



Rapport de Metamodelisation

Axel OLOUGOUNA, Qian WANG
Ayoub CHOUKRI, Minh Duy NGUYEN
Juin 2025

Table des matières

1	Introduction	3
2	Conditions d'expérience	3
2.1	Paramètres de Conception	3
2.2	Contraintes Opérationnelles	3
2.3	Paramètres Incertains	3
2.4	Cas d'études	4
3	Problème 1 - Métamodèle et optimisation	4
3.1	Présentation du problème	4
3.2	Création du métamodèle	4
3.2.1	Différents choix de modèles surrogates	5
3.2.1.1	RBFRegressor	5
3.2.1.2	MOERegressor	5
3.2.2	Création d'un domaine d'expériences	5
3.2.3	Coefficient de détermination R^2	5
3.3	Optimisation du métamodèle	5
3.4	Résultats obtenus	6
3.4.1	Use Case 1	6
3.4.2	Cas d'étude 2	7
4	Problème 2 - Modélisation par substitut et quantification des incertitudes	9
4.1	Approche proposée	9
4.2	Use Case 1	10
4.2.1	Construction du modèle de substitut	10
4.2.2	Coefficient de détermination R^2	10
4.2.3	Résultats de l'entraînement	11
4.2.4	Estimation des quantités d'intérêt	11
4.2.4.1	Génération des échantillons	11
4.2.4.2	Distribution des sorties du modèle de substitut	11
4.2.4.1	Analyse des distributions	12
4.2.4.3	Estimation des moyennes et variances et Coefficients de variation	12
4.2.5	Indices de Sobol	14
4.2.5.1	Indices de Sobol pour les sorties du modèle de substitut	14
4.3	Cas d'étude 2	16
4.3.1	Modèle de substitut pour le Use Case 2	16
4.3.2	Résultats de l'entraînement	16
4.3.3	Estimation des quantités d'intérêt	17
4.3.3.1	Analyse des distributions	17
4.3.3.1	Estimation des moyennes et variations et Coefficients de variation	18
4.3.3.2	Indices de Sobol pour le Use Case 2	18
4.4	Case with x_{opt}	20
4.4.1	Surrogate Model pour le Use Case 1 avec \mathbf{x}_{opt}	20
4.4.2	Estimation des quantités d'intérêt	20
4.4.2.1	Distribution	20
4.4.2.2	Estimation des moyennes et variations et Coefficients de variation	21
4.4.2.3	Indices de Sobol pour le cas x_{opt}	22
5	Problème 3	22
5.1	Méthode de développement de Taylor	23
5.2	Méthode par surrogate direct	25
5.3	Méthode par surrogate local itératif	26
5.3.1	Algorithme	26
5.4	Visualiation de l'aéronef optimisé par des modèles surrogate	29

1 Introduction

La conception d'un avion commercial est un défi multidisciplinaire nécessitant l'optimisation de nombreux paramètres sous des contraintes strictes de performance, de sécurité et de réglementation. Les simulations numériques, bien que précises, sont coûteuses en temps et en ressources, rendant leur usage intensif difficile dans un cadre industriel. Pour surmonter cette limitation, la métamodélisation offre une solution efficace en construisant des modèles surrogates capables d'approximer le comportement des simulateurs à un coût computationnel réduit.

Ce rapport s'intéresse à l'optimisation d'un avion de 150 passagers environ volant à Mach 0.78, en étudiant deux configurations de propulsion : kérosène/turbofan et hydrogène/turbofan. L'objectif principal est de minimiser la masse maximale au décollage (**mtom**) tout en respectant des contraintes opérationnelles, telles que la longueur de piste au décollage, la vitesse d'approche ou l'envergure. Les paramètres de conception incluent la poussée, le nombre de passagers, la surface alaire et le rapport d'aspect, tandis que des paramètres incertains, liés à l'aérodynamique, la propulsion et la structure, sont pris en compte dans les analyses avancées.

Le rapport est structuré comme suit : le **Problème 1** explore la construction et l'optimisation d'un métamodèle pour résoudre un problème d'optimisation déterministe. Le **Problème 2** introduit les incertitudes via des distributions de probabilité, tandis que le **Problème 3** propose trois méthodes (développement de Taylor, surrogate direct et surrogate local itératif) pour intégrer ces incertitudes dans l'optimisation.

2 Conditions d'expérience

On veut optimiser un avion d'environ 150 passagers qui vole à Mach 0.78 en utilisant différents moteurs et différents type de carburant. Pour cela, on va vouloir trouver les paramètres de conception qui minimise le maximum de la masse au décollage (**mtom**). Ci-après, la présentation des différents paramètres sur lesquels nous allons travailler.

2.1 Paramètres de Conception

Les paramètres de conception sont :

- La poussée statique maximale au niveau de la mer ($100 \text{ kN} \leq \text{slst} \leq 200 \text{ kN}$, par défaut : 150 kN),
- Le nombre de passagers ($120 \leq \text{n_pax} \leq 180$, par défaut : 150),
- La surface des ailes ($100 \text{ m}^2 \leq \text{area} \leq 200 \text{ m}^2$, par défaut : 180 m²),
- Le rapport d'aspect de l'aile ($5 \leq \text{ar} \leq 20$, par défaut : 9).

2.2 Contraintes Opérationnelles

Les contraintes opérationnelles sont :

- La longueur de piste au décollage ($\text{tofl} \leq 1900 \text{ m}$),
- La vitesse d'approche ($\text{vapp} \leq 135 \text{ kt}$),
- La vitesse verticale ($\text{vz} \geq 300 \text{ ft/min}$),
- L'envergure de l'aile ($\text{envergure} \leq 40 \text{ m}$),
- La longueur de l'aile ($\text{longueur} \leq 45 \text{ m}$),
- La marge de carburant ($\text{fm} \geq 0\%$).

2.3 Paramètres Incertains

Dans les problèmes 2 et 3, certains paramètres sont considérés comme incertains. Ces paramètres incertains sont modélisés comme des variables aléatoires définies par des distributions de probabilité :

Variable	Distribution	Type de carburant	Type de moteur
gi	$T(0.35, 0.4, 0.405)$	hydrogène_liquide	Tous
vi	$T(0.755, 0.800, 0.805)$	hydrogène_liquide	Tous
aef	$T(0.99, 1.0, 1.03)$	Tous	Tous
cef	$T(0.99, 1.0, 1.03)$	Tous	Tous
sef	$T(0.99, 1.0, 1.03)$	Tous	Tous
fc_pwd	$T(0.8, 1, 1.02)$	hydrogène_liquide	électroventilateur
bed	$U(400, 700)$	batterie	Tous

TABLE 1 – Paramètres incertains avec leurs distributions de probabilité, type de carburant et type de moteur. T (minimum, mode, maximum) représente la distribution triangulaire, et U (minimum, maximum) représente la distribution uniforme.

Voici un tableau récapitulatif des significations probables des variables incertaines, basé sur le contexte de la conception aéronautique :

Variable	Signification probable
<i>g1</i>	Facteur de gravité ou rapport de poussée spécifique pour <i>liquide_H2</i> .
<i>v1</i>	Vitesse critique ou croisière normalisée pour <i>liquide_H2</i> .
<i>sef</i>	Efficacité structurelle, applicable à tous.
<i>cef</i>	Efficacité du coefficient de traînée, applicable à tous.
<i>aef</i>	Efficacité aérodynamique globale, applicable à tous.
<i>fc_pwd</i>	Facteur de consommation par watt pour <i>liquide_H2</i> et electrofan.
<i>bed</i>	Capacité de la batterie en Wh/kg pour systèmes à batterie.

2.4 Cas d'études

On va prendre deux configurations pour étudier résoudre divers problèmes.

	Type de carburant	Type de moteur	Portée de conception (km)
Use Case 1	Kérosène	Turbofan	5500
Use Case 2	Hydrogène	Turbofan	5500

TABLE 2 – Cas d'utilisation avec leurs types de carburant, types de moteur et portée de conception.

3 Problème 1 - Métamodèle et optimisation

3.1 Présentation du problème

Le but de ce problème est de créer un métamodèle de $\hat{f} : x \mapsto \hat{f}(x) = f(x, u_{\text{default}})$ (x représentant l'ensemble des paramètres de conception et u_{default} la valeur par défaut pour les paramètres incertains) pour approximer l'objectif et les contraintes du problème de design. Ensuite, nous utiliserons le métamodèle pour résoudre un problème d'optimisation en faisant varier la plage d'entrée des paramètres.

3.2 Création du métamodèle

Étant donné que nous disposons de quatre variables d'entrée et que le simulateur en temps réel est coûteux en temps de calcul, nous nous limitons à 20 échantillons pour constituer notre métamodèle.

Pour construire notre métamodèle, on utilise la fonction `SurrogateDiscipline` de **GEMSEO** avec comme régresseur (**RBFRegressor**, **MOERRegressor**, etc).

3.2.1 Différents choix de modèles surrogates

3.2.1.1 RBFRegressor

Le régresseur **RBFRegressor** approxime la fonction cible $\hat{f}(x)$ par une combinaison linéaire de fonctions à base radiale gaussiennes :

$$\hat{f}(x) = \sum_{i=1}^N w_i \phi(\|x - c_i\|) + b,$$

où $\phi(r) = \exp\left(-\frac{r^2}{2\varepsilon^2}\right)$ est la fonction radiale, c_i sont les centres (points d'entraînement), w_i les poids, b un biais, et ε un paramètre de lissage.

3.2.1.2 MOERegressor

Le régresseur **MOERegressor** est une mixture de plusieurs modèles de régressions. Dans un premier temps, il crée K clusters avec les données en entrées. Puis il crée sur chaque cluster des modèles de régression. La sortie pour une entrée devient la somme pondérée des prédictions de chaque modèle de régression.

$$y = \sum_{k=1}^K w_k(x) f_k(x),$$

On peut avoir deux types de classification :

- La classification dure : dans ce cas, les poids sont définis par

$$w_k(x) = \mathbb{1}_{x \in C_k},$$

et on ne garde que l'apport de la classe où l'entrée est classifiée.

- La classification douce : dans ce cas, les poids sont définis par

$$w_k(x) = \mathbb{P}(x \in C_k)$$

où $\mathbb{P}(x \in C_k)$ est la probabilité que l'entrée appartienne au cluster k .

3.2.2 Création d'un domaine d'expériences

Un plan d'expériences basé sur un échantillonnage Latin Hypercube Sampling (LHS) optimisé est généré pour échantillonner l'espace des paramètres de conception défini par les paramètres `slst`, `n_pax`, `area`, et `ar`.

3.2.3 Coefficient de détermination R^2

Le coefficient de détermination R^2 mesure la qualité de l'ajustement d'un modèle en quantifiant la proportion de la variance totale des données expliquée par le modèle. Mathématiquement, il est défini comme suit :

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

où y_i est la valeur observée, \hat{y}_i est la valeur prédite par le modèle, \bar{y} est la moyenne des valeurs observées, et n est le nombre d'observations. Un R^2 proche de 1 indique que le modèle explique presque toute la variabilité des données, tandis qu'un R^2 proche de 0 indique un mauvais ajustement.

3.3 Optimisation du métamodèle

Une fois le métamodèle construit, on procède à l'optimisation en l'utilisant à la place du modèle direct car il est moins coûteux en temps de calcul et donc plus facile à optimiser. On crée donc un scénario en se basant sur le métamodèle. Une fois l'optimisation finie, on s'assure que les contraintes sont respectées sur le modèle direct.

3.4 Résultats obtenus

3.4.1 Use Case 1

Dans un premier temps on réalise le métamodèle sur UC1 en utilisant RBFRegressor. On obtient pour un $R^2 = 1$ pour les différentes variables du modèle sur les données d'entraînement. Sur les données de test (100 échantillons), on obtient un $R^2 = 0.88$ pour la variable **tofl** et un $R^2 \geq 0.94$ pour toutes les autres variables. Cela signifie que le métamodèle a bien approximé notre simulateur.

La figure 1 ci-dessous présente les résultats de l'optimisation du métamodèle pour l'UC 1.

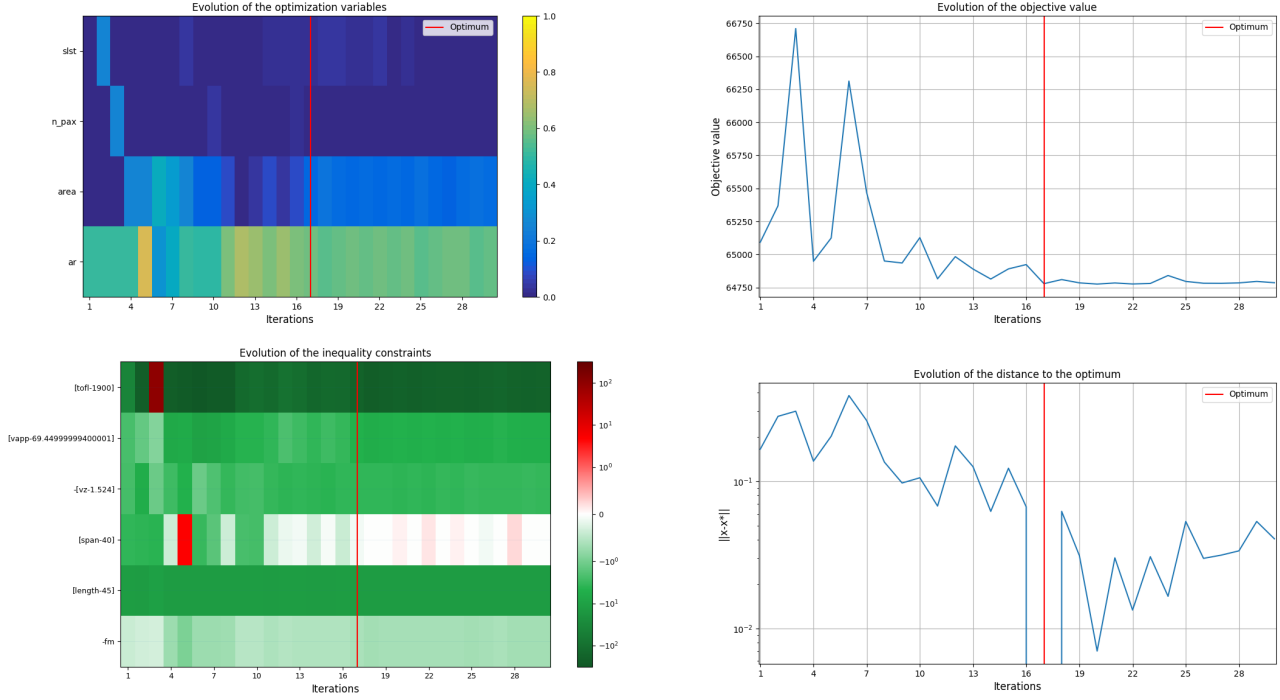


FIGURE 1 – Évolution de l'optimisation pour UC1

Les résultats de l'optimisation pour le cas UC1 montrent une convergence satisfaisante vers une solution optimale après 19 itérations.

