



INERIS



**Reconnaissance assistée de sites pollués par
l'utilisation conjointe de mesures rapides sur site
et de traitements géostatistiques – partie 2**

**Conception et validation d'une démarche itérative
de reconnaissance**

Rapport final REPERAGE

Avec le soutien de



Novembre 2008
RP FSSADEME2008001



INERIS



**Reconnaissance assistée de sites pollués par
l'utilisation conjointe de mesures rapides sur site
et de traitements géostatistiques – partie 2**

**Conception et validation d'une démarche itérative
de reconnaissance**

Rapport final REPERAGE

Hélène Demougeot-Renard
Laurence Haouche-Belkessam
Sébastien Denys

Novembre 2008
FSSADEME2008001

REDACTEURS	VERIFICATEURS	AUTORISATION DE DIFFUSION
H. Demougeot-Renard, L. Haouche-Belkessam, S. Denys	M.Garcia, L. Haouche-Belkessam	M. Garcia
		

Responsables du projet :

Pour FSS International r&d : Michel Garcia et H  l  ne Demougeot-Renard : non

Pour INERIS : S  bastien Denys et Karine TACK

Pour ISSeP : Marcel Lambert et Laurence Haouche-Belkessam

Confidentialit  

non	<input checked="" type="checkbox"/>
oui	<input type="checkbox"/>
dur��e	

N   de contrat ADEME : 06 72 C0015**Date du contrat :** 15/12/2006**Dur  e du contrat :** 26 mois**Nom du responsable ADEME :** Martine Louvrier**RELECTURES**

Nom	Organisme	Date
L. Haouche-Belkessam	ISSeP	11/2008
M. Garcia	FSS International r&d	12/2008

DIFFUSION

Destinataires	Nombre d'exemplaires du rapport version papier	Commentaires
ADEME	1	+ version pdf du rapport + CD-Rom des fichiers des donn��es et r��sultats
INERIS	1	+ version pdf du rapport + CD-Rom des fichiers des donn��es et r��sultats
ISSeP	1	+ version pdf du rapport + CD-Rom des fichiers des donn��es et r��sultats
FSS International r&d	1	+ version pdf du rapport + CD-Rom des fichiers des donn��es et r��sultats
GIS 3SP	1	+ version pdf du rapport + CD-Rom des fichiers des donn��es et r��sultats

Les donn  es et textes de ce rapport sont la co-propri  t   de FSS International r&d, de l'INERIS et de l'ISSeP. Ils ne peuvent   tre utilis  s    d'autres fins, sauf avec l'autorisation formelle de FSS International r&d, de l'INERIS et de l'ISSeP.

TABLE DES MATIERES

1	INTRODUCTION	8
1.1	Contexte	8
1.2	Objectifs	9
1.3	Les différentes étapes de travail	10
1.4	Objet et contenu du rapport final	12
2	PRESENTATION DE LA DEMARCHE DE COUPLAGE MESURES SUR SITE ET ANALYSE GEOSTATISTIQUE.....	13
2.1	Principes généraux de la démarche (Figure 1)	13
2.2	Présentation détaillée des étapes de la démarche	16
2.2.1	<i>Etape préliminaire de reconnaissance (étape 0)</i>	<i>16</i>
2.2.2	<i>Etapes de reconnaissance itérative.....</i>	<i>19</i>
2.2.3	<i>Critère d'arrêt de la reconnaissance et résultats finaux</i>	<i>21</i>
2.3	Points théoriques sur les conditions de modélisation	21
2.3.1	<i>Génération de teneurs de laboratoire.....</i>	<i>21</i>
2.3.2	<i>Classification des sols</i>	<i>22</i>
2.3.3	<i>Positionnement des points de mesure pour l'étape de reconnaissance suivante.....</i>	<i>22</i>
2.4	Le code développé pour la mettre en œuvre	24
2.5	Les conditions pratiques de mise en œuvre.....	24
3	TEST ET VALIDATION DE LA DEMARCHE SUR LES ENVIRONS DU SITE METALEUROP	26
3.1	Présentation du cas d'étude	26
3.2	Constitution de l'état de référence	28
3.2.1	<i>La campagne de terrain.....</i>	<i>28</i>
3.2.2	<i>Modélisation de l'état de référence.....</i>	<i>31</i>
3.3	Simulation de reconnaissances conduites par la démarche de couplage entre mesures FPXRF et analyse géostatistique.....	38
3.4	Confrontation des résultats de simulation à l'état de référence.....	42
3.5	Principales conclusions et retours d'expérience	43
3.5.1	<i>Validité de la démarche</i>	<i>43</i>
3.5.2	<i>Retours d'expérience</i>	<i>43</i>
4	MISE EN ŒUVRE « EN TEMPS REEL » DE LA DEMARCHE DE COUPLAGE SUR LE SITE HENRICOT	45
4.1	Présentation du site d'étude.....	45
4.2	Les conditions de mise en œuvre de la démarche.....	46
4.3	La campagne de terrain.....	47
4.4	Le traitement des données.....	48
4.4.1	<i>Etape préliminaire.....</i>	<i>48</i>
4.4.2	<i>Etapes itératives</i>	<i>49</i>
4.5	Les résultats obtenus	50
4.6	Principales conclusions et retours d'expérience	52
5	CONDITIONS D'INTEGRATION DE LA DEMARCHE DANS LES METHODOLOGIES FRANCAISE ET WALLONE DE GESTION DES SITES POLLUES	53
5.1	Méthodologie française	53
5.1.1	<i>Rappel sur les outils de gestion des sites et sols pollués en France</i>	<i>53</i>
5.1.2	<i>Application de la méthode dans le contexte français</i>	<i>53</i>
5.1.3	<i>Extrapolation à d'autres contaminants</i>	<i>54</i>
5.2	Méthodologie wallonne.....	55
6	ASPECTS FINANCIERS	56
6.1	Coûts de la campagne de reconnaissance avec mesures FPXRF sur MétalEurop	56
6.2	Coûts de la campagne de reconnaissance mettant en œuvre la démarche de couplage mesures FPXRF – analyse géostatistique sur Henricot II.....	56
7	CONCLUSIONS ET PERSPECTIVES	58
7.1.1	<i>Principaux résultats</i>	<i>59</i>
7.1.2	<i>Retours d'expérience</i>	<i>61</i>

7.1.3 Perspectives	61
--------------------------	----

LISTE DES TABLEAUX

Tableau 1 Statistiques élémentaires sur les teneurs en Pb mesurées par FPXRF in situ ou sur site et sur celles analysées en laboratoire pour le secteur d'étude MétalEurop.	31
Tableau 2 Caractéristiques des données écartées pour simuler la référence-labo du cas d'étude MétalEurop.	33
Tableau 3 Caractéristiques des données écartées pour simuler la référence-FPXRF in situ du cas d'étude MétalEurop.	34
Tableau 4 Statistiques élémentaires des teneurs en Pb des simulations référence-labo et référence-FPXRF in situ générées sur le secteur d'étude MétalEurop.	34
Tableau 5 Paramètres retenus pour calculer les indices de risque à partir des simulations des teneurs en Pb sur la couche 0-5 cm.	37
Tableau 6 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Volumes de sol classés « pollués », « sains » et « incertains » à chaque nouvelle étape de reconnaissance.	41
Tableau 7 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Conditions de modélisation à chaque nouvelle étape de reconnaissance.	50
Tableau 8 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Volumes de sol classés « pollués », « sains » et « incertains » à chaque nouvelle étape de reconnaissance.	52
Tableau 9 Coûts effectifs de la campagne de reconnaissance menée sur le cas d'étude MétalEurop pour constituer un état de référence (sans couplage avec un traitement géostatistique). Unité : € HT.	56
Tableau 10 Coûts effectifs de la campagne de reconnaissance menée sur le cas d'étude Henricot II.	57
Tableau 11 Statistiques élémentaires des teneurs en Pb de laboratoire disponibles sur le secteur d'étude MétalEurop.	88
Tableau 12 Statistiques élémentaires des teneurs en Pb de laboratoire disponibles sur le secteur d'étude MétalEurop.	92
Tableau 13 Coefficients de corrélation entre rapport de teneurs en Pb mesurées in situ par XRF ou au laboratoire par ICP et droites de régression.	97
Tableau 14 Teneur en Pb sur échantillon 0-5cm mesurée réellement au laboratoire et estimateur de ces teneurs calculé à partir de la relation entre les rapports de mesures sur échantillons 0-5cm et 0-20cm de terrain et de laboratoire. Les teneurs pour lesquelles les estimateurs génèrent une erreur de classification par rapport à un seuil réglementaire (200 et 550 mg/kg ms) apparaissent en gras.	99
Tableau 15 Statistiques élémentaires des teneurs en Pb mesurées in situ par XRF à la surface du sol (0 cm) - REPERAGE.	100
Tableau 16 Statistiques élémentaires des teneurs en Pb mesurées in situ par XRF à la surface du sol (0 cm) – OUTRAGE et REPERAGE.	100
Tableau 17 Identifiants et coordonnées (Lambert II) des points de mesures prévues pour l'étape itérative 5.	109

LISTE DES FIGURES

Figure 1 Schéma de la démarche de couplage de mesures sur site et d'un traitement géostatistique.	15
Figure 2 Loi de régression typique, entre analyses de laboratoire et mesures sur site, définie par une droite de régression moyenne (a) et une fonction polynômiale de l'écart-type du résidu autour de la droite de régression (b), écart-type qui tend à augmenter avec la teneur. Les paramètres de la loi peuvent être ajustés à partir des données collectées lors de l'étape préliminaire.	22
Figure 3 Schéma du principe de positionnement des nouveaux points de mesure dans la zone « incertaine ». Remarque : la grille d'échantillonnage et la grille de classification peuvent parfois avoir des directions différentes.	23
Figure 4 Localisation théorique des points d'échantillonnage et de mesure FPXRF dans le secteur d'étude MétalEurop. Fonds topographique IGN (base de données BD carto®). Lambert II étendu.	27

Figure 5 Localisation réelle des points d'échantillonnage et de mesure FPXRF dans le secteur d'étude MétalEurop (ortho-photos IGN).....	30
Figure 6 Etats de référence de la pollution moyenne en Pb sur les 20 premiers de sol du secteur d'étude MétalEurop. (a) teneurs de la référence-labo (b) teneurs de la référence FPXRF in situ (c) classification par rapport aux seuils d'action nationale Pb de la référence-labo (d) classification de la référence-FPXRF in situ.....	35
Figure 7 Résultats de l'application virtuelle de la démarche de couplage sur MétalEurop. Evolution de la classification des sols au fur et à mesure de l'acquisition de nouvelles mesures FPXRF.....	40
Figure 8 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Evolution des volumes de sols classés « contaminés », « sains » ou « incertains » au fur et à mesure de l'acquisition de nouvelles mesures FPXRF.....	42
Figure 9 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Confrontation de la classification des sols obtenue en fin de reconnaissance à la classification de l'état de référence de la contamination.....	43
Figure 10 Limite du secteur d'étude sur le site de l'ancienne usine sidérurgique Henricot II et localisation des 18 sondages prévus pour l'étape préliminaire de reconnaissance.....	46
Figure 11 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Evolution de la classification des sols au fur et à mesure de l'acquisition de nouvelles mesures FPXRF.....	51
Figure 12 Carte de localisation des données de Pb de laboratoire. Secteur d'étude MétalEurop.....	88
Figure 13 Variogramme expérimental des données de Pb de laboratoire et modèle ajusté. 10 pas de 300 m. Site d'étude MétalEurop.....	89
Figure 14 Simulation #2 SSG des teneurs en Pb de laboratoire sur le secteur d'étude MétalEurop ..	89
Figure 15 Simulation #2 SSG des teneurs en Pb de laboratoire après correction sur le secteur d'étude MétalEurop (gauche) et classification par rapport aux seuils d'action nationale Pb (droite).....	90
Figure 16 Relation entre moyenne des teneurs FPXRF in situ et des analyses ICP labo REPERAGE. Site d'étude MétalEurop.....	91
Figure 17 Relation entre moyenne des teneurs FPXRF in situ corrigées de la teneur en eau et des analyses ICP labo REPERAGE. Site d'étude MétalEurop.....	91
Figure 18 Carte des données de Pb FPXRF de terrain (moyenne de mesures FPXRF in situ par verticale). Secteur d'étude MétalEurop.....	92
Figure 19 Variogramme expérimental des données de Pb de terrain (moyenne de mesures FPXRF in situ par verticale) et modèle ajusté. 8 pas de 300 m.....	93
Figure 20 Simulation #2 des teneurs en Pb (moyenne des mesures FPXRF in situ par verticale) et classification par rapport aux seuils d'action nationale Pb (droite).....	94
Figure 21 Simulation #2 SSG des teneurs FPXRF en Pb après correction sur le secteur d'étude MétalEurop (gauche) et classification par rapport aux seuils d'action nationale Pb (droite).....	94
Figure 22 Diagrammes de dispersion entre le rapport des teneurs en Pb de laboratoire R_{labo} (ordonnée) et le rapport des teneurs en Pb mesurées in situ par XRF (abscisse). Ce dernier rapport est R_{XRF_moy} sur le graphe (a), R_{XRF_0cm} sur le graphe (b) et R_{XRF_5cm} sur le graphe (c). ...	98
Figure 23 Diagramme de dispersion entre teneur en Pb de laboratoire estimée (X) et teneur en Pb de laboratoire réellement mesurée (Y).....	99
Figure 24 Variogramme expérimental des données REPERAGE de Pb de terrain (mesures XRF in situ à 0 cm de profondeur) et modèle ajusté. 8 pas de 300 m.....	100
Figure 25 Variogramme expérimental des données OUTRAGE et REPERAGE de Pb de terrain (mesures XRF in situ à 0 cm de profondeur) et modèle ajusté. 8 pas de 300 m.....	101
Figure 26 Simulation de la contamination en Pb dans les 5 premiers de sol du secteur d'étude MétalEurop. (1) teneurs de laboratoire (2) teneurs mesurées par FPXRF (3) indices de risques calculés à partir des teneurs de laboratoire (4) indices de risque calculés à partir des mesures FPXRF.....	102
Figure 27 Carte de positionnement de 12 sondages supplémentaires pour l'étape itérative 5 de reconnaissance du cas d'étude Henricot.....	109

LISTE DES ANNEXES

Annexe 1 – Contenu du CD-Rom	63
Annexe 2 – Références bibliographiques	65
Annexe 3 – Descriptif des scripts constitutifs du code Matlab de mise en œuvre de la démarche de couplage mesures sur site - géostatistique	68
Annexe 4 – Conditions de simulations des états de référence laboratoire et FPXRF	87

Annexe 5 – Analyse exploratoire et variographique des données collectées sur la couche de sol 0-5cm	95
Annexe 6 – Document de préparation de l'étape préliminaire de reconnaissance sur Henricot II	103
Annexe 7 – Documentation de terrain fournie à chaque étape itérative de reconnaissance sur le cas d'étude Henricot II	108

1 INTRODUCTION

1.1 Contexte

La reconnaissance d'un site pollué effectuée avec des méthodes d'analyses conventionnelles est une étape longue et par conséquent coûteuse, qui incite à limiter le nombre de points d'échantillonnage. Le nombre et la position de ces points sont généralement définis à l'avance sans pouvoir être réajustés, pour mieux délimiter les sols pollués en cours de campagne de reconnaissance, compte tenu des délais longs d'obtention des résultats d'analyses de laboratoire. De telles pratiques vont à l'encontre d'une bonne précision d'estimation de l'état de pollution des sites. Pour un même phénomène et un même protocole de mesure, la précision est en effet généralement d'autant plus faible que le nombre de points de mesures est réduit. Cette situation est d'autant plus critique que l'erreur d'estimation ne fait pas forcément l'objet d'une quantification. Les erreurs d'estimation sont mal connues des praticiens. Les méthodes d'interpolation les plus utilisées, méthodes empiriques ou méthodes d'interpolation standards, ne permettent pas de quantifier l'incertitude d'estimation. Les erreurs d'estimation peuvent donc être élevées sans que les décideurs en aient conscience, et être davantage subies que choisies.

Une amélioration de la caractérisation des sites pollués, en termes de précision d'estimation et dans des conditions de délais et de coûts acceptables, peut cependant être espérée en mettant en œuvre deux types d'outils.

- Les outils de diagnostic rapide sur site qui permettent de quantifier certaines substances polluantes en un temps très court, à un coût généralement moindre que celui des coûts d'analyses en laboratoire et avec un résultat immédiatement disponible.
- Les méthodes géostatistiques développées initialement pour les besoins de l'industrie minière et qui sont adaptées au traitement des données de sites pollués. A partir d'un ensemble d'informations directes ou indirectes (auxiliaires), elles permettent en effet d'estimer des teneurs en polluants en tenant compte de la corrélation spatiale observée sur les données et de quantifier l'incertitude d'estimation. Leur utilisation pour cartographier les teneurs en polluants et estimer les quantités de terres à traiter se pratique de plus en plus fréquemment.

Le couplage des deux outils permet d'envisager la conduite de campagnes de reconnaissance optimisées. Il donne la possibilité de déterminer, au fur et à mesure de l'avancement de la reconnaissance, le nombre et la localisation de nouveaux points d'échantillonnage pour diminuer au mieux l'incertitude sur les estimations par rapport à des objectifs d'étude (par ex. cartographie des teneurs en polluants dans les sols, quantification des volumes de sols contaminés). Au-delà de l'aspect purement caractérisation, la méthodologie proposée ouvre la voie vers un outil de représentation cartographique optimisée de la pollution et de ses risques, outil qui améliore significativement l'aide à la décision et la communication auprès des différentes parties prenantes impliquées dans un dossier.

Si les fondements théoriques et expérimentaux existent pour les outils de diagnostic rapide et les méthodes géostatistiques, certains aspects de la démarche demandent encore à être approfondis et font l'objet du projet OUTRAGE - REPERAGE.

1.2 Objectifs

Le projet de recherche OUTRAGE – REPERAGE se fixe comme objectifs :

- de répondre aux interrogations soulevées par l'utilisation conjointe d'outils de diagnostic rapide sur site et de méthodes géostatistiques,
- de développer une méthodologie,
- de démontrer l'intérêt, mais aussi les limites, d'un tel couplage.

Il s'agit d'étudier, non seulement sa faisabilité technique, mais aussi son intérêt financier. Ce couplage permet-il d'améliorer la cartographie des teneurs en polluants à moindre coût de reconnaissance ? Permet-il de réduire les risques de mauvais classement des terres polluées par rapport à celles non polluées, et donc les risques environnementaux et les risques financiers associés ? Permet-il d'envisager une approche en temps réel de la cartographie de la pollution d'un site, en valorisant le caractère immédiat de la réponse des outils de diagnostic rapide sur site et la capacité de la géostatistique à intégrer des données diverses (directes et indirectes) dans les estimations et la quantification des incertitudes ?

Pour aborder convenablement le couplage des deux outils, le programme de recherche a été divisé en deux parties.

La première partie, dénommée OUTRAGE, a porté sur l'étude des conditions d'utilisation d'outils de diagnostic rapide sur site, sur le mode de comparaison et les corrélations entre les deux types d'analyse de laboratoire et sur site, et sur les approches géostatistiques possibles pour ce type de couplage. OUTRAGE a débuté en 2005 et s'est terminée en décembre 2006. Cette première partie a été réalisée en partenariat par FSS International r&d et le CNRSSP, et a fait l'objet d'un rapport final de synthèse des résultats¹.

La deuxième partie, dénommée REPERAGE, vise à concevoir une démarche couplant mesures sur site et traitement géostatistique pour conduire des campagnes de reconnaissance de façon rationnelle par rapport à des scénarios et des objectifs classiques de gestion de sites pollués. REPERAGE a démarré fin décembre 2006 pour une durée de 26 mois. Cette deuxième partie est réalisée par FSS International r&d en partenariat avec l'INERIS et l'ISSeP qui interviennent dans le cadre du Groupement d'Intérêt Scientifique 3SP.

Le projet REPERAGE soulève plusieurs défis.

- Pour être réaliste, la méthodologie de couplage doit tenir compte des contraintes de coûts, de délais et de terrain (ex. inaccessibilité de certaines zones à l'échantillonnage) d'une étude classique de site pollué. Pour que campagne de terrain et modélisation puissent s'enchaîner sans délai d'attente ou perte de temps pour l'opérateur de terrain ou

¹ Rapport CNRSSP/07/001 – FSSADEME2007001

le modélisateur, les données de terrain doivent pouvoir être intégrées facilement dans le modèle, et leur traitement géostatistique doit pouvoir être réalisé en un temps très court (quelques heures d'un jour de campagne à l'autre). L'opérateur de terrain quant à lui doit pouvoir disposer de moyens techniques et de documents qui lui facilitent le positionnement des points d'échantillonnage à chaque jour de campagne. Le développement de la démarche a donc consisté, non seulement à développer des programmes spécifiques de traitement géostatistique rapide des données, mais aussi à organiser la relation entre terrain et bureau pour un échange rapide et efficace des informations entre équipes de terrain et de bureau (géostatisticiens). Outre l'utilisation d'outils de mesure sur site (ex. Spectromètre Portable de Fluorescence X ou FPXRF), la relation entre terrain et bureau a été rendue possible grâce à l'utilisation de techniques de positionnement géographique (ex. Digital Global Positioning System ou DGPS) et des capacités informatiques actuelles (ordinateur portable, courriels, internet).

- La démarche doit pouvoir s'appliquer à une grande variété de situations de pollution de sols (ex. 2D ou 3D, mono ou multi-polluants) et de contextes réglementaires (ex. seuils de teneurs ou seuils de risques, seuils variant avec la profondeur ou la zone géographique). Les programmes de traitement géostatistique ont donc été conçus pour être le plus général possible.
- La démarche doit pouvoir être testée et validée avant mise en application en temps réel. Pour cela, un cas d'étude a été choisi pour constituer un état de référence d'une pollution des sols, c'est-à-dire une cartographie fine des teneurs en polluant qui puisse être considérée comme une « image » réaliste de l'état de pollution du site étudié. Cet état de référence a été utilisé pour mettre en œuvre virtuellement la démarche de couplage : des données ont été collectées fictivement sur la carte de référence selon un positionnement et un phasage dicté par la démarche de couplage (simulation d'une campagne de reconnaissance). La modélisation obtenue en fin de reconnaissance fictive a ensuite été confrontée à la carte de référence exhaustive comme moyen de vérification des performances de la démarche.
- Pour démontrer sa faisabilité et vérifier sa validité dans un autre contexte de pollution (méthode généralisable pouvant s'appliquer à des sites aux caractéristiques variées), la démarche se devait d'être mise en pratique sur un site réel dans un cadre non expérimental (dans des délais et des budgets typiques de diagnostics réels). Un deuxième cas d'étude a donc été choisi pour appliquer la démarche en temps réel.
- La démarche doit pouvoir être intégrée dans les méthodologies nationales et régionales de gestion des sites pollués des différents acteurs du projet. Une réflexion a donc été menée pour dégager des recommandations quant à l'utilisation de la démarche de couplage dans les cadres réglementaires français et wallon en vigueur.

1.3 Les différentes étapes de travail

Chronologiquement, le projet REPERAGE s'est déroulé en cinq grandes étapes.

a. Constitution d'un cas de pollution de sols de référence correspondant aux environs de MétalEurop (mai à novembre 2007)

Pour étudier les performances de l'approche de couplage sur un cas de pollution de sols de référence, il a été choisi les environs du site de l'ancienne usine MétalEurop, en région Nord de la France, sur lesquels l'ADEME développe des missions de recherche et de développement. Ce secteur est également celui sur lequel les essais de terrain du projet OUTRAGE ont été menés. Le site MétalEurop était un ancien centre de production de plomb, de zinc et d'alliages et sous-produits de ces métaux. Une campagne de terrain conduite en mai et juin 2007 a permis de collecter un ensemble de mesures sur site (FPXRF) et d'analyses de laboratoire (ICP) sur les sols de surface (0-20 cm). Ces données ont été utilisées pour générer par simulation conditionnelle un état réaliste de la répartition spatiale de la pollution au plomb sur la zone des plus fortes teneurs (> 200 mg/kg ms). Deux cartes de teneurs ont été simulées, l'une correspondant à des teneurs en plomb telles qu'analysées en laboratoire, l'autre à des teneurs en plomb telles que mesurées sur site par FPXRF. Ces deux cartes reproduisent les corrélations observées sur les données entre teneurs de laboratoire et teneurs FPXRF. Avec leur résolution très fine (32761 noeuds situés sur une grille 2D régulière de 25m x 25m), elles constituent un état de référence de la pollution au plomb sur les 20 premiers cm de sols.

b. Conception d'une démarche de couplage mesures sur site et analyse géostatistique (décembre à mars 2008)

Une approche générale de reconnaissance associant mesures sur site et exploitation géostatistique des données a été élaborée. Un code spécifique a été développé en langage et dans l'environnement Matlab (The Mathworks™) pour assurer sur une même plateforme informatique un traitement géostatistique rapide des données à chaque étape de reconnaissance et proposer un positionnement des données à collecter à l'étape suivante. Les modules (scripts) destinés à la classification des sols par rapport à des seuils réglementaires se fondent sur une démarche géostatistique éprouvée par FSS International r&d sur de nombreux cas d'étude réels. Les autres modules, et notamment ceux destinés au positionnement des données d'une phase d'échantillonnage à l'autre, sont totalement originaux. Le code comprend également des modules pour formater et importer rapidement des données acquises sur le terrain (DGPS, FPXRF) et pour fournir automatiquement à l'équipe de terrain des tableaux et des cartes de positionnement des données à acquérir lors de la phase de reconnaissance suivante.

c. Test et validation de la démarche de couplage (mars à mai 2008)

Le cas de référence de la pollution au plomb des environs de MétalEurop a été utilisé pour mettre en œuvre virtuellement (simuler sur ordinateur) la démarche de couplage et pour étudier sa validité. Des données ont ainsi été collectées virtuellement sur l'état de référence (grille de simulation) selon un positionnement, un phasage et des types de mesures (laboratoire ou FPXRF) dictés par l'approche proposée. La modélisation géostatistique obtenue en fin de reconnaissance (virtuelle) a ensuite été confrontée à la cartographie exhaustive de référence des teneurs en plomb (équivalents laboratoire).

d. Mise en pratique de la démarche sur le site Henricot II (juin et juillet 2008)

La démarche de reconnaissance établie et validée aux étapes précédentes a été mise en pratique en temps réel sur un site pollué wallon, Henricot II, situé sur la commune de Court-Saint-Etienne. Il s'agit d'un ancien centre sidérurgique d'une surface de 5.6 ha dont le sous-sol comporte des remblais de déchets industriels et de construction sur une épaisseur de l'ordre de 2 m. Une pollution organique et inorganique a été mise en évidence lors de différentes campagnes d'investigation. La démarche de couplage FPXRF et géostatistique a été mise en œuvre sur une partie du site pour une caractérisation de la pollution métallique. Elle a consisté en un jour de campagne préliminaire où des échantillons de sols ont été collectés pour une double analyse, sur site par FPXRF et au laboratoire, puis en 5 jours de campagne itérative basée uniquement sur des mesures sur site par FPXRF.

e. *Intégration de la démarche dans les méthodologies nationale et régionale (septembre-novembre 2008)*

Les conditions d'intégration de la démarche de reconnaissance proposée dans le cadre des méthodologies nationale ou régionale en France et en région wallonne ont été évaluées par les partenaires institutionnels du projet (INERIS et ISSeP).

1.4 Objet et contenu du rapport final

Le présent rapport constitue le rapport final du projet REPERAGE. Après une présentation de la démarche de couplage de mesures sur site et de traitement géostatistique dans le chapitre 2, les tests et la validation de la démarche sur un cas de référence tiré de la pollution au plomb des environs de MétalEurop sont exposés au chapitre 3. Les conditions et les résultats de la mise en œuvre en temps réel de la démarche sur le site Henricot II sont détaillés dans le chapitre 4, puis les conditions d'intégration de la démarche dans les méthodologies française et wallonne de gestion des sites pollués sont discutées au chapitre 5. Les coûts de réalisation des campagnes de reconnaissance sur les sites d'étude sont détaillés au chapitre 6. Le chapitre 7 est dédié aux conclusions et perspectives.

Un CD-Rom associé au rapport contient le fichier pdf du rapport, les fichiers xls des données collectées, ainsi que tous les films, tableaux et figures issus de l'application de la démarche aux deux cas d'étude MétalEurop et Henricot. Son contenu détaillé est listé en **Annexe 1**.

2 PRESENTATION DE LA DEMARCHE DE COUPLAGE MESURES SUR SITE ET ANALYSE GEOSTATISTIQUE

2.1 Principes généraux de la démarche (Figure 1)

La démarche de couplage de mesures sur site et analyse géostatistique s'inscrit dans le contexte d'étude suivant. La caractérisation d'un site pollué doit être menée pour classer les sols² par rapport à des seuils réglementaires dans les catégories « pollué » ou « sain ». Les seuils peuvent être des seuils de teneurs (teneurs limites en polluant au-delà desquelles les sols doivent être traités, mis en décharge ou valorisés) ou des seuils de risque (indices de risque ou excès de risque individuel, limites au-delà desquelles les sols sont considérés comme susceptibles de nuire à l'homme et son environnement). La collecte de données a pour objectif de réduire l'incertitude sur la classification des sols, dans certaines contraintes financières (budget limité), temporelles (délais de réalisation du diagnostic) et géographiques (zones inaccessibles à l'échantillonnage en présence de bâtiments, infrastructures, etc.).

La démarche proposée consiste donc à collecter un ensemble de données en plusieurs jours, les données étant utilisées d'un jour à l'autre pour mettre à jour la classification des sols et proposer un positionnement des nouvelles données dans les zones classées comme trop incertaines.

Les seuils réglementaires doivent être comparés à des analyses de laboratoire, alors que les données acquises en cours de campagne de reconnaissance, et qui doivent servir à améliorer la classification des sols, sont des mesures sur site qui présentent en général un écart avec les analyses de laboratoire. La démarche comporte donc une première étape, dite préliminaire, destinée à établir une loi de régression entre mesures sur site et analyses de laboratoire pour chacun des polluants d'intérêt pour le site d'étude (loi spécifique). Pour cela, un ensemble de mesures sur site et d'analyses de laboratoire sont effectuées sur les mêmes échantillons en veillant à disposer de données représentatives de toute la gamme de teneurs en polluants présente sur le site. Cette étape préliminaire vise aussi à modéliser la variabilité spatiale des polluants d'intérêt (modèles de co-variogrammes entre mesures sur site et analyses de laboratoire). Les modèles statistiques de corrélation ainsi obtenus (lois de régression et co-variogrammes) sont ceux à utiliser ensuite pour simuler la pollution par une méthode géostatistique et pour mettre à jour la classification des sols au cours de la campagne de reconnaissance en prenant en compte l'ensemble des données disponibles, c'est-à-dire les analyses de laboratoire effectuées en phase préliminaire et les mesures sur site en nombre croissant.

La rationalisation de la reconnaissance d'un site consiste, en théorie, à optimiser le phasage, le nombre et la localisation des points d'échantillonnage et de mesure. En pratique, il nous paraît cependant important que la rationalisation tienne compte des contraintes opérationnelles de terrain. Ainsi, la répartition spatiale de la pollution, la nature du sous-sol et les moyens mis en œuvre pour prélever les échantillons et réaliser les mesures sur site conditionnent le nombre maximal de mesures

² Par mailles ou blocs de dépollution.

qu'il est possible de réaliser à chaque phase (journée) de reconnaissance. Même si cela est théoriquement envisageable, le nombre de données à collecter à chaque phase ne peut donc en pratique pas être optimisé, mais seulement accepté par le modélisateur.

La rationalisation de la reconnaissance proposée consiste donc à positionner les données de la phase de reconnaissance $j+1$ dans la(les) zone(s) du modèle classée(s) comme « incertaine(s) »³ en fin de phase j . Dans le cas d'une modélisation 2D, les points sont placés selon un schéma régulier⁴ dans la zone incertaine, adapté à sa géométrie qui peut être complexe. Dans le cas d'une modélisation 3D, il est de la même façon possible de répartir au mieux un nombre fixé de données dans la zone incertaine, mais au risque de positionner un seul échantillon par verticale. Là encore, le terrain impose ses contraintes : on ne peut en effet pas imaginer en pratique réaliser un forage pour prélever un seul échantillon ! Pour des raisons économiques évidentes, dans le cas de pollutions 3D, il a donc été décidé de fixer a priori le nombre et la position des échantillons collectés dans chaque sondage et de développer un algorithme qui optimise uniquement la position des sondages dans la zone incertaine. Sur la verticale, les échantillons peuvent être positionnés systématiquement ou bien en fonction de la lithologie ou de tout autre critère.

Compte-tenu des ressources informatiques et des capacités des outils de terrain actuels, il a été jugé raisonnable de pouvoir mettre à jour la modélisation géostatistique à la fin de chaque journée de terrain. Nous considérons donc qu'une étape de reconnaissance itérative correspond à un jour de terrain. Le traitement géostatistique des données est assuré en quelques heures, de sorte que la position des données du jour suivant puisse être envoyée par courriel à l'équipe de terrain le soir même ou le lendemain matin avant le démarrage des travaux. Un traitement des données sur le terrain au fur et à mesure de leur acquisition (mise à jour du modèle après chaque nouvelle donnée) nécessiterait une automatisation plus poussée de la modélisation géostatistique et une présence du géostatisticien sur le site qui n'ont pas été envisagées pour le moment. Les positions optimales de n données supplémentaires étant différentes et généralement supérieures aux n positions que donnerait une optimisation à l'unité de l'emplacement des données, un tel traitement ne présente que très peu d'intérêt.

Un budget et un délai limitent en général l'ampleur de la reconnaissance des sites. Le module de positionnement des nouvelles données comporte donc des compteurs qui permettent d'alerter le géostatisticien en cas de dépassement des coûts et des délais imposés, et ajuste automatiquement le nombre de données qu'il est encore possible de collecter. De même, des zones sont parfois inaccessibles à l'échantillonnage, c'est pourquoi le code développé pour les besoins de la démarche permet de positionner les nouveaux points de sondages en dehors de telles zones.

³ Zone avec un risque trop élevé de classer les mailles dans les classes « contaminées » ou « saines ».

⁴ Avec dans certains cas quelques resserrements locaux pour capter la structure spatiale à petite distance.

