

# Optimisation des paramètres par simulation

---

## 4.1 Introduction

L'optimisation des paramètres par simulation est une solution possible pour le dimensionnement du SYSMFS en minimisant une fonction coût (Équation (85)). En principe, il s'agit d'une comparaison des solutions obtenues par la simulation qui partagent le même schéma conceptuel. En étudiant l'effet de certains paramètres sur la réponse obtenue, les meilleures solutions seront déterminées.

Les algorithmes d'optimisation offrent en général l'avantage de l'économie de temps en réduisant le nombre d'essais (simulations) nécessaires en vue d'obtenir une valeur optimale de la réponse. Ils conviennent normalement à un large éventail d'applications et ils utilisent des techniques diverses pour la recherche d'optimum : le gradient conjugué, la recherche stochastique, les stratégies d'évolution, etc.

Pendant une optimisation, le calcul de la fonction coût demande généralement beaucoup plus de temps de calcul que le temps requis pour déterminer les valeurs suivantes dans une itération d'optimisation ; ce principe est valable pour toutes les procédures d'optimisation dans les problèmes de simulation des bâtiments. En premier lieu, le gain de temps total d'une optimisation se fait donc au niveau de la simulation ; un nombre assez important des solutions est possible afin d'optimiser le déroulement de la simulation (type de solveur, hypothèse de modèle, etc.). Ensuite, c'est à l'algorithme d'optimisation de réduire efficacement le nombre des simulations en réalisant un avancement maîtrisé vers la valeur optimale de la fonction coût ; ce dernier devrait alors réduire le nombre des simulations inutiles, sans pour autant s'arrêter sur des optimums locaux.

La performance d'un algorithme d'optimisation prend en compte les valeurs optimales trouvées, le nombre des simulations effectuées par chaque méthode ainsi que toutes autres informations supplémentaires données par l'algorithme suite à l'optimisation.

Dans la suite de ce travail, deux algorithmes d'optimisation sont testées pour leurs performances : un algorithme hybride GPS-PSO et un algorithme basé sur les plans d'expériences. Le premier algorithme est déjà implémenté dans un code d'optimisation générique [22] tandis que pour le deuxième nous avons créé les procédures et l'implémentation. Chaque algorithme est testé avec une fonction de test et avec le modèle SYSMFS ; une comparaison de ces deux méthodes est également présentée dans ce chapitre.

### 4.1.1 Types des variables à optimiser

Les paramètres à optimiser peuvent être soit de type discret, soit de type continu (§1.2.2). Les panneaux solaires sont constitués de plusieurs modules interconnectés entre eux, chaque module ayant une surface prédéfinie par le fabricant. Nous considérons cependant que la surface totale des panneaux peut varier librement dans la gamme  $[5, 50] \text{ m}^2$  ; pour respecter la consigne précédente, la valeur obtenue de surface doit être alors arrondie à  $0,5 \text{ m}^2$  près.

Les ballons de stockage thermique sont, quant à eux, fabriqués principalement dans des séries standardisées de volumes ; il est possible cependant de commander des ballons sur mesure. En utilisant l'Équation (32), le volume du ballon peut être considéré comme une variable continue avec des valeurs prises dans la fourchette  $[25, 3000] \text{ L}$ .

### 4.1.2 Valeurs initiales des paramètres à optimiser

Certains algorithmes d'optimisation (ex. gradient conjugué) sont plus influencés par les valeurs initiales des paramètres à optimiser que d'autres algorithmes. Suivant le domaine d'application et l'algorithme employé, certaines valeurs initiales peuvent rendre la convergence plus rapide vers un optimum potentiel, qu'il soit local ou global. En revanche, d'autres valeurs initiales peuvent retarder la convergence en éloignant l'algorithme de la zone optimale pour certain temps avant que la convergence commence.

Une connaissance préalable du comportement de la fonction coût pour une application donnée pourrait éventuellement venir en aide dans le choix des valeurs initiales des paramètres à optimiser. Dans ce but, l'étude paramétrique menée précédemment avec la surface de réponse (Figure 3.20) est indispensable.

### 4.1.3 Méthodes de test pour les algorithmes d'optimisation étudiés

L'objectif principal de notre étude est de tester le concept d'optimisation par plans d'expériences pour le dimensionnement optimal des systèmes comme les SYSMFS. L'application de la méthode d'optimisation par plans d'expériences est une nouveauté car, selon notre connaissance, elle n'a pas été utilisée auparavant pour l'optimisation du dimensionnement par la simulation des systèmes solaires comme les SSC ou les SYSMFS.

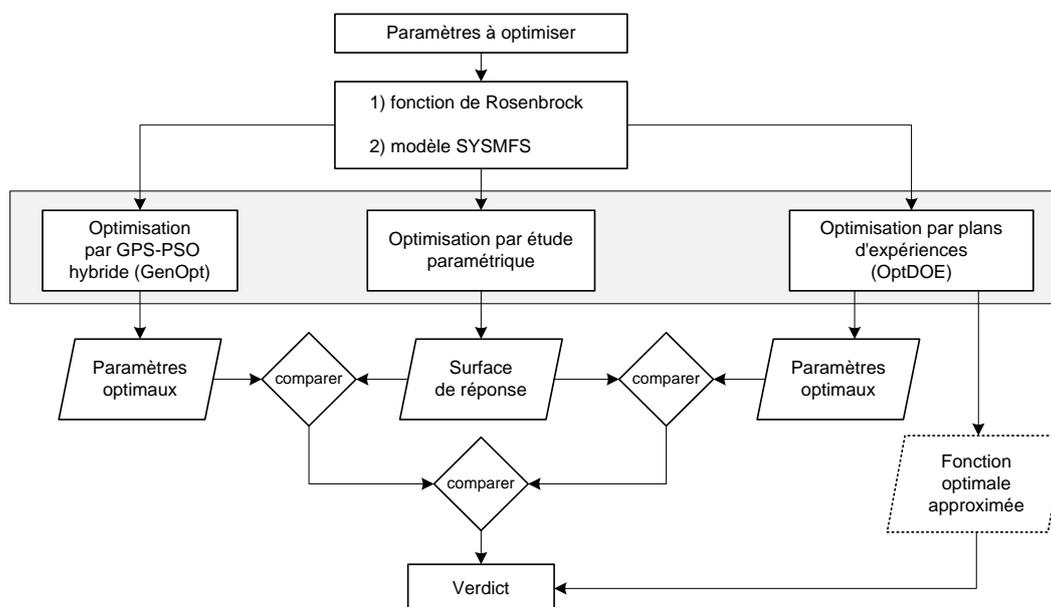


Figure 4.1 : Plan à suivre pour la comparaison de la performance des méthodologies d'optimisation

La Figure 4.1 donne une représentation du plan à suivre afin de comparer la performance de trois méthodes d'optimisation : étude paramétrique, algorithme hybride GPS-PSO et optimisation par plans d'expériences (OptDOE<sup>10</sup>).

Tableau 4.1 : Niveaux des difficultés des tests pour les algorithmes d'optimisation

Application	Niveau de difficulté	Caractéristiques
Fonction de Rosenbrock	très difficile	forme complexe, minimum global très bien assimilé
Modèle de SYSMFS	difficile	forme vallée, légère inclinaison (Figure 3.20)

<sup>10</sup> Acronyme pour l'optimisation par plans d'expériences ou pour l'optimiseur basé sur la même méthode

Le benchmark de ces méthodes est réalisé pour deux cas : une fonction de Rosenbrock et le modèle SYSMFS. De cette manière, nous validons la méthode OptDOE et nous comparons sa performance avec d'autres méthodes d'optimisation (Tableau 4.1).

#### 4.1.3.1 Fonction de Rosenbrock

Introduite par Rosenbrock en 1960, la fonction de Rosenbrock est une fonction non convexe de deux variables qui est souvent utilisée dans le domaine des problèmes d'optimisation pour tester la performance des algorithmes. Elle est définie par (Figure 4.2) :

$$f(x, y) = (1 - x^2) + 100(y - x^2)^2 \quad (87)$$

Le minimum global de la fonction de Rosenbrock se trouve à (1,1), où la valeur de la fonction est nulle. Ce minimum est situé à l'intérieur d'une vallée de forme parabolique longue et étroite. Les algorithmes d'optimisation peuvent facilement trouver la vallée, mais la convergence vers le minimum global est d'une difficulté considérable. Sur le contour de cette fonction (Figure 4.2, b), nous présentons les points initiaux que nous avons choisis et à partir desquels nous testons les algorithmes d'optimisation.

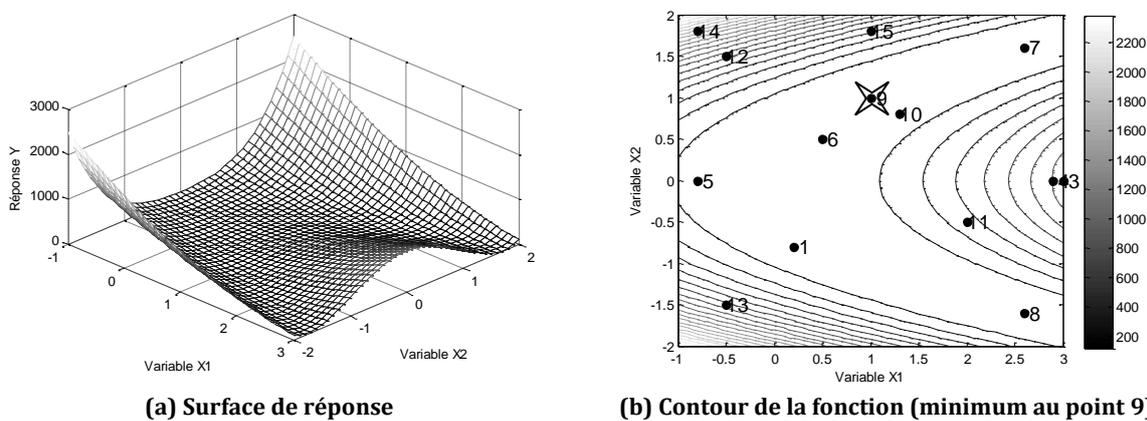


Figure 4.2 : Fonction de Rosenbrock

#### 4.1.3.2 Modèle de SYSMFS

La forme de la réponse du modèle SYSMFS a été présentée précédemment dans le Chapitre 3. Cette surface (Figure 3.20) constitue une base de comparaison pour les deux algorithmes d'optimisation qui seront appliqués ultérieurement au même modèle.

## 4.2 Optimisation par l'algorithme hybride GPS-PSO

Dans le 0, une brève description a été faite des deux algorithmes individuels qui constituent l'optimisation hybride GPS-PSO. Le logiciel d'optimisation générique GenOpt intègre plusieurs algorithmes d'optimisation différents. Pour les problèmes de simulations propres au bâtiment, l'algorithme hybride a été déjà utilisé dans la littérature avec un logiciel de simulation des bâtiments [80].

Le couplage entre notre modèle SYSMFS développé avec TRNSYS et le logiciel d'optimisation est relativement direct. La Figure 4.3 illustre le schéma de couplage entre les deux logiciels. L'échange des données entre les deux logiciels se fait avec des fichiers au format texte simple. Les résultats obtenus par cette méthode sont comparés avec ceux issus de l'étude paramétrique afin d'en savoir plus sur la performance de cette méthode d'optimisation.