- **Evolution de l'optimisation des variables :** Cette heatmap illustre l'évolution des variables d'optimisation (**slst**, **n_pax**, **ar**) sur les itérations. Les couleurs vont du bleu (valeurs basses) au jaune (valeurs hautes). On constate que les paramètres (**slst**, **n_pax**, **area**, **ar**) se stabilisent autour de la 17^{ème} itération, marquée par la ligne rouge indiquant l'optimum.
- **Evolution de la valeur de l'objectif :** La valeur objective (**mtom**) diminue progressivement de 78000 kg à environ 64000 kg, confirmant une réduction effective de la masse totale au décollage. Elle se stabilise après l'optimum à l'itération 24, marqué par une ligne rouge verticale.
- **Evolution des contraintes d'inégalités :** Ce graphique en heatmap montre l'évolution des contraintes (**tofl**, **vapp**, **vz**, **span**, **length**, **fm**) sur les itérations. Les contraintes d'inégalité restent majoritairement respectées (donc de couleur verte, inversement) au cours des itérations, bien qu'on constate quelques fois des pics temporaires (notamment pour **span** et **tofl**).
- **Evolution de la distance à l'optimum :** Cette représente la distance (en norme) à l'optimum au fil des itérations. La distance diminue progressivement, avec des fluctuations, indiquant une convergence vers l'optimum, marqué par une ligne rouge verticale à l'itération 19. La distance à l'optimum, initialement de l'ordre de 10^0 , chute à 10^{-2} vers la fin, validant la convergence de l'algorithme.

La vérification avec le modèle direct confirme la cohérence des résultats. Les paramètres après optimisation sont les suivants :

- **slst** : 102669.61243896
- **n_pax** : 120

— area : 114.20991753

— ar : 13.70699063

On affiche maintenant les avions par défaut et celui obtenu après optimisation.

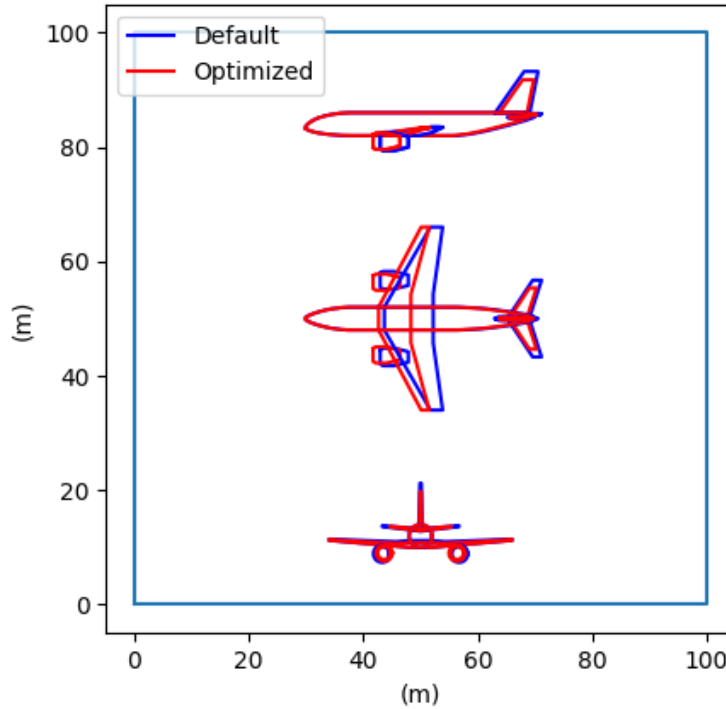


FIGURE 2 – Les différents avions

On constate que l'avion obtenu possède des dimensions plus petites que celui par défaut.

3.4.2 Cas d'étude 2

Dans un premier temps on utilise le RBFRegressor pour réaliser le métamodèle. Mais on se rend compte que les contraintes après optimisation ne sont pas vérifiées pour le modèle direct. On cherche donc le régresseur qui fonctionne et au final on choisit MOERRegressor avec une classification de type dure.

On obtient pour un $R^2 = 1$ pour les différentes variables du modèle sur les données d'entraînement. Sur les données de test (100 échantillons), on obtient un $R^2 = 0.88$ pour la variable `tof1` et un $R^2 \geq 0.94$ pour toutes les autres variables. Cela signifie que le métamodèle a bien approximé notre simulateur.

Il faut noter que le paramètre `tof1` est un paramètre dont le contrainte n'était pas vérifiée avec d'autres régresseurs.

La figure 3 ci-dessous présente les résultats de l'optimisation du métamodèle pour l'UC 2.

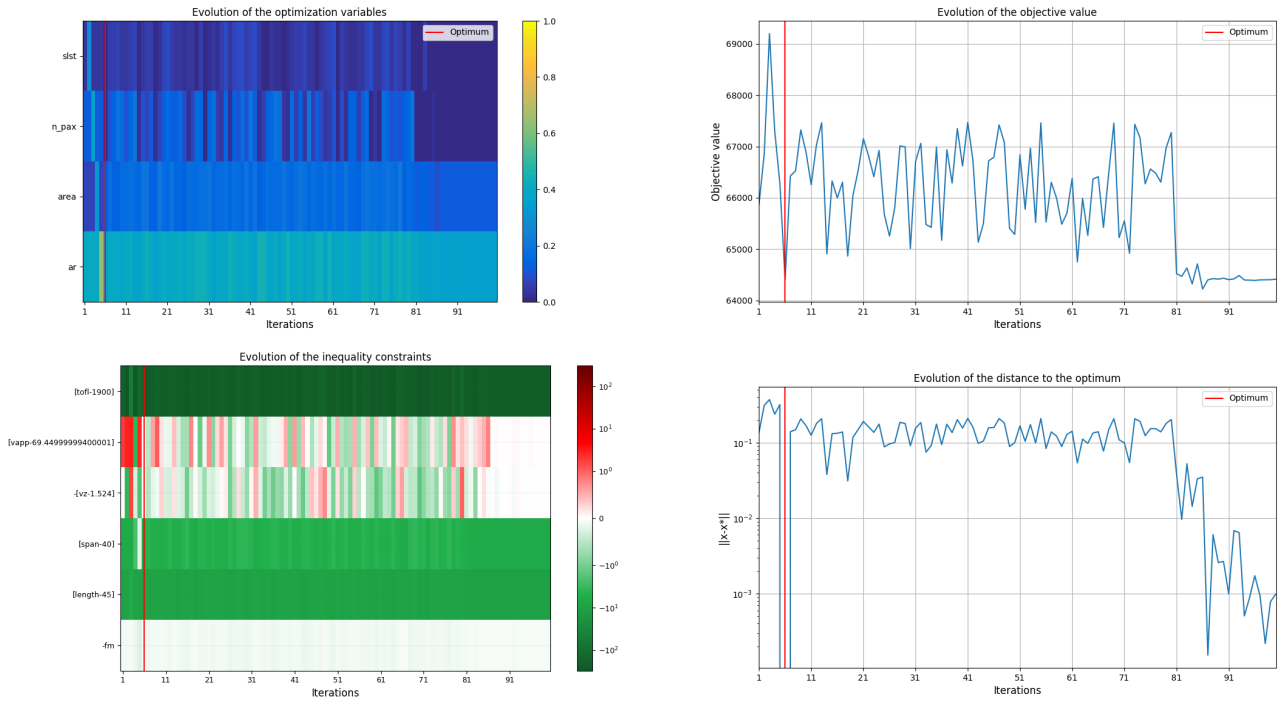


FIGURE 3 – Évolution de l’optimisation pour UC2

Les résultats de l’optimisation pour le cas UC1 montrent une convergence satisfaisante vers une solution optimale après 5 itérations.

- **Evolution de l’optimisation des variables :** Cette heatmap illustre l’évolution des variables d’optimisation (`slst`, `n_pax`, `ar`) sur les itérations. Les couleurs vont du bleu (valeurs basses) au jaune (valeurs hautes). On constate que les paramètres (`slst`, `n_pax`, `area`, `ar`) se stabilisent autour de la 5^{ème} itération, marquée par la ligne rouge indiquant l’optimum.
- **Evolution de la valeur de l’objectif :** La valeur objective (`mtom`) diminue progressivement de 69000 kg à environ 64000 kg, confirmant une réduction effective de la masse totale au décollage. Elle est atteinte à l’itération 5 mais se stabilise après l’itération 81.
- **Evolution des contraintes d’inégalités :** Ce graphique en heatmap montre l’évolution des contraintes (`tofi`, `vapp`, `vz`, `span`, `length`, `fm`) sur les itérations. Les contraintes d’inégalité restent majoritairement respectées au cours des itérations, bien qu’on constate quelques fois des pics temporaires (notamment pour `vapp` et `vz`).
- **Evolution de la distance à l’optimum :** Cette représente la distance(en norme) à l’optimum au fil des itérations (de 1 à 100). On constate qu’elle fluctue énormément augmentent encore même après avoir attendu l’optimum à la 5^{ème} itération.

La vérification avec le modèle direct confirme la cohérence des résultats. Les paramètres après optimisation sont les suivants :

- `slst` : 100000.
- `n_pax` : 120
- `area` : 112.48807812
- `ar` : 10.34990761

On a une `mtom` = 63402.36070368503. On affiche maintenant les avions par défaut et celui obtenu après optimisation.

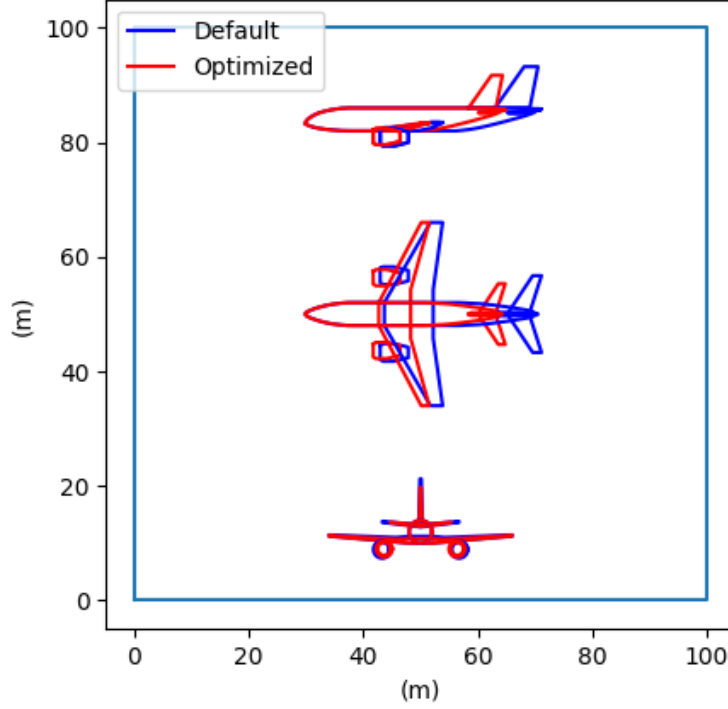


FIGURE 4 – Les différents avions

On constate que l’avion obtenu possède ici aussi des dimensions plus petites que celui par défaut.

4 Problème 2 - Modélisation par substitut et quantification des incertitudes

Dans cette partie, nous développerons une approche pour quantifier les incertitudes liées aux choix technologiques dans le processus de conception d’un avion. L’objectif est de créer des modèles de substitut pour approximer la fonction $u \mapsto f(\mathbf{x}, \mathbf{u})$, où \mathbf{x} représente les paramètres de conception et \mathbf{u} les paramètres incertains. Nous examinerons deux cas spécifiques : $f(\mathbf{x}_{\text{default}}, \mathbf{u})$ avec les paramètres de conception à leurs valeurs par défaut, et $f(\mathbf{x}_{\text{opt}}, \mathbf{u})$ avec les paramètres optimisés. Ces modèles seront utilisés pour propager les incertitudes et analyser leur impact.

4.1 Approche proposée

Notre méthodologie pour quantifier les incertitudes dans la conception d’un avion repose sur l’utilisation de la bibliothèque GEMSEO pour modéliser et analyser le système. Les étapes suivantes décrivent l’approche mise en œuvre :

1. Création de modèles de substitut :

- (a) **Création d’un plan d’expériences (DoE) :** Un plan d’expériences basé sur un échantillonnage Latin Hypercube Sampling (LHS) optimisé est généré pour échantillonner l’espace des paramètres incertains \mathbf{u} . Cela permet de collecter des données d’entraînement représentatives pour les modèles de substitut, en utilisant la fonction `sample_disciplines` de GEMSEO.
- (b) **Entraînement des modèles de substitut :** Des modèles de substitut sont construits pour $f(\mathbf{x}_{\text{default}}, \mathbf{u})$ et $f(\mathbf{x}_{\text{opt}}, \mathbf{u})$ à l’aide de la méthode `RBFRegressor` (Radial Basis Function Regressor) implémentée via la classe `SurrogateDiscipline` de GEMSEO, assurant une approximation précise des relations entrée-sortie.

2. Estimation des quantités d’intérêt :

- (a) **Échantillonnage Monte Carlo :** Les paramètres incertains \mathbf{u} , définis par leurs distributions de probabilité (par exemple, triangulaires pour $g1$, $v1$, sef , cef , aef , fc_pwd , et uniforme pour

bed), sont échantillonnés à l'aide de simulations Monte Carlo pour générer un grand nombre d'échantillons représentatifs.

- (b) **Estimation statistique** : À partir des échantillons, les modèles de substitut sont utilisés pour estimer les moyennes, variances, coefficients de variation et autres statistiques des sorties y (par exemple, `mtom`, `tofl`, `vz`, `fm`, `vapp`), permettant d'évaluer l'impact des incertitudes.

- 3. **Analyse de sensibilité** : Les indices de Sobol sont calculés pour quantifier l'influence de chaque paramètre incertain \mathbf{u} sur les sorties du système. Cette analyse est réalisée à l'aide des outils de GEMSEO, avec des visualisations générées via `sobol.plot` pour les variables `mtom`, `tofl`, `vz`, `fm`, et `vapp`, mettant en évidence les contributions relatives des paramètres incertains.

Cette approche, appuyée par les fonctionnalités de GEMSEO telles que `AutoPyDiscipline`, `DesignSpace`, et `MDOScenario`, permet une modélisation robuste des interactions multidisciplinaires et une quantification efficace des incertitudes dans le processus de conception aéronautique.

4.2 Use Case 1

Pour créer un modèle de substitut permettant d'approximer la fonction $u \mapsto f(\mathbf{x}, \mathbf{u})$, nous avons suivi une approche systématique utilisant la bibliothèque GEMSEO. Les étapes détaillées sont décrites ci-dessous, suivies d'une évaluation des performances du modèle.