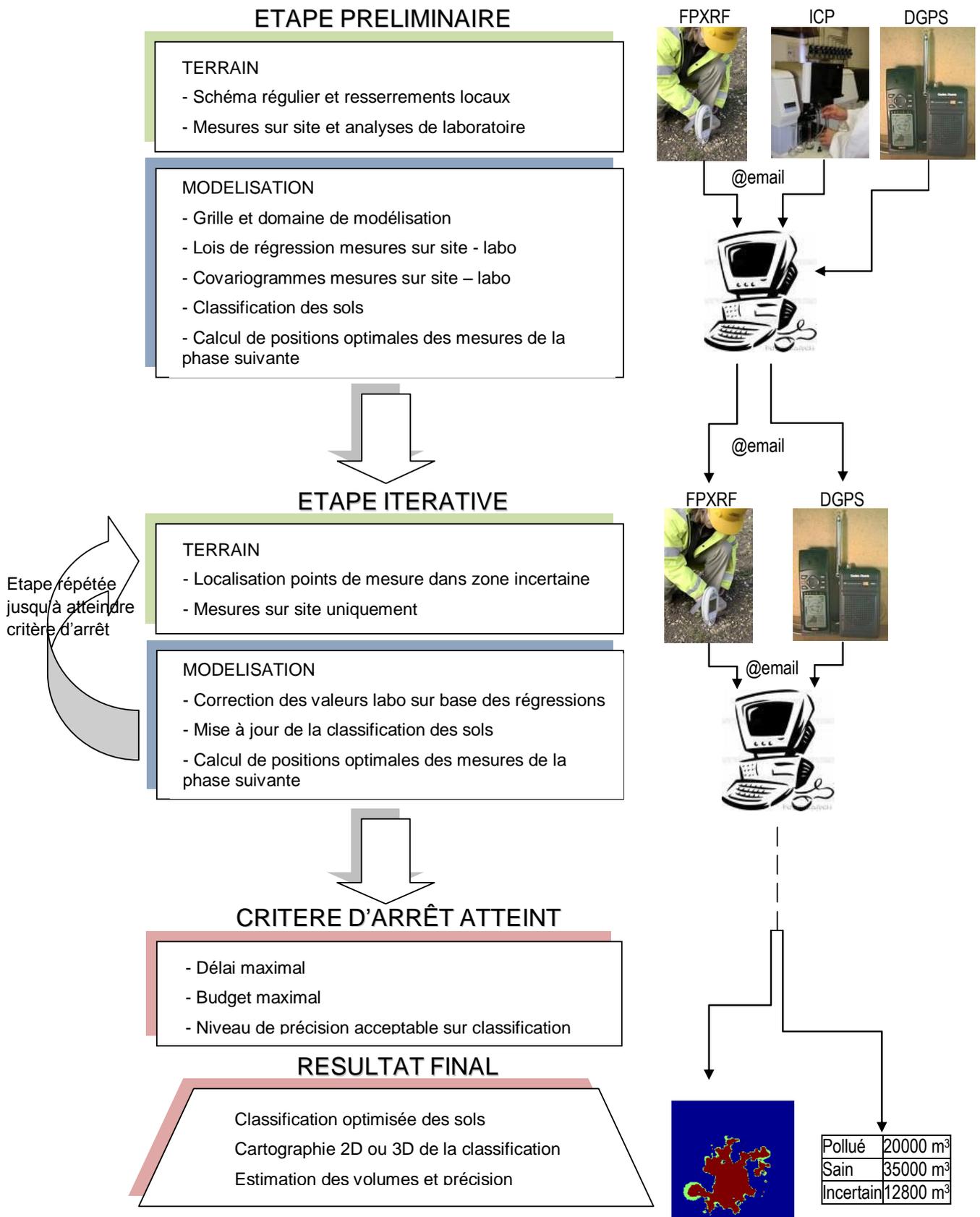


Figure 1 Schéma de la démarche de couplage de mesures sur site et d'un traitement géostatistique

2.2 Présentation détaillée des étapes de la démarche

La démarche consiste en une étape préliminaire (§ 2.2) et plusieurs étapes de reconnaissance itératives aboutissant à la mise à jour du modèle de classification des sols (§ 2.2.2) au fur et à mesure de la reconnaissance jusqu'à atteindre le(s) critère(s) d'arrêt (§ 2.2.3). Les calculs et modélisations réalisés à chaque étape sont détaillés dans les lignes suivantes.

2.2.1 Etape préliminaire de reconnaissance (étape 0)

L'étape préliminaire est une phase de reconnaissance préalable qui n'est pas soumise aux mêmes contraintes temporelles que les étapes suivantes et qui a plusieurs finalités.

- Etablir les lois de régression entre mesures sur site et analyses de laboratoire des polluants d'intérêt spécifiques au cas d'étude traité.
- Etablir un premier modèle dit de corégionalisation (variogrammes directs et croisés, encore appelés covariogrammes) des mesures sur site et des analyses de laboratoire pour tous les polluants d'intérêt.
- Simuler par une approche géostatistique adaptée la répartition spatiale des teneurs en polluants dans les sols.
- En déduire un premier modèle de classification des sols « sains », « pollués » et « incertains » dans le cadre réglementaire requis.
- Proposer un positionnement des points d'échantillonnage et de mesure pour la première étape de reconnaissance itérative dans la zone de sols classés « incertains ».

Les données sont collectées généralement selon un schéma d'échantillonnage approximativement régulier avec quelques resserrements locaux. Ce schéma doit être guidé par les informations documentaires (ex. historique des activités industrielles, contexte géologique) ou éventuellement de terrain disponibles sur le cas d'étude. Il doit aussi permettre dans la mesure du possible d'appréhender la continuité spatiale du phénomène de pollution.

Sur le terrain, les échantillons sont dupliqués selon un protocole opératoire qui garantisse la comparabilité des doublons.

Le traitement des données est soumis aux délais d'obtention des résultats d'analyse. L'analyse géostatistique doit à ce stade être approfondie pour garantir le bon déroulement de la reconnaissance itérative à suivre et la validité de la modélisation finale. Les délais imposés à la reconnaissance itérative ne permettent en effet plus ce type de réflexion.

Dans le détail, le traitement des données en phase préliminaire se décompose de la manière décrite ci-dessous. Le nom du module (script Matlab) dédié à chaque tâche est indiqué dans le titre du paragraphe entre parenthèses.

a. Préparation des données et construction du domaine de modélisation (Nomdusite_etape_0_prep.m)

- Les données FPXRF, DGPS et de laboratoire sont importées, formatées et sauvegardées.

- Si nécessaire, la moyenne de mesures répétées sur un même échantillon ou bien la moyenne de mesures prises en différents points d'un échantillon composite ou d'une zone donnée du site sont moyennées.
- Les seuils réglementaires sont importés et sauvegardés.
- Les limites du domaine d'étude, les limites des zones interdites à l'échantillonnage et les limites lithologiques sont importées et sauvegardées.
- Une grille fine de calcul (2D ou 3D) est construite dans les limites du domaine d'étude.
- Un fonds de plan est constitué et sauvegardé. Il peut s'agir d'un ensemble d'objets (lignes, polygones) représentant des limites de site et de zones, limites de bâtiments, cours d'eau ou voies de circulation. Il peut également être constitué d'une image (ex. format jpg), d'un plan ou carte topographique, ou bien encore d'une photographie.
- Les données collectées en phase préliminaire sont visualisées et sauvegardées.

b. Analyse exploratoire des données (Nomdusite_etape_0_explo.m)

- L'analyse exploratoire des données peut être menée grâce à un ensemble d'outils de calcul statistique et de représentations graphiques.
- Les distributions des teneurs en polluants étant en général dissymétriques, les données de teneur de chaque polluant subissent une transformation dite gaussienne pour que leur distribution suive une loi normale nécessaire à la méthode de simulation géostatistique utilisée (simulation séquentielle gaussienne). Dans la suite, nous noterons données NST (pour *normal score transform*) les données transformées.
- Après transformation des données, le nuage de corrélation entre mesures sur site et analyses de laboratoire est tracé et utilisé pour ajuster une loi de régression entre les deux types de mesures, et ce pour chaque polluant d'intérêt.

c. Analyse variographique des données (Nomdusite_etape_0_vario.m)

- Les variogrammes et covariogrammes entre polluants et entre mesures sur site et analyses de laboratoire sont calculés (sur données NST en général).
- Des modèles de variogrammes et covariogrammes y sont ajustés pour les besoins de la modélisation géostatistique.

d. CoSimulations Séquentielles Gaussiennes (Nomdusite_etape_0_simu.m)

- Dans les modèles de covariogrammes retenus, un ensemble de simulations des teneurs en polluants transformées (NST) sont générées sur la grille de calculn (2D ou 3D) recouvrant le domaine d'étude. Ces simulations sont conditionnées aux données, c'est-à-dire que chacune d'entre elles retrouve la valeur de la donnée⁵ à son point de collecte. Les simulations sont réalisées à partir de Matlab en appelant le code gcosim3d (three dimensional multiple variable

⁵ Valeur NST transformée.

multiGaussian conditional sequential simulation program, Gomez-Hernandez, J. and Srivastava, R., 1990).

- La transformation inverse des valeurs NST simulées permet de retrouver des teneurs en polluants simulées conformes à la distribution de probabilité (histogramme) observée sur les données mesurées.

e. Post-traitement des simulations (Nomdusite_etape_0_postraitement.m)

- A partir des simulations et des seuils réglementaires, la probabilité globale (c'est-à-dire pour tous les polluants d'intérêt) de dépassement des seuils est calculée pour chacune des mailles de la grille de calcul ou d'une grille plus grossière de dépollution. Une première classification des mailles peut alors être obtenue en classant comme « polluées » les mailles de probabilité supérieure ou égale à 50%, et comme « saines » les mailles de probabilité inférieure. Subsiste cependant dans ces deux grandes classes un risque d'erreur de classification non négligeable, notamment pour les mailles dont la probabilité globale de dépassement de seuils est proche de 50%.
- Les simulations sont alors à nouveau utilisées pour calculer, dans chaque catégorie « polluée » ou « saine », le risque d'erreur de classification. En fixant une valeur maximale acceptable de risque d'erreur de classification⁶, il est alors possible d'affiner la classification de départ en « pollué avec bon niveau de confiance », « pollué avec risque d'erreur de classification inacceptable », « sain avec bon niveau de confiance », « sain avec risque d'erreur de classification inacceptable ».
- Les mailles classées comme « polluées avec risque d'erreur de classification inacceptable » ou comme « saines avec risque d'erreur de classification inacceptable » constituent la zone dite « incertaine » dans laquelle sont positionnés les points d'échantillonnage et de mesure de la phase de reconnaissance suivante.
- La classification des sols est représentée sous forme d'une ou de plusieurs cartes, chacune d'entre elles correspondant à une couche de sols à une profondeur particulière.
- Les volumes des sols classés comme « pollués », « sains » et « incertains » sont calculés.

f. Positionnement des sondages et des échantillons de l'étape de reconnaissance suivante (Nomdusite_etape_0_echage_zoneincertaine.m (2D), Nomdusite_etape_0_echage_zoneincertaine_sondages.m (3D))

- Dans le cas d'une modélisation 3D, les sondages prévus pour la phase de reconnaissance suivante sont positionnés selon un schéma régulier dans la zone incertaine. L'anisotropie des mailles reflète l'anisotropie qui a pu être observée sur les modèles de variogrammes ajustés aux données. Leur positionnement est guidé par une grille grossière superposée à la grille de calcul, dont la dimension des mailles est fonction du nombre de sondages qu'il est prévu de réaliser le jour suivant, ainsi que des dimensions de la zone incertaine. Il en est déduit un

⁶ Risque d'erreur admissible qui peut se calculer selon certains critères financiers ou autres et doit être validé entre les différents acteurs du projet.

positionnement des points d'échantillonnage le long de tous les sondages, le nombre et l'espacement de ces points par sondage étant fixés au préalable.

- Dans le cas 2D, ce sont les points d'échantillonnage qui sont positionnés directement selon une grille grossière dont les dimensions des mailles dépendent du nombre de données à collecter le jour suivant et de l'ampleur de la zone incertaine.
- Un tableau Excel⁷ est constitué et sauvegardé avec les identifiants des points d'échantillonnage et de mesure et leurs coordonnées X, Y et Z dans le référentiel adopté. De même, la carte de positionnement des sondages et/ou des points de mesure sur fonds de plan est constituée et sauvegardée sous un format png pour les besoins de l'équipe de terrain.

2.2.2 Etapes de reconnaissance itérative

L'étape préliminaire de reconnaissance est suivie d'une série d'étapes de reconnaissance et de modélisation itératives au cours desquelles :

- Des mesures FPXRF sont réalisées à l'étape j dans la zone « incertaine » selon le schéma dicté par l'algorithme de positionnement des points d'échantillonnage appliqué aux données disponibles à la fin de l'étape j-1.
- Ces mesures viennent rejoindre les données déjà collectées et sont utilisées pour mettre à jour le modèle géostatistique de classification des sols.
- Un positionnement des points de mesure de l'étape de reconnaissance suivante (étape j+1) est proposé dans la zone « incertaine » ainsi redéfinie.

Au cours d'une campagne de reconnaissance, une étape correspond au minimum à une journée d'échantillonnage et de mesures sur sites.

La collecte de données et leur traitement géostatistique est répété jusqu'à ce que le critère d'arrêt de la reconnaissance soit atteint. A la fin de chaque étape de reconnaissance, les données sont transmises et traitées en quelques heures. Les tâches de traitement des données sont très proches de celles décrites au §2.2.1. Les principales différences sont détaillées ci-dessous.

a. Préparation des données et construction du domaine de modélisation (Nomdusite_étape_j_prep.m)

- Les limites du domaine d'étude et des zones interdites à l'échantillonnage ont été établies une fois pour toutes dans l'étape préliminaire, de même que la grille de calcul, le fonds de plan et la définition des seuils réglementaires. La préparation des données consiste donc simplement à importer et à formater les nouvelles données, puis à les intégrer à l'ensemble des données déjà existantes.

⁷ Microsoft™

b. Analyse exploratoire des données (Nomdusite_etape_j_explo.m)

- Des valeurs de laboratoire sont générées à partir des nouvelles mesures FPXRF en utilisant les lois de régression établies pour chaque polluant lors de l'étape de reconnaissance préliminaire. Ces valeurs de laboratoires générées aléatoirement en corrélation avec les mesures sur site servent exclusivement à affiner ou corriger les distributions de probabilité (histogrammes) des teneurs de laboratoire. Sans entrer dans le détail, ce traitement se justifie d'un point de vue statistique pour assurer la cohérence des données de laboratoire par rapport aux mesures sur site qui elles seules s'enrichissent au fur et à mesure.
- Pour tous les polluants d'intérêt, la distribution de teneurs de laboratoire corrigée est utilisée pour opérer à nouveau la transformation gaussienne des données de teneurs de laboratoire vraies et en tirer des données NST corrigées.
- Pour corriger les distributions de teneurs de laboratoire, il est également possible de calculer des poids de dégroupement sur les données de mesure sur site. La multiplication du nombre de mesures peut s'accompagner en effet d'un resserrement local des données qui peut introduire un biais d'échantillonnage et altérer les distributions des variables d'intérêt et leur variographie.

c. Analyse variographique des données (Nomdusite_etape_j_vario.m)

- Les variogrammes et covariogrammes des mesures sur site sont calculés et modélisés. L'évolution des portées des modèles de (co-)variogrammes au fur et à mesure de l'acquisition des mesures de terrain peut ainsi être suivie.
- A partir des résultats précédents, les portées des modèles de covariogrammes entre mesures sur site et analyses de laboratoire, utilisés pour la génération de simulations conditionnelles des teneurs (ou de variables de risque), peuvent si nécessaires être mises à jour.
- Les variogrammes et covariogrammes des analyses de laboratoire ne sont pas corrigés car les valeurs de laboratoire générées à partir de la loi de régression n'ont pas de positions spatiales valides.

d. CoSimulations Séquentielles Gaussiennes (Nomdusite_etape_j_simu.m)

- Des simulations séquentielles gaussiennes sont générées selon la procédure décrite au §2.2.1, item d, en utilisant les modèles de (co-)variogrammes, dont les portées ont éventuellement été mises à jour, et l'ensemble des données disponibles, y compris les mesures sur site nouvellement acquises, comme données de conditionnement.

e. Post-traitement des simulations (Nomdusite_etape_j_postraitement.m)

- Le post-traitement des simulations est réalisé selon la procédure décrite au § 2.2.1, item e.

**f. Positionnement des sondages et des échantillons de l'étape de reconnaissance suivante (Nomdusite_etape_j_echage_zoneincertaine.m (2D),
Nomdusite_etape_j_echage_zoneincertaine_sondages.m (3D))**

- Le positionnement des sondages et points d'échantillonnage est déterminé de manière identique à celle de l'étape préliminaire de reconnaissance.

2.2.3 Critère d'arrêt de la reconnaissance et résultats finaux

L'arrêt de la reconnaissance peut être décidé lorsque n'importe lequel des critères suivants est atteint.

- Coût maximal de la reconnaissance.
- Durée maximale de la reconnaissance.
- Niveau de précision sur la classification des sols jugé acceptable.

Le post-traitement des simulations obtenues avec l'ensemble des données collectées sur la totalité de la campagne de terrain fournit alors les cartes de classification finale des sols par rapport aux seuils réglementaires, ainsi que les volumes de sols associés aux différentes catégories « sains », « pollués » et « incertains ».

2.3 Points théoriques sur les conditions de modélisation

2.3.1 Génération de teneurs de laboratoire

La génération de teneurs de laboratoire à partir des lois de régression établies entre mesures sur site et analyses de laboratoire lors de l'étape préliminaire, permet de faire bénéficier les analyses de laboratoire de l'apport d'information des nouvelles mesures sur site. Les teneurs de laboratoire sont ainsi générées à partir des nouvelles mesures sur site pour mettre à jour leurs distributions (histogrammes) utilisées pour les transformations gaussiennes des (vraies) données de teneurs de laboratoire. En procédant de cette façon, les valeurs simulées se retrouvent à répercuter sur les histogrammes des teneurs de laboratoire des différents polluants les connaissances nouvelles acquises à partir des mesures sur site.

Dans le détail, les couples de mesures sur site et analyses de laboratoire obtenus en phase préliminaire permettent d'établir une loi de régression du type (**Figure 2**, graphe (a)) :

$$L_i = a_i \cdot M_i + b_i + \varepsilon_i \quad (1)$$

où L_i désigne la variable « analyse de laboratoire » pour le polluant i , M_i désigne la variable « mesure sur site » pour ce polluant sur un même échantillon, a_i et b_i sont les paramètres de la loi de régression et ε_i représente un résidu traduisant la dispersion des points (L_i, M_i) autour de la loi de régression moyenne. Le résidu est considéré comme une variable normale centrée d'écart-type σ_{ε_i} . Les mesures sur site de l'étape préliminaire permettent de calculer σ_{ε_i} par classes de valeurs de la mesure sur site, et d'ajuster une fonction polynômiale du type (**Figure 2**, graphe (b)) :

$$\sigma_{\varepsilon_i} = c_i \cdot |L_i - M_i|^{p_i} + d_i \cdot |L_i - M_i| \quad (2)$$

où c_i , d_i et p_i désignent les paramètres de la fonction polynômiale pour le polluant i .

Pour générer des valeurs de laboratoire à partir de nouvelles mesures sur site lors des étapes de reconnaissance itératives, il est procédé comme suit pour chaque nouvelle mesure sur site $m_{i\alpha}$.

- L'écart-type $\sigma_{\varepsilon_{i\alpha}}$ du résidu correspondant à la mesure $m_{i\alpha}$ est calculé à partir de l'équation (2).
- Une valeur de résidu $\varepsilon_{i\alpha}$ est tirée aléatoirement dans la distribution gaussienne centrée et d'écart-type $\sigma_{\varepsilon_{i\alpha}}$.
- La teneur de laboratoire correspondante $l_{i\alpha}$ est déduite de l'équation (1).

Il est à noter que le calage de la fonction polynômiale des résidus peut nécessiter quelques précautions pour borner convenablement les teneurs de laboratoire et éviter des valeurs négatives.

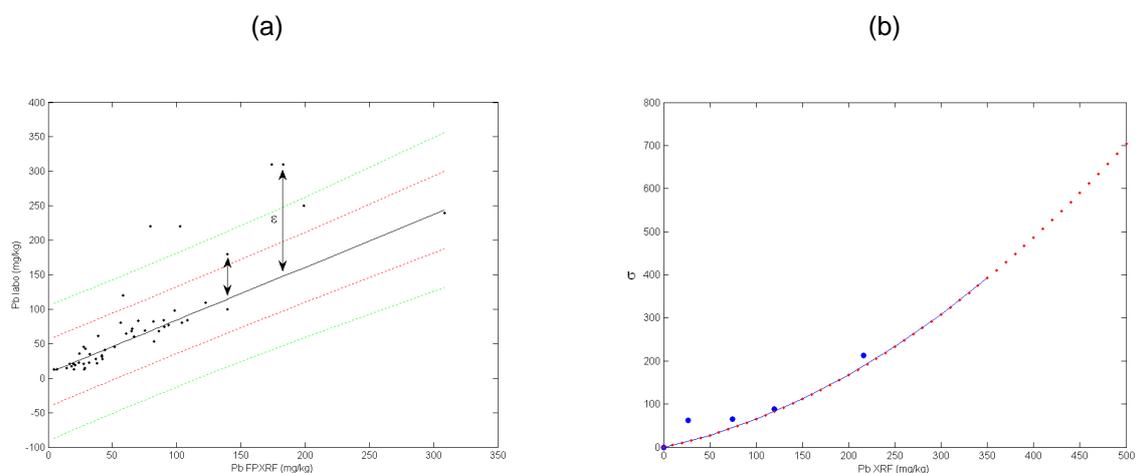


Figure 2 Loi de régression typique, entre analyses de laboratoire et mesures sur site, définie par une droite de régression moyenne (a) et une fonction polynômiale de l'écart-type du résidu autour de la droite de régression (b), écart-type qui tend à augmenter avec la teneur. Les paramètres de la loi peuvent être ajustés à partir des données collectées lors de l'étape préliminaire

2.3.2 Classification des sols

La démarche de classification des sols, par rapport à des seuils appliqués à un ensemble de simulations géostatistiques des teneurs en polluants, a été décrite à plusieurs reprises par FSS International r&d dans diverses publications. On se référera pour plus de détail au manuel GeoSiPol (Chilès et al., 2005) et au manuscrit de D'Or et al. « An integrated geostatistical approach for contaminated site and soil characterization » accepté pour publication dans un numéro spécial de Mathematical Geosciences.

2.3.3 Positionnement des points de mesure pour l'étape de reconnaissance suivante

Pour un évaluation 2D (en surface) de la pollution des sols, le positionnement des points de mesure dans la zone incertaine s'effectue de la manière suivante (**Figure 3**).

- Un nombre maximal de points de mesure N_i est fixé, ou plus justement dicté par les moyens de mesure mis en œuvre sur le site, pour la phase de reconnaissance suivante i .

- L'enveloppe 3D incluant la totalité des mailles de la grille de calcul classées comme « incertaines » est déterminée.
- Une grille régulière dite d'échantillonnage est superposée à la grille de classification des sols. Cette grille, généralement plus grossière que celle de classification, est constituée de mailles dont les dimensions dans les trois directions principales dépendent (1) du nombre de points de mesures supplémentaires prévu dans la phase de reconnaissance suivante, (2) de l'étendue de la zone incertaine et (3) de l'anisotropie observée sur les modèles de co-variogrammes (anisotropie dans la continuité spatiale des teneurs en polluants).
- Le nombre de mailles « incertaines » de la grille de classification incluses dans chacune des mailles de la grille d'échantillonnage est calculé.
- Les N_i mailles de la grille d'échantillonnage incluant les plus grands nombres de mailles « incertaines » de la grille de classification sont sélectionnées, et les centres de gravité des mailles incertaines à l'intérieur des mailles d'échantillonnage sélectionnées sont calculés. Ces centres de gravité déterminent alors les points de mesure de la phase de reconnaissance suivante.

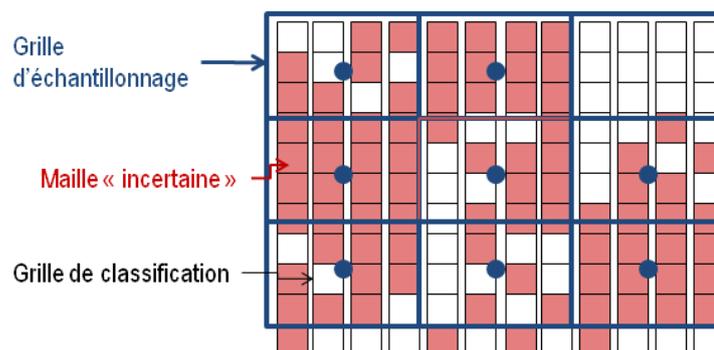


Figure 3 Schéma du principe de positionnement des nouveaux points de mesure dans la zone « incertaine ». Remarque : la grille d'échantillonnage et la grille de classification peuvent parfois avoir des directions différentes.

Pour une évaluation 3D de la pollution des sols, où il est nécessaire de « rentabiliser » les sondages en prélevant et analysant un minimum d'échantillons sur chacun d'entre eux, l'algorithme est légèrement différent.

- Le nombre de sondages N_i et le nombre de points d'échantillonnage et de mesure sont dictés par les moyens de creusement et de mesure mis en œuvre sur le site pour la phase de reconnaissance suivante i . Les échantillons analysés sur site sont distribués équitablement sur les sondages et répartis le long de chacun selon un schéma prédéfini (généralement régulier).
- L'extension latérale (dans le plan) de la zone incertaine, correspondant aux mailles classées comme « incertaines » dans la grille de classification, est déterminée.

- Une grille d'échantillonnage régulière et généralement plus grossière est superposée à la grille de classification. L'orientation de cette grille et les dimensions de ses mailles dépendent (1) du nombre de points de sondages supplémentaires prévu pour la phase de reconnaissance suivante, (2) de l'étendue de la zone incertaine et (3) de l'anisotropie dans le plan horizontal des modèles de co-variogrammes (anisotropie dans la continuité spatiale des teneurs). La dimension verticale des mailles de la grille d'échantillonnage correspond à l'espacement systématique fixé pour les points de mesure dans chaque sondage.
- Le nombre de mailles « incertaines » de la grille de classification incluses dans chaque maille de la grille d'échantillonnage est calculé, puis ces nombres sont additionnés sur la verticale pour capitaliser l'incertitude sur toute la hauteur de sols étudiée (incertitudes spatiales évaluées sur chaque pile de mailles de la grille d'échantillonnage).
- Les N_i piles de mailles de la grille d'échantillonnage incluant les plus grands nombres de mailles « incertaines » de la grille de classification sont sélectionnées. Comme en 2D, les centres de gravité des zones incertaines sont calculés à l'intérieur des N_i piles de mailles les plus incertaines. Ces centres de gravité déterminent les positions des sondages de la phase de reconnaissance suivante, sondages dans lesquels des échantillons de sols sont à prélever selon le schéma prédéfini.

2.4 Le code développé pour la mettre en œuvre

Le code développé pour recevoir et traiter les données et pour calculer les positions des points de mesure d'une étape à l'autre est constitué d'un ensemble de modules (scripts) développés en langage Matlab. Un script est dédié à chacune des étapes de calcul décrites au § 2.2, les noms des scripts étant indiqués entre parenthèses. Le contenu détaillé des scripts se trouve en **Annexe 3**.

Le code a été conçu le plus général possible pour être adaptable à la grande variété de situations rencontrées, à la fois en termes de pollution (répartition spatiale, type de pollution, etc.) et de cadre réglementaire (seuils de teneur ou de risque, seuils fonction de la profondeur ou de l'occupation des sols, etc.). Cela se traduit par de nombreux paramètres de modèle (ex. nombre et nom des polluants, nombre et nom des seuils, caractéristiques de la grille de calcul) et par la présence de modules optionnels (ex. application d'une fonction de transfert aux teneurs pour évaluer des risques sanitaires).

En outre, pour que le traitement géostatistique des données puisse être réalisé en quelques heures, les différentes étapes de calcul ont été automatisées au maximum. Ceci inclut la production et la sauvegarde automatique de résultats sous forme de fichiers Matlab, tableaux Excel (MicrosoftTM) ou de figures au format standard (png). Les étapes qui nécessitent l'expertise du géostatisticien (ex. variographie) restent cependant non automatisées.

2.5 Les conditions pratiques de mise en œuvre

Plusieurs conditions doivent être remplies pour que la démarche soit mise en œuvre correctement.

La campagne de terrain et les conditions de traitement des données doivent être soigneusement préparées pour limiter les imprévus. Ces tâches préparatoires rentrent dans l'étape préliminaire qui n'est pas soumise aux mêmes contraintes de temps que les étapes itératives. Elles incluent l'élaboration d'un protocole opératoire d'échantillonnage et de mesure, la définition des identifiants des prélèvements, le format et le type de fichier à transmettre en fin de journée au modélisateur et en début de journée à l'équipe de terrain, etc.

Les données doivent être collectées, sauvegardées et transmises dans des délais très courts au modélisateur qui n'est pas forcément sur place. Le positionnement des points de mesure s'effectue par DGPS et les mesures par un appareil de terrain portable (ex. FPXRF) permettant l'acquisition, la sauvegarde et la transmission des données. L'équipe de terrain doit avoir la possibilité de transmettre les données par internet : ordinateur portable et connexion internet, téléphone portable, etc.

La communication entre opérateur de terrain et modélisateur est importante. Il est conseillé de munir toute l'équipe de téléphones portables pendant la durée de la reconnaissance (y compris pendant la période de modélisation, qui s'effectue après 18h en général). Malgré tout le soin apporté à la préparation de la campagne, des imprévus peuvent en effet survenir : sondage localisé sur un terrain trop dur pour être foré, erreurs de sauvegarde des positions des sondages rendant impossible le traitement géostatistique, etc.

Le code de calcul permet de proposer un positionnement des points de mesure optimisé, compte-tenu des données disponibles sur l'état de pollution des sols à la fin d'une phase. L'équipe de terrain détient cependant un formidable outil de mesure qui, en quelques minutes, fournit une évaluation de la teneur en polluants. Cette connaissance immédiate de la mesure offre à l'opérateur la possibilité d'adapter la reconnaissance par exemple en prolongeant un sondage, si des teneurs élevées continuent d'être mesurées en profondeur, ou en rajoutant de nouveaux si la pollution atteint des zones initialement prédites comme non polluées. Il serait dommage de ne pas profiter de ce caractère immédiat de l'appareil de mesure en ne se permettant pas d'écart par rapport au plan d'échantillonnage et de mesure issu du code de calcul. Nous suggérons donc de laisser au libre arbitre de l'équipe de terrain une petite partie des mesures qu'il est possible de réaliser en une phase de reconnaissance. Une répartition du type 80% des mesures positionnées par le code de calcul et 20% positionnées au jugé sur le terrain nous semble par exemple satisfaisante.

La démarche proposée permet de rationaliser l'échantillonnage d'un site dont la pollution présente une certaine continuité spatiale. Elle sera inefficace si le phénomène de pollution n'est pas structuré, c'est-à-dire qu'il présente un effet de pépite pur (répartition purement aléatoire des teneurs). Sans continuité spatiale, la problématique devient une estimation globale ou par zones de l'état de pollution des sols et donc de leur classification comme pollués ou sains. Sans corrélations spatiales à exploiter, les méthodes géostatistiques ne présentent alors plus d'avantages particuliers par rapport à des méthodes statistiques plus classiques. En 1989, US EPA (**Annexe 2**) a développé une méthodologie statistique, fondée sur des tests d'hypothèses, qui pourrait être une voie de modélisation et d'optimisation de l'échantillonnage pour ce type de situation.

3 TEST ET VALIDATION DE LA DEMARCHE SUR LES ENVIRONS DU SITE METALEUROP

3.1 Présentation du cas d'étude

Le cas d'étude choisi pour tester et valider la démarche porte sur les environs du site de l'ancienne usine MétalEurop, en région Nord de la France, où l'ADEME conduit des missions d'étude mais aussi de recherche et développement. C'est également sur ce secteur que les essais de terrain du projet OUTRAGE ont été menés. Le site MétalEurop était un ancien centre de production de plomb, de zinc et d'alliages et sous-produits de ces métaux. La présence de plomb et de zinc dans les sols à des teneurs pouvant atteindre plusieurs milliers de mg/kg ms a été vérifiée au travers de nombreux travaux d'étude et de recherche.

Le choix du secteur d'étude s'est porté sur l'ensemble des sols de 0 à 20 cm de profondeur où les teneurs en Pb peuvent dépasser 200 mg/kg ms. Ce choix se fonde sur l'Action Nationale Plomb⁸ qui avait été mise en œuvre avant la définition des nouveaux outils de gestion des sites et sols pollués (voir § 5.1). Dans ce cadre, un guide d'orientation des actions à mener sur un site dont les sols sont potentiellement pollués par le plomb a été rédigé (INERIS, 2004). En milieu résidentiel, les actions à enclencher selon les niveaux de concentration en Pb sont les suivantes.

- **Teneur maximale entre (100 mg/kg ou fonds géochimique naturel) et 200 mg/kg.** Un niveau d'intervention dit 2a est requis. Dans un premier temps, un diagnostic approfondi, tel que défini dans le guide « Gestion des sites pollués » (MATE, BRGM, INERIS, 2000), doit être réalisé pour évaluer plus précisément la pollution de l'environnement. Cette caractérisation doit porter également sur d'autres métaux comme le cadmium, le mercure, l'arsenic et le chrome. Le diagnostic approfondi vise à confirmer ou infirmer la pollution de l'environnement par le plomb et à identifier les autres substances chimiques qu'il faudra étudier si nécessaire dans les phases suivantes. A l'issue de ce diagnostic approfondi et selon ses résultats, il peut être envisagé de poursuivre l'action par une Evaluation Détaillée des Risques (EDR, niveau 2b voire niveau 3).
- **Teneurs comprises entre 200 mg/kg et 550 mg/kg** Un des échantillons prélevés a une teneur comprise entre 200 et 550 mg/kg. Dans ce cas le niveau d'intervention 2b est atteint, la réalisation d'un diagnostic approfondi et d'une EDR doit être envisagée.
- **Teneurs supérieures à 550 mg/kg** Le niveau d'intervention est le niveau 3 qui doit conduire au déclenchement d'un diagnostic approfondi et d'une EDR ainsi qu'à la mise en place immédiate de mesures sanitaires.

Le secteur choisi est un carré d'une superficie d'environ 20 000 ha délimité par un trait noir sur la **Figure 4**. Ce secteur est occupé par des installations industrielles, des terrils, une gare d'eau sur le canal de la Haute Deûle, des axes routiers et ferroviaires et par les agglomérations de Courcelles, Noyelles-Godault et Evin-Malmaison. Le secteur présente également une activité agricole, même si

⁸ Action nationale française.

les cultures tendent progressivement à être remplacées par des friches ou des zones boisées. De nombreuses parcelles agricoles impropres à la culture ont été végétalisées à nouveau pour stabiliser la contamination.

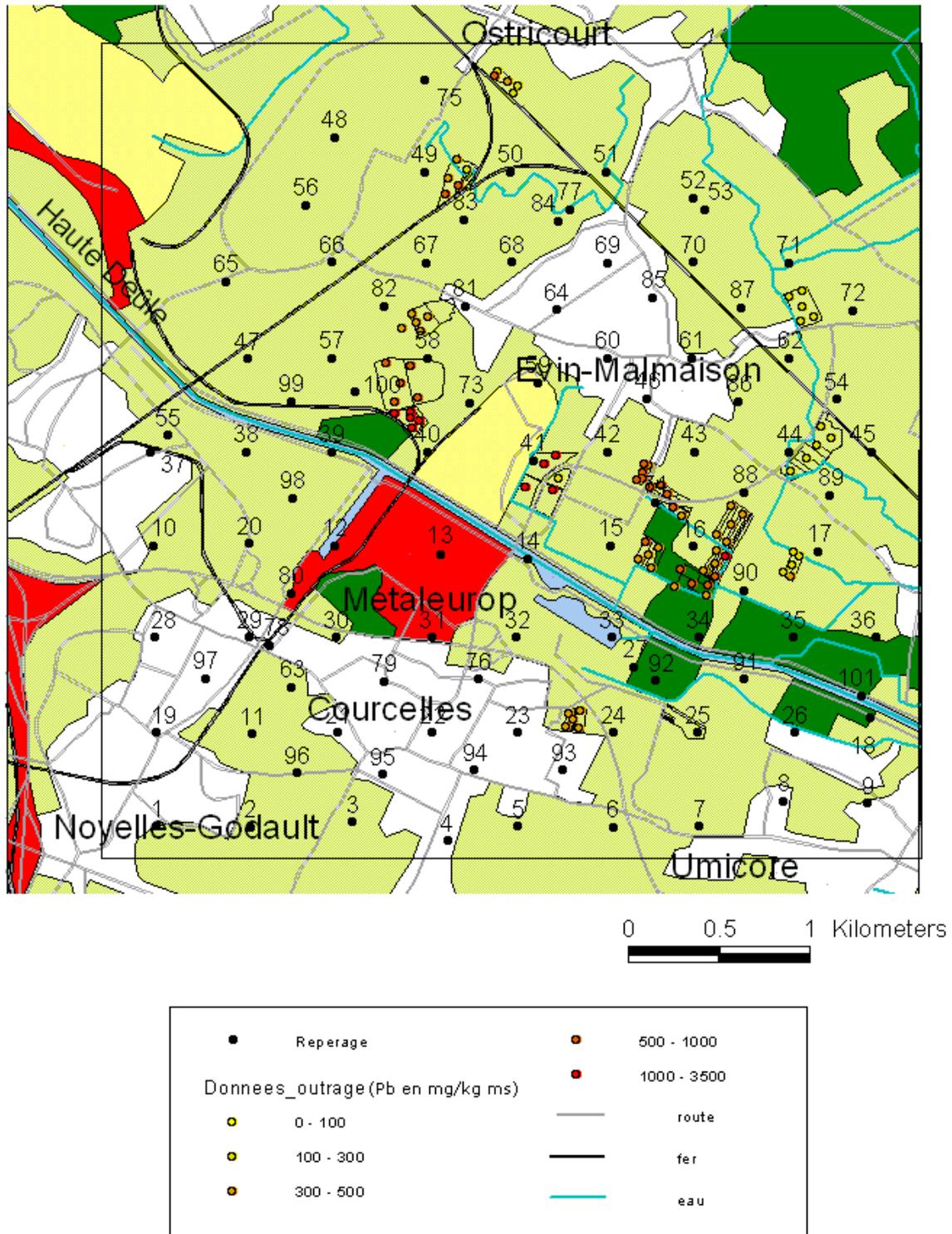


Figure 4 Localisation théorique des points d'échantillonnage et de mesure FPXRF dans le secteur d'étude MétalEurop. Fonds topographique IGN (base de données BD carto®). Lambert II étendu.

