrée$, $NbMaxVarCche_2 = 0, 6 \cdot NbVarEntrée$), moyenne et écart type des $RSM(i)$, et des $Seuil_x$, $x \in \{1, 2\}$
- Construction d'un classement des coefficients $RSM(i)$ du plus important au plus faible, par leur position (i) dans le modèle (RSM), puis stockage dans le vecteur *Classement*
- Construction de trois classements de longueur différentes et limités aux $Seuil_x$ définis, stockage dans les vecteurs *ClassementCche1*, *ClassementCche2*, de longueurs spécifiques *LongueurMax₁*, *LongueurMax₂*
- Début de boucle sur les couches (2 premières couches seulement) et initialisation d'un compteur (*Compteur*=1)
 - **while** ((*Longueur(ListeVar)* < *NbMaxVariablesCche_x*) & (*Compteur* < *LongueurMax_x*))
 - Identification du coefficient suivant (*ClassementCche_x*(*Compteur*))
 - Boucle d'identification du numéro du coefficient dans le modèle
 - Coefficient linéaire
 1. Identification du nom de la variable *NomVariable*
 2. Ajout à la liste *ListeVar*, si le *NomVariable* n'y est pas déjà inclus
 - Coefficient quadratique

1. Identification du nom de la variable *NomVariable*
 2. Ajout à la liste *ListeVar*, si le *NomVariable* n'y est pas déjà inclus
- Coefficient d'interaction
 1. Identification du nom de la variable *NomVariable1*
 2. Ajout à la liste *ListeVar*, si le *NomVariable1* n'y est pas déjà inclus
 3. Identification du nom de la variable *NomVariable2*
 4. Ajout à la liste *ListeVar*, si le *NomVariable2* n'y est pas déjà inclus
 - Incrémentation du compteur *Compteur* ++
 - Fin boucle **while**
 - Fin Algorithme de sélection de variables

Le critère sur les coefficients permet de restreindre le domaine des variables d'entrée à celles dont l'influence fait partie des plus fortes sur la variable de sortie. Le pire coefficient choisi pour la première couche reste dans les 10% des meilleurs contributeurs. De même, la dernière variable d'entrée choisie contribue, linéairement, grâce à une autre variable, ou quadratiquement à une variable de sortie de façon comparable. Il est possible que la limite sur les coefficients soit atteinte avant celle sur le nombre de variables : le processus ne peut chercher des variables dont l'influence relative sur une variable de sortie ne soit pas parmi les 10% les meilleures. Cette limitation est maintenue quelles que soient les listes de sélection pour les différentes variables de sortie. Les variables d'entrée de toutes les listes sont toutes parmi les 10% des meilleures contributrices à au moins une variable de sortie.

Le résultat de ce traitement est montré à travers les analyses multidisciplinaires des figures 3.2, 3.3 et 3.4. La dernière analyse multidisciplinaire (voir 3.4) est le problème d'optimisation original.

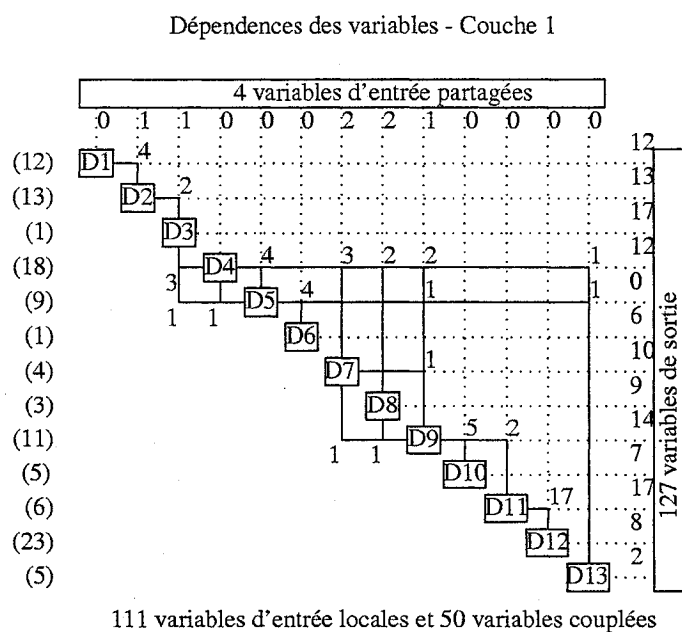


FIG. 3.2 Analyse multidisciplinaire exploitant la couche de variables 1

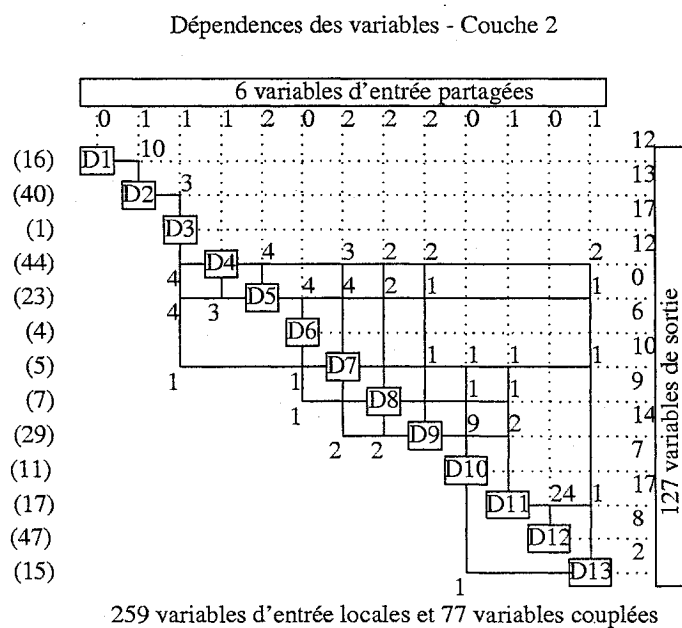


FIG. 3.3 Analyse multidisciplinaire exploitant la couche de variables 1 et 2

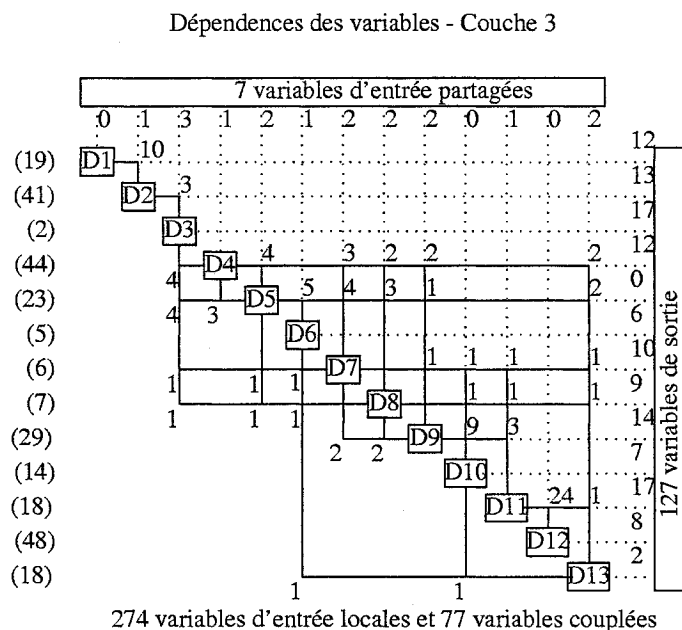


FIG. 3.4 Analyse multidisciplinaire exploitant la couche de variables 1, 2 et 3

3.5 Stabilité de l'espace de départ et contraintes

Les premières tentatives d'analyses de sensibilités menèrent à de nombreux *NaNs* ("Not a Number", dûs à une division par zéro ou à un calcul menant à un infini). Toute construction de surface de réponse est alors impossible.

Le contrôle visuel de la répartition des données a montré des problèmes sur la définition des variables. Par exemple, le rayon du moyeu peut avoir une valeur supérieure au rayon externe du compresseur basse pression. Le bon sens nous indique qu'un rayon de moyeu est toujours plus faible qu'un rayon externe, mais cette contrainte n'étant pas explicitée, l'analyse de sensibilité cherchait à tester toutes les valeurs autorisées par l'intervalle de définition.

Pour éviter ce problème, quatre options sont possibles :

- Créer une nouvelle contrainte qui garantit l'inégalité,
- Changer les bornes de définition afin que les intervalles soient disjoints et que l'incon-

sistance ne soit pas possible,

- Ajouter une fonction *if* dans la feuille Excel, qui vérifie l'inégalité et assigne une valeur consistante afin que le calcul soit cohérent,
- Ré-échantillonner les variables d'entrée jusqu'à l'obtention d'un nombre suffisant de points.

Pour alléger la fonction coût et conserver un processus le plus générique possible, la première solution doit être évitée. La seconde option n'est pas intéressante du point de vue de l'optimisation, car cela réduit énormément l'espace de conception, en forçant les variables à rester dans un voisinage de leur point de définition initial. Par ailleurs, la troisième solution n'est pas meilleure non plus. En effet, pour différentes valeurs inconsistantes, la feuille Excel réinterpréterait une valeur consistante et donnerait un résultat identique. Dans ce cas de figure, il n'y a pas de calcul de gradient possible, ni d'indication à l'analyse que le point testé est non valide, ni à quel point il est invalide.

Par ailleurs, tout changement de formulation sous une forme adimensionnelle ou relative est à exclure, car ces variables sont couplées : leurs valeurs dépendent des variables de sortie d'autres disciplines. Cette impasse n'advenant que pour les analyses de sensibilités, a été résolue par la simple recherche d'autres points valides (i.e., la quatrième solution proposée).

Les surfaces de réponses qui ont permis d'établir les MDA simplifiés, ont ainsi nécessité un mécanisme de ré-échantillonnage des valeurs des variables d'entrée. Ce processus itère jusqu'à l'obtention d'un nombre suffisant d'évaluations valides pour construire des surfaces de réponses, considérant le nombre de variables d'entrée.

CHAPITRE 4

ÉTUDE DES COUPLAGES

Ce chapitre explique comment certains couplages entre les disciplines ont été étudiés pour continuer à simplifier l'analyse multidisciplinaire issue de la sélection de variables. Des résultats issus de l'analyse de sensibilité sont exploités pour qualifier des liens de couplage ciblés. La répartition des valeurs des variables et l'influence de variables d'entrée sur des variables de sortie permettent de déterminer le couplage transmis d'une discipline à une autre.

4.1 Notion de transport de couplage

Considérons deux disciplines A et B , chacune pourvue de ses variables d'entrée (respectivement X_i et X_j), et de ses variables de sortie (respectivement X_j et X_k). Dans cet exemple, une des variables est couplée : la variable X_j est une variable de sortie de la discipline A et une variable d'entrée de la discipline B . En utilisant la définition de la sensibilité donnée dans le chapitre précédent, les termes $S_{j,i}$ et $S_{k,j}$ peuvent être élevés, si, pour la variable d'entrée variant dans tout son intervalle, la variable de sortie décrit son domaine de définition dans son intégralité, comme dans l'exemple de la figure 4.1. Pour la discipline A , lorsque la variable d'entrée X_i varie dans l'intégralité de son intervalle, la variable de sortie X_j varie sur 75% de son intervalle. La discipline B est telle que à l'intervalle de variation de la variable d'entrée X_j , correspond une variation de l'intégralité de l'intervalle pour la variable de sortie X_k . Toutefois, le couplage entre ces deux disciplines est limité. Dans le cas de la figure 4.1, la variable d'entrée de A (X_i) ne permet de faire varier la variable de sortie de B (X_k), à travers le couplage, que sur

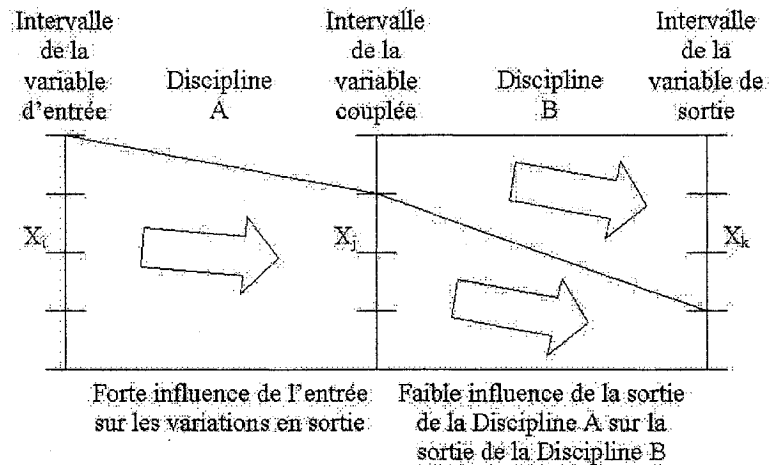


FIG. 4.1 Transport de couplage

25% de son intervalle. L'influence de cette variable de *A* est importante sur la variable d'entrée de *B*, mais finalement faible sur la variable de sortie de *B*. La cause de cette influence limitée est illustrée par la figure 4.1 : l'intervalle qui permet d'exercer la plus grande variation de la variable de sortie de *B* n'est pas atteint par la variable de *A*. Un critère sur les influences des sous-intervalles pourrait permettre de suspendre le lien de couplage, et donc de simplifier l'analyse multidisciplinaire.

4.2 Détermination de l'hypothèse de transport de couplage disjoint

L'identification du type de situation décrit en 4.1 requiert plusieurs analyses de sensibilités sur des sous-intervalles (4 dans l'exemple de la figure 4.1). Plusieurs conditions doivent être réunies pour suspendre le lien de couplage de l'analyse multidisciplinaire. La méthodologie de vérification est présentée dans les sections suivantes, dont chacune présente une condition suffisante, mais non nécessaire à la suspension du couplage.

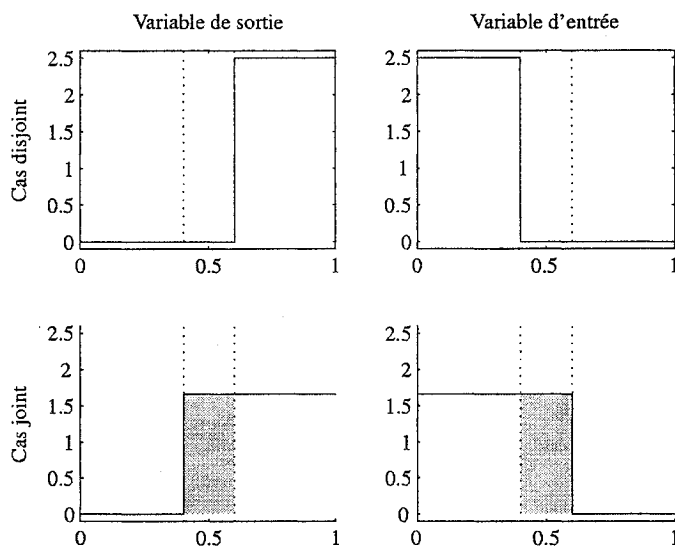


FIG. 4.2 Illustration des intervalles joints et disjoints

4.2.1 Répartition des valeurs des variables couplées

La première hypothèse à vérifier pour les variables couplées est la répartition des valeurs valides dans leur intervalle de variation, selon qu'elles soient variables d'entrée ou de sortie. En effet, les analyses disciplinaires peuvent mener à de nombreux problèmes de calculs, comme évoqué dans la section 3.5 "Stabilité de l'espace de départ" du chapitre précédent. Cette répartition montre comment, à partir d'une distribution uniforme des échantillons pour les variables d'entrée, les seuls points valides pour les analyses sont répartis.

