

5.1.1.5	Validation croisée	66
5.1.2	Krigeage.....	68
5.1.3	Simulations conditionnelles.....	78
5.1.3.1	Comparaison entre les SGCs et les SSCs	79
5.1.3.2	Validation statistique des résultats des simulations conditionnelles	83
5.1.3.3	Analyse des résidus de la moyenne de 500 simulations conditionnelles	84
5.2	Deuxième cas d'étude : aube d'une turbine hydraulique.....	91
5.2.1	Paramétrisation surfacique.....	91
5.2.1.1	Paramétrisation d'un maillage triangulaire.....	93
5.2.1.2	Paramétrisation de l'aube d'une turbine	98
5.2.2	Analyse variographique	101
5.2.3	Krigeage.....	107
5.2.4	Simulations stochastiques	111
CONCLUSION.....		119
RECOMMANDATIONS		121
LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES.....		127

LISTE DES TABLEAUX

	Page
Tableau 2-1	Comparaison entre les (SC) et (SNC).....37
Tableau 5-1	Comparaison entre les trois modèles : sphérique, gaussien et exponentiel65
Tableau 5-2	Coefficient de corrélation de Pearson entre les déformations théoriques et les trois types du krigeage69
Tableau 5-3	Comparaison des coefficients de corrélations de Pearson entre les KS, KO et KU71
Tableau 5-4	Régression linéaire simple entre le krigeage ordinaire et les simulations par des ÉFs74
Tableau 5-5	Intervalle de confiance des coefficients de prédiction75
Tableau 5-6	Coefficients de corrélations des simulations stochastiques81
Tableau 5-7	Régression linéaire simple entre les valeurs de déformations calculées par les ÉFs et la moyenne de 500 simulations conditionnelles86

LISTE DES FIGURES

		Page
Figure 1-1	Résumé des méthodes d'interpolation déterministes	5
Figure 1-2	Les méthodes d'interpolation probabilistes	8
Figure 2-1	Présentation schématique de l'incrémentation h	18
Figure 2-2	Traçage du semivariogramme expérimental.....	21
Figure 2-3	Les composantes d'un semivariogramme.....	22
Figure 2-4	Comparaison entre plusieurs modèles des semivariogrammes.....	24
Figure 2-5	Schéma mettant en évidence le krigeage,	36
Figure 2-6	Schéma explicatif de la simulation séquentielle	38
Figure 2-7	Algorithme des simulations séquentielles conditionnelles (SSCs).....	40
Figure 2-8	Algorithme des simulations gaussiennes conditionnelles (SGSs) - 1	42
Figure 2-9	Algorithme des simulations gaussiennes conditionnelles (SGSs) - 2.....	43
Figure 4-1	Méthodologie expérimentale	51
Figure 5-1	Éprouvettes à encoches (Arpin-Pont, 2012)	56
Figure 5-2	Déformation de l'éprouvette.....	56
Figure 5-3	Semivariogrammes pour toute, $\frac{1}{2}$ et $\frac{1}{4}$ de l'éprouvette.....	57
Figure 5-4	Transformation gaussienne	58
Figure 5-5	Comparaison entre les données primaires ÉFs, la transformée gaussienne et la transformée inverse.....	59
Figure 5-6	Comparaison entre la fonction de distribution cumulative des données primaires ÉFs et sa transformée inverse.	60
Figure 5-7	Variation de la direction de collecte des données.....	61
Figure 5-8	Variation de la direction de collecte des données $0^\circ : 10^\circ : 90^\circ$	62

Figure 5-9	Paires des données utilisées lors de la construction du variogramme expérimental.....	63
Figure 5-10	Nuage variographique et suppression des points atypiques.....	64
Figure 5-11	Modélisation du semivariogramme expérimental avec les semivariogrammes analytiques exponentiel, sphérique et gaussien	65
Figure 5-12	Ajustement du semivariogramme expérimental par la méthode des moindres carrés	66
Figure 5-13	Validation croisée : Comparaison entre les coefficients de corrélations entre les trois modèles.....	67
Figure 5-14	Krigeage simple, ordinaire et universel	68
Figure 5-15	Comparaison entre les <i>cdf</i> de trois types de krigeage avec celle de simulations par des éléments finis (<i>cdf</i> référentielle)	70
Figure 5-16	Les semivariogrammes des surfaces krigées KS, KO et KU.....	71
Figure 5-17	Variance de krigeage KS, KO et KU	72
Figure 5-18	Régression linéaire simple entre le krigeage ordinaire et les ÉFs	73
Figure 5-19	Analyse des résidus du krigeage ordinaire.....	76
Figure 5-20	Comparaison entre les déformations ÉFs et les cinq exemples de simulations conditionnelles choisies aléatoirement parmi l'ensemble de 500 simulations initiales.....	78
Figure 5-21	Comparaison entre la cartographie de la moyenne de 100, 250 et 500 simulations stochastiques.....	79
Figure 5-22	Comparaison entre la cartographie de la variance de 100, 250 et 500 simulations stochastiques.....	80
Figure 5-23	Graphique matriciel diagonal en fonction du nombre des simulations.....	81
Figure 5-24	Comparaison entre les semivariogrammes de 100, 250 et 500 simulations stochastiques SSCs et SGSs	82
Figure 5-25	Cartographie des déformations simulées par ÉFs et leur prédiction par les simulations séquentielles conditionnelles	83
Figure 5-26	Histogrammes des déformations calculées par la méthode ÉFs et leurs prédictions par les 500 simulations séquentielles conditionnelles.....	84

Figure 5-27	Histogramme des résidus de la moyenne de 500 simulations conditionnelles.....	85
Figure 5-28	Droite d'ajustement entre les estimations calculées par ÉFs et la moyenne de 500 simulations conditionnelles.....	86
Figure 5-29	Comparaison entre la variance du krigeage ordinaire et la variance de la moyenne de 500 simulations conditionnelles.....	87
Figure 5-30	Coefficient de variance de la variance moyenne de 500 simulations conditionnelles.....	88
Figure 5-31	Les semivariogrammes : Variation du nombre de jauges et leurs positions.....	90
Figure 5-32	Paramétrisation : Le modèle du ressort.....	95
Figure 5-33	Paramétrisation : angles du maillage triangulaire.....	96
Figure 5-34	Zone A de la turbine hydraulique.....	99
Figure 5-35	Paramétrisation de la zone A de l'aube.....	100
Figure 5-36	Transformation gaussienne de contraintes von-Mises de la zone A.....	101
Figure 5-37	Construction du semivariogramme expérimental suivant plusieurs directions.....	102
Figure 5-38	Construction du semivariogramme expérimental.....	103
Figure 5-39	Ajustement du semivariogramme par la méthode des moindres carrés.....	104
Figure 5-40	Validation croisée.....	105
Figure 5-41	Sensibilité du variogramme à la portée.....	105
Figure 5-42	Krigeage avec un variogramme sphérique Palier =0.15 et Portée= 0.64.....	106
Figure 5-43	Krigeage avec un variogramme sphérique Palier =0.15 et Portée= 0.34.....	106
Figure 5-44	Comparaison entre la cartographie du KS, KO et KU avec la référence ÉFs.....	107
Figure 5-45	Comparaison entre les histogrammes du KS, KO et KU.....	108

XVIII

Figure 5-46	Comparaison entre les variances des KS, KO et KU.....	108
Figure 5-47	Droite d'ajustement entre les contraintes simulées par les ÉFs et estimées par le krigeage ordinaire.....	109
Figure 5-48	Analyse des résidus du krigeage.....	110
Figure 5-49	Les semivariogrammes de la moyenne de 500 simulations, surface krigée et ÉFs.....	111
Figure 5-50	Droite d'ajustement entre la moyenne des 500 simulations conditionnelles et les résultats simulés par ÉFs.....	112
Figure 5-51	Comparaison entre la variance du krigeage et la variance de 500 simulations stochastiques.....	113
Figure 5-52	Comparaison des histogrammes des ÉFs et la moyenne de 500 simulations stochastiques.....	114
Figure 5-53	Comparaison entre les pdf et les cdf des ÉFs et la moyenne de 500 simulations stochastiques.....	114
Figure 5-54	CV de la moyenne de 500 simulations	115
Figure 5-55	Contraintes de von Mises en 3D ÉFs.....	116
Figure 5-56	Contraintes de von Mises en 3D : moyenne de 500 simulations stochastiques	116
Figure 5-57	Contraintes de von Mises 3D : variance de 500 simulations stochastiques	117

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

CTRIA6	Maillage surfacique composé de 6 nœuds utilisé par NX 8
ÉFs	Simulations par les méthodes des éléments finis
KO	Krigeage ordinaire
EG	Espace gaussien
KS	Krigeage simple
KU	Krigeage universel (Krigeage avec tendance)
IREQ	Institut de recherches d'Hydro Québec
NX 8	NX CAD Siemens PLM Software
SC	Simulations conditionnelles
SGeMS	Standford Geostatistical Modeling Software
SGCs	Simulations conditionnelles gaussiennes
SSCs	Simulations séquentielles conditionnelles
SNCs	Simulations stochastiques non conditionnelles
SSCs	Simulations séquentielles conditionnelles
TGI	Transformation gaussienne inverse

LISTE DES SYMBOLES ET UNITÉS DE MESURE

Abréviation	Signification
a	Portée du semivariogramme théorique
C_0	Effet de pépite du semivariogramme théorique
C_R	Covariance du résidu de la variable aléatoire
C	Covariance de la variable aléatoire
$ccdf$	<i>Conditionnal Cumulative Distribution Function</i>
cdf	Fonction de distribution cumulative
CV	Coefficient de variation en %
D	Domaine d'étude de la simulation
$E\{z\}$	Espérance de la variable aléatoire
$epsz$	Déformation suivant l'axe de traction de l'éprouvette
FA	Fonction aléatoire
h	Incrémentation lors de la recherche des paires de mesures
l	Numéro de la simulation conditionnelle
L	Nombre total des simulations conditionnelles
$n(h)$	Nombre des nœuds du maillage
p	Puissance (les méthodes barycentriques)
pdf	Fonction de distribution de probabilité
QQ	Graphique quantile-quantile
s_0	Localisation de la mesure à estimer
s_i	Localisation de la mesure expérimentale lors de l'interpolation
s_α	Localisation de la première mesure de la paire du point lors de la variographie expérimentale.
s_β	Localisation de la deuxième mesure de la paire du point lors de la variographie expérimentale.
p	Coefficient de la méthode <i>distances inverses</i>
VA	Variable aléatoire
$\hat{z}(s_0)$	Estimation de la variable aléatoire

$z(\mathbf{s}_i)$	Attribut à la position de la valeur observée
$Z(\mathbf{s}_\alpha)$	Attribut à la localisation \mathbf{s}_α
$Z_K^{SNC}(\mathbf{s}_0)$	Simulation non conditionnelle générée à partir du semivariogramme des données gaussiennes.
$Z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0)$	Premier conditionnement par krigeage de la simulation non conditionnelle
$z_K^{SC}(\mathbf{s}_0)$	Estimation de la variable aléatoire suite à un deuxième conditionnement par le krigeage.
μ_S	Unité de mesure de la déformation en <i>micro-strain</i>
λ_i	Poids de l' $i^{\text{ème}}$ observation expérimentale
$\mu(\mathbf{s})$	Dérive de la fonction aléatoire
$\epsilon(\mathbf{s})$	Bruit de la fonction aléatoire
σ^2	Variance de la fonction aléatoire
$\hat{\beta}_j$	Coefficient de régression
\mathbf{u}_j	Localisation du nœud de la grille de simulation
R_{ajust}^2	Coefficient de corrélation ajusté
\mathbf{u}_α	Localisation de l'attribut z sur la grille
$\gamma(\mathbf{h})$	Moitié de la différence de semivariance entre deux mesures expérimentales
$\hat{\gamma}(\mathbf{h})$	Estimation de la moitié de la semivariance expérimentale

INTRODUCTION

Problématique industrielle

La production de l'énergie électrique à partir de l'énergie hydraulique est assurée par des turbines hydroélectriques et des génératrices synchrones. Ces équipements sont coûteux de telle sorte que les utilisateurs les exploitent à la limite de leurs spécifications pour des raisons concurrentielles, économiques et environnementales. Ne pas avoir des données expérimentales précises pourra facilement réduire dans un premier temps l'estimation de la fiabilité résiduelle de la turbine hydroélectrique, et dans un second temps augmenter les risques de défaillances et engendrer des coûts d'exploitation supplémentaires suite à des arrêts pour des réparations. En revanche, le fait de mesurer plus précisément la charge appliquée permet de prédire avec un bon niveau de confiance le niveau de dommage et l'espérance de vie subsistante de la turbine. Ces deux paramètres seront par la suite les intrants clés dans un programme de gestion des risques.

Également, la variation, cyclique ou aléatoire, des besoins énergétiques (jour, saison, heures de pointe et la demande supplémentaire causée par l'ouverture du marché) a pour conséquence une utilisation non constante des turbines hydroélectriques. Ces modes de production intermittents engendrent une augmentation du nombre des régimes transitoires, non stationnaires : le démarrage (chargement) et l'arrêt (déchargement) de la turbine. Or, c'est précisément ces régimes qui peuvent induire le plus de dommage en fatigue. Ainsi, pour répondre aux besoins énergétiques quotidiens, les producteurs d'hydroélectricité augmentent considérablement le risque de défaillance de la turbine. En effet, la plus part des turbines ont été conçues lors de leur conception pour des cycles journaliers et non pas pour plusieurs cycles pendant la même journée. Le fait d'alterner le chargement et le déchargement favorise l'avancement des fissures et augmente par conséquent la probabilité de rupture de production de ces équipements suite à des endommagements.

Vu qu'on ne peut pas remplacer les anciennes turbines par des nouvelles qui peuvent accepter ces nouvelles variations, on doit prévoir la nouvelle durée de vie, avec un niveau de confiance acceptable tout en tenant compte des nouvelles conditions d'utilisations. Ainsi, nous pouvons adapter une nouvelle politique de remplacement ou de réparation de la turbine. En plus de cette utilisation non constante, le vieillissement des centrales hydroélectriques a un impact majeur sur leurs entretiens. Par conséquent, l'estimation du niveau de dommage et de la charge dynamique réelle appliquée (mesurée) seront primordiales pour élaborer un échéancier efficient pour l'entretien et gérer le risque d'indisponibilité des équipements lors des réparations ou de remplacements.

L'Institut de recherche d'Hydro Québec (IREQ) et Andritz Hydro Ltée ont entrepris depuis quelques années plusieurs projets de recherche et développement dans le but de relever ces défis. Un de ces projets, dans lequel mes travaux s'imbriquent, est le projet PrédIT (Prédiction de la dégradation et diagnostic intégré des turbines). C'est un projet qui a pour but de développer les outils de diagnostic et des méthodes optimisées pour estimer l'état opérationnel des turbines et, par conséquent, délimiter les arrêts imprévus en fonction de leur état réel.

Le développement, ainsi que la validation, de ces outils se basent sur une estimation de la fiabilité en fatigue en générant des modèles de prédiction de l'état d'endommagement de la turbine dans le but de planifier plus précisément la stratégie de réparation. Afin d'approfondir les recherches, plusieurs travaux ont déjà été réalisés pour caractériser la vie résiduelle. En tenant compte du diagramme de Kitagawa (Gagnon et al., 2013), deux paramètres importants peuvent influencer la prédiction de la fatigue : la taille des fissures et la charge dynamique appliquée. Plusieurs travaux traitent le côté métallurgique des fissures. Notre partie du projet s'intéressera à la détermination du deuxième paramètre : la charge appliquée.

Depuis 2008, dans le cadre du projet PrédIT, Szczota a modélisé l'historique des profils d'utilisation des turbine à partir d'un générateur de séquences synthétiques de chargement qui seront utilisées par la suite comme entrée pour le modèle de l'estimation de la vie

résiduelle de la turbine (Szczota, 2012). Ensuite, Gagnon a proposé à son tour une procédure de démarrage qui minimise les cycles de fatigue (LCF *Low Cycles Fatigues*) (Gagnon et al., 2010). Dans le même contexte de la quantification de la charge appliquée, Poirier a extrapolé les signaux de chargement d'une aube pour simuler les déformations mécaniques en régime permanent (Poirier, 2013). Finalement, J.Arpin-Pont a développé une méthode numérique d'évaluation des incertitudes des mesures par des jauges de déformation sur les aubes de la turbine. Ces derniers travaux seront exploités dans le présent projet lors de l'ajustement des modèles donnés par des simulations par des éléments finis ÉFs (Arpin-Pont, 2012).

Objectif de la recherche

Dans le cas d'une turbine hydraulique, on ne peut pas mesurer directement les contraintes réelles sur toutes les zones de l'aube. En fait, la prise des mesures est très coûteuse parce qu'elle engendre un arrêt de la turbine. Même si on immobilise la turbine lors des réparations, on est restreint spatialement et temporellement pour pouvoir installer des jauges de déformations. D'une part, les courtes durées réservées dans l'échéancier, pour qu'on puisse coller les jauges à des différents endroits de l'aube et d'une autre part, on ne pourra coller qu'un nombre limité de jauges dans des endroits qui ne sont pas forcément des localisations optimales cibles parfois inaccessibles. Ceci limite la collecte d'informations, car on n'aura qu'une faible partie de l'information désirée.

C'est pour ces raisons qu'on considère aujourd'hui des modèles numériques pour estimer la distribution de contraintes sur les aubes de la turbine. Ces modèles sont robustes pour simuler le comportement statique. En revanche, ils sont généralement moins adaptés pour prédire le comportement dynamique de la charge appliquée sur l'aube de la turbine. Pour valider et estimer les vraies contraintes supportées par les aubes de la turbine, on prend des mesures expérimentales à différents emplacements et on détermine les mesures restantes sur toute l'aube de la turbine par interpolation spatiale en utilisant le krigeage et les simulations stochastiques conditionnelles.

Le présent projet a pour but de développer un modèle mathématique pour interpoler les mesures expérimentales de la charge dynamique tout en utilisant les résultats des simulations par éléments finis comme source d'information à priori. L'hypothèse basique que nous postulons ici, est que la morphologie (le comportement) de la distribution spatiale des contraintes telles qu'estimées par ÉFs est très similaire à celle rencontrée expérimentalement. Par conséquent, on considère une variable spatiale dont elle est simulée théoriquement par des ÉFs partout dans l'espace et expérimentalement par sondage dans des points de mesures particuliers. On cherche donc à reconstituer la valeur expérimentale de cette variable en un point non sondé tout en déterminant l'erreur (incertitude) qu'on commet lors de l'interpolation spatiale.

Ce mémoire débutera par une revue de littérature de différentes méthodes d'interpolation spatiales déterministes ainsi que probabilistes. Le chapitre 2 présentera le cadre théorique. On développera les algorithmes du krigeage simple, ordinaire et universel ainsi que ceux des simulations stochastiques conditionnelles. Les deux chapitres suivants discuteront des hypothèses ainsi que la méthodologie expérimentale retenue et suivie. Finalement, le Chapitre 5 analysera les résultats des simulations sur une éprouvette à encoches et sur une aube d'une turbine hydraulique. Les conclusions, les recommandations, les perspectives de recherches ainsi que les limites de la méthode proposée clôturent ce mémoire.

CHAPITRE 1

REVUE DE LITTÉRATURE

L'objectif principal de ce mémoire consiste à effectuer l'interpolation spatiale des mesures expérimentales ainsi que la propagation de l'erreur engendrée. Ce chapitre présente une revue des principales méthodes d'interpolation spatiale distinguées en littérature. On distingue deux grandes classes : les méthodes déterministes et les méthodes probabilistes. Ces dernières se différencient par le concept du hasard et proposent des modèles probabilistes modélisant les erreurs aléatoires de prévision dans le but d'aider la prise des décisions lors des analyses des risques. Dans ce chapitre, on présente ces méthodes en détail tout en justifiant leurs avantages et inconvénients.

1.1 Méthodes déterministes

L'idée de base des méthodes déterministes est fondée sur l'observation suivante : les valeurs les plus proches les unes des autres doivent se ressembler plus et par conséquent reçoivent un poids important lors de la modélisation. La Figure 1-1 présente un résumé des principales méthodes déterministes.

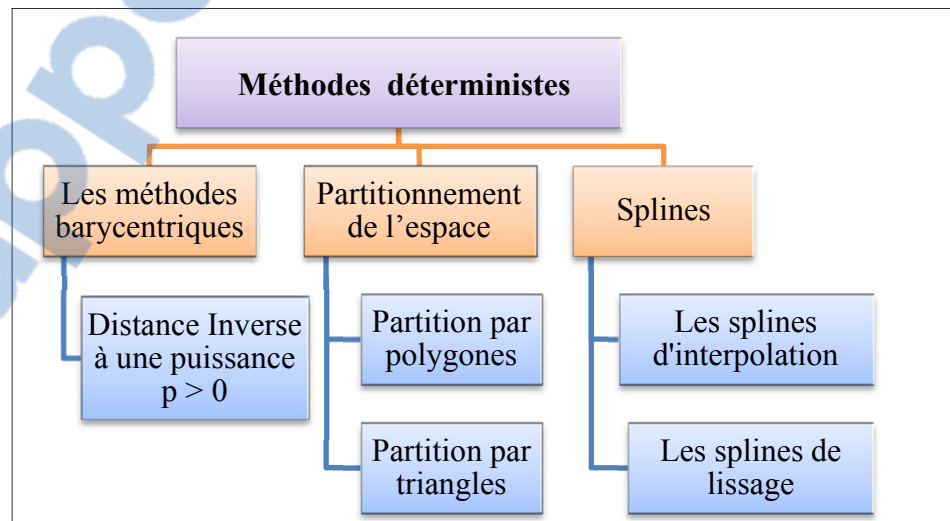


Figure 1-1 Résumé des méthodes d'interpolation déterministes

1.1.1 Méthodes barycentriques

Les méthodes barycentriques, dans d'autres versions on cite les moyennes mobiles (Ripley B.D 1981) ou encore l'approximation de Kernel (Christakos, 2001), sont très intuitives; elles estiment la valeur inconnue \hat{z} située à \mathbf{s}_0 par une simple moyenne pondérée des autres valeurs connues $z(\mathbf{s}_i)$:

$$\hat{z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i z(\mathbf{s}_i) \quad (1.1)$$

avec n est le nombre de données primaires, λ_i sont les poids qui dépendent de la distance euclidienne $h = \|\mathbf{h}\| = \|\mathbf{s}_i - \mathbf{s}_0\|$ entre la position de la valeur observée $z(\mathbf{s}_i)$ et la valeur à estimer $z(\mathbf{s}_0)$. Plus on s'éloigne de la valeur à estimer, plus le poids λ_i diminue.

Un cas particulier se présente pour calculer les λ_i : la méthode de l'inverse de distance à une puissance $p > 0$. On peut exprimer $\hat{z}(\mathbf{s}_0)$ comme une combinaison linéaire de plusieurs observations environnantes avec des poids inversement proportionnels au carré de la distance (Goovaerts, 2000b).

$$\hat{z}(\mathbf{s}_0) = \sum_{i=1}^{n_{\text{voisinage de } \mathbf{s}_0}} \frac{\frac{1}{\|\mathbf{s}_i - \mathbf{s}_0\|^p}}{\underbrace{\sum_{i=1}^{n_{\text{voisinage de } \mathbf{s}_0}} \frac{1}{\|\mathbf{s}_i - \mathbf{s}_0\|^p}}_{\lambda_i}} \times z(\mathbf{s}_i) \quad (1.2)$$

1.1.2 Méthodes d'interpolation par partitionnement de l'espace

Ces méthodes diffèrent des autres méthodes déterministes par l'utilisation d'une partition du champ d'études \mathbf{D} lors de la détermination des poids λ_i ainsi que le nombre nécessaire de mesures voisines du point estimé. Comme son nom l'indique, on doit positionner tout

d'abord les points cibles (les mesures expérimentales) dans l'espace, puis réaliser l'interpolation spatiale.

Pour le positionnement des points, il existe deux types de partitions d'échantillon en régions juxtaposées : une partition par polygones ou par triangles (Baillargeon, 2005).

Dans le premier cas, on définit un polygone pour chaque observation (mesure expérimentale) de façon que chaque point inclus, soit le plus proche du point de mesure que pour les autres polygones à côté. Une de ces méthodes est le partitionnement par polygone de Thiessen (Kergreis et Claude, 1991). Elle permet d'évaluer des moyennes spatiales sur un domaine à partir d'informations ponctuelles. La valeur de l'observation la plus proche est affectée au polygone défini par le vecteur \mathbf{s}_0 :

$$\hat{z}(\mathbf{s}_0) = z(\mathbf{s}_\alpha) \text{ avec } |\mathbf{s}_0 - \mathbf{s}_\beta| < |\mathbf{s}_0 - \mathbf{s}_\alpha| \quad \forall \alpha \neq \beta \quad (1.3)$$

En revanche, pour le deuxième cas, on aura une partition par triangulation où on découpe l'échantillon sous forme de triangles dont les sommets sont des points de mesure. Il suffit donc de relier le point de la mesure à estimer par les trois sommets du triangle qui le contient. Ensuite, le triangle principal est divisé en trois petits triangles. Comme on considère uniquement les valeurs mesurées qui interceptent le triangle de la valeur à estimer, le poids λ_i de chaque valeur mesurée est proportionnel à la surface d'intersection.

1.1.3 Splines

Les *splines* sont des fonctions qui minimisent l'énergie de flexion sous certaines contraintes d'ajustement (Dubrule, 1983). On distingue essentiellement deux types de splines : les splines d'interpolations qui obligent la courbe de passer par tous les points de mesures et les splines de lissages qui passent à proximité de toutes les mesures sans être forcées de passer par tous ces points (Matheron, 1975).

Les travaux de Hastie (Hastie, 1990) ont abouti à des fonctions qui permettent la représentation d'une spline dans l'espace bidimensionnel tout en contrôlant le passage à proximité des points des mesure de façon à minimiser la courbure. On note aussi l'existence de plusieurs variantes comme les splines multi-quadratiques de Hardy (Hardy, 1971) et les splines régularisées avec tension (Ripley B.D 1981). Finalement, on mentionne les splines laplaciennes étudiées également par Hastie et qui utilisent des sous-modèles linéaires pour effectuer une interpolation spatiale multivariable.

1.2 Méthodes probabilistes

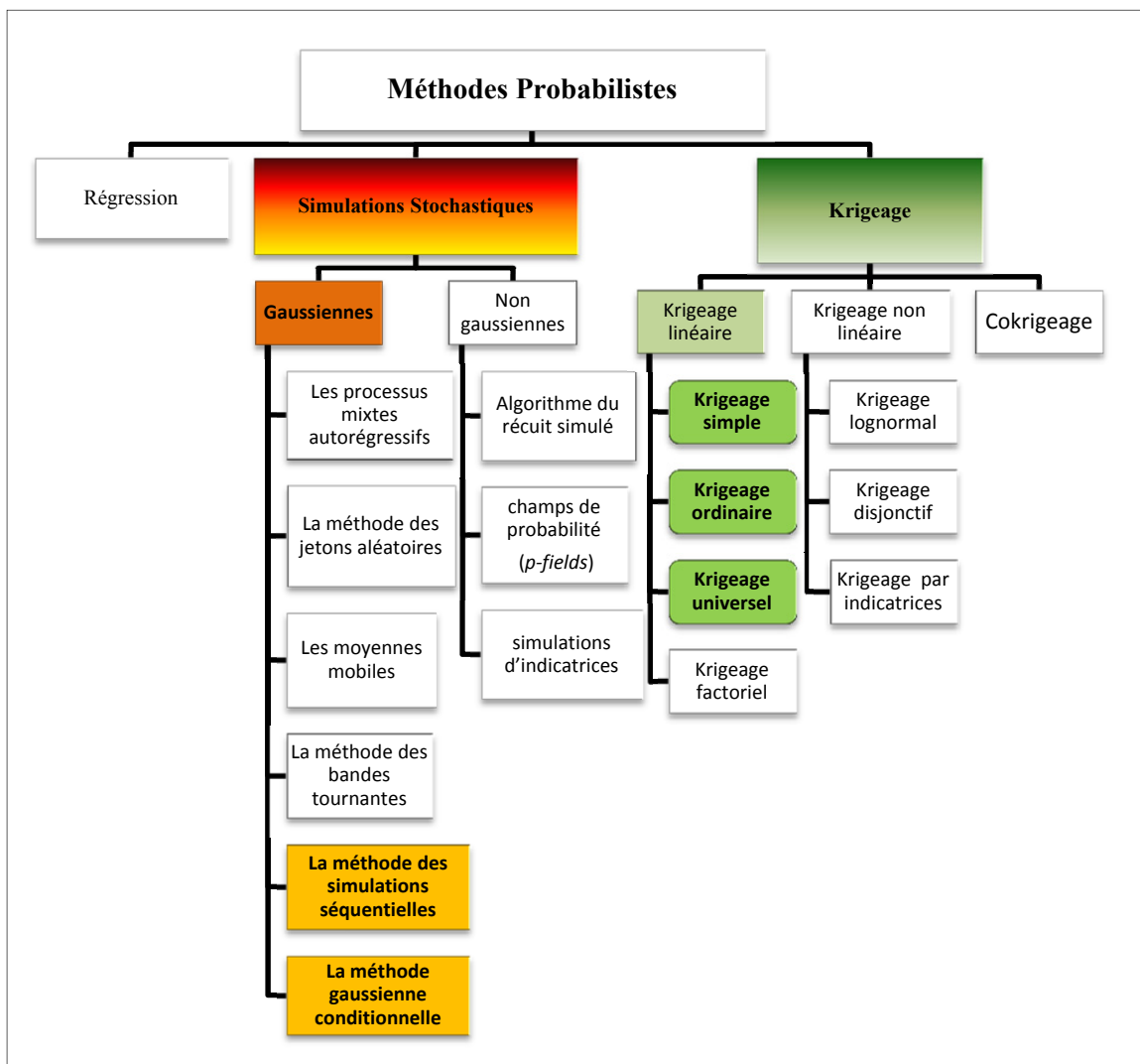


Figure 1-2 Les méthodes d'interpolation probabilistes

L'utilisation des modèles probabilistes, pour décrire le comportement d'une variable aléatoire permet d'évaluer l'incertitude de la valeur estimée. Parmi les méthodes recensées en littérature, on cite la régression classique, le krigeage et les simulations stochastiques, comme le montre la Figure 1-2.

1.2.1 Régression linéaire

La régression linéaire interpole des points en minimisant l'erreur quadratique entre un modèle a priori et les valeurs expérimentales, pour ajuster une surface par rapport aux valeurs mesurées. On suppose que la mesure est une variable aléatoire décomposée en deux parties :

$$Z(\mathbf{s}) = \underbrace{\mu(\mathbf{s})}_{\text{dérive}} + \underbrace{\epsilon(\mathbf{s})}_{\text{résidu}} \quad (1.4)$$

Avec $\mu(\mathbf{s})$ la valeur moyenne (la dérive) et $\epsilon(\mathbf{s})$ une fonction qui modélise les fluctuations autour de cette moyenne. Dans le cas de la régression classique, $\epsilon(\mathbf{s})$ est une fonction aléatoire normale d'espérance nulle et de variance constante. En d'autres termes, elle modélise, un bruit blanc présenté par des erreurs de mesure. La forme la plus utilisée est un polynôme de degré p :

$$\mu(\mathbf{s}) = \mu(x, y) = \sum_{l+k < p} \beta_{lk} x^l y^k \quad (1.5)$$

On estime les coefficients β_{lk} en minimisant une métrique de l'erreur par la méthode du maximum de vraisemblance ou les moindres carrés. On utilise une fonction de pondération proportionnelle au nombre des points du voisinage de façon à maximiser le poids des valeurs les plus proches.

1.2.2 Krigeage

Le krigeage est une méthode d'interpolation spatiale qui tient compte de la structure de dépendance spatiale des données (Trochu, 2003). Cette méthode a été proposée pour la première fois par l'ingénieur minier D.G. Krige (Krige, 1952), d'où découle le mot krigeage. Ensuite, des travaux ont été poursuivis pour développer toute la théorie du krigeage à l'École des Mines de Paris par le mathématicien français Matheron (Matheron, 1969). Pendant les années 1960-1970, cette méthode d'estimation était utilisée dans le domaine minier pour évaluer la teneur de gisements. Au cours des années 1980, on a utilisé le krigeage en cartographie numérique et en modélisation géométrique. À partir des années 1990, on avait commencé à appliquer le krigeage dans divers domaines : pétrolier, environnemental et mécanique (Bourgeois 2006).

Le krigeage se caractérise par une interpolation non biaisée tout en conservant une variance minimale. Les poids des coefficients de krigeage dépendent de la localisation des observations et surtout de leurs structures de dépendance spatiale. La structure de dépendance de la fonction aléatoire est déterminée à partir de l'étude variographique à travers le modèle du semivariogramme.

Plusieurs types de krigeage ont été étudiés en littérature (Cressie, 1990) : le krigeage simple (KS) avec une dérive constante et connue partout sur le domaine d'étude, le krigeage ordinaire (KO) avec une dérive constante et inconnue (supposée constante et connue juste au voisinage de la mesure à estimer) et finalement le krigeage universel (nommé aussi le krigeage avec tendance) (Bourennane et al., 1996) où la dérive n'est pas stationnaire.

De plus, on trouve le krigeage factoriel qui est utilisé pour séparer la composante de courte portée de celle de longue portée (Laporterie et al., 2000). Aussi, ce krigeage a été appliqué pour distinguer des anomalies géochimiques (Chilès et Delfiner, 2012) et pour filtrer les données radar et le renforcement des relations avec les caractéristiques du terrain et la topographie (Laporterie et al., 2000).

Il existe d'autres types de krigeage non linéaire qui appliquent des transformations non linéaires aux données initiales. On pourra citer par exemple le krigeage log-normal (Marcotte et Groleau, 1997) qui fait subir une transformation logarithmique aux données dans le but de travailler dans un espace quasi-linéaire. On trouve aussi le krigeage par indicatrice (Suro-Pérez et Journel, 1991) qui se distingue par une fonction de répartition représentée sous forme discrète. Chaque seuil de cette fonction de répartition est exprimé par une indicatrice et un variogramme différent, ce qui pourra augmenter la possibilité d'incohérence dans les modélisations puisqu'on perd de l'information en réduisant l'ensemble conditionnant. Dans ce cas, la modélisation du semivariogramme est complexe et entraîne par conséquent des simplifications abusives. Enfin, ce type de krigeage est utilisé pour des variables catégoriques où on contrôle des variables continues par des variables discrètes. Finalement, on trouve le krigeage disjonctif (Liao, 1990; Rivoirard, 1994) avec des transformations polynomiales spéciales et des lois bivariées isofactorielles.

Le krigeage peut aussi être utilisé dans des cas multi-variés qu'on désigne en littérature scientifique par le *cokrigeage* (Marcotte, 1991a). En effet, la variable estimée s'écrit sous la forme d'une combinaison linéaire pondérée à la variable primaire et des variables auxiliaires ou secondaires. (Goovaerts, 2000b) et (P. Bossier, 2011), montrent que pour estimer plus précisément la mesure recherchée dans le cas du cokrigeage, il faut minimiser l'erreur quadratique moyenne. En revanche, le cokrigeage fait appel à la construction d'un semivariogramme croisé, une tâche qui n'est pas évidente en général (Bourgault et Marcotte, 1991).

1.2.3 Simulations stochastiques

En variant la fonction aléatoire, on génère des simulations stochastiques qui restituent cette variabilité (Caers, 2011). Contrairement au krigeage qui n'utilise que les deux premiers moments (espérance et variogramme), la simulation va bien au-delà en simulant une fonction aléatoire entière. Les simulations sont nécessaires pour tout problème impliquant des transformations non linéaires des variables mesurées. En fait, elles sont intéressantes parce

qu'elles utilisent l'information inconnue à partir d'un ensemble à priori supposé connu. Il existe plusieurs techniques et algorithmes pour générer des simulations stochastiques conditionnelles. Parmi les méthodes utilisées, on distingue les approches qui nécessitent des transformations gaussiennes de celle non gaussiennes, qui ne nécessitent aucune transformation.

Parmi une panoplie de méthodes gaussiennes, certains génèrent des fonctions aléatoires conditionnées par les données mesurées en se basant sur la décomposition de Cholesky de la matrice de covariance, les moyennes mobiles, les méthodes autorégressives, les méthodes fréquentielles et la méthode des simulations séquentielles (Andre G. Journel 1998; Froidevaux, 2004; King, 2000).

La méthode de simulation matricielle (LU, Décomposition de Chlosky) est une méthode simple à programmer et elle est efficace pour simuler des petits champs d'étude. Elle est utilisée pour les simulations conditionnelles ainsi que non conditionnelles (Webster, 2007). Cette méthode est rapide, car elle nécessite une seule inversion de matrice. En revanche, sa matrice de covariance doit être non singulière pour pouvoir effectuer la décomposition de Cholesky (une mesure à simuler ne doit pas se coïncider avec une observation).

La méthode des bandes tournantes (Matheron, 1969) peut être vue comme une généralisation de la méthode spectrale utilisée par (Alfaro-Sironvalle 1979) pour simuler des champs gaussiens. Elle présente l'avantage de pouvoir s'appliquer à des champs aussi bien stationnaires qu'à des accroissements stationnaires, définis dans des espaces de dimension quelconque, et à des domaines de simulation de taille et de forme arbitraires. Malgré que cette méthode soit la plus efficace en espace mémoire et temps de calcul, elle nécessite un post conditionnement, une étude de l'anisotropie et la programmation de plusieurs structures imbriquées.

Les moyennes mobiles (Journel A.G., 1981) et l'analyse spectrale (Mantoglou et Wilson, 1982) sont des approches directes qui reposent sur la factorisation de la fonction de

covariance et sur le théorème de Bochner (Bochner 1960). Ces méthodes nécessitent la méthode des bandes tournantes (Matheron 1973) lors de la modélisation dans \mathcal{R}^n .

Une deuxième catégorie s'ajoute; la catégorie des méthodes géométriques telles que la méthode des jetons aléatoires (Alfaro-Sironvalle 1979; Journel A.G. , 1981) et la méthode des partitions aléatoires de l'espace (Matheron 1975). De plus, on peut citer les méthodes de type séries spatiales telles que les processus mixtes autorégressifs à moyennes mobiles (Box 1970).

Toutes ces méthodes ont un certain nombre de limitations notamment sur la détermination explicite de la décomposition de la covariance (Chilès 1977), le non-respect du comportement à l'origine des transformées de Fourier pour les méthodes d'analyse spectrale et la discrétisation des bandes tournantes et finalement la difficulté de détermination des paramètres du modèle (Graupe D., 1975).

D'autre part, il y a les simulations non gaussiennes, comme les méthodes de simulations d'indicatrices (Andre G. Journel 1998), les méthodes utilisant des champs de probabilité (p-fields) (Goovaerts, 2002; Journel et Ying, 2001) et l'algorithme du recuit simulé (Andre G. Journel 1998).

Les simulations de variables indicatrices (catégories) se rencontre souvent lorsqu'on a une variable discontinue où l'indicatrice peut représenter la présence ou l'absence de la variable mesurée. Dans notre cas d'étude, la variable est supposée continue, donc on ne considérera pas ce type de simulations. D'une autre part, la méthode de recuit simulé (Deutsch et Cockerham, 1994; van Groenigen et Stein, 1998) est une méthode itérative basée sur l'optimisation d'une fonction objective. Dans ce cas, on cherche à minimiser l'écart entre le variogramme théorique (respectivement la distribution théorique) et le variogramme simulé (respectivement la distribution simulée). Malgré que cette méthode permette de traiter les données non conditionnelles, elle a plusieurs inconvénients. Il est difficile de fixer la cédule

(valeur cible) et c'est une méthode lente, car on modifie juste un point à la fois lors de la simulation.

En conclusion de cette revue concise de la littérature, on utilisera uniquement quelques méthodes d'interpolation dans le cadre de ce projet et non pas toutes. En effet, les contraintes temporelles d'un projet de maîtrise dans un premier temps et par inadéquation d'application pour certaines méthodes dans un deuxième temps nous ont dicté nos choix de méthodes. Les méthodes déterministes sont des méthodes trop approximatives et ne donnent pas d'indication sur l'incertitude de la mesure prédite. Par conséquent, nous n'en avons retenu aucune. En revanche, des méthodes telles que le krigeage et les simulations stochastiques semblent être les plus adéquates et sont parmi les plus utilisées. Par conséquent, nous les avons retenues pour développer notre méthodologie.

CHAPITRE 2

CADRE THÉORIQUE

L'existence d'un modèle de dépendance spatiale permet l'estimation d'une variable dans une localisation non échantillonnée. Le Chapitre 2 présente le développement théorique du krigeage ainsi que les simulations stochastiques pour une variable continue dans l'espace. Tout d'abord, on présentera les modèles mathématiques des trois principaux types de krigeage : krigeage simple (KS), krigeage ordinaire (KO) et krigeage universel (KU) (krigeage avec un modèle de tendance) ainsi leurs algorithmes. Ensuite, on présentera les deux types de simulations conditionnelles pour déterminer les variances des estimations données par le krigeage.

2.1 Variable aléatoire et fonction aléatoire

Le concept de variable régionalisée se base sur une fonction aléatoire dans un domaine d'étude D de l'espace. En fait, on introduit le concept d'une variable aléatoire pour modéliser l'incertitude sur la mesure. Les valeurs de sortie possibles d'une variable aléatoire sont contrôlées par la fonction de distribution de probabilité *pdf*. Ce concept d'une fonction aléatoire est la base du modèle d'interpolation spatiale utilisé pour prédire les valeurs manquantes du domaine d'étude ainsi que leurs incertitudes. Une variable déterministe prend juste une valeur de sortie, par contre une valeur probabiliste peut prendre plusieurs valeurs de sortie avec une certaine probabilité attribuée à chacune d'elles.

Dans le cas discret, chaque valeur de sortie est attachée à une probabilité

$$p_i = \text{Prob}(Z = z_i)$$
$$\text{avec } \sum_{i=1}^n p_i = 1 \quad (2.1)$$

Pour une variable continue, la distribution de probabilité sera déterminée soit par la fonction de distribution cumulative *cdf* ou par la fonction de densité de probabilité *pdf* :

- La fonction de distribution cumulative *cdf* représente la probabilité que la variable aléatoire Z ne dépasse pas une valeur seuil z donnée par

$$\mathcal{F}(z) = \text{Prob}\{Z \leq z\} \quad (2.2)$$

- La fonction de densité de probabilité *pdf* est définie par la pente $f(z) = \frac{d\mathcal{F}(z)}{dz}$ quand \mathcal{F} est dérivable.

Ainsi à partir de la *cdf* (ou la *pdf*), on peut déterminer la probabilité de la valeur qu'une variable aléatoire peut prendre dans un intervalle $[a,b]$:

$$\text{Prob}\{a \leq Z \leq b\} = \mathcal{F}(b) - \mathcal{F}(a) = \int_a^b f(z) dz \quad (2.3)$$

La clé pour l'interprétation probabiliste d'une variable aléatoire est la modélisation de sa *cdf*. L'espérance $E\{z\} = m$, où la valeur la plus probable de la variable aléatoire, peut être considérée comme la valeur permettant de minimiser l'erreur par la méthode des moindres carrés. Cette moyenne peut être déterminée à partir de la *cdf* comme une moyenne arithmétique de L réalisations équiprobables

$$m = E\{z\} = \int_{z_{min}}^{z_{max}} z f(z) dz = \frac{1}{L} \sum_{l=1}^L z^{(l)} \quad (2.4)$$

La variance, un estimateur de la dispersion centrée sur la moyenne estimée m , sera donnée par

$$\begin{aligned}\sigma^2 &= \text{Var}\{Z\} = E\{(Z - m)^2\} = \int_{z_{\min}}^{z_{\max}} (z - m)^2 f(z) dz \\ &= \frac{1}{L} \sum_{l=1}^L (z^{(l)} - m)^2\end{aligned}\tag{2.5}$$

2.2 Semivariogramme

Le krigeage repose sur le fait que la nature n'est pas entièrement imprévisible : deux observations situées l'une proche de l'autre devraient en moyenne se ressembler davantage plus que deux observations éloignées, puisque la structure spatiale de la variable étudiée n'est ni chaotique ni aléatoire.

