



**HAL**  
open science

# Optimisation de réseaux de capteurs pour la caractérisation de source de rejets atmosphériques

Hamza Kouichi

► **To cite this version:**

Hamza Kouichi. Optimisation de réseaux de capteurs pour la caractérisation de source de rejets atmosphériques. Mécanique des fluides [physics.class-ph]. Université Paris Saclay, 2017. Français. NNT: . tel-01593834v1

**HAL Id: tel-01593834**

**<https://hal.science/tel-01593834v1>**

Submitted on 26 Sep 2017 (v1), last revised 4 Oct 2017 (v2)

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

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

NNT : 2017SACLE020

THESE DE DOCTORAT  
DE  
L'UNIVERSITE PARIS-SACLAY  
PREPAREE A  
“ L'UNIVERSITE D'EVRY-VAL- D'ESSONNE ”

ECOLE DOCTORALE N° 579  
Sciences mécaniques et énergétiques, matériaux et géosciences  
Spécialité de doctorat : Mécanique des fluides

Par

**M. Hamza Kouichi**

**Optimisation de réseaux de capteurs pour la  
caractérisation de source de rejets atmosphériques**

**Thèse présentée et soutenue à « Evry », le « 18/07/2017 » :**

**Composition du Jury :**

Mme, Martin-Demina, Irina	Professeure, Université Aix-Marseille	Président et examinatrice
M, Sharan, Maithili	Professeur, Indian Institute of Technology Delhi	Rapporteur
M, Soulhac, Lionel	Maitre de conférences-HDR, Ecole Centrale de Lyon	Rapporteur
M, Chpoun, Amer	Professeur, Université d'Evry Val d'Essonne	Directeur de thèse
M, Ngae, Pierre	Maitre de conférences, Université d'Evry Val d'Essonne	Co-directeur de thèse
M, Issartel, Jean Pierre	Docteur et chercheur associé, Université d'Evry Val d'Essonne	Membre invité

*Find a truly original idea. It is the only way I will ever distinguish myself.  
It is the only way I will ever matter...*

« John Forbes Nash - A Beautiful Mind »

## Résumé

L'objectif principal de cette étude est de définir les méthodes nécessaires pour optimiser un réseau de surveillance conçu pour la *caractérisation* de source de rejets atmosphériques. L'optimisation consiste ici à déterminer le nombre et les positions optimales de capteurs à déployer afin de répondre à ce type de besoin. Dans ce contexte, l'optimisation est réalisée pour la première fois par un couplage entre la technique d'inversion de données dite de « renormalisation » et des algorithmes d'optimisation métaheuristiques. La méthode d'inversion a été en premier lieu évaluée pour la caractérisation de source ponctuelle, et a permis ensuite, de définir des critères d'optimalité pour la conception des réseaux. Dans cette étude, le processus d'optimisation a été évalué dans le cadre d'expériences réalisées en terrain plat sans obstacles (DYCE) et en milieu urbain idéalisé (MUST). Trois problématiques ont été définies et testées sur ces expériences. Elles concernent (i) la détermination de la taille optimale d'un réseau permettant de *caractériser* une source de pollution, où une fonction coût (erreurs normalisées), traduisant l'écart entre les observations et les données modélisées, a été minimisée; (ii) la conception optimale d'un réseau permettant de *caractériser* une source ponctuelle inconnue, pour une condition météorologique particulière. Dans ce contexte, une fonction coût entropique a été maximisée afin d'augmenter la quantité d'information fournie par le réseau; (iii) la détermination d'un réseau optimal permettant de *caractériser* une source ponctuelle inconnue pour des configurations météorologiques multiples. Pour ce faire, une fonction coût entropique généralisée, que nous avons définie, a été maximisée. Pour ces trois problématiques, l'optimisation est assurée dans le cadre d'une approche d'optimisation combinatoire. La détermination de la taille optimale d'un réseau (problématique 1) s'est révélée particulièrement sensible aux différentes conditions expérimentales (hauteur et débit de la source, conditions de stabilité, vitesse et direction du vent, etc.). Nous avons noté pour ces expériences, que les performances des réseaux sont meilleures dans le cadre d'une dispersion sur terrain plat comparativement aux milieux urbains. Nous avons également montré que différentes architectures de réseaux pouvaient converger vers le même optimum (approché ou global). Pour la caractérisation de sources inconnues (problématiques 2 et 3), les fonctions coûts entropiques se sont avérées robustes et ont permis d'obtenir des réseaux optimaux performants (de tailles raisonnables) capables de caractériser différentes sources pour une ou plusieurs conditions météorologiques.

**Mots clefs :** *Optimisation, renormalisation, champs adjoints, modèle gaussien, CFD, caractérisation de source, réseau de capteurs, fonction objectif, erreurs normalisées, entropie de réseau, information, entropie généralisée.*

---



## Abstract

The main objective of this study is to define the methods required to optimize a monitoring network designed for atmospheric source characterization. The optimization consists in determining the optimal number and locations of sensors to be deployed in order to respond to such needs. In this context, the optimization is performed for the first time by a coupling between the data inversion technique named "renormalization" and the metaheuristic optimization algorithms. At first, the inversion method was evaluated for a point source, and then have allowed to define optimality criteria for networks design. In this study, the optimization process was evaluated in experiments carried out in flat terrain without obstacles (DYCE) and in an idealized urban environment (MUST). Three problems were defined and tested based on these experiments. These problems concern (i) the determination of the optimal network size for source characterization, for which a cost function (standard errors) estimating the gap between observations and modeled data, has been minimized; (ii) the optimal design of a network to retrieve an unknown point source for a particular meteorological condition. In this context, an entropy cost function has been maximized in order to increase the information's amount provided by the network; (iii) the determination of an optimal network to reconstruct an unknown point source for multiple meteorological configurations. For this purpose, a generalized entropic cost function that we have defined, has been maximized. For these all problems, optimization is ensured within the framework of a combinatorial optimization approach. The determination of the optimal network size (problem 1) was highly sensitive to experimental conditions (source height and intensity, stability conditions, wind speed and direction, etc.). We have noted that the networks performance is better for a dispersion on flat terrain compared to the urban environments. We have also shown that different networks architectures can converge towards the same optimum (approximate or global). For unknown sources reconstruction (problems 2 and 3), the entropic cost functions have proven to be robust and allowed to obtain optimal networks (for reasonable sizes) capable of characterizing different sources for one or multiple meteorological conditions.

**Keywords:** *Optimization, renormalization, adjoint fields, Gaussian model, CFD, source characterization, sensor network, objective function, normalized errors, network entropy, information, generalized entropy.*

---

## Table des matières

Nomenclature.....	I
Liste des figures .....	III
Liste des tableaux .....	VI
Introduction générale.....	VIII
CHAPITRE 1 LES RESEAUX DE CAPTEURS ET LES ENJEUX « ENVIRONNEMENTAUX, INDUSTRIELS, MILITAIRES ET SECURITAIRES .....	1
1.1.Introduction.....	3
1.2. Contexte environnemental.....	3
1.2.1. Générateurs de risques .....	4
1.2.2. Conséquences possibles.....	7
1.2.3. Enjeux et besoins .....	9
1.3. Contexte industriel.....	11
1.3.1. Générateurs de risques .....	12
1.3.2. Conséquences possibles.....	13
1.3.3. Enjeux et les besoins.....	14
1.4. Contexte militaire et sécuritaire.....	16
1.4.1. Générateurs de risques .....	16
1.4.2. Effets possibles.....	21
1.4.3. Enjeux et besoins .....	21
1.5. Conclusion du chapitre.....	22
« Bibliographie chapitre 1 ».....	25
CHAPITRE 2 OPTIMISATION DE RESEAUX DE CAPTEURS.....	28
2.1. Introduction.....	30
2.2. Expression du besoin .....	30
2.3. Les questions clés .....	31
2.4. Définition des zones étudiées .....	33
2.5. Définition du Probleme d'optimisation .....	36
2.5.1. Enoncé du problème .....	36
2.5.2. Les critères d'optimalité .....	37
2.5.2.1. Les critères basés sur la notion d'information.....	37

---

2.5.2.2. Les critères basés sur la notion de couverture.....	39
2.5.2.3. Autres critères .....	42
2.5.3. Choix de l’algorithme d’optimisation.....	44
2.5.3.1. L’algorithme du Recuit Simulé.....	46
2.5.3.1.1. Principe .....	46
2.5.3.1.2. Algorithme .....	47
2.5.3.1.3. Paramétrage.....	48
2.5.3.1.4. Avantages et inconvénients .....	50
2.5.3.2. L’Algorithme Génétique.....	51
2.5.3.2.1. Principe .....	51
2.5.3.2.2. Algorithme .....	52
2.5.3.2.3. Paramétrage.....	54
2.5.3.2.4. Avantages et inconvénients .....	55
2.6. Conclusion du chapitre.....	56
« Bibliographie chapitre 2 ».....	57

### CHAPITRE 3 OPTIMISATION DE RESEAUX DE CAPTEURS POUR

LA CARACTERISATION DE SOURCE.....	64
3.1. Introduction.....	67
3.2. Rappel de l’objectif et hypotheses simplificatrices .....	67
3.3. Problème direct VS Problème inverse .....	68
3.4. Approche basée source VS approche basée récepteur .....	69
3.4.1. Approche basée source .....	70
3.4.2. Approche basée récepteurs .....	71
3.4.3. Avantages de l’approche basée récepteurs dans le cadre des problèmes inverses.....	74
3.5. Choix de la méthode d’inversion de données .....	76
3.5.1. Aperçu des techniques existantes .....	76
3.5.2. La méthode d’assimilation de données renormalisées .....	77
3.5.2.1. Historique de la méthode .....	77
3.5.2.2. Formalisme mathématique .....	78
3.5.2.3. Formulation discrète du problème .....	80
3.5.2.3.1. Technique de renormalisation discrète .....	81
3.5.2.3.2. Proposition d’un algorithme pour le calcul de la matrice de pondération .....	81
3.5.2.3.3. Estimation de la source.....	85

---

3.5.3. Modélisation des rétropanaches.....	87
3.5.3.1. Modélisation analytique et « CFD ».....	87
a. Approche analytique.....	87
b. Approche CFD.....	90
3.6. Evaluation de la methode d'inversion.....	92
3.6.1. Description du projet DYCE .....	92
3.6.2. Mise en œuvre de la méthode de renormalisation .....	96
3.6.3. Résultats et discussions .....	96
3.7. Utilisation de la méthode de renormalisation pour définir des critères d'optimalité .....	99
3.7.1. La méthode de renormalisation dans le cadre d'une approche paramétrique.....	99
3.7.2. La méthode de renormalisation dans le cadre d'une approche entropique .....	101
3.7.3. Définition des problématiques à étudier.....	103
3.7.3.1. Problème 1 « Détermination du nombre optimal de capteurs pour la caractérisation d'une source » .....	104
3.7.3.2. Problème 2 «Déploiement en urgence d'un réseau mobile sans information a priori sur la source ».....	104
3.7.3.3. Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source ».....	104
3.8. Conclusion du chapitre.....	105
« Bibliographie chapitre 3 ».....	106
 CHAPITRE 4 EVALUATION DU PROCESSUS D'OPTIMISATION .....	111
4.1. Introduction.....	114
4.2. Approche analytique « Expérience DYCE » .....	115
4.2.1. Définition du problème d'optimisation .....	115
4.2.1.1. Description des calculs .....	116
4.2.1.1.1. Echelle de complexité.....	116
4.2.1.1.2. Couplage entre l'algorithme de renormalisation et l'algorithme d'optimisation.....	117
4.2.2. Problème 1 «Détermination du nombre optimal de capteurs pour la caractérisation d'une source » .....	118
4.2.2.1. Paramétrage.....	119
4.2.2.1.1. Recuit simulé.....	119
4.2.2.1.2. Algorithme génétique .....	120

---

4.2.2.2. Résultats et discussions .....	120
4.2.2.3. Conclusion à l'issue du problème 1 .....	125
4.2.3. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source » .....	126
4.2.3.1. Paramétrage.....	126
4.2.3.2. Résultats et discussions .....	126
4.2.3.3. Conclusion à l'issue du problème 2 .....	129
4.3. Approche CFD « Expérience MUST » .....	129
4.3.1. Description des expériences.....	130
4.3.2. Modélisation inverse de (Kumar et al, 2015) .....	132
4.3.2.1. Vérification de la simulation inverse .....	134
4.3.3. Définition du problème d'optimisation .....	135
4.3.3.1. Description des calculs .....	135
4.3.3.1.1. Echelle de complexité.....	135
4.3.3.1.2. Couplage entre l'algorithme d'optimisation l'algorithme de renormalisation et les « champs adjoints CFD » .....	135
4.3.4. Problème 1 « Détermination du nombre optimal de capteurs pour la caractérisation de source » .....	136
4.3.4.1. Paramétrage.....	136
4.3.4.2. Résultats et discussions .....	136
4.3.4.3. Conclusion à l'issue du problème 1 .....	149
4.3.5. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source ».....	150
4.3.5.1. Paramétrage.....	150
4.3.5.2. Résultats et discussions .....	150
4.3.5.3. Conclusion à l'issue du problème 2 .....	154
4.3.6. Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source ».....	154
4.3.6.1. Paramétrage.....	155
4.3.6.2. Résultats et discussions .....	155
4.3.6.3. Conclusion à l'issue du problème 3 .....	158
4.4. Conclusion du chapitre.....	159
« Bibliographie chapitre 4 ».....	161

---

Conclusion générale et perspectives.....	162
Annexe 1 .....	166
Annexe 2 .....	178
Annexe 3 .....	185
Annexe 4 .....	192

## Nomenclature

<b>Symbole</b>	<b>Signification</b>
$D_p$	Zone à protéger
$D$	Zone d'intérêt
$D_e$	Environnement poche de $D_p$
$D_d$	Zone de danger
$D_i$	Zone d'implantation
$p$	Position potentielle
$G_I$	Ensemble des positions potentielles
$m$	Positions instrumentées $\leftrightarrow$ nombre de capteurs $\leftrightarrow$ nombre de mesures
$R$	Ensemble des positions instrumentées (réseau de $m$ capteurs)
$p-m$	Positions non-instrumentées
$N_I$	Ensemble des positions non-instrumentés
$k$	Nombre entier
$\mathbf{u}_k$	Positions avec $k = 1 \dots p$
$i, j, k$	Nombres entiers
$\mathbf{u}_i$	Positions avec $i = 1 \dots m$
$Z_I$	Champ mesuré
$g$	Densité de probabilité
$R_o$	Réseau optimal
$C_m^p$	Combinaisons possibles
$\mathbf{Z}_R$	Champ mesuré sur les sites instrumentés
$\mathbf{Z}_N$	Champ mesuré sur les sites non-instrumentés
$S$	Entropie
$I$	Information : indice d'information de Shannon
$E$	Espérance mathématique
$\mathbf{C}_{Z_R}$	Matrice de covariance
$a$	Sensibilité d'un capteur
$A$	Matrice de sensibilité
$d$	Distance
$\mathbf{M}_c$	Matrice de couverture
$a_{min}$	Seuil de sensibilité minimale
$m_j$	Le nombre de capteurs qui sont « sensibles » à un point $x_j$
$n_c$	Le nombre de points du domaine couvert par un réseau
$m_d$	Le nombre total de détections potentielles réalisables par un réseau
$D$	La moyenne des distances entre des sites candidats et un réseau
$D_m$	Dimension fractale
$\alpha$	Nombre réel
$J$	Fonction objectif (coût)
$\bar{c}_k$	Le champ extrapolé au point $k$
$c_k$	Le champ simulé au point $k$
$T$	Température
$T_0$	Température de départ
$T_{i+1}$	Température à l'itération $i+1$
$T_i$	Température à l'itération $i$
$E$	Energie
$K$	Constante de Boltzmann
$T_f$	Température d'arrêt
$P$	Aléas
$P_0$	Probabilité de départ
$\theta$	Facteur de décroissance
$r_a$	Ratio
$L$	Longueur d'un palier de température

$L_{MO}$	Longueur de Monin-Obukhov
$T_p$	Taille d'une population
$Ng$	Nombre de génération
$T_c$	Taux de croisement
$T_m$	Taux de mutation
$q_s$	Debit d'une source
$x_s, y_s, z_s$	Position d'une source
$C$	Concentration
$\Omega$	Domaine
$t$	Temps
$\pi$	Fonction d'échantillonnage
$\mu$	Mesure
$\epsilon$	Erreur de mesure
$\mathbf{O}$	Matrice d'observation
$N$	Nombre de points du maillage
$s$	Source
$u, v, w$	Vitesse respectivement dans le sens des x,y et z
$K_x, K_y, K_z$	Diffusivité turbulente respectivement dans le sens des x,y et z
$q$	Débit
$x_0, y_0, z_0$	Position estimée de la source
$q_0$	Débit estimé de la source
$s_0$	Source estimée
$s_{//}$	Composante parallèle de la source
$s_{\perp}$	Composante perpendiculaire de la source
$\lambda$	Vecteur
$\mathbf{H}$	Matrice de Gram
$f$	Fonction de pondération
$s_{//f}$	Composante parallèle de la source pondérée
$\mathbf{H}_f$	Matrice de Gram pondérée par la fonction $f$
$\varphi$	Fonction de visibilité
$s_{//\varphi}$	Source optimale estimée
$\mathbf{H}_\varphi$	Matrice de Gram optimale pondérée
$\mathbf{W}$	Matrice de pondération
$\mathbf{G}$	Opérateur inverse
$\mathbf{H}_w$	Matrice de Gram pondérée par la matrice $\mathbf{W}$
$\omega$	Réel
$\sigma_x, \sigma_y, \sigma_z$	Ecart-types de la dispersion
$M$	Quantité totale de produit libéré
$x_c, y_c, z_c$	Position centre du bouffé
$a, b, c$	Coefficients dépendants des conditions de stabilité
$f_y, f_z$	Fonctions sans dimensions
$T_y, T_z$	Echelles de temps Lagrangiennes de la turbulence
$H_c$	Hauteur de couche limite
$E_{H_w^{-1}}$	Résidus normalisés



## Liste des figures

Figure 1.1 Schématisation de la dispersion atmosphérique de rejets issus d'une usine .....	5
Figure 1.2 Emissions des oxydes d'azote issus du trafic routier en France en 2004 .....	5
Figure 1.3 Emissions des PM <sub>10</sub> issues des activités d'élevage en France en 2004.....	7
Figure 1.4 Nombre de mois de perte d'espérance de vie -moyenne dans l'UE due aux particules fines (PM <sub>2,5</sub> ) .....	8
Figure 1.5 Schématisation d'un scénario de dispersion d'un produit toxique suite à un accident industriel.....	12
Figure 1.6 Schématisation d'une agression par le biais de la dispersion atmosphérique .....	16
Figure 2.1 Cas de figure d'optimisation de réseaux de capteurs.....	30
Figure 2.2 Questions clés auxquelles la phase d'avant-projet doit répondre .....	32
Figure 2.3 Zones d'études pour un scénario d'agression terroriste .....	34
Figure 2.4 Zones d'études pour un scénario d'accident à l'intérieur d'un site industriel. ....	34
Figure 2.5 Représentation symbolique d'un réseau, d'après (Bocquet, 2006).....	35
Figure 2.6 Processus générale d'optimisation de réseaux de surveillance.....	37
Figure 2.7 Classification des méthodes d'optimisation.....	45
Figure 2.8 Principe de fonctionnement de l'algorithme du recuit simulé.....	48
Figure 2.9 Principe de construction d'une population formée par $k$ configurations de $m$ capteurs.....	52
Figure 2.10 Principe du croisement entre deux configurations de capteurs.....	53
Figure 2.11 Principe de la mutation d'une configuration de capteurs.....	53
Figure 2.12 Principe général de l'algorithme génétique .....	54
Figure 3.1 Problème direct et inverse .....	68
Figure 3.2 Rétropanaches, issus de six capteurs, pour une météo donnée .....	72
Figure 3.3 Approches basée source et basée récepteurs pour la résolution d'un problème inverse à l'aide d'une méthode paramétrique .....	75
Figure 3.4 Algorithme pour le calcul de la fonction de pondération optimale $W$ .....	82
Figure 3.5 Temps de calcul de la fonction de pondération optimale .....	84
Figure 3.6 Principe de fonctionnement de l'algorithme d'inversion.....	86
Figure 3.7 Schématisation, dans le plan xy et à l'échelle de la soufflerie, de la configuration expérimentale.....	92
Figure 3.8 Disposition des quatre FFIDs dans la soufflerie atmosphérique .....	93
Figure 3.9 Ecart types (a) latéraux $\sigma Y/H$ et (b) verticaux $\sigma Z/H$ (Rudd et al, 2012).....	94

Figure 3.10 Résultats obtenus : Ecart types (a) latéraux $\sigma_Y/H$ et (b) verticaux $\sigma_Z/H$ .....	94
Figure 3.11 Résultats obtenus dans le cadre des « Single FFID expériences » .....	95
Figure 3.12 Cas 3 : Fonction de visibilité $\varphi(\mathbf{x})$ , (échelle logarithmique) et $s_{//\varphi}(\mathbf{x})$ .....	97
Figure 3.13 Cas 10 : Fonction de visibilité $\varphi(\mathbf{x})$ , (échelle logarithmique) et $s_{//\varphi}(\mathbf{x})$ .....	97
Figure 3.14 Résidus normalisés cas 3 et cas 10 .....	101
Figure 4.1 Représentation des approches et des problématiques évaluées .....	114
Figure 4.2 Schéma de l'expérience de DYCE.....	116
Figure 4.3 Evolution du nombre de combinaisons en fonction du nombre de capteurs .....	117
Figure 4.4 Couplage entre l'algorithme d'optimisation et l'algorithme de renormalisation.....	118
Figure 4.5 Shémas de perturbation testés .....	120
Figure 4.6 Configuration optimale ( $m=4$ ) obtenue par le recuit simulé et évolution de la fonction objectif.....	121
Figure 4.7 Configuration optimale ( $m=4$ ) obtenue par l'algorithme génétique et évolution de la fonction objectif « fitness » .....	122
Figure 4.8 Configurations optimales ( $m=3, 6, 9$ et $12$ ) obtenues par le recuit simulé et par l'algorithme génétique.....	123
Figure 4.9 Erreurs d'estimations de la position et du débit de la source par le recuit simulé et par l'algorithme génétique.....	124
Figure 4.10 Configuration optimale ( $m=4$ ) et évolution de la fonction objectif.....	127
Figure 4.11 Réseaux optimaux ( $m=3, 5, 9$ et $13$ ) obtenus par le recuit simulé.....	127
Figure 4.12 Erreurs d'estimation de la position et du débit de la source .....	128
Figure 4.13 Photographies de l'expérience de MUST (Mock Urban Setting Test) .....	130
Figure 4.14. Schéma de l'expérience de MUST avec emplacement des 40 capteurs, des caissons métalliques et des sources dans le domaine d'étude.....	131
Figure 4.15 Maillage des deux domaines pour les calculs CFD.....	133
Figure 4.16 Evolution du nombre des combinaisons en fonction du nombre de capteurs.....	135
Figure 4.17 Couplage entre le recuit simulé, l'algorithme de renormalisation et les champs adjoints CFD.....	136
Figure 4.18 Configurations optimales pour $m=10$ et $m=13$ « Essai 6 ».....	137
Figure 4.19 Fonctions de visibilité $\varphi(\mathbf{x})$ des réseaux optimaux $m= 10$ à $m=13$ et les sources $s_{//\varphi}(\mathbf{x})$ pour $m=10$ et $m=13$ .....	137
Figure 4.20 Erreurs d'estimation de la position et du débit de la source « Essai 6 » .....	138
Figure 4.21 Configurations optimales pour $m=10$ et $m=13$ « Essai 8 ».....	140

Figure 4.22 Fonctions de visibilité $\varphi(\mathbf{x})$ des réseaux optimaux $m=10$ à $m=13$ et les sources $s_{//\varphi}(\mathbf{x})$ pour $m=10$ et $m=13$ .....	140
Figure 4.23 Erreurs d'estimation de la position et du débit de la source « Essai 8 » .....	141
Figure 4.24 Configurations optimales pour $m=10$ et $m=13$ « Essai 14 » .....	143
Figure 4.25 Fonctions de visibilité $\varphi(\mathbf{x})$ des réseaux optimaux $m=10$ à $m=13$ et les sources $s_{//\varphi}(\mathbf{x})$ pour $m=10$ et $m=13$ .....	144
Figure 4.26 Erreurs d'estimation de la position et du débit de la source « Essai 14 » .....	145
Figure 4.27 Configurations optimales pour $m=10$ et $m=13$ « Essai 19 » .....	146
Figure 4.28 Fonctions de visibilité $\varphi(\mathbf{x})$ des réseaux optimaux $m=10$ à $m=13$ et les sources $s_{//\varphi}(\mathbf{x})$ pour $m=10$ et $m=13$ .....	147
Figure 4.29 Erreurs d'estimation de la position et du débit de la source « Essai 19 » .....	148
Figure 4.30 Configurations optimales $m=10$ et $m=12$ « Essai 6 » .....	151
Figure 4.31 Configurations optimales $m=10$ et $m=12$ « Essai 8 » .....	152
Figure 4.32 Configurations optimales $m=10$ et $m=12$ « Essai 14 » .....	152
Figure 4.33 Configurations optimales $m=10$ et $m=12$ « Essai 19 » .....	153
Figure 4.34 Configurations optimales $m=8$ , $m=10$ et $m=12$ .....	156
Figure 4.35 Réseaux optimaux $m=10$ VS réseaux (Kumar et al, 2015) : Erreurs de localisation $El(m)$ et erreurs d'estimation du débit $Eq$ de la source .....	158

## Liste des tableaux

Tableau 1.1 Poids du secteur industriel dans les émissions polluantes .....	6
Tableau 3.1 Nombre d'itérations pour obtenir $W$ .....	85
Tableau 3.2 Formulations de Briggs des écarts types de dispersion en milieu rural (Briggs-rural) pour des distances comprises entre 100 et 10 000 m .....	89
Tableau 3.3 Evaluation du modèle directe.....	95
Tableau 3.4 Résultats pour les onze configurations .....	98
Tableau 4.1 Paramétrage du recuit simulé « problème 1 ».....	119
Tableau 4.2 Paramétrage de l'algorithme génétique « problème 1» .....	120
Tableau 4.3 Matrice des structures communes « problème 1 ».....	124
Tableau 4.4 Paramétrage du recuit simulé « problème 2 ».....	126
Tableau 4.5 Matrice des structures communes « problème 2 ».....	129
Tableau 4.6 Caractéristiques des expériences sélectionnées .....	131
Tableau 4.7 Résultats obtenus dans le cadre des travaux de (Kumar et al, 2015) .....	134
Tableau 4.8 « All Measurements » (Kumar et al, 2015), VS « Résultats de nos simulations » .....	134
Tableau 4.9 Paramétrage recuit simulé « problème 1 ».....	136
Tableau 4.10 Essai 6 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015).....	138
Tableau 4.11 Matrice des structures communes « Essai 6 » .....	139
Tableau 4.12 Essai 8 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015).....	141
Tableau 4.13 Matrice des structures communes « Essai 8 » .....	142
Tableau 4.14. Essai 8 : Erreurs obtenues par technique d'optimisation hybride.....	143
Tableau 4.15 Essai 14 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015) .....	145
Tableau 4.16 Matrice des structures communes « Essai 14 » .....	146
Tableau 4.17 Essai 19 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015) .....	148
Tableau 4.18 Matrice des structures communes « Essai 14 » .....	149
Tableau 4.19 Paramétrage recuit simulé « Problème 2 ».....	150
Tableau 4.20 Essai 6 : Erreurs obtenues pour le problème 2 .....	151
Tableau 4.21 Essai 8 : Erreurs obtenues pour le problème 2 .....	152
Tableau 4.22 Essai 14 : Erreurs obtenues pour le problème 2 .....	153

Tableau 4.23 Essai 19 : Erreurs obtenues pour le problème 2 .....	154
Tableau 4.24 Paramétrage recuit simulé « problème 3 ».....	155
Tableau 4.25 Caractérisation des 4 sources : Nos résultats VS les résultats de (Kumar et al, 2015).....	157

# INTRODUCTION GENERALE

Des méthodes de détection de rejets atmosphériques et de *caractérisation* ou de reconstruction de sources en champs proches ont été développées depuis de nombreuses années. Elles s'appuient sur l'utilisation de mesures de concentrations effectuées par des réseaux de capteurs. Ces méthodes font partie de l'ensemble des moyens de surveillance de la contamination de l'air par des agents NRBC (Nucléaire, Radiologique, Biologique et Chimique). Elles sont utilisées pour la protection des sites sensibles face à des contaminations hostiles (une agression terroriste, etc.), accidentelles (une fuite de gaz sur un site industriel, etc.), chroniques (la contamination issue d'une zone industrielle, etc.) ou occasionnelles (la contamination d'origine routière pour des conditions météorologiques spécifiques, pics de pollution, etc.). Les sites sensibles peuvent être, par exemple, des installations militaires sujettes à d'éventuelles agressions extérieures ou des zones d'habitation situées à proximité directe d'installations industrielles classées, telles que les sites SEVESO (à hauts risques), les installations nucléaires civiles, etc. Nous pouvons également envisager le cas où la contamination atmosphérique trouve son origine à l'intérieur du site sensible, par exemple lorsqu'un accident se produit au cœur d'un site industriel de production de produits dangereux ou qu'une attaque terroriste vise un lieu public (gare, quartiers touristiques, zones commerciales, centres d'affaires, etc.).

Dans tous ces cas, il est utile pour les industriels, les autorités locales, les responsables de la sécurité, les autorités militaires, de disposer d'une méthode de positionnement des détecteurs sur la zone sensible et/ou à ses abords de façon à *détecter* un incident potentiel et éventuellement de pouvoir *caractériser* son origine. La *caractérisation* consiste, au minimum, à *localiser* la source de contamination et à *estimer* son intensité. Pour répondre à ce besoin, il est primordial que la conception des réseaux de capteurs soit optimale afin de fournir des données qui permettent une utilisation efficace des méthodes existantes. Dans ce travail, nous nous intéresserons plus particulièrement à l'une d'entre elles : l'assimilation de données renormalisées. Cette méthode, dite de « renormalisation », est une technique déterministe, facilement utilisable en situation opérationnelle.

### **Objectif de la thèse**

L'objectif de cette thèse est de définir les moyens nécessaires pour optimiser un réseau de surveillance conçu pour la *caractérisation* de sources de rejets atmosphériques. L'optimisation consiste à déterminer les emplacements optimaux des capteurs afin de répondre à ce besoin.

Nous envisageons, également, d'optimiser le nombre de capteurs à déployer. Les principales questions auxquelles ce travail tente d'apporter des réponses sont les suivantes :

1. Quel critère mathématique permet d'évaluer la performance d'une configuration de réseau utilisée pour la *caractérisation* de source ? Ce critère, nommé « critère d'optimalité », doit être défini en fonction de la méthode d'inversion choisie.
2. Une fois le critère choisi, quelle démarche utiliser pour trouver la configuration optimale d'un réseau ?
3. Comment configurer, paramétrer, mettre en œuvre, évaluer les méthodes et les algorithmes retenus.

Pour atteindre cet objectif et répondre aux questions précédentes, ce travail de thèse a été réalisé sous un ensemble d'hypothèses :

- La zone sensible à protéger s'étend sur quelques kilomètres au plus,
- La zone sensible et son environnement proche forment un ensemble faiblement bâti avec au plus quelques dizaines de bâtiments,
- Le contaminant détecté est considéré émis à l'extérieur des locaux. Ceci limite les scénarios des accidents ou des attaques considérés. Souvent, une libération accidentelle trouve sa cause à l'intérieur d'un bâtiment. Par cette hypothèse nous évitons ainsi d'avoir à modéliser la dispersion intérieure.
- La source de contamination est supposée a priori statique, ponctuelle et localisée à des hauteurs peu importantes. Egalement, nous ne considérons pas le cas où la contamination provient de plusieurs sources distinctes.

Les hypothèses précédentes permettent de simplifier les configurations réalistes souvent très complexes. Les faibles dimensions de la zone sensible et de son environnement ont deux conséquences :

- Les temps caractéristiques de la dispersion sont relativement courts et les conditions météorologiques varient peu. Nous évaluerons donc les configurations des réseaux pour des « situations » météorologiques stationnaires.
- Chaque situation peut-être décrite à partir d'un faible nombre de données telles que la vitesse moyenne et la direction du vent, la classe de stabilité atmosphérique.



## Originalité de la thèse

Une importante étude bibliographique, synthétisée dans ce manuscrit, a montré que : (i) Le problème de l'optimisation des réseaux de surveillance de la contamination de l'air et le problème de la reconstitution de source (problème inverse en dispersion atmosphérique) ont tous deux été largement abordés dans la littérature. (ii) Ces deux problèmes sont généralement traités séparément et n'ont été que très rarement considérés ensemble. Dans cette étude ils seront considérés conjointement.

De plus, cette étude s'inscrit parmi plusieurs travaux de recherche qui s'effectuent autour de la méthode d'assimilation de données renormalisées. Cette technique est ici utilisée pour la première fois dans le cadre de problèmes d'optimisation de réseaux de capteurs. Nous montrerons qu'elle permet naturellement de définir des critères d'optimalité.

Du point de vue opérationnel, la méthodologie que nous proposons est suffisamment générale pour permettre de concevoir des réseaux efficaces dans des situations variées. Elle peut être, par exemple, utilisée comme un outil d'aide à la décision dans des situations de crise où les autorités doivent évacuer une zone à risque, neutraliser une attaque, localiser une fuite toxique, etc.

## Plan de la thèse

Ce manuscrit comporte quatre chapitres :

Dans le premier, nous réalisons un inventaire (non exhaustif) des besoins liés à la surveillance de la contamination de l'air par des agents NRBC. Pour cela nous avons distingué les besoins environnementaux, industriels, militaires et sécuritaires. Ce chapitre permet d'identifier les besoins spécifiques de chaque domaine et de bien préciser lesquels seront abordés dans cette thèse.

Dans le second, nous définissons le problème d'optimisation des réseaux de capteurs puis nous présentons les critères d'optimalité les plus couramment utilisés dans la littérature. Nous proposons ensuite une méthodologie pour le résoudre et en particulier les moyens algorithmiques à mettre en œuvre afin d'obtenir une solution optimale.

Le troisième chapitre est consacré à la description de l'objectif de la *caractérisation* de source. Pour cela, dans un premier temps, le problème d'Estimation du Terme Source (STE problem, en anglais) est défini. Dans un deuxième temps, les différentes approches permettant de le

résoudre sont présentées. Une distinction entre les approches « basées sources » et « basées récepteurs » est proposée. Le reste du chapitre est consacré à la présentation, à l'analyse, à la mise en œuvre et à l'évaluation de la méthode de renormalisation, retenue pour ce travail. Nous montrons comment cette méthode nous a permis de définir des critères d'optimalité originaux. Enfin, trois types de problèmes sont définis. L'ensemble de la démarche d'optimisation proposée sera évalué sur ces derniers.

Le quatrième chapitre est consacré à cette évaluation. Nous présentons les résultats obtenus sur des cas concrets recouvrant les différents types de problèmes envisagés. Les modèles de dispersion atmosphérique utilisés pour résoudre le problème inverse sont soit des modèles analytiques, soit des modèles issus de la mécanique des fluides numérique (Computational Fluid Dynamics (CFD) en anglais). Les résultats de ces évaluations sont analysés, synthétisés et critiqués.

**CHAPITRE 1 :**  
**LES RESEAUX DE CAPTEURS ET LES ENJEUX**  
**ENVIRONNEMENTAUX, INDUSTRIELS,**  
**MILITAIRES ET SECURITAIRES**

**CHAPITRE 1**  
**LES RESEAUX DE CAPTEURS ET LES ENJEUX**  
**« ENVIRONNEMENTAUX, INDUSTRIELS, MILITAIRES ET**  
**SECURITAIRES**

**Plan du chapitre 1**

CHAPITRE 1 LES RESEAUX DE CAPTEURS ET LES ENJEUX « ENVIRONNEMENTAUX, INDUSTRIELS, MILITAIRES ET SECURITAIRES .....	1
1.1. Introduction.....	3
1.2. Contexte environnemental.....	3
1.2.1. Générateurs de risques .....	4
1.2.2. Conséquences possibles.....	7
1.2.3. Enjeux et besoins .....	9
1.3. Contexte industriel.....	11
1.3.1. Générateurs de risques .....	12
1.3.2. Conséquences possibles.....	13
1.3.3. Enjeux et les besoins.....	14
1.4. Contexte militaire et sécuritaire .....	16
1.4.1. Générateurs de risques .....	16
1.4.2. Effets possibles.....	21
1.4.3. Enjeux et besoins .....	21
1.5. Conclusion du chapitre.....	22
« Bibliographie chapitre 1 ».....	25

## 1.1. INTRODUCTION

L'objectif de ce premier chapitre est de faire un inventaire (le plus exhaustif possible) des besoins et des enjeux liés à la contamination de l'air par des agents NRBC, dans les contextes environnementaux, industriels, militaires et sécuritaires. Ce travail doit nous permettre d'identifier les besoins spécifiques, en termes de configuration de réseaux de capteurs, de chaque domaine et de mettre en évidence ceux qui relèvent de cette étude.

Nous présentons ces besoins en nous focalisant à la fois sur les aspects préventifs (anticipation des risques par une détection précoce, etc.) et sur les aspects opérationnels (gestion des situations d'urgence, etc.). En premier lieu, nous définissons les générateurs de risque probables, nous détaillons également leurs effets possibles. Enfin, nous définissons les enjeux spécifiques à chaque domaine et nous identifions ceux auxquels nous tenterons d'apporter une solution.

## 1.2. CONTEXTE ENVIRONNEMENTAL

En matière d'environnement, la contamination de l'air relève de la problématique générale de la pollution atmosphérique. Contrairement à l'eau, qui est traitée avant d'être consommée par l'ensemble de la population, l'air que l'on respire n'est pas filtré et les conséquences de sa pollution ont un caractère irréversible. A titre d'exemple, l'air que l'on respire sur le chemin du travail, ou celui que respirent les enfants en se dirigeant vers leurs écoles chaque jour, ne peut pas être filtré ou traité. La pollution, quelle que soit son origine, a des impacts négatifs sur les humains, mais également sur la faune et la flore.

Les problèmes environnementaux peuvent apparaître au niveau local (émissions nocives en champ proche) mais aussi au niveau mondial (champ global) et avoir une incidence sur le changement climatique (problème de réchauffement lié à l'accroissement de la concentration des gaz à effet de serre) dans l'atmosphère terrestre (*Ministère de l'Environnement, 2006*) Cependant, dans ce travail, seuls les aspects locaux seront pris en compte.

Il est clair que, dans un contexte environnemental, la contamination chronique de l'air à l'échelle locale influe directement sur la santé des individus. Ainsi, il est indispensable que les autorités locales ou régionales (particulièrement au niveau des grandes métropoles) veillent sur le respect des seuils réglementaires définis pour préserver la qualité de l'air et signalent et pénalisent tout, éventuel, dépassement des normes. Dans ce cadre, le

déploiement de réseaux de surveillance adapté est indispensable (*RECORD, 2013*). La figure 1.1 schématise une source de pollution industrielle, menaçant son entourage proche.

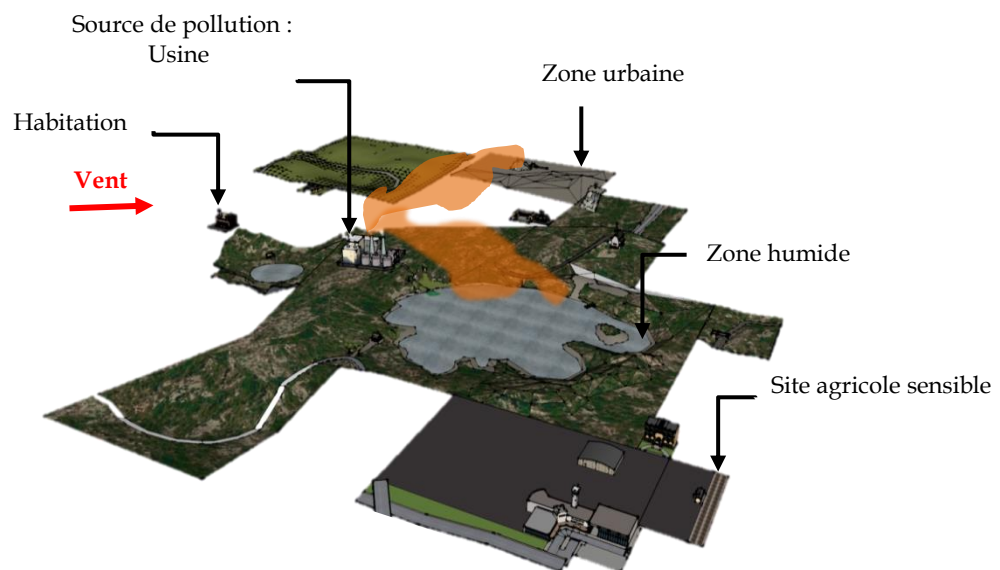


Figure 1.1. Schématisation de la dispersion atmosphérique de rejets issus d'une usine

### 1.2.1. Générateurs de risques

En général, la contamination conduisant à la pollution de l'air a deux origines possibles : la première naturelle, la seconde humaine (*PREV'AIR*). Parmi les sources naturelles de pollution les volcans qui rejettent des fumées chargées en dioxyde de soufre,  $SO_2$ , et qui dégagent de la chaleur favorable à la formation d'oxyde d'azote,  $NO_x$ . Il est à noter que de façon plus subjective, les gaz dégagés peuvent avoir une odeur gênante pour la population (*Grahn, 2015*). Les sources de pollution artificielles proviennent de l'ensemble des activités humaines et engendrent une modification de la composition naturelle de l'air. A ce stade, nous évoquons, essentiellement, les pollutions issues des transports, des activités industrielles et des habitations.

- Pollution liée aux transports

Cette pollution provient principalement des émissions des gaz d'échappement des véhicules à moteurs (véhicules particuliers, camions, bus, etc.). Cette forme de pollution est quasi-permanente et présente des pics très dangereux (horaires de forte affluence, conditions météorologiques défavorables).

Les polluants émis par ces équipements sont généralement classés en deux catégories : les polluants réglementés (PPR) et les non réglementés (PNR) (*Pillot et al, 2006*). Les PPR sont des polluants dont la mesure est indispensable pour l'homologation des véhicules et regroupent : les oxydes d'azote,  $\text{NO}_x$ , le dioxyde de carbone,  $\text{CO}_2$ , l'oxyde de carbone, CO, le dioxyde de soufre  $\text{SO}_2$ , le benzène, le plomb et les Matières Particulaires (PM). Par contre, la mesure des PNR n'est pas obligatoire. Parmi eux nous trouvons le protoxyde d'azote,  $\text{N}_2\text{O}$ , les aldéhydes, les cétones, etc. Les concentrations dans l'air des gaz cités ci-dessus se mesurent couramment, en temps réel (*IBGE-LRE, 2006-2008*), à l'aide de capteurs et d'analyseurs automatiques qui peuvent utiliser divers principes selon les polluants (variation de conductivité, chromatographie, spectrométrie, chimiluminescence, etc.). Les particules fines se mesurent, elles, après une analyse d'un échantillon d'air, qui permet d'obtenir une information très précise sur la quantité de particules et sur leurs tailles. Sur la figure 1.2, nous pouvons distinguer les quantités des  $\text{NO}_x$  (en tonnes) issues du trafic routier en France pendant l'année 2004.

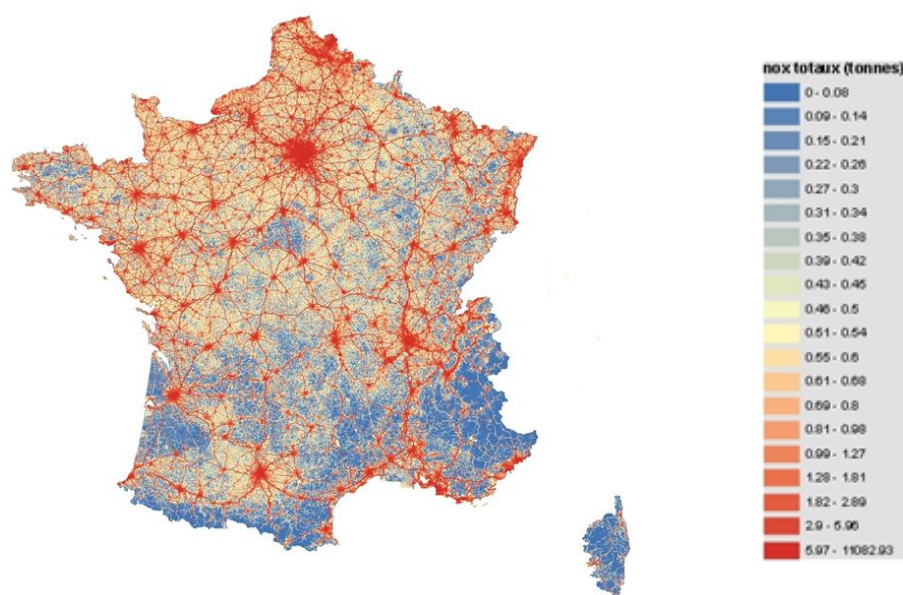


Figure 1.2. Emissions des oxydes d'azote issus du trafic routier en France en 2004. (*INS, 2004*)

- Pollution d'origine industrielle

La pollution de l'air est présente à toutes les étapes d'un processus industriel ordinaire: en amont (lors de l'extraction et du transport des matières premières), au cours de la transformation de ces matières et en aval (traitement des déchets et des résidus générés). Les pollutions sont de nature variée et leurs dangers dépendent de la branche industrielle dont

elles sont issues. L'atteinte à l'environnement peut provenir d'émissions radioactives ou gazeuses chargées d'espèces chimiques nocives (parmi lesquelles on retrouve celles déjà émises par les véhicules) ou de particules en suspension, etc. Les rejets chargés de résidus solides peuvent être composés d'imbrûlés carbonés, d'agents de traitement des gaz qui n'ont pas réagi, de cendres volantes captées sur le parcours des gaz éjectés, d'émissions solides (poussières) rejetées en sortie de cheminées, etc (*CITEPA*). Ces particules se mesurent avec des capteurs à diffraction laser déterminant leurs fractions massique et leurs tailles exactes.

Cependant, l'identification précise de la composition de l'ensemble de particules en suspension nécessite la connaissance du processus industriel et des matières manipulées. Il en est de même pour l'identification des espèces nocives, émises sous forme gazeuses : elles nécessitent un inventaire au cas par cas. Par exemple, pour des industries manipulant du charbon, des rejets gazeux sont émis dans l'atmosphère dès son extraction et son stockage (le méthane,  $\text{CH}_4$ , contenu dans les charbons s'échappe lorsque ceux-ci sont mis en contact avec l'atmosphère) et se poursuivent lors de son utilisation (combustion) (*Bobin et al, 2005*). Un autre exemple concerne les industries chimiques qui produisent des acides tels que l'acide sulfurique  $\text{H}_2\text{SO}_4$  ou nitrique  $\text{HNO}_3$ , et dont les effluents gazeux peuvent être chargés en  $\text{SO}_2$ ,  $\text{SO}_3$ , ou en  $\text{NO}_x$ , etc (*AWAC*). Le tableau 1.1 montre la part du secteur industriel dans les émissions de certains polluants ( $\text{SO}_2$ ,  $\text{NO}_x$ , COVNM : composés organiques volatils non méthaniques, CO et le  $\text{CO}_2$ ) en France en l'an 2000.

Polluant	Part de l'industrie en France
$\text{SO}_2$	81 %
$\text{NO}_x$	22 %
COVNM	20 %
CO	20 %
$\text{CO}_2$	35 %

Tableau 1.1. Poids du secteur industriel dans les émissions polluantes (*CITEPA, 2000*)

- Pollution liée aux habitations

On considère souvent, à tort, que la pollution de l'air par le secteur résidentiel a un faible impact sur l'environnement (*CITEPA, 2017*). Nous signalons que, dans ce cadre, on distingue la pollution de l'air intérieur (pollution des environnements clos comme les habitations, les établissements médico-sociaux, etc.) et la pollution de l'air extérieur générée par les habitations et les résidences. Cette dernière provient des pratiques liées à la vie courante, comme par exemple, le conditionnement et le refroidissement de l'air, le chauffage des habitations, la cuisson de la nourriture, les feux de cheminée ou de jardin, etc.



Les principaux polluants émis par le secteur domestique sont des particules dont la majorité est issue de la combustion du bois, notamment du fait de l'utilisation d'équipements individuels anciens, peu performants ou à cause de pratiques non optimisées (foyer ouvert ou mal isolé).

Il est intéressant de noter que certaines classifications (*PREV'AIR*) de générateurs de risque de pollution de l'air considèrent que l'agriculture est un quatrième secteur à prendre en compte. Les principales sources sont les activités liées à l'élevage (stockage et épandage des déchets des animaux), au pâturage, aux pratiques liées aux cultures (épandage d'engrais chimiques, etc.). Sur la figure 1.3, nous pouvons distinguer les quantités (en kg) des particules PM<sub>10</sub> (particules en suspension dans l'air d'un diamètre aérodynamique inférieur à 10 micromètres) issues des activités d'élevage en France pendant l'année 2004. D'autres sources sont les équipements de combustion et de réfrigération, les machines agricoles, etc. Pour ce secteur, les principaux rejets sont les NO<sub>x</sub>, l'ammoniac NH<sub>3</sub> et les particules fines PM<sub>10</sub>, PM<sub>2,5</sub>, etc.

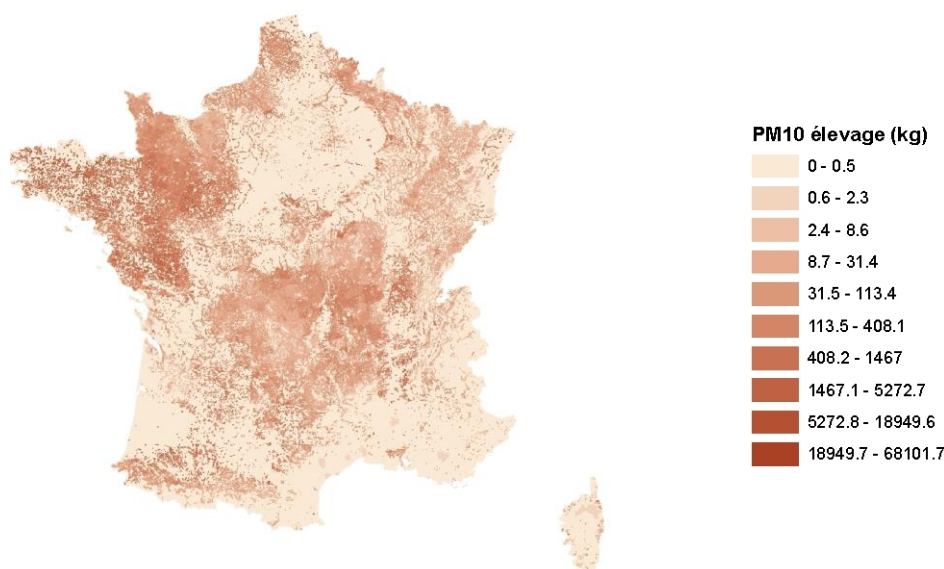


Figure 1.3. Emissions des PM<sub>10</sub> issues des activités d'élevage en France en 2004. (*INS, 2004*)

### 1.2.2. Conséquences possibles

Les effets possibles de l'altération des niveaux de qualité et de pureté de l'air peuvent être regroupés en trois familles: les répercussions sur l'environnement, les effets sur l'être humain (*AIRPARIF-a*) et enfin les effets économiques.