4.2.1 Construction du modèle de substitut

1. **Définition de l'espace des paramètres incertains** : Un espace de paramètres incertains \mathbf{u} est défini à l'aide de la classe `MyUncertainSpace`, héritant de `ParameterSpace` de GEMSEO. Trois paramètres incertains (`cef`, `aef`, `sef`) sont modélisés par des distributions triangulaires avec les paramètres suivants : `minimum = 0.99`, `mode = 1`, `maximum = 1.03`, utilisant `OTTriangularDistribution` d'OpenTURNS.
2. **Plan d'expériences (DoE)** : Un ensemble d'entraînement est généré en échantillonnant l'espace des paramètres incertains avec la méthode LHS OPT (`OT_LHS_OPT`), produisant 20 échantillons pour les sorties `mtom`, `tofl`, `vz`, `span`, `length`, `fm`, et `vapp`. Ces échantillons sont obtenus via la fonction `sample_disciplines` de GEMSEO, avec une graine fixée à 38 pour garantir la reproductibilité.
3. **Entraînement du modèle** : Le modèle de substitut est construit à l'aide de la classe `SurrogateDiscipline` avec la méthode `RBFRegressor` (Radial Basis Function Regressor), configurée avec un paramètre `epsilon = 100`. Ce modèle est entraîné sur l'ensemble de données généré, permettant une approximation des relations entre les paramètres incertains et les sorties.
4. **Évaluation des performances** : Un ensemble de test est généré avec 200 échantillons en utilisant un plan factoriel complet (`OT_FULLFACT`), avec une graine fixée à 318, pour évaluer la précision du modèle de substitut sur des données indépendantes.

4.2.2 Coefficient de détermination R^2

Dans notre cas, les valeurs de R^2 sont calculées pour les ensembles d'entraînement et de test à l'aide de la classe `R2Measure` de GEMSEO. Les résultats sont visualisés sous forme de diagrammes en barres pour chaque sortie (`mtom`, `tofl`, `vz`, `span`, `length`, `fm`, `vapp`), comparant les performances sur les données d'entraînement et de test.

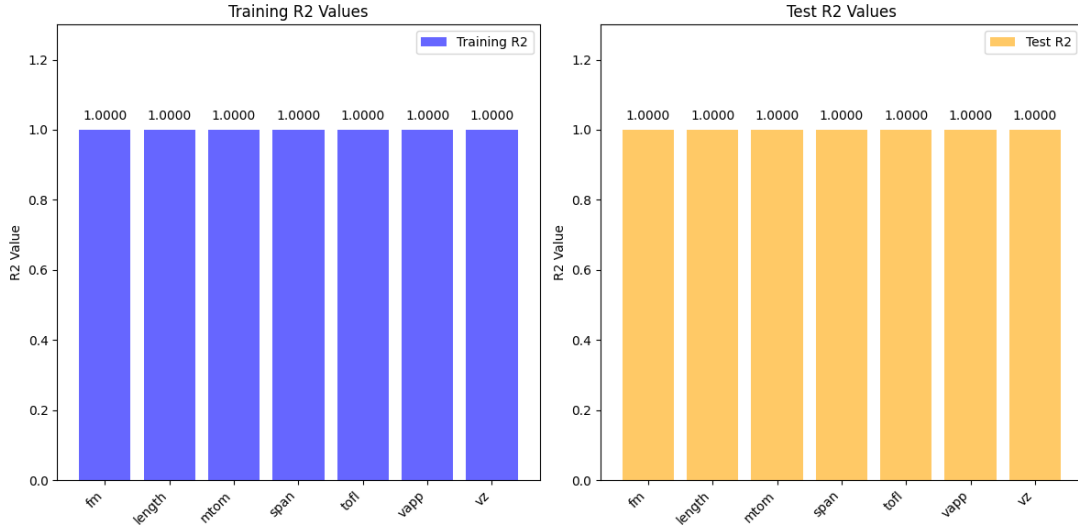


FIGURE 5 – Comparaison des valeurs de R^2 pour l’entraînement (à gauche) et le test (à droite) pour chaque sortie.

4.2.3 Résultats de l’entraînement

Le modèle de substitut a été entraîné avec succès en utilisant seulement 20 points d’échantillonnage, atteignant des valeurs de R^2 proches de 1 pour presque toutes les sorties (**mtom**, **tofl**, **vz**, **span**, **length**, **fm**, **vapp**). Ces résultats indiquent une excellente capacité du modèle à capturer les relations entre les paramètres incertains et les sorties, même avec un ensemble d’entraînement limité.

4.2.4 Estimation des quantités d’intérêt

Dans cette partie, on s’intéresse à l’estimation des quantités d’intérêt comme la moyenne, la variance, le coefficient de variation et les quantiles des sorties du modèle de substitut. Et ce, en injectant les lois de probabilité des paramètres incertains dans le modèle.

Ces estimations vont permettre de quantifier les incertitudes liées aux choix technologiques dans le processus de conception d’un avion.

4.2.4.1 Génération des échantillons

Pour générer les échantillons des paramètres incertains, nous avons utilisé la méthode Monte Carlo. Les distributions de probabilité des paramètres incertains **u** sont définies comme suit :

- *sef* : Distribution triangulaire $T(0.99, 1.1, 1.03)$
- *cef* : Distribution triangulaire $T(0.99, 1.1, 1.03)$
- *aef* : Distribution triangulaire $T(0.99, 1.1, 1.03)$

Nous avons généré 20000 échantillons pour chaque paramètre incertain en utilisant la méthode Monte Carlo. Ces échantillons sont stockés dans un tableau de taille 5000×3 , où chaque colonne correspond à un paramètre incertain.

Une idée intéressante serait de visualiser la distribution des échantillons générés pour chaque paramètre incertain, et aussi de visualiser la distribution des sorties du modèle de substitut pour les échantillons générés. Cela permettrait de mieux comprendre la variabilité des sorties en fonction des incertitudes des paramètres.

4.2.4.2 Distribution des sorties du modèle de substitut

Pour visualiser la distribution des sorties du modèle de substitut, nous avons utilisé les échantillons générés précédemment. Chaque sortie du modèle de substitut est calculée en injectant les échantillons des paramètres incertains dans le modèle. Les sorties considérées sont **mtom**, **tofl**, **vz**, **span**, **length**, **fm**, et **vapp**.

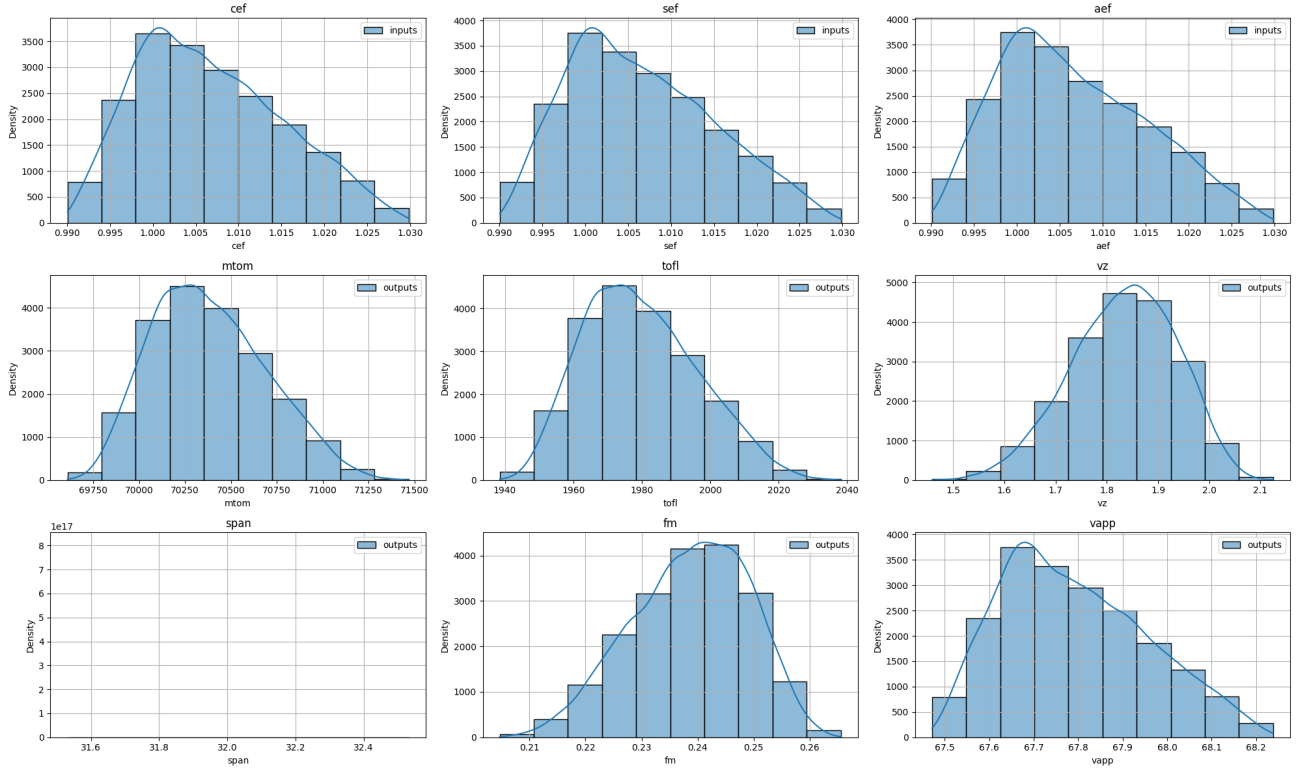


FIGURE 6 – Distribution des échantillons générés pour chaque paramètre incertain (en haut) et distribution des sorties du modèle de substitut (en bas) pour les 20000 échantillons.

4.2.4.1 Analyse des distributions La figure ci-dessous présente, dans les trois premiers graphiques, les distributions de probabilité des paramètres incertains injectés dans le modèle de substitut : **cef**, **sef**, et **aef**. Ces paramètres suivent tous une loi triangulaire définie par $T(0.99, 1.1, 1.03)$, avec un mode à 1.03. Cette distribution reflète une incertitude modérée autour d’une valeur nominale, avec une concentration importante des échantillons autour de ce mode.

Les autres histogrammes représentent les distributions des sorties du modèle de substitut obtenues via simulation Monte Carlo. On peut faire les observations suivantes :

- **mtom** (masse maximale au décollage) : distribution légèrement asymétrique à droite, concentrée entre 69750 et 71500. Cela indique une variabilité notable, avec un risque de valeurs élevées non négligeable.
- **tofl** (longueur de piste au décollage) : distribution centrée, presque symétrique, autour de 1900. La variabilité est faible, ce qui traduit une bonne stabilité face aux incertitudes.
- **vz** (vitesse ascensionnelle) : distribution symétrique et bien concentrée autour de 1.7 à 1.9. Cela traduit une réponse régulière du modèle et peu sensible aux incertitudes d’entrée.
- **span** (envergure) : distribution quasi-plate, indiquant une variable probablement constante ou très peu influencée par les paramètres incertains. Cela peut refléter une hypothèse de conception fixant cette grandeur.
- **fm** (figure of merit) : distribution légèrement asymétrique vers la gauche, centrée vers 0.24.
- **vapp** (vitesse d’approche) : distribution légèrement asymétrique à droite, concentrée entre 67.5 et 68.2. Cela montre une certaine sensibilité à l’incertitude, mais sans dérive excessive.

4.2.4.3 Estimation des moyennes et variances et Coefficients de variation

Le coefficient de variation (CV) d’une variable X_i est défini comme l’écart-type $\sqrt{\text{Var}(X_i)}$ divisé par la valeur absolue de l’espérance $|\mathbb{E}(X_i)|$, exprimé mathématiquement par :

$$CV = \frac{\sqrt{\text{Var}(X_i)}}{|\mathbb{E}(X_i)|} \quad (1)$$

et

TABLE 3 – Paramètres de l'aéronef : moyenne, variance et coefficient de variation

Paramètre	Moyenne	Variance	CV (%)
MTOM (kg)	7.04×10^4	9.67×10^4	0.442
TOFL (m)	1.98×10^3	2.77×10^2	0.841
Vz (m s^{-1})	1.83	1.06×10^{-2}	5.609
Span (m)	3.20×10^1	5.05×10^{-29}	0.000
Length (m)	3.70×10^1	0.00	0.000
FM	2.39×10^{-1}	1.10×10^{-4}	4.389
Vapp (kn)	6.78×10^1	2.66×10^{-2}	0.241
CEF	1.01	7.18×10^{-5}	0.842
AEF	1.01	7.24×10^{-5}	0.845
SEF	1.01	7.15×10^{-5}	0.840

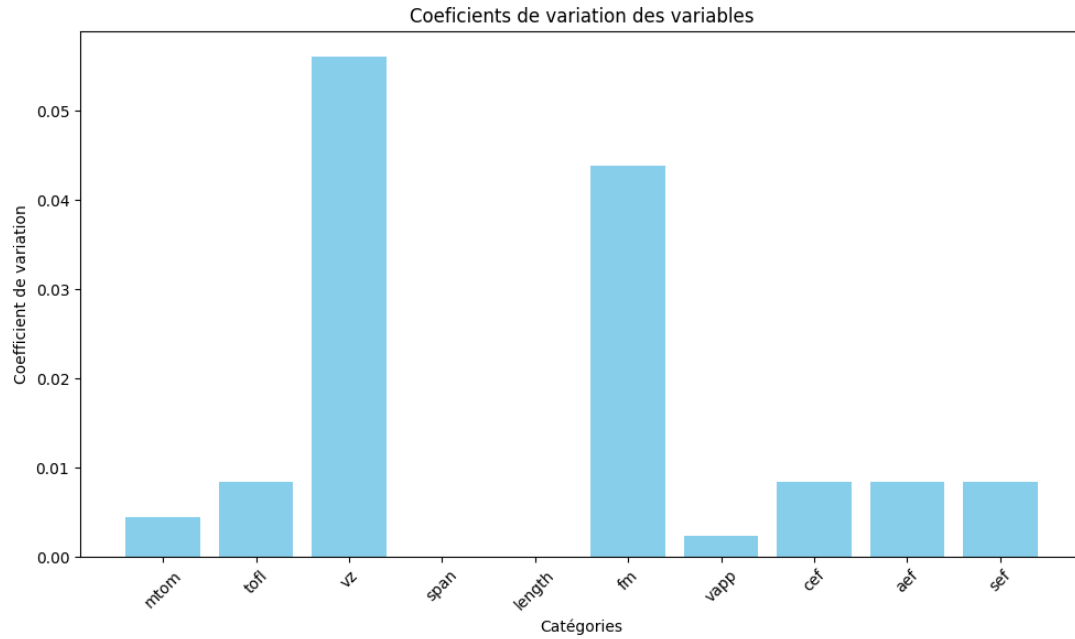


FIGURE 7 – Coefficients de variation des différentes variables

On remarque que la vitesse verticale (**vz**) , ainsi que la marge de carburant (**fm**) ont les coefficients de variation les plus élevés (5%, 4.8%). Ce qui pourrait poser des problèmes dans la vérification des contraintes opérationnelles. Pour les autres variables, le coefficient de variation reste dans l'ordre de 1%