Par définition, un semivariogramme expérimental est un graphique montrant comment la moitié des différences au carré entre des paires des observations, varie en fonction de la distance qui les sépare. En général, on s'attend à voir des accroissements des variances sur des courtes distances puis une variance stable (égale à la variance globale de l'échantillon) sur des distances plus longues.

Le semivariogramme expérimental est un outil pour l'analyse statistique spatiale des données en se basant sur la mesure de la dissimilarité. On part tout d'abord d'un plan d'implantation (le domaine \mathbf{D} et les n points de mesures expérimentales), on détermine la valeur de la variable $z(\mathbf{s})$ aux différents n endroits de mesures et on calcule la variance entre chaque paire $[z(\mathbf{s}_\alpha) ; z(\mathbf{s}_\beta), \forall \alpha \in \mathbf{D}, \forall \beta \in \mathbf{D}]$. Pour chaque paire de mesures, les valeurs sont séparées spatialement par un vecteur d'incrément $\mathbf{h} = \mathbf{s}_\alpha - \mathbf{s}_\beta$ comme le montre la Figure 2-1.

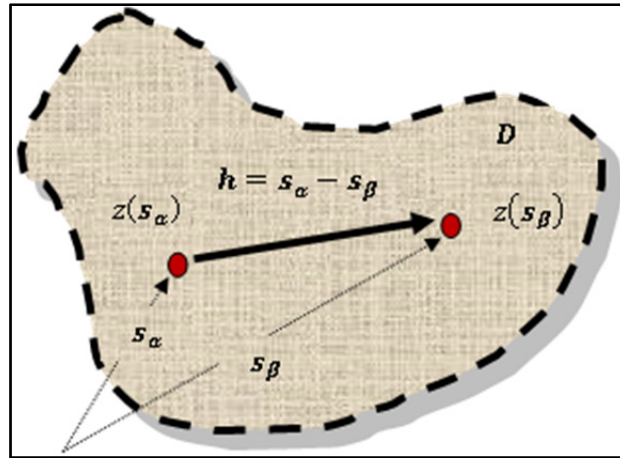


Figure 2-1 Présentation schématique de l'incrément \mathbf{h}

En partant de la définition de la variance,

$$\begin{aligned}
 \text{Var} (Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})) &= E [\{Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})\}^2] \\
 &= E [Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})]^2 \\
 &= \underbrace{E[Z(\mathbf{s})]}_{=0} - \underbrace{E[Z(\mathbf{s} + \mathbf{h})]}_{=0}
 \end{aligned} \tag{2.6}$$

et sachant que la fonction aléatoire vérifie l'hypothèse intrinsèque, on pourra définir :

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{Var} (Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})) \tag{2.7}$$

D'après (2.6) et (2.7), (Webster, 2007), on déduit que,

$$\gamma(\mathbf{h}) = E [\{Z(\mathbf{s}) - Z(\mathbf{s} + \mathbf{h})\}^2] \tag{2.8}$$

Ainsi, on définit l'expression caractérisant le semivariogramme expérimental qui sera proportionnelle à la demi-variance de la différence entre les mesures expérimentales par :

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{(i,j) \in N(h)} [z(\mathbf{s}_i) - z(\mathbf{s}_j)]^2 \quad (2.9)$$

avec $N(\mathbf{h}) = \{(i, j) \mid \forall i \neq j, \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$ l'ensemble des paires des points distincts.

Il est important de mentionner que l'équation (2.9) est sensible aux valeurs aberrantes et en particulier les valeurs extrêmes, ce qui rend cet estimateur moins robuste, car il n'est pas capable de reproduire ces dernières. Ce problème est résolu par des estimateurs probabilistes qui ne seront pas traités dans ce mémoire et pour plus de détails il suffit de consulter (Cressie, 1993). Dans notre cas d'étude, on se limitera à des transformations gaussiennes, pour assurer l'hypothèse de la stationnarité de la fonction aléatoire. En effet, on réalisera dans un premier temps une transformation dans un espace gaussien où on effectuera nos simulations puis une transformation inverse permet de replacer l'échantillon dans l'espace réel initial. Par cette transformation gaussienne, on réduit la variance globale et par la même occasion l'impact des valeurs extrêmes comparativement aux autres.

On désigne « analyse variographique » la déduction du semivariogramme empirique (théorique) à partir du semivariogramme expérimental. Un semivariogramme est une fonction paire, positive et bornée. C'est une fonction définie positive conditionnelle et vérifie l'équation suivante (Gooverts, 1997):

$$\begin{aligned} \forall k \in N, \forall v_1, \dots, v_k \in R^k \sum_{i=1}^k v_i &= 0, \\ \forall \mathbf{s}_1, \dots, \mathbf{s}_k \in D^k, & \\ \sum_{i=1}^n \sum_{j=1}^n v_i v_j \gamma(\mathbf{s}_i - \mathbf{s}_j) &\leq 0 \end{aligned} \quad (2.10)$$

2.2.1 Construction du semivariogramme expérimental

Après avoir collecté les mesures expérimentales dans le domaine D , on pourra initier la construction du semivariogramme expérimental. Pour ce faire, on rassemble les mesures de variances dans des classes en fonction de la distance et de l'angle d'orientation de la droite qui lie chaque paire de mesures expérimentales (Figure 2-2). La notation vectorielle tient compte de la dépendance de la longueur de l'incrément \mathbf{h} et de l'orientation de la ligne passante par la paire des points qui la relie.

À partir du plan d'implantation de n points de mesures, on pourra représenter une nuée variographique de $n \times \frac{n-1}{2}$ points. Le résultat sera représenté en fonction du module de la séparation vectorielle \mathbf{h} , ensuite le nuage de points sera divisé des classes en fonction du pas de calcul (incrément) $h = \|\mathbf{h}\|$. Finalement chaque classe de points sera représentée par la moyenne de toute la classe et le semivariogramme expérimental est obtenu tout simplement en reliant l'ensemble de ces moyennes tel que montré à la Figure 2-2.

L'estimation de la partie manquante entre les points représentatifs des moyennes des incréments h , engendre la connaissance du semivariogramme à n'importe quel incrément arbitraire h . D'où la nécessité de la modélisation du semivariogramme expérimental et la représentation de ce dernier par un modèle analytique qui permet de déterminer les valeurs du semivariogramme $\gamma(\mathbf{h})$ pour n'importe quel incrément et non uniquement ceux qui sont définis par les mesures expérimentales. De plus, le modèle analytique est un moyen de filtrer le bruit causé par le semivariogramme expérimental généré entre autre par l'imperfection des mesures et parfois par le manque de ces données (Trauth, 2010).

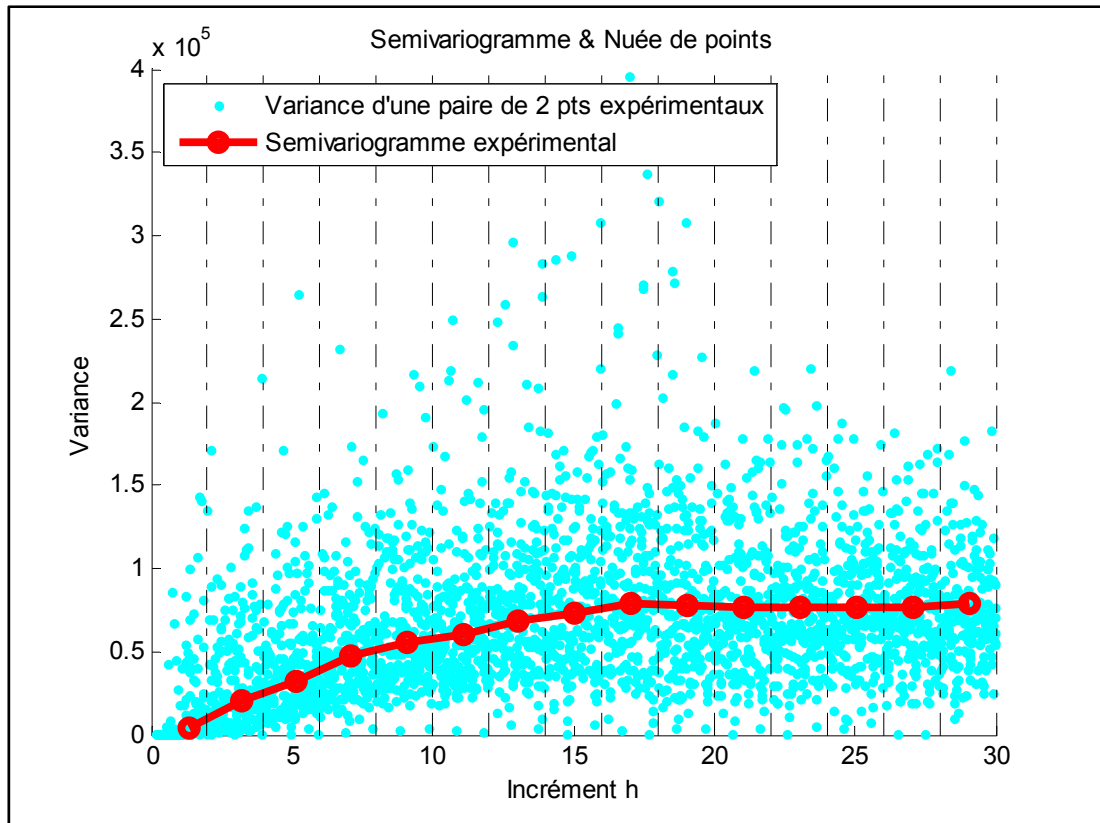


Figure 2-2 Traçage du semivariogramme expérimental

2.2.2 Modélisation du semivariogramme expérimental

Le modèle analytique doit garantir une variance positive pour n'importe quelle combinaison des points de l'échantillon en passant par (ou proche de) la séquence des points formant le semivariogramme expérimental. Néanmoins, si on utilise uniquement les valeurs données par le semivariogramme expérimental non ajusté lors de la résolution du système des équations de krigeage, on risque d'avoir des variances négatives et par conséquent les estimations des valeurs de la variable étudiée dans des points non échantillonnés seront invalides.

Le modèle du semivariogramme est une expression analytique d'une fonction qui tente de passer par le maximum des points du semivariogramme expérimental en respectant la condition de minimisation par les moindres carrés et la positivité conditionnelle.

Pour chaque semivariogramme présentant un seuil (un palier) on peut définir :

- Une portée a : La distance à partir de laquelle deux observations ne se ressembleront plus en moyenne et qu'il n'y aura aucune corrélation qui les relie linéairement (covariance nulle). À partir de cet endroit, on pourra déduire la variance de la variable aléatoire.
- Palier $C+C_0$: Cette partie du semivariogramme définit la variance moyenne de la variable aléatoire. Il est atteint généralement au niveau de la portée.
- Effet de pépité C_0 : Ceci représente la variation à très courte échelle et les erreurs lors de la prise de la mesure expérimentale. L'effet de pépité se présente comme une discontinuité à l'origine du semivariogramme, car par définition $\gamma(0) = 0$ lorsque $h = 0$ et $\gamma(h) = C_0$ lorsque $h = 0^+$.

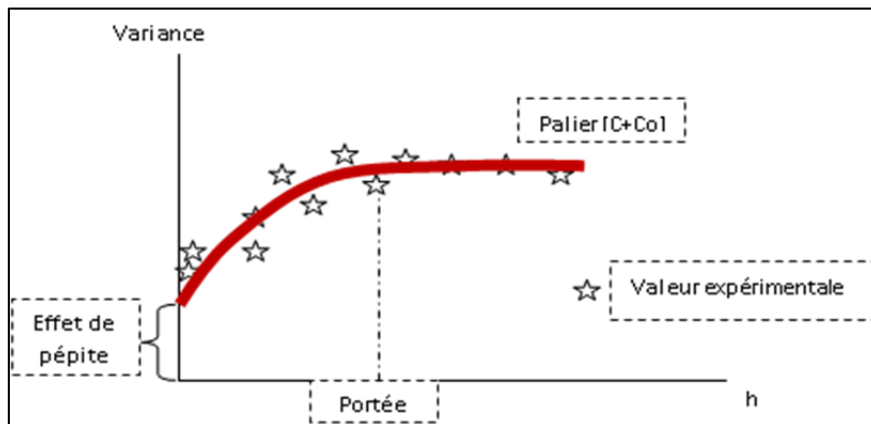


Figure 2-3 Les composantes d'un semivariogramme

Il existe un nombre important de modèles de semivariogramme, mais dans se mémoire on se limitera aux modèles suivants :

- Modèle gaussien : $\gamma(h) = C [1 - \exp(-3 (\frac{h}{a})^2)]$
- Modèle exponentiel : $\gamma(h) = C [1 - \exp(-3 \frac{h}{a})]$
- Modèle sphérique : $\gamma(h) = \begin{cases} C [\frac{3h}{2a} - \frac{1}{2} (\frac{h}{a})^3] & \text{si } 0 < h < a \\ C & \text{si } h \geq a \end{cases}$

Ces modèles ainsi que leurs expressions analytiques sont détaillés dans les trois livres de références en géostatistiques (Chilès et Delfiner, 2012; Deutsch et Journel, 1998; Gooverts, 1997; Journel A.G. , 1981).

Le comportement à l'origine est crucial pour le choix du modèle de semivariogramme théorique lors de sa modélisation. Ce comportement reflète le degré de régularité spatiale de la variable à estimer. On trouve généralement 3 types : parabolique, linéaire et discontinu. Pour le premier cas, ceci s'explique par une grande régularité spatiale (aspect lisse), pour le deuxième cas, on a un comportement moins régulier, mais continu et dérivable, et pour le comportement discontinu, on aura un saut à l'origine et ceci est expliqué par l'absence partielle ou totale de corrélation entre deux valeurs proches.

Comme le montre la Figure 2-4, le modèle sphérique atteint le palier exactement au point de la portée. En revanche les modèles exponentiel et gaussien rejoignent la portée asymptotiquement quand la variance atteint 95% de la valeur du palier. De plus, le modèle gaussien a un comportement parabolique à l'origine qui représente le lissage de la propriété krigée. Les modèles sphérique et exponentiel montrent à leurs tours un comportement quasi-linéaire à l'origine, ces deux modèles peuvent être utilisés lorsqu'il s'agit d'une propriété trop variable sur un petit champ de variation.

Généralement, une anomalie (non-homogénéité) peut être détectée si on a une dissimilarité avec des courtes distances (modèle d'un semivariogramme avec un effet de pépite). En effet, pour la plupart de modèles, on observe une dissimilarité moyenne qui augmente quand l'espacement entre les paires de l'échantillon augmente à son tour contrairement au modèle

effet de pépite qui décolle brutalement à partir de l'origine et montre une forte dissimilarité avec des faibles distances de séparation.

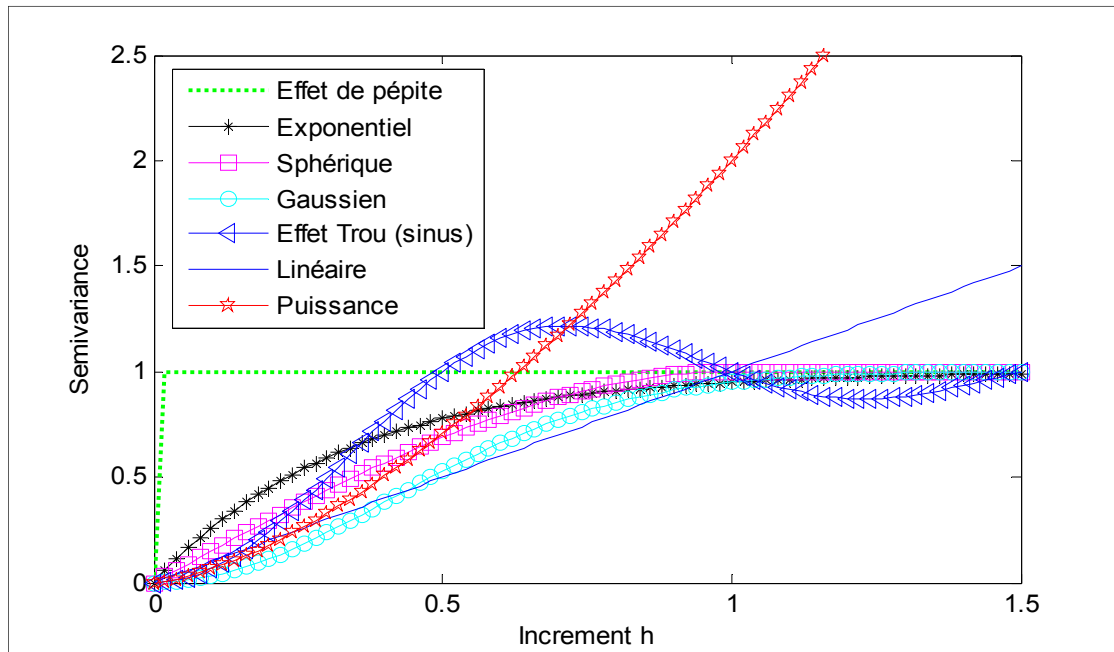


Figure 2-4 Comparaison entre plusieurs modèles des semivariogrammes

2.3 Krigeage

Le krigeage est une méthode optimale d'interpolation basée sur la régression des mesures observées autour de la mesure à estimer et dont le poids de chacune sera déterminé en se basant sur sa covariance spatiale avec les autres mesures. L'estimation par krigeage permet de prédire les mesures manquantes et permet de donner une idée sur l'allure de la propagation de l'erreur locale de la prédiction.

Tous les algorithmes de simulation énumérés dans le chapitre précédent estiment la valeur en se basant sur le moyennage des valeurs voisines du point cible sans minimiser l'erreur sur les poids de la moyenne pondérée. En fait, on assigne le poids suivant le principe : plus on s'éloigne de la valeur recherchée plus le poids diminue. En revanche, le krigeage tient compte de toutes les mesures de l'échantillon lors de l'estimation tout en respectant une variance d'estimation minimale (Gratton, 2002).

Toutes les versions de krigeage reposent sur une régression linéaire généralisée :

$$\hat{Z}(\mathbf{s}) - m(\mathbf{s}) = \sum_{\alpha=1}^{n(\mathbf{s})} \lambda_{\alpha} [Z(\mathbf{s}_{\alpha}) - m(\mathbf{s}_{\alpha})] \quad (2.11)$$

avec $Z(\mathbf{s}_{\alpha})$ est une variable aléatoire qui représente l'attribut à la localisation \mathbf{s}_{α} , $m(\mathbf{s}) = E\{Z(\mathbf{s})\}$ est son espérance et $\hat{Z}(\mathbf{s})$ son estimateur linéaire.

Une fois qu'on a modélisé le variogramme expérimental, on peut calculer la variance d'estimation associée à cet estimateur et déterminer les λ_{α} qui minimisent la variance d'estimation tout en assurant une absence de biais.

2.3.1 Krigeage simple (KS)

L'estimateur de la mesure à la localisation cible \mathbf{s}_0 est donné par la combinaison linéaire des observations des autres localisations et la moyenne m (connue et constante sur tout le domaine d'étude D).

$$\hat{Z}(\mathbf{s}_0) = m + \sum_{i=1}^n \lambda_i (Z(\mathbf{s}_i) - m) \quad (2.12)$$

Le but est de déterminer les poids λ_i qui minimisent l'erreur quadratique moyenne de la variance d'estimation (Stein, 1999) définit en \mathbf{s}_0 par :

$$\sigma_{SK}^2(\mathbf{s}_0) = \frac{Cov(\mathbf{s}_0, \mathbf{s}_0)}{Var(\mathbf{s}_0)} + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov(\mathbf{s}_i, \mathbf{s}_j) - 2 \sum_{i=1}^n \lambda_i Cov(\mathbf{s}_i, \mathbf{s}_0) \quad (2.13)$$

Les poids optimaux de l'estimateur (2.12) seront représentés par le vecteur Λ :

$$\Lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}_{(n)} \quad (2.17)$$

3) On résout le système linéaire suivant :

$$\mathbf{C}_{(n \times n)} \Lambda_{(n \times 1)} = \mathbf{c}_{(n \times 1)} \quad (2.18)$$

Étant donné que la moyenne du krigeage simple est constante, on définit le vecteur suivant :

$$Y = \begin{pmatrix} s_1 - m \\ s_2 - m \\ \vdots \\ s_n - m \end{pmatrix}_{(n)} \quad (2.19)$$

4) À partir de ce qui précède, on pourra déduire la valeur de l'estimation ainsi que la variance du krigeage à n'importe quel point du domaine \mathbf{D} :

$$\widehat{Z}_{SK}(\mathbf{s}_0) = m + Y' \Lambda = m + Y' \mathbf{C}^{-1} \mathbf{c} \quad (2.20)$$

$$\sigma_{SK}^2(\mathbf{s}_0) = Cov(0) - \mathbf{c}' \Lambda = Cov(0) - \mathbf{c}' \mathbf{C}^{-1} \mathbf{c} \quad (2.21)$$

2.3.2 Krigeage ordinaire (KO)

On a vu que le krigeage simple requiert la connaissance de la moyenne pour déterminer les poids de l'estimateur et pour estimer sa variance. En revanche, le krigeage ordinaire franchi ce requit. En réarrangeant l'expression de l'estimateur du krigeage simple, on obtient :

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{s}_i) + m \left(1 - \sum_{i=1}^n \lambda_i \right) \quad (2.22)$$

À partir de cette nouvelle forme, on observe que si la somme des poids sera égale à l'unité, alors l'estimateur est indépendant de la moyenne et c'est ici qu'on aperçoit l'élégance de ce type de krigeage. Par conséquence, en enlevant la moyenne de la formule de l'estimateur, tout ce qui suit notamment pour le calcul des poids et la variance sera indépendant de la moyenne.

La résolution d'un système de krigeage ordinaire est un problème d'optimisation. On doit trouver les poids en minimisant l'erreur quadratique moyenne. La résolution de ce problème d'optimisation se réduit à une résolution de la fonction de Lagrange définit par :

$$L(\lambda_1, \lambda_2, \dots, \lambda_k; \mu) = \sigma_{OK}^2(\mathbf{s}_0) + 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (2.23)$$

où μ est le multiplicateur de Lagrange.

Tout en assurant que le semivariogramme est défini négatif (une covariance définit positive), l'unique variance minimale de l'estimation sera la solution de ce système :

$$\left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i = 1 \\ \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_1) - \mu = \gamma(\mathbf{s}_0, \mathbf{s}_1) \\ \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_2) - \mu = \gamma(\mathbf{s}_0, \mathbf{s}_2) \\ \vdots \\ \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_n) - \mu = \gamma(\mathbf{s}_0, \mathbf{s}_n) \end{array} \right. \quad (2.24)$$

En multipliant ce système d'équations par les poids respectifs λ_i , et en réalisant une simple sommation on obtiendra :

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(\mathbf{s}_i, \mathbf{s}_j) = \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_0) + \mu \quad (2.25)$$

L'erreur quadratique moyenne du krigeage ordinaire sera donnée par :

$$\sigma_{KO}^2(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_0) - \mu \quad (2.26)$$

Algorithme du krigeage ordinaire KO)

1. On calcule chaque élément de la matrice de covariance \mathbf{G} définit par :

$$\mathbf{G} = \begin{bmatrix} \gamma(\mathbf{s}_1, \mathbf{s}_1) & \gamma(\mathbf{s}_2, \mathbf{s}_1) & \cdots & \gamma(\mathbf{s}_n, \mathbf{s}_1) & 1 \\ \gamma(\mathbf{s}_1, \mathbf{s}_2) & \gamma(\mathbf{s}_2, \mathbf{s}_2) & \cdots & \gamma(\mathbf{s}_n, \mathbf{s}_2) & 1 \\ \vdots & \vdots & \ddots & \vdots & \cdots \\ \gamma(\mathbf{s}_1, \mathbf{s}_n) & \gamma(\mathbf{s}_2, \mathbf{s}_n) & \cdots & \gamma(\mathbf{s}_n, \mathbf{s}_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}_{(n+1) \times (n+1)} \quad (2.27)$$

2. On calcule ensuite la semivariance entre le point à estimer et chaque point de l'échantillon des mesures, on définit le vecteur \mathbf{g} par :

$$\mathbf{g} = \begin{pmatrix} \gamma(\mathbf{s}_0, \mathbf{s}_1) \\ \gamma(\mathbf{s}_0, \mathbf{s}_2) \\ \vdots \\ \gamma(\mathbf{s}_0, \mathbf{s}_n) \\ 1 \end{pmatrix}_{(n+1)} \quad (2.28)$$

Les poids optimaux de l'estimateur seront représentés par le vecteur \mathbf{W} :

$$\mathbf{W} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ -\mu \end{pmatrix}_{(n+1)} \quad (2.29)$$

3. On résout le système linéaire suivant :

$$\mathbf{GW} = \mathbf{g} \quad (2.30)$$

On définit l'ensemble des variables aléatoires $Z(\mathbf{s}_i)$ aux sites des observations qui formeront un vecteur aléatoire Z :

$$\mathbf{Z} = \begin{pmatrix} Z(\mathbf{s}_1) \\ Z(\mathbf{s}_2) \\ \vdots \\ Z(\mathbf{s}_n) \\ 0 \end{pmatrix}_{(n+1)} \quad (2.31)$$

4. L'estimation ainsi que sa variance sont :

$$\widehat{Z}_{OK}(\mathbf{s}_0) = \mathbf{Z}'\mathbf{W} = m + \mathbf{Z}'\mathbf{G}^{-1}\mathbf{g} \quad (2.32)$$

$$\sigma_{OK}^2(\mathbf{s}_0) = \mathbf{g}'\mathbf{W} = \mathbf{g}'\mathbf{G}^{-1}\mathbf{g} \quad (2.33)$$

2.3.3 Krigeage universel (KU)

L'estimation locale de la moyenne avec le krigeage ordinaire, considère implicitement une modélisation d'une fonction aléatoire stationnaire juste au voisinage du point à estimer. Dans certains cas, supposer que l'hypothèse de stationnarité soit respectée, n'est pas une approche appropriée même pour un voisinage petit. D'où l'idée du krigeage universel (krigeage par tendance) où la tendance locale du voisinage est modélisée comme une fonction qui varie de façon lisse (Armstrong, 1998; Matheron, 1969).

Le but du krigeage universel, comme pour le KS et le KO, est de trouver les poids λ_i optimaux. Sachant que l'estimateur est non biaisé, qu'il a une erreur quadratique moyenne minimale et en se basant sur le développement mathématique détaillé (Hillier et Lieberman, 2010) en utilisant les multiplicateurs de Lagrange, on définit la fonction lagrangienne suivante pour le krigeage universel.

$$L(\lambda_i; \mu) = \sigma^2(\mathbf{s}_0) + 2\mu_0 \left(\sum_{i=1}^k \lambda_i - 1 \right) + 2 \sum_{l=1}^n \mu_l \left(\sum_{i=1}^k \lambda_i f_l(\mathbf{s}_i) - f_l(\mathbf{s}_0) \right) \quad (2.34)$$

où μ_l sont les multiplicateurs de Lagrange et λ_i les poids de l'estimateur qui produisent la variance d'estimation minimale. Ceux-ci seront obtenus en solutionnant le système (2.35).

$$\left\{ \begin{array}{l}
\sum_{i=1}^k \lambda_i \gamma_R(\mathbf{s}_i, \mathbf{s}_1) - \mu_0 - \sum_{i=1}^n \mu_i f_i(\mathbf{s}_1) = \gamma_R(\mathbf{s}_0, \mathbf{s}_1) \\
\sum_{i=1}^k \lambda_i \gamma_R(\mathbf{s}_i, \mathbf{s}_2) - \mu_0 - \sum_{i=1}^n \mu_i f_i(\mathbf{s}_2) = \gamma_R(\mathbf{s}_0, \mathbf{s}_2) \\
\vdots \\
\sum_{i=1}^k \lambda_i \gamma_R(\mathbf{s}_i, \mathbf{s}_k) - \mu_0 - \sum_{i=1}^n \mu_i f_i(\mathbf{s}_k) = \gamma_R(\mathbf{s}_0, \mathbf{s}_k) \\
\sum_{i=1}^n \lambda_i = 1 \\
\sum_{i=1}^k \lambda_i f_1(\mathbf{s}) = f_1(\mathbf{s}) \\
\sum_{i=1}^k \lambda_i f_2(\mathbf{s}) = f_2(\mathbf{s}) \\
\vdots \\
\sum_{i=1}^k \lambda_i f_n(\mathbf{s}) = f_n(\mathbf{s})
\end{array} \right. \quad (2.35)$$

On peut réécrire ce système en fonction de la covariance du résidu pour obtenir :

$$\left\{ \begin{array}{l}
\sum_{\beta=1}^n \lambda_{\beta}^{(KT)}(\mathbf{s}) \mathbf{C}_R(\mathbf{s}_{\beta} - \mathbf{s}_{\alpha}) + \sum_{k=1}^K \mu_k(\mathbf{s}) \mathbf{f}_k(\mathbf{s}_{\alpha}) = \mathbf{C}_R(\mathbf{s} - \mathbf{s}_{\alpha}) \\
\sum_{\beta=1}^n \lambda_{\beta}^{(KT)}(\mathbf{s}) \mathbf{f}_k(\mathbf{s}_{\alpha}) = \mathbf{f}_k(\mathbf{s}) \quad ; \\
\sum_{\beta=1}^n \lambda_{\beta}^{(KT)}(\mathbf{s}) = \mathbf{1}
\end{array} \right. \quad (2.36)$$

avec $\mathbf{C}_R(\cdot)$ la covariance du résidu. La variance d'erreur minimale sera donnée par :

$$\sigma_{0K}^2(\mathbf{s}) = \mathbf{C}_R(\mathbf{0}) - \sum_{\alpha=1}^n \lambda_{\alpha}^{KT} \mathbf{C}_R(\mathbf{s}_{\alpha} - \mathbf{s}) - \sum_{k=0}^K \mu_k^{KT} \mathbf{f}_k(\mathbf{s}) \quad (2.37)$$

On remarque avec les expressions (2.35), (2.36) et (2.37); qu'on ne s'occupe plus de la fonction de covariance de la fonction aléatoire qui représente l'attribut principal mais de son résidu. Ainsi, le krigeage avec un modèle de tendance requiert d'une part la détermination des fonctions de tendance $f_k(\mathbf{s})$ et de la fonction de covariance C_R .

Modélisation de la tendance

Les fonctions décrivant la tendance $f_k(\mathbf{s})$ sont liées à la physique du problème étudié. Le calcul de la covariance résiduelle n'est pas évident car on a les attributs (valeurs primaires) et non pas les valeurs des résidus pour tracer le semivariogramme expérimental γ_R et déduire par la suite les covariances résiduelles C_R . En revanche les deux semivariogrammes sont liés ensemble par la relation suivante :

$$2\gamma(\mathbf{h}) = E \left[\left(\underbrace{Z(\mathbf{s})}_{m(\mathbf{s})+R(\mathbf{s})} - \underbrace{Z(\mathbf{s}+\mathbf{h})}_{m(\mathbf{s}+\mathbf{h})+R(\mathbf{s}+\mathbf{h})} \right)^2 \right] \quad (2.38)$$

et en appliquant la définition de l'espérance mathématique, on déduit que,

$$\gamma(\mathbf{h}) = \gamma_R(\mathbf{h}) + \frac{1}{2} \left(\underbrace{m(\mathbf{s}) - m(\mathbf{s}+\mathbf{h})}_{\text{inconnu}} \right)^2 \quad (2.39)$$

Algorithme du Krigeage universel(KU)

On considère la matrice A défini par :

$$\mathbf{A} = \begin{bmatrix} \gamma_R(\mathbf{s}_1, \mathbf{s}_1) & \gamma_R(\mathbf{s}_2, \mathbf{s}_1) & \cdots & \gamma_R(\mathbf{s}_n, \mathbf{s}_1) & 1 & f_1(\mathbf{s}_1) & f_2(\mathbf{s}_1) & \cdots & f_n(\mathbf{s}_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_R(\mathbf{s}_1, \mathbf{s}_n) & \gamma_R(\mathbf{s}_2, \mathbf{s}_n) & \cdots & \gamma_R(\mathbf{s}_n, \mathbf{s}_n) & 1 & f_2(\mathbf{s}_2) & f_2(\mathbf{s}_n) & \cdots & f_n(\mathbf{s}_n) \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 & 0 \\ f_1(\mathbf{s}_1) & f_1(\mathbf{s}_2) & \cdots & f_1(\mathbf{s}_n) & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n(\mathbf{s}_1) & f_n(\mathbf{s}_2) & \cdots & f_n(\mathbf{s}_n) & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} (n+1) \\ \times \\ (n+1) \end{matrix} \quad (2.40)$$

le vecteur \mathbf{X} et la matrice \mathbf{a} sont définis par :

$$\mathbf{X} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}'_{(n+k)} ; \quad \mathbf{a} = \begin{pmatrix} \gamma_R(\mathbf{s}_0, \mathbf{s}_1) \\ \vdots \\ \gamma_R(\mathbf{s}_0, \mathbf{s}_k) \\ 1 \\ f_1(\mathbf{s}_0) \\ \vdots \\ f_n(\mathbf{s}_0) \end{pmatrix}_{(n+k)} \quad (2.41)$$

1. Calculer les coefficients de la matrice \mathbf{A} et du vecteur \mathbf{a} .
2. Dédire ensuite les termes du vecteur \mathbf{X} en résolvant $\mathbf{A} \mathbf{X} = \mathbf{a}$.
3. Une fois qu'on trouve les poids λ_i , on pourra estimer la valeur $\forall \mathbf{s}_0$

$$\widehat{Z}_{KU}(\mathbf{s}_0) = \sum_{i=1}^N \lambda_i Z(\mathbf{s}_i) \quad (2.42)$$

4. La variance d'estimation par krigeage universel sera donnée par :

$$\sigma_{KU}^2(\mathbf{s}_0) = \mathbf{X}' \mathbf{a} \quad (2.43)$$

2.4 Simulations stochastiques

Au lieu d'une seule carte d'estimations locales par le krigeage, les simulations stochastiques génèrent L cartes (L réalisations) équiprobables. Chaque point du domaine de l'étude aura donc L valeurs à partir desquelles on pourra calculer la carte de variance.

En supposant que $z(\mathbf{s}_0)$, la variable régionalisée, une réalisation particulière d'une fonction aléatoire continue dans l'espace et obéissent à une structure spatiale continue caractérisée par un semivariogramme, on se place volontairement dans une optique probabiliste. Les simulations stochastiques produisent une estimation de l'incertitude sur la variation spatiale de la mesure. En effet, la moyenne des simulations stochastiques fournit un exemple de ce que pourrait être la réalité.

Les réalisations sont statistiquement équivalentes entre elles; elles ont la même espérance, même fonction de covariance et le même histogramme. De plus, elles sont statistiquement indépendantes entre elles (une covariance croisée nulle entre elles). En construisant ces simulations stochastiques, on ne cherche plus à trouver LA réalisation qui incarne la réalité, mais à générer une famille de réalisations ayant les mêmes propriétés statistiques.

Simulations stochastiques non conditionnelles

Plusieurs algorithmes existent pour simuler non conditionnellement des fonctions aléatoires. On pourra citer par exemple la méthode de décomposition matricielle, les moyennes mobiles discrètes, les méthodes spectrales, les bandes tournantes et les méthodes séquentielles (Jean-Paul Chilès, 2012). Les simulations stochastiques non conditionnelles (*SNC*) possèdent une similitude purement statistique. Aux points de mesures, les valeurs des simulations peuvent différer des valeurs expérimentales. Ceci se traduit par :

$$Z^{SNC}(\mathbf{s}_0) \neq Z^{exp}(\mathbf{s}_0) , \mathbf{s}_0 \in \mathbf{D} \quad (2.44)$$

Simulations stochastiques conditionnelles

Dans le cadre des simulations stochastiques conditionnelles, une simulation est une réalisation d'un modèle probabiliste alors que si elle est conditionnelle elle doit respecter les mesures des valeurs échantillonnées. Autrement dit, la simulation doit obligatoirement passer par les points de mesures : la valeur simulée doit être égale à la valeur mesurée.

$$Z^{SC}(\mathbf{s}_0) = Z^{exp}(\mathbf{s}_0) , \quad \forall \mathbf{s}_0 \in \mathbf{D} \quad (2.45)$$

Les simulations conditionnelles sont utiles pour contrecarrer le lissage généré par le krigeage.

Chaque simulation conditionnelle a pour objectifs de :

- ✓ Reproduire le même histogramme des données primaires;
- ✓ Reproduire le modèle de continuité spatiale utilisé (le semivariogramme).

Le Tableau 2-1 représente une comparaison entre les simulations conditionnelles et non conditionnelles en comparant l'histogramme, le semivariogramme et les données primaires.

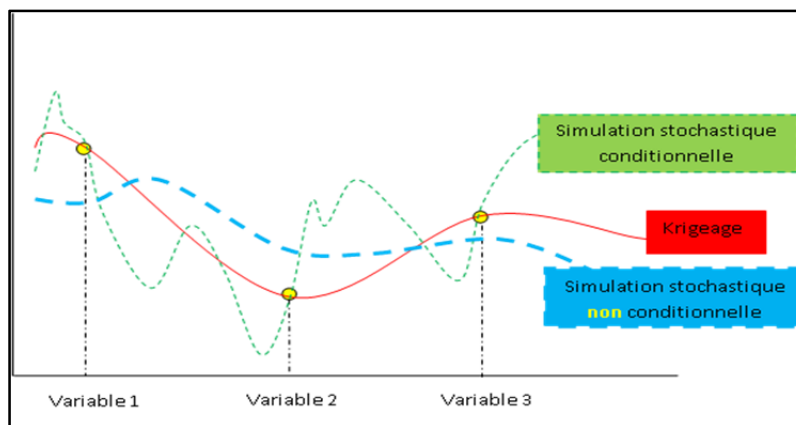


Figure 2-5 Schéma mettant en évidence le krigeage, les simulations stochastiques conditionnelle et non conditionnelle

Tableau 2-1 Comparaison entre les (SC) et (SNC)

	Simulations Conditionnelles (SC)	Simulations Non Conditionnelles(SNC)
Histogramme	☺ Oui	☺ Oui
Semivariogramme	☺ Oui, mais un mélange entre théorique et expérimental	☺ Oui
Données originales	☺ Oui	⊗ Non

2.4.1 Simulations séquentielles conditionnelles (SSCs)

La simulation séquentielle utilise la moyenne et la variance du krigeage pour générer une fonction de distribution cumulative conditionnelle (Nowak et Verly, 2005). Comme son nom l'indique, c'est une fonction de distribution cumulative conditionnée par les valeurs voisines les plus proches incluses dans la zone de recherche (les données primaires) et par les valeurs simulées cumulativement lors de la simulation.

Les méthodes de la géostatistique non linéaire, notamment les simulations stochastiques séquentielles, supposent une distribution de fréquences de la variable étudiée suivant un modèle gaussien. En réalité, il est souvent très rare de rencontrer des distributions parfaitement gaussiennes. Par conséquent, on procèdera à une transformation de la variable réelle pour la ramener à une variable gaussienne. Les simulations séquentielles conditionnelles se basent principalement sur le théorème de Bayes (Gaetan, 2008) et les simulations stochastiques de Monte-Carlo (Ankenman, Nelson et Staum, 2010; Gaetan, 2010).

La simulation est dite conditionnelle si elle reproduit les valeurs de l'échantillon lors de l'estimation.

$$z^l(\mathbf{s}_\alpha) = z(\mathbf{s}_\alpha), \quad l = 1 \dots L, \quad \alpha = 1, \dots, n \quad (2.46)$$

avec l le numéro de la réalisation et \mathbf{s}_α est la localisation de l'attribut z sur la grille.

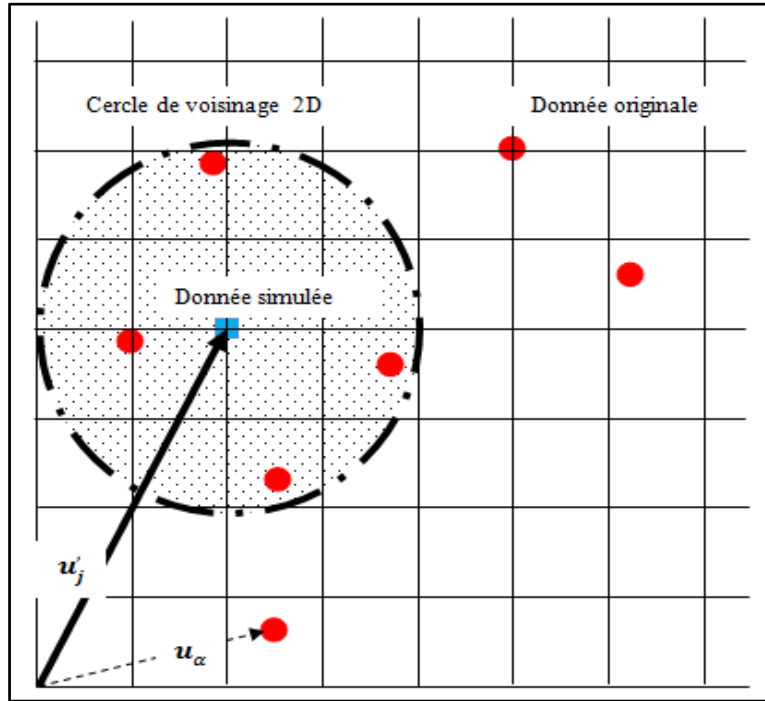


Figure 2-6 Schéma explicatif de la simulation séquentielle conditionnelle d'une donnée sur un nœud de la grille

La fonction aléatoire représente un ensemble des variables aléatoires $\{Z(\mathbf{s}_j), j = 1, \dots, N\}$ qui sera défini sur les N nœuds par le vecteur de localisation \mathbf{s}_j (Figure 2-6). L'objectif de la simulation séquentielle conditionnelle est de générer plusieurs réalisations conjointes (*joint realizations*) de N variables aléatoires définies par $(z^l(\mathbf{s}_j), j = 1, \dots, N, l = 1 \dots L)$ conditionnelles aux valeurs originales (valeurs expérimentales) $\{z(\mathbf{u}_\alpha) \alpha = 1, \dots, n\}$. Ceci nous donne une N -points *ccdf* qui modélisent l'incertitude conjointe (*joint uncertainty*) en N localisations en utilisant l'égalité suivante :

$$\underbrace{F(u'_1, \dots, u'_N; z_1 \dots z_N | (\mathbf{n}))}_{N \text{ points } ccdf} = Prob(Z(u'_1) \leq z_1, \dots, Z(u'_N) \leq z_N | (\mathbf{n})) \quad (2.47)$$

En utilisant le théorème de Bayes, la *ccdf* de N points est le produit de N *ccdf* de chaque point.

$$\begin{aligned}
 F(\mathbf{s}'_1, \dots, \mathbf{s}'_N; z_1 \dots z_N | (\mathbf{n})) \\
 &= F(\mathbf{s}'_N; z_N | (n + N - 1)) \\
 &\times F(\mathbf{s}'_{N-1}; z_{N-1} | (n + N - 2) \times \dots \times F(\mathbf{s}'_1; z_1 | (\mathbf{n})))
 \end{aligned} \tag{2.48}$$

Cette décomposition permet à la réalisation de se faire en N étapes séquentielles successives, d'où la nomenclature de la simulation séquentielle conditionnelle.

Pour le premier vecteur, première localisation \mathbf{u}'_1 , la fonction de distribution cumulative conditionnelle *ccdf* sera conditionnée par les données primaires incluses dans le cercle de voisinage :

$$F(\mathbf{s}'_1; z | (\mathbf{n})) = Prob (Z(\mathbf{s}'_1) \leq z | (\mathbf{n})) \tag{2.49}$$

Ensuite, la valeur obtenue dans la localisation définie par \mathbf{s}'_1 lors de la première simulation se rajoute à l'ensemble des valeurs conditionnelles pour le reste des simulations et ainsi de suite. Cette procédure se refait jusqu'à la simulation de tous les nœuds N de la grille pour terminer une première réalisation. Une deuxième réalisation débute ensuite et doit suivre une séquence différente lors du passage par les nœuds de la grille.