La figure 4.2 montre des densités de répartition des points valides dans leur intervalle de variation. L'abscisse est l'intervalle de définition adimensionnel ($[0,1]$). Les valeurs en ordonnée sont telles que l'intégrale de la fonction de répartition soit de 1. Ces répartition sont particulières pour représenter les cas limites. Le cas disjoint montre une variable de sortie dont les valeurs permettent une analyse valide pour seulement une portion supérieure de l'intervalle (40% des plus fortes valeurs). La même variable présente une densité de répartition en entrée exclusivement dans la portion inférieure de l'inter-

valle (40% des valeurs les plus faibles). Un couplage basé sur cette variable ne pourrait transmettre aucune influence. En effet, la discipline A prenant en entrée cette variable hypothétique en sortie d'une autre discipline B ne pourrait recevoir que des valeurs qui conduisent à des incohérences (i.e., NaN) dans la discipline A . Un cas moins extrême est illustré dans cette même figure 4.2, avec le cas joint. Une autre variable hypothétique de couplage pourrait présenter une faible portion de son intervalle en commun comme variable de sortie et variable d'entrée. Dans ce cas, un couplage est transmissible d'une discipline (qui a cette variable en sortie) à une autre discipline (qui possède cette variable en entrée).

Le but de cette première hypothèse de vérification du couplage est de mesurer la portion d'intervalle en commun, ou encore la proportion de points qui permet le transport du couplage. Pour une proportion de points inférieure à un critère arbitraire de 10%, la transmission de couplage peut être considérée minime, et le couplage reliant les deux disciplines faible. Dans l'exemple de couplage illustré (cas joint), la longueur de l'intervalle de recouvrement est de 20%. La proportion de points concernés est alors l'intégrale de surface commune entre les deux graphes (la partie grisée). Ceci signifie que le couplage est transmis à travers le lien par seulement 30% des points. Suivant le seuil de sélection choisi, le couplage peut être considéré faible ou fort, et une suspension du lien pourrait en découler.

La figure 4.3 montre ainsi un exemple de densité de répartition de deux variables couplées, tour à tour variable de sortie et d'entrée pour deux disciplines distinctes. Comme variable de sortie, les valeurs issues d'analyses valides de *CentRtipLE* présentent une répartition relativement uniforme. Comme variable d'entrée, *CentRtipLE* conduit à plus d'analyses valides lorsque ses valeurs sont proches de la borne supérieure de son intervalle de définition. L'effet n'est pas autant significatif que dans les cas extrêmes, puisque dans ces cas de figure, aucune portion d'intervalle n'est invalide, même si certaines portions sont moins probables. La transmission du couplage entre la discipline 5 (où *CentRtipLE* est une variable de sortie) et la discipline 3 (où *CentRtipLE* est une

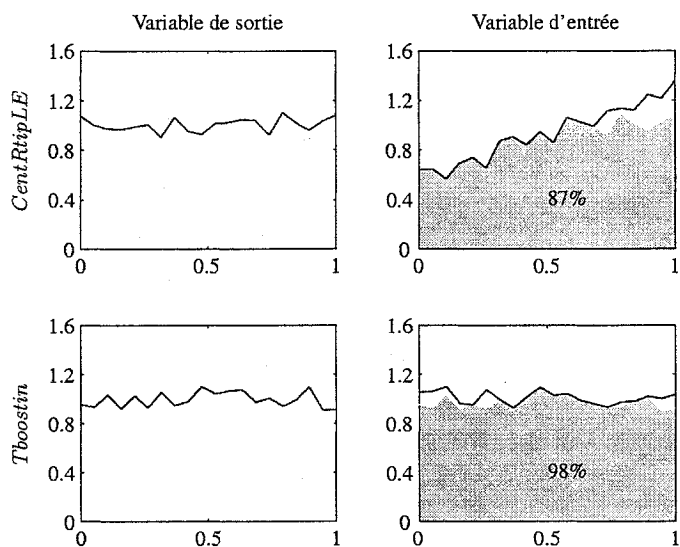


FIG. 4.3 Répartition des valeurs sur les intervalles de référence

variable d'entrée) n'est pas complète (le lecteur pourra se référer à la figure 3.2). En moyenne, 87% des valeurs de sortie de la discipline 5 donneront une analyse consistante de la discipline 3. Pour la variable *Tboostin*, la densité de répartition des valeurs dans les intervalles entraîne une meilleure transmission du couplage. Pour ces deux variables, la seule considération de la densité de répartition dans les intervalles ne suffit pas pour conclure sur l'influence du couplage.

Cette première étape permet de vérifier l'éventuelle incompatibilité des transmissions de couplage, à travers la répartition des points valides pour les analyses.

4.2.2 Intervalle d'influence

Cette section cherche à identifier quels sous-intervalles des variables d'entrée contribuent le plus à une variation des variables de sortie. Ces sous-intervalles d'influence permettent d'identifier la situation de la discipline *B* illustrée dans la figure 4.1.

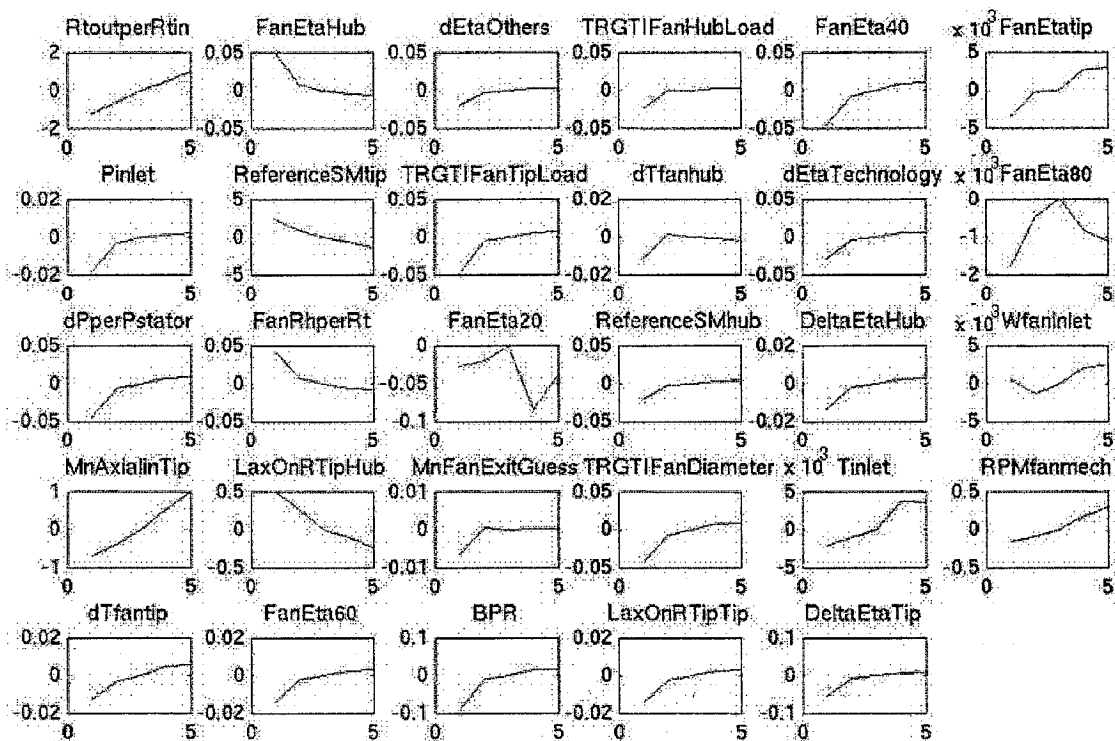
Exploitant le modèle déterminé par la surface de réponse, les pondérations des termes linéaires sont interprétés comme des coefficients directeurs de fonctions affines. Si une surface de réponse est construite pour chaque sous-intervalle, plusieurs coefficients directeurs sont obtenus, et décrivent une tendance de variation d'une variable de sortie par rapport à la variable d'entrée considérée.

Le processus consiste à :

- obtenir une RSM pour des variables d'entrée relativement à une variable de sortie,
- déterminer par la même analyse, une RSM sur des sous-intervalles distincts,
- construire la tendance des variables d'entrée sur l'intervalle au complet à partir des coefficients directeurs, pondérations linéaires dans les surfaces de réponses sur les sous-intervalles.

Toutefois, ces surfaces de réponses ne donnent pas une bonne représentation statistique des tendances. En effet, pour produire la figure 4.4, les variables ont été considérées dans le même sous-intervalle. La figure 4.5 illustre les sous-intervalles utilisés sous la dénomination "Non Recommandé". La première des quatre surfaces de réponse représente ainsi les tendances des variables d'entrée lorsque celles-ci ne varient toutes que dans le premier quart de leur intervalle de définition. Afin d'avoir une meilleure représentation statistique, la variation linéaire d'une variable d'entrée sur un sous-intervalle doit prendre en compte la variation des autres variables sur l'intégralité de leur intervalle, comme illustré par le schéma "Recommandé" sur la figure 4.5. Il s'agit de construire une surface de réponse par variable d'entrée et par sous-intervalle considéré.

Les tendances de 29 variables d'entrée à travers quatre sous-intervalles sur une variable de sortie, nécessitent $29 \times 4 = 116$ surfaces de réponses. Ce résultat est valable pour une



Trends of FanRinTip as a function of the input Variables

FIG. 4.4 Tendances de variables d'entrée par rapport à une variable de sortie de la discipline 1

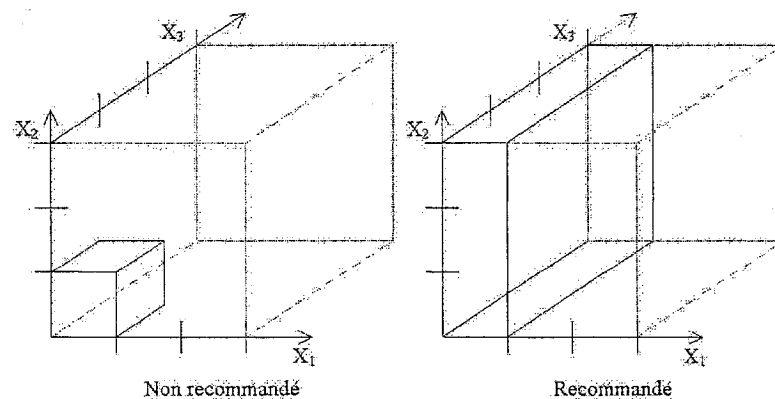


FIG. 4.5 Découpages de l'espace pour déterminer les surfaces de réponses de X_1

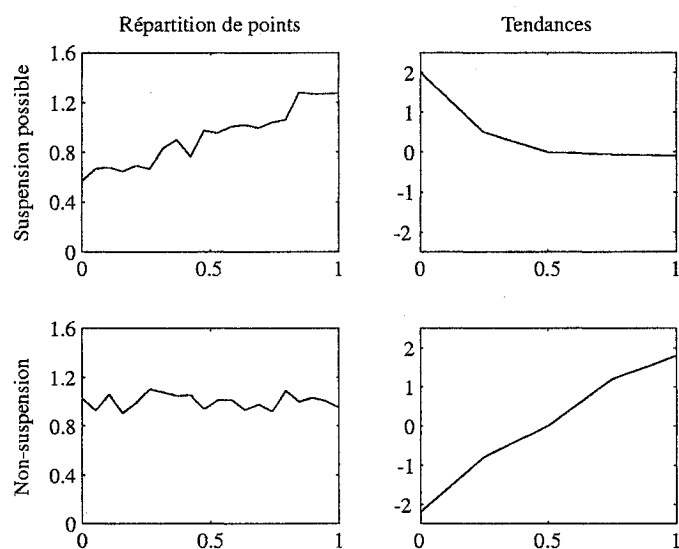


FIG. 4.6 Analyse croisée des répartitions des valeurs des variables et des intervalles d'influence

variable de sortie. Ces mêmes calculs ont besoin d'être repris pour toutes les variables de sortie qui sont des variables couplées soit 77. Toutefois, étant donné le nombre de calculs requis, l'analyse s'est restreinte aux liens de couplages les plus problématiques. D'après les travaux de Rogers et al. (1996), ces couplages problématiques sont les liens de retour dans les analyses multidisciplinaires (dans la figure 3.2 (page 50), ces couplages sont ceux reliant la discipline 9 aux disciplines 8 et 7, ou celles reliant 5 à 4 et 3). Les précautions prises pour prendre en compte une plus grande partie de l'espace n'ont toutefois pas beaucoup modifié les résultats initiaux. Toutefois, l'estimateur d'erreur n'a pas été pris en compte, étant donné l'important volume de données. La figure 4.4 donne des résultats très proches des tendances à travers tout l'espace, même si les tendances qu'elle représente devraient être considérées locales.

4.2.3 Croisement des informations

La figure 4.6 présente le croisement d'information qui pourrait permettre de suspendre un couplage. Les graphiques de gauche montrent des répartitions de valeurs de variables d'entrée pour mener à des analyses valides, et leurs tendances (succession de quatre coefficients directeurs à travers l'intervalle de définition) sur une variable de sortie dans les graphiques de droite. Le cas de suspension possible illustre une situation où la répartition des valeurs des variables est déséquilibrée : les valeurs de variables d'entrée menant à une analyse valide sont préférablement dans la partie supérieure de l'intervalle. Toutefois, le graphique montrant les tendances sur l'intervalle illustre une situation où la variable d'entrée a plus d'influence sur une variable de sortie lorsque la valeur de la variable d'entrée se situe dans la partie inférieure de l'intervalle. La proportion de points pouvant avoir une forte influence sur une variable de sortie est donc faible. Mais la plupart des cas analysés pour tenter de retirer les couplages de retour problématiques présentent des situations comme celles du cas de Non-suspension. À travers le graphe de la tendance, ce cas montre que la variable d'entrée a une influence forte sur la variable de sortie sur une grande partie, voire sur l'intégralité de l'intervalle. De plus, la répartition des valeurs de variables d'entrée menant à une analyse valide est plutôt uniforme sur l'intervalle. L'hypothèse de transport de couplage disjoint ne peut pas être prouvée, et le lien de couplage ne peut être suspendu.

4.3 Résultats de l'analyse

L'information de la répartition des valeurs des variables de sortie, croisée avec les données des tendances qui donnent les sous-intervalles contributeurs ne permet pas de retirer les couplages les plus problématiques. Après des calculs supplémentaires sur d'autres variables, des liens de couplage avant auraient pu être suspendus, mais cette opération

n'a pas été réalisée, pour plusieurs raisons :

- Les liens de couplage avant ne posent pas de problème particulier dans un problème d'optimisation, comparativement à des liens de retour d'après Rogers & Bloebaum (1994).
- Retirer de tels liens se concrétise en enlevant du système les contraintes de consistance qui y sont reliées. Ces contraintes sont gérées dans une décomposition multi-niveaux par l'optimiseur maître. La principale difficulté est d'empêcher le système de tenir compte d'une contrainte, alors que les variables concernées existent toujours lors de la définition du problème d'optimisation.
- Cet artifice est simple à réaliser à la main, une fois que le problème est défini, mais beaucoup plus délicat à implémenter dans le cadre d'une automatisation.
- Trop peu de liens (3) à retirer ne justifient pas une telle modification du code.

Cette méthode fut une tentative de caractériser de façon absolue un couplage, avec l'information sur des variables, et non relativement aux autres liens de couplage interdisciplinaires. En effet, les approches de Sobieszczanski-Sobieski (1990), Alyaout et al. (2005), Rogers (1996a), Rogers et al. (1996) et Rogers (1996b) requièrent le calcul des sensibilités de tous les couplages. Un classement par critères statistiques (seuils pondérés par la moyenne et l'écart type) peut alors différencier relativement tous les couplages. Ces approches permettent de simplifier indéfiniment un MDA puisque les couplages les plus faibles et les plus éloignés de la moyenne sont suspendus. En réappliquant la même approche sur un MDA simplifié, il existera toujours des liens moins importants que tous les autres. Or, le MDA considéré est déjà simplifié par un traitement sur les variables. De plus, l'approche proposée de caractérisation particulière du couplage tentait de montrer que le couplage pouvait dépendre d'une portion de l'intervalle des variables (sous-intervalles d'influence) au lieu de l'intervalle dans son intégralité.