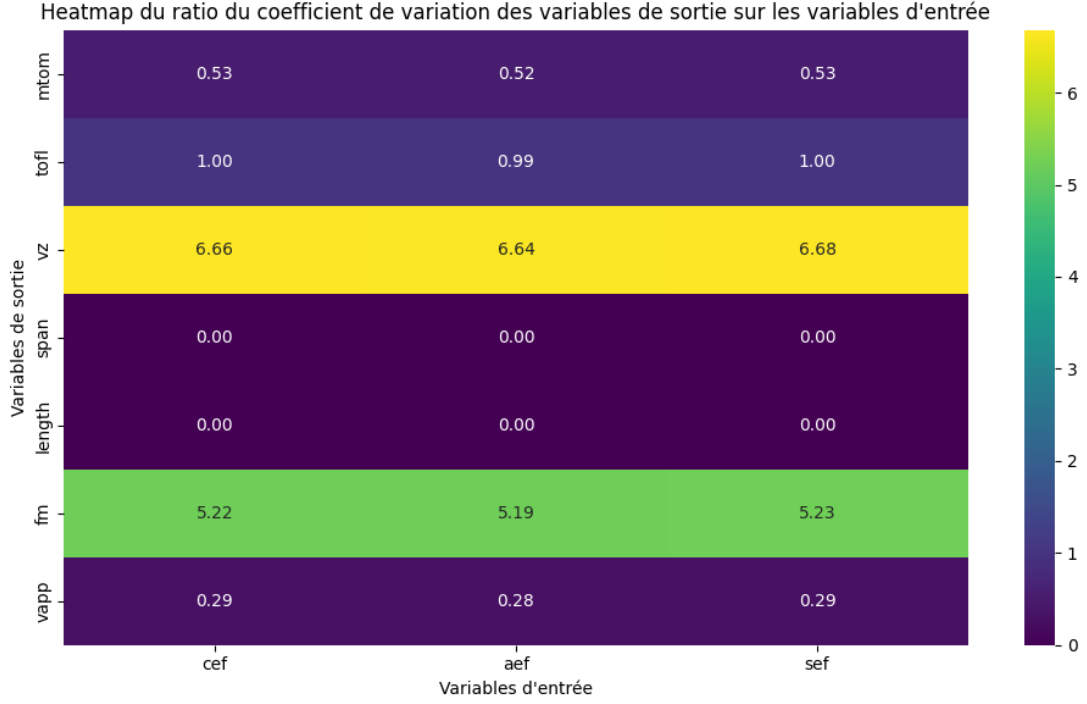


FIGURE 8 – Heatmap du ratio du coefficient de variation

4.2.5 Indices de Sobol

Les indices de Sobol sont un outil essentiel en analyse de sensibilité, permettant de quantifier la contribution de chaque variable d'entrée à la variance de la sortie d'un modèle. Pour une fonction $f(\mathbf{x})$ avec un vecteur d'entrée $\mathbf{x} = (x_1, x_2, \dots, x_d)$, la variance totale V de la sortie peut être décomposée en contributions des effets principaux des variables et de leurs interactions :

$$V = \sum_{i=1}^d V_i + \sum_{1 \leq i < j \leq d} V_{ij} + \dots + V_{1,2,\dots,d}, \quad \text{avec} \quad V_i = \text{Var}[\mathbb{E}[Y | x_i]], \quad V_{ij} = \text{Var}[\mathbb{E}[Y | x_i, x_j]] - V_i - V_j.$$

Ici, V_i représente la variance due à l'effet principal de la variable x_i , tandis que des termes comme V_{ij} capturent les interactions entre x_i et x_j .

L'indice de Sobol de premier ordre pour une variable x_i est défini comme :

$$S_i = \frac{V_i}{V} = \frac{\text{Var}[\mathbb{E}[Y | x_i]]}{\text{Var}[Y]}.$$

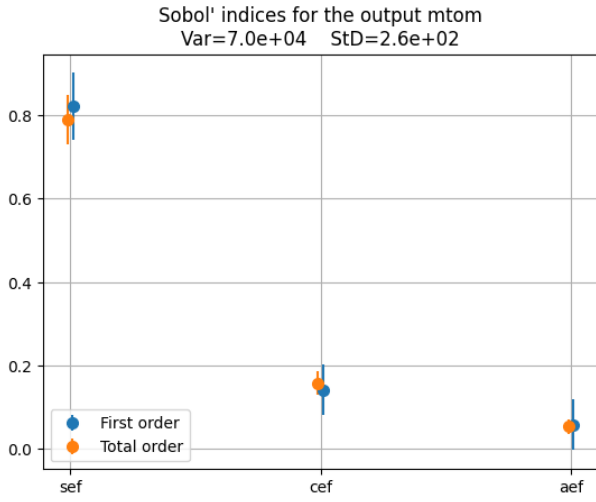
Cet indice mesure la contribution directe de x_i à la variance totale, sans les interactions. Pour évaluer l'influence globale de x_i , y compris ses interactions avec les autres variables, on utilise l'indice de Sobol total S_{T_i} , donné par :

$$S_{T_i} = 1 - \frac{V_{-i}}{V} = \sum_{E \text{ tel que } i \in E} \frac{V_E}{V}$$

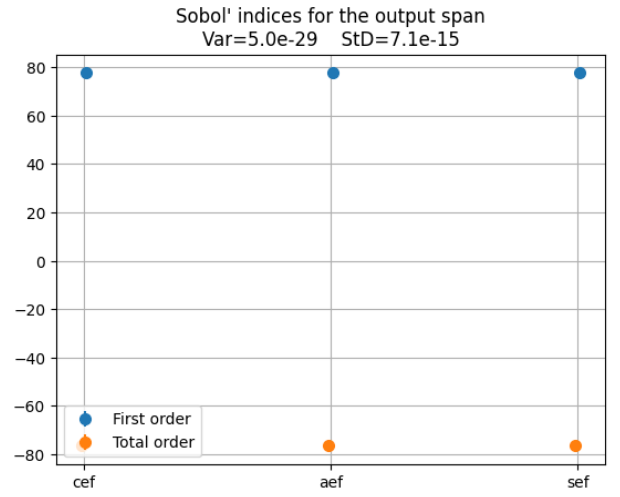
où V_{-i} est la variance lorsque x_i est fixée et que toutes les autres variables varient. Les indices de Sobol sont normalisés, la somme des indices de premier ordre et des termes d'interaction valant 1, ce qui en fait un outil idéal pour classer l'importance des variables dans des modèles complexes.

4.2.5.1 Indices de Sobol pour les sorties du modèle de substitut

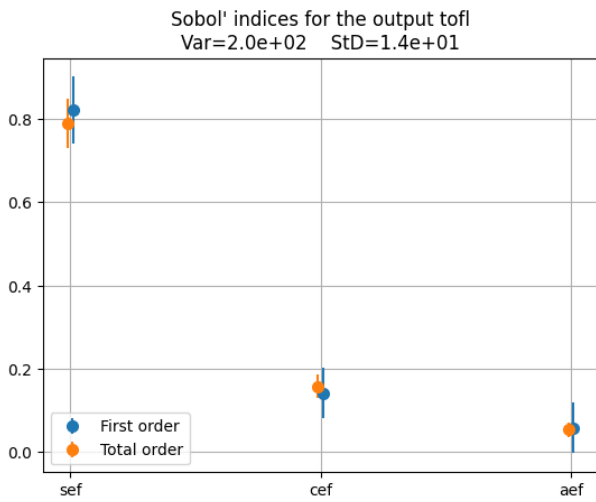
Les indices de Sobol sont calculés pour les sorties du modèle de substitut afin d'évaluer l'importance relative des paramètres incertains sur chaque sortie. Les résultats sont présentés dans la figure ci-dessous, où chaque barre représente l'indice de Sobol pour un paramètre donné et une sortie spécifique.



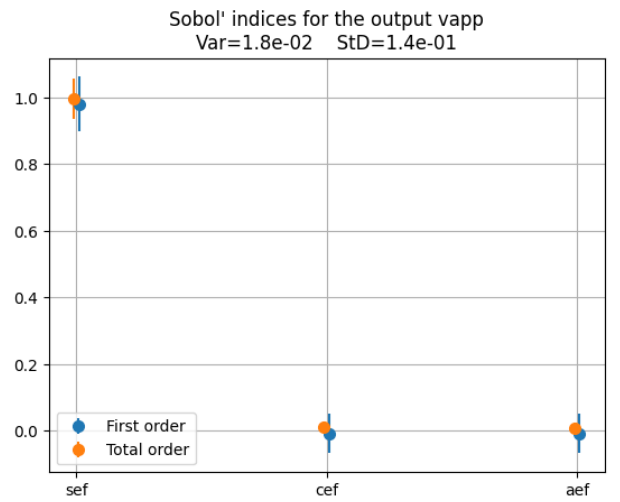
(a) MTOM



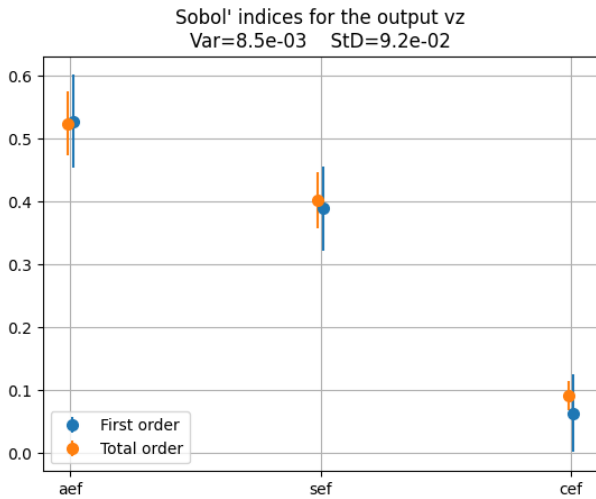
(b) Span



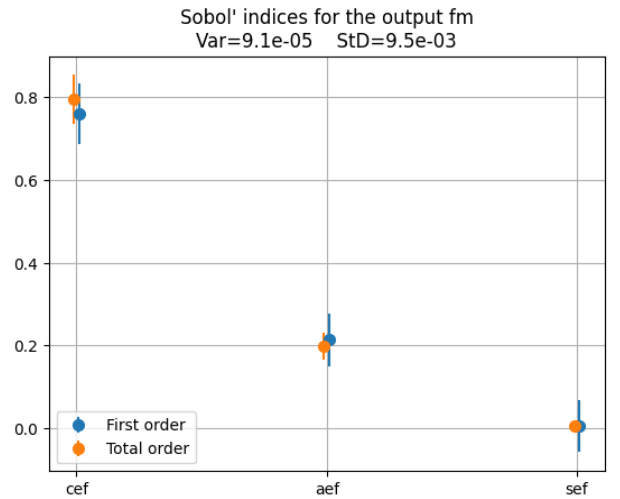
(c) TOFL



(d) Vapp



(e) Vz



(f) FM

FIGURE 9 – Indices de Sobol pour différentes sorties du modèle de substitut.

- Les sorties (MTOM, Span, TOFL, Vz, Vapp) dépendent principalement des paramètres set, ctf, aef et aet, comme indiqué par les indices de Sobol.
- Pour MTOM, l'indice de Sobol total est significativement plus élevé que l'indice de premier ordre, suggérant une forte interaction entre les paramètres.
- Pour Span, les indices totaux (Var=5.0e-29, STD=7.1e-15) dépassent les indices de premier ordre,

indiquant des effets d'interaction notables.

- TOFL montre une dépendance marquée aux indices totaux (Var=2.0e-02, STD=1.4e-01), avec des interactions plus importantes que les effets individuels.
- Vz et Vapp présentent des indices totaux modérés (Var=9.5e-03, STD=9.5e-03 pour Vapp), mais les différences avec les indices de premier ordre restent mineures, suggérant une dépendance plus directe.
- Globalement, les indices de Sobol totaux sont souvent supérieurs aux indices de premier ordre, révélant que les interactions entre paramètres jouent un rôle clé dans la variabilité des sorties.

4.3 Cas d'étude 2

Le Use Case 2 est lié à l'étude de l'avion à hydrogène. Trois nouvelles variables incertaines ont été introduites : rapport de poussée spécifique ($g1$), vitesse de croisière ($v1$) et facteur de charge (fc_pwd).

Ces variables sont modélisées par des distributions triangulaires, avec les paramètres suivants :

Variable	Distribution	Type de carburant	Type de moteur
$g1$	$T(0.35, 0.4, 0.405)$	liquide_H2	All
$v1$	$T(0.755, 0.800, 0.805)$	liquide_H2	All
sef	$T(0.99, 1.1, 1.03)$	All	All
cef	$T(0.99, 1.1, 1.03)$	All	All
aef	$T(0.99, 1.1, 1.03)$	All	All
fc_pwd	$T(0.8, 1.1, 1.02)$	liquide_H2	electrofan
bed	$U(400, 700)$	battery	All

4.3.1 Modèle de substitut pour le Use Case 2

Les mêmes étapes que pour le Use Case 1 ont été suivies pour créer un modèle de substitut permettant d'approximer la fonction $u \mapsto f(\mathbf{x}, \mathbf{u})$ pour le Use Case 2. Les paramètres incertains ont été définis, un plan d'expériences a été généré, et le modèle de substitut a été entraîné en utilisant la classe `SurrogateDiscipline` avec la méthode `RBFR regressor`.

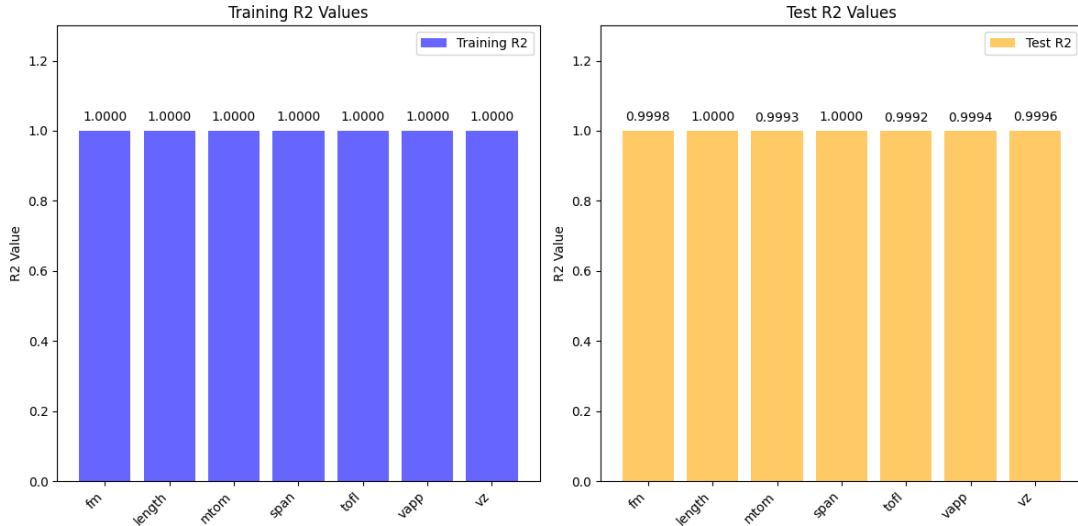


FIGURE 10 – Comparaison des valeurs de R^2 pour l'entraînement (à gauche) et le test (à droite) pour chaque sortie.

