HAL
open science

L'incertitude en conception : formalisation, estimation

Gilles Pujol, Rodolphe Le Riche, Olivier Roustant, Xavier Bay

To cite this version:

Gilles Pujol, Rodolphe Le Riche, Olivier Roustant, Xavier Bay. L'incertitude en conception : formalisation, estimation. Optimisation Multidisciplinaire de Systèmes Mécaniques 2 : réduction de modèles, robustesse, fiabilité, réalisations logicielles, Hermes publisher, 2009, Mécanique et Ingénierie des Matériaux. hal-00409905

HAL Id: hal-00409905

https://hal.science/hal-00409905v1

Submitted on 12 Apr 2019

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

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

Chapitre 1

L'incertitude en Conception : Formalisation, Estimation

« *Optimization is actually just the opposite of robustness* »
J. Marczyk [MAR 00]

1.1. Introduction

Intégrer les incertitudes dans le processus de conception est une pratique bien ancrée chez les ingénieurs. Cela va du dimensionnement de systèmes pour des valeurs critiques, de l'utilisation de coefficients de sécurité, aux techniques plus évoluées du calcul de fiabilité. Le but est de concevoir un système ayant *statistiquement* les meilleures performances, ces dernières étant susceptibles de varier à cause des incertitudes. Par exemple, on peut rechercher des performances faiblement sensibles aux incertitudes (critère de *robustesse des performances*). On peut aussi vouloir ne pas dépasser un seuil de performances minimales avec une probabilité donnée (critère de *fiabilité des performances*). De plus, les problèmes de conception sont toujours des problèmes d'optimisation sous contraintes. En présence d'incertitudes, on va vouloir qu'elles soient vérifiées avec une grande probabilité (critère de *fiabilité du système*).

Les critères de robustesse et de fiabilité sont généralement antagonistes au critère de performance. On présentera au paragraphe 1.2 une structure simple composée de deux barres, qui, optimisée sans tenir compte des incertitudes, a un taux de fiabilité inférieur à 50%. On comprend qu'il s'agit de faire un compromis : de combien est-on prêt à dégrader les performances pour quel niveau de robustesse ou de fiabilité ?

Chapitre rédigé par Gilles PUJOL et Rodolphe LE RICHE et Olivier ROUSTANT et Xavier BAY.

Ainsi, comme le souligne Egorov [EGO 02], le problème de conception en présence d'incertitudes est fondamentalement un problème multi-critères.

L'optimisation en présence d'incertitudes concerne deux aspects : l'optimisation, et l'estimation de l'incertitude. Elle en cumule surtout les difficultés. L'objectif de ce chapitre est double :

1) formaliser l'incertitude et voir comment elle peut s'intégrer dans un problème d'optimisation ; nous présentons d'abord un exemple simple de mécanique des structures (paragraphe 1.2), puis nous considérons le problème de manière plus générale (paragraphe 1.3) ;

2) présenter des méthodes d'estimation qui pourront être utilisées dans une boucle d'optimisation ; ces méthodes sont de type Monte Carlo (paragraphe 1.4) ou font appel à construction de modèles approchés, appelés couramment *méta-modèles* (paragraphe 1.5).

Ainsi, bien que le propos concerne essentiellement le volet incertitude, on n'oublie jamais qu'au final, le but est d'optimiser.

REMARQUE.– Les programmes ayant servi pour la conception de ce chapitre (résultats numériques, figures) ont été réalisés avec le logiciel R [R D 08] et sont téléchargeables à l'adresse
`http://www.emse.fr/~roustant/deuxbarres.zip`

1.2. Le cas test des deux barres

Le cas test des deux barres est un exemple simple de mécanique des structures qui servira de fil conducteur tout au long du chapitre.

1.2.1. *Modèle déterministe*

Le modèle des deux barres [JIN 03] décrit une structure composée de deux tubes creux de mêmes dimensions, chacun fixé au sol, et reliés entre-eux à leur sommet, comme représenté sur la figure 1.1.

On applique une force F dirigée vers le bas sur le sommet de la structure. On cherche d, le diamètre de la section des tubes, et L, la longueur des tubes, de façon à minimiser le volume total V. La structure a deux modes de défaillance, selon que la contrainte normale s dépasse la limite d'élasticité s_{max} ou la contrainte critique de flambage s_{crit}. Le problème de conception s'écrit donc :

$$\begin{cases} \min\limits_{d,L} V \\ \text{tel que } s \leq s_{\text{max}} \\ \text{et } s \leq s_{\text{crit}} \end{cases} \tag{1.1}$$

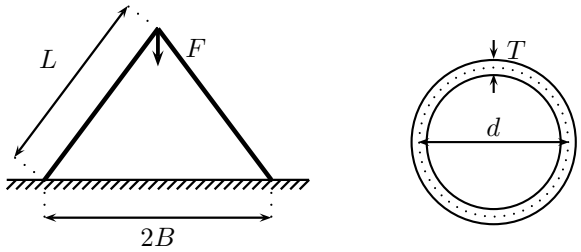


Figure 1.1. *Le cas test des deux barres. À gauche, la structure des deux barres, à droite, la section en coupe d'une barre.*

Paramètre		Valeur nominale	Écart-type (C.V.)
d	diamètre de la section des tubes (mm)	20 – 80	1 ($<$ 5%)
L	longueur des tubes (mm)	800 – 1200	5 ($<$ 1%)
F	force appliquée (N)	150000	30000 (20%)
E	module d'Young (N/mm^2)	210000	21000 (10%)
B	demi-largeur de la structure (mm)	750	5 ($<$ 1%)
T	épaisseur de la section des tubes (mm)	2,5	0,1 (4%)
s_{max}	limite d'élasticité (N/mm^2)	400	- -

Tableau 1.1. *Paramètres du modèle des deux barres. Les perturbations autour des valeurs nominales sont gaussiennes, chaque paramètre indépendamment des autres (C.V. = coefficient de variation = écart-type / moyenne). Source : [HUA 07].*

Le modèle est suffisamment simple pour que les grandeurs d'intérêt V, s et s_{crit} s'écrivent analytiquement :

$$V = 2\pi T d L \tag{1.2}$$

$$s = \frac{FL}{2\pi T d \sqrt{L^2 - B^2}} \tag{1.3}$$

$$s_{\text{crit}} = \frac{\pi^2 E (T^2 + d^2)}{8L^2} \tag{1.4}$$

Les variables du modèle et leurs valeurs sont résumées dans les deux premières colonnes du tableau 1.1 (la troisième servant au paragraphe 1.2.2).

La figure 1.2 représente les lignes de niveau du volume V en fonction des paramètres d et L, les autres paramètres étant fixés à leurs valeurs nominales, avec en

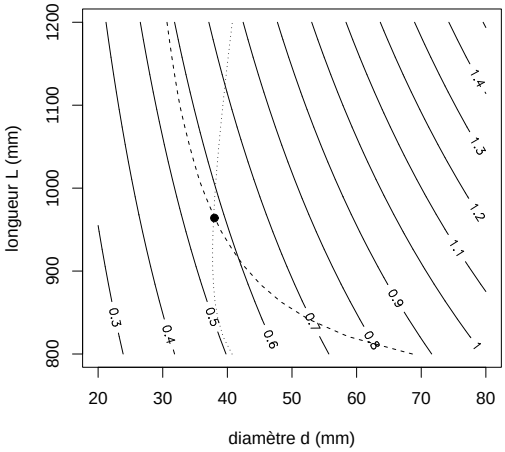


Figure 1.2. *Lignes de niveau pour le volume V de la structure (en dm^3). En pointillés, les lignes d'état critique pour les contraintes (traits : $s = s_{max}$, points : $s = s_{crit}$), avec défaillance à gauche de chaque ligne. L'optimum ($d^* = 38, L^* = 964$) est représenté par un point.*

superposition, les lignes d'état critique $s = s_{\text{max}}$ et $s = s_{\text{crit}}$. L'optimum est à la rencontre de ces deux lignes : c'est le point $d^* = 38$, $L^* = 964$ (en mm entiers). Les sorties du modèle en ce point sont[1] :

```
> twobars(c(38, 964))
          V        s    scrit
1 0.5754141 399.9931 404.3131
```

Le volume est exprimé en dm^3.

1.2.2. *Incertitude associée*

On considère maintenant de l'incertitude sur les paramètres du modèle. Cette incertitude est de type gaussienne, centrée sur la valeur nominale de chaque paramètre, et ses écarts-types sont indiqués dans la troisième colonne du tableau 1.1 (l'incertitude

1. Les lignes qui suivent montrent l'évaluation des commandes utilisées en R (voir la remarque sur R page 2). D'autres extraits de la session R sont donnés dans ce paragraphe.

sur chaque paramètre étant indépendante des autres). Généralement, les conditions d'utilisation du système représentent la plus grande source d'incertitude ; c'est bien le cas ici avec le chargement F (20% en coefficient de variation).

Lorsqu'on propage l'incertitude dans le modèle pour l'optimum du problème déterministe (1.1), les quantités V, s et s_{crit} deviennent aléatoires. Le résumé statistique du volume[2] est le suivant :

```
> summary(V)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.4776  0.5561  0.5752  0.5753  0.5939  0.6874
```

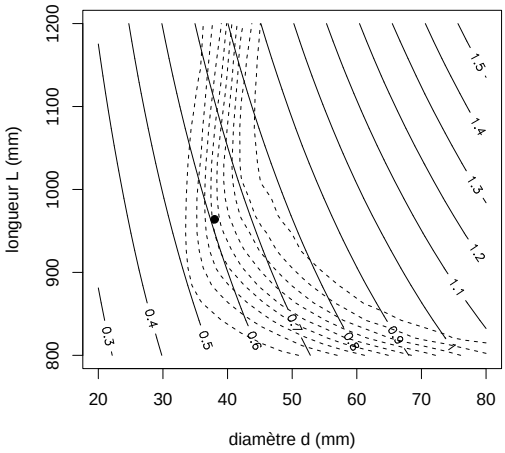
La moyenne coïncide avec les performances du problème sans incertitude. La dispersion autour de cette moyenne est importante, environ 20% de chaque côté. On ne satisfait donc pas le critère de robustesse. Concernant le taux de fiabilité du système, c'est-à-dire la probabilité que les deux contraintes soient vérifiées,

```
> mean(s < smax  &  s < scrit)
[1] 0.429
```

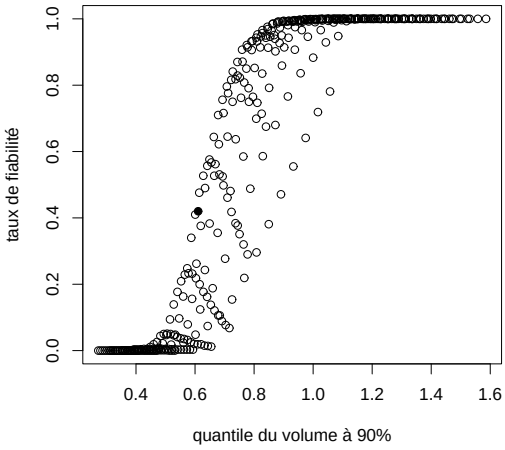
Il y a défaillance du système dans plus de la moitié des cas. Ce design n'est donc pas fiable du tout !

L'étude que nous venons de réaliser concerne un seul design $x^* = (d^*, L^*)$. Les figures 1.3 représentent l'évolution de deux critères en fonction de d et L : le quantile à 90% du volume et le taux de fiabilité. En 1.3(a) sont représentées les lignes de niveau pour les deux critères. Le domaine des designs acceptables n'est plus délimité par une frontière comme sur la figure 1.2, mais par un "front probabiliste". En 1.3(b), le taux de fiabilité est représenté en fonction du quantile. On note que ces critères sont bien antagonistes, et on peut voir se dessiner un front de Pareto. La solution cherchée est sur ce front de Pareto, en fonction du compromis que l'on souhaite réaliser. Noter que dans ce cas précis, l'optimum du problème déterministe (1.1) correspond bien à une solution optimale du problème avec incertitudes, mais pour un compromis en général innacceptable : améliorer les performances au détriment de la fiabilité.