CHAPITRE 5

PROCESSUS GLOBAL D'OPTIMISATION : APPLICATION

L'analyse de sensibilité des variables et l'étude des couplages sont replacés dans ce chapitre par rapport au processus global d'optimisation. Dans la deuxième section, la stratégie est appliquée sur un exemple simple, pour déduire certaines conditions nécessaires au problème pour que l'application de la stratégie soit efficace. Enfin, une analyse des caractéristiques du système PMDO permet un premier jugement sur la pertinence d'application de cette stratégie sur le problème industriel.

5.1 Stratégie globale d'optimisation

Les phases préalables à l'optimisation globale du système PMDO ont consisté en :

- Une analyse de sensibilité permettant d'établir 3 classes de variables d'entrée d'influence décroissante sur les variables de sortie. Cette analyse a donné trois analyses multidisciplinaires, de complexité croissante.
- Une étude des couplages pour tenter de simplifier les couplages les plus problématiques de l'analyse multidisciplinaire. Cette étude n'a toutefois pas permis de retirer de liens dans l'analyse multidisciplinaire la plus simplifiée.

Par la suite, plusieurs étapes d'optimisation sont conduites avec les différentes couches de variables sélectionnées. La figure 5.1 résume les deux phases d'analyse de sensibilité et d'étude des couplages préalables à l'optimisation pour la discipline 1, avec 3 couches de variables. La même procédure est appliquée aux autres disciplines. L'analyse des couplages se place à la fin de l'étape d'analyse, pour tenter de simplifier les liens problématiques, s'ils existent.

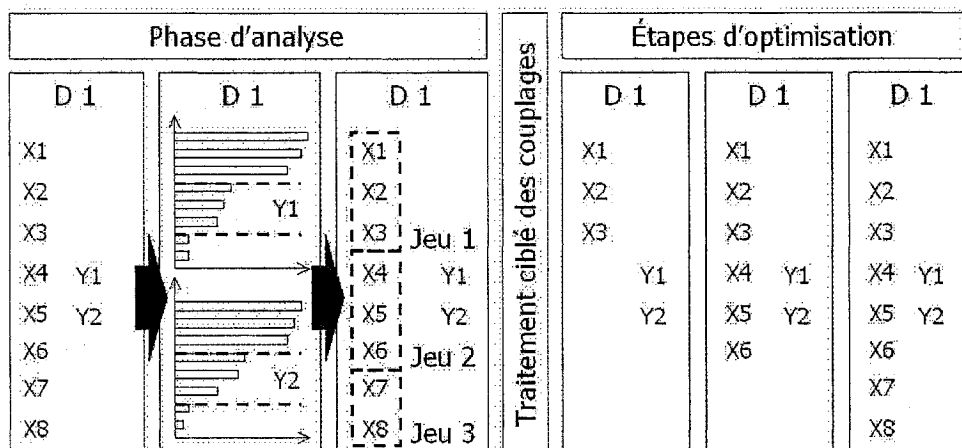


FIG. 5.1 Processus général d'optimisation, avec 3 couches de variables et une discipline

5.2 Illustration par un cas test

5.2.1 Formulation

La stratégie d'optimisation par couches de variables successives est appliquée sur la fonction coût décrite par le polynôme suivant :

$$f(X) = 1000X_1^2 + 300X_2^2 + 0.5(X_3 - 0.5)^2 + 0.2(X_4 + 0.5)^2 + 0.0002(X_5 - 0.3)^2 + 0.0005(X_6 + 0.2)^2 \quad (5.1)$$

La figure 5.2 montre les caractéristiques de la surface de réponse pour le cas test sur les variables ($X_1, X_2, X_3, X_4, X_5, X_6$). Le lecteur pourra se référer à la figure 3.1 (page 44) pour retrouver les types de coefficients. Dans la situation du cas test, aucun terme d'interaction n'apparaît. Les lignes pointillées horizontales délimitent les couches de variables très distinctes étant donnés les ordres de grandeur des coefficients du polynôme :

- Couche 1 : X_1, X_2
- Couche 2 : X_3, X_4
- Couche 3 : X_5, X_6

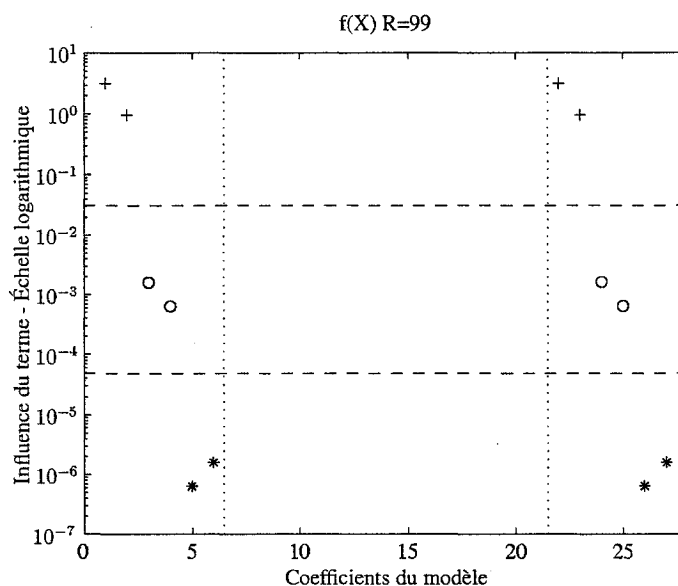


FIG. 5.2 Surface de réponse pour la fonction cas test

5.2.2 Historique et résultats d'optimisation

L'historique d'optimisation sous une forme logarithmique est donnée par la figure 5.3. La figure ne montre pas toutes les évaluations de la fonction coût. Ainsi, les calculs de gradient, et de recherche de ligne n'apparaissent pas. Le trait plein est l'historique de l'optimisation menée par la méthode BFGS¹. Deux plateaux permettent de distinguer trois étapes dans cette optimisation. Le premier plateau est obtenu par l'optimisation sur les variables (X_1, X_2) . Des itérations sont dépensées sur des optimisations majoritairement dirigées par les variables (X_3, X_4) , pour parvenir au deuxième plateau. Enfin, l'optimisation parvient à diminuer la fonction coût en itérant sur le dernier ensemble de variables (X_5, X_6) . En trait pointillé, l'optimisation par couches franchit 3 étapes séparées par les traits verticaux. Le processus global consomme moins d'itérations en exploitant la stratégie par couches, comparativement à une méthode BFGS classique, même si plusieurs itérations sont comparables.

¹La méthode BFGS est une variante de la méthode de Newton, qui s'appuie sur un calcul de gradient pour déterminer des directions de recherche.

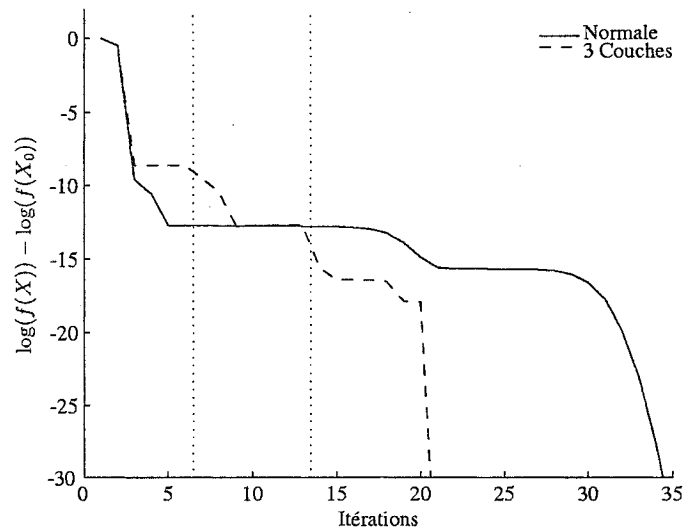


FIG. 5.3 Historique d'optimisation : valeurs de la fonction coût à la fin de chaque itération

L'optimum de la fonction est unique et peut être déterminé analytiquement. Pour des valeurs de $\{0;0;0,5;-0,5;0,3;-0,2\}$ correspondant respectivement aux variables d'entrée $\{X_1, X_2, X_3, X_4, X_5, X_6\}$, l'optimum de la fonction est atteint et $f(X) = 0$. Les résultats intermédiaires donnent plus d'information quantitatives :

- L'optimisation classique se termine en 38 itérations, et un itéré final convergé à une valeur de $1,293310^{-19}$.
- La fin de l'optimisation de la deuxième couche de variables ((X_1, X_2) et (X_3, X_4)) permet, en 13 itérations, de déterminer un optimum de $8,4810^{-4}$.
- La fin de l'optimisation de la troisième couche de variables, permet, en 25 itérations, de déterminer un optimum de $7,196110^{-29}$.

Ces chiffres permettent de montrer, sur cet exemple, que :

- Avec moins de la moitié des itérations, la méthodologie proposée atteint une bonne approximation de l'optimum. En effet, la fin de l'optimisation qui tient compte de deux couches de variables consomme beaucoup moins d'itérations pour un résultat donnant une bonne approximation.

- Avec un nombre comparable d’itérations, la stratégie permet d’obtenir un résultat beaucoup plus précis. En effet, la fin de l’étape d’optimisation contenant les trois couches de variables parvient à une valeur de fonction coût de l’ordre de 10^{-29} , comparativement à 10^{-19} , pour l’optimisation classique.

Des fonctions polynômes dotés de coefficients de pondérations d’ordres de grandeur plus proches ont également été testés. L’économie d’itérations pour parvenir à l’optimum diminue alors rapidement.

5.2.3 Conditions nécessaire d’efficacité

La stratégie fonctionne efficacement sur cet exemple, parce que plusieurs conditions sur la définition du problème sont remplies. Pour que l’application de cette méthode à d’autres problèmes soit pertinente, les caractéristiques suivantes doivent être vérifiées :

- Les résultats de l’analyse de sensibilité doivent être quantitativement fiables. Sur le cas test, il est simple de vérifier les coefficients du modèle de surface de réponse, puisque la formulation initiale ne comporte que des termes que le modèle de surface de réponse est capable d’identifier. Dans les autres cas, seul le coefficient de détermination permet de juger de la qualité de l’approximation construite. Pour le cas test, cette mesure donne un excellent modèle (coefficient de détermination $R = 99$).
- Les ordres de grandeurs des coefficients devraient permettre d’établir facilement des critères pour séparer les variables en couches. En effet, les fonctions dont les coefficients de pondérations avaient des ordres de grandeur plus rapprochés ne permettaient pas d’obtenir des économies d’itérations aussi intéressantes, comparativement au cas de la fonction définie par l’équation 5.1.

Ces conditions sont nécessaires pour que la stratégie soit efficace. Si ces deux conditions ne sont pas simultanément réunies, un nombre plus élevé de calculs devra très probablement être effectué pour parvenir à la même solution, comparativement à une optimisation classique n’utilisant aucune couche de variables. En effet, la méthode proposée termine

le processus avec une étape comprenant le problème initial, sans aucune suspension de variable. Le but de la stratégie est de débiter cette dernière étape à partir d'un point plus proche de l'optimum que le point initial du processus.

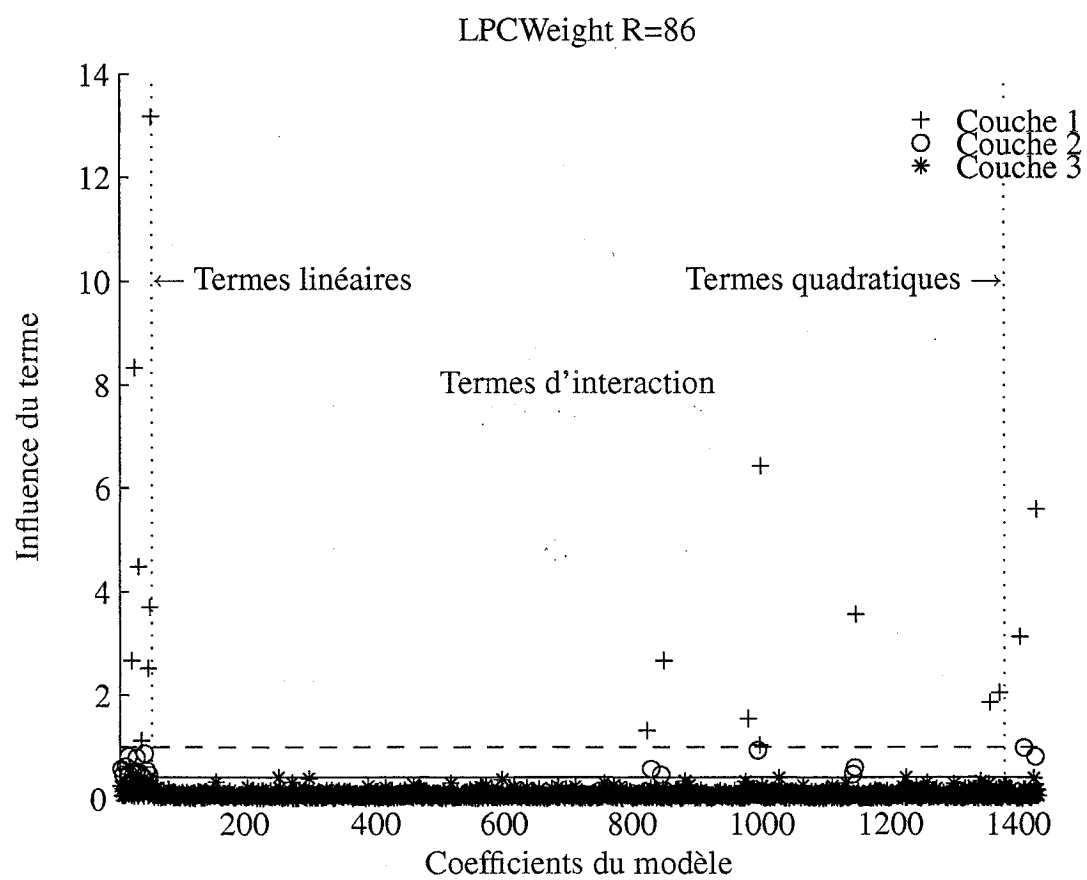
CHAPITRE 6

RÉSULTATS DES OPTIMISATIONS DU SYSTÈME PMDO

Ce chapitre présente dans la première section l'application de la stratégie au système PMDO, et caractérise le problème d'optimisation multidisciplinaire par rapport au problème idéal pour lequel la méthodologie proposée est efficace. La deuxième section expose les résultats des optimisations en appliquant la stratégie avec Isight, l'optimiseur intégré au prototype. Pour fins de comparaison, un optimiseur à gradient a également été intégré au prototype et les résultats figurent dans la troisième section. Enfin, une critique de la stratégie est développée dans la quatrième section et plusieurs pistes pour un développement ultérieur sont exposées dans la dernière section.

6.1 Caractéristiques du système PMDO

La figure 6.1 montre un exemple représentatif des caractéristiques des surfaces de réponses obtenues dans l'analyse de sensibilité. Les pondérations ne sont pas aussi clairement identifiées que dans le cas du problème test. De plus, la différence d'amplitude entre les coefficients du modèle reste limitée, ce qui éloigne PMDO du problème idéal pour être résolu par cette méthode. Ceci n'implique pas que la stratégie ne fonctionnera pas, mais que l'économie d'itérations recherchée ne sera probablement pas aussi intéressante que dans l'exemple du cas-test.



6.2 Optimisations avec Isight

Le principe de la stratégie est fonctionnel. Plusieurs routines Matlab permettent le tri des variables, l'analyse de sensibilité, et le classement qui en découle. Les optimisations menées avec Isight utilisent un optimiseur à gradient NLPQL, calculé par différences finie.

6.2.1 Fonction coût

Le problème d'optimisation configuré automatiquement au moyen de fichiers dynamiques comporte des contraintes définies pour assurer la consistance des couplages. Certaines variables de sortie sont définies comme des résidus dont la convergence détermine la résolution de couplages internes aux disciplines. Dans Isight, des bornes sévères sont définies comme contraintes, en plus d'un objectif de minimisation, pour s'assurer d'une pénalisation si ces couplages internes ne sont pas bien résolus. L'objectif de minimisation global porte donc sur la somme du poids et des volumes des composantes.