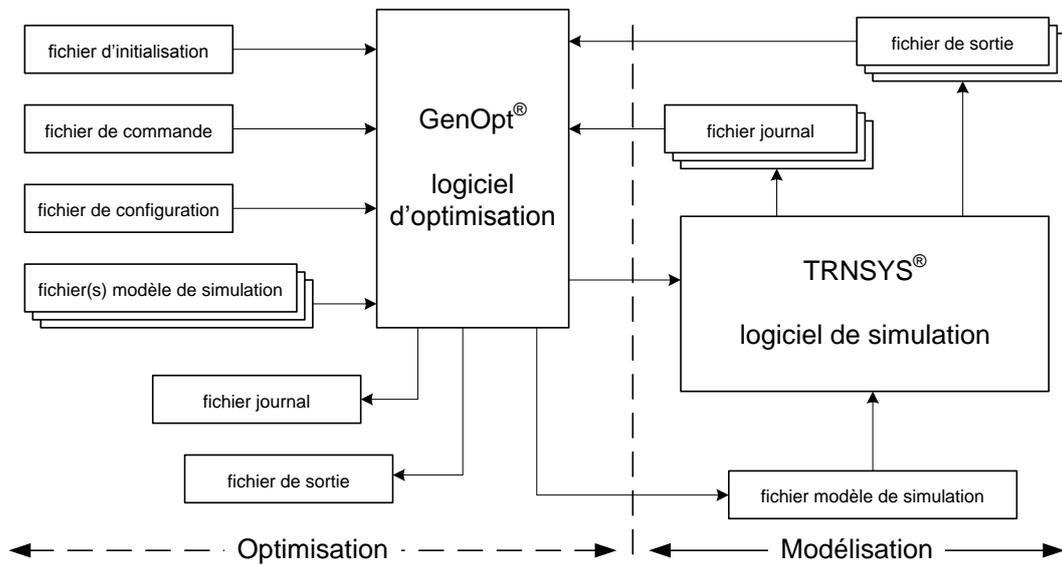


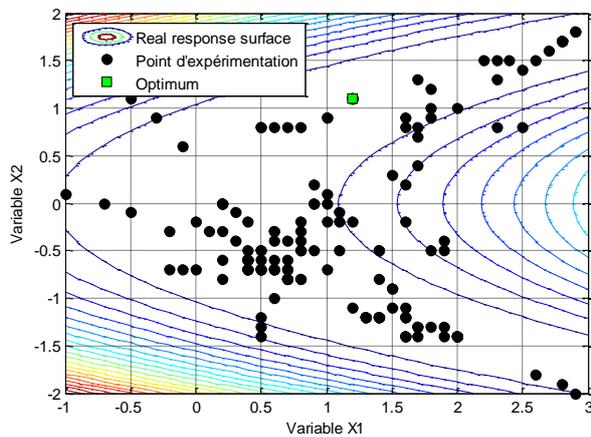
Figure 4.3 : Schéma technologique de la liaison GenOpt-TRNSYS

#### 4.2.1 Optimisation par l'algorithme hybride GPS-PSO de la fonction de Rosenbrock

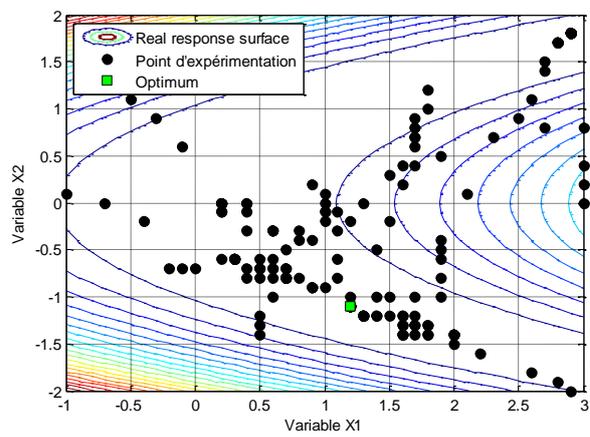
Le Tableau 4.2, présente les cas d'optimisation de la fonction de Rosenbrock par l'algorithme hybride GPS-PSO. D'après ces essais, l'algorithme hybride GPS-PSO n'a pas réussi à trouver le minimum global à l'exception du cas numéro 9 où le point de départ est le minimum lui même. Concernant le nombre de générations, nous l'avons fixé à 10, ce qui a fait que le nombre de simulations est toujours égal à 163. Les Figures .4 à .10, illustrent les simulations effectuées par l'algorithme hybride GPS-PSO.

Tableau 4.2 : Optimisation par l'algorithme hybride GPS-PSO de la fonction de Rosenbrock

Test	Variable X1	Variable X2	Remarque sur le point de départ	Variable X1 optimale	Variable X2 optimale	Réponse Y optimale	Simulations
1	0,2	-0,8	-	1,2	1,1	0,05	163
2	2	2,5	hors limites			Pas de réponse	
3	3	0	hors limites	1,2	-1,1	0,05	163
4	2,9	0	voisinage de frontière	1,2	-1,1	0,05	163
5	-0,8	0	voisinage de frontière	1,2	-1,1	0,05	163
6	0,5	0,5	-	1,2	-1,1	0,05	163
7	2,6	1,6	-	1,2	-1,1	0,05	163
8	2,6	-1,6	-	1	-1	0	163
9	1	1	point minimum	1	1	0	163
10	1,3	0,8	proche de minimum	1,2	-1,1	0,05	163
11	2	-0,5	proche de minimum	1,2	-1,1	0,05	163
12	-0,5	1,5	-	1,2	-1,1	0,05	163
13	-0,5	-1,5	-	1,2	-1,1	0,05	163
14	-0,8	1,8	-	1,2	-1,1	0,05	163
15	1	1,8	-	1,2	-1,1	0,05	163

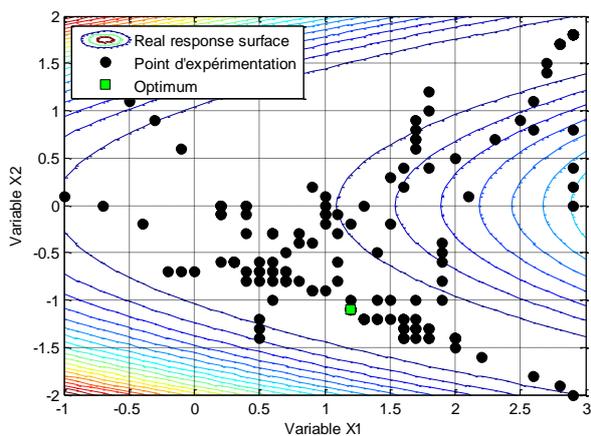


cas 1

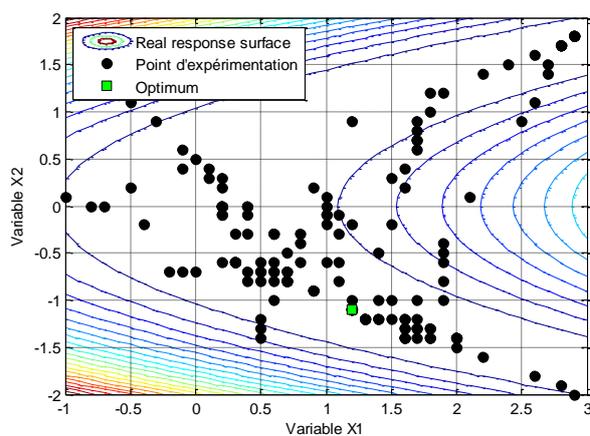


cas 3

Figure 4.4 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 1 et 3)

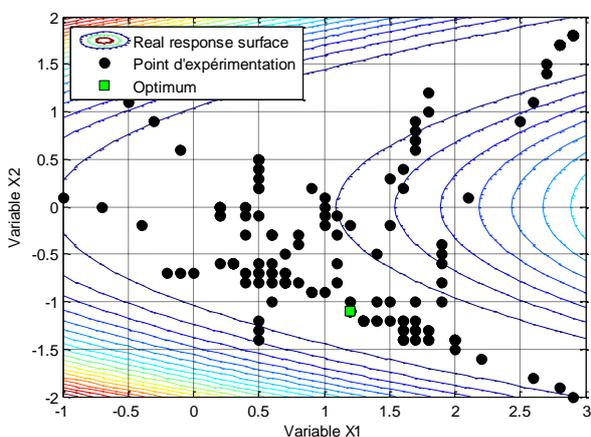


cas 4

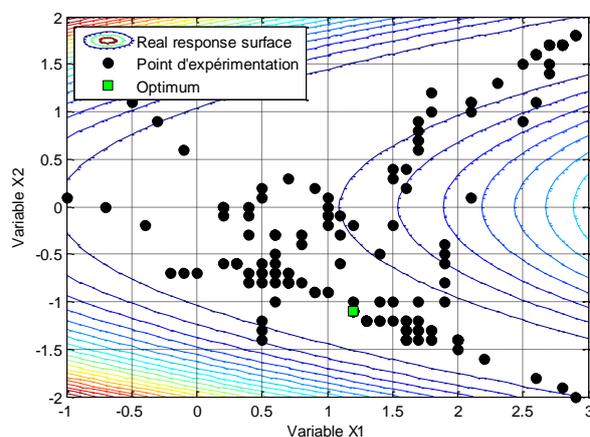


cas 5

Figure 4.5 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 4 et 5)

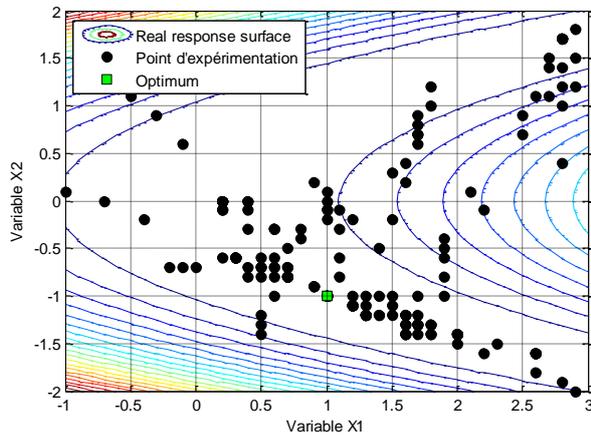


cas 6

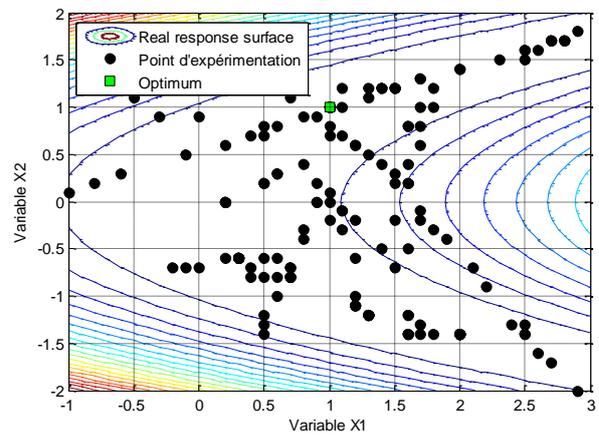


cas 7

Figure 4.6 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 6 et 7)

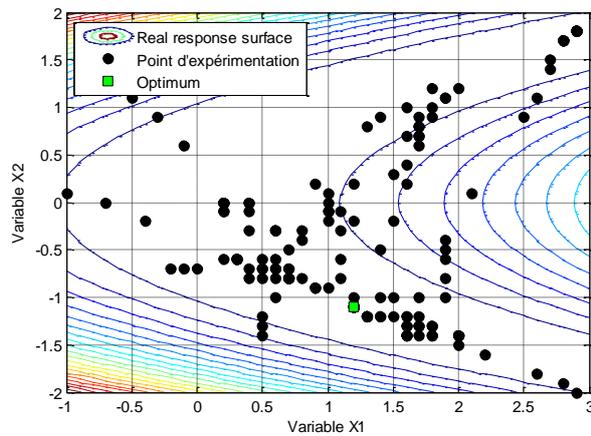


cas 8

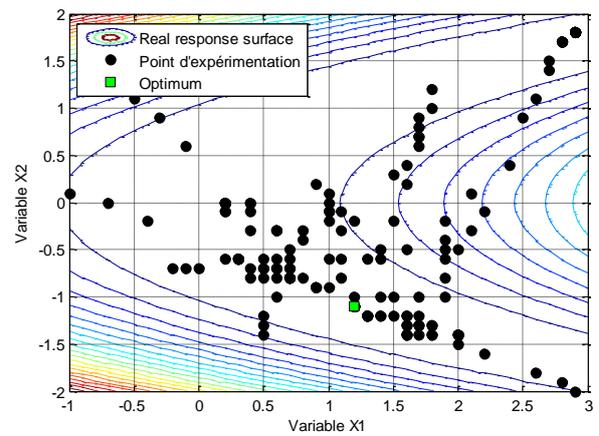


cas 9

Figure 4.7 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 8 et 9)

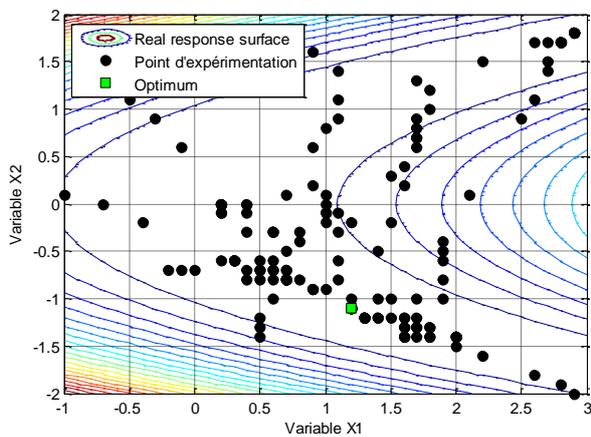


cas 10

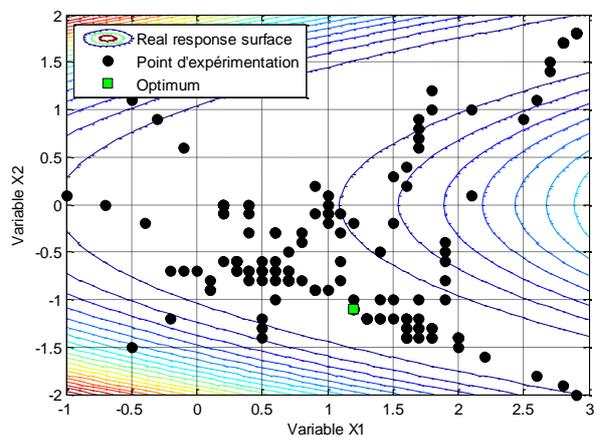


cas 11

Figure 4.8 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 10 et 11)



cas 12



cas 13

Figure 4.9 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 12 et 13)