3.2 Constitution de l'état de référence

L'état de référence est une cartographie fine des teneurs en plomb dans les sols de surface sur le secteur d'étude. Il est le résultat d'une modélisation géostatistique à partir de données de pollution du sol (mesures FPXRF et analyses de laboratoire) issues du projet REPERAGE mais également du projet OUTRAGE et d'autres projets menés par l'Institut Supérieur d'Agriculture de Douai (ISA).

3.2.1 La campagne de terrain

3.2.1.a Conditions générales

Une campagne de terrain a été menée durant les deux dernières semaines du mois de juin et début juillet 2007 (voir. rapport intermédiaire REPERAGE, octobre 2007). Elle avait été précédée d'un travail préparatoire de positionnement de 101 points d'échantillonnage et de mesure sur le secteur d'étude, de la production de documents cartographiques et de tableaux des coordonnées des points pour l'équipe de terrain, et de prises de contact avec les propriétaires et exploitants des parcelles concernées par la campagne.

Le déroulement de la campagne a été perturbé par des épisodes orageux qui ont rendu impossible la prise de mesures FPXRF à certains moments. Par ailleurs, des points de mesure, qui avaient été localisés initialement sur base documentaire, se sont révélés inaccessibles sur le terrain et ont dû être décalés ou bien abandonnés. En d'autres endroits, les sols étaient trop durs pour être collectés à la pelle à main. Le nombre final de points renseignés (95) est donc légèrement inférieur au nombre de points de mesure prévu initialement (101), ce qui n'a généré en rien la constitution d'un état de référence et la validation de la démarche.

Deux à trois personnes ont assuré la réalisation des divers prélèvements et mesures sur le terrain. Les coordonnées des points de prélèvement et mesure sur site ont été déterminées à l'aide d'un DGPS Géoexplorer 2005 XT TerraSyncPro. Les teneurs en Pb ont été mesurées à l'aide d'un FPXRF de la marque Niton.

La durée des analyses FPXRF a été fixée à 2 minutes. Ce temps d'analyse découle des travaux effectués dans le cadre du projet OUTRAGE sur l'influence de la durée de mesure sur sa précision. Sur les sols du secteur environnant MétalEurop, il a en effet été démontré que cette durée permettait d'atteindre une fidélité de la mesure du Pb, exprimée sous la forme d'un coefficient de variation sur triplicata d'environ 3 %, et que cette fidélité était d'autant plus grande (coefficient de variation d'autant plus petit) que les teneurs étaient élevées⁹. Cette fidélité est jugée satisfaisante puisqu'elle est supérieure à la fidélité classiquement recherchée pour les analyses effectuées en laboratoire.

L'appareil FPXRF a été utilisé en mode de calibration dit standard des paramètres fondamentaux.

3.2.1.b Mesures FPXRF in situ

Les teneurs en Pb ont été évaluées in situ sur les 20 premiers cm de sol en 95 points géographiques de coordonnées (X, Y) différentes, en appliquant le protocole opératoire établi dans le cadre du projet OUTRAGE (voir rapport final OUTRAGE, juillet 2007) :

⁹ Voir rapport final OUTRAGE, p. 71, pour une définition précise de la fidélité d'une mesure.

- retrait du couvert végétal (s'il existe),
- positionnement de l'appareil sur le sol en place tous les 5 cm, de 0 à 20 cm, pour obtenir 5 mesures sur la verticale.

Cette procédure permet d'appréhender la variabilité verticale de la contamination en plomb et de calculer un estimateur de la teneur en plomb moyenne sur les 20 premiers cm de sol¹⁰. Un total de 417 mesures FPXRF a ainsi été réalisé.

3.2.1.c *Echantillonnage du sol et analyse ICP-OES*

Sur les 95 points géographiques, des échantillons de sol représentatifs des premiers 20 cm ont été collectés, conditionnés et transportés pour une analyse par ICP-OES après minéralisation totale (NFX 31-147, acide fluorhydrique, acide perchlorique). Ces analyses ont été réalisées par l'INERIS. Les échantillons ont également fait l'objet d'une mesure de la teneur en eau qui est un paramètre très influant sur la réponse de l'appareil FPXRF.

En outre, une soixantaine de ces points répartis sur l'ensemble du secteur d'étude ont donné lieu également à un échantillonnage spécifique de la couche de terrain la plus superficielle, de 0 à 5 cm de profondeur. Vingt de ces échantillons ont été analysés au laboratoire par ICP-OES pour quantification du Pb. Ces données supplémentaires, non prévues dans le programme de travail initial du projet REPERAGE, étaient destinées à tester la démarche de couplage entre mesures sur site et géostatistique dans le cadre d'une évaluation du risque sanitaire par rapport à la couche de sol la plus superficielle comprise entre 0 et 5 cm. Cette évaluation du risque porte sur l'exposition à la pollution du sol par contact cutané, ingestion ou éventuellement inhalation par envol de poussières (voir §3.2.2.d). La variable modélisée dans ce cas n'est pas la teneur en polluant mais un indice de risque ou un excès de risque individuel.

3.2.1.d *Mesures FPXRF ex situ*

Après homogénéisation manuelle grossière, les 95 échantillons représentatifs de la couche de sol 0-20 cm ont fait l'objet d'une analyse FPXRF. Pour apprécier la fidélité de ces analyses, trois mesures ont été effectuées en des points différents sur chaque échantillon de sol homogénéisé. Un total de 95 mesures FPXRF en triplicata a ainsi été réalisé.

3.2.1.e *Bilan*

Au terme de cette campagne, les données disponibles issues de mesures par FPXRF sont nombreuses et peuvent se répartir en plusieurs groupes qui ont permis de tester la démarche dans différentes conditions.

- Mesures in situ réalisées tous les 5 cm, de 0 à 20 cm. Elles renseignent sur la variation verticale des teneurs en métaux et peuvent être moyennées pour obtenir des valeurs représentatives de la teneur en Pb sur les 20 premiers cm.
- Mesures sur site de la teneur en Pb sur des échantillons homogénéisés représentatifs des 20 premiers cm de sol.

¹⁰ Dans le cadre du projet OUTRAGE, il a été démontré que la moyenne des 5 mesures in situ constitue un estimateur satisfaisant de la teneur moyenne sur les 20 premiers cm de sol.

- Mesures sur site de la teneur en Pb sur des échantillons représentatifs des 5 premiers cm de sol.

A ces données se rajoutent les analyses de laboratoire par ICP-OES, effectuées par l'INERIS sur des échantillons représentatifs des 20 ou 5 premiers cm de sol, et les lois de régression qui ont été établies entre ces analyses et les mesures FPXRF in situ et sur site.

Le positionnement réel des points d'échantillonnage et de mesure est représenté à la **Figure 5**.

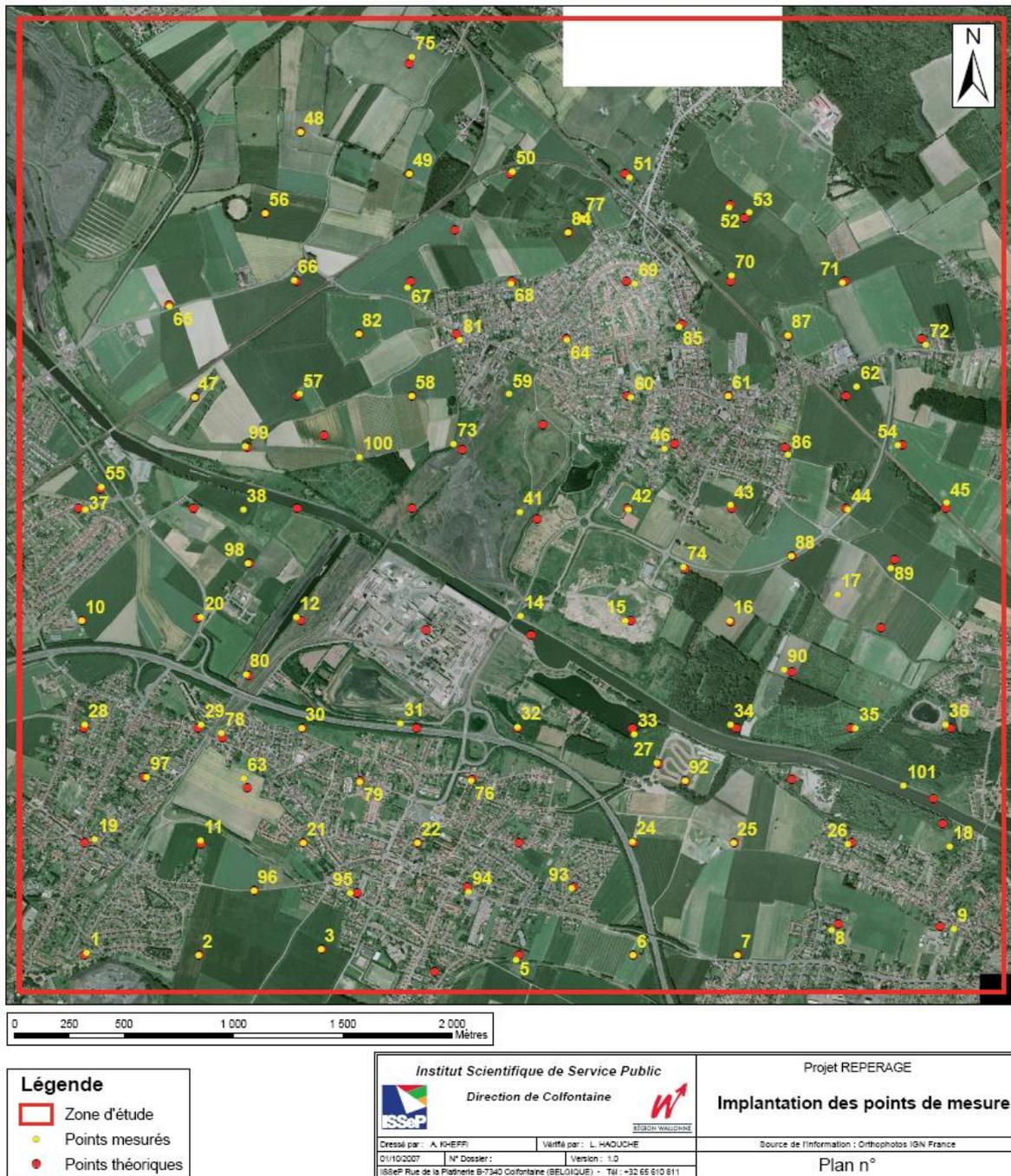


Figure 5 Localisation réelle des points d'échantillonnage et de mesure FPXRF dans le secteur d'étude MétalEurop (ortho-photos IGN).

Un résumé statistique des teneurs en Pb mesurées sur le terrain ou au laboratoire sur les différents échantillons de sol se trouve dans le **Tableau 1**.

Tableau 1 Statistiques élémentaires sur les teneurs en Pb mesurées par FPXRF in situ ou sur site et sur celles analysées en laboratoire pour le secteur d'étude MétalEurop.

	Nb	Moyenne	Ecart-type	Minimum	Maximum	Médiane	Coefficient de variation
	-	mg/kg ms	mg/kg ms	mg/kg ms	mg/kg ms	mg/kg ms	–
FPXRF in situ Mesures prises tous les 5 cm, de 0 à 20 cm	417	599	1604	24	23985	225	2.68
FPXRF sur site Représentatif sols 0-20 cm Moyennes sur triplicata	95	536	996	33	5994	227	1.86
ICP-OES Représentatif sols 0-20 cm	95	552	1201	30	8754	223	2.18
ICP-OES Représentatif sols 0-5 cm	20	1265	2444	106	10683	379	1.93

Les coûts effectifs de réalisation de la campagne de reconnaissance sur le secteur d'étude MétalEurop sont fournis au § 6.1.

3.2.2 Modélisation de l'état de référence

La démonstration de la démarche requiert de disposer des cartes suivantes pour pouvoir simuler des campagnes de reconnaissance par ordinateur avec l'obtention de mesures FPXRF sur site ou d'analyses de laboratoire en n'importe quel point du site.

- Une cartographie fine des teneurs en Pb analysées au laboratoire par une méthode d'analyse reconnue. Cette cartographie se veut être la plus proche possible de l'état de contamination vrai du secteur d'étude afin de pouvoir être utilisée pour évaluer les performances de la démarche de couplage en comparant les modèles géostatistiques à la solution de référence connue. Elle sert bien sûr aussi à fournir les données d'analyses de laboratoire à l'étape préliminaire de la démarche où la loi de régression entre mesures sur site et analyses de laboratoire doit être établie sur des échantillons de sol dont le nombre et la position sont à définir. Cette cartographie est désignée dans la suite du texte par **référence-labo**.
- Une cartographie fine des teneurs en Pb mesurées sur le terrain par FPXRF. Cette cartographie doit être corrélée à la première en reproduisant la corrélation observée sur les couples de données FPXRF et laboratoire. Elle est utilisée pour « simuler » la collecte de mesures FPXRF aux positions déterminées à chaque phase de la démarche de couplage. Cette cartographie est dénommée **référence-FPXRF** dans la suite du texte.

Les conditions de comparaison des mesures FPXRF aux analyses de laboratoire sont variables en fonction du mode opératoire choisi et des caractéristiques du sol et de la contamination. Ainsi, la mesure FPXRF peut être réalisée sur site, sur des échantillons de sol, ou bien in situ sur le sol en place ; dans ce dernier cas, il peut être nécessaire de calculer la moyenne de plusieurs valeurs pour obtenir une valeur représentative du support de l'information (§ 3.2.1.b). Lorsque la teneur en eau dans les sols est trop importante, des corrections de la mesure FPXRF peuvent aussi être nécessaires pour améliorer la corrélation avec les analyses de laboratoire.

Il existe de ce fait plusieurs façons de définir une cartographie des mesures FPXRF en considérant des mesures in situ ou sur site, avec ou sans correction de la teneur en eau. Le choix s'est porté ici sur la constitution d'une **référence-FPXRF in situ**.

La totalité des données disponibles ont été utilisées pour constituer les cartographies de référence, après avoir vérifié que les données issues des différentes campagnes et projets constituaient des populations statistiquement homogènes.

- La **Référence-labo** a été modélisée à partir de 338 analyses de laboratoire¹¹ par ICP-MS ou ICP-OES sur des échantillons composites représentatifs des 20 premiers cm de sol.
- La **Référence-FPXRF in situ** repose sur 758 mesures FPXRF in situ réalisées au cours des projets OUTRAGE et REPERAGE. Les moyennes des 5 mesures prises tous les 5 cm, entre 0 et 20 cm de profondeur, ont été calculées en chacun des 80 points du projet OUTRAGE et des 95 points du projet REPERAGE. Ce sont ainsi 175 valeurs, réparties sur le secteur d'étude, qui ont servi à cartographier les mesures FPXRF.

Les tâches effectuées pour modéliser ces deux solutions de référence ont été les suivantes.

- Analyse exploratoire des données disponibles, avec notamment étude de la corrélation entre mesures sur site et analyses de laboratoire.
- Analyse variographique des données.
- Cosimulation des teneurs en Pb correspondant à des analyses de laboratoire et à des mesures FPXRF.

Les conditions de modélisation des solutions de référence sont détaillées en **Annexe 4** et résumées dans les paragraphes ci-dessous.

3.2.2.a Constitution de la référence-labo

La référence-labo a été obtenue en simulant les teneurs en Pb par la méthode de simulation séquentielle gaussienne (SSG, voir par ex. Chilès, 1999). Cette méthode permet de générer des cartes de teneur réalistes dans le sens où elles reproduisent à la fois les teneurs mesurées (conditionnement du modèle), leur distribution de probabilité (histogramme) et leur corrélation spatiale (variogramme). La distribution des teneurs n'étant pas gaussienne, les données doivent être transformées avant simulation (transformation gaussienne), le modèle de corrélation spatiale à reproduire doit être établi sur les données transformées et les valeurs simulées doivent faire l'objet d'une transformation inverse.

La simulation a subi deux adaptations pour réduire le « bruit » (forte hétérogénéité à petite distance) ainsi que des anomalies visuelles observées initialement.

- L'effet de pépite du modèle de variogramme ajusté au variogramme expérimental des teneurs en Pb a été supprimé. Cet effet de pépite peut traduire soit des erreurs de mesures ou d'échantillonnage, soit une continuité spatiale à petite échelle qui n'a pas pu être mise en évidence à partir des données disponibles. En éliminant (filtrant) cet effet, nous nous sommes

¹¹ 235 analyses ISA, 10 analyses OUTRAGE, 95 analyses REPERAGE.

attachés à modéliser la continuité spatiale de la pollution à grande échelle (influence de l'usine sur plusieurs km en raison de la propagation de poussières par l'atmosphère), et non à sa variabilité locale à très petite échelle (dizaine ou centaine de mètres). Ce effet en effet .

- Deux données collectées dans le cadre de REPERAGE ont été éliminées de l'ensemble des données utilisé pour conditionner le modèle (**Tableau 2**). Il s'agit des points 41 et 59 qui présentent des teneurs en Pb anormalement faibles par rapport aux valeurs mesurées sur des points voisins. Ils correspondent à des matériaux solides qui ont été déplacés de leur lieu d'origine (ex. remblais de démolition, terril de matériaux issus de l'usine) et qui de ce fait ne reflètent pas la répartition générale de la contamination liée à une dissémination aérienne à partir de l'usine MétalEurop.

Tableau 2 Caractéristiques des données écartées pour simuler la référence-labo du cas d'étude MétalEurop.

ID	Pb – labo mg/kg ms	Occupation	(X,Y) m, Lambert II
59	37	Friche près ancien puits de mine – remblais de démolition	(648845, 2605026)
41	111	Haut d'un terril – terre noire – beaucoup de cailloux	(648895, 2604481)

3.2.2.b Constitution de la référence-FPXRF in situ

La référence-FPXRF in situ a été obtenue avec la même méthode de simulation que précédemment (référence-labo) mais en prenant en compte la corrélation spatiale entre teneurs en Pb FPXRF et teneurs en Pb de laboratoire ainsi que la carte de référence-labo simulée précédemment et introduite comme information indirecte (auxiliaire). Le choix de simuler d'abord les teneurs de laboratoire, puis celles issues des mesures PFXRF en les corrélant aux premières, se justifie par le plus grand de données disponibles pour les teneurs de laboratoire. Sans entrer dans le détail, la méthode de simulation utilisée est une méthode dite de cosimulation séquentielle gaussienne colocalisée (voir par ex. Chilès, 1999). Elle a consisté à simuler la variable principale, ici la moyenne des mesures in situ sur les 20 premiers cm de sol (préalablement transformée en distribution gaussienne), à partir d'une variable secondaire, ici les valeurs de laboratoire simulées précédemment par SSG pour constituer la référence-labo, en reproduisant la corrélation observée sur les données entre mesures sur site et analyses de laboratoire. Comme précédemment, chaque type de données (teneurs FPXRF et teneurs de laboratoire) a dû faire l'objet d'une transformation gaussienne avant simulation et les modèles statistiques à reproduire ont été établis sur les données transformées. Ces modèles incluent le coefficient de corrélation entre les deux types de teneur et qui s'élève à 0.9 sur les données transformées.

Comme la référence-labo, la référence-FPXRF in situ a subi deux adaptations.

- Suppression de l'effet de pépite du modèle de variogramme.
- Suppression de deux mesures FPXRF (**Tableau 3**) dont les valeurs apparaissaient anormalement basses, et qui se sont avérées correspondre à des matériaux déplacés de leur lieu d'origine.

Tableau 3 Caractéristiques des données écartées pour simuler la référence-FPXRF in situ du cas d'étude MétalEurop.

ID	Pb – FPXRF in situ mg/kg ms	Occupation	(X,Y) m, Lambert II
59	59	Friche près de l'ancien puits de mine – remblais de démolition	(648845,2605026)
12	105	Bras de déchargement la Deûle – talus le long de la berge	(647873,2604000)

A la différence de ce qui avait été trouvé sur les données du projet OUTRAGE, l'étude approfondie des données du projet REPERAGE a démontré que la corrélation entre la mesure FPXRF sur site (moyenne de 5 mesures tous les 5 cm) et l'analyse de laboratoire sur échantillon composite n'était pas augmentée en corrigeant les mesures sur site de l'effet de la teneur en eau. Les données du projet OUTRAGE étant moins nombreuses que celles du projet REPERAGE, la validité de la correction peut être remise en question ou du moins son caractère général. Il ressort de ce résultat que la corrélation entre mesures ne peut pas être rendue universelle mais doit rester spécifique à chaque site et qu'il est nécessaire de pouvoir l'estimer convenablement en s'appuyant sur un nombre suffisant de couples de données (mesures sur site, analyses de laboratoire). Un minimum de 20 couples semble souhaitable. Aucune correction de la teneur en eau des mesures FPXRF n'a donc été appliquée ici.

3.2.2.c Résultats

La **Figure 6** montre les cartes de teneur en Pb simulées correspondant aux deux états références labo et FPXRF ainsi que les cartes de classification du sol par rapport aux seuils d'action nationale de 200 et 550 mg/kg ms (voir § 3.1). Le **Tableau 4** présente les statistiques élémentaires sur ces deux simulations de référence.

Tableau 4 Statistiques élémentaires des teneurs en Pb des simulations référence-labo et référence-FPXRF in situ générées sur le secteur d'étude MétalEurop.

	Nb	Moyenne	Ecart-type	Minimum	Maximum	Médiane	Coefficient de variation
	-	mg/kg ms	mg/kg ms	mg/kg ms	mg/kg ms	mg/kg ms	-
Référence-FPXRF in situ	32761	556	1067	33	8232	218	1.92
Référence-labo	32761	440	688	24	8754	241	1.56

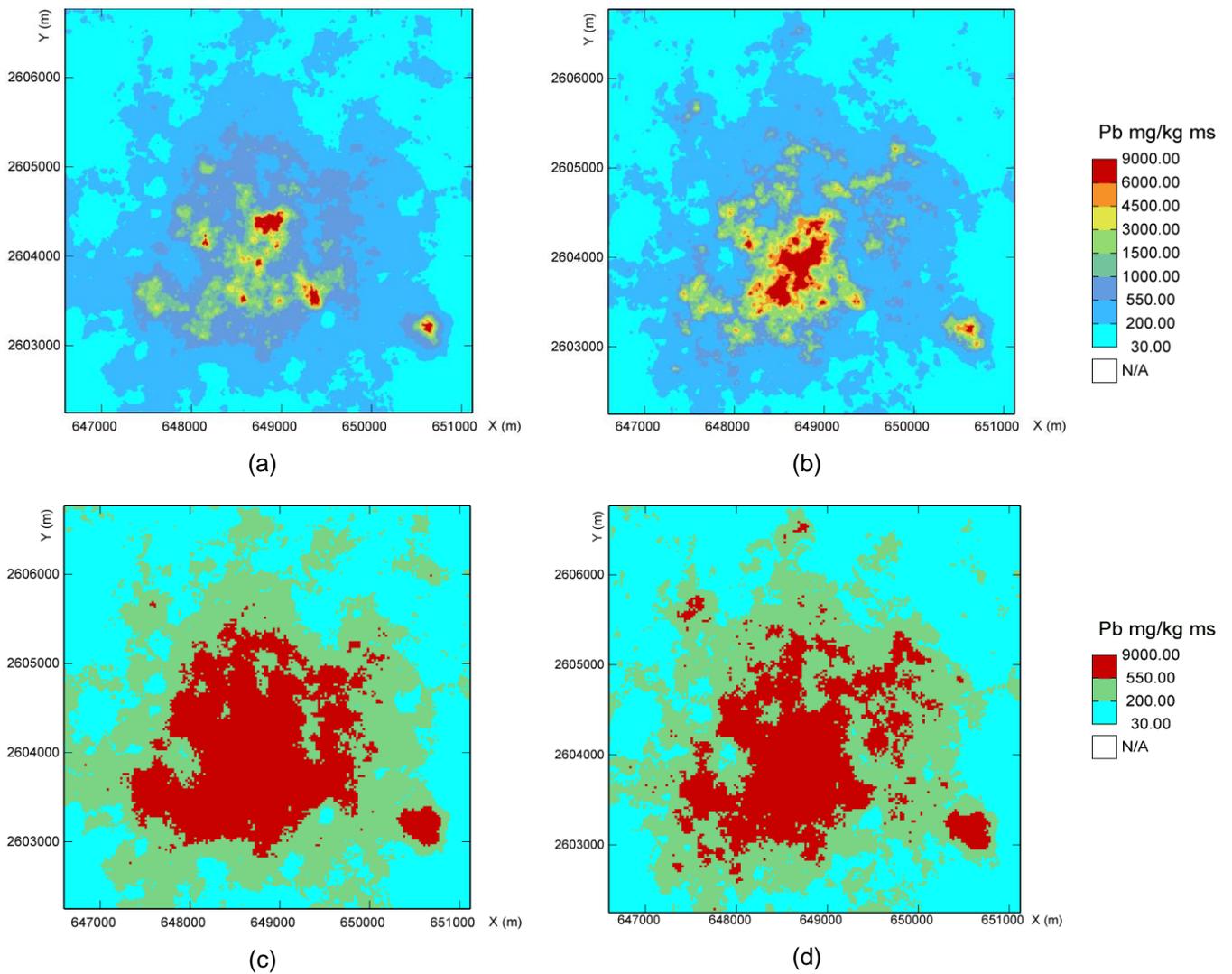


Figure 6 Etats de référence de la pollution moyenne en Pb sur les 20 premiers de sol du secteur d'étude MétalEurop. (a) teneurs de la référence-labo (b) teneurs de la référence FPXRF in situ (c) classification par rapport aux seuils d'action nationale Pb de la référence-labo (d) classification de la référence-FPXRF in situ.

3.2.2.d Constitution de la référence dans un cadre risque

En zone résidentielle ou récréative, la couche de sol qui présente un risque sanitaire par contact cutané, ingestion ou inhalation, est la couche de surface (épaisseur 5 cm). En zone agricole, le risque par contamination de la chaîne alimentaire (cultures, bétail) porte sur une couche de sol plus épaisse d'environ 20 cm. Dans un tel cadre d'évaluation des risques sanitaires, la constitution d'un état de référence de la pollution des sols sur une hauteur de 20 cm est nécessaire mais pas suffisante. Un état de référence sur la hauteur 0-5 cm est également requis.

La campagne de terrain menée dans le cadre de REPERAGE a été mise à profit pour collecter un petit nombre (20) d'échantillons représentatifs de la couche de sol 0-5 cm et les analyser au laboratoire par ICP-OES (voir § 3.2.1.c). Les 20 échantillons ont été sélectionnés parmi les 59 existants sur la base des critères suivants.

- Elimination des points d'échantillonnage pour lesquels des mesures FPXRF in situ sont manquantes sur la verticale. Cette situation nuit en effet à l'établissement de la relation entre types d'analyses.
- Présence d'échantillons pour toutes les classes de teneurs observées.
- Présence d'échantillons provenant de points où, au moyen des mesures FPXRF in situ, différentes situations de variabilité verticale ont été observées : grande variabilité sans tendance nette, décroissance avec la profondeur, croissance avec la profondeur, teneur plus forte en surface qu'en profondeur.
- Présence d'échantillons provenant de zones d'occupation variée : agricole, urbaine, bois, voie de communication, ancienne zone d'activité de MétalEurop.

Ces valeurs de laboratoire ne sont cependant pas suffisamment nombreuses pour qu'elles puissent, seules, permettre de générer un état de référence-labo de la couche 0-5 cm.

Une analyse exploratoire des données a donc été menée pour évaluer s'il existait une relation entre les données des couches 0-5 cm et 0-20 cm qui permette d'estimer, avec une bonne précision, des valeurs de laboratoire pour la couche 0-5 cm (voir **Annexe 5**). Les rapports entre mesures FPXRF des couches 0-5 cm et 0-20 cm d'une part et entre analyses de laboratoire des deux mêmes couches d'autre part se sont révélés être corrélés, leur relation pouvant être modélisée par une droite de régression. Des valeurs de laboratoire représentatives de la couche 0-5 cm ont donc été estimées à partir de cette droite de régression. Elles se sont révélées être proches des valeurs de laboratoire vraies des 20 points d'échantillonnage disponibles sur le site d'étude (erreurs de classification¹² de 15%). Dans ces conditions, il nous a semblé raisonnable d'utiliser cette relation pour compléter le jeu de données en estimant une valeur d'analyse de laboratoire pour les 40 échantillons de sol collectés sur la couche 0-5 cm, analysés sur le terrain à l'aide du FPXRF, mais non analysés au laboratoire.

¹² Par rapport aux seuils réglementaires 200 et 550 mg/kg ms.

Les mesures FPXRF et les analyses de laboratoire représentatives des teneurs en Pb de la couche 0-5 cm atteint ainsi un nombre de 50 et sont réparties de manière assez homogène sur l'ensemble du secteur d'étude. L'analyse variographique de ces données a mis en évidence une continuité spatiale à grande échelle, sans qu'une distinction puisse être faite entre sols remaniés (par des pratiques culturales en général) et sols non remaniés, ni en fonction des types d'usage (voir **Annexe 5**).

Une simulation des teneurs en Pb de laboratoire et mesurées sur site sur la couche 0-5 cm ont été générées selon les mêmes procédures que celles décrites aux § 3.2.2.a et 3.2.2.b pour la couche de sol 0-20 cm. Ces cartographies 2D sont présentées en **Annexe 5**.

Les variables d'intérêt dans un cadre risque ne sont cependant pas les teneurs en polluants, mais un indice de risque (IR) ou bien un excès de risque individuel (ERI). A titre d'illustration, on considère ici que la variable d'intérêt est un indice de risque calculé pour un risque par ingestion de sol ou de poussières dans un scénario chronique. Les paramètres retenus sont synthétisés dans le **Tableau 5**.

Tableau 5 Paramètres retenus pour calculer les indices de risque à partir des simulations des teneurs en Pb sur la couche 0-5 cm.

DI	mg/kg.j	Dose d'exposition par ingestion de sol ou de poussières contenant du sol	DI = Cs*Qs/P
P	kg	Poids du récepteur	15 kg pour un enfant de 6 ans
Ts	mg/kg	Teneur en polluant dans le sol ou les poussières contenant du sol	Teneur en Pb simulée
Qs	kg sol /j	Quantités de sol ou de poussières ingérées par jour	150x10 ⁻⁶ kg/j Rq. : En première approche, il est supposé que l'enfant passe son temps entier sur le site, d'où une fréquence égale à 1.
IR	-	Indice de risque	IR = DI/ VTR
VTR	mg/kg.j	Valeur Toxicologique de Référence	3.5x10 ⁻³ mg/kg.j

La fonction de transfert appliquée à chacune des simulations pour obtenir une cartographie des indices de risque est la suivante :

$$IR(x, y, z) = \frac{T_s(x, y, z) \cdot Q_s}{P \cdot VTR} \quad (3)$$

$$IR(x, y, z) = T_s(x, y, z) \cdot 0.0029 \quad (4)$$

Nous obtenons ainsi une référence-labo et une référence-FPXRF in situ de la variable IR sur l'ensemble du secteur d'étude. Le seuil IR=1 correspond à un seuil de teneur en Pb de 350 mg/kg ms. Ces cartographies 2D se trouvent en **Annexe 5**.

3.3 Simulation de reconnaissances conduites par la démarche de couplage entre mesures FPXRF et analyse géostatistique

La démarche de reconnaissance a été mise en œuvre de manière fictive (simulation sur ordinateur) à l'aide des cartes de référence-labo et référence-FPXRF in situ en se plaçant dans les conditions suivantes.

- L'objectif de la reconnaissance est de délimiter la zone où les teneurs en Pb dépassent le seuil de dépollution de 550 g/kg ms. Le volume de sol à dépolluer¹³ et la classification des sols doivent être estimés avec un niveau de confiance minimal de 80% : les mailles sont classées comme « incertaines » lorsque la probabilité de dépassement du seuil réglementaire est supérieure à 20% (au-dessous le sol est admis comme sain) et inférieure à 80% (au-dessus le sol est admis comme pollué).
- On considère que le nombre maximal de mesures sur site pouvant être réalisées en un jour de terrain est égal à 50.
- Aucune limite financière n'est fixée, et on considère en première approche qu'il n'y a pas de zones inaccessibles à l'échantillonnage.
- L'étape de reconnaissance préliminaire se fonde sur la collecte de données en 81 points répartis sur l'ensemble du domaine d'étude selon un schéma de grille régulier de 500 m x 500 m. Ce nombre de points de mesures semble grand, mais il est à rapporter à l'étendue particulièrement importante du secteur d'étude (14 000 ha). La durée estimée de cette étape est de 2 jours.
- A chaque étape itérative, les points de mesure FPXRF supplémentaires sont positionnés à partir d'une carte de classification des sols plus restrictive que celle que l'on cherche à atteindre : les mailles sont classées « incertaines » lorsque leur probabilité globale de dépasser les seuils réglementaires est inférieure à 95% (voir justification au § 3.5).
- La reconnaissance est arrêtée lorsque le niveau de précision sur la classification des sols dans les catégories « pollué » et « sain » est jugée acceptable. La reconnaissance simulée sur ordinateur a été ainsi stoppée au cinquième jour des étapes itératives.

Deux modules spécifiques ont été développés dans Matlab pour collecter les données sur les cartes de référence selon le schéma dicté par la démarche (*Nomdusite_etape_0_echagregulier_ref.m* et *Nomdusite_etape_j_echage_ref.m*).

L'évolution de la classification des sols au fur et à mesure de l'acquisition des données est visible sur les cartes de la **Figure 7** où chaque ligne correspond à une étape de reconnaissance. La colonne de droite de la figure présente les cartes de classification des sols dans les catégories « pollué », « non pollué »¹⁴ et « incertain » pour un niveau de confiance de 80%. Les données de laboratoire et de terrain collectées virtuellement dans l'étape préliminaire sont représentées dans

¹³ Ensemble des mailles dont les teneurs en Pb sont supérieures au seuil.

¹⁴ Equivalent à « sain »

les colonnes de gauche et centrale de la première ligne de la figure. Les données FPXRF collectées dans les zones incertaines de la classification, pour un niveau de confiance de 95%, sont représentées dans les colonnes de gauche et centrales des 5 lignes suivantes du tableau. Ces figures permettent de constater visuellement une réduction de la zone incertaine autour de la tache centrale de sols classés comme « pollués », ainsi qu'une délimitation affinée du contour de cette tache.