2. Pour un échantillon de 1000 tirages, ce résumé est composé des plus faible et plus grande valeures des échantillons, `Min` et `Max`, de la moyenne et de la médiane, `Mean` et `Median`, ainsi que des premier et troisième quartiles, `1st Qu.` et `3rd Qu.`, respectivement. Les premier et troisième quartiles sont les quantiles à 25% et 75% de la variable aléatoire V. Le quantile à α% d'une variable V, q_α, est défini par $\mathrm{P}(V \leq q_\alpha) = \alpha$.



(a) Lignes de niveau du quantile à 90% du volume (trait plein) et du taux de fiabilité du système (pointillés)



(b) Quantile à 90% du volume vs. taux de fiabilité

Figure 1.3. *Propagation d'incertitude sur le modèle des deux barres. Les critères probabilistes sont évalués sur une grille* (d, L) *de taille* 20×20 *par méthode de Monte Carlo CRN (voir paragraphe 1.4.2), avec* $n = 1000$ *tirage à chaque fois, soit au total 400000 évaluations du modèle. L'optimum du problème déterministe est repéré sur les deux figures par un point noir.*

REMARQUE.– On a repris le modèle des deux barres tel qu'il est présenté dans la littérature ([JIN 03, HUA 07]), à la différence que l'on étudie la longueur de tube L à la place de la hauteur de la structure $H = \sqrt{L^2 - B^2}$; en effet, il nous semble plus exact de considérer une incertitude, ici gaussienne, sur un paramètre physique comme L. Pour bien faire les choses, deux autres points pourraient être améliorés. Premièrement, puisqu'on prend en compte une incertitude sur le matériau à travers le module d'Young E, il faudrait aussi en considérer une sur la limite d'élasticité $s_{\max}$ (voir tableau 1.1) ; ces quantités sont a priori non indépendantes. Ensuite, pour être réaliste, l'incertitude doit être propagée indépendamment sur les deux barres car ce sont deux système physiques indépendants ; ce point est étudié au paragraphe 1.7.

1.3. Formulation du problème de conception en présence d'incertitude

À travers l'exemple précédent, on a pu voir quelques caractéristiques de la conception en présence d'incertitude. Dans ce paragraphe, le problème est formulé de manière générale.

1.3.1. *Paramétrisation contrôle-bruit*

On distingue souvent deux types d'incertitudes : l'incertitude de paramètres, et l'incertitude de modèles. L'incertitude de paramètres est due aux aléas ou au manque de connaissance sur certains paramètres du modèle. L'incertitude de modèle est due aux approximations associées à la modélisation et/ou aux résolutions numériques. Cette distinction est en fait sans objet, car, dans la pratique, les modèles ne sont manipulés que par l'intermédiaire de paramètres. C'est pourquoi on ne considèrera que de l'incertitude de paramètres, sans perte de généralité.

On note y le modèle numérique, et θ le vecteur des paramètres de ce modèle. Dans les problèmes avec incertitude, les paramètres θ dépendent de deux types de variables : les *variables de conception*, regroupées dans le vecteur x, et les *variables de bruit*, regroupées dans le vecteur U. L'évaluation du modèle se note donc (voir aussi figure 1.4) :

$$y(\theta(x, U))$$

et plus simplement $y(x, U)$. Le vecteur x est un vecteur de nombres pris dans un espace $\mathcal{X}$, généralement du type $\prod_{i=1}^{d}[a_i, b_i]$, tandis que le vecteur U est un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathrm{P})$. On supposera que la loi du vecteur aléatoire U ne dépend pas de x. Cette hypothèse est nécessaire pour pouvoir distinguer le contrôle du bruit dans le modèle, et n'est généralement pas restrictive, comme illustré dans l'exemple suivant.

EXEMPLE.– Un paramètre θ est soumis à une incertitude gaussienne, et on contrôle sa valeur nominale ainsi que sa classe de qualité, c'est-à-dire sa moyenne et son écart-type. Dans ce cas on écrira $x = (m, \sigma)$, et $\theta = m + \sigma U$ avec $U \sim \mathcal{N}(0, 1)$.

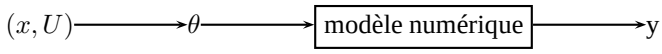


Figure 1.4. *La chaîne de calcul : paramètres de conception x, variables de bruit U, paramètres θ du modèle numérique, y sa réponse.*

Dans la pratique, on rencontre très souvent les deux cas particuliers suivants :

1) contrôle bruité : $\theta = x + U$. Par exemple, x décrit la géométrie nominale d'un objet, et $U \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$ représente l'incertitude liées à la fabrication. Ce cas est typique de la conception fiabiliste (voir par ex. [BAT 06]).

2) bruit exogène au contrôle : $\theta = (x, U)$. Par exemple, si x représente toujours la géométrie d'un objet, U peut représenter dans ce cas les incertitudes liées aux conditions d'utilisation, par exemple les chargements. Ce cas est typique de la conception robuste (voir par ex. [CHE 05]).

Dans le cas des deux barres, les paramètres du modèle sont $\theta = (d, L, F, E, B, T)$. Les variables de contrôle sont les valeurs nominales des paramètres d et L : on note $x = (\bar{d}, \bar{L})$. Contrairement au deux cas particuliers précédents, l'incertitude concerne ici aussi bien des paramètres de contrôle que d'autres paramètres du modèle. On note $U = (\varepsilon_d, \varepsilon_L, \varepsilon_F, \varepsilon_E, \varepsilon_B, \varepsilon_T)$ le vecteur aléatoire, de composantes gaussiennes indépendantes, de paramètres donnés dans le tableau 1.1. Avec ces notations, l'application $\theta(x, U)$ est définie par :

$$d = \bar{d} + \varepsilon_d,\ L = \bar{L} + \varepsilon_L,\ F = \varepsilon_F,\ E = \varepsilon_E,\ B = \varepsilon_B,\ T = \varepsilon_T \tag{1.5}$$

REMARQUE.– La notation $y(x, U)$ sous-entend que la variabilité du phénomène est explicitement donnée au modèle numérique par l'utilisateur, c'est-à-dire que l'on a effectivement une fonction $y(x, u)$ dont on peut évaluer la valeur pour une réalisation $u = U(\omega)$ donnée. Les cas où la variabilité est intrinsèque au modèle numérique – que l'on noteraient plus volontier sous la forme d'un processus stochastique $Y(x)$ – sont beaucoup plus rares et a priori exclus.

1.3.2. *L'incertitude dans le problème d'optimisation*

Une fois l'incertitude définie, il reste à l'intégrer dans le problème d'optimisation. Les problèmes d'optimisation déterministes les plus généraux s'écrivent

$$\begin{cases} \min\limits_{\theta} f_i(\theta)\ , \quad i = 1, \ldots, p \\ \text{avec } \theta \in \Theta\ , \\ \text{tel que } g_j(\theta) \leq 0\ , \quad j = 1, \ldots, q \\ \text{et } h_k(\theta) = 0\ , \quad k = 1, \ldots, r \end{cases} \tag{1.6}$$

où Θ est l'espace de recherche. Les f_i, g_i et h_i sont appelés critères d'optimisation. Dans ce chapitre, on ne considèrera que le cas $p = 1$, et on notera plus simplement

f l'unique fonction objectif. Tous les critères d'optimisation peuvent être vus comme des calculs effectués à partir des réponses du modèle $y(\theta)$. Ainsi, $f(\theta)$ est l'écriture contractée de $f(\theta, y(\theta))$, et il en est de même pour les autres critères.

REMARQUE.– Cette écriture a pour inconvénient de cacher le modèle numérique y qui consomme l'essentiel du temps de calcul. C'est pourtant un élément important à prendre en compte dans l'optimisation. Typiquement, les critères peuvent avoir des coûts numériques différents. Par exemple, dans le cas où les fonctions coût sont moins chères à évaluer que les contraintes, la stratégie courante est de construire un méta-modèle pour les contraintes.

Les incertitudes se propagent sur les critères d'optimisation f, g_j et h_k. On les considère alors via des statistiques : moyenne, variance, ...

1.3.2.1. *Cas de la fonction objectif*

Les statistiques les plus usitées pour la fonction objectif sont :

– la moyenne,

$$m(x) := \mathrm{E}(f(x, U)) = \int_\Omega f(x, u)\, \mathrm{P}_U(\mathrm{d}u) \tag{1.7}$$

– la variance,

$$v(x) := \mathrm{var}(f(x, U)) = \mathrm{E}\big((f(x, U) - m(x))^2\big) \tag{1.8}$$

– le α-quantile,

$$q_\alpha(x) \text{ tel que } \mathrm{P}(f(x, U) \leq q_\alpha(x)) = \alpha \tag{1.9}$$

On cherche bien entendu à minimiser ces critères. Ils vont être reformulés au paragraphe suivant (1.3.2.2) pour prendre en compte la dépendance avec les contraintes d'inégalité.

REMARQUE.– Le quantile (1.9) est connu en finance sous le nom de *value-at-risk* (VaR). Voir [AND 04] pour d'autres mesures de risques utilisées en finance.

1.3.2.2. *Cas des contraintes d'inégalité*

Soit $\mathcal{G}(x)$ l'évènement aléatoire "toutes les contraintes d'inégalité sont respectées",

$$\mathcal{G}(x) = \bigcap_{j=1}^{q} \{g_j(x, U) \leq 0\} \tag{1.10}$$

Une première approche naïve serait de chercher à maximiser la probabilité d'être dans l'espace des réalisables :

$$\max_x \mathrm{P}(\mathcal{G}(x)) \tag{1.11}$$

mais, comme on l'a vu avec l'exemple des deux barres au paragraphe 1.2.2, ceci est antinomique au critère de performance sur la fonction coût : plus le taux de fiabilité est grand, moins bonnes sont les performances. Le plus souvent, le cahier des charges impose le compromis en fixant le taux de fiabilité minimal β. Le critère (1.11) se récrit alors comme une contrainte d'inégalité :

$$\mathrm{P}(\mathcal{G}(x)) \geq \beta \tag{1.12}$$

REMARQUE.– Dans la littérature, il est presque toujours considéré un niveau de fiabilité différent par critère (voir par ex. [KIM 06], [FON 07]) :

$$\mathrm{P}(g_j(x, U) \leq 0) \geq \beta_j \tag{1.13}$$

En faisant ainsi, on ne garantit le taux de fiabilité du système, i.e. la probabilité que toutes les contraintes soient vérifiées, que si les $g_j(x, U)$ sont indépendants (par rapport à U).

Reconsidérons maintenant les critères sur la performance (1.7), (1.8) et (1.9). Ces critères incluent *toutes* les réalisations $f(x, U(\omega))$ pour $\omega \in \Omega$, c'est-à-dire aussi bien celles qui respectent les contraintes que celles qui ne les respectent pas. Or, ce que l'on cherche, c'est améliorer les performances pour les réalisations qui respectent les contraintes ; celles qui ne les respectent pas, on souhaite juste les éviter – ce que dit exactement (1.12). Pour prendre en compte cela, on peut conditionner les critères de performance par rapport à l'évènement $\mathcal{G}(x)$:

$$m^C(x) := \mathrm{E}(f(x, U) \mid \mathcal{G}(x)) \tag{1.14}$$

$$v^C(x) := \mathrm{var}(f(x, U) \mid \mathcal{G}(x)) \tag{1.15}$$

$$q_\alpha^C(x) \text{ tel que } \mathrm{P}(f(x, U) \leq q_\alpha^C(x) \mid \mathcal{G}(x)) = \alpha \tag{1.16}$$

Ces critères conditionnels étant plus difficiles à estimer, nous ne les utiliserons pas dans la suite.

1.3.2.3. *Cas des contraintes d'égalité*

L'incertitude est difficilement compatible avec les "vraies" contraintes d'égalité, car un évènement du type $h_k(x, U) = 0$ est de probabilité nulle. Très souvent, on peut s'affranchir des contraintes d'égalité en les sortant de la formulation [FON 07]. En effet, en conception, les contraintes d'égalité sont souvent d'origine physique et traduisent un état d'équilibre ; ces égalités font partie du modèle plus que du problème d'optimisation.