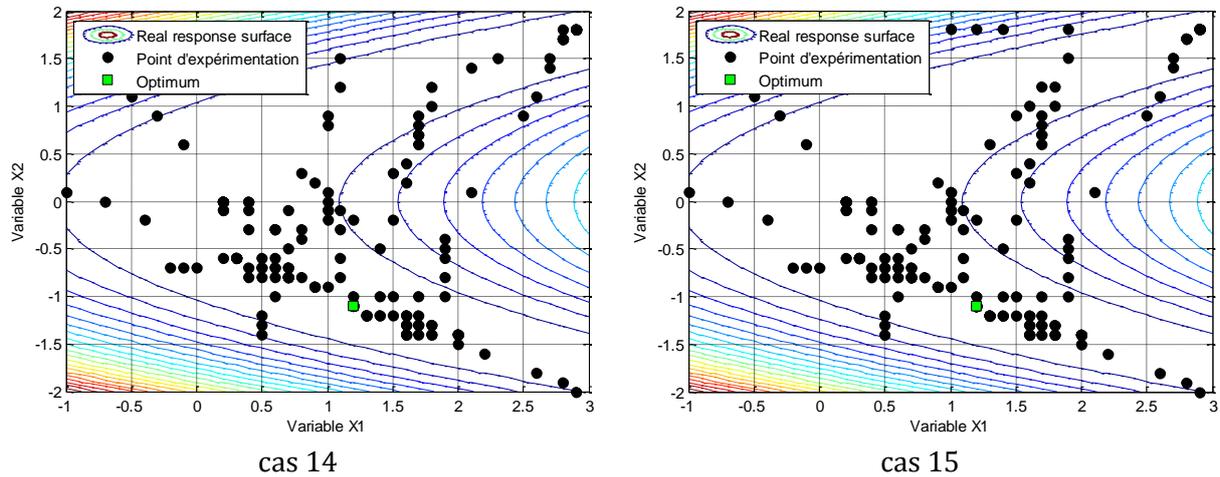


Figure 4.10 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO appliqué à la fonction de Rosenbrock (cas 14 et 15)

#### 4.2.2 Optimisation par l'algorithme hybride GPS-PSO du modèle SYSMFS

D'une manière similaire au test précédent, l'algorithme hybride GPS-PSO est appliqué au modèle SYSMFS pour vérifier l'effet des points initiaux sur les résultats obtenus. Dans ce contexte, nous avons choisi quatre points dans le domaine du problème distribués autour de la zone de la vallée et dans sa direction (Figure 4.11).

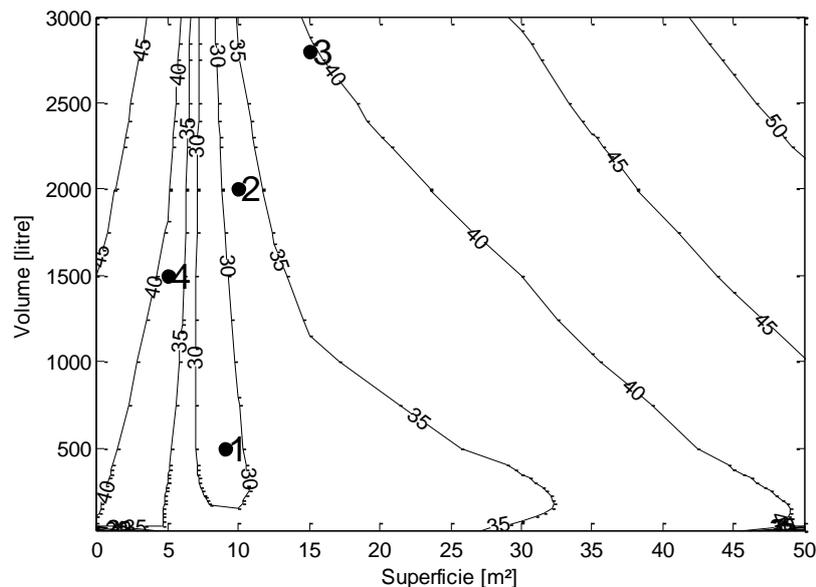


Figure 4.11 : Points initiaux pour l'optimisation du SYSMFS

D'après les cas GEN1, GEN2, GEN3 et GEN4 montrés dans le Tableau 4.3, l'algorithme hybride GPS-PSO ne semble pas largement affecté par les valeurs initiales des paramètres d'entrées. Dans le test GEN2 où le point de départ est relativement proche du minimum global, l'optimum obtenu a été le plus éloigné de celui que l'on attendait. La nature stochastique de l'algorithme peut expliquer en partie cette insensibilité à la position du point de départ. Les optimums obtenus sont néanmoins distribués dans une zone relativement limitée comme le montre le Tableau 4.3.

Lors de l'optimisation du SYSMFS par l'algorithme hybride GPS-PSO, il a été remarqué que l'optimisation a été effectuée principalement par la phase PSO ; la phase GPS postérieure à PSO n'améliore apparemment pas les résultats de la phase PSO. Vu que c'est une version locale de GPS, GPS reçoit uniquement les valeurs des variables continues pour qu'elles soient optimisées

dans un sous-domaine de l'espace du problème. Pendant nos tests, le nombre de simulations reste néanmoins 59% plus petit que celui de l'étude paramétrique.

Tableau 4.3 : Optimisation par l'algorithme hybride GPS-PSO du SYSMFS

Algorithme	Test	Superficie initiale	Volume initiale	Superficie optimale	Volume optimale	Coût optimal	Écart du coût de PARI	Simulations	Type d'optimum
		(m <sup>2</sup> )	(litre)	(m <sup>2</sup> )	(litre)	(k€)	(%)		
<b>Paramétrique</b>	PARI	-	-	7,7	1500	27	-	401	
<b>Hybride GPS-PSO</b>	GEN1	9	500	7,5	1425	27,2	0,8	163	-
	GEN2	10	2000	7,5	2150	27,3	1,3	163	-
	GEN3	15	2800	7,5	1900	27,1	0,5	163	-
	GEN4	5	1500	8	1425	27,3	1	163	-

Les Figures 4.12, 4.13, .14 et .15, montrent la distribution des points expérimentaux où l'algorithme hybride GPS-PSO a effectué ses simulations dans les quatre cas (GEN1, GEN2, GEN3 et GEN4).

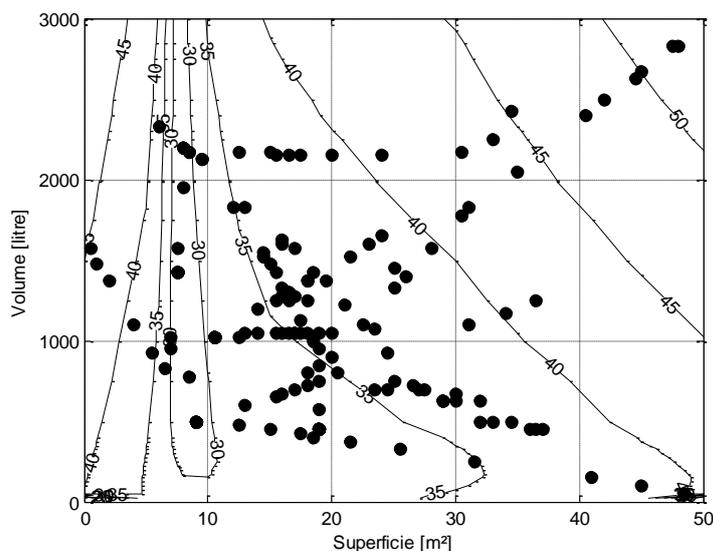


Figure 4.12 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO dans le cas GEN1

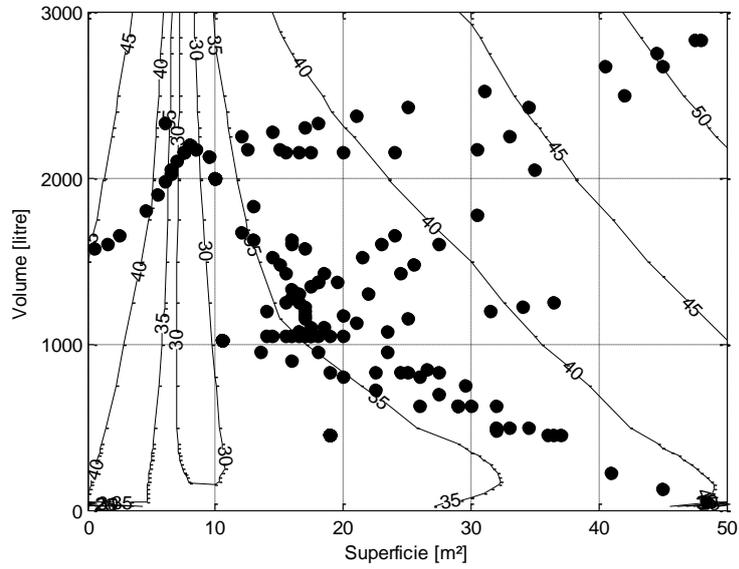


Figure 4.13 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO dans le cas GEN2

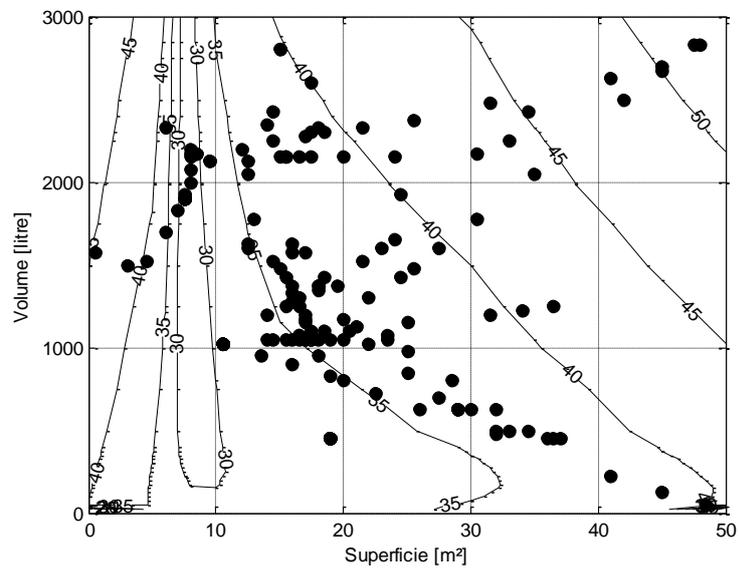


Figure 4.14 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO dans le cas GEN3

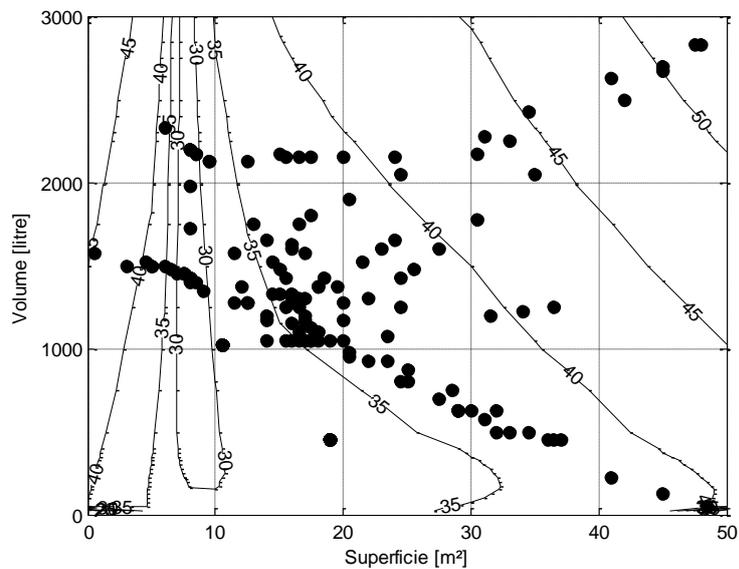


Figure 4.15 : Points d'expérimentations effectués par l'algorithme hybride GPS-PSO dans le cas GEN4

### 4.3 Optimisation par plans d'expériences

La solution approximée de la fonction coût (§1.2.3.2) associée aux simulations des bâtiments rendent inapproprié l'usage des techniques d'optimisation de type gradient pur sans mesure adéquate pour détecter les optimums locaux où elles peuvent se piéger. Nous proposons donc une méthodologie d'optimisation basée sur les plans d'expériences, ce qui permet de donner des valeurs optimales et une estimation approximée de la fonction coût dans le voisinage de la solution optimale [81].

L'idée de cette approche d'optimisation se résume à planifier des expérimentations qui ont pour but d'estimer les paramètres à optimiser et de déterminer un modèle qui se rapproche de la fonction coût dans la zone optimale [82], [83]. Ce modèle est ensuite utilisé pour une reproduction du comportement du système original étudié dans l'espace de validité du modèle (un sous-domaine de l'espace du problème).

#### 4.3.1 Méthode d'optimisation par plans d'expériences

La procédure d'optimisation par les plans d'expériences consiste en deux étapes majeures : l'amélioration de la solution actuelle et la détermination de l'optimum [84]. Nous développons au cours de ce travail un optimiseur numérique basé sur les plans d'expériences que nous appelons OptDOE.

Par définition, un plan d'expériences est effectué localement sur un sous-domaine de l'espace du problème. La procédure ne comporte pas en elle-même une méthode de déplacement dans cet espace à la recherche d'optimum ; l'amélioration de la solution actuelle nécessite donc une méthode d'avancement.