- Conséquences environnementales

Au-delà des risques sanitaires, la pollution atmosphérique a un immense impact environnemental (*Ministère de la Transition écologique et solidaire, 2017*). Une forte densité d'émissions de polluants (surtout les pollutions récurrentes) entraîne des dégâts considérables sur les écosystèmes. Nous pouvons assister, par exemple, à l'adynamie des forêts, à la diminution de la croissance des végétaux, à la fonte des glaces polaires, à la réduction des territoires et des ressources nécessaires à la survie de la faune, à la contamination de la chaîne alimentaire, à la baisse des rendements de cultures, etc.

- Conséquences humaines

Sur l'être humain, la mauvaise qualité de l'air a un impact à court et à long terme (*Ministère de la Transition écologique et solidaire, 2017*). La santé des individus fréquemment exposés, particulièrement les personnes sensibles comme les personnes asthmatiques et les personnes âgées, est fortement menacée. En effet, vivre dans un milieu pollué provoque des effets malheureusement fatals, réduit considérablement l'espérance de vie (*AIRPARIF-a*) et entraîne un accroissement des maladies respiratoires et cardiovasculaires et donc une diminution de l'espérance de vie, etc. Sur la figure 1.4, nous pouvons examiner le nombre de mois de perte d'espérance de vie -moyenne dans l'Union Européenne due aux particules fines  $PM_{2,5}$ .

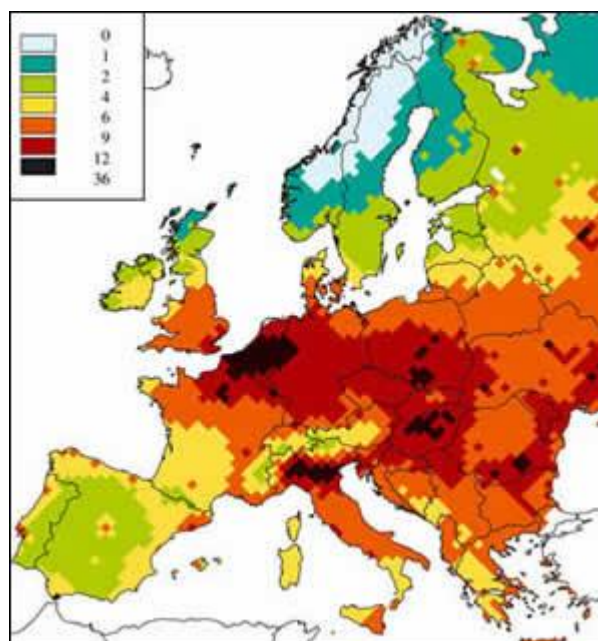


Figure 1.4. Nombre de mois de perte d'espérance de vie -moyenne dans l'UE due aux particules fines ( $PM_{2,5}$ ). (*IIASA*)

- Conséquences économiques

La pollution entraîne également des conséquences économiques. De nombreuses études à l'échelle européenne et mondiale (*AÏCHI, 2015*) constatent que le dépassement du niveau moyen annuel de pollution coûte, chaque année, des milliards d'euros en dépenses de santé, en pertes de journées de travail pour cause de maladie, en coûts associés à la perte de qualité de vie et en baisse d'espérance de vie. Selon le ministère de l'Ecologie, du Développement durable et de l'Energie Français, en 2012, les coûts de santé liés à la pollution de l'air, en France, s'élevaient annuellement entre 20 à 30 Mds €, ce qui correspond à environ 400 à 500 €/an/habitant.

Pour les organismes industriels, les répercussions économiques se traduisent, également, sous forme de taxes environnementales à payer par tonnes d'émission. Par exemple, l'émission de dioxyde de soufre SO<sub>2</sub> coûte de 5600 à 16 000 € par tonne. L'émission des particules fines PM<sub>2,5</sub> coûte 26 000 à 75 000 € par tonne (*CAFE, 2005*).

### 1.2.3. Enjeux et besoins

Selon Président du Conseil National de l'Air (*Saddier, 2011*), la qualité de l'air représente l'un des enjeux majeurs du XXIe siècle. En effet, le combat contre la pollution atmosphérique est à la fois un défi pour la santé publique et l'environnement. Nous avons précédemment (partie 1.2.2) donné un aperçu des différents impacts de la pollution de l'air : sur la vie quotidienne de l'être humain, sur le climat, la faune et la flore. Nous présentons maintenant les besoins spécifiques des acteurs nationaux et régionaux liés à la prévention et à la gestion de crises.

- Aspects préventifs

Les aspects préventifs regroupent toutes les mesures de fond permettant d'éliminer ou de réduire le problème de la pollution atmosphérique. D'un point de vue pratique, la meilleure stratégie de prévention des risques est la prévention à la source. Pour ce faire, des réglementations, des normes et des taxes (spécifiques à chaque type de polluants) sont fixées par les autorités pour imposer la réduction des rejets de polluants dans l'air (*Directive européenne n°2008/50/CE, 2008*). Il est utile de noter que certaines actions préventives peuvent aller jusqu'à l'interdiction du rejet de certains polluants ou à l'imposition d'instruments spécifiques à la lutte contre la pollution de l'air. Par exemple, les autorités peuvent imposer aux industriels et aux particuliers de s'équiper, à leurs charges d'outils de traitement des

rejets fortement coûteux. On peut citer les filtres à particules (FAP) pour les voitures à moteur diesel ou les filtres électrostatiques pour les cimenteries, etc.

Pour veiller au respect et au non violation des réglementations, des contrôles périodiques doivent être effectués et des plans de surveillance de la qualité de l'air doivent être adoptés. Les plans de surveillance peuvent être assurés par des réseaux de mesures conçus et architecturés spatialement sur une zone d'intérêt pour répondre à cet objectif. Par exemple en région PACA, 53 stations de mesure surveillent les concentrations en SO<sub>2</sub>. En cas de dépassement des seuils, le dispositif STERNES (Système Temporaire d'Encadrement Réglementaire Normatif des Emissions Soufrées) ([DRIRE P. ACA, 2004](#)) est déclenché afin de limiter les rejets soufrés. Lors du déclenchement du dispositif les industriels sont contraints de respecter des quotas d'émissions plus faibles, notamment en utilisant des combustibles moins soufrés. Les entreprises concernées par ce dispositif sont au nombre de 11 sur les Bouches-du-Rhône (SMP chimie, SMP raffinerie, BP raffinerie, Naphtachimie, EDF Ponteau, TOTAL, SOLLAC, ESSO raffinerie, CABOT, LAFARGE Aluminates et SETCM). Il est possible d'envisager une évolution de ce type de dispositif en utilisant, par exemple, les données issues du réseau de mesure afin de permettre de *localiser* et *d'identifier* le, ou les, acteur(s) le(s) plus impliqué(s) dans l'épisode de pollution.

- Aspects opérationnels

Les aspects opérationnels regroupent les procédures et les efforts mis en service pour *informer* et *alerter*, en cas de pic de pollution et pour mettre en œuvre un plan d'urgence adéquat. Le niveau de pollution est mesuré par des réseaux de capteurs qui surveillent la qualité de l'air en temps réel et détectent le dépassement du seuil critique d'un ou de plusieurs polluants. Par exemple, pour surveiller la qualité de l'air, l'association de surveillance de la qualité de l'air en Ile-de-France AIRPARIF dispose d'environ 70 stations de mesure ([AIRPARIF-b](#)). Ce réseau est constitué de plus d'une cinquantaine de stations automatiques permanentes et de nombreuses stations semi-permanentes positionnées à proximité du trafic. Elles sont réparties sur un rayon de 100 km autour de Paris et mesurent la qualité de l'air respiré par la population (plus de 11 millions d'habitants dans toute la région). Les mesures issues de tels réseaux peuvent être associées à différents types de modèles de qualité de l'air (modèles gaussiens, modèles eulériens) qui simulent les champs de concentrations sur la zone concernée en fonction de mesures météorologiques et en utilisant des bases de données d'émissions. Lorsqu'un risque imminent pour la santé est

diagnostiqué, des procédures adaptées sont mises en place : règles de circulation alternée pour les véhicules, réductions exceptionnelles des vitesses de circulation, gratuité du stationnement résidentiel, interdiction de circulation sur une ou sur plusieurs routes fortement polluantes, mise à l'arrêt d'usines, etc. Des consignes sont également adressées à la population, surtout aux individus sensibles, afin d'adopter un comportement adapté (éviter les actions physiques en extérieur, favoriser un circuit propre, suivre des traitements médicaux à visée respiratoire, etc.) ou de participer au processus de limitation de la pollution (réduire les consommations d'énergie, privilégier les transports en commun, etc.).

A un niveau décisionnel, un réseau de capteurs permet *d'informer* les autorités de la nécessité de réduire les émissions de polluants. Il permet aussi de *valider* et de *recaler* les modèles de qualité de l'air. Dans ce contexte, l'information du dépassement des seuils critiques peut être complétée par la *localisation* et la *quantification* de la source principale de pollution. En effet, il est impératif de *vérifier* que les données utilisées en entrée des modèles de simulation sont correctes. Ce besoin concerne directement notre étude. Nous tenterons donc de proposer des solutions pour l'optimisation de réseaux de détection en vue de la *localisation* et de la *quantification* des sources de pollution atmosphérique.

### 1.3. CONTEXTE INDUSTRIEL

Sur les sites industriels, le risque lié à la contamination de l'air par des agents NRBC est toujours présent. Par exemple, un accident au cœur d'un site industriel de production ou de manipulation de produits dangereux peut entraîner directement et/ou indirectement des conséquences immédiates graves pour le personnel à l'intérieur du site, pour les populations avoisinantes et pour les biens et l'environnement (*ARIA*). De ce fait, l'aspect sécuritaire face à ce phénomène est une des préoccupations majeures des industriels. Il est indispensable pour eux d'être efficace et rapidement opérationnel lors de sinistres auxquels ils sont susceptibles de faire face.

La figure 1.5 schématise un scénario d'accident sur un site industriel conduisant à la dispersion d'un gaz toxique mettant en danger la vie du personnel.

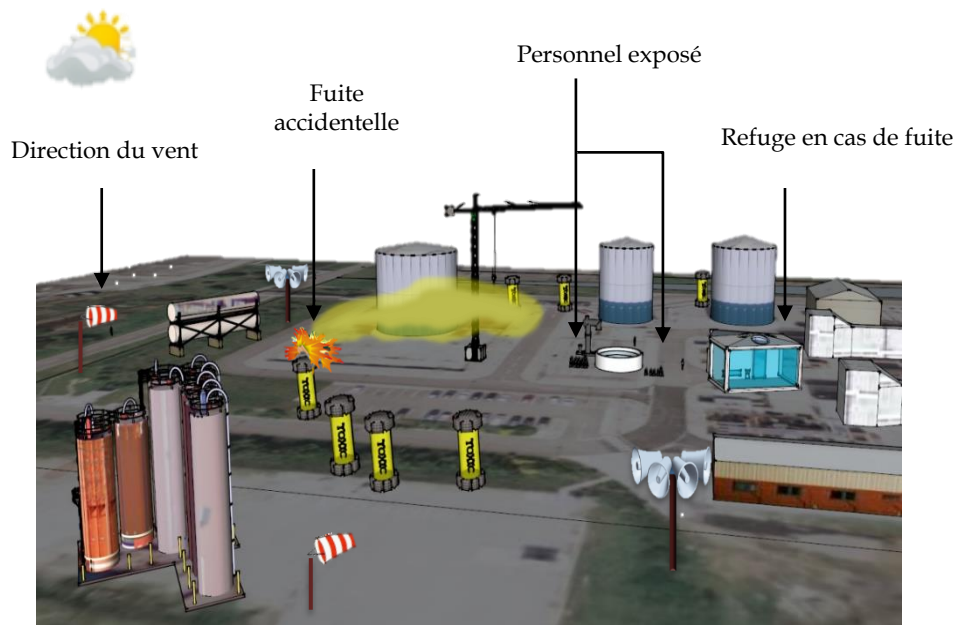


Figure 1.5. Schématisation d'un scénario de dispersion d'un produit toxique suite à un accident industriel

### 1.3.1. Générateurs de risques

Dans les milieux industriels, les scénarios de contamination de l'air sont nombreux et dépendent de la nature de l'activité, de la nature des produits présents (toxiques, inflammables, explosifs, etc.), des installations (matériaux, modes de stockage, etc.), du facteur humain (négligence, erreur d'appréciation, etc.) et aussi des phénomènes extérieurs comme les inondations, les séismes, les accidents d'industrie voisine, etc.

La variété des scénarios, et surtout la possibilité que plusieurs accidents puissent avoir lieu d'une manière simultanée ou enchaînée, rend complexe l'identification des générateurs de risques. Nous avons choisi de les regrouper en trois grandes familles, suivant le type d'activité industrielle. Nous distinguons : les industries chimiques, les industries pétrochimiques et l'industrie nucléaire civile.

- Risques Chimiques

L'activité de l'industrie chimique consiste à fabriquer des produits chimiques de base, des produits chimiques intermédiaires et des produits finis comme ceux destinés à l'agroalimentaire (notamment les engrais), les produits pharmaceutiques, les produits de consommation courante (alcool, eau de javel, etc.), les polymères, etc.



En cas d'accidents, les rejets peuvent être des substances toxiques et/ou corrosives tels que le sulfure d'hydrogène, H<sub>2</sub>S, l'ammoniac NH<sub>3</sub>, l'acide chlorhydrique, HCL, le dioxyde de soufre SO<sub>2</sub>, l'acétylène, le chlore, des vapeurs d'acides sulfurique ou phosphorique ou nitrique, des substances inflammables comme l'hydrogène H<sub>2</sub>, des substances inertes ou peu nocives comme l'argon, Ar, et enfin des oxydants tel que le O<sub>2</sub>.

- Risques pétrochimiques

L'activité de l'industrie pétrochimique consiste à produire l'ensemble des produits dérivés du pétrole comme l'essence, les goudrons, les lubrifiants, le gaz de pétrole liquéfié (GPL), etc. En cas d'incidents, les rejets sont essentiellement des produits fortement inflammables tel que le gaz naturel, le méthane, le propane, l'acétylène, des vapeurs de fuel de kérosène, etc.

- Risques nucléaires

L'industrie nucléaire civile a pour objectif la production d'énergie électrique à partir de l'énergie nucléaire. Cette industrie est potentiellement une source de contamination accidentelle de l'air par des éléments radioactifs. Les combustibles nucléaires utilisés sont des matières fissiles telles que l'uranium et le plutonium. Dans ce cadre, un risque est également présent lors de l'extraction minière de l'élément radioactif, sur le trajet de son transport, au cours de sa préparation (enrichissement) et pendant la phase de son stockage (avant et après utilisation).

### 1.3.2. Conséquences possibles

Les conséquences d'un accident dans un milieu industriel peuvent être classées selon les typologies d'effets: effets toxiques, effets thermiques (dispersion de produits inflammables conduisant à des combustions et/ou des explosions) et effets mécaniques (ondes de chocs engendrant des conséquences graves sur l'environnement proche et sur les personnes) ([Gouvernement.fr](http://Gouvernement.fr)), mais peuvent être classées également d'une façon plus générale, en conséquences sur l'environnement, sur l'humain et sur l'économie :

- Conséquences environnementales

Un accident industriel peut avoir des répercussions importantes sur les écosystèmes ([INERIS, 2015](http://INERIS)). Nous pouvons assister à une contamination de la faune et de la flore pouvant avoir un impact sanitaire (pollution d'une nappe phréatique, etc.). Dans ce cadre, Un exemple d'accident assez désastreux est celui de Fukushima ([IRSN, 2016](http://IRSN)).

- Conséquences humaines

Une fuite de substance toxique peut, par inhalation, par contact avec la peau ou les yeux, ou par l'ingestion, provoquer de graves lésions (et même causer le décès) sur les personnes physiques directement ou indirectement exposées aux conséquences de l'accident (par exemple la catastrophe de Bhopal en 1984 (*Gautier, Y*) . Les effets peuvent être, par exemple, un gonflement aigu du poumon, une atteinte du système nerveux ou des brûlures chimiques cutanées ou oculaires.

- Conséquences économiques

Un accident industriel majeur peut altérer l'outil économique d'une zone. En plus de la possible destruction des bâtiments, des routes ou des voies de chemin de fer voisines du lieu de l'accident, la zone peut être isolée pour une période donnée. Dans ce cas, les conséquences économiques peuvent être désastreuses (*IRSN*).

### 1.3.3. Enjeux et les besoins

La surveillance de la contamination de l'air dans les milieux industriels est un besoin capital et nécessite des stratégies de surveillance optimisées (*Pineau, 1995*):

- Enjeux préventifs

Les enjeux préventifs, dans le cadre de la contamination accidentelle de l'air par des agents NRBC concernent les stratégies de *surveillance* et de *mise en alerte* en cas d'incident. Le but est de détecter les fuites, les accidents et les anomalies liés aux activités industrielles et de déclencher les opérations de réparation, d'évacuation, de secours et de sauvetage (selon la gravité de la situation) le plus rapidement possible. Dans ce cadre, le besoin que les industriels associent aux réseaux de mesures est l'*alerte précoce*. Un réseau de capteurs est d'autant plus efficace qu'il informe tôt et, par conséquent, laisse une large marge de manœuvre pour lancer les plans d'urgence. Un autre type d'enjeu concerne la *surveillance* des niveaux de concentration émis par le site (contrôle réglementé des émissions) et, éventuellement, le dégagement de responsabilité du site si un accident se produit ailleurs. A plus grande échelle, on peut noter l'existence de réseau OPERA-AIR (prélèvement d'air sur filtre fixe) qui surveille la radioactivité des aérosols atmosphériques sur le territoire français. Constitué de plus d'une quarantaine de stations, dont 32 à proximité des installations nucléaires, il permet de *décélérer* et de *caractériser* toute élévation anormale de la radioactivité de l'air.



- Enjeux opérationnels

Les enjeux opérationnels en cas de contamination accidentelle de l'air sont principalement associés aux mesures d'urgence à appliquer en cas d'accident et dépendent du type d'industrie concernée. Par exemple, il y a en France près de 1205 installations classées SEVESO (dont 657 « seuil haut ») (*Radisson, 2015*) pour lesquels le risque technologique doit être diagnostiqué et maîtrisé. En Europe, le texte de loi (*Directives Seveso 1,2 et 3*) n'impose pas directement, pour l'instant, de mise en place de réseaux de surveillance de la contamination accidentelle de l'air. Cependant certains grands industriels du domaine chimique et pétrochimique ont déjà doté leurs installations majeures de réseaux de mesures. A terme, et compte tenu du vieillissement des installations en France et en Europe, il est envisagé un renforcement de la législation pour la *surveillance* des installations qui peut se traduire par l'imposition de réseaux de mesures pour la *détection* et *l'identification*, sur tous les sites à risques. Par conséquent, les besoins sont nombreux et peuvent concerner l'intérieur ou l'extérieur du site.

Le premier besoin, déjà évoqué ci-dessus est la *détection*. Le deuxième besoin est *l'identification*, qui est nécessaire pour certaines usines complexes qui manipulent différents produits dangereux ou inflammables (les raffineries par exemple). Dans ce cas, il s'agit *d'identifier* le ou les produits dispersés. Le troisième besoin, qui concerne directement notre étude, est la *localisation* précise de la source surtout pour les usines étendues sur une grande superficie. Il est également impératif de *quantifier* l'intensité des émissions afin de pouvoir *estimer* les quantités émises. La *localisation* permet l'application efficace de mesures correctives (neutralisation). *L'estimation* des quantités dispersées permet de soutenir les opérations d'évaluation des risques (niveaux d'exposition, action médicale, etc.). Une fois la source localisée et quantifiée, il est également possible de faire de la *prévision*, à l'aide de modèles de dispersion atmosphérique, du transport et de la dispersion des produits concernés. L'objectif est de déterminer les zones susceptibles d'être impactées et de lancer les actions adaptées comme, par exemple, l'évacuation des zones urbaines en cas de dispersion accidentelle de radionucléides à partir d'une centrale nucléaire civile (*Abida, 2010*). Après l'accident, il peut être également utile de vérifier que le site n'est plus impacté et de *reconstituer* l'émission à l'aide des mesures effectuées afin de déterminer les zones qui ont été réellement impactées.

## 1.4. CONTEXTE MILITAIRE ET SECURITAIRE

La force du vent peut être utilisée comme support d'armes redoutables. Ainsi, des produits chimiques ou biologiques dispersés dans l'atmosphère peuvent porter atteinte aux êtres humains, aux installations de défense ou encore aux sources de ravitaillement.

Même si ce type d'agression reste rare, le danger est réel et toujours présent sur les champs de bataille et dans les pays ciblés par des agressions terroristes (*European Commission, 2014*).

La Figure 1.6 schématise une attaque chimique contre une ville.

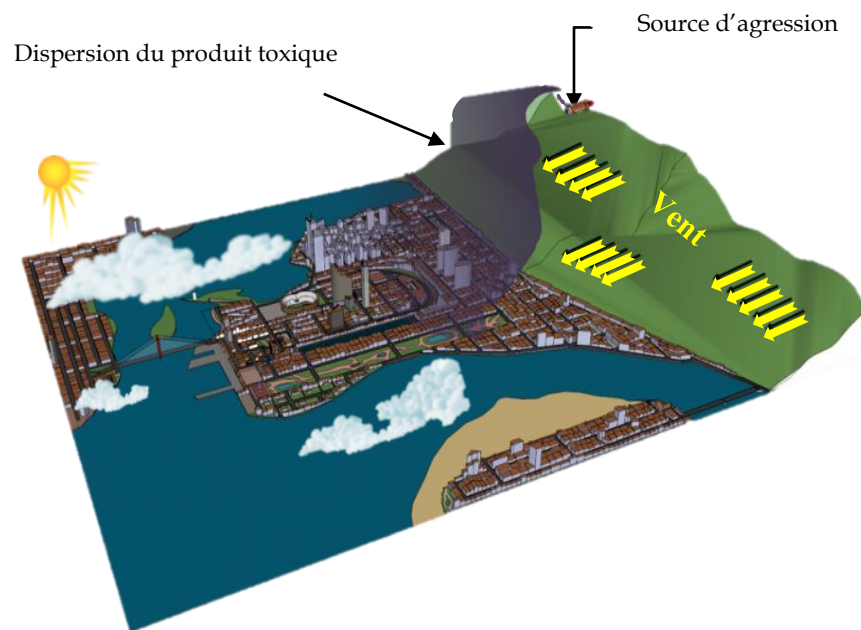


Figure 1.6. Schématisation d'une agression par le biais de la dispersion atmosphérique

### 1.4.1. Générateurs de risques

Dans le contexte militaire, la contamination de l'air par des agents NRBC n'est pas accidentelle, comme dans le contexte industriel, mais prend plutôt la forme d'agressions planifiées. Généralement, les risques NRBC sont liés à une pratique "non conventionnelle" des combats ou à des actions de type terroriste. Nous donnons ci-dessous un aperçu rapide des armes de types nucléaires et radiologiques. Ensuite, nous décrivons plus en détail les caractéristiques des armes chimiques ou biologiques car ces dernières se basent sur la dispersion atmosphérique et dépendent essentiellement des conditions météorologiques (*SIPRI Yearbook, 2004*).

- Risques nucléaires et radiologiques

Les armes nucléaires (N) sont composées de dispositifs explosifs nucléaires et de leurs vecteurs. Ces dispositifs explosifs utilisent des réactions nucléaires auto-entretenues qui transforment la structure des atomes et dégagent une énergie considérable. Les armes radiologiques (R) prennent, elles, généralement la forme de bombes entourées de matériaux radioactifs destinés à être répandus en poussière lors de l'explosion (Dirty Bombs). Cette explosion n'a pas l'intensité thermique et mécanique d'une bombe conventionnelle, mais répand autour d'elle des éléments radioactifs qui auront des effets à long terme.

- Risques biologiques

Les armes biologiques (B) sont considérées comme des armes de destruction massive, tout comme les armes nucléaires. Ces armes sont plus faciles à fabriquer que les armes nucléaires coûtent beaucoup moins cher et peuvent proliférer assez facilement. Opérationnellement, les agents biologiques peuvent être dispersés par des munitions de dispersion, des munitions explosives ou des munitions de pulvérisation.

D'après le bureau des affaires du désarmement des Nations Unis : « les munitions de dispersion utilisent des générateurs d'aérosol, montés sur des véhicules, qui relâchent des poudres d'une granulométrie précise. Il est généralement difficile de préparer les agents sous cette forme, mais, lorsqu'ils sont prêts, leur diffusion par dispersion est relativement simple et malheureusement efficace. Les munitions explosives utilisant la force induite par la détonation pour disperser l'agent sur la cible. Elles ne sont pas très efficaces car l'effet de l'explosion peut désactiver instantanément la plupart, voire la totalité, de l'agent. Quant aux munitions de pulvérisation, elles dispersent à partir d'une source fixe l'agent biologique sous la forme d'un nuage aérosol invisible composé de particules microscopiques ».

Toujours selon les nations unis, les agents biologiques susceptibles d'être utilisés dans des armes sont classés en cinq catégories (*Riche & Binder, 1983*).

1. Les bactéries, ce sont des micro-organismes généralement cultivés dans des milieux artificiels. Différentes bactéries sont susceptibles d'être utilisées dans des armes biologiques : *Bacillus anthracis*, *Vibrio cholerae*, *Pasteurella tularensis*, *Salmonella typhi*, etc.

2. Les virus, ce sont des micro-organismes nettement plus petits que les bactéries et ils ne peuvent se développer que dans des cellules vivantes. Leurs effets peuvent être amplifiés par une mutation naturelle ou par une modification génétique. Différents agents sont susceptibles d'être utilisés dans des armes biologiques : le virus d'Ebola, le virus de Hantaan, le virus de la fièvre de la vallée du Rift et le virus de la fièvre jaune, etc.
3. Les rickettsies, ce sont des micro-organismes dont la structure est similaire à celle des bactéries. Elles ont besoin, comme les virus, de cellules vivantes pour se développer. Différentes rickettsies sont susceptibles d'être utilisées dans des armes biologiques : *Rickettsia prowasecki*, *Rickettsia rickettsii*, *Coxiella burnetii*, etc.
4. Les champignons, ce sont des micro-organismes qui se nourrissent de matière organique. Ils ne sont généralement pas nocifs pour les êtres humains ni les animaux, mais peuvent être nuisibles pour les plantes. Différents agents sont susceptibles d'être utilisés pour des armes biologiques : *Helminthosporium oryzae*, *Microcyclus ulei*, *Puccinia graminis*, *Colletotrichum kahawae*, etc.
5. Les toxines ([Mollaret, 2002](#)), ce sont des substances toxiques; elles sont des produits ou des dérivés d'animaux, de plantes ou de micro-organismes. À la différence des autres agents biologiques, les toxines ne sont pas des organismes vivants et ne peuvent donc se reproduire. Différentes toxines sont susceptibles d'être utilisées dans des armes biologiques : toxines botuliques, aflatoxines, entérotoxines (*Staphylococcus aureus*), saxitoxines, etc.

Aujourd'hui, les équipements de détection et d'identification d'agents biologiques, pour les forces armées et les acteurs de la sécurité civile, sont nombreux. Dans le rapport ([OTAN, 2005](#)) nous pouvons retrouver un aperçu détaillé des technologies utilisées. A titre d'exemple, on peut citer :

- a. Les détecteurs à distance : Plusieurs technologies, comme celles du radar Doppler, du lidar (radar optique) ou la spectroscopie par claquage induit par éclair laser (LIBS). Ces détecteurs utilisent les ondes radio ou des techniques de réflectance à la lumière pour détecter la présence d'agents pathogènes dans l'atmosphère.
- b. Les détecteurs en champs proche. Ce sont des détecteurs soit spécifiques (qui donnent une spécification de l'agent) soit non spécifiques (détection de présence sans identification). Une des techniques utilisée par ces capteurs est la spectrométrie de masse.

- Risques chimiques

Les armes chimiques utilisent les propriétés toxiques des substances chimiques pour blesser ou pour tuer. Elles sont, elles aussi, considérées comme des armes de destruction massive. Elles coûtent beaucoup moins cher, sont plus faciles à mettre au point, à fabriquer et à entretenir que les armes nucléaires. Elles sont donc particulièrement redoutables lorsqu'elles sont utilisées par des terroristes contre des populations civiles.

Sur le plan militaire, au cours de l'histoire moderne, elles ont surtout été utilisées (et produites de façon industrielle) à partir de la première guerre mondiale. Elles ont fait l'objet d'une interdiction par une convention, de son nom officiel « Convention sur l'interdiction de la mise au point, de la fabrication, du stockage et de l'usage des armes chimiques et sur leur destruction (CIAC) » entrée en vigueur en 1997 et de nombreux pays ont détruit une grande partie de leurs stocks.

En général, les agents toxiques sont dispersés sous la forme de gaz, d'aérosols ou de liquides. Pour les attaques à distance, des gaz et des aérosols comprenant des particules de diamètre bien déterminé sont utilisés. Les substances libérées sont emportées par le vent et attaquent le système respiratoire. Pour les attaques qui consistent à relâcher la substance directement sur la cible, la forme liquide est utilisée.

Dans les différents modes d'attaque, la granulométrie est une caractéristique déterminante : les particules trop petites sont emportées par le vent et manquent leur cible tandis que celles trop grosses ne sont pas transportées ou atteignent leur cible mais n'agissent pas de la façon désirée (non-pénétration dans la peau par exemple, etc.).

D'après le bureau des affaires du désarmement des Nations Unis, les produits toxiques utilisés dans la fabrication d'armes chimiques peuvent être classés selon leurs effets :

1. Les suffocants. Ils attaquent les yeux et les voies respiratoires. Ce type d'agents est particulièrement nocif pour les poumons qui se remplissent progressivement de liquide et gonflent de façon à ce que le sang ne soit plus être alimenté en oxygène, provoquant l'étouffement progressive puis la mort.
2. Les agents hémotoxiques. Ils bloquent l'échange d'oxygène entre les globules rouges et le tissu cellulaire. Ils sont fatals et agissent très rapidement.
3. Les agents vésicants. Ils provoquent, de graves brûlures sur la peau, sur les yeux et sur les poumons. L'exposition aux agents vésicants déclenche des douleurs et des lésions immédiates, et peut entraîner la mort par asphyxie.

4. Les agents neurotoxiques. Ils agissent sur la transmission des impulsions nerveuses dans le système nerveux et classés comme les produits chimiques toxiques les plus puissants. Ils sont généralement incolores et inodores, et peuvent être facilement absorbés par le système respiratoire, les yeux, la peau et le tube digestif. Ils sont extrêmement toxiques et sont fatals même si l'exposition n'a été que très brève.
5. Les agents incapacitants. Ils n'ont que des effets corporels temporaires qui disparaissent généralement assez rapidement et n'entraînent pas de lésions ni la mort, mais les personnes visées sont incapables de conduire leurs activités habituelles.
6. Les agents neutralisants. Ils n'entraînent que des effets physiologiques temporaires comme des troubles de la vue ou de la respiration et ne provoquent généralement pas de lésions graves.
7. Les toxines. Elles sont des poisons produits par des organismes vivants ou leurs équivalents de synthèse. Elles sont extrêmement toxiques et peuvent être fatales.

De point de vue technologique, aujourd'hui, un certain nombre de techniques sont disponibles pour détecter les agents chimiques toxiques (*INERIS, 2009*).

A titre d'exemple, parmi les appareils de mesure les plus utilisés, on peut citer :

- a. Les détecteurs optiques, dans lesquels un rayon lumineux est utilisé comme sonde afin d'examiner les caractéristiques optiques du gaz (diffraction, raies d'absorption, etc.). Par exemple, la caméra de télédétection (en temps réel) de gaz par infrarouge « Second Sight » (*Bertin instrument*) est une caméra de détection de gaz à distance, utile pour la surveillance de population ou la protection des infrastructures critiques.
- b. Les détecteurs basés sur la spectrométrie et la chromatographie. Les capteurs utilisés peuvent être à photoionisation (PID), à spectroscopie de mobilité ionique (IMS) ou à ionisation de flamme (FID). Parmi les appareils qui utilisent ces technologies, on peut citer le moniteur de vapeurs chimiques « CAM », le détecteur de gaz toxiques « AP2C, AP4C, » (*Mullota, 2010*), etc.
- c. Les détecteurs électrochimiques. Les capteurs utilisés sont soit potentiométriques, conductimétries ou ampérométriques.

### 1.4.2. Effets possibles

- Nucléaire

L'effet principal des armes nucléaires est la destruction à grande échelle par onde de choc et effet de souffle. Les radiations émises lors d'une explosion provoquent quant à elles des effets dévastateurs, à plus ou moins long terme, sur les personnes irradiées.

- Radiologique

L'effet principal des armes de type (R) n'est pas de détruire, mais de contaminer une zone géographique et les personnes exposées par des radiations directes et/ou l'ingestion et/ou l'inhalation de matériaux radioactifs.

- Biologique

Les effets des armes biologiques dépendent de nombreux facteurs et notamment du type et de la qualité des agents utilisés, des conditions environnementales, de la diffusion et de la sensibilité de la cible. Chaque type d'agent provoque donc des effets différents.

Dans un contexte militaire, les armes biologiques peuvent être utilisées pour frapper de grandes cibles relativement statiques, comme les bases de missiles, les postes de commandement et de contrôle, les installations logistiques, etc. Contre les populations civiles, les armes biologiques peuvent provoquer une épidémie de très grande échelle, contaminer les sources d'eau et d'alimentation. Elles peuvent également servir à des actes terroristes pour tuer ou affaiblir des êtres humains ou des animaux.

- Chimique

Les effets des armes chimiques dépendent de plusieurs facteurs : de l'efficacité de la dispersion, des conditions météorologiques et du niveau de défense de la cible visée, etc. Si, par exemple, des substances toxiques sont dispersées de manière efficace sur une cible, leur impact peut entraîner la contamination des matériels et des terrains. Elle peut entraîner la mise hors de combat des personnels (situation de guerre) ou la mort de civils (actions terroristes).

### 1.4.3. Enjeux et besoins

Face à la contamination, volontaire, de l'air par des agents NRBC (ou l'être humain est, directement ou indirectement, la cible) les enjeux sont importants.



- Les enjeux préventifs

Les autorités de défense sont appelées à protéger leurs forces et leurs équipements en situation de combat mais également à assurer la sécurité des sites militaires ou civils et des installations vitales comme les sources d'approvisionnement en eau, etc. L'aspect préventif, dans ce cadre, consiste à mettre en œuvre des stratégies de *surveillance* et de *mise en alerte* en cas d'attaque. Le but est de déclencher les opérations de riposte, d'évacuation, de secours et de sauvetage, le plus rapidement possible. Comme le notent (*Hills, 2001*) et (*Primmerman, 2000*) ces stratégies doivent s'appuyer sur des données issues de capteurs mis en réseaux.

- Les enjeux opérationnels

Dans le cas d'une agression, surtout par des agents de type chimique ou biologique, aucune information à priori sur la source de l'attaque n'est disponible. Aussi, d'un point de vue opérationnel, il est capital de pouvoir : *identifier* (ou de confirmer l'identification) les agents utilisées dans l'action hostile, *localiser* (en espace et en temps) et/ou *caractériser* la source (*quantifier* le débit, le type d'émission continue/transitoire). Pour les deux derniers points, notre étude trouve une justification importante.

Pour le cas des armes nucléaires, les efforts pour réduire les risques relèvent du traité de non-prolifération. Le besoin essentiel dans ce cadre est de pouvoir *détecter* les essais non déclarés (*Hourdin et Issartel, 2000*). Mais l'échelle à laquelle se fait ressentir ce besoin ne concerne pas notre étude.

## 1.5. CONCLUSION DU CHAPITRE

À travers ce chapitre, nous avons réalisé un inventaire non exhaustif des enjeux et des besoins liés à la contamination de l'air par des agents NRBC, dans les contextes environnementaux, industriels, militaires et sécuritaires. Nous avons identifié des besoins spécifiques et communs, en termes de réseaux de surveillance, et nous avons mis en évidence ceux qui relèvent de notre étude.

Les réseaux de surveillance de la contamination de l'air peuvent être utiles pour répondre à des enjeux stratégiques. La majorité d'entre eux est actuellement utilisé pour :

- *Surveiller* ou *observer*, au sens large, un site, une région, en mesurant à intervalles réguliers des concentrations d'espèces données. L'objectif est d'observer l'évolution des



concentrations afin de pouvoir dresser des constats, par exemple, en cas de dépassement de valeur limite (pollution routière, rejets d'une installation industrielle, etc.).

- *Détecter* des fuites, accidents, anomalies, attaques ou, de façon plus générale, tous rejets inhabituels afin d'*alerter* au plus tôt, en cas de dépassement des seuils vis à vis de normes fixées. Activer les mesures d'urgence (mesure d'urgence, évacuation, mise à l'abri et protection, riposte, etc.), tout en limitant les fausses alarmes. C'est la stratégie dite du « detect to warn ».

Une grande partie des réseaux existants permet aussi de :

- *Prédire* ou *reconstruire* spatialement, à un instant donné, un champ de concentration à partir des mesures collectées par le réseau (déterminer les zones qui ont été effectivement impactées par un rejet, estimer le niveau de pollution dans une région, etc.). La reconstruction peut être réalisée directement à partir des mesures (à l'aide de techniques d'estimation linéaire, comme le Krigeage, etc.). Pour *prédire*, à court terme, l'évolution spatiotemporelle du champ de concentration, il est nécessaire d'utiliser des méthodes plus élaborées, comme l'assimilation des données. Cette dernière utilise, en plus des mesures, des modèles numériques et des données statistiques sur les champs étudiés. Dans ce cas, on cherche à ce que l'erreur commise par l'estimation sur les zones où il n'y a pas de capteurs soit la plus faible possible.

Un besoin plus spécifique, notamment cité par ([Primmerman, 2000](#)) consiste à :

- *Identifier* (ou *confirmer*) les résultats d'une première identification d'agents toxiques afin de devancer les effets de la contamination. C'est la stratégie dite du « Detect to Treat ». L'objectif est de déclencher, le plus rapidement possible, un processus permettant de traiter les personnes contaminées avant même que les symptômes d'une infection ne se manifestent.

Mais les deux enjeux qui seront particulièrement étudiés dans cette étude sont :

- *Localiser* et/ou *caractériser* la source de contamination. Le premier objectif est de situer la source afin de mettre en œuvre les mesures appropriées pour résoudre l'incident (intervention adaptée sur zone, fermeture d'un axe routier ou d'une usine, neutralisation d'une attaque, etc.). Le second objectif est de *quantifier* les flux rejetés (en masse par unité de temps) mais aussi de déterminer si le rejet est transitoire ou continu.

## CHAPITRE 1 :

### Les réseaux de capteurs et les enjeux environnementaux, industriels, militaires et sécuritaires

---

*« L'objectif général du présent travail est donc le développement d'un processus permettant de positionner des capteurs sur un site donné de façon à former un réseau de surveillance optimal en termes de localisation et de caractérisation de la source d'émission. Pour cela il sera nécessaire d'inverser les données fournies par le réseau. Il faudra donc utiliser (i) un modèle physique capable de représenter la dispersion des agents observés et (ii) une méthode adaptée pour l'inversion des données.*

*Le challenge consiste à ce que les mesures réalisées par le réseau optimal fournissent, une fois inversées, les données les plus pertinentes possible »*

« BIBLIOGRAPHIE CHAPITRE 1 »

Abida, Rachid, (2010). Construction optimale de réseaux fixes et mobiles pour la surveillance opérationnelle des rejets accidentels atmosphériques. Modélisation et simulation. Ecole des Ponts ParisTech. Français. <https://pastel.archives-ouvertes.fr/pastel-00638050>.

AÏCHI, L, (2015). Rapport n° 610 (2014-2015), fait au nom de la CE coût économique et financier de la pollution de l'air, déposé le 8 juillet 2015. <http://www.senat.fr/rap/r14-610-2/r14-610-21.pdf>

AIRPARIF-a. <https://www.airparif.asso.fr/pollution/effets-de-la-pollution-generalites>

AIRPARIF-b. <http://www.airparif.asso.fr/methodes-surveillance/reseau-mesure>

ARIA. <https://www.aria.developpement-durable.gouv.fr/en-cas-daccident/echelle-europeenne-des-accidents-industriels/>

AWAC : Agence Wallonne de l'Air et du Climat. Les émissions atmosphériques acidifiantes de l'industrie. <http://environnement.wallonie.be/enviroentreprises/pages/etatenviindustrie.asp?doc=syn-ind-aci>

Bertin instrument.

[https://bertin.fr/sites/default/files/media/bertin/media/brochures/pdf/Brochure\\_SecondSight\\_FR\\_WEB.pdf](https://bertin.fr/sites/default/files/media/bertin/media/brochures/pdf/Brochure_SecondSight_FR_WEB.pdf)

Bobin, J-L., Nifenecker, H., Stéphan C., (2005). L'énergie de demain bilan et perspectives. Pages 226-228.

CAFE, (2005). Estimation du Clean Air for Europe.

CITEPA, (2000). <https://www.citepa.org/fr/activites/inventaires-des-emissions/secten>

CITEPA, (2017). <https://www.citepa.org/fr/air-et-climat/analyse-sectorielle/tertiaire>

Directive européenne n°2008/50/CE, (2008). <https://www.airparif.asso.fr/reglementation/normes-europeennes>

Directives Seveso (1, 2 et 3). <http://www.installationsclassees.developpement-durable.gouv.fr/Risques-accidentels.html>

DRIRE P. ACA, (2004). Principales entreprises ayant des rejets dans l'air.

[http://www.paca.developpement-durable.gouv.fr/IMG/pdf/EI04chair\\_cle039229.pdf](http://www.paca.developpement-durable.gouv.fr/IMG/pdf/EI04chair_cle039229.pdf)

European Commission, (2014). OMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS. On a new EU approach to the detection and mitigation of CBRN-E risks. Brussels, 5.5.2014 COM(2014) 247 final.

Gouvernement.fr. <http://www.gouvernement.fr/risques/accident-industriel>

GAUTIER.Y. « BHOPAL ACCIDENT CHIMIQUE DE (2-3 décembre 1984) », Encyclopædia Universalis. <http://www.universalis.fr/encyclopedie/accident-chimique-de-bhopal/>

Grahn, H., von Schoenberg, P., Brännströma, N., (2015). What's that smell? Hydrogen sulphide transport from Bardarbunga to Scandinavia. Journal of Volcanology and Geothermal Research.

Hills, R, (2001). Sensing for danger: correlated sensor networks," in Lawrence Livermore National Laboratory, vol. 1, pp. 11-17.

Hourdin, F. et Issartel, J. -P., (2000). Sub-surface nuclear tests monitoring through the CTBT Xenon Network. AN AGU Journal Volume 27, Issue 151, Pages 2245-2248. DOI: 10.1029/1999GL010909.

IBGE-LRE, (2006-2008). Rapport de la Qualité de l'Air.

IIASA : International Institute for Applied Systems Analysis.

<https://www.airparif.asso.fr/pollution/effets-de-la-pollution-generalites>

INERIS, (2009). - Principes et techniques pour la détection des gaz. RAPPORT D'ÉTUDE 26/01/2009. N° INERIS-DRA-08-86244-13727B

INERIS, (2015). Sécurité industrielle\_Evaluer la gravité des conséquences environnementales d'un accident\_Eléments de méthodologie.

<http://www.ineris.fr/centredoc/dp-conseq-enviro-def-1433839136.pdf>

INS : inventaire national spatialisé, (2004). <http://www2.prevoir.org/content/origine-et-sources-de-pollution>

IRSN, (2016). Fukushima en 2016. Impact environnemental en 2016 de l'accident de Fukushima Daiichi.[http://www.irsn.fr/FR/connaissances/Installations\\_nucleaires/Les-accidents-nucleaires/accident-fukushima-2011/fukushima-2016/Pages/2-impact-environnement-accident-nucleaire-fukushima-2016.aspx#.WZyQwChJbIU](http://www.irsn.fr/FR/connaissances/Installations_nucleaires/Les-accidents-nucleaires/accident-fukushima-2011/fukushima-2016/Pages/2-impact-environnement-accident-nucleaire-fukushima-2016.aspx#.WZyQwChJbIU)

IRSN. Coût économique des accidents nucléaires. Le coût économique pour deux scénarios d'accident. [http://www.irsn.fr/FR/connaissances/Installations\\_nucleaires/Les-accidents-nucleaires/cout-economique-accident/Pages/2-cout-economique-pour-2-scenarios.aspx#.WZyQBihJbIU](http://www.irsn.fr/FR/connaissances/Installations_nucleaires/Les-accidents-nucleaires/cout-economique-accident/Pages/2-cout-economique-pour-2-scenarios.aspx#.WZyQBihJbIU)

Ministère de la Transition écologique et solidaire. <https://www.ecologique-solidaire.gouv.fr/pollution-lair-origines-situation-et-impacts>

Ministère de l'Environnement, (2006). Brochure sur les économies d'énergies et le changement climatique.

[http://www.environnement.public.lu/air\\_bruit/publications/Campagne\\_changement\\_climatique\\_ff/CCC\\_F.pdf](http://www.environnement.public.lu/air_bruit/publications/Campagne_changement_climatique_ff/CCC_F.pdf)

Mollaret, H. H. (2002). L'arme biologique - bactéries, virus et terrorisme, Plon.

Mullota, J.-U., Bousqueta, A., Burnatb, P., (2010). L'AP4C : caractéristiques d'un nouvel appareil de détection des toxiques chimiques de guerre et de certains toxiques industriels. Médecine et armées 38(5):459-464.

OTAN, (2005). Détection chimique, biologique, radiologique ou nucléaire (CBRN) : aperçu des technologies. Lord Jopling (Royome-uni) rapporteur special.

Pillot, D., Hervé Guegan, L., Paturel, F., Cazier, J.C., Déchaux, Evelyne Combet, J.-M., André, R., Vidon, P., Tassel, A., Mercier, et al. (2006). Emissions unitaires de métaux et de polluants non réglementés des voitures particulières équipées ou non de filtre à particules. Rapport de recherche. <https://hal.archives-ouvertes.fr/hal-00545906>.

Pineau, Jean-Philippe, (1995). Analyse de risques d'installations industrielles. Journée "Protection de l'Environnement, Amiens, France. <https://hal-ineris.ccsd.cnrs.fr/ineris-00971953>.

PREV' AIR. L'air en France aujourd'hui et demain. <http://www2.prevoir.org/content/origine-et-sources-de-pollution>

## CHAPITRE 1 :

### Les réseaux de capteurs et les enjeux environnementaux, industriels, militaires et sécuritaires

---

Primmerman, C, (2000). Detection of Biological Agents. Lincoln Laboratory Journal, Volume 12, Number 1.

Radisson. L, (2015). ICPE : les dernières statistiques disponibles.

<https://www.actu-environnement.com/ae/news/ICPE-statistiques-chiffres-cles-inspection-23816.php4>

RECORD, (2013). Méthodes de mesurage des polluants rejetés à l'atmosphère ou dans les eaux. Surveillance des rejets de polluants organiques et inorganiques dans les milieux fluides. Métrologies existantes et en développement, 2013, 548 p, n°11-0142/1A.

Riche, D., & Binder, P., (1983). Les armes chimiques et biologiques. L'Archipel 2011.

Saddier. M, (2011). La lutte contre la pollution atmosphérique, enjeu majeur au cœur des politiques environnementales depuis 20 ans.

[http://lodel.irevues.inist.fr/pollution-atmospherique/docannexe/file/473/edito\\_211.pdf](http://lodel.irevues.inist.fr/pollution-atmospherique/docannexe/file/473/edito_211.pdf)

SIPRI Yearbook, (2004). Armaments, disarmament and international security, Oxford Press University, 2004, p.561

**CHAPITRE 2 :**  
**OPTIMISATION DE RESEAUX**  
**DE CAPTEURS**

## CHAPITRE 2

### OPTIMISATION DE RESEAUX DE CAPTEURS

#### Plan du chapitre 2

CHAPITRE 2 OPTIMISATION DE RESEAUX DE CAPTEURS.....	28
2.1. Introduction.....	30
2.2. Expression du besoin .....	30
2.3. Les questions clés .....	31
2.4. Définition des zones étudiées .....	33
2.5. Définition du Probleme d'optimisation .....	36
2.5.1. Enoncé du problème .....	36
2.5.2. Les critères d'optimalité .....	37
2.5.2.1. Les critères basés sur la notion d'information.....	37
2.5.2.2. Les critères basés sur la notion de couverture.....	39
2.5.2.3. Autres critères .....	42
2.5.3. Choix de l'algorithme d'optimisation.....	44
2.5.3.1. L'algorithme du Recuit Simulé.....	46
2.5.3.1.1. Principe .....	46
2.5.3.1.2. Algorithme .....	47
2.5.3.1.3. Paramétrage .....	48
2.5.3.1.4. Avantages et inconvénients .....	50
2.5.3.2. L'Algorithme Génétique.....	51
2.5.3.2.1. Principe .....	51
2.5.3.2.2. Algorithme .....	52
2.5.3.2.3. Paramétrage .....	54
2.5.3.2.4. Avantages et inconvénients .....	55
2.6. Conclusion du chapitre.....	56
« Bibliographie chapitre 2 ».....	57

## 2.1. INTRODUCTION

Ce chapitre, qui s'appuie sur une importante étude bibliographique, est consacré dans un premier temps à la définition générale du problème d'optimisation. L'objectif est de définir une méthodologie à suivre afin d'assurer une conception optimale de réseaux de surveillance. Nous présentons notamment les critères les plus couramment utilisés pour définir une conception optimale (« optimal design » en anglais). Cette partie nous permet de formaliser la problématique à laquelle cette thèse apporte réponse. Dans un second temps, nous présenterons les moyens algorithmiques qui permettront d'obtenir les solutions optimales.

## 2.2. EXPRESSION DU BESOIN

L'optimisation des réseaux de surveillance peut répondre à plusieurs besoins, comme le montre la figure 2.1.

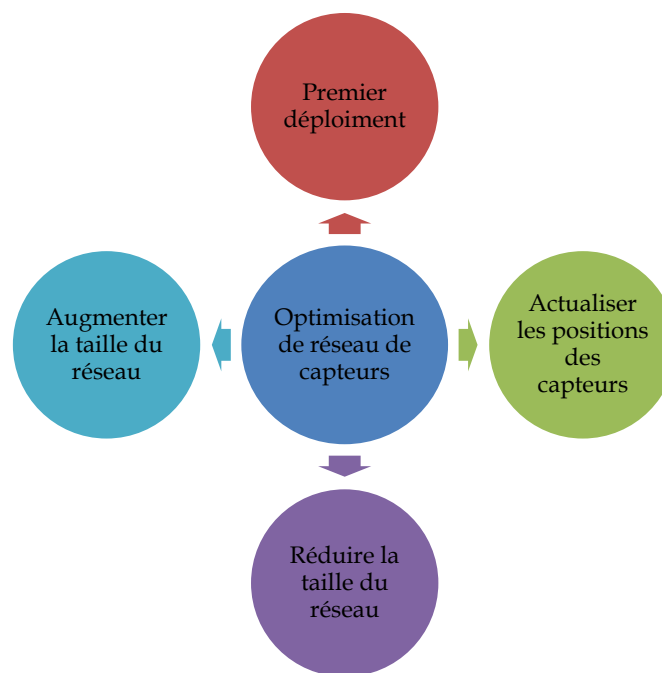


Figure 2.1. Cas de figure d'optimisation de réseaux de capteurs

Nous pouvons être amenés à créer un nouveau réseau de toute pièce (à partir d'un cahier des charges utilisateur) ou à modifier la configuration spatiale d'un réseau déjà existant en déplaçant les capteurs (Abida et Bocquet, 2009). Nous pouvons également être amenés à réduire un réseau déjà existant (Nychka et Saltzman, 1998 ; Hourdin et Issartel, 2000 ; Abida, 2010) ou à élargir sa taille. Ce dernier cas correspond à celui d'un réseau qui aurait été configuré à un instant donné et dont les utilisateurs souhaiteraient élargir la taille afin de



répondre à de nouveaux besoins de détection (par exemple : augmentation de la taille du domaine à surveiller). Dans notre étude, nous nous intéresserons principalement au cas de premier déploiement (création d'un réseau à partir d'un cahier des charges).