6.2.2 Historique d'optimisation

L'historique (voir figure 6.2) montre la succession des optimisations. Chaque nouvelle couche a pour point de départ le meilleur point de la couche qui le précède. Le graphique présente toutes les évaluations de la fonction coût, soit toutes les composantes du gradient et les recherches de ligne.

L'utilisation de couches permet de trouver un meilleur point qu'une optimisation classique, en un temps de calcul rallongé. En effet, l'optimisation classique permet de di-

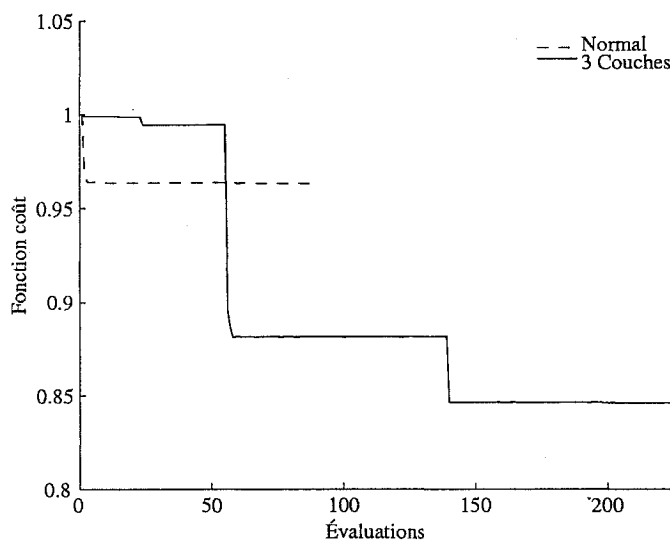


FIG. 6.2 Historique de la fonction coût

minuer la fonction coût d'environ 4%. L'optimisation de toutes les variables dans la stratégie par couches permet de diminuer la fonction coût dans les mêmes proportions, tout en ayant profité des réductions des couches précédentes, soit une diminution totale de 16%. La première et deuxième couche de variables ont permis de diminuer la fonction coût dans une proportion supérieure à l'optimisation classique, probablement parce que le nombre de variables limité a évité à l'optimiseur de se perdre avec un trop grand nombre de variables.

L'application de la stratégie d'optimisation est toutefois un succès limité à cause du nombre d'itérations plus élevé. En effet, le nombre d'itérations requises pour l'ensemble de la stratégie est plus grand que le nombre d'itérations requises pour une optimisation classique considérant toutes les variables. Ensuite, pour chaque étape, seulement un calcul de gradient et une recherche de ligne a été effectuée, soit une itération de l'optimiseur maître complétée. L'optimisation s'est arrêtée malgré plusieurs réglages de critères de descente et d'arrêt. Ces critères et la configuration de l'optimiseur sont identiques aux

disciplines, sur lesquelles des optimisations permettent des itérations normales. Il semble que les problèmes rencontrés par l'optimiseur soient dûs à plusieurs raisons :

- Une très faible variation des valeurs des variables du problème maître engendre souvent l'apparition d'erreurs dans l'exécution des disciplines, même si les optimisations semblent bien se dérouler.
- Les fichiers de journal des optimisations disciplinaires montrent souvent des matrices mal conditionnées, d'où des directions de recherche non pertinentes.
- Enfin, les recherches linéaires requièrent trop d'itérations : la fonction coût ne varie que trop peu pour trouver un nouveau point. La modification de la limite du nombre d'itérations de recherche linéaire n'aurait pour effet que de rallonger un processus d'optimisation qui semble déjà perdu.

6.3 Implémentation d'un optimiseur à gradient

Devant ces limitations observées dans Isight, le remplacement du module d'optimisation a été tenté avec un programme en C exécutable. Suivant le même protocole de communication que Isight, ce code s'est avéré capable de communiquer plus rapidement avec le JDM qui gère l'exécution des calculs.

6.3.1 Fonction coût utilisée

Toutes les contraintes gérées auparavant par Isight doivent être gérées manuellement dans le cas de ce nouvel optimiseur. Une méthode de pénalisation des contraintes a été choisie pour intégrer la consistance des couplages entre les disciplines ainsi que la convergence des résidus internes aux disciplines. Afin d'accorder autant d'importance à la minimisation du poids et du volume qu'à la résolution des couplages, les contraintes sont pondérées afin que leurs variations soient du même ordre de grandeur que les vo-

lumes et les poids des composantes.

Ainsi, les valeurs des fonctions coûts dans cette section ne peuvent être comparées aux valeurs dans la section précédente issues d'optimisation dans Isight, puisque la fonction coût a une définition différente. Les pondérations utilisées permettent de comparer les variations des variables de sortie sur une base commune et de contribuer à la fonction coût dans une proportion semblable. Ainsi, les contraintes de couplage interdisciplinaires sont pondérées de 100, les poids et les volumes des composantes sont pondérées de 1, et les résidus des disciplines sont pondérées par 100.

6.3.2 Réglages de l'algorithme d'optimisation

L'utilisation d'un optimiseur à gradient en remplacement de Isight a nécessité l'introduction de nouveaux termes dans la fonction coût. Toutefois, l'introduction de ce module a permis des améliorations dans l'algorithme d'optimisation, pour traiter plus efficacement le problème. Cette section présente plusieurs fonctionnalités ajoutées à un code pour permettre d'optimiser le système PMDO.

L'optimiseur devrait gérer un problème adimensionnel : les variables d'entrée par rapport à leur intervalle, et les variables de sortie par rapport à un intervalle typique de variation. Au cours de l'optimisation, la région de l'optimum se précise, et l'adimensionnalisation doit exploiter ce raffinement en exploitant un intervalle de variation restreint pour les variables de sortie, afin d'accentuer les améliorations ou dépréciations de la fonction coût. Dans le problème de PMDO, l'adimensionnalisation de la fonction coût est ainsi déterminée à partir des meilleurs points de l'itération précédente.

La détermination des directions de recherche est capitale à la réussite d'une optimisation par algorithmes à gradients. Tout problème dans l'exécution d'une discipline doit

pouvoir être géré par l'optimiseur. Un temps non négligeable fut perdu en raison de la mauvaise gestion de valeurs singulières par Isight. Des mécanismes furent alors développés pour que la lecture de telles valeurs dans la base de données soit traduite en un nombre élevé mais non infini pour traduire l'infaisabilité de telle exécution. Mais même dans ce cas de figure, Isight utilisant une différence avant de calcul du gradient, toute valeur extrême des composantes entraîne une mauvaise détermination de direction de recherche. La direction de recherche associée à un calcul invalide devrait être nulle. La correction de la valeur de la fonction coût devrait alors être faite soit par un mécanisme de filtrage dans le calcul du gradient (comme dans le cas de l'optimiseur à gradient intégré), soit par un remplacement de la valeur élevée par la composante initiale du gradient, afin que la direction associée soit nulle. Toutefois, dans le cadre de la recherche linéaire, la valeur élevée doit être conservée, afin que l'algorithme continue à chercher un pas vers le point optimal.

Enfin, le calcul des directions de recherche dans les composantes du gradient doit également, pour être efficace, comporter un mécanisme d'adaptation du pas de différentiation pour conserver l'hypothèse de petite variation. En effet, même si le pas de différentiation est relatif à l'intervalle (variables adimensionnées), un tel changement peut provoquer une forte variation de la fonction coût, suivant la variable considérée. Le calcul des composantes doit alors tenter d'autres pas de différentiations jusqu'à obtenir des variations acceptables de la fonction coût. Dans le cas de PMDO, ce pas de différentiation relatif est réduit jusqu'à ce que la variation dans les composantes du gradient soit inférieure à une fraction de la meilleure valeur de l'itération précédente :

- Meilleure valeur issue de l'itération précédente $BestF = f(X)$
- Pas de différentiation par défaut relatif à l'intervalle de la composante : $PasDiff = 0.01 \cdot (X_{i,sup} - X_{i,inf})$
- Calcul de la composante $f(X + \Delta X_i)$, $\Delta X_i = PasDiff \cdot (X_{i,sup} - X_{i,inf})$

- **while** $\frac{f(X + \Delta X_i) - BestF}{BestF} \leq k_1 \cdot BestF$
- $PasDiff = \frac{PasDiff}{2}$
- Calcul de la composante $f(X + \Delta X_i)$, $\Delta X_i = PasDiff \cdot (X_{i,sup} - X_{i,inf})$
- Fin boucle *while*

6.3.3 Calcul du gradient

En toute rigueur, un calcul de gradient par différences finies au niveau du problème maître nécessite plusieurs sous-optimisations pour déterminer les valeurs de la fonction coût pour chaque composante. Une possibilité, pour gagner du temps, consiste à calculer ce gradient du problème maître avec des approximations. En effet, N variables d'entrée au problème maître requièrent au moins $N + 1$ évaluations de la fonction coût pour une différence avant d'ordre 1, soit $N + 1$ résultats des disciplines. L'approximation consiste à utiliser des résultats d'exécutions de disciplines plutôt que des résultats de sous-optimisations. Cette approximation reste valable si la sensibilité du sous-problème aux variables globales est faible.

Les figures 6.3 et 6.4 montrent le résultat d'une itération d'optimisation du problème maître, soit un calcul de gradient par différences centrées (à 84 variables, soit 168 évaluations) et une recherche de ligne (9 évaluations), utilisant respectivement des résultats de sous-optimisations des disciplines et des résultats d'exécutions. La recherche linéaire est toujours réalisée avec des sous-optimisations des disciplines. Les graphiques montrent que la fonction coût est soumise à de fortes discontinuités. En effet, plusieurs composantes du gradient montrent des valeurs supérieures à 65000. Les valeurs singulières des valeurs de sortie (telles que *NaN*, ou infini (*infy*)) sont traduites sous forme numérique (200000, valeur numérique supérieure à tout résultat issu d'une analyse valide). Pendant le calcul des composantes du gradient, toutes les variations significatives de la

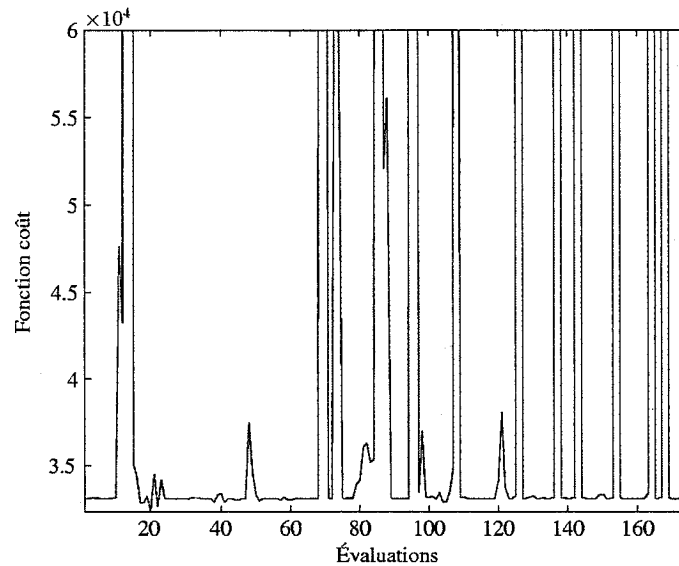


FIG. 6.3 Historique de la fonction coût : un calcul de gradient avec optimisation des disciplines et une recherche linéaire

fonction coût sont positives : aucune variable globale (gérée par l'optimiseur maître) ne permet une diminution importante de la fonction coût. Toutefois, les deux optimisations convergent vers le même point (dernière évaluation de la fonction coût, à la fin de la recherche linéaire).

Les historiques (figures 6.3 et 6.4) montrent deux choses :

- L'approximation du calcul de gradient (par exécutions disciplinaires au lieu de sous-optimisations) est pertinente. En effet, les calculs montrent que les mêmes directions de recherche sont déterminées, et que les mêmes points sont trouvés par l'optimisation, à la fin de l'itération. La réduction de poids est identique.
- Les variables de l'optimiseur maître ont une influence très forte sur les disciplines, et une faible variation de celles-ci rendent les disciplines infaisables, comme le montrent les fortes discontinuités des figures 6.3 et 6.4.

De plus, ces deux optimisations révèlent que peu d'améliorations sont envisageables à travers la modification des valeurs des variables du problème maître. En effet, une

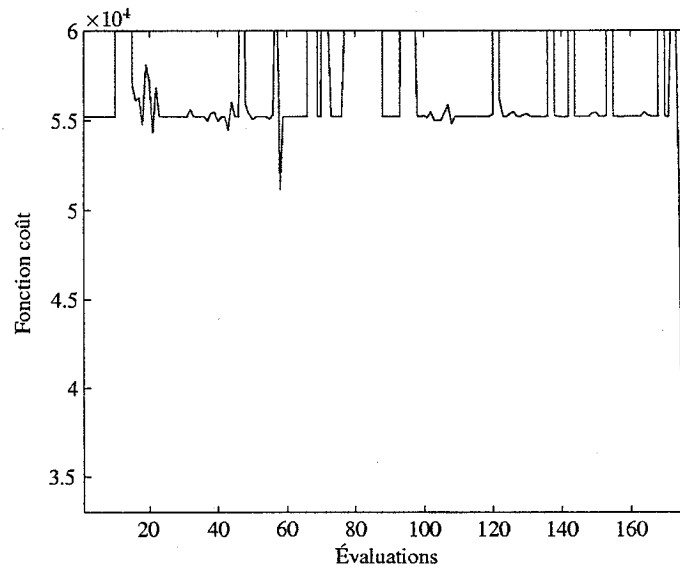


FIG. 6.4 Historique de la fonction coût : un calcul de gradient avec exécution des disciplines et une recherche linéaire

optimisation limitée aux variables locales des disciplines à partir du point de départ, sans aucun changement des variables globales du problème maître, permet de trouver un point optimal dont la fonction coût est très similaire à celle déterminée après une itération du problème d'optimisation complet.

6.3.4 Optimisations avec les couches de variables

La section précédente a montré que l'optimisation du problème maître ne pouvait pas mener à des résultats pertinents. En effet, les sous-optimisations disciplinaires sur les variables locales semblent être capables d'atteindre les meilleurs réductions possibles, sans modifications des valeurs des variables globales de l'optimiseur maître. En effet, les figures 6.3 et 6.4 montrent que le point obtenu par la recherche linéaire de l'optimiseur maître n'est pas meilleur que certains points déterminés par une sous-optimisation des disciplines.

Pour déterminer des réductions de fonction coût obtenues par les optimisations précédentes, mais avec un processus plus court, les variables globales de l'optimiseur maître vont rester fixées, et les optimisations sont menées sur les disciplines avec les variables locales seulement. Toutefois, si seulement ces disciplines sont optimisées, plus aucun mécanisme ne peut assurer la consistance des couplages. En effet, les variables locales ont une influence sur certaines variables de sortie, potentiellement couplées. Or les variables couplées, lorsqu'elles sont variables d'entrée, sont globales (gérées par l'optimiseur maître). La consistance des valeurs des variables couplées étant assurée seulement par la tâche maître dans la décomposition DO, la divergence des contraintes de couplage est plus que probable. Un mécanisme postérieur à l'optimisation doit s'assurer de la résolution des couplages. Ce mécanisme consiste à injecter les valeurs des variables couplées obtenues en sortie, comme nouvelles valeurs d'entrée des disciplines concernées. L'exécution des disciplines mène à la définition d'un écart entre la valeur des variables couplées injectée en entrée et à leurs nouvelles valeurs obtenues en sortie. La répétition du processus permet de réduire les inconsistances de couplage au prix d'une augmentation de la fonction coût. Dès que la somme des inconsistances du problème courant est inférieure à la somme des inconsistances du point initial, le processus est terminé.