Une fois couplé avec une méthode d'avancement appropriée, le plan pourrait ainsi se déplacer dans l'espace du problème suivant la direction de minimisation de la fonction coût. La direction de minimisation est déterminée après chaque plan, seulement et seulement si, un modèle de premier ordre peut être correctement calé dans le domaine d'étude du plan.

Les modèles obtenus de chaque plan sont considérés comme locaux. Afin de faciliter la compréhension et la manipulation des modèles locaux, les paramètres à optimiser peuvent être réduits, eux aussi, à une échelle entre  $[-1,1]$  en utilisant l'Équation (25).

L'estimation d'un modèle local requiert normalement un nombre d'expérimentations ; ce dernier doit être au moins égal au nombre des paramètres pour que le problème soit déterminé. Les expérimentations physiques sont, quant à elles, souvent susceptibles d'erreurs arbitraires. Des essais supplémentaires sont alors envisagés afin de réduire l'effet de ces erreurs. Ce qui n'est pas le cas des simulations numériques qui ont la faculté d'être reproductible à l'identique.

Afin de mieux illustrer la démarche de la méthode d'optimisation par les plans d'expériences, un problème exemplaire d'optimisation d'une fonction coût à deux paramètres d'entrée est présenté dans la Figure 4.16.

Dans un premier temps, le point initial de départ et la taille du plan local ( $d_1, d_2$ ) sont choisis en fonction des propriétés physiques du système étudié et du domaine d'application. Les valeurs initiales choisies doivent être dans l'espace du problème pour pouvoir constituer le premier plan factoriel. Ensuite, un plan factoriel centré à cinq points est fait, on tente alors de faire passer un modèle (ex. Équation (88)) de premier ordre par les points du plan factoriel afin de représenter la surface de réponse locale.

Ensuite, un test de manque d'ajustement<sup>11</sup> (test de Fisher) doit être effectué afin que le modèle soit bien ajusté aux données à partir desquelles il a été généré [81]. Dans le cas où le modèle est juste, on peut déduire la direction de minimisation de la fonction coût comme étant la direction du gradient du modèle calé de premier ordre. Sinon, un autre modèle d'ordre plus élevé est alors recherché pour représenter cette partie de la surface de réponse.

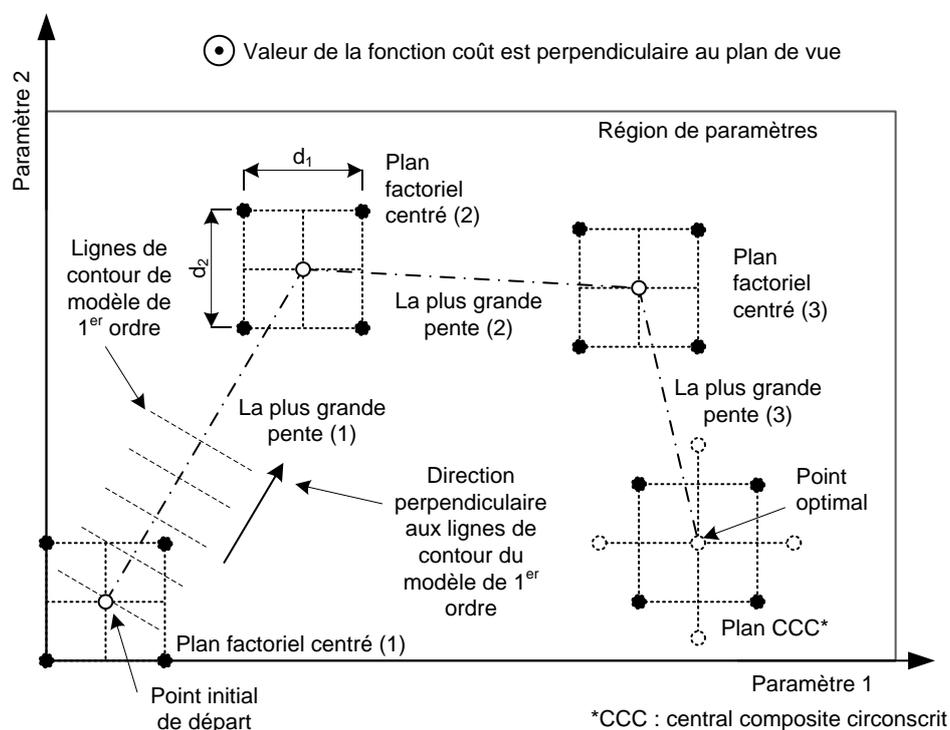


Figure 4.16 : Optimisation par les plans d'expériences d'un problème exemplaire à deux paramètres

Si le modèle de premier ordre est juste, un avancement de recherche d'optimum peut ensuite être initié dans la direction de minimisation et avec un pas donné. On vérifie que la valeur fonction coût diminue après chaque pas en comparant sa valeur avec celle de l'itération précédente. Au moment où la valeur de la fonction de coût ne s'améliore plus, on attend un changement de direction de minimisation ou une discontinuité dans cette fonction ; un modèle de premier ordre est tenté en vue de la détermination d'une nouvelle direction s'il y en a.

Cette modalité d'avancement repose, pour partie, sur le suivi de la direction du gradient de la fonction optimisée, mais les modèles obtenus par les plans d'expériences aident à comprendre le type de la surface de réponse et à trouver une nouvelle direction à partir d'analyse du modèle de premier ordre. Cette mesure aide à surmonter le problème hérité d'une méthode de recherche basée uniquement sur le gradient et qui peut piéger l'avancement de l'algorithme d'optimisation dans des optimums locaux.

Si la tentative d'ajustement de modèle de premier ordre échoue, un modèle de deuxième ordre peut être calé. Le test de manque d'ajustement proposé (test de Fisher) demande que les moyennes des populations analysées ne soient pas très différentes pour que l'analyse de variance soit utilisable [85]. Dans le cas contraire, il se peut que le modèle de deuxième ordre ne passe pas le test statistique du manque d'ajustement ; nous proposons alors soit de changer la taille des valeurs de  $(d_1, d_2)$  qui délimitent la zone du plan local de base (Figure 4.16) ou bien de reprendre l'optimisation avec un nouveau point de départ. Avec la première proposition, une

<sup>11</sup> Synonyme de « lack-of-fit » en anglais

version adaptative de la méthodologie peut être investiguée ultérieurement avec des contraintes sur les  $(d_1, d_2)$ .

Finalement, si le modèle de deuxième ordre obtenu est juste, il peut alors représenter la fonction coût approximée dans la zone optimale. Une analyse postérieure de ce modèle donne le type du point d'optimum (maximum, minimum ou point col<sup>12</sup>) et la forme de la surface dans cette zone.

#### 4.3.1.1 Modèles de régression

Dans un procédé des plans d'expériences (DOE<sup>13</sup>), les modèles locaux sont calés avec des fonctions polynomiales. Un modèle local doit représenter, avec une certaine précision, une portion de la surface de réponse empirique d'origine (ex. Figure 3.20). La fonction coût à optimiser est donnée ainsi sous la forme de polynômes qui approximent par une série de Taylor la fonction coût réelle compliquée ; seuls les termes au premier et au deuxième ordre sont retenus [86].

Suivant la forme choisie du polynôme, les fonctions coût obtenues sont linaires :

$$y = \beta_0 + \beta_1 P_1 + \beta_2 P_2 \quad (88)$$

ou linaires avec interactions :

$$y = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_{12} P_1 P_2 \quad (89)$$

ou quadratiques pures

$$y = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_{11} P_1^2 + \beta_{22} P_2^2 \quad (90)$$

ou quadratiques avec interactions :

$$y = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_{11} P_1^2 + \beta_{22} P_2^2 + \beta_{12} P_1 P_2 \quad (91)$$

où  $y$  est la réponse du modèle,  $P_1, P_2$  sont les deux paramètres du modèle de régression (paramètres à optimiser),  $\beta_1, \beta_2$  sont les coefficients linaires du polynôme du modèle de régression,  $\beta_{11}, \beta_{22}$  sont les coefficients quadratiques du polynôme du modèle de régression,  $\beta_{12}$  est le coefficient d'interaction du polynôme du modèle de régression.

Pour cette étude, les paramètres à optimiser  $P_1, P_2$  représentent la surface du panneau solaire et le volume du stockage thermique du ballon solaire principal. Les coefficients  $\beta_1, \beta_2, \beta_{11}, \beta_{22}, \beta_{12}$  sont à identifier pour minimiser les erreurs entre les modèles polynômiaux et les résultats de la simulation [86].

Pour chacun des modèles précédents (Équations (88), (89), (90) et (91)), plusieurs plans standards sont disponibles. L'optimiseur peut tenter de caler plusieurs modèles linaires puis des modèles quadratiques au même point. Par souci d'économie de temps de calcul, il sera profitable de choisir des plans qui sont augmentables. Un plan factoriel centré peut être augmenté, par exemple, au plan central composite circonscrit (CCC) en ajoutant aux points factoriels des points (c. à d. des simulations) supplémentaires comme il a été déjà démontré dans la Figure 1.17 au 0.

<sup>12</sup> Point col ou point selle désigne un point auquel les dérivées d'une fonction s'annulent, sans pour autant que le point soit un minimum ou un maximum local (« saddle point » en anglais).

<sup>13</sup> Acronyme pour la méthode des plans d'expériences ou « Design of Experiments » en anglais.

#### 4.3.1.2 Identification des coefficients

Les modèles présentés précédemment dans les Équations (88), (89), (90) et (91), peuvent être exprimés sous la forme matricielle :

$$\mathbf{X}\beta = \mathbf{y} \quad (92)$$

où  $\mathbf{X}$  est la matrice d'information,  $\beta$  est le vecteur des coefficients,  $\mathbf{y}$  est le vecteur des sorties.

Le vecteur des sorties s'écrit :

$$\mathbf{y} = [y_1 \quad \cdots \quad y_m]^T \quad (93)$$

où  $m$  est le nombre d'expérimentations.

Le vecteur des coefficients s'écrit :

$$\beta = [\beta_1 \quad \cdots \quad \beta_n]^T \quad (94)$$

La matrice d'information est :

$$\mathbf{X} = \begin{bmatrix} f_1(u_1) & \cdots & f_n(u_1) \\ \vdots & \ddots & \vdots \\ f_1(u_m) & \cdots & f_n(u_m) \end{bmatrix} \quad (95)$$

où  $n$  est le nombre des paramètres d'entrée.

Dans le cas où le système d'équations (92) est surdéterminer ( $n > m$ ), il n'a pas de solution. Pour que le système ait une solution, on ajoute un vecteur d'erreur à l'Équation (92) :

$$\mathbf{X}\beta + \mathbf{e} = \mathbf{y} \quad (96)$$

où  $\mathbf{e}$  est le vecteur d'erreur.

Dans ce cas, le problème des moindres carrés revient à trouver  $\mathbf{b} = \beta$  qui minimise la somme des carrés des erreurs :

$$E(\beta) = \sum_{i=1}^m \left( y_i - \sum_{j=1}^n \beta_j f_j(x_i) \right)^2 \quad (97)$$

$$E(\beta) = (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) \quad (98)$$

$$E(\beta) = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) \quad (99)$$

$$E(\beta) = \mathbf{e}^T \mathbf{e} \quad (100)$$

où  $\hat{\mathbf{y}}$  est la sortie approximée donnée comme :

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} \quad (101)$$

Si le modèle est linéaire et  $\mathbf{X}^T \mathbf{X}$  est non singulière, alors  $\mathbf{b}$  est unique et donné par :

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (102)$$

### 4.3.1.3 Test de manque d'ajustement

Souvent quand l'ajustement des modèles aux données se fait par la méthode des moindres carrés, le test de manque d'ajustement aide à vérifier que le modèle proposé est bien calé aux données.

L'hypothèse nulle présume que les écarts type<sup>14</sup> du modèle et celui de ses données sont égaux :

$$H_0: \sigma_1 = \sigma_2 \quad (103)$$

où  $\sigma_1$  est l'écart type des résultats du modèle de régression,  $\sigma_2$  est l'écart type des données.

Autrement dit, on cherche à prouver qu'une liaison existe entre  $y$  et  $P_1, P_2$  (Équations (88), (89), (90) et (91)). L'hypothèse nulle (Équation (103)) est rejetée si :

$$F > F_{\alpha, k-1, m} \quad (104)$$

où  $F$  est un ratio donné par Équation (105) (Loi de Fisher<sup>15</sup>),  $F_{\alpha, k-1, m}$  est la valeur critique de  $F$ ,  $\alpha$  est le niveau de signification (ex. 5%),  $k$  est le nombre de paramètres (ex. 2),  $m$  est le nombre d'expérimentations (ex. 5 pour un plan factoriel centré).

Le ratio  $F$  peut être donné par [85]:

$$F = \frac{\frac{SS_{Tr}}{k-1}}{\frac{SS_E}{m-k}} \quad (105)$$

où  $SS_{Tr}$  est la somme des carrés de manque d'ajustement,  $SS_E$  est la somme des carrés d'erreur pure,  $k$  est le nombre de paramètres,  $m$  est le nombre d'expérimentations (ex. 5 pour un plan factoriel centré).