EXEMPLE.– Une classe importante de problèmes s'écrit sous la forme

$$\begin{cases} \min\limits_{\theta} f(\theta, y(\theta)) \\ \text{avec } \theta \in \Theta \ , \\ \text{tel que } g_j(\theta, y(\theta)) \leq 0 \ , \quad j = 1, \ldots, q \end{cases} \tag{1.17}$$

où le modèle $y(\theta)$ est la solution du système d'équations

$$h_k(\theta, y(\theta)) = 0 \ , \quad k = 1, \ldots, r \tag{1.18}$$

C'est le cas par exemple lorsque le modèle est un code élément fini. En mécanique du solide, (1.18) exprime l'égalité des efforts internes et externes.

REMARQUE.– Les méthodes SAND (*simultaneous analysis and design*) [HAF 85] constituent la principale exception à cette règle. Elles consistent à résoudre un problème de type (1.17) en réintégrant les égalités (1.18) dans le problème d'optimisation.

1.3.3. *Problème de conception robuste*

Les formulations de type conception robuste ajoutent au problème de performance maximale un impératif de stabilité de la performance vis-à-vis des incertitudes. La stabilité de la performance des solutions peut être ajoutée (par ex.) sous la forme d'une variance maximum. Enfin, lorsque la stabilité de la solution est un critère de conception aussi important que sa performance, la conception robuste est abordée comme un problème multi-critères :

$$\begin{cases} \min\limits_{x} \mathrm{E}(f(x,U)) \\ \text{et } \min\limits_{x} \mathrm{var}(f(x,U)) \\ \text{avec } x \in \mathcal{X} \\ \text{tel que } \mathrm{P}(\mathcal{G}(x)) \geq 1 - \beta \end{cases} \tag{1.19}$$

Le compromis moyenne-variance peut être tranché en introduisant un multiplicateur de Lagrange λ, et le critère à minimiser devient (voir par ex. [JIN 04]) :

$$f_\lambda(x) = \mathrm{E}(f(x,U)) + \lambda\sqrt{\mathrm{var}(f(x,U))} \tag{1.20}$$

On notera ici le présupposé gaussien. En effet, si $f(x,U)$ est une variable aléatoire gaussienne, alors (1.20) en est un quantile.

Les "vrais" problèmes de conception robuste sont très particuliers et nécessitent que les critères moyenne et variance soient antagonistes. Ainsi, le cas des deux barres n'en est pas un car, comme on peut le voir sur les figures 1.6(a) et 1.6(b) (p. 17) , la moyenne et la variance évoluent dans le même sens.

EXEMPLE.– En s'inspirant de [FON 07], on peut modifier le cas test des deux barres pour en faire un problème de conception robuste. Il suffit pour cela :

1) d'inclure un contrôle sur l'incertitude : on rejette les barres de section trop grande, par exemple en les faisant passer dans un gabarit de diamètre $d + \Delta_d$ où Δ_d, la tolérance sur le diamètre d, est une variable de contrôle supplémentaire (cela revient en fait à tronquer la loi de ε_d) ;

2) d'ajouter un critère économique lié à ce contrôle dans la fonction coût : le volume V calculé est le volume de la structure, auquel s'ajoute celui de toutes les barres rejetées.

Dans cet exemple, plus on souhaitera contrôler l'incertitude (i.e. Δ_d faible), et plus on rejettera de barres pour fabriquer le structure, donc plus elle coûtera cher. Ici il y a bien antagonisme entre la robustesse et les performances.

1.3.4. *Problème de conception fiabiliste*

Dans l'objectif de fiabilité, on va chercher à *garantir* le niveau du "plafond" des performances, ce qui se traduit par le critère sur le α-quantile :

$$\begin{cases} \min\limits_{x} q_\alpha(x) \\ \text{avec } x \in \mathcal{X} \\ \text{tel que } \mathrm{P}(\mathcal{G}(x)) \geq 1 - \beta \end{cases} \tag{1.21}$$

Contrairement au problème de conception robuste, les performances à l'optimum ne sont pas nécessairement stables vis-à-vis de l'incertitude. Comme vu précedemment sur l'exemple des deux barres (fig. 1.3(b)), le quantile de la performance et le taux de fiabilité du système sont généralement antagonistes.

1.3.5. *Remarque : problèmes en boucle fermée*

On a supposé de manière implicite que l'incertitude est "en boucle ouverte" (désigné en anglais sous les termes "*feed-forward*" ou "*here and now*" selon les communautés), c'est-à-dire que l'on cherche un design x^* quelle que soit les réalisations $U(\omega_1)$, $U(\omega_2)$, ... C'est tout-à-fait naturel puisque la conception de l'objet est antérieure à sa réalisation (et donc à celle de l'incertitude).

Dans d'autres types de problèmes, par exemple en décision (par ex. optimisation de la distribution d'électricité), ou en calibration de modèle, l'incertitude se matérialise avant la résolution du problème d'optimisation. La formulation pertinente est alors dite "en boucle fermée" ("*feed-back*" ou "*wait and see*", voir [DIW 02]), et consiste à résoudre le problème pour des réalisations $U(\omega)$ quelconques :

$$\begin{cases} \min\limits_{x} f(x, U(\omega)) \\ \text{avec } x \in \mathcal{X} \ , \\ \text{tel que } g_j(x, U(\omega)) \leq 0 \ , \quad j = 1, \ldots, q \\ \text{et } h_k(x) = 0 \ , \quad k = 1, \ldots, r \end{cases} \tag{1.22}$$

ce qui revient à cherche la loi de la variable aléatoire $X^*(\omega)$, solution du problème (1.22) pour tous les $\omega \in \Omega$. Cela est difficile dans le cas général, car cela revient à résoudre

autant de problèmes d'optimisation que de réalisations de l'incertitude $U(\omega)$ suffisantes pour estimer la loi (ou les moments) de la variable aléatoire $X^*(\omega)$.

1.4. Estimation de l'incertitude par des méthodes de Monte Carlo

Pour résoudre les problèmes de conception en présence d'incertitudes, l'approche la plus simple consiste à imbriquer l'estimation des critères probabilistes dans une boucle d'optimisation classique. Il ne s'agit pas ici de discuter de l'(in)efficacité de cette approche, mais de présenter quelques méthodes d'estimation pouvant être utilisées dans cette optique. Plus précisément, nous présentons dans les paragraphes suivants des méthodes de type Monte Carlo, c'est-à-dire qui utilisent des nombres pseudo-aléatoires, et qui ne requièrent pas explicitement la construction d'un méta-modèle (voir pour cela le paragraphe 1.5).

1.4.1. *Estimateurs empiriques*

1.4.1.1. *Estimation ponctuelle*

Pour alléger les notations, on note Z une variable aléatoire quelconque dont on veut estimer l'une ou l'autre des quantités suivantes :

$$m = \mathrm{E}(Z) \tag{1.23}$$

$$v = \mathrm{var}(Z) \tag{1.24}$$

$$s \text{ tel que } \mathrm{P}(Z \leq s) = \alpha \tag{1.25}$$

$$p = \mathrm{P}(Z \leq 0) \tag{1.26}$$

Dans les problèmes considérés, (1.23), (1.24), (1.25) correspondent aux critères sur la fonction objectif (1.7), (1.8), (1.9) avec $Z = f(x, U)$; (1.26) correspond au taux de fiabilité (1.12) avec $Z = \prod_{j=1}^{q} g_j(x, U)$.

Soit $z_1, \ldots, z_n$ la réalisation d'un n-échantillon de la variable aléatoire Z. Les estimateurs empiriques de (1.23) – (1.26) sont :

$$\hat{m} = \frac{1}{n} \sum_{i=1}^{n} z_i \tag{1.27}$$

$$\hat{v} = \frac{1}{n-1} \sum_{i=1}^{n} (z_i - \hat{m})^2 \tag{1.28}$$

$$\hat{s} = z_{(\lfloor n\alpha \rfloor)} \tag{1.29}$$

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} 1_{]-\infty,0]}(z_i) \tag{1.30}$$

où la notation $z_{(\lfloor n\alpha \rfloor)}$ désigne la $\lfloor n\alpha \rfloor^{\text{e}}$ plus petite valeur parmi les n valeurs $z_1, \ldots, z_n$ ($\lfloor . \rfloor$ désignant la partie entière inférieure).

1.4.1.2. *Estimation par intervalles*

Pour quantifier l'incertitude associée aux estimations précédentes, on peut calculer des intervalles de confiance, c'est-à-dire des intervalles qui ont une probabilité donnée, disons $1-\gamma$, de contenir les valeurs théoriques (1.23) – (1.26).

Pour la moyenne et la variance, en supposant que Z est gaussienne, on a les intervalles de confiance suivants :

$$m \in \left[\hat{m} - t_{\gamma/2}\sqrt{\frac{\hat{v}}{n}},\ \hat{m} + t_{\gamma/2}\sqrt{\frac{\hat{v}}{n}}\right] \quad \text{avec proba } 1-\gamma \tag{1.31}$$

$$v \in \left[\frac{(n-1)\hat{v}}{\chi_2^2},\ \frac{(n-1)\hat{v}}{\chi_1^2}\right] \quad \text{avec proba } 1-\gamma \tag{1.32}$$

où $t_{\gamma/2}$ est le quantile d'ordre $\gamma/2$ d'une loi de Student à $n-1$ degrés de libertés, et χ_1^2 (resp. χ_2^2) est le quantile d'ordre $1-\gamma/2$ (resp. $\gamma/2$) d'une loi du χ^2 à $n-1$ degrés de libertés. Par exemple, pour $n = 100$ et $\gamma = 0.2$, on a $t_{\gamma/2} \simeq -1.29$, $\chi_1^2 \simeq 81$ et $\chi_2^2 \simeq 117$.

REMARQUE.– Si Z ne peut pas raisonnablement être supposée gaussienne, la technique du bootstrap [EFR 93] permet alors d'estimer les intervalles de confiance.

Pour le quantile s, un intervalle de confiance est donnée par la formule de Wilks [WIL 41] :

$$\mathrm{P}(s \in [z_{(i_1)}, z_{(i_2)}]) = \sum_{k=i_1}^{i_2-1} C_n^k \alpha^k (1-\alpha)^{n-k} \tag{1.33}$$

Il s'agit donc de trouver les indices i_1 et i_2 les plus proches tels que la somme (1.33) soit supérieure ou égale à $1-\gamma$. Par exemple, pour $n = 100$, $\alpha = 0.9$, l'intervalle $[z_{(87)}, z_{(95)}]$ a une probabilité légèrement supérieure à 0.8.

Pour la probabilité p, et pour n grand (typiquement np et $n(1-p)$ supérieurs à 5), on a l'intervalle

$$p \in \left[\hat{p} - u_{\gamma/2}\sqrt{\frac{p(1-p)}{n}},\ \hat{p} + u_{\gamma/2}\sqrt{\frac{p(1-p)}{n}}\right] \quad \text{avec proba } 1-\gamma \tag{1.34}$$

où $u_{\gamma/2}$ est le quantile d'ordre $\gamma/2$ d'une loi $\mathcal{N}(0,1)$. Par exemple, pour $\gamma = 0.2$, $u_{\gamma/2} \simeq -1.28$.

REMARQUE.– Tous ces résultats sont classiques et peuvent se trouver (à l'exception de la formule de Wilks) dans [SAP 90].

pour $i = 1, \ldots, N$ **faire**
  Déterminer le point x_i (selon la méthode d'optimisation);
  Tirer un n-échantillon $u_i^1, \ldots, u_i^n$ de loi U;
  pour $l = 1, \ldots, n$ **faire**
    $y_i^l := f(x_i, u_i^l)$;
    si $\prod_{j=1}^q g_j(x_i, u_i^l) \leq 0$ **alors** $z_i^l := 1$ **sinon** $z_i^l := 0$;
  fin
  $\hat{m}_i := \frac{1}{n} \sum_{l=1}^n y_i^l$;
  $\hat{v}_i := \frac{1}{n-1} \sum_{l=1}^n (y_i^l - \hat{m}_i)^2$;
  $\hat{s}_i := y_i^{n\alpha:n}$;
  $\hat{p}_i := \frac{1}{n} \sum_{l=1}^n z_i^l$;
fin

Algorithme 1.1 : Algorithme d'estimation par méthode de Monte Carlo simple, imbriquée dans une boucle d'optimisation.