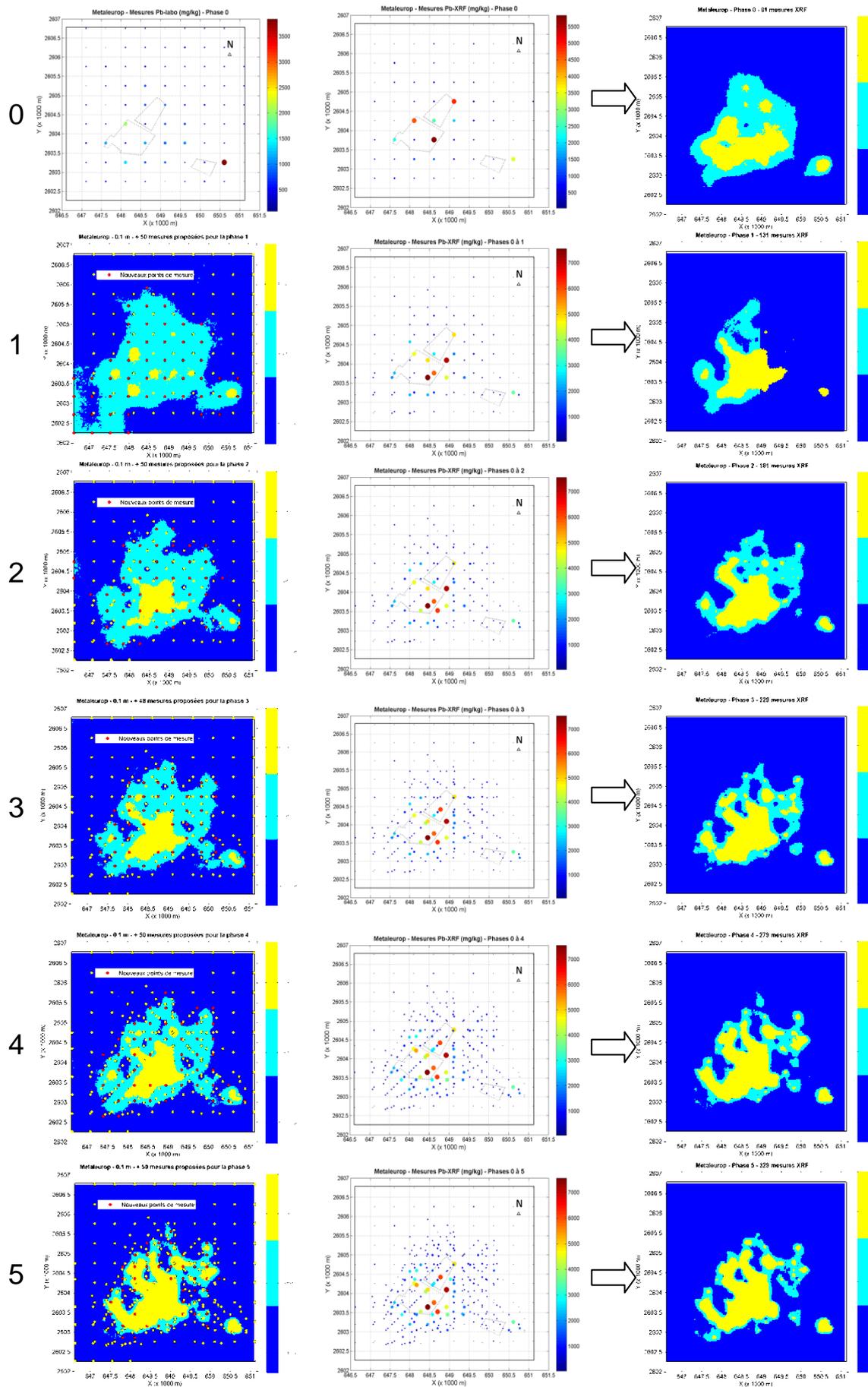


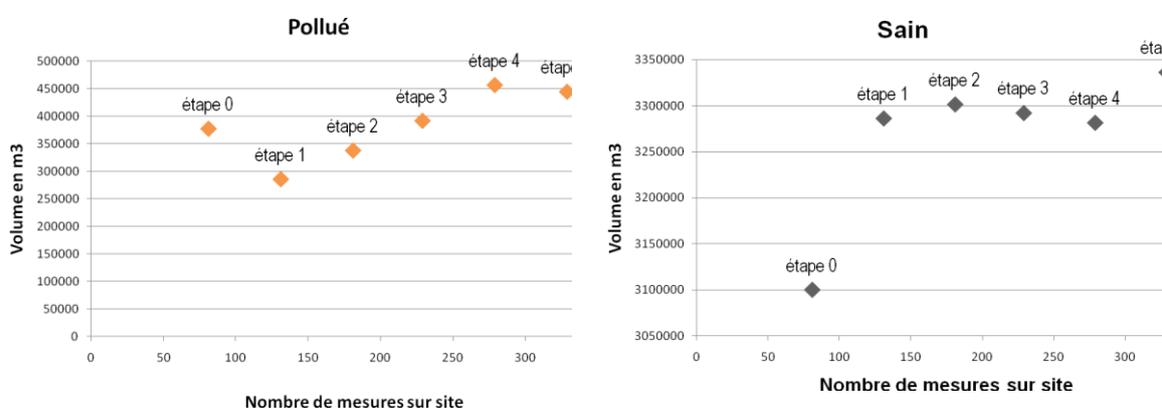
Figure 7 Résultats de l'application virtuelle de la démarche de couplage sur MetalEurop. Evolution de la classification des sols au fur et à mesure de l'acquisition de nouvelles mesures FPXRF.

Ces résultats visuels demandent cependant à être confirmés par l'estimation des volumes de sol associés aux différentes classes. Le **Tableau 6** résume les volumes calculés à chaque nouvelle étape de reconnaissance. Il montre que le volume de sol classé « incertain » diminue de moitié entre la phase préliminaire (étape 0) et l'étape itérative 5, passant de 618 000 à 314 250 m³. Le volume de sol classé « pollué » avec un niveau de confiance $\geq 80\%$ augmente de 18% entre ces deux phases, passant de 377 125 à 444 500 m³. Le volume de sol classé « sain » avec un niveau de confiance $\geq 80\%$ augmente quant à lui de 8%. Cette évolution apparemment « modeste » de l'estimation du volume de sol « sain » est toute relative. Elle correspond en réalité à un volume supplémentaire de sol « sain » de 236 375 m³ en fin de phase 5, volume qui semble faible lorsqu'il est rapporté au volume total, particulièrement grand, de la zone classée « saine » (volume total estimé à 3 336 375 m³ en fin de phase 5).

Tableau 6 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Volumes de sol classés « pollués », « sains » et « incertains » à chaque nouvelle étape de reconnaissance.

Étape	Nombre de données par étape FPXRF / labo	Nombre de données cumulées FPXRF	Pollué m ³	Sain m ³	Incertain m ³
0	81 / 81	81	377125	3100000	618000
1	50 / 0	131	285125	3286250	523750
2	50 / 0	181	337500	3301500	456125
3	48 / 0	229	391500	3292125	411500
4	50 / 0	279	456625	3281500	357000
5	50 / 0	329	444500	3336375	314250

Le graphe (haut, gauche) de la **Figure 8** représente les volumes de sols « pollués » estimés au fur et à mesure de l'acquisition de nouvelles données. Il montre un palier à partir de l'étape itérative 4 de la reconnaissance (279 mesures FPXRF). Ce palier se manifeste dès l'étape 1 sur les volumes de sols « sains » (graphe haut, droite) et est légèrement marqué sur les volumes de sol « incertains ». L'atteinte de ces paliers nous laisse penser qu'un apport supplémentaire d'information n'augmentera pas significativement la connaissance de l'extension de la zone « polluée », et a déterminé l'arrêt de la reconnaissance fictive.



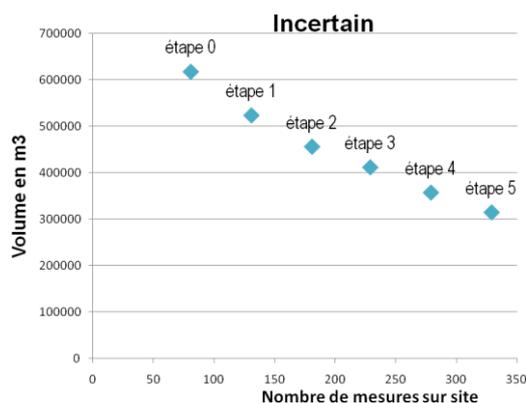


Figure 8 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Evolution des volumes de sols classés « contaminés », « sains » ou « incertains » au fur et à mesure de l'acquisition de nouvelles mesures FPXRF

Au bilan, pour un total de 81 analyses de laboratoire et 329 mesures FPXRF collectées en 1 jour de campagne préliminaire de reconnaissance et 5 jours de campagne itérative, la zone du secteur d'étude classée comme « incertaine » se trouve diminuée de manière très significative, et la délimitation de la zone « polluée » paraît satisfaisante. L'état de référence de la contamination nous permet d'approfondir cette première évaluation, en comparant la modélisation obtenue au bout de la cinquième étape itérative à la référence-labo, tel que décrit dans le paragraphe suivant.

3.4 Confrontation des résultats de simulation à l'état de référence

Le seuil de 550 mg/kg ms a été appliqué à la référence-labo pour classer l'ensemble de ses mailles dans les catégories « pollué » et « sain ». La cartographie de la classification de la référence-labo est superposée à la cartographie de la classification des sols issue de la modélisation de l'étape 5 de reconnaissance. Le résultat de cette superposition est présenté à la **Figure 9**. Les zones correctement classées par le modèle apparaissent en orange et jaune, les secteurs incorrectement classés figurent en vert clair et bleu ciel, tandis que les zones classées incertaines (risques d'erreur de classification $\geq 20\%$) sont représentées en bleu soutenu. Visuellement, les zones correctement classées apparaissent constituer l'essentiel du domaine d'étude.

Ce résultat est confirmé par la comparaison des volumes réels de sols classés « pollués » et « sains » de la référence-labo, aux volumes estimés par modélisation géostatistique en fin d'étape 5 (voir tableau de la **Figure 9**). Le modèle classe correctement 86% du volume total du secteur d'étude. Les 14% de sols restant se répartissent entre 6% de sols classés à tort dans les catégories « pollué » ou « sain » et 8% de sols classés « incertains ».

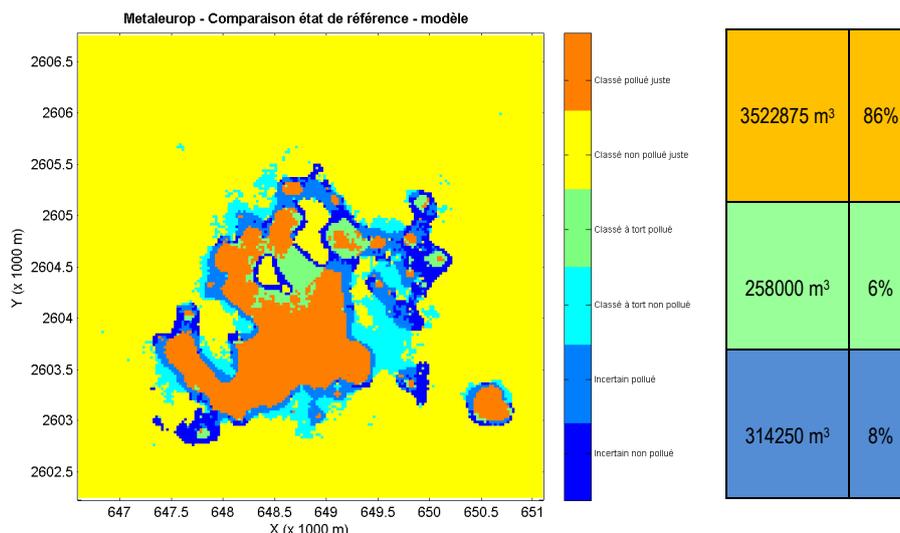


Figure 9 Résultats de l'application virtuelle de la démarche de couplage sur les environs de l'usine MétalEurop. Confrontation de la classification des sols obtenue en fin de reconnaissance à la classification de l'état de référence de la contamination.

3.5 Principales conclusions et retours d'expérience

3.5.1 Validité de la démarche

L'application virtuelle de la démarche dans des conditions réalistes de reconnaissance confirme son intérêt pour optimiser les diagnostics de sites. Sur la base de 2 jours d'étape préliminaire et de 5 jours de campagne itérative, on obtient :

- Une délimitation fine de la zone classée « polluée » avec un niveau de confiance $\geq 80\%$
- Une réduction de moitié de la zone « incertaine » en terme de volumes

La démarche est validée par la confrontation du modèle obtenu en fin de reconnaissance à la classification « vraie » des sols (référence-labo) : le modèle classe correctement l'essentiel du domaine d'étude (86%). Le nombre de mesures FPXRF qui peut être réalisé en un temps très réduit est important : on estime à 248 le nombre de mesures réalisables en 5 jours de campagne itérative.

3.5.2 Retours d'expérience

A côté de ces résultats généraux sur sa validité, un certain nombre d'enseignements sur les conditions de mise en œuvre de la démarche de couplage peuvent être tirés des tests réalisés sur les environs du site MétalEurop.

3.5.2.a Importance de l'étape préliminaire

L'étape préliminaire joue ici un rôle déterminant dans la suite de la démarche. C'est en effet à ce stade que les principales taches de pollution sont détectées. L'algorithme de positionnement des mesures collectées lors des étapes itératives suivantes est conçu initialement pour préciser la délimitation et l'extension de ces taches, mais pas pour en déceler de nouvelles. Les tests réalisés sur le cas d'étude MétalEurop ont montré qu'un maillage trop lâche lors de l'étape préliminaire pouvait conduire à « rater » la petite zone « polluée » se trouvant dans le secteur sud-est de la carte (voir

Figure 9). Le maillage de la zone d'étude doit donc être défini soigneusement, en rapport avec la taille minimale attendue des taches de pollution. Des informations à ce sujet peuvent généralement être tirées de l'investigation historique du site. Lorsque l'étape préliminaire s'avère insuffisante pour détecter les principales taches, l'algorithme de positionnement des mesures des étapes itératives suivantes peut être adapté pour servir à la fois à identifier les taches et à délimiter leur extension : c'est ce qui a été fait lors de la campagne de reconnaissance réelle du site Henricot II (voir § 4.4.2).

C'est aussi au stade de l'étape préliminaire que les modèles de covariogrammes des données de laboratoire et de terrain sont ajustés. Les mesures sur site acquises au fur et à mesure des étapes de reconnaissance suivantes permettent d'adapter éventuellement les portées des modèles, mais pas leur palier. Les lois de régression entre valeurs de laboratoire et mesures de terrain sont quant à elles modélisées définitivement à ce stade, pour être ensuite utilisées pour déterminer des teneurs équivalentes à des analyses de laboratoire à partir de mesures sur site nouvellement acquises. Le nombre de couples de données (analyses de laboratoire, mesures sur site) doit donc être choisi pour qu'une fiabilité suffisante soit attribuée à l'ajustement des modèles de covariogramme et des lois de régression. Alors que le nombre permettant d'assurer un ajustement correct de modèles de structure spatiale dépend des caractéristiques de la contamination, un nombre minimum de 20 couples de données représentatives de la gamme des teneurs observées sur le site peut être avancé pour un ajustement correct de lois de régression.

3.5.2.b *Influence du resserrement des mesures dans la zone « incertaine »*

La procédure de reconnaissance initialement mise en œuvre sur le cas d'étude MétalEurop consistait à positionner les mesures sur site nouvellement acquises dans les zones « incertaines » calculées pour le niveau de confiance recherché ($\geq 80\%$). Les tests effectués ont cependant révélé que cette procédure ne permettait pas de réduire la proportion de sols classés incorrectement, même si l'effort de reconnaissance (multiplication des points de mesures sur site) était conséquent. Il est apparu plus efficace au final de positionner les points sur une carte de classification des sols plus restrictive (niveau de confiance $\geq 95\%$), ce qui revient à disposer des points dans- et de part et d'autre de la zone « incertaine » pour le niveau de confiance recherché. Cette procédure aide à préciser le contour de la zone classée « polluée ».

Les tests réalisés sur le cas d'étude MétalEurop ont montré également que le resserrement des points de mesures dans une zone préférentielle (la zone « incertaine ») perturbait la fonction de transformation des distributions des données en distributions gaussiennes (fonction de transformation par la méthode des scores normaux) ainsi que les covariogrammes (augmentation anormale des portées des modèles de covariogrammes). Le premier effet a été corrigé en effectuant un dégroupement¹⁵ sur les mesures sur site. Le deuxième effet n'a pas été répercuté sur la classification des sols en ne mettant volontairement pas à jour la portée des modèles de covariogrammes à partir de la deuxième étape itérative de reconnaissance.

¹⁵ Procédure consistant à attribuer une pondération à chacune des données de la distribution, cette pondération étant d'autant plus faible que la donnée se trouve à petite distance d'autres données (données en clusters).

3.5.2.c Choix du critère d'arrêt de la reconnaissance

Le secteur d'étude MétalEurop se caractérise par une tache centrale de grande dimension de sol « pollué » dont la circonférence, aux nombreuses digitations, est très grande. Cette spécificité conduit à la situation où un effort important de reconnaissance (multiplication du nombre de données) est nécessaire pour arriver à une réduction finalement modeste de l'ampleur de la zone « incertaine » englobant la zone « polluée » (voir **Figure 7**). C'est cette situation, caractéristique de ce site et qui ne se retrouvera pas forcément dans d'autres cas d'étude, qui justifie l'arrêt de la reconnaissance fictive à l'étape itérative 5.

4 MISE EN ŒUVRE « EN TEMPS REEL » DE LA DEMARCHE DE COUPLAGE SUR LE SITE HENRICOT

4.1 Présentation du site d'étude

Le site sélectionné pour mettre en œuvre la démarche de couplage « en temps réel » est une partie de l'ancienne usine sidérurgique Henricot II de Court Saint Etienne (5.6 ha), située non loin du confluent de la Dyle et de la Thyle, en Brabant wallon, Belgique. L'ISSeP a eu pour mission de mener des études visant à déterminer les risques environnementaux et à définir un plan d'assainissement pour ce site. Ces études ont mis en évidence une contamination de double nature, organique et inorganique, dans les remblais d'épaisseur moyenne 2 m. En l'absence d'arrêté d'exécution relatif à la réhabilitation de friches industrielles¹⁶ à la date de réalisation des études, la nécessité d'assainir le site a été évaluée à partir d'un arrêté wallon relatif à l'assainissement des stations-service¹⁷. Cet arrêté définit des valeurs seuils des teneurs en polluants dans le sol et le sous-sol.

La démarche de couplage a été mise en œuvre à titre expérimental, après que le plan d'assainissement ait été défini, l'ISSeP n'ayant pas l'intention d'exploiter les résultats de cette reconnaissance dans le cadre des mesures de gestion du site, mais percevant un intérêt scientifique et méthodologique dans cette expérimentation.

Le secteur d'étude (**Figure 10**) correspond aux sols dont certaines teneurs en métaux se sont révélées dépasser ces valeurs seuils lors d'investigations menées par l'ISSeP. Ce domaine exclut les zones de forte contamination en hydrocarbures. Les métaux retenus pour les besoins de la démonstration sont le Cu, le Ni, le Pb et le Zn.

¹⁶ Arrêté d'exécution du décret relatif à l'assainissement des sols pollués et aux sites d'activités économiques à réhabiliter du 1^{er} avril 2004

¹⁷ AGW du 4 mars 1999 : arrêté du gouvernement wallon relatif à l'assainissement des stations-service.

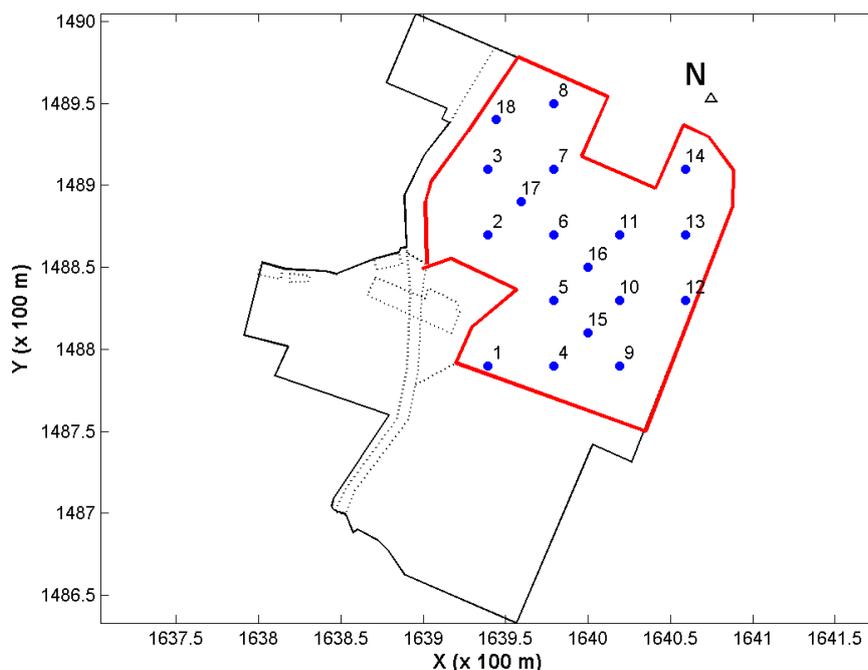


Figure 10 Limite du secteur d'étude sur le site de l'ancienne usine sidérurgique Henricot II et localisation des 18 sondages prévus pour l'étape préliminaire de reconnaissance

4.2 Les conditions de mise en œuvre de la démarche

La démarche de couplage a été mise en œuvre dans un cadre dépollution, où l'objectif final de la reconnaissance est de classer les sols par rapport à des seuils de teneurs en métaux au-delà desquels ils doivent être assainis. Les seuils appliqués sont les suivants (en mg/kg ms).

Cu	Ni	Pb	Zn
100	100	100	150

Compte-tenu de la faible épaisseur de terrains à traverser, il a été décidé de recourir à des sondages réalisés à l'aide d'une foreuse de type « Geotool ». Même si elle ne permet pas une description précise de la stratigraphie, cette foreuse est très facile à manier et répond aux besoins de l'étude, à savoir réaliser un nombre non négligeable de sondages en une journée. Le nombre maximal de sondages pouvant être réalisé avec ce type de machine a été estimé à 10, avec possibilité de collecter 3 à 5 mesures par sondage, soit un total de 30 à 50 mesures sur site FPXRF par jour.

Les appareils FPXRF et DGPS utilisés sont les mêmes que ceux utilisés lors de la campagne de terrain menée sur le secteur d'étude MétalEurop (voir § 3.2.1). La durée de mesure sur site a été fixée à 2 mn, et 3 mesures sur site ont été prises en 3 endroits différents de chaque échantillon, puis moyennées.

La quantification des métaux Cu, Ni, Pb et Zn a été réalisée au laboratoire de l'INERIS, pour l'étape préliminaire seulement, dans les mêmes conditions que celles observées lors de la campagne de terrain sur le cas d'étude MétalEurop.

L'étape préliminaire de reconnaissance a été planifiée à partir de la documentation disponible sur le site. Un plan de localisation des sondages et des points de mesure et d'échantillonnage a été établi, ainsi que les tableaux des identifiants et des coordonnées Lambert II de ces points. Le plan de reconnaissance de l'étape préliminaire inclut 18 sondages répartis selon une maille régulière 40 m x 40 m et 3 sondages plus rapprochés. Chaque sondage a fait l'objet de 3 échantillons de sol destinés à une double analyse FPXRF – ICP-OES, ces échantillons étant représentatifs d'une hauteur de 25 cm de sol. Les 3 échantillons sont séparés dans les sondages d'une hauteur de 75 cm. Un document de préparation de l'étape préliminaire a été fourni à l'ensemble des acteurs du projet (voir **Annexe 6**).

On attribua à la reconnaissance l'objectif de réduire l'incertitude sur la classification des sols à un niveau acceptable. Dans un premier temps, on chercha à classer les sols dans les catégories « pollué » et « sain » avec un niveau de confiance minimum de 70%, mais cette objectif fut revu à la baisse au fur et à mesure de l'acquisition de nouvelles données lorsque la contamination des sols se révéla très hétérogène (voir § 4.4.2).

Le budget de la reconnaissance était limité à celui fixé dans la convention de recherche liant l'ADEME et les partenaires du projet (voir § 6.2 pour les coûts effectifs).

La durée nécessaire à la campagne itérative de reconnaissance fut estimée à 5 jours en fin d'étape préliminaire. Cette estimation s'avérait nécessaire pour planifier la disponibilité de l'équipe de terrain et vérifier que le budget attribué à la reconnaissance était respecté (demande de devis à la société de forage notamment).

4.3 La campagne de terrain

La campagne de terrain a consisté en une très longue journée d'étape préliminaire (19 juin 2008) où 18 sondages furent foncés pour prélèvements et mesures sur 52 échantillons de remblais, et en 5 jours d'étapes itératives de reconnaissance, du 7 au 11 juillet 2008, durant lesquels 216 mesures sur site ont été réalisées sur 47 sondages différents. En moyenne, 43 mesures FPXRF ont pu être réalisées chaque jour (**Tableau 7**).

Durant la phase de reconnaissance itérative, la journée de terrain débutait vers 8h00. Le DGPS était utilisé dès le début de la journée pour marquer la position des sondages indiquée sur le plan et dans le tableau fournis par les géostatisticiens. Les sondages, prélèvements et mesures sur site étaient ensuite réalisés à ces emplacements par une équipe de trois personnes. Entre 17h et 18h, les résultats des mesures DGPS et FPXRF de la journée étaient transmis par courriel sous forme de tableaux Excel aux géostatisticiens. Les données étaient traitées durant la soirée, puis les positions des nouveaux sondages et points de mesure pour la journée suivante étaient transmises à l'équipe de terrain par courriel sous forme de tableaux Excel et de carte. Un exemple de la documentation fournie à l'équipe de terrain est présenté à l'**Annexe 7**.

Bien que le nombre de sondages qu'il était estimé possible d'implanter en une journée fût de 8 à 10, la position de 10 à 12 sondages était fournie au début de chaque nouvelle étape de reconnaissance. L'équipe de terrain disposait ainsi d'un nombre suffisant d'ouvrages pour remplacer

l'un d'entre eux en cas de refus sur terrain induré ou infrastructure. Les principaux incidents survenus lors des travaux de terrain ont consisté en effet en des refus de sondages à certaines profondeurs.

Durant les étapes itératives, 5 échantillons de sol étaient prélevés dans chacun des sondages, entre 0 et 2 m de profondeur, pour analyse des métaux à l'aide du FPXRF. Le nombre de mesures par sondage a été augmenté en fin d'étape préliminaire (passage de 3 à 5 mesures par sondage) pour améliorer la connaissance de la continuité spatiale de la contamination sur la verticale. Les mesures étaient donc prises sur des échantillons représentatifs d'une hauteur de 25 cm de sol, aux profondeurs 0.225 m, 0.575 m, 0.925 m, 1.275 m et 1.625 m.

4.4 Le traitement des données

4.4.1 Etape préliminaire

Les données de terrain et de laboratoire de l'étape préliminaire ont été traitées en 3-4 jours. Les conditions de modélisation ont été définies à ce stade :

- La grille de calcul est constituée de 77952 mailles de dimensions 2 m x 2 m x 0.25 m. La grille comporte 7 mailles par verticale, de 0 à 1.75 m de profondeur.
- L'analyse exploratoire des données a mis en évidence une corrélation entre les éléments Cu, Ni, Pb et Zn (coefficients de corrélation de Spearman allant de 0.61 à 0.80). De même, une corrélation a été observée entre les mesures FPXRF et les analyses ICP-OES de laboratoire (coefficients de corrélation de Spearman allant de 0.55 à 0.95). De plus, l'analyse variographique a mis en évidence une structure spatiale sur les 4 éléments, avec une moindre continuité sur les éléments Cu et Ni que sur les éléments Pb et Zn. La portée des phénomènes est plus petite sur la verticale que dans le plan horizontal. L'ordre de cosimulation des teneurs en métaux tient compte de ces observations : les types de données les mieux corrélées sont simulés en priorité. Les teneurs en Zn analysées au laboratoire et mesurées sur le terrain sont cosimulées dans le modèle de covariogramme des teneurs en Zn de laboratoire et de terrain. Puis les teneurs en Pb de laboratoire sont cosimulées dans le modèle de covariogramme des teneurs en Pb et en Zn de laboratoire, puis les teneurs en Cu de laboratoire dans le modèle de covariogramme des teneurs en Cu et en Pb de laboratoire, puis les teneurs en Ni dans le modèle de covariogramme des teneurs en Ni et en Cu de laboratoire.
- Une loi de régression a été ajustée aux données de laboratoire et de terrain et la relation entre l'écart-type de résidu et les mesures FPXRF (voir § 2.3.1) a été modélisée pour chacun des 4 métaux.
- Sur la base des résultats obtenus sur le cas d'étude MétalEurop (voir § 3.5.2.b), la classification globale des sols par rapport aux seuils définis pour les 4 métaux a été calculée pour un niveau de confiance $\geq 70\%$, tandis que le positionnement de 10 nouveaux sondages a été réalisé sur la base d'une classification des sols plus sévère, pour un niveau de confiance $\geq 80\%$.

4.4.2 Etapes itératives

Les données des étapes itératives ont été traitées en 7h le premier soir, 5h en fin d'étapes 2 et 3, 3h en fin d'étape 4 et 2h en fin d'étape 5. Le traitement des données « en temps réel » a nécessité un certain temps d'adaptation, ce qui explique une meilleure efficacité en fin qu'en début de reconnaissance. Un incident survenu le premier soir (erreur dans les coordonnées des points de mesure prises sur le terrain) explique également le temps plus long consacré au traitement des données de l'étape 1.

Le traitement des données fut principalement marqué par la nécessité d'adapter les conditions de modélisation et de positionnement des nouveaux sondages au fur et à mesure de leur acquisition (**Tableau 7**).

- L'apport d'informations supplémentaires conduisit en effet à corriger les modèles de covariogrammes. Les portées horizontales et verticales des modèles se révélèrent nettement plus courtes que ce qui avait pu être déduit des données de la campagne de terrain préliminaire (étape 0). Dans le plan horizontal, selon le polluant, les portées furent réduites de 33-46 m à 9-15 m ; sur la verticale, les portées furent réduites de 1-6 m à 0.8-1.2 m.
- La contamination se révéla progressivement comme présentant une forte variabilité à petite échelle (de fortes teneurs peuvent côtoyer de faibles teneurs, même séparées de courtes distances) qui ne permettait pas d'envisager de classer les sols avec le même niveau de confiance que celui recherché sur le cas d'étude MétalEurop (caractérisé par une forte continuité spatiale à grande échelle), sauf effort démesuré en termes de reconnaissance. Le niveau de confiance sur la classification des sols fut donc revu à la baisse au fur et à mesure de l'acquisition des données, pour se fixer à $\geq 55\%$.
- Le phénomène apparaît cependant structuré, avec plusieurs petits « spots » de fortes teneurs répartis dans l'ensemble du domaine d'étude. Il apparut alors nécessaire d'adapter l'algorithme de positionnement des sondages. En l'état, l'algorithme définit l'enveloppe incluant toutes les mailles de la grille de classification classées « incertaines », et place les sondages à l'intérieur de cette enveloppe selon un schéma régulier. Cette enveloppe est très grande lorsque les mailles « incertaines » se répartissent sur tout le domaine d'étude, ce qui est le cas avec une disposition de la contamination en « spots » dans tout le domaine. Les nouveaux points de mesures sont alors positionnés dans cette grande enveloppe sans qu'ils puissent aider à préciser la délimitation des « spots ». La localisation des points de sondage a donc été modifiée à partir de l'étape itérative 4 : on définit une grille fine de positionnement des points de sondages, et on sélectionne parmi les mailles de cette grille fine, celles qui contiennent le plus de mailles « incertaines » de la grille de classification (La taille des mailles de la grille de positionnement des sondages ne dépend alors plus du nombre de sondages à positionner) (voir § 2.3.3). On obtient de cette façon un positionnement des nouveaux sondages autour des « spots » de pollution.

Tableau 7 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Conditions de modélisation à chaque nouvelle étape de reconnaissance

Etape	Sondages par étape	Sondages cumulés	Données par étape FPXRF / labo	Données cumulées FPXRF	Portées horizontales m	Portées verticales m	Localisation des sondages	NC Localisation
0	18	18	52 / 52	52	33-46	1-6	Mailles larges	80%
1	8	26	40 / 0	92	17-30	0.6-1	Mailles larges	60%
2	9	35	40 / 0	132	8-15	0.8-1.5	Mailles larges	57.5%
3	10	45	50 / 0	182	9-15	0.8-1.5	Mailles larges	55%
4	10	55	44 / 0	226	9-15	0.8-1.2	Petites mailles	55%
5	10	65	42 / 0	268	9-15	0.8-1.2	Petites mailles	55%

4.5 Les résultats obtenus

L'évolution de la classification des sols dans les catégories « contaminé » et « sain » est présentée pour la couche de sol 0-0.25 m dans la colonne de droite de la **Figure 11**. La position des données collectées à chaque étape de reconnaissance est montrée dans la colonne de gauche de cette figure sur carte de classification des sols et dans la colonne du centre pour l'élément Zn.

Les codes de calcul Matlab fournissent des cartes des données pour chacune des 7 couches de sol de 0.25 m, de 0 à 1.75 m de profondeur, pour chacun des métaux Cu, Ni, Pb et Zn. La totalité des cartes produites lors de la campagne de reconnaissance se trouvent sauvegardées dans le CD-Rom fourni avec le rapport.

Ces cartes mettent en évidence une répartition des sols dont les teneurs dépassent les valeurs seuils en une dizaine de petits « spots » répartis sur l'ensemble du secteur d'étude.