La somme des carrés de manque d'ajustement,  $SS_{Tr}$ , est donnée par [81] :

$$SS_{Tr} = \sum_{i=1}^k n_i (\bar{y}_i - \hat{y}_i)^2 \quad (106)$$

où  $k$  est le nombre de paramètres,  $n_i$  est le nombre d'expérimentations au niveau  $i$  du paramètre  $P_i$ ,  $\bar{y}_i$  est la moyenne du  $i^{\text{ème}}$  set,  $\hat{y}_i$  est la moyenne globale pour toutes les expérimentations.

La somme des carrés d'erreur pure,  $SS_E$ , est donnée par [81] :

$$SS_E = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (107)$$

où  $k$  est le nombre de paramètres,  $y_{ij}$  est  $j^{\text{ème}}$  la réponse au paramètre  $P_i$ ,  $\bar{y}_i$  est la moyenne du  $i^{\text{ème}}$  set,  $n_i$  est le nombre d'expérimentations au niveau  $i$  du paramètre  $P_i$ .

La valeur de  $F_{\alpha, k-1, m}$  peut être déduit du Tableau 4.12 dans Annexe VII.

<sup>14</sup> en statistique, c'est une mesure la dispersion d'une série de valeurs autour de leur moyenne.

<sup>15</sup> F-distribution ou Loi de Fisher en statistiques, est une loi de probabilité continue en statistique.

#### 4.3.1.4 Déplacement dans l'espace du problème : direction de recherche

À partir d'un modèle juste du premier ordre, une direction de minimisation peut être déduite pour se déplacer à la recherche d'optimum dans l'espace du problème. Ce qui revient, en autres termes, à trouver la direction et le pas de déplacement.

Considérons un problème de minimisation de la fonction coût :

$$y = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \dots + \beta_k P_k \quad (108)$$

qui est sujet à la contrainte par la condition :

$$\rho^2 \geq \sum_{i=1}^k P_i^2 \quad (109)$$

où  $y$  est la réponse de la fonction coût,  $P_1, \dots, P_k$  sont les paramètres,  $\beta_1, \dots, \beta_k$  sont les coefficients du polynôme,  $\rho$  est la distance de déplacement.

En utilisant l'approche du multiplicateur de Lagrange, un terme de pénalité  $\lambda$  est ajouté aux solutions qui ne satisfont pas la contrainte de l'équation (109) [87] :

$$L = \sum_{i=1}^k \beta_i P_i + \lambda \left( \sum_{i=1}^k P_i^2 - \rho^2 \right) \quad (110)$$

où  $\lambda$  est le multiplicateur de Lagrange.

Pour un problème de minimisation, le terme  $L$  précédent est minimisé en prenant ces dérivées partielles par rapport à  $P$  :

$$\frac{\partial L}{\partial P} = \sum_{i=1}^k \beta_i + 2 \lambda \sum_{i=1}^k P_i = 0 \quad (111)$$

et par rapport à  $\lambda$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^k P_i^2 - \rho^2 = 0 \quad (112)$$

La solution des Équations (111) et (112) est [42] :

$$P_i^* = -\rho \frac{\beta_i}{\sqrt{\sum_{j=1}^k \beta_j^2}}; i = 1, 2, \dots, k \quad (113)$$

Donnant une valeur incrémentale au terme  $\rho = (1, 2, 3, \dots)$  qui est multiplié par le gradient (le deuxième terme à droite de l'Équation (113)), le processus est conduit suivant la direction de la minimisation de la fonction coût (Équation (108)).

#### 4.3.1.5 Détermination d'optimum et son type

Dès que le modèle de deuxième ordre est disponible, une analyse s'effectue afin de connaître la position du point optimum où les dérivées partielles de la fonction coût approximée seront nulles. Le point d'optimum obtenu ainsi est appelé point stationnaire ; c'est un point où la réponse est maximale, minimale ou un point col.

Le modèle de deuxième ordre exprimé dans l'Équation (91) peut s'écrire sous la notation matricielle suivante :

$$\hat{y} = \beta_0 + \mathbf{x}^T \mathbf{b} + \mathbf{x}^T \mathbf{B} \mathbf{x} \quad (114)$$

Où le vecteur des variables est :

$$\mathbf{x} = [x_1 \quad \cdots \quad x_n]^T \quad (115)$$

le vecteur des coefficients de premier ordre est :

$$\mathbf{b} = [\beta_1 \quad \cdots \quad \beta_n]^T \quad (116)$$

et la matrice symétrique des coefficients de deuxième ordre est :

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1n}/2 \\ \vdots & \ddots & \vdots \\ \cdot & \cdots & \beta_{nn} \end{bmatrix} \quad (117)$$

La dérivée du modèle (114) par rapport au vecteur  $\mathbf{x}$  est nulle au point stationnaire :

$$\frac{\partial \hat{y}}{\partial \mathbf{x}} = \mathbf{b} + 2 \mathbf{B} \mathbf{x} = 0 \quad (118)$$

La résolution de ce dernier donnera le point stationnaire :

$$\mathbf{x}_s = -\frac{1}{2} \mathbf{B}^{-1} \mathbf{b} \quad (119)$$

et la réponse prédite au point stationnaire :

$$\hat{y}_s = \beta_0 + \frac{1}{2} \mathbf{x}_s^T \mathbf{b} \quad (120)$$

Pour connaître le type d'optimum trouvé au point stationnaire, il suffit d'obtenir les valeurs propres de la matrice  $\mathbf{B}$  (Équation (117)) [81] :

$$\mathbf{B} \mathbf{x} = \lambda \mathbf{x} \quad (121)$$

et pour le cas où les  $\{\lambda_i\}$  ont :

- des signes positifs, alors  $\mathbf{x}_s$  est un point de réponse minimale ;
- des signes négatifs, alors  $\mathbf{x}_s$  est un point de réponse maximale ;
- des signes variés, alors  $\mathbf{x}_s$  est un point de réponse de type point col.

Le modèle de deuxième ordre obtenu peut être courbé dans une ou plusieurs directions. Une analyse des dérivées partielles secondes notées sous la forme de la matrice Hessienne permet de donner plus d'informations sur cette courbature. La matrice Hessienne du modèle exprimé dans l'Équation (91) est donnée par :

$$H(y) = \begin{bmatrix} \frac{\partial^2 \hat{y}}{\partial P_1^2} & \frac{\partial^2 \hat{y}}{\partial P_1 \partial P_2} \\ \frac{\partial^2 \hat{y}}{\partial P_1 \partial P_2} & \frac{\partial^2 \hat{y}}{\partial P_2^2} \end{bmatrix} \quad (122)$$

La comparaison des termes diagonaux dans la matrice précédente indique l'intensité de la courbature de la surface de réponse dans la zone de validité du modèle et pour chaque paramètre. Des méthodes graphiques et les contours peuvent être également employées afin d'étudier la forme de la surface de réponse et pour prédire avec certitude le type du point optimum [81].

### 4.3.1.6 Optimiseur OptDOE

Pour rendre l'optimisation par simulation et par plans d'expériences possible avec le modèle SYSMFS, un optimiseur OptDOE a été développé au cours de ce travail. OptDOE emploie trois actions principales qui sont effectuées en cascade (Figure 4.17) : un plan d'expériences avec un modèle de premier ordre, un déplacement par une méthode de la plus grande pente et un plan d'expériences avec un modèle de deuxième ordre. Des actions supplémentaires interviennent occasionnellement afin de vérifier les points où les simulations se dérouleront, pour vérifier la précision et pour déterminer le type d'optimum obtenu.

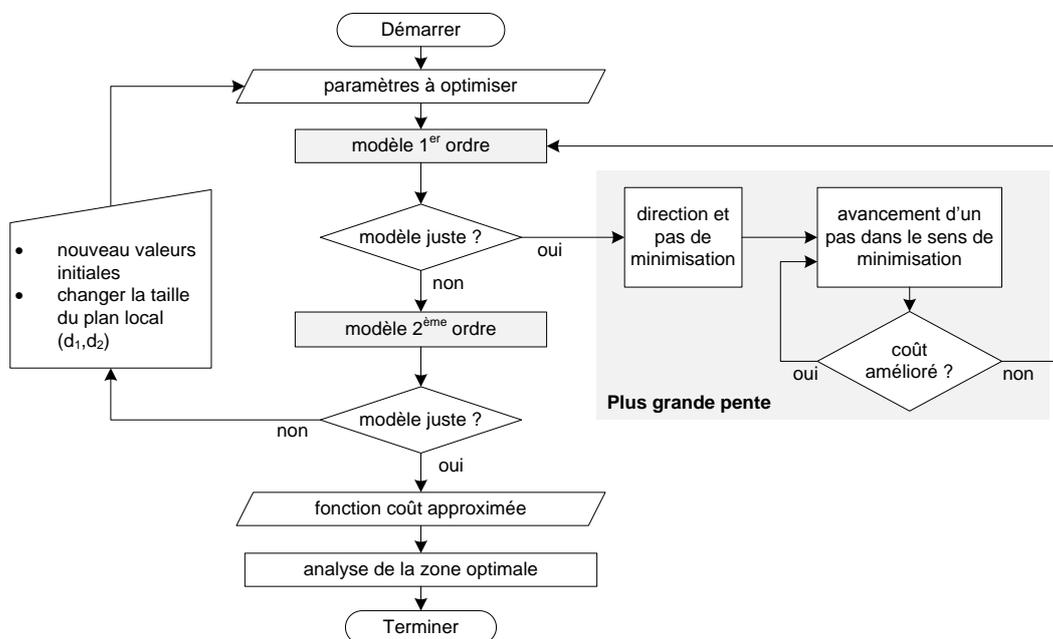


Figure 4.17 : Organigramme de l'optimiseur OptDOE

Techniquement parlant, l'algorithme de l'optimiseur OptDOE est écrit en MATLAB [77], [88]. Il dispose de sa propre base de données dans laquelle il enregistre les valeurs de la fonction coût pour chaque simulation. Ainsi, il peut passer la valeur de la fonction coût d'une simulation précédente sans la refaire, ce qui économise considérablement du temps de calcul.

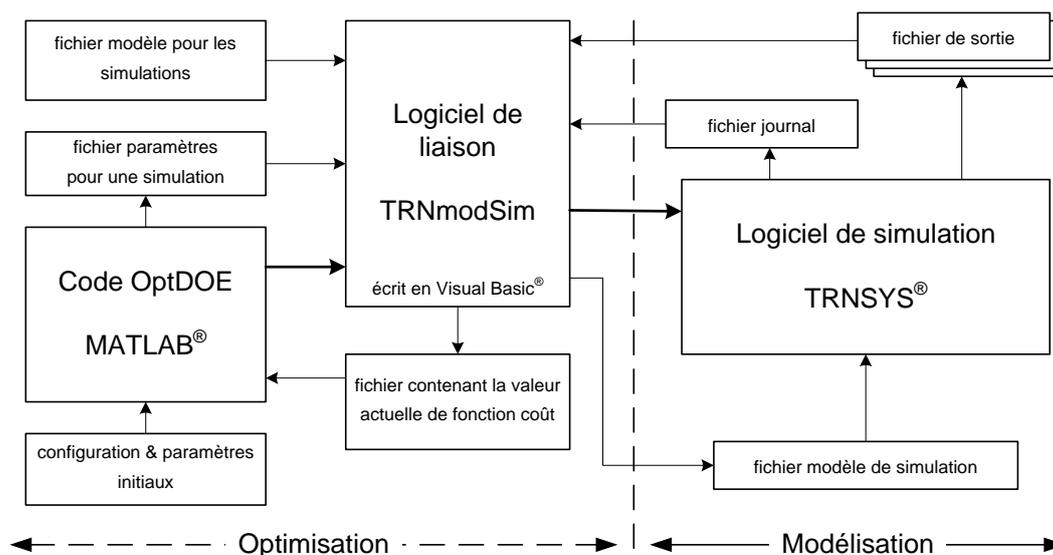


Figure 4.18 : Schéma technologique de la liaison OptDOE-TRNSYS

Le transfert des données en amont et aval entre les deux logiciels (simulation TRNSYS et optimisation MATLAB) se fait par le moyen d'un outil de liaison, développé en Visual Basic [89]. La Figure 4.18 montre le schéma de la liaison OptDOE-TRNSYS.

### 4.3.2 Optimisation par plans d'expériences de la fonction de Rosenbrock

L'application de l'optimisation par plans d'expériences (OptDOE) sur la fonction de référence de Rosenbrock va nous révéler les points forts et faibles de la méthode ; le Tableau 4.4 donne un aperçu de ces tests.