Le tableau 6.1 montre les résultats d'optimisation avec les différentes couches de variables locales. Le point de départ est le même pour toutes les expériences. Le Test 1 est une optimisation n'utilisant que la première couche de variables. Le Test 2 utilise les variables de la couche 2, mais part du même point de départ. Enfin le Test 3 est une optimisation des sous-disciplines utilisant toutes les variables. Le premier résultat est le point initial, comme base de comparaison.

Le tableau de résultats montre que l'optimisation des couches 2 et 3 mène à une valeur de fonction coût supérieure à l'optimisation de la couche 1. Un trop grand nombre de variables à gérer pourrait expliquer que l'optimiseur ne soit pas parvenu à déterminer

	Moteur de référence	Test 1	Test 2	Test 3
Soufflante	1183,34	1183,34	1183,34	1183,34
Compresseur basse pression	358,94	334,33	334,33	330,21
Compresseur haute pression	548,80	511,45	511,45	511,45
Chambre de combustion	112,64	109,57	109,57	109,57
Turbine haute pression	354,59	129,01	139,56	139,56
Turbine basse pression	619,99	590,02	595,31	595,31
Mélangeur	60,84	60,84	60,84	60,84
Sortie	150,07	150,07	150,07	150,07
Boîte de vitesse	160,21	160,21	160,21	160,21
Autres composants	306,17	306,17	306,17	306,17
Marge	206,65	206,65	206,65	206,65
Nacelle	1808,10	1808,10	1808,10	1808,10
Poids Total	5870,35	5549,77	5565,61	5561,50

TAB. 6.1 Répartition du poids selon les composantes pour les 3 cas

des directions de recherche précises. En effet, l'optimisation de la couche 1 est suffisante pour obtenir un résultat équivalent et mène à une réduction plus importante de la fonction coût en comparaison d'une optimisation classique faisant intervenir toutes les variables locales. Ce constat permet de tirer deux remarques :

- L'analyse de sensibilité réalisée sur les disciplines est pertinente : l'optimisation sur ces seules variables locales permet de déterminer un point dont la fonction coût est équivalente à celle obtenue lors d'une optimisation sur toutes les variables des disciplines.
- L'application de la stratégie par couches sur ce problème n'est pas pertinente, ou du moins non efficace, à partir du point initial fourni. En effet, la minimisation de la fonction coût est maximale dès l'optimisation de la couche 1.

Les résultats de la section précédente (Optimisation avec Isight) montraient que chaque couche permettait une amélioration de la fonction coût. Dans la section courante, seules les variables locales sont optimisées, tandis que dans le cas de Isight, les variables globales étaient également prises en compte.

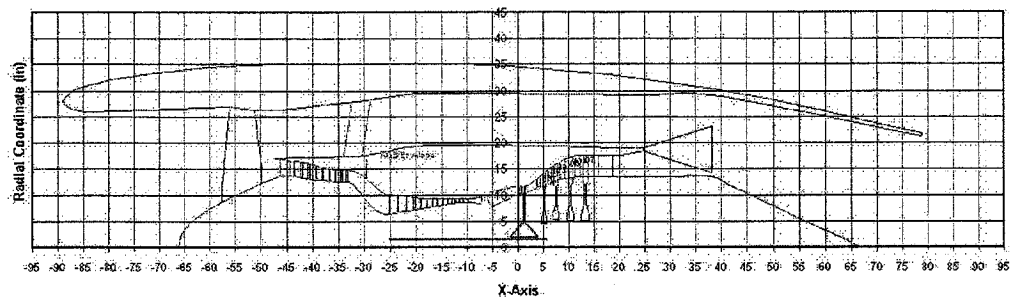


FIG. 6.6 Coupe du moteur : Optimisation de la couche 1, Poids total 5550 livres

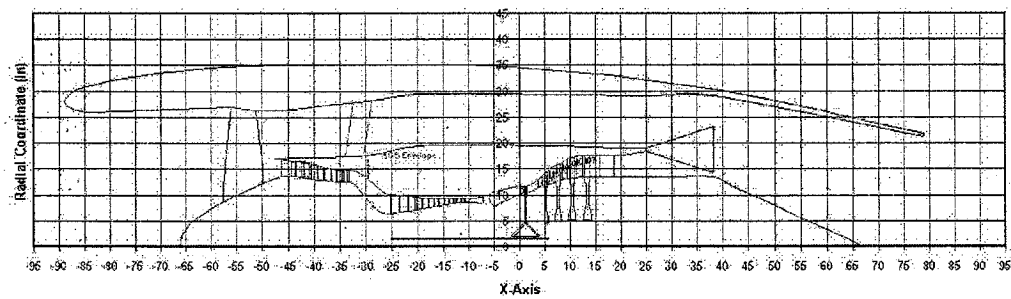


FIG. 6.7 Coupe du moteur : Optimisation de la couche 2, Poids total 5566 livres

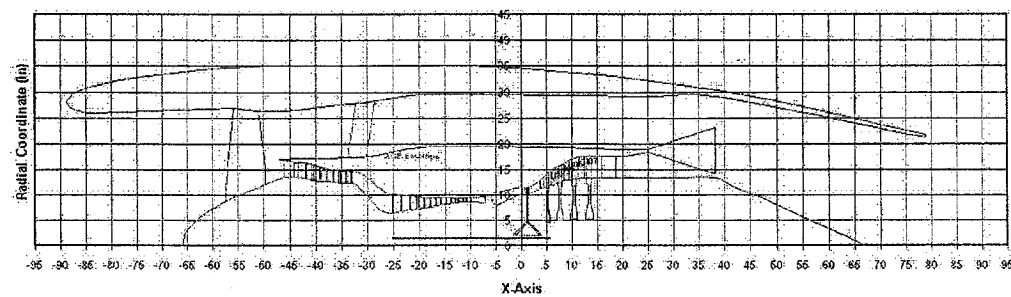


FIG. 6.8 Coupe du moteur : Optimisation de la couche 3, Poids total 5561 livres

6.3.5 Bilan de l'optimisation par gradients

L'historique de l'optimisation du problème maître (figures 6.3 et 6.4) a montré que l'amélioration du point initial n'est pas possible par les variables globales de la tâche maître. La stratégie d'optimisation par couches a donc été testée sur les sous-optimisations disciplinaires. La première couche a permis de déterminer toute l'amélioration possible, à partir du point de départ donné. Il n'est donc d'aucun intérêt de poursuivre l'optimisation en utilisant les autres couches de variables. L'analyse de sensibilité est très efficace, puisqu'elle a permis de déterminer l'amélioration possible du point de design dans un temps réduit proportionnellement au nombre de variables locales :

- Couche 1 : 111 variables locales pour une optimisation en 3 minutes 9 secondes.
- Couche 2 : 259 variables locales pour une optimisation en 5 minutes 14 secondes.
- Couche 3 : 274 variables locales pour une optimisation en 5 minutes 36 secondes.

Les optimisations menées avec Isight requéraient des temps de calcul beaucoup plus longs, de l'ordre de plusieurs jours, mais les variables et le problème d'optimisation ne sont pas les mêmes.

6.4 Discussion critique des résultats obtenus

6.4.1 Réduction de poids

L'objectif d'optimisation est donc rempli :

- Considérant la turbomachine sans la turbine haute pression, le poids a été réduit de près de 2%.
- Le processus d'optimisation est réalisé dans un temps court, considérant le nombre de variables prises en compte et les performances du prototype PMDO.

Certaines contraintes non explicitées mènent à des problèmes mineurs de définition de géométrie. La réduction de poids semble intéressante pour une application en industrie : un poids réaliste est atteint dans un temps très raisonnable.

La méthodologie a permis de déterminer une méthode rapide pour trouver l'optimum local du point fourni. En effet, le temps nécessaire à toutes les analyses de sensibilité pour l'étude des variables, en excluant l'analyse des couplages qui peut s'avérer délicate, est d'approximativement 24 heures. Une optimisation séquentielle de chacune des disciplines nécessite suivant la couche entre 6 (avec toutes les variables) et 3 minutes (avec seulement les variables de la première couche). Par ailleurs, le temps de résolution du couplage est négligeable en comparaison du temps d'optimisation. En effet, deux ou trois itérations semblent suffire, et nécessitent moins d'une dizaine de secondes. Un calcul de gradient par différences centrées de l'optimiseur maître nécessitant des optimisations disciplinaires pour chaque point évalué, nécessite alors entre 1008 et 504 minutes pour 84 variables globales. Dans un cas similaire à celui étudié, au delà de trois calculs de gradients, la différence est supérieure au temps d'analyse des variables. Pour obtenir une approximation de l'optimum qui, d'après la pertinence de l'analyse de sensibilité, soit fiable, l'utilisation de la première couche de variables est alors suffisante, et permet d'économiser un temps non négligeable.

6.4.2 Problème d'optimisation

La démarche a permis de déterminer que :

- L'optimisation des variables du problème maître ne permettait pas d'améliorer le point de design fourni, probablement parce que le point initial était déjà optimisé par rapport à ces variables.
- L'optimisation du problème au niveau disciplinaire uniquement montrait la pertinence de l'analyse de sensibilité, et du classement des variables sélectionnées.
- La résolution des couplages dans une phase post-optimisation, n'influe pas sur les

dimensions des composants, ni leur poids, mais uniquement sur les couplages internes des disciplines.

La stratégie proposée, dans sa configuration actuelle, avec les critères et paramètres d'optimisation ainsi réglés, ne permet pas d'améliorer le point optimal trouvé (i.e., la réduction de poids maximale trouvée), ni de diminuer le temps de résolution à partir du point initial fourni. De plus, la résolution des couplages internes aux disciplines et externes (entre les disciplines) est meilleure que celle de la géométrie fournie : la pertinence du résultat trouvé n'est donc pas à remettre en cause. La section suivante présente les voies d'améliorations possibles.

Sans modifications de la stratégie, une géométrie plus performante en terme de poids des composantes pourrait être déterminée en débutant l'optimisation à partir d'un autre point initial faisable différent de celui fourni afin d'explorer une autre région de l'espace de conception. Mais cette tâche est, selon les ingénieurs de P&WC, ardue, car beaucoup de contraintes limitent les possibilités, bien que l'espace de conception soit large. La méthodologie appliquée habituellement chez P&WC est d'exploiter un nombre de variables très restreint, comme le cycle thermodynamique, et d'autres variables de haut niveau, que le présent projet avait choisi de fixer pour faciliter la tâche de l'optimiseur.

6.5 Orientation des travaux futurs

Le but de cette section est d'identifier les méthodes proposées dans la stratégie qui se sont montrées efficaces dans le cas de figure étudié (considérant le problème industriel et le point de départ pour l'optimisation), et de présenter des voies de développements pour améliorer la démarche utilisée.

6.5.1 Analyse de sensibilité

L'analyse de sensibilité avait pour objectif de déterminer les variables d'entrée qui ont une influence significative sur les variables de sortie constituant la fonction coût. Les critères utilisés permettent de s'assurer que la liste constituée reste restreinte à un nombre limité de variables.

Les résultats ont montré que la méthodologie développée est performante pour une sélection d'un nombre restreint de variables influentes. En effet, la couche 1 est suffisante pour déterminer un optimum proche de celui trouvé en considérant toutes les variables. Cette méthode peut donc être utilisée comme un outil d'analyse de sensibilité pertinent, en dehors de la stratégie globale d'optimisation. Toutefois, la pertinence de la sélection des autres couches n'a pu être vérifiée, et dans le cadre de la stratégie et de l'optimisation multi-étapes, certains facteurs de cette méthode pourraient être étudiés pour améliorer ses performances :

- La modification des seuils aurait certainement un impact sur le degré de simplification du problème. Toutefois, des compromis sont alors à déterminer entre la simplification acceptée et le temps de calcul. Une étude intéressante à mener déterminerait le temps nécessaire de résolution du problème en fonction du degré de simplification de la première couche.
- Il serait également pertinent de valider les différences d'influences des coefficients du modèle de surface de réponse utilisé. En effet, un terme d'interaction ou quadratique pourrait être plus important qu'un terme linéaire dans le point de vue d'une optimisation à gradient, en raison de ses dérivées non constantes. De plus, les coefficients d'interaction donnent leur influence à deux variables. Connaître les influences spécifiques des coefficients quadratiques, d'interactions et des linéaires, permettrait de déterminer si une mise à l'échelle serait profitable pour améliorer la sélection des variables. Toutefois, cette valorisation n'aurait d'impact que pour un algorithme d'optimisation à gradient.

6.5.2 Étude des couplages

Ce traitement avait pour objectif de simplifier davantage l'analyse multidisciplinaire. L'utilisation de méthodes proposées par Rogers et al. (1996) et Sobieski & Kroo (1996) aurait conduit inévitablement à une simplification de liens de couplages. La méthodologie proposée avait pour but de déterminer des liens de couplage incompatibles par rapport aux répartitions de valeurs de variables dans leur intervalle de définition. Cette méthode de détermination proposée est absolue pour un couplage et non relative à l'intensité des autres couplages du système. Elle permettrait à un utilisateur à s'intéresser à des couplages spécifiques, les plus problématiques (liens de retour), pour tenter de les retirer temporairement du système, sans avoir à effectuer une étude de tous les couplages en présence, et les comparer par la suite.

La pertinence de cette analyse reste toutefois non prouvée. En effet, d'une part, le problème simple sur lequel la stratégie d'optimisation par couches a été testée avec succès, n'a pas permis de vérifier l'utilité de cette méthode, puisque cette formulation analytique ne comportait aucun couplage. D'autre part, les couplages problématiques des liens de retour n'ont pas pu être retirés de l'analyse multidisciplinaire du système PMDO. Ces couplages étaient peut être trop intenses pour être suspendus de l'analyse multidisciplinaire. Toutefois, certains aspects seraient à prendre en considération si une nouvelle étude devait être entreprise :

- Dans l'étude, des surfaces de réponses prenant en compte des termes linéaires, d'interaction, et quadratiques sont déterminés. Mais seuls les termes linéaires sont étudiés lors de l'analyse des couplages. Une nouvelle étude devrait se limiter à des modèles de surface de réponse utilisant des termes linéaires, à moins d'interpréter les autres termes par rapport à une tendance à travers plusieurs intervalles.
- Les données récupérées des modèles furent utilisées sans égard au coefficient de détermination (l'erreur d'approximation du modèle). Il est donc possible que des modèles très approximatifs furent exploités. D'autant plus que seuls les termes linéaires ont été

récupérés, sans considérer la variation des variables de sortie expliquée par les termes d'interaction et quadratiques. Il serait pertinent de gérer les cas où l'approximation soit tellement mauvaise que le modèle construit ne peut être exploité pour réaliser l'étude.

- Des critères quantitatifs plus précis devraient être définis pour déterminer l'incompatibilité d'intervalles disjoints ou faiblement recouverts entre 2 disciplines.

CONCLUSION

Le projet PMDO répondait à l'objectif de MOSAIC qui consistait en l'intégration d'outils d'analyse et d'optimisation dans des systèmes multidisciplinaires pour la conception de composants aéronautiques. Le mémoire présente une stratégie de résolution du problème d'optimisation multidisciplinaire fourni par le partenaire industriel Pratt & Whitney Canada, s'appuyant sur le prototype établi par M. Prado. Les méthodes composant cette stratégie se concentrent sur deux aspects fondamentaux des analyses multidisciplinaires : l'analyse de sensibilité des variables et l'étude des couplages. L'idée maîtresse de la stratégie est l'établissement d'une optimisation en trois étapes, pour un rapprochement progressif vers la région de l'optimum.