Ainsi, cette simulation séquentielle a été conditionnée à l'aide d'un seul krigeage pour estimer les valeurs manquantes des nœuds de la grille durant chaque réalisation. En revanche, la simulation gaussienne conditionnelle, comme on expliquera par la suite dans chapitre, nécessitera 2 krigeages lors du conditionnement.

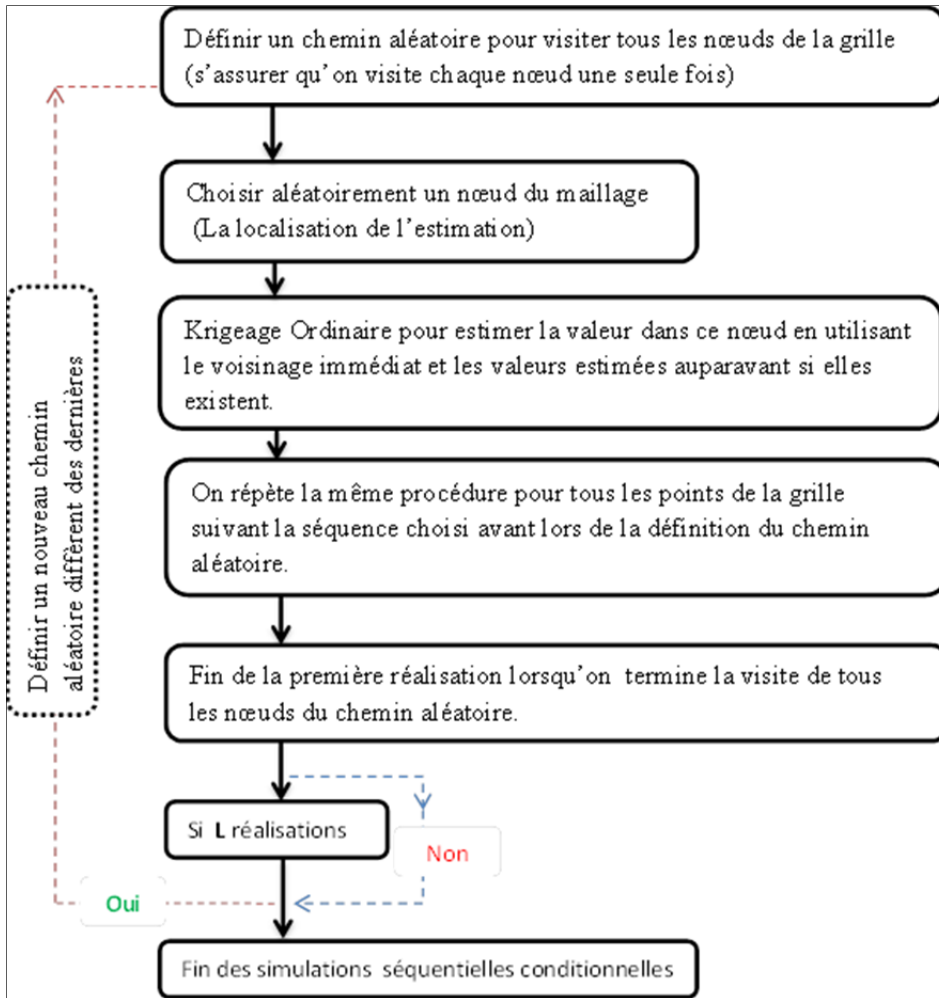


Figure 2-7 Algorithme des simulations séquentielles conditionnelles (SSCs)

2.4.2 Simulations gaussiennes conditionnelles (SGCs)

Le conditionnement pour ce type de simulations se résout à l'aide d'un double krigeage. On part principalement d'une base de données (les données expérimentales $(z^{exp}(\mathbf{s}_i), i = 1, \dots, n)$) et d'une structure de corrélation spatiale (le semivariogramme des variables gaussiennes) dans le but d'estimer la valeur réelle $\hat{z}(\mathbf{s}_0)$, $\forall \mathbf{s}_0 \in \mathbf{D}$.

Comme $\hat{z}(\mathbf{s}_0)$ est une estimation, alors on générera une erreur définie par $[Z(\mathbf{s}_0) - \hat{Z}(\mathbf{s}_0)]$:

$$\underbrace{\widehat{z(u)}}_{\text{inconnue}} = \underbrace{\widehat{z(u)}}_{\widehat{z(u)}} + \underbrace{[z(u) - \widehat{z(u)}]}_{\text{inconnue}} \quad (2.50)$$

Dans le cas d'une fonction aléatoire gaussienne, le résidu en tout point $[Z(\mathbf{s}_0) - \widehat{Z}(\mathbf{s}_0)]$, est en principe indépendant des valeurs aux points des données. Ainsi, le résidu est indépendant de n'importe quelle combinaison linéaire des données, notamment n'importe quelle estimation par le krigeage. L'astuce est donc de remplacer le résidu inconnu par un résidu simulé ayant exactement la même structure spatiale. Ce dernier sera simulé par un processus aléatoire : une simulation non conditionnelle générée à partir du semivariogramme des données gaussiennes, qui sera notée par $z_K^{SNC}(\mathbf{s}_0)$.

Ensuite, on estime par krigeage la valeur $z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0)$ du champ généré sur des valeurs prises par la simulation stochastique non conditionnelle aux points de mesure \mathbf{s}_i et on calcule en tout point, le résidu du krigeage qui sera estimé par $z_K^{SNC}(\mathbf{s}_0) - z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0)$. Enfin on substitue le résidu inconnu dans l'équation précédente $[z(\mathbf{s}_0) - \widehat{z}(\mathbf{s}_0)]$ par le nouveau calculé $z_K^{SNC}(\mathbf{s}_0) - z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0)$.

Finalement, la nouvelle équation sera donnée par (Chilès et Delfiner, 2012)

$$z_K^{SC}(\mathbf{s}_0) = \widehat{z}(\mathbf{s}_0) + \left[z_K^{SNC}(\mathbf{s}_0) - z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0) \right] \quad (2.51)$$

Comme le krigeage est un interpolateur exact, alors la variance aux points de mesures est nulle. $\forall \mathbf{s}_i, i = 1, \dots, n$, le résidu $z_K^{SNC}(\mathbf{s}_0) - z\left(\widehat{\frac{SNC}{K}}\right)(\mathbf{s}_0) = 0$ d'où on retrouve bien la valeur conditionnelle $z_K^{SC}(\mathbf{s}_0) = \widehat{z}(\mathbf{s}_0)$.

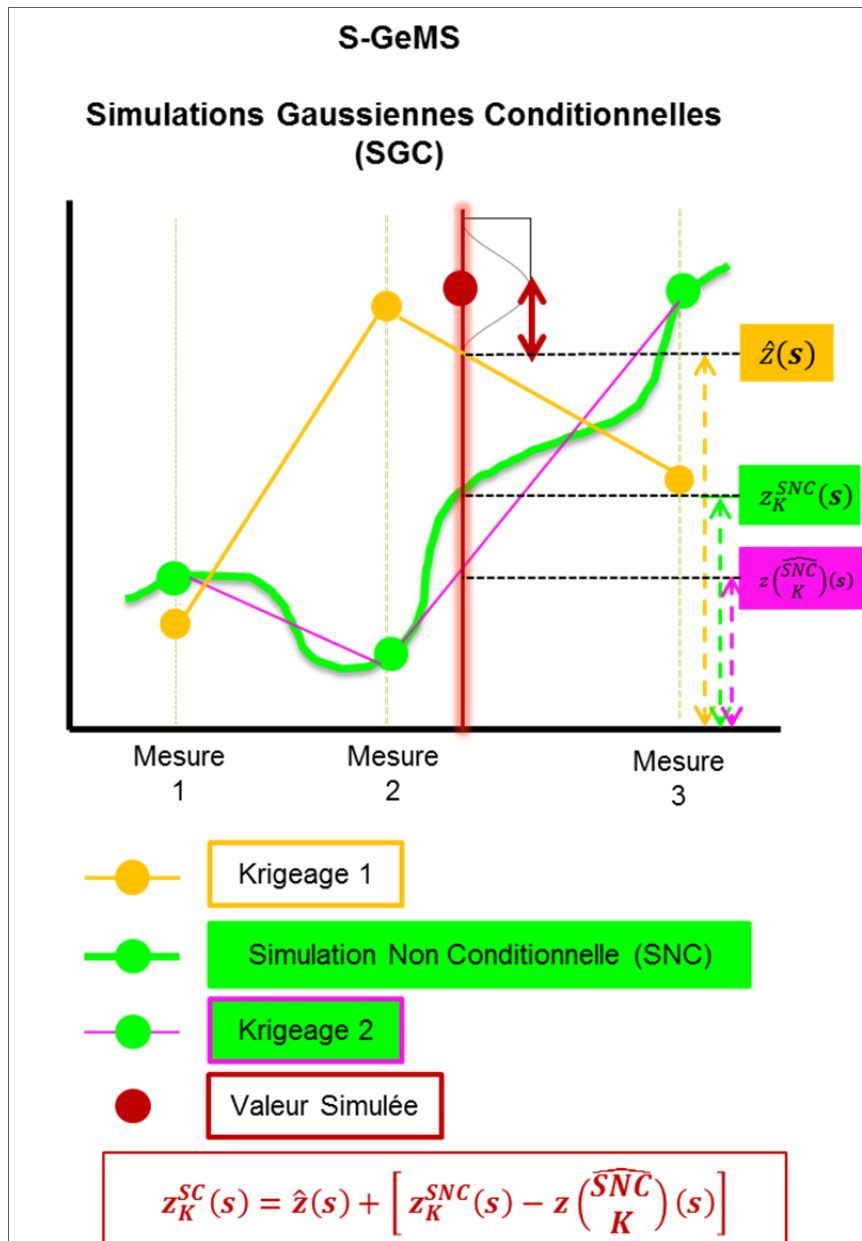


Figure 2-8 Algorithme des simulations gaussiennes conditionnelles (SGSs) - 1

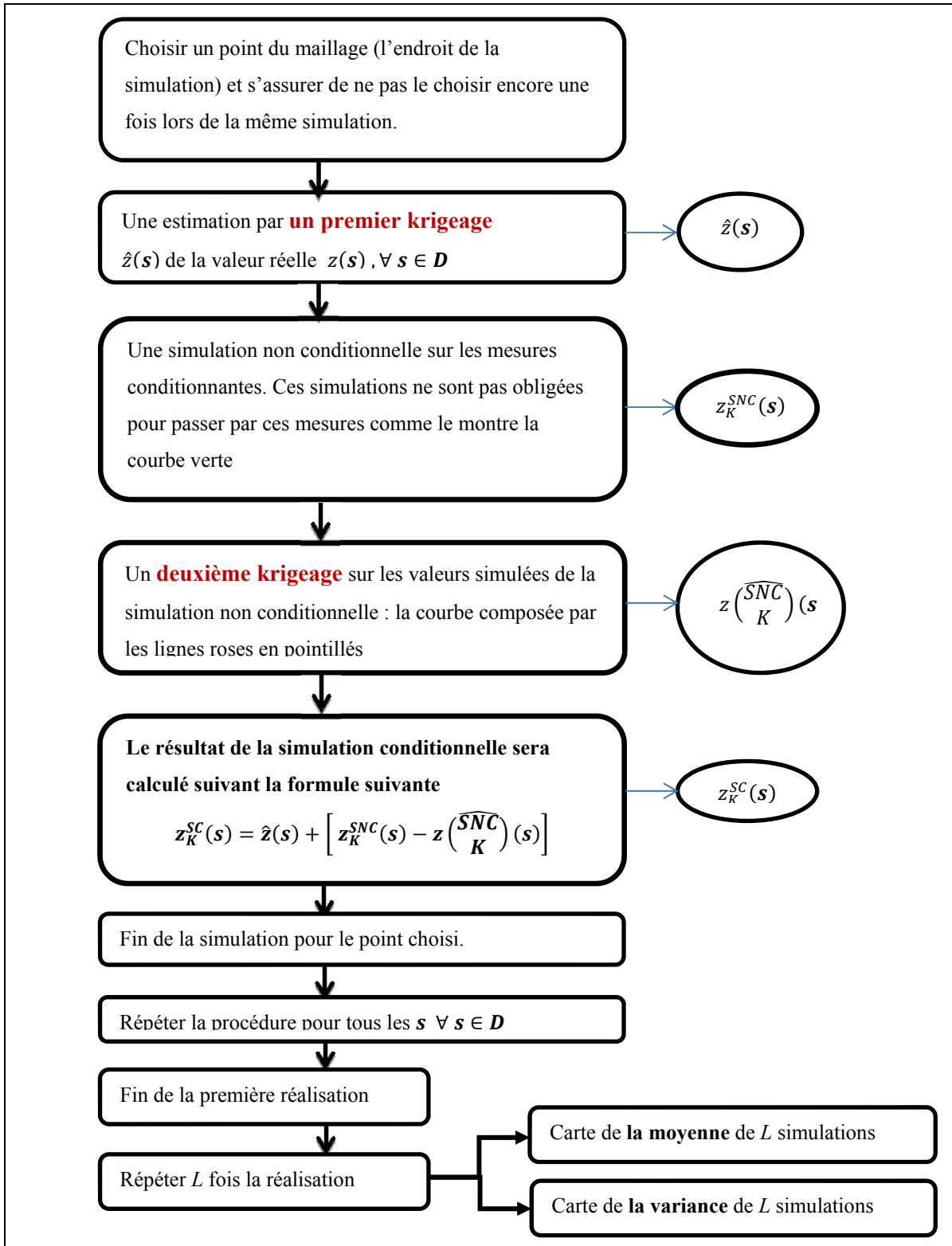


Figure 2-9 Algorithme des simulations gaussiennes conditionnelles (SGSs) – 2

CHAPITRE 3

HYPOTHÈSES RETENUES

Comme mentionné précédemment lors du krigeage et des simulations stochastiques conditionnelles, on cherche à estimer adéquatement les deux premiers moments statistiques (espérance et variance) de la variable aléatoire. On doit postuler certaines hypothèses qui supposent une certaine régularité et une homogénéité des champs de mesures à simuler.

Si les données primaires (utilisées pour la construction du semivariogramme expérimental) ne respectent pas l'hypothèse de normalité, ces dernières subiront une transformation gaussienne dans le but de minimiser la variance et l'influence des valeurs atypiques si elles existent.

Dans ce projet, le gradient des déformations ne causera aucune restriction de la continuité spatiale, car il n'y a pas de changement brutal catégorique de la tendance de la variable à estimer ; on traitera juste les variables régionalisées continues (et non pas les variables catégoriques).

La construction du semivariogramme expérimental se base sur les valeurs calculées par les simulations par des éléments finis ÉFs, car on ne pourra pas générer un semivariogramme à partir d'un nombre minime de mesures expérimentales. Ainsi, le fait d'appliquer le semivariogramme déduit à partir des calculs par les simulations par des éléments finis doit en principe préserver la variance en se basant sur la mécanique des milieux continus.

Il est important de noter que le semivariogramme recherché doit respecter l'hypothèse d'ergodicité. En effet, on a supposé au début que les données primaires (les observations données par les simulations ÉFs) ne sont qu'une réalisation d'une fonction aléatoire, ce qui est insuffisant pour déduire la fonction aléatoire en tant que telle. Alors, l'hypothèse

ergodique stipule qu'une réalisation sera suffisante pour déduire les propriétés de l'ensemble de toutes les réalisations possibles.

Pour construire un semivariogramme expérimental représentatif, on doit tenir compte des points suivants (Journel et Huijbregts, 1978) :

- La précision des estimés est proportionnelle au nombre des paires $n(h)$ par incrément h . Arbitrairement, on suggère de s'assurer qu'on estime 30 paires de points pour chaque incrément h afin de distinguer le type de semivariogramme empirique à utiliser lors de la modélisation.
- On limite les incréments h du semivariogramme expérimental à une distance extrême approximativement égale à la moitié de dimension de l'échantillon à analyser.

L'application du krigeage basée sur un modèle d'une fonction aléatoire (FA) requiert un minimum d'hypothèses sans lesquelles on ne pourra pas trouver les poids optimaux du système de krigeage (On présente un rappel pour les hypothèses utilisées et pour plus de détails, se référer au chapitre suivant).

- On désigne par $z(\mathbf{s})$, la localisation spatiale euclidienne du point de mesure (observation) et on suppose que l'échantillonnage (l'ensemble des observations) est une réalisation d'une fonction aléatoire.
- La fonction aléatoire est stationnaire de deuxième ordre, c.-à-d. que les moments statistiques qui impliquent une ou deux variables sont invariants par translations spatiales et ne dépendent que de la distance euclidienne.

En respectant la stationnarité de deuxième ordre, on utilisera des variogrammes bornés lors de la modélisation.

Autrement dit, il faut que la structure spatiale vérifie l'équation suivante

$$\lim_{h \rightarrow \infty} Cov(h) = 0 \quad (3.1)$$

La fonction de covariance doit être définie positive et seul les modèles paramétriques de variogramme (gaussien, exponentiel et sphérique) sont utilisables. Lors de la modélisation, on ne doit pas interpoler linéairement entre les points du variogramme, pas non plus d'effectuer des ajustement polynomial et pas d'estimation non paramétrique (par exemple la méthode à noyau). Si le variogramme est borné alors la covariance existe et ce dernier respecte la stationnarité de second ordre (Journel A.G. , 1981).

On considère que la moyenne est constante et connue sur tout le domaine d'étude uniquement pour le krigeage simple (KS), constante et inconnue (connue juste au voisinage de l'estimation) pour le krigeage ordinaire (KO). En revanche pour le krigeage universel (KU), la moyenne est déterministe et inconnue. Toutefois, elle peut être modélisée par une combinaison linéaire de fonctions analytiques. Idéalement ces fonctions, qui définissent la tendance, doivent être spécifiques à la physique du problème. En absence d'une interprétation physique, la dérive est modélisée par un polynôme à faible degré ≤ 2 (Deutsch et Journel, 1998).

L'ensemble des déformations choisies aléatoirement parmi toutes les valeurs calculées par les simulations ÉFs symbolisera l'échantillon des valeurs expérimentales et est représentatif du domaine d'étude **D**.

CHAPITRE 4

MÉTHODOLOGIE EXPÉRIMENTALE

Dans cette partie du mémoire, on présentera la méthodologie expérimentale utilisée pour aboutir à l'objectif du projet. On commence tout d'abord par une définition du domaine d'étude D puis par un examen exploratoire des données extraites à partir des simulations par la méthode des ÉFs. On analyse les statistiques descriptives dans le but de déterminer d'une part, la moyenne arithmétique qui sera utilisée lors du krigeage simple (KS), et d'une autre part, la présence des valeurs atypiques. De plus, ces données 'théoriques' (ÉFs) sont utilisées pour générer le variogramme expérimental vu que nous possédons un nombre suffisamment grand des données expérimentales.

4.1 Paramétrisation surfacique

Les données sont extraites des valeurs calculées par des ÉFs au niveau des nœuds d'un maillage surfacique. Si ce dernier n'est pas bidimensionnel, il faut paramétrer le maillage tridimensionnel. En effet, dans le cas de l'aube de la turbine, on a besoin d'un maillage surfacique triangulaire : c'est-à-dire, la liste des positions des sommets des triangles et la liste de connectivités. On a transformé le maillage de type CTRIA6 (un maillage surfacique avec 6 nœuds de connectivité pour chaque maille) en un autre maillage triangulaire avec 3 sommets tout en respectant les liaisons et la position des nœuds principaux (on éliminera juste les nœuds moyens : *midnodes*). Une fois qu'on réarrange les surfaces du nouveau maillage triangulaire, on commence la paramétrisation surfacique. On a déterminé les nouvelles coordonnées (u,v) pour chaque sommet du maillage initial (x,y,z) , en utilisant une fonction bijective qui fixe la frontière lors de la paramétrisation en résolvant la condition de la frontière de type Neumann (Hormann et al., 2007) .

4.2 Transformation gaussienne

Cette transformation est primordiale. Généralement, les données ne se présentent pas sous la forme d'une distribution symétrique. On a construit la fonction de distribution cumulative représentative de ces données. La définition théorique de la distribution est donnée par $F(z) = P[Z \leq z]$. Comme la distribution est calculée à partir d'un échantillon discret de valeur, la correction de continuité est donnée par $F_{est}(z) = \frac{(P_{est}[Z \leq z] + P_{est}[Z < z])}{2}$ avec $P_{est}[\cdot]$ étant la probabilité estimée à partir des fréquences.

4.3 Construction du semivariogramme expérimental

Une fois qu'on a réalisé la transformation gaussienne, on pourra commencer l'analyse variographique. La première étape consiste à choisir la zone de l'étude du domaine D . Cette zone doit comporter un gradient de la variabilité spatiale de la propriété à estimer.

On commence dans un premier temps la construction du variogramme omnidirectionnel (le modèle isotropique sera le premier candidat), puis suivant les directions principales Nord-Sud et Est-West. Si on a une idée sur la direction de variabilité de la propriété à estimer, on pourra construire le semivariogramme directionnel suivant des directions bien précises. La fonction principale MatLab® permet de choisir n'importe quelle direction d'étude et même de fixer des tolérances sur les angles choisies.

Lorsqu'on construit le semivariogramme expérimental, on se trouve face à des paires de points atypiques qui peuvent influencer la modélisation de ce dernier. On doit utiliser une fonction lasso² et éliminer ces points avant de commencer l'ajustement.

² C'est une fonction de lissage des contours de sélection à l'aide d'un contour progressif choisi aléatoirement et manuellement par l'utilisateur. Elle permet d'éliminer les points atypiques du semivariogramme expérimental.

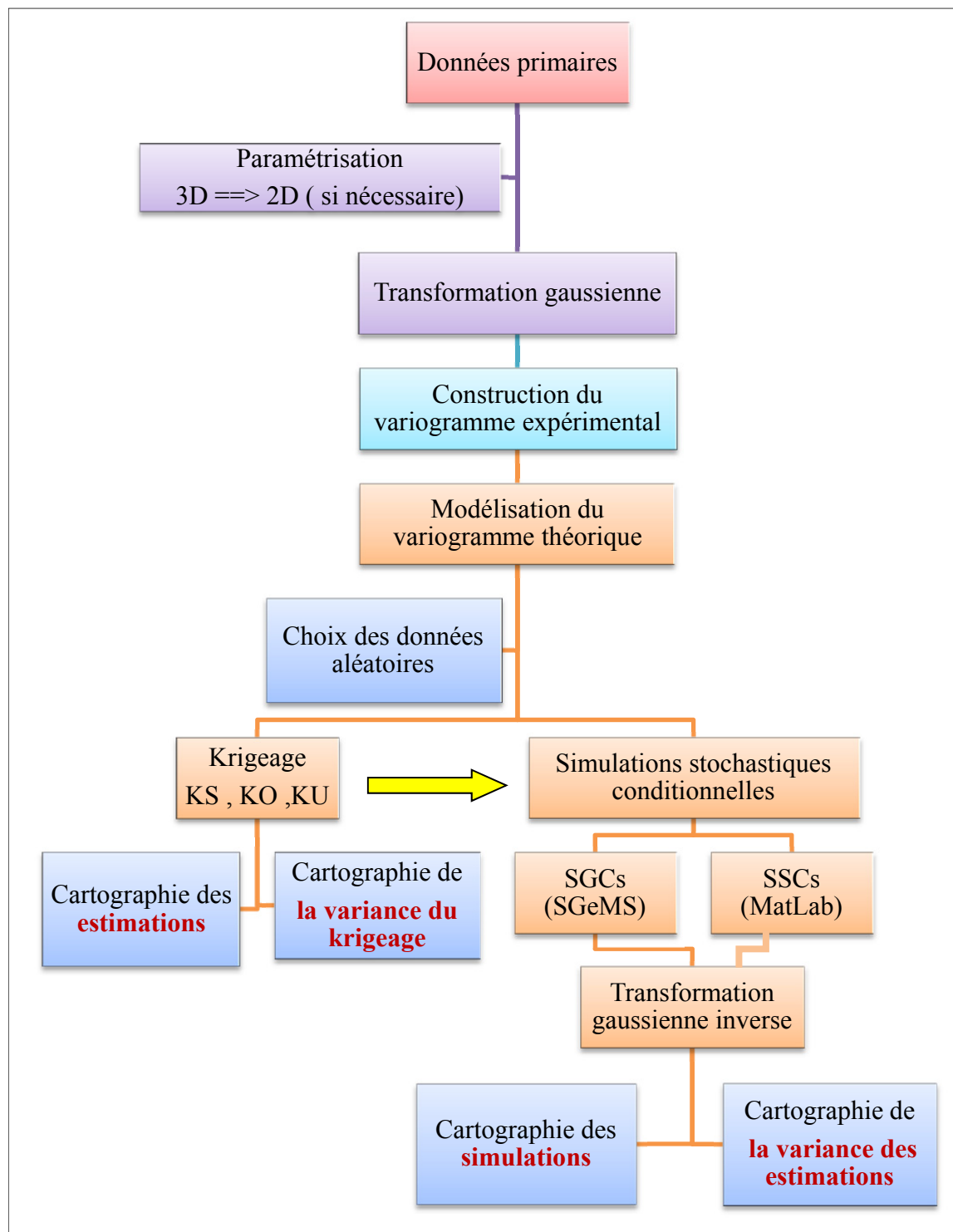


Figure 4-1 Méthodologie expérimentale

Une fois que le variogramme expérimental est construit, il faudra l'ajuster et le valider afin d'obtenir le variogramme empirique (variogramme théorique). En fait, on ajuste le variogramme par la méthode de moindres carrés et on déduit le type du modèle théorique (sphérique, gaussien ou exponentiel), l'effet de pépite, la portée ainsi que le palier du variogramme. Une fois qu'on a ajusté les paramètres du variogramme, on utilisera la technique de la validation croisée ainsi qu'une analyse de sensibilité pour les valider.

4.4 Krigeage

La deuxième partie de ce travail est le krigeage proprement dit. On choisit aléatoirement un échantillon de données parmi les valeurs des résultats des simulations ÉFs.

On a choisi trois types de krigeage : Krigeage simple (KS), Krigeage ordinaire (KO) et Krigeage universel (KU). Pour chaque type, on génère une carte d'estimation des déformations à partir d'un échantillon de N points choisis aléatoirement et une carte approximative de l'allure de la variance du krigeage pour tous les points du maillage. Normalement, si notre calcul est bon, on doit avoir la même carte de krigeage pour les 3 types, mais pas nécessairement la même carte de variance à cause des hypothèses intrinsèques à chaque type de krigeage. En se basant sur cette dernière carte, on garde l'estimateur qui possède moins de variance que les autres. Finalement, pour réaliser le krigeage, on doit choisir des mesures (des données primaires) qui joueront les rôles des mesures expérimentales. Il est fortement recommandé de choisir aléatoirement les points de mesure de telle façon à éviter l'effet de regroupement (*clustering*) et à couvrir le mieux possible le domaine d'étude.

Pour valider les résultats, une première étape consiste à comparer la cartographie des prédictions et celle des déformations calculées par les simulations ÉFs. Une ressemblance pourra être authentifiée en utilisant plusieurs outils statistiques (comparer les histogrammes, les semivariogrammes des surfaces krigées, les QQ-plot, une analyse des résidus, etc.). Plus particulièrement, une régression linéaire simple entre les valeurs estimées et les valeurs de

références ÉFs peut être effectuée pour estimer le degré de corrélation entre les deux vecteurs de valeurs. À ce stade, il est important de s'assurer que le krigeage donne des bonnes estimations car il sera utilisé par la suite pour conditionner les simulations stochastiques.

4.5 Simulations stochastiques

Deux (2) types de simulations seront réalisés tel que comme mentionné précédemment. Pour les simulations séquentielles conditionnelles (SSCs), on génère une moyenne de simulations ainsi qu'une cartographie de propagation de l'incertitude sur les mesures. Pour confirmer le résultat, on fait appel à des simulations gaussiennes conditionnelles (SGCs) en utilisant le logiciel gratuit SGeMS. La comparaison entre les deux simulations permet de valider le résultat final. Une fois que les simulations sont achevées, on réalise une transformation inverse pour chaque simulation. Finalement, on déduit la moyenne ainsi que la variance de ces simulations. Les mêmes méthodes statistiques de validation pour le krigeage seront utilisées lors de simulations stochastiques

La validation des modèles mathématiques, krigeage et simulations stochastiques, est purement statistiques. Cette vérification analysera les résidus de la fonction aléatoire. En fait, on commencera avec une régression linéaire entre les données primaires ÉFs et les valeurs prédites puis une analyse des valeurs résiduelles en fonction des valeurs ajustées et de l'ordre des simulations.

CHAPITRE 5

RÉSULTATS ET INTERPRÉTATIONS

Ce chapitre présente l'application de notre méthodologie (krigeage jumelé à des simulations stochastiques conditionnelles) sur un quart d'éprouvette à encoches d'une part et sur une zone de l'aube hydraulique d'autre part. L'objectif principal consiste à interpoler des mesures de déformations à partir d'un échantillon de données choisies aléatoirement des valeurs calculées par des simulations par la méthode des ÉFs. En effet, on cherche à définir un modèle mathématique qui représentera une variable expérimentale à partir d'un échantillon de mesures dans le but d'aboutir à une cartographie complète des mesures expérimentales ainsi que leurs variances.

On commence par le recueil et l'analyse préalable des valeurs calculées par la méthode des éléments finis. Dans notre cas, on aura des variables continues qui ne présentent pas des données atypiques mais qui nécessiteront une transformation gaussienne afin d'amortir l'effet du gradient et respecter la condition de stationnarité. Une autre manipulation des données primaires se présentera lors de l'extraction des données du maillage. En effet, le maillage surfacique de l'éprouvette est bidimensionnel (2D) contrairement à celui de l'aube qui sera tridimensionnel (3D). On discutera la paramétrisation surfacique de l'aube et la façon optimale pour transférer les données primaires dans un espace bidimensionnel tout en minimisant la distorsion du maillage.

Pour les deux cas d'étude, on suivra la même stratégie. On débutera par une analyse variographique afin de construire le variogramme expérimental et déduire le modèle du variogramme empirique (théorique). Ce dernier sera ajusté par la méthode des moindres carrés et confirmé par une validation croisée. Une fois le bon modèle est obtenu, on débutera le krigeage qui sera à son tour utilisé pour conditionner les simulations stochastiques tel qu'on a mentionné précédemment lors de l'étude théorique. Finalement, plusieurs tests et analyses statistiques seront effectués pour valider ce dernier modèle.

5.1 Premier cas d'étude : l'éprouvette à encoches

L'éprouvette à encoches, dont les caractéristiques techniques sont déjà détaillées dans le mémoire de J. Arpin-Pont (Arpin-Pont, 2012), a fait l'objet du premier cas d'étude

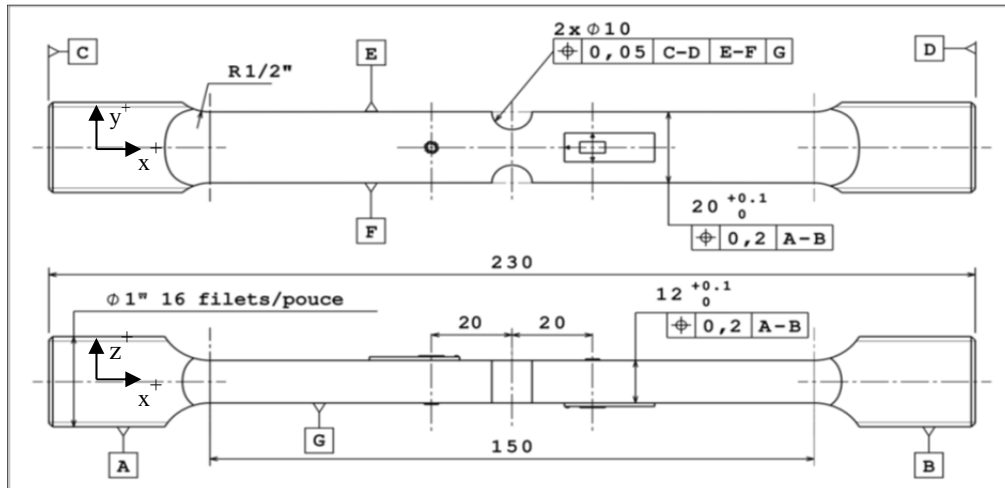


Figure 5-1 Éprouvettes à encoches (Arpin-Pont, 2012)

5.1.1 Analyse variographique

5.1.1.1 Zone d'étude

Dans un premier temps, la Figure 5-2 montre une symétrie axiale suivant les deux principaux axes (axe de traction et de compression). Cette symétrie, permet de réduire la zone d'étude et d'alléger le système linéaire du krigeage en diminuant le nombre des équations à résoudre et par conséquent le temps de calcul surtout lors des simulations stochastiques.

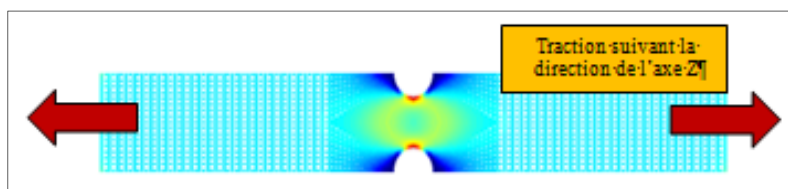


Figure 5-2 Déformation de l'éprouvette

On a tracé les trois semivariogrammes respectivement de toute, la moitié et le quart la surface de l'éprouvette à la Figure 5-3 dans le but de repérer une éventuelle différence significative de la zone d'étude qui pourra refléter un changement de la structure de dépendance spatiale.

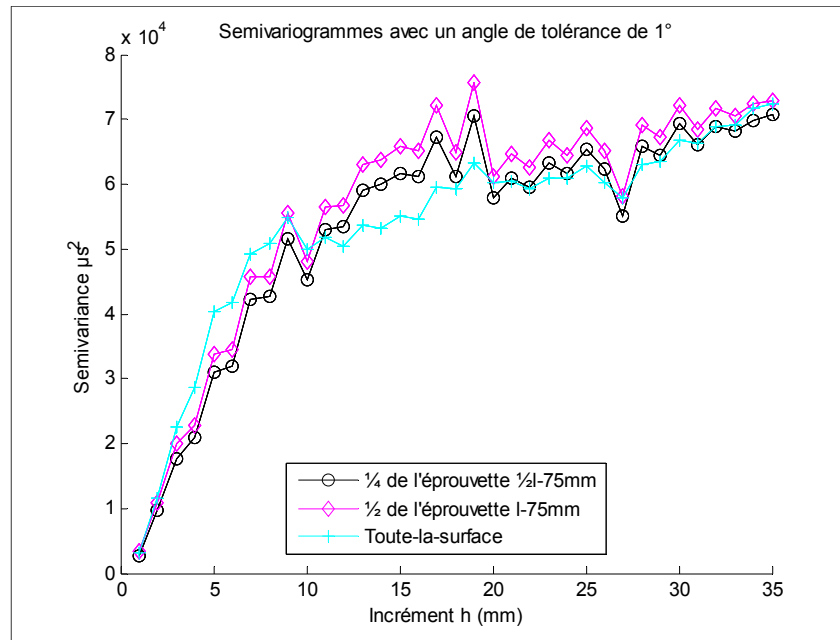


Figure 5-3 Semivariogrammes pour toute, 1/2 et 1/4 de l'éprouvette

On constate que les trois semivariogrammes représentent la même structure de variabilité spatiale. Ils ont le même pallier ainsi que la même portée. Ainsi, on gardera le 1/4 de l'éprouvette pour le reste des analyses de ce cas d'étude.

5.1.1.2 Transformation gaussienne

L'histogramme des déformations mesurées sur l'éprouvette est représenté par une distribution asymétrique avec quelques valeurs extrêmes (un coefficient d'asymétrie = 0.18). Dans le but de respecter l'hypothèse de stationnarité de second ordre, l'approche gaussienne nécessite une transformation gaussienne avant de débiter le krigeage. En effet, les deux types de simulations stochastiques exigent dans un premier temps une transformation des données primaires et dans un deuxième temps une transformation inverse pour revenir de l'espace

gaussien à l'espace réel. La théorie de la transformation gaussienne est bien détaillée dans plusieurs références en géostatistiques notamment (Gooverts, 1997; Journel A.G. , 1981).

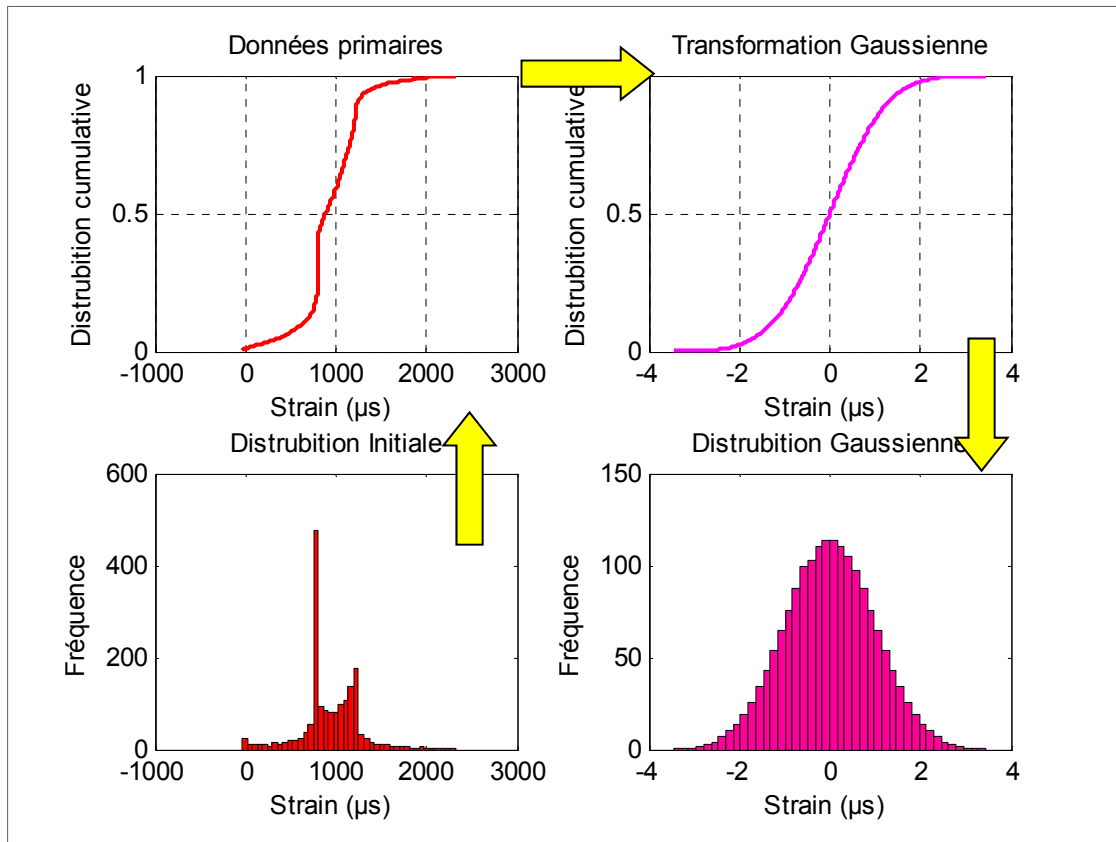


Figure 5-4 Transformation Gaussienne

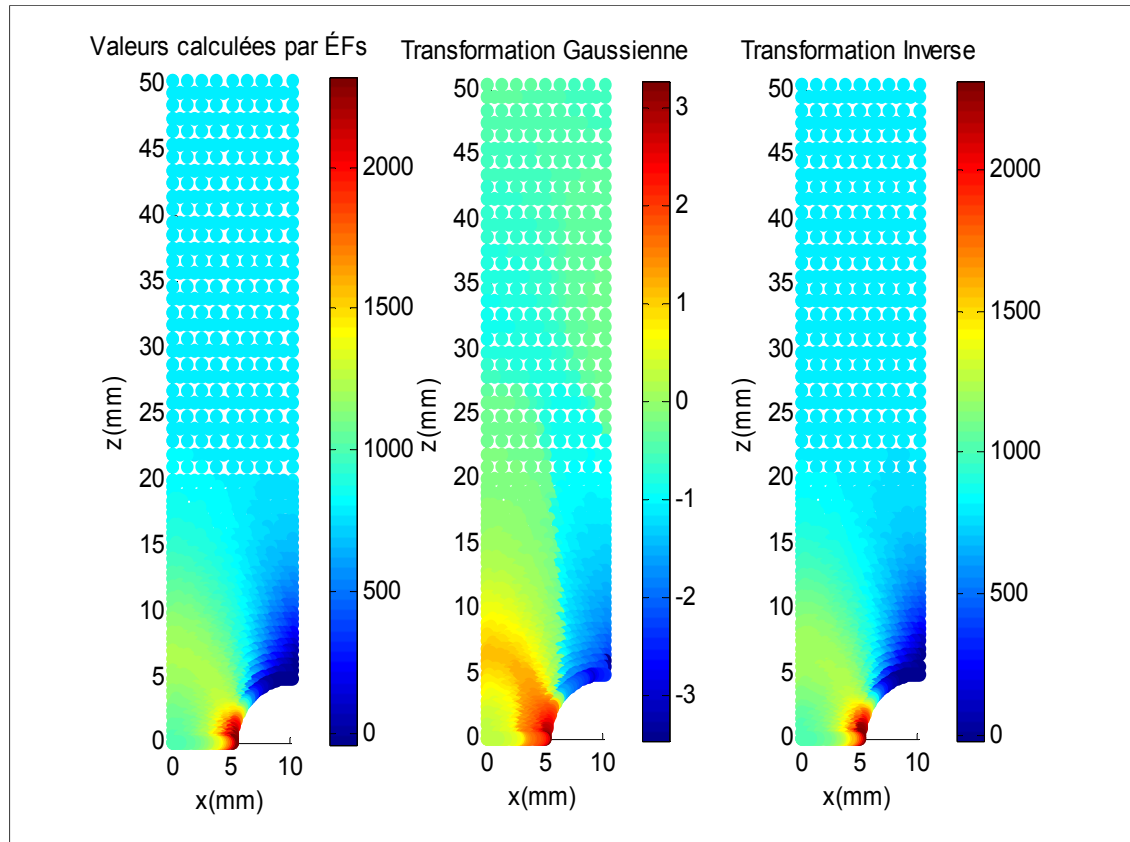


Figure 5-5 Comparaison entre les données primaires ÉFs, la transformée gaussienne et la transformée inverse

La Figure 5-5 présente le champ de déformation de l'éprouvette calculé par des simulations ÉFs puis l'allure de ce dernier dans un espace gaussien. Le but de la transformation gaussienne est d'amortir le gradient au niveau de l'encoche. Le champ de déformation est beaucoup plus lisse dans l'espace gaussien, ce qui nous permet de bien visualiser la stationnarité du domaine d'étude. Finalement, tout à droite, on présente le champ de déformation après le retour dans l'espace réel. Ce dernier est semblable à celui des déformations simulées par des ÉFs. La Figure 5-6 montre une parfaite superposition entre les deux fonctions de distribution cumulative des données primaires ÉFs et celles de la transformation inverse.