4.3.2 Résultats de l'entraînement

Le modèle de substitut a été entraîné avec succès en utilisant seulement 20 points d'échantillonnage, atteignant des valeurs de R^2 proches de 1 pour presque toutes les sorties (**mtom**, **tofl**, **vz**, **span**, **length**, **fm**, **vapp**). Ces résultats indiquent une excellente capacité du modèle à capturer les relations entre les paramètres incertains et les sorties, même avec un ensemble d'entraînement limité.

4.3.3 Estimation des quantités d'intérêt

Pour le Use Case 2, nous avons également généré des échantillons des paramètres incertains en utilisant la méthode Monte Carlo. Les distributions de probabilité des paramètres incertains sont les mêmes que celles définies précédemment, avec l'ajout des nouvelles variables $g1$ et $v1$.

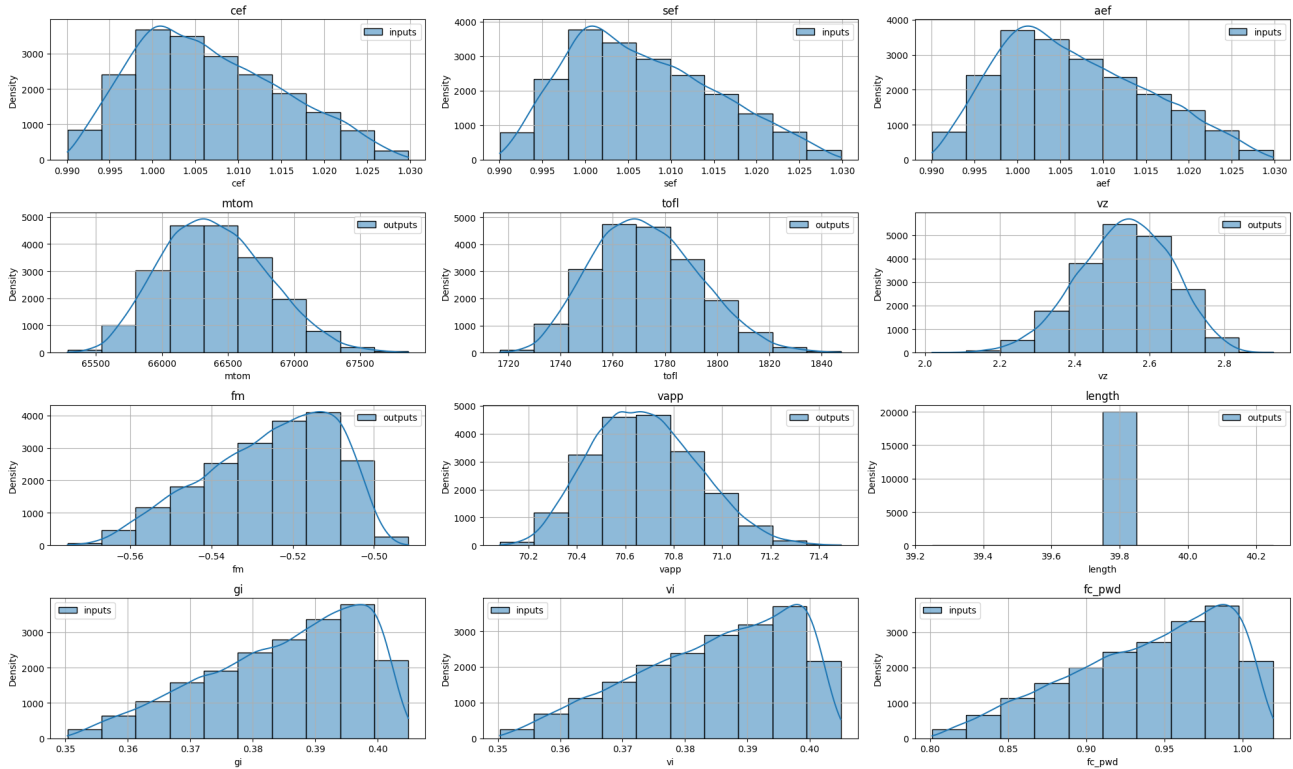


FIGURE 11 – Distribution des échantillons générés pour chaque paramètre incertain (en haut) et distribution des sorties du modèle de substitut (en bas) pour les 20000 échantillons.

4.3.3.1 Analyse des distributions La figure ci-dessous présente, dans les trois premiers graphiques, les distributions de probabilité des paramètres incertains injectés dans le modèle de substitut : **cef**, **sef**, et **aef**. Ces paramètres suivent tous une loi triangulaire définie par $T(0.99, 1.1, 1.03)$, avec un mode à 1.03. Cette distribution reflète une incertitude modérée autour d'une valeur nominale de 1.03, avec une concentration importante des échantillons autour de ce mode et des queues s'étendant jusqu'à 0.99 et 1.1, indiquant une variabilité contrôlée.

Les autres histogrammes représentent les distributions des sorties du modèle de substitut obtenues via simulation Monte Carlo. On peut faire les observations suivantes :

- **mtom** (masse maximale au décollage) : Distribution légèrement asymétrique à droite, centrée entre 66000 et 67000, avec une queue s'étendant jusqu'à 72000. Cette variabilité notable suggère un risque de valeurs élevées non négligeable, probablement influencée par les incertitudes des paramètres d'entrée.
- **tofl** (longueur de piste au décollage) : Distribution presque symétrique, centrée autour de 1780 à 1900, avec une densité maximale proche de 1800. La faible dispersion indique une bonne stabilité du modèle face aux incertitudes d'entrée.
- **vz** (vitesse ascensionnelle) : Distribution symétrique et bien concentrée entre 1.7 et 1.9, avec un pic autour de 1.8. Cette régularité reflète une réponse robuste et peu sensible aux variations des paramètres incertains.
- **fm** (figure of merit) : Distribution gaussienne centrée autour de -0.52 à -0.50, avec une densité maximale près de -0.51. La forme étroite et symétrique indique une stabilité élevée du critère de performance malgré les incertitudes.
- **vapp** (vitesse d'approche) : Distribution légèrement asymétrique à droite, concentrée entre 70.2 et 71.2, avec un pic autour de 70.8. Cette légère dissymétrie suggère une sensibilité modérée aux incertitudes, sans dérive excessive.

- **length** (longueur de l'avion) : Distribution symétrique et très étroite, centrée autour de 39.6 à 39.8, avec une densité maximale près de 39.7. La faible variabilité témoigne d'une excellente prédictibilité.
- **gi** : Distribution légèrement asymétrique à droite, concentrée entre 0.36 et 0.40, avec un pic autour de 0.38. Cette variabilité modérée indique une influence mesurée des paramètres incertains.
- **vi** : Distribution symétrique et bien concentrée entre 0.37 et 0.39, avec un pic autour de 0.38. Cette stabilité suggère une faible sensibilité aux variations d'entrée.
- **fc_pwd** : Distribution symétrique et centrée entre 0.90 et 0.95, avec un pic autour de 0.92. La faible dispersion reflète une robustesse et une prédictibilité élevées.

4.3.3.1 Estimation des moyennes et variations et Coefficients de variation

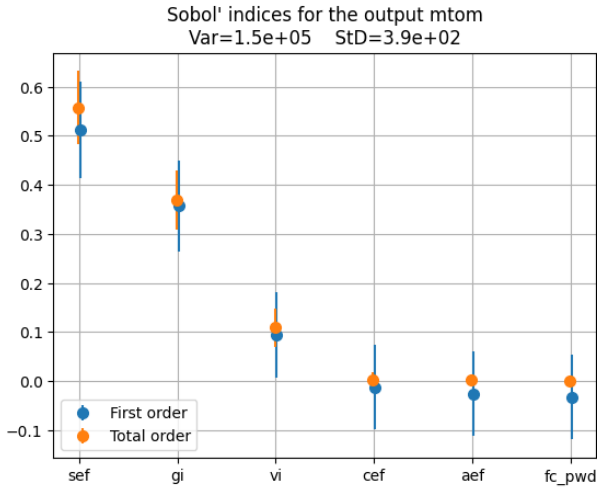
De la même façon que pour le Use Case 1, nous avons calculé les moyennes, variances et coefficients de variation, l'inter-quartile pour les sorties du modèle de substitut. Les résultats sont présentés dans le tableau ci-dessous :

TABLE 4 – Paramètres de l'aéronef : moyennes, variances, coefficients de variation et intervalles interquartiles corrigés

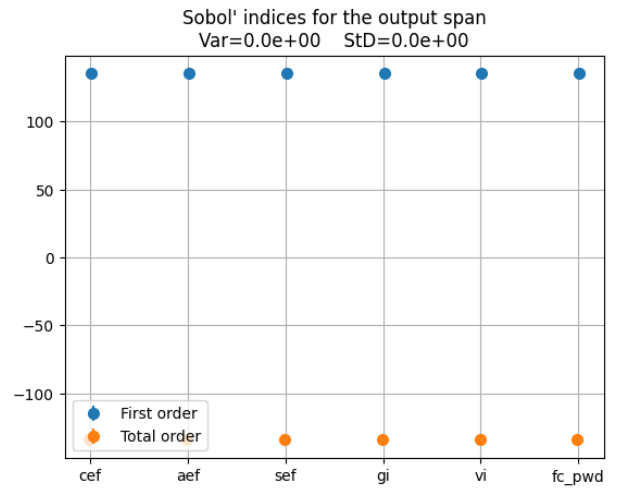
Paramètre	Moyenne	Variance	CV (%)	IQ
MTOM (kg)	6.6402×10^4	1.5904×10^5	0.601	0.0042
TOFL (m)	1.7728×10^3	4.0639×10^2	1.137	0.0080
Vz (m s^{-1})	2.534	1.5410×10^{-2}	4.900	0.0341
Span (m)	3.2031×10^1	5.0487×10^{-29}	0.000	0.0000
Length (m)	3.975×10^1	0.0000	0.000	0.0000
FM	-5.2553×10^{-12}	5.221×10^{-4}	3.022	-0.0228
Vapp (kn)	7.0681×10^1	4.7853×10^{-2}	0.309	0.0022
CEF	1.0066	7.1774×10^{-5}	0.842	0.0063
AEF	1.0067	7.2565×10^{-5}	0.846	0.0064
SEF	1.0067	7.1899×10^{-5}	0.842	0.0063
gi	3.8517×10^{-1}	1.5231×10^{-4}	3.204	0.0245
vi	3.8484×10^{-1}	1.5525×10^{-4}	3.238	0.0252
fc_pwd	9.3976×10^{-12}	2.4755×10^{-3}	5.294	0.0409

4.3.3.2 Indices de Sobol pour le Use Case 2

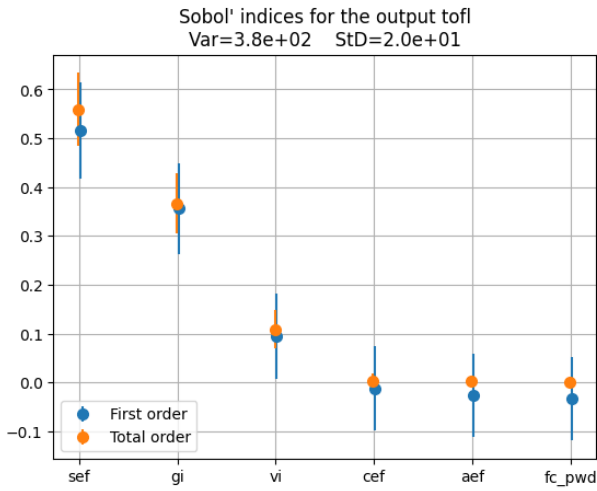
Les indices de Sobol sont calculés pour les sorties du modèle de substitut afin d'évaluer l'importance relative des paramètres incertains sur chaque sortie. Les résultats sont présentés dans la figure ci-dessous, où chaque barre représente l'indice de Sobol pour un paramètre donné et une sortie spécifique.



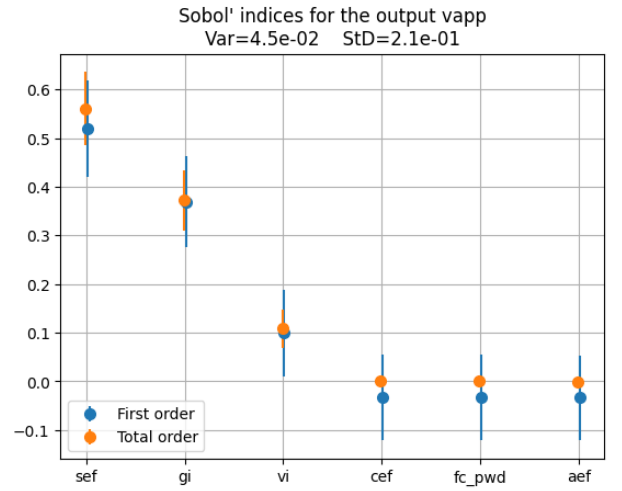
(a) MTOM



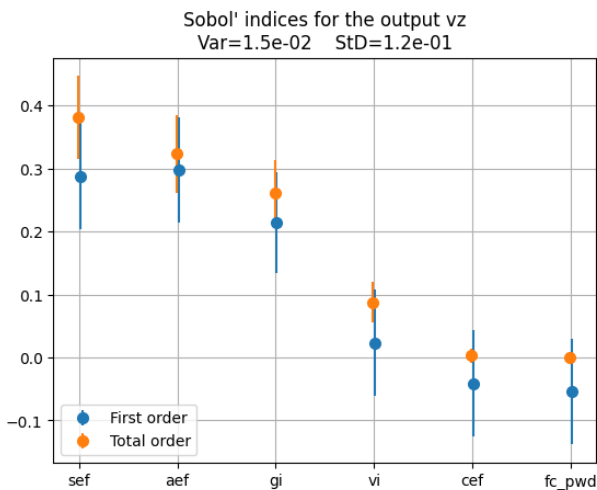
(b) Span



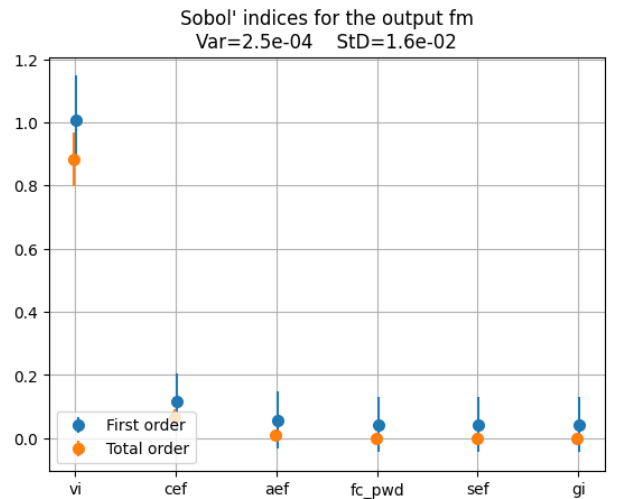
(c) TOFL



(d) Vapp



(e) Vz



(f) FM

FIGURE 12 – Indices de Sobol pour différentes sorties du modèle de substitut.