Les volumes de sols estimés à chacune des étapes de la reconnaissance sont présentés dans le **Tableau 8**. Les volumes des différences classes « pollué », « sain » et « incertain » ne sont pas comparables d'une étape de reconnaissance à l'autre, puisqu'ils correspondent à des niveaux de confiance sur la classification et des modèles de covariogrammes différents. Au terme d'une semaine de reconnaissance itérative, le volume de sol classé « pollué » avec un niveau de confiance $\geq 55\%$ est égal à 3384 m³ (8% du volume total du domaine d'étude), le volume de sol classé « sain » avec ce même niveau de confiance est égal à 36523 m³ (81% du volume total du domaine d'étude) et le volume de sol classé « incertain » est égal à 5047 m³ (11% du volume total),

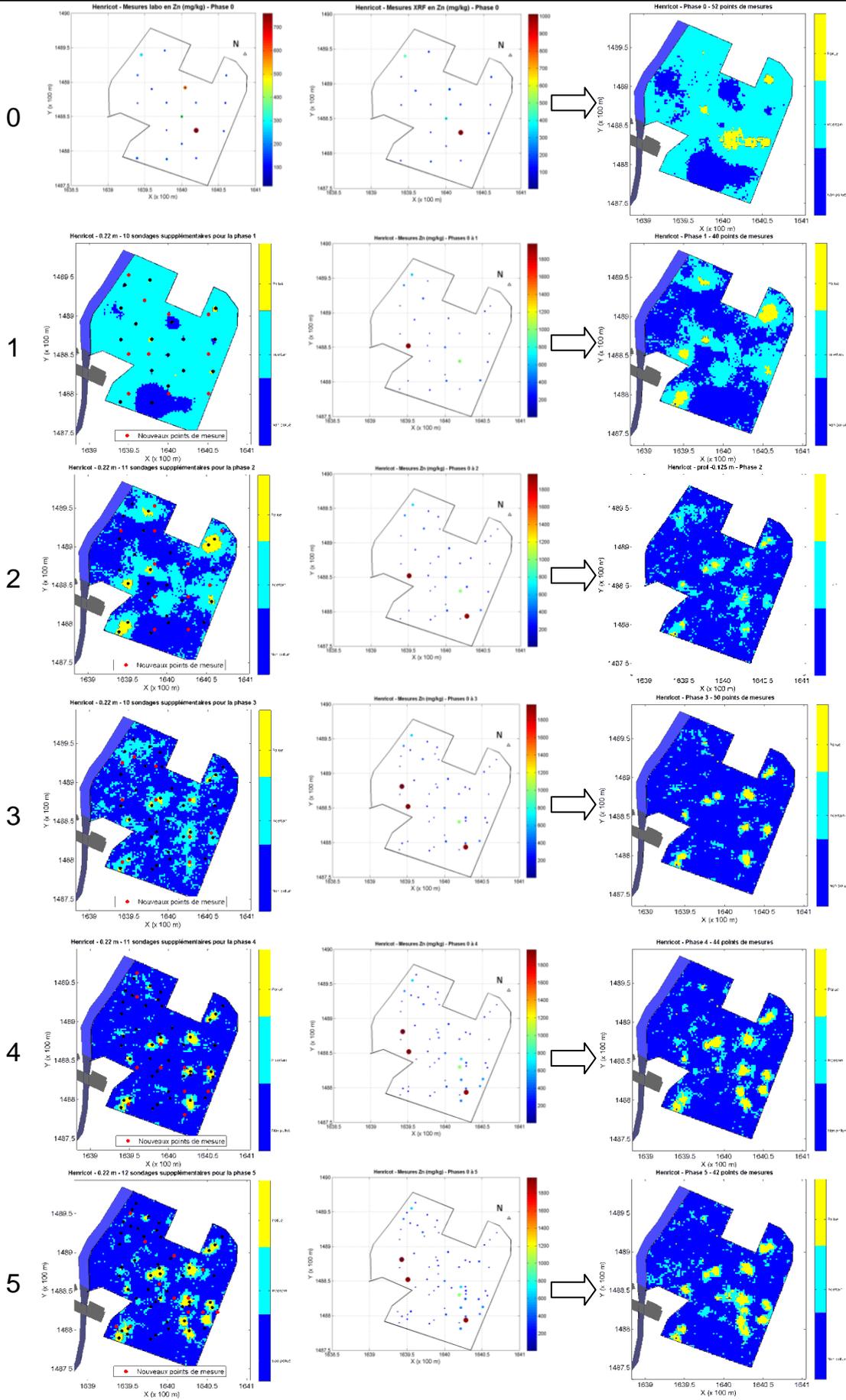


Figure 11 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Evolution de la classification des sols au fur et à mesure de l'acquisition de nouvelles mesures FPXRF

Tableau 8 Résultats de l'application réelle de la démarche de couplage sur le site Henricot II. Volumes de sol classés « pollués », « sains » et « incertains » à chaque nouvelle étape de reconnaissance

Étape	Sondages p. étape	Sondages cumulés	Données p. étape FPXRF / labo	Données cumulées FPXRF	NC Classif.	Pollué m ³	Sain m ³	Incertain m ³
0	18	18	52 / 52	52	70%	3616	17813	23525
1	8	26	40 / -	92	60%	3186	27428	14340
2	9	35	40 / -	132	57.5%	1796	34596	8562
3	10	45	50 / -	182	55%	2060	38760	4134
4	10	55	44 / -	226	55%	2996	36542	5416
5	10	65	42 / -	268	55%	3384	36523	5047

4.6 Principales conclusions et retours d'expérience

La démarche de couplage mise en œuvre en temps réel sur le cas d'étude Henricot II a permis de réaliser 47 sondages et 216 mesures FPXRF en 5 jours continus de campagne, et 65 sondages, 52 analyses de laboratoire et 268 mesures FPXRF en 6 jours de campagne si l'on inclut l'étape de reconnaissance préliminaire. L'estimation, sur le cas d'étude MétalEurop, du nombre de mesures sur site pouvant être réalisées de manière journalière (30 à 50 mesures FPXRF par jour), est confirmé sur le cas d'étude Henricot, malgré des conditions d'échantillonnage différentes (grand domaine et sols de surface sur le cas d'étude MétalEurop, petit domaine et remblais sur 2 m d'épaisseur sur le cas d'étude Henricot II).

L'apport de la démarche de couplage à la classification des sols est différent de celui qui avait pu être prévu à partir des tests réalisés sur le cas MétalEurop. La reconnaissance itérative aboutit à une meilleure connaissance de la répartition spatiale de la contamination dans les remblais. Il s'agit de plusieurs « spots » de contamination de dimensions restreintes, correspondant à un faible volume de sol, estimé à 3384 m³ (pour un niveau de confiance $\geq 55\%$). Mais du fait d'un phénomène particulièrement hétérogène, il ne peut être attendu de réduction importante de l'incertitude sur la classification sans effort de reconnaissance démesuré. Les objectifs en termes de précision sur la classification ont donc été revus à la baisse en cours de reconnaissance pour s'adapter aux caractéristiques spécifiques de ce site. L'acquisition de données supplémentaires ne permet pas d'observer dans ce contexte de réduction d'incertitude sur la classification, puisque les conditions de modélisation sont modifiées d'une étape à l'autre.

Ces résultats ne retirent en rien l'intérêt de la démarche pour déterminer la répartition spatiale de la pollution en un temps très court. Ils montrent de plus l'importance de pouvoir adapter la démarche aux caractéristiques spécifiques de chaque site, même si l'essentiel du traitement géostatistique des données a été automatisé. Dans le cas du site Henricot II, des solutions ont été trouvées en cours de reconnaissance pour resserrer les points de mesure autour des « spots » de contamination, ce qui ne pouvait être obtenu avec l'algorithme de calcul initial.

5 CONDITIONS D'INTEGRATION DE LA DEMARCHE DANS LES METHODOLOGIES FRANCAISE ET WALLONE DE GESTION DES SITES POLLUES

5.1 Méthodologie française

5.1.1 Rappel sur les outils de gestion des sites et sols pollués en France

La circulaire du 8 février 2007 définit les modalités de gestion des sites et sols pollués. Deux outils ont été mis en place : l'interprétation de l'état des milieux et le plan de gestion.

Le plan de gestion s'applique sur l'emprise du site contaminé. Il consiste à supprimer les sources de contamination puis les voies de transfert entre les sources et les populations qui fréquentent ou fréquenteront le site. Dans le cas où le plan de gestion ne permet pas de supprimer la totalité de la source, alors une analyse des risques résiduels (ARR) doit être conduite. L'ARR consiste à démontrer que les expositions résiduelles qui subsistent à l'issue du plan de gestion n'induisent pas de niveaux de risque inacceptables pour les populations.

L'interprétation de l'état des milieux (IEM) a pour objectif d'étudier l'impact éventuel d'un site contaminé sur son environnement. Elle consiste à vérifier, pour un usage donné d'un milieu, la compatibilité entre l'état de ce milieu et l'usage qui en est fait. Elle se fonde en priorité sur la mesure des concentrations des polluants dans les différents milieux auxquels les populations sont exposées : eaux, sols, légumes par exemple. Les concentrations ainsi mesurées sont alors comparées à des concentrations de référence qui sont :

- Soit les valeurs réglementaires dans le cas où elles existent
- Soit les valeurs de référence caractérisées dans un environnement témoin, non impacté

Dans le cas où les concentrations mesurées dans l'environnement sont comparables aux concentrations de référence, alors l'état du milieu est déclaré compatible avec les usages. Dans le cas contraire, un plan de gestion doit être mis en place.

Ces deux outils se fondent sur un schéma conceptuel qui synthétise les sources de contamination, les voies de transfert et les enjeux présents sur le site ou aux alentours du site.

5.1.2 Application de la méthode dans le contexte français

Différentes applications de la méthode développée dans le cadre du projet REPERAGE sont envisageables :

- **Lors du diagnostic de site** : le couplage entre fluorescence X portative et géostatistique permet, à partir d'un nombre limité de points, d'avoir une connaissance approfondie des concentrations en métaux présents dans les sols. Le couplage proposé dans le cadre du projet REPERAGE permet en outre une démarche itérative en un temps relativement court, ce que ne permettent pas les campagnes d'échantillonnage classique où seuls des prélèvements pour envois en laboratoire

sont réalisés. Ceci implique bien entendu une bonne coordination entre les opérateurs sur le terrain et les personnes qui réalisent la modélisation géostatistique.

- **Lors du plan de gestion** : la technique proposée ici permet de quantifier les volumes de terres contaminés. Ce point est crucial en vue de quantifier au mieux les coûts de réhabilitation mise en œuvre dans le plan de gestion.
- **Quantification des incertitudes** : cette phase est indispensable dans toute démarche de gestion des sites et sols pollués. Toutefois, elle est rarement abordée faute de technique robuste permettant la quantification des incertitudes. Du fait de l'hétérogénéité des sols, des incertitudes majeures sont associées à leur échantillonnage. La géostatistique couplée à un système d'information géographique permet d'avoir une représentation spatialisée des incertitudes associées aux teneurs prédites. Le couplage proposé ici permet une quantification des incertitudes liées à une étape majeure qu'est le diagnostic de site et en particulier l'échantillonnage de sols. En ce sens la méthode proposée ici permet une avancée majeure dans la quantification des incertitudes liées au diagnostic de site. La prise de décision en sera ainsi facilitée.
- **Communication** : La communication est une étape fondamentale de la gestion des sites et sols contaminés. Elle se fait à différents niveaux : entre experts, avec les administrations déconcentrées de l'état ou avec les riverains des sites. La méthode développée dans le cadre du projet REPERAGE permet une représentation spatialisée des teneurs en contaminants sur le site ou aux alentours du site. En outre, comme mentionné plus haut, une représentation spatiale des incertitudes associées aux concentrations prédites est également réalisée. Ce mode de représentation n'est actuellement que peu employé lors d'opérations de communication. Ceci est notamment lié au fait que les techniques actuellement employées ne permettent pas d'extrapoler les concentrations mesurées sur un échantillon prélevé en un point du site. En proposant une représentation cartographique, la technique développée ici contribuera ainsi à faciliter la communication entre les différentes parties prenantes de la gestion d'un site pollué.

5.1.3 Extrapolation à d'autres contaminants

Dans le cas du projet REPERAGE, une application a été proposée à partir d'un appareil permettant de caractériser les métaux dans les sols. Il est tout à fait envisageable de proposer une méthode équivalente pour des appareils de terrain permettant de caractériser les pollutions organiques. L'INERIS a récemment réalisé un recensement des appareils de mesures in situ (Burguet, 2008). Ainsi la spectrométrie Raman et la chromatographie phase gazeuse couplée à la spectrométrie de masse sont deux techniques qui peuvent s'appliquer à l'analyse in situ ou on site des contaminants organiques et qui pourraient être couplées à la géostatistique.

5.2 Méthodologie wallone

5.2.1 Rappels sur les outils de gestion des sols pollués en Wallonie

Actuellement en Wallonie, les sols sont gérés conformément au « décret déchets » du 27/06/1996 et à l'arrêté du gouvernement wallon du 04/03/1999 (modifié le 17/7/2003) concernant l'implantation et l'exploitation des stations service. Ces textes définissent trois catégories de concentrations en contaminants potentiels : des valeurs de référence (VR), des valeurs seuil (VS) et des valeurs d'intervention (VI). Les valeurs de VS et VI dépendent de l'affectation du terrain considéré. Quatre types d'affectation sont définis.

La « directive déchets » du 21/10/2008 n'inclut pas les sols pollués, d'où un vide juridique qu'il faut rapidement combler.

Le « décret sol » existe depuis 2004 mais des arrêtés d'exécution manquaient pour son application. Le parlement wallon vient d'approuver le projet de décret le 3 décembre 2008. Ce décret fait une distinction entre une pollution historique (avant avril 2007) et une pollution nouvelle. Dans le cas d'une pollution historique (antérieure à avril 2007), l'assainissement n'est pas obligatoire si on arrive à prouver que cette pollution ne présente pas de menace pour la santé humaine. Par contre, dans le cas d'une pollution récente, les décisions d'assainissement s'appuient sur les dépassements de valeurs normatives (VR, VS, VI).

La méthodologie générale d'approche d'un site ou d'un terrain quel qu'il soit consiste en :

- une étude d'orientation réalisée sur base d'historique et d'investigations analytiques « légères » ;
- une étude de caractérisation pour déterminer l'ampleur de la pollution et le risque potentiel que cela constitue, si l'étude d'orientation conclut à une pollution avérée ;
- une phase d'assainissement ;
- le suivi du site après « travaux », le cas échéant.

5.2.2 Application de la démarche dans le contexte wallon

L'utilisation de la démarche du projet REPERAGE, c'est-à-dire des mesures in situ de contaminants couplées à un traitement géostatistique des données peut s'inscrire, dans la démarche de gestion des sols en Wallonie à différents niveaux :

- Lors de l'étude d'orientation pour se faire une idée rapide du niveau de concentration, de sa répartition et de l'erreur associée à cette répartition.
- De la même façon, lors de l'étude de caractérisation où là il faut affiner la localisation des polluants et estimer les volumes impactés.
- Lors de l'information du public. Les cartes de répartition de la contamination pourraient s'avérer être de bons outils de communication.

6 ASPECTS FINANCIERS

6.1 Coûts de la campagne de reconnaissance avec mesures FPXRF sur MétalEurop

Le coût total de la campagne de reconnaissance menée sur le cas d'étude MétalEurop pour constituer un état de référence de la pollution des sols est de **35907 € HT**. Cette campagne a été réalisée sans outil de forage mécanique (utilisation de pelle et truelle pour creuser les fouilles), avec de grandes distances à parcourir entre chaque point de mesure (nécessité d'utiliser un véhicule). La campagne a été menée sans mettre en œuvre la démarche de couplage entre mesures sur site et traitement géostatistique. Le coût se décompose en différents postes précisés dans le **Tableau 9**.

Tableau 9 Coûts effectifs de la campagne de reconnaissance menée sur le cas d'étude MétalEurop pour constituer un état de référence (sans couplage avec un traitement géostatistique). Unité : € HT

	Personnel	DGPS	FPXRF	Laboratoire	Divers	SOUS-TOTAUX
Préparation de la campagne	4400					4400
Géoréférencement de 100 points	1500	1857				
Prélèvement, conditionnement et transport de 115 échantillons de sol	6000	6000			2000	20007
Réalisation de 513 mesures sur site			2650			
Réalisation de 115 analyses ICP (5 métaux)				11500		11500
TOTAL						35907

6.2 Coûts de la campagne de reconnaissance mettant en œuvre la démarche de couplage mesures FPXRF – analyse géostatistique sur Henricot II

Le coût total de la campagne de reconnaissance menée sur le cas d'étude Henricot II en se fondant sur la démarche de couplage des mesures sur site et du traitement géostatistique est de **38305 € HT**. Ce coût inclut la réalisation de 65 sondages de 2 m de profondeur à l'aide d'une foreuse mécanique ainsi que le travail des géostatisticiens. Le coût de la campagne de reconnaissance est de **30605 € HT** si l'on exclut le travail des géostatisticiens. La décomposition du coût total de la campagne se trouve dans le **Tableau 10**.

Tableau 10 Coûts effectifs de la campagne de reconnaissance menée sur le cas d'étude Henricot II avec la démarche de couplage des mesures sur site et du traitement géostatistique. Unité : € HT

	Personnel Terrain	Personnel Modélisation	DGPS	FPXRF	Forages	Labo	Divers	SOUS-TOTAUX
Préparation de la campagne	3000	1400						4400
Etape préliminaire	Géoréférencement de 18 sondages	300	371					5401
	Réalisation de 18 sondages 2 m				1400			
	Prélèvement, conditionnement et transport de 52 échantillons de sol	1800					1000	
	Réalisation de 52 mesures sur site			530				5200
	Réalisation de 52 analyses ICP (5 métaux)					5200		
	Traitement géostatistique		2800					2800
Etapas itératives	Géoréférencement de 47 sondages	750	929					17004
	Réalisation de 47 sondages 2 m				7000			
	Prélèvement, conditionnement et transport de 216 échantillons de sol	6000					1000	
	Réalisation de 216 mesures sur site			1325				3500
	Traitement géostatistique		3500					
TOTAL								38305
TOTAL sans géostatistique								30605

7 CONCLUSIONS ET PERSPECTIVES

Le projet de recherche REPERAGE, débuté en décembre 2006 et finalisé en janvier 2008, a consisté à développer une démarche de caractérisation des sites pollués faisant le couplage entre mesures sur site et traitements géostatistiques. La démarche permet d'intégrer des mesures sur site au fur et à mesure de leur acquisition sur le terrain dans un modèle de répartition spatiale de la pollution, et de proposer un positionnement optimisé des points de mesure dans les zones apparaissant comme « incertaines » dans le modèle, c'est-à-dire là où la classification des sols par rapport à des seuils réglementaires (ex. seuils de teneurs, seuils de risque), n'apparaît pas suffisamment fiable (probabilité trop faible de dépasser les seuils). On obtient ainsi en un temps très court une cartographie optimisée de la classification des sols et une estimation des volumes et tonnages de sols associés, le tout assorti d'une précision d'estimation.

Dans le détail, le travail de recherche a consisté à :

1. Développer une méthodologie de couplage réaliste, qui tienne compte des contraintes de coûts, de délais et de terrain d'une étude classique de site pollué, et qui soit suffisamment générale pour être adaptée à une grande variété de situations rencontrées sur les friches industrielles (2D ou 3D, mono- ou multi-polluants),
2. Développer en langage Matlab (The Mathworks™) un code spécifique réalisant en un temps très court la totalité des opérations de préparation et de traitement des données nécessaires à la modélisation géostatistique et au positionnement des nouvelles mesures d'une phase de reconnaissance à l'autre,
3. Organiser la relation entre l'équipe de terrain et l'équipe de modélisation pour que campagne de terrain et traitement des données puissent s'enchaîner sans perte de temps. L'échange rapide et efficace des informations entre le terrain et le bureau est rendu possible grâce à l'utilisation des capacités de stockage des données des outils de mesure sur site et des DGPS, et par l'emploi des moyens de communication actuels (ordinateur portable, courriels, internet).
4. Tester et valider la démarche de couplage sur un premier cas d'étude (20 ha, 2D, mono-polluant) avant mise en application en temps réel. Les environs de l'ancienne usine MétalEurop (Nord de la France) ont été choisis pour constituer des cartographies de référence de la contamination au plomb à partir d'un grand nombre de mesures de terrain et de laboratoire et d'une modélisation géostatistique (cosimulation séquentielle gaussienne des teneurs en Pb mesurées au laboratoire et à l'aide d'un FPXRF). Ces cartographies de référence ont été utilisées pour mettre en œuvre virtuellement la démarche de couplage et vérifier ses performances par confrontation de la modélisation obtenue en fin de reconnaissance virtuelle à la carte de référence des analyses de laboratoire.

5. Mettre en œuvre la démarche « en temps réel » sur un deuxième cas d'étude (5.6 ha, 3D, multi-polluants) ayant des caractéristiques très différentes du premier. Il s'agit de l'ancienne usine sidérurgique Henricot II (Wallonie) dont les remblais présentent sur une épaisseur de 2 m des teneurs en Cu, Ni, Pb et Zn qui dépassent les valeurs-seuils de référence.
6. Etudier les conditions d'intégration de la démarche dans les méthodologies française et wallonne de gestion des sites pollués.

La démarche développée comporte une étape préliminaire où des mesures sur site et des analyses de laboratoire sont réalisées sur un ensemble d'échantillons répartis de manière assez régulière dans le domaine d'étude. Ces données sont utilisées pour établir des lois de régression entre mesures sur site et analyses de laboratoire et des modèles de covariogramme nécessaires à la modélisation géostatistique, puis pour proposer un positionnement optimisé des points de mesures de l'étape de reconnaissance suivante. L'étape préliminaire est suivie d'une série d'étapes de reconnaissance itératives où des mesures sur site sont positionnées dans les zones classées comme « incertaines » et utilisées pour mettre à jour le modèle de classification des sols. La reconnaissance est arrêtée lorsque le critère d'arrêt (coût maximal, durée maximale de la reconnaissance, niveau de précision recherché) est atteint.

7.1 Principaux résultats

Le projet REPERAGE permet de montrer, sur sites réels et dans des conditions réalistes tant financières que temporelles, dans quelles conditions il est possible de coupler des mesures sur site et une modélisation géostatistique pour obtenir, en un temps record et à des coûts concurrentiels, une cartographie fine de l'état de contamination des sols et des volumes et tonnages associés. Dans le détail, les résultats acquis par ce projet sont les suivants :

- La démarche permet d'acquérir un grand nombre de mesures sur site en un temps très court tout en optimisant leur positionnement. A titre d'illustration, sur le cas d'étude MétalEurop, on estime à 248 le nombre de mesures pouvant être réalisées en 5 jours de campagne itérative. Ces données permettent de classer correctement 86% du volume du domaine d'étude dans les catégories « pollué » ou « sain » ; sur le cas d'étude Henricot II, 47 sondages et 216 mesures FPXRF ont été effectivement réalisés en 5 jours continus de campagne, et 65 sondages, 52 analyses de laboratoire et 268 mesures FPXRF ont été réalisés en 6 jours de campagne si l'on y inclut l'étape de reconnaissance préliminaire. Ces données permettent de classer 89% du volume du domaine d'étude dans les catégories « pollué » ou « sain » avec un niveau de confiance $\geq 55\%$.
- Le coût de la campagne de reconnaissance menée sur le cas d'étude MétalEurop en vue de constituer un état de référence de la pollution (115 analyses de laboratoire, 513 mesures FPXRF) est de 35907 € HT. Le coût de la campagne de reconnaissance menée sur le cas d'étude Henricot II en mettant en œuvre la démarche de couplage « en temps réel », avec modélisation géostatistique (52 analyses de laboratoire, 268 mesures FPXRF, 65 sondages de 2 m), est égal à 38305 € HT. Le travail de modélisation géostatistique représente 20% du coût global (7700 € HT).

- La démarche permet d'acquérir une bonne connaissance de la répartition spatiale de la pollution dans les sols. Cet apport est très net sur le cas d'étude Henricot II. La contamination révèle sa répartition en multiples « spots » de petite taille au fur et à mesure de l'acquisition des données. L'apport est moins frappant sur le cas d'étude MétalEurop, car dès le début de la campagne de reconnaissance, la tache centrale de grande dimension des sols classés comme « pollués » est reconnue. Les données collectées progressivement servent alors à préciser son pourtour, fortement digité. Dans les deux cas, la reconnaissance se passe en deux temps : le premier temps revient à détecter les taches de pollution, tandis que le deuxième permet de délimiter et d'évaluer leur extension. La détection des taches de pollution est assurée par l'étape préliminaire pour le cas d'étude MétalEurop, et par l'ensemble de l'étape préliminaire et des trois premiers jours de reconnaissance itérative pour le cas d'étude Henricot II (après adaptation de l'algorithme de positionnement des points de mesure). La délimitation des taches de pollution est assurée par les étapes de reconnaissance itératives (en appliquant l'algorithme original de positionnement des points de mesure).
- L'hétérogénéité de répartition de la pollution peut être très différente d'un site à l'autre, selon l'origine et l'évolution de la contamination. Les deux sites d'essai en témoignent : le cas d'étude MétalEurop est caractéristique d'un phénomène très continu, avec des hétérogénéités locales liées à l'activité humaine ; le cas d'étude Henricot II est typique d'un phénomène très hétérogène, peu structuré. Les tests réalisés dans le cadre du projet REPERAGE mettent en évidence la nécessité d'adapter les exigences en termes de précision d'estimation à la situation rencontrée. Ainsi, il est possible d'atteindre un niveau de précision $\geq 80\%$ sur la cartographie et les volumes de sol du cas d'étude MétalEurop, alors qu'il est illusoire de vouloir atteindre un tel niveau sauf effort démesuré de reconnaissance sur le cas Henricot II. Le niveau de précision recherché a été finalement fixé à $\geq 55\%$ pour ce dernier site.
- La réduction des risques d'erreurs de classification des sols au fur et à mesure de l'acquisition de données est vérifiée sur le cas d'étude MétalEurop. Elle ne peut en revanche être clairement vérifiée sur le cas d'étude Henricot II, puisque les niveaux de confiance recherchés ont été revus à la baisse au fur et à mesure de l'acquisition des données lorsque le phénomène de pollution s'est révélé fortement hétérogène.
- La démarche de couplage de mesures sur site et d'une analyse géostatistique peut être appliquée à différents stades des méthodologies française et wallonne de gestion des sites pollués : en France, la démarche permettrait de réaliser des **diagnostics de site** en des temps plus courts que ne le permettent des campagnes classiques avec analyses des sols au laboratoire. Lors **du plan de gestion**, une telle démarche permettrait d'estimer les volumes de terres contaminées, et par suite les coûts de réhabilitation. Les méthodes géostatistiques offrent en outre la possibilité de **quantifier les incertitudes** et d'en avoir une représentation spatialisée. Enfin, les cartographies issues de la démarche apparaissent intéressantes en tant qu'**outil de communication**. En Wallonie, la démarche de couplage pourrait être intégrée à une **étude d'orientation** ou à une

étude de caractérisation. Elle apparaît également comme un bon moyen pour faciliter la **communication** entre les parties prenantes.

7.2 Retours d'expérience

Le développement de la démarche et son essai de manière fictive ou réelle sur deux cas d'étude ont montré que certaines conditions pratiques doivent être remplies pour que la démarche soit mise en œuvre correctement.

- La campagne de terrain et les conditions de traitement des données doivent être soigneusement préparées pour limiter les imprévus.
- Les données doivent être collectées, sauvegardées et transmises dans des délais très courts au modélisateur. L'équipe de terrain doit donc avoir la possibilité de transmettre les données par internet de manière journalière (ordinateur portable, connexion internet).
- Une bonne coordination entre les opérateurs de terrain et les géostatisticiens est essentielle. Les deux parties doivent pouvoir se joindre nuit et jour par téléphone. Il est conseillé pour cela de munir l'ensemble de l'équipe de téléphones portables pendant toute la durée de la campagne itérative.
- Même si le code de calcul permet de proposer un positionnement optimisé des points de mesure, il est important de laisser la possibilité à l'opérateur de terrain de valoriser le caractère immédiat de la réponse de l'appareil de mesures, en adaptant une partie de la reconnaissance aux résultats trouvés en cours de campagne (ex. prolonger un sondage lorsque des teneurs élevées sont mesurées en profondeur). Nous suggérons par exemple de laisser la possibilité à l'opérateur de terrain de positionner au jugé 10 à 20% des mesures journalières.
- Pour que le traitement géostatistique des données puisse être réalisé en quelques heures, les différentes étapes de calcul sont automatisées au maximum dans le code de calcul, à l'exception de celles qui nécessitent l'expertise des modélisateurs.

7.3 Perspectives

Le projet REPERAGE aboutit à une démarche et à un code de calcul qui sont directement opérationnels. La suite logique à donner à ce travail consiste donc à mettre en œuvre la démarche sur d'autres sites dans des conditions réelles.

Quelques points pourraient cependant mériter un approfondissement.

- La démarche est applicable en théorie à tout type de mesures sur site. Dans la pratique, il est cependant nécessaire d'adapter le mode de comparaison entre mesures sur site et analyses de laboratoire au type de méthodes et d'appareils. La démarche pourrait ainsi être testée avec des appareils de terrain permettant de caractériser des pollutions organiques (ex. spectrométrie de Raman et chromatographie en phase gazeuse couplée à la spectrométrie de masse).

- La démarche a été testée et validée sur deux cas d'étude dans un contexte de dépollution (classification des sols par rapport à des seuils de teneurs). La démarche pourrait être également testée dans un cadre d'évaluation des risques (classification des sols par rapport à un seuil de risque), en utilisant les données du projet REPERAGE. Une cartographie de référence des indices de risque par ingestion de sol ou de poussières dans un scénario chronique a été constituée dans cette optique pour le cas d'étude MétalEurop, mais sans pouvoir aller plus loin dans les temps impartis au projet REPERAGE¹⁸.
- Les cartographies de référence des teneurs en plomb établies pour le cas d'étude MétalEurop pourraient encore être exploitées pour comparer d'autres scénarios de reconnaissance que ceux qui ont été présentés ici.

¹⁸ *L'évaluation de la démarche dans un cadre risques sanitaires est hors cadre de la convention de recherche REPERAGE.*

Annexe 1 – Contenu du CD-Rom

Nom du fichier ou répertoire	Description du contenu du fichier ou répertoire
Rapfinal_reperage_janv08.pdf	Fichier du rapport final REPERAGE
MetalEurop	Répertoire contenant tous les fichiers produits lors de la mise en œuvre de la démarche sur le cas d'étude MétalEurop : données brutes, figures, tableaux, animations.
Henricot	Répertoire contenant tous les fichiers produits lors de la mise en œuvre de la démarche sur le cas d'étude Henricot II : données brutes, figures, tableaux, animations.

Annexe 2 – Références bibliographiques

Burguet N. Expertise post-accidentelle : appareil de mesures in situ et recensement des laboratoires d'intervention en France. Rapport de stage INERIS / Université d'Orsay; 2008, 36 pp.

Chilès, J.P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*; John Wiley & Sons: New York, 1999.

Chilès J.P. ; Demougeot-Renard H. ; Fouquet De C. ; D'Or D. ; Garcia M., Le Loc'h G. ; Jeannée N. ; Peraudin, J.J. *Géostatistique appliquée aux sites et sols pollués. Manuel méthodologique et exemples d'application*. Document GeoSiPol disponible sous <http://www.geosipol.org>; 2005.

Demougeot-Renard, H.; De Fouquet, C. Geostatistical approach for assessing soil volumes requiring remediation: Validation using lead-polluted soils underlying a former smelting works. *Environ. Sci. Technol.*, **2004**, *38*, 5120–5126.

Demougeot-Renard, H.; Fouquet, C.; Renard, P. Forecasting the number of soil samples required to reduce remediation cost uncertainty. *Journal of Environmental Quality*, **2004**, *33*, 1694–1702.

Demougeot-Renard, H. ; Steyer, E. Reconnaissance assistée de sites pollués par utilisation conjointe de mesures de terrain et de traitements géostatistiques – partie 1. Rapport final OUTRAGE ; Rapport CNRSSP/07/01 et FSSADEME2007001; 2007.

D'Or, D. Towards a real-time multi-phase sampling strategy optimization. In *Proceedings of the Fifth European Conference on Geostatistics for Environmental Applications*; Renard, P., Demougeot-Renard, H., Froidevaux, R., Eds; Springer, 2005.

D'Or, D.; Demougeot-Renard, H.; Garcia, M. Geostatistics for contaminated sites and soils: some pending questions. In *geoEnV VI – Geostatistics for Environmental Applications*; Soares, M. P. A., Dimitrakopoulos, R., Eds; Quantitative geology and geostatistics (Springer), 2008.

D'Or, D.; Demougeot-Renard, H., Garcia, M. An integrated geostatistical approach for contaminated site and soil characterization. *Accepted in October 2008 for publication in Mathematical Geosciences*.

EPA. *Methods for evaluating the attainment of cleanup standards. Volume 1: Soils and solid media*. Tech. Rep. EPA/230-02-89-042, United States Environmental Protection Agency. Document disponible sous www.clu-in.org/download/stats/vol1soils.pdf; 1989.

EURACHEM. *Guide Eurachem /CITAC. Quantifier l'incertitude dans les mesures analytiques. 2^{ème} édition*. Ellision, S.L.R., Rosslein, M, Williams, A. Eds; 2000.

Gomez-Hernandez, J.J.; Srivastava, R.M. ISIM3D. An ANSI-C three dimensional multiple indicator conditional simulation program. *Computers and Geosciences*, **1994**, *16* (4), 395-440.

Haouche-Belkessam, L ; Demougeot-Renard, H. Reconnaissance assistée de sites pollués par l'utilisation jointe de mesures rapides sur site et de traitements géostatistiques – partie 2. Rapport intermédiaire REPERAGE ; Rapport FSSADEME2007002 / RP1-RAP-07-02289LHA ; 2007.

INERIS. *Guide pour l'orientation des actions à mettre en œuvre autour d'un site dont les sols sont potentiellement pollués par le plomb*. Rapport INERIS-DRC-57282/DESP-R030f disponible sous www.ineris.fr; 2004.