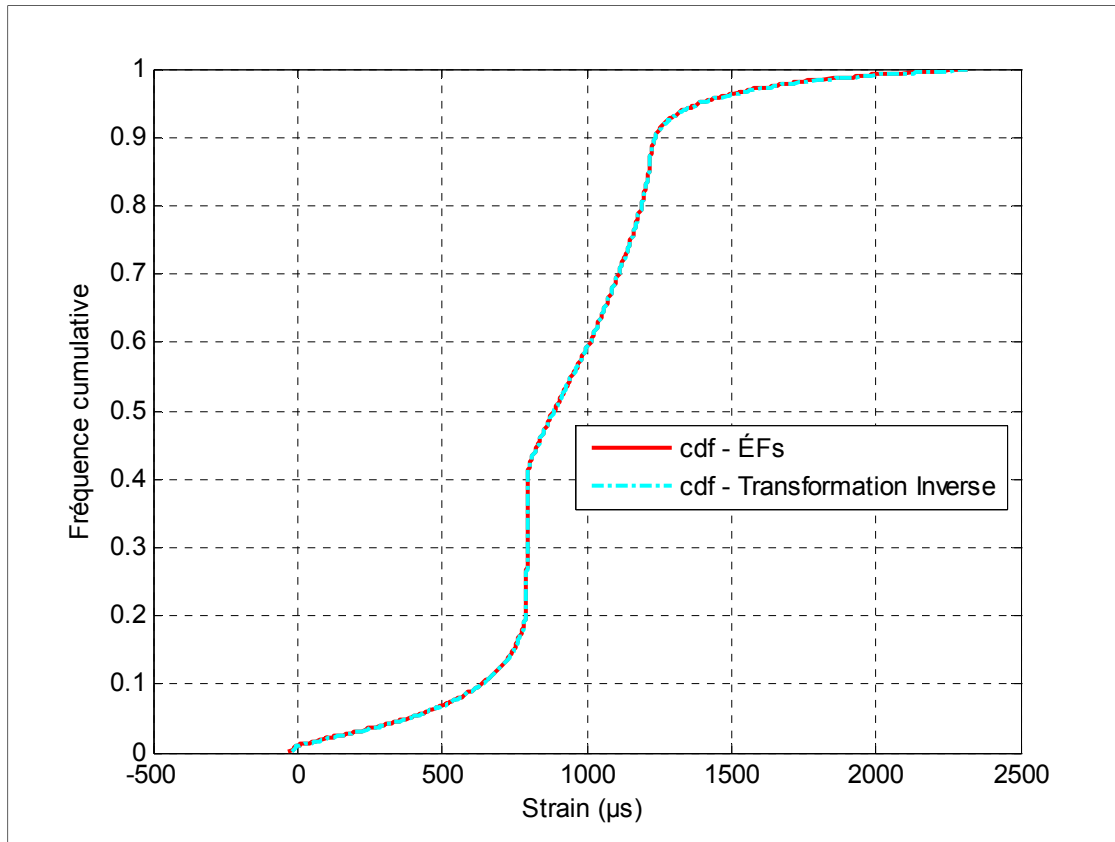


Figure 5-6 Comparaison entre la fonction de distribution cumulative des données primaires ÉFs et sa transformée inverse.

5.1.1.3 Influence de la direction de recherche des paires des données

Un paramètre essentiel intervient lors de la recherche des paires des mesures pour construire le semivariogramme expérimental : l'influence de la direction de collecte des points. En effet, comme on a mentionné dans la partie théorique que le nombre optimal pour la construction est de 30 paires pour chaque classe (Journel et Huijbregts, 1978), et étant donné que le nombre de données générées par des simulations par ÉFs est de loin supérieur à ce requis comme le montre la Figure 5-5, on a choisi de les collecter dans quelques directions privilégiées sachant que la traction de l'éprouvette est uni axiale. On définit l'angle nul à la direction x^+ (voir Figure 5-1) et le sens de l'orientation inversement aux aiguilles d'une montre.

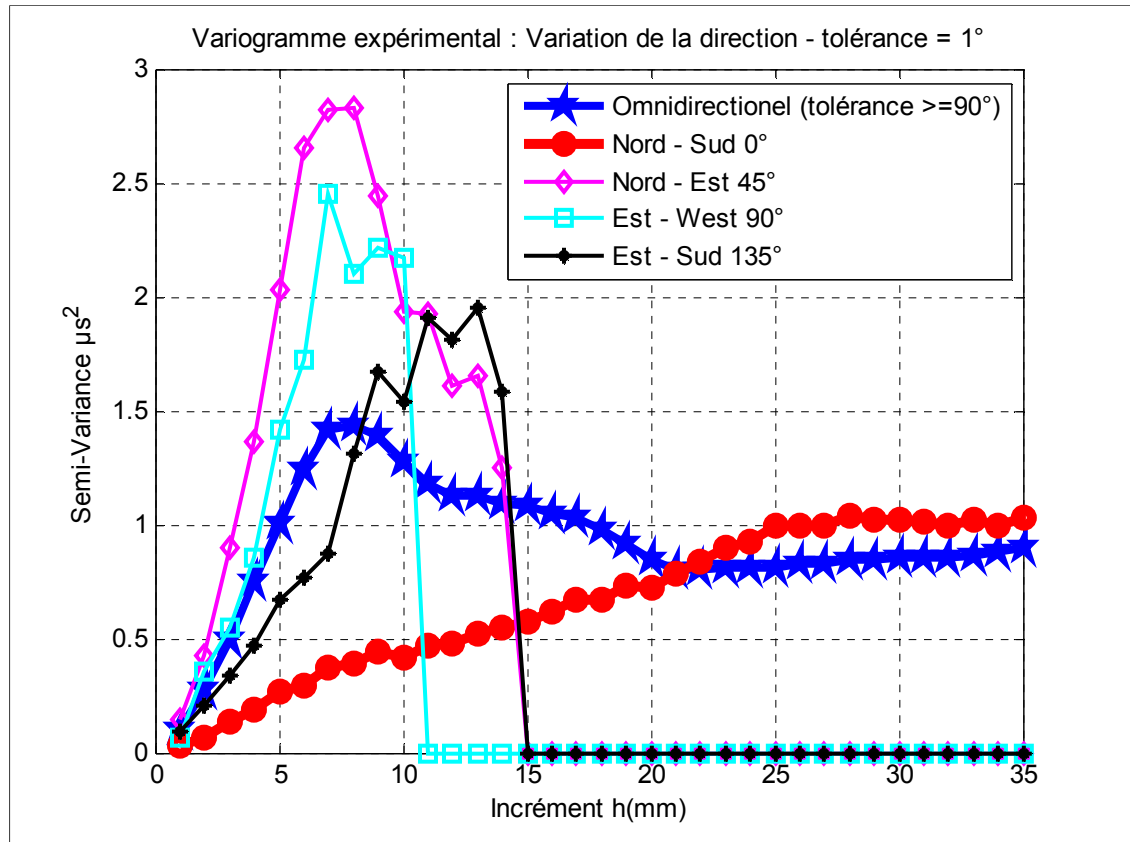


Figure 5-7 Variation de la direction de collecte des données

Dans un premier temps, on a construit les variogrammes dans directions principales suivantes :

- Direction Nord-Sud (x^+) : c'est la direction qui correspond à celle de la traction de l'éprouvette.
- Direction Est-West (y^+) : c'est la direction perpendiculaire à la direction de traction.
- Direction Nord-Est : c'est une direction oblique et qui fait un angle de 45° avec l'axe principal majeur de l'éprouvette
- Direction Est-Sud : C'est la direction perpendiculaire à la direction oblique précédente.
- Finalement, un variogramme omnidirectionnel avec un angle de tolérance supérieur à 90° .

Perpendiculairement à l'axe de traction, le semivariogramme (courbe en bleu ciel) montre un pic puis une chute brutale. Ceci est expliqué par le fait qu'on a une variabilité spatiale importante de la déformation suivant cet axe, ce qui exclut dès le début cette direction d'étude. Le variogramme omnidirectionnel ne présente pas lui non plus l'allure standard d'un variogramme typique : il y a une bosse avant d'atteindre le palier. Ce variogramme pourra être modélisé par un modèle effet trou sinus (Goovaerts, 2000a). En revanche, dans notre cas d'étude, on sait que la déformation suit la direction de traction de l'éprouvette, ce qui privilégie la construction du variogramme dans cette direction. D'ailleurs, le variogramme dans la direction Nord_Sud (direction de l'axe x) montre une structure normale avec un décollage ordinaire puis une stabilisation qui montre une variance constante.

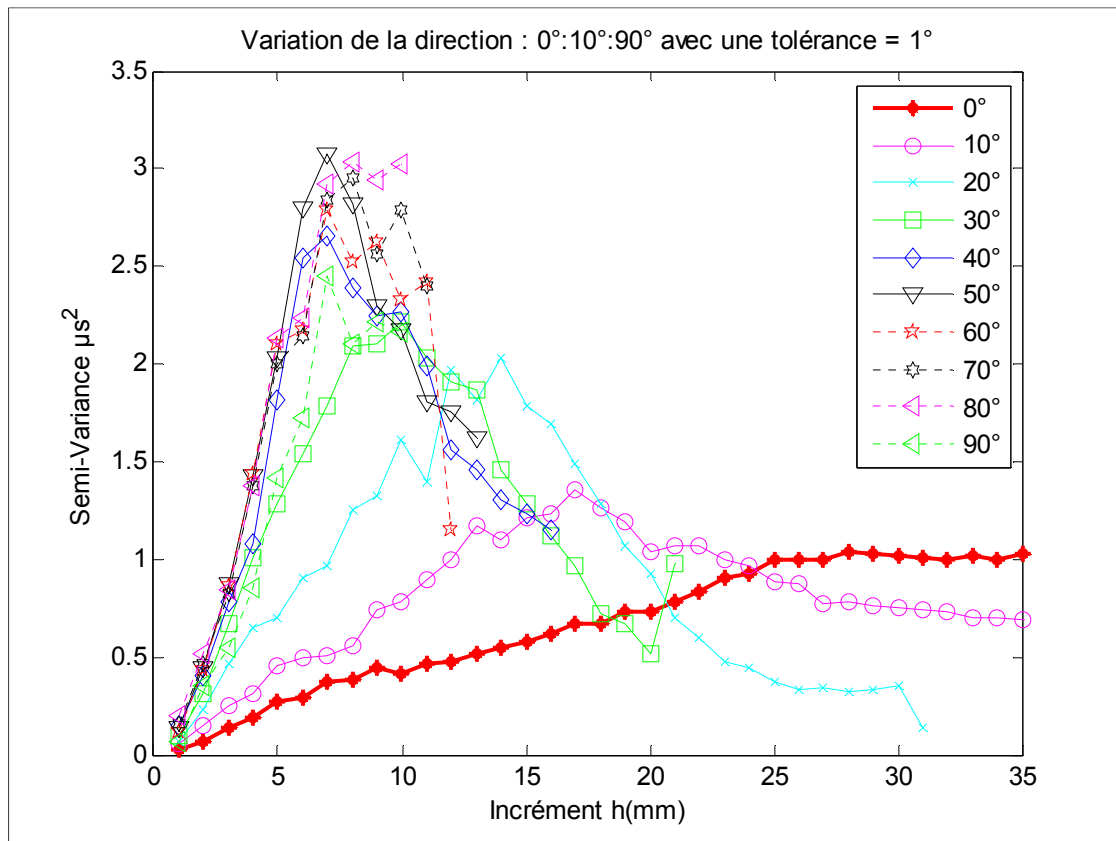


Figure 5-8 Variation de la direction de collecte des données 0° : 10° : 90°

Ensuite, on a tracé de tracer le variogramme dans la direction parallèle à l'axe de traction mais en variant la direction en passant d'un angle allant de 0° à 90° . La Figure 5-8 montre que plus cet angle augmente, plus une bosse apparait et déforme le semivariogramme. Ceci devient critique surtout, tel que a mentionné précédemment, que la variance doit être définie positive si non le variogramme perd sa crédibilité (Gooverts, 1997).

Une fois qu'on a déterminé la direction préférentielle de l'étude, on analysera plus en détail le variogramme expérimental afin de mieux définir ces paramètres (portée et palier). Une première étape consiste à construire son nuage variographique. Ce nuage nous permet de détecter les paires de données atypiques qui peuvent nuire à la bonne modélisation de ce variogramme. La Figure 5-9 montre croissance du variogramme après sa stabilisation. Ceci est dû à des paires de données atypiques statistiquement non significatives.

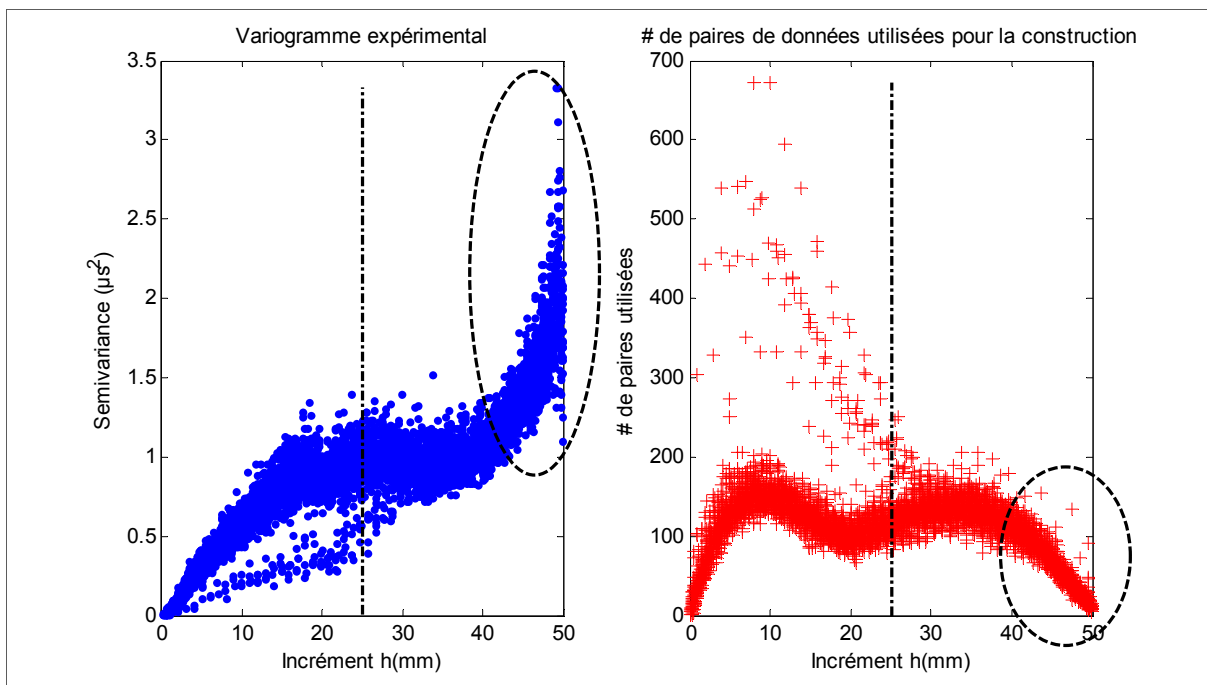


Figure 5-9 Paires des données utilisées lors de la construction du variogramme expérimental

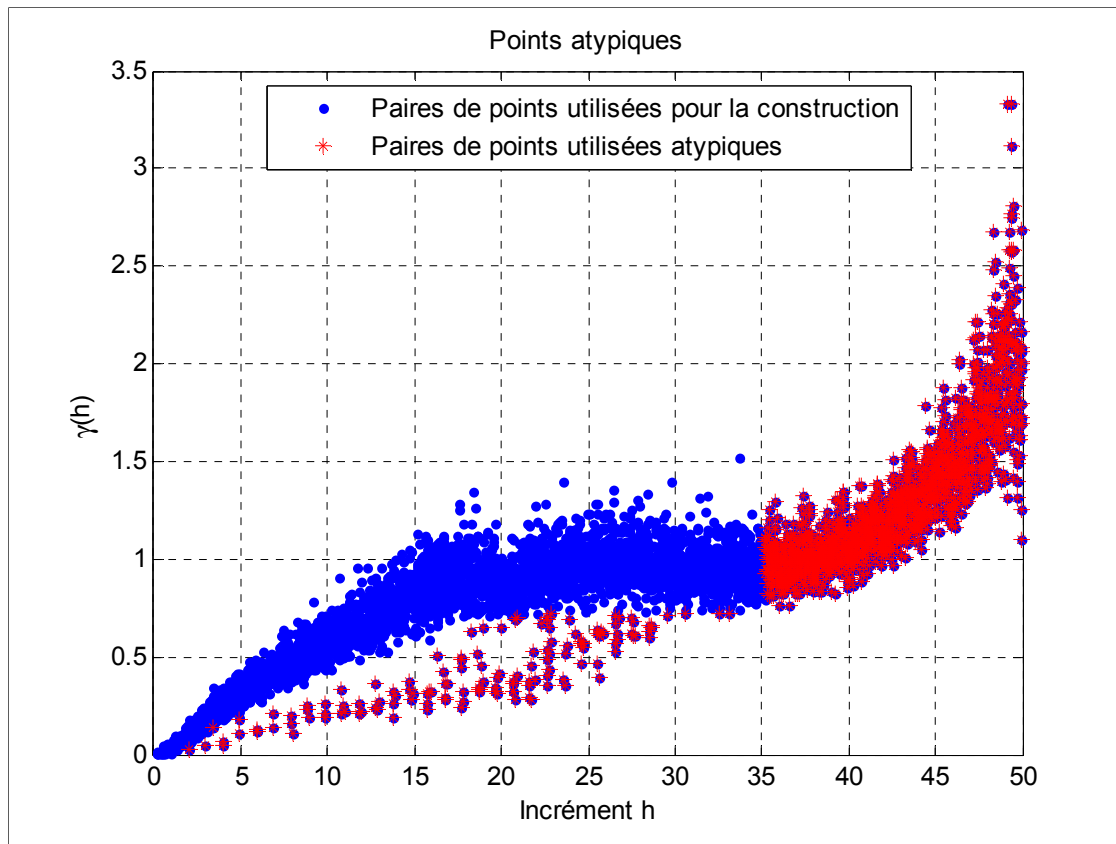


Figure 5-10 Nuage variographique et suppression des points atypiques

La Figure 5-10 montre l'échantillon des points à supprimer lors de la modélisation de ce variogramme. En plus de l'élimination de hauts incréments $h \geq 35$ mm, l'endroit où le variogramme décolle pour la deuxième fois, on supprime l'ensemble inférieur qui sous-estime la variance.

5.1.1.4 Ajustement du semivariogramme théorique

Une fois qu'on a construit le variogramme expérimental, on doit estimer ses paramètres par la méthode des moindres carrés (Cressie, 1985). En effet, on doit lui faire correspondre parmi les variogrammes théoriques, le plus représentatif de sa structure. Les résultats de trois ajustements sont présentés dans le Tableau 5-1.

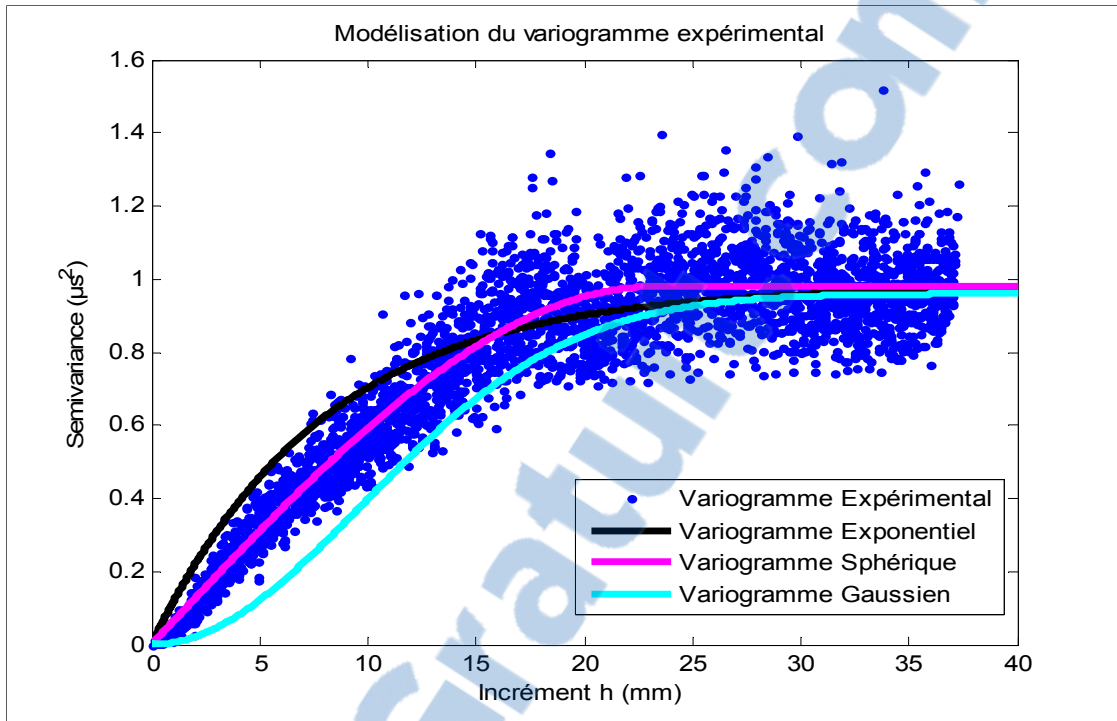


Figure 5-11 Modélisation du semivariogramme expérimental avec les semivariogrammes analytiques exponentiel, sphérique et gaussien

A première vue, la Figure 5-11 montre que le variogramme sphérique l'emporte comparativement aux deux autres. En effet, il s'ajuste le mieux avec l'ensemble des points du nuage variographique (voir Figure 5-12). Les deux autres ont pratiquement la même portée que le sphérique mais leurs premières parties (le début du variogramme) sont moins représentatives des faibles incréments h .

Tableau 5-1 Comparaison entre les trois modèles : sphérique, gaussien et exponentiel

Modèle	Effet de pépité	Portée	Palier
Sphérique	0	23.33 mm	$0.9755 \mu s^2$
Gaussien	0	25.69 mm	$0.9601 \mu s^2$
Exponentiel	0	23.80 mm	$0.9803 \mu s^2$

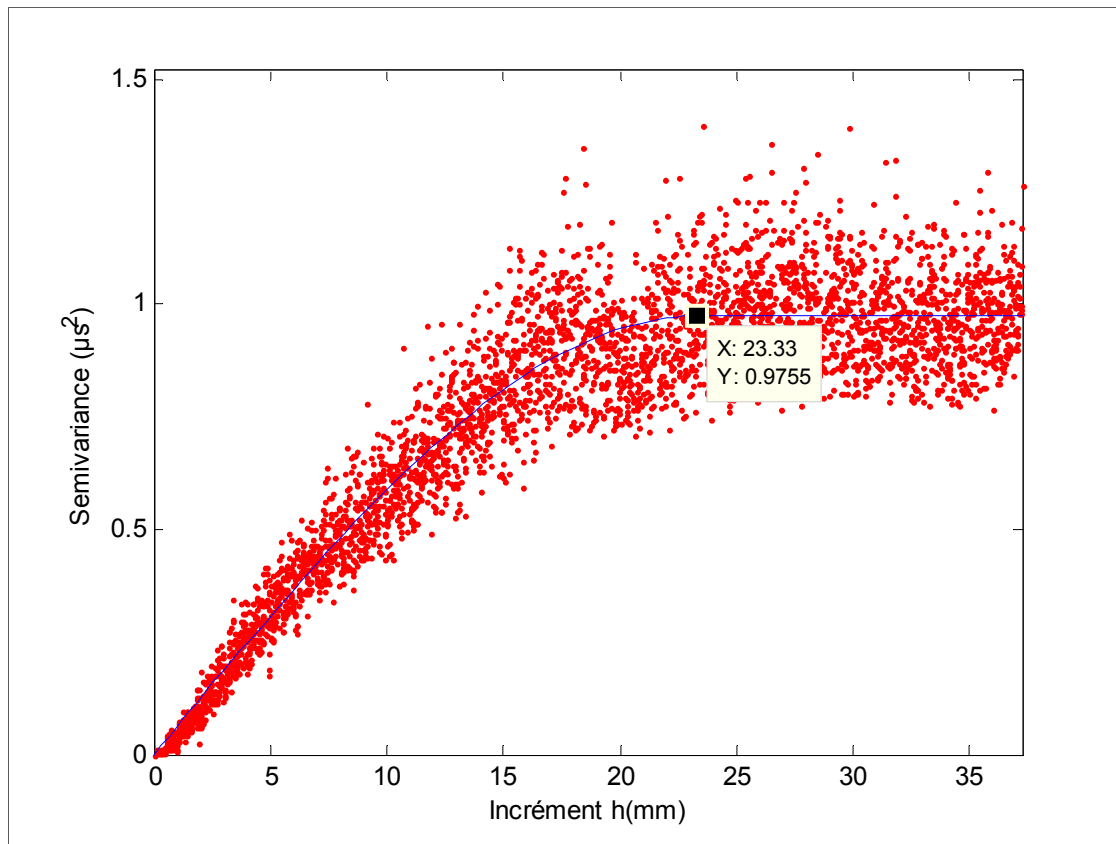


Figure 5-12 Ajustement du semivariogramme expérimental par la méthode des moindres carrés

5.1.1.5 Validation croisée

En général, plus qu'on alimente l'échantillon par des données plus le modèle prédictif sera robuste. En revanche, la validation croisée utilise le concept inverse : elle élimine temporairement une donnée (ou plusieurs à la fois) puis estime sa valeur par krigeage à l'aide du reste de l'échantillon et du modèle de variogramme ajusté par la méthode des moindres carrés. Ce processus sera répété pour plusieurs points (sous ensemble) de l'échantillon.

Dans notre cas d'étude, on a supprimé des échantillons des mesures (non pas juste une seule donnée) à chaque itération de façon aléatoire. En effet, on choisit aléatoirement un ratio de 1.03% de l'échantillon initial (18 données sur un total de 1746) et on estime le reste des 1721 mesures par un krigeage ordinaire en utilisant les paramètres optimisés par la méthode des

moindres carrés. Cette opération est réalisée pour les trois modèles empiriques sphérique, gaussien et exponentiel. À chaque fois on calcule le coefficient de corrélation linéaire ajusté entre les valeurs estimées par le krigeage et les valeurs initiales calculées par des ÉFs. Ensuite, on garde la position des points choisis qui correspondent au coefficient de corrélation le plus élevé, dans le but d'optimiser la position des données principales lors du krigeage et les simulations stochastiques conditionnelles par la suite.

La Figure 5-13 montre les histogrammes des trois modèles pour 1000 itérations de validation croisée. Il est clair que le modèle sphérique est plus représentatif comparativement aux deux autres.

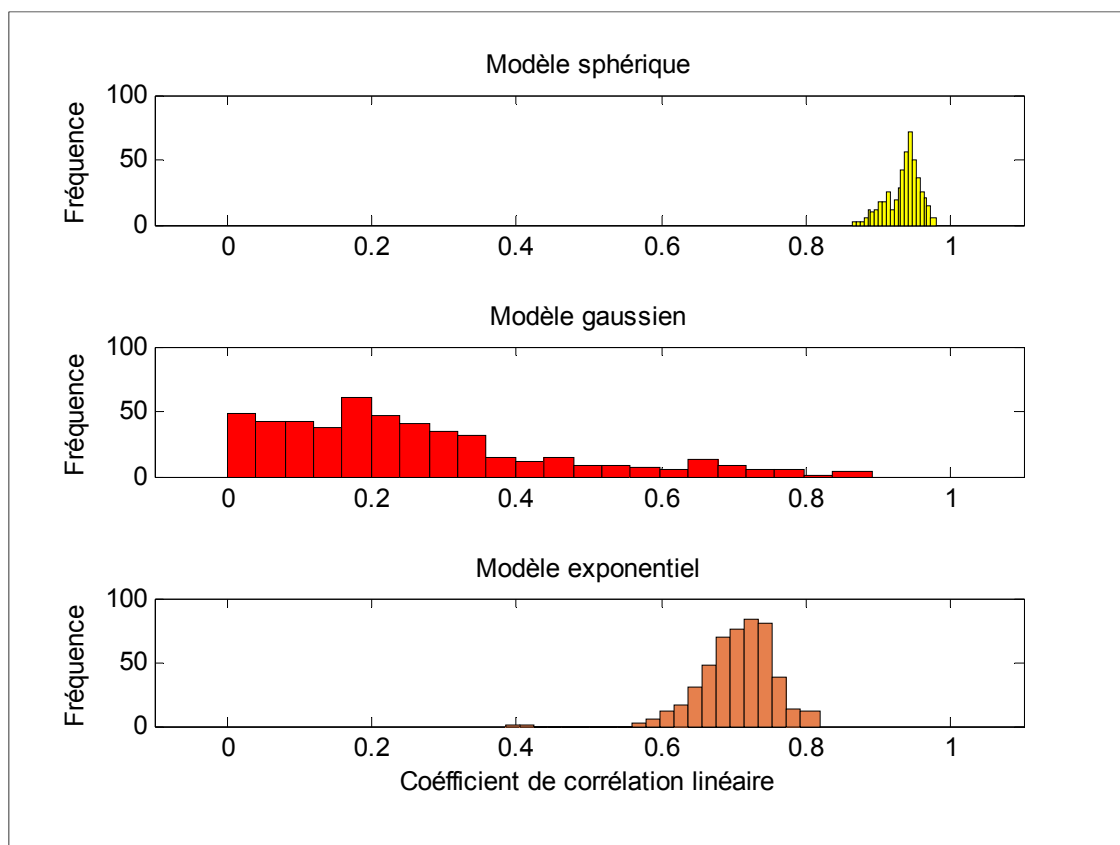


Figure 5-13 Validation croisée : Comparaison entre les coefficients de corrélations entre les trois modèles

Une fois qu'on a déterminé le variogramme analytique qui modélise la variabilité spatiale de la déformation suivant l'axe de traction, qu'on l'a ajusté par la méthode des moindres carrés et qu'on l'a validé par la validation croisée, on pourra confirmer la robustesse de ses paramètres et commencer le krigeage et les simulations stochastiques.

5.1.2 Krigeage

Dans cette partie du mémoire, on comparera les estimations par le krigeage simple (KS), ordinaire (KO) et universel (KU) dans le but de choisir le plus représentatif de la variabilité spatiale pour l'utiliser après lors du conditionnement des simulations stochastiques.

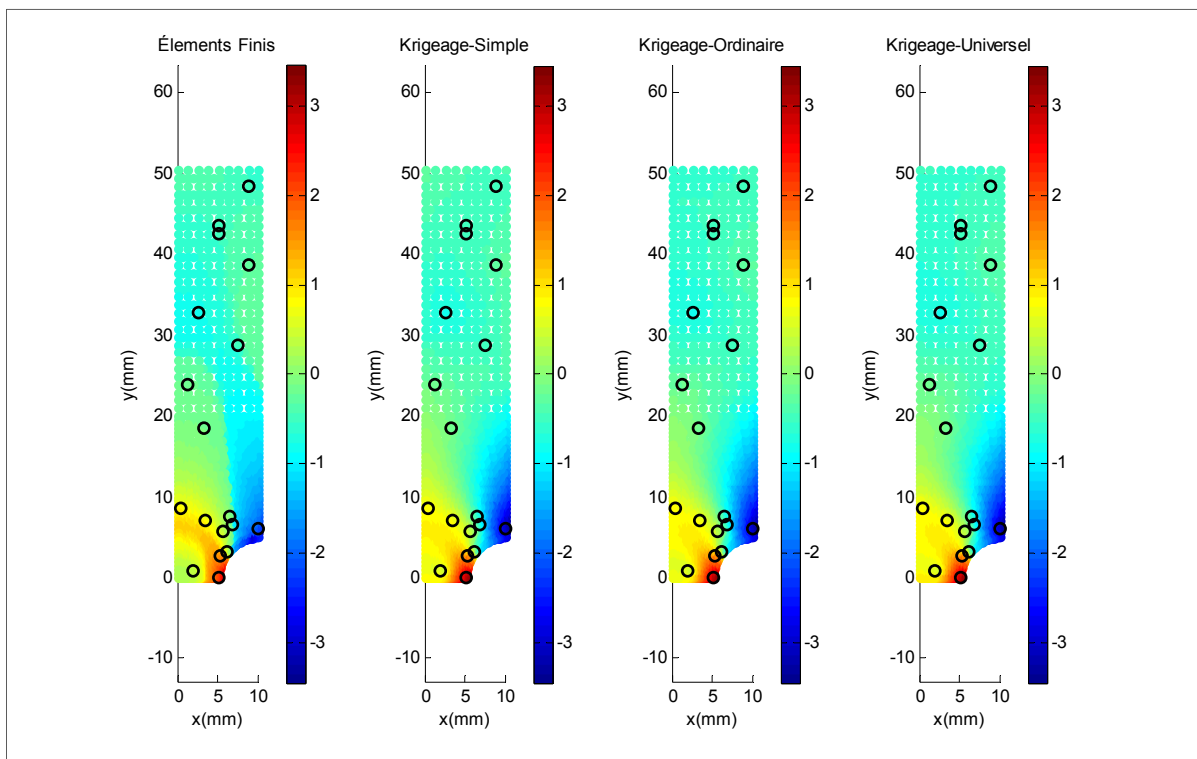


Figure 5-14 Krigeage simple, ordinaire et universel

La Figure 5-14 montre l'estimation de la déformation par les trois types de krigeage. Les données principales utilisées pour ces estimations sont représentées par des cercles noirs et sont choisies de façon aléatoire. Il est important de noter que le krigeage est capable de

reproduire la cartographie de la déformation donnée par les éléments finis avec un effet de lissage. Cette vraisemblance entre chaque type de krigeage et les déformations théoriques calculées par les simulations par ÉFs (figure complètement à gauche), sera validée par plusieurs tests statistiques.

Tableau 5-2 Coefficient de corrélation de Pearson entre les déformations théoriques et les trois types du krigeage

	Éléments Finis	Krigeage Simple	Krigeage Ordinaire
Krigeage Simple	0.974 0.000		
Krigeage Ordinaire	0.976 0.000	0.996 0.000	
Krigeage Univers	0.973 0.000	0.996 0.000	0.999 0.000
Cell Contents: Pearson rho P-Value			

Tableau 5-2 Le tableau ci-dessus montre un fort coefficient de corrélation linéaire entre les déformations théoriques calculées par les ÉFs et les trois types de krigeage. La nullité de la *p value* implique qu'on ne peut pas rejeter l'hypothèse que le coefficient de Pearson soit significatif.

Une deuxième méthode de validation consiste à comparer les fonctions cumulatives de probabilité de chaque estimateur avec celles des valeurs théoriques. La Figure 5-15 compare les 3 *cdf* des krigeages et montre une certaine ressemblance avec la *cdf* des ÉFs.

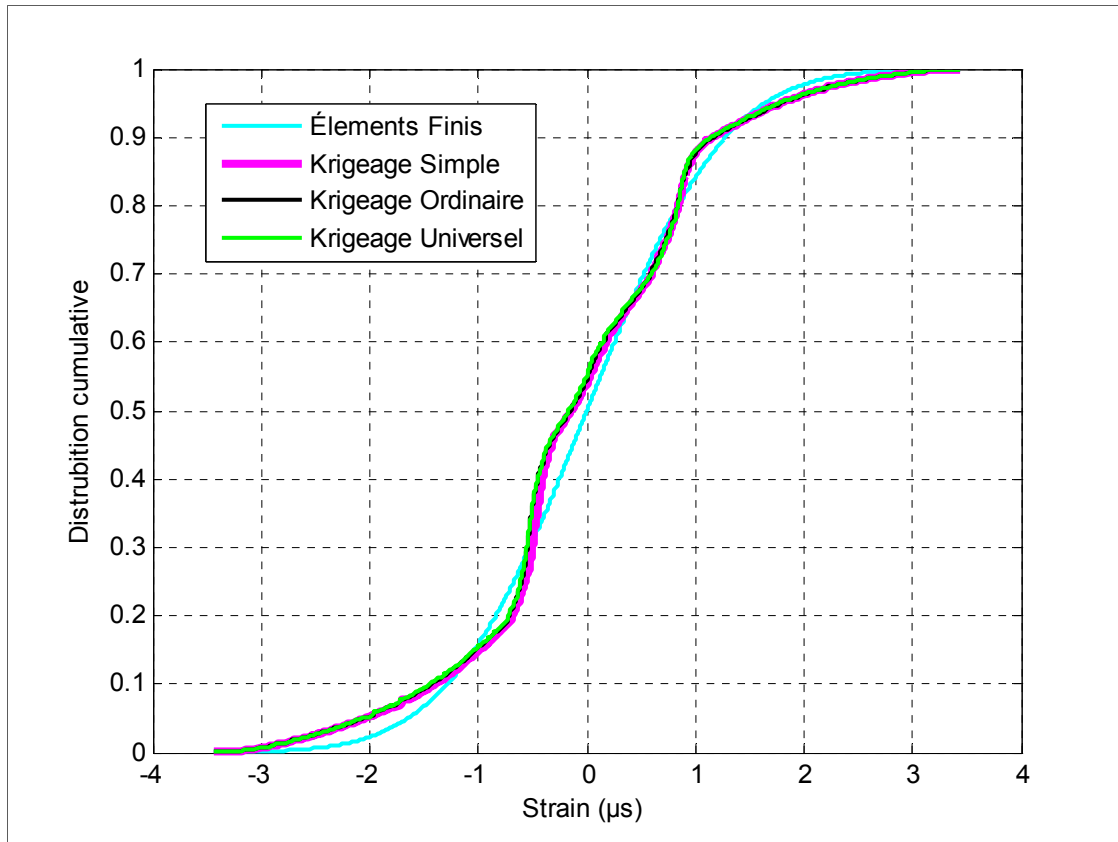


Figure 5-15 Comparaison entre les *cdf* de trois types de krigage avec celle de simulations par des éléments finis (*cdf* référentielle)

Une troisième validation du krigage est présentée par les semivariogrammes de la Figure 5-16, où on remarque une reproduction de celui des déformations calculées par les éléments finis (la référence). Ceci montre que statistiquement, les surfaces krigées se comportent comme la carte référentielle ÉFs, ce qui explique qu'on est capable de prédire les mesures manquantes en utilisant uniquement le krigage dans un premier temps mais on ne peut pas quantifier l'erreur sans le passage par des simulations stochastiques. On remarque aussi une petite variation de la portée des trois variogrammes du krigage suite à l'effet de lissage qui ramène les valeurs évaluées à la moyenne de l'estimateur ce qui, augmentera la variance en d'autres termes. Finalement, on a laissé le variogramme complet pour montrer que même à des grandes valeurs d'incrémentations, ce dernier est capable de reproduire la même variabilité du variogramme référentiel.

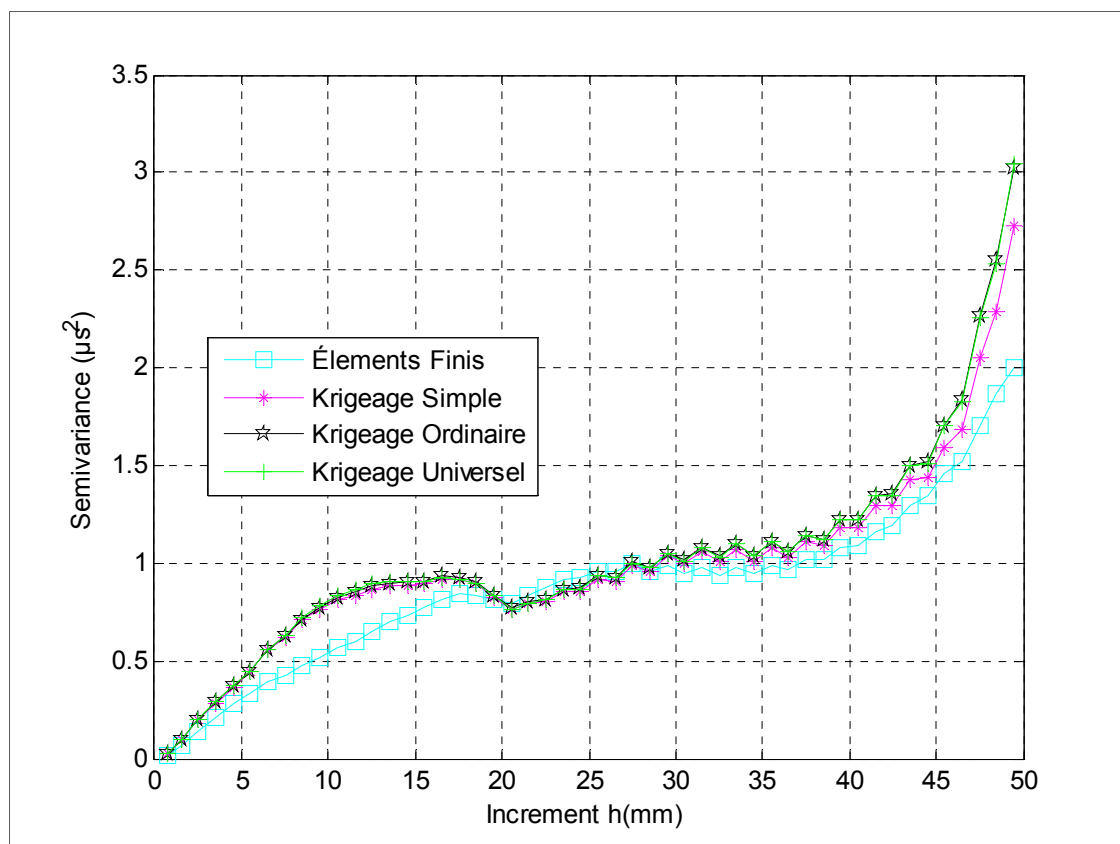


Figure 5-16 Les semivariogrammes des surfaces krigées KS, KO et KU

Tableau 5-3 Comparaison des coefficients de corrélations de Pearson entre les KS, KO et KU

Krigeage	Variance S	R-Sq	R-Sq(adj)
Simple	0.2712	93.9%	93.9%
Ordinaire	0.2677	94.1%	94.1%
Universel	0.2748	93.8%	93.8%

La Figure 5-17 montre la carte de la variance du krigeage pour les trois types. Elle ne représente pas l'erreur réelle sur la valeur estimée à cause de l'effet du lissage mais elle donne une idée sur l'allure de la propagation de l'erreur de prédiction. La valeur la plus probable de l'erreur sera simulée par la suite en utilisant les simulations stochastiques.

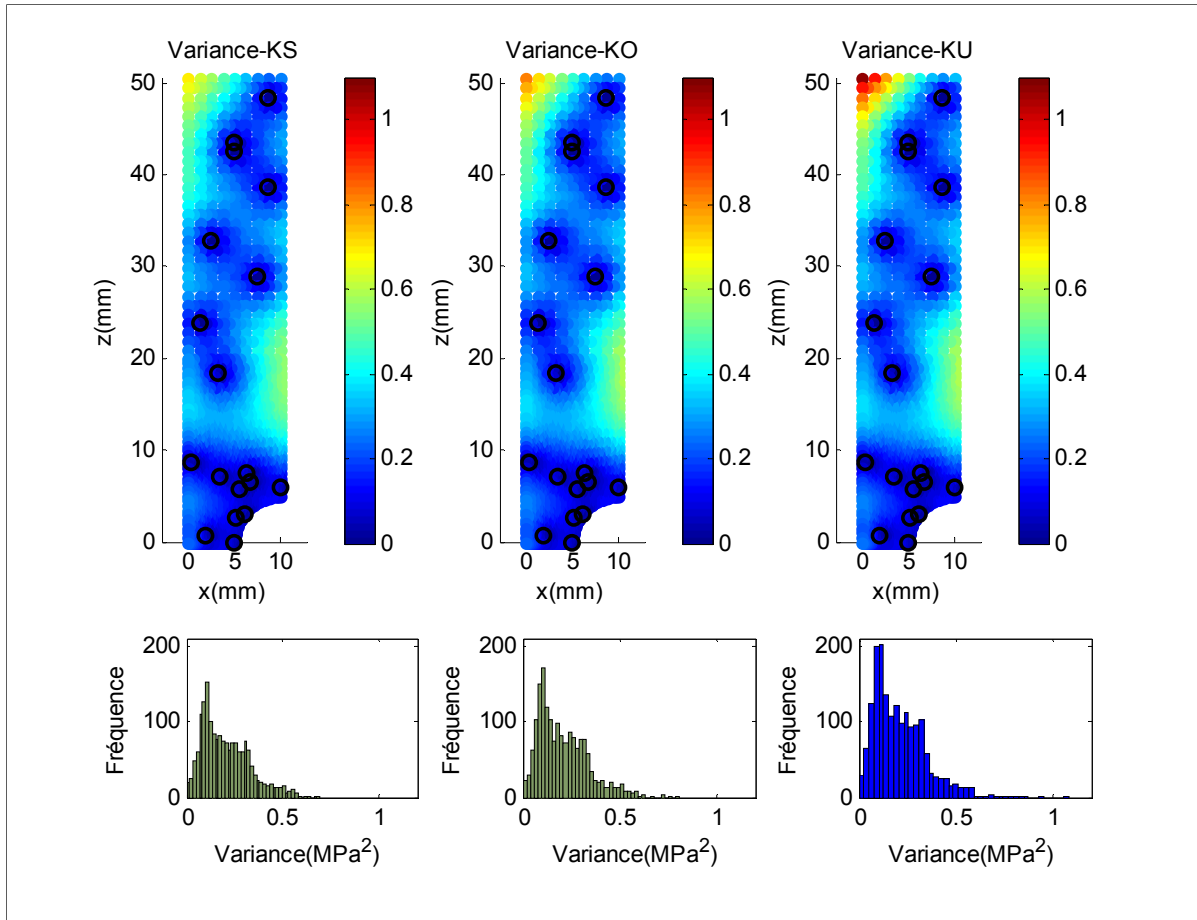


Figure 5-17 Variance de krigeage KS, KO et KU

En somme, comme l'estimateur est plus précis au sens où la variance de l'erreur d'estimation soit minimale et en tenant compte de la Figure 5-17 et le Tableau 5-3 on gardera le krigeage ordinaire pour le reste du rapport et pour le conditionnement des simulations stochastiques³.