La conception optimale consiste alors à déterminer la configuration spatiale la plus performante (i) d'un nombre prédéterminé de capteurs ou (ii) d'un nombre indéterminé de capteurs, ce nombre étant également une inconnue du problème d'optimisation.

Dans notre étude, nous envisageons d'examiner les deux cas. L'optimisation d'un nombre connu de capteurs, plus simple, sera privilégiée mais nous tenterons d'apporter des éléments de réponse au problème de la détermination du nombre optimal de détecteurs à utiliser.

### 2.3. LES QUESTIONS CLES

Le processus de « conception optimale » d'un réseau de surveillance comporte en premier lieu une phase d'avant-projet au cours de laquelle des questions fondamentales doivent trouver réponse. Tout d'abord, il faut définir précisément l'objectif assigné au réseau. Sera-t-il utilisé uniquement pour la détection et la mise en alerte, pour l'identification, pour la caractérisation et la localisation ? La réponse à cette première question est importante puisqu'elle conditionnera entièrement le choix d'un critère quantitatif qui permettra d'évaluer la performance du réseau lors de l'étape d'optimisation.

Il faut également définir la nature des agents surveillés (espèces chimiques, biologiques, particules, etc.), les seuils d'alerte et la technologie qui permettra d'obtenir les concentrations dans l'air de ces agents (gamme des capteurs ? techniques de mesures ? plages de mesure ?).

Il faut aussi étudier la réalité du terrain et définir précisément la zone à surveiller, ou à protéger puis déterminer la faisabilité du déploiement du réseau dans, ou autour, de cette zone. L'objectif est ensuite de déterminer une zone de déploiement à l'intérieur de laquelle il faut choisir les positions potentielles des capteurs. Ces positions peuvent dépendre d'exigences techniques de mise en œuvre (les capteurs doivent-ils être montés sur des poteaux ? à quelle hauteur ?) mais aussi d'exigences légales (peut-on installer un détecteur sur le toit d'un bâtiment public ?).

Lorsque l'on souhaite concevoir de façon optimale un réseau de surveillance, plusieurs réponses doivent donc être apportées à des questions clés. Ces différentes réponses fournissent les éléments nécessaires à la définition du problème d'optimisation. Ceci est résumé sur la figure 2.2.

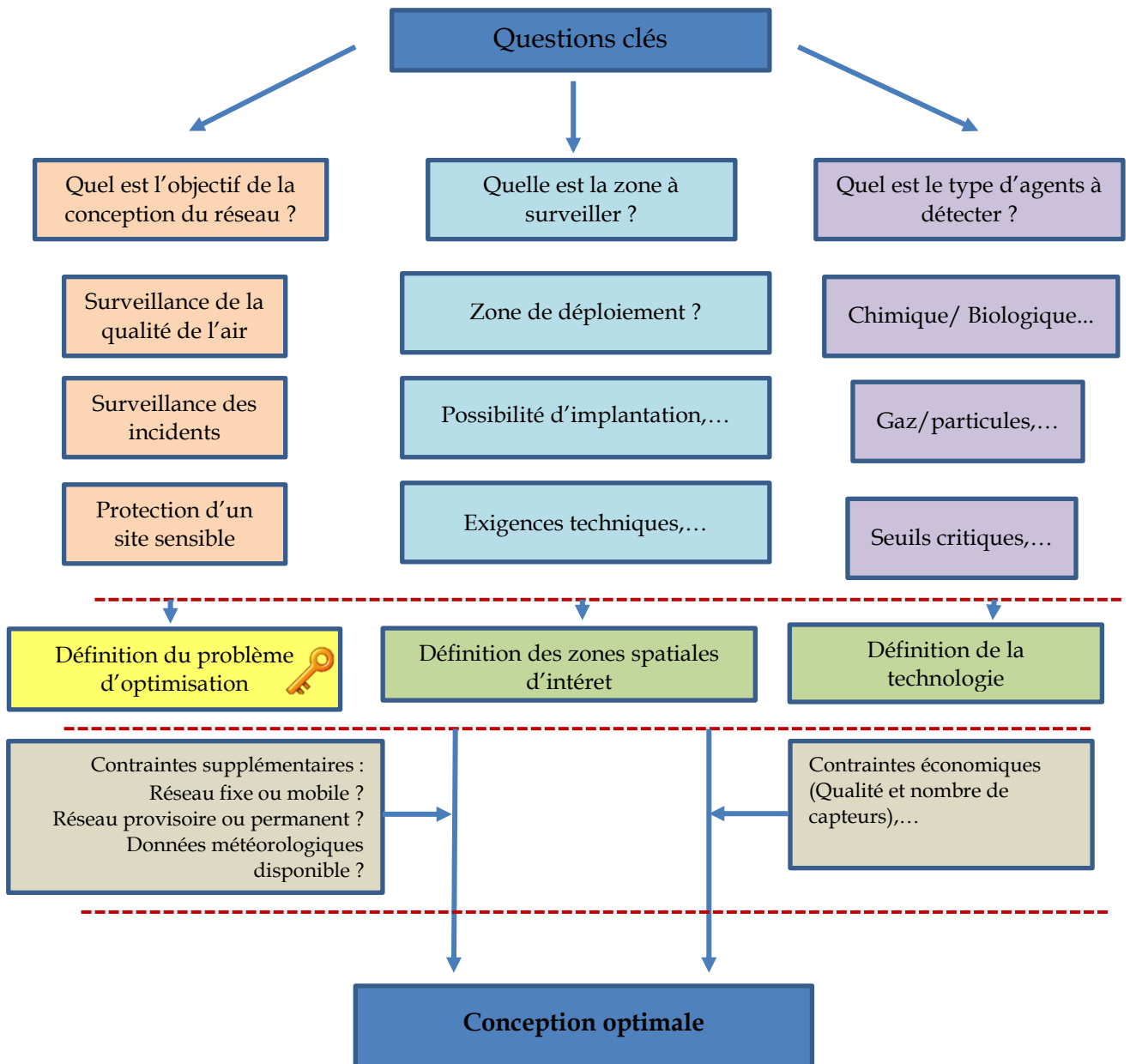


Figure 2.2. Questions clés auxquelles la phase d'avant-projet doit répondre<sup>1</sup>.

Il est utile de noter que dans notre travail nous étudierons deux types de réseaux : des réseaux fixes et permanents qui devront garder leur efficacité dans des conditions météorologiques variées et des réseaux mobiles déployés dans des situations d'urgence durant lesquelles les paramètres météorologiques sont connus en temps réel. Nous reviendrons ultérieurement sur la problématique liée à la météorologie.

Notre but ici n'est pas d'examiner tous les scénarios possibles mais de proposer une méthodologie suffisamment générale qui pourra apporter des éléments de réponse dans de nombreux cas. Des informations techniques sur les capteurs ayant été données dans le

<sup>1</sup> Proposition d'un ensemble de questions pour assurer la conception optimale d'un réseau de surveillance des rejets atmosphériques. La clé désigne la phase à laquelle se travail de thèse apporte réponse.

premier chapitre, nous n'aborderons pas ici la partie « Définition de la technologie ». Nous nous concentrerons sur les parties « Définition des zones spatiales d'intérêt » et « Définition des critères d'optimalité ».

## 2.4. DEFINITION DES ZONES ETUDIEES

La détermination des zones d'étude est une étape clé dans le processus d'optimisation. En effet, leur définition oriente la façon par laquelle le réseau de capteurs sera optimisé. Nous distinguons :

- La zone à protéger notée  $D_p$ . Elle correspond au domaine qui fait l'objet de la protection par le réseau.
- La zone d'intérêt notée  $D$ , qui regroupe la zone à protéger et son environnement proche, noté  $D_e = D - D_p$ .
- La zone de danger  $D_d$  ( $D_d \subset D$ ) qui représente la zone à partir desquelles une émission peut atteindre  $D_p$ . Cette zone peut varier en fonction des conditions météorologiques.
- La zone d'implantation  $D_i$  ( $D_i \subset D$ ) qui représente l'espace dans lequel des capteurs peuvent être placés.

Sur les figures 2.3 et 2.4 nous illustrons ces définitions par deux exemples concrets. Nous montrons, ainsi, le lien étroit entre le scénario pour lequel le réseau est conçu, l'objectif dédié à la conception et les zones d'étude.

Sur la figure 2.3, les différentes zones d'études sont schématisées pour un scénario d'agression de type terroriste contre une installation militaire (*Base militaire de Rosnay, Indre, France*). Nous identifions, ainsi, ces zones pour concevoir un réseau chargé de caractériser la source (*localiser* et/ou *estimer* l'intensité de la source) et pour concevoir un réseau chargé d'*alerter* en cas d'attaque. Pour le premier cas, la zone d'implantation  $D_i$  et la zone d'intérêt  $D$  sont équivalentes. En effet,  $D_i$  englobe la zone à protéger  $D_p$  (la base militaire) et son environnement proche  $D_e$ . Enfin la zone de danger  $D_d$  est située à l'extérieure de  $D_p$  et fait partie de  $D_e$ . Pour le deuxième cas, la zone d'implantation  $D_i$  appartient à  $D_e$  tout comme  $D_d$ .

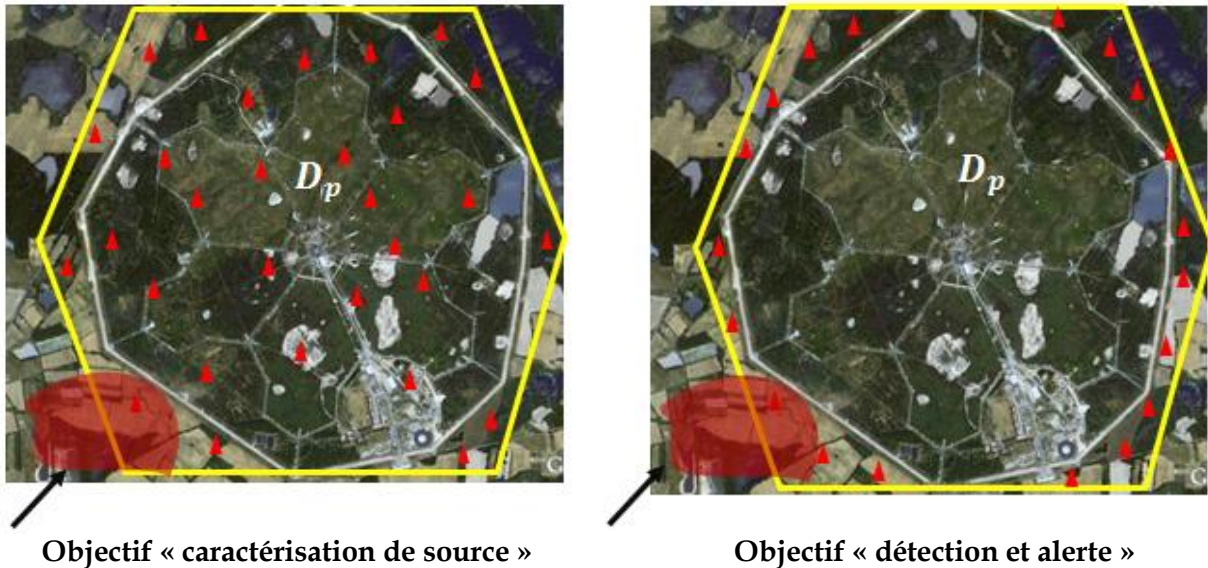


Figure 2.3. Zones d'études pour un scénario d'agression terroriste. La schématisation regroupe : les capteurs (triangle), la zone d'implantation (contour jaune) la zone à protéger (contour blanc), la zone de danger (espace rouge) et la direction du vent (flèche). (*Google Earth*)

Sur la figure 2.4, les différentes zones d'études sont schématisées pour un scénario d'accident à l'intérieur de site industriel (*Usine Sidech, Belgique*). De la même façon, nous identifions, ces zones pour concevoir un réseau chargé d'estimer l'intensité de la source et pour concevoir un réseau chargé d'alerter en cas d'un accident. Dans les deux cas, la zone à protéger  $D_p$ , la zone de danger  $D_d$  et la zone d'implantation  $D_i$  sont équivalentes, en effet,  $D_e$  est inexistante ( $D_e = \emptyset$ ). Notons que ce scénario est différent de celui de la protection d'une zone d'habitation se trouvant à proximité d'un site industriel ( $D_d$ ). Dans ce cas, la zone à protéger se trouve alors à l'extérieur du site industriel ( $D_p \subset D_e$ ) et des capteurs peuvent être installés également à proximité des habitations ( $D_i \subset D_e$ ).

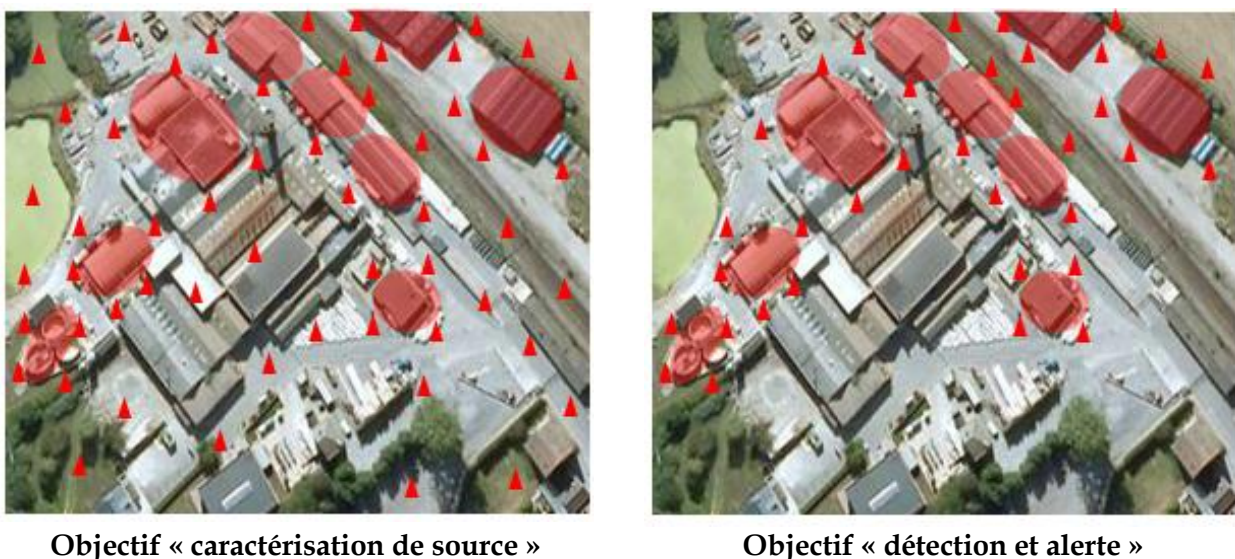


Figure 2.4. Zones d'études pour un scénario d'accident à l'intérieur d'un site industriel. La schématisation regroupe : les capteurs (triangles) et les sources potentielles (espace rouge).

Quel que soit le scénario envisagé, dans la zone d'implantation, nous considérons  $p$  emplacements potentiels parmi lesquels nous cherchons à sélectionner les  $m$  « meilleurs emplacements » pour répondre aux objectifs fixés au réseau.

L'ensemble formé par les  $p$  emplacements potentiels est appelé  $G_I$ . L'ensemble des  $m$  positions instrumentées est noté  $R$ , l'ensemble des  $p-m$  positions non instrumentées est noté  $N_I$  (voir figure 2.5). On a donc un réseau global  $G_I$  formé par un ensemble de  $p$  positions potentielles et dans lequel on distingue un sous-réseau instrumenté  $R$  ( $m$  capteurs) et un sous-réseau non instrumenté  $N_I$  ( $p-m$  positions). Dans notre étude, ces différentes zones seront généralement considérées comme des espaces bidimensionnels discrets.

On peut donc définir  $G_I$  comme étant un ensemble de  $p$  positions  $\mathbf{u}_k = (x_k, y_k)$  avec  $k = 1 \dots p$ .

$$G_I = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\} \subset D_i \quad (2.1)$$

Il est important de noter que ces  $p$  positions peuvent être définies intuitivement (par l'expérience, en fonction de contraintes techniques, etc.) ou peuvent être déterminées par une étude préliminaire spécifique (Kouichi et al 2016), en cohérence avec le scénario d'étude (surveillance d'accident dans une zone industrielle, surveillance de zones sensibles, etc.).

L'ensemble  $R$  peut lui être défini comme étant l'ensemble des  $m$  positions  $\mathbf{u}_i = (x_i, y_i)$  avec  $i = 1 \dots m$ .

$$R = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\} \subset G_I \quad (2.2)$$

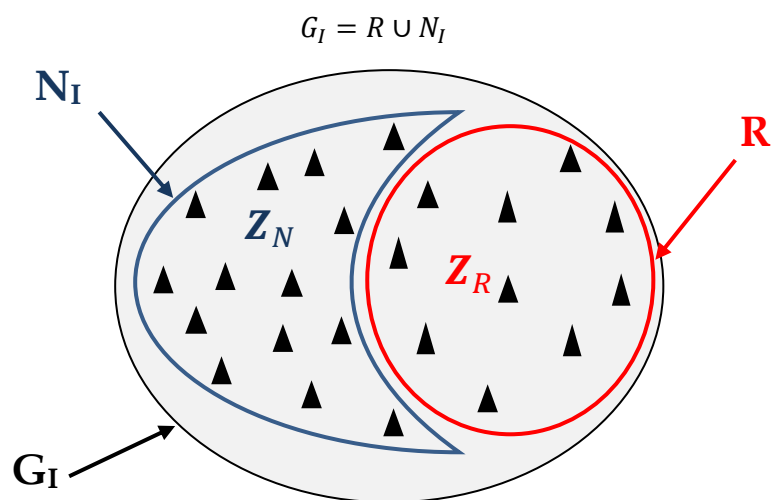


Figure 2.5. Représentation symbolique d'un réseau, d'après (Bocquet, 2006),  $Z$  désigne le champ mesuré



## 2.5. DEFINITION DU PROBLEME D'OPTIMISATION

### 2.5.1. Enoncé du problème

Le travail de conception du réseau consiste à sélectionner  $m$  emplacements parmi les  $p$  prédéterminés, c'est-à-dire de choisir le « meilleur » sous-ensemble  $R$  parmi tous les sous-ensembles réalisables. Ce « meilleur » sous-ensemble est nommé « réseau optimal » et est noté

$$R_o = \{\mathbf{u}_1^o, \mathbf{u}_2^o, \dots, \mathbf{u}_m^o\} \subset G_I \quad (2.3)$$

La recherche du sous-ensemble  $R_o$  est basée sur la minimisation (ou la maximisation) d'une fonction objectif  $J$  qui traduit, de façon quantitative, la qualité du réseau. Ce type de problème est dit « *Problème d'Optimisation Combinatoire* » : il consiste à trouver, parmi  $C_m^p$  configurations possibles, la configuration  $R_o \subset G_I$  telle que

$$J(R_o) = \min_{R \subset G_I} J(R) \quad (2.4)$$

Fondamentalement, le terme fonction objectif est utilisé en optimisation mathématique pour désigner une fonction qui sert de critère pour déterminer la meilleure solution à un problème d'optimisation. Le but est alors de minimiser ou de maximiser cette fonction pour trouver un optimum. Dans le cadre de la conception optimale des réseaux de surveillance de rejets atmosphériques, cette fonction doit permettre d'évaluer la qualité de son architecture. Sa définition repose donc sur une analyse fine de l'objectif fixé au réseau. Ainsi, tout en restant très général, nous pouvons identifier trois phases dans le processus d'optimisation, comme le schématise la figure 2.6.

La définition du but pour lequel le réseau doit être optimisé est la première étape de tout processus d'optimisation. La conception est alors définie selon les objectifs associés au réseau. Compte tenu de la variété et de la pluralité des objectifs (*Reay et SwiftHook, 1979*), il est indispensable d'identifier chaque problème distinctement.

Dans le premier chapitre nous avons donné un large aperçu sur l'utilité des réseaux de mesure et nous avons résumé leurs objectifs opérationnels au niveau de la conclusion. Les principaux objectifs identifiés ont été :

- La *surveillance* et la *détection*,
- La *prédiction* ou la *reconstruction*,
- La *localisation* ou la *caractérisation*.

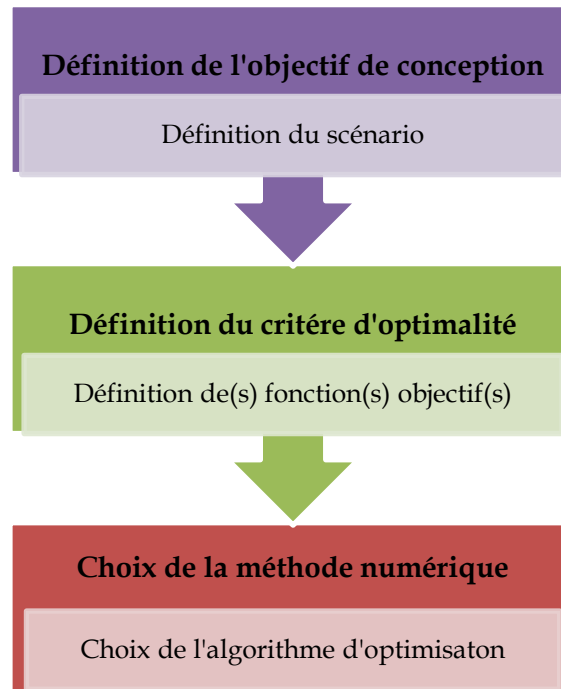


Figure 2.6. Processus générale d'optimisation de réseaux de surveillance

Comme l'explique (Abida, 2010), d'un point de vue théorique, pour traduire ces objectifs informels en terme mathématiques, des notions issues de différentes disciplines peuvent être employées. Ces domaines sont, par exemple, la géostatistique (notion de remplissage spatial), la thermodynamique (notion d'entropie), et la télécommunication (notion d'information, notion de couverture, etc.), les plans d'expériences (notion d'Optimal Experimental Design), etc.

Ci-dessous, nous présentons une liste non exhaustive des critères utilisés pour répondre aux objectifs rappelés ci-dessus.

## 2.5.2. Les critères d'optimalité

### 2.5.2.1. Les critères basés sur la notion d'information

La première approche permettant de définir des fonctions objectifs, s'inspire de la théorie de l'information développée en 1948 par Claude Shannon. Ce dernier est le premier à avoir défini l'information comme étant une grandeur mesurable (Shannon, 1948). Dans ce cas, le réseau est considéré optimal lorsqu'il possède un indice d'information maximal. Cet indice mesure le gain en information apporté par les  $m$  sites instrumentés aux  $p-m$  sites non instrumentés. Ainsi, les sites instrumentés sont choisis de telle sorte que les  $m$  mesures

fournissent le maximum d'information sur les valeurs du champ observé au niveau des  $p-m$  sites non instrumentés. Le champ aléatoire observé est appelé  $\mathbf{Z} \in \mathbb{R}^p$ . Il peut être partitionné en deux vecteurs :  $\mathbf{Z}_R \in \mathbb{R}^m$  mesuré sur les sites instrumentés et  $\mathbf{Z}_N \in \mathbb{R}^{p-m}$  défini sur les sites non instrumentés. L'incertitude sur le champ  $\mathbf{Z}$  est caractérisée par l'entropie de sa fonction de densité de probabilité  $S(\mathbf{Z})$  (Caselton et Husain, 1980 ; Le et Zidek, 1994 ; Zidek et al, 2000) et par l'information totale qu'il contient  $I(\mathbf{Z}) = -S(\mathbf{Z})$ . L'idée consiste à choisir les  $m$  sites instrumentés de sorte que la prédiction du champ total basé sur ces  $m$  mesures fournisse le maximum possible d'information sur les  $p-m$  sites non instrumentés. La mesure du gain en information sur  $\mathbf{Z}_N$  en mesurant  $\mathbf{Z}_R$  est appelée l'indice d'information de Shannon. Il est donné par (Bernardo, 1975)

$$I(\mathbf{Z}_N, \mathbf{Z}_R) = I(\mathbf{Z}_N | \mathbf{Z}_R) - I(\mathbf{Z}_N) \quad (2.5)$$

Où  $I(\mathbf{Z}_N | \mathbf{Z}_R)$  est l'information conditionnelle de  $\mathbf{Z}_N$  connaissant  $\mathbf{Z}_R$  (connaissance à posteriori obtenue sur les sites  $N_I$  à partir des mesures réalisées sur les sites  $R$ ) et  $I(\mathbf{Z}_N)$  est l'information à priori sur les mesures réalisées par  $N_I$ . L'idée de base des méthodes basées sur la théorie de l'information consiste à supposer que les processus  $\mathbf{Z}, \mathbf{Z}_N, \mathbf{Z}_R$  sont Gaussiens. Ainsi  $I(\mathbf{Z}_N, \mathbf{Z}_R)$  peut être exprimé en fonction des matrices de covariance de ces processus. Le réseau optimal est obtenu en maximisant l'indice d'information.

La seconde approche repose sur la notion d'entropie. L'origine des recherches sur cette théorie remonte à la fin du XIXe siècle et aux travaux de Boltzmann (en thermique et mathématiques). La notion d'entropie est définie à partir de la densité de probabilité de  $\mathbf{Z}$ , notée  $g(\mathbf{Z})$ , telle que  $S(\mathbf{Z}) = E\{-\log(g(\mathbf{Z}))\}$  où  $E\{.\}$  est l'espérance mathématique. Comme le remarque (Abida, 2010), les méthodes basées sur l'entropie cherchent elles aussi à maximiser l'information fournie par le réseau de capteurs mais en minimisant l'entropie conditionnelle  $S(\mathbf{Z}_N | \mathbf{Z}_R)$ . Cette quantité mesure l'incertitude résiduelle restante après avoir observé  $\mathbf{Z}_R$ . (Abida, 2010) montre que l'entropie totale du système est :

$$S(\mathbf{Z}_N, \mathbf{Z}_R) = S(\mathbf{Z}_N | \mathbf{Z}_R) + S(\mathbf{Z}_R) \quad (2.6)$$

Or l'entropie totale étant une quantité fixe, minimiser l'entropie conditionnelle est équivalent à maximiser l'entropie sur les sites instrumentés  $S(\mathbf{Z}_R)$ . Si  $\mathbf{Z}_R$  est Gaussien, avec une matrice de covariance  $\mathbf{C}_{\mathbf{Z}_R}$ , le réseau optimal est celui qui maximise le déterminant de cette matrice

$$R_o = \arg \max \left( \frac{1}{2} \det(\mathbf{C}_{\mathbf{Z}_R}(R)) \right) \quad (2.7)$$



Ce type de critère, basé sur la théorie de l'information, peut être utilisé pour concevoir des réseaux dont l'objectif est de caractériser une source d'émissions à partir de mesures de concentrations. Mais il faut déterminer à priori une loi de répartition de l'erreur de mesure (cadre bayésien), ce qui peut biaiser le processus d'optimisation.

### 2.5.2.2. Les critères basés sur la notion de couverture

Dans le domaine de la conception de réseau de capteurs, un objectif qui est souvent visé est celui de la « bonne couverture spatiale ». En fonction du domaine étudié, cette notion n'a pas tout à fait la même signification : D'une part, dans les travaux très généraux de [\(Dhillon et Chakrabarty, 2003 ; Altinel et al, 2008\)](#) ou [\(Nazrul Alam et Haas, 2006\)](#), chaque capteur possède une « zone de couverture » (souvent une sphère en 3D, un disque en 2D, etc.), ou une « zone de détection ». Il est considéré que le réseau assure une couverture complète si chaque point de la zone surveillée se trouve dans au moins une de ces « zones de couverture ». Si la zone est considérée comme étant un domaine continu (à deux dimensions) il est parfois possible de résoudre le problème géométriquement [\(Rourke, 1987\)](#). De façon plus générale, si la zone est discrétisée, des modèles mathématiques spécifiques sont utilisés. Cette notion de couverture a été étendue au K-coverage (chaque position du domaine surveillé couverte par au moins K capteurs) par [\(Wu et al, 2008 ; Mini et al, 2011\)](#) ou [\(Zhou et al, 2004\)](#). [\(Ghosh et Das, 2008\)](#) mentionnent que la couverture peut être représentée soit de façon binaire (la capacité de détection est supposée uniforme, un point couvert reçoit la valeur 1, un point en dehors de la zone de couverture reçoit la valeur 0) soit à l'aide de modèles plus réalistes prenant en compte la diminution de la précision de détection avec la distance [\(Wang et Wang, 2011\)](#).

De façon très simplifiée, nous proposons ici un formalisme mathématique qui exploite cette notion de couverture. On considère qu'un point du domaine discrétisé ( $\mathbf{x} \in G_j$ ) est couvert par le réseau s'il se trouve dans la zone de couverture d'au moins un des capteurs situé sur une position instrumentée ( $\mathbf{u} \in R$ ). Considérons que le domaine comporte  $j=1..N$  points et que le réseau comporte  $i$  capteurs, positionnée en  $m$  positions  $u_i$ . La sensibilité du capteur  $i$  au point  $x_j$  peut-être caractérisée par une fonction de sensibilité, qui décroît lorsque la distance  $d(\mathbf{x}, \mathbf{u})$  augmente. Nous noterons  $a_i(\mathbf{x})$  la sensibilité du capteur  $i$  au point  $\mathbf{x}$  (nous reviendrons ultérieurement sur cette notion très importante (partie 3.5.3). On peut ainsi construire une matrice de sensibilité  $\mathbf{A} \in \mathbb{R}^{m \times N}$ . Les  $i=1..m$  lignes de cette matrice représentent les sensibilités des  $i$  vecteurs aux  $N$  points  $\mathbf{x}_j$ .

On peut généralement considérer qu'un point  $x_j$  est couvert par le capteur  $i$  si  $a_i(x_j)$  est supérieure à un seuil de détection que nous noterons  $a_{min}$  (on suppose que ce seuil est le même pour tous les capteurs). On peut ainsi définir une matrice de couverture notée  $\mathbf{M}_c \in \mathbb{R}^{m \times N}$  telle que

$$[\mathbf{M}_c]_{ij} = \begin{cases} 1 & \text{si } [A]_{ij} \geq a_{min} \\ 0 & \text{sinon} \end{cases} \quad (2.8)$$

Le point  $x_j$  est donc couvert par le réseau si

$$m_j = \sum_{i=1}^m [\mathbf{M}_c]_{ij} \geq 1 \quad (2.9)$$

Notons que  $m_j$  est le nombre de capteurs qui sont « sensibles » au point  $x_j$ . De même, le nombre de points du domaine couvert par le réseau est

$$n_c = \sum_{j=1}^n \max_i [\mathbf{M}_c]_{ij} \quad (2.10)$$

Le nombre total de détections potentielles réalisables par le réseau est

$$m_d = \sum_{j=1}^n m_j \quad (2.11)$$

Les paramètres définis ci-dessus peuvent être utilisés pour définir des critères d'optimalité ou des contraintes d'optimisation, basées sur la notion de couverture. On peut par exemple maximiser le nombre de points surveillés par le réseau de capteurs. La configuration optimale est obtenue en maximisant le critère  $n_c$ .

$$R_o = \arg \max_{R \subset G_I} n_c(R) \quad (2.12)$$

Notons que des contraintes additionnelles complètent souvent ce type de problème d'optimisation. Elles peuvent concerner le nombre de capteurs sensibles en chaque point du domaine (*K-optimality*), la connectivité (afin de garantir la transmissibilité des informations recueillies vers une base) voir ([Shakkottai et al. 2005](#) ; [Liu et al. 2006](#) ; [Bai et al. 2006](#) ; [Ammari et Das, 2012](#)), la consommation d'énergie et la durée de vie du réseau ([Zhao et Gurusamy, 2008](#) ; [Idrees et al, 2015](#)).

Ce type de critère, basé sur la notion de couverture, peut être utilisé pour concevoir des réseaux dont l'objectif est de caractériser une source d'émissions à partir de mesures de concentrations atmosphériques. Mais il faut définir de façon précise la notion de sensibilité,

ce qui sera fait à la partie 3.4.2. D'autre part, le vocable « couverture » est aussi souvent associé à la notion de remplissage spatial. Dans ce cas, un réseau assure une bonne couverture lorsque ses capteurs sont repartis efficacement sur l'ensemble de la région d'intérêt.

Les méthodes qui garantissent une bonne répartition spatiale des capteurs (*Nychka and Saltzman, 1998*) n'utilisent ni modèle physique ni notion de zone de détection mais font intervenir les distances entre les points du domaine discrétisé ( $\mathbf{x} \in G_I$ ) et les positions instrumentées ( $\mathbf{u} \in R$ ). Par exemple, pour assurer la répartition uniforme des capteurs dans un domaine, la moyenne des distances entre les sites candidats et le réseau est donnée par

$$D(R) = \left( \frac{1}{p-m} \sum_{\mathbf{x} \in G_I} d(\mathbf{x}, R)^\alpha \right)^{1/\alpha} \text{ pour } \alpha > 0 \quad (2.13)$$

Lorsque  $\alpha = 1$  le critère de remplissage spatial  $D(R)$  est dit uniforme et la distance d'un point  $\mathbf{x}$  non instrumenté au réseau est défini par  $d(\mathbf{x}, R) = \min_{\mathbf{u} \in R} \|\mathbf{x} - \mathbf{u}\|$ . Le réseau optimal  $R_o$ , pour le remplissage spatial uniforme (*Abida, 2010*), est celui qui minimise la quantité  $D(R)$ :

$$R_o = \arg \min_{R \subset G_I} D(R) \quad (2.14)$$

(*Abida, 2010*) cite d'autres critères de remplissage spatial permettant de concevoir des réseaux optimaux répondant à divers objectifs (comme, par exemple, l'étalement des points du réseau les uns par rapport aux autres).

D'autres études utilisent la notion de dimension fractale pour définir un critère géométrique de remplissage spatial. Cette dimension, notée  $D_m$ , introduite par (*Lovejoy et al. 1989*) et (*Raes et al. 1991*) caractérise l'hétérogénéité spatiale du réseau. En deux dimensions, l'écart  $2 - D_m$  mesure « le degré d'irrégularité spatiale » du réseau. (*Di Zio et al, 2004*) définissent un réseau optimal comme étant :

$$R_o = \arg \min_{R \subset G_I} (2 - D_m(R)) \quad (2.15)$$

De façon générale, ces critères basés sur la notion de remplissage spatial, peuvent être très intéressants lorsque l'objectif fixé au réseau est uniquement la détection et/ou la surveillance. (*Abida et al, 2008*) les ont utilisés pour la construction optimale de réseaux fixes et mobiles chargés de la surveillance opérationnelle des rejets atmosphériques accidentels. Le projet consistait en l'optimisation d'un réseau automatisé de télésurveillance des aérosols radioactifs dans l'air, DESCARTES, utilisé par la direction de l'environnement et de l'intervention de l'IRSN.

### 2.5.2.3. *Autres critères*

- **Critères basés sur la notion de probabilité de détection**

Dans le cadre du traité de non-prolifération nucléaire, et afin de détecter la dispersion atmosphérique de radionucléides (Radium) à l'échelle mondiale, un réseau de 80 stations est utilisé (International Monitoring System), voir ([Mason et al 1995](#)). Ce réseau a été conçu afin de détecter une détonation nucléaire de 1 kilotonne de TNT dans un intervalle de temps de 14 jours, avec une probabilité de 90%. Le critère de conception du réseau a donc été sa « capacité de détection » déterminée à l'aide de Modèles de Transports Atmosphériques (ATM), pour des scénarios correspondant à des conditions météorologiques courantes. Parmi ces 80 stations, 40 ont été choisies pour détecter également les isotopes du Xénon. Les stations non retenues ont été celles dont l'absence dégradait le moins les probabilités de détection du réseau. C'est dans ce cadre qu'([Hourdin et Issartel, 2000](#)) ont adopté une « approche basée récepteurs » afin de définir la notion de visibilité d'un capteur. Cette visibilité correspond à la sensibilité du capteur à des sources potentielles situées en tous points du domaine. Dans le cadre d'une approche Eulérienne, cette sensibilité est calculée à partir de panaches dispersés à rebours dans le temps à partir de la position du détecteur. En utilisant les seuils cités précédemment, et pour plusieurs conditions météorologiques, des cartes de probabilités de détection pour des sources potentielles en tous points de la planète ont été tracées. De même le nombre moyen de stations indépendantes détectant chaque source potentielle (critère assez proche du terme  $m_j$  de l'équation 2.9) est indiqué. Nous avons choisi d'utiliser cette approche pour définir la notion de « couverture d'un capteur ». Cette notion sera étendue à la notion de « visibilité d'un réseau » au chapitre 3.

- **Critères basés sur la notion de réduction de l'impact**

Il est utile de noter que la problématique d'optimisation de réseaux de capteurs n'est pas spécifique à la surveillance de la contamination de l'air par des agents NRBC. Nous trouvons, dans la littérature, des travaux intéressants dans d'autres domaines. Par exemple, ([Berry et al, 2006-a](#)) ont étudié la conception d'un réseau de capteurs ayant pour but la réduction de l'impact d'attaques sur les réseaux de distribution d'eau. Le problème de conception est nommé SPP (Sensors Placement Problem) et la mise en alerte CWS (Contaminant Warning System). L'idée consiste à se baser sur des scénarios d'attaques, dont chacune a une probabilité d'occurrence bien déterminée. L'impact de chaque attaque potentielle est déterminé par un logiciel de simulation (EPANET).

Cette idée a été notamment reprise par (*Watson J-P et al, 2004*) et (*Krause et al, 2008*), toujours dans le cadre de la protection de réseau de distribution d'eau, et par (*Hamel et al, 2010*) dans le cadre de la prévention d'attaques RBC en champ proche. Dans cette dernière étude, sur un site donné, des scénarios d'attaques sont définis et leur probabilité d'occurrence déterminée à partir des conditions météorologiques. Pour chaque scénario, la source est un point du domaine et un modèle basé sur la mécanique des fluides numérique (CFD en anglais) est utilisé pour calculer la quantité de produit toxique délivrée avant que le réseau ne détecte l'attaque. Le réseau optimal est celui pour lequel la quantité de produit toxique « consommée » (moyennée sur le nombre total d'attaque) est la plus faible possible. Pour notre travail, nous retiendrons l'idée d'utiliser les conditions météorologiques pour définir une probabilité d'occurrence. Par contre, la notion de « consommation » de produit toxiques, basée sur des modèles qui décrivent le mouvement de la cible pendant l'attaque (foules ou de population), ne nous semble pas adaptée à notre étude.

- **Critères basés sur la notion d'Optimal Experimental Design**

La théorie de « l'Optimal Experimental Design » a principalement été établie par (*Kiefer et al, 1960*) afin de déterminer des plans d'expériences optimaux. Comme l'explique (*Abida, 2010*), le lien entre cette théorie et la conception de réseaux a été établie par (*Fedorov, 1972 ; Silvey, 1980 ; Fedorov et Müller, 1988, 1989*). Lorsque les mesures sont utilisées pour la prédiction d'un champ physique aléatoire (par régression linéaire), la confiance que l'on porte aux résultats de la prédiction est caractérisée par une matrice d'information dite de Fisher. Les propriétés de cette matrice permettent de caractériser l'incertitude associée aux résultats. C'est cette erreur qui peut être minimisée, suivant plusieurs critères (*D-optimality, A-optimality, E-optimality, etc.*), lors de la conception d'un réseau.

- **Critères basés sur les erreurs d'estimation**

(*Abida et al, 2008*) et (*Saunier et al, 2009*) ont travaillé sur la conception d'un réseau de mesure des concentrations de radionucléide à l'échelle de la France. L'objectif était d'extrapoler ces mesures, à l'aide de méthodes d'interpolation spatiale (Krigage simple, méthode du point le plus proche). La capacité du réseau à fournir des données adaptées a été évaluée quantitativement à travers une fonction coût, qui mesurait les écarts entre les champs de concentrations extrapolés (au niveau du sol) et des concentrations de référence simulées, pour  $q$  scénarios, à l'aide du modèle Polair3D. La configuration optimale était celle qui minimisait une fonction coût (Hölder norm) du type :

$$J_\alpha = \left( \frac{1}{N} \sum_{i=1}^q \sum_{k=1}^N |\bar{c}_k^i - c_k^i|^\alpha \right)^{1/\alpha} \quad (2.16)$$

Où  $\bar{c}_k$  est le champ extrapolé au point  $k$  et  $c_k$  le champ simulé au même point. Le paramètre  $\alpha$  peut être ajusté afin de minimiser l'importance des fortes concentrations. Pour  $\alpha = 2$ , on retrouve la norme  $L_2$ . Une approche assez similaire avait été utilisée par (Jiang et al, 2007) pour la détection d'accidents nucléaires à l'échelle régionale. Leur méthode consistait à simuler, à l'aide d'un modèle, de nombreux scénarios de dispersion afin d'obtenir des champs de concentrations de référence. Ensuite, ces champs étaient comparés à des estimations obtenues par interpolation des données mesurées par le réseau. Le choix optimal des emplacements des détecteurs (mobiles) était celui qui minimisait les fausses alertes (faux positifs : concentrations, après estimation surestimées, faux négatif : concentrations, après estimation, sous-estimées).

Nous notons que ce type de critère est efficace lorsque l'on cherche à minimiser l'écart entre des données de référence et des données estimées.

- **Critères basés sur la minimisation de l'incertitude**

Une partie des travaux de (Abida et Bocquet, 2009) montrent comment utiliser les mesures collectées par le réseau et un modèle numérique de dispersion, à travers un schéma d'assimilation de données, pour prévoir l'évolution du panache dispersé. Le réseau surveillance est conçu en utilisant toute l'information disponible, afin de minimiser l'incertitude dans la prévision.

Dans ce travail, une technique de reconstruction séquentielle du panache dispersé, est assurée en couplant la modélisation inverse à une stratégie de ciblage des observations. Cette technique permet de prendre en compte correctement l'incertitude sur les champs météorologiques, qui est identifiée comme la deuxième source d'incertitude après l'incertitude sur le terme source.

### 2.5.3. Choix de l'algorithme d'optimisation

Quelle que soit la fonction objectif  $J$  retenue, résoudre un problème d'optimisation consiste à déterminer la « meilleure solution » possible à ce problème, cette dernière correspond à l'« optimum global ». Ainsi, le choix de la méthode d'optimisation dépend de la façon dont le problème est approché (espace de recherche continu ou discret).

Dans la littérature, nous trouvons de nombreuses classifications des techniques d'optimisation, par exemple celle proposée par (Talbi, 2009) ou celle proposée par (Dréo et al, 2003). Nous présentons, cette dernière sur la figure 2.7.

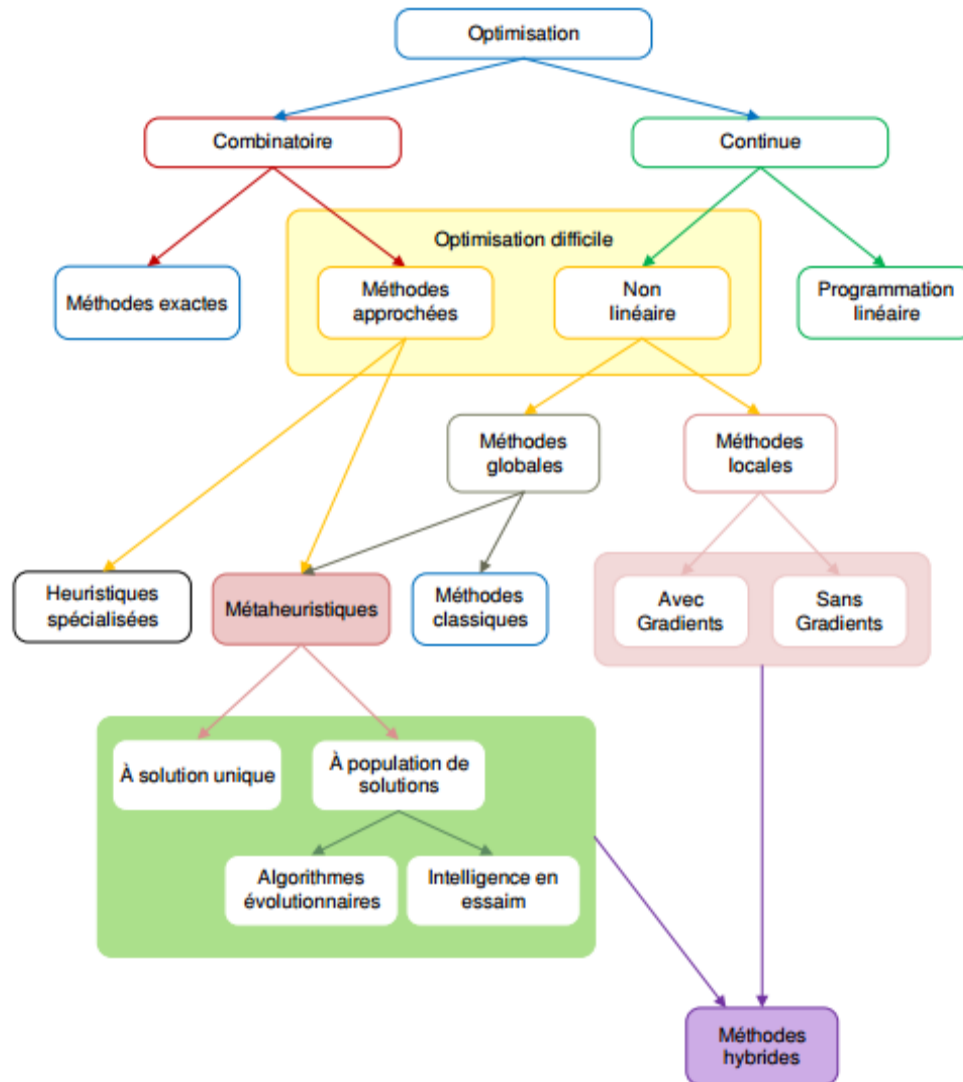


Figure 2.7. Classification des méthodes d'optimisation (Dréo et al, 2003)

Nous rappelons que le problème d'optimisation, que nous avons défini à la partie 2.5, consiste à choisir  $m$  emplacements parmi  $p$  positions potentielles prédéterminées, c'est-à-dire à trouver la solution optimale, parmi  $C_m^p$  combinaisons possibles, telle que

$$R_o = \arg \min(J(R)) \quad (2.17)$$

Ainsi, l'algorithme d'optimisation est l'outil qui va permettre d'obtenir les solutions numériques à ce problème. À priori, trouver une configuration optimale parmi un ensemble fini de configurations possibles est un problème simple, il suffit de tester toutes les configurations et de les comparer les unes aux autres pour en extraire la meilleure. Mais, en



pratique, le nombre de configurations possible est généralement trop important et une telle étude est impossible. Cette complexité a été soulignée par (Guttorp et al, 1993) et (Le et al, 2003). (Ko et al, 1995) ont montré que le problème d'optimisation d'un réseau de mesure est « NP-difficile » (NP : Non-déterministe Polynomial). Ceci signifie que les algorithmes qui évaluent toutes les solutions possibles (méthodes dites exactes ou déterministes) sont inutilisables dans notre problème car elles ne permettent pas d'obtenir la solution en un temps raisonnable. Par contre, dans la littérature, on trouve plusieurs algorithmes basés sur la recherche stochastique (recherche par exploration aléatoire) qui fournissent des solutions approchés, réalisables, de bonne qualité (et optimales dans les meilleurs cas) au problème défini au paragraphe 2.5.1. Parmi ces algorithmes, nous citons : la recombinaison de chemin « PR » relinking) (Glover, 1977), les algorithmes de colonies de fourmis (Coloni, 1991 ; Dorigo, 1992), les algorithmes par essaims particulaires « PSO » (Kennedy et Eberhart, 1995), la procédure de recherche gloutonne aléatoire adaptative « GRASP » (Feo et Resende, 1989), les algorithmes à estimation de distribution « EDA » (Mühlenbein et Paafß, 1996), les systèmes immunitaires artificiels « AIS » (Farmer et al, 1986), la recherche avec tabous « TS » (Glover, 1986), l'algorithme de « Kiefer-Wolfowitz » (Kiefer et Wolfowitz, 1952), le recuit simulé « SA », les algorithmes génétiques « GA », etc.

Parmi toutes ces approches, les plus utilisées sont l'algorithme probabiliste dit du « recuit simulé » (Simulated Annealing, SA, en anglais) (Nourani et Andresen, 1998 ; Nunes, 2004 ; Abida et al, 2008 ; Abida et bocquet, 2009 ; Saunier et al, 2009) car efficaces tout en étant relativement simple à mettre en œuvre, et l'algorithme évolutionnaire ou évolutionniste dit « génétique » (Reed et al, 2000 ; Biellia et al, 2002 ; Wu et al, 2005 ; Farshbaf et Feizi-Derakhshi, 2009 ; Türk et al 2012 ; Potuzak, 2014). Ces deux algorithmes sont de la famille « métaheuristique » et se basent sur deux techniques différentes de recherche d'optimum. Le recuit simulé se base sur l'évolution itérative d'une solution unique et l'algorithme génétique manipule en parallèle toute une population de solutions.

Nous présentons ici ces deux algorithmes, qui seront utilisés dans le quatrième chapitre.

### 2.5.3.1. L'algorithme du Recuit Simulé

#### 2.5.3.1.1. Principe

Le recuit simulé est un algorithme proposé suite aux travaux de (Metropolis et al, 1953) et (Kirkpatrick et al. 1983). C'est une technique de recherche probabiliste inspirée d'un phénomène naturel. En métallurgie, le procédé du recuit consiste à chauffer, graduellement,



le métal et à le refroidir d'une façon contrôlée, dans le but de modifier ses caractéristiques. En phase de refroidissement les particules qui forment le métal s'arrangent entre elles de façon à ce que les groupes qui possèdent le plus faible niveau d'énergie soient privilégiés. Tant que la température du métal est assez importante, les atomes trouvent l'énergie pour changer de groupe. En revanche, quand la température baisse, les atomes manquent de ressources pour passer vers des groupes plus instables. A un niveau de température  $T$ , le passage d'un groupe d'atomes à un niveau d'énergie supérieur (avec un écart  $dE$ ) est accepté avec une probabilité donnée par

$$\exp\left(-\frac{dE}{kT}\right) \quad (2.18)$$

où  $k$  est la constante de Boltzmann.

Lorsque la température diminue, cette probabilité privilégie les états de moindre énergie. Physiquement, cette probabilité signifie que la chance pour un groupe d'atomes de passer à une configuration de niveau d'énergie plus élevée sera d'autant plus grande que la différence d'énergie sera faible et que la température sera élevée.