1.4.1.3. *Illustration sur le cas test des deux barres*

Pour la résolution du problème d'optimisation, il faut imbriquer les estimations dans la boucle d'optimisation, comme décrit par l'algorithme 1.1 (estimations ponctuelles). Bien entendu, le coût de cette approche est exhorbitant : il est de $N \times n$, où N est le nombre de pas d'optimisation, et n est la taille des échantillons utilisés.

Le but de ce chapitre étant d'étudier seulement l'estimation de l'incertitude, dans la suite nous allons estimer les critères pour des x_i choisis a priori, c'est-à-dire non guidés par une procédure d'optimisation.

Dans un premier temps, nous considérons une grille régulière dans l'espace des x pour explorer toutes les solutions (ce qui est bien entendu irréalisable en pratique !). La figure 1.5 représent les lignes de niveau pour les critères du problème de conception fiabiliste (1.21) (le quantile et le taux de fiabilité). On constate que les critère sont bruités à cause de la variabilité non négligeable des estimations. Ce bruit peut être réduit en augmentant la taille n des échantillons, mais le coût de la méthode augmente encore : si l'on double la taille des échantillons, c'est $2N$ évaluations des fonctions coût en plus qu'il faut réaliser. Ainsi, avec l'approche de Monte Carlo, il faudra "apprendre à vivre avec le bruit" (voir discussion plus loin dans ce paragraphe).

Dans un deuxième temps, les estimation sont réalisées sur une trajectoire particulière $x_1, \ldots, x_n$, représentée par les cercles sur la figure 1.5 (de haut en bas). Tous les critères sont estimés (moyenne, variance, probabilité, taux de fiabilité), ainsi que leurs intervalles de confiance. Les résultats sont représentés sur la figure 1.6. Plusieurs choses (bien connues) sont à noter :

– l'estimation de la moyenne et du quantile sont de bonnes qualités ;

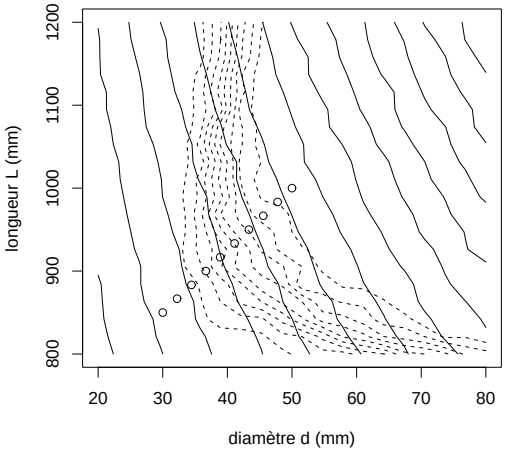


Figure 1.5. *Estimation par Monte Carlo.*

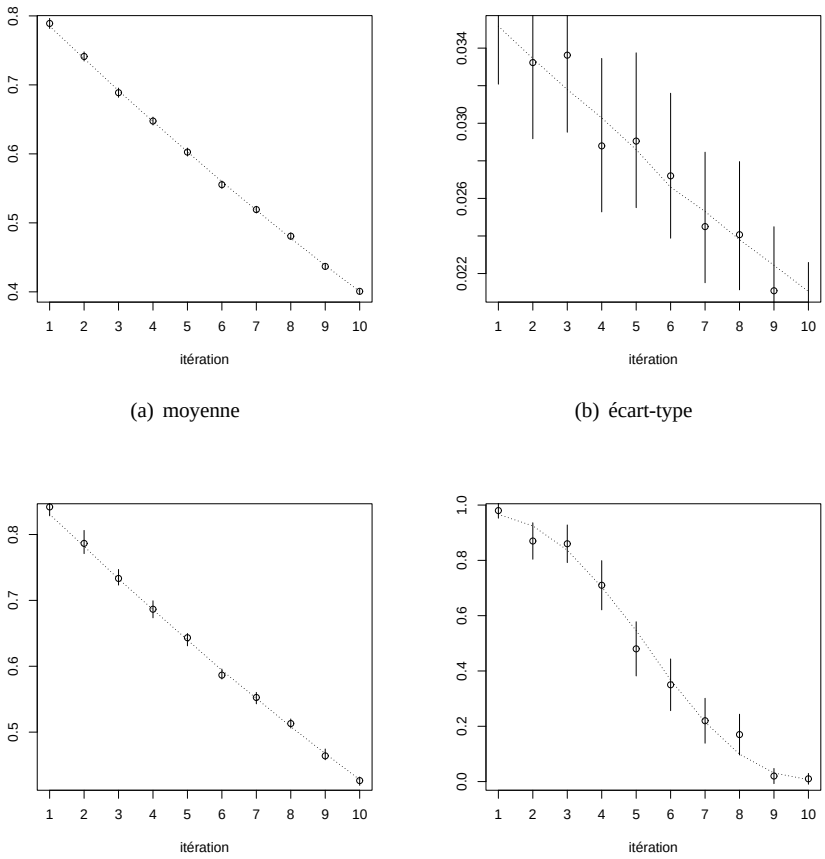
– au contraire, l'estimation de la variance est très mauvaise ;

– la qualité de l'estimation du taux de fiabilité dépend de sa valeur (la variance de l'estimateur vaut $p(1-p)/n$). Ainsi, dans les cas de taux de fiabilités élevés qui nous intéressent, la variance d'estimation est assez faible.

Pour résumer, le problème de conception fiabiliste (1.21) ne fait intervenir que des critères "faciles" à estimer (quantile, taux de fiabilité élevé), comparé au problème de conception robuste (1.19) qui, lui, fait intervenir la variance.

La variabilité des estimations peut perturber le déroulement de l'algorithme d'optimisation. L'algorithme d'optimisation doit donc être choisi en fonction du bruit, donc de la taille d'échantillon n. Pour un n grand, on pourra utiliser des méthodes d'optimisation classiques (par exemple des méthodes de type gradient), tandis qu'un n petit nécessitera l'emploi d'algorithmes robustes au bruit. Il est important ici de ne pas confondre fonction bruitée avec fonction non lisse. L'optimisation dite "non lisse" (voir par ex. Kelley [KEL 99]) recours souvent à des méthodes dont le comportement est déterministe, comme la méthode de Nelder et Mead [][3]. De par leur nature déterministe, ces méthodes supportent très mal les échecs, pourtant inévitables en paysage aléatoire. Ainsi, on veillera plutôt à utiliser des méthodes d'optimisation stochastiques,

3. Voir aussi le chapitre Scilab.



(a) moyenne (b) écart-type

(c) quantile (d) taux de fiabilité

Figure 1.6. *Estimation par Monte Carlo.*

comme par exemple les stratégies évolutionnaire, qui, par nature, tolèrent un certain taux d'échec dans leur progression.

1.4.2. *Stratégie "Common Random Numbers" (CRN)*

1.4.2.1. *Principe*

La stratégie *Common Random Numbers* [YAN 91] (littéralement les nombres aléatoires communs) est une légère variante de la méthode de Monte Carlo simple présentée précédemment. Elle consiste à utiliser toujours le même échantillon $u_1, \ldots, u_n$

Tirer un n-échantillon $u_1, \ldots, u_n$ de loi U;
pour $i = 1, \ldots, N$ **faire**
  Déterminer le point x_i (selon la méthode d'optimisation);
  pour $l = 1, \ldots, n$ **faire**
    $y_i^l := f(x_i, u_l)$;
    si $\prod_{j=1}^{q} g_j(x_i, u_l) \leq 0$ **alors** $z_i^l := 1$ **sinon** $z_i^l := 0$;
  fin
  $\hat{m}_i := \frac{1}{n} \sum_{l=1}^{n} y_i^l$;
  $\hat{v}_i := \frac{1}{n-1} \sum_{l=1}^{n} (y_i^l - \hat{m}_i)^2$;
  $\hat{s}_i := y_i^{n\alpha:n}$;
  $\hat{p}_i := \frac{1}{n} \sum_{l=1}^{n} z_i^l$;
fin

Algorithme 1.2 : Algorithme d'estimation par méthode de Monte Carlo CRN, imbriquée dans une boucle d'optimisation.

pour les estimations aux différents points de la trajectoire d'optimisation $x_1, \ldots, x_N$ (voir algorithme 1.2).

Bien que l'échantillon $u_1, \ldots, u_n$ soit toujours le même, le coût numérique n'est pas allégé, car la variable x, elle, change. En prenant l'exemple de la moyenne, pour estimer tous les $\hat{m}_i$, il faut évaluer $y_i^l = f(x_i, u_l)$ pour $i = 1, \ldots, N$, $l = 1, \ldots, n$ (cf algorithme 1.2), soit $N \times n$ évaluations comme auparavant.

En revanche, il n'y a plus de bruit d'estimation d'un x_i à l'autre. En effet, toujours sur l'exemple de la moyenne, l'estimateur peut s'écrire

$$\hat{m}(x) = \frac{1}{n} \sum_{l=1}^{n} f(x, u_l) \tag{1.35}$$

Si quelque soit u, la fonction $x \mapsto f(x, u)$ est régulière (e.g. continue, $\mathcal{C}^1$, ...), il en sera de même pour la fonction $\hat{m}(x)$. On peut faire le même raisonnement pour les estimateurs $\hat{v}(x)$, $\hat{s}_i(x)$ et $\hat{p}(x)$. Ainsi, avec la stratégie d'estimation CRN, on peut utiliser un algorithme d'optimisation standard (c'est-à-dire pour une fonction régulière). Attention, le fait qu'il n'y ait plus de bruit d'un x_i à l'autre ne signifie pas que les estimations sont plus précises : elles ont exactement la même précision qu'avec la méthode de Monte Carlo simple.

Il y a dans la méthode CRN un certain danger, car les estimations $\hat{m}(x_1)$, ..., $\hat{m}(x_N)$ ne sont plus indépendantes, mais corrélées. Un biais initial, dû à un mauvais échantillon, aura tendance à persister tout au long de l'optimisation. Il est donc très important de bien choisir l'échantillon de départ $u_1, \ldots, u_n$. Par exemple, on pourra utiliser une suite à faible discrépance (méthode de quasi-Monte Carlo, voir [CAF 98]) ou un échantillon de type Hypercube Latin ([MCK 79]).

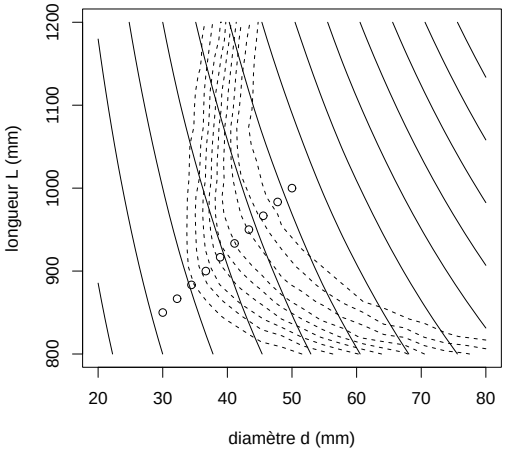


Figure 1.7. *stratégie CRN*

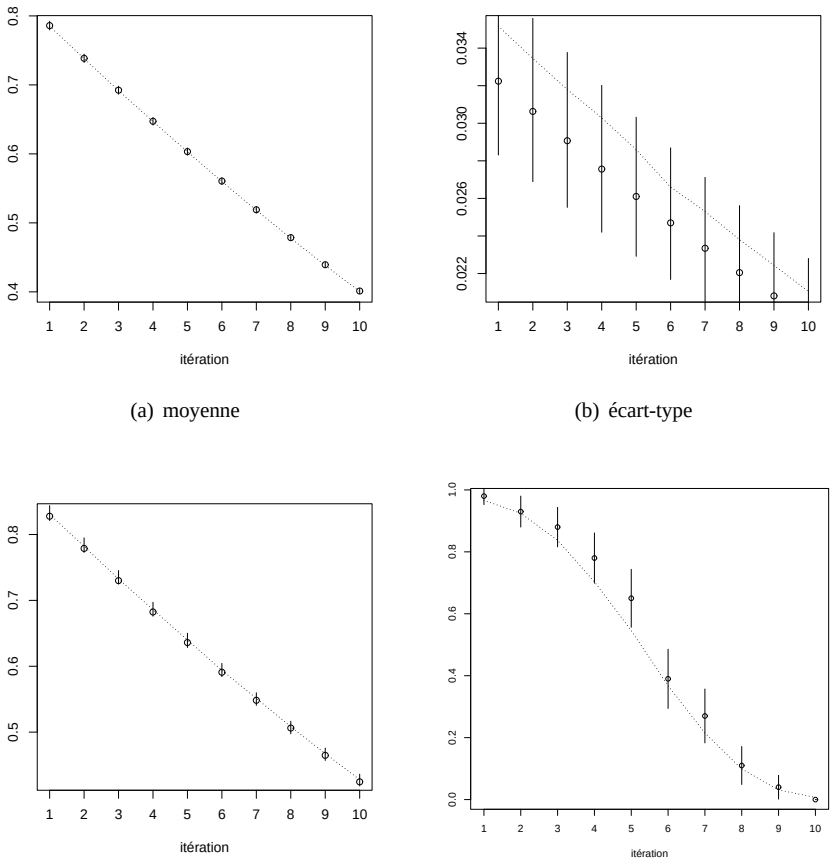
1.4.2.2. *Illustration sur le cas test des deux barres*

Comme au paragraphe 1.4.1.3, on a estimé les critères quantile du volume et taux de fiabilité sur une grille régulière (figure 1.7), ainsi que sur la trajectoire particulière repérée sur cette figure par les cercles (figure 1.8). On constate effectivement qu'il n'y a plus de bruit. Ainsi, une méthode d'optimisation classique ne rencontrerait aucun problème de convergence.