- Les sorties (MTOM, Span, TOFL, Vapp, Vz, FM) présentent des sensibilités variées vis-à-vis des paramètres **sef**, **gi**, **vi**, **cef**, **aef** et **fc_pwd**.
- Pour MTOM, le paramètre **sef** domine clairement la contribution à la variance, suivi par **gi**, tandis que les autres paramètres ont un impact négligeable.
- Pour **Span**, les indices de Sobol sont incohérents (valeurs extrêmes positives et négatives), ce qui suggère

une sortie constante ou une absence de sensibilité significative.

- **TOFL** et **Vapp** présentent une forte dépendance aux paramètres **sef** et **gi**, avec des indices totaux systématiquement supérieurs aux indices de premier ordre, indiquant l'existence d'interactions notables entre ces paramètres.
- Pour **Vz**, l'influence est plus répartie : **sef** et **aef** contribuent de manière comparable, suivis par **gi** et **vi**.
- Enfin, la sortie **FM** est presque entièrement déterminée par le paramètre **vi**, qui présente un indice total très élevé, traduisant une dépendance quasi exclusive sans contribution notable des autres paramètres.
- Dans l'ensemble, la majorité des sorties sont fortement influencées par un petit nombre de paramètres dominants (**sef**, **gi**, **vi**), et les différences entre indices totaux et de premier ordre révèlent l'importance des effets d'interaction pour certaines sorties.

4.4 Case with x_{opt}

Ici, nous cherchons à effectuer les mêmes étapes que précédemment, mais en fixant \mathbf{x} à \mathbf{x}_{opt} obtenu dans le Problème 1 à l'aide d'un modèle de substitut. Par souci de concision, nous présentons uniquement le Use Case 1. Un nouveau modèle de substitut a été entraîné, et un espace d'incertitude a été généré pour estimer des quantités telles que les moyennes, les variances, les coefficients de variation, et autres métriques pertinentes.

4.4.1 Surrogate Model pour le Use Case 1 avec \mathbf{x}_{opt}

Pour évaluer l'impact des incertitudes sur les sorties du modèle lorsque \mathbf{x} est fixé à \mathbf{x}_{opt} , nous avons entraîné un nouveau modèle de substitut pour le Use Case 1 en suivant la même méthodologie que précédemment, en utilisant la classe **SurrogateDiscipline** avec la méthode **RBFRegressor**. Le modèle de substitut approxime la fonction $\mathbf{u} \mapsto f(\mathbf{x}_{\text{opt}}, \mathbf{u})$, où \mathbf{u} représente les paramètres incertains (**cef**, **sef**, **aef**, etc.) définis par leurs distributions de probabilité respectives (par exemple, distributions triangulaires $T(0.99, 1.1, 1.03)$).

4.4.2 Estimation des quantités d'intérêt

4.4.2.1 Distrubtion

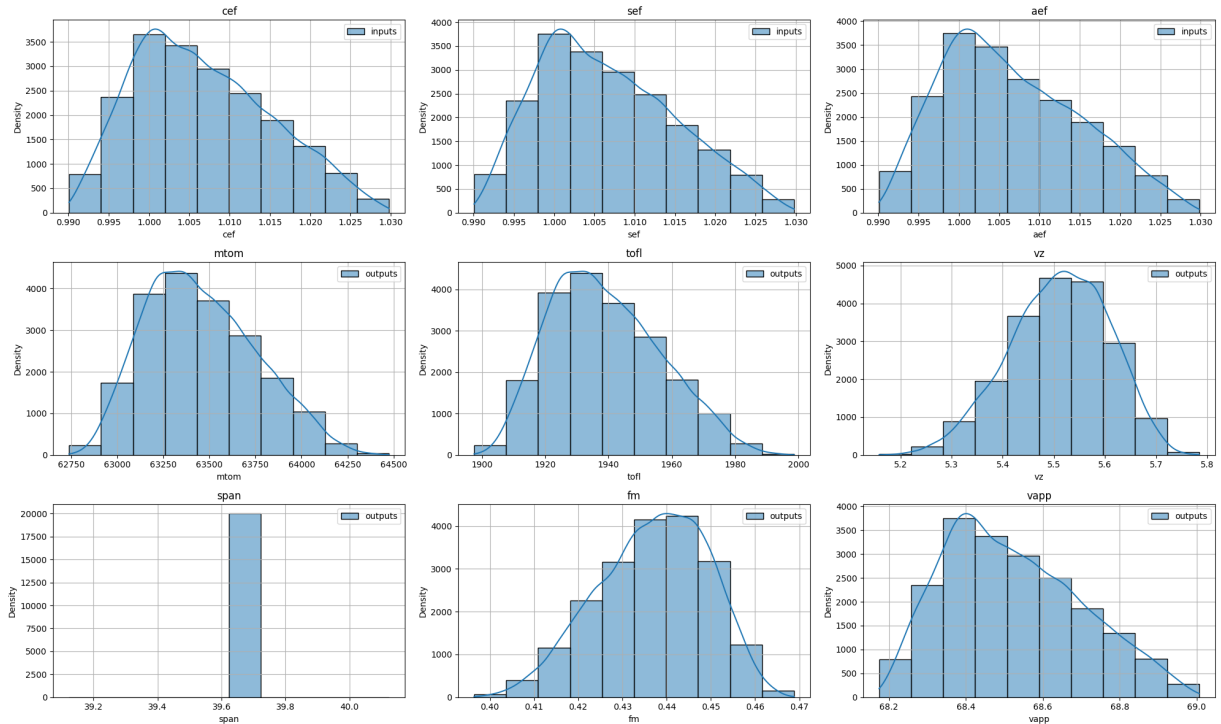


FIGURE 13 – Distribution des échantillons générés pour chaque paramètre incertain (en haut) et distribution des sorties du modèle de substitut (en bas) pour les 20000 échantillons.

4.4.2.2 Estimation des moyennes et variations et Coefficients de variation

TABLE 5 – Paramètres de l’aéronef : moyennes, variances, coefficients de variation et intervalles interquartiles corrigés

Paramètre	Moyenne	Variance	CV (%)	IQ
MTOM (kg)	6.3457×10^4	8.9973×10^4	0.473	0.0034
TOFL (m)	1.9391×10^3	3.0361×10^2	0.899	0.0065
Vz (m s^{-1})	5.5103	9.4520×10^{-3}	1.764	0.0124
Span (m)	3.9623×10^1	0.0000	0.000	0.0000
Length (m)	3.2000×10^1	5.0487×10^{-29}	0.000	0.0000
FM	4.3691×10^{-1}	1.5517×10^{-4}	2.851	0.0206
Vapp (kn)	6.8523×10^1	3.1324×10^{-2}	0.258	0.0019
CEF	1.0067	7.1781×10^{-5}	0.842	0.0063
AEF	1.0066	7.2384×10^{-5}	0.845	0.0064
SEF	1.0067	7.1464×10^{-5}	0.839	0.0063

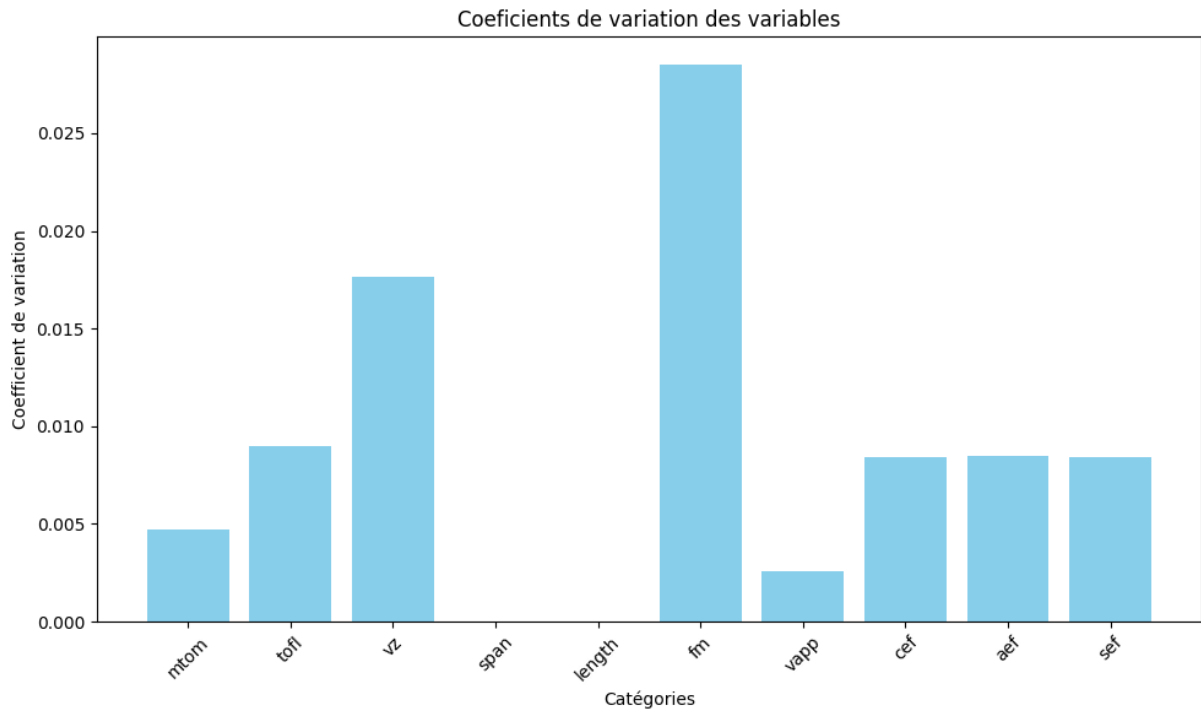
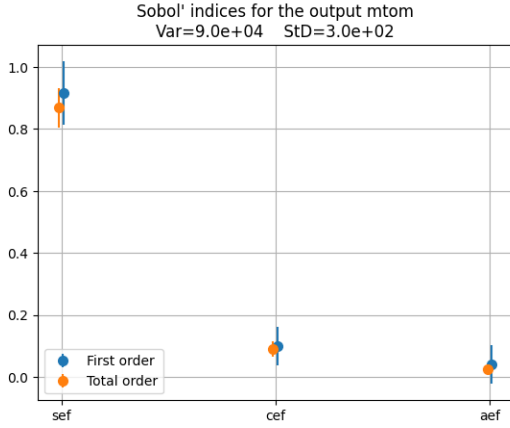
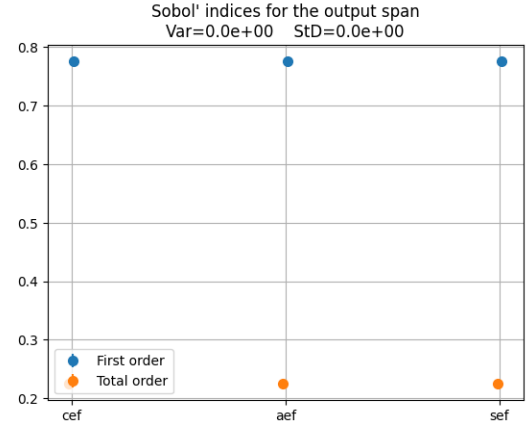


FIGURE 14 – Coefficients de variation des différentes variables

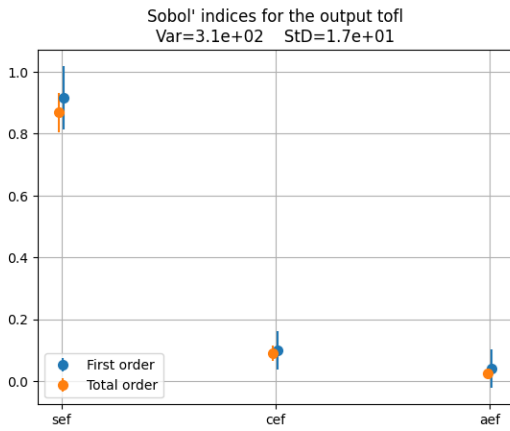
4.4.2.3 Indices de Sobol pour le cas x_{opt}



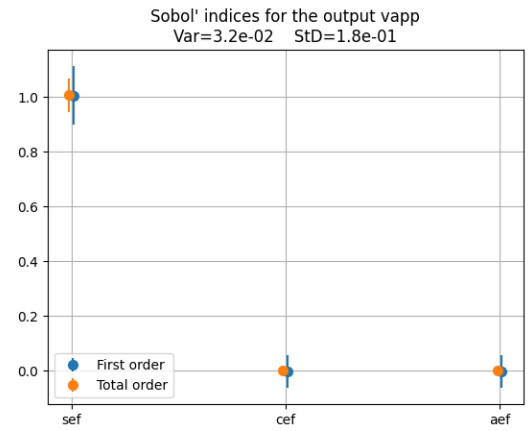
(a) MTOM



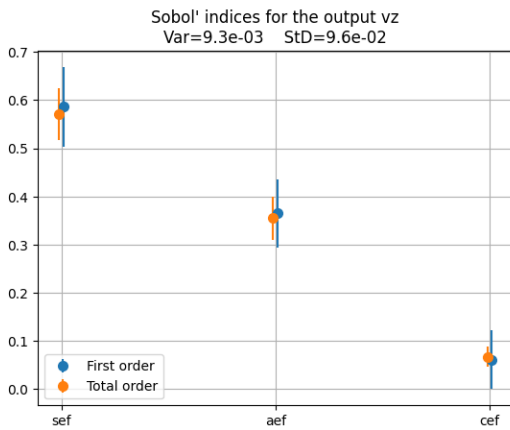
(b) Span



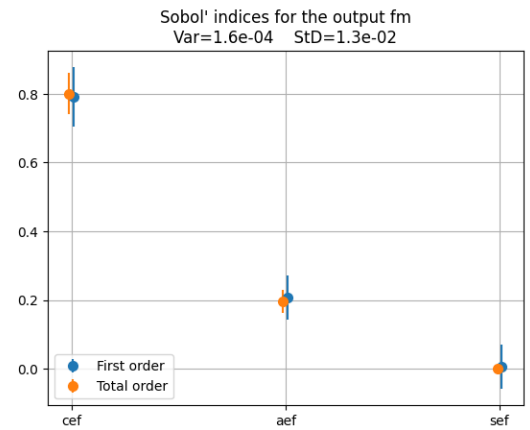
(c) TOFL



(d) Vapp



(e) Vz



(f) FM

FIGURE 15 – Indices de Sobol pour différentes sorties du modèle de substitut.