³ Voir aussi le dernier chapitre de ce mémoire : Les recommandations, pour plus de justification du KO

Étude statistique des valeurs estimées par le krigeage ordinaire

Le Tableau 5-4 présente l'équation de régression linéaire (la droite d'ajustement) entre le krigeage ordinaire et la méthode ÉFs ainsi qu'une ANOVA, les coefficients de corrélation et une estimation de l'écart type de l'erreur du modèle.

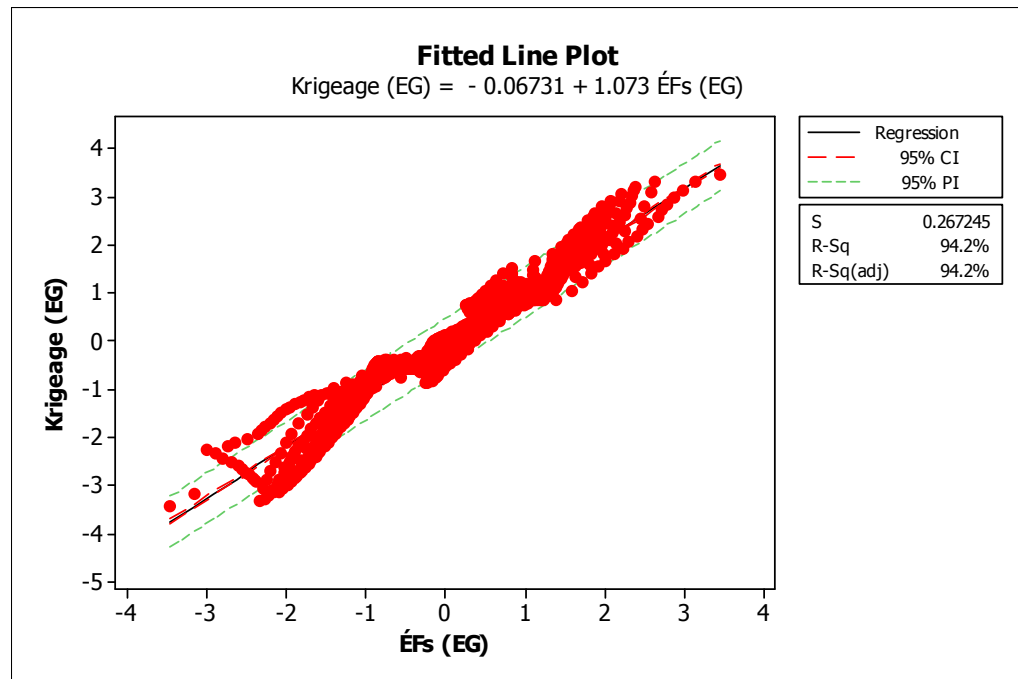


Figure 5-18 Régression linéaire simple entre le krigeage ordinaire et les ÉFs

Interprétation

Cette analyse de régression linéaire simple, voir Figure 5-18, génère une équation pour décrire la relation statistique entre l'estimation par krigeage et les données de référence calculées par les ÉFs. La valeur de R_{ajust}^2 indique que le krigeage explique 94.2% de la variance de déformation de l'éprouvette calculée par les ÉFs. Ce coefficient indique à priori que le modèle semble être adapté aux données et présente une capacité de prévision adaptée. Les résultats de la régression, Tableau 5-4 , indiquent que les coefficients de la droite de régression sont significatifs en raison de leurs faibles valeurs de *p-value* inférieure au risque du premier ordre $\alpha=5\%$ ce qui implique qu'il y a au moins un coefficient différent de zéro.

Tableau 5-4 Régression linéaire simple entre le krigeage ordinaire et les simulations par des ÉFs

```

The regression equation is
Krigeage (EG) = - 0.0673 + 1.07 ÉFs (EG)

Predictor      Coef      SE Coef      T      P      VIF
Constant      -0.067311  0.006396    -10.52  0.000
ÉFs (EG)      1.07317   0.00640     167.73  0.000  1.000

S = 0.267245   R-Sq = 94.2%   R-Sq(adj) = 94.2%

PRESS = 124.992   R-Sq(pred) = 94.14%

Analysis of Variance

Source          DF      SS      MS      F      P
Regression      1    2009.4  2009.4  28134.31  0.000
Residual Error  1744    124.6    0.1
Total          1745    2133.9

```

Ce tableau fournit plusieurs informations : il fournit l'estimation par la méthode des moindres carrés de la constante $\hat{\beta}_0$ et la pente $\hat{\beta}_1$ de la droite d'ajustement.

Test d'hypothèse pour la pente de régression

La troisième colonne de sortie T fournit des statistiques de tests. Dans la régression linéaire, on souhaite tester la signification de la pente de la droite.

- **H0 : hypothèse nulle = [Le coefficient de la pente de la droite β_1 est nul (il n'y a aucun rapport explicatif entre les variables explicatives et dépendantes)]**
- **H1 : hypothèse alternative = [$\beta_1 < 0$ ou $\beta_1 > 0$ ou $\beta_1 \neq 0$].**

La statistique de test t est égale à l'estimation du paramètre de la pente de la droite divisée par son écart type. Ce rapport suit une loi de Student avec $n-2=1744$ degrés de liberté.

Pour chaque coefficient de régression $\hat{\beta}_j ; j \in \{0,1\}$, on construit un intervalle de confiance en utilisant la formule suivante

$$IC_{100 \times (1-\alpha)\%} = \left[\hat{\beta}_j \pm \underbrace{t_{\alpha, n-p-1}}_{\substack{\text{valeur à partir de} \\ \text{la table de Student}}} \times s(\hat{\beta}_j) \right] \quad (5.1)$$

- $IC_{\text{pente}(95)\%} = 1.07317 \pm 1.96 * 0.00640 = [1.060626 ; 1.085714]$
- $IC_{\text{constante}(95)\%} = -0.067311 \pm 1.96 * 0.006396 = [-0.07985 ; -0.05477]$

L'intervalle de confiance des coefficients de l'équation de régression permet de conclure sur l'existence ou non d'un lien entre la variable explicative correspondante et la réponse en tenant compte des autres variables explicatives ajustées par le modèle.

- Si l'intervalle de confiance du coefficient $\hat{\beta}_j$ exclut la valeur zéro, on conclut une association linéaire significative entre la variable et la réponse.
- Si l'intervalle de confiance inclut la valeur zéro, on conclut qu'on tenant compte des autres explications, la variable en question n'apporte aucune information supplémentaire utile pour la prédiction de la réponse.

Comme le montre le Tableau 5-5, tous les intervalles de confiance n'incluent pas la valeur nulle, ce qui implique qu'ils sont tous significatifs pour ce modèle de régression linéaire.

Tableau 5-5 Intervalle de confiance des coefficients de prédiction

Coefficient de la droite d'ajustement entre le krigeage ordinaire et les ÉFs		Intervalle de confiance @ 95%	
		minimum	maximum
Constante $\hat{\beta}_0$	- 0.06731	- 0.07985	- 0.05477
Pente $\hat{\beta}_1$	1.07317	1.060626	1.085714

La table de l'analyse de variance montre une valeur de $F = 28134.31 > \hat{F}_{théorique}$, c'est un critère de la bonne signification globale de la régression avec une *p-value* nulle.

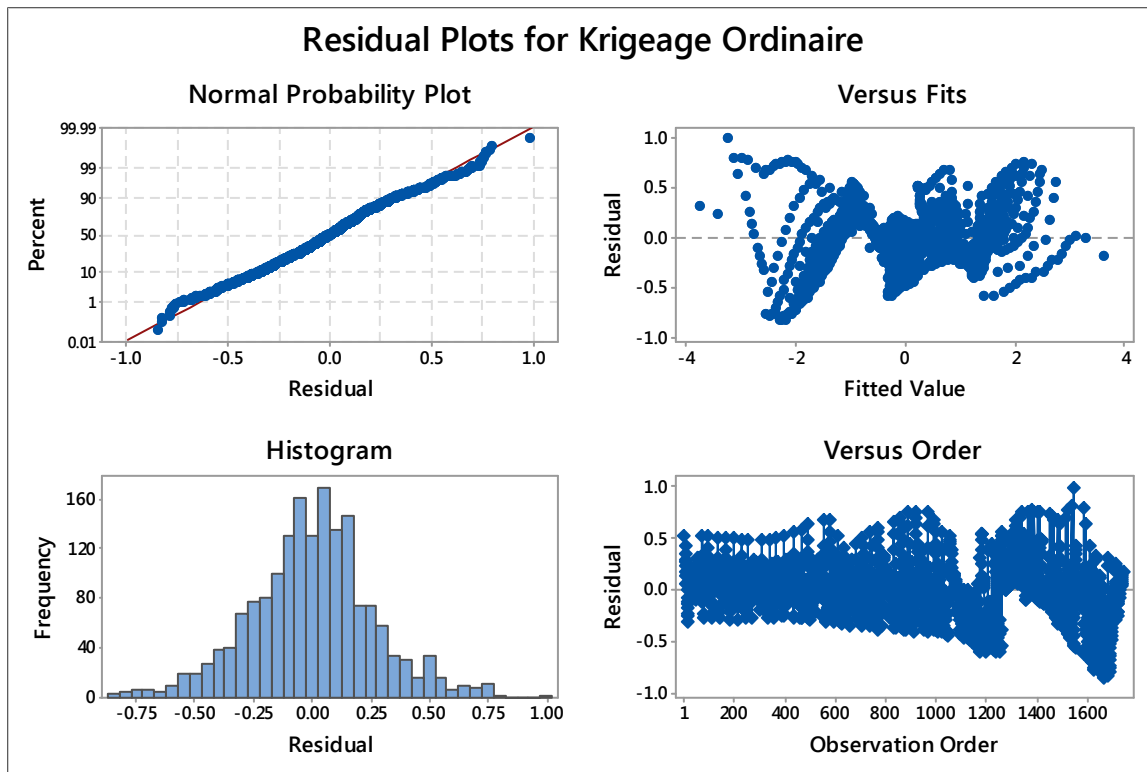


Figure 5-19 Analyse des résidus du krigeage ordinaire

Le Tableau 5-4 et la Figure 5-19 montrent que le krigeage est un bon estimateur pour prédire les mesures manquantes des déformations à partir d'uniquement 18 valeurs calculées par des simulations ÉFs. L'analyse de résidus est primordiale pour confirmer que cet estimateur est représentatif des déformations de l'éprouvette.

Histogramme des valeurs résiduelles

L'histogramme des résidus normalisés se présente sous la forme de la cloche typique d'une distribution gaussienne. Le pic de la distribution est concentré au alentour de la valeur nulle, signe de l'absence d'un biais systématique en moyenne. On a pratiquement une égalité entre la médiane (0.00842) et la moyenne (0.00004). La distribution est quasi symétrique, ce qui s'approche à la distribution gaussienne.

Droite de Henri des valeurs résiduelles

Le graphe de la droite de Henri des valeurs résiduelles forme une droite dans l'espace des quantiles et seconde l'hypothèse que les valeurs résiduelles suivent une loi normale. Il y a quelques points qui s'écartent de l'ensemble vers l'extrémité, ceci peut être expliqué par la présence des valeurs atypiques. Les décollements des points vers les extrémités de la droite de Henri, permet de déceler les écarts à la tendance linéaire, la situation d'hétérogénéité de variance de l'erreur ou les valeurs atypiques. Le test de normalité des résidus en utilisant la statistique d'Anderson-Darling ($AD=1.796$) confirme qu'on peut rejeter l'hypothèse nulle (la distribution des résidus est normale) puisque la *p-value* < 0.005 est inférieure au risque $\alpha=5\%$. En effet, le test ne donne pas la bonne décision car on a un nombre très élevé de données (1746) et plus ce dernier augmente plus le test doit ajuster strictement les deux distributions cumulatives de probabilité d'une façon exacte et le moindre dépassement annule la normalité.

Valeurs résiduelles en fonction des valeurs ajustées

On a une présence d'une répartition aléatoire des valeurs résiduelles de deux côtés de l'axe des abscisses. On remarque l'éloignement de quelques points qui pourront être des valeurs atypiques. De plus, on n'a pas une prédominance des valeurs résiduelles positives ou négatives. En effet, on n'a pas une corrélation entre la valeur absolue des résidus avec les valeurs de l'axe des abscisses, il n'y a pas d'hétéroscédasticité (nuage de point en tromblon).

Valeurs résiduelles en fonction de l'ordre

Le quatrième diagramme (en bas à droite) de la Figure 5-19, présente les valeurs résiduelles dans l'ordre où les données ont été collectées et ne montre pas des erreurs non aléatoires en particulier pour les fonctions temporelles. Ce graphe ne permet pas d'identifier une dérive apparente en fonction du temps ou bien une corrélation entre les erreurs.

5.1.3 Simulations conditionnelles

Les simulations sont considérées conditionnelles si elles reproduisent à chaque itération, d'une part les 18 données choisies précédemment, et d'une autre part, l'histogramme de valeurs théoriques ÉFs. La Figure 5-20 montre une comparaison entre cinq simulations choisies aléatoirement parmi un ensemble de 500 simulations séquentielles conditionnelles. Les simulations reproduisent les valeurs encadrées par des cercles noirs car elles sont les valeurs expérimentales conditionnantes comme le montre la Figure 5-20 .

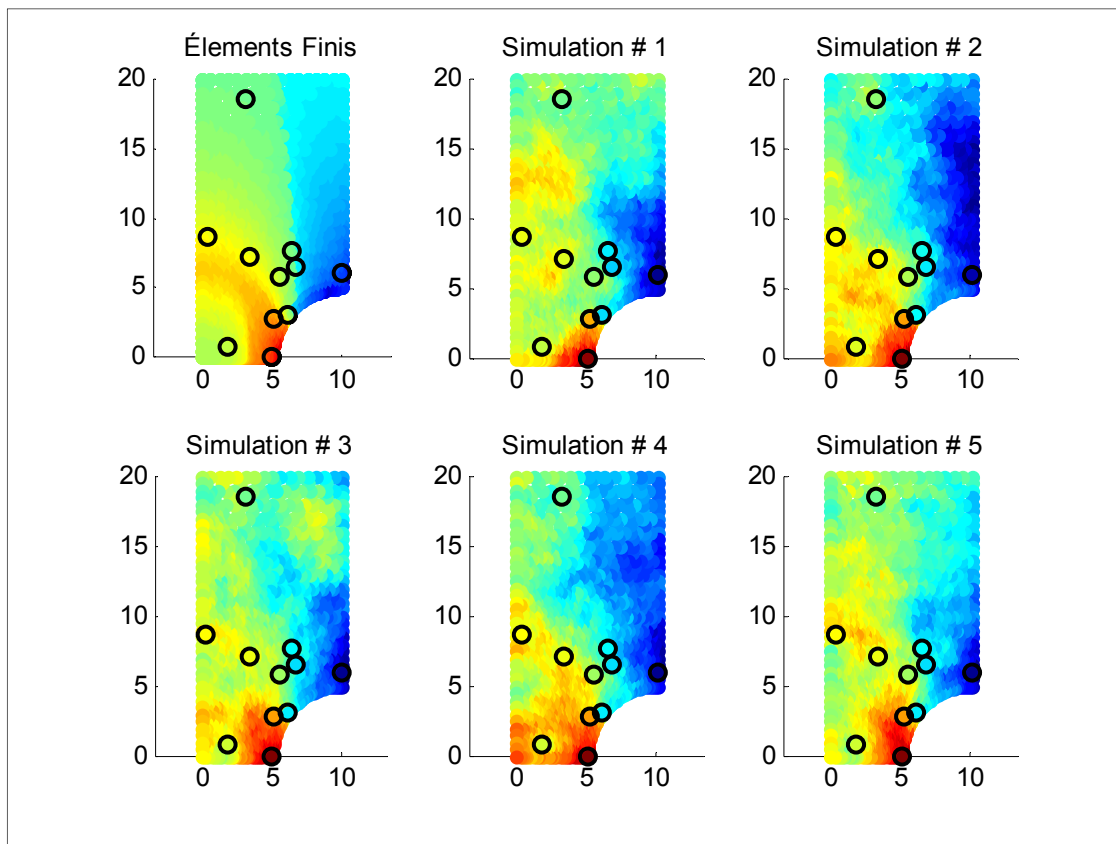


Figure 5-20 Comparaison entre les déformations ÉFs et les cinq exemples de simulations conditionnelles choisies aléatoirement parmi l'ensemble de 500 simulations initiales

5.1.3.1 Comparaison entre les SGCs et les SSCs

Variation du nombre de simulations

Le but de cette partie du rapport est dans un premier temps de valider les simulations stochastiques en comparant les deux algorithmes discutés précédemment : les simulations séquentielles conditionnelles SSCs réalisées en utilisant les fonctions MatLab® discutées en méthodologie et les simulations gaussiennes conditionnelles SGCs en utilisant le logiciel SGeMS (*Stanford Geostatistical Modeling Software*). D'une autre part, on veut prouver qu'on a déjà utilisé un nombre suffisant de simulations stochastiques pour quantifier la variance de l'estimateur en comparant respectivement les résultats de 100, 250 et 500 simulations.

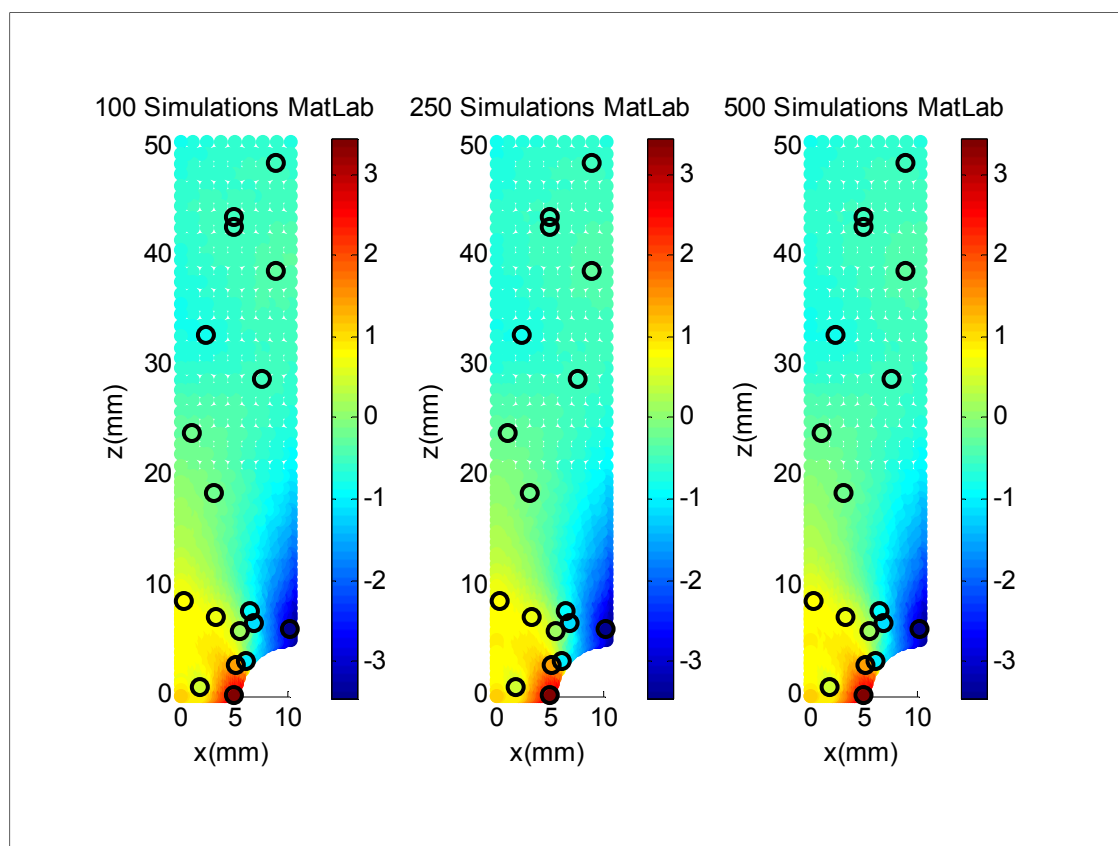


Figure 5-21 Comparaison entre la cartographie de la moyenne de 100, 250 et 500 simulations stochastiques

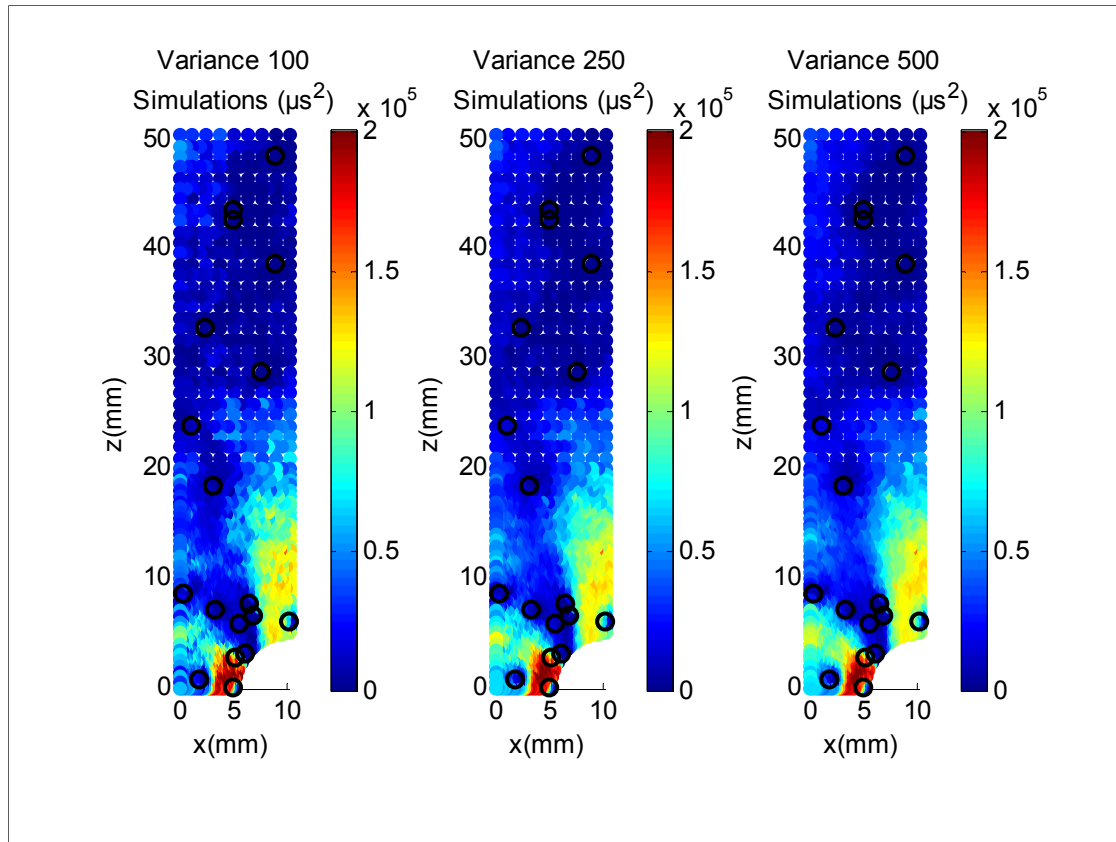


Figure 5-22 Comparaison entre la cartographie de la variance de 100, 250 et 500 simulations stochastiques

La Figure 5-21 montre une vraie ressemblance entre les trois cartes. Le nombre de simulation n'a pas d'influence sur l'estimation. En effet, on n'a pas besoin de réaliser une analyse de convergence pour ce cas, car les trois cartes montrent qu'on a déjà atteint un état stationnaire et que le fait d'augmenter le nombre de simulations ne modifie pratiquement pas le résultat final. Ceci se confirme par la suite (Figure 5-22) par les trois cartes de variance (représentées dans un espace gaussien) montrant la même allure. De plus, la Figure 5-23 montre une excellente corrélation entre les deux types de simulations et même entre différents nombres de simulations. Les coefficients de corrélation linéaire de Pearson confirment une excellente concordance entre les différentes parties du graphique matriciel présentées par la Figure 5-23 (les *p-value* sont toutes nulles).

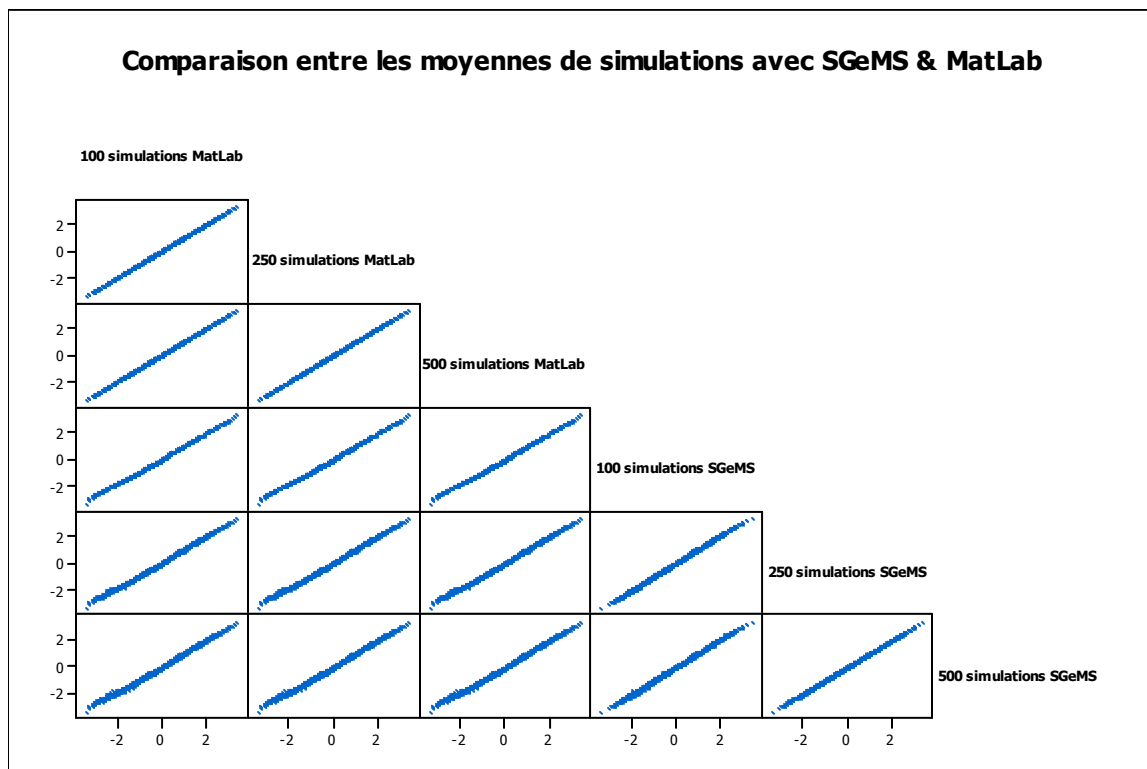


Figure 5-23 Graphique matriciel diagonal en fonction du nombre des simulations

Tableau 5-6 Coefficients de corrélations des simulations stochastiques

Nombre Simulations		SSCs (MatLab®)			SGSs (SGeMS)		
		100	250	500	100	250	500
SSCs (MatLab®)	100						
	250	≈1.000					
	500	≈1.000	≈1.000				
SGSs (SGeMS)	100	0.999	0.999	0.999			
	250	0.998	0.998	0.998	0.999		
	500	0.998	0.998	0.998	0.999	≈1.000	

Les coefficients de corrélation pour les 6 moyennes des simulations sont excellents et sont supérieurs à 99.8% dans tous les cas avec un *p-value* quasi nul.

De plus, cette affirmation se confirme clairement avec le traçage des 3 semivariogrammes de chaque type de simulations. Ces semivariogrammes présentent la même structure de dépendance spatiale entre elles, ce qui confirme leurs équivalences. La déviation, proche de la portée des semivariogrammes, comparativement au semivariogramme original s'explique par l'effet de lissage du krigeage. Vu que cet effet a tendance à ramener les estimations vers la moyenne, on aura donc une diminution de la portée comparativement à celle donnée par la cartographie des éléments finis.

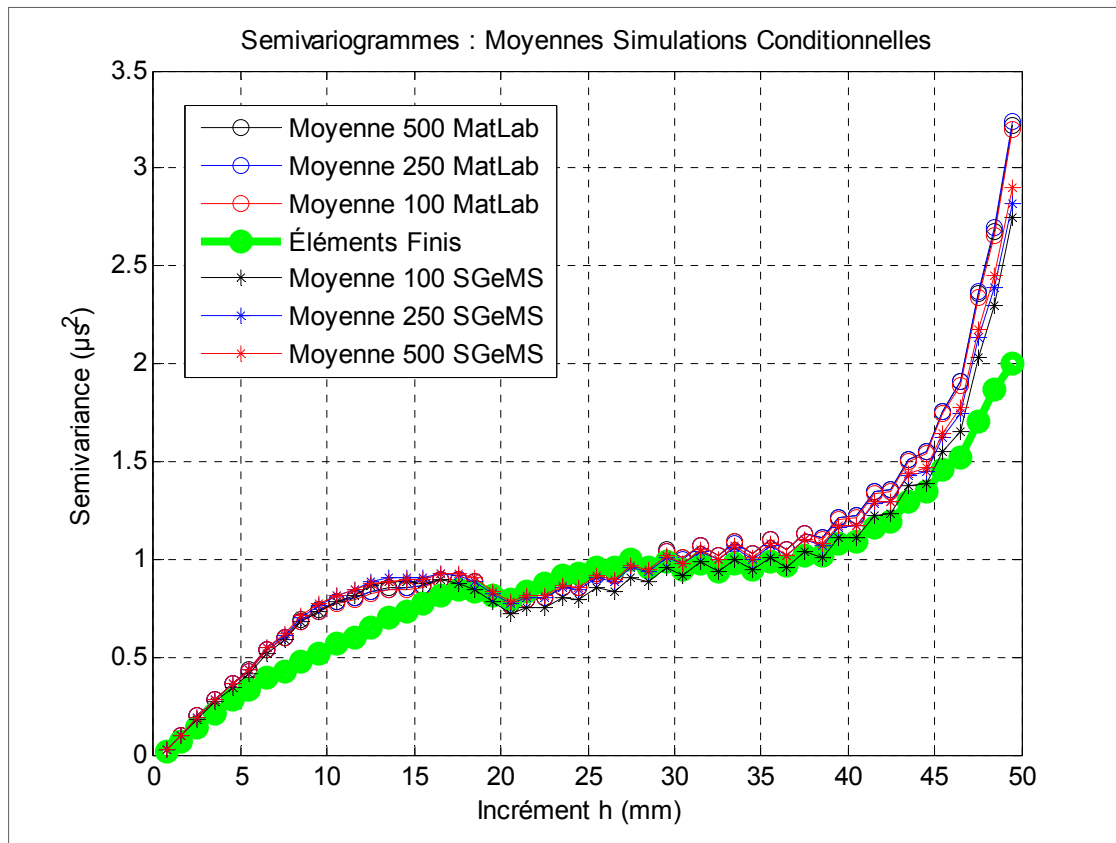


Figure 5-24 Comparaison entre les semivariogrammes de 100, 250 et 500 simulations stochastiques SSCs et SGSs

5.1.3.2 Validation statistique des résultats des simulations conditionnelles

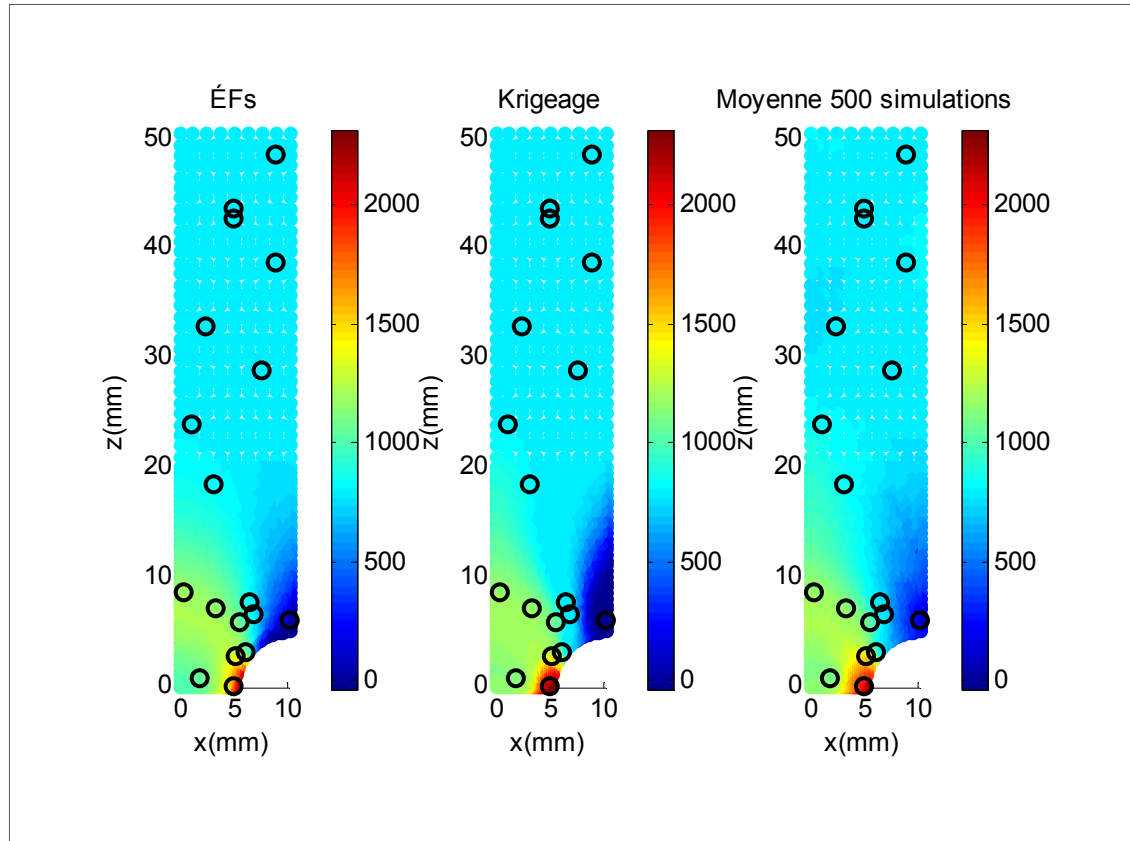


Figure 5-25 Cartographie des déformations simulées par ÉFs et leur prédiction par les simulations séquentielles conditionnelles

La cartographie de la moyenne des simulations stochastiques est semblable à celle de la référence ÉFs. L'effet de lissage est toujours présent surtout proche des zones de tension et compression au niveau de l'encoche. Ceci est tout à fait justifiable car les simulations sont conditionnées par le krigage. La Figure 5-26 montre d'avantage cette corrélation, avec deux histogrammes semblables et un nuage de corrélation rétréci sur la première bissectrice.

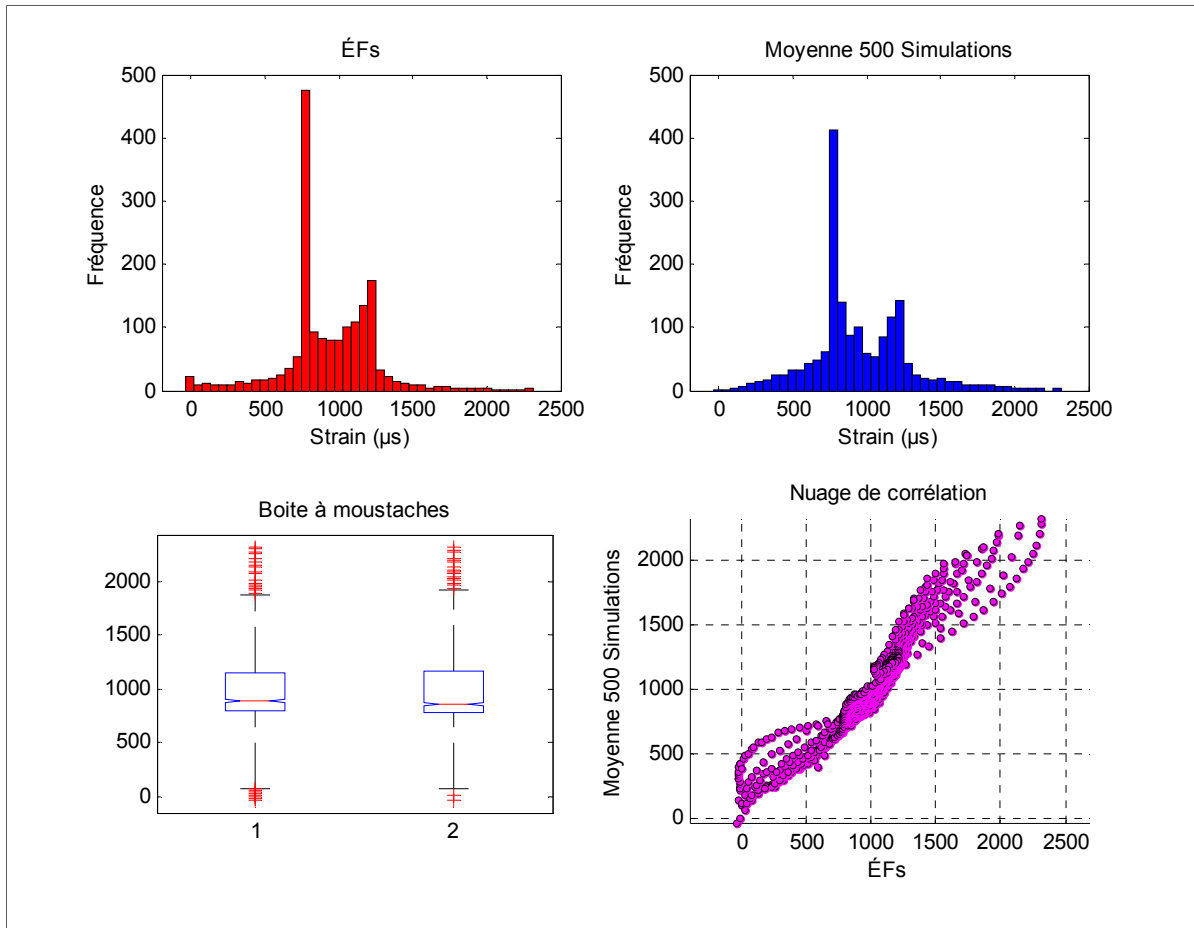


Figure 5-26 Histogrammes des déformations calculées par la méthode ÉFs et leurs prédictions par les 500 simulations séquentielles conditionnelles

5.1.3.3 Analyse des résidus de la moyenne de 500 simulations conditionnelles

L'analyse des résidus montre un maxima proche de la valeur nulle, ce qui prouve que le modèle utilisé est un estimateur consistant et qu'il est capable de reproduire les déformations de l'éprouvette. La moyenne des résidus est égale à $0.0 \mu\text{S}$ avec un intervalle de confiance de $[-4.055 ; 4.055]$ à 95% avec une médiane de $-15.21 \mu\text{S}$ et un intervalle de confiance de $[-16.76 ; -13.93]$ à 95%. Finalement, le premier quartile est égal à $-36.56 \mu\text{S}$ et le 3^{ème} quartile est de $8.01 \mu\text{S}$, ce qui montre encore une fois une robustesse du modèle des simulations stochastiques.

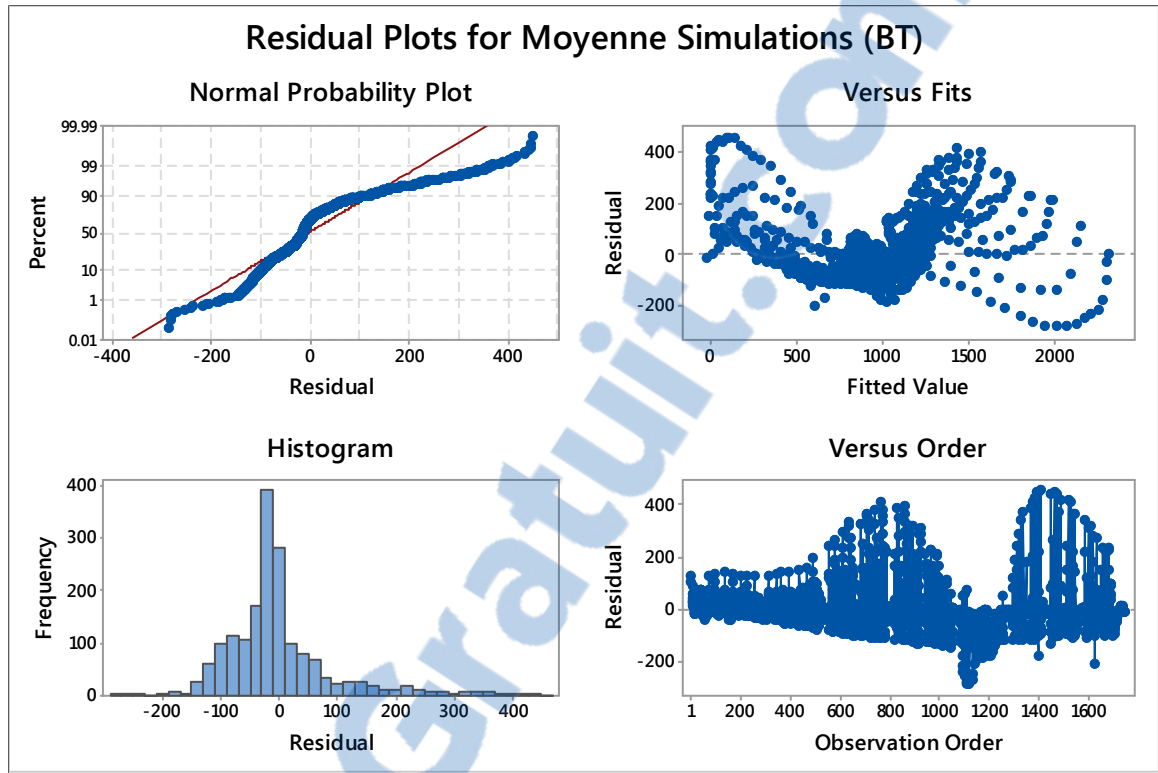


Figure 5-27 Histogramme des résidus de la moyenne de 500 simulations conditionnelles.

Les graphiques ci-dessus montrent qu'on respecte les hypothèses des moindres carrés et qu'on peut générer des estimations de coefficient sans biais avec une variance minimale. Le Tableau 5-7 présente le modèle avec les valeurs de coefficients estimés ainsi que des valeurs plus précises notamment les coefficients de corrélation et une estimation de l'écart type de l'erreur du modèle.