#### 2.5.3.1.2. *Algorithme*

La méthode du recuit simulé se base sur la règle dite de Métropolis (*Metropolis et al. 1953*). Le principe, dans le cadre de l'optimisation de réseaux de capteurs, est le suivant : nous définissons une fonction coût  $J$  qui décrit la performance des réseaux et nous partons d'une configuration de capteurs de départ (qui peut être aléatoire). Nous déterminons une température de départ  $T_0$ , d'arrêt  $T_f$ , un schéma de refroidissement de la température (plus d'informations seront données ci-dessous) et un nombre d'itérations par palier (de température) si nécessaire.

Dans ce qui suit nous décrivons l'algorithme du recuit simulé pour un schéma de refroidissement par palier. Ainsi, à chaque palier, on réalise un nombre fixe de modification de la configuration du réseau. Si une modification augmente (ou diminue)  $J$  (cela dépend de la fonction choisie), on calcule sa probabilité d'acceptation qui dépend de la température courante selon la fonction  $\exp\left(\frac{-\Delta J}{T}\right)$ . On tire ensuite aléatoirement un nombre entre 0 et 1. Si ce nombre est inférieur à la probabilité, on accepte cette modification.

Après avoir réalisé tous les itérations sur un palier, la température est baissée et l'on recommence, tant que la température d'arrêt n'est pas atteinte. La figure 2.8 schématise cet algorithme.

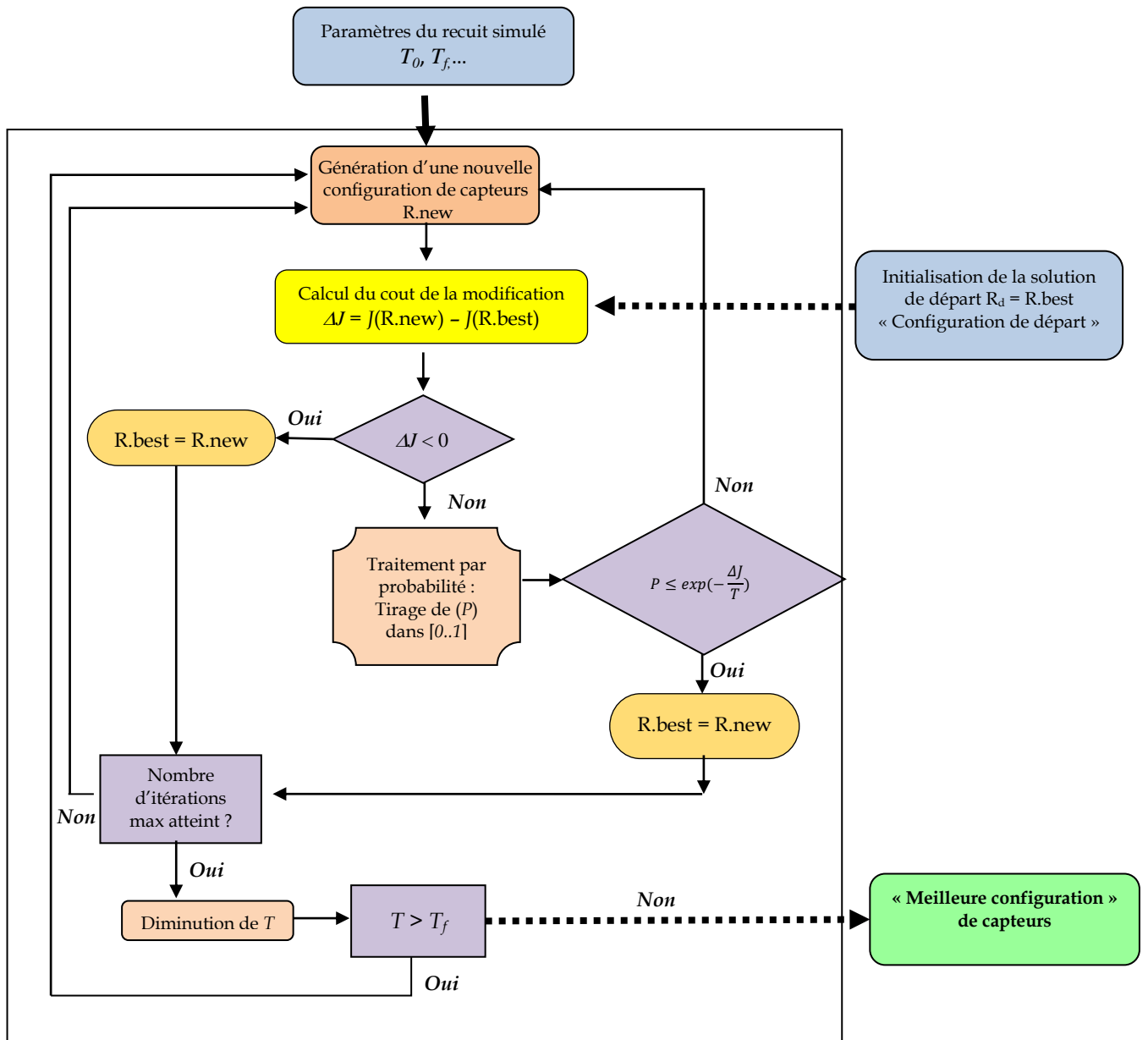


Figure 2.8. Principe de fonctionnement de l’algorithme du recuit simulé

### 2.5.3.1.3. Paramétrage

Dans la littérature, nous trouvons des méthodes générales pour paramétrer le recuit simulé (Siarry, 2014). Néanmoins, ces méthodes ne sont pas forcément les plus efficaces. Nous présentons, ci-dessous, une méthode standard de paramétrage du recuit simulé :

- Température de départ

Au début des itérations, la recherche est supposée éloignée de l’optimum, en effet, il est plus convenable de choisir une probabilité d’acceptation élevée. (Kirkpatrick et al, 1983) ont

proposé une probabilité moyenne  $P_0 = 0,8$ . Pour déterminer la température de départ, une première recherche peut être effectuée. Ainsi, pour un échantillon de configurations de capteurs (aléatoires), tous les accroissements de la fonction objectif  $\delta J^+$  sont acceptés. Enfin, en déterminant la moyenne  $\overline{\delta J^+}$  et en utilisant la règle de Métropolis, la relation entre la probabilité et la température de départ est donnée par

$$P_0 = \exp\left(\frac{-\overline{\delta J^+}}{T_0}\right) \quad (2.19)$$

Enfin la température de départ est obtenue par

$$T_0 = \frac{-\overline{\delta J^+}}{\log(0,8)} \quad (2.20)$$

Cette méthode pour la détermination de la température de départ a des faiblesses, du fait qu'elle dépend étroitement de la taille de l'échantillon choisi. Aussi, la probabilité d'acceptation  $P_0$  élevée engendre, dans certain cas, la perte d'itérations inutilement sur l'ensemble des itérations et augmente ainsi le temps de calcul.

- Schéma de refroidissement

Dans la littérature, plusieurs modèles ont été proposés pour simuler la décroissance de la température. Ces modèles sont, souvent, nommés schémas de refroidissement (« cooling schedule » en anglais). De manière générale, ces schémas peuvent être classés en trois catégories : décroissance non-monotone (la température décroît à chaque itération avec la possibilité d'accroissements occasionnelles), décroissance continue (la température décroît à chaque itération) et décroissance par paliers (la température est conservée constante pendant un nombre d'itérations donné, et décroît ainsi par paliers).

(*Nourani et Andresen, 1998*) ont donné un aperçu détaillé sur les stratégies de refroidissement les plus importantes (exponentielle, logarithmique, et linéaire) et ont effectué une étude comparative de la performance du recuit simulé en utilisant ces schémas de refroidissement. Ils ont ainsi conclu que le schéma de refroidissement exponentiel est le plus performant. Egalement, d'après (*Abida, 2010*), le schéma de décroissance de la température, le plus simple et d'ailleurs le plus utilisé, suit une loi exponentielle.

Ce schéma relie les températures au niveau de deux paliers successifs  $i$  et  $i+1$  par :

$$T_{i+1} = \theta T_i \quad (2.21)$$

Où  $\theta \in [0,1]$  est le facteur de décroissance.

- Température d'arrêt

Pour arrêter les itérations, aucune règle générale n'est proposée dans la littérature. Ainsi, il est possible d'arrêter le calcul lorsqu'aucune amélioration dans la fonction coût n'est observée durant un nombre conséquent de combinaisons ou lorsque la température d'arrêt  $T_f$  est inférieure à une fraction donnée  $r_a$  de la température de départ  $T_0$ .  $T_f$  peut-être, donc, déterminée à partir de la relation suivante

$$T_f = r_a \times T_0 \quad (2.22)$$

- Longueur d'un palier de température

Ce paramètre, que nous notons  $L$ , est utilisé lorsque le schéma de refroidissement est par paliers (décroissance de la température par paliers). La longueur d'un palier est le nombre d'itération à effectuer à chaque niveau de température. Dans la littérature aucune règle générale n'est proposée pour déterminer une longueur convenable. Ce nombre est généralement fixe et choisi d'une façon proportionnelle à la taille du problème.

#### 2.5.3.1.4. *Avantages et inconvénients*

Les avantages du recuit simulé sont nombreux :

- Il est efficace pour résoudre les problèmes d'optimisation combinatoire souvent de grandes tailles,
- Il est relativement simple à programmer et à mettre en œuvre,
- Il propose des solutions de bonne qualité car à haute température, il est capable de s'extraire des minima locaux (en acceptant occasionnellement une augmentation de la fonction coût) (*Kirkpatrick et al, 1983*).
- Il offre une souplesse d'emploi puisque nous pouvons y inclure facilement de nouvelles étapes intermédiaires,
- Il peut être arrêté lorsque l'utilisateur le souhaite.

L'inconvénient majeur du recuit simulé est qu'à basse température, il peut se trouver piégé dans un minimum local et il lui est impossible d'en sortir seul. Pour résoudre ce genre de problème, plusieurs solutions ont été proposées, nous citons par exemple :

- L'acceptation d'une brusque remontée de la température, de temps en temps, pour relancer la recherche sur d'autres configurations plus éloignées.

- La limitation de la descente de la température en imposant une valeur minimale au-delà de laquelle le palier de température n'est plus changé.

D'autres inconvénients sont : le temps de calcul qui peut devenir excessif dans certains cas et le paramétrage (température de départ, température d'arrêt, etc.) très empirique, qui peut influencer la qualité de la solution, le recuit simulé est aussi sensible à la configuration initiale choisie. Notons qu'il est important d'admettre que la diminution de la température doit être suffisamment lente et qu'un nombre important de palier doit être respecté.

### 2.5.3.2. L'Algorithme Génétique

L'algorithme génétique (AG) proposé par (Holland, 1975) est une méthode qui appartient à la famille des algorithmes évolutionnaires (« evolutionary computation » en anglais). Il est basé sur une analogie avec les mécanismes d'évolution de la nature (Fogel, 1963; Fogel et al, 1966 ; Rechenberg, 1973 ; Schwefel, 1977) et repose sur un codage des variables en structures chromosomiques (génotypique). (De Jong, 1975) été le premier à avoir orienté l'utilisation des AG dans le cadre de la résolution des problèmes d'optimisation (De Jong, 1992). Généralement, un algorithme génétique peut être définie par (Lerman et Ngouenet, 1995 ; Novkovic et Sverko, 1997) :

- *Individu ou chromosome* : une solution potentielle du problème,
- *Population* : un ensemble de chromosomes ou de points de l'espace de recherche,
- « *Fitness* » : la fonction objectif à maximiser ou minimiser,
- *Environnement* : l'espace de recherche.

#### 2.5.3.2.1. Principe

L'algorithme génétique s'appuie sur la théorie Darwinienne (Darwin, 1859) de l'évolution génétique d'une population. Dans la nature, les individus d'une population se trouvent en compétition, par exemple lors de la recherche de nourriture ou pour la reproduction. D'après la théorie de l'évolution, les individus les plus adaptés survivent majoritairement et ont une descendance. Par suite, les caractéristiques génétiques de ces individus sont transmises à la génération suivante, qui voit potentiellement sa performance augmenter. Le patrimoine génétique de la population s'améliore ainsi de générations en générations.

Un modèle de cette évolution a inspiré l'algorithme génétique : un procédé de codage est introduit, de manière à générer une structure génétique à partir des variables du problème

d'optimisation. Ensuite, « des opérateurs » (sélection, croisement et mutation) permettant l'évolution du patrimoine génétique de la population d'une génération à l'autre sont définis. Ces opérateurs font largement appel à des lois aléatoires. La performance d'un individu, qui détermine sa capacité à survivre, est liée à la valeur de la fonction de coût  $J$ . La procédure d'optimisation consiste alors à simuler l'évolution d'une population, durant un certain nombre de générations, jusqu'à la détermination d'un individu optimal.

### 2.5.3.2.2. Algorithmes

Le principe, dans le cadre de l'optimisation de réseaux de capteurs, est le suivant : nous commençons le processus par générer une population initiale formée par  $k$  configurations de  $m$  capteurs (aléatoires).  $k$  représente, ainsi, la taille de la population et chaque configuration représente un chromosome formé de  $m$  gènes (voir figure 2.9). Dans notre étude, nous utilisons la méthode de codage (transformer les gènes des chromosomes) réel qui consiste à représenter les variables du problème d'une façon naturelle. Cette méthode est la plus récente et la plus efficace (Michalewicz, 1992 ; Michalewicz et al, 1994) vu l'absence de l'étape de transcodage (retour vers le réel), mais dans la littérature nous trouvons d'autres les méthodes de codage dite « binaire » et « gray » (Glover et al, 2003 ; Dréo et al 2006).

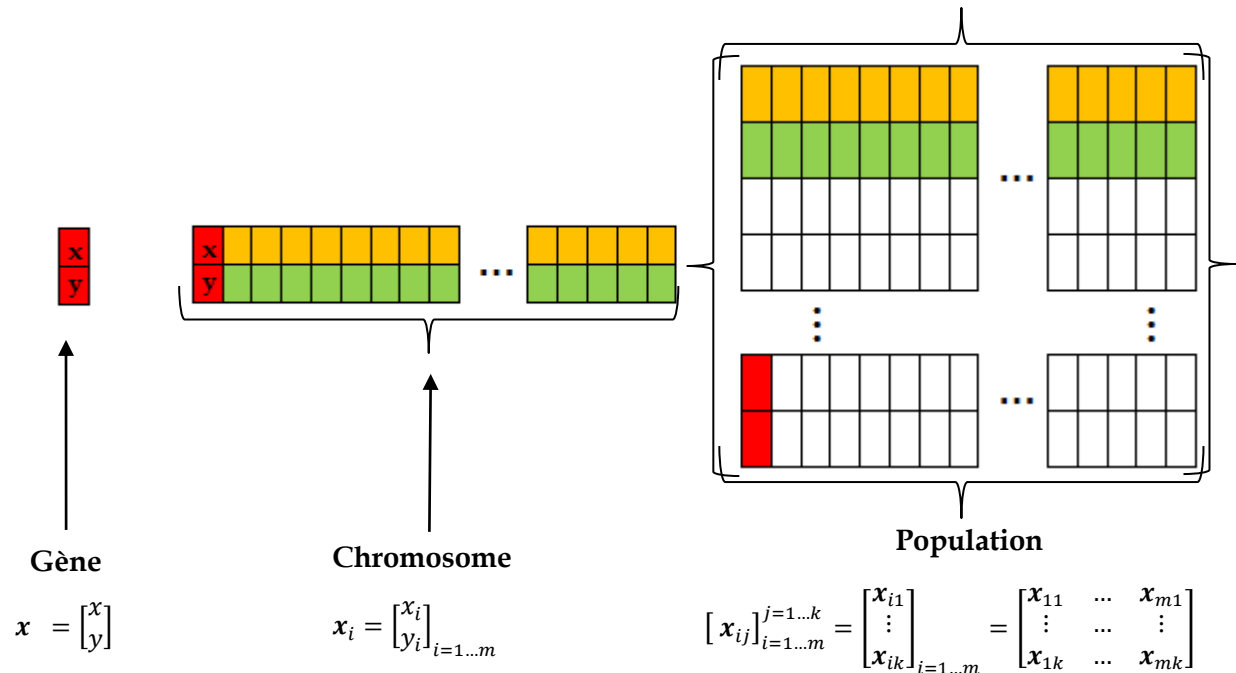


Figure 2.9. Principe de construction d'une population formée par  $k$  configurations de  $m$  capteurs

Après avoir construit la population initiale, on évalue la qualité (dit « fitness ») de chaque configuration par la fonction objectif  $J$ . Cela revient à estimer la probabilité de sélection des différentes configurations. Une fois toute la population évaluée, on sélectionne les

configurations qui deviendront parents (à conserver dans la prochaine génération). On trouve, dans la littérature, de nombreuses méthodes de sélection (*Goldberg, 1989, Bourazza, 2006, Merzouk, 2007 ; Talbi, 2009 ; Boullagui, 2011*): par roulette, par tournoi, « steady-state », etc. Dans notre cas, nous choisissons la technique de sélection par rang qui consiste à trier d'abord la population par fitness (en fonction des performances). Ensuite, nous attribuons à chaque configuration de capteurs un rang en fonction de la valeur de sa fonction coût. Ainsi la mauvaise configuration aura le rang 1 jusqu'au meilleur chromosome qui aura le rang  $k$  : les parents de chaque population auront donc le rang  $k$  et  $k - 1$ . Puis nous passons aux opérations d'hybridation qui permettent de passer d'une génération à une autre. Ces opérations sont soit le croisement (à 1 point, à  $k$  point, uniforme, etc.), qui consiste en une opération d'accouplement entre les deux meilleures configurations appartenant à une génération (voir figure 2.10), soit la mutation, qui consiste à modifier un ou plusieurs capteurs (gène) d'une configuration (chromosome), (voir figure 2.11) (des informations sont disponibles sur les différents schémas de croisement et de mutation (*Siarry, 2014*)).

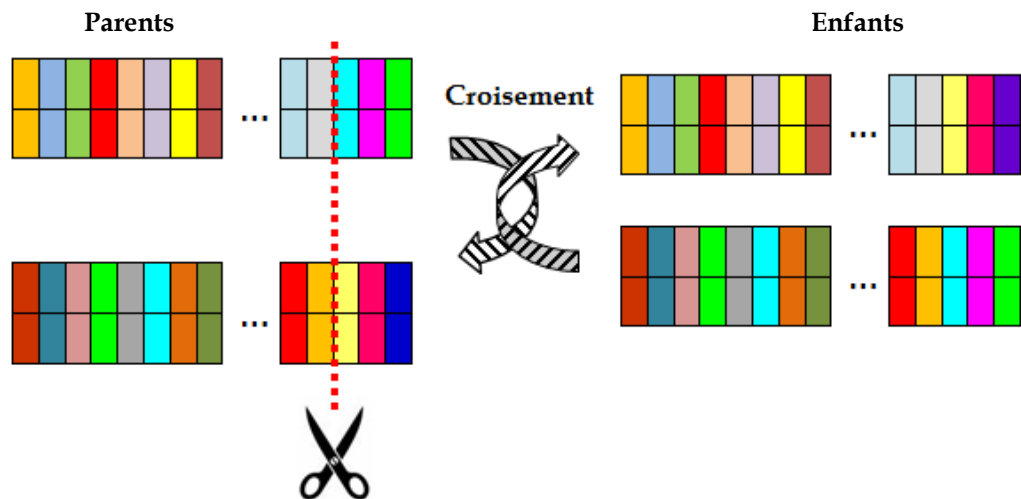


Figure 2.10. Principe du croisement entre deux configurations de capteurs

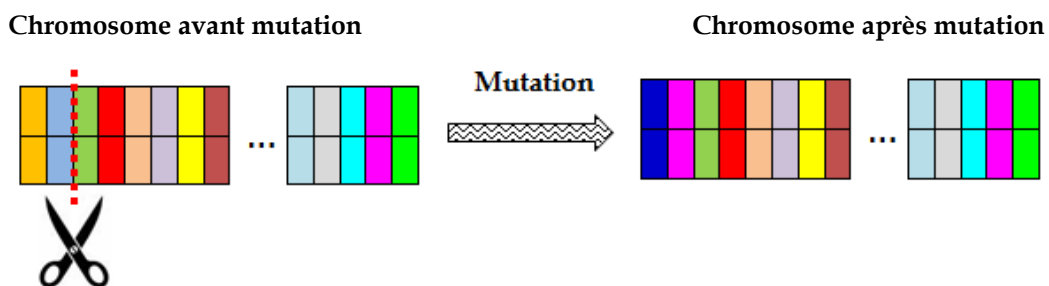


Figure 2.11. Principe de la mutation d'une configuration de capteurs

Enfin, ces nouvelles configurations (hybrides) seront mélangées avec d'autres configurations choisies aléatoirement pour former une nouvelle population. Pendant la création de cette

nouvelle population, il y a de grandes chances que les configurations hybrides ne soient pas sélectionnées comme étant des parents. La méthode d'élitisme peut remédier à ce problème. Elle consiste à utiliser un ou plusieurs des meilleurs chromosomes pour former une partie de la nouvelle population. Ensuite, on génère le reste de la population d'une manière aléatoire. Cette méthode améliore l'efficacité de l'algorithme, car elle permet de conserver les meilleures solutions. Enfin, pour arrêter le processus itératif nous avons besoin d'un test permettant de juger la qualité de la population vis à vis des objectifs. Nous pouvons aussi l'arrêter au bout d'un nombre de générations laissé à l'appréciation de l'utilisateur. Sur la figure 2.12 est schématisé le principe général de fonctionnement de l'algorithme génétique.

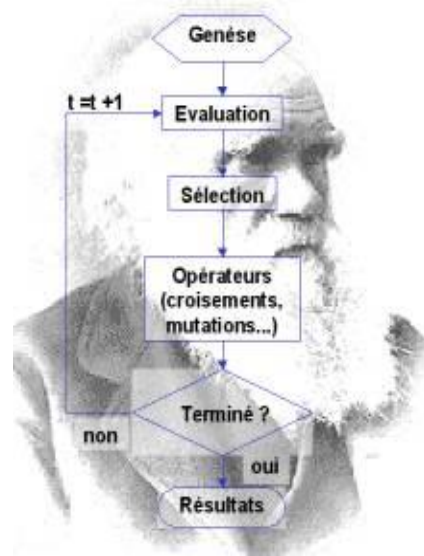


Figure 2.12. Principe général de l'algorithme génétique (Magnin, 2006) <sup>2</sup>

### 2.5.3.2.3. Paramétrage

Les paramètres à fixer pour assurer le fonctionnement de l'algorithme génétique sont : la taille  $k$  d'une population, le nombre de générations (pour fixer la fin du processus d'évolution) et le taux (probabilité) de mutation ou de croisement. Dans la littérature, il n'existe pas de paramétrage universel adapté à la résolution de tous les problèmes qui peuvent être posés. D'une façon générale, il n'est pas possible de trouver des solutions convenables en moins de 10 générations et au bout de 500 générations, les solutions n'évoluent plus (Mezhoud, 2012) (ceci n'est qu'un ordre de grandeur, tout dépend du problème à résoudre).

<sup>2</sup> La photo de fond est celle de Charles Darwin. Elle est associée à l'Algorithme Génétique car il est inspiré du concept de « sélection naturelle » élaboré par Charles Darwin. Le vocabulaire employé dans cet algorithme est directement calqué sur celui de la théorie de l'évolution et de la génétique.



Dans la littérature, des valeurs sont souvent utilisées à titre indicatif (*Goldberg, 1988 ; Novkovic et Sverko, 1997 ; Yildizoglu et Vallée, 2004*):

- La taille d'une population  $T_p$  se situe entre 30 et 50 individus,
- Le nombre de génération  $N_g$  est entre 500 et 1000,
- Le taux de croisement  $T_c$  est généralement entre 0,5 et 0,9,
- Le taux de mutation  $T_m$  pour l'obtention de bons résultats est de l'ordre de 0,01.

#### 2.5.3.2.4. *Avantages et inconvénients*

Les avantages de l'algorithme génétique sont nombreux :

- Il parvient à trouver de bonnes solutions à des problèmes très complexes,
- Il n'introduit pas de grandeur supplémentaire dans le processus de recherche comme le font les algorithmes probabilistes,
- « Il permet de traiter des problèmes où la fonction à optimiser ne présente aucune propriété de continuité ou de dérivabilité, par exemple » (*Bouallagui, 2011*).

Parmi les inconvénients de cet algorithme on peut noter :

- Il est coûteux en temps de calcul (vitesse de convergence lente) car :
  - Pour sélectionner les meilleurs chromosomes il doit évaluer l'intégralité de chaque population (cas de sélection par rang).
  - Il doit effectuer des opérations supplémentaires (croisement, mutation). Lorsque des probabilités « inadéquates » de croisement et mutation sont utilisées, le temps de calcul peut augmenter considérablement.
  - Un temps additionnel peut être consacré au codage et au transcodage des chromosomes.
- Il n'est pas trivial à mettre en œuvre.
- Dans la littérature, on ne trouve pas de méthode standard pour la définition du critère d'arrêt et des différents paramètres nécessaires à son fonctionnement.
- Certaines étapes ne sont pas réalisables lorsque les chromosomes sont de petite taille (par exemple, pour le croisement, les individus ne peuvent pas transmettre assez d'information génétique ce qui génère une ressemblance des descendance). Nous reviendrons sur cet inconvénient dans le chapitre 4.

## 2.6. CONCLUSION DU CHAPITRE

Tout au long de ce chapitre, nous avons cerné les différents aspects et méthodologies qui permettent la conception optimale d'un réseau de capteurs. En rapport avec les objectifs définis au premier chapitre, nous avons donné un aperçu non exhaustif des critères fondamentaux utilisés pour la conception optimale. Dans le cadre de cette étude, nous retiendrons principalement les critères basés sur la notion d'information, sur la notion d'erreur d'estimation et sur la notion de couverture. Cette dernière est particulièrement importante et nous l'utiliserons pour décrire le concept de sensibilité (défini par ([Hourdin et Issartel, 2000](#))). Notons que la sensibilité d'un capteur fluctue en fonction des conditions météorologiques. Ceci doit impérativement être pris en compte dans le processus de conception, soit sous forme de scénarios correspondant à différentes conditions, soit sous forme probabiliste si les caractéristiques météorologiques du site sont connues.

Nous avons également défini les zones spatiales que nous étudierons comme étant des espaces discrets comportant un nombre limité de positions potentielles parmi lesquelles nous sommes amenés à sélectionner les meilleures pour implanter les capteurs de notre réseau optimal. Ce problème est qualifié dans la littérature de « *NP-difficile* ». Pour contourner cette difficulté, nous avons pu identifier des moyens algorithmiques adéquats permettant d'aboutir à des solutions concrètes. Effectivement, nous avons sélectionné parmi plusieurs méthodes possibles, l'algorithme probabiliste du recuit simulé et l'algorithme évolutionnaire génétique. Après avoir expliqué le principe et le schéma de fonctionnement de ces algorithmes nous avons indiqué leurs principaux avantages et inconvénients. Pour des raisons pratiques nous favoriserons l'algorithme du recuit simulé. En effet, lors de sa mise en œuvre, nous nous sommes aperçus que l'algorithme génétique rencontrait des difficultés pour les réseaux de petite taille (faible nombre de capteurs). Ce problème vient du fait qu'une configuration de capteurs de petite taille ne fournit pas assez d'informations génétiques, ce qui bloque l'opération d'hybridation. Nous reviendrons en détails sur ce problème dans le chapitre 4.

Ainsi, nous avons défini une méthodologie généralisable à tout problème d'optimisation de réseaux de surveillance. Cette dernière comporte quatre phases principales :

1. La définition de l'objectif de conception,
2. La définition du critère d'optimalité,
3. La définition des zones d'études,
4. Le choix de l'algorithme d'optimisation.

« **BIBLIOGRAPHIE CHAPITRE 2** »

Abida, R, (2010). Construction optimale de réseaux fixes et mobiles pour la surveillance opérationnelle des rejets accidentels atmosphériques. Modélisation et simulation. Ecole des Ponts ParisTech. Français. <https://pastel.archives-ouvertes.fr/pastel-00638050>.

Abida, R, et Bocquet, M., (2009). Targeting of observations for accidental atmospheric release monitoring. *Atmospheric Environment*, 43, 6312–6327, doi: 10.1016/j.atmosenv.2009.09.029.

Abida, R., Bocquet, M., Vercauteren, N., Isnard, O., (2008). Design of a monitoring network over France in case of a radiological accidental release. *Atmospheric Environment*, 42, 5205–5219, doi: 10.1016/j.atmosenv.2008.02.065.

Altinel, I.K., Aras, N., Güney, E., Ersoy, C., (2008). Binary integer programming formulation and heuristics for differentiated coverage in heterogeneous sensor networks. *Computer Networks*, 52: 2419–2431.

Ammari, H. M. et Das, S. K., (2012). Centralized and clustered k-coverage protocols for wireless sensor networks. *IEEE Transactions On Computers*, 61:118–133.

Bai, X., Kumar, S., Xuan, D., Yun, Z., Lai, T. H., (2006). Deploying wireless sensors to achieve both coverage and connectivity. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, pages 131–142.

Berry, J., Hart, W., Phillips, C. A., and Watson, J. P., (2006-a). A facility location approach to sensor placement optimization. 8th Annual Symp.on Water Distribution Systems Analysis, Cincinnati, Ohio, Environmental and Water Resources Institute of ASCE (EWRI of ASCE), New York.

Bernardo, J. M., (1975). Information theory and decision making. *Theories of Decision in Practice*, White, D. J, and Bowen, K. C (Eds). London : Hodder and Stoughton, 247-251.

Biellia, M., Caramiab, M., Carotenutoc, P., (2002). Genetic algorithms in bus network optimization. *Science direct* Volume 10, Issue 1, Pages 19–34. doi.org/10.1016/S0968-090X(00)00048-6.

Bocquet, M, (2006). Construction optimale de réseaux de mesure application à la surveillance des polluants aériens. Notes de cours de l'École Nationale Supérieure des Techniques Avancées ParisTech. CERIA, Ecole des Ponts ParisTech Université Paris-Est et INRIA.

Bouallagui, S, (2011). Techniques d'optimisation déterministe et stochastique pour la résolution de problèmes difficiles en cryptologie. *Mathématiques générales [math.GM]*. INSA de Rouen, 2010. Français. <https://tel.archives-ouvertes.fr/tel-00557912>.

Bourazza, S., (2006). "Variantes d'algorithmes génétiques appliquées aux problèmes d'ordonnancement", Thèse de doctorat - Université du Havre. Français. <https://tel.archives-ouvertes.fr/tel-00126292v2>.

Caselton, W. et Husain, T., (1980). Hydrologic networks: information transmission. *J. Water Resources Planning and Management*, 106, 503–529.

Colorni, A., Dorigo M., Maniezzo, V., (1991) Distributed Optimization by Ant Colonies, actes de la première conférence européenne sur la vie artificielle, Paris, France, Elsevier Publishing, 134-142.

Darwin, C, R, (1859). On the Origin of Species by means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life, John Murray, London.

De Jong, K.A., (1975). An analysis of the behaviour of a class of genetic adaptive systems, Doctoral Dissertation, University of Michigan, Ann Arbor, Michigna.

De Jong, K.A., (1992). Are genetic algorithms function optimizers? In: Manner, R. & Manderick, B. (Eds.) Parallel Problem-Solving from Nature, 2, Elsevier Science Publishers, Amsterdam, 3-13.

Dhillon, S., et Chakrabarty, K., (2003). Sensor placement for effective coverage and surveillance in distributed sensor networks, IEEE Wireless Communications and Networking.

Di Zio, S., Fontanella, L., Ippoliti, L., (2004). Optimal spatial sampling schemes for environmental surveys. Environmental and Ecological Statistics 11, 397-414.

DORIGO, M, (1992). Optimization, Learning and Natural Algorithms. Thèse de doctorat, Politecnico di Milano, Italy.

Dréo, J., Petrowski, A., Taillard, É. D., Siarry. P., (2003). Métaheuristiques pour l'optimisation difficile. Eyrolles (Editions), November 2003. ISBN 2212113684.

Dréo, J., Pétrowski, A., Siarry, P., Taillard, E., (2006). Metaheuristics for Hard Optimization Methods and Case Studies.

Farmer, J. D., Packard, N. H., Perelson, A. S., (1986).The immune system, adaptation, and machine learning. Phys. D, 2(1-3): 187-204. ISSN 0167-2789.

Farshbaf, M. et Feizi-Darakhshi, M., (2009). "Multi-objective Optimization of Graph Partitioning using Genetic Algorithms". Third International Conference on Advanced Engineering Computing and Applications in Sciences pp. 1-6.

Fedorov, V, V, (1972). Theory of Optimal Experiments. Academic Press, New York.

Fedorov, V. V., et Müller, W. G., (1988). Two approaches in optimization of observing networks. Optimal Design and Analysis of Experiments, Dodge, Y., and Fedorov, V. V, and Wynn, H. P., New York, : North-Holland

Fedorov, V. V. et Müller, W. G., (1989). Comparison of two approaches in the optimal design of an observation network'. Statistics, 20, 339-351.

Feo, T. A., et Resende, M. G. C., (1989). A probabilistic heuristic for a computationally difficult set covering problem. Operations Research Letters, 8(2): 67-71.

Fogel, L, J, (1963). Biotechnology: Concepts and Applications, Prentice-Halls, Englewood Cliffs, NJ.

Fogel, L.J., Owens, A.J., Walsh, M.J., (1966). *Artificial Intelligence through Simulated Evolution*, John Wiley, New York.

Ghosh, A. et Das, S. K., (2008). Coverage and connectivity issues in wireless sensor networks : A survey. *Pervasive and Mobile Computing*, 4 :303–334.

Glover, F, (1977). Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8(1):156–166.

Glover, F, (1986). Future paths for integer programming and links to artificial intelligence. *Computers and Operations Research*, 13(5): 533–549, 1986.

Glover, Fred W., Kochenberger, Gary A. (2003). *Handbook of Metaheuristics*.

Goldberg, D., E. et Holland J., H. (1988). "Genetic Algorithms and Machine Learning", *Machine Learning*, vol. 3, pp.95-99.

Goldberg, D, (1989). *Genetic Algorithm In Search, Optimization And Machine Learning*, Addison-Wesley.

Guttorp, P., Le, N. D., Sampson, P. D., Zidek, J. V., (1993). Using entropy in the redesign of an environmental monitoring network. *Multivariate Environmental Statistics*, Eds, Patil, G. P., Rao, C. R., New York : North Holland/Elsevier Science, 175-202.

Hamel, D., Chwastek, M., Garcia, M., Farouk, B., Kam, M., Dandekar, K. R., (2010). Sensor Placement for Urban Homeland Security Applications. *International Journal of Distributed Sensor Networks*, vol. 2010, Article ID 859263, p15.

Holland, J, H, (1975). *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Michigan.

Hourdin, F. et Issartel, J. -P., (2000). Sub-surface nuclear tests monitoring through the CTBT Xenon Network. *AN AGU Journal Volume 27, Issue 151, Pages 2245–2248. DOI: 10.1029/1999GL010909.*

Idrees, A. K., Deschinkel, K., Salomon, M., Couturier, R., (2015). Perimeter-based coverage optimization to improve lifetime in wireless sensor networks. *Journal Engineering Optimization Volume 48, 2016 - Issue 11 Pages 1951-1972.*

Jiang, Z., de Bruin, S., Heuvelink, G.B.M., Twenhofel, C.J.W., (2007). Optimization of mobile radioactivity monitoring networks, *Fifth International Symposium on Spatial Data Quality*, Enschede, The Netherlands, 2007-06-13/2007-06-15.

Kennedy, J. et Eberhart, R., (1995). Particle Swarm Optimization, In *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, vol. 4, 1995.

Kiefer, J. et Wolfowitz, J., (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12, 363–366.

Kiefer, J, et Wolfowitz, J., (1952). Stochastic estimation of the maximum of a regression function *Ann. Math. Statistics*, 23:462--466.

- Kirkpatrick, S., Gelatt Jr, C.D., Vecchi, M.P., (1983). Optimization by simulated annealing. *Science* 220 (4598), 671-680.
- Ko, C.-W., Lee, J., Queyranne, M., (1995). An exact algorithm for maximum entropy ampling. *Operations Research*, 43, 684-691.
- Kouichi, H., Turbelin, G., Ngae, P., Feiz, A. A., Barbosa, E., & Chpoun, A., (2016). Optimization of sensor networks for the estimation of atmospheric pollutants sources 24th International Conference on Modelling, Monitoring and Management of Air Pollution.
- Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., Faloutsos, C., (2008). Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks. *Journal of Water Resources Planning and Management (JWRPM)*.
- Lerman, I., et Ngouenet, F., (1995). Algorithmes génétiques séquentiels et parallèles pour une représentation affine des proximités, Rapport de Recherche de l'INRIA Rennes - Projet REPCO 2570, INRIA
- Le, N. D., Sun, L., Zidek, J. V., (2003). Designing networks for monitoring multivariate environmental fields using data with monotone pattern. Tech. rep., Statistical and Applied Mathematical Sciences Institute, RTP, NC.
- Le, N. D., et Zidek, J. V., (1994). Network designs for monitoring multivariate random spatial fields. *Recent. Adv. in Statist. and Prob, Eds., Vilaplana, J. P., Puri, M. L., 191-206.*
- Liu, C., Wu, K., Xiao, Y., Sun, B., (2006). Random coverage with guaranteed connectivity: Joint scheduling for wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 17 :562-575.
- Lovejoy, S., Schertzer, D., Ladoy, P., (1989). Fractal characterization of inhomogeneous geophysical measuring networks. *Nature*, 319, 43 - 44.
- Magnin, V., (2006). Optimisation et algorithmes génétiques. Cours en ligne de Polytech'Lille. <http://magnin.plil.net/spip.php?article45>
- Mason, L. R., et Bohlin, J. B., (1995). Network optimization of a radionuclide monitoring system for the comprehensive nuclear test ban treaty. Tech. rep., Pacific-Sierra Research Corporation, PSR Technical Report 2585.
- Merzouk, S, E, (2007). Problème de dimensionnement de lots et de livraisons : application au cas d'une chaîne logistique. Thèse de doctorat - Université de Technologie de Belfort-Montbéliard et Université de Franche-Comté. <http://indexation.univ-fcomte.fr/nuxeo/site/esupversions/ac855fab-3457-4b83-8c66-59aa9ced27e1>.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21, 1087-1092.
- Mezhoud, N., (2012). Une approche hybride pour la segmentation d'images Application aux images médicales. Rapport de thèse de doctorat. Page 31.



<http://bu.umc.edu.dz/theses/informatique/MEZ6197.pdf>

Michalewicz, Z., (1992). Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag.

Michalewicz, Z., Logan, T., Swaminathan, S., (1994). Evolutionary operators for continuous convex parameters spaces, in A. Sebald & L. Fogel, eds, 'Proceedings of the 3rd Annual Conference on Evolutionary Programming', World Scientific Publishing, pp. 84-107.

Mini, S., Udgata, S. K., Sabat, S. L., (2011). Artificial bee colony based sensor deployment algorithm for target coverage problem in 3-d terrain. In Proceedings of the 7th International Conference on Distributed Computing and Internet Technology, pages 313-324.

Mühlenbein, H., et Paaß, G., (1996). From recombination of genes to the estimation of distributions i. binary parameters. In Proceedings of the 4th International Conference on Parallel Problem Solving from Nature, PPSN IV, pp. 178-187, London, UK. Springer-Verlag. ISBN 3-540-61723-X.

Nazrul Alam, S. M., et Haas, Z. J., (2006). Coverage and connectivity in three-dimensional networks. In Proceedings of the 12th annual International Conference on Mobile computing and networking, pages 346-357.

Novkovic, S., et Sverko, D., (1997). Genetic waste and the role of diversity in genetic algorithm simulations. Working Paper. Saint Mary's University. Canada.

Nourani, Y. et Andresen, B., (1998). A comparison of simulated annealing cooling strategies. Journal of Physics A General Physics 31(41):8373-8385 ·DOI: 10.1088/0305-4470/31/41/011.

Nunes, L. M., Paralta, E., Cunha, M. C., Ribeiro, L., (2004). Groundwater nitrate monitoring network optimization with missing data. Water Resour. Res, 40.

Nychka, D. et Saltzman, N., (1998). Design of air quality monitoring networks, chapter 4 of case studies in environmental statistics. In: Nychka, D., Piegorsch, W., Cox, L.H. (Eds.), Springer Lecture Notes in Statistics, Number 132. Springer Verlag, New-York, pp. 51-76.

Potuzak, T., (2014). "Time Requirements of Optimization of a Genetic Algorithm for Road Traffic Network Division Using a Distributed Genetic Algorithm". <em>Issues and Challenges in Artificial Intelligence Studies in Computational Intelligence</em> vol. 559 pp. 155-166.

Raes, F., De Cort, M., Graziani, G., (1991). The multi-fractal nature of radioactivity deposition on soil after the chernobyl accident. Health physics, 61, 271-274.

Reay, J. S. S. et Swift-Hook, D. T., (1979). The philosophy of monitoring [and discussion]. Phil. Trans. of the Royal Soc A, 290, 609-623. DOI: 10.1098/rsta.1979.0017.

Rechenberg, I., (1973). Evolutionsstrategie : Optimierung technischer Systeme nach Prinzipien der biologischen Evolution, Frommann-Holzboog Verlag, Stuttgart. (2nd Edition 1993).

Reed, P., Minsker, B., Valocchi, A. J., (2000). Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation. Water Resour. Res, 36, 3731-3741.

Rourke, J, O, (1987). Art gallery theorems and algorithms. Oxford University Press, Inc, Oxford, 1092.

Saunier, O., Bocquet, M., Mathieu, A., Isnard, O., (2009). Model reduction via principal component truncation for the optimal design of atmospheric monitoring networks. Atmospheric Environment 43. 4940–4950.

Schwefel, H,-P, (1977). Numerische Optimierung von Computer-modellen mittels der Evolutionsstrategie, Birkhauser Verlag, Basel. (English edition : Numerical Optimization of Computer Models, John Wiley & Sons, Chichester, 1981).

Siarry, P, (2014). Métaheuristiques Recuits simulé, recherche avec tabous, recherche à voisinages variables, méthodes GRASP, algorithmes évolutionnaires, fourmis artificielles, essais particuliers et autres méthodes d'optimisation. Date de parution : 06/03/2014 EAN13 : 9782212139297 Diffusé par Geodif.

Shakkottai, S., Srikant, R., Shroff, N., (2005). Unreliable sensor grids: Coverage, connectivity and diameter. Ad Hoc Networks, 3 :702–716.

Shannon .C.E. et Weaver. W., (1949). The mathematical Theory of communication, University of Illinois, Urbana III.

Silvey, S, D, (1980). Optimal design. Chapman and Hall, London.

Talbi, E-G., (2009). Metaheuristics: From Design to Implementation. Wiley Publishing.

Türk, S., Liu, Y., , Radeke, R., Lehnert R., (2012). Network Migration Optimization Using Genetic Algorithms. Information and Communication Technologies pp 112-123. 18th EUNICE/ IFIP WG 6.2, 6.6 International Conference, EUNICE 2012, Budapest, Hungary, August 29-31, 2012.

Wang, X. et Wang, S., (2011). Hierarchical deployment optimization for wireless sensor networks. IEEE Transactions on mobile computing, 10 :1028–1041.

Watson, J.-P., Greenberg, H. J., Hart, W. E., (2004). A multipleobjective analysis of sensor placement optimization in water networks. Proc., World Water and Environmental Resources Conf., ASCE, Reston Va.

Wu, J., Zheng, C., Chien, C. C., (2005). Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions. Journal of Contaminant Hydrology, 77, 41–65.

Wu, Y., Li, M., Cai, Z., Zhu, E., (2008). A distributed algorithm to approximate node-weighted minimum  $\alpha$ -connected  $(\theta, k)$ -coverage in dense sensor networks. In Proceedings of the 2nd annual international workshop on Frontiers in Algorithmics, pages 221–232.

Yildizoglu, M. et Vallée. T., (2004). Présentation des algorithmes génétiques et de leurs applications en économie. Revue d'Economie Politique, Editions Dalloz, 2004, 114, pp.711-745. <hal-00125103>.



Zidek, J. V., Sun, W., Le, N. D., (2000). Designing and integrating composite networks for monitoring multivariate (gaussian) pollution fields. *Journal of Applied Statistics*, 49, 63–79.

Zhao, Q., et Gurusamy, M., (2008). Lifetime maximization for connected target coverage in wireless sensor networks. *IEEE/ACM Transactions Networking*, 16 :1378–1391.

Zhou, Z., Das, S., Gupta, H., (2004). Connected k-coverage problem in sensor networks. In *Proceedings of the 13th International Conference on Computer Communications and Networks*, pages 373–378.

**CHAPITRE 3 :**  
**OPTIMISATION DE RESEAUX DE CAPTEURS**  
**POUR LA CARACTERISATION DE SOURCE**

## CHAPITRE 3

### OPTIMISATION DE RESEAUX DE CAPTEURS POUR LA CARACTERISATION DE SOURCE

#### Plan du chapitre 3

CHAPITRE 3 OPTIMISATION DE RESEAUX DE CAPTEURS POUR LA CARACTERISATION DE SOURCE.....	64
3.1. Introduction.....	67
3.2. Rappel de l'objectif et hypotheses simplificatrices .....	67
3.3. Problème direct VS Problème inverse .....	68
3.4. Approche basée source VS approche basée récepteurs .....	69
3.4.1. Approche basée source .....	70
3.4.2. Approche basée récepteurs .....	71
3.4.3. Avantages de l'approche basée récepteurs dans le cadre des problèmes inverses.....	74
3.5. Choix de la méthode d'inversion de données .....	76
3.5.1. Aperçu des techniques existantes .....	76
3.5.2. La méthode d'assimilation de données renormalisées .....	77
3.5.2.1. Historique de la méthode.....	77
3.5.2.2. Formalisme mathématique .....	78
3.5.2.3. Formulation discrète du problème .....	80
3.5.2.3.1. Technique de renormalisation discrète .....	81
3.5.2.3.2. Proposition d'un algorithme pour le calcul de la matrice de pondération.....	81
3.5.2.3.3. Estimation de la source.....	85
3.5.3. Modélisation des rétropanaches.....	87
3.5.3.1. Modélisation analytique et « CFD ».....	87
a. Approche analytique.....	87
b. Approche CFD .....	90
3.6. Evaluation de la methode d'inversion .....	92
3.6.1. Description du projet DYCE .....	92
3.6.2. Mise en œuvre de la méthode de renormalisation .....	96
3.6.3. Résultats et discussions .....	96
3.7. Utilisation de la méthode de renormalisation pour définir des critères d'optimalité.....	99

3.7.1. La méthode de renormalisation dans le cadre d'une approche paramétrique.....	99
3.7.2. La méthode de renormalisation dans le cadre d'une approche entropique .....	101
3.7.3. Définition des problématiques à étudier.....	103
3.7.3.1. Problème 1 « Détermination du nombre optimal de capteurs pour la caractérisation d'une source » .....	104
3.7.3.2. Problème 2 «Déploiement en urgence d'un réseau mobile sans information a priori sur la source » .....	104
3.7.3.3. Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source » .....	104
3.8.Conclusion du chapitre.....	105
« Bibliographie chapitre 3 ».....	106

### 3.1. INTRODUCTION

Tout au long de ce chapitre, nous décrivons une stratégie nommée « *détecter pour caractériser* », (« *detect to characterize* » en anglais) dédiée à la conception de réseaux de capteurs dont l'objectif principal est la *caractérisation* d'une source d'émission à partir des mesures de concentration issues du réseau. C'est un problème inverse, appelé problème d'Estimation du Terme Source (Source Term Estimation problem, en Anglais). Pour le résoudre, de nombreuses techniques existent, il est donc nécessaire d'introduire, dans ce chapitre, de nouveaux concepts théoriques/mathématiques liés à la définition de ce type de problème. Il est également impératif de donner un aperçu des techniques de modélisation de la dispersion atmosphérique. Ensuite nous présenterons l'approche retenue et nous proposerons un algorithme efficace pour la mettre en œuvre. Nous montrerons également qu'elle fournit des éléments permettant de définir des fonctions objectifs pouvant être utilisées dans le processus de conception optimale décrit au précédant chapitre. L'approche sera également validée à partir de données expérimentales. Enfin, plusieurs problèmes seront proposés afin d'évaluer le processus d'optimisation de réseaux.

### 3.2. RAPPEL DE L'OBJECTIF ET HYPOTHESES SIMPLIFICATRICES

Dans cette étude nous cherchons à concevoir des réseaux de capteurs dont l'objectif est de « *caractériser* » une source. Par *caractériser* nous entendons « *localiser* » spatialement et temporellement la source d'émission et « *quantifier* » son débit. Dans des situations opérationnelles, un réseau conçu de façon optimale doit permettre la mise en œuvre efficace d'actions appropriées afin de résoudre l'incident (intervention sur zone, neutralisation de l'attaque, etc.). En effet, cette étude est réalisée sous un ensemble d'hypothèses simplificatrices :

- Les zones d'intérêts s'étendent sur quelques kilomètres au plus (champ proche),
- Les zones à surveiller et leur environnement proche (zone rurale, semi-urbaine ou site industriel) forment un ensemble supposé plat et peu rugueux (faiblement bâti avec au plus quelques dizaines de bâtiments),
- Les rejets intentionnels ou accidentels, pris en compte dans notre étude, ont de fortes chances d'être des rejets localisés (fuite d'une installation industrielle, attaque terroriste, etc.) dont la source peut être considérée comme étant ponctuelle et statique,

- Afin de réduire la complexité, les rejets seront souvent supposés continus et stationnaires. Cette dernière hypothèse permet de simplifier les équations qui seront présentées dans ce chapitre. En effet, il faut bien noter que le processus défini, ainsi que la méthode d'inversion retenue, restent valables en cas de rejets transitoires.
- Enfin, nous considérerons que les conditions météorologiques sont connues ou peuvent être reconstituées à partir de données statistiques.

### 3.3. PROBLEME DIRECT VS PROBLEME INVERSE

Nous avons indiqué que le problème d'Estimation du Terme Source est un problème inverse. Sur le plan fondamental, il est utile de rappeler la différence entre un problème dit « direct » et un problème dit « inverse » en dispersion atmosphérique. La figure 3.1 schématise ces deux types de problèmes.