Tableau 4.4 : Résultats de la validation d'OptDOE avec la fonction de Rosenbrock

Test	Variable X1	Variable X2	Remarque sur le point de départ	Optimum trouvé à la simulation n°	Variable X1 optimale	Variable X2 optimale	Réponse Y optimale	Type d'optimum trouvé
1	0,2	-0,8	-	6	0,3	-0,6	1,31	point col
2	2	2,5	hors limites		pas de réponse			-
3	3	0	hors limites		pas de réponse			-
4	2,9	0	voisinage de frontière	34	0	0	1	†
5	-0,8	0	voisinage de frontière	13	0	0	1	-
6	0,5	0,5	-	3	0,4	0,6	0,52	point col
7	2,6	1,6	-	1	2,6	1,6	2,72	†
8	2,6	-1,6	-	1	2,6	-1,6	2,72	†
9	1	1	point minimum	1	1	1	0	†
10	1,3	0,8	proche de minimum	point stationnaire hors limites				-
11	2	-0,5	proche de minimum	17	1,5	-1,2	0,61	point col
12	-0,5	1,5	-	21	0,1	0,3	0,82	point col
13	-0,5	-1,5	-	21	0,1	-0,3	0,82	point col
14	-0,8	1,8	-	32	0	0	1	†
15	1	1,8	-	point stationnaire hors limites				-

(†) Le modèle de deuxième ordre ne peut pas présenter les données.

D'après les tests précédents, l'OptDOE arrive bien à trouver une valeur améliorée en comparaison avec le point de départ. Il suit correctement le chemin de minimisation jusqu'à la vallée ; mais après, il s'arrête directement donnant une valeur finale sans se déplacer plus loin dans l'espace plat de la vallée. Ce comportement peut être expliqué par le fait que le fond de la vallée dans la fonction Rosenbrock a une très faible inclinaison ce qui rend difficile de détecter la direction d'inclinaison dans certains endroits de cette vallée.

Dans la pratique, il est technologiquement avantageux de trouver une telle zone car, en effet, cette zone représente une plage de variation des paramètres où le coût du système ne change que très peu. Une zone similaire a été déjà obtenue pour le SYSMFS lors de son étude paramétrique (Figure 3.20).

De plus, OptDOE a l'avantage supplémentaire de décrire le point optimum trouvé. Dans le Tableau 4.4, l'analyse du modèle de deuxième ordre donne un indice sur le type d'optimum trouvé, ce qui permet de savoir, par exemple, si l'optimum était un point col (ex.

test 11, Tableau 4.4). Dans ce dernier cas, certaines combinaisons des paramètres dans la zone optimales pourront augmenter le coût comme le montre la Figure 4.19).

Dans le Tableau 4.4, il y avait des cas où le modèle de deuxième ordre obtenu ne pouvait pas représenter ces données (cas 4, 7, 8, 9 et 14). Il se peut parfois que les populations (simulations) aient des moyennes très diversifiées et que l'analyse de variance utilisée ne peut pas être effectuée correctement [85] ; nous considérons alors que l'optimisation n'aboutit pas au modèle prédictif dans ces cas.

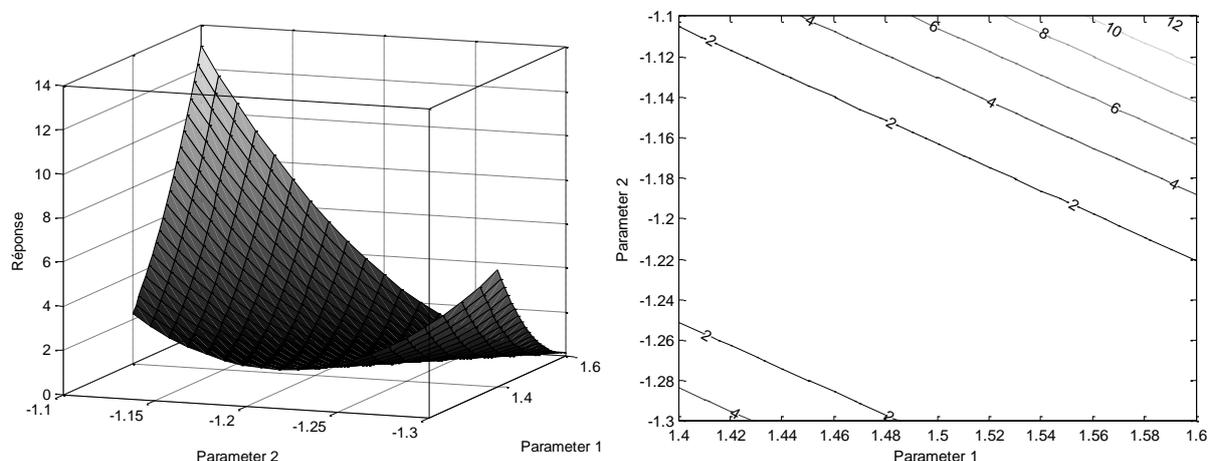


Figure 4.19 : Modèle de deuxième ordre pour le Test 11 dans le Tableau 4.4

Finalement, les points trouvés se diversifient largement suivant la position de leurs points de départ. Le vrai minimum est obtenu uniquement dans le test 9 où le point de départ est proche de l'optimum lui-même. Pour un point de départ situé directement dans la vallée (tests 1 et 6), le déplacement de l'algorithme reste limité.

Les Figures .20 à 4.24, présentent une illustration des résultats présentés dans le Tableau 4.4.

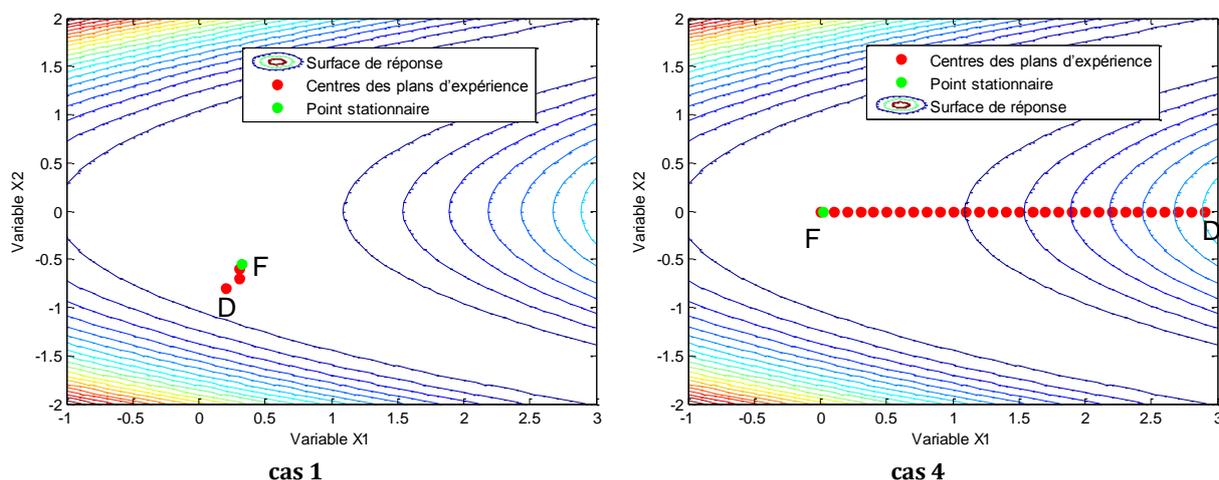
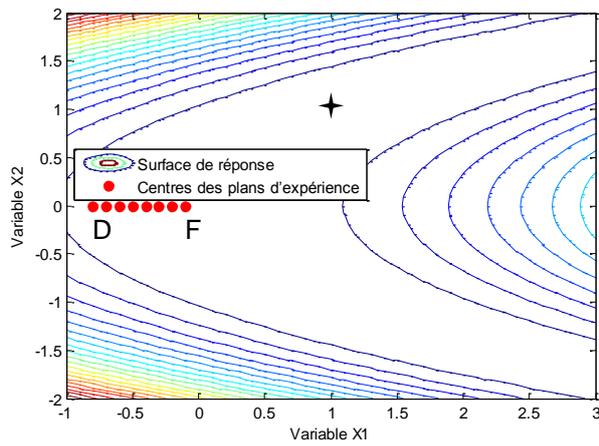
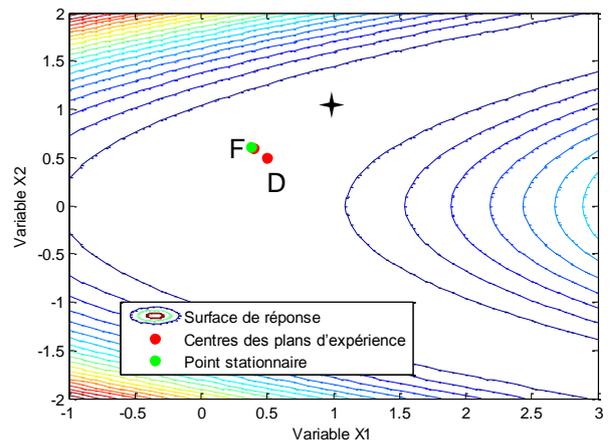


Figure 4.20 : Validation d'OptDOE avec la fonction de Rosenbrock (cas 1 et 4)

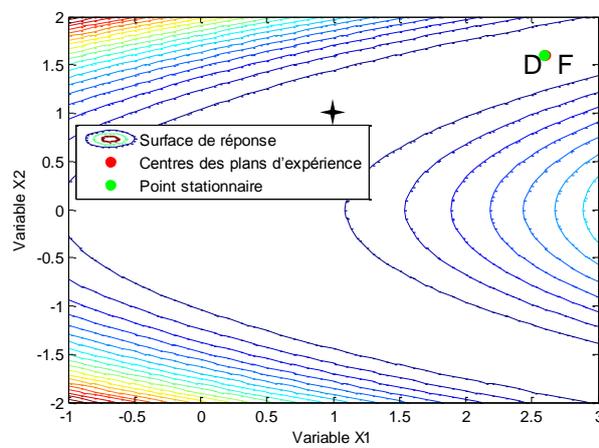


cas 5

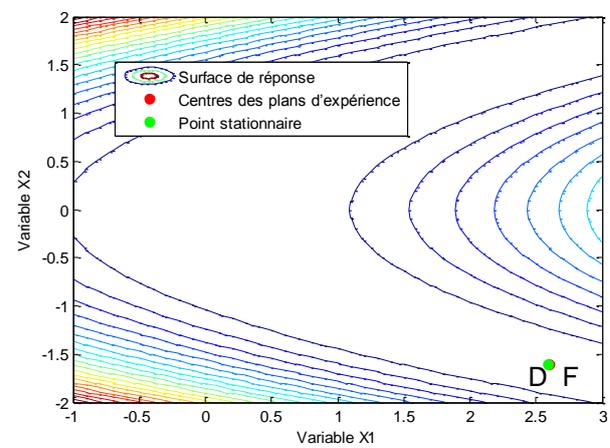


cas 6

Figure 4.21 : Validation d'OptDOE avec la fonction de Rosenbrock (cas 5 et 6)

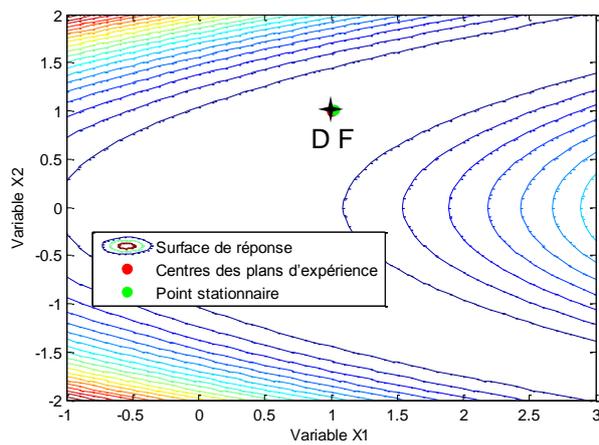


cas 7

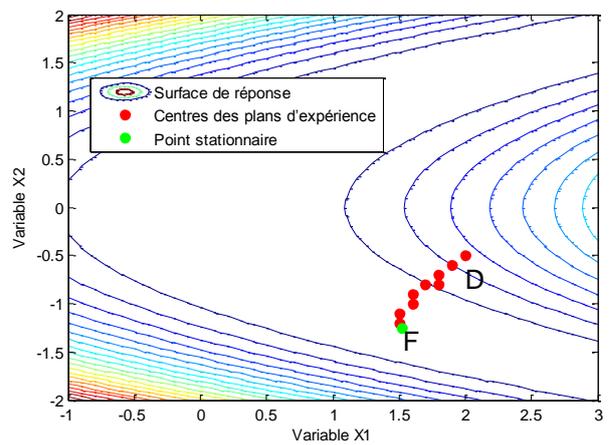


cas 8

Figure 4.22 : Validation d'OptDOE avec la fonction de Rosenbrock (cas 7 et 8)



cas 9



cas 11

Figure 4.23 : Validation d'OptDOE avec la fonction de Rosenbrock (cas 9 et 11)

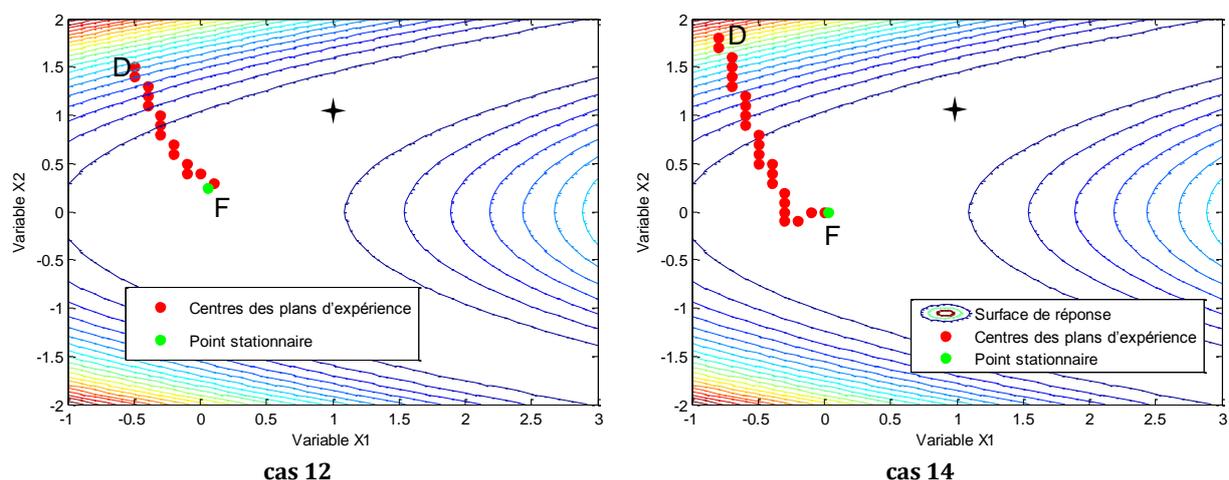


Figure 4.24 : Validation d'OptDOE avec la fonction de Rosenbrock (cas 12 et 14)

Le test avec la fonction de Rosenbrock nous a permis de valider le code d'OptDOE, au moins, d'une manière générique. Une application d'OptDOE sur le modèle SYSMFS va révéler si cette méthode est appropriée pour le dimensionnement optimal du SYSMFS.