L'intervalle de confiance des coefficients de l'équation de régression permet de conclure sur l'existence ou non d'un lien entre la variable explicative correspondante et la réponse en tenant compte des autres variables explicatives ajustées par le modèle.

$$\circ IC_{pente(95)\%} = 0.995104 \pm 1.96 \times 0.007181 = [0.9882 ; 1.0092]$$

$$\circ IC_{constante(95)\%} = 16.323 \pm 1.96 \times 7.118 = [2.3717 ; 30.2743]$$

Tableau 5-7 Régression linéaire simple entre les valeurs de déformations calculées par les ÉFs et la moyenne de 500 simulations conditionnelles

The regression equation is
Moyenne des simulations (TGI) = 16.3 + 0.995 ÉFs

Predictor	Coef	SE Coef	T	P	VIF
Constant	16.323	7.118	2.29	0.022	
ÉFs	0.995104	0.007181	138.58	0.000	1.000

S = 96.7665 R-Sq = 91.7% R-Sq(adj) = 91.7%

PRESS = 16415875 R-Sq(pred) = 91.63%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	179818605	179818605	19203.68	0.000
Residual Error	1744	16330396	9364		
Total	1745	196149001			

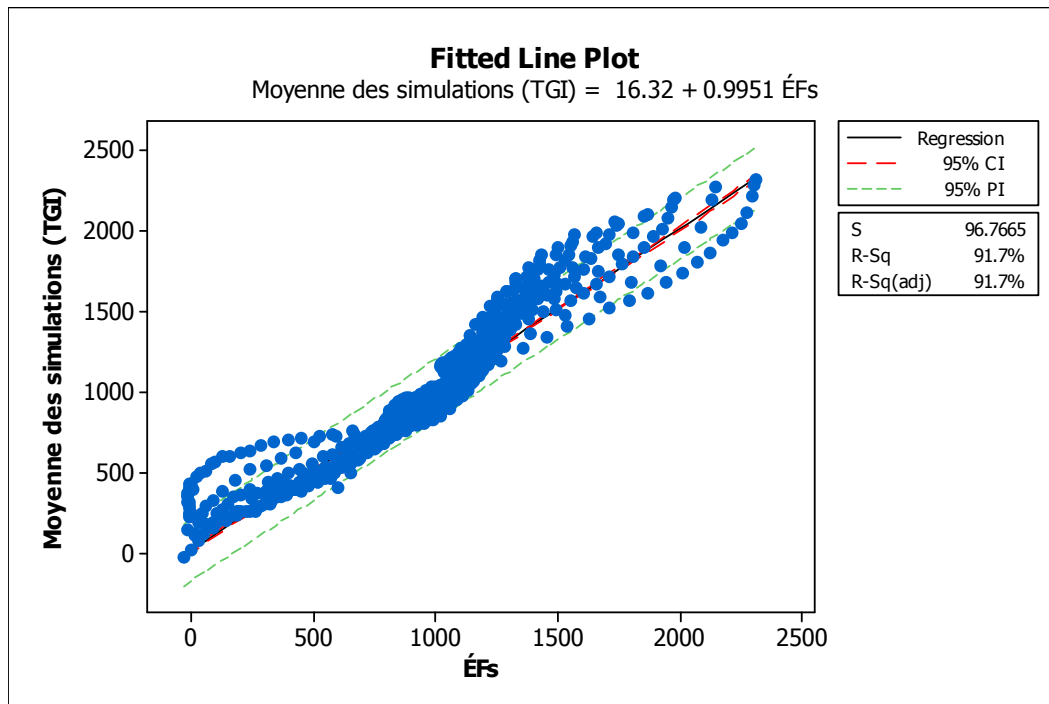


Figure 5-28 Droite d'ajustement entre les estimations calculées par ÉFs et la moyenne de 500 simulations conditionnelles

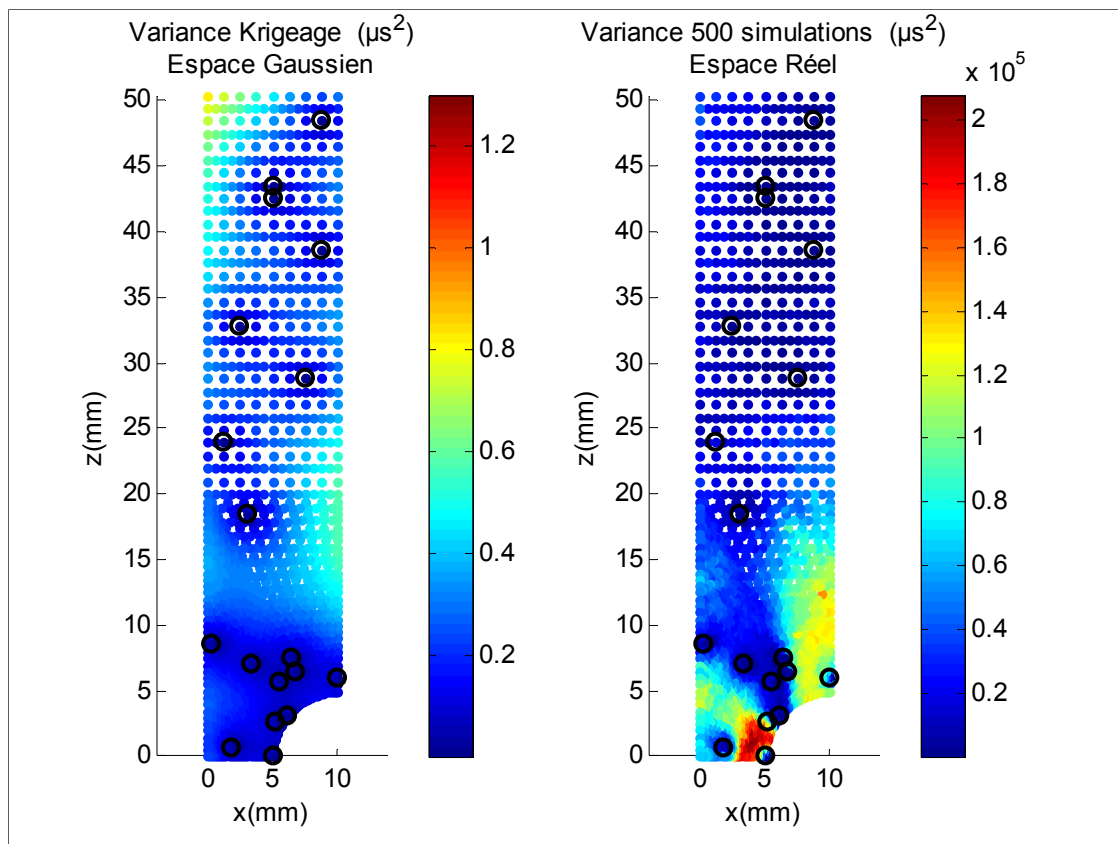


Figure 5-29 Comparaison entre la variance du krigeage ordinaire et la variance de la moyenne de 500 simulations conditionnelles

La Figure 5-29 montre une comparaison entre la variance du krigeage simple et la variance de la moyenne de 500 simulations conditionnelles. On a indiqué précédemment que la variance de krigeage ne donne pas une bonne estimation de l'erreur mais juste une allure de sa propagation dans l'espace; plus qu'on s'éloigne du point de mesure plus la variance s'intensifie.

Il est clair qu'avec les simulations, on est plus précis en calculant cette variance, car le degré d'intensité de l'erreur dépasse l'effet de lissage causé par le krigeage comme le montre les zones de compression et de tension de l'éprouvette. Plus on s'éloigne de la mesure et plus l'erreur s'intensifie rapidement comparativement au krigeage qui est incapable d'expliquer par l'effet du lissage.

Le coefficient de variation (CV), exprimé en pourcentage, permet de mesurer la dispersion relative des déformations de l'éprouvette nœud par nœud. Il compare la variabilité de deux séries : espérance de la déformation représentée par la moyenne de 500 simulations conditionnelles ainsi que son écart type en chaque point du maillage. Le CV est défini comme le rapport de l'écart type sur la moyenne. Il donne une idée de l'homogénéité de la série de mesure. Généralement, si la valeur est inférieure à 15%, on dit que les données sont homogènes, sinon ils seront considérées comme hétérogènes. Plus ce coefficient est élevé et plus la dispersion autour de la moyenne est grande. Plus la valeur de ce coefficient est faible, plus l'estimation est précise.

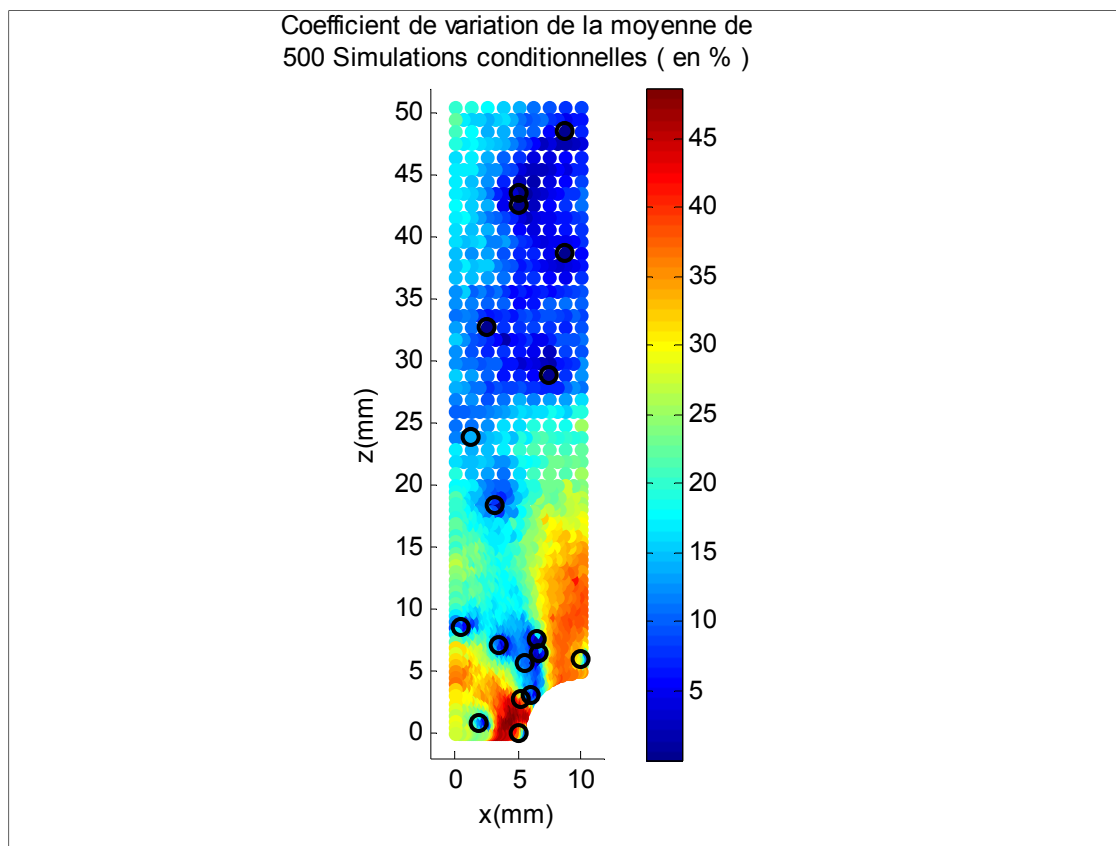


Figure 5-30 Coefficient de variance de la variance moyenne de 500 simulations conditionnelles

Dans notre cas d'étude, on remarque que le modèle montre un fort CV proche des zones où la compression et la tension sont extrêmes. Ceci est dû principalement au manque de mesures

dans ces régions et que le gradient est très brutal au niveau de l'arc de l'encoche. La région entre les deux possède un coefficient de variance qui ne dépasse pas 15%. Ainsi, le nombre et la position des jauges de mesures des déformations jouent un rôle essentiel sur la qualité de la prédiction surtout lorsque le gradient devient plus critique proche de la zone de l'encoche. Plus qu'on s'éloigne de la mesure, plus que le CV augmente et la qualité de la prédiction des déformations sera réduite.

La Figure 5-31 montre l'impact du nombre des jauges de déformations utilisées ainsi que l'influence de leurs positions. En effet, à chaque itération sur un total de 100, on fixe le nombre de jauge à placer tout en s'assurant de coller une sur l'arête où on a une compression maximale et une autre au centre de l'encoche où règne une déformation maximale. Le reste des jauges est reparti aléatoirement sur l'éprouvette. Dans le cas où on colle juste 3 jauges de déformation, on remarque qu'il est impossible de trouver une bonne localisation pour détecter la variabilité du semivariogramme expérimental construit à partir des déformations des ÉFs. De plus, même si on ajoute 2 jauges, on n'améliore pratiquement pas la situation. En revanche, on commence à repérer le semivariogramme des ÉFs à partir de 7 jauges, et plus que le nombre de jauge augmente, plus on aura de chance de reproduire la variabilité des déformations calculées par les éléments finis.

On ne cherche pas à optimiser les deux paramètres (nombre et positions de jauges) dans ce mémoire, mais à illustrer leurs impacts sur la structure de semivariogramme. Comme on a mentionné précédemment, le semivariogramme est la représentation statistique de la variabilité spatiale des déformations. Plus le semivariogramme de la surface krigée (respectivement celui de la moyenne des simulations conditionnelles) est confondu avec celui des ÉFs, plus l'estimation de la déformation de l'éprouvette est considérée bonne.

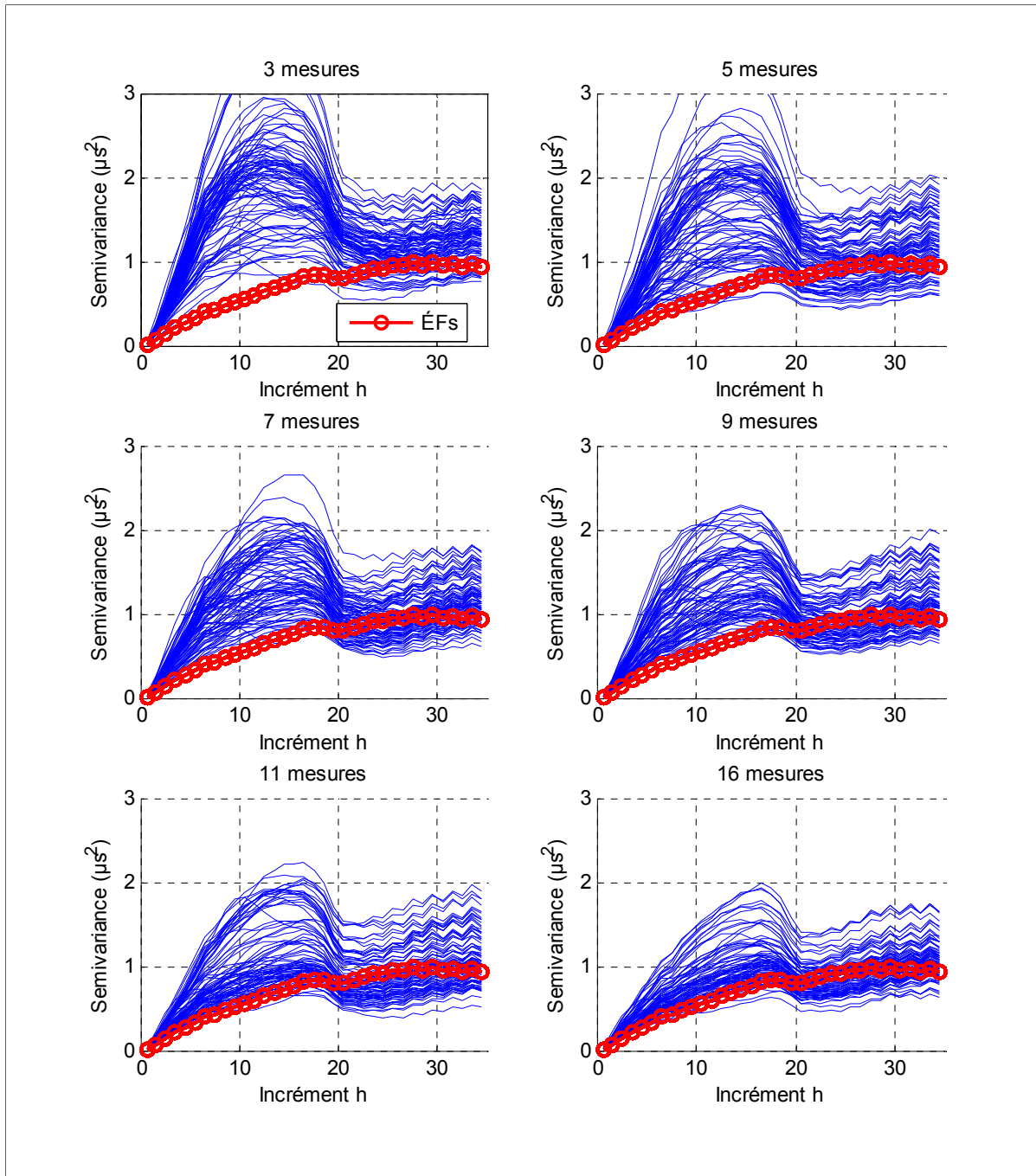


Figure 5-31 Les semivariogrammes :
Variation du nombre de jauges et leurs positions

5.2 Deuxième cas d'étude : aube d'une turbine hydraulique

Cette partie traitera le deuxième cas d'étude : une aube d'une turbine hydraulique. D'abord on part d'un maillage de la surface 3D de l'aube et on essaiera ensuite de le transformer en 2D en utilisant la paramétrisation barycentrique. Ensuite, on appliquera la même procédure du chapitre précédent avec le krigeage et les simulations stochastiques conditionnelles.

5.2.1 Paramétrisation surfacique

Dans un espace euclidien, une surface $S \subset \mathcal{R}^3$ peut être décrite mathématiquement par une représentation implicite, explicite ou paramétrique. On expliquera la paramétrisation bidimensionnelle en utilisant des méthodes linéaires. On distingue deux catégories de paramétrisation : les approches discrètes qui facilitent la manipulation locale de la topologie de la surface, et les approches continues qui soutiennent les calculs numériques mais qui restent sensibles à la continuité mathématique des surfaces à paramétrer.

Les approches continues permettent de paramétrer les surfaces par des fonctions mathématiques polynomiales. Généralement, lors de la paramétrisation, on est amené à utiliser des méthodes qui permettent de présenter mathématiquement des géométries arbitraires. On cite comme exemple de ces méthodes : les surfaces de Bézier (Bezier, 1987), les *B-splines* (Farin, 1990) et les *NURBS* (Piegl, 1997). Ces méthodes échantillonnent les données et remplacent la surface par un certain nombre de points de contrôle.

Une première méthode de paramétrisation continue consiste à interpoler entre les points de contrôle, en revanche ceci introduit des discontinuités. En effet, après la paramétrisation, les discontinuités provoquent des variations brusques et des effets de facettes aux niveaux des plans tangents. Une exigence lors de la paramétrisation de la surface consiste donc à respecter la continuité du premier ordre (tangentes) et de deuxième ordre (points d'inflexion). Toutes ces méthodes utilisent des polynômes de degrés inférieurs d'une unité au nombre de points de contrôle.

Ces surfaces ne passent pas par les points de contrôle. Elles se caractérisent par un contrôle global (donc un manque de contrôle local) puisque la modification d'un point redéfinit toute la surface. Si on modifie le pas de maillage localement, il faudra recalculer de nouveau l'équation caractéristique de la surface de Bézier (respectivement les *B-Splines* et les *NURBS*). Ce contrôle global rend difficile une mise au point interactive. Le degré des polynômes dépend de la forme à paramétrer, plus le nombre de point de contrôle est élevé plus la manipulation devient compliquée.

Malheureusement, ces méthodes sont connues pour engendrer des déformations. Néanmoins, les approches discrètes, où on représente la surface tridimensionnelle par l'assemblage d'éléments tel que les sommets ainsi que les arrêtes des triangles du maillage, tiennent compte de ce problème de distorsion.

Plusieurs approches ont été utilisées pour résoudre le problème de distorsion. Des transformations géométriques simples (Bier et Sloan, 1986) et (Peachey, 1985), des méthodes de dépliages (Samek, Slean et Weghorst, 1986) et des méthodes d'optimisation globale (Krishnamurthy et Levoy, 1996) mais ces dernières méthodes nécessitent beaucoup d'interventions avec un grand nombre de courbes iso-paramétriques. On a aussi essayé de minimiser la norme du tenseur de déformation de Green Lagrange mais cette méthode perd le contrôle sur la nature et la répartition des déformations. Une approche récente fait appel à des applications barycentriques (Hormann et al., 2007). Le problème des déformations est ainsi résolu en remplaçant les isobarycentres par des combinaisons convexes en utilisant des coefficients qui minimisent les distorsions.

Dans notre cas spécifique et pour mieux saisir l'impact de la distorsion, on utilisera, une paramétrisation linéaire. En effet, les dérivées partielles de la paramétrisation sont des vecteurs dans l'espace qui définissent les plans tangents à la surface au nœud du maillage considéré. En considérant juste les deux premiers termes d'une série de Taylor, on pourra approximer raisonnablement la paramétrisation par une fonction linéaire (Hormann et al., 2007).

Le but est donc de trouver la fonction de mise à plat qui minimise la distorsion. Puisque la fonction est bijective, la carte bidimensionnelle sera unique. Deux choix peuvent se présenter avant de réaliser la paramétrisation : fixer ou libérer la bordure de la carte bidimensionnelle. Puisque le maillage est plus raffiné dans les zones de concentration de contraintes de l'aube, on fixera la bordure pour mieux contrôler les distorsions dans ces zones.

La géométrie différentielle permet de préserver les angles si la transformation est conforme, les aires si la transformation est authalique et les longueurs si cette dernière est isométrique. La question qui se pose : est-ce qu'on pourra conserver plus qu'une propriété lors de la transformation ? La réponse est malheureusement négative. Carl Friedrich Gauss a prouvé que la paramétrisation des surfaces générales engendre une distorsion des angles ou des aires du maillage et dans la majorité des cas les deux en même temps. Il a prouvé aussi que les seules surfaces qui peuvent être paramétrées sans distorsions sont celles qui ont une courbure gaussienne nulle, ou ce qu'on appelle parfois des surfaces développables (Hormann et al. 2007).

5.2.1.1 Paramétrisation d'un maillage triangulaire

Une paramétrisation est une fonction bijective entre une surface $S \subset \mathcal{R}^3$ et un domaine de paramétrisation planaire $\Omega \subset \mathcal{R}^2$. Pour que ce dernier existe, il faut que S et Ω aient la même topologie.

Notation (pour la suite du développement mathématique) :

- On note, dans \mathcal{R}^3 , un sommet par $\mathbf{p} = (x, y, z)$ et dans \mathcal{R}^2 $\mathbf{u} = (u, v)$.
- Un maillage triangulaire d'une surface tridimensionnelle $S_{\mathcal{T}} \subset \mathcal{R}^3$ est composé d'un ensemble de triangles surfaciques $\mathcal{T} = \{\mathbf{T}_1, \dots, \mathbf{T}_m\}$ qui s'interceptent à des arêtes communs $\mathcal{E} = \{\mathbf{E}_1, \dots, \mathbf{E}_l\}$ et des sommets $\mathcal{V} = \left\{ \underbrace{\mathbf{p}_1, \dots, \mathbf{p}_n}_{\text{sommet à interne}} \right\} + \left\{ \underbrace{\mathbf{p}_{n+1}, \dots, \mathbf{p}_{n+b}}_{\text{Sommets sur la frontière}} \right\}$

- Deux sommets distincts $\mathbf{p}_i, \mathbf{p}_j \in \mathcal{V}$ sont dits des voisins si les deux extrêmes de l'arrête $E = [\mathbf{p}_i, \mathbf{p}_j] \in \mathcal{E}$ et on note pour chaque $\mathbf{p}_i \in \mathcal{V}$ l'ensemble $N_i = \{j : [\mathbf{p}_i, \mathbf{p}_j] \in \mathcal{E}\}$, l'ensemble des indices de tous les voisins de \mathbf{p}_i .

L'idée principale de la paramétrisation d'un maillage triangulaire se base sur le modèle physique suivant. On remplace l'arrête d'un triangle (la liaison entre deux sommets) par un ressort. Si on fixe la frontière de ce réseau de ressorts, l'intérieur de ce dernier se relaxera dans une configuration énergétiquement convenable et par conséquence elle sera la paramétrisation voulue.

La première étape lors de la construction d'un maillage barycentrique est la paramétrisation des sommets de la frontière. La façon la plus simple est de projeter ces sommets sur un plan dans le sens des moindres carrés. Pour notre cas d'étude, la zone A, cette projection est une bonne approche car la topologie est quasi planaire dans cette région. Dans le cas où la topologie de la frontière est plus complexe, on doit tenir compte du choix de la forme du domaine de paramétrisation (Shi et Van Albada, 2007) et de la distribution des sommets paramétrés proche de la frontière. La procédure la plus utilisée en littérature lors du choix de la distribution est la longueur de corde (*chordlength*) (Boissonnat et al., 2011).

L'énergie d'un ressort qui lie les deux sommets \mathbf{p}_i et \mathbf{p}_j est donnée par l'équation suivante :

$$E = \frac{1}{2} \mathbf{D}_{ij} \mathbf{s}_{ij}^2 \quad (5.2)$$

Avec $\mathbf{D}_{ij} > 0$ est la constante du ressort et $\mathbf{s}_{ij} = \|\mathbf{u}_i - \mathbf{u}_j\|$ est la longueur du ressort. $\mathbf{D}_{ij} = \mathbf{D}_{ji}$ sont les constantes d'élasticité des ressorts liant les sommets du maillage. La vraie question qui se pose : comment calculer la constante d'élasticité?

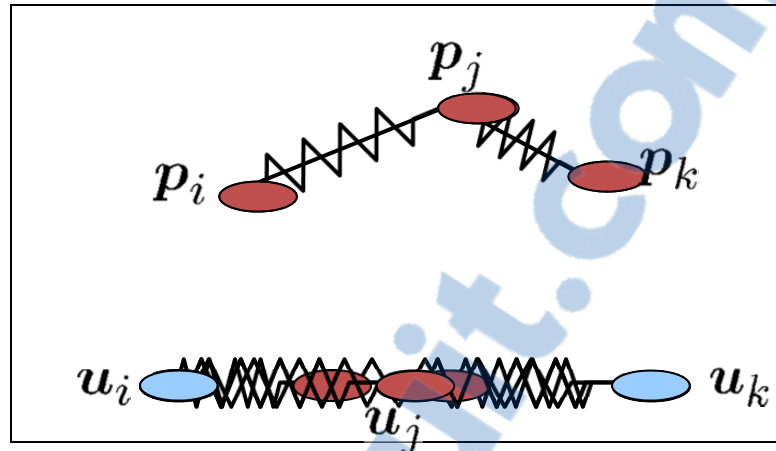


Figure 5-32 Paramétrisation : Le modèle du ressort

Plusieurs approches ont été mentionnées en littérature

- La constante de ressort est proportionnelle à la longueur de l'arrête du triangle du maillage (Floater et Reimers, 2001). L'inconvénient de cette méthode est de ne pas respecter le requis de minimisation discuté précédemment.
- Coordonnées *Wachspress* (Dikshit et Ojha, 2003) :

$$D_{ij} = \left(\frac{\cotan \alpha_{ji} + \cotan \beta_{ij}}{r_{ij}^2} \right) \quad (5.3)$$

- Coordonnées harmoniques discrètes (Pinkall et Polthier, 1993)

$$D_{ij} = (\cotan \gamma_{ij} + \cotan \gamma_{ji}) \quad (5.4)$$

- Coordonnées valeur-moyenne (*MeanValueCoordinate*)

$$D_{ij} = \left(\frac{\tan \frac{\alpha_{ji}}{2} + \tan \frac{\beta_{ij}}{2}}{r_{ij}} \right) \quad (5.5)$$

- La méthode de type éléments finis

$$\mathbf{D}_{ij} = \frac{L_{i,k_1}^2 + L_{j,k_1}^2 + L_{i,j}^2}{\text{Aire}(i,j,k_1)} + \frac{L_{i,k_2}^2 + L_{j,k_2}^2 + L_{i,j}^2}{\text{Aire}(i,j,k_2)} \quad (5.6)$$

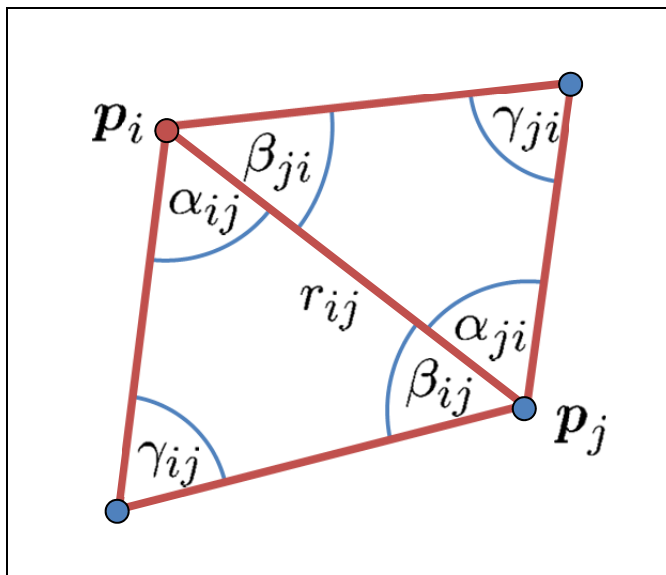


Figure 5-33 Paramétrisation : angles du maillage triangulaire

En généralisant ce processus sur tous les sommets du maillage triangulaire, on obtient l'énergie totale :

$$E = \frac{1}{2} \sum_{i=1}^{n+b} \sum_{j \in N_i} \mathbf{D}_{ij} \|\mathbf{u}_i - \mathbf{u}_j\|^2 \quad (5.7)$$

Afin de trouver l'énergie convenable de la relaxation des ressorts, on minimise l'équation précédente par une première dérivée partielle :

$$\frac{\partial E}{\partial \mathbf{u}_i} = \sum_{j \in \mathbf{N}_i} \mathbf{D}_{ij}(\mathbf{u}_i - \mathbf{u}_j) \quad (5.8)$$

on obtient le minimum de E lorsque :

$$\sum_{j \in \mathbf{N}_i} \mathbf{D}_{ij}(\mathbf{u}_i) = \sum_{j \in \mathbf{N}_i} \mathbf{D}_{ij}(\mathbf{u}_j) \quad (5.9)$$

L'équation précédente est équivalente à dire que chaque sommet paramétrique intérieur est une combinaison affine de ses voisins, ce qui se traduit par :

$$\mathbf{u}_i = \sum_{j \in \mathbf{N}_i} \lambda_{ij} \mathbf{u}_j \quad (5.10)$$

avec λ_{ij} qui sont des coefficients normalisés et dont la somme est égale à l'unité.

$$\lambda_{ij} = \frac{\mathbf{D}_{ij}}{\sum_{j \in \mathbf{N}_i} \mathbf{D}_{ij}} \text{ avec } \sum_{j \in \mathbf{N}_i} \lambda_{i,j} = 1 \quad (5.11)$$

Par une simple séparation entre les sommets paramétriques de l'intérieur et ceux de la bordure (Hormann et al., 2007), on obtient :

$$\underbrace{\mathbf{u}_i - \sum_{j \in \mathbf{N}_i, j \leq n} \lambda_{ij} \mathbf{u}_j}_{\substack{\text{Sommets paramétriques} \\ \text{inconnus}}} = \underbrace{\sum_{j \in \mathbf{N}_i, j > n} \lambda_{ij} \mathbf{u}_j}_{\substack{\text{Sommets paramétriques} \\ \text{fixés (la bordure)}}} = \bar{\mathbf{u}}_i \quad (5.12)$$

Ce qui nous amène à résoudre le système linéaire suivant :

$$\begin{cases} \mathbb{A} U = \bar{U} \\ \mathbb{A} V = \bar{V} \end{cases} \quad (5.13)$$

Avec $U = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ et $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ qui sont les vecteurs colonnes des coordonnées inconnues et $\bar{U} = (\bar{u}_1, \dots, \bar{u}_n)$ et $\bar{V} = (\bar{v}_1, \dots, \bar{v}_n)$ qui sont les vecteurs colonnes des sommets de la bordure avec les coefficients :

$$\bar{u}_i = \sum_{j \in \wp_i, j > n} \lambda_{ij} \mathbf{u}_j \quad \text{et} \quad \bar{v}_i = \sum_{j \in \wp_i, j > n} \lambda_{ij} \mathbf{v}_j \quad (5.14)$$

et $\mathbb{A} = (a_{ij})_{i,j=1,\dots,n}$ est une matrice symétrique avec :

$$a_{ij} = \begin{cases} 1 & \text{si } i = j \\ -\lambda_{ij} & \text{si } j \in \wp_i \\ 0 & \text{si non} \end{cases} \quad (5.15)$$

Sous forme matricielle, le système linéaire précédent s'écrit comme :

$$\begin{pmatrix} 1 & \cdots & \cdots & -\lambda_{ij} \\ \vdots & \ddots & \cdots & \vdots \\ \vdots & \cdots & 1 & \vdots \\ -\lambda_{ji} & \cdots & \cdots & 1 \end{pmatrix}_{(n \times n)} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}_{(n \times 1)} = \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \vdots \\ \bar{u}_n \end{pmatrix}_{(n \times 1)}$$

$$\begin{pmatrix} 1 & \cdots & \cdots & -\lambda_{ij} \\ \vdots & \ddots & \cdots & \vdots \\ \vdots & \cdots & 1 & \vdots \\ -\lambda_{ji} & \cdots & \cdots & 1 \end{pmatrix}_{(n \times n)} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_{(n \times 1)} = \begin{pmatrix} \bar{v}_1 \\ \bar{v}_2 \\ \vdots \\ \bar{v}_n \end{pmatrix}_{(n \times 1)} \quad (5.16)$$

5.2.1.2 Paramétrisation de l'aube d'une turbine

Le maillage de l'aube de la turbine a été réalisé en utilisant le logiciel de simulation Siemens NX8. Après élimination de nœuds moyens (*Midnodes mesh*) de maillage surfacique, on garde

un maillage surfacique triangulaire composé par les nœuds principaux. Une des zones de concentration des contraintes est la zone A définie sur la Figure 5-34. Cette zone est proche de l'extrémité de l'aube, la jonction soudée entre l'aube et la ceinture de la turbine. En effet, on s'intéresse à cette partie de l'aube puisqu'il est possible d'y placer une rosette pour récolter les données de la charge dynamique appliquée.

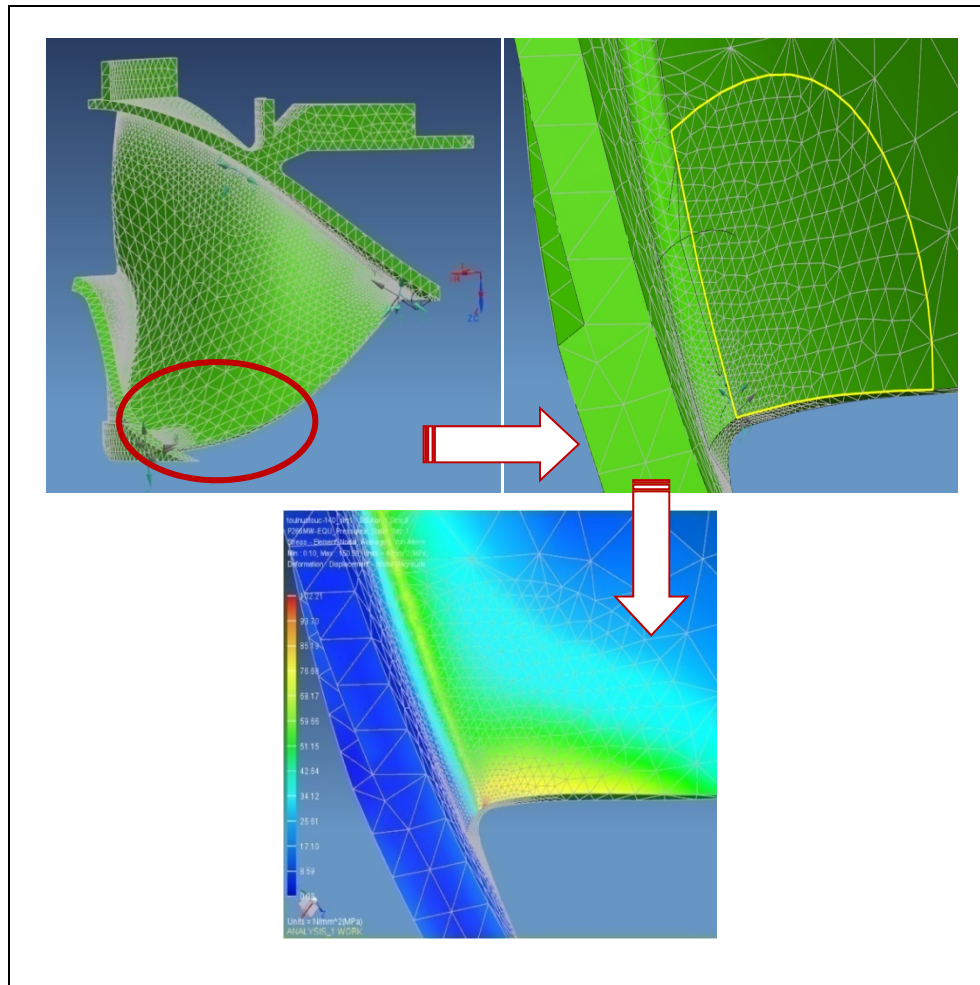


Figure 5-34 Zone A de la turbine hydraulique

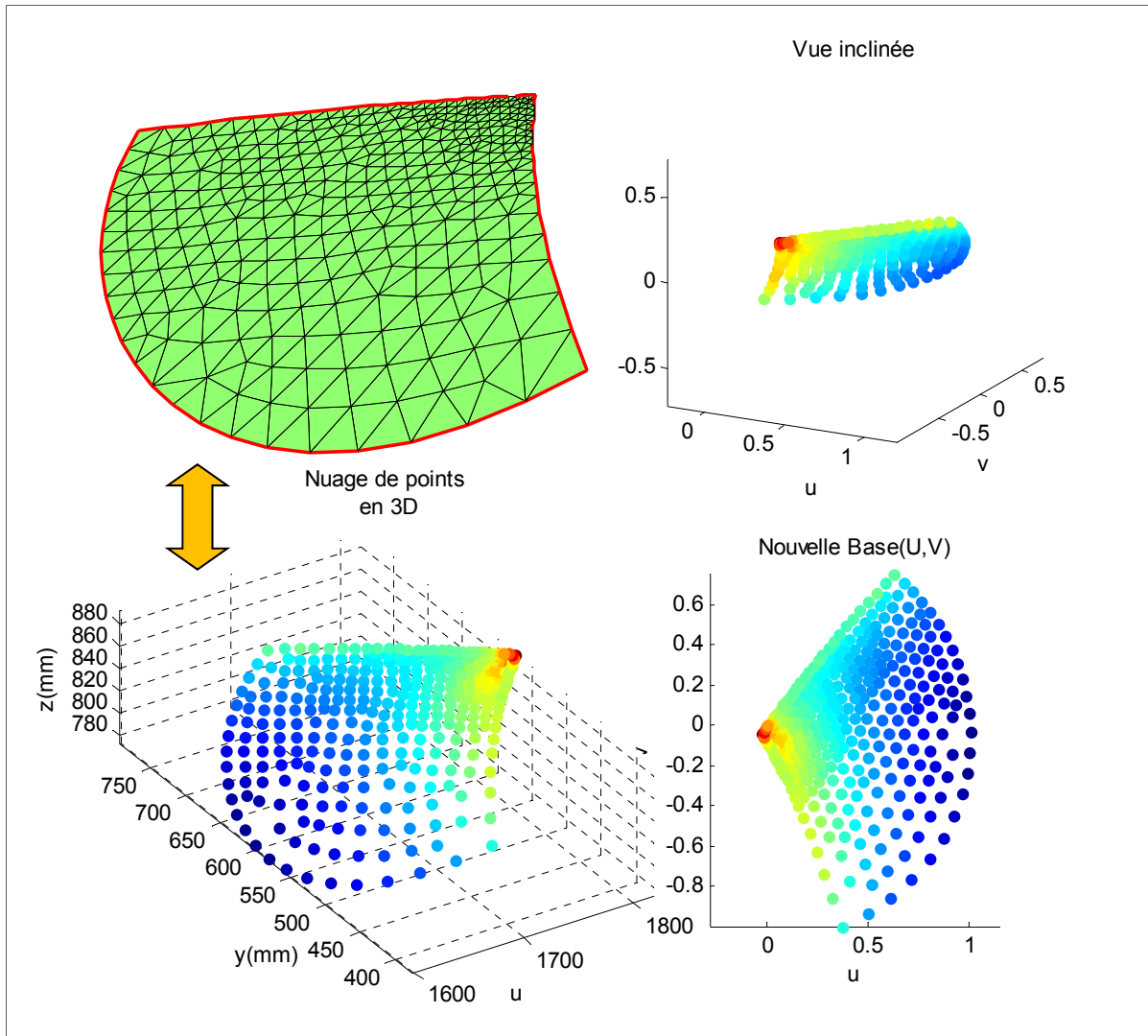


Figure 5-35 Paramétrisation de la zone A de l'aube

La paramétrisation de la zone A est présentée par la Figure 5-35. Complètement en haut à gauche, on a le maillage original (surface en 3D) extrait à partir des résultats des simulations par éléments finis NX8. Juste en bas, ce maillage est configuré avec la fonction MatLab® (code principal) afin d'avoir la liste finale des sommets des triangles ainsi que la liste de connectivités. Complètement en bas, on a la paramétrisation de la zone A et pour confirmer la bi-dimensionnalité, on a ajouté une vue inclinée qui valide ce qui précède.

5.2.2 Analyse variographique

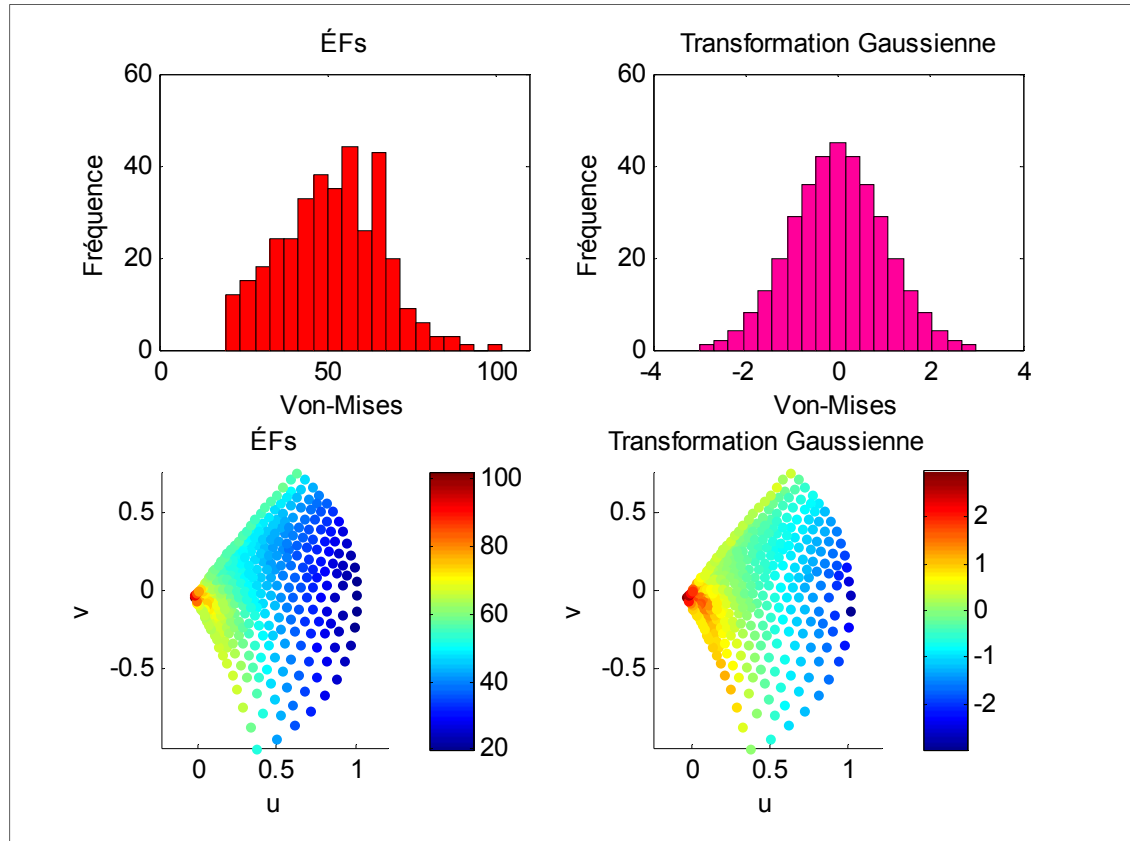


Figure 5-36 Transformation gaussienne de contraintes von-Mises de la zone A

Tel que mentionné précédemment, une transformation gaussienne est primordiale avant de commencer l'analyse variographique. L'historgramme de valeurs primaires de von-Mises montre une distribution asymétrique avec un coefficient de coefficient d'aplatissement négatif de -0.28 loin de la valeur nulle pour une distribution gaussienne.