**Annexe 3 – Descriptif des scripts constitutifs du
code Matlab de mise en œuvre de la démarche de
couplage mesures sur site - géostatistique**

ETAPE PRELIMINAIRE (ETAPE 0)

1 PRÉPARATION DES DONNÉES ET CONSTRUCTION DU DOMAINE DE MODÉLISATION Nomdusite_etape_0_prep.m

ENTRÉES

- **Fichiers des données brutes préparés par le géostatisticien**
 - o XRF_0_Nomdusite.xls
 - o DGPS_0_Nomdusite.xls
 - o labo_0_Nomdusite.xls
 - o seuils_Nomdusite.xls
 - o domaine_Nomdusite.xls
 - o image_site_Nomdusite.jpg (bmp)
- **Fonctions Matlab appelées par le programme**
 - o Reperage_build_3D_grid.m
 - o Visu3d.m
 - o Imagedusite.m

ETAPES DE CALCUL

- **Données FPXRF** réception, import et formatage + calcul de la moyenne des mesures répétées sur même échantillon
- **Données DGPS**
 - o réception, import et formatage
 - o validation de la position des sondages (par comparaison de la position mesurée à la position théorique fournie par le géostatisticien à l'opérateur de terrain)
 - o attribution des coordonnées (X,Y) des sondage à chacun des points de mesure
- **Données laboratoire** réception, import et formatage
- **Sauvegarde** données et nom des variables
- **Seuils de référence** réception, import, formatage, sauvegarde
- **Domaine d'étude et zones inaccessibles à l'échantillonnage** réception, import, formatage, et sauvegarde des limites du domaine d'étude et des zones inaccessibles à l'échantillonnage
- **Grille fine de modélisation** création, contrôle visuel, sauvegarde de la grille
- **Fonds de plan** visualisation et sauvegarde
- **Cartographie 2D et 3D des données**
 - o visualisation 2D et 3D des données sur image du site
 - o sauvegarde des figures

SORTIES

- **Données Matlab**
 - o **database_0_Nomdusite.mat** variables (polluants) et noms des variables
 - o **database_seuils_Nomdusite.mat** seuils et noms des seuils
 - o **database_domaine_Nomdusite.mat** limites et noms des limites de domaine et zones inaccessibles
 - o **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
- **Données Excel**
 - o **database_XRF_0_Nomdusite.xls** (variables) + **database_XRF_0_Nomdusite_nom.xls** (nom des variables)
 - o **database_labo_0_Nomdusite.xls** (variables) + **database_labo_0_Nomdusite_nom.xls** (nom des variables)

- **Figures Matlab**
 - o Nomdusite_fondeplan.fig
 - o Carte_Donnees_Phase_0_Nomdusite.fig
 - o Carto3D_Donnees_Phase_0_Nomdusite.Fig
 - o Imagedusite_Donnees_Phase_0_Nomdusite.fig
- **Figures p/ rapports, exposés**
 - o Nomdusite_fondeplan.png
 - o Carte_Donnees_Phase_0_Nomdusite.png
 - o Carto3D_Donnees_Phase_0_Nomdusite.png
 - o Imagedusite_Donnees_Phase_0_Nomdusite.png

2 ANALYSE EXPLORATOIRE DES DONNEES **Nomdusite_etape_0_explo.m**

Ensemble d'outils de calcul et de représentation graphique utiles à une première analyse statistique des données + transformation gaussienne des distributions (si dissymétriques) + loi de régression avec résidu entre données FPXRF et labo

ENTRÉES

- **Données Matlab**
 - o **database_0_Nomdusite.mat** variables et noms des variables
 - o **database_seuils_Nomdusite.mat** seuils et noms des seuils
- **Fonctions Matlab appelées par le programme**
 - o Statelem.m
 - o Nst_transform.m
 - o Correlation.m
 - o Regres.m

ÉTAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Chargement des données + choix des polluants (optionnel)**
- **Statistiques élémentaires des données FPXRF + sauvegardes**
- **Statistiques élémentaires des données labo + sauvegardes**
- **Histogrammes des données FPXRF avec visualisation des seuils par polluant sauvegardes + figures**
- **Proportions de données FPXRF supérieures aux seuils par polluant sauvegardes**
- **Histogrammes des données labo avec visualisation des seuils par polluant sauvegardes + figures**
- **Proportions de données labo supérieures aux seuils par polluant sauvegardes**
- **Cartographie 2D des données FPXRF** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 3D des données FPXRF** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 2D des données labo** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 3D des données labo** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Variation des données FPXRF sur la verticale** par polluant – sauvegarde des figures
- **Variation des données labo sur la verticale** par polluant – sauvegarde des figures
- **Transformation des données FPXRF et des seuils par NST (Normal Score Transform)**
 - o Calculs par polluant
 - o Histogrammes des données et seuils transformés (vérification)
 - o Sauvegarde des données
- **Transformation des données labo et des seuils par NST**
 - o Calculs par polluant

- Histogrammes des données et seuils transformés (vérification)
- Sauvegarde des données
- **Corrélations entre données brutes FPXRF et labo par polluant** Calcul des corrélations de Pearson et de Spearman (rang) + sauvegarde
- **Loi de régression robuste entre données brutes FPXRF et labo par polluant** figures + sauvegarde
- **Figures de comparaison entre loi de régression robuste et par moindres carrés**
- **Calcul des std des résidus par classe de teneurs FPXRF et modélisation de la relation entre std et les teneurs par polluant**
 - Calcul des std des résidus par classe de teneurs sur données disponibles
 - Nuage de corrélation entre moyennes des teneurs par classe et std des résidus
 - Ajustement d'une fonction polynômiale et extrapolation
 - Sauvegarde des paramètres de la fonction et des figures
- **Corrélations entre données transformées NST FPXRF et labo par polluant** Calcul des corrélations de Pearson et de Spearman (rang) + sauvegarde
- **Loi de régression robuste entre données transformées NST FPXRF et labo par polluant** figures + sauvegarde
- **Corrélations entre polluants sur données brutes FPXRF** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données transformées NST FPXRF** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données brutes labo** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données transformées NST labo** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde

SORTIES

- **Données Matlab**
 - **anaexplo_0_Nomdusite.mat** statistiques, proportions des valeurs sup seuils, données transformées NST, corrélations FPXRF – labo, paramètres de la loi de régression entre données brutes, paramètres de la fonction polynômiale modélisant la relation std résidus et teneurs en polluant, corrélations FPXRF transformées NST – labo transformées NST, paramètres de la loi de régression entre données transformées NST, corrélations entre polluants FPXRF brutes, corrélations entre polluants FPXRF NST transformées, corrélations entre polluants labo brutes, corrélations entre polluants labo NST transformées
- **Données Excel (plusieurs feuilles par fichier)**
 - **stat_XRF_0_Nomdusite.xls** statistiques des données FPXRF, proportions des valeurs sup seuils
 - **stat_lab0_0_Nomdusite.xls** statistiques des données labo, proportions des valeurs sup seuils
 - **correlation_XRFlabo_0_Nomdusite.xls** corrélations entre données brutes
 - **correlation_XRFlabo_nst_0_Nomdusite.xls** corrélations entre données transformées NST
 - **correlation_XRF_0_Nomdusite.xls** corrélations entre polluants FPXRF données brutes, FPXRF données NST transformées
 - **correlation_lab0_0_Nomdusite.xls** corrélations entre polluants labo données brutes, labo données NST transformées
- **Figures Matlab**
 - **Hist_nomdupolluant_XRF_0_Nomdusite.fig** histogramme des données brutes FPXRF et seuils
 - **Hist_nomdupolluant_lab0_0_Nomdusite.fig** histogramme des données brutes labo et seuils
 - **Carto2D_XRF_nomdupolluant_0_Nomdusite.fig**
 - **Carto3D_XRF_nomdupolluant_0_Nomdusite.fig**
 - **Carto2D_lab0_nomdupolluant_0_Nomdusite.fig**
 - **Carto3D_lab0_nomdupolluant_0_Nomdusite.fig**
 - **Vertic_XRF_nomdupolluant_0_Nomdusite.fig**
 - **Vertic_lab0_nomdupolluant_0_Nomdusite.fig**
 - **Hist_NST_nomdupolluant_XRF_0_Nomdusite.fig**

- Hist_NST_nomdupolluant_lab0_0_Nomdusite.fig
 - Regres_nomdupolluant_XRFlabo_0_Nomdusite.fig
 - Stdres_XRF_nomdupolluant_XRFlabo_0_Nomdusite.fig
 - Regres_nomdupolluant_XRFlabo_nst_0_Nomdusite.fig
 - Matrice_correl_pearson_XRF_0_Nomdusite.fig
 - Matrice_correl_spearman_XRF_0_Nomdusite.fig
 - Matrice_correl_pearson_XRF_nst_0_Nomdusite.fig
 - Matrice_correl_spearman_XRF_nst_0_Nomdusite.fig
 - Matrice_correl_pearson_lab0_0_Nomdusite.fig
 - Matrice_correl_spearman_lab0_0_Nomdusite.fig
 - Matrice_correl_pearson_lab0_nst_0_Nomdusite.fig
 - Matrice_correl_spearman_lab0_nst_0_Nomdusite.fig
- **Figures p/ rapports, exposés**
- Hist_nomdupolluant_XRF_0_Nomdusite.png histogramme des données brutes FPXRF et seuils
 - Hist_nomdupolluant_lab0_0_Nomdusite.png histogramme des données brutes labo et seuils
 - Carto2D_XRF_nomdupolluant_0_Nomdusite.png
 - Carto3D_XRF_nomdupolluant_0_Nomdusite.png
 - Carto2D_lab0_nomdupolluant_0_Nomdusite.png
 - Carto3D_lab0_nomdupolluant_0_Nomdusite.png
 - Vertic_XRF_nomdupolluant_0_Nomdusite.png
 - Vertic_lab0_nomdupolluant_0_Nomdusite.png
 - Hist_NST_nomdupolluant_XRF_0_Nomdusite.png
 - Hist_NST_nomdupolluant_lab0_0_Nomdusite.png
 - Regres_nomdupolluant_XRFlabo_0_Nomdusite.png
 - Stdres_XRF_nomdupolluant_XRFlabo_0_Nomdusite.png
 - Regres_nomdupolluant_XRFlabo_nst_0_Nomdusite.png
 - Matrice_correl_pearson_XRF_0_Nomdusite.png
 - Matrice_correl_spearman_XRF_0_Nomdusite.png
 - Matrice_correl_pearson_XRF_nst_0_Nomdusite.png
 - Matrice_correl_spearman_XRF_nst_0_Nomdusite.png
 - Matrice_correl_pearson_lab0_0_Nomdusite.png
 - Matrice_correl_spearman_lab0_0_Nomdusite.png
 - Matrice_correl_pearson_lab0_nst_0_Nomdusite.png
 - Matrice_correl_spearman_lab0_nst_0_Nomdusite.png

3 ANALYSE VARIOGRAPHIQUE **Nomdusite_etape_0_vario.m**

ENTRÉES

- **Données Matlab**
 - **anaexplo_0_Nomdusite.mat** données FPXRF et labo NST transformées (selon choix des variables à analyser)
 - **database_0_Nomdusite.mat** coordonnées des données et éventuellement données brutes (selon choix des variables à analyser)
 - **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
 - **anavario_0_Henridot.mat** paramètres des fonctions géostatistiques constitutives des modèles de (co-) variogrammes
- **Fonctions Matlab appelées par le programme**
 - Vario3D.m fonction de la BMElib (startup)
 - Modelfit.m BMElib
 - Modelplot.m BMElib
 - Coregfit.m BMElib

ETAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Analyse variographique des données FPXRF non corrélées**
 - o Calcul des variogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes
 - o Figures + sauvegardes
- **Analyse variographique des données FPXRF corrélées**
 - o Calcul des variogrammes et covariogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes et covariogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes et covariogrammes
 - o Figures + sauvegardes
- **Analyse variographique des données labo non corrélées**
 - o Calcul des variogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes
 - o Figures + sauvegardes
- **Analyse variographique des données labo corrélées**
 - o Calcul des variogrammes et covariogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes et covariogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes et covariogrammes
 - o Figures + sauvegardes
- **Analyse variographique des données FPXRF et labo**
 - o Calcul des variogrammes et covariogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes et covariogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes et covariogrammes
 - o Figures + sauvegardes

SORTIES

- **Données Matlab**
 - o **anavario_0_Nomdusite.mat** variogrammes expérimentaux (valeur en fonction de la distance de séparation des points de données, nombre de couples de points), paramètres des modèles de variogrammes (type de modèles élémentaires, portées, paliers)
- **Figures Matlab**
 - o Vario_exp_omni_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_exp_horiz_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_exp_horizsn_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_exp_vert_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_model_omni_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_model_horiz_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_model_horizsn_nomdupolluant_XRF_0_Nomdusite.fig
 - o Vario_model_vert_nomdupolluant_XRF_0_Nomdusite.fig
 - o etc...
- **Figures p/ rapports, exposés**
 - o Vario_exp_omni_nomdupolluant_XRF_0_Nomdusite.png
 - o Vario_exp_horiz_nomdupolluant_XRF_0_Nomdusite.png
 - o Vario_exp_horizsn_nomdupolluant_XRF_0_Nomdusite.png
 - o Vario_exp_vert_nomdupolluant_XRF_0_Nomdusite.png
 - o Vario_model_omni_nomdupolluant_XRF_0_Nomdusite.png
 - o Vario_model_horiz_nomdupolluant_XRF_0_Nomdusite.png

- Vario_model_horizhsn_nomdupolluant_XRF_0_Nomdusite.png
- Vario_model_vert_nomdupolluant_XRF_0_Nomdusite.png
- etc....

4 SIMULATION DES TENEURS EN POLLUANTS **Nomdusite_etape_0_simu.m**

Génération de simulations des teneurs en polluants à partir du programme gcosim3d

ENTRÉES

- **Fichiers Matlab issus de **Nomdusite_etape_0_prep.m** et **Nomdusite_etape_0_explo.m****
 - **database_0_Nomdusite.mat** données brutes (selon choix des variables à analyser)
 - **anaexplo_0_Nomdusite.mat** données NST transformées (selon choix des variables à analyser)
 - **anavario_0_Nomdusite.mat** paramètres des covariogrammes des mesures sur site et analyses labo des polluants d'intérêt
 - **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
- **Fonctions Matlab appelées dans le programme**
 - Exportvariogcosim3d_Nomdusite.m
 - Exportgeogcosim3d.m
 - Visu3d.m

ETAPES DE CALCUL

- **Export des données conditionnantes (constitution du fichier gcosim3d.dat)**
 - Vérification de la cohérence entre variables des données conditionnantes et des variogrammes
 - Modification des coordonnées de la grille de modélisation et des données (translation)
 - Création du fichier gcosim3d.dat + sauvegarde
- **Export des modèles de (co-)variogrammes ((constitution du fichier gcosim3d.var)**
 - Ecriture et sauvegarde du fichier gcosim3d.var et sauvegarde dans un fichier Matlab
- **Export des conditions de modélisation (constitution du fichier gcosim3d.geo)**
 - Ecriture et sauvegarde du fichier gcosim3d.geo et sauvegarde dans un fichier Matlab
- **Exécution de gcosim3d.exe**
- **Import et sauvegarde des résultats de simulation dans Matlab**
- **Backtransformation des résultats de simulation + sauvegarde**
- **Vérification des résultats**
 - Visualisation d'une ou plusieurs simulations.
 - Visualisation de la moyenne par maille des valeurs simulées

SORTIES

- **Fichiers d'entrée de gcosim3d**
 - Gcosim3d.dat, gcosim3d.var, gcosim3d.geo
- **Données Matlab**
 - Export_vario_gcosim3d_0_Nomdusite.mat paramètres précisant les conditions de simulation
 - Export_geo_gcosim3d_0_Nomdusite.mat paramètres précisant les conditions de simulation
 - Simu_Nomdusite_0_nomdupolluant simulations des variables brutes et NST transformées

5 POST-TRAITEMENT DES SIMULATIONS **Nomdusite_etape_0_postraitement.m**

Traitement des simulations pour calculer la zone incertaine

ENTRÉES

- **Données Matlab**
 - o **database_seuils_Nomdusite.mat** seuils
 - o **database_domaine_Nomdusite.mat** limite de site et de zones inaccessibles à l'échantillonnage
 - o **database_0_Nomdusite** nom des variables ayant été simulées
 - o **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
- **Fonctions Matlab appelées dans le programme**
 - o Build_zOAGrid.m
 - o Classif_simu_pol.m
 - o Soilclassification.m
 - o Cartoclassifsoil.m
 - o Visu3d.m
 - o Bicolorclassif.m
 - o Estimvolumincert.m
 - o Estimvolumincert_tricolor.m

ÉTAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Construction de la matrice des seuils** Lecture des seuils de teneurs et attribution des seuils aux nœuds de la grille
- **Classification des sols par simulation et par polluant**
- **Classification des sols par rapport à un ou plusieurs polluants**
 - o Calcul des probabilités de dépassement des seuils et des risques d'erreurs de classification
 - o Cartographie de la classification des sols
- **Classification des sols par rapport à tous les polluants**
 - o Cartographie de la classification globale (pollué – sain – coupure à proba 0.5)
 - o Cartographie des risques d'erreurs de classification
 - o **Définition des seuils de risque**
 - o **Cartographie des sols en pollué-certain, pollué-incertain, sain-certain, sain-incertain selon ces seuils de risque**
 - o **Estimation des volumes de chacune des 4 catégories**
- **Sauvegarde des résultats**

SORTIES

- **Données Matlab**
 - o **Postraitement_0_Nomdusite.mat** matrice des seuils de teneurs, indicatrices obtenues après seuillage des simulations de teneurs, variable de classification des sols par rapport à une sélection de polluants, variable de classification par rapport à toutes les variables, probabilité globale qu'une maille soit polluée, probabilité globale qu'une maille soit saine, risque d'erreur de classification à tort des sols comme pollués, risque d'erreur de classification à tort des sols comme sains, variable de classification des sols par filière (sain, pollué, incertain) une fois des seuils de risque d'erreur fixés, volumes de sol par filière
- **Tableaux Excel**
 - o VolumIncertain_0_Nomdusite.xls
- **Figures Matlab**
 - o Classif_prof_profondeurdela couche_0_Nomdusite.fig Carte de classification des sols pollué-sain par rapport à une sélection de polluants
 - o Classif_globale_profondeurdela couche_0_Nomdusite.fig Carte de classification des sols pollué-sain par rapport à tous les polluants

- Erreur_classif_prof_profondeurdela couche_0_Nomdusite.fig Carte des risques d'erreur de classification des sols classés pollués-sains (proba=0.5) par rapport à tous les polluants
- **ZoneIncertaine_profondeurdelacouche_0_Nomdusite.fig** Carte des sols classés comme sains-pollués-incertains
- **Figures p/ rapports, exposés**
 - Classif_prof_profondeurdela couche_0_Nomdusite.png Carte de classification des sols pollué-sain par rapport à une sélection de polluants
 - Classif_globale_profondeurdela couche_0_Nomdusite.png Carte de classification des sols pollué-sain par rapport à tous les polluants
 - Erreur_classif_prof_profondeurdela couche_0_Nomdusite.png Carte des risques d'erreur de classification des sols classés pollués-sains (proba=0.5) par rapport à tous les polluants
 - **ZoneIncertaine_profondeurdelacouche_0_Nomdusite.png** Carte des sols classés comme sains-pollués-incertains

6 POSITIONNEMENT DES SONDAGES ET DES ECHANTILLONS DE L'ETAPE SUIVANTE **Nomdusite_etape_0_echage_zonincertaine_sondages.m**

Positionnement des sondages et des échantillons de sol dans la zone incertaine. On positionne les sondages selon un maillage rectangulaire régulier dans la zone incertaine, en respectant l'anisotropie des portées de corrélation. Les échantillons sont localisés à intervalles réguliers et systématiques dans les sondages. Le principe est de construire une grille d'échantillonnage que l'on superpose à la grille fine de calcul, dont les mailles reflètent l'anisotropie des variogrammes, et de sélectionner les mailles de cette grille d'échantillonnage qui renferment le plus grand nombre de mailles classées incertaines de la grille fine. On sélectionne autant de mailles que de sondages nécessaires à la prochaine étape de reconnaissance, et on positionne le sondage au centre de la maille.

ENTRÉES

- **Données Matlab**
 - **compteurs_Nomdusite.mat** compteurs du nombre de mesures, nombre de sondages, etc...mis à jour au fur et à mesure des étapes de reconnaissance
 - **database_0_Nomdusite** coordonnées des variables, nom du site
 - **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
 - **postraitement_0_Nomdusite** variable de classification des sols pollué-sain-incertain
 - **anavario_0_Nomdusite.mat** portées de corrélation des modèles de variogramme
- **Fonctions Matlab appelées dans le programme**
 - Visu3d.m

ÉTAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Mise à jour des compteurs et adaptation du nombre de mesures**
 - Calcul du nombre d'échantillons (et de mesures XRF) réalisées, y compris cette nouvelle étape, et comparaison au nombre maximal de mesures possible.
 - Calcul du nombre de jours de campagne, y compris cette nouvelle étape, et comparaison au nombre maximal de jours de campagne possible.
 - Calcul du coût de campagne, y compris cette nouvelle étape, et comparaison au coût maximal.
 - Le nombre de mesures est adapté (limité) pour ne pas dépasser ces valeurs maximales
- **Calcul des dimensions des mailles de la grille d'échantillonnage que l'on superpose à la grille fine de calcul**
 - **Calcul de la surface de la zone incertaine (incluant toutes les mailles incertaines de toutes les profondeurs)**
 - **Calcul de la surface horizontale de la maille associée à tout sondage complémentaire**
 - **Calcul des dimensions horizontales de cette maille en fonction de l'ellipse des portées de corrélation.**
 - **La dimension verticale de la maille associée à chaque échantillon dépend de l'espacement systématique fixé sur la verticale pour ces échantillons/mesures XRF**

- Création d'une grille d'échantillonnage sur tout le domaine d'étude
- Détermination du nombre de mailles à échantillonner de la grille d'échantillonnage
- Calcul des coordonnées des échantillons de tous les sondages
- Attribution d'un identifiant aux échantillons
- Cartographie de la position des nouveaux échantillons proposés
 - o 3D superposée à la classification des sols pollué-sain-incertain
 - o 2D par couche superposée à la classification des sols pollué-sain-incertain
 - o 2D par couche superposée aux fonds de plan
- Cartographie de la position des nouveaux sondages proposés avec leur identifiant
- Tableau destiné à l'opérateur de terrain avec
 - o Identifiant des échantillons, coordonnées (x,y,z) des échantillons

SORTIES

- **Données Matlab**
 - o **compteurs_Nomdusite.mat** nombre de mesures, nombre de jours de campagne, coût de la campagne mis à jour
 - o **echagezincertaine_0_Nomdusite.mat** nombre d'échantillons (mesures XRF), nombre de sondages de l'étape de campagne suivante, surface de la zone incertaine, surface d'une maille de la grille d'échantillonnage, dimensions de cette maille, caractéristiques de la grille fine de calcul, caractéristiques du centre de gravité de la grille fine, caractéristiques de la grille grossière d'échantillonnage (origine, angle par rapport au système de coordonnées choisi, dimension des mailles, nombre de mailles), **coordonnées des nouveaux sondages, coordonnées des nouveaux échantillons, identifiant des échantillons**
- **Tableaux Excel**
 - o **Echagepropose_phase_1_Nomdusite.xls** identifiants et coordonnées des nouveaux échantillons
- **Figures Matlab**
 - o Echagepropose_3D_phase_1_Nomdusite.fig
 - o Echagepropose_zincert_profondeurdelacouche_phase_1_Nomdusite.fig
 - o Echagepropose_profondeurdelacouche_phase_1_Nomdusite.fig fond de plan
 - o Sondagepropose_phase_1_Nomdusite.fig
 - o Nomdusite_echagepreliminaire_avecnumeros.fig
- **Figures p/ rapports, exposés**
 - o Echagepropose_3D_phase_1_Nomdusite.png
 - o Echagepropose_zincert_profondeurdelacouche_phase_1_Nomdusite.png
 - o Echagepropose_profondeurdelacouche_phase_1_Nomdusite.png fond de plan
 - o Sondagepropose_phase_1_Nomdusite.png
 - o Nomdusite_echagepreliminaire_avecnumeros.png

ETAPE ITERATIVE (ETAPE J)

7 IMPORT ET PRÉPARATION DES DONNÉES COLLECTÉES LORS DE L'ETAPE J Nomdusite_etape_j_prep.m

ENTRÉES

- **Fichiers des données brutes envoyées par l'équipe de terrain, et préparées par le géostatisticien**
 - o XRF_j_Nomdusite.xls
 - o DGPS_j_Nomdusite.xls
- **Fonctions Matlab appelées par le programme**
 - o Visu3d.m
 - o Imagedusite.m

ETAPES DE CALCUL

- **Données FPXRF** réception, import et formatage + calcul de la moyenne des mesures répétées sur même échantillon
- **Données DGPS**
 - o réception, import et formatage
 - o validation de la position des sondages (par comparaison de la position mesurée à la position théorique fournie par le géostatisticien à l'opérateur de terrain)
 - o attribution des coordonnées (X,Y) des sondage à chacun des points de mesure
- **Génération de valeurs labo** à partir de la loi de régression établie lors de l'étape préliminaire
- **Sauvegarde** données et nom des variables
- **Cartographie 2D et 3D des données de l'étape j**
 - o visualisation 2D et 3D des données sur image du site
 - o sauvegarde des figures
- **Sauvegarde** de toutes les données collectées depuis l'étape préliminaire dans un même fichier
- **Cartographie 2D et 3D de toutes les données depuis l'étape préliminaire**
 - o visualisation des données avec symboles différents pour chaque étape
 - o sauvegarde des figures

SORTIES

- **Données Matlab**
 - o **database_j_Nomdusite.mat** variables (polluants) et noms des variables de l'étape j
 - o **datasetot_j_Nomdusite.mat** variables (polluants) et noms des variables de toutes les données collectées de l'étape 0 à l'étape j
- **Données Excel**
 - o **database_XRF_j_Nomdusite.xls** (variables) + **database_XRF_j_Nomdusite_nom.xls** (nom des variables)
 - o **database_labo_j_Nomdusite.xls** (variables) + **database_labo_j_Nomdusite_nom.xls** (nom des variables)
- **Figures Matlab**
 - o Nomdusite_fondeplan.fig
 - o Carte_Donnees_Phase_j_Nomdusite.fig
 - o Carto3D_Donnees_Phase_j_Nomdusite. fig
 - o Histo_labo_genere_Nomdupolluant_j_Nomdusite.fig Histogramme des valeurs générées
 - o SD_XRF_labo_genere_Nomdupolluant_j_Nomdusite.fig Nuage de corrélation entre mesures FPXRF et valeurs labo générées

- Carte_Toutes_Donnees_Phase_j_Nomdusite.fig Carte de toutes les données en différenciant les données des différentes étapes
-
- **Figures p/ rapports, exposés**
 - Nomdusite_fondeplan.png
 - Carte_Donnees_Phase_j_Nomdusite.png
 - Carto3D_Donnees_Phase_j_Nomdusite.png
 - Histo_labo_genere_Nomdupolluant_j_Nomdusite.png Histogramme des valeurs générées
 - SD_XRF_labo_genere_Nomdupolluant_j_Nomdusite.png Nuage de corrélation entre mesures FPXRF et valeurs labo générées
 - Carte_Toutes_Donnees_Phase_j_Nomdusite.png Carte de toutes les données en différenciant les données des différentes étapes

8 ANALYSE EXPLORATOIRE DES DONNEES **Nomdusite_etape_j_explo.m**

Ensemble d'outils de calcul et de représentation graphique utiles à une première analyse statistique des données + transformation gaussienne des distributions (si dissymétriques) + loi de régression avec résidu entre données FPXRF et labo

ENTRÉES

- **Données Matlab**
 - **datasetot_j_Nomdusite.mat** variables et noms des variables
 - **database_seuils_Nomdusite.mat** seuils et noms des seuils
- **Fonctions Matlab appelées par le programme**
 - Statelem.m
 - Nst_transform.m
 - Correlation.m
 - Regres.m

ÉTAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Chargement des données + choix des polluants (optionnel)**
- **Statistiques élémentaires des données FPXRF + sauvegardes**
- **Statistiques élémentaires des données labo + sauvegardes**
- **Histogrammes des données FPXRF avec visualisation des seuils par polluant sauvegardes + figures**
- **Proportions de données FPXRF supérieures aux seuils par polluant sauvegardes**
- **Histogrammes des données labo avec visualisation des seuils par polluant sauvegardes + figures**
- **Proportions de données labo supérieures aux seuils par polluant sauvegardes**
- **Cartographie 2D des données FPXRF** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 3D des données FPXRF** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 2D des données labo** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Cartographie 3D des données labo** échelle de couleur selon niveau de teneur – une carte par polluant – sauvegarde des figures
- **Variation des données FPXRF sur la verticale** par polluant – sauvegarde des figures
- **Variation des données labo sur la verticale** par polluant – sauvegarde des figures
- **Transformation des données FPXRF et des seuils par NST (Normal Score Transform)**
 - **Calculs par polluant**
 - **Histogrammes des données et seuils transformés (vérification)**
 - **Sauvegarde des données**

- **Transformation des données labo et des seuils par NST**
 - o Calculs par polluant
 - o Histogrammes des données et seuils transformés (vérification)
 - o Sauvegarde des données
- **Corrélations entre données brutes FPXRF et labo par polluant** Calcul des corrélations de Pearson et de Spearman (rang) + sauvegarde
- **Loi de régression robuste entre données brutes FPXRF et labo par polluant** figures + sauvegarde
- **Figures de comparaison entre loi de régression robuste et par moindres carrés**
- **Calcul des std des résidus par classe de teneurs FPXRF et modélisation de la relation entre std et les teneurs par polluant**
 - o Calcul des std des résidus par classe de teneurs sur données disponibles
 - o Nuage de corrélation entre moyennes des teneurs par classe et std des résidus
 - o Ajustement d'une fonction polynômiale et extrapolation
 - o Sauvegarde des paramètres de la fonction et des figures
- **Corrélations entre données transformées NST FPXRF et labo par polluant** Calcul des corrélations de Pearson et de Spearman (rang) + sauvegarde
- **Loi de régression robuste entre données transformées NST FPXRF et labo par polluant** figures + sauvegarde
- **Corrélations entre polluants sur données brutes FPXRF** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données transformées NST FPXRF** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données brutes labo** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde
- **Corrélations entre polluants sur données transformées NST labo** Calcul des corrélations de Pearson et de Spearman (rang) + figures des matrices de corrélation + sauvegarde

SORTIES

- **Données Matlab**
 - o **anaexplor_j_Nomdusite.mat** statistiques, proportions des valeurs sup seuils, données transformées NST, corrélations FPXRF – labo, paramètres de la loi de régression entre données brutes, paramètres de la fonction polynômiale modélisant la relation std résidus et teneurs en polluant, corrélations FPXRF transformées NST – labo transformées NST, paramètres de la loi de régression entre données transformées NST, corrélations entre polluants FPXRF brutes, corrélations entre polluants FPXRF NST transformées, corrélations entre polluants labo brutes, corrélations entre polluants labo NST transformées
- **Données Excel (plusieurs feuilles par fichier)**
 - o **stat_XRF_j_Nomdusite.xls** statistiques des données FPXRF, proportions des valeurs sup seuils
 - o **stat_labor_j_Nomdusite.xls** statistiques des données labo, proportions des valeurs sup seuils
 - o **correlation_XRFlabor_j_Nomdusite.xls** corrélations entre données brutes
 - o **correlation_XRFlabor_nst_j_Nomdusite.xls** corrélations entre données transformées NST
 - o **correlation_XRF_j_Nomdusite.xls** corrélations entre polluants FPXRF données brutes, FPXRF données NST transformées
 - o **correlation_labor_j_Nomdusite.xls** corrélations entre polluants labo données brutes, labo données NST transformées
- **Figures Matlab**
 - o **Hist_nomdupolluant_XRF_j_Nomdusite.fig** histogramme des données brutes FPXRF et seuils
 - o **Hist_nomdupolluant_labor_j_Nomdusite.fig** histogramme des données brutes labo et seuils
 - o **Carto2D_XRF_nomdupolluant_j_Nomdusite.fig**
 - o **Carto3D_XRF_nomdupolluant_j_Nomdusite.fig**
 - o **Carto2D_labor_nomdupolluant_j_Nomdusite.fig**
 - o **Carto3D_labor_nomdupolluant_j_Nomdusite.fig**
 - o **Vertic_XRF_nomdupolluant_j_Nomdusite.fig**

- Vertic_lab0_nomdupolluant_j_Nomdusite.fig
 - Hist_NST_nomdupolluant_XRF_j_Nomdusite.fig
 - Hist_NST_nomdupolluant_lab0_j_Nomdusite.fig
 - Regres_nomdupolluant_XRFlab0_j_Nomdusite.fig
 - Stdres_XRF_nomdupolluant_XRFlab0_j_Nomdusite.fig
 - Regres_nomdupolluant_XRFlab0_nst_j_Nomdusite.fig
 - Matrice_correl_pearson_XRF_j_Nomdusite.fig
 - Matrice_correl_spearman_XRF_j_Nomdusite.fig
 - Matrice_correl_pearson_XRF_nst_j_Nomdusite.fig
 - Matrice_correl_spearman_XRF_nst_j_Nomdusite.fig
 - Matrice_correl_pearson_lab0_j_Nomdusite.fig
 - Matrice_correl_spearman_lab0_j_Nomdusite.fig
 - Matrice_correl_pearson_lab0_nst_j_Nomdusite.fig
 - Matrice_correl_spearman_lab0_nst_j_Nomdusite.fig
- **Figures p/ rapports, exposés**
- Hist_nomdupolluant_XRF_j_Nomdusite.png histogramme des données brutes FPXRF et seuils
 - Hist_nomdupolluant_lab0_j_Nomdusite.png histogramme des données brutes lab0 et seuils
 - Carto2D_XRF_nomdupolluant_j_Nomdusite.png
 - Carto3D_XRF_nomdupolluant_j_Nomdusite.png
 - Carto2D_lab0_nomdupolluant_j_Nomdusite.png
 - Carto3D_lab0_nomdupolluant_j_Nomdusite.png
 - Vertic_XRF_nomdupolluant_j_Nomdusite.png
 - Vertic_lab0_nomdupolluant_j_Nomdusite.png
 - Hist_NST_nomdupolluant_XRF_j_Nomdusite.png
 - Hist_NST_nomdupolluant_lab0_j_Nomdusite.png
 - Regres_nomdupolluant_XRFlab0_j_Nomdusite.png
 - Stdres_XRF_nomdupolluant_XRFlab0_j_Nomdusite.png
 - Regres_nomdupolluant_XRFlab0_nst_j_Nomdusite.png
 - Matrice_correl_pearson_XRF_j_Nomdusite.png
 - Matrice_correl_spearman_XRF_j_Nomdusite.png
 - Matrice_correl_pearson_XRF_nst_j_Nomdusite.png
 - Matrice_correl_spearman_XRF_nst_j_Nomdusite.png
 - Matrice_correl_pearson_lab0_j_Nomdusite.png
 - Matrice_correl_spearman_lab0_j_Nomdusite.png
 - Matrice_correl_pearson_lab0_nst_j_Nomdusite.png
 - Matrice_correl_spearman_lab0_nst_j_Nomdusite.png

9 ANALYSE VARIOGRAPHIQUE **Nomdusite_etape_j_vario.m**

ENTRÉES

- **Données Matlab**
 - o **anaexplo_j_Nomdusite.mat** données FPXRF et labo NST transformées (selon choix des variables à analyser)
 - o **databasetot_j_Nomdusite.mat** coordonnées des données et éventuellement données brutes (selon choix des variables à analyser)
 - o **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
 - o **anavario_j_Henridot.mat** paramètres des fonctions gigognes constitutives des modèles de (co-) variogrammes
- **Fonctions Matlab appelées par le programme**
 - o Vario3D.m fonction de la BMElib (startup)
 - o Modelfit.m BMElib
 - o Modelplot.m BMElib
 - o Coregfit.m BMElib

ETAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Analyse variographique des données FPXRF non corrélées**
 - o Calcul des variogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes
 - o Figures + sauvegardes
- **Analyse variographique des données FPXRF corrélées**
 - o Calcul des variogrammes et covariogrammes expérimentaux directionnels horizontaux et verticaux
 - o Calcul des variogrammes et covariogrammes expérimentaux omnidirectionnels
 - o Ajustement de modèles de variogrammes et covariogrammes
 - o Figures + sauvegardes

SORTIES

- **Données Matlab**
 - o **anavario_j_Nomdusite.mat** variogrammes expérimentaux (valeur en fonction de la distance de séparation des points de données, nombre de couples de points), paramètres des modèles de variogrammes (type de modèles élémentaires, portées, paliers)
- **Figures Matlab**
 - o Vario_exp_omni_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_exp_horiz_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_exp_horizsn_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_exp_vert_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_model_omni_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_model_horizh_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_model_horizsn_nomdupolluant_XRF_j_Nomdusite.fig
 - o Vario_model_vert_nomdupolluant_XRF_j_Nomdusite.fig
 - o etc...
- **Figures p/ rapports, exposés**
 - o Vario_exp_omni_nomdupolluant_XRF_j_Nomdusite.png
 - o Vario_exp_horiz_nomdupolluant_XRF_j_Nomdusite.png
 - o Vario_exp_horizsn_nomdupolluant_XRF_j_Nomdusite.png
 - o Vario_exp_vert_nomdupolluant_XRF_j_Nomdusite.png
 - o Vario_model_omni_nomdupolluant_XRF_j_Nomdusite.png

- Vario_model_horizh_nomdupolluant_XRF_j_Nomdusite. png
- Vario_model_horizhsn_nomdupolluant_XRF_j_Nomdusite. png
- Vario_model_vert_nomdupolluant_XRF_j_Nomdusite. png
- etc....