L'analyse de sensibilité, basée sur la construction d'une surface de réponse, et l'étude de ses coefficients de pondération liées aux variables d'entrée, a permis de sélectionner un nombre réduit de variables influentes sur des variables de sortie constituant la fonction coût. Trois représentations du problème d'optimisation, de complexité croissante ont alors été construits à partir d'un classement de variables d'influence décroissante. Cette méthode de sélection est pertinente. En effet, l'optimum déterminé avec le problème simplifié est comparable à celui trouvé avec le problème original, pour un temps de résolution fortement réduit. La méthode pourrait être encore améliorée par une analyse des coefficients en fonction de leur type de contribution. La sélection de variables peut également être plus stricte pour agir sur le degré de simplification du premier problème, et rendre la stratégie d'optimisation globale potentiellement plus efficace.

L'étude proposée des couplages permet un examen spécifique de liens dans l'analyse multidisciplinaire. Les disciplines nécessitant des ensembles de valeurs de variables

d'entrée consistants, le transport du couplage peut être perturbé par ces répartitions singulières dans l'espace de conception. Par ailleurs, l'analyse multidisciplinaire la plus simplifiée présentait encore des couplages qui conservaient une complexité élevée au problème d'optimisation.

L'application de cette étude au problème PMDO a montré que cette méthode nécessitait des évolutions pour être efficace. Néanmoins, l'étude menée sur les liens de couplages a utilisé de nombreuses approximations. Par ailleurs, plusieurs aspects laissent une marge d'amélioration à la pertinence de cette méthode. Son application sur d'autres cas de figure permettrait de déterminer si cette méthode de caractérisation est viable.

La stratégie globale d'optimisation fut établie pour exploiter des algorithmes d'optimisation à gradients. Plusieurs recommandations sont formulées pour l'utilisation de ce type d'algorithme telles que l'adimensionalisation du problème, la gestion des erreurs d'exécution des analyses, et un pas de différentiation adaptatif. Ces aspects vérifiés permettent principalement un calcul des directions de recherche pertinent, pour accélérer la convergence vers le point optimal. Si les codes commerciaux n'incluent pas ces fonctionnalités, il est en général plus simple de construire un module externe capable de gérer ces opérations. De telles modifications aux optimiseurs dont les codes sources sont disponibles sont potentiellement plus simples, au détriment de la simplicité d'utilisation de logiciels commerciaux.

L'idée majeure de ce processus d'optimisation n'a pu être testée que sur une formulation analytique simple, et non sur le problème d'optimisation proposé par le partenaire industriel. Toutefois, cette approche est inspirée par les méthodologies employées par P&WC. En effet, les tentatives d'optimisation ont montré, sur le point initial fourni, que le moteur semblait déjà optimisé par rapport aux variables du problème maître, les variables locales associées aux disciplines étant très probablement fixées. Seule la première partie

(l'analyse de sensibilité) de cette stratégie a pu être validée par l'application au problème d'optimisation industriel.

Les expériences menées avec le logiciel Isight montrent toutefois que le processus est fonctionnel et que son automatisation est réalisable. L'application de la stratégie au problème a mené à un succès mitigé, en raison principalement du point initial, partiellement optimisé. Ceci a induit un temps de résolution triple du temps de résolution normal. Toutefois, la réduction de la fonction coût plus importante démontre l'intérêt de déterminer un nombre réduit de variables influentes. En effet, l'optimiseur parvient à déterminer des directions de recherche pertinentes pour remplir l'objectif de minimisation.

La stratégie présentée est efficace pour déterminer des optimums locaux, plus ou moins éloignés du point duquel le processus est lancé. Afin de rendre cette stratégie globale et adaptable à d'autres problèmes complexes, une méthodologie pourrait être intégrée afin de déterminer des points de départ intéressants, sans forcément utiliser d'algorithme d'optimisation (Sasena et al., 2002, Assanis et al., 1999). Le but recherché, dans ce cas de figure, est d'utiliser l'optimisation seulement pour améliorer un point de départ.

Enfin, si les résultats finaux souffrent de certaines anomalies géométriques, c'est en raison de la capacité des optimiseurs d'exploiter les failles mathématiques des analyses. Toutefois, la géométrie optimale de turbomachine déterminée est, aux dires des ingénieurs de Pratt & Whitney Canada, probablement la meilleure qu'il est possible d'obtenir, compte tenu des outils d'analyses utilisés, des variables fixées et du point de départ fourni.

Grâce à l'application d'une partie de cette stratégie, le problème d'optimisation actuel fourni par P&WC a pu être résolu en à peine plus de 3 minutes, contre plus d'une semaine dans le cadre d'une optimisation classique (utilisant une méthode de décomposition, mais sans couches de variables). Les applications futures inclueront plusieurs

points de départ, pour permettre la comparaison de différents concepts, un nombre plus réduit de variables fixées, pour explorer un espace de conception plus large, et des outils d'analyse plus élaborés. Cette complexité croissante ne peut qu'augmenter le temps de résolution. Même si P&WC a acquis un produit commercial pour faciliter l'intégration d'analyses, et que l'entreprise dispose de ressources de calcul importantes, l'utilisation d'une telle stratégie permet d'utiliser efficacement de tels moyens. Ce motoriste pourra alors explorer plus rapidement des concepts de turbomachine plus efficaces et plus puissants, et être plus compétitif sur un marché très concurrentiel.

BIBLIOGRAPHIE

- Abbeel, P. & Ng, A. (2004). Learning first-order markov models of control. *Proceedings of Neural Information Processing Systems, Vancouver, Canada*.
- Aluja-Banet, T. (1998). *Analyse en composantes principales*. Saint-Mandé (France) : CISIA-CERESTA.
- Alyaqout, S., Papalambros, P., & Ulsoy, A. (2005). Quantification and use of system coupling in decomposed design optimization problems. *Proceedings of IMECE2005, ASME International Mechanical Engineering Congress and Exposition, Orlando, Florida, USA*.
- Assanis, D., Delagrammatikas, G., & Michelena, M. (1999). Optimization approach to hybrid electric propulsion system design. *Journal of Mechanics of Structures and Machines, Automotive Research Center Special Edition Issue, Ed. E. J. Haug, 27 (4)* 393-421.
- Azarm, S. (1994). Reduction method with system analysis for multiobjective optimization-based design. *Structural Optimization, 7(1-2)*, 47-54.
- Azarm, S. & Eschenauer, H. (1993). A minimax reduction method for multi-objective decomposition-based design optimization. *Structural and Multidisciplinary Optimization, 6(2)*, 94-98.
- Buche, D., Schraudolph, N., & Koumoutsakos, P. (2004). Accelerating evolutionary algorithms with gaussian process fitness function models. *IEEE transactions of systems, man, and cybernetics*.
- Buchsbaum, A. L., Fowler, G. S., & Giancarlo, R. (2003). Improving table compression with combinatorial optimization. *Journal of the Association for Computing Machinery*

50(6), 825–851. Tiré de

<http://citeseer.ist.psu.edu/buchsbaum02improving.html>.

Chipman, J. & Winker, P. (2000). Optimal industrial classification : An application to the german industrial classification system. Econometric Society World Congress 2000 Contributed Papers 0522, Econometric Society. Tiré de

<http://ideas.repec.org/p/ecm/wc2000/0522.html>.

Dengiz, B. & Alabas, C. (2000). Simulation optimization using tabu search. *Proceedings of the 2000 Winter Simulation Conference*. <http://www.informs-cs.org/wsc00papers/prog00.htm>

Diedrich, J. (1990). *Artificial Neural Network : concept learning*. Los Alamitos, California : IEEE Computer Society Press.

Egorov, L., Kretinin, G., & Leshchenko, I. (1999). The multilevel optimization of complex engineering systems. *ISSMO, Short paper proceedings of 3rd WCSMO, New York*, pp.414-417.

Fathy, H., Reyer, J., & Papalambros, P. (2001). On the coupling between the plant and controller optimizer problems. *Proceedings of the American Control Conference, Arlington, VA*.

Jarvis, R. & Goodacre, R. (2005). Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data. *Bioinformatics*, 21(7), 860–868.

Kleijnen, J., Bettonvil, B., & Persson, F. (2003). Finding the important factors in large discrete-event simulation : sequential bifurcation and its applications. Discussion Paper 104, Tilburg University, Center for Economic Research. Tiré de <http://ideas.repec.org/p/dgr/kubcen/2003104.html>.

- Kott, G. & Gabriele, G. (1999). Performance and reliability of a new mixed variable optimization approach : The maximum distance distribution method. *WCSMO3, Niagara Falls, New York, May 17-21*.
- Kozma, R., Voicu, H., & Wong, D. (2003). Dynamical neural network algorithm for autonomous learning and navigation control. *IEEE Transactions 2003*.
- Martin, A. & Quinn, K. (1995). A review of discrete optimization algorithms. *The Political Methodologist*, 7(2), 6–10.
- McKay, M., Beckman, R., & Conover, W. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1), 55–61.
- Moerland, P. & Fiesler, E. (1996). Hardware-friendly learning algorithms for neural networks : An overview. *Proceedings of MicroNeuro*.
- Montano, J. & Palmer, A. (2003). Numeric sensitivity analysis applied to feedforward neural networks. *Neural Computation and Applications*, 12, 119–125.
- Montgomery, D. (2000). *Design and Analysis of Experiments, 5th edition*. New York : John Wiley.
- Mookerjee, V. & Mannino, M. (1997). Sequential decision models for expert system optimization. *IEEE Transactions on knowledge and data engineering*, 9(5), 675–687
- Morelle, P. (1999). New strategies for local/global optimization. *Proceedings of Optimization in Industry-II, Banff, Alberta, Canada*.
- Nye, J. (2004). Genetic algorithms : Explanations and implementation tradeoffs. Tiré de <http://www.devmaster.net/articles/genetic-algorithms2/>.

- Papalambros, P. & Michelena, N. (2001). Trends and challenges in design optimization. *Optimal Design Laboratory*. Tiré de
<http://citeseer.ist.psu.edu/papalambros00trends.html>.
- Prado, P. (2005). Développement d'un système d'optimisation multidisciplinaire flexible et adaptable à différentes méthodes de décomposition. Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, Québec, Canada.
- Resit, U., Lepsch, J., & McMillin, M. (1998). Response surface model building and multidisciplinary optimization using overdetermined d-optimal designs. *NASA Langley Technical Reports*. Tiré de
<http://techreports.larc.nasa.gov/dspace/jsp/handle/2002/13092>.
- Reyer, J. & Papalambros, P. (1999). Combined optimal design and control with application to an electric dc motor. *Journal of Mechanical Design*, 124, 183–191.
- Reyer, J. & Papalambros, P. (2000). An investigation into modeling and solution strategies for optimal design and control. *ASME 2000 Design Engineering Technical Conferences and Computers and Information Engineering Conferences*.
- Rogers, J. (1996a). DeMAID/GA an enhanced design manager's aid for intelligent decomposition. *American Institute of Aeronautics and Astronautics, NASA Langley Research Center*. Tiré de
<http://techreports.larc.nasa.gov/dspace/jsp/handle/2002/15438>.
- Rogers, J. (1996b). DeMAID/GA user's guide design manager's aid for intelligent decomposition with a genetic algorithm. *NASA Langley Research Center*. Tiré de
<http://techreports.larc.nasa.gov/dspace/jsp/handle/2002/10940>

- Rogers, J. L. & Bloebaum, C. L. (1994). Ordering design tasks based on coupling strengths. *Fifth AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Panama City, Florida. AIAA 94-4326.
- Rogers, J. L., McCulley, C. M., & Bloebaum, C. L. (1996). Integrating a Genetic Algorithm into a knowledge-based system for ordering complex design processes. Technical Memorandum 110247, NASA, Hampton, Virginia.
- Sasena, M., Papalambros, P., & Goovaerts, P. (2002). Global optimization of problems with disconnected feasible regions via surrogate modeling. *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Atlanta, Georgia*. AIAA-2002-5573 Tiré de
http://ode.engin.umich.edu/publications/papers/2002/AIAA_2002_5573.pdf
- Siah, E., Papalambros, P., & Volakis, J. (1999). Parameter optimization using the divided rectangles global algorithm with kriging interpolation surrogate modeling. *Journal of Mechanical Design*, 124, 183-191.
- Sobieski, I. & Kroo, I. (1996). Aircraft design using Collaborative Optimization. *AIAA 34th Aerospace Sciences Meeting and Exhibit*, AIAA 96-0715.
- Sobieszczanski-Sobieski, J. (1990). Sensitivity of complex, internally coupled systems. *American Institute of Aeronautics and Astronautics Journal*, 28(1), 153-160.
- Spinellis, D. & Papadopoulos, C. (1999). Production line buffer allocation : Genetic algorithm versus simulated annealing. *Second International Aegean Conference on the Analysis and Modelling of Manufacturing Systems, Greece, Tinos*.
- Tribes, C. & Trépanier, J.-Y. (2006). Distributed optimizations using optimal solution sensitivity. *11th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimi-*

zation, Portsmouth, VA, AIAA-2006-7065, (pp. 1-25). American Institute of Aeronautics and Astronautics, Inc.

Trocine, L. & Malone, L. (2000). Finding important independent variables through screening designs : A comparison of methods. Tiré de
<http://citeseer.ist.psu.edu/521280.html>.

Trocine, L. & Malone, L. (2001). Experimental design and analysis : an overview of newer, advanced screening methods for the initial phase in an experimental design. *WSC '01 : Proceedings of the 33rd conference on Winter simulation*, (pp. 169-178). Washington, DC, USA : IEEE Computer Society.

Turcotte, J. (2004). Intégration et optimisation multidisciplinaires d'une turbine à gaz au stade préliminaire de conception. Mémoire de maîtrise, École Polytechnique de Montréal, Montréal, Québec, Canada.

Unal, R. & Dean, E. (1991). Taguchi approach to design optimization for quality and cost : An overview. *Proceedings of the International Society of Parametric Analysts*.

Wagner, S., Affenzeller, M., & Schragl, D. (2004). Traps and dangers when modelling problems for genetic algorithms. *Austrian Society for Cybernetic Studies, Cybernetics and Systems, 1*, 79-84. Tiré de
<http://citeseer.ist.psu.edu/wagner04traps.html>.

Wan, H., Ankenman, B., & Nelson, B. (2003). Controlled sequential bifurcation : a new factor-screening method for discrete-event simulation. *WSC '03 : Proceedings of the 35th conference on Winter simulation*, (pp. 565-573). Winter Simulation Conference.

ANNEXE I

INTÉGRATION DES DISCIPLINES

Le système PMDO est une plateforme modulaire, flexible et reconfigurable en fonction des outils que l'utilisateur souhaite exploiter. Cette modularité est toutefois soumise à des caractéristiques spécifiques pour l'intégration de composants. Les disciplines fournies par Pratt & Whitney Canada sont un exemple de ces modules dont l'intégration était nécessaire à l'optimisation d'une turbomachine.

I.1 Outil fourni par P&WC et outils nécessaires pour PMDO

L'outil fourni par P&WC est un classeur Excel (*FlowpathV22.xls*), composé de différentes feuilles de calcul. Chaque composant de la turbomachine peut être analysé par des feuilles de calcul spécifiques (Compresseur basse pression, haute pression, Turbine, chambre de combustion, etc). Ces feuilles sont reliées les unes aux autres par des paramètres, références d'une feuille à une autre. Ces références inter-feuilles correspondent aux variables de couplages : des résultats (variables de sortie) d'une feuille sont exploités comme donnés (variables d'entrée) au sein d'une autre feuille. Par ailleurs, chaque feuille dispose de ses propres variables (variables locales). Des feuilles supplémentaires dans le classeur regroupent certaines variables d'entrée communes à plusieurs disciplines (variables partagées).

Le prototype exploite des disciplines indépendantes (fichiers Excel distincts), et non un fichier unique. Le Serveur JDM (figure 1.1) est programmé pour exécuter des analyses à partir de valeurs de variables d'entrée, et pour mettre à jour une base de données spécifique à chaque discipline. Le lecteur pourra se référer à la section Exécution d'une

discipline de l'introduction pour de plus amples détails.