5 Problème 3

Dans cette troisième partie, comme dans la réalité, nous souhaitons intégrer les incertitudes dans notre problème d'optimisation. Nous choisissons un critère où l'objectif est de minimiser l'espérance de la masse, en considérant que toutes les paramètres (u) et variables de sortie deviennent aléatoires.

Afin de mieux respecter les contraintes malgré la présence d'incertitudes, nous ajoutons également un terme correspondant à deux fois l'écart-type, appliqué sur l'espérance des contraintes. Cela signifie que si les

variables des contraintes suivent une loi normale, alors ces contraintes seront satisfaites avec une probabilité de 95%.

Le problème initial devient donc le suivant :

$$\begin{cases} \min_x \mathbb{E}[f(x, U)] \\ \text{s.c. } \mathbb{E}[g(x, U)] + 2S[g(x, U)] \leq 0 \end{cases} \quad (2)$$

Nous allons ainsi proposer trois méthodes, impliquant des charges de calcul différentes, pour résoudre ce problème de manière numérique.

5.1 Méthode de développement de Taylor

Par le développement de Taylor :

$$f(x, U) \approx f(x, \mathbb{E}[U]) + \frac{\partial f(x, \mathbb{E}[U])}{\partial U}(U - \mathbb{E}[U]) \quad (3)$$

On prend l'espérance, en rendant compte que $\mathbb{E}[U - \mathbb{E}[U]] = 0$ donc $\mathbb{E}[f(x, U)] \approx f(x, \mathbb{E}[U])$ et $\mathbb{E}[g(x, U)] \approx g(x, \mathbb{E}[U])$. On a donc

$$\begin{aligned} S[g(x, U)] &= \mathbb{E}[g(x, U) - \mathbb{E}[g(x, U)]]^2 \\ &\approx \mathbb{E}[g(x, U) - g(x, \mathbb{E}[U])]^2 \\ &= \mathbb{E}\left[\frac{\partial g(x, \mathbb{E}[U])}{\partial U}(U - \mathbb{E}[U])\right]^2 \\ &= \left(\frac{\partial g(x, \mathbb{E}[U])}{\partial U}\right)^2 \mathbb{E}[U - \mathbb{E}[U]]^2 \\ &= \left(\frac{\partial g(x, \mathbb{E}[U])}{\partial U}\right)^2 \text{Var}(U) \end{aligned}$$

Le problème d'optimisation se ramène à la version linéarisée :

$$\begin{cases} \min_x f(x, \mathbb{E}[U]) \\ \text{s.c. } g(x, \mathbb{E}[U]) + 2\frac{\partial g(x, \mathbb{E}[U])}{\partial U}\sqrt{\text{Var}(U)} \leq 0 \end{cases}$$

Nous obtenons donc le résultat suivant.

- Cas d'usage 1 :

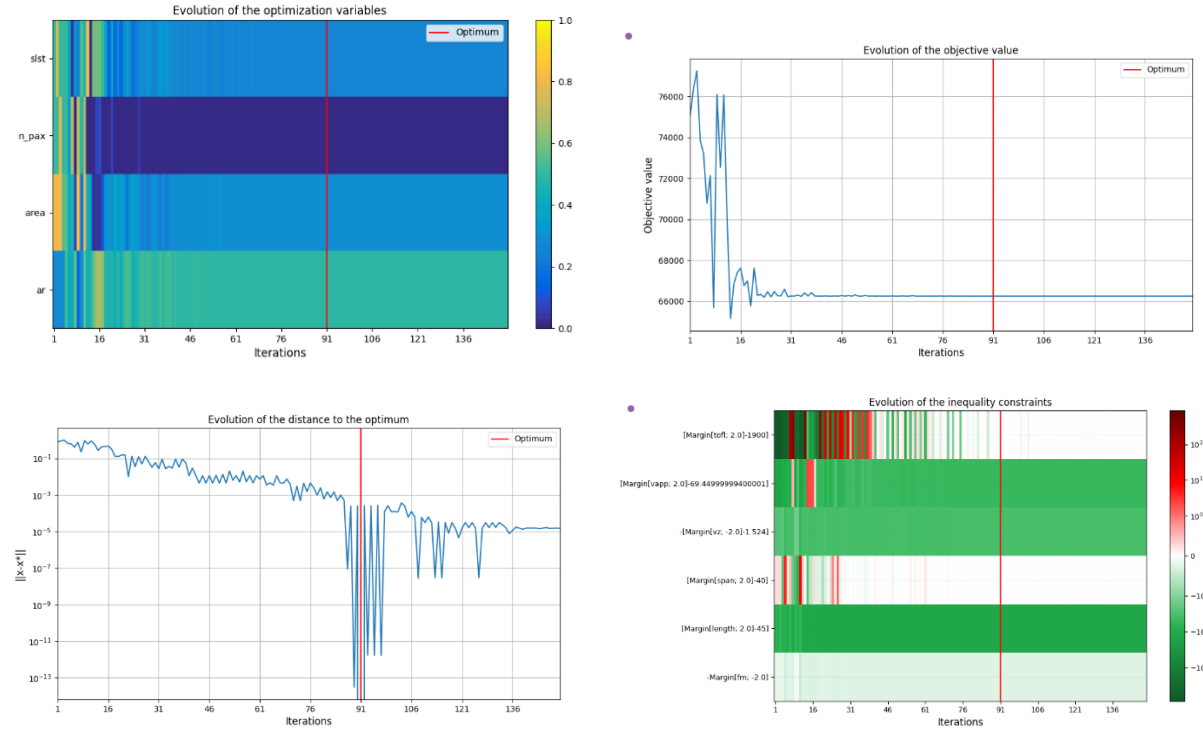


FIGURE 16 – Evolution d’optimisation du cas 1 par la méthode de développement de Taylor

- Cas d’usage 2 :

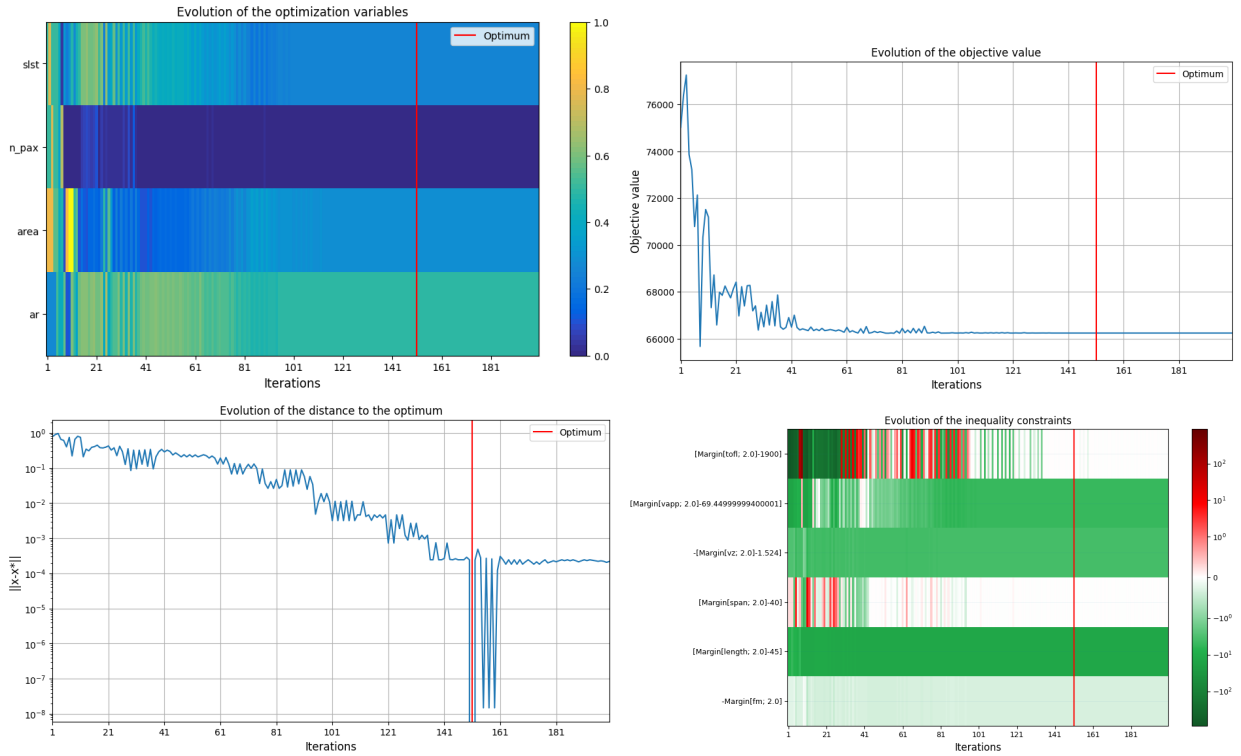


FIGURE 17 – Evolution d’optimisation du cas 2 par la méthode de développement de Taylor

L’optimisation a été correctement réalisée et la solution a bien convergé. Nous avons vérifié cela dans le scénario réel de la discipline concernée, et toutes les contraintes sont respectées.

5.2 Méthode par surrogate direct

De manière similaire au problème 1, notre première idée consiste à construire un modèle surrogate direct \hat{f} prenant à la fois x et u comme variables d'entrée, et renvoyant l'objectif et les contraintes. Une fois ce modèle surrogate validé, l'optimisation peut être réalisée sur x en utilisant une approche de Monte Carlo sur u .

On peut alors reformuler le problème sous la forme suivante :

$$\begin{cases} \min_x \frac{1}{N} \sum_{i=1}^N \hat{f}(x, u_i) \\ \text{s.c. } \frac{1}{N} \sum_{i=1}^N \hat{g}(x, u_i) + 2S[\hat{g}(x, u_{1:N})] \leq 0 \\ \|(f, g) - (\hat{f}, \hat{g})\| < \epsilon \end{cases}$$

où l'écart-type est défini par :

$$S[\hat{g}(x, u_{1:N})] = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \left(\hat{g}(x, u_i) - \frac{1}{N} \sum_{j=1}^N \hat{g}(x, u_j) \right)^2}$$

Cette approche permet ainsi de prendre en compte la variabilité induite par u dans l'évaluation des contraintes, via la moyenne et l'écart-type estimés à partir d'un échantillonnage Monte Carlo.

Comme \hat{f} est un modèle substitut surrogate, l'évaluation sur ce modèle est relativement peu coûteuse. Il est donc possible de choisir une taille d'échantillon N suffisamment grande, ce qui permet à la moyenne empirique de bien approcher l'espérance théorique.

Cependant, une difficulté subsiste dans des situations réalistes où la dimension de la variable d'entrée x est souvent élevée. De plus, la construction d'un modèle substitut nécessite généralement l'ajout de variables supplémentaires, telles que u , ce qui augmente encore la dimension du problème. Étant donné que l'évaluation du modèle direct est particulièrement coûteuse, il n'est généralement pas possible de générer un grand volume de données d'entraînement. Dans ce contexte, il devient difficile pour le modèle substitut de représenter fidèlement le comportement du modèle direct.

- Cas d'usage 2

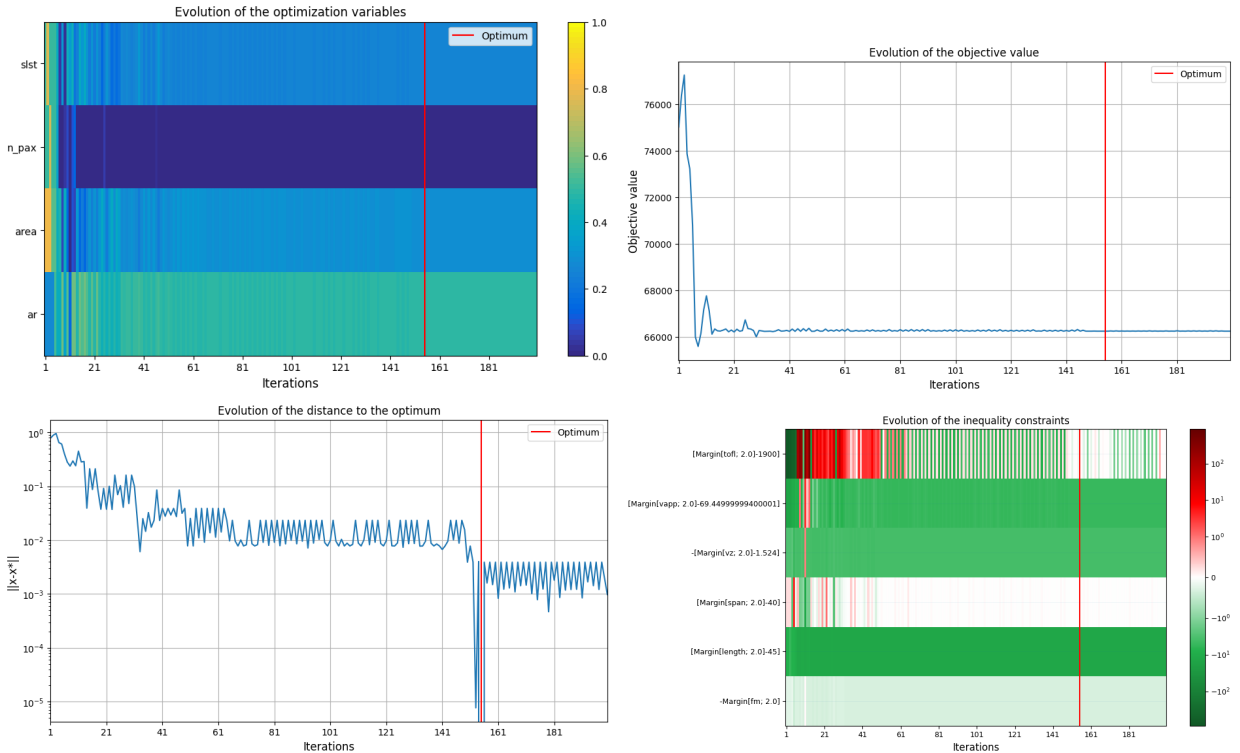


FIGURE 18 – Evolution d'optimisation du cas d'usage 2 par la méthode de surrogate direct

Malheureusement, malgré le changement des différentes méthodes de régression, lorsque nous testons les résultats de ces méthodes sur le modèle réel, la contrainte TOFL n'est toujours pas respectée.

```

INFO - 15:37:29: Solution:
WARNING - 15:37:29: The solution is not feasible.
INFO - 15:37:29: Objective: 63700.75098336771
INFO - 15:37:29: Standardized constraints:
INFO - 15:37:29: -Margin[fm; -2.0] = [-0.25183269]
INFO - 15:37:29: -[Margin[vz; -2.0]-1.524] = [-3.40922554]
INFO - 15:37:29: [Margin[length; 2.0]-45] = [-13.]
INFO - 15:37:29: [Margin[span; 2.0]-40] = [-2.50381337]
INFO - 15:37:29: [Margin[tofl; 2.0]-1900] = [1025.47544372]
INFO - 15:37:29: [Margin[vapp; 2.0]-69.44999999400001] = [2.42229642]

```

FIGURE 19 – Contraint tofl n'est pas vérifié

La raison pourrait être que le R^2 du modèle surrogate n'est pas parfait (0,93), et que la solution tombe dans une zone de mauvaise prédiction. Cela reste à vérifier.