### 4.3.3 Optimisation de SYSMFS par les plans d'expériences

#### 4.3.3.1 Taille des plans d'expériences locaux ( $d_1, d_2$ )

Nous avons déjà proposé de se baser sur les propriétés physiques des composants à optimiser (§4.3.1) pour déterminer la taille du plan factoriel de base ( $d_1, d_2$  dans la Figure 4.16). Le choix de ( $d_1, d_2$ ) se fait alors avec l'aide des données des fabricants du panneau solaire et du ballon de stockage thermique. La taille minimale des modules qui constituent le panneau entier peut être prise pour la valeur de ( $d_1$ ) (ex. 0,5 m<sup>2</sup>). Pour le volume de ballon, ( $d_2$ ) sera donc pris égal à l'itération minimale d'incrément du volume (ex. 25 L).

Tableau 4.5 : Optimisation du SYSMFS (modèle antérieur) par plans d'expériences

Algorithme	Test	Superficie initiale	Volume initial	Superficie optimale	Volume optimal	Coût optimal	Écart du coût de PAR1	Simulations	Type d'optimum
		(m <sup>2</sup> )	(litre)	(m <sup>2</sup> )	(litre)	(k€)	(%)		
<b>Paramétrique</b>	PAR1	-	-	7,7	1500	27	-	401	
<b>OptDOE</b>	DOE1	9	500	8,1	479	28,1	4,2	15	S
	DOE2	10	2000	7,7	1990	27,1	0,4	18	M
	DOE3	15	2800	7,7	2777	27,4	1,5	28	M
	DOE4	5	1500	7,8	1493	27,1	0,2	18	M
<b>OptDOE (<math>d_1, d_2</math>) plus large</b>	DOE1L	9	500	9	498	28,9	7,3	8	†
	DOE2L	10	2000	8,4	1983	28,9	6,9	14	M
	DOE3L	15	2800	15,3	2986	43,9	62,8	8	X
	DOE4L	5	1500	8,7	1554	29,1	7,8	15	†

(†) Le modèle de deuxième ordre ne peut pas présenter ses données, (M) Type d'optimum est un minimum, (X) Type d'optimum est un maximum, (S) Type d'optimum est un point col.

Nous avons voulu cependant tester l'effet de la taille du plan local sur le déroulement et les résultats de l'optimisation. Nous considérons alors les dimensions ( $d_1, d_2$ ) des modèles locaux linéaires et quadratiques comme des paramètres de calibrage pour l'OptDOE. Pour déterminer leurs effets, des expérimentations ont été menées avec les valeurs de ( $d_1, d_2$ ) que nous proposons (0,5 m<sup>2</sup>, 25 L) et celle plus grandes (2 m<sup>2</sup>, 100 L) ; une comparaison de ces résultats est présentée dans le Tableau 4.5.

D'après ces tests (Tableau 4.5), les valeurs proposées de ( $d_1, d_2$ ) (DOE1 à DOE4) demandent plus de simulations par rapport aux tests menés avec des valeurs plus grandes (DOE1L à DOE4L) ; mais elles donnent en revanche un écart plus petit du coût de test PAR1 ce qui implique une précision améliorée. Pour la suite de ce travail, nous considérons uniquement les tests menés avec les valeurs proposées de ( $d_1, d_2$ ) (c.-à-d. 0,5 m<sup>2</sup>, 25 L). Nous considérons également que le test DOE4, qui donne l'écart le plus petit de 0,2%, comme étant la meilleure solution trouvée par la méthode OptDOE.

Les modèles de deuxième ordre dans les tests DOE1L et DOE4L ne réussissent pas le test de manque d'ajustement et ils ne représentent pas en conséquence leurs données. L'analyse du point optimum trouvé par le test DOE3L a révélé que ce point est une maximisation ce qui rend les résultats de ce test complètement faux.

#### 4.3.3.2 Chemin suivi par la méthode d'optimisation par plans d'expériences

L'optimisation du dimensionnement du modèle SYSMFS final (Figure 3.10) est faite avec l'OptDOE. Contrairement à la méthode d'optimisation stochastique GPS-PSO, un chemin peut être tracé suivant le déplacement de la méthode OptDOE. Les Figures 4.25 et 4.26 montrent ce chemin avec les points de chaque plan d'expérimentation local des tests DOE1 à DOE4 du Tableau 4.5.

La numérotation des points illustrés dans les Figures 4.25 et 4.26 correspond à : (1) point centre du plan factoriel ou celui de Box-Wilson (CCC), (2) point d'extrémité du plan factoriel, (4) point de déplacement dans la direction de la plus grande pente (direction de minimisation), (5) point étoile du plan de Box-Wilson (CCC), (6) point d'optimum minimum, (7) point d'optimum maximum, (8) point col.

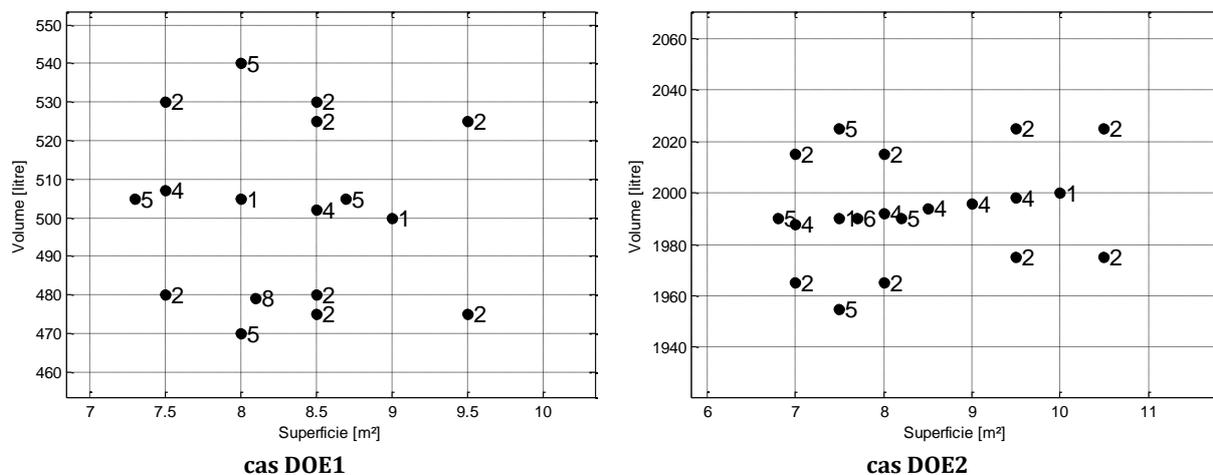


Figure 4.25 : Points d'expérimentations effectués par OptDOE dans les cas DOE1 et DOE2

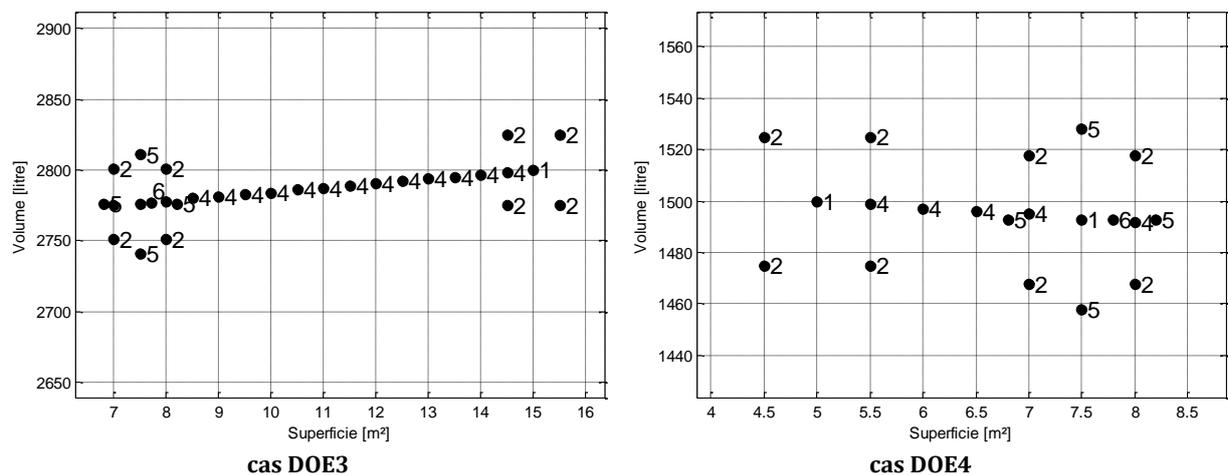


Figure 4.26 : Points d'expérimentations effectués par OptDOE dans le cas DOE3 et DOE4

#### 4.3.3.3 Comparaison des résultats d'optimisation du SYSMFS avec les deux algorithmes

L'application des méthodes d'optimisation sur un cas pratique comme un système solaire peut révéler les potentiels de chaque méthode. Même si une méthode d'optimisation se présente comme étant plus précise mathématiquement, dans la pratique, elle ne doit pas être coûteuse en temps. Une méthode d'une précision acceptable pour le domaine d'application et qui demande moins de simulations peut sembler donc intéressante.

Tableau 4.6 : Comparaison de l'optimisation du SYSMFS par plans d'expériences et l'algorithme hybride GPS-PSO

Algorithme	Test	Superficie initiale	Volume initial	Superficie optimale	Volume optimal	Coût optimal	Écart du coût de PAR1	Simulations	Type d'optimum
		(m <sup>2</sup> )	(litre)	(m <sup>2</sup> )	(litre)	(k€)	(%)		
<b>Paramétrique</b>	PAR1	-	-	7,7	1500	27	-	401	
<b>OptDOE</b>	DOE1	9	500	8,1	479	28,1	4,2	15	S
	DOE2	10	2000	7,7	1990	27,1	0,4	18	M
	DOE3	15	2800	7,7	2777	27,4	1,5	28	M
	DOE4	5	1500	7,8	1493	27,1	0,2	18	M
	DOE5	2	50	8	488	28,2	4,3	53	S
<b>Hybride GPS-PSO</b>	GEN1	9	500	7,5	1425	27,2	0,8	163	-
	GEN2	10	2000	7,5	2150	27,3	1,3	163	-
	GEN3	15	2800	7,5	1900	27,1	0,5	163	-
	GEN4	5	1500	8	1425	27,3	1	163	-

(†) Le modèle de deuxième ordre ne peut pas présenter ses données, (M) Type d'optimum est un minimum, (X) Type d'optimum est un maximum, (S) Type d'optimum est un point col.

Le Tableau 4.6 présente une comparaison des résultats de l'application des trois méthodes d'optimisation sur le même modèle SYSMFS. Comparant GPS-PSO à OptDOE, les meilleurs résultats sont ceux des tests DOE4 et GEN3, qui donne les coûts les plus proches de celui de

l'étude paramétrique du test PAR1. La comparaison démontre également la différence de l'effet du choix des valeurs initiales sur l'optimum obtenu entre GPS-PSO et OptDOE. L'écart d'erreur de la méthode OptDOE se réduit considérablement lorsque les valeurs initiales se trouvent suffisamment proche de l'optimum.