La première étape d'intégration est donc de scinder le classeur Excel fourni (*FlowpathV22.xls*), en plusieurs fichiers, chacun représentant une discipline (*DisciplineX.xls*, avec $X, X \in [1..13]$ représentant le numéro de la discipline). Le JDM est ensuite en mesure d'exploiter les calculs effectués dans les feuilles à partir des informations sur les variables de la feuille.

I.2 Principe d'ajout d'une variable

Le JDM ne perçoit que les variables lors de l'exécution d'une discipline. Au lancement du prototype, le JDM lit les informations de configuration dans un fichier au format ASCII qui contient les caractéristiques de toutes les variables d'une discipline : le fichier de propriétés (*DisciplineX.properties*).

Dans le principe, l'ajout d'une variable à une discipline est une opération simple. 15 lignes décrivent les caractéristiques d'une variable :

- `x1.key=x1` donne l'identifiant unique par lequel la variable sera appelée.
- `x1.name=x1` permet de décrire la variable par un nom plus explicite.
- `x1.location = B2` décrit la localisation dans la feuille Excel de la variable.
- `x1.discipline=Discipline1` procure le nom de la discipline auquel la variable appartient.
- `x1.mode=input` représente le statut de la variable : entrée (input) ou sortie (output).
- `x1.type=real` détaille le type de variable : entier (integer) ou réel (real).
- `x1.category=local` est la catégorie par défaut de toutes les variables. Ce champ est utilisé par le Serveur dans une phase de pré-traitement, pour déterminer si la variable est locale (local), partagée (shared) ou couplée (coupled).
- `x1.value=1` est la valeur initiale par défaut mais spécifique de la variable.

- `x1.unit=m` exprime l'unité dans laquelle la valeur de la variable est donnée.
- `x1.objective=%` correspond à l'objectif d'optimisation recherché : minimisation (-1), maximisation (1) ou sans objectif pour cette variable (%).
- `x1.target=%` détermine une valeur cible à atteindre pour la variable : nombre réel ou sans valeur cible (%).
- `x1.lowerbound=0.01` explicite la borne inférieure de l'intervalle de validité de la variable.
- `x1.upperbound=10` exprime la borne supérieure de l'intervalle de validité de la variable.
- `x1.sendCoordinator=0` précise si la valeur de la variable doit être communiquée à un composant en particulier durant l'optimisation (la tâche maître dans une optimisation à deux niveaux).
- `x1.IsParameter=0` déclare la variable comme fixe. Les composants d'optimisation peuvent toutefois choisir de la modifier.

Dans un fichier de propriétés, ces lignes décrivent ainsi toutes les variables d'une discipline, qu'elles soient des variables d'entrée ou de sortie. Le statut de variable locale, partagée ou couplées n'a de sens que lorsque les fichiers de propriétés d'au moins deux disciplines sont considérés. L'étape de pré-traitement des variables, réalisée par le JDM, consiste en le recoupement de ces informations, pour déterminer ce statut. La convention de nom (key) est alors critique pour assurer une cohérence des statuts. Les 557 variables sont toutes décrites par ces 15 lignes de texte à travers les 13 fichiers de propriétés, parfois plusieurs fois pour toutes les variables couplées (variables d'entrée et de sortie de disciplines sont distinctes et doivent figurer dans le fichier de propriétés de la discipline auxquelles elles appartiennent).

	A	B	C	D	E	F
1	FAN AERODYNAMIC DESIGN:					
2						
3	PWC Fan for engine cycle					
4						
5						
6	OPERATING CONDITIONS:					
7						
8	FAN INLET:					
9						
10	P_inlet	in	(psia)	3.9524		
11	delta	in		0.268870748		
12	T_inlet	in	(R)	440.05		
13	theta	in		0.848419997		
14	RPM_fan_mech		(rpm)	5228.585386		
15	RPM_corr		(rpm)	5676.475354		
16	W_fan_inlet	tot	(lbs/sec)	181.84		
17	W_in_corr	tot	(lbs/sec)	622.9473103		
18	BPR			5.3762		
19	Fan_Rh/Rt	in		0.320		
20	Rt_out/Rt_in			0.9801	Denis Walch: RtoutperRtin	
21	Rh/Rt	out		0.463		
22	Diameter	in	(in)	53.95		
23	Efficiency_stg	ad		0.9057	<= fr. P. Townsend's fan eff spreadsheet	

FIG. I.1 Extrait de la feuille Excel d'analyse de l'hélice

I.3 Variables Locales

L'intégration d'une nouvelle variable locale est simple. Considérant la feuille décrivant la discipline 1, (figure I.1), la case D20 contient une valeur de variable d'entrée. Les informations présentes dans la feuille permettent de définir souvent simplement le nom de la variable (*RtoutperRtin*), la valeur de la variable (0,98), et les autres caractéristiques requises par le fichier de propriétés.

```

RtoutperRtin.key=RtoutperRtin
RtoutperRtin.name=RtoutperRtin
RtoutperRtin.location = D20
RtoutperRtin.discipline=Discipline1
RtoutperRtin.mode=input
RtoutperRtin.type=real
RtoutperRtin.category=local
RtoutperRtin.value=0.98
RtoutperRtin.unit=%
RtoutperRtin.objective=%

```

H9	=LPCIR63							
	A	B	C	D	E	F	G	H
1	S-DUCT: MEANLINE SPLINE							
2								
3	TEST CASE							
4								
5	AERODYNAMICS				Recommended Limit			
6	Average Slope				Max. ECA : Mach in			
7	$\Delta R / \Delta X$ (deg)				(deg)			
8	-0.700				39.489428			
9	-35.000				-1.8			
10	Overall				Overall			
11								

	A	B	C	D	E	R	S
1							
2	AXIAL COMPRESSOR DESIGN:						
3							
4	Axial Boost Stages 1 to 10-Stage Generic						
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							
46							
47							
48							
49							
50							
51							
52							
53							
54							
55							
56							
57							
58							
59							
60							
61							
62							
63							
64							
65							

FIG. I.2 Référence à une autre discipline

```

RtoutperRtin.target=%
RtoutperRtin.lowerbound=0.95
RtoutperRtin.upperbound=1.0
RtoutperRtin.sendCoordinator=0
RtoutperRtin.IsParameter=0

```

I.4 Variables couplées

Le processus est comparable lorsqu'une cellule nécessite des données de feuilles tierces. Dans le cas d'une référence à une discipline intégrée au prototype, une nouvelle variable de sortie doit être créée pour chaque référence à ces feuilles, en cohérence avec la création d'une variable d'entrée dans la feuille courante. La figure I.2 montre une telle situation. Les fichiers de propriétés des disciplines concernées doivent alors être éditées en conséquence :

– Fichier de propriétés de la Discipline 3 (*S-Duct*) :

```
LPCMnBoost15.key=LPCMnBoost15
LPCMnBoost15.location=H9
LPCMnBoost15.mode=input
...
LPCMnBoost15.value=0.40251301
...
LPCMnBoost15.lowerbound=0.2
LPCMnBoost15.upperbound=0.6
...
```

– Fichier de propriétés de la Discipline 2 (*LPC, Axial Compressor*) :

```
LPCMnBoost15.key=LPCMnBoost15
LPCMnBoost15.location=R63
LPCMnBoost15.mode=output
...
LPCMnBoost15.value=0.40251301
...
LPCMnBoost15.lowerbound=0.2
LPCMnBoost15.upperbound=0.6
...
```

I.5 Références à des feuilles non intégrées

Le processus se complexifie lorsque la référence pointe vers des feuilles qui ne sont pas intégrées au prototype (figure I.3). En effet, le classeur fourni par P&WC comporte 33 onglets ou feuilles de calcul. 18 de ces feuilles ont été intégrées dans PMDO sous la forme de 13 fichiers indépendants : certaines feuilles ont été fusionnées. Toutefois, plusieurs feuilles sont utilisées pour des calculs intermédiaires récurrents. Les étapes de calculs ont donc dû être transférées dans chaque feuille requérant ce type de données. De plus, ces feuilles effectuant des calculs intermédiaires nécessitent également des données d'autres feuilles intégrées. La figure I.4 montre la discipline 3 (*S-Duct*) qui

E18: Plot Data Lo Spool1558

A	B	C	D	E	F	G	H	I	J	K	L	
S-DUCT: MEANLINE SPLINE & AREA RULING												
TEST CASE												
AERODYNAMICS						Recommended Limit		MECHANICAL				
Average Slope				ECA		Max. ECA		Mach in		Min. Length		
$\Delta R / \Delta X$ (deg)				(deg)		(deg)				ΔX min		
0.700				-35.000		-1.8		39.493428		0.40		
Overall				Overall								
ΔX				ΔR (mean)								
8.000				-5.433								
DUCT FLOWPATH SIZING												
NOTE: Length is Driven by MECHANICAL Requirements												
		ΔX	ΔR (mean)	X (mean)	R (mean)	Local Slope	Area	Area Ratio	Height: H	Hv. axial	Hr. radial	
		(in)	(in)	(in)	(in)	$\Delta R / \Delta X$	(sq-in)		(in)	(in)	(in)	
18	INLET	0%	0.000	0.000	34.75102	1.81818	0	188.6033	1	2.321	0.000	2.321
19		5%	0.400	-0.033	37.352	13.586	-0.084	198.2811	0.998378	2.323	-0.194	2.315
20		10%	0.800	-0.133	33.952	13.466	-0.249	197.9502	0.996757	2.336	-0.584	2.267
21		15%	1.200	-0.296	33.652	13.322	-0.408	197.6375	0.995137	2.361	-0.892	2.186
22		20%	1.600	-0.519	33.152	13.089	-0.567	197.3161	0.993619	2.397	-1.186	2.095
23		25%	2.000	-0.796	32.752	12.823	-0.692	196.995	0.991902	2.445	-1.391	2.011
24		30%	2.400	-1.120	32.352	12.498	-0.810	196.6741	0.990286	2.504	-1.577	1.946

FIG. I.3 Référence à une feuille non intégrée

a de nombreuses références à ces feuilles non intégrées (*Plot Data Hi-Spool* et *Plot Data Lo-Spool*), qui elle-même ont des références vers d'autres disciplines qui sont intégrées.

I.6 Bilan

Cette procédure est longue et fastidieuse. De plus, les possibilités d'erreurs sont nombreuses entre les erreurs de localisation dans les feuilles Excel, les numéros des disciplines auxquelles les variables appartiennent, surtout lorsqu'il s'agit d'intégrer des variables couplées déjà intégrées au prototype. Ensuite, la convention de nom doit être très stricte, avec des caractères interdits, et une différenciation majuscules-minuscule rigoureuse. Enfin, il est délicat d'effectuer des vérifications autrement que par une re-lecture des fichiers. Un outil de vérification aurait pu permettre de détecter les erreurs plus rapidement : des erreurs ont encore été identifiées plus d'un an après le début de la phase d'intégration.

Toutefois, la procédure est identique pour des solutions commerciales, telles que le logiciel utilisé par Pratt & Whitney Canada. En effet, une solution logicielle pourrait permettre de déterminer la localisation des variables dans une feuille Excel, mais il revient toujours à l'utilisateur expérimenté de définir au moins une valeur initiale et des bornes

	A	B	C	D	E	F	G	H	I	J
1	Configuration Front Sheet				1 = Turbofan 0 = Turboshift / Turbo prop					
2	Fan yes_1_or_no_0				1					
3										
4	S-DUCT: MEANLINE SPLINE & AREA RULING									
5										
	R	S	T	U	V	W				
1	Plot Data Lo Spool			Plot Data Hi Spool						
2	Duct Out Rip	10.0541684	T55	Comp. Rip Stage 1	10.0541684	C6				
3	Duct Out Rhub	6.316984131	T56	Comp. Rhub Stage	6.316984131	C7				
4	Duct In Rhub	12.45764956	S56	Comp. Law Stage 1	-26.75192361	C8				
5	Duct In Rip	14.77871021	S55	Gap	-7.251987542	F41				
6	Law (neg.) Stg Out	-34.75192361	S58							
7	Law (neg.) Duct Out	-26.75192361	T58							
8	Aero Front Sheet Hi Spool			HPC						
9	HPC Stages		7 C19	Rip Stage 1	10.0541684	D133				
10	Centrifugal rear stage		0 C21	Rhub Stage 1	6.316984131	D134				
11	Combustor type		0 C25	LPC						
12	Aero Front Sheet Lo Spool			Rip stage Out	14.77871021	S133				
13	Inter Duct Law/H	3.44670012	C26	Rhub Stage Out	12.45764956	S134				
14				LPC Min Boost 15	0.40251301	R63				
				LPC H out stg 15	2.321060644	R121				

FIG. I.4 Calculs transférés dans la discipline

de définition réalistes pour les variables : choisir une variation de plus ou moins un pourcentage défini restreindra l'espace de conception à un voisinage du point initial.

ANNEXE II

REMPLACEMENT DE FONCTION SPLINES

Les outils de dimensionnement analytique sous formes de feuilles Excel confiées par le département Advanced Engineering de P&WC nécessitent des fonctions Splines spécifiques. Mais la plate-forme d'intégration ne peut gérer ces fonctions macros dont la formulation est contenue dans un fichier séparé.

II.1 Problème posé

Le but de ces splines est de présenter un chemin pour la veine gazeuse lissé à travers plusieurs composantes telles que le conduit menant du compresseur basse pression au compresseur haute pression, ou de toute portion de conduit entre deux étages de turbine. Le lecteur pourra se référer aux coupes de moteurs dans la section des résultats (figures 6.6, 6.7 et 6.8). L'utilisation d'une formulation générique et non itérative est nécessaire pour être exploitable par les feuilles de calcul et réactualiser la géométrie.

II.2 Formulations possibles

Une courbe doit donc être déterminée afin de lisser le tracé d'un conduit d'élargissement défini dans les feuilles Excel par 5 points (figure II.1). La courbe tracée est la courbe idéale à atteindre. Les deux points extrêmes ont une ordonnée identique.

Une première idée pour une courbe de lissage de ces points consiste en l'utilisation de

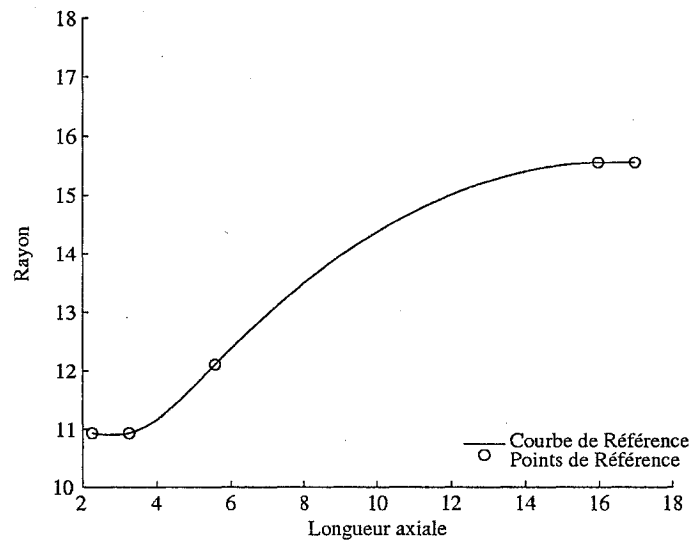


FIG. II.1 Définition du problème de lissage

polynômes du second ordre, du type :

$$f(x) = ax^2 + bx + c \quad (\text{II.1})$$

Avec 3 inconnues à déterminer par fonction, et 4 splines nécessaires, le problème est défini par 12 inconnues. Les coordonnées de 5 points donnent 8 contraintes (1 contrainte par point extrême, et deux contraintes pour chacun des 3 points milieux) et l'imposition de la continuité de la dérivée aux 3 points milieux donne 3 contraintes supplémentaires. Mais il devient alors impossible de rajouter l'imposition de dérivées nulles aux deux extrémités. En effet, les contraintes des coordonnées des points et la continuité des tangentes donnent déjà 11 contraintes. Il n'est possible que d'ajouter une seule fois l'imposition de dérivées nulles aux bornes extérieures.