Sur les figures 1.8(b) (écart-type du volume) et 1.8(d) (taux de fiabilité), on peut constater le phénomène de biais évoqué précédemment. Pour l'écart-type du volume, les estimations sont toujours sous-estimées, tandis qu'elles sont toujours sur-estimées dans le cas du taux de fiabilité. Cela vient de l'échantillon de départ qui a été tiré aléatoirement. Un meilleur échantillon aurait permis d'éviter ce phénomène indésirable.

1.4.3. *Méthode de Robbins-Monro*

La méthode de Robbins-Monro fait partie de la famille des algorithmes adaptatifs, appelés par les probabilistes "approximations stochastiques" (voir Benveniste et al. [BEN 87]). Les algorithmes adaptatifs les plus connus sont : les estimation récursive en régression linéaire, *Least-Mean-Squares* (LMS) et *Recursive-Mean-Squares* (RMS), ainsi que le filtre de Kalman (voir par ex. [SPA 03]). La motivation essentielle de tous ces algorithmes est d'estimer une quantité qui varie au cours du temps. Dans



(a) moyenne (b) écart-type

(c) quantile (d) taux de fiabilité

Figure 1.8. *Estimation par Monte Carlo.*

notre contexte, ce sont les statistiques à estimer (moyenne, variance, quantile, taux de fiabilité), qui varient au cours de la procédure d'optimisation.

1.4.3.1. *Principe*

Soit le problème inverse suivant :

$$\text{trouver } \theta^* \text{ tel que } \mathrm{E}(\phi(\theta, U)) = 0 \tag{1.36}$$

où $\phi(\theta, u)$ est une fonction donnée. Ce problème correspond tout à fait à l'estimation des critères qui nous intéressent :

– la moyenne :

$$\theta = m \quad \text{et} \quad \phi(\theta, U) = m - f(x, U) \tag{1.37}$$

– la variance :

$$\theta = \begin{pmatrix} m \\ v \end{pmatrix} \quad \text{et} \quad \phi(\theta, U) = \begin{pmatrix} m - f(x, U) \\ v - (m - f(x, U))^2 \end{pmatrix} \tag{1.38}$$

– le α-quantile :

$$\theta = s \quad \text{et} \quad \phi(\theta, U) = 1_{\{f(x,U) \leq s\}} - \alpha \tag{1.39}$$

– le taux de fiabilité :

$$\theta = p \quad \text{et} \quad \phi(\theta, U) = p - 1_{\{\prod_{j=1}^{q} g_j(x,U) \leq 0\}} \tag{1.40}$$

La méthode de Robbins-Monro [ROB 51] permet de résoudre un problème du type (1.36) de manière itérative : on construit une suite $\theta_1, \theta_2, \ldots$ par le schéma de récurrence

$$\begin{array}{l} \theta_0 \text{ donné} \\ \theta_k = \theta_{k-1} - a_k \phi(\theta_{k-1}, u_k), \quad k \geq 1 \end{array} \tag{1.41}$$

où a_k est une suite décroissante vers zéro telle que $\sum_{k=1}^{\infty} a_k = \infty$ et $\sum_{k=1}^{\infty} a_k^2 < \infty$. Par exemple, $a_k = ak^{-b}$ avec $a > 0$ et $0.5 < b \leq 1$. La décroissance de cette suite permet la convergence de l'algorithme, mais elle ne doit pas être trop rapide pour ne pas provoquer une convergence prématurée. Avec une telle suite, et sous certaines hypothèses complémentaires, la suite θ_n converge vers θ^* solution de (1.36) (voir par ex. [SPA 03]).

EXEMPLE.– Écrivons l'estimateur de la moyenne empirique sous forme récursive :

$$\bar{X}_k = \frac{1}{k} \sum_{j=1}^{k} X_j \tag{1.42}$$

$$= \frac{k-1}{k} \bar{X}_{k-1} + \frac{1}{k} X_k \tag{1.43}$$

$$= \bar{X}_{k-1} - \frac{1}{k} \left(\bar{X}_{k-1} - X_k \right) \tag{1.44}$$

On retrouve bien la forme (1.41) avec $a_k = 1/k$ et $\phi(\theta, u) = \theta - u$, comme pour (1.37).

La méthode d'estimation de Robbins-Monro peut être intégrée dans une boucle d'optimisation de la même manière que précédemment, comme décrit pas l'algorithme 1.3. À chaque pas d'optimisation, la valeur initiale pour l'estimation (le θ_0) est la valeur estimée lors du pas précédent. Ainsi, on peut espérer que, si les pas d'optimisation sont rapprochés (typiquement une méthode d'optimisation locale), les critères à estimer soient peut différents, leur mise à jour par la méthode de Robbins-Monro ne nécessitant alors que peu d'itérations.

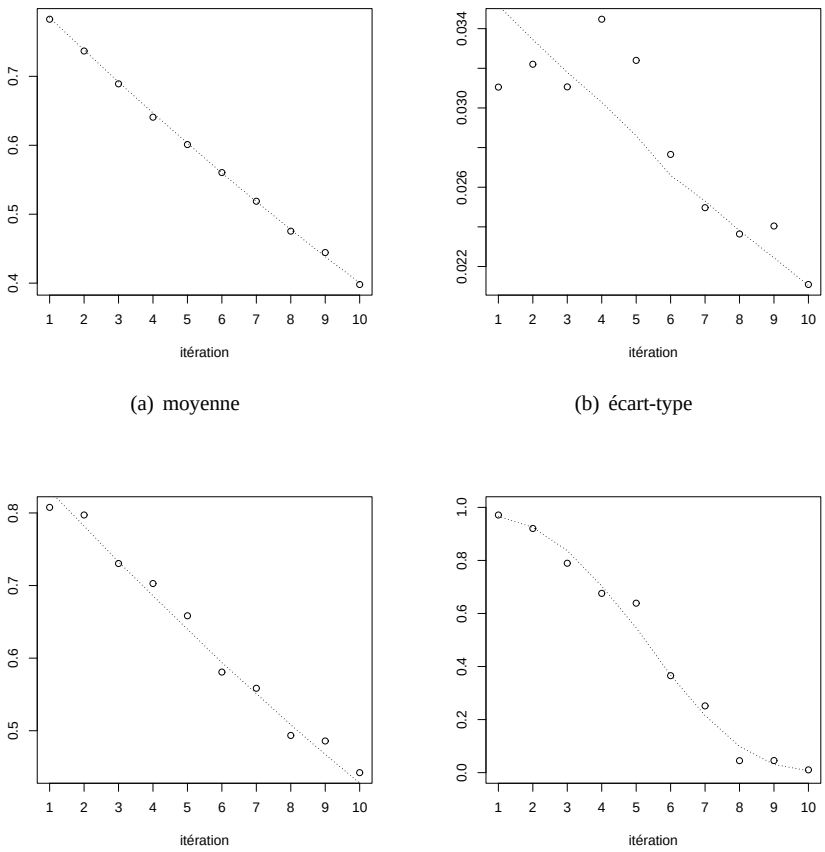
Données : Valeurs initiales $\hat{m}_0^n$, $\hat{v}_0^n$, $\hat{s}_0^n$ et $\hat{p}_0^n$
pour $i = 1, \ldots, N$ **faire**
  Déterminer le point x_i (selon la méthode d'optimisation);
  Initialisation : $\hat{m}_i^0 := \hat{m}_{i-1}^n$, $\hat{v}_i^0 := \hat{v}_{i-1}^n$, $\hat{s}_i^0 := \hat{s}_{i-1}^n$, $\hat{p}_i^0 := \hat{p}_{i-1}^n$;
  pour $l = 1, \ldots, n$ **faire**
    Tirer aléatoirement un point u_l, réalisation de la loi de U;
    $y_i^l := f(x_i, u_l)$;
    $\hat{m}_i^l := \hat{m}_i^{l-1} - a_{l-1}\left(\hat{m}_i^{l-1} - y_i^l\right)$;
    $\hat{v}_i^l := \hat{v}_i^{l-1} - a_{l-1}\left(\hat{v}_i^{l-1} - (\hat{m}_i^l - y_i^l)^2\right)$;
    $\hat{s}_i^l := \hat{s}_i^{l-1} - a_{l-1}\left(1_{]-\infty,0]}(y_i^l - \hat{s}_i^{l-1}) - \alpha\right)$;
    si $\prod_{j=1}^q g_j(x_i, u_l) \leq 0$ **alors** $z_i^l := 1$ **sinon** $z_i^l := 0$;
    $\hat{p}_i^l := \hat{p}_i^{l-1} - a_{l-1}\left(\hat{p}_i^{l-1} - 1_{]-\infty,0]}(z_i^l)\right)$;
  fin
fin

Algorithme 1.3 : Algorithme d'estimation par méthode de Robbins Monro, imbriquée dans une boucle d'optimisation.

1.4.3.2. *Illustration sur le cas test des deux barres*

L'algorithme 1.3 a été implémenté et appliqué sur le cas test des deux barres, sur la même trajectoire x_1, …, x_n que précédemment (cf paragraphes 1.4.1.3 et 1.4.2.2). À chaque pas d'optimisation, 100 itérations de la méthode de Robbins-Monro sont effectuées. Les valeurs initialiales des critères pour la première itération sont choisis très proches de valeurs réelles : $\hat{m}_0^n = 0.78$, $\hat{v}_0^n = 0.036$, $\hat{s}_0^n = 0.82$ et $\hat{p}_0^n = 0.97$ (dans la pratique, on pourra faire une première estimation par la méthode de Monte Carlo simple pour obtenir ces valeurs). Pour la suite (a_k), on a choisi $a_k = k^{-0.8}$. Les estimations des critères à chaque pas d'optimisation sont représentés sur les figures 1.9(a) – 1.9(d). Ces estimations, à part celles de l'écart-type, sont de bonne qualité. Noter que la méthode ne permet pas d'estimer d'intervalles de confiances.

L'intérêt de cette méthode par rapport à la méthode de Monte Carlo simple (paragraphe 1.4.1) peut ne pas être évident, puisque sur cet exemple, nous avons choisi un coût numérique identique. Cependant, que ce coût est "débrayable" d'un pas d'optimisation à l'autre : on peut espérer que la mise à jour d'une estimation demande moins d'itérations qu'une estimation complète, plus encore si la valeur de départ est proche de la valeur à estimer.