Pour pouvoir s'approcher suffisamment de l'optimum, il est indispensable d'avoir des informations sur la forme de la fonction coût du système étudié (surface de réponse). Or, cette situation idéale n'est pas toujours possible surtout avec des nouvelles conceptions dont le comportement n'est pas encore bien étudié. Nous avons pensé cependant à prendre en compte la nature physique réelle du système à optimiser ; il est donc plus approprié de démarrer l'optimisation au point le plus proche possible d'origine (ex. à 0 m<sup>2</sup> et à 0 L). Le test DOE5 (Tableau 4.6) prend en compte cette hypothèse mais la marge d'erreur et le nombre de simulations ont effectivement augmenté par rapport aux cas précédents (DOE1 à DOE4) ; le test DOE5 abouti néanmoins sur un volume de ballon de stockage thermique qui semble plus couramment rencontré dans les installations solaires domestiques et qui est deux fois moins grand que celui du DOE4. Il est donc intéressant d'analyser le modèle obtenu par le test DOE4 et dans la lumière des valeurs du test DOE5.

#### 4.3.3.4 Analyse de la fonction optimale donnée par l'optimisation par plans d'expériences

L'optimisation par plans d'expériences permet d'améliorer le dimensionnement d'un schéma donné (ex. SYSMFS) en donnant les valeurs optimales des paramètres dimensionnant. En outre, cette méthode a également l'avantage majeur de révéler la zone d'intérêt dans le voisinage des valeurs optimales trouvées. Le test d'optimisation DOE4 (Tableau 4.6) fournit une description de la fonction coût approximée dans la zone optimale et de la forme polynômiale suivante :

$$\hat{y} = 27270 - 1249.1 x_1 + 11.34 x_2 - 21.3 x_1 x_2 + 1156.7 x_1^2 + 84.13 x_2^2 \quad (123)$$

$$x_1 = \frac{A_{panu} - 7.5}{0.7}, x_2 = \frac{V_{bln} - 1493}{35}$$

où  $V_{bln}$  est le volume du ballon (L),  $A_{panu}$  est la surface du panneau solaire (m<sup>2</sup>).

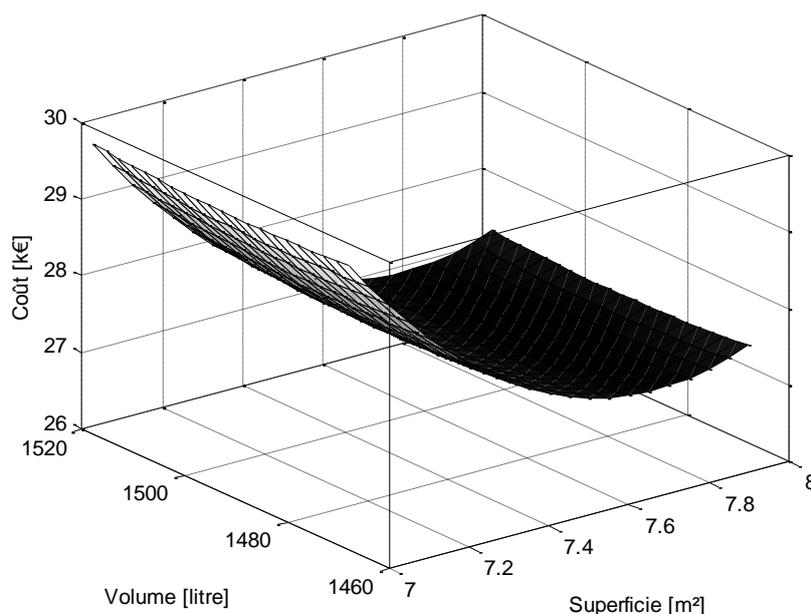


Figure 4.27 : Fonction coût approximée par la forme polynômiale (Équation (123))

Une méthode pour valider le modèle obtenu consiste à comparer les valeurs prédites par les réponses des données qui ont généré ce modèle. Nous trouvons ainsi que cette valeur est petite comparée aux coûts de l'ordre de k€ obtenus précédemment (Tableau 4.6) :

$$Err_{max} = \max(|\hat{y} - y|) = 208 \text{ €} \quad (124)$$

où  $Err_{max}$  est l'erreur maximale entre la valeur réelle et celle prédite (€),  $\hat{y}$  est la réponse prédite par le modèle (Équation (123)) (€),  $y$  est la réponse réelle du système à partir de la simulation dans ce cas.

Le modèle est, quant à lui, valable uniquement dans la zone d'optimum :  $[6,8, 8,2] \text{ m}^2$  et  $[1458, 1528] \text{ L}$ . Les Figures .27 et .28 représentent graphiquement le modèle de deuxième ordre calé autour de l'optimum  $(7,8, 1493)$  ; on y voit bien la forme de vallée que nous avons vu précédemment dans la surface de réponse originale (Figure 3.20).

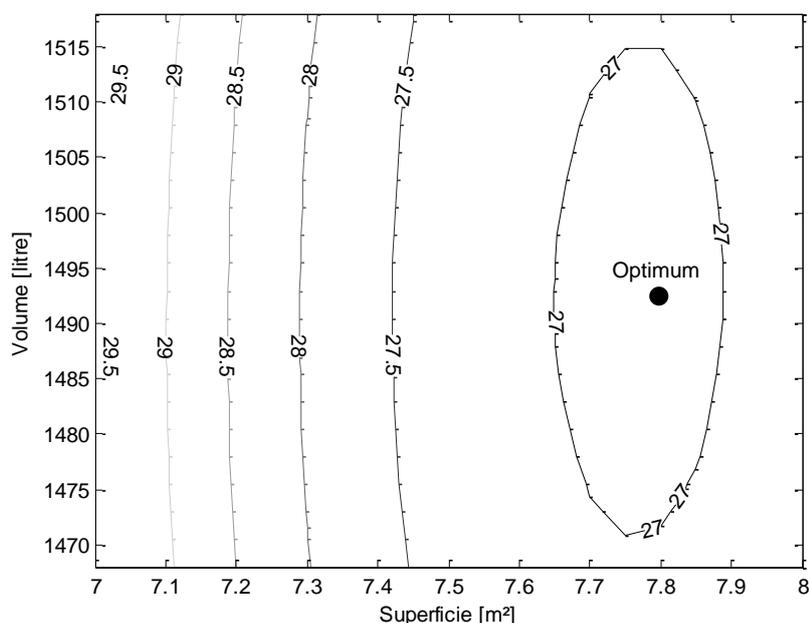


Figure 4.28 : Fonction coût approximée par la forme polynômiale (Équation (123))

La variation de la surface et du volume n'est pas assez grande pour la fonction obtenue ; elle peut cependant fournir analytiquement quelques informations sur la surface (Figure 4.27). Dans ce contexte, nous pouvons noter la matrice Hessienne pour le modèle obtenu :

$$H = \begin{bmatrix} \frac{\partial^2 \hat{y}}{\partial x_1^2} & \frac{\partial^2 \hat{y}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \hat{y}}{\partial x_1 \partial x_2} & \frac{\partial^2 \hat{y}}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 \times 1156.7 & -21.3 \\ -21.3 & 2 \times 84.13 \end{bmatrix} \quad (125)$$

Les vecteurs propres de la Hessienne sont :

$$v_1 = [-1 \quad 0,01]^T, v_2 = [-0,01 \quad -1]^T \quad (126)$$

Ces vecteurs indiquent la direction de la courbure de la surface de réponse :  $v_1$  montre une direction dans le sens de  $x_1$ , tandis que  $v_2$  montre une direction dans le sens de  $x_2$ . Les valeurs des courbures sont les valeurs propres de la Hessienne (Équation (125)) :

$$\lambda_1 = 2313, \lambda_2 = 168$$

(127)

La courbure principale est donc dans la direction de  $v_1$  ce qui est presque dans la même direction que l'axe de  $x_1$ , la surface du panneau solaire. La deuxième valeur propre  $\lambda_2$  montre également que la courbure est plus petite dans la direction du deuxième vecteur propre,  $v_2$ , qui est dans la direction du volume de ballon de stockage thermique.

Le fond de la vallée de la surface de réponse (voir Figures .20 et .27) s'incline d'une façon différente dans la direction de deux paramètres, surface et volume. Ce fait empêche le bon calage d'un modèle de premier ordre au fond de la vallée et une direction de descente ne peut pas être alors détectée. L'analyse de la Hessienne donne cependant une direction dans laquelle le coût ne varie pas pour une large variation du volume et une baisse dans la valeur du volume de ballon n'aura pas ainsi un effet signifiant sur la valeur du coût total. Cette conclusion est technologiquement importante puisqu'on cherche toujours à diminuer la taille des équipements comme le ballon de stockage thermique.

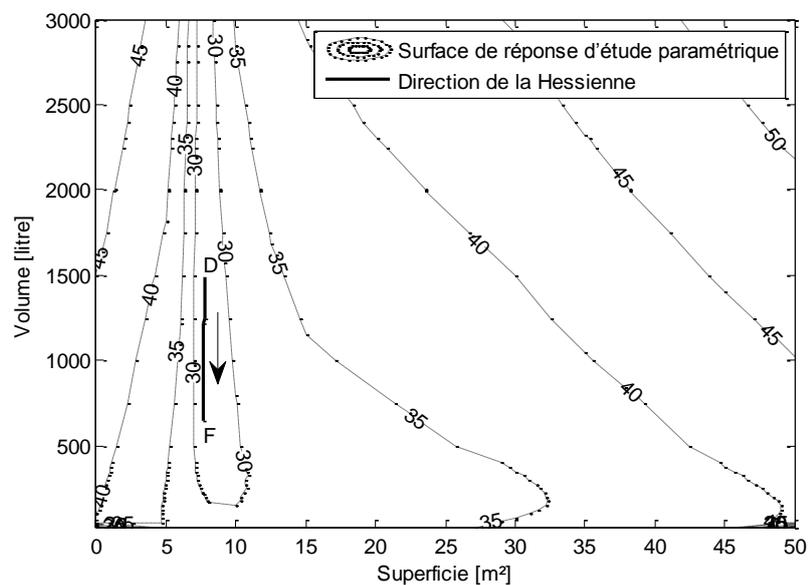


Figure 4.29 : Suivi de la direction de la Hessienne

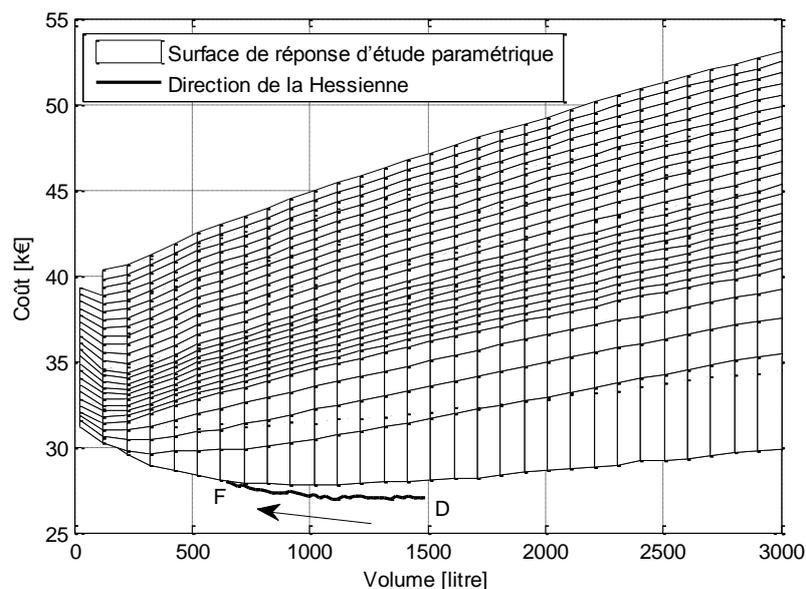


Figure 4.30 : Suivi de la direction de la Hessienne superposé sur la surface de réponse

Un intérêt peut être alors apporté à ce dernier point et le processus d'optimisation peut alors reprendre du point optimal suivant la direction de l'inclinaison la plus faible déduite par l'analyse de la Hessienne. Pour cette étude, nous ne considérons que les valeurs du volume les plus petites par rapport à la valeur optimale (direction négative de l'axe  $x_2$ ) ; nous cherchons ainsi à donner la plage de variation optimale du volume dans laquelle la variation du coût total reste dans une tolérance prédéfinie (ex. 3%). Les Figures 4.29 et 4.30 montrent le suivi de la direction déduite par la Hessienne avec une tolérance de 3% de variation du coût.

Nous pouvons désormais constater que l'optimisation du SYSMFS par les plans d'expériences (OptDOE) aboutit sur un point optimum, un modèle prédictif dans la zone optimale et une plage de valeurs pour le paramètre dont sa variation modifie le coût du système très légèrement. Ces informations sont un plus par rapport aux méthodes d'optimisation qui donnent uniquement les valeurs optimisées sans aucune information supplémentaire sur la zone optimale. C'est également un avantage lorsque la surface de réponse d'un système physique étudié forme une vallée comme c'est le cas pour le SYSMFS (des résultats supplémentaires des modélisations sur des modèles similaires sont présentés dans l'annexe ).

Pour le SYSMFS étudié et optimisé, une zone optimale de surface et de volume et une plage de variation du volume offrent une grande manœuvrabilité afin de déterminer, avec les contraintes du site d'installation, le dimensionnement optimal qui garantisse le moindre coût total avec une bonne performance énergétique.