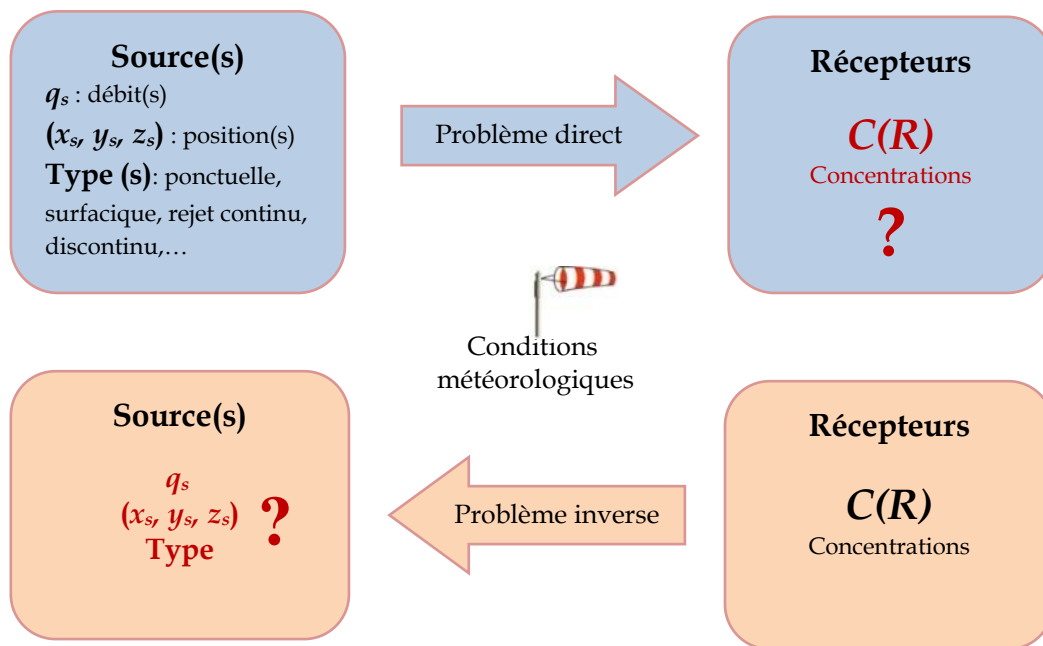


Figure 3.1. problème direct et inverse

La résolution du problème direct, consiste à prédire les concentrations ( $C$ ) au niveau du réseau ( $R$ ) de récepteurs connaissant les caractéristiques de la source (sa géométrie, sa position dans l'espace, son débit) et les conditions météorologiques. Le problème inverse consiste lui à « retrouver » l'origine des observations réalisées par le réseau. À partir d'un ensemble fini de mesures de concentrations récoltées à l'aide de  $m$  capteurs répartis sur le domaine d'étude, l'objectif est d'identifier au mieux, en espace et en temps, la zone géographique et la date des rejets (ainsi que son intensité, la hauteur du rejet, etc.). Ici, la source est, a priori, totalement inconnue. Le champ mesuré est le champ de concentration de

l'espèce surveillée par le réseau. Les  $m$  mesures réalisées, à un instant  $t$  donné, peuvent être assemblées dans un vecteur, dit vecteur de mesures, que nous noterons  $\boldsymbol{\mu} \in \mathbb{R}^m$  (dans le cas général du chapitre 2, un tel vecteur avait été noté  $\mathbf{Z}_R$ ).

$$\boldsymbol{\mu}(t) = \begin{pmatrix} \mu_1(t) \\ \mu_2(t) \\ \vdots \\ \mu_m(t) \end{pmatrix} \quad (3.1)$$

Pour interpréter les concentrations mesurées en divers points, et à divers moments, par un réseau de surveillance, il faut définir la relation qui existe entre la source et les mesures. Nous supposons ici, de façon classique, que la source peut être représentée, à un instant donné, par une fonction, dite fonction source, notée  $s(\mathbf{x}, t)$ . Cette fonction décrit spatialement et temporellement la répartition du taux d'émission d'une source émettant dans le domaine spatio-temporel  $\Omega = D \times [0, T]$ . Pour des conditions météorologiques données, le problème inverse consiste donc à trouver, dans  $\Omega$ , la fonction  $s(\mathbf{x}, t)$  qui explique au mieux les mesures  $\boldsymbol{\mu}(t)$ . Ce cas général correspond à une approche dite « non paramétrique » (le terme source est défini en tout point du domaine). Si, a priori, on peut supposer que la source est caractérisée par un nombre fini de paramètres, alors l'approche est dite « paramétrique » et le problème consiste à identifier ces paramètres. C'est le cas, par exemple, d'une source ponctuelle qui émet de façon continue. Cette dernière est caractérisée par sa position dans l'espace  $\mathbf{x}_s$  et par son débit  $q_s$  et l'on a  $s(\mathbf{x}) = q_s \delta(\mathbf{x} - \mathbf{x}_s)$ .

### 3.4. APPROCHE BASEE SOURCE VS APPROCHE BASEE RECEPTEURS

Comme nous l'avons déjà évoqué ci-dessus, pour interpréter les concentrations mesurées en divers points, et à divers moments, par un réseau de surveillance, il faut définir la relation qui existe entre la source et les observations. Pour cela, il est possible d'adopter soit une approche orientée « source » soit une approche orientée « récepteurs » (Ullasz, 1993). Ces deux alternatives sont illustrées en détail sur la figure 3.3. En effet, une mesure peut s'écrire :

- Soit en fonction du champ de concentration  $C(\mathbf{x}, t)$  (approche basée source) :

$$\mu(\mathbf{u}_i, t) = \mu_i(t) = \int_{\Omega} C(\mathbf{x}, t) \pi_i(\mathbf{x}, t) d\mathbf{x} dt + \epsilon_i \quad (3.2)$$

Où  $\pi_i$  est une fonction d'échantillonnage qui décrit, où et quand, la  $i^{\text{ème}}$  mesure a été réalisée et  $\epsilon_i$  est l'erreur de mesure. Notons que le champ de concentration dépend explicitement de la fonction source.

- Soit en fonction de la source  $s(\mathbf{x}, t)$  (approche basée récepteurs)

$$\mu(\mathbf{u}_i, t) = \mu_i(t) = \int_{\Omega} s(\mathbf{x}, t) a_i(\mathbf{x}, t) d\mathbf{x} dt + \epsilon_i \quad (3.3)$$

Où  $a_i(\mathbf{x}, t)$  exprime, pour un vent donné, la sensibilité du  $i^{\text{ème}}$  capteur vis à vis d'une émission unitaire localisée au point  $\mathbf{x}$  et à l'instant  $t$ .

Sans perte de généralité, dans le cas stationnaire, si le domaine étudié  $\Omega = D$  est discrétisé et comporte  $j=1 \dots N$  points  $\mathbf{x}_j(x_j, y_j, z_j)$ , le vecteur de mesure peut s'écrire :

- Soit en fonction du vecteur de concentration  $\mathbf{C} \in \mathbb{R}^N$  (approche basée source) :

$$\boldsymbol{\mu} = \mathbf{O}\mathbf{C} + \boldsymbol{\epsilon} \quad (3.4)$$

Où  $\mathbf{O}$  est la matrice d'observation, de dimension  $m \times N$ , et  $\boldsymbol{\epsilon} \in \mathbb{R}^m$  est le vecteur des erreurs de mesure.

- Soit en fonction du vecteur d'émission  $\mathbf{s} \in \mathbb{R}^N$  (approche basée récepteurs)

$$\boldsymbol{\mu} = \mathbf{A}\mathbf{s} + \boldsymbol{\epsilon} \quad (3.5)$$

Où  $\mathbf{A}$ , de dimension  $m \times N$ , est la matrice de sensibilité.

Les  $m$  composantes du vecteur de mesures sont les valeurs des concentrations au niveau des  $m$  points du maillage (positions des capteurs). Les  $m$  lignes de la matrice de sensibilité contiennent les  $N$  valeurs des fonctions de sensibilité au niveau des  $N$  points du maillage. Dans l'équation (3.4), les éléments de volume ont été intégrés à la matrice d'observation, dont les  $m$  lignes contiennent les  $N$  valeurs des fonctions d'échantillonnage au niveau des  $N$  points du maillage. Dans l'équation (3.5) les éléments de volume ont été intégrés au vecteur source, dont les  $N$  composantes représentent l'intensité de la source au niveau des  $N$  points du maillage. Notons que les équations (3.3) à (3.5) peuvent être utilisées aussi bien pour résoudre des problèmes directs que des problèmes inverses.

### 3.4.1. Approche basée source

Dans cette approche, qui est majoritairement utilisée pour résoudre des problèmes directs (c'est-à-dire pour estimer des mesures lorsque le champ d'émission est connu), la mesure est exprimée à partir de la distribution spatiale et temporelle de la concentration  $C(\mathbf{x}, t)$  en



utilisant l'expression (3.2) (ou (3.4) pour le cas discret). Cette dernière est calculée comme solution d'une équation du type

$$L(C) = s \quad (3.6)$$

Où  $s = s(\mathbf{x}, t)$  décrit la distribution spatiale et temporelle de la source. Dans un cadre Lagrangien,  $L$  est un opérateur qui représente une équation différentielle stochastique de type Langevin (*Flesh et al, 1995,1996 ; Wilson et Sawford, 1996 ; Stohl et Seibert, 1998 ; Seibert, 2001 ; Stohl, 2005 ; Stohl et al, 2009*). Dans un cadre Eulérien,  $L$  est un opérateur linéaire qui représente l'équation de transport diffusion, associée aux conditions aux limites appropriées (*Pudykiewicz, 1998*).  $C(\mathbf{x}, t)$  est la solution de :

$$\frac{\partial C}{\partial t} + \mathbf{U} \cdot \nabla C - \nabla \cdot (\mathbf{K} \nabla C) = s \quad (3.7)$$

Où  $\mathbf{U}(u, v, w)$  est le vecteur vitesse du vent et  $\nabla \cdot (\mathbf{K} \nabla C)$  est la diffusion turbulente avec  $\mathbf{K} = (K_x, K_y, K_z)$  la diffusivité turbulente.

Cette équation peut être résolue de différentes façons. On trouve des modèles analytiques (principalement Gaussiens) qui fournissent directement les concentrations en tous points du domaine. L'équation peut aussi être discrétisée puis résolue, conjointement aux équations donnant l'évolution de  $\mathbf{U}$ . Nous reviendrons sur ce point au paragraphe 3.5.3.a.

Notons que, si pour une mesure réalisée au point  $\mathbf{u}_i$  et au temps  $t_k$ , la fonction d'échantillonnage est de la forme  $\pi_i = \delta(\mathbf{x} - \mathbf{u}_i) \delta(t - t_k)$  où  $\delta$  est la fonction de Dirac, nous pouvons écrire directement  $\mu_i(t_k) = C(\mathbf{u}_i, t_k) + \epsilon_i$ .

### 3.4.2. Approche basée récepteurs

Il est souvent plus pratique (notamment pour les problèmes inverses, voir la figure 3.3) d'opter pour une approche orientée « récepteurs » (*Robertson, 2004*). La mesure est alors exprimée à l'aide de fonctions de sensibilité, notées  $a_i(\mathbf{x}, t)$  calculées pour chaque capteur. Ces fonctions expriment, pour un vent donné, la sensibilité de la mesure (valeur que mesurera un capteur) vis à vis d'une émission unitaire localisée au point  $\mathbf{x}$  et à l'instant  $t$  (si une source est placée au point  $\mathbf{x}$  et émettait avec un débit égal à l'unité). Ces fonctions peuvent être construites soit en résolvant  $N$  fois un problème direct (c'est-à-dire en plaçant successivement une source « unité » en chaque point du domaine, ce qui est évidemment extrêmement coûteux en temps de calcul) soit en résolvant une équation du type

$$L^*(a_i) = \pi_i \quad (3.8)$$

$L^*$  est un opérateur linéaire qui, dans un cadre Eulérien, représente une équation similaire (aux signes près) à l'équation de transport-diffusion (3.7) (Hourdin et al, 1999 ; Wotawa et al, 2003 ; Issartel et Baverel, 2002, 2003 ; Hourdin et al, 2005 ; Hourdin et Talagrand, 2006) mais en utilisant la fonction d'échantillonnage comme source :

$$-\frac{\partial a_i}{\partial t} - \mathbf{U} \cdot \nabla a_i - \nabla \cdot (\mathbf{K} \nabla a_i) = \pi_i \quad (3.9)$$

Nous voyons, par analogie avec l'équation (3.7), que  $a_i(\mathbf{x}, t)$  correspond à une pseudo-concentration dispersée « à rebours », c'est-à-dire calculée en remontant le vent et le temps à partir de la position du  $i^{\text{ème}}$  détecteur. Dans le cas stationnaire cette fonction est donc aussi souvent nommée « rétropanache ». Elle dépend du champ de vent considéré, des caractéristiques du site mais elle est totalement indépendante de la source.

Une mesure est exprimée à partir d'une fonction de sensibilité en utilisant l'expression (3.3) (ou (3.5) pour le cas discret). Dans le cas d'une émission ponctuelle constante d'intensité  $q_s$ , localisée au point  $\mathbf{x}_0$ , nous avons directement

$$\mu_i = q_s a_i(\mathbf{x}_0) \quad (3.10)$$

La figure 3.2 montre ces fonctions pour un réseau de six capteurs et pour une direction et vitesse de vent données.

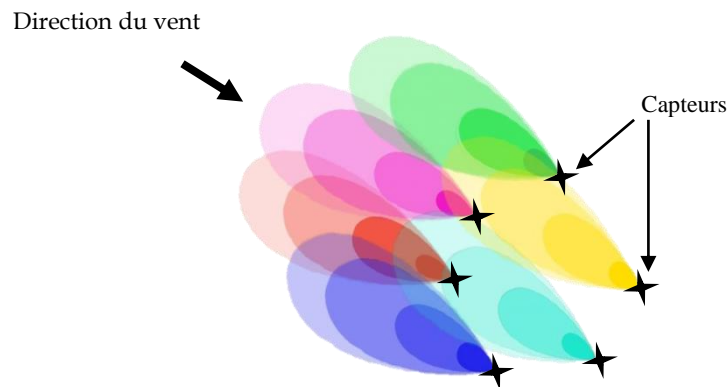


Figure 3.2. Rétropanaches, issus de six capteurs, pour une météo donnée

Notons que cette notion a été initialement proposée, dans le cadre Eulérien, par (Pudykiewicz, 1998) puis par (Hourdin et al, 1999) dans le cadre de l'inversion de concentration de traceur. L'idée consistait à faire « tourner » un modèle de transport-diffusion « à l'envers » dans l'espace et le temps afin de générer des rétropanaches. (Keats et al. 2007), signalent qu'un rétropanache définit une « zone d'influence », c'est-à-dire une zone dans laquelle une source potentielle peut contribuer à générer une mesure non nulle au niveau du capteur considéré.

On voit donc qu'un intérêt des fonctions  $a_i(\mathbf{x}, t)$  est qu'elles permettent de définir une notion similaire à la notion de couverture (évoquée au chapitre 2) : si on note  $C_{min}$  le seuil à partir duquel les capteurs détectent des concentrations, la zone pour laquelle  $a_i(\mathbf{x}, t) > C_{min}$  est la zone dans laquelle des sources potentielles, dont le débit est au minimum égale à une unité de masse par seconde ( $1 \text{ u/s}$ ), produiront des mesures non nulles au niveau du  $i^{\text{ème}}$  capteur. De même, la zone plus étendue pour laquelle  $a_i(\mathbf{x}, t) > C_{min}/10$ , correspond à la zone dans laquelle des sources potentielles, dont le débit est au minimum égale à  $10 \text{ u/s}$ , génèrent des mesures non nulles. On peut donc faire l'analogie entre ces zones et les zones de couverture définies au chapitre 2.

- **Notion d'adjoint**

De façon générale plusieurs auteurs ([Pudykiewicz, 1998](#) ; [Bocquet, 2005-a](#) ; [Hourdin et Talagrand, 2006](#) ; [Issartel et al, 2007](#)) ont montré que les équations de la mesure (3.3) et (3.4) correspondent à un produit scalaire défini comme :

$$(\zeta, Y) = \int_{\Omega} \zeta Y d\mathbf{x}dt \quad (3.11)$$

Ainsi, pour l'approche basée source, la mesure correspond au produit scalaire de la concentration et de la fonction d'échantillonnage :

$$\mu_i = \int_{\Omega} C(\mathbf{x}, t) \pi_i(\mathbf{x}, t) d\mathbf{x}dt = (C, \pi_i) \quad (3.12)$$

Pour l'approche basée récepteurs, la mesure correspond au produit scalaire du taux d'émission et de la fonction de sensibilité :

$$\mu_i = \int_{\Omega} s(\mathbf{x}, t) a_i(\mathbf{x}, t) d\mathbf{x}dt = (s, a_i) \quad (3.13)$$

Nous avons donc, en utilisant les équations (3.6) et (3.8) :

$$\mu_i = (C, \pi_i) = (s, a_i) \text{ soit } (C, L^*(a_i)) = (L(C), a_i) \quad (3.14)$$

Dans ce cas, on dit que les opérateurs  $L$  et  $L^*$  sont adjoints. L'équation (3.9) est donc également appelée « équation adjointe » et les fonctions  $a_i(\mathbf{x}, t)$  « fonctions adjointes ». Ce concept a été décrit en détail par ([Marchuk, 1995](#)) et ([Hourdin et Talagrand, 2006](#)). Notons que le terme de diffusion apparaît sous la même forme dans  $L$  et  $L^*$  (il est dit auto-adjoint). Ceci permet d'interpréter les fonctions  $a_i(\mathbf{x}, t)$  comme étant des champs de pseudo-concentrations dispersés à rebours dans l'espace et dans le temps, et ce à partir des récepteurs. Cette interprétation n'est plus possible en présence de phénomènes non linéaires, non auto-adjoints, qui peuvent intervenir dans le processus de dispersion ([Chow et al, 2008](#)).

### 3.4.3. Avantages de l'approche basée récepteurs dans le cadre des problèmes inverses

Résoudre un problème inverse à l'aide d'une approche paramétrique consiste, généralement, à minimiser une fonction coût notée  $J$  qui exprime l'écart entre le vecteur de mesures observées  $\boldsymbol{\mu}$  et modélisées  $\boldsymbol{\mu}_{mod}$ . Comme illustré sur la figure 3.3, évaluer  $J$  avec une approche basée source consiste à :

- Résoudre l'équation de transport diffusion (3.7), dont la solution dépend des paramètres recherchés.
- Résoudre l'équation d'observation (3.4) afin d'évaluer les concentrations au niveau des capteurs.
- Comparer l'observation calculée à celle mesurée afin de trouver une estimation optimale des paramètres inconnus (*Thomson et al, 2007*).

Par contre, avec une approche basée récepteurs, évaluer  $J$  consiste à :

- Résoudre l'équation adjointe (autant de fois qu'il y a de récepteurs).
- Générer des mesures, en fonction des paramètres recherchés.
- Comparer les mesures prédites à celles observées afin de trouver une estimation optimale des paramètres inconnus.

Les premières étapes de ces approches (les résolutions respectives de l'équation de transport diffusion et des équations adjointes) sont les plus coûteuses en temps de calcul. Or, dans le cas d'une approche basée source, cette étape fait partie de la boucle d'optimisation : à chaque nouvelle évaluation de la source, il faut résoudre l'équation de transport diffusion.

Dans le cas d'une approche basée récepteurs, l'équation adjointe est résolue autant de fois qu'il y a de récepteurs. Mais cette étape est totalement indépendante de la boucle d'optimisation : les termes sources que l'on cherche à identifier n'interviennent pas dans ces équations. Ainsi, les fonctions adjointes calculées peuvent être directement utilisées pour rechercher les paramètres optimaux : la boucle sera donc moins coûteuse en temps de calcul. Par contre, il convient de préciser que les algorithmes utilisés pour rechercher l'estimation

optimale peuvent être similaires dans les deux approches. Les méthodes utilisées ne se distinguent les unes des autres que par la forme de la fonctionnelle à minimiser.

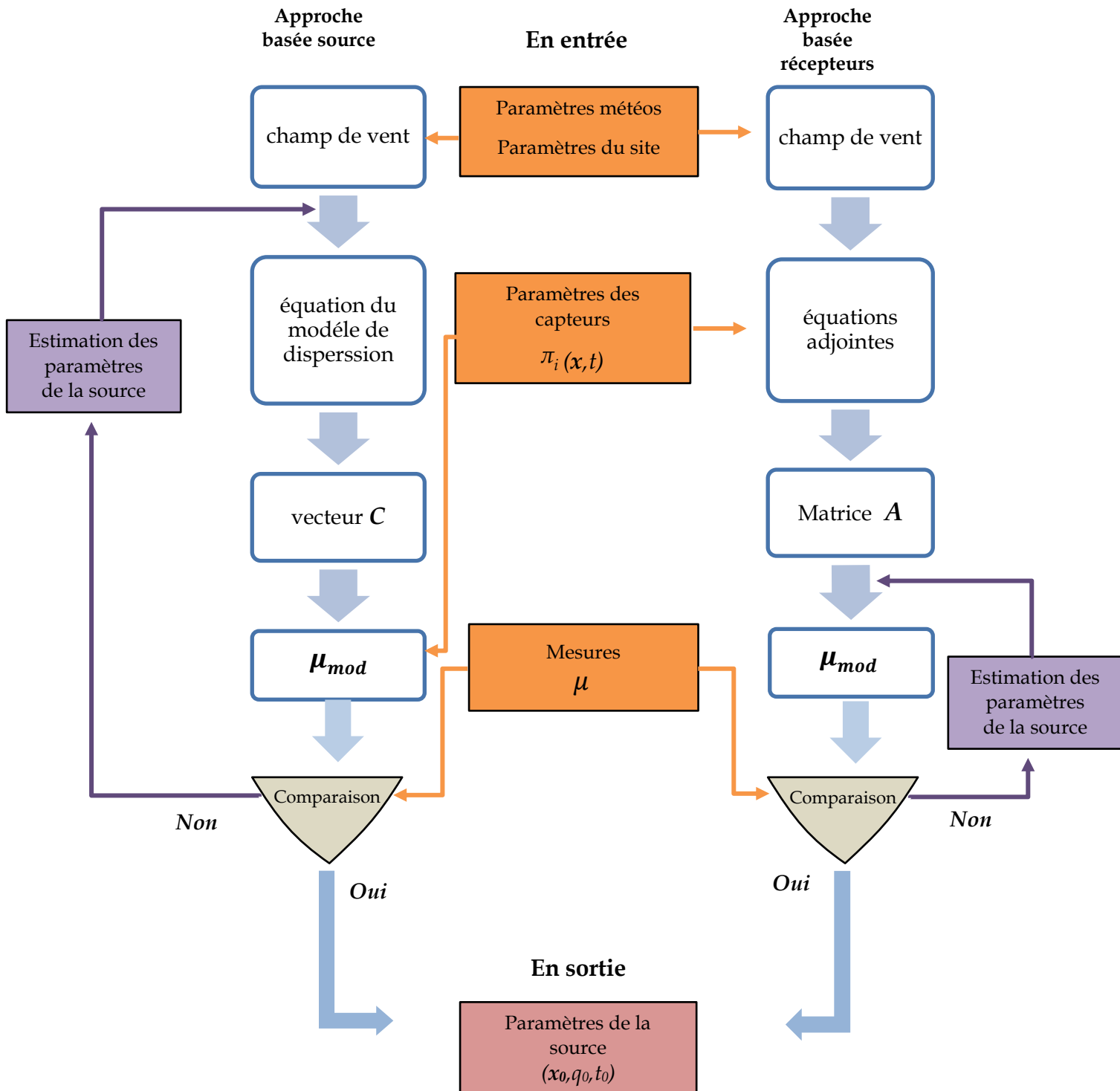


Figure 3.3. Approches basée source et basée récepteurs pour la résolution d'un problème inverse à l'aide d'une méthode paramétrique

### 3.5. CHOIX DE LA METHODE D'INVERSION DE DONNEES

#### 3.5.1. Aperçu des techniques existantes

Dans plusieurs travaux on trouve des inventaires assez exhaustifs des méthodes existantes. Par exemple, (*Winiarek, 2014*) classe les méthodes en deux catégories : les méthodes paramétriques et les méthodes non paramétriques. Cette distinction a déjà été évoquée plus haut. (*Liu et Zhai, 2007*) opposent les méthodes basées source (forward) aux méthodes basées récepteurs (backward). Cette distinction a également été évoquée au paragraphe précédent. Dans ce dernier article, mais aussi dans ceux de (*Rao, 2007 ; Yee, 2012*), les méthodes déterministes (deterministic optimization methods) sont opposées aux méthodes stochastiques (stochastic Bayesian methods).

Les méthodes probabilistes considèrent que les paramètres de la source et les erreurs de mesure sont des variables aléatoires et leur associent, à priori, des densités de probabilités. On peut citer les méthodes basées : sur la théorie de l'estimation bayésienne (*Delle Monache et al, 2008 ; Yee et al, 2008*), sur les algorithmes de Monte Carlo par chaînes de Markov (MCMC) (*Gilks et al, 1996 ; Gamerman et Lopes ; 2006*) et sur divers algorithmes d'échantillonnage stochastique (*Keats et al, 2009*). Les principaux inconvénients de ces méthodes viennent du fait qu'elles nécessitent l'usage de probabilités à priori, souvent difficiles à évaluer. D'autre part, elles nécessitent des calculs complexes, donc coûteux en temps. Ceux-ci les rendent difficiles à mettre en œuvre dans un cadre opérationnel.

Les méthodes déterministes, souvent inspirées de méthodes d'optimisation, emploient des fonctions coûts (ou objectifs) pour évaluer l'écart entre les concentrations observées et les concentrations modélisées. Elles se basent sur un processus itératif pour minimiser cet écart. Mathématiquement, la majorité d'entre elles minimisent une fonction notée  $J(s)$ . Nous avons alors:

$$s_o = arg[\min(J(s))] \quad (3.15)$$

Parmi les méthodes déterministes nous trouvons, par exemple, celles basées sur la technique des moindres carrés (*Michalis et Panayiotou, 2005 ; Davoine et Bocquet 2007 ; Sharan et al 2012-a ; Issartel et al, 2012 ; Singh et Rani, 2014, Ben Salem, 2014*).

### 3.5.2. La méthode d'assimilation de données renormalisées

Parmi les méthodes inverses, la méthode d'assimilation de données renormalisées (ou méthode de renormalisation) est la technique que nous avons retenue pour cette étude. *La renormalisation est une approche « orientée récepteurs » qui permet d'estimer aussi bien des sources ponctuelles que des sources distribuées. Ceci permet de l'utiliser dans des situations opérationnelles variées, lorsque les sources sont ponctuelles (accident industriel, attaque terroriste, etc.) ou réparties (pollution automobile et industrielle, etc.). Un autre avantage de cette méthode est qu'elle permet, une fois les concentrations observées aux niveaux des capteurs, d'obtenir une estimation très rapide de la source. Enfin, elle introduit le concept de visibilité qui permet d'évaluer globalement la qualité d'un réseau (ce qui est très utile pour définir un critère d'optimalité). Ce sont ces caractéristiques qui justifient son utilisation ici.*

La méthode d'assimilation de données renormalisées a été définie par (Issartel, 2005). C'est une approche déterministe orientée récepteur qui exploite la linéarité entre les mesures et la source. Cette technique s'appuie sur l'utilisation des rétropanaches et définit une fonction de pondération  $\varphi$ , appelée fonction de visibilité, afin de ne pas surinterpréter les informations issues des zones situées à proximité des capteurs. Initialement, la technique est non paramétrique et vise à reconstruire une source continue  $s(x, t)$  (Issartel et al, 2007). Dans ce cas, les équations sont données sous la forme intégrale. Après discrétisation du domaine, on obtient un système d'équations linéaires que l'on peut écrire sous forme matricielle. (Turbelin et al, 2014) ont décrit en détail les formulations mathématiques continues et discrètes de la méthode d'assimilation de données renormalisées.

#### 3.5.2.1. Historique de la méthode

Au départ, la méthode de renormalisation a été définie par (Issartel, 2005) pour estimer des sources distribuées. Puis (Sharan et al. 2009, 2012-b ; Issartel et al, 2012) ont montré que cette technique était particulièrement efficace pour l'estimation de sources ponctuelles. Il est utile de rappeler que dans de nombreux problèmes réels, on peut souvent considérer, à priori, que la source est ponctuelle. Par exemple, les fuites à l'origine de contaminations industrielles, accidentelles, peuvent être considérées, à l'échelle de la zone d'intérêt, comme étant des phénomènes localisés ponctuels. Il en est de même dans des situations d'agressions de type terroriste où la source n'occupe qu'une partie infinitésimale de l'espace qui l'entoure. Il est donc justifié de considérer ce type de sources comme étant ponctuelles. Notons que, plus

récemment, la méthode a été étendue pour estimer plusieurs sources ponctuelles émettant simultanément (*Singh et al, 2013*).

### 3.5.2.2. Formalisme mathématique

Nous avons déjà vu, équation (3.13), que pour un rejet atmosphérique, la relation « source-récepteur » est supposée linéaire. Sans perte de généralité, dans le cas d'un rejet continu, elle peut être exprimée sous la forme :

$$\mu_i = \int_{\Omega} s(\mathbf{x}) a_i(\mathbf{x}) d\mathbf{x} = (s, a_i) \quad (3.16)$$

Le terme source, dans le domaine  $\Omega$ , est représenté par la fonction  $s(\mathbf{x})$  qui représente le flux d'émission en chaque point du domaine. Les  $i = 1 \dots m$  mesures sont regroupées dans le vecteur de mesures  $\boldsymbol{\mu}$  de dimension  $m$ .  $a_i(\mathbf{x})$  est la fonction adjointe, ou rétropanache, qui évalue la sensibilité du  $i$ -ème capteur en chaque point de  $\Omega$ . Comme nous l'avons déjà indiqué, pour calculer ces rétropanaches (qui correspondent à des pseudo-concentrations), on peut résoudre des problèmes de dispersion « à rebours ». Pour chaque capteur, on calcule le panache issu d'une source fictive, de débit unitaire, qui émet depuis l'emplacement du capteur et qui se disperse dans le sens inverse du sens réel du vent.

Résoudre le problème inverse consiste à déterminer la fonction  $s(\mathbf{x})$  à partir de l'équation (3.16) (équation intégrale semi discrète du premier type). Ce type de problème est dit mal posé (on ne dispose que d'un nombre fini de mesures) et possède une infinité de solution. Pour obtenir une solution unique, le terme source  $s(\mathbf{x})$  peut-être décomposé en deux composantes :

La première,  $s_{//}(\mathbf{x})$ , parallèle aux fonctions adjointes, c.à.d  $s_{//}(\mathbf{x}) = \sum_{i=1}^m \lambda_i a_i(\mathbf{x})$ , et la seconde  $s_{\perp}(\mathbf{x})$  perpendiculaire aux fonctions adjointes, ainsi le produit scalaire  $(s_{\perp}(\mathbf{x}), a_i(\mathbf{x})) = 0$ , (*Sharan et al, 2009*) :

$$s = s_{//} + s_{\perp} \quad (3.17)$$

La composante perpendiculaire ne contribue pas aux mesures, contrairement à la composante parallèle. En regroupant les paramètres  $\lambda_i$  dans un vecteur  $\boldsymbol{\lambda}$ , et les fonctions  $a_i(\mathbf{x})$  dans un vecteur  $\mathbf{a}(\mathbf{x}) \in \mathbb{R}^m$ , on a  $s_{//}(\mathbf{x}) = \boldsymbol{\lambda}^T \mathbf{a}(\mathbf{x})$ , où «  $T$  » désigne la transposition, et  $\boldsymbol{\mu} = \int_{\Omega} s_{//}(\mathbf{x}) \mathbf{a}(\mathbf{x}) d\mathbf{x}$ . On a donc  $\boldsymbol{\mu} = \int_{\Omega} \mathbf{a}(\mathbf{x}) \mathbf{a}^T(\mathbf{x}) d\mathbf{x} \boldsymbol{\lambda}$  soit

$$\boldsymbol{\lambda} = \mathbf{H}^{-1} \boldsymbol{\mu} \text{ où } H_{ij} = \int_{\Omega} a_i(\mathbf{x}) a_j(\mathbf{x}) d\mathbf{x} = (a_i(\mathbf{x}), a_j(\mathbf{x})) \quad (3.18)$$



La matrice ( $m \times m$ ) symétrique  $\mathbf{H}$  est la matrice de Gram. Une solution au problème inverse est donc

$$s_{//}(\mathbf{x}) = \boldsymbol{\mu}^T \mathbf{H}^{-1} \mathbf{a}(\mathbf{x}) \text{ ou } s_{//}(\mathbf{x}) = \mathbf{a}(\mathbf{x})^T \mathbf{H}^{-1} \boldsymbol{\mu} \quad (3.19)$$

Mais (Issartel, 2011) a montré que les sources estimées à partir de l'équation (3.19) étaient concentrées dans les régions situées au voisinage des capteurs. Ces artefacts ont été expliqués par les valeurs singulières que prennent les rétropanaches au niveau des capteurs. Pour éliminer ces artefacts, le processus de renormalisation a été défini. Ce dernier fait appel à une fonction de pondération notée  $f(\mathbf{x})$  dont le but est de réduire les singularités aux niveaux des capteurs. Après pondération l'équation (3.16) devient

$$\mu_i = \int_{\Omega} s(\mathbf{x}) \frac{a_i(\mathbf{x})}{f(\mathbf{x})} f(\mathbf{x}) d\mathbf{x} \quad (3.20)$$

L'estimation du terme source est obtenue par

$$s_{//f} = \boldsymbol{\mu}^T \mathbf{H}_f^{-1} \frac{\mathbf{a}(\mathbf{x})}{f(\mathbf{x})} \quad (3.21)$$

Où  $\mathbf{H}_f$  est la matrice de Gram pondérée dont les termes sont donnés par

$$H_{f_{ij}} = \int_{\Omega} \frac{a_i(\mathbf{x})}{f(\mathbf{x})} \frac{a_j(\mathbf{x})}{f(\mathbf{x})} f(\mathbf{x}) d\mathbf{x} \quad (3.22)$$

Le choix de la fonction de pondération  $f(\mathbf{x})$  est une étape importante. En effet, cette fonction doit être la plus neutre possible afin de ne pas apporter dans le problème trop d'information subjective. (Issartel et al., 2005), puis (Busch, 2015), a démontré que la fonction de pondération optimale est celle qui correspond à un critère d'entropie minimale : c'est la fonction pour laquelle le déterminant de la matrice de Gram pondérée  $\det(\mathbf{H}_f)$  est minimum. La fonction optimale qui vérifie ce critère, est nommée fonction de visibilité et est notée  $\varphi$ . Elle vérifie trois conditions :

$$(a) \varphi(\mathbf{x}) > 0 \quad (b) \int_{\Omega} \varphi(\mathbf{x}) d\mathbf{x} = m \quad (c) \frac{\mathbf{a}^T(\mathbf{x})}{\varphi(\mathbf{x})} \mathbf{H}_{\varphi}^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})} = 1 \quad (3.23)$$

Ces trois conditions, sont les conditions dites de renormalisation (Issartel et al, 2007). Il faut que : (a) la fonction de pondération optimale soit positive, (b) la somme des poids de  $\varphi$  soit égale au nombre de mesures  $m$  et enfin que  $\varphi$  vérifie l'équation (c).

La fonction  $\varphi(\mathbf{x})$  est aussi appelée « fonction de visibilité » car elle fournit des informations pertinentes sur les capacités globales de détection du réseau. En effet, les zones où les valeurs de  $\varphi(\mathbf{x})$  sont élevées sont les zones situées près des capteurs et couvertes par plusieurs d'entre eux. A l'inverse, les zones où les valeurs de  $\varphi(\mathbf{x})$  sont faibles sont les zones situées

très en amont du réseau (dans les zones non couvertes,  $\varphi(\mathbf{x}) = 0$ ). L'estimation optimale de la source est donc donnée par

$$s_{//\varphi}(\mathbf{x}) = \boldsymbol{\mu}^T \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})} \quad (3.24)$$

Une des propriétés intéressante est que, si la source est ponctuelle, le maximum de l'estimation correspond au point d'émission. En effet, dans le cas d'une émission ponctuelle d'intensité  $q_s$  depuis un point  $\mathbf{x}_0$ , on a  $s(\mathbf{x}) = q\delta(\mathbf{x} - \mathbf{x}_0)$ . Les mesures associées sont alors  $\boldsymbol{\mu} = q\mathbf{a}(\mathbf{x}_0)$ . En substituant cette expression dans l'équation (3.24), on obtient :

$$s_{//\varphi}(\mathbf{x}) = q\varphi(\mathbf{x}_0) \frac{\mathbf{a}^T(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})} \quad (3.25)$$

Or d'après l'inégalité de Cauchy-Schwartz

$$\left| \frac{\mathbf{a}^T(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})} \right| \leq \sqrt{\frac{\mathbf{a}^T(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x}_s)}{\varphi(\mathbf{x}_s)}} \sqrt{\frac{\mathbf{a}^T(\mathbf{x})}{\varphi(\mathbf{x})} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})}} \quad (3.26)$$

Avec la condition de renormalisation donnée par l'équation (3.23) on a  $\left| \frac{\mathbf{a}^T(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x})}{\varphi(\mathbf{x})} \right| \leq 1$  et  $\frac{\mathbf{a}^T(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} \mathbf{H}_\varphi^{-1} \frac{\mathbf{a}(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)} = 1$ . La fonction  $s_{//\varphi}(\mathbf{x})$  a donc pour maximum  $s_{//\varphi}(\mathbf{x}_0) = q\varphi(\mathbf{x}_0)$ . On remarque également que  $q_0 = \frac{s_{//\varphi}(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)}$ .

### 3.5.2.3. Formulation discrète du problème

Comme nous l'avons vu au paragraphe 3.4.1, le problème peut être discrétisé afin d'être résolu numériquement. Pour cela :

- Le domaine  $\Omega$  est discrétisé en  $j = 1..N$  points.
- Les valeurs des fonctions adjointes sont calculées pour les points  $\mathbf{x}_j = (x_j, y_j, z_j)$ .
- On construit ainsi les  $m$  lignes de la matrice de sensibilité  $\mathbf{A}$ .

Le système linéaire à résoudre est alors :

$$\boldsymbol{\mu} = \mathbf{A}\mathbf{s} \quad (3.27)$$

Le vecteur des mesures  $\boldsymbol{\mu} \in \mathbb{R}^m$  est donné par

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \quad (3.28)$$

$\mathbf{s} \in \mathbb{R}^N$  est le vecteur source tel que :

$$\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad (3.29)$$

Et  $\mathbf{A} \in \mathbb{R}^{m \times N}$  est la matrice de sensibilité telle que :

$$A_{ij} = a_i(\mathbf{x}_j) \quad (3.30)$$

Cette matrice peut s'écrire également sous la forme  $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_j \cdots \mathbf{a}_N]$  où les  $j = 1 \dots N$  vecteurs  $\mathbf{a}_j \in \mathbb{R}^m$  sont les vecteurs des fonctions adjointes  $\mathbf{a}(\mathbf{x}_j)$ . Le problème consiste donc à trouver les  $N$  composantes du vecteur  $\mathbf{s}$  connaissant les  $m$  composantes du vecteur  $\boldsymbol{\mu}$ . Comme nous avons  $m < N$ , le système est mal posé (système sous-déterminé) et possède une infinité de solutions.

### 3.5.2.3.1. Technique de renormalisation discrète

Soit  $\mathbf{W} \in \mathbb{R}^{N \times N}$  la matrice de pondération diagonale telle que  $\mathbf{W} = \text{diag}(w_1 \dots w_N)$ , (Turbelin et al., 2014) ont montré que la solution discrète renormalisée est donnée par

$$\mathbf{s}_{//} = \mathbf{G}\boldsymbol{\mu} \quad (3.31)$$

Où  $\mathbf{G} \in \mathbb{R}^{N \times m}$  est un opérateur inverse qui correspond à :

$$\mathbf{G} = \mathbf{W}^{-1} \mathbf{A}^T (\mathbf{A} \mathbf{W}^{-1} \mathbf{A}^T)^{-1}$$

Les composantes de la matrice  $\mathbf{W}$  vérifient les conditions de renormalisation discrètes :

$$(a) w_j > 0 \quad (b) \sum_{j=1}^N w_j = m \quad (c) w_j^2 = \mathbf{a}_j^T \mathbf{H}_w^{-1} \mathbf{a}_j \quad (3.32)$$

On remarque que la matrice de Gram est donnée par la relation :  $\mathbf{H}_w = \mathbf{A} \mathbf{W}^{-1} \mathbf{A}^T$  et que l'on a  $\mathbf{H}_\varphi = \mathbf{H}_w d\mathbf{x}^2$ . On peut également montrer que la matrice  $\mathbf{G} = \mathbf{W}^{-1} \mathbf{A}^T \mathbf{H}_w^{-1}$  correspond au pseudo-inverse, pondéré par la matrice  $\mathbf{W}$ , de la matrice  $\mathbf{A}$ .

### 3.5.2.3.2. Proposition d'un algorithme pour le calcul de la matrice de pondération

Les premiers articles dans lesquels la méthode de renormalisation a été décrite, traitaient des problèmes continus. (Issartel et al, 2007) proposaient d'utiliser un algorithme itératif pour calculer la fonction  $\varphi(\mathbf{x})$  :

$$f_{P+1}(\mathbf{x}) = f_p \sqrt{(f_p^{-1} \mathbf{a}_p^T(\mathbf{x}) \mathbf{H}_{f_p}^{-1} f_p^{-1} \mathbf{a}_p(\mathbf{x}))} \quad (3.33)$$

Dans le cadre de ce travail, nous avons adapté cet algorithme pour calculer les composantes de la matrice  $\mathbf{W}$ . Pour cela, nous avons remarqué que, d'après la condition (3.32), les poids optimaux  $w_j$  sont les termes diagonaux de la matrice de résolution, de dimension  $(N \times N)$ , telle que  $\mathbf{R} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{H}_w^{-1}\mathbf{A}$ . De même, lorsque la condition (3.32) est respectée, les termes diagonaux de la matrice  $(N \times N)$   $\mathbf{R}_w = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{H}_w^{-1}\mathbf{A}\mathbf{W}^{-1}$  sont égaux à l'unité. De plus, on peut montrer que, d'après les propriétés de la matrice  $\mathbf{R}$ , on a  $trace(\mathbf{R}) = \sum_{j=1}^N w_j = m$ . A l'aide de ces définitions, la matrice de pondération optimale peut être obtenue en réalisant une série d'opérations matricielles et/ou vectorielles (voir figure 3.4)

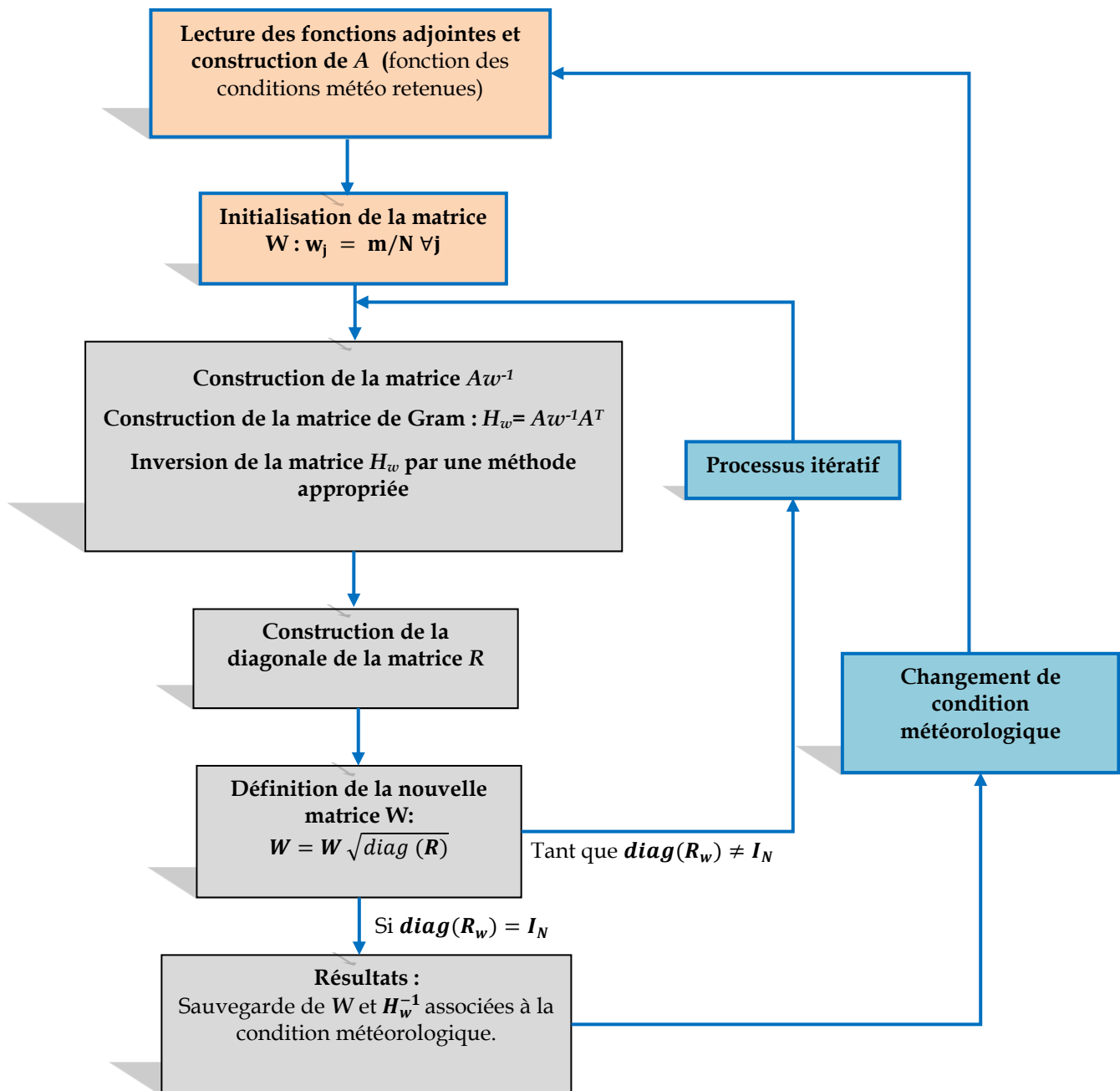


Figure 3.4. Algorithme pour le calcul de la fonction de pondération optimale  $\mathbf{W}$

De façon plus formelle on peut présenter cet algorithme de la façon suivante :

**Input :**  $A \in \mathbb{R}^{m \times N}$

1.  $N = \text{columns}[A]$
2.  $W = m/N I_N$
3. While  $p_{\min} \Rightarrow 0.99999$  et  $p_{\max} \leq 1.00001$
4.  $H_w = AW^{-1}A^T$
5. for  $j = 1$  to  $N$
6.  $a_j = A(j)$
7.  $p(j) = a_j^T H_w^{-1} a_j w_j^{-2}$
8. End for
9.  $p_{\min} = \min(p)$
10.  $p_{\max} = \max(p)$
11.  $W = W^*(\text{diag}[p])^{-1/2}$
12. End while

**Output :**  $W$

L'algorithme est composé de deux parties distinctes dont la complexité dépend :

- Du nombre de capteurs (c'est-à-dire de  $m$ ).
- Du nombre de mailles (c'est à dire de  $N$ ).

La première partie concerne la construction et l'inversion de la matrice de Gram pondérée.

Cette partie comprend :

- $N$  divisions scalaires de vecteurs de dimension  $m$  (les colonnes de la matrice  $A$  par les termes de la diagonale de  $W^{-1}$ ). La complexité est d'ordre  $(m.N)$ .
- La transposition d'une matrice  $[m \times N]$ . La complexité est d'ordre  $(m.N)$ .
- La multiplication de matrices  $[m \times N]$  &  $[N \times m]$ . La complexité est d'ordre  $(m^2N)$ .
- L'inversion d'une matrice  $[m \times m]$ . La complexité est d'ordre  $(m^3)$

Le temps de calcul  $t_1$  pour cette partie sera donc principalement fonction de  $m^2N$  (voir figure 3.5).

La seconde partie concerne la construction des termes diagonaux de la matrice  $\mathbf{R}$ , la vérification de la condition de convergence et la construction de la nouvelle matrice de pondération. Cette partie comprend :

- La multiplication de matrices  $[m \times N]$  &  $[N \times m]$ . La complexité est d'ordre  $(m^2N)$ .
- $N$  multiplications scalaires de vecteur de dimension  $m$ . La complexité est d'ordre  $(m.N)$ .
- $N$  additions de  $m$  scalaires. La complexité est d'ordre  $(m.N)$ .
- Recherches de min/max sur des vecteurs de dimension  $N$ . La complexité est d'ordre  $(N)$ .
- La multiplication scalaire de vecteurs de dimension  $N$ . La complexité est d'ordre  $(N)$ .

Le temps de calcul  $t_2$  pour cette partie sera donc principalement fonction de  $m^2N$  (voir figure 3.5). Ceci est vérifié sur nos tests lorsque  $m$  et  $N$  varient.

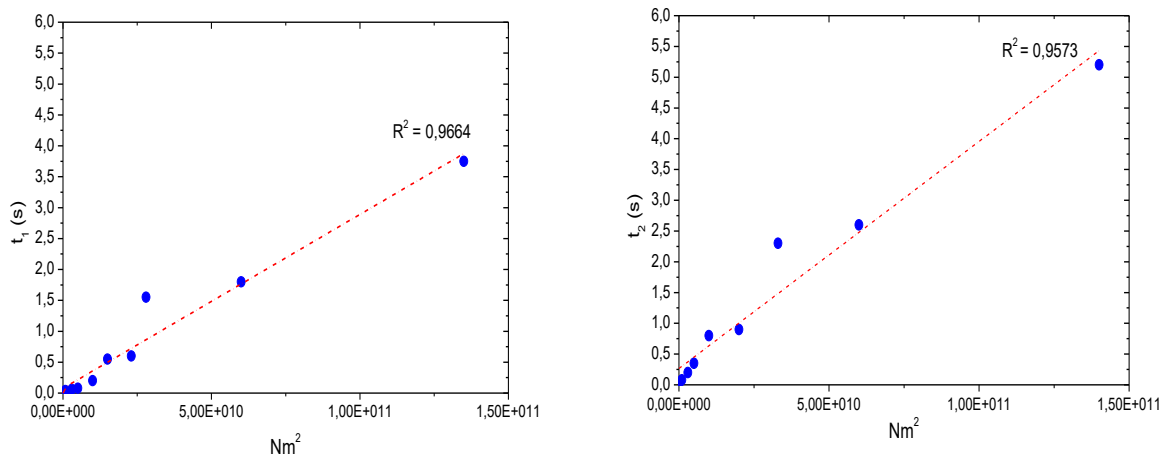


Figure 3.5. temps de calcul de la fonction de pondération optimale

L'algorithme répète ces opérations à chaque nouvelle itération en considérant la matrice de pondération obtenue à l'itération précédente. Notons que la complexité des méthodes disponibles pour inverser la matrice  $\mathbf{H}_w$  (méthode de Gauss-Jordan, décomposition LU, etc.) est d'ordre  $m^3$  et qu'il est rare de travailler avec plus de 100 capteurs. Pour des matrices  $[100 \times 100]$  les temps d'inversion sont de l'ordre du dixième de seconde. Il n'est donc pas primordial actuellement de chercher à optimiser ce temps. L'algorithme a été programmé sous Matlab (voir annexe 3).

Pour optimiser cette fonction :

- Les termes diagonaux de la matrice  $\mathbf{W}$  sont initialisés avec la valeur  $m/N$ , afin de vérifier la condition (3.32).
- Les termes diagonaux de la matrice  $\mathbf{W}$  ont été stockés dans un vecteur  $\mathbf{w}$  de dimension  $N$ .
- La fonction `bsxfun(@rdivide,  $\mathbf{A}^t$ ,  $\mathbf{w}$ )` (*bsx* signifiant “binary singleton expansion”) a été utilisée pour renormaliser chaque ligne de  $\mathbf{A}$ . Cette fonction réalise des divisions scalaires, éléments-par-éléments, des lignes de  $\mathbf{A}^t$  par  $\mathbf{w}$ .
- La fonction d’inversion classique de Matlab, `inv()`, basée sur la décomposition LU de la librairie LAPACK a été utilisée pour l’inversion de  $\mathbf{H}_w$ .
- Seuls les termes diagonaux de la matrice  $\mathbf{R}_w$  sont calculés et stockés dans un vecteur colonne. Chaque terme correspond au produit scalaire des lignes de la matrice  $\mathbf{W}^{-1}\mathbf{A}^t\mathbf{H}_w^{-1}$  et des colonnes de  $\mathbf{A}\mathbf{W}^{-1}$ .