La transformation gaussienne a permis d'amortir le gradient de contraintes et de l'étaler sur un espace plus large afin de respecter la stationnarité. La distribution obtenue a un coefficient d'aplatissement égale à -0.04, une moyenne nulle avec un intervalle de confiance $[-0.104 ; +0.104]$ et une variance proche de l'unité égale à 0.999 $[0.931 ; 1,079]$ @95%.

La Figure 5-37 montre les 4 variogrammes expérimentaux dans les directions Nord-Sud, Est-West et Nord-Est (oblique, à 45°) ainsi que le variogramme omnidirectionnel. On remarque que dans la direction 90° , le variogramme suit un comportement d'un modèle de puissance. Ceci s'explique par la continuité du gradient de déformation le long de cette direction sur toute la zone d'étude comme le montre la Figure 5-37. Dès qu'on se déplace de l'horizontale à la verticale (Nord-Sud), on minimise la variance du variogramme et de plus en plus le palier apparaît. En effet, on agrandissant des débuts de ces variogrammes (zone de décollage), le variogramme construit dans la direction Nord-Sud montre un palier et une constance de la variance contrairement aux autres directions. Ainsi, on décide de garder ce variogramme et cette direction d'étude surtout que la Figure 5-37 montre un étalement de gradient dans cette direction. Ce variogramme sera ajusté par la méthode des moindres carrés et approuvé par la suite par une validation croisée.

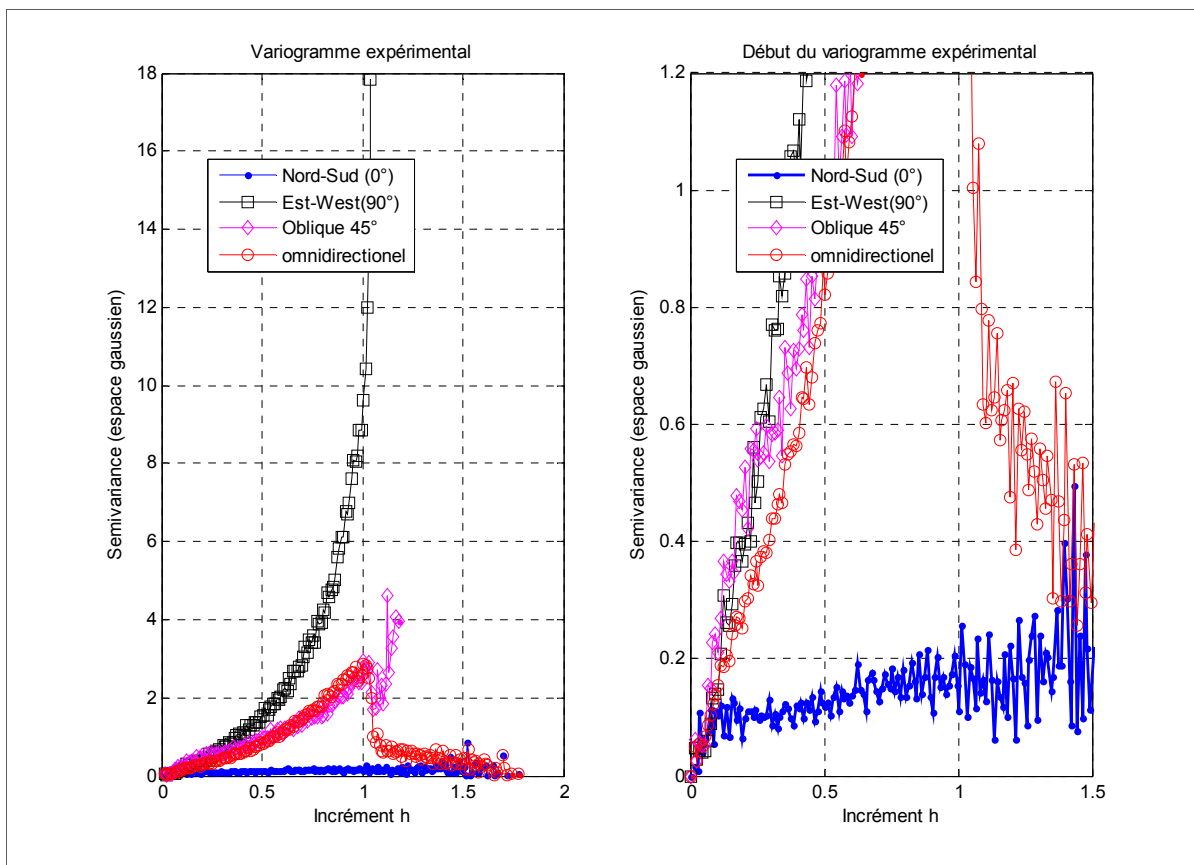


Figure 5-37 Construction du semivariogramme expérimental suivant plusieurs directions

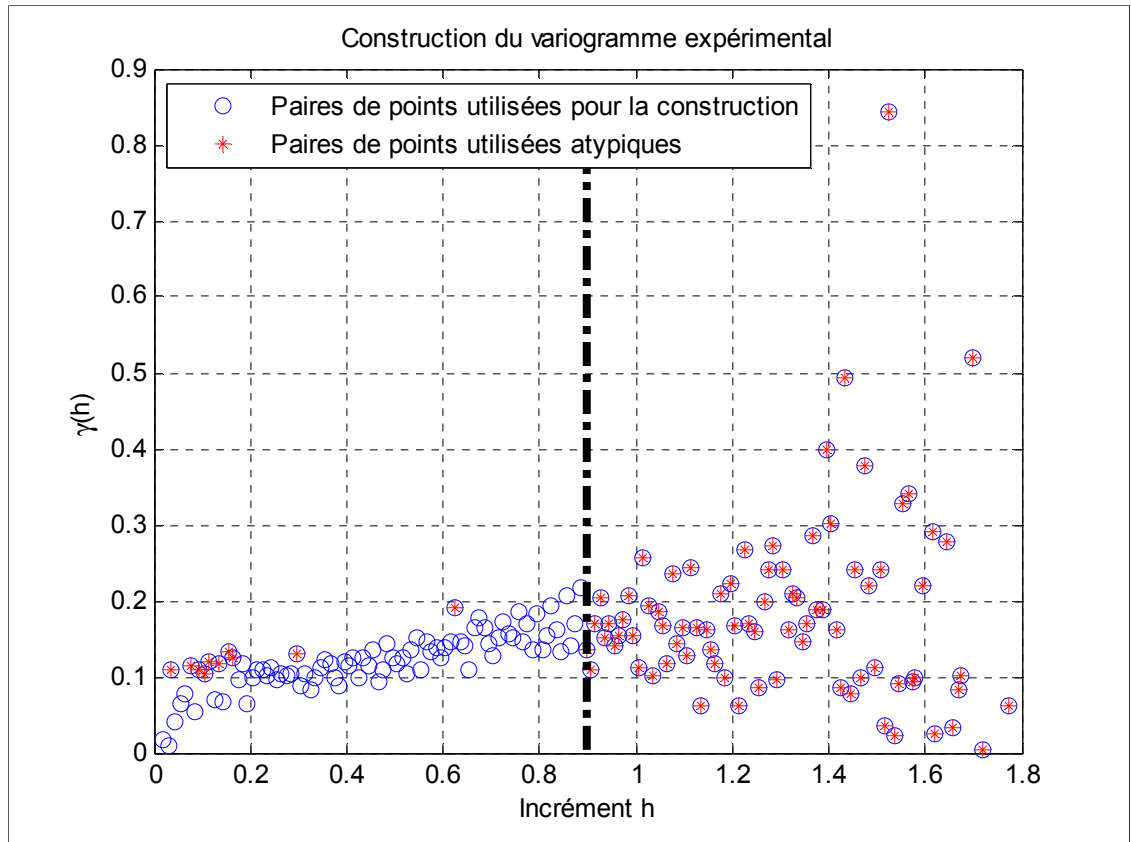


Figure 5-38 Construction du semivariogramme expérimental dans la direction Nord-Sud

Le semivariogramme est construit dans la direction Nord-Sud avec un angle de tolérance de 10° . En effet, plus cet angle augmente plus le variogramme a tendance d'avoir l'allure du semivariogramme omnidirectionnel et de suivre le comportement d'un modèle puissance. Vu qu'on a un nombre important de valeurs pour la construction (355 mesures sur la zone A), on juge que c'est suffisant de garder juste une ouverture d'angle de collecte de 10° dans la direction de 0° (Nord-Sud) comme le montre la Figure 5-37.

Les cercles bleus représentent les couples des paires des données conservées pour la construction et les restes (avec des étoiles rouges) sont considérés comme des valeurs atypiques. En effet, la présence d'une concentration de ces paires lors du décollage sature le variogramme avant sa stabilisation. Il faudra éliminer toutes les paires de points dans la

deuxième moitié car elles sont statistiquement non significatives et parce que le variogramme montre son palier avant même d'attendre cette zone.

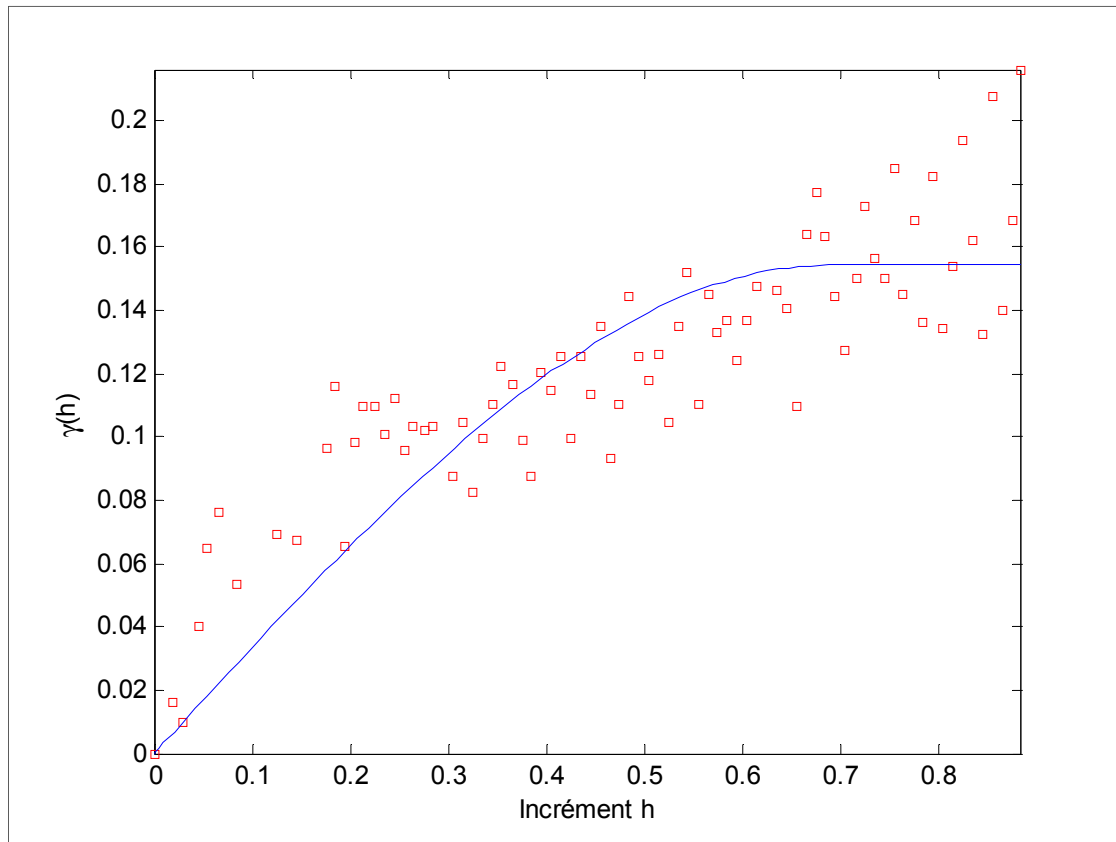


Figure 5-39 Ajustement du semivariogramme par la méthode des moindres carrés

L'ajustement du semivariogramme montre que le modèle sphérique sans effet de pépite est le plus convenable dans ce cas avec un palier de l'ordre de 0.15 et une portée de 0.64 mm, Figure 5-39 et Figure 5-41.

La validation croisée (illustrée à la Figure 5-40) montre à son tour que le modèle sphérique est plus performant que les deux autres malgré que les trois aient été ajustés par la méthode des moindres carrés et que pour chaque itération on garde les paramètres optimisés de chaque modèle tout en changeant la position de 12 points de mesures.

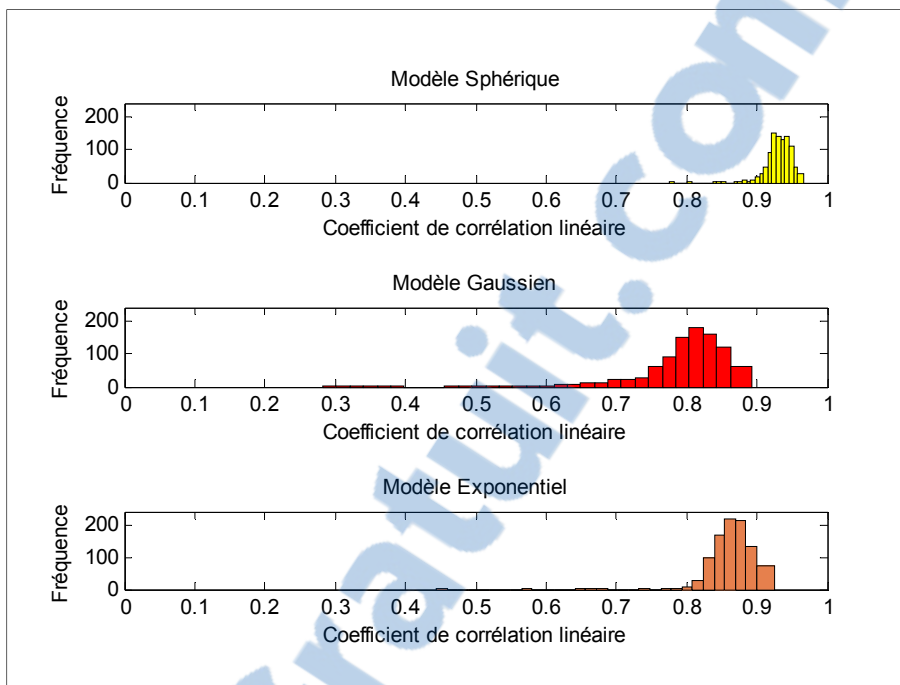


Figure 5-40 Validation croisée

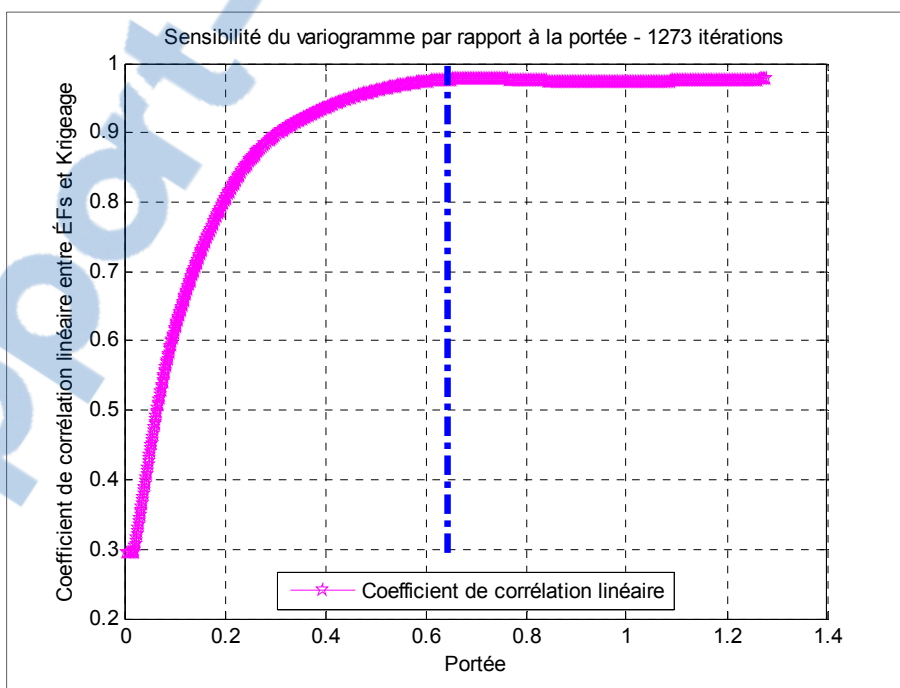


Figure 5-41 Sensibilité du variogramme à la portée

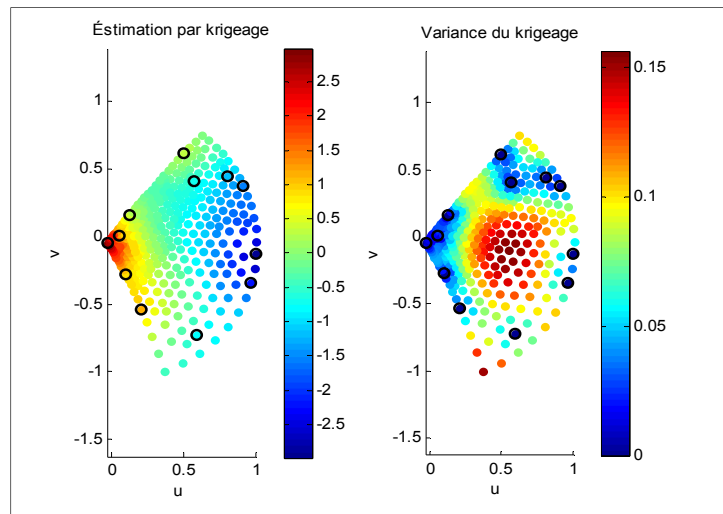


Figure 5-42 Krigage avec un variogramme sphérique
Palier =0.15 et Portée= 0.64

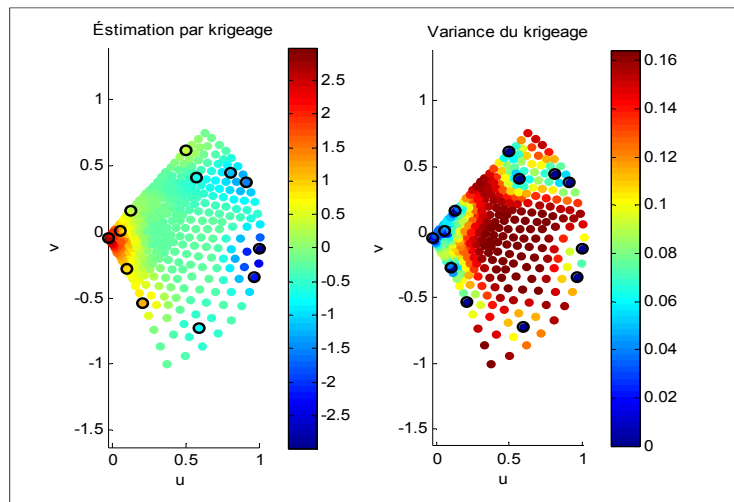


Figure 5-43 Krigage avec un variogramme sphérique
Palier =0.15 et Portée= 0.34

La Figure 5-41 montre que le variogramme est sensible à la portée. En effet, plus la portée est inférieure à la valeur ajustée par les moindres carrés (0.6428), plus le coefficient de corrélation linéaire diminue. La variance du krigeage sera plus importante comme le montre la Figure 5-42 et la Figure 5-43 .

5.2.3 Krigeage

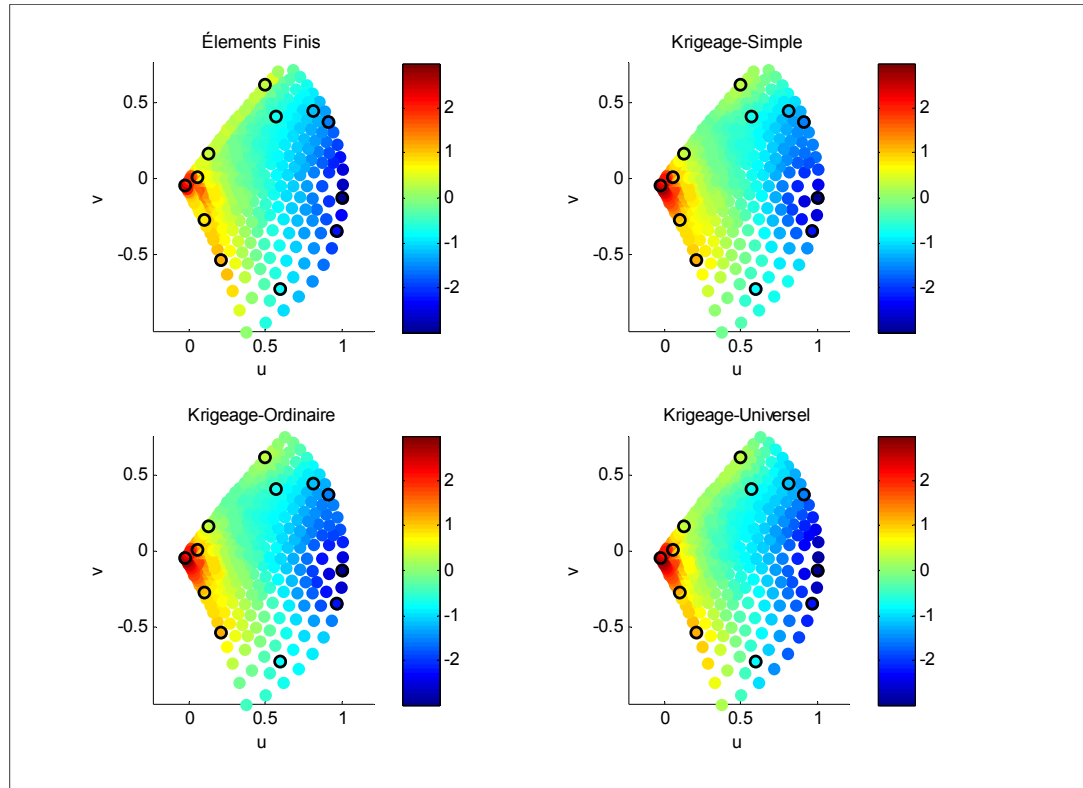


Figure 5-44 Comparaison entre la cartographie du KS, KO et KU avec la référence ÉFs

La figure ci-dessus montre une ressemblance entre les trois cartographies de krigeage et la carte référentielle ÉFs. Ceci se confirme par les histogrammes de la Figure 5-45. Les cartes de la variance du krigeage montrent que le (KO) et (KS) ont pratiquement la même variance. Néanmoins le krigeage universel a une variance plus importante qui s'accroît lorsqu'on s'éloigne de points de mesures. Tel que mentionné précédemment, on ne retient pas le krigeage universel (KU) car c'est un estimateur qui minimise moins la variance comparativement aux deux autres, Figure 5-46. Comme le (KO) considère la moyenne localement (au voisinage du point à estimer) contrairement au KS qui la suppose connue sur tout le domaine d'étude et qu'il est difficile de la connaître pratiquement, on gardera le krigeage ordinaire pour le reste de l'analyse.

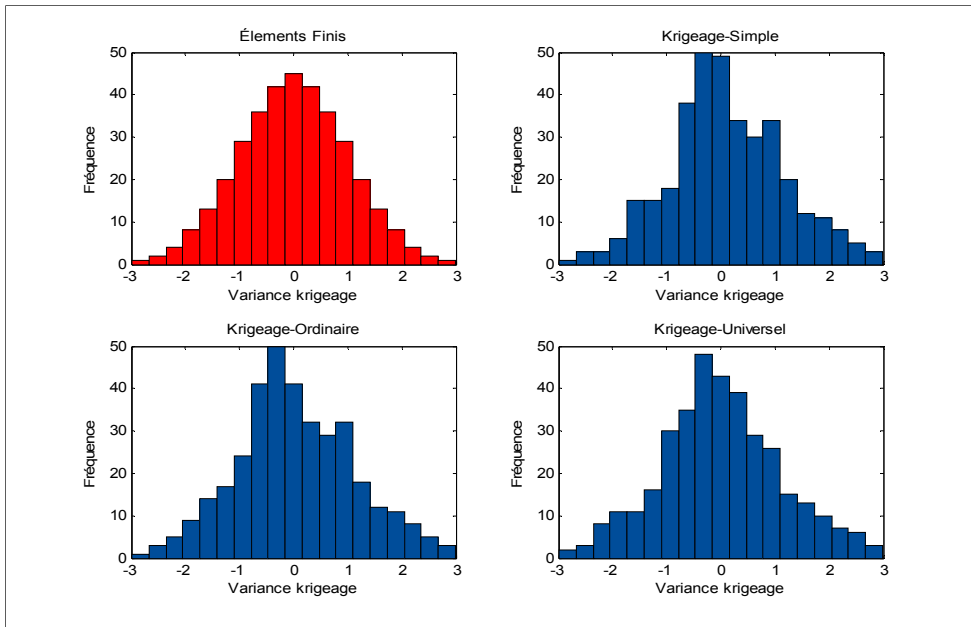


Figure 5-45 Comparaison entre les histogrammes du KS, KO et KU

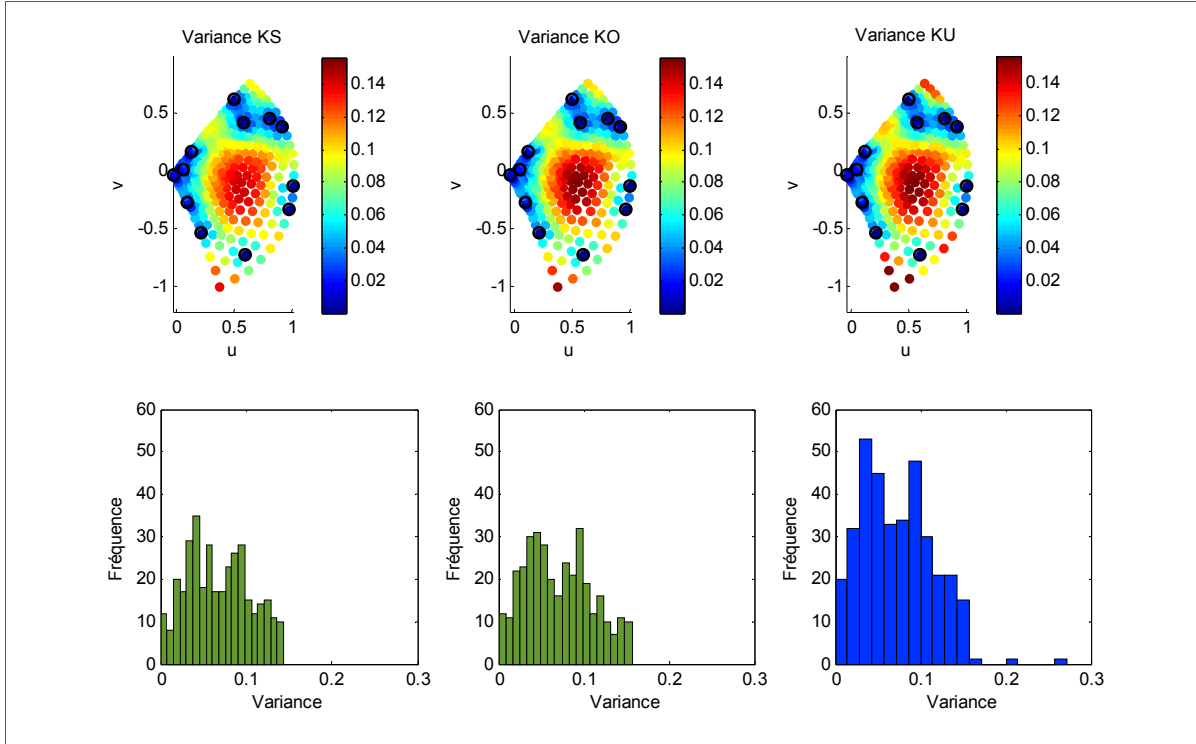


Figure 5-46 Comparaison entre les variances des KS, KO et KU

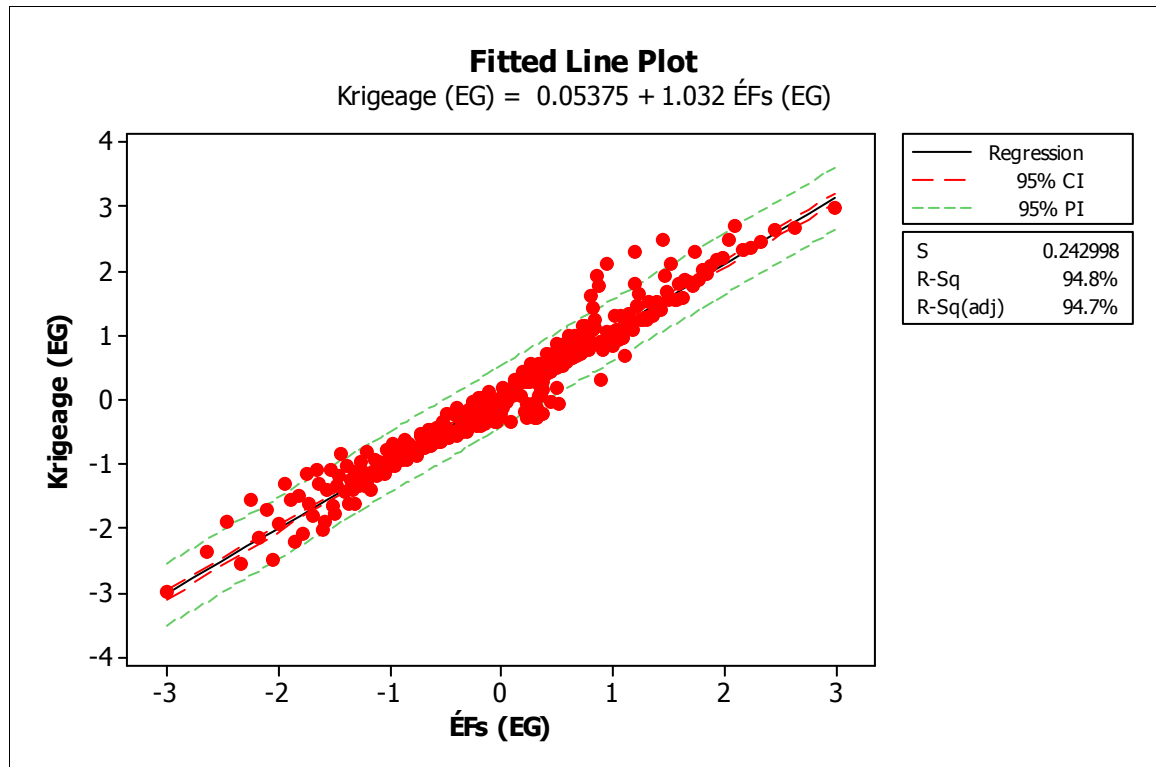


Figure 5-47 Droite d'ajustement entre les contraintes simulées par les ÉFs et estimées par le krigeage ordinaire

La régression linéaire simple montre une forte corrélation entre la variable explicative ÉFs et la variable aléatoire à expliquer (krigeage), Figure 5-47. Le nuage de points est pratiquement entièrement englobé dans l'intervalle de prédiction @95%. Cette corrélation montre une pente égale à 1.032 (très proche de l'unité) et une intersection à l'origine égale à 0.05375 (proche de la valeur nulle, donc un biais non significatif). L'analyse de corrélation évalue la qualité de cette régression avec un coefficient de corrélation ajusté égal à 94.7%. Ceci permet de conclure que le modèle du krigeage peut être utilisé pour des applications prédictives.

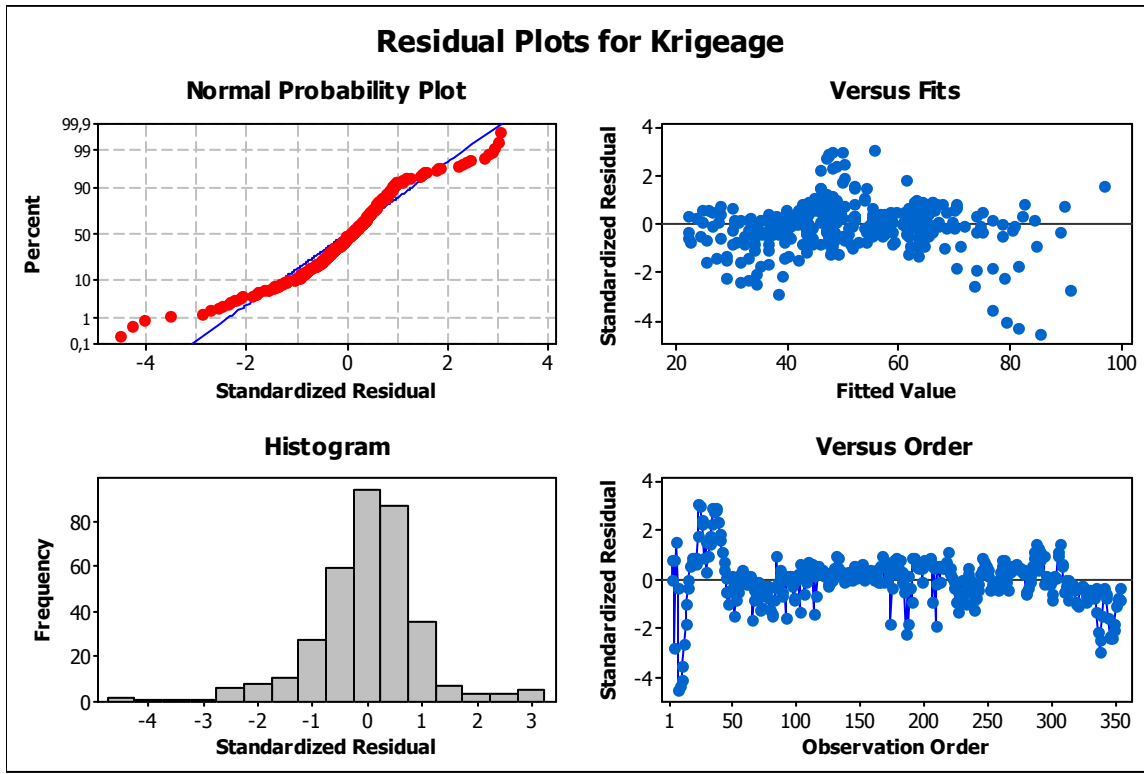


Figure 5-48 Analyse des résidus du krigage

Selon la Figure 5-48, le modèle de régression montre qu'on n'est pas capable de prédire l'erreur pour une observation donnée. En effet, les résidus ne montrent pas une zone de concentration haute ou basse, ils sont centrés sur la valeur nulle sur toute la plage des valeurs ajustées. Les résidus adjacents ne sont pas corrélés entre eux (une absence d'auto corrélation).

Ainsi, le modèle du variogramme sera apte à conditionner les simulations stochastiques puisque les résidus sont d'espérance nulle, la variance est quasi constante et est indépendante des valeurs, les résidus sont deux à deux indépendants et ne dépendent pas de la valeur des données.

5.2.4 Simulations stochastiques

La Figure 5-49 montre les variogrammes de toutes les simulations stochastiques. La moyenne de ces derniers, le variogramme de la surface krigée et le variogramme référentiel présentent la même allure et ont la même portée (indiquée par la ligne jaune). La variance du variogramme de la moyenne des simulations est plus importante comparativement aux deux autres puisque à chaque simulation, on perturbe d'une manière différente la structure de dépendance spatiale de la variable à estimer dans le but de modéliser le hasard de cette variable aléatoire.

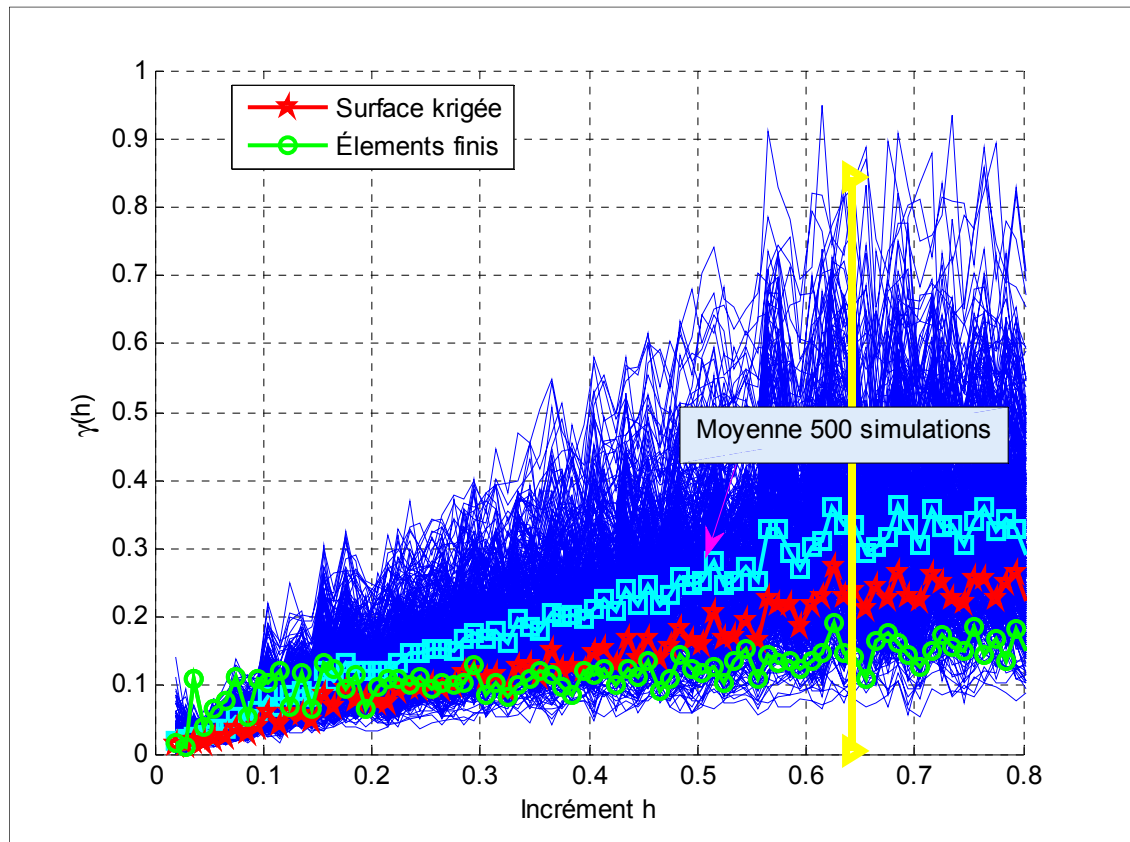


Figure 5-49 Les semivariogrammes de la moyenne de 500 simulations, surface krigée et EFs

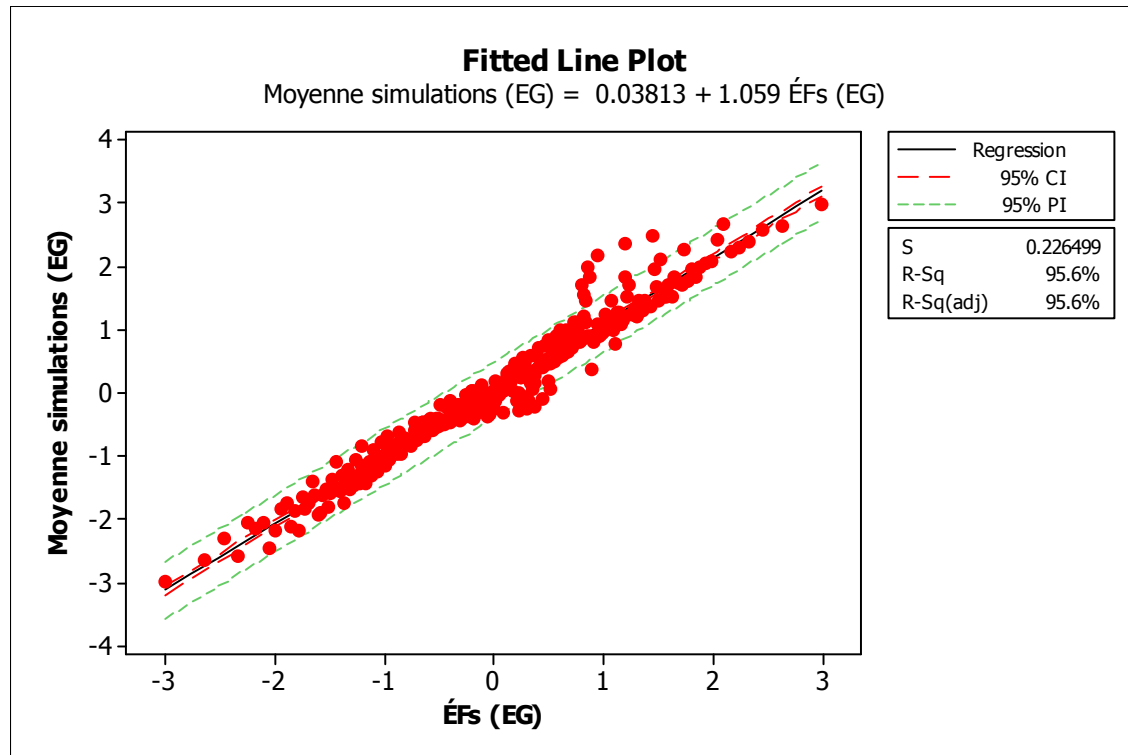


Figure 5-50 Droite d'ajustement entre la moyenne des 500 simulations conditionnelles et les résultats simulés par ÉFs

Le coefficient de corrélation montre que les valeurs de contraintes calculées par des simulations ÉFs entrent pour 95.6% dans la variation de la moyenne de 500 simulations stochastiques conditionnelles. Un examen de la figure ci-dessus montre que les données sont réparties d'une manière aléatoire autour de la ligne de régression, ce qui implique l'absence d'inadéquation systématique de l'ajustement. Les limites de confiance de 95% (IC de 95%) pour la moyenne des 500 simulations ainsi que les limites de prévision de 95% (IP de 95%) englobent pratiquement toutes les données. Le modèle prédictif est assez bon puisque la pente est pratiquement égale à l'unité et la constante à l'origine est nulle. Le résultat est valide en passant d'un espace gaussien à l'espace réel.