10 SIMULATION DES TENEURS EN POLLUANTS **Nomdusite_etape_j_simu.m**

Génération de simulations des teneurs en polluants à partir du programme gcosim3d

ENTRÉES

- **Fichiers Matlab issus de **Nomdusite_etape_j_prep.m** et **Nomdusite_etape_j_explo.m****
 - **datasatot_j_Nomdusite.mat** données brutes (selon choix des variables à analyser)
 - **anaexplo_j_Nomdusite.mat** données NST transformées (selon choix des variables à analyser)
 - **anavario_0_Nomdusite.mat** paramètres des covariogrammes des mesures sur site et analyses labo des polluants d'intérêt
 - **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
- **Fonctions Matlab appelées dans le programme**
 - Exportvariogcosim3d_Nomdusite.m
 - Exportgeogcosim3d.m
 - Visu3d.m

ETAPES DE CALCUL

- **Export des données conditionnantes (constitution du fichier gcosim3d.dat)**
 - Vérification de la cohérence entre variables des données conditionnantes et des variogrammes
 - Modification des coordonnées de la grille de modélisation et des données (translation)
 - Création du fichier gcosim3d.dat + sauvegarde
- **Export des modèles de (co-)variogrammes ((constitution du fichier gcosim3d.var)**
 - Ecriture et sauvegarde du fichier gcosim3d.var et sauvegarde dans un fichier Matlab
 - Les paramètres des modèles de co-variogrammes sont ceux trouvés lors de l'étape préliminaire, avec éventuellement une adaptation des paliers en fonction des modèles de variogrammes trouvés sur les mesures FPXRF avec toutes les données, de l'étape 0 à l'étape j.
- **Export des conditions de modélisation (constitution du fichier gcosim3d.geo)**
 - Ecriture et sauvegarde du fichier gcosim3d.geo et sauvegarde dans un fichier Matlab
- **Exécution de gcosim3d.exe**
- **Import et sauvegarde des résultats de simulation dans Matlab**
- **Backtransformation des résultats de simulation + sauvegarde**
- **Vérification des résultats**
 - Visualisation d'une ou plusieurs simulations.
 - Visualisation de la moyenne par maille des valeurs simulées

SORTIES

- **Fichiers d'entrée de gcosim3d**
 - Gcosim3d.dat, gcosim3d.var, gcosim3d.geo
- **Données Matlab**
 - Export_vario_gcosim3d_j_Nomdusite.mat paramètres précisant les conditions de simulation
 - Export_geo_gcosim3d_j_Nomdusite.mat paramètres précisant les conditions de simulation
 - Simu_Nomdusite_j_nomdupolluant simulations des variables brutes et NST transformées

11 POST-TRAITEMENT DES SIMULATIONS **Nomdusite_etape_j_postraitement.m**

Traitement des simulations pour calculer la zone incertaine

ENTRÉES

- **Données Matlab**
 - o **database_seuils_Nomdusite.mat** seuils
 - o **database_domaine_Nomdusite.mat** limite de site et de zones inaccessibles à l'échantillonnage
 - o **databasetot_j_Nomdusite** nom des variables ayant été simulées
 - o **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
- **Fonctions Matlab appelées dans le programme**
 - o Build_zOAGrid.m
 - o Classif_simu_pol.m
 - o Soilclassification.m
 - o Cartoclassifsoil.m
 - o Visu3d.m
 - o Bicolorclassif.m
 - o Estimvolumincert.m
 - o Estimvolumincert_tricolor.m

ÉTAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Construction de la matrice des seuils** Lecture des seuils de teneurs et attribution des seuils aux nœuds de la grille
- **Classification des sols par simulation et par polluant**
- **Classification des sols par rapport à un ou plusieurs polluants**
 - o Calcul des probabilités de dépassement des seuils et des risques d'erreurs de classification
 - o Cartographie de la classification des sols
- **Classification des sols par rapport à tous les polluants**
 - o Cartographie de la classification globale (pollué – sain – coupure à proba j.5)
 - o Cartographie des risques d'erreurs de classification
 - o **Définition des seuils de risque**
 - o **Cartographie des sols en pollué-certain, pollué-incertain, sain-certain, sain-incertain selon ces seuils de risque**
 - o **Estimation des volumes de chacune des 4 catégories**
- **Sauvegarde des résultats**

SORTIES

- **Données Matlab**
 - o **Postraitement_j_Nomdusite.mat** matrice des seuils de teneurs, indicatrices obtenues après seuillage des simulations de teneurs, variable de classification des sols par rapport à une sélection de polluants, variable de classification par rapport à toutes les variables, probabilité globale qu'une maille soit polluée, probabilité globale qu'une maille soit saine, risque d'erreur de classification à tort des sols comme pollués, risque d'erreur de classification à tort des sols comme sains, variable de classification des sols par filière (sain, pollué, incertain) une fois des seuils de risque d'erreur fixés, volumes de sol par filière
- **Tableaux Excel**
 - o **Volumelncertain_j_Nomdusite.xls**
- **Figures Matlab**
 - o **Classif_prof_profondeurdela couche_j_Nomdusite.fig** Carte de classification des sols pollué-sain par rapport à une sélection de polluants
 - o **Classif_globale_profondeurdela couche_j_Nomdusite.fig** Carte de classification des sols pollué-sain par rapport à tous les polluants

- Erreur_classif_prof_profondeurdela couche_j_Nomdusite.fig Carte des risques d'erreur de classification des sols classés pollués-sains (proba=j.5) par rapport à tous les polluants
- **ZoneIncertaine_profondeurdelacouche_j_Nomdusite.fig** Carte des sols classés comme sains-pollués-incertains
- **Figures p/ rapports, exposés**
 - Classif_prof_profondeurdela couche_j_Nomdusite.png Carte de classification des sols pollué-sain par rapport à une sélection de polluants
 - Classif_globale_profondeurdela couche_j_Nomdusite.png Carte de classification des sols pollué-sain par rapport à tous les polluants
 - Erreur_classif_prof_profondeurdela couche_j_Nomdusite.png Carte des risques d'erreur de classification des sols classés pollués-sains (proba=j.5) par rapport à tous les polluants
 - **ZoneIncertaine_profondeurdelacouche_j_Nomdusite.png** Carte des sols classés comme sains-pollués-incertains

12 POSITIONNEMENT DES SONDAGES ET DES ECHANTILLONS DE L'ETAPE SUIVANTE

Nomdusite_etape_j_echage_zonincertaine_sondages.m

Positionnement des sondages et des échantillons de sol dans la zone incertaine. On positionne les sondages selon un maillage rectangulaire régulier dans la zone incertaine, en respectant l'anisotropie des portées de corrélation. Les échantillons sont localisés à intervalles réguliers et systématiques dans les sondages. Le principe est de construire une grille d'échantillonnage que l'on superpose à la grille fine de calcul, dont les mailles reflètent l'anisotropie des variogrammes, et de sélectionner les mailles de cette grille d'échantillonnage qui renferment le plus grand nombre de mailles classées incertaines de la grille fine. On sélectionne autant de mailles que de sondages nécessaires à la prochaine étape de reconnaissance, et on positionne le sondage au centre de la maille.

ENTRÉES

- **Données Matlab**
 - **compteurs_Nomdusite.mat** compteurs du nombre de mesures, nombre de sondages, etc...mis à jour au fur et à mesure des étapes de reconnaissance
 - **databasetot_j_Nomdusite** coordonnées des variables, nom du site
 - **defgrid3d_0_Nomdusite.mat** grille fine de modélisation
 - **postraitement_j_Nomdusite** variable de classification des sols pollué-sain-incertain
 - **anavario_0_Nomdusite.mat** portées de corrélation des modèles de variogramme (avec éventuellement adaptation des paliers en fonction des nouvelles données acquises)
- **Fonctions Matlab appelées dans le programme**
 - Visu3d.m

ETAPES DE CALCUL (NON OBLIGATOIRES OU **OBLIGATOIRES**)

- **Mise à jour des compteurs et adaptation du nombre de mesures**
 - Calcul du nombre d'échantillons (et de mesures XRF) réalisées, y compris cette nouvelle étape, et comparaison au nombre maximal de mesures possible.
 - Calcul du nombre de jours de campagne, y compris cette nouvelle étape, et comparaison au nombre maximal de jours de campagne possible.
 - Calcul du coût de campagne, y compris cette nouvelle étape, et comparaison au coût maximal.
 - Le nombre de mesures est adapté (limité) pour ne pas dépasser ces valeurs maximales
- **Calcul des dimensions des mailles de la grille d'échantillonnage que l'on superpose à la grille fine de calcul**
 - **Calcul de la surface de la zone incertaine (incluant toutes les mailles incertaines de toutes les profondeurs)**
 - **Calcul de la surface horizontale de la maille associée à tout sondage complémentaire**
 - **Calcul des dimensions horizontales de cette maille en fonction de l'ellipse des portées de corrélation.**

- La dimension verticale de la maille associée à chaque échantillon dépend de l'espacement systématique fixé sur la verticale pour ces échantillons/mesures XRF
- Création d'une grille d'échantillonnage sur tout le domaine d'étude
- Détermination du nombre de mailles à échantillonner de la grille d'échantillonnage
- Calcul des coordonnées des échantillons de tous les sondages
- Attribution d'un identifiant aux échantillons
- Cartographie de la position des nouveaux échantillons proposés
 - 3D superposée à la classification des sols pollué-sain-incertain
 - 2D par couche superposée à la classification des sols pollué-sain-incertain
 - 2D par couche superposée aux fonds de plan
- Cartographie de la position des nouveaux sondages proposés avec leur identifiant
- Tableau destiné à l'opérateur de terrain avec
 - Identifiant des échantillons, coordonnées (x,y,z) des échantillons

SORTIES

- **Données Matlab**
 - **compteurs_Nomdusite.mat** nombre de mesures, nombre de jours de campagne, coût de la campagne mis à jour
 - **echagezincertaine_j_Nomdusite.mat** nombre d'échantillons (mesures XRF), nombre de sondages de l'étape de campagne suivante, surface de la zone incertaine, surface d'une maille de la grille d'échantillonnage, dimensions de cette maille, caractéristiques de la grille fine de calcul, caractéristiques du centre de gravité de la grille fine, caractéristiques de la grille grossière d'échantillonnage (origine, angle par rapport au système de coordonnées choisi, dimension des mailles, nombre de mailles), **coordonnées des nouveaux sondages, coordonnées des nouveaux échantillons, identifiant des échantillons**
- **Tableaux Excel**
 - **Echagepropose_phase_j+1_Nomdusite.xls** identifiants et coordonnées des nouveaux échantillons
- **Figures Matlab**
 - Echagepropose_3D_phase_j+1_Nomdusite.fig
 - Echagepropose_zincert_profondeurdelacouche_phase_j+1_Nomdusite.fig
 - Echagepropose_profondeurdelacouche_phase_j+1_Nomdusite.fig fond de plan
 - Sondagepropose_phase_j+1_Nomdusite.fig
- **Figures p/ rapports, exposés**
 - Echagepropose_3D_phase_j+1_Nomdusite.png
 - Echagepropose_zincert_profondeurdelacouche_phase_j+1_Nomdusite.png
 - Echagepropose_profondeurdelacouche_phase_j+1_Nomdusite.png fond de plan
 - Sondagepropose_phase_j+1_Nomdusite.png

Annexe 4 – Conditions de simulations des états de référence laboratoire et FPXRF

1 CONSTITUTION DE LA RÉFÉRENCE-LABO

Une simulation des teneurs en Pb mesurées par ICP au laboratoire sur échantillon composite (représentatif des 20 premiers cm de sol) est générée à partir des données ISA (235), OUTRAGE (10) et REPERAGE (95).

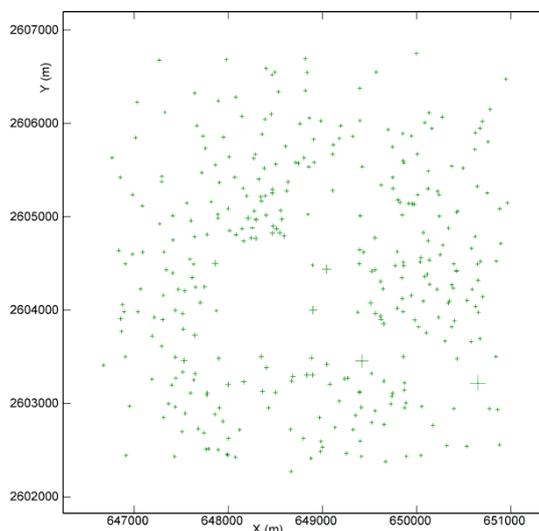


Figure 12 Carte de localisation des données de Pb de laboratoire. Secteur d'étude MétalEurop

La carte de la **Figure 12** montre quelques regroupements des données. Elles font donc l'objet d'un dégroupement (fenêtre 100 m x 100 m). Les statistiques calculées sur les données dégroupées se trouvent dans la deuxième ligne du tableau suivant.

Tableau 11 Statistiques élémentaires des teneurs en Pb de laboratoire disponibles sur le secteur d'étude MétalEurop

	Nombre	Minimum	Maximum	Moyenne	Ecart-type	CV
Non dégroupées	338	30	8754	456	706	1.55
Dégroupées	338	30	8754	451	741	1.64

La distribution des données étant très asymétrique, les données font l'objet d'une transformation en distribution gaussienne. Une fonction d'anamorphose gaussienne qui permet d'utiliser les données dégroupées est utilisée (30 polynômes d'Hermite, méthode d'inversion empirique).

Le variogramme omnidirectionnel des données dégroupées transformées est calculé avec 10 pas de 300 m (tolérance sur la distance 50%). La fonction ajustée est constituée d'un effet de pépité et d'un modèle sphérique de portée 1434 m (**Figure 13**). La variance totale du modèle est de 1.065.

$$\gamma(h) = 0.165EP + 0.900Sph(1434m)$$

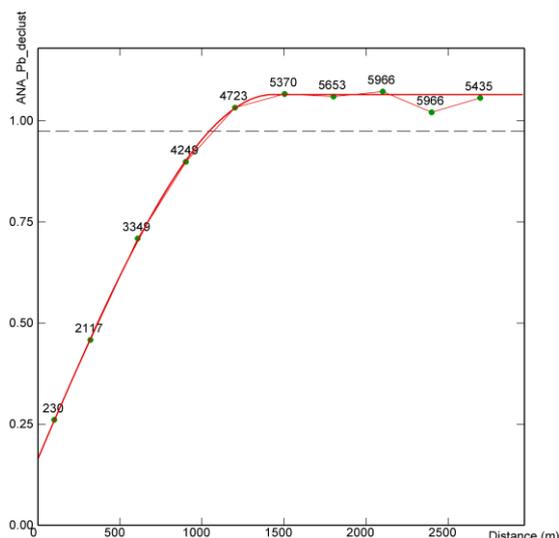


Figure 13 Variogramme expérimental des données de Pb de laboratoire et modèle ajusté. 10 pas de 300 m. Site d'étude MétalEurop

Une grille de modélisation de $181 \times 181 = 32761$ mailles de $25\text{m} \times 25\text{m}$ est construite (origine de la grille de coordonnées $(X_0, Y_0) = (646610\text{m}, 2602260\text{m})$ dans le système de coordonnées Lambert II étendu).

Un ensemble de 10 simulations ponctuelles des teneurs en Pb est généré par la méthode des Simulations Séquentielles Gaussiennes (SSG) dans les conditions suivantes :

- Voisinage glissant de rayon 2000 m, nombre minimal de points dans le voisinage : 1, nombre de secteurs angulaires : 4, nombre optimal de points par secteur angulaire : 4, distance minimale entre points : 5.
- Nombre de nœuds déjà simulés : 10.
- Seed : 423141.
- Backtransformation demandée.
- Temps calcul : 2'33".

La simulation #2 (**Figure 14**) est choisie arbitrairement pour constituer l'état de référence du Pb mesuré au laboratoire.

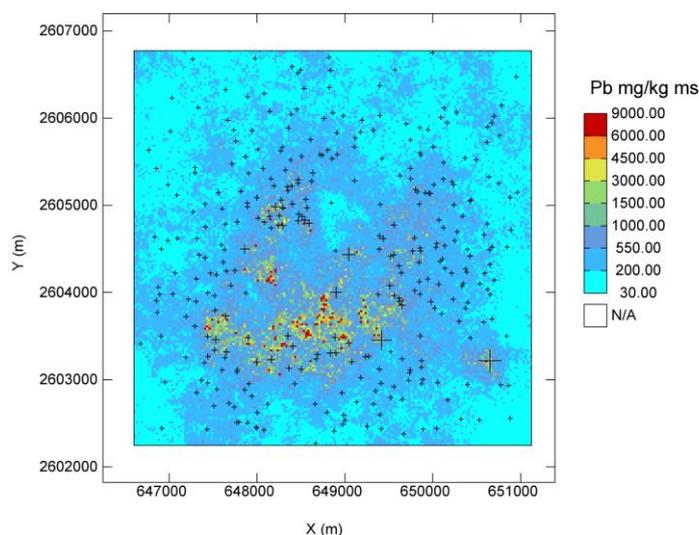


Figure 14 Simulation #2 SSG des teneurs en Pb de laboratoire sur le secteur d'étude MétalEurop

Cette simulation est adaptée par :

- Suppression de l'effet de pépite dans le modèle de variogramme : $\gamma(h) = 0.900Sph(1434m)$
- Suppression de deux points responsables de taches de teneurs anormalement faibles.

ID	Campagne	Pb - INERIS	Occupation	(X,Y)
59	REPERAGE	37 mg/kg ms	Friche près ancien puits de mine – remblais de démolition	(648845,2605026)
41	REPERAGE	111 mg/kg ms	Haut d'un terail – terre noire – bcp de cailloux	(648895,2604481)

La simulation finalement obtenue pour constituer l'état de référence des teneurs en Pb de laboratoire sur la couche de sol 0-20 cm est représentée à la **Figure 15**.

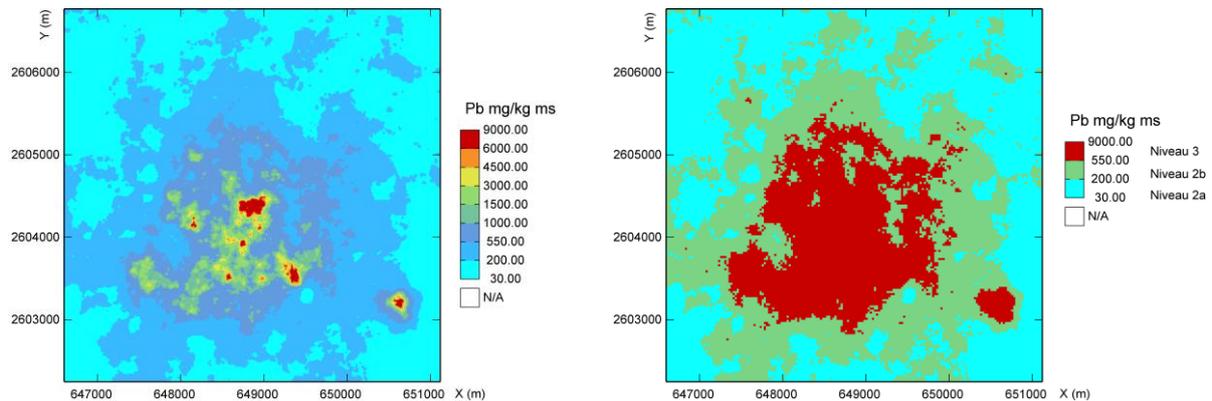


Figure 15 Simulation #2 SSG des teneurs en Pb de laboratoire après correction sur le secteur d'étude MétalEurop (gauche) et classification par rapport aux seuils d'action nationale Pb (droite)

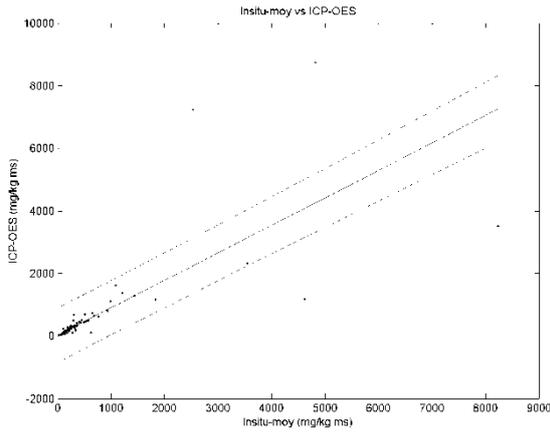
2 ETUDE DE LA CORRELATION ENTRE MESURES SUR SITE ET DE LABORATOIRE

Du fait de l'effet de support, les mesures FPXRF in situ ne peuvent être directement comparées aux mesures de laboratoire effectuées sur un échantillon composite (représentatif d'une hauteur de sol de 20 cm). Un estimateur de la mesure de laboratoire est donc calculé à partir des mesures in situ réalisées tous les 5 cm sur la verticale en chaque sondage, de 0 à 20 cm de profondeur :

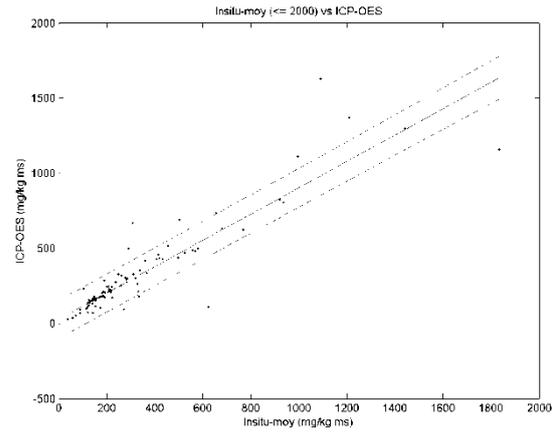
$$M^* = \frac{1}{5} \cdot \sum_{i=1}^5 X(i), \text{ où } X(i) \text{ désigne la mesure FPXRF in situ à la profondeur } [(i-1) \cdot 5] \text{ cm.}$$

Dans le cadre d'OUTRAGE, la relation qui avait été observée entre les mesures du Pb FPXRF in situ et les mesures ICP du laboratoire du BRGM (10 mesures en triplicata) pouvait être modélisée par une loi de régression linéaire robuste (sans correction de la teneur en eau, cette relation est : $Y = 1.03 X + 65.47$ avec $Y = \text{labo}$, $\rho = 0.9984$). Le coefficient de variation calculé sur les triplicatas des mesures FPXRF ex situ et des mesures ICP ne dépassait pas 10% pour une durée de mesure FPXRF de 8 mn.

La relation entre la moyenne des teneurs mesurées in situ par FPXRF et l'analyse de laboratoire par ICP au laboratoire de l'INERIS est étudiée également dans le cadre du projet REPERAGE (**Figure 16**). Les couples de mesures à notre disposition dans le cadre de REPERAGE (95 couples) sont plus nombreux que dans le cadre d'OUTRAGE, ce qui permet de préciser ou d'ajuster la relation constatée sur les données d'OUTRAGE.



Toutes données :
 $Y = 0.88 X + 26.85$ (Y : labo)
 $\rho = 0.71$



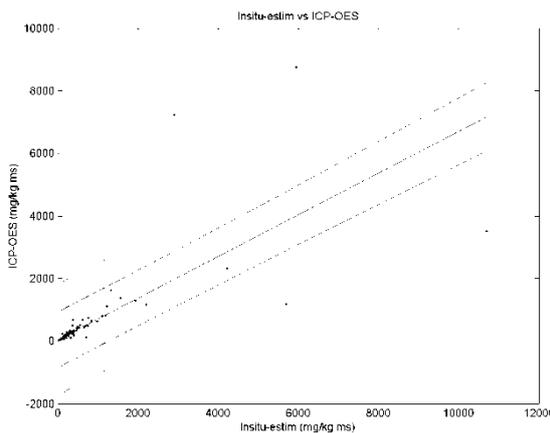
Sélection des moyennes in situ ≤ 2000 :
 $Y = 0.88 X + 27.01$ (Y : labo)
 $\rho = 0.91$

Figure 16 Relation entre moyenne des teneurs FPXRF in situ et des analyses ICP labo REPERAGE. Site d'étude MétalEurop

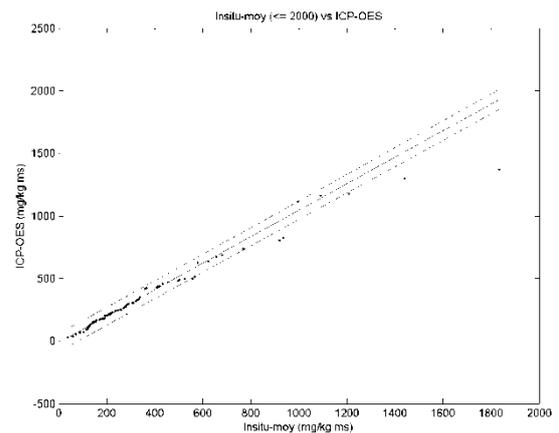
Ces graphes aboutissent aux constats suivants :

- La corrélation entre les deux types de mesures pour toutes les données est médiocre ($\rho = 0.71$). La corrélation est bonne pour des moyennes de teneurs in situ inférieures à 2000 mg/kg ms ($\rho = 0.91$).
- Les teneurs en Pb de laboratoire sont sur - estimées par les mesures FPXRF in situ.
- La dispersion des valeurs autour de la loi de régression est nettement plus grande aux plus fortes teneurs ($X \geq 1000$ mg/kg ms).

Si les moyennes des teneurs FPXRF in situ sont corrigées de la teneur en eau du sol (α) de la façon suivante : $L^* = \frac{1}{1-\alpha} \cdot \frac{1}{5} \sum_{i=1}^5 X(i)$, les relations entre les estimateurs L^* et les analyses de laboratoire sont modifiées ainsi.



Toutes données :
 $Y = 0.67 X + 37.82$ (Y : labo)
 $\rho = 0.69$



Sélection des moyennes in situ ≤ 2000 :
 $Y = 0.67 X + 37.92$ (Y : labo)
 $\rho = 0.92$

Figure 17 Relation entre moyenne des teneurs FPXRF in situ corrigées de la teneur en eau et des analyses ICP labo REPERAGE. Site d'étude MétalEurop

Les teneurs en Pb de laboratoire sont davantage sur - estimées par les mesures FPXRF in situ corrigées de la teneur en eau que par les simples moyennes des mesures FPXRF in situ. Les coefficients de corrélation sont comparables.

De plus, la fidélité des mesures FPXRF a été évaluée par le laboratoire INERIS sur échantillon composite à partir de triplicatas. Le coefficient de variation des triplicatas apparaît au minimum égal à 1%, au maximum égal à 20% et en moyenne égal à 6%. La fidélité des mesures FPXRF pour une durée de 2 mn est donc moins bonne que celle observée sur OUTRAGE (CV < 10%).

Au bilan, on décide d'utiliser la moyenne des mesures in situ sans correction de la teneur en eau comme estimateur des valeurs de laboratoire, et on considère que la corrélation de ces estimateurs avec les analyses de laboratoire est de 0.70.

3 CONSTITUTION DE LA RÉFÉRENCE-FPXRF IN SITU

L'état de référence des teneurs en Pb mesurées par FPXRF est constitué par cosimulation colocalisée, où la moyenne des mesures FPXRF in situ en Pb (après transformation gaussienne) constitue la variable principale et la simulation # 2 des teneurs en Pb mesurées au laboratoire constitue la variable secondaire (transformées gaussiennes des teneurs).

On utilise à la fois les données FPXRF du projet OUTRAGE (80) et celles du projet REPERAGE (95), soit un total de 175 données, pour renseigner la variable principale (**Figure 18**).

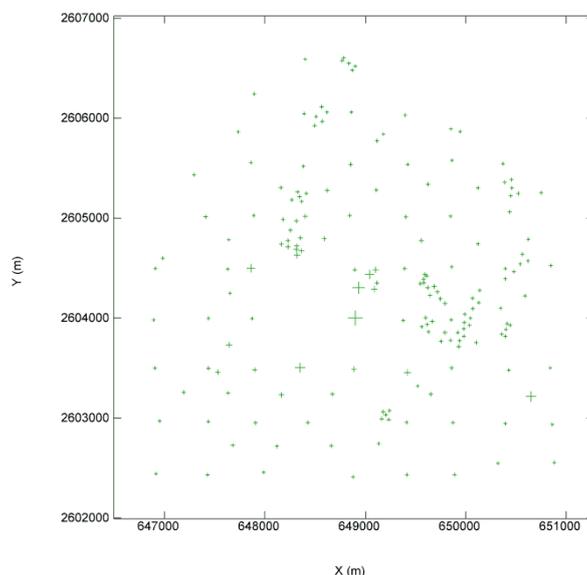


Figure 18 Carte des données de Pb FPXRF de terrain (moyenne de mesures FPXRF in situ par verticale). Secteur d'étude MétalEurop

Les données sont dégroupées (fenêtre 200 m x 200 m). Les statistiques des données non dégroupées et dégroupées sont résumées au **Tableau 12**.

Tableau 12 Statistiques élémentaires des teneurs en Pb de laboratoire disponibles sur le secteur d'étude MétalEurop

	Nombre	Minimum	Maximum	Moyenne	Ecart-type	CV
Non dégroupées	175	38	8232	576	1000	1.74
Dégroupées	175	38	8232	615	1189	1.93

Les données dégroupées font l'objet d'une transformation par anamorphose gaussienne (30 polynômes d'Hermite, méthode d'inversion empirique).

Même s'il n'est pas utilisé dans le cadre de la cosimulation colocalisée, un variogramme est calculé et modélisé pour ces données transformées pour apprécier la répartition spatiale des données FPXRF in

situ disponibles. Le modèle ajusté est composé d'un effet de pépité et d'un modèle exponentiel de portée 2215 m (**Figure 19**). La variance totale du modèle est de 1.32.

$$\gamma(h) = 0.020EP + 1.300Exp(2215m)$$

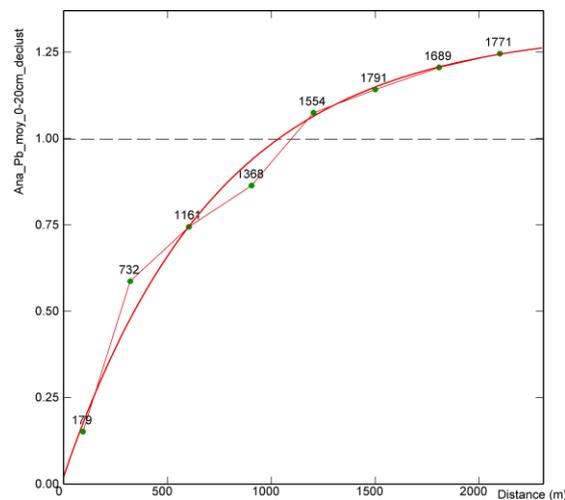


Figure 19 Variogramme expérimental des données de Pb de terrain (moyenne de mesures FPXRF in situ par verticale) et modèle ajusté. 8 pas de 300 m

La même grille de modélisation que celle utilisée pour constituer l'état de référence des teneurs en Pb de laboratoire est utilisée pour simuler les teneurs en Pb mesurées in situ par FPXRF.

Un total de 10 simulations ponctuelles des teneurs en Pb est généré par la méthode des Cosimulations colocalisées dans les conditions suivantes :

- Variable principale : données transformées par la fonction d'anamorphose et dégroupées des moyennes des mesures FPXRF in situ par verticale.
- Variable secondaire : valeurs simulées non back transformées des teneurs en Pb de laboratoire.
- Modèle de variogramme de la variable secondaire : modèle de variogramme ajusté aux données de laboratoire transformées et dégroupées (§ 1).
- Voisinage : identique au voisinage utilisé pour simuler les teneurs en Pb de laboratoire (§ 1).
- Rapport entre la variance de la variable secondaire et la variance de la variable principale (distributions gaussiennes) : on se base sur les valeurs des paliers des modèles de variogramme ajustés aux données pour en déduire un rapport < 1 (plus grande variabilité sur les données FPXRF que sur les données de laboratoire). $r = 1.07/1.32 = 0.81$.
- Coefficient de corrélation entre données de la variable secondaire et données de la variable principale : le coefficient de corrélation entre les données REPERAGE transformées est égal à 0.91. 0.90 est donc la valeur retenue.
- Proportion d'effet de pépité de la variance totale de la variable secondaire : 15%.
- Back-transformation demandée.
- Seed : 423141.
- Nombre de bandes tournantes : 100.

La simulation #2 retenue est présentée figure suivante.

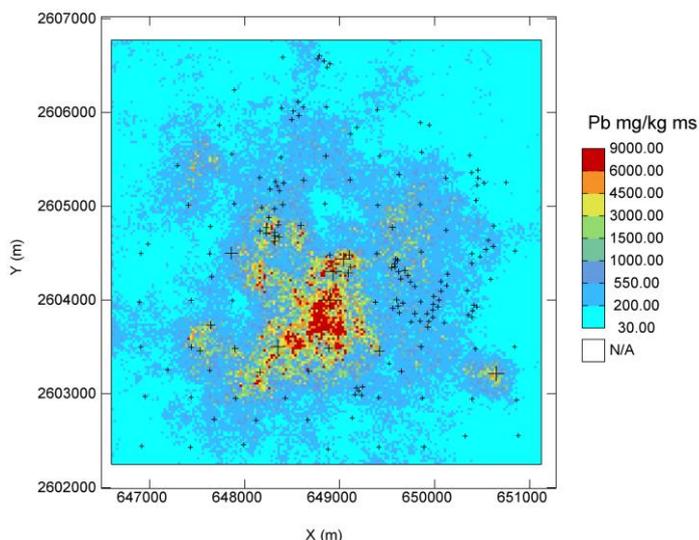


Figure 20 Simulation #2 des teneurs en Pb (moyenne des mesures FPXRF in situ par verticale) et classification par rapport aux seuils d'action nationale Pb (droite)

Cette simulation est adaptée par suppression de deux points responsables de taches de teneurs anormalement faibles.

ID	Pb – FPXRF in situ mg/kg ms	Occupation	(X,Y)
59	59	Friche près ancien puits de mine – remblais de démolition	(648845,2605026)
12	105	Bras déchargement Deûle – talus le long de la berge	(647873,2604000)

La simulation finalement obtenue pour constituer l'état de référence des teneurs en Pb de laboratoire sur la couche de sol 0-20 cm est représentée à la **Figure 21**.

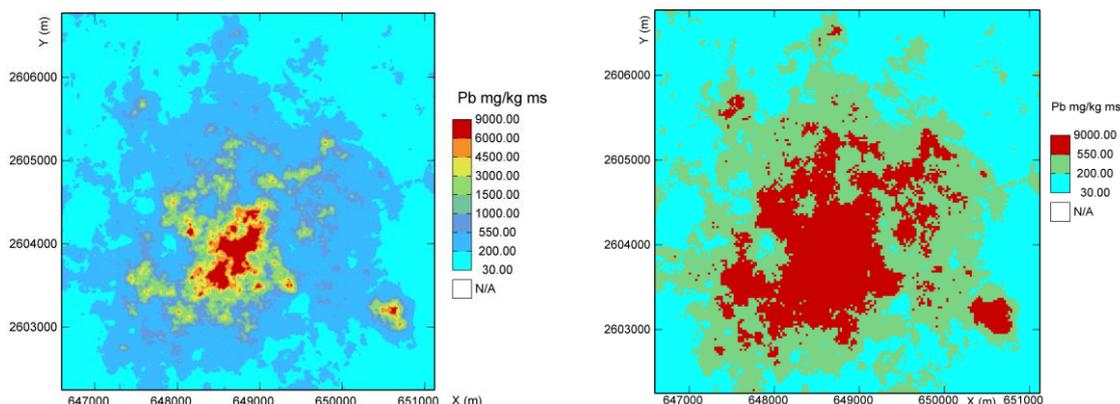


Figure 21 Simulation #2 SSG des teneurs FPXRF en Pb après correction sur le secteur d'étude MétalEurop (gauche) et classification par rapport aux seuils d'action nationale Pb (droite)

Annexe 5 – Analyse exploratoire et variographique des données collectées sur la couche de sol 0-5cm

En zone résidentielle ou récréative, la couche de sol qui présente un risque sanitaire par contact cutané, ingestion ou inhalation, est la couche de surface (épaisseur 5 cm). En zone agricole, la couche de sol présentant un risque par contamination de la chaîne alimentaire (cultures, bétail) est une couche plus épaisse, de l'ordre de 20 cm. Lorsqu'on se place dans le cadre d'une évaluation des risques sanitaires, la constitution d'un état de référence de la contamination des sols sur une hauteur de 20 cm ne suffit pas. Un état de référence sur la hauteur 0-5 cm est également nécessaire.

La constitution d'un état de référence de la couche 0-5 cm demande cependant une réflexion préalable sur les données disponibles. Les données de laboratoire collectées dans le cadre de REPERAGE pour renseigner la pollution de la couche 0-5 cm sont en effet moins nombreuses que celles renseignant la pollution de la couche 0-20 cm.

- 59 points de la zone d'étude ont fait l'objet d'un échantillonnage représentatif de la couche 0-5 cm (sur un total de 95 points de mesure).
- Pour le moment, seuls 20 de ces échantillons ont été analysés par ICP pour quantification du Pb.