Le nombre d’itérations nécessaire pour obtenir  $\mathbf{W}$  dépend de la précision choisie (`pre`). On obtient :

Précision	1.00E-03	1.00E-05	1.00E-10
Nb Itérations	6	13	29

Tableau 3.1. Nombre d’itérations pour obtenir  $\mathbf{W}$

Avec un ordinateur Intel® Core™ i7-4790 CPU@3,60 GHz 64 bits, 16 Go Ram, pour  $N=1,5E+06$  et  $m=50$  et  $pre=1,0E-5$  le temps de calcul est d’environ 50s. C’est-à-dire 13 itérations avec, à chaque itération, 1,5s pour calculer et inverser  $\mathbf{H}_w$ , puis 2,5 s pour estimer les poids.

Notons que pour un site donné, ces calculs peuvent être réalisés pour différentes conditions météorologique. Les résultats peuvent être stockés dans une base de données et réutilisés en conditions opérationnelles.

### 3.5.2.3.3. Estimation de la source

La construction du vecteur source par la formule (3.31) et la recherche de son maximum dépendent du nombre de détecteurs et du nombre de points du maillage pour lesquelles, la visibilité est non nulle, c’est-à-dire  $\varphi(\mathbf{x}) > \omega$ , avec  $\omega$  suffisamment petit. Le calcul de  $\mathbf{s}_{//}$  comprend :

- $m$  divisions scalaires termes à termes de vecteurs de dimension  $N$ . La complexité est d’ordre  $(m.N)$ .

- La transposition d'une matrice  $[m \times N]$ . La complexité est d'ordre  $(m.N)$ .
- La multiplication de matrices  $[N \times m]$  &  $[m \times 1]$ . La complexité est d'ordre  $(N.m)$ .
- La recherche du max sur un vecteur de dimension  $N$ . La complexité est d'ordre  $(N)$ .

La complexité de la construction de la fonction source est donc d'ordre  $(m.N)$ . L'intensité de la source, une fois la localisation estimée, est obtenue par une simple division  $q_0 = \frac{s_{//}(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)}$ . Le schéma de fonctionnement du processus global est présenté sur la figure 3.6.

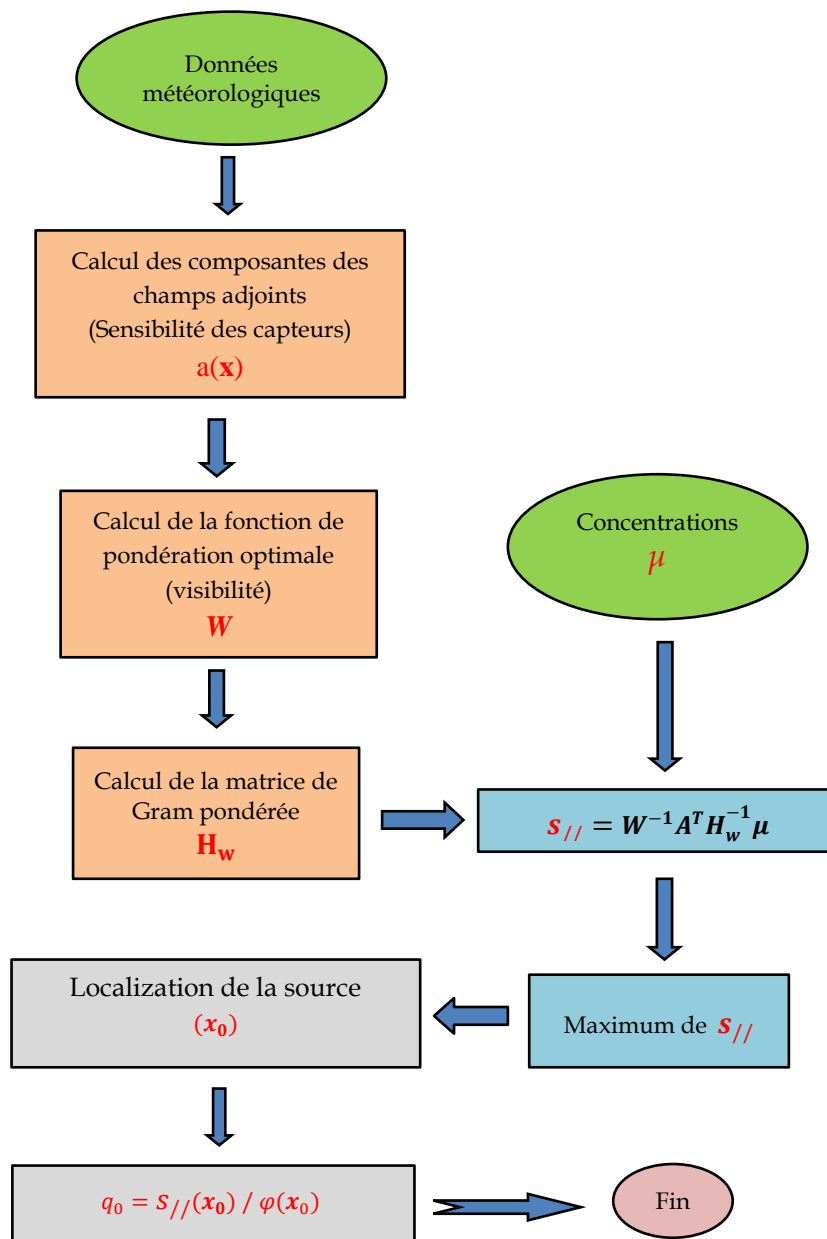


Figure 3.6. Principe de fonctionnement de l'algorithme d'inversion



### 3.5.3. Modélisation des rétropanaches

Le rétropanache issu du capteur  $i$  est obtenu en résolvant l'équation (3.9). Si on compare cette équation au modèle direct (3.7), on remarque que le champ de vent est inversé et que les champs qui décrivent la turbulence sont conservés. L'idée couramment utilisée consiste donc à utiliser les outils de résolution du modèle direct et d'inverser le champ de vent, et le temps, pour calculer les rétropanaches (*Hourdin et al, 2006*). Dans le cadre d'une approche Eulérienne, les méthodes de résolution peuvent être classés en deux grandes familles : les approches analytiques (souvent gaussiennes) et les approches basées sur la mécanique des fluides numérique (en anglais, CFD, Computational Fluid Dynamics). Précisons que lorsque des modèles CFD sont utilisés, certains schémas numériques ne sont pas forcément compatibles avec l'idée qu'une simple inversion des signes suffise à obtenir le champ adjoint (*Hourdin et Talagrand, 2006 ; Winiarek, 2014 ; DISCARD, 2015*). Dans ce cas, il est nécessaire de réaliser une phase préliminaire de « vérification de dualité » en simulant la dispersion de traceurs à l'aide d'une approche basée source et d'une approche basée récepteurs et de vérifier la proximité des résultats.

#### 3.5.3.1. Modélisation analytique et « CFD »

##### a. Approche analytique

L'approche analytique consiste à rechercher des solutions de l'équation (3.7) en faisant un certain nombre d'hypothèses simplificatrices. Dans le cas de rejets instantanés ou continus, et si le terrain est plat, on peut supposer que le champ de concentration en aval de la source a une distribution Gaussienne dans les directions verticales et horizontales. Les valeurs des concentrations dépendent alors de la distance vis-à-vis de la source, de son débit, de sa hauteur, de la stabilité atmosphérique, de la vitesse moyenne et de la direction du vent (*Hanna, 1982*). Dans le cas d'un rejet continu, à partir d'une source située en  $(x_s, y_s, z_s)$ , le champ de concentration moyen à l'intérieur du panache est donnée par

$$\bar{c}(x, y, z) = \frac{q_s}{2\pi\bar{U}\sigma_y\sigma_z} \exp\left(-\frac{(y - y_s)^2}{2\sigma_y^2}\right) \left[ \exp\left(-\frac{(z - z_s)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z + z_s)^2}{2\sigma_z^2}\right) \right] \quad (3.34)$$

Où :

$\bar{c}$  : concentration ,

$\bar{U}$  : vitesse moyenne du vent supposée uniforme,

$q_s$  : débit de la source,

$\sigma_y, \sigma_z$  : écarts-types de la dispersion.

Les écarts-types  $\sigma_y$  et  $\sigma_z$ , dits paramètres de dispersion, dépendent des caractéristiques de l'écoulement, en particulier de l'état de turbulence donc de la stabilité atmosphérique et de la rugosité du site. Ces modèles sont simples à mettre en œuvre, le processus de calcul est rapide ce qui les rends adaptés à des situations opérationnelles. Cependant, ils dépendent fortement du choix des paramètres de dispersion (on trouve dans la littérature de nombreux schémas de paramétrisation) et ne sont valables ni en présence d'obstacles ni lorsque le vent est très faible. De plus, ils sont limités aux rejets de débits constants et pour lesquels l'écoulement est uniforme dans l'espace et constant dans le temps.

Dans le cas d'un rejet instantané et lorsque le champ de vitesse et la turbulence varient avec le temps, on peut utiliser des modèles de bouffées. Pour ces modèles le rejet peut être discrétisé en une succession de rejets instantanés (*Demaël, 2007*). Le pas de temps entre deux rejets est défini de telle sorte que soient reproduits correctement la cinétique d'évolution du rejet et des conditions météorologiques et que le panache ainsi décrit ait un aspect continu (*Perkins et al. 2002*).

Pour un rejet instantané localisé près d'une surface réfléchissante, la formulation à bouffées est de la forme

$$\bar{C}(x, y, z) = \frac{M}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_z} \exp \left[ -\frac{(x - x_c)^2}{2\sigma_x^2} - \frac{(y - y_c)^2}{2\sigma_y^2} \right] \left[ \exp\left(-\frac{(z - z_c)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z + z_c)^2}{2\sigma_z^2}\right) \right] \quad (3.35)$$

Où :

$M$  : quantité totale de produit libéré,

$x_c$ ,  $y_c$  et  $z_c$  : coordonnées du centre de la bouffée qui évoluent en fonction du champ de vent,

$\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$  : écarts-types de la dispersion, par rapport à la localisation de la bouffée à un instant donné.

Notons qu'un rejet continu peut être assimilé à une succession de rejets instantanés. Ces modèles permettent donc aussi de traiter des rejets continus dont le débit varie dans le temps. Leur qualité dépend également de la paramétrisation utilisée. Le choix des bons écarts-types est donc capital.

Les paramètres de dispersion sont généralement exprimés sous forme de fonctions dépendant de la distance à la source ou de la durée du transport. La forme de ces fonctions peut dépendre de la stabilité de l'atmosphère, de la rugosité du site, des caractéristiques de la turbulence, etc. Les modèles présentés ci-dessous ne sont pas exhaustifs et leur application ne se limite pas uniquement aux modèles gaussiens en régime stationnaire (d'autres modèles sont regroupés en annexe I). Les écarts types de (Briggs, 1973) sont le fruit d'une combinaison de trois formulations (Pasquill, 1961), Tennessee Valley Authority (TVA) et Brookhaven National Laboratory) et sont issus de l'analyse de campagnes expérimentales menées dans les années 1950 et 1960 sur des zones rurales, plates à faible rugosité (Barad, 1958 ; Carpenter et al, 1971 ; Singer et al, 1966). Ces travaux ont abouti à des formulations qui sont encore aujourd'hui largement utilisées. Les écarts types sont exprimés sous la forme

$$\sigma_{y,z}(x) = ax(1 + bx)^c \quad (3.36)$$

Où :

$x$  : distance par rapport à la source,

$a$ ,  $b$ , et  $c$  : coefficients dépendants des conditions de stabilité (classification de Pasquill-Gifford).

Le tableau 3.2 donne la formulation des écarts types de dispersion en milieu rural (valable jusqu'à 10 km) :

Classe de Pasquill	$\sigma_y$ (m)	$\sigma_z$ (m)
A	$0.22x(1+0.0001x)^{-1/2}$	$0.22x$
B	$0.16x(1+0.0001x)^{-1/2}$	$0.12x$
C	$0.11x(1+0.0001x)^{-1/2}$	$0.08x(1+0.0002x)^{-1/2}$
D	$0.08x(1+0.0001x)^{-1/2}$	$0.06x(1+0.0015x)^{-1/2}$
E	$0.06x(1+0.0001x)^{-1/2}$	$0.03x(1+0.0003x)^{-1}$
F	$0.04x(1+0.0002x)^{-1/2}$	$0.016x(1+0.0003x)^{-1}$

Tableau 3.2. Formulations de Briggs des écarts types de dispersion en milieu rural (Briggs-rural) pour des distances comprises entre 100 et 10 000 m

Une autre approche consiste à relier les paramètres de dispersion aux caractéristiques de la turbulence du vent. Dans ce cas, les écarts types sont exprimés en fonction du temps de transport  $t = x/U$ , sous la forme

$$\sigma_{y,z}(t) = \sigma_{v,w} t f_{y,z}(t/T_{y,z}) \quad (3.37)$$

$\sigma_v$  et  $\sigma_w$  sont les écarts types des fluctuations latérales et verticales du vent.  $f_y$  et  $f_z$  sont des fonctions sans dimension qui dépendent des échelles de temps Lagrangiennes de la turbulence,  $T_y$  et  $T_z$ . Les caractéristiques de la turbulence ( $\sigma_v$ ,  $\sigma_w$ ,  $T_y$  et  $T_z$ ) peuvent être

obtenues à partir de mesures à haute fréquence de la vitesse du vent où déduites à partir de formulations empiriques. On trouve dans la littérature de nombreuses expressions pour les fonctions  $f_{y,z}$ . Après comparaison, (*Gryning et al, 1987*) ont conclu que celles proposées par (*Draxler, 1976*) étaient les plus efficaces :

$$f_y = \frac{1}{1 + \left(\frac{t}{2T_y}\right)^{0.5}} \quad \text{et} \quad f_z = \frac{1}{1 + \left(\frac{t}{2T_z}\right)^{0.5}} \quad (3.38)$$

avec  $T_z = 30s$  pour les conditions stables,  $T_z = 300s$  pour les conditions instables et  $T_y = 200s$  pour les sources situées au niveau du sol, quelle que soit la stabilité.

### b. Approche CFD

L'approche CFD est une approche eulérienne, tridimensionnelle où la résolution de l'équation de transport-diffusion (équation 3.39) est réalisée à partir d'un champ de vent calculé en résolvant les équations de Navier-Stokes (*Bird et al, 2002*) :

$$\frac{\partial C}{\partial t} - \mathbf{U} \cdot \nabla C - \nabla \cdot (\mathbf{K} \nabla C) = s \quad (3.39)$$

où  $\mathbf{U}(\mathbf{x}, t)$  est le champ de vent moyen.  $\nabla$  représente l'opérateur gradient,  $\mathbf{K}$  la matrice des coefficients de diffusivité turbulente et enfin  $s = s(\mathbf{x}, t)$  la fonction source. Nous rappelons que dans l'approche orientée source<sup>1</sup>,  $s(\mathbf{x}, t)$  est supposée connue en espace  $\mathbf{x} = (x, y, z)$  et en temps  $t$  et le champ de concentration  $C(\mathbf{x}, t)$  du gaz traceur dans l'air est déterminé par la résolution de cette équation (3.39).

Pour résoudre ces équations (Navier-Stokes et transport-diffusion), le vent est assimilé à un écoulement de fluide incompressible turbulent. Afin d'obtenir le champ de concentration, il faut donc :

---

<sup>1</sup> Il est à noter que cette approche CFD peut également être utilisée dans le cadre d'une approche basée récepteurs. La différence avec l'approche source vient du fait que le calcul CFD concerne uniquement les champs adjoints. Pour ce faire, les calculs pour les champs adjoints peuvent être réalisés par simple inversion du champ de vent dans le domaine de calcul, et par intégration à rebours de l'équation de transport. Nous rappelons qu'il y a deux manières d'obtenir la solution adjointe (*Winiarek, 2014*) : (1) La première consiste à prendre l'adjoint du problème continu et ensuite le discrétiser. Il s'agit en effet du modèle direct pour lequel est effectué l'inversion du temps, et du vent. Les autres processus sont inchangés. Cette approche fait une approximation de l'ordre de l'erreur de représentativité qui découle de la discrétisation. (2) Dans la deuxième approche on prend l'adjoint du problème déjà discrétisé. Nous avons choisi dans le cadre de cette thèse cette deuxième approche puisque les résultats des simulations de la dispersion atmosphérique proviennent de l'exécution d'un modèle direct, fluidyn-PANACHE (*Kumar et al, 2015*). L'exigence de compatibilité impose le choix d'adjoint du modèle utilisé qui est discret.

1. Définir la géométrie du domaine d'étude et obtenir le maillage correspondant (Pré-processing).
2. Définir les conditions initiales et aux frontières (Pré-processing).
3. Résoudre les équations de Navier-Stokes et l'équation de transport diffusion discrétisées à l'aide d'un solveur adapté. Il existe plusieurs méthodes de discrétisation (différence finis, volume finis et éléments finis) et plusieurs approches pour prendre en compte la turbulence (*Ferziger and Peric, 2002*).
4. Afficher et visualiser les résultats (Post-processing).

L'approche CFD-3D peut être utilisée pour prédire la dispersion atmosphérique dans des configurations complexes intégrant des bâtiments et du relief. Pour simuler des champs de vent turbulents, plusieurs approches sont envisageables (*Wynngaard, 2010*) :

- Les approches dites moyennées (RANS) « Reynolds Average Navier-Stokes » ont pour but la résolution des équations de Navier-Stokes moyennées (au sens de Reynolds). Ce sont, généralement, les plus utilisées. L'idée est de résoudre les équations régissant l'écoulement moyen et de modéliser la turbulence (à l'aide de modèles de fermeture, c'est-à-dire d'équations spécifiques).
- Les approches dites simulation des grandes échelles, (LES) « Large Eddy Simulation », sont fondées sur la résolution explicite des équations de Navier-Stokes y compris pour les plus grandes échelles de la turbulence. La partie de la turbulence non résolue explicitement (petites structures) est modélisée statistiquement (modèles sous-maille).
- Les approches dites Simulation Numérique Directe (DNS) « Direct Numerical Simulation », consistent à résoudre explicitement les équations de Navier-Stokes pour toutes les échelles de la turbulence. Cette méthode est la plus proche de la réalité vu qu'elle prend en considération toutes les échelles possibles, ce qui la rend extrêmement coûteuse en temps de calcul. Le coût est d'autant plus important que le domaine de calcul est étendu et complexe.

L'approche CFD (en particulier l'approche moyennée, RANS) est de plus en plus utilisée dans le cadre de la dispersion atmosphérique. Elle permet de prendre en compte la complexité des champs de vitesse sur des terrains complexes et en présence d'obstacles. Cependant, il n'est pas encore envisageable de l'utiliser en temps réel. De plus, son utilisation demande une bonne maîtrise des outils de simulation et l'utilisation de bonnes pratiques (*INERIS, 2015*).

### 3.6. EVALUATION DE LA METHODE D'INVERSION

L'objectif de cette partie est d'évaluer la méthode inverse et le nouvel algorithme proposé au paragraphe 3.5.2.3.2. Pour cela nous avons utilisé des résultats expérimentaux obtenus en 2011 dans le cadre du projet DYCE (DYnamic deployment planning for monitoring of ChEmical leaks using an ad-hoc sensor network). Le principal objectif de ce projet était d'étudier l'influence de la configuration d'un réseau sur ses performances, notamment en matière d'inversion de données. Aussi, en plus de l'évaluation de la méthode, ces résultats nous permettront d'examiner l'influence de la position des capteurs sur l'estimation des paramètres de la source.

#### 3.6.1. Description du projet DYCE

Dans le cadre du projet Dyce ([Lepley et al. 2011](#)) des mesures expérimentales ont été réalisées au sein de la soufflerie atmosphérique de l'ENFLO « Environmental Flow Research Centre » à l'Université de Surrey au Royaume-Uni. Ces mesures ont été utilisées par ([Rudd et al. 2012](#)) pour étudier la problématique de l'identification des paramètres de la source. A partir d'un nombre de mesures limité (4 capteurs), un modèle inverse inspiré de ([Rodgers, 2000](#)) a permis d'estimer l'intensité et la position de la source. Ce modèle minimise une fonction coût qui représente la norme de l'écart entre les vecteurs de mesures observé et modélisé, pondérée par la variance des mesures. L'approche est « basée source » et la méthode d'optimisation est celle de Gauss-Newton. Les mesures expérimentales ont été collectées au cours d'une série d'expériences de dispersion réalisée en soufflerie. Les expériences consistaient à disperser un mélange de propane/air sur des périodes de 15 min. La proportion de propane dans le mélange était de 1,5 % pour un débit volumique total de  $7,5 \cdot 10^{-7} \text{ m}^3/\text{s}$ . Dans l'ensemble des expériences, la source, ponctuelle, était placée en  $x_s = 0$  et  $y_s = 0$  avec  $x$  parallèle à la direction du vent et  $y$  dans le sens de la largeur comme schématisé sur la figure 3.7.

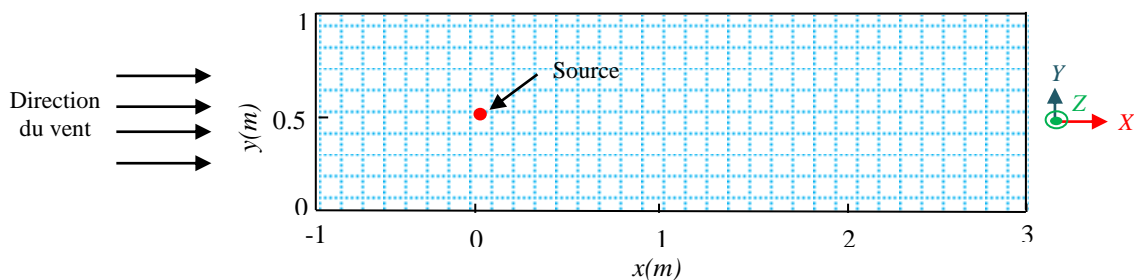


Figure 3.7. Schématisation, dans le plan  $xy$  et à l'échelle de la soufflerie, de la configuration expérimentale ([Rudd et al. 2012](#)). La source d'émission est représentée par le disque rouge.

La hauteur de la source était de  $0.1 H_c$ , avec  $H_c$  la hauteur de la couche limite (la rugosité de surface dans la soufflerie a permis de maintenir une couche limite de  $1 m$ ). La vitesse du vent à la hauteur  $H_c$ , mesurée à l'aide d'un anémomètre à ultrasons; était de  $U_{H_c} = 2,5 m s^{-1}$ .

Pour le problème inverse, 4 détecteurs situés à la même hauteur que la source ont été employés. Il s'agissait de détecteurs à Ionisation de Flamme à Réponse Rapide (FFID, Fast response Flame Ionisation Detectors). Onze différentes configurations de réseau ont été testées (figure 3.8). Nous remarquons que les configurations 1 à 5 possèdent aux moins 3 capteurs alignés perpendiculairement à la direction du vent, tandis que les configurations 6 à 11, sont des configurations de type « grille » (« box array configurations »)

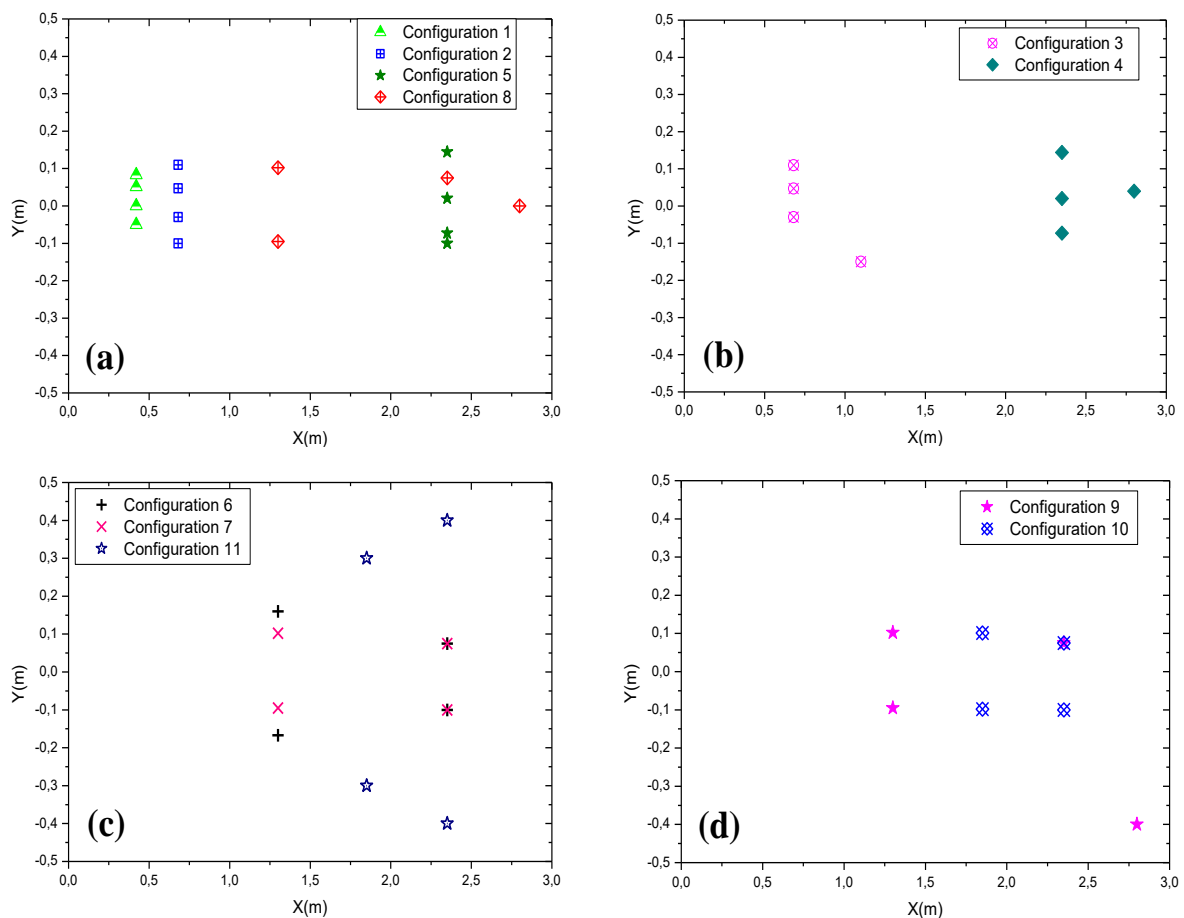


Figure 3.8. Disposition des quatre FFIDs dans la soufflerie atmosphérique (WT scale).

- (a) Configurations 1, 2, 5 et 8 (triangle en vert clair, carreau en bleu, étoile en vert foncé),  
 (b) configurations 3 et 4 (cercle rose barrée et losange Bleu-vert plein), (c) configurations 6, 7 et 11  
 (plus bleu, croix rouge et étoile creuse bleu foncé) et (d) configurations 9 et 10 (étoile rose pleine et losange bleu barré). (d'après *Rudd et al. 2012*)

Afin de minimiser les erreurs de représentativité, (*Rudd et al. 2012*) ont modélisé la dispersion à l'aide d'un modèle de panache gaussien (*Pasquill et Smith, 1983*) spécialement paramétré pour cette expérience. La concentration en tout point situé à la hauteur d'émission ( $H_c$ ) est donnée par :

$$C(x, y, z) = \frac{q_s}{\pi u \sigma_y \sigma_z} \exp\left(-\frac{(y - y_s)^2}{2 \sigma_y^2}\right) \quad (3.40)$$

(Rudd et al. 2012) n'ont pas fourni les écarts types  $\sigma_y$  et  $\sigma_z$  (figure 3.9) donc dans ce qui suit nous proposons de les rétablir.  $\sigma_y$  et  $\sigma_z$  ont été ajustés à partir d'une première série de 178 mesures réalisées à l'aide d'un seul FFID « Single FFID expériences » déplacé dans toute la soufflerie. Pour ces expériences, le même mélange propane-air (1,5% de propane) a été relâché avec un débit  $q_s = 3 \text{ l/min}$ . La hauteur des mesures et de la source était également égale à  $0,1H$ , et la vitesse du vent à  $U_H = 2,5 \text{ ms}^{-1}$ .

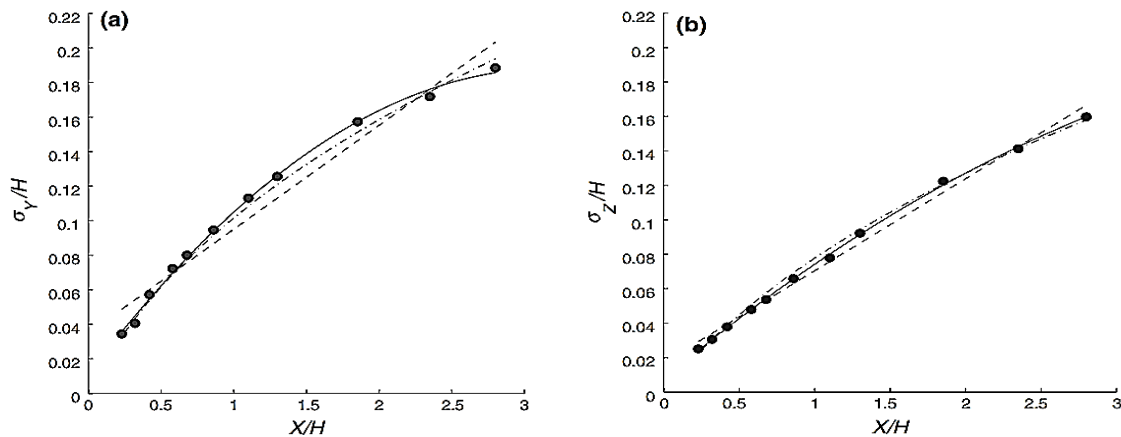


Figure 3.9. Ecart types (a) latéraux  $\sigma_y/H$  et (b) verticaux  $\sigma_z/H$  (Rudd et al, 2012)

### 3.6.2. Vérification du modèle direct

(Rudd et al, 2012) ont montré que les paramètres de dispersion suivent des lois polynomiales du second ordre (figure 3.9). Ils ont aussi montré que le modèle donné par l'équation 3.40 produit les meilleurs résultats pour  $u = 1,3 \text{ ms}^{-1}$  soit  $U_H/u = 1,92$ . Nous avons reproduit ces lois polynomiales (figure 3.10) en se basant sur les données de « Single FFID expériences ».

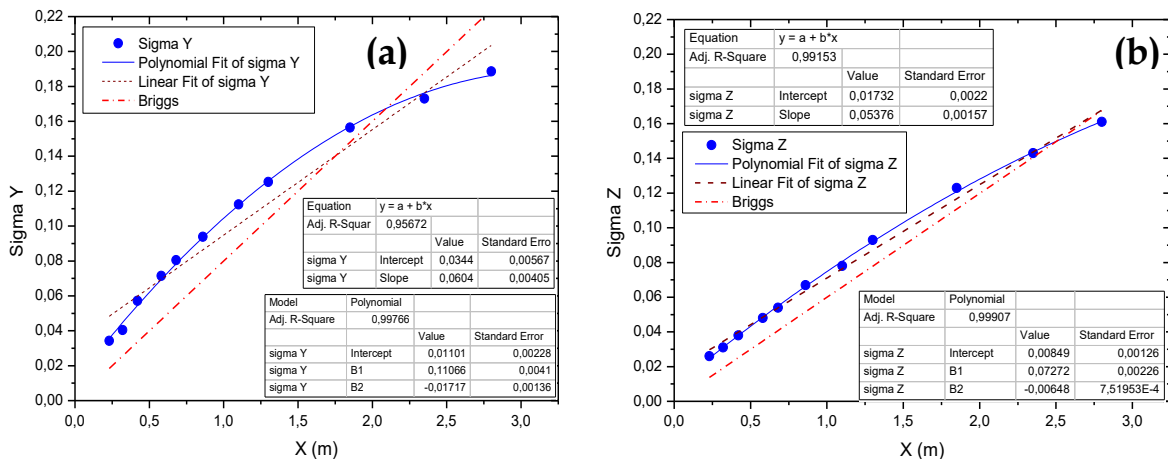


Figure 3.10. Résultats obtenus : Ecart types (a) latéraux  $\sigma_y/H$  et (b) verticaux  $\sigma_z/H$ .



Une fois les écarts types ajustés, nous avons programmé le modèle de dispersion (équation 3.40) et nous l'avons évalué en utilisant les résultats des « Single FFID expériences ». Sur la figure 3.11 nous présentons les résultats sous forme d'un diagramme de dispersion et nous les comparons à ceux obtenus par (Rudd *et al*, 2012). Notons que les concentrations mesurées, sont exprimées en [ppm] et sont ensuite adimensionnées pour obtenir  $C^*$ .

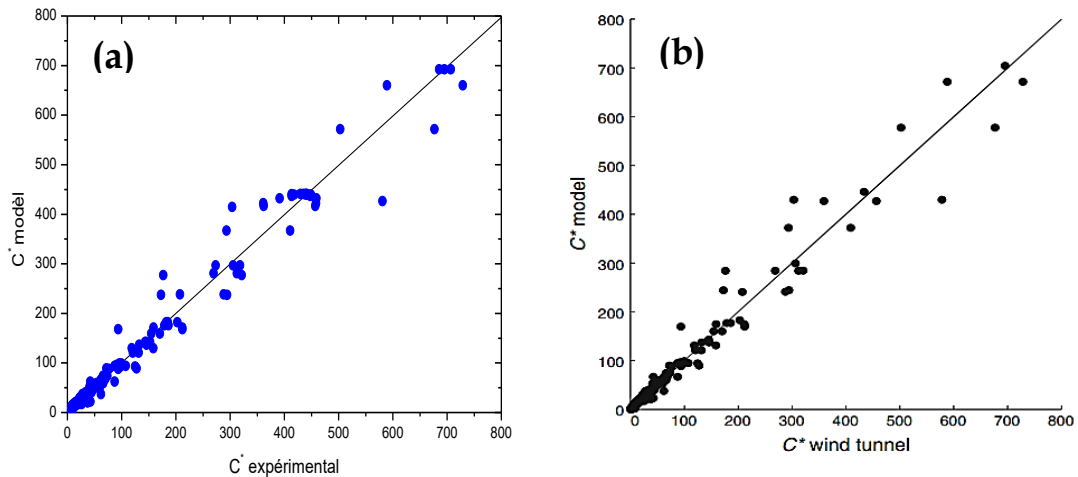


Figure 3.11. Résultats obtenus dans le cadre des « Single FFID expériences » : (a) les résultats de notre étude et (b) ceux de (Rudd *et al*, 2012).

En complément, nous avons calculé les statistiques proposées par (Hanna *et Chang*, 2012) afin d'évaluer les performances de ce modèle (FB : Fractional Bias, NMSE : Normalized Mean Square Error, MG : Geometric Mean bias, VG : Geometric variance, FAC2 : Fraction of predictions within a factor of two of observations, R : Correlation coefficient). Nous trouvons ainsi :

	Valeur	Modèle parfait	Valeurs admises
<b>FB</b>	0,001	0	$ FB  \leq 0,3$
<b>NMSE</b>	0.0394	0	$NMSE \leq 3$
<b>MG</b>	0.9981	1	
<b>VG</b>	1	1	
<b>FAC2</b>	0,9326	1	$FAC2 \geq 0,5$
<b>NAD</b>	0.0051	0	$NAD \leq 0,3$
<b>R</b>	0.9826	1	

Tableau 3.3. Evaluation du modèle directe

Nous voyons que le modèle obtient de très bons résultats et que son utilisation dans le cadre de l'inversion des données minimisera autant que possible les erreurs de représentativité.

### 3.6.3. Mise en œuvre de la méthode de renormalisation

Les calculs pour les 11 configurations de réseau, ont été réalisés à l'échelle de la soufflerie, sur un domaine rectangulaire de dimensions  $L \times l = 4 \times 1 \text{ m}$ . La position de la source est  $x_s = 1 \text{ m}$  et  $y_s = 0,5 \text{ m}$ . Le domaine de calcul a été discrétisé en  $2000 \times 500$  mailles, soit des éléments de longueur  $dx = dy = 0,002 \text{ m}$ .

L'algorithme de calcul est décrit en détail dans le paragraphe 3.5.2.3.2. La première étape consiste à calculer, à l'aide du modèle (3.41) et en inversant le sens de l'écoulement, les composantes de la matrice de sensibilité  $\mathbf{A}$ . La seconde étape consiste à calculer les poids optimaux  $w(\mathbf{x})$  à l'aide de l'algorithme présenté au paragraphe 3.5.2.3.2. Dans, la troisième étape, les données mesurées sur les 15 minutes sont moyennées puis utilisées pour évaluer  $s_{//}(\mathbf{x})$  en tout point du domaine de calcul.  $s_{//}(\mathbf{x})$  est tracée sur une carte bidimensionnelle et la position estimée de la source ( $\mathbf{x}_0$ ), qui se situe au niveau de la valeur maximale de  $s_{//}(\mathbf{x})$  est indiquée. Finalement le débit est estimé par  $\frac{s_{//}(\mathbf{x}_0)}{\varphi(\mathbf{x}_0)}$ .

Les onze configurations (Figure 5) ont été testées. Sur un machine standard, de type Intel® Core™ i7-4790 CPU @3.60 GHz and 16 GB RAM, le temps de calcul a été de  $\sim 15 \text{ s}$  /configuration.

### 3.6.4. Résultats et discussions

Tout d'abord, nous présentons les résultats globaux pour des expériences caractéristiques (configuration #3, pour un réseau « aligné » et configuration #10, pour un réseau « en grille »), les résultats des 9 autres configurations sont regroupées en Annexe II.

Sur les figures 3.12(i) et 3.13(i) sont représentées les fonctions de visibilité des réseaux. Les fortes valeurs sont représentées en blanc et sont localisées au voisinage de chaque capteur. Au fur et à mesure que l'on s'éloigne des capteurs la visibilité diminue (couleur grise). Les zones non couvertes par le réseau sont représentées en noir.

Les cartes représentant  $s_{//\varphi}(\mathbf{x})$  sont affichées sur les figures 3.12(ii) et 3.13(ii). Sur ces dernières, nous avons indiqué la position réelle de la source et son estimation.

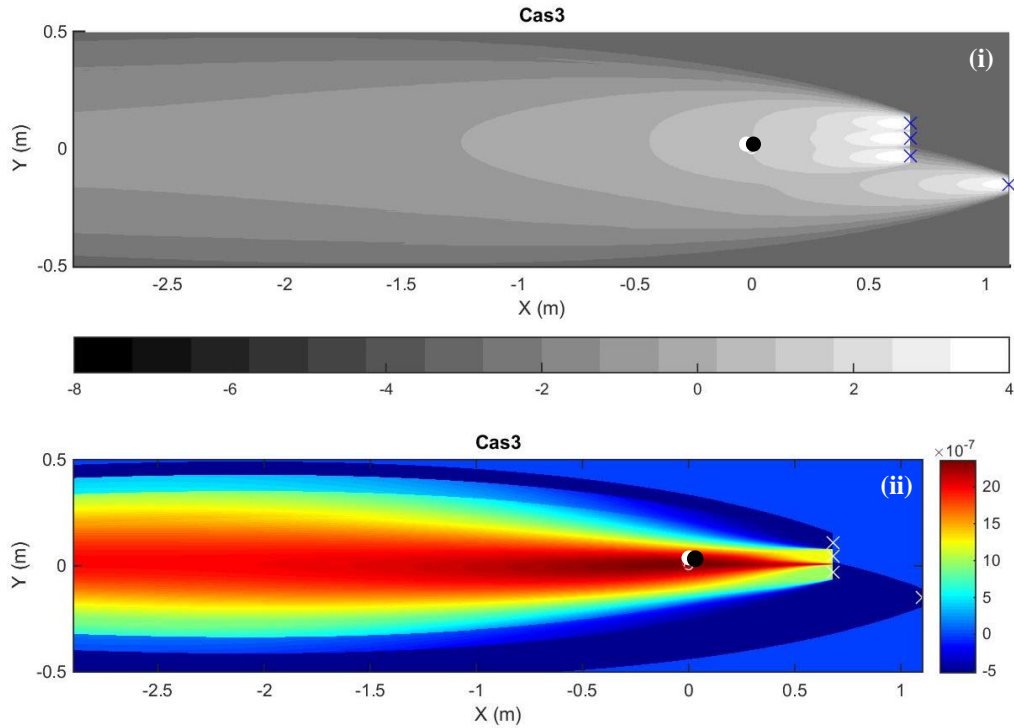


Figure 3.12. Cas 3 : (i) Fonction de visibilité  $\varphi(\mathbf{x})$ , (échelle logarithmique) (ii)  $s_{//\varphi}(\mathbf{x})$   
 La position réelle de la source correspond au disque blanc, la position estimée au disque noir. Les capteurs sont symbolisés par des croix.

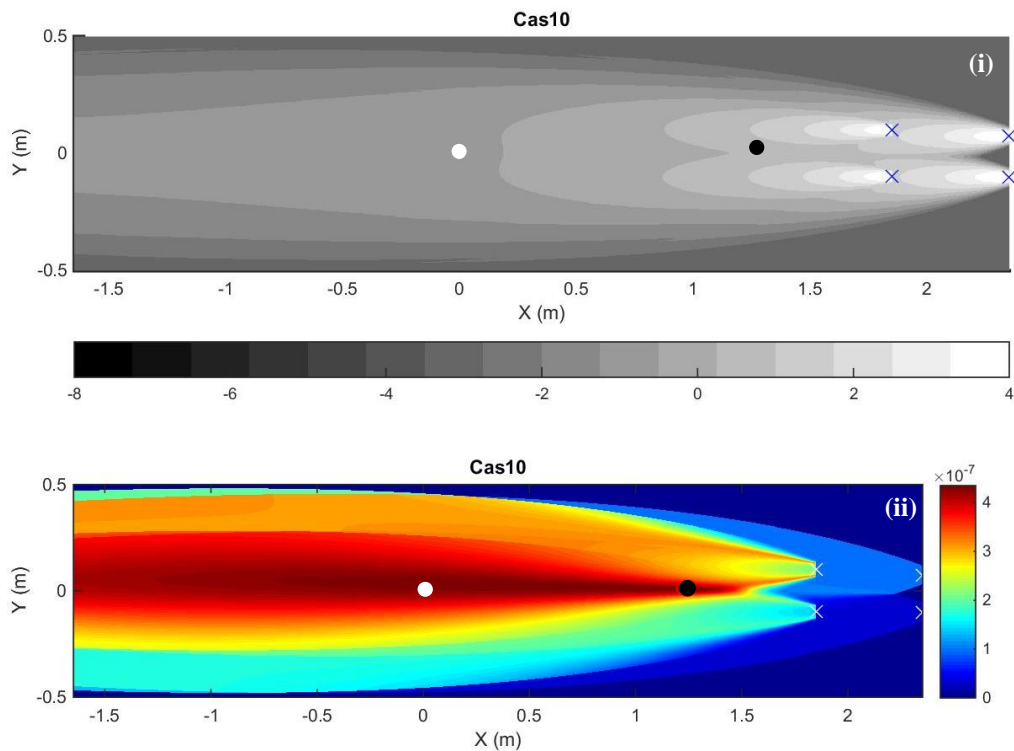


Figure 3.13. Cas 10 : (i) Fonction de visibilité  $\varphi(\mathbf{x})$ , (échelle logarithmique) (ii)  $s_{//\varphi}(\mathbf{x})$   
 La position réelle de la source correspond au disque blanc, la position estimée au disque noir. Les capteurs sont symbolisés par des croix.

Dans le tableau 3.4, nous indiquons : les erreurs de localisation  $\Delta x = x_s - x_0$  et  $\Delta y = y_s - y_0$ , le débit estimé  $q_0$  et l'erreur sur son estimation  $\Delta q = \left| \frac{q_s - q_0}{q_s} \right|$  sachant que le débit de la source est de  $8.9210^{-7}$  kg/s et  $(x_s, y_s) = (0,0)$ .

Configuration	$\Delta x$ (m)	$\Delta y$ (m)	$q_0 \times 10^{-7}$ (Kg/s)	$\Delta q$ (%)
1	- 0,02	0,01	9,59	7,5
2	- 0,02	0,008	9,05	1,4
3	- 0,01	0,01	8,85	0,7
4	- 0,41	0,04	6,09	31
5	- 0,13	0,04	7,22	19
6	- 0,86	0,04	11,25	26
7	0,19	0,02	6,91	22
8	0,8	0,008	4,87	45
9	0,10	0,022	5,78	35
10	1,21	0,008	3,24	63
11	- 0,08	0,032	6,21	30

Tableau 3.4. Résultats pour les onze configurations

Nous remarquons en premier lieu que la méthode fonctionne et que l'algorithme converge dans tous les cas. Les estimations obtenues sont généralement acceptables. Pour toutes les configurations, l'erreur  $\Delta y$  est très faible et varie peu (entre  $0,008$  m et  $0,04$  m) contrairement à l'erreur  $\Delta x$  qui varie de  $0,01$  à  $0,86$  m. En se basant sur la performance des onze configurations (à la fois  $\Delta q$  et  $\Delta x$ ) les configurations peuvent être regroupées en trois classes : (a) configurations 1, 2 et 3 (b) configurations 5, 7, 9 et 11 (c) configurations 4, 6, 8 et 10. Le premier groupe est le plus performant : les paramètres de la source sont estimés avec de meilleures précisions, l'erreur de reconstitution du débit est comprise entre  $0,7$  % et  $7,5$  % et l'erreur de l'estimation de la position,  $|\Delta x|$ , est entre  $0,01$  m et  $0,02$  m. Il est utile de noter que pour ce premier groupe l'estimation de la position de la source est en amont de sa vraie position ( $\Delta x < 0$ ). Pour la seconde classe les erreurs sont plus importantes : l'erreur d'estimation du débit est entre  $19$  % et  $35$  % alors que  $|\Delta x|$  est entre  $0,08$  m et  $0,19$  m. Dans ce groupe, deux configurations (5 et 11) retrouvent une source en amont de la vraie position ( $\Delta x < 0$ ). Le dernier groupe est le moins performant :  $\Delta q$  est compris entre  $26$  % et  $63$  % et  $|\Delta x|$  varie de  $0,41$  m et  $1,21$  m. Pour cette troisième classe la configuration 6 sous-estime la source.

La reconstitution des paramètres de la source par la méthode d'assimilation des données renormalisées, pour les onze configurations du projet DYCE, confirme les résultats obtenus

par (Rudd et al, 2012) : les meilleurs réseaux sont ceux « en lignes » et situés proches de la source (1,2 et 3 à une distance  $< 1 m$  de la vraie source). Ceci peut être formulé en utilisant la notion de visibilité (figures 3.12(i) et 3.13(i)) : les paramètres de la source sont mieux reconstitués lorsqu'elle se trouve dans une zone de forte visibilité.

La conclusion la plus importante pour notre étude est que la configuration spatiale du réseau est un paramètre déterminant dans les problèmes d'estimation de sources. En effet, malgré un modèle parfaitement ajusté, des données peu bruitées et l'absence d'obstacles, on observe de fortes variations dans les estimations. Ceci justifie la mise en œuvre d'une stratégie d'optimisation de réseaux de capteurs.

### 3.7. UTILISATION DE LA METHODE DE RENORMALISATION POUR DEFINIR DES CRITERES D'OPTIMALITE

Dans cette partie, nous analysons plus précisément la méthode de renormalisation afin de pouvoir l'utiliser pour définir des critères d'optimalité et définir les fonctions objectifs spécifiques qui seront minimisées dans un processus de conception de réseaux. Nous définirons ensuite des problématiques pour lesquels nous évaluerons ce processus d'optimisation.

#### 3.7.1. La méthode de renormalisation dans le cadre d'une approche paramétrique

Nous avons vu que, souvent, il était possible de supposer, a priori, que la source est seulement caractérisée par un nombre fini de paramètres. C'est le cas, par exemple, pour une source ponctuelle qui émet de façon continue. Les paramètres à identifier sont sa position  $\mathbf{x}_s$  et son débit  $q_s$ . De façon générale, ce problème peut être résolu en minimisant une fonction coût notée  $J$  qui exprime l'écart entre le vecteur de mesure observé  $\boldsymbol{\mu}$  et modélisé  $\boldsymbol{\mu}_{mod} = q\mathbf{a}(\mathbf{x})$ . Cet écart est représenté par la norme de résidus  $\mathbf{r}(\mathbf{x}) = \boldsymbol{\mu} - \boldsymbol{\mu}_{mod}(\mathbf{x})$ . La théorie de la renormalisation suggère d'utiliser  $\mathbf{H}_w^{-1}$  pour définir cette norme. La fonction de coût est alors décrite par

$$J(q, \mathbf{x}) = \frac{1}{2} \|\boldsymbol{\mu} - q\mathbf{a}(\mathbf{x})\|_{\mathbf{H}_w^{-1}}^2 = \frac{1}{2} (\boldsymbol{\mu} - q\mathbf{a}(\mathbf{x}))^T \mathbf{H}_w^{-1} (\boldsymbol{\mu} - q\mathbf{a}(\mathbf{x})) \quad (3.41)$$

Pour un  $\mathbf{x}$  donné, le débit  $q_0$  qui minimise la fonction coût doit vérifier

$$(a) \left. \frac{\partial J(q, \mathbf{x})}{\partial q} \right|_{q=q_0} = 0 \quad (b) \left. \frac{\partial^2 J(q, \mathbf{x})}{\partial q^2} \right|_{q=q_0} > 0 \quad (3.42)$$

Soit

$$\frac{\partial J(q, \mathbf{x})}{\partial q} = -\mathbf{a}(\mathbf{x})^T \mathbf{H}_w^{-1} (\boldsymbol{\mu} - q \mathbf{a}(\mathbf{x})) \quad (3.43)$$

et

$$\frac{\partial^2 J(q, \mathbf{x})}{\partial q^2} = \mathbf{a}(\mathbf{x})^T \mathbf{H}_w^{-1} \mathbf{a}(\mathbf{x}) \quad (3.44)$$

En tenant compte de la condition de renormalisation donnée par l'équation (3.32),  $w^2(\mathbf{x}) = \mathbf{a}(\mathbf{x})^T \mathbf{H}_w^{-1} \mathbf{a}(\mathbf{x})$ , la dérivée seconde de la fonction coût devient

$$\frac{\partial^2 J(q, \mathbf{x})}{\partial q^2} = w^2(\mathbf{x}) \quad (3.45)$$

La condition (b) est donc bien vérifiée. Toujours en utilisant les conditions de renormalisation, nous déterminons finalement  $q_0$  à l'aide de (3.42.a) et (3.44)

$$\left. \frac{\partial J(q, \mathbf{x})}{\partial q} \right|_{q=q_0} = 0 \Rightarrow q_0(\mathbf{x}) = \frac{\mathbf{a}(\mathbf{x})^T \mathbf{H}_w^{-1} \boldsymbol{\mu}}{w^2(\mathbf{x})} \quad (3.46)$$

On remarque que le terme de droite correspond à  $\frac{s_{//}(\mathbf{x})}{w(\mathbf{x})}$ . En remplaçant  $q$  par cette expression dans (3.42), la fonction coût devient

$$J(\mathbf{x}) = \frac{1}{2} \left\| \boldsymbol{\mu} - \frac{s_{//}(\mathbf{x})}{w(\mathbf{x})} \mathbf{a}(\mathbf{x}) \right\|_{\mathbf{H}_w^{-1}}^2 = \frac{1}{2} (\boldsymbol{\mu}^T \mathbf{H}_w^{-1} \boldsymbol{\mu} - s_{//}^2(\mathbf{x})) \quad (3.47)$$

Résoudre le problème paramétrique consiste donc à minimiser la fonction coût  $J(\mathbf{x})$ . Puisque  $\boldsymbol{\mu}^T \mathbf{H}_w^{-1} \boldsymbol{\mu} = \|\boldsymbol{\mu}\|_{\mathbf{H}_w^{-1}}^2 > 0$ , cela revient à maximiser la quantité  $s_{//}(\mathbf{x}) = \frac{\mathbf{a}(\mathbf{x})^T \mathbf{H}_w^{-1} \boldsymbol{\mu}}{w(\mathbf{x})}$ . La position de la source est donc

$$\mathbf{x}_o = \arg \max_{\mathbf{x} \in \Omega} (s_{//}(\mathbf{x})) \text{ ou } \mathbf{x}_o = \arg \min_{\mathbf{x} \in \Omega} (J(\mathbf{x})) \quad (3.48)$$

Notons qu'en divisant chaque terme de l'équation (3.48) par  $\|\boldsymbol{\mu}\|_{\mathbf{H}_w^{-1}}^2$  on obtient

$$E_{\mathbf{H}_w^{-1}}(\mathbf{x}) = \frac{2J(\mathbf{x})}{\|\boldsymbol{\mu}\|_{\mathbf{H}_w^{-1}}^2} = 1 - \frac{s_{//}^2(\mathbf{x})}{\|\boldsymbol{\mu}\|_{\mathbf{H}_w^{-1}}^2} \quad (3.49)$$

Cette nouvelle fonction coût  $E_{\mathbf{H}_w^{-1}}$  (résidus normalisés) présente l'avantage d'avoir des valeurs qui varient entre 0 et 1, ce qui la rend plus « lisible » et facilement « interprétable » que  $s_{//\varphi}(\mathbf{x})$ , voir figure 3.14.