(a) moyenne (b) écart-type

(c) quantile (d) taux de fiabilité

Figure 1.9. *Estimation par Monte Carlo.*

1.5. Utilisation d'un méta-modèle

1.5.1. *Différentes approches*

Comme on l'a vu au paragraphe précédent, le coût des méthodes de Monte Carlo ainsi que le bruit d'estimation qui en résulte les rend difficilement utilisables dans une boucle d'optimisation. Le recours à des modèles approchés estimés grâce à des simulation, appelés méta-modèles, est la solution la plus communément utilisée.

On peut distinguer trois approches différences dans la manière d'utiliser un méta-modèle en optimisation avec incertitudes.

1.5.1.1. *Méta-modèle estimé dans la boucle d'optimisation*

Il s'agit, à l'intérieur de la boucle d'optimisaton, de remplacer la boucle d'estimation de l'incertitude par la construction d'un méta-modèle. Ainsi, à chaque pas d'optimisation x_i, on construit un méta-modèle pour les fonctions

$$u \mapsto f(x_i, u) \tag{1.45}$$

$$u \mapsto g_j(x_i, u), \quad j = 1 \ldots, q \tag{1.46}$$

Le coût de cette approche est toujours en $N \times n$, où n est le nombre de points nécessaire à l'estimation des méta-modèles. Cependant, n est beaucoup plus petit que pour les méthodes de Monte Carlo.

Cette approche est présentée sur une application dans le chapitre [DUV 08].

D'autres applications sont données par Kim et al. [KIM 06] et Lucor et al. [LUC 07].

1.5.1.2. *Méthodologie des surfaces de réponses*

Les méthodes de Taguchi [TAG 04] utilisent des plans d'expériences pour la conception robuste. Ces méthodes ont connu un très grand succès, si bien que dans certaines communauté, on confonde "conception robuste" et "méthodes de Taguchi". Cependant, bien que les concepts à la base de ces méthodes soient pertinents, leur application s'est révélée être malhabile, comme l'ont montré Welch et al. [WEL 90], et Shomaker et al. [SHO 91, SHO 93]. Pour pallier à ces défauts, ils ont proposé d'appliquer la méthodologie des surfaces de réponse (RSM : *Response Surface Methodology*) dans le cadre de la conception robuste.

REMARQUE.– Cette approche correspond généralement à un problème dans lequel il n'y a pas de contraintes g_j (donc pas de taux de fiabilité considéré), et formulé en moyenne - variance. C'est le cas le plus simple en conception robuste.

Premièrement, on construit un plan d'expériences dit croisé, car obtenu à partir de deux plans d'expériences, l'un en x, $\mathcal{P}_x = (x_i)_{i=1,\ldots,N}$, et l'un en u, $\mathcal{P}_u = (u_l)_{l=1,\ldots,n}$. Le plan croisé $\mathcal{P}_{x\times u}$ est obtenu en faisant correspondre chaque x_i à chaque u_j, soit $n \times N$ points. Cela correspond en fait à la double boucle bien connue, mais il faut noter que les points sont répartis sur tout le domaine de variation (espace des points (x, u)), et non pas localement (i.e. en certains x_i) comme auparavant. Le but est ici d'explorer le phénomène dans sa globalité.

En notant $y_i^l = f(x_i, u_l)$ la valeur de la fonction coût aux points du plan $\mathcal{P}_{x \times u}$, les critères sur la moyenne et la variance sont estimés par

$$m_i = \frac{1}{n} \sum_{l=1}^{n} y_i^l \tag{1.47}$$

$$v_i = \frac{1}{n-1} \sum_{l=1}^{n} (y_i^l - m_i)^2 \tag{1.48}$$

Ensuite, on estime des modèles de régression pour ces critère, typiquement des modèles quadratiques :

$$m_i \approx \sum_{j=1}^{d} \alpha_i x_{ij} + \sum_{j=1}^{d} \sum_{j'=1}^{d} \beta_{jj'} x_{ij} x_{ij'} \tag{1.49}$$

$$v_i \approx \sum_{j=1}^{d} \gamma_i x_{ij} + \sum_{j=1}^{d} \sum_{j'=1}^{d} \delta_{jj'} x_{ij} x_{ij'} \tag{1.50}$$

Cette approche est connue sous le nom de modélisation duale (voir Vining et Myers [VIN 90], Bates et al. [BAT 06]).

L'optimisation est ensuite réalisé sur ces modèles de régression. L'approche est en générale séquentielle : on raffine le plan d'expériences dans la zone où l'optimum a été identifié et on itère...

REMARQUE.– Les méta-modèles sur la moyenne et sur la variance sont ici estimés indépendamment. Or souvent, l'évolution de la moyenne et de la variance est liée, comme par exemple notre cas des deux barres (voir figures 1.6(a) et 1.6(b)). Les modèles joints (Mc Cullagh et al. [MCC 89]) permettent de prendre en compte cette dépendance.

1.5.1.3. *Modélisation des fonctions coût*

Cette approche est assez semblable de la précédente puisqu'elle consiste à construire un plan d'expériences global sur l'espace des points (x, u), plan à partir duquel est estimé un ou plusieurs méta-modèles. Cependant, elle diffère dans la manière de construire le plan d'expériences et le méta-modèle :

– le plan d'expériences n'est plus "croisé" mais quelconque :

$$\mathcal{P} = (x_i, u_i)_{i=1,\ldots,n} \tag{1.51}$$

– les méta-modèles ne modélisent pas les critères à optimiser, mais les fonctions coût

$$(x, u) \mapsto f(x, u) \tag{1.52}$$

$$(x, u) \mapsto g_j(x, u), \quad j = 1, \ldots, q \tag{1.53}$$

L'optimisation est ensuite réalisée en remplaçant les fonctions coût par les méta-modèles estimés. Les critère probabilistes (moyenne, variance, ...) sont estimé sur le méta-modèle soit par méthode de Monte Carlo, soit analytiquement (Chen et al. [CHE 05]). Les méta-modèles gagneront à être ré-estimés tout au long de l'optimisation avec des techniques de planification adaptative (Lehman et al. [LEH 04]).

De manière quasi générale, le type de méta-modèle utilisé dans cette approche est un modèle de krigeage (par exemple, les références citées précédemment, et aussi Jin et al. [JIN 03], Apley et al.[APL 06], Bates et al. [BAT 06]). Dans la suite, après avoir rappelé le modèle de krigeage, nous l'utiliserons pour modéliser les fonction coût du cas test des deux barre.

1.5.2. *Modèle de krigeage*

On considère une fonction $y(x)$ que l'on veut modéliser.

REMARQUE.– Par soucis de clareté, les notations de ce paragraphe ne sont pas cohérentes avec le début ou la suite du chapitre, mais il n'y a pas d'ambiguïté.

Dans le modèle de krigeage, on considère la fonction y comme une réalisation d'un processus gaussien

$$Y(x) = m(x) + Z(x) \tag{1.54}$$

ce qui signifie que

$$y(x) = Y_{(\omega)}(x) \tag{1.55}$$

pour un $\omega \in \Omega$ particulier (mais inconnu).

La fonction $m(x)$ est un terme de tendance qui peut être quelconque, mais généralement donnée comme une combinaison linéaire de fonctions formant une base orthogonale :

$$m(x) = \beta_0 + \sum_i \beta_i f_i(x) \tag{1.56}$$

Le processus stochastique $Z(x)$ est un processus gaussien stationnaire, c'est-à-dire, centré ($\mathrm{E}(Z(x)) = 0\ \forall x$) et dont la covariance entre deux points ne dépend que de leur distance :

$$\mathrm{cov}(Z(x), Z(x')) = \sigma^2 r(x - x') \tag{1.57}$$

La fonction de covariance traduit de façon probabiliste une hypothèse de régularité : les valeurs $Y(x)$ et $Y(x')$ ont d'autant plus de chances d'être proches que les points x et x' le sont. La fonction de covariance généralement utilisée est la fonction de covariance gaussienne :

$$\mathrm{cov}(Z(x), Z(x')) = \sigma^2 \exp\left(-\sum\left(\frac{x_j - x'_j}{\phi_j}\right)^2\right) \tag{1.58}$$

Les paramètres ϕ_j, homogènes à une longueur, peuvent s'interpréter comme des paramètres de fréquence spatiale : une petite valeur permettra des petites variations dans la j^{e} direction, tandis qu'une grande valeur imposera un comportement beaucoup plus lisse.

Étant donné un plan d'expériences $\mathcal{P} = (x_i)_{i=1,\ldots,n}$ et les réponses correspondantes $y_i = y(x_i)$, la fonction $y(x)$ peut être approchée par la moyenne de toutes les trajectoires du processus $Y(x)$ qui passent par ces points :

$$\hat{y}(x) = \mathrm{E}(Y(x) \mid Y(x_i) = y_i, i = 1, \ldots, n) \tag{1.59}$$

$\hat{y}(x)$ s'appelle le *prédicteur de krigeage*. Ce prédicteur s'écrit comme combinaison linéaire des observations y_i (voir par ex Santner et al. [SAN 03]).

La formule (1.59) est donnée pour des paramètres σ^2, β_i et ϕ_j connus. Dans la pratique, ces paramètres sont inconnus. Ils sont estimés, par exemple par maximum de vraisemblance ([SAN 03]). Ces estimations sont ensuite remplacées dans l'expression du prédicteur, qui est alors utilisé comme méta-modèle.

1.5.3. *Illustration sur le cas test des deux barres*

Si on applique telle quelle la méthodologie du paragraphe 1.5.1.3 au cas test des deux barres, on doit construire les trois méta-modèles suivants :

$$\hat{V} \approx \quad (x, u) \mapsto V(x, u) \tag{1.60}$$

$$\hat{s} \approx \quad (x, u) \mapsto s(x, u) \tag{1.61}$$

$$\widehat{s_{\text{crit}}} \approx \quad (x, u) \mapsto s_{\text{crit}}(x, u) \tag{1.62}$$

$$\tag{1.63}$$

Nous allons voir que ce n'est pas la meilleure manière de construire les méta-modèles, ni pour les entrées, ni pour les sorties.

1.5.3.1. *Entrées : sur quel espace de facteurs construire le méta-modèle ?*

On rappelle que $x = (\bar{d}, \bar{L})$, $U = (\varepsilon_d, \varepsilon_L, \varepsilon_F, \varepsilon_E, \varepsilon_B, \varepsilon_T)$, et les paramètres θ du modèle sont donnés par

$$d = \bar{d} + \varepsilon_d,\ L = \bar{L} + \varepsilon_L,\ F = \varepsilon_F,\ E = \varepsilon_E,\ B = \varepsilon_B,\ T = \varepsilon_T \tag{1.64}$$

Séparons les variables de bruit en deux groupes : $U = (U_C, U_N)$, où $U_c = (\varepsilon_d, \varepsilon_L)$ sont des perturbations de variables de contrôle, et $U_N = (\varepsilon_F, \varepsilon_E, \varepsilon_B, \varepsilon_T)$ introduisent un bruit exogène au contrôle. Ainsi, les paramètres du simulateur sont donnés par la relation

$$\theta = (x + U_C, U_N) \tag{1.65}$$

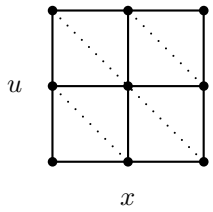


Figure 1.10. *Plan factoriel* 2^3 *(2 facteurs,* x *et* u*, trois niveaux chaque). Les points reliés par des pointillés sont tels que* $\theta = x + u$ *(5 valeurs différentes de* θ*).*

Faire un plan d'expériences pour les variables (x, u) revient à travailler sur l'espace produit $\mathcal{X} \times \mathcal{U}$. Dans cet espace, on a la relation

$$y(x, (u_C, u_N)) = y(x', (u'_C, u'_N)) \text{ dès que } x + u_C = x' + u'_C \text{ et } u_N = u'_N \quad (1.66)$$

avec $y = V$, s, ou s_{crit}. Or, cette relation (qui est une relation d'équivalence), n'est pas prise en compte lors de la construction du méta-modèle, ce qui a deux inconvénients :

1) On travaille sur un espace de dimension plus grande que la dimension effective du problème. Ici, les variables (x, u) évoluent dans huit dimensions, alors qu'il n'y a que six paramètres dans le modèle. Ainsi, une partie de l'information contenue dans le plan d'expériences est gaspillée pour reconstruire la relation (1.66), pourtant très bien connue. En prenant un exemple simple d'une seule variable x et d'une seule variable u, si on utilise un plan factoriel à trois niveaux pour chaque facteur, soit neuf points on va faire quatre calculs inutiles car il y a seulement cinq valeurs différences de $\theta = x + u$ (voir figure 1.10).