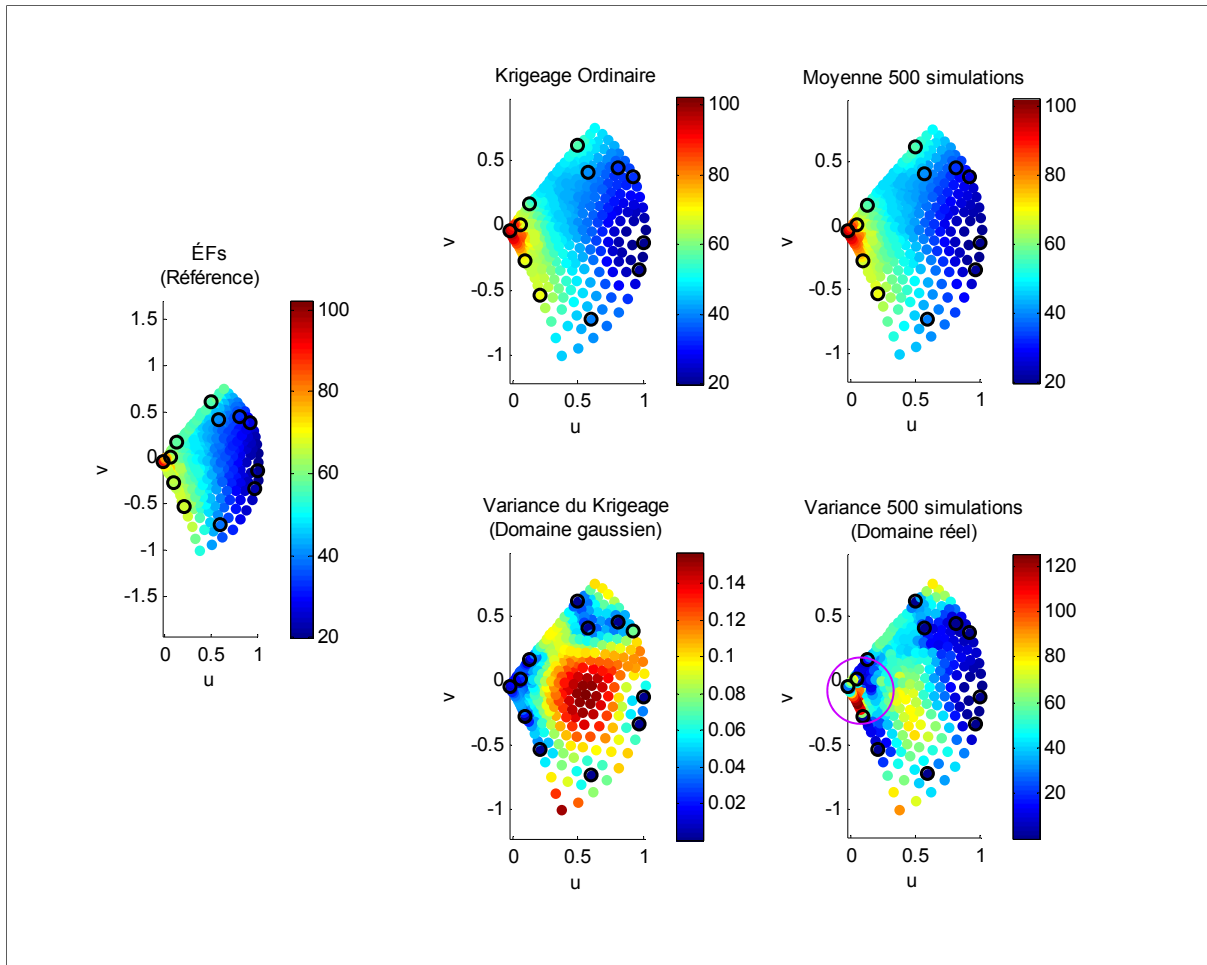


Figure 5-51 Comparaison entre la variance du krigeage et la variance de 500 simulations stochastiques

On résume ce chapitre sur la Figure 5-51. La distribution de la moyenne de 500 simulations conditionnelles est semblable à celle estimée par le krigeage. En revanche, la cartographie de la variance est plus significative que celle du krigeage, et ceci est attribuable à l'effet du lissage intrinsèque à la méthode. En effet, proche de la zone de concentration de contrainte, la moyenne de simulations montre un fort effet de lissage dans l'espace réel (figure en haut complètement à droite). On a donc plus de risque de commettre une erreur lors de l'estimation par krigeage, ce qui sera ajusté par une augmentation de la variance par la suite dans cette zone.

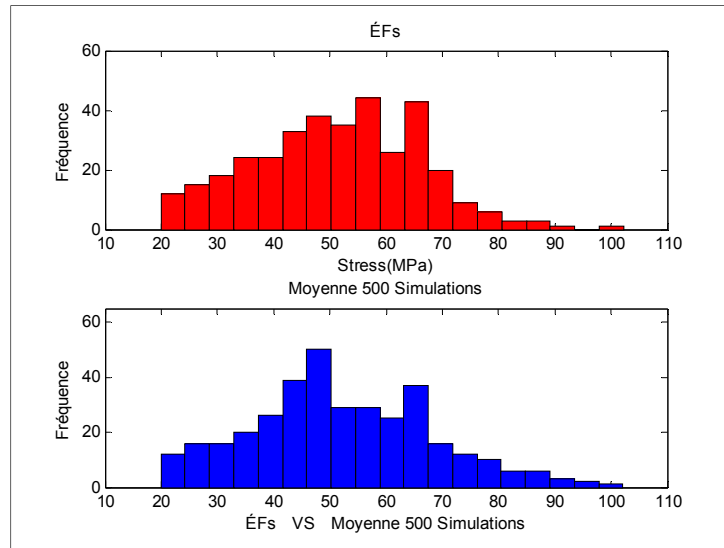


Figure 5-52 Comparaison des histogrammes des ÉFs et la moyenne de 500 simulations stochastiques

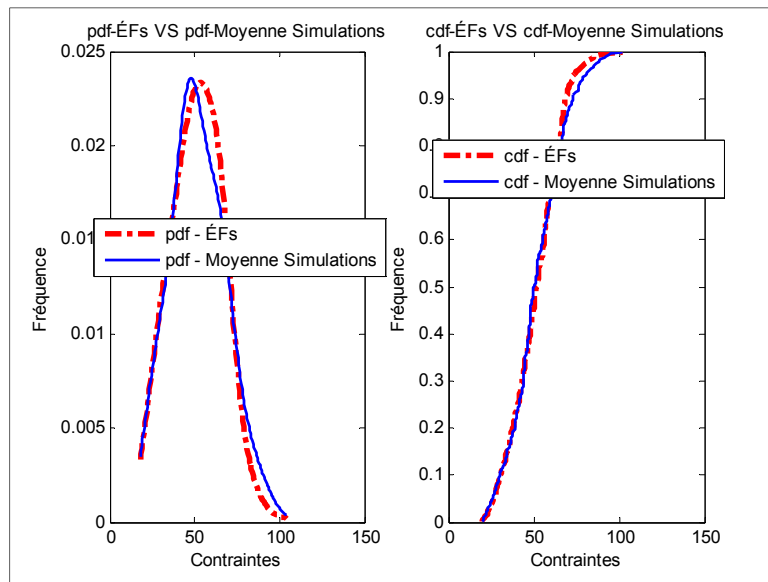


Figure 5-53 Comparaison entre les *pdf* et les *cdf* des ÉFs et la moyenne de 500 simulations stochastiques

La Figure 5-52 montre une analogie de l'histogramme des données primaires ÉFs avec celui de la moyenne des simulations stochastiques. Aussi, les courbes des *pdf* et *cdf* montre une bonne corrélation entre les deux comme le montre la Figure 5-53.

Encore une fois, le coefficient de variation (CV) confirme la forte variance de la mesure dans la région proche de la zone de concentration de contrainte avec une valeur qui dépasse 20% comme le montre la Figure 5-54.

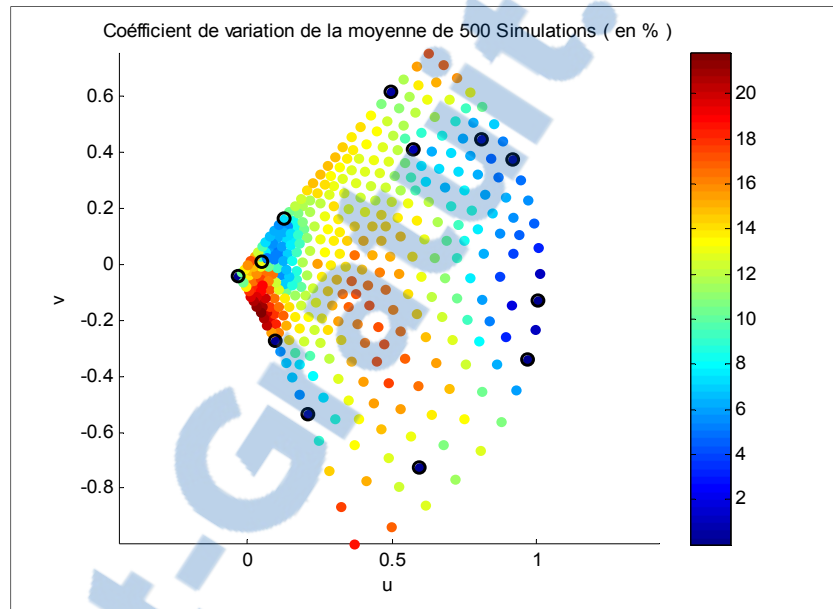


Figure 5-54 CV de la moyenne de 500 simulations

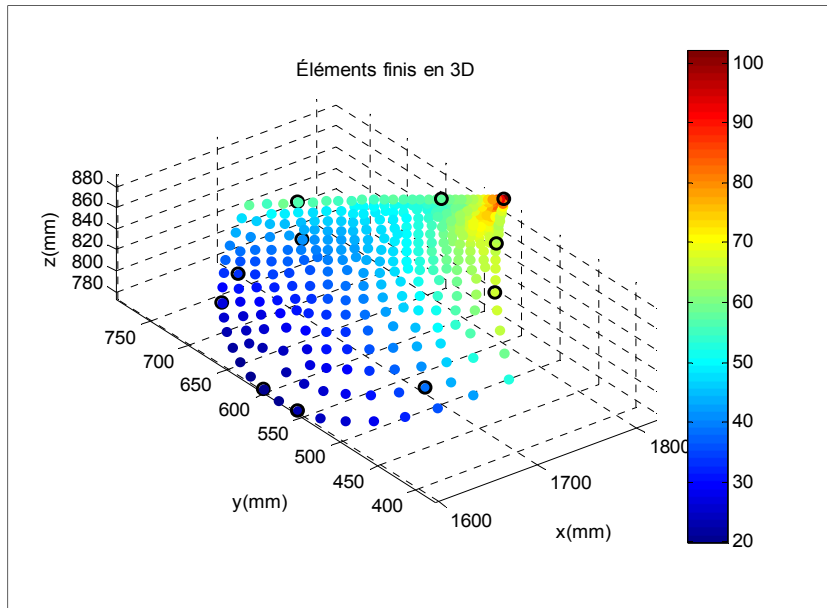


Figure 5-55 Contraintes de von Mises en 3D (ÉFs)

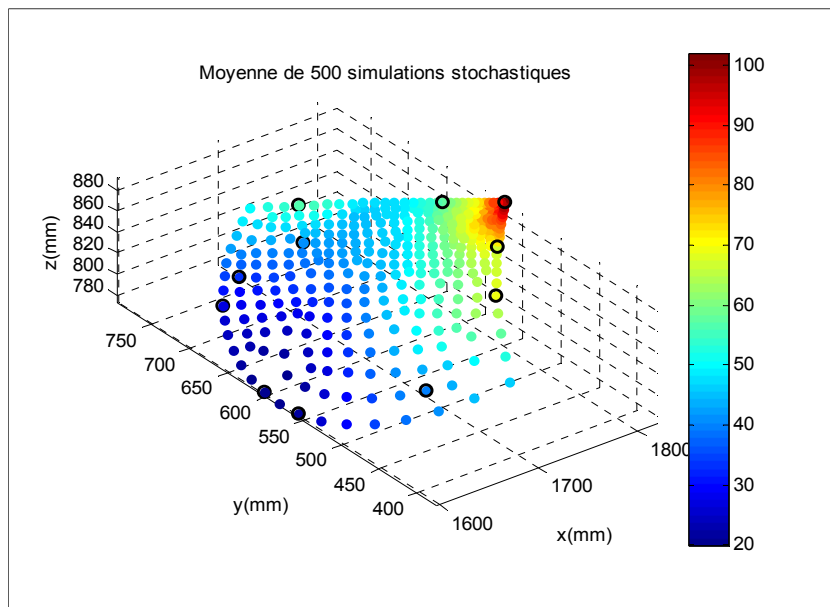


Figure 5-56 Contraintes de von Mises en 3D : moyenne de 500 simulations stochastiques

Arrivant à la dernière étape de la méthodologie expérimentale, on doit revenir de l'espace bidimensionnel à la surface tridimensionnelle. Les Figure 5-55 et Figure 5-56 montre les contraintes en 3D et incarne la fiabilité de la transformation bijective. Finalement, la Figure 5-57 montre la variance de simulations stochastiques tout en mettant l'accent sur la zone d'incertitude maximale proche de la zone de concentration de contraintes.

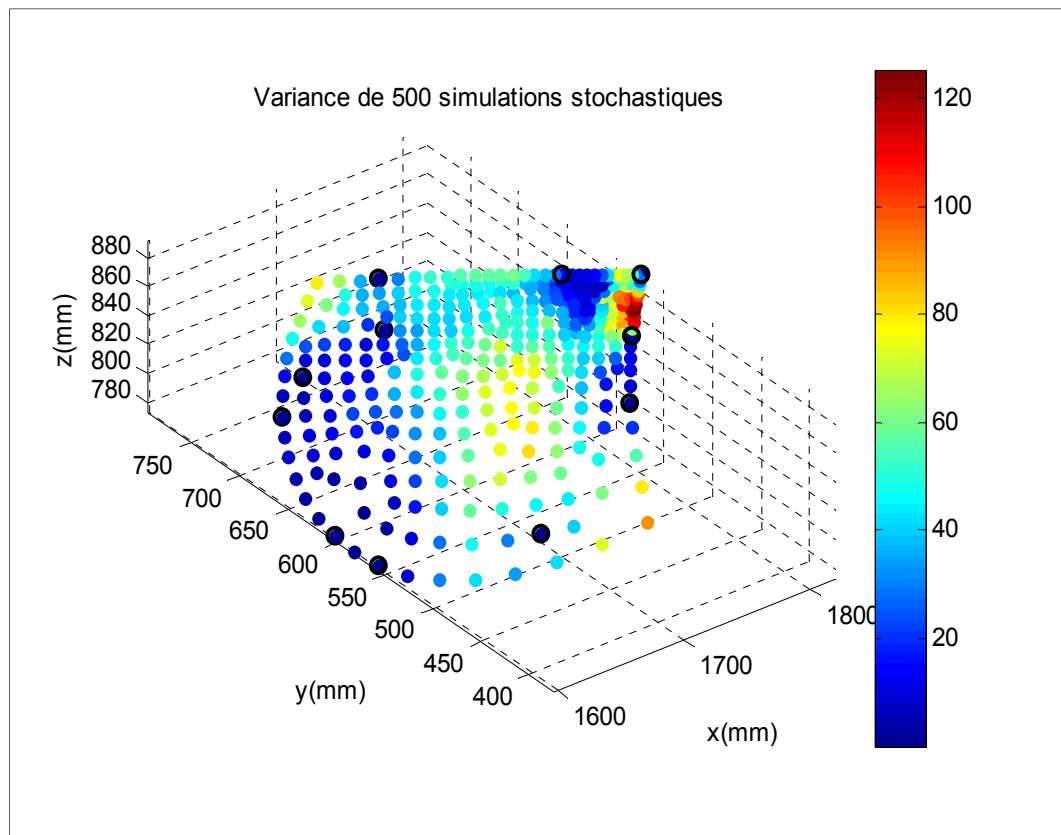


Figure 5-57 Contraintes de von Mises en 3D : variance de 500 simulations stochastiques

CONCLUSION

Tel que mentionné précédemment, notre objectif vise, dans un premier temps, à interpoler spatialement des mesures expérimentales et, dans un deuxième temps, à déterminer la propagation spatiale de l'erreur de ces mesures. Pour réaliser ces deux tâches et pour estimer les mesures manquantes, on a eu recours au krigeage, interpolateur linéaire, non biaisé et garant d'une variance de prédiction minimale. De plus, les simulations stochastiques conditionnelles nous ont permis de quantifier l'incertitude associée aux mesures manquantes tout en compensant l'effet du lissage.

La robustesse du modèle de krigeage employé repose sur le variogramme qui caractérise la variabilité spatiale des mesures. En effet, on a réalisé une analyse variographique afin de définir les paramètres optimaux du modèle retenu, et ce avant de commencer le krigeage et les simulations conditionnelles. Pour les deux cas d'étude présentés dans ce mémoire, on a conclu que le krigeage ordinaire est le plus adéquat. Aussi, on a démontré que les simulations séquentielles conditionnelles (SSCs) réalisées à l'aide de nos programmes MatLab[®] procurent des résultats quantitativement et qualitativement équivalents aux simulations gaussiennes conditionnelles (SGSs) réalisées avec le logiciel SGeMS.

Les simulations stochastiques conditionnelles ont permis de cartographier l'incertitude des mesures d'une façon plus réaliste que le krigeage, faute à son effet de lissage, qui représente mieux les variations réelles surtout dans le cas où le gradient de variation est important.

Le krigeage et les simulations stochastiques ont montré une dépendance aux paramètres du modèle du variogramme utilisé. Il est donc important de bien analyser et de valider les paramètres du variogramme, surtout sa portée comme on a montré pour l'aube de la turbine. Pour ce faire, on a ajusté le modèle en utilisant la méthode des moindres carrées puis la validation croisée. Si le variogramme n'est pas optimal, les estimations ainsi que leurs variances ne seront plus valides. De plus, il est nécessaire de bien choisir une configuration

(nombre et position) des points représentatifs du domaine de l'étude afin d'améliorer la fiabilité des estimations et de minimiser leurs variances.

Plusieurs méthodes statistiques ont été utilisées pour valider les résultats des estimations par krigeage ainsi que la moyenne des simulations conditionnelles. Les modèles théoriques ont montré leurs aptitudes pour reproduire les résultats référentiels des simulations par ÉFs. En effet, à partir d'un échantillon de l'ordre de 1% du domaine d'étude, on a pu trouver de forts coefficients de corrélation entre les valeurs théoriques ÉFs et les estimations. De plus, les analyses des résidus, ont montré une bonne capacité prédictive des estimateurs.

En somme, nous prétendons que la mission de ce projet a été accomplie avec succès. On est parvenu à trouver un modèle mathématique qui permet de déterminer les variances suite à une propagation spatiale des mesures. Néanmoins, on ne peut pas l'automatiser puisqu'un mauvais ajustement du variogramme engendrera invariablement une approximation déficiente de l'erreur. Ainsi, cette approche nécessite de l'utilisateur un suivi méthodique et attentif de la stratégie détaillée dans le chapitre qui traite de la méthodologie expérimentale, et d'avoir une certaine subjectivité lors de la modélisation du variogramme et même lors du traitement statistique des résultats.

Pour clore ce mémoire, ce travail m'a permis de valider un slogan que j'ai tenté d'appliquer durant toute la durée de mes recherches « Les catastrophes sont rarement le fruit de hasard ». J'ai joué avec les simulations stochastiques pour simuler plusieurs scénarios de la réalité afin de déterminer les incertitudes des mesures ainsi que leur propagation spatiale, et par conséquent, éviter les complications qui peuvent être engendrées suite à des mauvaises estimations.

RECOMMANDATIONS

Suite à notre étude, plusieurs recommandations peuvent être formulées. On commencera tout d'abord par des pistes d'améliorations qu'on pourra considérer puis par des travaux futurs qui peuvent améliorer la capacité prédictive de notre modèle d'estimation.

Dans un premier temps, on recommande d'utiliser le krigeage ordinaire lors du conditionnement car c'est un estimateur qui a minimisé la variance et qui est basé, structurellement parlant, sur la connaissance de la moyenne juste au voisinage du point cible, contrairement au krigeage simple qui la suppose constante partout sur le domaine d'étude ou le krigeage universel qui a affiché une variance nettement supérieure.

Le choix du voisinage lors du krigeage ordinaire doit être assez grand et doit inclure un nombre suffisant des mesures car le modèle variographique doit impérativement refléter la dépendance spatiale. On recommande aussi de tester plusieurs tailles du voisinage par la technique de validation croisée et de sélectionner la plus optimale lors du conditionnement du krigeage ou des simulations stochastiques.

Dans un deuxième temps, la variance de la prédiction dépend de la position et du nombre des mesures utilisées comme on a vu précédemment. Il est essentiel d'optimiser ces deux paramètres avant de démarrer le krigeage ou les simulations stochastiques. On recommande de suivre les recommandations de plusieurs travaux qui abordent cette problématique avant de commencer l'optimisation (Andronikov, Davidson et Spiers, 2000; Brus et Heuvelink, 2007; Heriawan et Koike, 2008).

- Ajustement du variogramme expérimental

L'ajustement du variogramme expérimental est une étape critique et délicate. En effet, le variogramme décrit la continuité des structures spatiales et la régularité du phénomène étudié. Dans le but d'améliorer la modélisation (ou la paramétrisation) du variogramme

expérimental, on a vu qu'on a retenu l'ajustement au sens des moindres carrés. La Figure 5-37 semble montrer un variogramme avec une structure emboîtée dans certaines directions. On parle donc dans ce cas d'un variogramme composé, ou ce qu'on appelle en littérature un variogramme gigogne. En effet, pour certaines directions d'étude, on peut distinguer plusieurs paliers de variation qui seront en rapport étroit avec la morphologie de la variable à estimer. Peu importe, le variogramme sera décomposé en plusieurs structures déduites des différentes fonctions de covariances avec plusieurs portées et paliers. Cette technique de décomposition est proche des méthodes d'analyse spectrale très populaires en prospection géophysique (Serra, 1968). On recommande d'étudier et d'investiguer le potentiel de ce genre de variogramme surtout qu'il semble réaliser un certain filtrage des données primaires.

On s'est limité dans ce mémoire juste à des modèles de variogrammes bornés et des types de krigeage linéaire. On a utilisé les variogrammes élémentaires notamment le modèle sphérique, exponentiel et gaussien, qui sont les plus employés dans la littérature du domaine. On recommandera de tester les autres modèles de variogrammes empiriques lors de la modélisation. D. Marcotte présente un répertoire des modèles empiriques dépassant la trentaine (Marcotte, 2010).

De plus, on pourra passer à un krigeage non linéaire en étudiant des combinaisons non linéaires de la fonction aléatoire tout en s'assurant de réaliser l'anamorphose gaussienne de la distribution des données primaires comme étape préalable (Lantuejoul, 2002) . Dans certaines circonstances, les méthodes linéaires peuvent atteindre leurs limites et les estimateurs linéaires ne seront plus optimaux surtout si on traite des variables présentant une distribution très asymétrique. Afin de garder une variance de krigeage minimale et réaliste, pour certains cas d'étude, on pourra examiner différentes approches notamment le krigeage disjonctif qui tient compte de la relation entre plusieurs variables (Rivoirard, 1991).

- Zone d'accroissement du variogramme

Tel que vu précédemment, la construction d'un variogramme expérimental dans notre cas spécifique se base principalement sur le maillage utilisé lors des simulations par ÉFs. Une première recommandation consiste à utiliser un maillage non régulier avec un pas variable. Le fait de raffiner le maillage, améliore la définition de la zone du décollage (faibles incréments) et par conséquent, évite le risque de l'amplification de l'effet de pépite.

Plus pratiquement, et lors de l'extraction des résultats des simulations par ÉFs, on a vu qu'on a transformé le maillage de sorte à maintenir uniquement un maillage triangulaire. On pourra réinjecter les nœuds moyens et considérer en plus les 4 triangles secondaires (chaque triangle sera formé par un sommet principal et deux nœuds moyens). Ainsi ce raffinement du maillage, permet d'augmenter le nombre des données utilisées lors de la construction du variogramme expérimental et par conséquent la fiabilité de la zone d'accroissement lors de la modélisation théorique.

- Variogramme des incertitudes

Les erreurs de mesure qui affectent la valeur à estimer peuvent être soit des erreurs d'implantation (la mesure n'est pas connue avec précision), soit des erreurs d'échantillonnage. L'erreur expérimentale sera représentée par l'effet de pépite comme on a mentionné précédemment. En fait, l'effet de pépite représente la variabilité entre deux mesures effectuées à des emplacements infiniment proches. Normalement cette variation doit être presque nulle vu la continuité de l'espace. Cet effet de pépite prend en compte la variabilité de l'instrument de mesure en partie, son erreur systématique.

Nos travaux ont réussi à déterminer l'incertitude locale d'une rosette (respectivement une jauge de déformation) collée ou soudée sur l'aube de la turbine. On a pu quantifier cette incertitude grâce à la méthodologie proposée par Arpin-Pont *et al.* MNEIM (Arpin-Pont, 2012). On pourra dans un premier temps simuler l'erreur locale dans plusieurs localisations

du domaine d'étude D et utiliser un effet de pépite moyen puis d'y déduire le variogramme expérimental correspondant à ces incertitudes. Ensuite, à partir de ce variogramme, on utilisera le krigeage et les simulations stochastiques pour déduire la propagation de l'incertitude expérimentale à chaque nœud du maillage. Finalement, en combinant les valeurs obtenues par éléments finis ainsi que l'erreur calculée précédemment, on aura la valeur expérimentale la plus représentative de la réalité.

- Cokrigeage

Vu qu'on a une information abondante à partir des simulations par ÉF, on pourra avoir recours à des informations auxiliaires lors de l'estimation de la variable cible. Le cokrigeage permet d'améliorer les estimations obtenues par le krigeage en utilisant l'information fournie par d'autres variables secondaires (Marcotte, 1991b). En géostatistiques par exemple, et lors de la modélisation de transmissivité d'un aquifère, on considère les transmissivités dans un modèle d'éléments finis comme variable principale et les conductivités hydrauliques obtenues par *slug-test* comme variable secondaires. On peut même combiner plusieurs variables secondaires notamment, les tests de pompage, les courbes granulométriques et la capacité spécifiques de puits. Dans notre cas d'étude, on pourra considérer la déformation de l'aube comme une variable primaire et la vitesse de l'écoulement de l'eau sur la surface de l'aube comme une variable secondaire. Une telle approche suppose que des corrélations pertinentes existent entre la variable principale et les variables auxiliaires. En effet, pour chaque nœud du maillage, on a accès à des déformations ainsi que des contraintes dans plusieurs directions. On pourra combiner ses informations pour améliorer la capacité prédictive de l'estimateur.

On doit tenir compte dans un premier temps de la dimension de variabilité spatiale, car certains de ses propriétés physiques (mécaniques) peuvent opérer à de petites distances contrairement à d'autres à des grandes distances. Dans un deuxième temps, comme pour la plupart des approches où toutes les sources de variation sont mixées (analyse de corrélation, analyse en composantes principale, régression linéaire multiple,...), les filtrages des

différentes composantes conduisent à une amélioration des relations entre les variables étudiées. Finalement, le co-krigeage se base sur les mêmes concepts théoriques que le krigeage. La différence majeure entre les deux, réside dans la construction du variogramme croisé expérimental.

- Anisotropie

Le variogramme dépend uniquement de la distance d'incrément. Le vecteur de déplacement lors de la recherche des paires de données contient une information sur la norme de l'incrément ainsi que son orientation. Si le variogramme dépend de cette orientation, il est anisotrope. On distingue deux catégories d'anisotropie : zonale et géométrique. Comme on doit construire plusieurs variogrammes directionnels au début de l'analyse variographique, on pourra détecter et déduire une éventuelle anisotropie. Si on a une anisotropie géométrique, les variogrammes des différentes directions auront le même palier mais pas la même portée. Afin de résoudre cette anomalie, on doit effectuer des transformations linéaires des coordonnées spatiales et des rotations suivant les directions de la plus petite et plus grande continuité (Deutsch et Journel, 1998).

Ainsi on pourra étudier cette anisotropie avant de modéliser le variogramme expérimental car si ce dernier est fiable, l'estimateur sera optimal. Cette technique (analyse de l'anisotropie) adapte l'échantillonnage en augmentant la densité dans la direction de plus faible portée. Par conséquent, pour le même coût d'échantillonnage, on peut obtenir des informations beaucoup plus précises si on ajuste la stratégie d'échantillonnage (positions et nombre de jauges de déformation) à l'anisotropie.

- Analyse bayésienne : SIPPI

Avec cette approche, l'estimateur se base sur un modèle *a priori* pour déterminer la distribution *a posteriori* suivant la relation suivante

$$P(\text{modèle} | \text{données}) \propto P(\text{données} | \text{modèle}) \times P(\text{modèle}) \quad (5.17)$$

Plusieurs méthodes peuvent déterminer la distribution de probabilité à posteriori notamment l'algorithme de Métropolis Hasting et la méthode de rejet (Hansen et al., 2013; Mejer Hansen et al., 2013). Les deux algorithmes permettent d'engendrer indirectement des observations à partir d'une distribution à priori. On part d'une loi instrumentale *proposal distribution* qui n'est pas toujours facile à définir (la distribution des estimations données par le krigeage sera le modèle à priori) et on déduit les estimations à posteriori. On recommande de combiner nos fonctions MatLab® et la librairie SIPPI⁴.

Le modèle bayésien typique en interpolation spatiale qui effectue un krigeage en utilisant fonction de covariance (un variogramme) avec des paramètres inconnus. Généralement, ces paramètres doivent respecter une information a priori liée à la tendance (la moyenne) de la fonction aléatoire. Un des plus grands avantages de l'approche bayésienne est la perceptive conditionnelle sans être obligé de répéter des expériences, condition souvent difficile à reproduire expérimentalement (Baillargeon, 2005). C'est un processus cumulatif, chaque donnée injectée, sera considérée comme une information conditionnelle supplémentaire, et sera utilisée pour réaliser une interpolation plus reproductrice de la réalité (mesure expérimentale).

Ainsi, en combinant ces recommandations avec le travail de ce mémoire, nous conjecturons qu'on pourra réduire d'avantage l'incertitude calculée par notre modèle.

⁴<http://sourceforge.net/p/sippi/wiki/Home/>

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Alfaro-Sironvalle , M. 1979. « Etude de la robustesse des simulations des fonctions aléatoires ». Fontainebleau, E.N.S.M.P. , 161 p.
- Andre G. Journel , Clayton V.Deutsch (364). 1998. *GSLIB: Geostatistical Software Library and User's Guide (Applied Geostatistics)*, 2. New York: Cambridge University Press.
- Andronikov, Sergei V., Donald A. Davidson et Ronald B. Spiers. 2000. « Variability in contamination by heavy metals: Sampling implications ». *Water, Air, and Soil Pollution*, vol. 120, n° 1-2, p. 29-45.
- Ankenman, Bruce, Barry L. Nelson et Jeremy Staum. 2010. « Stochastic Kriging for Simulation Metamodeling ». *Operations Research*, vol. 58, n° 2, p. 371-382.
- Armstrong, Margaret. 1998. *Basic linear geostatistics*. Berlin [u.a.: Springer.
- Arpin-Pont, Jeremy. 2012. « Methode de determination des incertitudes de mesures par jauges de deformation ». M.Eng. Ann Arbor, Ecole de Technologie Superieure (Canada), 154 p. In ProQuest Dissertations & Theses Full Text.
- Baillargeon, Sophie. 2005. « Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations ». Université de Laval Québec, 137 p.
- Bezier, Pierre. 1987. *Courbes et surfaces*, 2e ed. rev. et augm.. Paris: Hermès, 221 p., [8] p. de pl. p.
- Bier, E. A., et K. R. Sloan. 1986. « Two-Part Texture Mappings ». *Computer Graphics and Applications, IEEE*, vol. 6, n° 9, p. 40-53.
- Bochner , S. (161 p). 1960. *Harmonic analysis and the theory of probability*. University of California Press.
- Boissonnat, J.D., P. Chenin, A. Cohen, C. Gout, T. Lyche, M.L. Mazure et L. Schumaker. 2011. *Curves and Surfaces: 7th International Conference, Avignon, France, June 24-30, 2010, Revised Selected Papers*. Springer.
- Bourennane, H., D. King, P. ChÉRY et A. Bruand. 1996. « Improving the kriging of a soil variable using slope gradient as external drift
- Bourgault, Gilles, et Denis Marcotte. 1991. « Multivariable variogram and its application to the linear model of coregionalization ». *Mathematical Geology*, vol. 23, n° 7, p. 899-928.

- Bourgeois , Jocelyn. 2006. « Le krigeage comme méthode d'approximation pour l'optimisation ». École polytechnique de Montréal, 155 p.
- Box , G.E.P. , Jenkins , G.M. 1970. *Time series analysis, forecasting and control*. London: Holden Day, 553 p.
- Brus, Dick J., et Gerard B. M. Heuvelink. 2007. « Optimization of sample patterns for universal kriging of environmental variables ». *Geoderma*, vol. 138, n° 1-2, p. 86-95.
- Caers, Jef. 2011. « Modeling uncertainty in the earth sciences ». < <http://site.ebrary.com/id/10483226> >.
- Chilès , J.P. . 1977. « Thèse de docteur ingénieur : Géostatistique des phénomènes non stationnaires dans le plan ». Nancy, 152 p.
- Chilès, J.P., et P. Delfiner. 2012. *Geostatistics: Modeling Spatial Uncertainty*. Wiley.
- Christakos, George Bogaert Patrick Serre Marc L. 2001. *Temporal GIS : advanced functions for field based applications ; with 23 tables*. Berlin [u.a.]: Springer.
- Cressie, N.A.C. 1993. *Statistics for spatial data*. J. Wiley.
- Cressie, Noel. 1985. « Fitting variogram models by weighted least squares ». *Journal of the International Association for Mathematical Geology*, vol. 17, n° 5, p. 563-586.
- Cressie, Noel. 1990. « The origins of kriging ». *Mathematical Geology*, vol. 22, n° 3, p. 239-252.
- Deutsch, Clayton V., et A.G. Journel. 1998. *Gslib: Geostatistical Software Library and User's Guide*. Oxford University Press, Incorporated.
- Deutsch, ClaytonV, et PerryW Cockerham. 1994. « Practical considerations in the application of simulated annealing to stochastic simulation ». *Mathematical Geology*, vol. 26, n° 1, p. 67-82.
- Dikshit, H. P., et A. Ojha. 2003. « A simple subdivision formula for quadrilateral Wachspress patches ». *Computer Aided Geometric Design*, vol. 20, n° 7, p. 395-399.
- Dubrulle, Olivier. 1983. « Two methods with different objectives: Splines and kriging ». *Journal of the International Association for Mathematical Geology*, vol. 15, n° 2, p. 245-257.
- Farin, Gerald E. 1990. *Curves and surfaces for computer aided geometric design : a practical guide*, 2nd ed.. San Diego, Calif.: Academic Press, xvii, 444 p., [4] p. de pl. p.

- Floater, Michael S., et Martin Reimers. 2001. « Meshless parameterization and surface reconstruction ». *Computer Aided Geometric Design*, vol. 18, n° 2, p. 77-92.
- Froidevaux, R. 2004. « Sequential Updating Simulation ». In *geoENV IV — Geostatistics for Environmental Applications*, sous la dir. de Sanchez-Vila, Xavier, Jesus Carrera et JoséJaime Gómez-Hernández. Vol. 13, p. 307-318. Coll. « Quantitative Geology and Geostatistics »: Springer Netherlands. < http://dx.doi.org/10.1007/1-4020-2115-1_26 >.
- Gaetan, Carlo Guyon Xavier. 2008. « Modélisation et statistique spatiales ». < <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=367442> >.
- Gaetan, Carlo Guyon Xavier Bleakley Kevin. 2010. *Spatial statistics and modeling*. New York; London: Springer.
- Gagnon, M., S. A. Tahan, P. Bocher et D. Thibault. 2010. « Impact of Startup Scheme on Francis Runner life Expectancy ». In *25th IAHR Symposium on Hydraulic Machinery and Systems, 20-24 Sept. 2010*. (UK) Vol. 12, p. 012107 (8 pp.). Coll. « IOP Conf. Ser., Earth Environ. Sci. (UK) »: IOP Publishing Ltd. < <http://dx.doi.org/10.1088/1755-1315/12/1/012107> >.
- Gagnon, Martin, Antoine Tahan, Philippe Bocher et Denis Thibault. 2013. « A probabilistic model for the onset of High Cycle Fatigue (HCF) crack propagation: Application to hydroelectric turbine runner ». *International Journal of Fatigue*, vol. 47, n° 0, p. 300-307.
- Goovaerts, P. 2000a. « Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall ». *Journal of hydrology*, vol. 228, n° 1, p. 113-129.
- Goovaerts, P. 2000b. « Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall ». *Journal of Hydrology*, vol. 228, n° 1–2, p. 113-129.
- Goovaerts, P. 2002. « Geostatistical modelling of spatial uncertainty using p -field simulation with conditional probability fields ». *International Journal of Geographical Information Science*, vol. 16, n° 2, p. 167-178.
- Gooverts, P. 1997. « Geostatistics for natural resources evaluation: Oxford University Press ». *New York*.
- Gratton, Yves. 2002. « LE KRIGEAGE : LA MÉTHODE OPTIMALE D'INTERPOLATION SPATIALE ». *Les articles de l'Institut d'Analyse Géographique*.

- Graupe D., Krause D.G. , Moore J.B. 1975. « Identification of autoregressive moving-average parameters of time series ». *I.E.E.E Trans. Automat. Contr. AC.* 20, p. 104–107.
- Hansen, Thomas Mejer, Knud Skou Cordua, Majken Caroline Looms et Klaus Mosegaard. 2013. « SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 2—Application to crosshole GPR tomography ». *Computers & Geosciences*, vol. 52, n° 0, p. 481-492.
- Hardy, Rolland L. 1971. « Multiquadric equations of topography and other irregular surfaces ». *Journal of Geophysical Research*, vol. 76, n° 8, p. 1905-1915.
- Hastie, T. J. , Tibshirani, R. J. 1990. *Generalized Additive Models*, 1. Chapman and Hall, 352 p.
- Heriawan, Mohamad Nur, et Katsuaki Koike. 2008. « Uncertainty assessment of coal tonnage by spatial modeling of seam distribution and coal quality ». *International Journal of Coal Geology*, vol. 76, n° 3, p. 217-226.
- Hillier, F.S., et G.J. Lieberman. 2010. *Introduction to Operations Research*. McGraw-Hill Higher Education.
- Hormann, Kai, Bruno L, #233, vy et Alla Sheffer. 2007. « Mesh parameterization: theory and practice <h3>Video files associated with this course are available from the citation page</h3> ». In *ACM SIGGRAPH 2007 courses*. (San Diego, California), p. 1. 1281510: ACM.
- Jean-Paul Chilès, Pierre Delfiner. 2012. *Geostatistics: Modeling Spatial Uncertainty, 2nd Edition*. New York: John Wiley & Sons 734 p.
- Journel, A.G., et C. Huijbregts. 1978. *Mining geostatistics*. Academic Press.
- Journel A.G. , Huijbregts Ch.J. (600). 1981. *Mining geostatistics*. London: Academic Press.
- Journel, AndreG, et Zhanjun Ying. 2001. « The Theoretical Links Between Sequential Gaussian Simulation, Gaussian Truncated Simulation, and Probability Field Simulation ». *Mathematical Geology*, vol. 33, n° 1, p. 31-40.
- Kergreis, A., et J. Claude. 1991. *Utilisation rationnelle de l'eau des petits bassins versants en zone aride*. John Libbey Eurotext.
- King, SusanL. 2000. « Sequential Gaussian simulation vs. simulated annealing for locating pockets of high-value commercial trees in Pennsylvania ». *Annals of Operations Research*, vol. 95, n° 1-4, p. 117-203.

- Krige, D.G. 1952. « A statistical approach to some basic mine valuation problems on the Witwatersrand ». *J. of Chem., Metal. and Mining Soc. of South Africa*, p. 119-139.
- Krishnamurthy, Venkat, et Marc Levoy. 1996. « Fitting smooth surfaces to dense polygon meshes ». In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. p. 313-324. 237270: ACM.
- Lantuejoul, C. 2002. *Geostatistical Simulation: Models and Algorithms*. Springer.
- Laporterie, Florence, Olivier Amram, Guy Flouzat, Ewa Pilich et Michel Gay. 2000. « Data fusion thanks to an improved morphological pyramid approach: Comparison loop on simulated images and application to SPOT 4 data ». In *2000 International Geoscience and Remote Sensing Symposium (IGARSS 2000), July 24, 2000 - July 28, 2000*. (Honolulu, HI, USA) Vol. 5, p. 2117-2119. Coll. « International Geoscience and Remote Sensing Symposium (IGARSS) »: IEEE.
- Liao, Hongtao. 1990. « Estimation des reserves recuperables de gisements d'or : comparaison entre krigeage disjonctif et krigeage des indicatrices ». <http://www.theses.fr/1990ORLE2013>.
- Mantoglou, Aristotelis, et John L. Wilson. 1982. « The Turning Bands Method for simulation of random fields using line generation by a spectral method ». *Water Resources Research*, vol. 18, n° 5, p. 1379-1394.
- Marcotte, D. 1991a. « Cokriging with Matlab ». *Computers & Geosciences*, vol. 17, n° 9, p. 1265-80.
- Marcotte, D. 2010. « Cours GML6402 : Géostatistique ». < <http://www.groupe.polymtl.ca/geo/marcotte/gml6402.html> >.
- Marcotte, D., et P. Groleau. 1997. « A Simple and Robust Lognormal Estimator ». *Mathematical Geology*, vol. 29, n° 8, p. 993-1008.
- Marcotte, Denis. 1991b. « Cokriging with Matlab ». *Computers and Geosciences*, vol. 17, n° 9, p. 1265-1280.
- Matheron, G (261). 1975. *Random sets and integral geometry*. New York: Wiley and Sons.
- Matheron, G. 1973. « The intrinsic random functions and their applications ». *Adv. Appl. Prob.*, vol. Vol 5, p. 439-468.
- Matheron, G. 1975. *Splines and Kriging: Their Formal Equivalence*.
- Matheron, Georges (83). 1969. *Le Krigeage universel*. Paris: Ecole nationale supérieure des mines de Paris.

- Mejer Hansen, Thomas, Knud Skou Cordua, Majken Caroline Looms et Klaus Mosegaard. 2013. « SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 1—Methodology ». *Computers & Geosciences*, vol. 52, n° 0, p. 470-480.
- Nowak, Marek, et Georges Verly. 2005. « The Practice of Sequential Gaussian Simulation ». In *Geostatistics Banff 2004*, sous la dir. de Leuangthong, Oy, et Clayton V Deutsch. Vol. 14, p. 387-398. Coll. « Quantitative Geology and Geostatistics »: Springer Netherlands. < http://dx.doi.org/10.1007/978-1-4020-3610-1_39 >.
- Peachey, Darwyn R. 1985. « Solid texturing of complex surfaces ». *SIGGRAPH Comput. Graph.*, vol. 19, n° 3, p. 279-286.
- Piegl, Les A. 1997. *The NURBS book*, 2nd ed.. Berlin / New York: Springer, xiv, 646 p. p.
- Pinkall, Ulrich, et Konrad Polthier. 1993. « Computing Discrete Minimal Surfaces and Their Conjugates ». *Experimental Mathematics*, vol. 2, n° 1, p. 15-36.
- Poirier, Marc. 2013. « MODÉLISATION ET SIMULATION DU COMPORTEMENT DYNAMIQUE DES AUBES DE TURBINES HYDROÉLECTRIQUES ». M.Ing. Ann Arbor, Ecole de Technologie Supérieure (Canada), 223 p. In ProQuest Dissertations & Theses Full Text.
- Ripley B.D 1981. *Spatial statistics*. New York: John Wiley & Sons Inc.
- Rivoirard, J. 1991. *Introduction au krigeage disjonctif et à la géostatistique non-linéaire: cours de géostatistique non-linéaire*. ENSMP.
- Rivoirard, Jacques. 1994. *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford; New York: Clarendon Press ; Oxford University Press.
- Samek, Marcel, Cheryl Slean et Hank Weghorst. 1986. « Texture mapping and distortion in digital graphics ». *The Visual Computer*, vol. 2, n° 5, p. 313-320.
- Serra, J. 1968. « Les structures gigognes: morphologie mathématique et interprétation métallogénique ». *Mineralium Deposita*, vol. 3, n° 2, p. 135-154.
- Shi, Y., et G.D. van Albada. 2007. *Computational Science - ICCS 2007: 7th International Conference, Beijing China, May 27-30, 2007, Proceedings*. Springer.
- Stein, Michael Leonard. 1999. *Interpolation of spatial data : some theory for kriging*. New York, NY [u.a.]: Springer.
- Suro-Pérez, V., et A. G. Journel. 1991. « Indicator principal component kriging ». *Mathematical Geology*, vol. 23, n° 5, p. 759-788.

- Szczota, Mickael. 2012. « Modelisation de l'historique d'operation de groupes turbine-alternateur ». M.Ing. Ann Arbor, Ecole de Technologie Superieure (Canada), 130 p. In ProQuest Dissertations & Theses Full Text.
- Trauth, Martin H. Gebbers Robin Marwan Norbert Sillmann Elisabeth. 2010. « MATLAB® recipes for earth sciences ». < <http://public.eblib.com/EBLPublic/PublicView.do?ptiID=645372> >.
- Trochu, F. 2003. *Krigeage en CAO et FAO : notes de cours*. École polytechnique de Montréal.
- van Groenigen, J. W., et A. Stein. 1998. « Constrained Optimization of Spatial Sampling using Continuous Simulated Annealing ». *J. Environ. Qual.*, vol. 27, n° 5, p. 1078-1086.
- Webster, R. Oliver M. A. 2007. « Geostatistics for environmental scientists ». < <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=213965> >.