Les données de Pb XRF renseignant sur la pollution de la couche 0-5 cm sont constituées de 95 mesures in situ réalisées à 0 cm de profondeur (après avoir enlevé le couvert végétal) et de 95 mesures in situ réalisées à 5 cm de profondeur.

Deux principales questions se posent :

- Sur la base des 20 points pour lesquels nous disposons à la fois de mesures XRF in situ de terrain et de mesures ICP de laboratoire représentatives des deux couches 0-5 cm et 0-20 cm, est-il possible d'établir une relation entre ces mesures qui permette ensuite, pour les 39 échantillons collectés sur la couche 0-5 cm mais non analysés au laboratoire, d'en déduire (d'estimer) la valeur de l'analyse ICP ? Ce qui permettrait de faire des économies sur cette partie du programme analytique qui n'avait pas été prévue au budget initial.
- Les teneurs en Pb mesurées sur la couche 0-5 cm présentent-elles une continuité spatiale ? La distribution des mesures XRF in situ de surface met en effet en évidence une plus grande variabilité que celle constatée sur les mesures XRF représentatives de la couche 0-20 cm, ce qui peut laisser craindre un phénomène moins structuré en surface.

Ce document décrit l'analyse exploratoire et variographique menée sur les données pour répondre à ces questions.

1 RELATION ENTRE TENEURS SUR LA COUCHE 0-5 CM ET TENEURS SUR LA COUCHE 0-20 CM

Les données disponibles ont été utilisées pour calculer le rapport entre la teneur mesurée sur la couche 0-5 cm et la teneur mesurée sur la couche 0-20 cm pour les mesures réalisées in situ par XRF et les mesures de laboratoire.

Trois rapports ont été calculés à partir des mesures XRF (rapports calculés pour 20 points) :

- Calcul de la moyenne des 2 mesures XRF in situ prises à 0 cm et à 5 cm, puis calcul du rapport entre cette moyenne et la moyenne des 5 mesures XRF in situ prises tous les 5 cm sur 20 cm de hauteur : R_{XRF_moy} . La moyenne des 2 valeurs XRF prises à 0 et à 5 cm est considérée a priori comme la valeur pouvant être la plus proche de l'analyse de laboratoire menée sur un échantillon de sol représentatif des 5 premiers cm de sol.
- Calcul du rapport entre la mesure XRF in situ prise à 0 cm et la moyenne des 5 mesures XRF R_{XRF_0cm} .
- Calcul du rapport entre la mesure XRF in situ prise à 5 cm et la moyenne des 5 mesures XRF R_{XRF_5cm} .

Un seul rapport peut être et a été calculé à partir des mesures de laboratoire (rapport calculé pour 95 points):

- Calcul du rapport entre l'analyse ICP sur échantillon représentatif de 5 cm et l'analyse ICP sur échantillon représentatif de 20 cm R_{labo} .

Les rapports des mesures XRF ont été comparés au rapport des analyses de laboratoire R_{labo} en traçant leur diagramme de dispersion, en calculant leur coefficient de corrélation et en ajustant une droite de régression linéaire (**Tableau 13** et **Figure 22**). Seuls 20 couples de points sont disponibles pour ces comparaisons.

Ces résultats montrent une bonne corrélation entre les rapports R_{labo} et R_{XRF_0cm} ($\rho = 0.87$). Contre toute attente, cette corrélation est meilleure que celle constatée entre R_{labo} et R_{XRF_moy} ($\rho = 0.76$). Les rapports R_{labo} et R_{XRF_5cm} sont corrélés négativement.

Le rapport R_{XRF_moy} sous-estime largement le rapport R_{labo} tandis que le rapport R_{XRF_0cm} le sur-estime légèrement.

Tableau 13 Coefficients de corrélation entre rapport de teneurs en Pb mesurées in situ par XRF ou au laboratoire par ICP et droites de régression

	Coefficients corrélation	Droites de régression
(R_{labo}, R_{XRF_moy})	0.76	$R_{labo}^* = 1.78 R_{XRF_moy} - 0.60$
(R_{labo}, R_{XRF_0cm})	0.87	$R_{labo}^* = 0.90 R_{XRF_0cm} + 0.29$
(R_{labo}, R_{XRF_5cm})	-0.32	$R_{labo}^* = -0.09 R_{XRF_5cm} + 1.26$

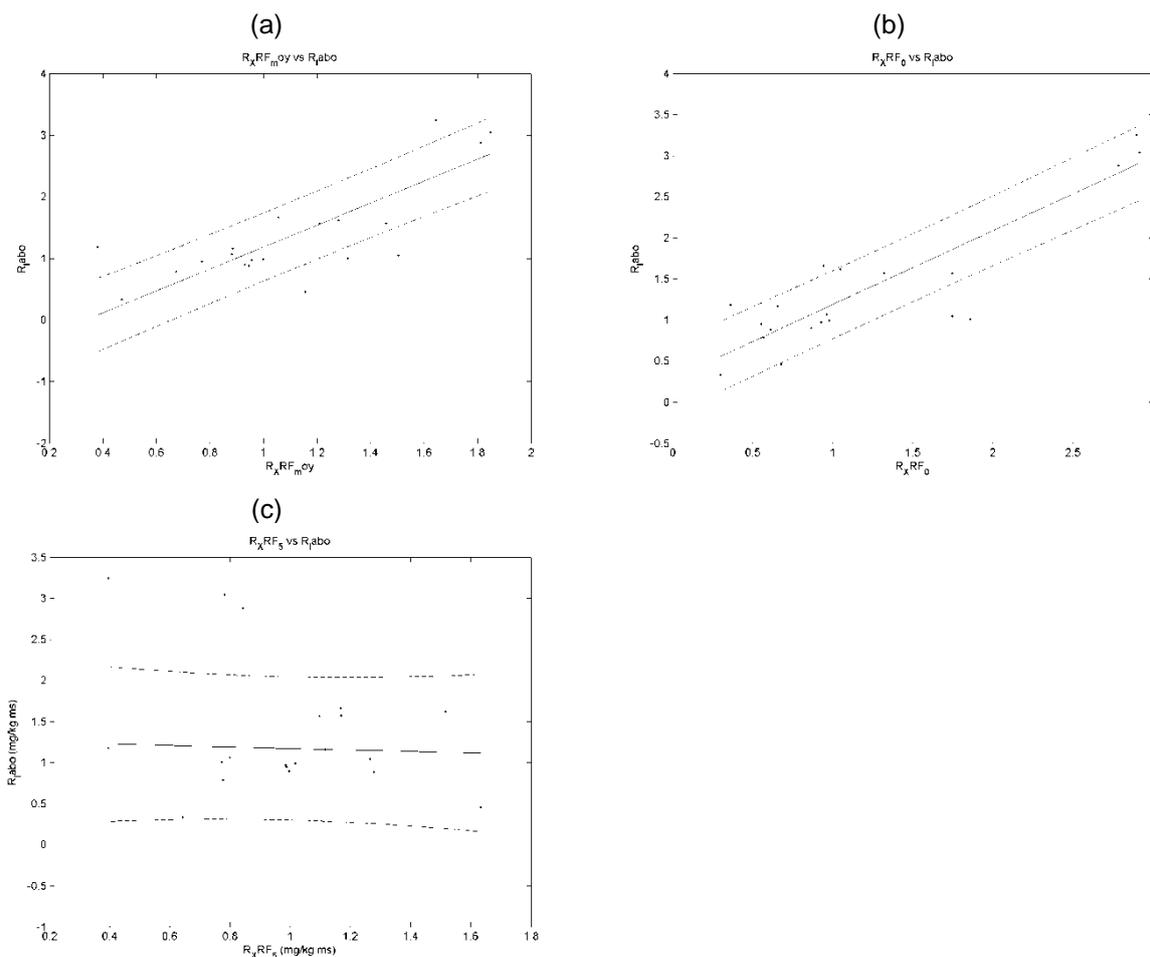


Figure 22 Diagrammes de dispersion entre le rapport des teneurs en Pb de laboratoire R_{lab0} (ordonnée) et le rapport des teneurs en Pb mesurées in situ par XRF (abscisse). Ce dernier rapport est R_{XRF_moy} sur le graphe (a), R_{XRF_0cm} sur le graphe (b) et R_{XRF_5cm} sur le graphe (c).

La loi de régression ajustée aux rapports (R_{lab0} , R_{XRF_0cm}) est utilisée pour calculer un estimateur du rapport R_{lab0} , représenté par R_{lab0}^* . Un estimateur $Pb_{lab0_0-5cm}^{est}$ de la teneur de laboratoire est alors calculé en multipliant R_{lab0}^* par la teneur Pb_{lab0_0-5cm} de laboratoire mesurée sur échantillon représentatif de 5 cm de sol :

$$Pb_{lab0_0-5cm}^{est} = R_{lab0}^* \times Pb_{lab0_0-5cm}$$

Ces teneurs estimées et les teneurs réellement mesurées au laboratoire sur échantillons 0-5 cm sont ensuite comparées, globalement en traçant leur diagramme de dispersion et en calculant un coefficient de corrélation et une loi de régression, et individuellement en comparant les valeurs estimées et réelles point par point.

Les points du diagramme de dispersion apparaissent très peu dispersés (**Figure 23**). Le coefficient de corrélation est très élevé, égal à 0.997. Les caractéristiques de la droite de régression sont les suivantes :

$$Pb_{lab0_0-5cm}^* = 0.95 Pb_{lab0_0-5cm}^{est} + 26.50$$

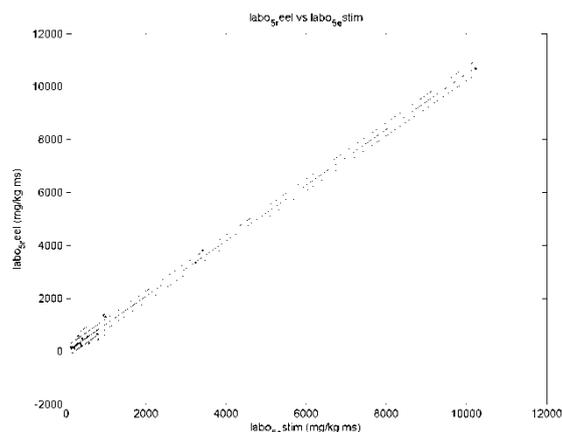


Figure 23 Diagramme de dispersion entre teneur en Pb de laboratoire estimée (X) et teneur en Pb de laboratoire réellement mesurée (Y)

La comparaison individuelle des 20 couples de teneurs (**Tableau 14**) montre que 3 valeurs estimées sur 20 (15%) peuvent susciter une erreur de classification des sols par rapport au seuil de 550 mg/kg ms (niveau d'intervention 3).

Tableau 14 Teneur en Pb sur échantillon 0-5cm mesurée réellement au laboratoire et estimateur de ces teneurs calculé à partir de la relation entre les rapports de mesures sur échantillons 0-5cm et 0-20cm de terrain et de laboratoire. Les teneurs pour lesquelles les estimateurs génèrent une erreur de classification par rapport à un seuil réglementaire (200 et 550 mg/kg ms) apparaissent en gras

ID	$Pb_{labo. 0-5cm}$	$Pb_{labo. 0-5cm}^{est}$
1	152	143
12	106	208
14	10683	10222
21	457	811
29	228	383
30	1372	936
31	3825	3410
45	157	118
46	1306	991
47	278	283
49	174	201
51	228	247
58	661	782
67	301	355
80	3345	3244
81	559	530
92	579	301
94	293	570
99	110	130
100	488	404

2 ETUDE DE LA CONTINUITÉ SPATIALE DES TENEURS EN Pb MESURÉES IN SITU SUR LA COUCHE 0-5 CM

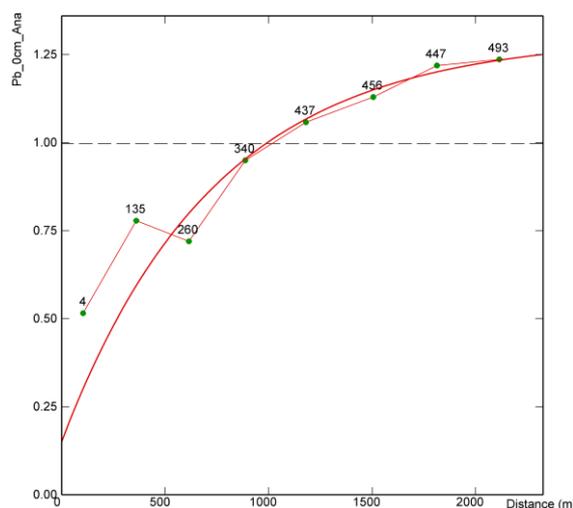
Les statistiques élémentaires des teneurs en Pb mesurées par XRF in situ en surface de sol (0 cm) sont résumées au **Tableau 15**.

Tableau 15 Statistiques élémentaires des teneurs en Pb mesurées in situ par XRF à la surface du sol (0 cm) - REPERAGE

Nombre	Minimum	Maximum	Moyenne	Ecart-type	CV
95	44	23985	826	2817	3.41

La distribution étant fortement asymétrique positive, les données sont transformées par une fonction d'anamorphose gaussienne (méthode d'inversion empirique, 30 polynômes d'Hermite). Les données transformées montrent une continuité spatiale qui peut être modélisée par la fonction :

$$\gamma(h) = 0.150EP + 1.150Exp(2215m)$$

**Figure 24** Variogramme expérimental des données REPERAGE de Pb de terrain (mesures XRF in situ à 0 cm de profondeur) et modèle ajusté. 8 pas de 300 m

Cette information issue de REPERAGE peut être complétée par les données d'OUTRAGE. On dispose ainsi de 175 données sur la zone d'étude, qui permettent de préciser la continuité spatiale observée uniquement sur les données de REPERAGE.

Tableau 16 Statistiques élémentaires des teneurs en Pb mesurées in situ par XRF à la surface du sol (0 cm) – OUTRAGE et REPERAGE

Nombre	Minimum	Maximum	Moyenne	Ecart-type	CV
175	44	23985	882	2753	3.12

Le variogramme expérimental est calculé sur des données dégroupées (fenêtre 20m x 200m) et transformées par anamorphose gaussienne. Le modèle ajusté a les caractéristiques suivantes :

$$\gamma(h) = 0.070EP + 1.010Exp(2123m)$$

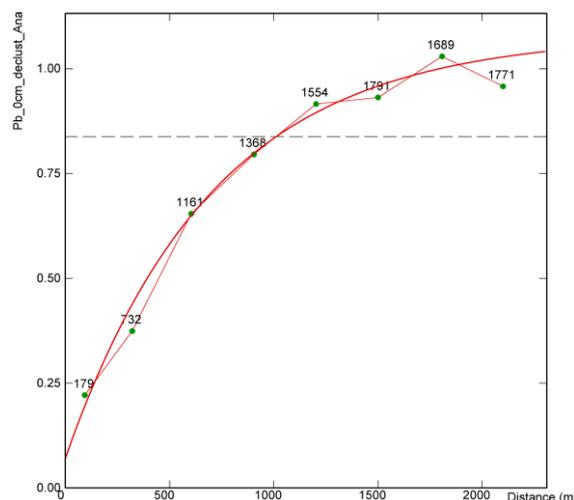


Figure 25 Variogramme expérimental des données OUTRAGE et REPERAGE de Pb de terrain (mesures XRF in situ à 0 cm de profondeur) et modèle ajusté. 8 pas de 300 m

On peut s'attendre à une différence de continuité spatiale lorsque les sols ont été remaniés ou non. Cependant les données disponibles ne permettent pas d'étudier cette différence : les données des deux classes « sols remaniés » et « sols non remaniés » sont trop peu nombreuses et surtout mal réparties sur la zone d'étude.

3 BILAN DE L'ANALYSE EXPLORATOIRE ET VARIOGRAPHIQUE

Il existe une relation entre les rapports des mesures XRF et des mesures de laboratoire représentatives des sols de surface et de la moyenne des sols sur 20 cm de hauteur. Cette relation semble pouvoir être utilisée pour déduire des données disponibles les teneurs de laboratoire pour les échantillons de sol représentatifs de la hauteur 0-5 cm.

Les données OUTRAGE et REPERAGE permettent de modéliser la distribution spatiale des teneurs en Pb de la couche de sol de surface, sans faire de distinction entre sols remaniés et sols non remaniés. De la même façon que pour la couche de sol 0-20 cm, il est donc possible de simuler un état de référence de la pollution de surface (0 cm) pour des mesures XRF in situ ou pour des mesures de laboratoire.

4 SIMULATION DES TENEURS EN PB SUR LA COUCHE 0-5 CM ET CALCUL DES INDICES DE RISQUE ASSOCIES

Une simulation des teneurs en Pb de laboratoire et des teneurs en Pb mesurées par FPXRF sur la couche de sol 0-5 cm a été générée selon les mêmes procédures que celles utilisées pour la couche de sol 0-20 cm. Les deux cartes ainsi simulées sont représentées à la **Figure 26**, ligne du haut. L'application de la fonction de transfert présentée au § 3.2.2.d à ces simulations permet d'en déduire une cartographie de référence des indices de risques (**Figure 26**, ligne du bas).

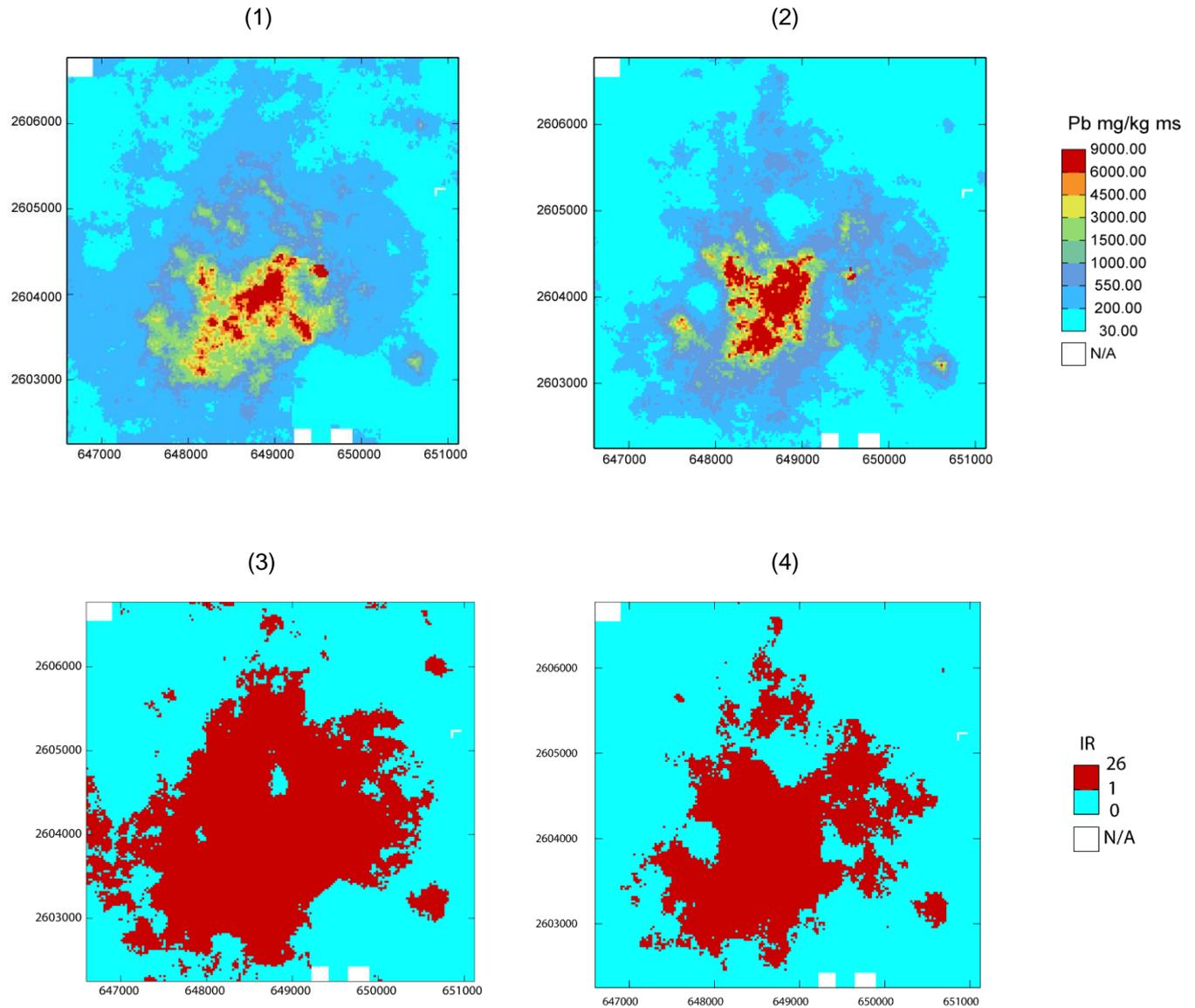


Figure 26 Simulation de la contamination en Pb dans les 5 premiers de sol du secteur d'étude MétalEurop. (1) teneurs de laboratoire (2) teneurs mesurées par FPXRF (3) indices de risques calculés à partir des teneurs de laboratoire (4) indices de risque calculés à partir des mesures FPXRF

Annexe 6 – Document de préparation de l'étape préliminaire de reconnaissance sur Henricot II

Document de travail

Sujet : Projet REPERAGE. Proposition de plan préliminaire d'échantillonnage et de mesure sur le site Henricot

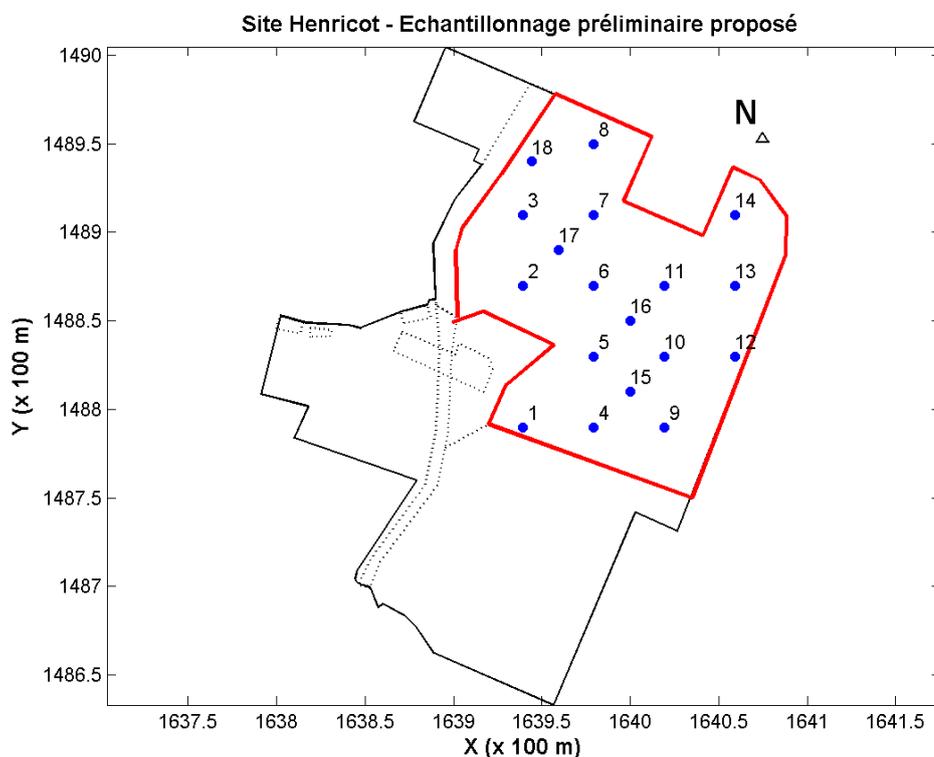
Date : 24/04/2008

A l'attention de : Laurence Haouche-Belkessam (ISSeP) et Sébastien Denys (INERIS)

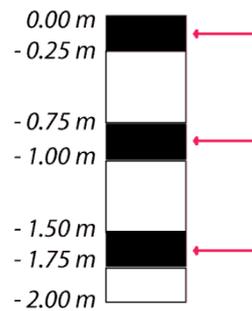
De la part de : Hélène Demougeot-Renard

cc : Dimitri D'Or et Michel Garcia

- Domaine d'investigation : partie Nord du site, en dehors de la zone centrale où ont été trouvés des hydrocarbures flottants (zone délimitée en rouge sur le plan)



- Profondeur d'investigation : sur toute la hauteur de remblais, sans dépasser le niveau piézométrique.
- Nombre de sondages : 18 (15 répartis selon une maille régulière 40 m x 40 m + 3 plus rapprochés) (voir plan ci-dessus)
- Mesures XRF par sondage : L'objectif est de réaliser le plus grand nombre de mesures XRF possible sur chaque hauteur de sondage, le minimum étant 3 mesures par sondage. Dans l'hypothèse où les remblais hors d'eau ont une hauteur de 2 m, ces 3 mesures minimales seront prises tous les 75 cm sur des échantillons composites représentatifs de 25 cm (voir schéma ci-après). L'espacement entre les mesures sera cependant réduit si le temps et le personnel attribués à la campagne le permettent. Remarque importante : le support des mesures XRF (25 cm) doit être impérativement respecté.



- Nombre total de mesures XRF : au minimum $18 \times 3 = 54$.
- Nombre total d'analyses de laboratoire : 54. Si le nombre total d'échantillons collectés s'avérait supérieur à 54, un sous-ensemble d'échantillons sera constitué en sélectionnant 3 échantillons par sondage, un en surface, un au milieu et un au fond.
- Remarque importante : il est nécessaire de ne pas placer les nouveaux sondages là où des sondages ou des fosses ont déjà été creusés lors des campagnes précédentes. Déplacer le point de sondage théorique de quelques dm ou m si nécessaire.
- Identifiants et coordonnées des échantillons : ils sont indiqués dans le tableau ci-dessous et dans le tableau excel *Echage_preliminaire_Henricot.xls*. Ces tableaux ont été établis dans l'hypothèse où 3 échantillons sont collectés tous les 75 cm sur chacun des 18 sondages ; ils devront être adaptés en fonction des mesures et échantillons qui auront réellement été collectés sur le terrain. Remarque importante : la notation des identifiants des échantillons (« numéro.profondeurencmavec4chiffres ») est à respecter strictement au moment de leur saisie dans l'appareil XRF, le DGPS, et sur le bordereau d'analyse de laboratoire.

Tableau des identifiants et des coordonnées des échantillons

SAMPLE	X	Y	Z
1,0125	163939,447	148790	-0,125
1,0875	163939,45	148790,00	-0,875
1,1625	163939,45	148790,00	-1,625
2,0125	163939,45	148870,00	-0,125
2,0875	163939,45	148870,00	-0,875
2,1625	163939,45	148870,00	-1,625
3,0125	163939,45	148910,00	-0,125
3,0875	163939,45	148910,00	-0,875
3,1625	163939,45	148910,00	-1,625
4,0125	163979,45	148790,00	-0,125
4,0875	163979,45	148790,00	-0,875
4,1625	163979,45	148790,00	-1,625
5,0125	163979,45	148830,00	-0,125
5,0875	163979,45	148830,00	-0,875
5,1625	163979,45	148830,00	-1,625
6,0125	163979,45	148870,00	-0,125
6,0875	163979,45	148870,00	-0,875
6,1625	163979,45	148870,00	-1,625
7,0125	163979,45	148910,00	-0,125
7,0875	163979,45	148910,00	-0,875
7,1625	163979,45	148910,00	-1,625
8,0125	163979,45	148950,00	-0,125
8,0875	163979,45	148950,00	-0,875
8,1625	163979,45	148950,00	-1,625
9,0125	164019,45	148790,00	-0,125
9,0875	164019,45	148790,00	-0,875
9,1625	164019,45	148790,00	-1,625
10,0125	164019,45	148830,00	-0,125
10,0875	164019,45	148830,00	-0,875
10,1625	164019,45	148830,00	-1,625
11,0125	164019,45	148870,00	-0,125
11,0875	164019,45	148870,00	-0,875
11,1625	164019,45	148870,00	-1,625
12,0125	164059,45	148830,00	-0,125
12,0875	164059,45	148830,00	-0,875
12,1625	164059,45	148830,00	-1,625
13,0125	164059,45	148870,00	-0,125
13,0875	164059,45	148870,00	-0,875
13,1625	164059,45	148870,00	-1,625
14,0125	164059,45	148910,00	-0,125
14,0875	164059,45	148910,00	-0,875
14,1625	164059,45	148910,00	-1,625
15,0125	164000,00	148810,00	-0,125
15,0875	164000,00	148810,00	-0,875
15,1625	164000,00	148810,00	-1,625
16,0125	164000,00	148850,00	-0,125
16,0875	164000,00	148850,00	-0,875
16,1625	164000,00	148850,00	-1,625
17,0125	163959,45	148890,00	-0,125
17,0875	163959,45	148890,00	-0,875

SAMPLE	X	Y	Z
17,1625	163959,45	148890,00	-1,625
18,0125	163944,45	148940,00	-0,125
18,0875	163944,45	148940,00	-0,875
18,1625	163944,45	148940,00	-1,625

**Annexe 7 – Documentation de terrain fournie à
chaque étape itérative de reconnaissance sur le cas
d'étude Henricot II**

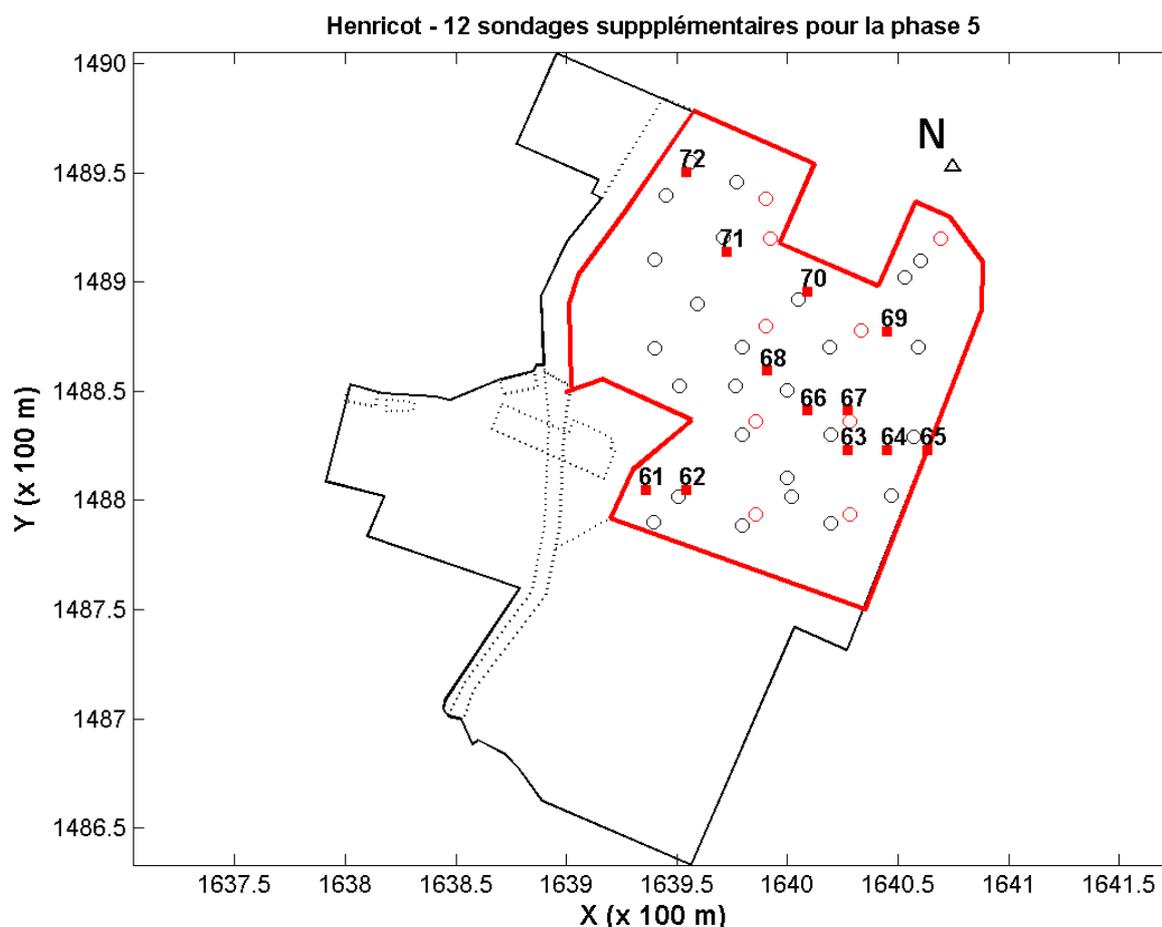


Figure 27 Carte de positionnement de 12 sondages supplémentaires pour l'étape itérative 5 de reconnaissance du cas d'étude Henricot

Tableau 17 Identifiants et coordonnées (Lambert II) des points de mesures prévues pour l'étape itérative 5

SAMPLE	X	Y	Z
61,0225	163936,88	148805,65	-0,225
61,0575	163936,88	148805,65	-0,575
61,0925	163936,88	148805,65	-0,925
61,1275	163936,88	148805,65	-1,275
61,1625	163936,88	148805,65	-1,625
62,0225	163955,09	148805,65	-0,225
62,0575	163955,09	148805,65	-0,575
62,0925	163955,09	148805,65	-0,925
62,1275	163955,09	148805,65	-1,275
62,1625	163955,09	148805,65	-1,625
63,0225	164027,95	148823,86	-0,225
63,0575	164027,95	148823,86	-0,575
63,0925	164027,95	148823,86	-0,925
63,1275	164027,95	148823,86	-1,275
63,1625	164027,95	148823,86	-1,625
64,0225	164046,17	148823,86	-0,225

SAMPLE	X	Y	Z
64,0575	164046,17	148823,86	-0,575
64,0925	164046,17	148823,86	-0,925
64,1275	164046,17	148823,86	-1,275
64,1625	164046,17	148823,86	-1,625
65,0225	164064,39	148823,86	-0,225
65,0575	164064,39	148823,86	-0,575
65,0925	164064,39	148823,86	-0,925
65,1275	164064,39	148823,86	-1,275
65,1625	164064,39	148823,86	-1,625
66,0225	164009,74	148842,08	-0,225
66,0575	164009,74	148842,08	-0,575
66,0925	164009,74	148842,08	-0,925
66,1275	164009,74	148842,08	-1,275
66,1625	164009,74	148842,08	-1,625
67,0225	164027,95	148842,08	-0,225
67,0575	164027,95	148842,08	-0,575
67,0925	164027,95	148842,08	-0,925
67,1275	164027,95	148842,08	-1,275
67,1625	164027,95	148842,08	-1,625
68,0225	163991,52	148860,29	-0,225
68,0575	163991,52	148860,29	-0,575
68,0925	163991,52	148860,29	-0,925
68,1275	163991,52	148860,29	-1,275
68,1625	163991,52	148860,29	-1,625
69,0225	164046,17	148878,51	-0,225
69,0575	164046,17	148878,51	-0,575
69,0925	164046,17	148878,51	-0,925
69,1275	164046,17	148878,51	-1,275
69,1625	164046,17	148878,51	-1,625
70,0225	164009,74	148896,72	-0,225
70,0575	164009,74	148896,72	-0,575
70,0925	164009,74	148896,72	-0,925
70,1275	164009,74	148896,72	-1,275
70,1625	164009,74	148896,72	-1,625
71,0225	163973,31	148914,94	-0,225
71,0575	163973,31	148914,94	-0,575
71,0925	163973,31	148914,94	-0,925
71,1275	163973,31	148914,94	-1,275
71,1625	163973,31	148914,94	-1,625
72,0225	163955,09	148951,37	-0,225
72,0575	163955,09	148951,37	-0,575
72,0925	163955,09	148951,37	-0,925
72,1275	163955,09	148951,37	-1,275
72,1625	163955,09	148951,37	-1,625