2) Un bon échantillonage dans l'espace $\mathcal{X} \times \mathcal{U}$ ne se traduit pas par un bon échantillonage dans l'espace de paramètres Θ. En reprenant l'exemple précédent, si on simule uniformément le carré de façon à remplir l'espace, la loi de $\theta = x + u$ est alors triangulaire. Il y aura donc une surabondance de simulations autour de paramètres nominaux, et une pénurie pour des jeux de paramètres en bord de domaine.

Il vaut donc mieux, tout simplement, construire un plan d'expériences pour les paramètres θ, à partir duquel seront estimés les méta-modèles :

$$\hat{V} \approx \quad \theta \mapsto V(\theta) \quad (1.67)$$

$$\hat{s} \approx \quad \theta \mapsto s(\theta) \quad (1.68)$$

$$\widehat{s_{\text{crit}}} \approx \quad \theta \mapsto s_{\text{crit}}(\theta) \quad (1.69)$$

Cependant, cette approche souffre encore d'un problème, concernant les réponses cette fois, comme nous allons le voir immédiatement.

1.5.3.2. *Sorties : le problème de la modélisation multi-réponses*

En utilisant le logiciel R, un plan d'expériences de type Hypercube Latin (LHS) de $n = 60$ points (en dimension 6) a été généré. Ce plan a été optimisé pour avoir de bonnes propriétés de remplissage de l'espace, au sens d'un critère "maximin"[4]. Des modèles de krigeage ont ensuite été estimés[5] (covariance gaussienne, tendance linéaire sans interactions) pour construire les approximations (1.67), (1.68) et (1.69). La validation de ces trois méta-modèles est réalisée sur une base de test de $N = 100$ points. Les résultats de cette validation sont représentés sur les figures 1.11(a), 1.11(b) et 1.11(c). On peut observer que l'approximation est précise pour V et s_{crit}, mais pas pour s.

La contrainte s intervient dans les deux contraintes d'inégalité du problème d'optimisation :

$$s - s_{\text{max}} \leq 0 \tag{1.70}$$

$$s - s_{\text{crit}} \leq 0 \tag{1.71}$$

Si on utilise les méta-modèles estimés précédemment, c'est-à-dire que l'on remplace s par $\hat{s}$ et s_{crit} par $\widehat{s_{\text{crit}}}$, on considère alors les approximations suivante :

$$\hat{s} - s_{\text{max}} \leq 0 \tag{1.72}$$

$$\hat{s} - \widehat{s_{\text{crit}}} \leq 0 \tag{1.73}$$

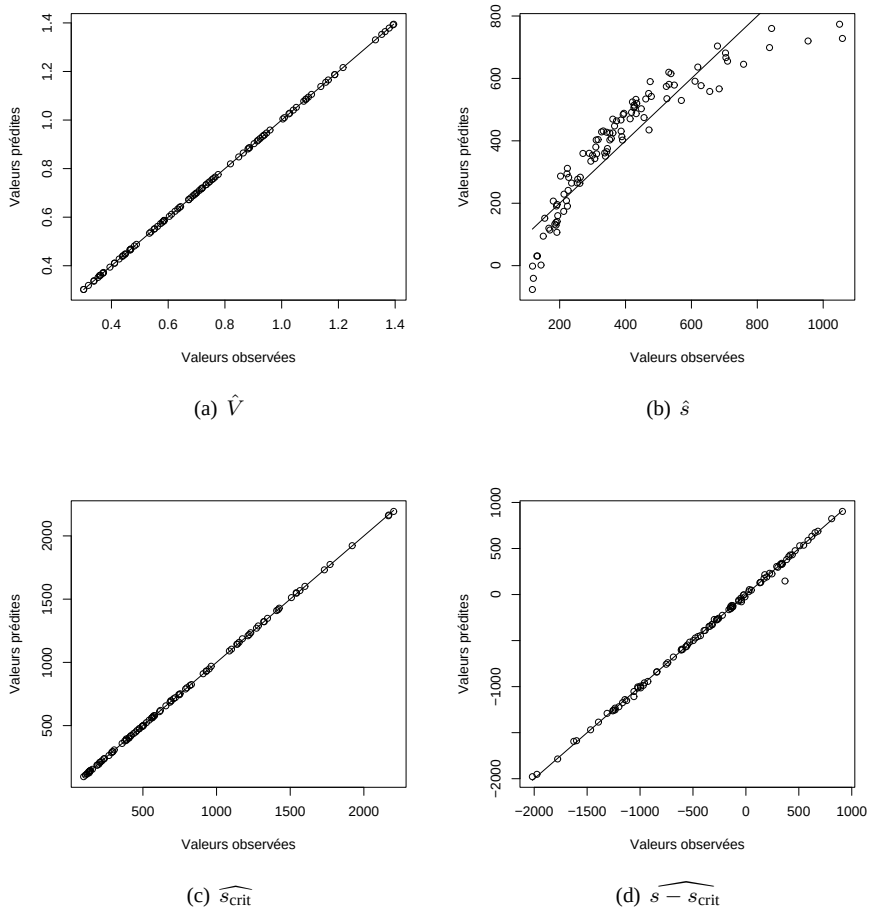
pour évaluer la ruine ou non de la structure. Ces approximations sont représentées respectivement sur les figures 1.12(a) et 1.13(a). On constate que les méta-modèles ne permettent pas de correctement représenter les contraintes. Il ne permettront donc pas d'évaluer la fiabilité de la structure des deux barres.

Construisons maintenant un méta-modèle directement sur la seconde contrainte d'inégalité du problème d'optimisation, méta-modèle noté $\widehat{s - s_{\text{crit}}}$. Comme on peut le voir sur le graphique de validation 1.11(d), ce méta-modèle est de très bonne qualité. Cependant, il reste toujours le problème de la première contrainte d'inégalité. La seconde contrainte d'inégalité va ici être d'une aide précieuse : grâce à elle, on peut récrire la première contrainte en ne faisant intervenir que des quantités que nous savons méta-modéliser :

$$s - s_{\text{max}} = s - s_{\text{crit}} + s_{\text{crit}} - s_{\text{max}} \approx \widehat{s - s_{\text{crit}}} + \widehat{s_{\text{crit}}} - s_{\text{max}} \tag{1.74}$$

4. fonction `maximinLHS` du package R `lhs`.

5. Package R `DiceKriging`.



(a) $\hat{V}$ (b) $\hat{s}$

(c) $\widehat{s_{\text{crit}}}$ (d) $\widehat{s - s_{\text{crit}}}$

Figure 1.11. *Validation des méta-modèles sur une base de test.*

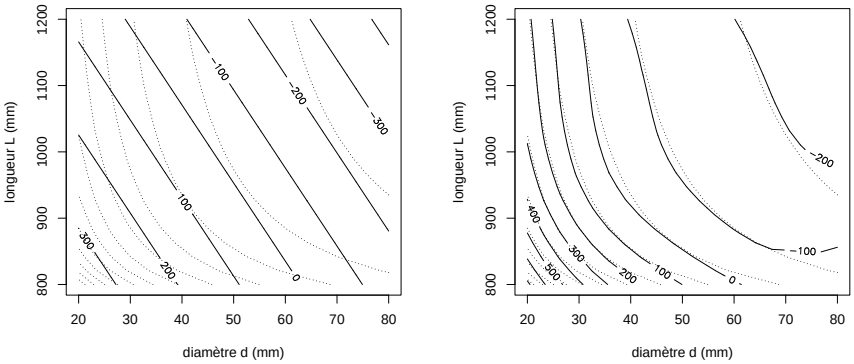
Ainsi, la ruine ou non de la structure est évalué grâce aux approximations suivantes :

$$\widehat{s - s_{\text{crit}}} + \widehat{s_{\text{crit}}} - s_{\text{max}} \leq 0 \tag{1.75}$$

$$\widehat{s - s_{\text{crit}}} \leq 0 \tag{1.76}$$

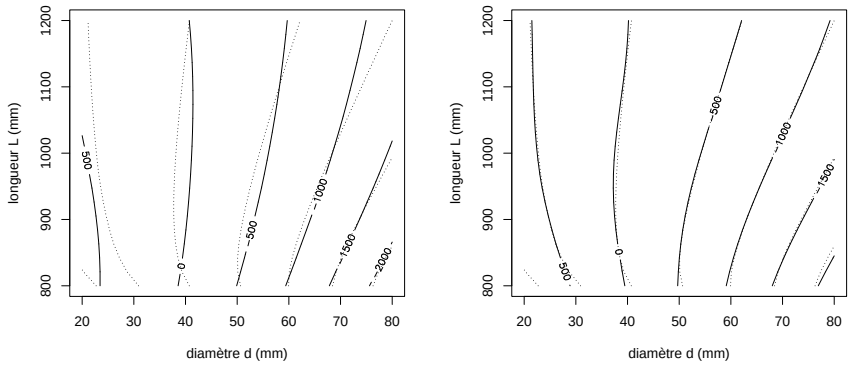
On peut voir sur les figures 1.13(a) et 1.13(b) que ces approximations représentent assez fidèlement les contraintes réelles du problème.

REMARQUE.– Les bons résultats pour la première contrainte d'inégalité s'expliquent d'une part par le fait que les approximations $\widehat{s - s_{\text{crit}}}$ et $\widehat{s_{\text{crit}}}$ sont de bonne qualité, et



(a) approximation par $\hat{s} - s_{\max}$ (b) approximation par $\widehat{s - s_{\text{crit}}} + \widehat{s_{\text{crit}}} - s_{\max}$

Figure 1.12. *Approximation de la première contrainte ($s - s_{max}$)*



(a) approximation par $\hat{s} - \widehat{s_{\text{crit}}}$ (b) approximation par $\widehat{s - s_{\text{crit}}}$

Figure 1.13. *Approximation de la deuxième contrainte ($s - s_{crit}$)*

d'autre part par le fait qu'elles sont négativement corrélées (ainsi, la variance de la somme est plus faible que la somme des variances), ce qui est ici assez intuitif.

Maintenant que nous avons de bons méta-modèles, nous pouvons nous en servir pour estimer les critères statistiques du problème de conception fiabiliste (quantile, taux de fiabilité). Ces critères sont calculés sur les méta-modèles par la méthode de

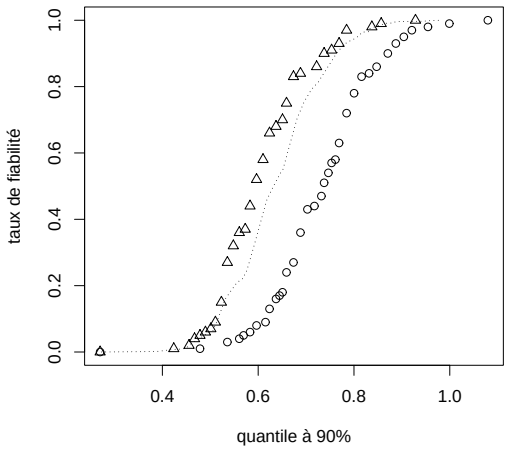


Figure 1.14. *Front de Pareto quantile – taux de fiabilité réel (en pointillé), estimé avec l'approche* (1.72)–(1.73) *(ronds), et avec l'approche* (1.75)–(1.76) *(triangles).*

Monte Carlo avec des échantillons de taille $n = 100$, avec la stratégie d'échantillonnage CRN présentée au paragraphe 1.4.2 pour ne pas avoir de bruit d'estimation. Ces estimations sont réalisées pour tous les points x sur une grille régulière. La figure 1.14 représente le front de Pareto quantile – taux de fiabilité. On voit que la stratégie de méta-modélisation choisie respecte assez fidèlement le front de Pareto, et donc permet de choisir le *bon* compromis du problème de conception fiabiliste. À l'inverse, on constate que l'approche de méta-modélisation naturelle, qui consiste à approximer directement les sorties du modèle conduit à une sous-évaluation du taux de fiabilité : ainsi, pour un taux de fiabilité donné.