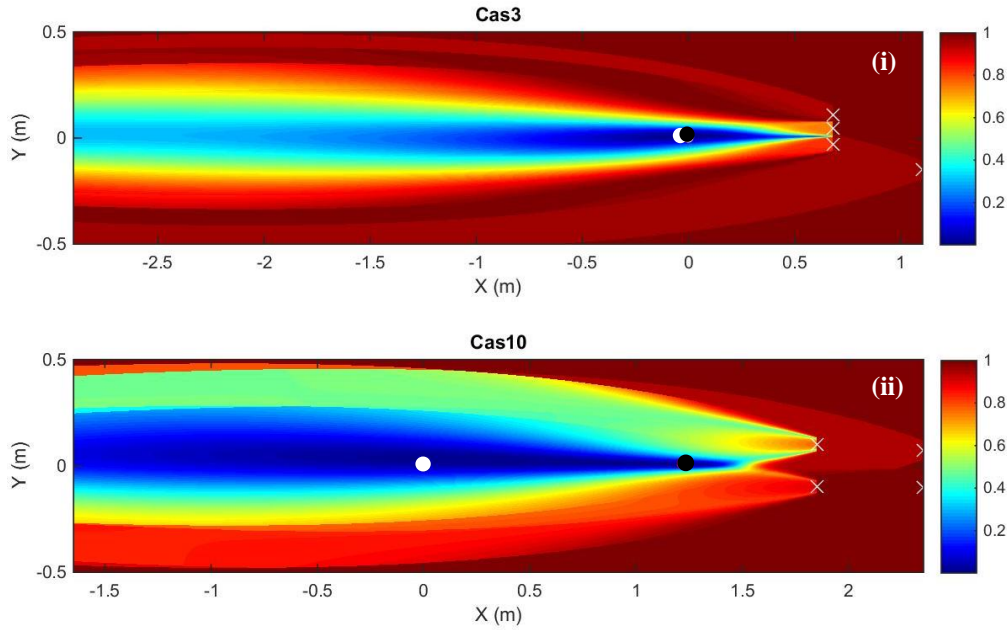


Figure 3.14. Résidus normalisés (i) Cas 3, (ii) Cas 10. La position réelle de la source correspond au disque blanc, la position estimée au disque noir. Les capteurs sont symbolisés par des croix.

La valeur  $J_R(\mathbf{x}_o)$  caractérise, pour un réseau donné  $R$ , l'écart entre le vecteur de mesures observé et le vecteur modélisé en utilisant la position et le débit estimés :

$$J_R(\mathbf{x}_o) = \frac{1}{2} \|\boldsymbol{\mu}_R - \boldsymbol{\mu}_{mod}(\mathbf{x}_o)\|_{H_w^{-1}}^2 \quad (3.50)$$

C'est donc une caractéristique qui pourra être utilisée pour comparer la qualité de plusieurs réseaux (formant un ensemble  $G_I$ ) dans une démarche d'optimisation. On pourra ainsi utiliser un critère de type « erreur d'estimation », de la même famille que ceux évoqués dans le second chapitre (équation 2.11), mais avec une norme différente, et choisir le réseau optimal tel que

$$R_o = \arg \min_{R \in G_I} J_R(\mathbf{x}_o) \quad (3.51)$$

### 3.7.2. La méthode de renormalisation dans le cadre d'une approche entropique

Le choix d'une fonction de pondération pour la résolution du problème inverse doit être le plus neutre possible afin de ne pas introduire trop d'information subjective. Ce critère a été formulé sous une forme entropique par (Issartel, 2005) et (Busch, 2015) : l'information fournie par les mesures dépend du choix de la fonction de pondération  $f$  et ne doit pas être

surestimée. Minimiser l'information revient à minimiser l'entropie des mesures  $H(\boldsymbol{\mu})$ . En faisant l'hypothèse que les mesures  $\mu_i$  avec  $i = 1 \dots m$  suivent une distribution Gaussienne, avec une matrice de covariance donnée par  $\mathbf{H}_f$ , (Issartel, 2005) a montré que l'entropie des mesures est

$$H(\boldsymbol{\mu}) = \frac{m}{2} \log(2\pi) + \frac{1}{2} \log \det \mathbf{H}_f \quad (3.52)$$

Comme nous l'avons évoqué au paragraphe 3.5.2.2, parmi toutes les matrices  $\mathbf{H}_f$  possibles, la matrice  $\mathbf{H}_\varphi$  (pour laquelle  $f$  vérifie les conditions de renormalisation) est celle dont le déterminant est minimum. L'entropie est minimale, et donc l'information moyenne fournie par les mesures est minimale, (c'est-à-dire la plus neutre possible) lorsque  $f = \varphi$ . Mais cette démonstration a été critiquée par ses auteurs. En effet, elle repose sur des hypothèses fortes : les mesures ont, a priori, une distribution gaussienne et leur matrice de corrélation est  $\mathbf{H}_f$ . Aussi, (Busch, 2015) a préféré s'appuyer sur la définition de l'entropie de Boltzmann pour obtenir un résultat similaire. En mécanique statistique, l'entropie d'un gaz, notée  $S$ , représente le logarithme du nombre de micro-états associé à un macro-état donné. (Busch, 2015) a donc défini l'entropie comme étant le logarithme du nombre de sources possibles. Le gain d'entropie est alors la différence entre l'entropie calculé après avoir réalisé les mesures et l'entropie calculée avant d'avoir réalisé les mesures.

$$\Delta S = S_{\text{après}} - S_{\text{avant}} \quad (3.53)$$

Afin de ne pas surestimer l'information fournie par les mesures, ce gain d'entropie (de variation négative) doit être maximisé. Or, une fois que les mesures ont été réalisées et que le vecteur  $\boldsymbol{\mu}$  est connu, la seule source possible est  $s_{//f}(\mathbf{x}) = \mathbf{a}(\mathbf{x})^T \mathbf{H}_f^{-1} \frac{\boldsymbol{\mu}}{f(\mathbf{x})}$ . L'entropie  $S_{\text{après}}$  est donc égale à  $\log(1) = 0$ . Maximiser  $\Delta S$  revient donc à minimiser  $S_{\text{avant}}$ . Mais, avant d'avoir réalisé les mesures, toutes les sources engendrées par le vecteur  $\frac{\mathbf{a}(\mathbf{x})}{f(\mathbf{x})}$  sont possibles et l'entropie est égale à l'infini.

(Busch, 2015) montre qu'en réalité le nombre de mesures que peut réaliser un capteur n'est pas infini. En effet, en général, un capteur a une précision finie et le résultat d'une mesure est nécessairement discret. Ainsi le nombre de mesures dans le monde réel est  $k_i \in \mathbb{Z}$ . L'espace de sources estimables est donc un espace discret défini par  $s_{//f} \in \{\frac{\mathbf{a}^T}{f} \mathbf{H}_f^{-1} \mathbf{k} \boldsymbol{\delta}, \mathbf{k} \in \mathbb{Z}^m\}$ . Il montre, également, que cet espace discret de dimension  $m$  est un hypercube constitué de cellules identiques de volume  $V_{\text{cell}} = \frac{\prod_i \delta_i}{\sqrt{\det \mathbf{H}_f}}$ . Le nombre de source estimables à l'intérieur



d'un volume  $V$  infiniment grand, est donc  $\frac{V}{V_{cell}}$ . L'entropie avant mesure (logarithme du nombre de sources estimables) est donc défini par

$$S_{avant} = \log\left(\frac{V}{\prod_i \delta_i}\right) + \frac{1}{2} \log \det \mathbf{H}_f \quad (3.54)$$

Minimiser cette entropie revient donc minimiser  $\det(\mathbf{H}_f)$ . Ceci correspond à une fonction de pondération optimale  $f = \varphi$  qui vérifie les conditions de renormalisation.

En plus de caractériser la fonction  $\varphi$ , les relations (3.51) et (3.49) montrent qu'un réseau de capteurs  $R$ , est caractérisé par une critère entropique  $S(R) = \frac{1}{2} \log \det \mathbf{H}_\varphi(R)$ . Ce critère pourra être utilisé pour comparer la qualité de plusieurs réseaux (formant un ensemble  $G_i$ ) dans une démarche d'optimisation. On pourra ainsi utiliser un critère basé sur l'entropie, proche de ceux évoqués dans le second chapitre (équation 2.18). Le réseau optimal sera celui pour lequel  $\det \mathbf{H}_\varphi$  sera le plus grand

$$R_o = \arg \max \frac{1}{2} \log \det \mathbf{H}_\varphi(R) \quad (3.55)$$

Le critère ci-dessus peut être utilisé pour comparer des configurations de réseaux pour une configuration météorologique donnée. Or sur un site d'implantation, les conditions météorologiques (vitesses et directions des vents) varient et sont, généralement, connues statistiquement. Chacune des  $k$  conditions est caractérisée par sa probabilité d'apparition  $P_i$  avec  $\sum_{i=1}^k P_i = 1$ . La performance d'un réseau doit donc être évaluée globalement pour l'ensemble des  $k$  conditions et c'est donc le critère entropique global qui doit être pris en compte.

L'entropie globale pour des conditions météorologiques multiples, regroupe l'entropie qui caractérise l'incertitude sur l'apparition des vents  $S_v = -\sum_{i=1}^k P_i \log(P_i)$  et l'entropie qui caractérise l'incertitude sur les sources estimables  $S_e = \sum_{i=1}^k P_i S_i(R)$ . Ainsi l'entropie globale est donnée par

$$S_g = -\sum_{i=1}^k P_i \log(P_i) + \sum_{i=1}^k P_i S_i(R) \quad (3.56)$$

Où  $S_i(R)$  est logarithme du nombre de sources estimables pour une condition  $i$ .

### 3.7.3. Définition des problématiques à étudier

Dans cette étude, trois problèmes sont proposés afin d'évaluer le processus d'optimisation de réseau, en utilisant un des critères défini ci-dessus.

3.7.3.1. *Problème 1 «Détermination du nombre optimal de capteurs pour la caractérisation d'une source»*

Dans le premier problème, Le but est de déterminer la taille optimale d'un réseau en vue de la caractérisation de la source. Les conditions météorologiques correspondent à une situation particulière et le réseau est appelé à estimer la position et le débit de la source d'une façon précise. Dans ce cas, on pourra utiliser un critère de type « erreur d'estimation », comme celui de l'équation (3.48).

3.7.3.2. *Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source »*

Pour le deuxième problème le réseau est également rapidement déployé sur le site mais aucune information n'est disponible sur la source. L'objectif est d'obtenir une estimation de l'intensité et l'emplacement de la source qui peut être positionnée n'importe où dans un large domaine. Ce problème correspond, par exemple, aux cas d'agressions terroristes ou de fuites sur des sites vastes sites industriels (raffineries complexes). Dans ces situations, les conditions météorologiques peuvent être connues instantanément mais les sources restent difficilement identifiable a priori. Le réseau est donc déployé rapidement pour estimer la position de la source afin de la neutraliser. Dans ce cas, on pourra utiliser un critère de type « entropique», comme celui de l'équation (3.52), afin de configurer un réseau le plus informatif possible.

3.7.3.3. *Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source »*

Pour le dernier problème, le réseau est permanent et fixe. Il est conçu pour protéger un site donné contre des agressions extérieures ou pour caractériser des contaminations internes accidentelles. La source peut être localisée n'importe où dans la zone de danger et aucun élément ne permet de présager de son intensité. L'objectif est de concevoir un réseau optimal qui conserve son efficacité lorsque les conditions météorologiques varient. On suppose que ces conditions météorologiques (vitesses et directions du vent) sont statistiquement connues. Ce problème correspond, par exemple, à la configuration de réseaux de défense chargés de protéger des sites sensibles (installations militaires, zones de stockage, réserves d'eau etc...) contre les agressions de type chimique et/ou biologique. Le réseau doit assurer une

surveillance efficace dans toutes les circonstances météorologiques Dans ce cas, on pourra utiliser un critère de type « entropique moyen », comme celui de l'équation (3.53).

### 3.8. CONCLUSION DU CHAPITRE

Dans ce chapitre, nous avons présenté les différentes approches disponibles pour résoudre le problème d'estimation des sources. Nous avons montré l'avantage d'utiliser les « approches basées récepteurs » et en particulier la méthode dite de renormalisation. Ainsi, une large partie du chapitre a été consacrée à la présentation et à l'évaluation de cette méthode. Nous avons détaillé le formalisme mathématique de la méthode et avons contribué à son développement en proposant un algorithme de renormalisation rapide et efficace. Nous avons ensuite montré que cette méthode donne de très bon résultats sur des données expérimentales obtenues afin d'évaluer la configuration de réseaux de mesures (projet DYCE). Nous avons concrètement mis en évidence l'influence de l'architecture des réseaux sur la précision de l'estimation des paramètres de la source (position et débit). La dernière partie du chapitre a été consacrée à la définition de critères d'optimalité à partir de notions issues de la méthode de renormalisation. Nous avons enfin définis trois problématiques qui seront testées dans le quatrième chapitre :

- Le premier problème concerne l'optimisation de la taille et de l'emplacement des capteurs dans des situations météorologiques particulières. Dans ce cas, nous proposons de minimiser une « erreur d'estimation ».
- Le deuxième problème concerne l'optimisation de réseaux mobiles déployés dans des situations d'urgence mais pour lesquelles on ne dispose d'aucune information a priori sur la localisation de la source. Pour définir un critère d'optimalité, nous nous sommes basés sur l'interprétation entropique de la méthode de renormalisation.
- Le dernier problème concerne l'optimisation de réseaux permanents et fixes qui doivent être efficaces dans différentes conditions météorologiques. Par conséquent, aucune information a priori de la source n'est possible et l'ensemble de données constitue des données météorologiques sous forme statistique (rose de vent annuelle). Pour répondre à cette problématique, nous proposons de maximiser une fonction objectif basée sur une entropie moyenne.

« BIBLIOGRAPHIE CHAPITRE 3 »

Abida, R. and Bocquet, M., (2009). Targeting of observations for accidental atmospheric release monitoring. *Atmospheric Environment*, 43 (40), pp. 6312-6327. doi.org/10.1016/j.atmosenv.2009.09.029

Barad, M.L, (1958). Project prairie Grass, a field program in diffusion. *Geophys.Res. Papers*, N° 59 vols I and II, report AFCRC-TR-58-235, U.S. Air Force Cambridge Research Center.

Ben Salem, N, (2014). Modélisation directe et inverse de la dispersion atmosphérique en milieux complexes. Ecole centrale de Lyon. <http://www.theses.fr/2014ECDL0023>

Bird, R.H., Stewart, W.E. and Lightfoot, E.N. (2002). *Transport Phenomena*, Ed. John Wiley & Sons, 2002.

Bocquet, M, (2005-a). Grid resolution dependence in the reconstruction of an atmospheric tracer source. *Nonlinear Processes in Geophys.* 12: 219-233. doi:10.5194/npg-12-219-2005.

Briggs, G.A, (1973). Diffusion estimation of small emissions. Contribution No.79. Atmospheric Turbulence and Diffusion Laboratory, Oak Ridge, TN.

Busch, X, (2015). Renormalized data assimilation : an entropic approach. Deuxième colloque sur la Modélisation de la dispersion atmosphérique, Identification de source, Evry, France.

Carpenter, S., Montgomery, T.L., Leavitt, J.M., Colbaugh, W.C., Thomas, F.W., (1971). Principal plume dispersion models : TVA power plants. *J. Air Pollut. Cont. Assoc.*, 21, 491-495.

Chang, J. et Hanna S., (2012). Acceptance criteria for urban dispersion model evaluation, *Meteorology and Atmospheric Physics* 2012 ; 116: 133. doi:10.1007/s00703-011-0177-1.

Chow, J. C., Doraiswamy, P., Watson, J.G., Chen, L-W. A., Ho, S. S. H., Sodeman, D. A., (2008). Advances in integrated and continuous measurements for particle mass and chemical composition. *J. Air Waste Manage. Assoc.* 58, 141-163.

Davoine, X., et Bocquet, M., (2007). Inverse modelling-based reconstruction of the Chernobyl source term available for long-range transport. *Atmos. Chem. Phys.* 7, 1549-1564.

Demaël, E., (2007). Modélisation de la dispersion atmosphérique en milieu complexe et incertitudes associées. Thèse de doctorat de l'École Nationale des Ponts et Chaussées. [http://cerea.enpc.fr/fich/these\\_demaël.pdf](http://cerea.enpc.fr/fich/these_demaël.pdf)

Delle Monache, L., et al. (2008). Bayesian inference and Markov chain Monte Carlo sampling to reconstruct a contaminant source on a continental scale, *J. Appl. Meteorol. Climatol.*, 47, 2600-2613.

DISCARD, (2015). Régime d'Appui aux Pme pour l'Innovation Duale (RAPID). Lot N°6. Application aux jeux de données expérimentales MUST (Mock Urban Setting Test) pour l'étude de la performance.

Draxler, R. (1976). Determination of atmospheric diffusion parameters. *Atmos. Environ.*, 10, 99-105.

Ferziger, J. H., and Peric, M., (2002). *Computational Methods for Fluid Dynamics*, Springer, Berlin, doi:10.1007/978-3-642-56026-2.

Flesch, T.K., Wilson, J.D., Yee, E., (1995). Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol.* 34, 1320-1332.  
Garratt, J.R., 1990. The internal boundary layer. *Boundary-Lay. Meteorol.* 50, 171-203.

Flesch, T.K., (1996). The footprint for flux measurements, from backward Lagrangian stochastic models. *Boundary-Lay. Meteorol.* 78, 399-404.

Gamerman, D., et Lopes, H. F., (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Taylor and Francis, London, 323 pp.

Gilks, W. R., Richardson, S., Spiegelhalter, D. J., (1996). *Markov Chain Monte Carlo in Practice*, Chapman and Hall/CRC, London, 486 pp.

Gryning, S.E., van Ulden, A.A.M., Irwin, J.S., Sievertsen, D., (1987). Applied dispersion modelling based on meteorological scaling parameters. *Atmos. Environ.* 21, 79-89.

Hanna, S.R., Briggs, G.A., Hosker, R.P. Jr., (1982). « Handbook on atmospheric diffusion ». Technical Information Center. U.S. Department of Energy.

Hourdin, F, et Talagrand O., (2006). Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parametrization of subgrid-scale transport. *Q. J. R. Meteorol. Soc.* 132: 585-603.

Hourdin F, Idelkadi A, Talagrand O (2005). Eulerian backtracking of atmospheric tracers: II numerical aspects, *QJRMS* (in press).

Hourdin, F., Issartel, J.-P., Cabrit, B., Idelkadi, A., (1999). Reciprocity of atmospheric transport of trace species. *C.R.A.S.*, 329, 623-628.

INERIS (2015). Guide de Bonnes Pratiques pour la réalisation de modélisations 3D pour des scénarios de dispersion atmosphérique en situation accidentelle. 180 p., [http://www.ineris.fr/aida/sites/default/files/gesdoc/86009/Guide\\_Bonnes\\_Pratiques.pdf](http://www.ineris.fr/aida/sites/default/files/gesdoc/86009/Guide_Bonnes_Pratiques.pdf)

Issartel, J.-P. et Baverel, J., (2002). Adjoint backtracking for the verification of the Comprehensive Test Ban Treaty. *Atmospheric Chemistry and Physics Discussions*, European Geosciences Union, 2002, 2 (6), pp.2133-2150. <hal-00327815>.

Issartel, J.-P. et Baverel, J., (2003). Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmos. Chem. Phys.* 3: 475-486. doi:10.5194/acp-3-475-2003.

Issartel, J.-P., (2005). Emergence of a tracer source from air concentration measurements: a new strategy for linear assimilation, *Atmos. Chem. Phys.* 5, 1, 249-273, doi:10.5194/acp-5-249-2005.

Issartel, J-P., Sharan, M., Modani, M., (2007). An inversion technique to retrieve the source of a tracer with an application to synthetic satellite measurements, *Proc. Royal Society A*, 463, 2863-2886, doi:10.1098/rspa.2007.1877.

Issartel, J-P., Sharan, M., Singh, S.K., (2012). Identification of a point of release by use of optimally weighted least squares. *Pure Appl. Geophys.*, 169(3), 467-482, doi:10.1007/s00024-011-0381-4.

Keats, A., Yee, E., Lien, F-S., (2007). Bayesian inference for source determination with applications to a complex urban environment. *Atmospheric Environment* 41(3):465-479. DOI: 10.1016/j.atmosenv.2006.08.044.

Keats, W. A, 2009. Bayesian Inference for Source Determination in the Atmospheric Environment, Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, Canada.

Lepley, J. J., Lloyd, D., Robins, A., Wilks, A., Rudd, A., Belcher, S., (2011). Dynamic sensor deployment for the monitoring of chemical releases in urban environments (DYCE), *Proc. of SPIE*, 8018(12), 1-11, doi:10.1117/12.883373.

Liu, X., and Z. Zhai (2007). Inverse modeling methods for indoor airborne pollutant tracking: Literature review and fundamentals, *Indoor Air*, 17(6), 419-438, doi:10.1111/j.1600-0668.2007.00497.x.

Marchuk, G.I, (1995). Adjoint Equations and Analysis of Complex Systems, Kluwer. Book Mathematics and Its Applications Volume 295.

Michaelides, M. P. et Panayiotou, C. G., (2005). Plume Source Position Estimation Using Sensor Networks. *Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation.*

Pasquill, F, (1961). Estimation of the dispersion of windborne material. *Meteorol. Mag.*, 90, 3349. 1, 24, 60.

Pasquill, F. et Smith, F.B., (1983). Atmospheric diffusion-study of dispersion of windborne material from industrial and other sources. *Ellis Horwood Series in Environmental Science*, Chichester, 320 pp.

Perkins, R., Soulhac, L., and Mejean, P., (2002), Modélisation de la dispersion des émissions atmosphériques d'un site industriel - Vers un guide de l'utilisateur. 1ère partie: État de l'art. Association R.E.C.O.R.D edn. Laboratoire de Mécanique des Fluides et d'Acoustique. UMR CNRS 5509 - UCBL1 - ECL. 18, 19, 24, 57

Pudykiewicz, J, (1998). Application of adjoint tracer transport equations for evaluating source parameters. *Atmospheric Environment*, 32, 17, 3039-3050.

Rao, K, S, (2007). Source estimation methods for atmospheric dispersion. *Atmos. Environ.* 41,6964-6973. doi:10.1016/j.atmosenv.2007.04.064.

Robertson, L. Persson C., (1994). Attempts to apply four dimensional data assimilation of radiological data using the adjoint technique. *Radiat. Prot. Dosim.* 50: 333-337.



Rodgers, C.D, (2000). Inverse methods for atmospheric sounding: theory and practice. World Scientific Publishing Co. Pte. Ltd., Singapore, ISBN 978-981-02-2740-1, 85 pp.

Rudd, A. C., Belcher, S. E., Robins, A. G., Lepley, J. J., (2012). An inverse method for determining source characteristics for emergency response applications, *Boundary Layer Meteorol.*, 144(1), 1-20.

Seibert, P, (2001). Inverse modelling with a Lagrangian particle dispersion model: application to point releases over limited time intervals. In: *Air Pollution Modeling and its Application XIV*, GryningSE, SchiermeierFA (eds). Plenum Press: 381-389.

Sharan, Issartel, M., J-P., Singh, S. K., Kumar, P., (2009). An inversion technique for the retrieval of single point emissions from atmospheric concentration measurements, *Proc. R. Soc. A*, 65, 2069-2088, doi:10.1098/rspa.2008.0402.

Sharan, M., Singh, S.K., Issartel, J-P., (2012-a). Least Square Data Assimilation for Identification of the Point Source Emissions. *Pure Appl. Geophys.* 169, 483-497.

Sharan, M., J.-P. Issartel, and S. K. Singh (2012-b), A point-source reconstruction from concentration measurements in low-wind stable conditions, *Q. J. R. Meteorol. Soc.*, 138, 1884-1894, doi:10.1002/qj.1921.

Singer, I.A. et Smith, M.E., (1966). Atmospheric dispersion at Brookhaven National Laboratory. *Int. J. Air. Water Pollut.* 10, 125-135.

Singh, S. K., Sharan, M., Issartel J-P., (2013). Inverse modelling for identification of multiple-point releases from atmospheric concentration measurements, *Boundary Layer Meteorol.*, 146, 277-295, doi:10.1007/s10546-012-9765-y.

Singh, S.K. et Rani R., (2014). A least-squares inversion technique for identification of a point release: Application to Fusion Field Trials 2007 Atmospheric environment 92, 104-117.

Stohl, A. et Seibert, P., (1998). Accuracy of trajectories as determined from the conservation of meteorological tracers. *Q. J. R. Meteor. Soc.*, 125, 1465-1484 (No. 549, July 1998 Part A).

Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G., (2005). Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.*, 5, 2461-2474, 2005 [www.atmos-chem-phys.org/acp/5/2461/](http://www.atmos-chem-phys.org/acp/5/2461/) SRef-ID: 1680-7324/acp/2005-5-2461 European Geosciences Union

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Grealley, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., (2009). An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, *Atmos. Chem. Phys.*, 9, 1597-1620, doi:10.5194/acp-9-1597-2009, 2009.

Thomson, L. C., Hirst, B., Gibson G., Gillespiec, S., Jonathanc, P., Skeldona, K. D., Padgetta, M. J., (2007). An improved algorithm for locating a gas source using inverse methods, *Atmospheric Environment*, vol. 41, no. 6, pp. 1128-1134. doi.org/10.1016/j.atmosenv.2006.10.003.

Turbelin, G., Singh, S. K., Issartel J-P., (2014). Reconstructing source terms from atmospheric concentration measurements: Optimality analysis of an inversion technique, *J. Adv. Model. Earth Syst.*, 6, 1244-1255, doi:10.1002/2014MS000385.

Uliasz, M, (1993). The atmospheric mesoscale dispersion modeling system (MDMS). *J. Appl. Meteor.*, 32, 139-149.

Vollmer, M. K., Weiss, R. F., and Yokouchi, Y., (2009). An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, *Atmos. Chem. Phys.*, 9, 1597-1620, doi:10.5194/acp-9-1597-2009.

Wilson J. D. et Sawford, B. L., (1996). Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. *Boundary-Layer Meteorology* 78, 191-210.

Winiarek, V, (2014). Dispersion atmosphérique et modélisation inverse pour la reconstruction de sources accidentelles de polluants. Université Paris-Est, 2014. Français. <https://pastel.archives-ouvertes.fr/tel-01004505>.

Wotawa, G., De Geera, L-E., Deniera, P., Kalinowskia, M., Toivonena, H., D'Amours, R., Desiatod, F., Issartel, J-P., Langerg, M., Seiberth, P., (2003). Atmospheric transport modelling in support of CTBT verification – Overview and basic concepts. *Atmospheric Environment.*, Volume 37, Issue 18 Pages 2529-2537. doi.org/10.1016/S1352-2310(03)00154-7.

Wyngaard, John C., (2010). *Turbulence in the Atmosphere*, Cambridge University Press, Cambridge, UK.

Yee, E., Lien, F. S., Keats, A., D'Amours, R., (2008). Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. *J. Wind Eng. Ind. Aerodyn.*, 96, 1805-1816.

Yee, E, (2012). Inverse dispersion for an unknown number of sources : Model selection and uncertainty analysis, *ISRN Appl. Math.*, 2012, 1- 20, doi:10.5402/2012/465320.



**CHAPITRE 4 :  
EVALUATION DU PROCESSUS  
D'OPTIMISATION**

## CHAPITRE 4 EVALUATION DU PROCESSUS D'OPTIMISATION

### Plan du chapitre

CHAPITRE 4 EVALUATION DU PROCESSUS D'OPTIMISATION .....	111
4.1. Introduction.....	114
4.2. Approche analytique « Expérience DYCE » .....	115
4.2.1. Définition du problème d'optimisation .....	115
4.2.1.1. Description des calculs .....	116
4.2.1.1.1. Echelle de complexité.....	116
4.2.1.1.2. Couplage entre l'algorithme de renormalisation et l'algorithme d'optimisation.....	117
4.2.2. Problème 1 «Détermination du nombre optimal de capteurs pour la caractérisation d'une source » .....	118
4.2.2.1. Paramétrage .....	119
4.2.2.1.1. Recuit simulé.....	119
4.2.2.1.2. Algorithme génétique .....	120
4.2.2.2. Résultats et discussions .....	120
4.2.2.3. Conclusion à l'issue du problème 1 .....	125
4.2.3. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source » .....	126
4.2.3.1. Paramétrage .....	126
4.2.3.2. Résultats et discussions .....	126
4.2.3.3. Conclusion à l'issue du problème 2.....	129
4.3. Approche CFD « Expérience MUST » .....	129
4.3.1. Description des expériences.....	130
4.3.2. Modélisation inverse de (Kumar et al, 2015) .....	132
4.3.2.1. Vérification de la simulation inverse .....	134
4.3.3. Définition du problème d'optimisation .....	135
4.3.3.1. Description des calculs .....	135
4.3.3.1.1. Echelle de complexité.....	135
4.3.3.1.2. Couplage entre l'algorithme d'optimisation l'algorithme de renormalisation et les « champs adjoints CFD » .....	135
4.3.4. Problème 1 « Détermination du nombre optimal de capteurs pour	

la caractérisation de source » .....	136
4.3.4.1. Paramétrage .....	136
4.3.4.2. Résultats et discussions .....	136
4.3.4.3. Conclusion à l'issue du problème 1 .....	149
4.3.5. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source ».....	150
4.3.5.1. Paramétrage .....	150
4.3.5.2. Résultats et discussions .....	150
4.3.5.3. Conclusion à l'issue du problème 2 .....	154
4.3.6. Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source » .....	154
4.3.6.1. Paramétrage .....	155
4.3.6.2. Résultats et discussions .....	155
4.3.6.3. Conclusion à l'issue du problème 3 .....	158
4.4. Conclusion du chapitre.....	159
« Bibliographie chapitre 4 ».....	161

### 4.1. INTRODUCTION

Tout au long de ce chapitre, nous présentons l'évaluation du processus d'optimisation défini au chapitre 2 et complété au chapitre 3. Les trois problématiques (définies dans la partie 3.7.3) sont testées dans le cadre d'une approche analytique et d'une approche CFD. L'objectif est d'identifier les paramètres de la source (position  $(x_s, y_s)$  et intensité  $q_s$ ). La figure 4.1 représente un résumé des deux approches ainsi que les différentes problématiques considérées.

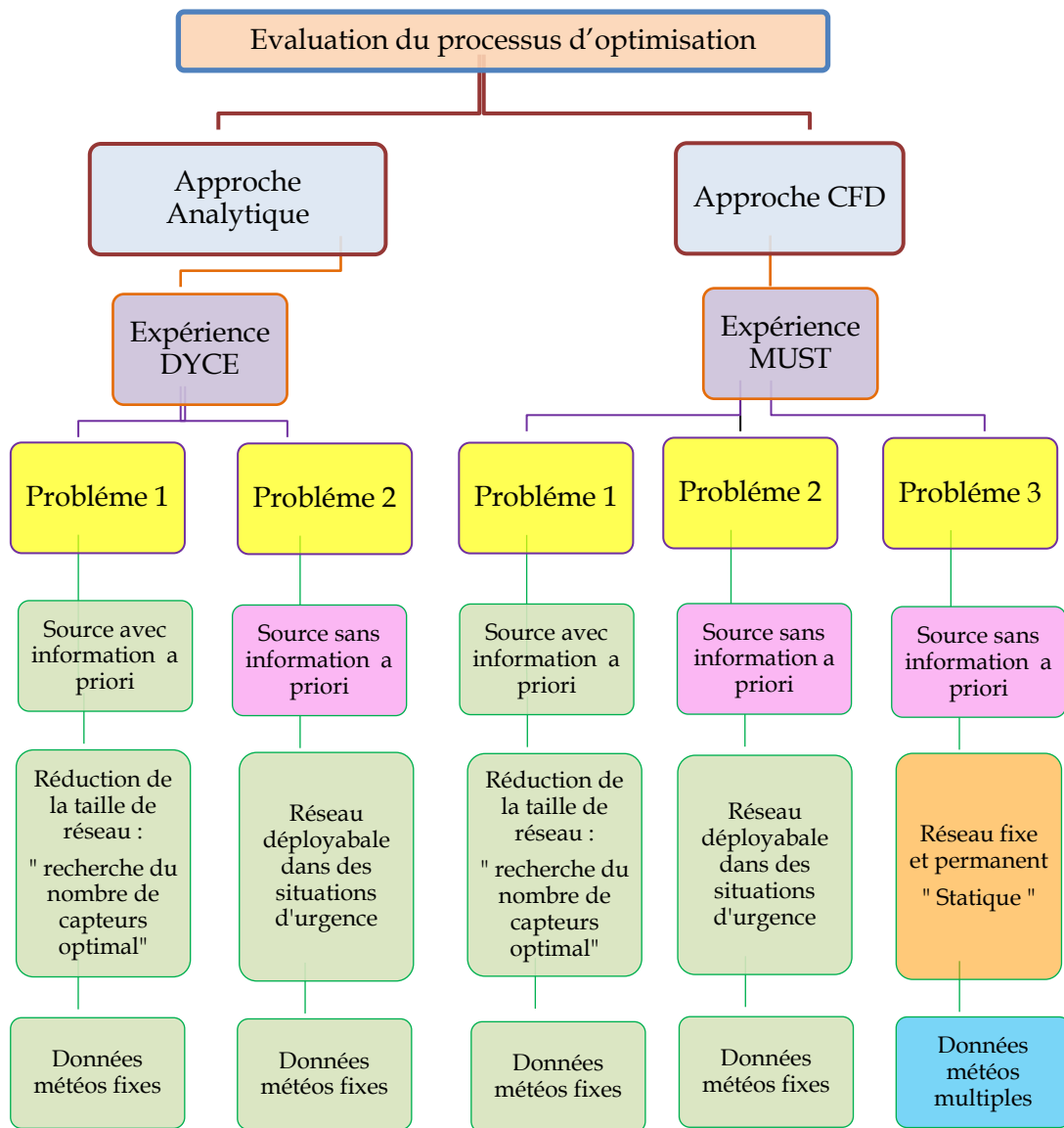


Figure 4.1. Représentation des approches et des problématiques évaluées

L'approche analytique est retenue lorsque les domaines de dispersion sont plats et sans obstacles. Ainsi, les champs adjoints sont obtenus à l'aide de modèles de type « Gaussiens ».

En revanche, lorsque les domaines de dispersion ont une forte rugosité (présence d'obstacles), une approche CFD est utilisée. On génère ainsi, des écoulements de vents plus réalistes mais la modélisation des champs adjoints devient plus complexe (modèle CFD adjoint). Les données des expériences DYCE et de MUST servent de support à la validation des approches analytique et CFD. Ces approches ont été illustrées à travers trois problèmes : (i) « *Problème 1* », l'objectif est de déterminer la taille optimale d'un réseau en vue de caractériser la source d'émission dans une situation météorologique particulière. (ii) « *Problème 2* » l'objectif est la conception optimale d'un réseau capable d'estimer l'intensité et l'emplacement de la source dans un large domaine, pour une condition météorologique particulière. Cette configuration peut correspondre au déploiement en urgence d'un réseau mobile sans connaissance a priori de la source. (iii) « *Problème 3* » l'objectif est d'obtenir un réseau statique optimal capable de caractériser la source dans un large domaine, pour des conditions météorologiques multiples. Cette configuration peut correspondre à la conception optimale d'un réseau statique sans connaissance a priori de la source.

## 4.2. APPROCHE ANALYTIQUE « EXPERIENCE DYCE »

Dans cette approche, le domaine de dispersion est plat et non rugueux. La modélisation des champs adjoint est basée sur des modèles de type « Gaussien » stationnaire. Des données expérimentales issues de l'expérience DYCE sont exploitées pour évaluer la méthodologie d'optimisation. L'expérience de Dyce (*Lepley et al. 2011*), présentée dans la partie 3.6.1, a été réalisée au centre de recherche ENFLO à l'Université de Surrey au Royaume Uni. Il s'agit de la dispersion du gaz propane dans une soufflerie.

### 4.2.1. Définition du problème d'optimisation

La région d'intérêt notée  $\Omega$ , est une section rectangulaire de la soufflerie (voir figure 4.2). Dans  $\Omega$ , 27 positions potentielles ont été considérées par l'expérimentateur. Ces positions sont déduites des configurations testées, dans le cadre du projet DYCE, pour l'inversion des données par le modèle de (*Rodgers, 2000*). L'ensemble  $G_I$  des positions potentielles  $\mathbf{x}_i = (x_i, y_i)_{i=1...27}$  est ainsi défini par

$$G_I = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{27}\} \subset \Omega$$

Dans un premier temps, nous cherchons les « meilleurs réseaux » de tailles  $m$  qui varient de 3 à 13, c'est à dire que nous cherchons à déterminer les  $m$  positions optimales  $\mathbf{x}_i^o = (x_i^o, y_i^o)_{i=1...m}$ .

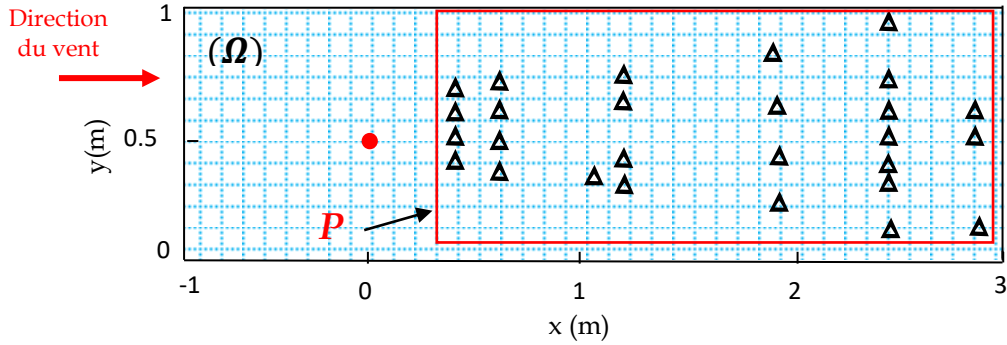


Figure 4.2. Schéma de l'expérience de DYCE, (disque) la source, (triangles) les 27 positions potentielles

Le choix du nombre de capteurs  $m$  reste arbitraire ; dans notre étude nous avons choisi une taille minimale de 3 capteurs car l'objectif de chaque réseau est d'estimer trois paramètres de la source (position spatiale  $(x_s, y_s)$  et intensité  $q_s$ ).

Pour la taille maximale nous avons choisis 13 capteurs, elle correspond à la moitié du domaine de recherche (27 positions potentielles). Chacun des 10 « réseaux optimaux » est donc défini par

$$R_{0m} = \{x_1^o, x_2^o, \dots, x_m^o\} \subset G_I$$

Les réseaux optimaux (de tailles  $m=3$  à  $m=13$ ) sont obtenus par un processus de recherche combinatoire en se basant sur deux fonctions coûts  $Jp_1$  et  $Jp_2$  que nous évoquerons ci-dessous. Ces fonctions sont spécifiques aux deux problèmes définis dans le chapitre 3 et rappelés dans l'introduction de ce chapitre (*i* et *ii*). Par exemple, pour le « problème 1 », les réseaux optimaux sont définis par

$$R_{0m} = \arg \min_{R_m \subset G_I} Jp_1(R_m)$$

Une fois les différents réseaux obtenus, nous examinons a posteriori l'erreur commise dans l'estimation de la source connue ( $x_s=0$  m,  $y_s=0,5$  m et  $q_s=8,925.10^{-7}$  kg/s) pour évaluer leur efficacité.

#### 4.2.1.1. Description des calculs

##### 4.2.1.1.1. Echelle de complexité

Comme introduit ci-dessus, il faut choisir  $m = 3, \dots, 13$  capteurs parmi 27 positions potentielles. Ainsi, le nombre de combinaisons possibles varie de  $C_{27}^3$  à  $C_{27}^{13}$ .

La figure 4.3 montre l'évolution de la complexité du problème d'optimisation combinatoire en fonction du nombre de capteurs.

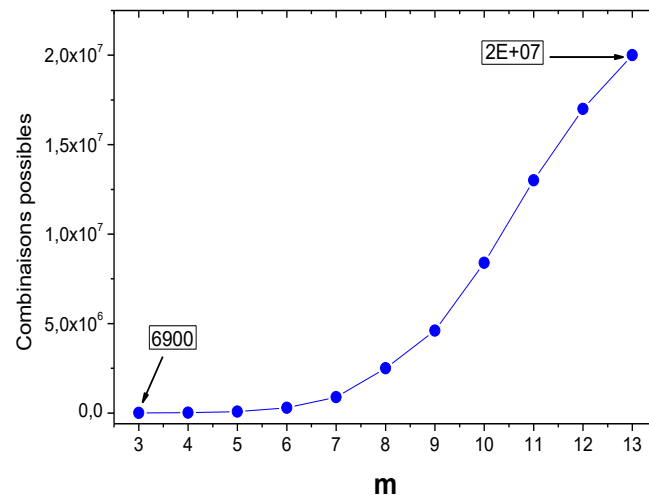


Figure 4.3. Evolution du nombre de combinaisons en fonction du nombre de capteurs

Comme nous pouvons le constater, le nombre de combinaisons évolue d'une façon considérable lorsque la taille du réseau augmente. Ainsi, les combinaisons varient de 6900 pour 3 capteurs à  $2 \cdot 10^7$  pour 13 capteurs.

L'idée pour résoudre ce genre de problème d'optimisation, qualifié de « *NP-difficile* », consiste à utiliser des méthodes approchées et non des méthodes exactes. Nous rappelons que notre choix s'est porté sur l'algorithme du « recuit simulé » et sur l'algorithme « génétique » en raison de la nature particulière de chacun d'entre eux (voir 2.5.3)

#### 4.2.1.1.2. Couplage entre l'algorithme de renormalisation et l'algorithme d'optimisation

Nous expliquons ici le processus de recherche du réseau optimal. L'optimisation est élaborée par un couplage entre l'algorithme d'optimisation (recuit simulé ou génétique) et l'algorithme de renormalisation. Une fois le paramétrage effectué, le code d'optimisation communique les positions des  $m$  capteurs ( $\mathbf{x}_i = (x_i, y_i)_{i=1\dots m}$ ) au code d'inversion. Celui-ci lui retourne la fonction objectif  $J(R_m)$ . La qualité de cette fonction est ainsi examinée par le code d'optimisation, qui au besoin régénère un nouvel ensemble de  $m$  capteurs dans le cadre d'une optimisation combinatoire.

Ce processus d'échange entre les deux algorithmes se fait à chaque itération (ou population) jusqu'à la convergence. Le principe de couplage est schématisé sur la figure 4.4.

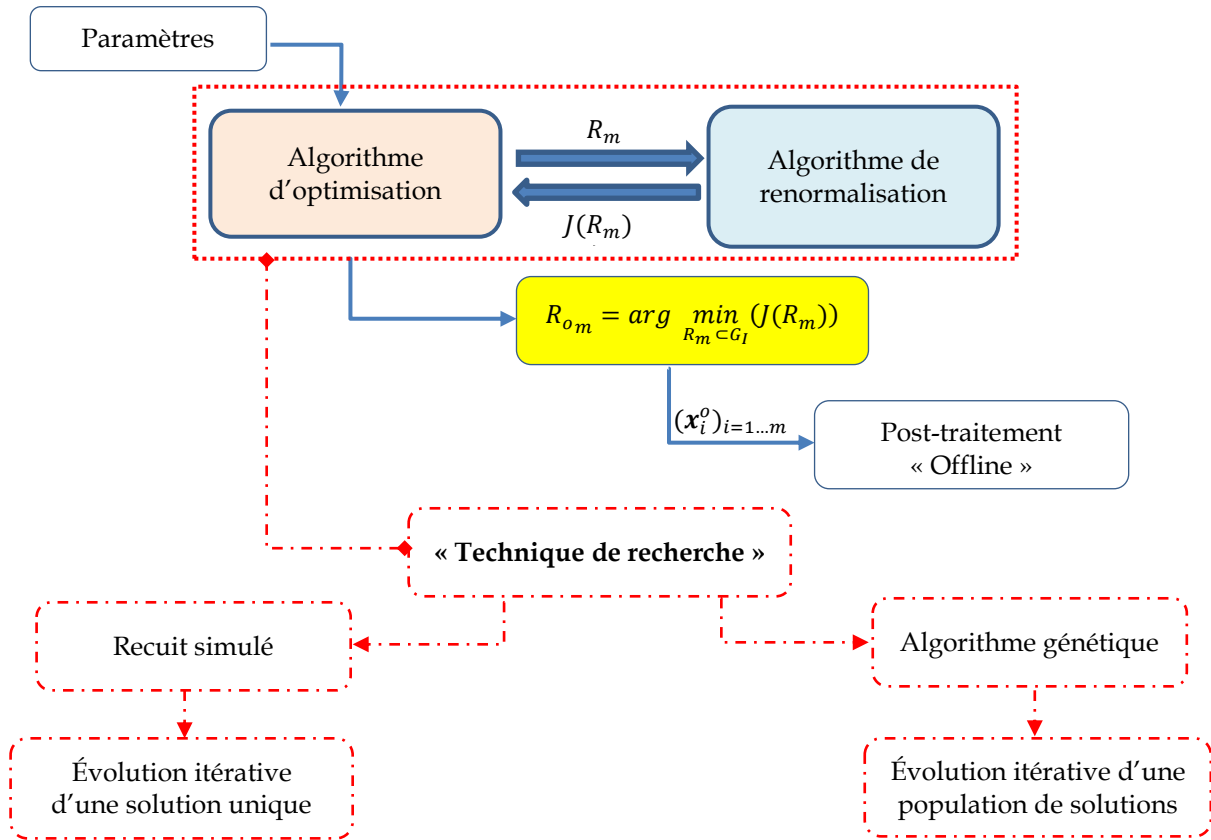


Figure 4.4. Couplage entre l’algorithme d’optimisation et l’algorithme de renormalisation

Il est à noter que, pour le calcul d’optimisation, nous avons réduit le maillage utilisé pour l’évaluation de la méthode de renormalisation. Dans la partie 3.6.2, les calculs ont été réalisés à l’échelle de la soufflerie, sur un domaine de dimensions  $L \times l = 4 \times 1 \text{ m}$  discrétisé en  $2000 \times 500$  mailles soit des éléments de longueur  $dx = dy = 0,002 \text{ m}$ , mais pour le calcul de l’optimisation nous avons discrétisé le domaine en  $400 \times 100$  mailles soit des éléments de longueur  $dx = dy = 0,01 \text{ m}$ . Ainsi, nous réduisons la taille de la matrice des champs adjoints  $A$  et le temps de calcul/configuration de 15s à 1s (assez long dans le processus d’optimisation).

#### 4.2.2. Problème 1 «Détermination du nombre optimal de capteurs pour la caractérisation d’une source »

Le premier problème introduit au niveau du paragraphe 3.7.3.1 est un cas d’étude pour des conditions météorologiques fixes (direction et vitesse de vent) et pour une source supposée inconnue (position et débit). L’objectif est de déterminer le nombre minimal de capteurs qui



permet de reconstituer cette source. Pour obtenir les réseaux optimaux  $R_{0m}$ , la fonction objectif à minimiser (résidus normalisés définis dans la partie 3.7.1) est

$$Jp_1(R_m) = 1 - \frac{s_{//w}^2(x_0)}{\|\mu\|_{H_w^{-1}}^2}$$

Où  $\|\mu\|_{H_w^{-1}}^2 = \mu^T H_w^{-1} \mu$

Nous rappelons qu'une fois que les réseaux  $R_{m=3,\dots,13}$  sont obtenus, nous évaluerons leur efficacité a posteriori (reconstitution de la source ( $x_s=0$  m,  $y_s=0,5$  m et  $q_s=8,925.10^{-7}$ kg/s)).

#### 4.2.2.1. Paramétrage

##### 4.2.2.1.1. Recuit simulé

Conformément aux recommandations de la littérature (partie 2.5.3.1.3.), nous avons utilisé le schéma de refroidissement exponentiel. Ainsi, la température décroît par palier. Suite à une étape de recherche préliminaire, les paramètres retenus pour le recuit simulé sont regroupés dans le tableau 4.1 et sont respectivement : la température de départ  $T_0$ , la température d'arrêt  $T_f$ , le nombre d'itérations par palier de température  $L$ , le facteur de refroidissement  $\theta$ . On rappelle que  $m$  est le nombre de capteurs (nombre de mesures).

$m$	$T_0$	$T_f$	$L$	$\theta$
3	$6.10^{-3}$	$1.10^{-13}$	50	0,9
4	$6.10^{-3}$	$1.10^{-13}$	50	0,9
5	$6.10^{-3}$	$1.10^{-13}$	50	0,9
6	$6.10^{-3}$	$1.10^{-13}$	70	0,9
7	$6.10^{-3}$	$1.10^{-13}$	70	0,9
8	$6.10^{-3}$	$1.10^{-13}$	100	0,9
9	$6.10^{-3}$	$1.10^{-13}$	100	0,9
10	$6.10^{-3}$	$1.10^{-13}$	100	0,9
11	$6.10^{-3}$	$1.10^{-13}$	120	0,9
12	$6.10^{-3}$	$1.10^{-13}$	120	0,9
13	$6.10^{-3}$	$1.10^{-13}$	120	0,9

Tableau 4.1. Paramétrage du recuit simulé « problème 1 »

- Choix de la solution voisine (schéma de perturbation)

Un dernier point utile à souligner, est comment générer la solution voisine, ou encore, comment perturber la solution à une itération donnée. Nous avons testé cinq schémas de perturbation pour déterminer la meilleure méthode. La figure 4.5 regroupe les cinq schémas testés. Nous avons retenus le schéma n°5. Il consiste à changer aléatoirement le nombre de capteurs à chaque itération. Par cette méthode on explore plus largement l'espace de solutions possibles. De plus, on diminue la probabilité d'être piégé dans un minimum local.