L'utilisation de polynômes du troisième degré permet de résoudre cette lacune :

$$f(x) = ax^3 + bx^2 + cx + d \quad (\text{II.2})$$

Avec 4 inconnues par fonction, le nombre d'équations disponibles monte à 16. Les

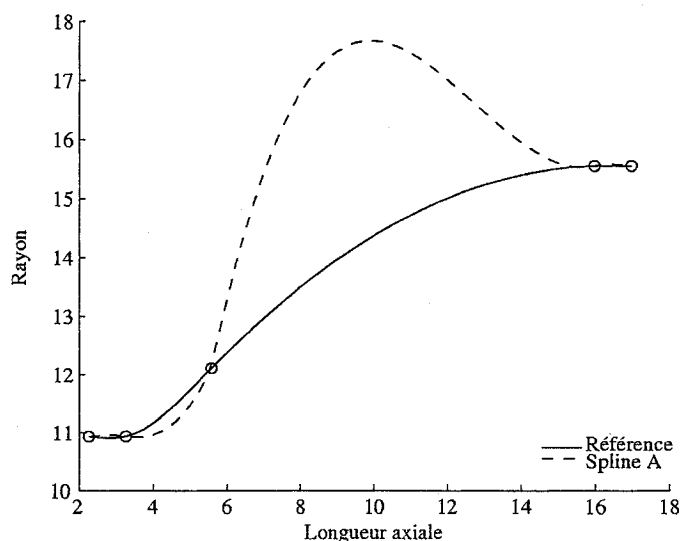


FIG. II.2 Lissage avec 4 splines de degré 3

contraintes sont les mêmes que précédemment, avec les coordonnées des points (8 contraintes), et la continuité des tangences (3 contraintes) ajoutent 11 équations. Ajoutant les dérivées premières nulles aux extrémités du domaine de définition, le nombre total de contraintes est de 13. Le système devient complet en ajoutant une continuité de courbure aux points milieux (3). Le résultat de cette formulation est illustrée par la figure II.2. La spline A comporte trop d'oscillations importantes pour donner un lissage pertinent des données. De plus, l'inversion analytique de matrice donne des formules exécutant la longueur permise dans les cellules des feuilles de calcul. Une opération manuelle consistant à scinder une somme de termes en deux sous-cellules permet de contourner la limitation logicielle. Toutefois, cette formulation devant être transposée pour une dizaine de remplacements à travers les feuilles, une telle procédure est à proscrire.

En interprétant les points extrémaux d'ordonnée identique comme une imposition de dérivée aux extrémités nulles, la formulation par polynômes de degré 3 peut être réduite à deux splines, reliant les 3 points milieux. En effet, cette formulation donne 8 inconnues à déterminer. En s'appuyant sur les contraintes des coordonnées des points (4 contraintes), la continuité de la tangence au point central (1 contrainte), et la continuité de la courbure

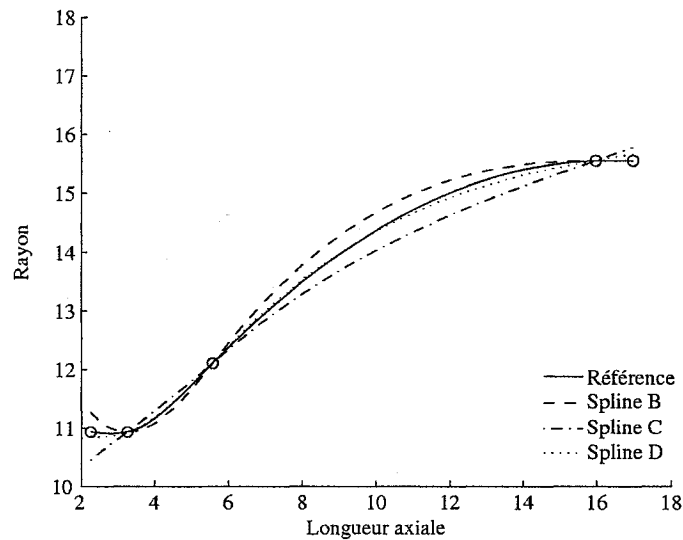


FIG. II.3 Lissage avec 2 splines de degré 3

(1 contrainte), le système nécessite encore 2 équations pour être pleinement déterminé.

Les figures II.3 et II.4 montrent des splines avec différentes contraintes imposées aux points extrêmes. La spline B est ainsi contrainte par l'imposition de dérivées premières nulles aux extrémités. Cette spline présente le même défaut que dans le cas de la spline A : une oscillation trop importante par rapport à la courbe de référence. L'annulation des dérivées seconde est imposée sur la spline C, et permet d'obtenir une amplitude réduite des oscillations. Mais la courbe est alors trop tendue par rapport à la courbe de référence. La spline D est ainsi une simple tentative de moyenner les deux courbes B et C, pour obtenir un lissage plus proche de la courbe de référence. Mais l'erreur de représentation est encore insatisfaisante.

Des contraintes ont été imposées de façon mixte sur les splines de la figure II.4. La dérivée seconde au point initial et la dérivée première sont contraintes à une valeur nulle pour la spline E. La dérivée première est imposée à une valeur faible non nulle au point initial et à une valeur nulle pour la spline F. Les résultats sont meilleurs mais l'erreur d'approximation reste grande.

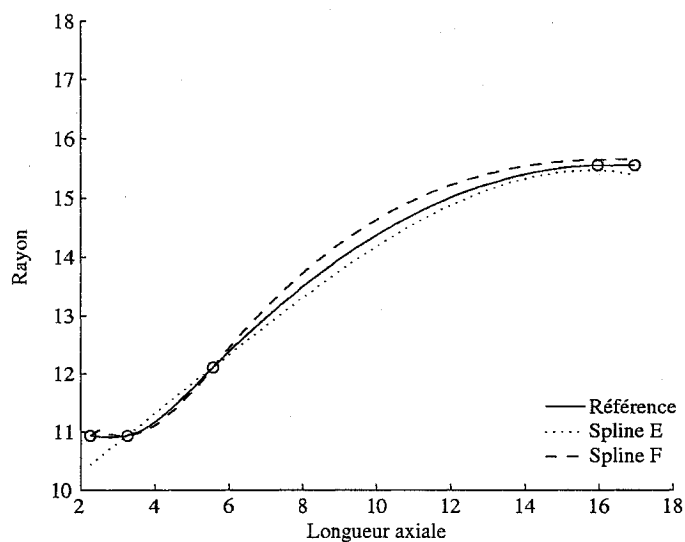


FIG. II.4 Lissage avec 2 splines de degré 3, contraintes mixtes

II.3 Formulation spécifique

En utilisant des hypothèses sur la géométrie, le problème de lissage peut être formulé de manière plus subtile. En effet, les opérations de lissage effectuées sur les composantes sont toujours effectuées sur des formes similaires. L'élément qui pose un problème le plus souvent est la représentation de l'entrée du conduit d'élargissement, tandis que la sortie est, sur les cas étudiés, en général moins délicate à traiter.

Une formulation originale a permis de déterminer une courbe de lissage dont l'erreur d'approximation est satisfaisante. L'astuce a consisté à utiliser 3 polynômes de degré 2, et un polynôme de degré 1 en sortie. La figure II.5 montre que cette formulation ne respecte pas une condition limite de dérivée nulle à l'origine (mathématiquement non applicable : système sur-déterminé). Toutefois, l'erreur est limitée grâce au doublement de points à l'entrée. La spline G montre une bonne approximation de la courbe de référence.

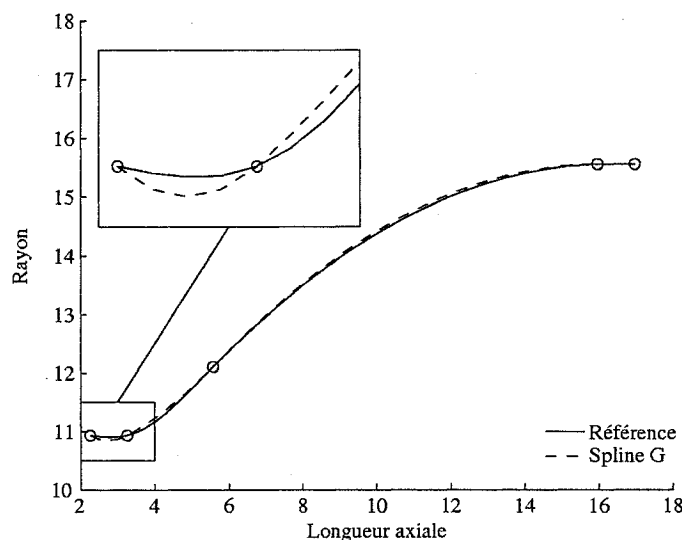


FIG. II.5 Lissage avec une formulation spécifique

II.4 Conséquence du remplacement

La formulation choisie reste une approximation de la courbe de référence, mais c'est la formulation qui présentait l'erreur de représentation la plus faible. Ce remplacement était toutefois nécessaire pour utiliser la discipline considérée (turbine basse pression). En effet, cette discipline a de nombreux liens de couplage avec d'autres disciplines (turbine haute pression et arbre de transmission). Ne pas intégrer la turbine basse pression entraîne le maintien de nombreuses variables à une valeur fixe des turbine haute pression, et de l'arbre de transmission. Ces dernières composantes possèdent ensuite d'autres couplages avec d'autres disciplines et pourraient devoir fixer des variables supplémentaires.

Le résultat final de cette approximation ne peut pas être négligé. En effet, la variable *MaxFlareAngle*, qui détermine l'angle des ailettes de la turbine basse pression, se retrouve surestimé. Le calcul effectué par la disciplines originale de P&WC donne une valeur de 34 degrés. Pour le même point de conception donné (valeur de toutes les autres variables maintenu à une valeur constante), l'approximation construite par des splines ré-

sulte en un angle de 41 degrés, soit une augmentation de 20%. Ce changement a toutefois été validé par P&WC. En effet, cette variable n'est pas utilisée par une autre discipline. L'ingénieur consulté a conseillé d'indiquer dans les feuilles de calcul que cette valeur est une surestimation de la valeur réelle.

Ces splines servent principalement à déterminer certaines coordonnées intermédiaires aux étages des compresseurs et turbines, non directement disponibles. En effet, un chemin de veine gazeuse plus réaliste est obtenu si les géométries inter-étages sont lissées par ce type de spline, plutôt que par une simple interpolation linéaire.

ANNEXE III

MÉTHODE DE DÉCOMPOSITION DO

Dans un souci de clarté, et pour familiariser le lecteur avec une méthode de décomposition, cette section présente comment, sur une formulation analytique, un problème est décomposé suivant la méthode DO. Soit le système d'équations suivant :

$$\begin{cases} F_1 &= (x_1^2 + 2x_1 + 1) + \sum_{i=2}^6 (2x_i^2 - 2x_i x_{i-1} + 1) \\ r_1 &= \frac{F_1}{1 + \beta \cdot r_2} \\ F_2 &= \sum_{i=7}^{12} (2x_i^2 - 2x_i x_{i-1} + 1) \\ r_2 &= F_2 \sqrt{1 + \beta \cdot r_1} \end{cases} \quad (\text{III.1})$$

où r_1 et r_2 sont des variables couplées. La forme initiale du problème d'optimisation est :

$$\min_{x_1, \dots, x_{12}} r_1 + r_2 \quad (\text{III.2})$$

La décomposition DO sépare le problème d'optimisation en un problème maître, et deux sous-problèmes d'optimisation. Le problème maître s'écrit :

$$\begin{cases} \min_{\{s,t\} \in \Gamma} & r_1 + r_2 \\ \text{s.t. :} & r_1 - \text{CopyIr}_1 = 0, \\ & r_2 - \text{CopyIr}_2 = 0 \end{cases} \quad (\text{III.3})$$

avec $s = \{x_6, x_{12}\}$, $t = \{\text{CopyIr}_1, \text{CopyIr}_2\}$, $\Gamma = \{x \in \mathbb{R}^m | 0 \leq x_i \leq 10, i \in \{1, \dots, 12\}, -10 \leq \text{CopyIr}_j \leq 10, j \in \{1, 2\}\}$. Une duplication des variables couplées permet de retirer le couplage des calculs de r_1 et r_2 . Les deux sous-problèmes ont leurs

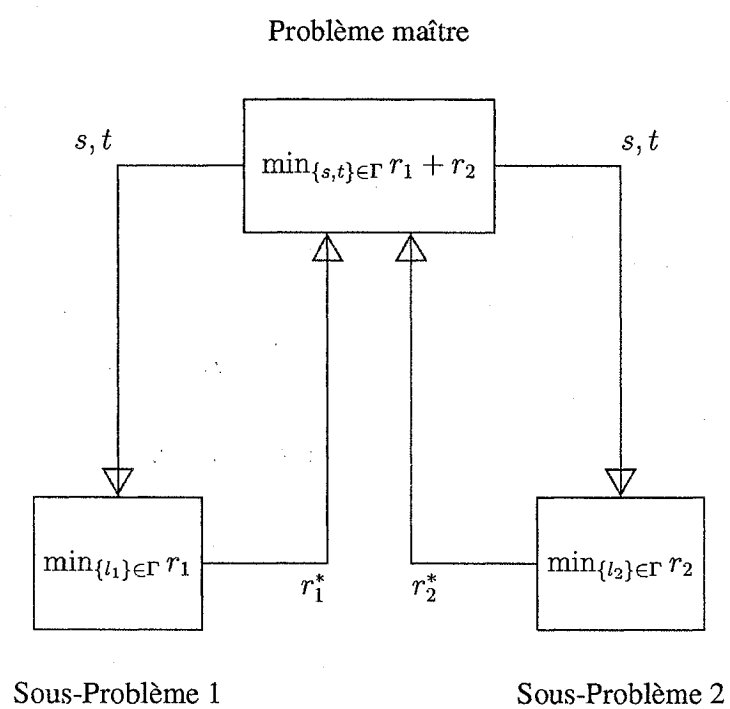


FIG. III.1 Décomposition DO

propres variables et fonction coût. Ainsi, le premier sous-problème s'écrit :

$$\min_{\{l_1\} \in \Gamma} r_1 \quad (\text{III.4})$$

avec $l_1 = \{x_1, x_2, x_3, x_4, x_5\}$, $\Gamma = \{x \in \mathbb{R}^m | 0 \leq x_i \leq 10, i \in [1 \dots 5]\}$. Le deuxième sous-problème est :

$$\min_{\{l_2\} \in \Gamma} r_2 \quad (\text{III.5})$$

avec $l_2 = \{x_7, x_8, x_9, x_{10}, x_{11}\}$, $\Gamma = \{x \in \mathbb{R}^m | 0 \leq x_i \leq 10, i \in [7 \dots 11]\}$.

Le problème maître choisit des valeurs pour les variables partagées ($s = \{x_6, x_{12}\}$) et couplées ($t = \{CopyIr_1, CopyIr_2\}$). Les sous-optimiseurs déterminent alors les variables locales qui leur permettent de réduire leur propre fonction coût (i.e., $r_j, j \in \{1, 2\}$).

En fonction de ces réponses, le problème maître modifie les variables globales afin de poursuivre l'optimisation. D'un problème d'optimisation initial comportant 10 variables locales, 2 partagées et 2 couplées, DO établit un problème maître gérant 2 variables partagées et 2 couplées, et deux sous-optimisations gérant chacune 5 variables locales. DO permet ainsi de scinder une lourde tâche d'optimisation en terme de nombre de variables, en un système hiérarchisé de plusieurs sous-tâches avec un nombre réduit de variables à gérer. Même si le nombre total de variables est augmenté de copies des variables couplées, la charge pour chaque problème d'optimisation est moindre, et rend l'optimisation plus "créative" (Tribes & Trépanier (2006)) au niveau local.