1.6. Conclusion

Dans ce chapitre, nous avons tenté de fournir quelques repères généraux pour aborder un problème connu pour être difficile. Nous espérons avoir pu faire *ressentir* des choses qui nous semblent essentielles, comme la différence entre un problème de

conception robuste et un problème de conception fiabiliste, ou la nécessité de considérer le lien entre la méthode d'estimation de l'incertitude et l'algorithme d'optimisation[6]... À ce stade, le lecteur peut se sentir assez démuni, car aucune méthode permettant de résoudre le problème de conception en présence d'incertitudes n'a été présentée. Ce sentiment de frustration pourra être comblé, par exemple, avec la lecture du chapitre 8 du livre de Keane et Nair [KEA 05], qui font une revue exhaustive des méthodes couramment appliquées en ingénierie. Un point de vue plus théorique, par Beyer et al. [BEY 07], fait le complément.

Et pour finir, à titre d'annexe, nous expliquons pourquoi tous les résultats que nous avons présentés sur l'exemple des deux barres sont... complètement faux !

1.7. Annexe : incertitudes et symétries

Comme mentionné au paragraphe 1.2.2 (dernière remarque), propager l'incertitude sur le modèle des deux barres tel que nous l'avons présenté (et tel qu'il est utilisé dans la littérature) n'a pas grand sens physique. Les deux barres étant deux systèmes physiques indépendants, il faut considérer deux longueurs L_1 et L_2, une pour chaque barre, deux modules d'Young E_1 et E_2, ... ce qui revient à refaire le modèle. d'incertitudes.

REMARQUE.– Dans la pratique on n'a jamais la possibilité de refaire la modélisation comme on va le faire dans la suite. Ce paragraphe n'est dont pas un exemple, mais une illustration d'un problème important – et souvent négligé – concernant la propagation d'incertitudes.

Une fois dédoublé, le modèle des deux barres s'écrit :

$$V = \pi(T_1 d_1 L_1 + T_2 d_2 L_2) \tag{1.77}$$

$$s_1 = \frac{F B_2 L_1}{2\pi T_1 d_1 B \sqrt{L_1^2 - B_1^2}} \tag{1.78}$$

$$s_2 = \frac{F B_1 L_2}{2\pi T_2 d_2 B \sqrt{L_2^2 - B_2^2}} \tag{1.79}$$

$$s_{1,\text{crit}} = \frac{\pi^2 E_1 (T_1^2 + d_1^2)}{8 L_1^2} \tag{1.80}$$

$$s_{2,\text{crit}} = \frac{\pi^2 E_2 (T_2^2 + d_2^2)}{8 L_2^2} \tag{1.81}$$

6. Malgrès son évidence, cet aspect est souvent négligé dans la littérature, mais on pourra par exemple le rencontrer dans [].

où $B_1 = B + \frac{L_1^2 - L_2^2}{4B}$ et $B_2 = B + \frac{L_2^2 - L_1^2}{4B}$ (remarquer que $B_1 + B_2 = 2B$).

Les paramètres de conception sont toujours les valeurs nominales $\bar{d}$ et $\bar{L}$, identiques pour les deux barres. Le problème de conception (1.1) devient alors :

$$\begin{cases} \min\limits_{\bar{d},\bar{L}} V \\ \text{tel que } s_1 \leq s_{\text{max}} \\ \text{et } s_1 \leq s_{1,\text{crit}} \\ \text{et } s_2 \leq s_{\text{max}} \\ \text{et } s_2 \leq s_{2,\text{crit}} \end{cases} \tag{1.82}$$

Comme en 1.2.2, mais avec des lois indépendantes pour chaque barre, l'incertitude est propagée pour le design optimal sans incertitude (c'est-à-dire $\bar{d}^* = 38, \bar{L}^* = 964$). Pour le volume de la structure on a :

```
> summary(V)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.5043  0.5618  0.5751  0.5753  0.5884  0.6534
```

Les variations sont plus faible amplitude qu'avant (comparer avec les résultats de la page 5). En effet, comme les deux barres sont indépendantes, les incertitudes se compensent (par ex. une barre plus longue associée à une barre plus courte). En ce qui concerne le taux de fiabilité du système :

```
> mean(s1 < smax  &  s1 < scrit1  &  s2 < smax  &  s2 < scrit2)
[1] 0.3441
```

Le taux de fiabilité est encore plus faible qu'avant (34% contre 43% auparavant). En effet, on a maintenant quatre contraintes à réaliser simultanément au lieu de deux.

Ainsi, quand on utilise un modèle utilisant des symétries, ce qui est le cas de (1.1), on a tendance à sur-estimer l'incertitude en sortie, ce qui est plutôt conservatif, mais aussi à sur-estimer le taux de fiabilité (ici de plus de 10%), ce qui est nettement plus inquiétant.

REMARQUE.– L'incertitude associée à un système dépend en fait du procédé de fabrication. Nous avons considéré que les deux barres avaient tous leurs paramètres indépendants. Cela peut par exemple correspondre à une situation où un fournisseur livre les barres pré-découpées, sans renseigner les conditions de fabrication. Supposons maintenant que les deux barres proviennent du même tube que l'on coupe en

deux. Alors on peut considérer les paramètres d_i, T_i, et E_i en commun ; en revanche, on a toujours deux longeurs L_1 et L_2. Supposons au contraire que les deux barres proviennent de deux tubes différents que l'on aligne l'un à côté de l'autre et que l'on découpe dans le même geste. Alors les longueurs L_1 et L_2 seront identiques, et les paramètres d_i, T_i et E_i indépendants.

1.8. Bibliographie

[AND 04] ANDRIEU L., Optimisation sous contrainte en probabilité, PhD thesis, École Nationale des Ponts et Chaussées, 2004.

[APL 06] APLEY D. W., LIU J., CHEN W., « Understanding the Effects of Model Uncertainty in Robust Design With Computer Experiments », *Journal of Mechanical Design*, vol. 128, p. 945–958, 2006.

[BAT 06] BATES R. A., KENNET R. S., STEINBERG D. M., WYNN H. P., « Achieving robust design from computer simulations », *Quality Technology & Quantitative Management*, vol. 3, p. 161–177, 2006.

[BEN 87] BENVENISTE A., METIVIER M., PRIOURET P., *Algorithmes adaptatifs et approximations stochastiques*, Masson, 1987.

[BEY 07] BEYER H. G., SENDHOFF B., « Robust optimization - A comprehensive survey », *Comput. Methods Appl. Mech. Engrg.*, vol. 196, p. 3190 – 3218, 2007.

[CAF 98] CAFLISCH R. E., « Monte Carlo and quasi-Monte Carlo methods », *Acta Numerica*, vol. 7, p. 1–49, 1998.

[CHE 05] CHEN W., JIN R., SUDJIANTO A., « Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty », *ASME Journal of Mechanical Design*, vol. 127, p. 875 – 886, 2005.

[DIW 02] DIWEKAR U., « Optimization Under Uncertainty : An Overview », *SIAG/OPT Views-and-News*, vol. 13, p. 1–8, 2002.

[DUV 08] DUVIGNEAU R., MARTINELLI M., CHANDRA-SHEKARAPPA P., « *Optimisation multidisciplinaire en mécanique : couplages, multiniveaux et robustesse* », Chapitre Estimation d'incertitude en aérodynamique, Mécanique et Ingénierie des Matériaux, Hermès, 2008.

[EFR 93] EFRON B., TIBSHIRANI R. J., *An introduction to the bootstrap*, Chapman & Hall, 1993.

[EGO 02] EGOROV I. N., KRETININ G. V., LESHCHENKO I. A., « How to execute robust design optimization », *AIAA*, 2002.

[FON 07] FONSECA J. R., FRISWELL M. I., LEES A. W., « Efficient robust design via Monte Carlo sample reweighting », *Int. J. Numer. Meth. Engng.*, 2007.

[HAF 85] HAFTKA R. T., « Simultaneous Analysis and Design », *AIAA J.*, vol. 23, n°7, p. 1099–1103, 1985.

[HUA 07] HUANG B., DU X., « Analytical robustness assessment for robust design », *Structural and multidisciplinary optimization*, vol. 34, p. 123–137, 2007.

[JIN 03] JIN R., DU X., CHEN W., « The Use of Metamodeling Techniques for Optimization Under Uncertainty », *Journal of Structural & Multidisciplinary Optimization*, vol. 25, p. 99–116, 2003.

[JIN 04] JIN R., CHEN W., SUDJIANTO A., « Analytical metamodel-based global sensitivity analysis and uncertainty propagation for robust design », *SAE Congress*, 2004.

[KEA 05] KEANE A. J., NAIR P. B., « Design in the presence of uncertainty », *Computational Approaches for Aerospace Design*, Chapitre 8, Wiley, 2005.

[KEL 99] KELLEY T. C., « Optimization of noisy functions », *Iterative methods for optimization*, Chapitre 6, 7, 8, SIAM, 1999.

[KIM 06] KIM N. H., WANG H., QUEIPO N. V., « Adaptive Reduction of Design Variables Using Global Sensitivity in Reliability-Based Optimization », *International Journal of Reliability and Safety*, vol. 1, p. 102 – 119, 2006.

[LEH 04] LEHMAN J. S., SANTNER T. J., NOTZ W. I., « Designing computer experiments to determine robust control variables », *Statistica Sinica*, vol. 14, p. 571–590, 2004.

[LUC 07] LUCOR D., ENAUX C., JOURDREN H., SAGAUT P., « Stochastic design optimization : Application to reacting flows », *Computer methods in applied mechanics and engineering*, vol. 196, p. 5047–5062, 2007.

[MAR 00] MARCZYK J., « Stochastic Multidisciplinary Improvement : Beyond Optimization », *8th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization, AIAA-2000-4929*, 2000.

[MCC 89] MCCULLAGH P., NELDER J. A., *Generalized Linear Models*, Chapman & Hall, 1989.

[MCK 79] MCKAY M. D., BECKMAN R. J., CONOVER W. J., « A comparison of three methods for selecting values of input variables in the analysis of output from a computer code », *Technometrics*, vol. 21, p. 239–245, 1979.

[R D 08] R DEVELOPMENT CORE TEAM, R : A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2008, ISBN 3-900051-07-0, URL `http://www.R-project.org`.

[ROB 51] ROBBINS H., MONRO S., « A Stochastic Approximation Method », *Annals of mathematical statistics*, vol. 22, p. 400–407, 1951.

[SAN 03] SANTNER T. J., WILLIAMS B. J., NOTZ W. I., *The design and analysis of computer experiments*, Springer, 2003.

[SAP 90] SAPORTA G., *Probabilités, analyse des données et statistiques*, Technip, 1990.

[SHO 91] SHOEMAKER A. C., TSUI K. L., WU C. F. J., « Economical experimentation methods for robust parameter design », *Technometrics*, vol. 33, p. 415–427, 1991.

[SHO 93] SHOEMAKER A. C., TSUI K., « Response model analysis for robust design », *Communications in Statistics – Simulations*, vol. 22, p. 1037–1064, 1993.

[SPA 03] SPALL J. C., *Introduction to stochastic search and optimization*, Wiley, 2003.

[TAG 04] TAGUCHI G., CHOWDHURY S., WU Y., *Taguchi's Quality Engineering Handbook*, Wiley, 2004.

[VIN 90] VINING G. G., MYERS R. H., « Combining Taguchi and Response Surface Philosophies : A Dual Response Approach », *Journal of Quality Technology*, vol. 22, p. 38–45, 1990.

[WEL 90] WELCH W. J., YU T. K., KANG S., SACKS J., « Computer experiments for quality control by parameter design », *Journal of Quality Technology*, vol. 22, p. 15–22, 1990.

[WIL 41] WILKS S. S., « Determination of Sample Sizes for Setting Tolerance Limits », *Annals of mathematical statistics*, vol. 12, n°1, p. 91–96, 1941.

[YAN 91] YANG W.-N., L. B. L. N. B., « Using common random numbers and control variates in multiple-comparison procedures », *Operations Research*, vol. 34, p. 583–591, 1991.