5.3 Méthode par surrogate local itératif

Dans la continuité des approches précédentes, nous proposons une méthode par surrogate local itératif visant à surmonter certaines limitations des méthodes antérieures. Contrairement à l'approximation linéaire de la méthode de Taylor, qui peut être trop simplificatrice, et à la méthode par surrogate direct, qui pâtit de la complexité liée à la forte dimensionnalité de x et U , cette approche itérative reconstruit un modèle surrogate à chaque itération. Cela permet d'affiner progressivement la représentation du modèle en se focalisant sur les régions pertinentes de l'espace des paramètres. L'algorithme de cette méthode est détaillé dans le tableau 1.

Plutôt que de construire un modèle surrogate global intégrant simultanément x et u comme variables d'entrée, cette méthode adopte une stratégie séquentielle. À chaque itération k , un modèle surrogate local $\hat{f}_{x_k}(u)$ est construit en fixant la variable x à x_k . Ce modèle approxime localement la fonction objectif f en fonction de la variable aléatoire u .

La valeur de f pour x_k est ensuite estimée par échantillonnage Monte Carlo de u à l'aide du surrogate. La valeur suivante, x_{k+1} , est déterminée en minimisant cette approximation locale. Ce processus itératif se poursuit jusqu'à satisfaire un critère de convergence ou atteindre un nombre maximal d'itérations.

Cette approche réduit la quantité de données nécessaires à chaque itération et permet de mieux maîtriser le coût computationnel, notamment pour les problèmes de haute dimension.

5.3.1 Algorithme.

Voici le pseudo-code de cette approche séquentielle :

Algorithm 1 Optimisation séquentielle avec surrogate local

- 1: Initialiser $x_0 = x_{\text{défaut}}$
 - 2: **while** $k < \text{MaxIter}$ && critère de convergence ne atteint pas **do**
 - 3: Échantillonner $u_1, \dots, u_N \sim P(u)$
 - 4: Construire un modèle surrogate $\hat{f}_{x_k}(u)$ en fixant $x = x_k$
 - 5: Calculer $x_{k+1} = \arg \min_x \mathbb{E}_u [\hat{f}_x(u)]$ par moyenne Monte Carlo
 - 6: **end while**
 - 7: **return** x_{k+1}
-

Cette méthode présente plusieurs avantages. Elle nécessite moins de données d'apprentissage à chaque itération, car le modèle substitut est construit localement pour une valeur donnée de x . Elle est ainsi moins sensible à la dimension de la variable d'entrée x , ce qui la rend bien adaptée aux problèmes de forte dimension où l'évaluation du modèle direct est coûteuse. De plus, le coût d'apprentissage et d'optimisation reste réduit à chaque itération. Cette approche séquentielle permet également d'ajuster dynamiquement le surrogate en

fonction de l'avancement de l'optimisation, ce qui peut améliorer la stabilité du processus et la qualité des solutions obtenues, notamment lorsque les ressources de calcul sont limitées.

En revanche, cette méthode présente certaines limitations. Elle peut nécessiter un nombre important d'itérations pour converger, en particulier si la fonction objective est fortement non linéaire ou si elle comporte plusieurs minima locaux. Par ailleurs, les interactions entre x et u sont prises en compte de manière plus approximative, puisque le surrogate est localisé autour de x_k et ne modélise pas globalement la dépendance de f à x . La qualité de la solution obtenue dépend donc directement de la précision locale des surrogates successifs.

Nous obtenons donc le résultat suivant.

- Cas d'usage 1 :

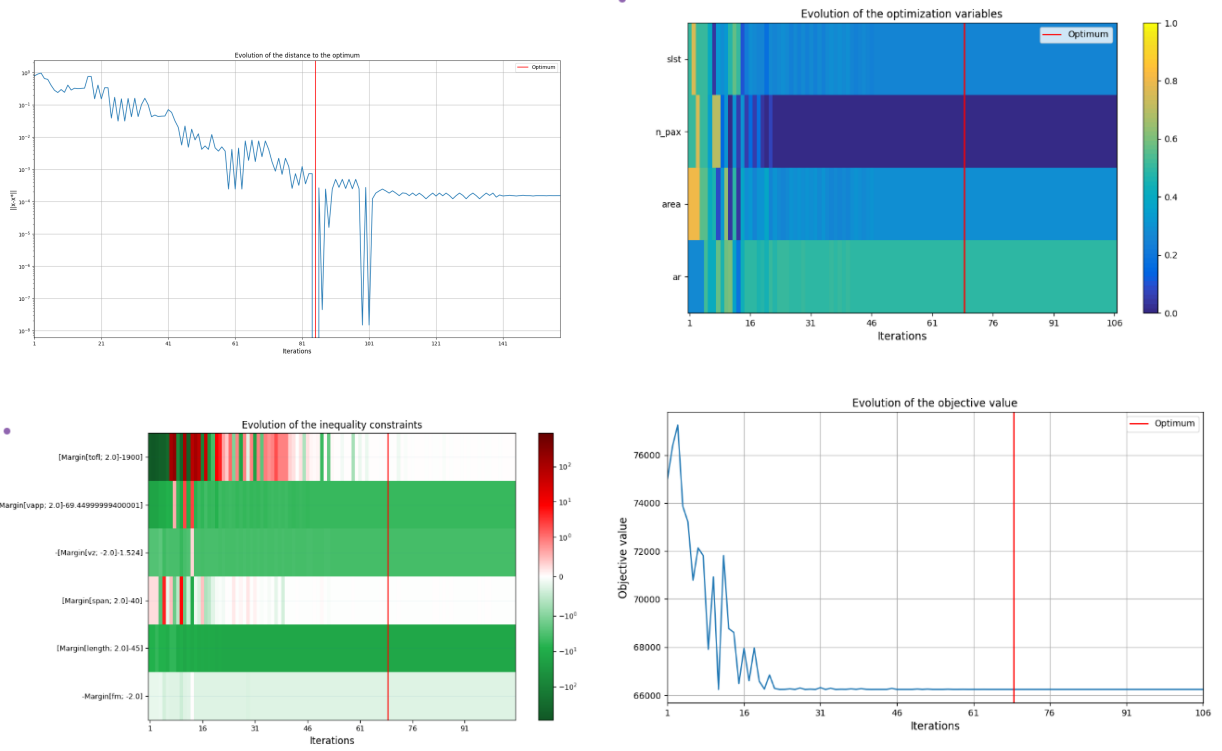


FIGURE 20 – Evolution d'optimisation du cas d'usage 1 par la méthode de surrogate modèle itératif

- Cas d'usage 2 :

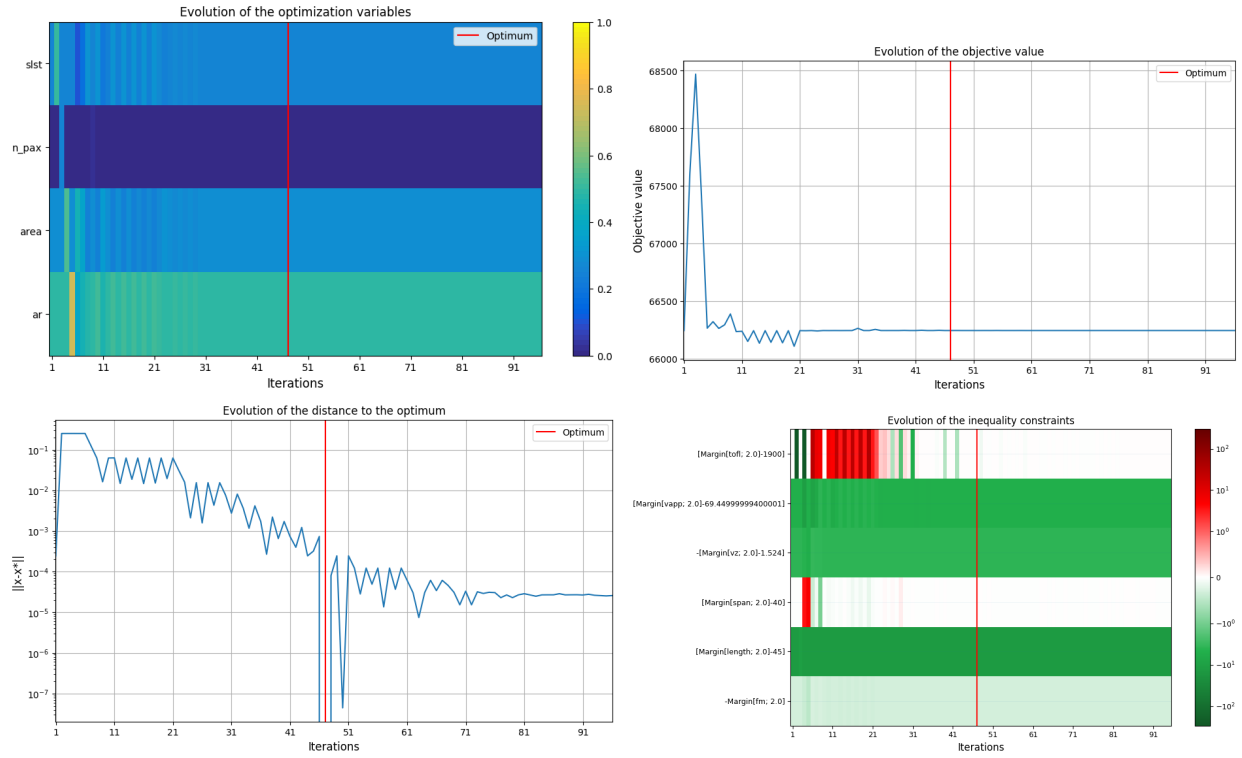


FIGURE 21 – Evolution d’optimisation du cas d’usage 2 par la méthode de surrogate modèle itératif

L’optimisation a été correctement réalisée et la solution obtenue est très proche de celle de la première méthode. Nous avons vérifié ce résultat dans un scénario réel propre à la discipline concernée, et toutes les contraintes sont respectées.

5.4 Visualisation de l'aéronef optimisé par des modèles surrogate

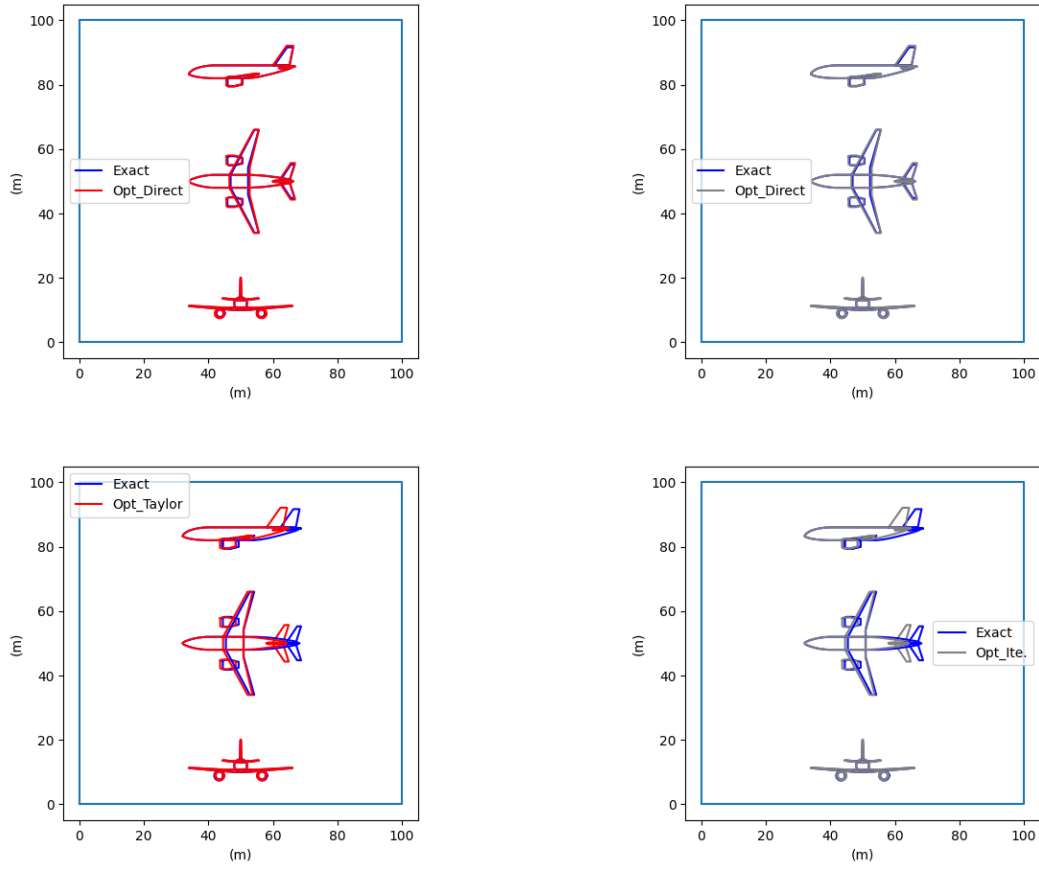


FIGURE 22 – Top : Use case 1. Bottom : Use case 2

6 Conclusion

Dans ce rapport, nous avons abordé l'optimisation d'un avion de 150 passagers environ, en considérant différents types de carburant et de moteurs, avec pour objectif principal de minimiser la masse maximale au décollage (**mtom**) tout en respectant les contraintes opérationnelles. À travers les trois problèmes étudiés, nous avons développé et comparé diverses approches de métamodélisation et d'optimisation, en tenant compte des incertitudes inhérentes à la conception aéronautique.

Dans le **Problème 1**, nous avons construit des métamodèles basés sur **RBFR regressor** et **MOER regressor** à partir d'un échantillonnage limité. Le **Problème 2** a introduit la prise en compte des incertitudes, modélisées par des distributions de probabilité pour les paramètres incertains. Enfin, dans le **Problème 3**, nous avons proposé trois méthodes pour intégrer les incertitudes dans l'optimisation : le développement de Taylor, le surrogate direct et le surrogate local itératif.