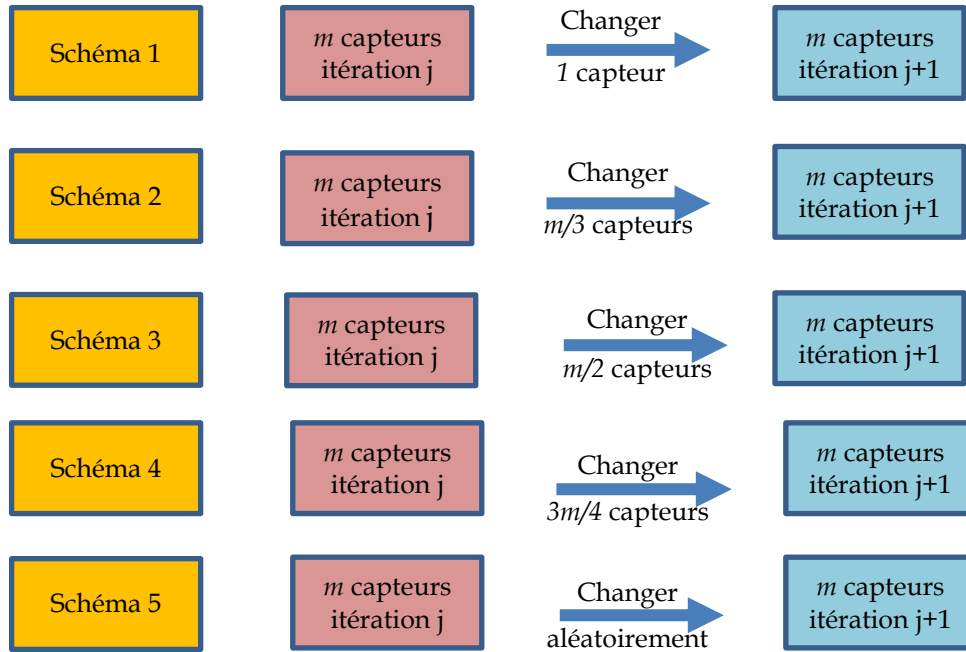


Figure 4.5. Shémas de perturbation testés

4.2.2.1.2. *Algorithme génétique*

Les paramètres retenus pour le paramétrage de l’algorithme génétique sont regroupés dans le tableau 4.2, et sont respectivement : la taille d’une population  $Tp$ , le nombre de génération  $Ng$ , le taux de croisement  $Tc$  et le taux de mutation  $Tm$ .

$m$	$Tp$	$Ng$	$Tc$	$Tm$
3	30	500	aléatoire	1 capteur/chromosome
4	30	500	aléatoire	1 capteur/chromosome
5	30	500	aléatoire	1 capteur/chromosome
6	30	500	aléatoire	1 capteur/chromosome
7	40	500	aléatoire	1 capteur/chromosome
8	40	500	aléatoire	1 capteur/chromosome
9	40	700	aléatoire	1 capteur/chromosome
10	40	700	aléatoire	1 capteur/chromosome
11	50	1000	aléatoire	1 capteur/chromosome
12	50	1000	aléatoire	1 capteur/chromosome
13	50	1000	aléatoire	1 capteur/chromosome

Tableau 4.2. Paramétrage de l’algorithme génétique « problème 1 »

4.2.2.2. *Résultats et discussions*

Nous rappelons que la vitesse et la direction du vent sont fixes ( $U=1,3$  m/s et l’angle  $\theta = 0^\circ$  par rapport à la direction des  $x>0$ , voir la figure 4.2), et que la source est connue ( $x_s=0$  m,  $y_s=0,5$  m et  $q_s=8,925.10^{-7}$  kg/s). L’idée consiste à examiner l’influence du nombre des capteurs sur l’erreur d’estimation des paramètres de la source. Pour ce faire, nous déterminons les

configurations optimales composées de 3 à 13 capteurs par l'algorithme génétique ainsi que par le recuit simulé.

#### 4.2.2.2.1. Recuit simulé

Tout d'abord nous analysons la configuration optimale, de taille  $m=4$ , sélectionnée par le recuit simulé (figure 4.6.a) ainsi que l'évolution de la fonction objectif pendant le processus de recherche (figure 4.6.b). Les 4 capteurs de la configuration optimale sont ceux situés plus proches de la source. Ceci est en cohérence avec l'interprétation physique de la fonction de visibilité d'un réseau. En effet, les capteurs sous le vent, proches de la position de la source, ont une meilleure visibilité que ceux éloignés. Ils permettent donc de mieux caractériser la source (voir figure 3.13)

Pour l'évolution de la fonction objectif (résidus normalisés), nous constatons qu'au début des itérations (lorsque la température est élevée) le processus de recherche n'est pas assez sélectif vu que l'erreur évolue dans un grand intervalle de  $6,2 \cdot 10^{-3}$  à  $13 \cdot 10^{-3}$ . Au fur et à mesure que les itérations augmentent (la température baisse) le recuit devient plus sélectif et l'erreur baisse et se stabilise jusqu'à atteindre la valeur minimale  $\approx 6 \cdot 10^{-3}$  sur plus de 7000 itérations (7000 configurations de capteurs pour lesquelles la solution ne s'améliore pas).

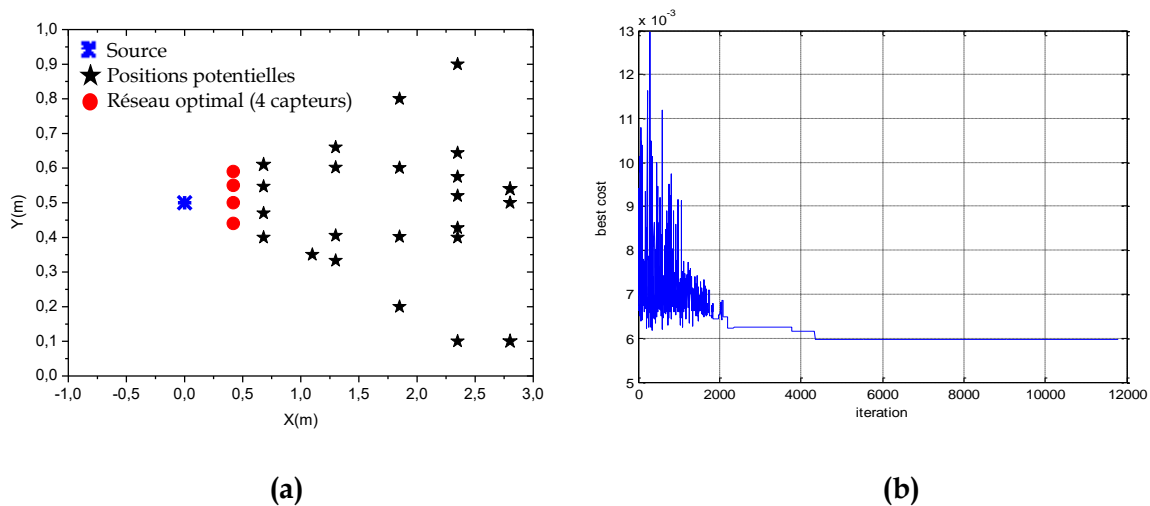


Figure 4.6. (a) Configuration optimale ( $m=4$ ) obtenue par le recuit simulé et (b) évolution de la fonction objectif

#### 4.2.2.2.2. Algorithme génétique

Pour l'algorithme génétique, la configuration optimale ( $m=4$ ) ainsi que l'erreur minimale par génération (fitness) sont représentées sur les figures 4.7.a et 4.7.b. On note, que le réseau optimal est le même que celui obtenu par le recuit simulé (voir figures 4.6.a et 4.7.a).

En ce qui concerne la fonction objectif, nous examinons l'erreur minimale (résidus normalisés) pour chaque génération (étoile) et le meilleur coût pour toutes les générations (trait). En effet, le meilleur coût de cette solution est identique à celui proposé par le recuit simulé (voir figures 4.6.b et 4.7.b).

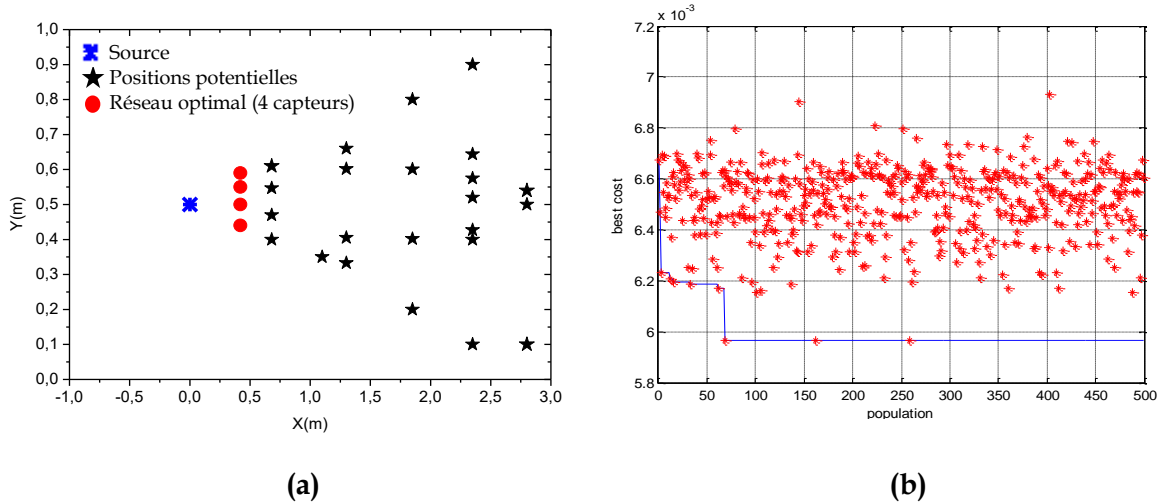


Figure 4.7. (a) Configuration optimale ( $m=4$ ) obtenue par l'algorithme génétique et (b) évolution de la fonction objectif « fitness »

#### 4.2.2.2.3. Recuit simulé VS algorithme génétique

Dans ce paragraphe, nous comparons les résultats obtenus par les deux algorithmes. Dans un premier temps, nous présentons une sélection de différentes configurations optimales ( $m=3, 6, 9$  et  $12$ ) (voir figure 4.8). Les réseaux optimaux, dont le nombre de capteurs varie de 3 à 13, ainsi que l'évolution de la fonction objectif spécifique à chaque réseau sont regroupées dans l'annexe IV.

Dans un second temps, afin d'effectuer une analyse plus fine, nous définissons une matrice dite « Matrice des Structures Communes » (MSC). Cette matrice montre le nombre de capteurs communs aux réseaux optimaux obtenus par les deux algorithmes.

Dans un troisième temps, pour effectuer une analyse quantitative, nous présentons les résultats de la caractérisation de la source connue ( $(x_s ; y_s) = (0 ; 0,5)$  et  $q_s=8,92.10^{-7}$  kg/s) par les réseaux optimaux (voir figures 4.9.a, 4.9.b, 4.9.c et 4.9.d). Ces résultats représentent l'erreur de localisation  $El$  en (m), exprimée sous la forme d'une distance euclidienne entre la position de la vraie source  $(x_s, y_s)$  et sa position estimée  $(x_0, y_0)$  et l'erreur de l'estimation du débit  $Eq$  exprimée sous la forme d'un rapport entre le débit de la vraie source  $q_s$  et le débit estimé  $q_0$  :  $El = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2}$  et  $Eq = q_0/q_s$ .

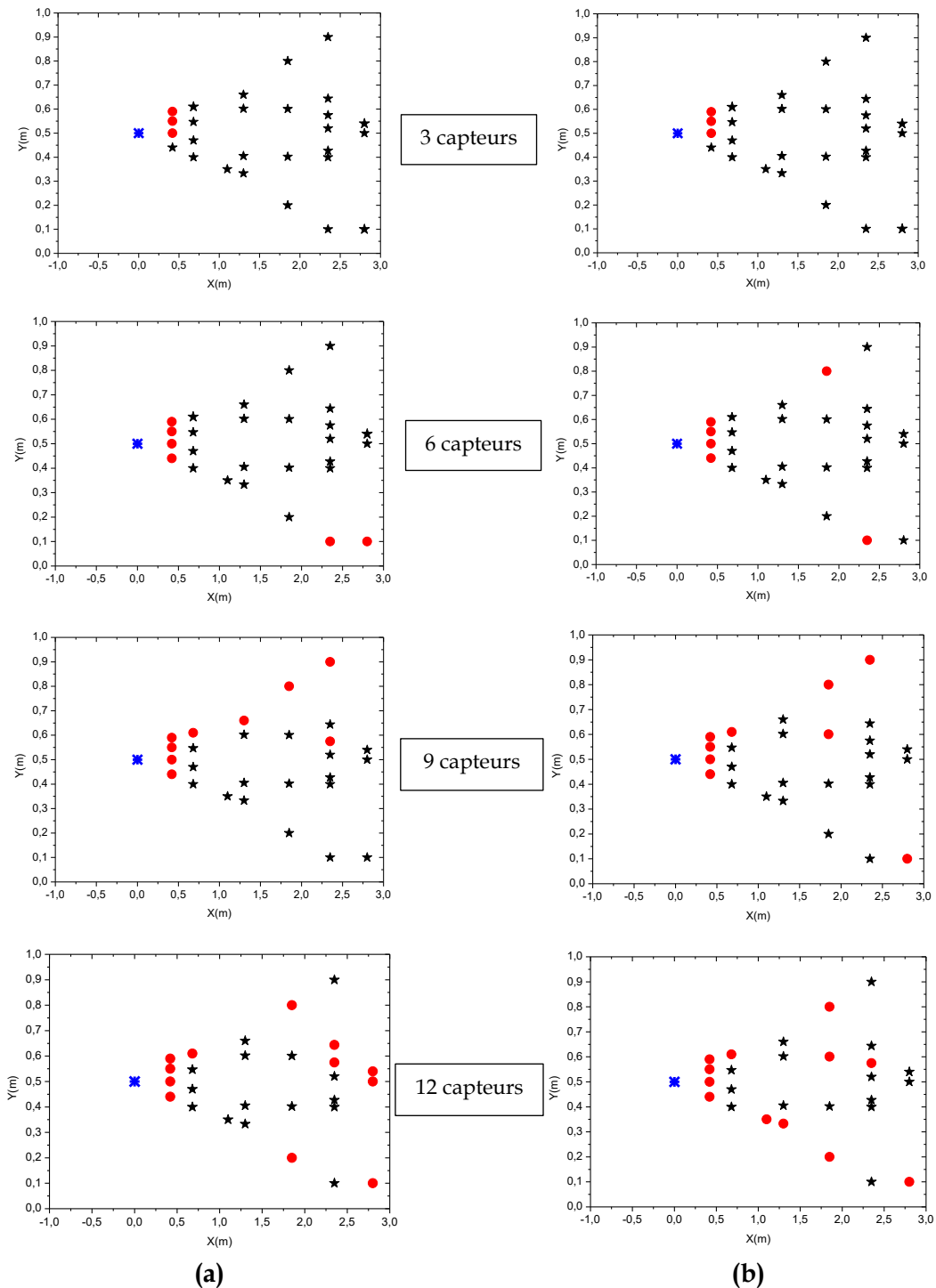


Figure 4.8. Configurations optimales ( $m=3, 6, 9$  et  $12$ ) obtenues par (a) le recuit simulé et par (b) l'algorithme génétique

Qualitativement, la MSC nous permet de constater que les réseaux obtenus par les deux algorithmes ont des structures semblables et sont de même qualité. Ceci se manifeste également sur les niveaux des fonctions coûts (voir annexe IV). Nous remarquons, également, que les capteurs les plus proches de la source ( $m=3$  et  $m=4$ ) sont communs à toutes les configurations optimales.

		Algorithme génétique											
		$m$	3	4	5	6	7	8	9	10	11	12	13
Recuit simulé	3	3	3	3	3	3	3	3	3	3	3	3	3
	4	3	4	4	4	4	4	4	4	4	4	4	4
	5	3	4	4	5	4	4	4	4	5	5	5	5
	6	3	4	4	5	5	6	5	6	5	5	5	5
	7	3	4	4	4	6	6	6	6	6	6	6	6
	8	3	4	4	4	6	6	7	6	7	6	7	7
	9	3	4	4	5	5	5	7	7	8	7	7	7
	10	3	4	4	5	5	5	8	6	9	8	7	7
	11	3	4	5	5	7	6	8	8	9	9	9	9
	12	3	4	5	5	7	7	7	8	10	9	9	9
	13	3	4	5	5	7	7	8	9	10	10	11	11

Tableau 4.3. Matrice des structures communes « problème 1 »

Quantitativement, nous remarquons que les erreurs induites par les solutions optimales, sont faibles :  $El \sim 0,014\ m$  et  $Eq \sim 1,03$ , ceci signifie que la source est reconstituée avec précision. Nous notons que la valeur de  $El$  est constante quel que soit le nombre de capteurs et que la valeur de  $Eq$  augmente faiblement (de l'ordre des millièmes) avec le nombre de capteurs. Ainsi pour cette expérience le nombre optimal de capteurs est  $m=3$  car au-delà, on n'améliore pas les capacités de caractérisation de la source.

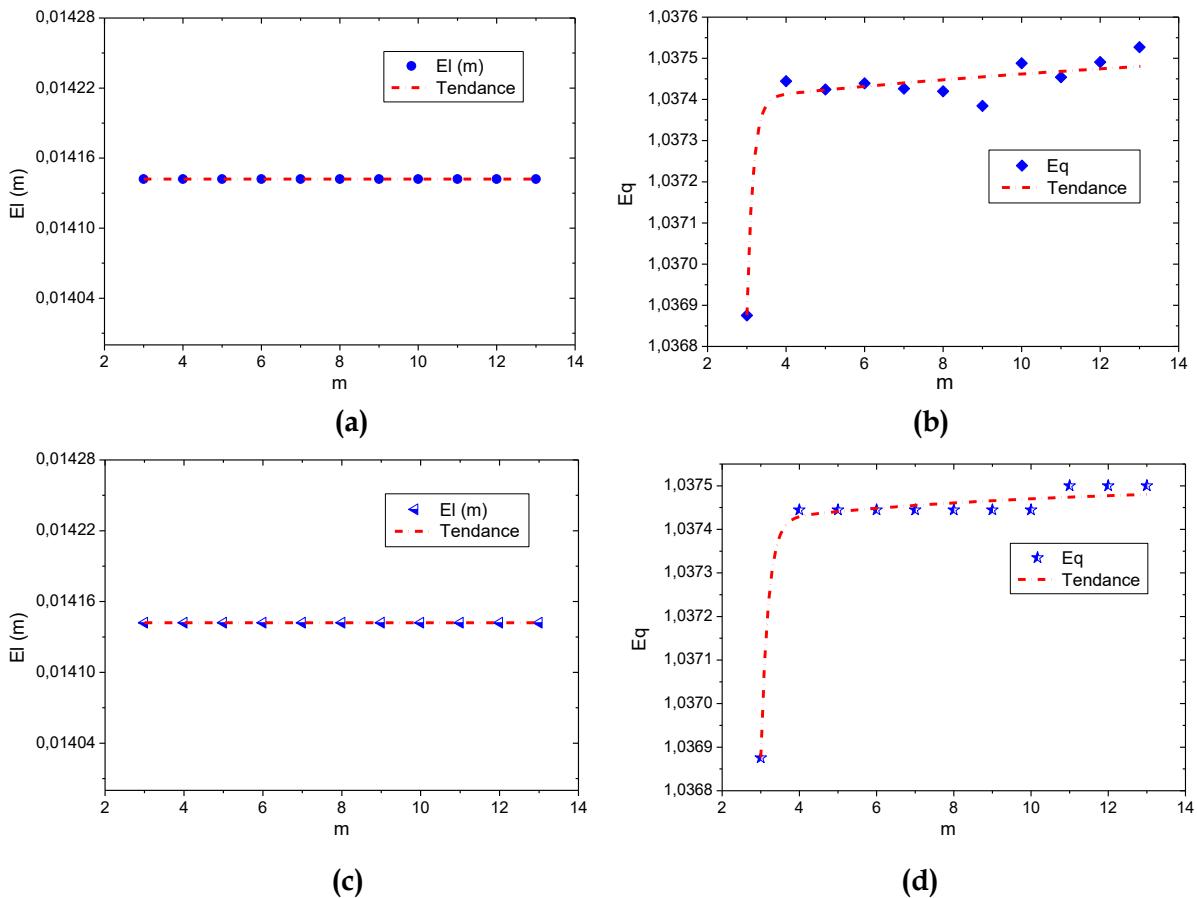


Figure 4.9. Erreurs d'estimations de (a) la position et (b) du débit de la source par le recuit simulé. Erreurs d'estimations de (c) la position et (d) du débit de la source par l'algorithme génétique.

Il est utile de noter que les résultats quantitatifs sont spécifiques à l'expérience DYCE car pour ce projet les paramètres de dispersion latérale et verticale sont ajustés spécifiquement à cette étude (voir 3.6.2), les mesures sont effectuées d'une façon précise et la vitesse de vent ainsi que la hauteur de la couche limite dans la soufflerie sont déterminées de façon à obtenir les meilleurs résultats possibles.

#### 4.2.2.3. Conclusion à l'issue du problème 1

Les différents résultats obtenus dans le cadre du « problème 1 » de l'approche analytique ont montré que l'on peut déterminer la taille optimale d'un réseau en vue de la caractérisation précise de la source sur des terrains supposés plats et faiblement rugueux. Cette démarche a consisté à évaluer les performances de différentes configurations optimales, le nombre optimal est ainsi déterminé pour le réseau produisant la meilleure estimation des paramètres de la source.

L'analyse de la MSC a montré que si un réseau optimal est inclus dans un réseau de taille plus importante cela ne conduit pas nécessairement à une amélioration de la performance du réseau résultant. Cette constatation peut être affinée par une analyse de sensibilité et d'incertitude dans l'estimation des paramètres de la source.

Pour les expériences considérées, on note que le minimum de la fonction coût (résidus normalisés) correspond systématiquement à une meilleure caractérisation de la source. On note également que pour un réseau optimal, une estimation précise de la position et du débit de la source sont simultanément obtenues.

En ce qui concerne l'efficacité des algorithmes d'optimisation (évolutionnaire et probabiliste), dont les techniques de recherche sont différentes, nous pouvons conclure qu'ils proposent des réseaux optimaux de même qualité pour des domaines de recherche assez large. Par conséquent, pour la suite de nos calculs nous utilisons uniquement l'algorithme du recuit simulé. Ce choix est également motivé par les difficultés de mise en œuvre de l'algorithme génétique lorsque le nombre de capteurs est faible (par exemple pour  $m=4$  et  $m=5$ ). Ainsi, lorsque les parents sont sélectionnés, l'opération de croisement est souvent bloquée car chaque descendant (enfant) doit avoir des gènes différents. Comme les parents sont sélectionnés par rapport à leur qualité (sélection par rang) ils sont souvent de structure semblable, donc leur croisement est irréalisable (pas assez d'information génétique à transmettre).

### 4.2.3. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source »

Pour ce problème, les données météorologiques sont fixes et la source est supposée inconnue et peut se situer n'importe où dans le domaine. Le réseau optimal dans ce cas, doit assurer une surveillance efficace de tout le domaine et doit permettre la caractérisation de la source (position et débit). Le critère adapté à ce type de scénario est l'entropie (voir partie 3.7.2).

L'entropie du réseau évalue la quantité d'information fournie par le réseau pour estimer une source inconnue. Pour cette configuration nous n'avons aucune information a priori sur la source donc le meilleur réseau (formé de  $m$  capteurs) est celui qui contient une quantité d'information maximale. Nous rappelons que l'expression de la fonction objectif, à maximiser, est de la forme :

$$Jp_2 = \frac{1}{2} \log(\det(\mathbf{H}_\varphi))$$

#### 4.2.3.1. Paramétrage

Pour ce problème le nombre de combinaisons possibles reste inchangé et les fonctions coûts sont différentes de celles du « problème 1 » (les résidus normalisés). Ceci suscite un paramétrage spécifique du recuit simulé. Les paramètres retenus sont regroupés dans le tableau 4.3.

$m$	$T_0$	$T_f$	$L$	$\theta$
3	10	1.10 <sup>-6</sup>	50	0,9
4	10	1.10 <sup>-6</sup>	50	0,9
5	10	1.10 <sup>-6</sup>	50	0,9
6	10	1.10 <sup>-6</sup>	70	0,9
7	10	1.10 <sup>-6</sup>	80	0,9
8	10	1.10 <sup>-6</sup>	100	0,9
9	10	1.10 <sup>-6</sup>	100	0,9
10	10	1.10 <sup>-7</sup>	100	0,9
11	10	1.10 <sup>-7</sup>	120	0,9
12	10	1.10 <sup>-7</sup>	120	0,9
13	10	1.10 <sup>-7</sup>	120	0,9

Tableau 4.4. Paramétrage du recuit simulé « problème 2 »

#### 4.2.3.2. Résultats et discussions

Tout d'abord, nous déterminons les configurations optimales composées de 3 à 13 capteurs par le recuit simulé. Une fois les différents réseaux obtenus, nous examinons a posteriori



l'erreur commise dans l'estimation de la source connue ( $x_s = 0\text{ m}$ ,  $y_s = 0,5\text{ m}$  et  $q_s = 8,925 \cdot 10^{-7} \text{ kg/s}$ ) pour évaluer leur efficacité.

Nous présentons sur la figure 4.10.a la configuration optimale de 4 capteurs et sur la figure 4.10.b l'évolution de la fonction objectif. La totalité des configurations est regroupée en annexe IV.

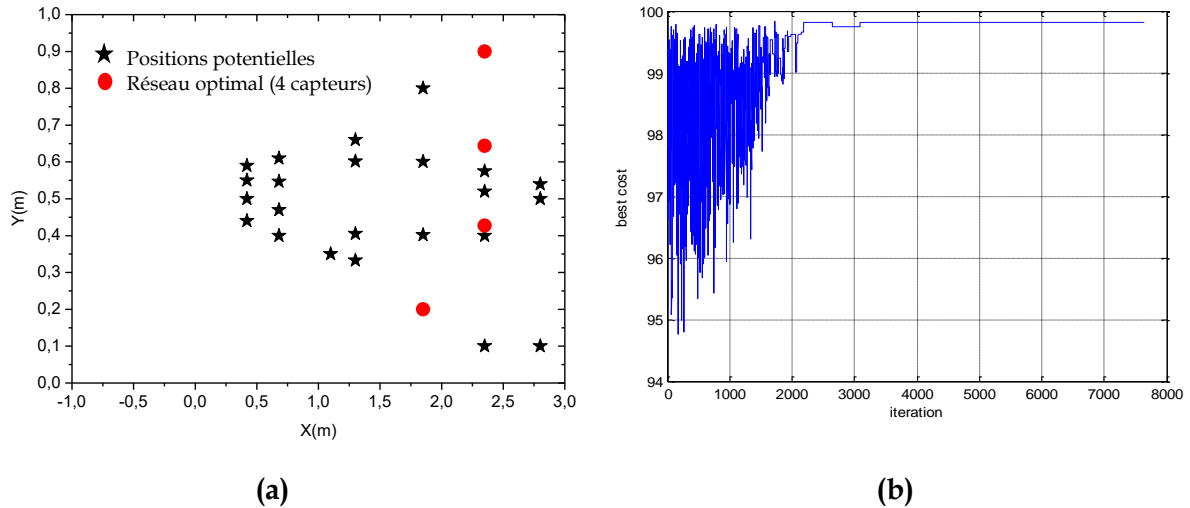


Figure 4.10. (a) Configuration optimale ( $m=4$ ) et (b) évolution de la fonction objectif

Nous présentons également, sur la figure 4.11 un échantillon des réseaux optimaux ( $m=3,5,9$  et 13) obtenus par le recuit simulé.

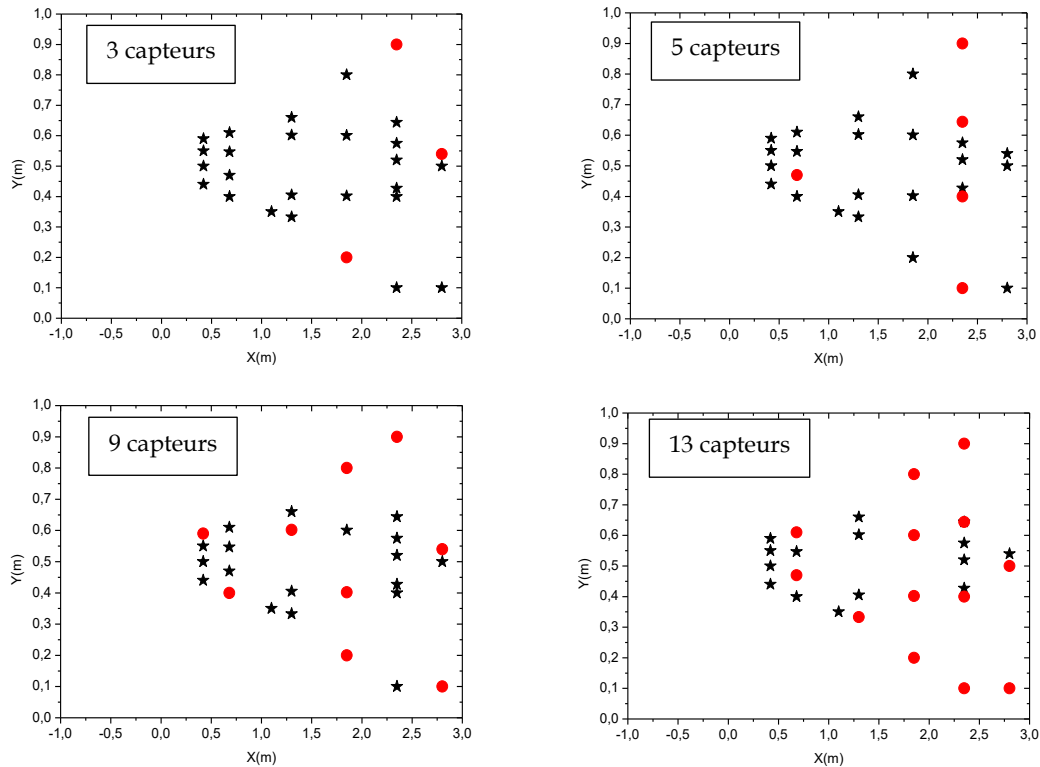


Figure 4.11. Réseaux optimaux ( $m=3, 5, 9$  et 13) obtenus par le recuit simulé

Qualitativement, nous notons que les différentes configurations optimales sont dispersées dans l'espace formé par les positions potentielles (contrairement aux réseaux optimaux du « problème 1 »). Ainsi les réseaux retenus ont tendance à couvrir l'ensemble du domaine.

Pour la suite, à partir des différentes configurations obtenues et des données expérimentales nous évaluons les erreurs produites dans la reconstitution de la source connue. Les résultats sont présentés sur la figure 4.12.

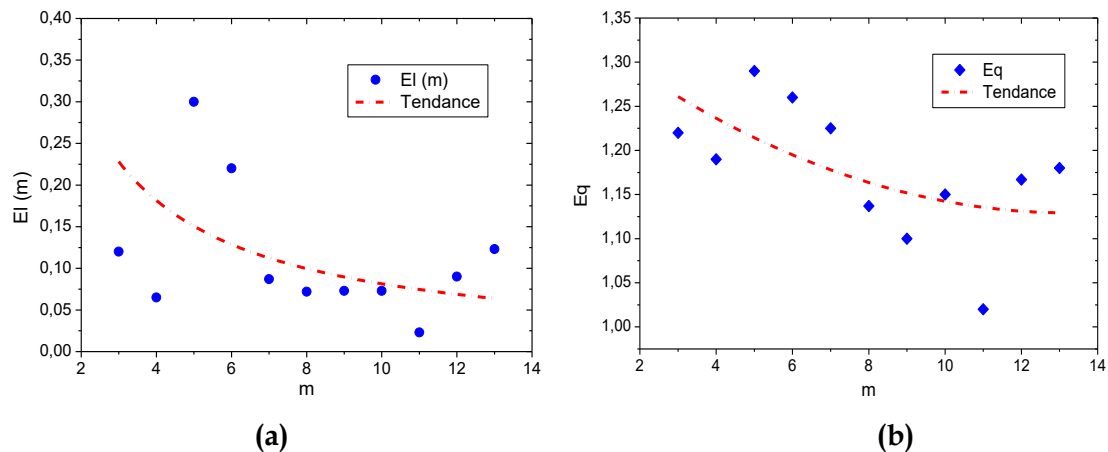


Figure 4.12. Erreurs d'estimation de (a) la position et (b) du débit de la source

Pour l'ensemble des configurations optimales on observe des résultats assez homogènes pour l'estimation du débit et de la position de la source. On note globalement une surestimation du débit,  $Eq$  varie ainsi de 1 à 1,3. Pour ce qui concerne les erreurs de localisation, la situation est plus contrastée, ainsi  $El$  varie de 0,025 m à 0,3 m. A l'exception des réseaux de tailles  $m=5$  et  $m=6$ , les performances des différents réseaux sont semblables. Toutefois pour cette expérience le meilleur réseau est de taille  $m = 11$ . Il est à noter que pour cette expérience, les réseaux basés sur l'entropie sont moins performants que ceux basés sur les résidus normalisés (problème 1), ce qui est logique.

Pour une analyse plus fine nous dressons la MSC de toutes les configurations testées (voir ci-dessous). Cette dernière montre que les réseaux optimaux avec  $m=8, 9, 10$  et  $12$  incluent le réseau de taille  $m=3$ . Il en résulte une amélioration de la caractérisation de la source, donc pour ces configurations, l'information complémentaire apportée au réseau ( $m=3$ ) améliore le processus d'inversion. On note également que le réseau de taille  $m=4$  est inclus dans les réseaux de taille  $m=6$  et  $m=7$ . Pour ces configurations, l'information complémentaire apportée au réseau ( $m=4$ ) détériore la caractérisation de la source. Cette dernière analyse montre que l'information rajoutée par l'instrumentation de nouvelles positions n'est pas

toujours utile à la reconstitution de la source, ce qui montre que le problème d'optimisation concerne à la fois le nombre et les positions des capteurs.

<i>m</i>	3	4	5	6	7	8	9	10	11	12	13
3	3	2	1	2	2	3	3	3	2	3	2
4	-	4	2	4	4	2	2	3	3	3	3
5	-	-	5	2	2	2	1	2	3	4	5
6	-	-	-	6	6	3	3	4	4	5	4
7	-	-	-	-	7	4	3	5	3	5	4
8	-	-	-	-	-	8	5	5	5	5	5
9	-	-	-	-	-	-	9	7	5	6	5
10	-	-	-	-	-	-	-	10	5	9	6
11	-	-	-	-	-	-	-	-	11	9	9
12	-	-	-	-	-	-	-	-	-	12	10
13	-	-	-	-	-	-	-	-	-	-	13

Tableau 4.5. Matrice des structures communes

#### 4.2.3.3. Conclusion à l'issue du problème 2

Les différents résultats obtenus dans le cadre du « problème 2 » de l'approche analytique ont montré que l'on peut déterminer des réseaux optimaux sur la base de l'information contenue dans chacun d'entre eux. En effet, l'entropie maximisée produit des réseaux dont la structure est dispersée sur tout le domaine. Ainsi, la dispersion spatiale des capteurs permet de surveiller des zones larges.

L'analyse de la MSC a montré que les réseaux optimaux ont de fortes ressemblances. Nous avons également noté que si on ajoute des capteurs à une structure de base d'un réseau optimal cela ne conduit pas nécessairement à des réseaux plus grands ayant de bonnes performances. Dans certains cas l'information ajoutée par les positions complémentaires améliore la précision du réseau et dans d'autres cas elle la dégrade.

Pour les expériences considérées, on a noté que pour un réseau optimal basé sur le concept d'entropie, l'estimation précise de la position et du débit de la source est obtenue simultanément. Egalement, on a noté que la fonction coût entropique est moins performante que celle basée sur les résidus normalisés surtout pour les réseaux de petite taille.

### 4.3. APPROCHE CFD « EXPERIENCE MUST »

En milieu urbain, les modèles de dispersion de type Gaussiens ne sont pas efficaces car ils ne prennent pas en compte les obstacles. Une approche de simulation plus réaliste est indispensable pour analyser toute la complexité de la dispersion des polluants ainsi que l'estimation de source. Dans cette partie, les champs adjoints sont modélisés par une

approche CFD. Pour évaluer notre processus d'optimisation nous nous basons sur les données de l'expérience MUST « MOCK URBAN STETTING TEST » (Biltoft, 2001 ; Yee, E, et Biltoft, 2004). C'est une expérience de dispersion du propylène  $C_3H_6$  à échelle réduite organisée par la Defense Threat Reduction Agency (DTRA) en septembre 2001. L'objectif de cette expérience était de mieux caractériser le transport et la dispersion des panaches autour des bâtiments. Ainsi, les campagnes expérimentales ont été réalisées sur un terrain bâti idéalisé, présenté sur la figure 4.13.



Figures 4.13. Photographies de l'expérience de MUST (Mock Urban Setting Test) (Biltoft, 2001)

### 4.3.1. Description des expériences

Les expériences ont été menées sur un domaine d'environ  $200 \times 200 m^2$ . La canopée urbaine est représentée par un quadrillage de containers. Ces derniers sont disposés selon 10 lignes et 12 colonnes sur un terrain de l'armée dans le désert de l'Utah aux États-Unis. Chaque caisson métallique a pour dimensions 2,54 m de haut, 12,2 m de long et 2,42 m de large. L'espacement entre les lignes est de 12,9 m, alors que les colonnes sont séparées par une distance de 7,9 m. La figure 4.14 représente la répartition des obstacles dans le domaine et l'emplacement des sources pour les essais réalisés. L'expérience consiste en 63 rejets d'un gaz non réactif (propylène  $C_3H_6$ , dont la masse volumique est de  $1,75 kg.m^{-3}$ ) pour une température de  $20^\circ C$ . Différentes conditions de vent (angle, vitesse) ainsi que différentes positions pour les rejets de gaz (à l'intérieur ou à l'extérieur de la canopée, à différentes hauteurs) ont été considérées. Ces émissions de gaz ont été effectuées dans des conditions de stratifications stable et neutre à l'extérieur de la canopée. Chaque expérience de rejet dure 15 minutes, période pendant laquelle des mesures de concentration sont effectuées. Ces mesures sont réalisées par 40 détecteurs à photo-ionisation (PIDs). Ces capteurs ont été positionnés sur quatre lignes horizontales à 1,6 m de hauteur. En outre, 8 détecteurs supplémentaires sont placés sur un mât central à diverses altitudes ( $z=1, 2, 4, 6, 8, 10, 12, \text{ et } 16 m$ ). Enfin, 8 anémomètres ont été répartis sur quatre mâts à différentes altitudes pour caractériser la turbulence du vent.

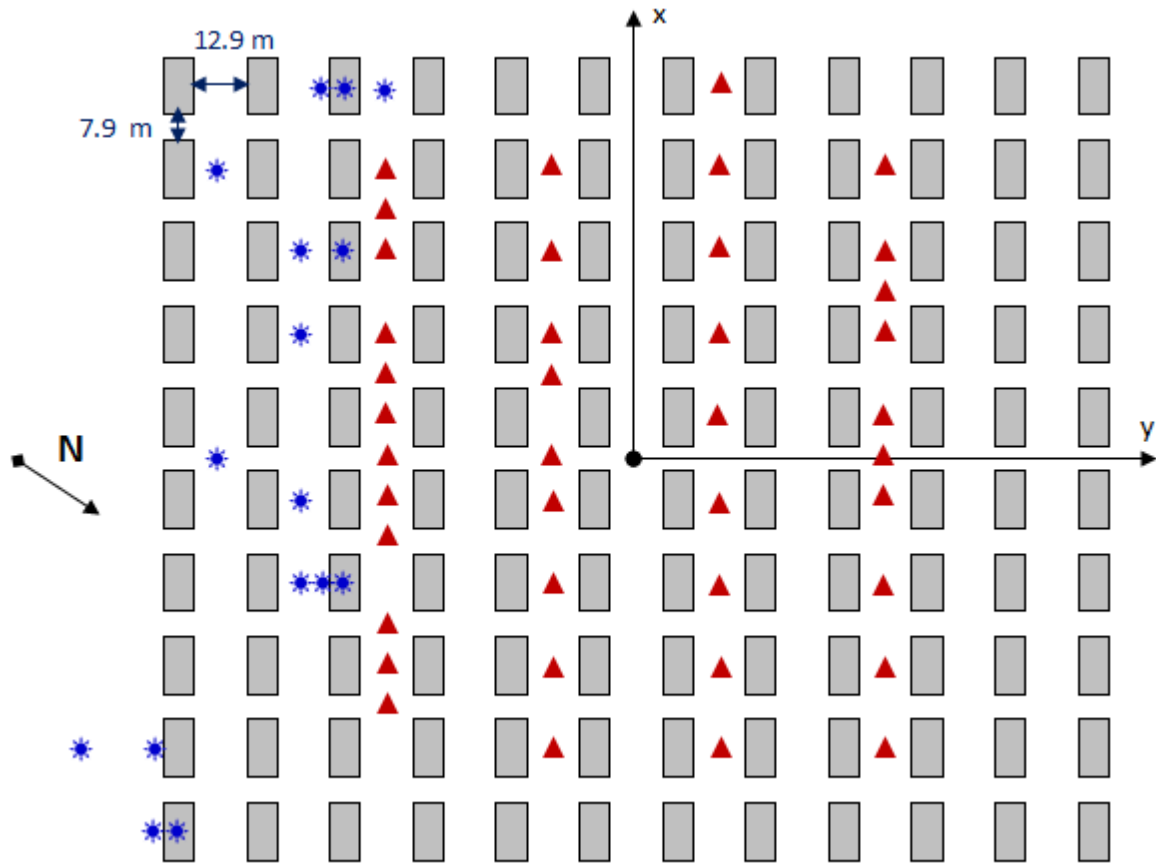


Figure 4.14. Schéma de l'expérience de MUST avec emplacement des 40 capteurs (triangle), des caissons métalliques (rectangle) et des sources (étoile) dans le domaine d'étude.

Pour évaluer notre processus d'optimisation nous utilisons les données issues de 4 expériences. Les caractéristiques de ces expériences sont regroupées dans le tableau 4.4.

	Expérience 1 « Essai 6 »	Expérience 2 « Essai 8 »	Expérience 3 « Essai 14 »	Expérience 4 « Essai 19 »
$x_s (m)$	-60,820	-54,713	-53,836	-84,852
$y_s (m)$	25,647	-36,181	89,740	-72,863
$z_s (m)$	0,15	1,8	2,6	1,3
$q_s (l/min)$	200	200	225	225
$U (m/s)$	1,89	2,68	5,02	2,98
$\Theta (degré)$	-10°, Nord-Nord-Est	30°, Nord-Ouest	-39°, Nord-Est	43°, Nord-Est
<b>Atmosphère</b>	très stable $0 < L_{MO} < 10 m$	modérément stable $100 < L_{MO} < 200 m$	faiblement stable $200 < L_{MO} < 1000 m$	modérément stable $100 < L_{MO} < 200 m$
$L_{MO}$	7,7	150	240	130
$u^*$	0,16	0,35	0,66	0,39
$Z_{capteurs} (m)$	1,60	1,60	1,60	1,60

Tableau 4.6. Caractéristiques des expériences sélectionnées

A partir du tableau ci-dessus, on peut constater que les sources considérées sont localisées à des positions et des hauteurs différentes. On constate aussi que les expériences sont réalisées pour des vitesses de vents comprises entre 1,89 m/s et 5,02 m/s. Ainsi, les conditions

météorologiques sont également variées. En revanche, les capteurs sont à des hauteurs fixes. On rappelle que  $L_{MO}$  correspond à la longueur de Monin-Obukov.

### 4.3.2. Modélisation inverse de (Kumar et al, 2015)

Sur la base des données expérimentales de l'expérience MUST, une étude a été réalisée par (Kumar et al, 2015) dans le cadre du projet DISCARD (projet de recherche mené en collaboration entre le laboratoire LMEE, la société Fluidyn et la Direction Générale de l'Armement DGA). Ainsi, des simulations directes et inverses ont été effectuées pour 20 cas tests. Cette étude a été la première à évaluer la méthode de renormalisation dans les milieux urbains. La modélisation de la dispersion directe ainsi que la génération des rétropanaches pour chaque essai ont été réalisées par le modèle CFD Fluidyn-PANACHE (Fluidyn-PANACHE, 2010). Celui-ci est spécifiquement dédié à l'étude des écoulements atmosphériques et de la dispersion de polluants (gazeux ou particulaires) sur les sites industriels ou urbains. Le mouvement de l'air est simulé par la résolution des équations tridimensionnelles de Navier-Stokes en régime turbulent qui reposent sur une méthode de volumes finis avec un schéma centré de second ordre. L'évolution de la concentration des espèces rejetées est obtenue par la résolution de l'équation d'advection-diffusion. Le modèle de turbulence sélectionné est le modèle  $k-\varepsilon$  standard (voir annexe III). Les rétro-panaches sont générés en utilisant un modèle d'approximation par discrétisation de l'équation 3D d'advection-diffusion adjointe.

La discrétisation spatiale est réalisée dans deux domaines de calcul imbriqués. Ainsi, un sous-domaine de dimensions  $250 \times 225 \text{ m}^2$ , avec une hauteur de  $100 \text{ m}$  a été créé afin de positionner tous les containers comme dans l'expérience. Ce sous-domaine est contenu dans un domaine de calcul plus étendu, dont les dimensions sont  $800 \times 800 \text{ m}^2$  avec une hauteur de  $400 \text{ m}$  pour assurer une transition douce de l'écoulement entre les bords du domaine et la zone d'obstacles. Pour l'ensemble, un maillage non-structuré a été généré, avec un raffinement au niveau de la zone urbanisée (le sous-domaine) (voir figure 4.15). Ainsi, 2849276 mailles ont permis une description suffisante des phénomènes physiques avec des temps de calcul raisonnables. Pour le calcul des rétropanaches issus des 40 capteurs pour chaque essai de l'expérience MUST, les conditions aux limites en entrée sont celles avec un profil de vent de (Gryning et al, 2007), sauf pour l'expérience 1 (Essai 6), dont le profil de vent est celui de (Beljaars and Holtslag, 1991) (le profil d'entrée de (Gryning et al, 2007) est non applicable pour des situations atmosphérique très stable) (voir annexe III). L'air est modélisé



comme de l'air humide avec des propriétés de mélange d'air sec et de vapeur d'eau. Le modèle thermodynamique du mélange de gaz suit une loi idéale (Ideal Gas Law en anglais).

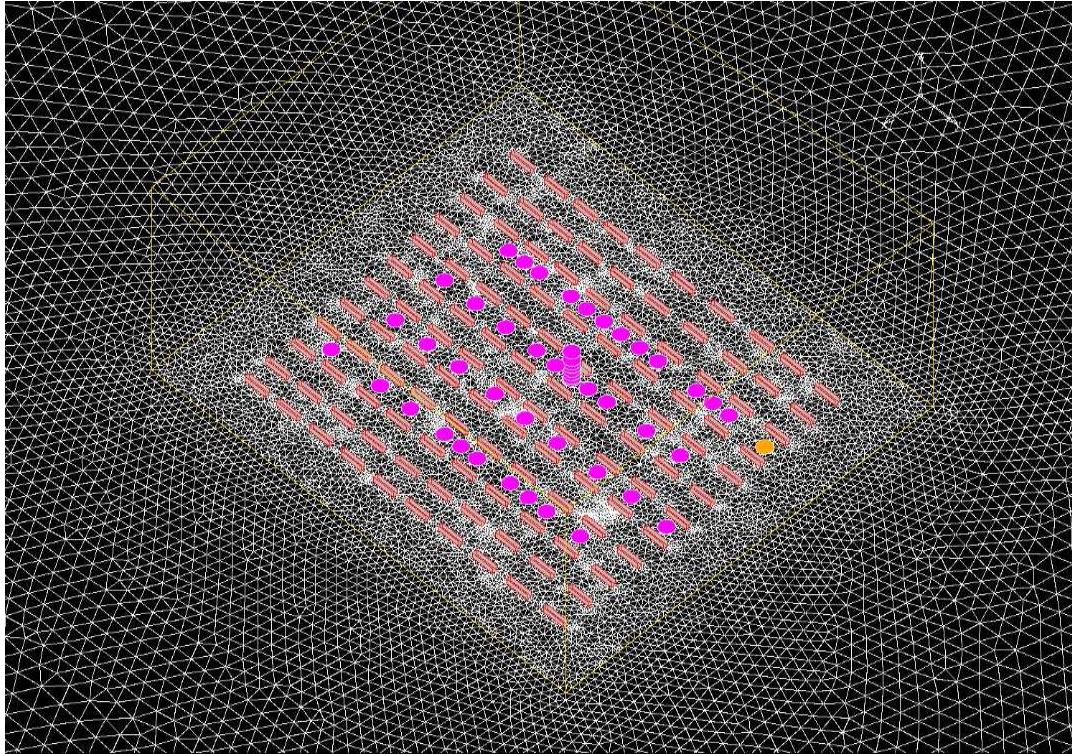


Figure 4.15. Maillage des deux domaines pour les calculs CFD (DISCARD, 2015)

Comme indiqué sur le tableau 4.5 chaque source correspond à un essai donné. Le nombre total d'essais dans cette étude est de 20. Le nombre de mailles est le même pour tous les essais et les calculs ont été réalisés en régime stationnaire.

Les résultats de la reconstitution des 20 sources par la méthode d'assimilation de données renormalisées ont été obtenus par (Kumar et al, 2015). Nous nous baserons sur ces résultats pour évaluer la performance des réseaux optimaux. Le tableau 4.5 regroupe les erreurs  $E_q$  et  $E_l$  commises dans l'estimation du débit et de la position de la source :  $E_q = q_0/q_s$  et  $E_l = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2}$ .

Pour chaque expérience, deux réseaux ont été évalués, le réseau formé de 40 capteurs nommé « All Measurements » et le réseau comportant les capteurs dont la concentration mesurée est différente de zéro nommé « Nonzero Measurements ». Ce dernier est considéré dans le but d'examiner l'effet des mesures nulles. (Kumar et al, 2015) ont montré que l'exclusion des capteurs dont la concentration est nulle, a un effet négligeable sur l'estimation de la source. Ils ajoutent qu'il est néanmoins important de les maintenir pour réduire l'incertitude lors de la reconstitution de la source. Ainsi, pour les deux réseaux considérés,  $m$  représente le nombre de capteurs (voir tableau 4.7).

No. Essai	Nom Essai (JJJhhmm)	« All Measurements »			« Nonzero Measurements »		
		$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$
1	2640138	40	3,3	0,92	34	3,3	0,92
2	2640246	40	42,9	4,01	28	42,9	4,01
3	2671852	40	10,8	1,17	30	10,8	1,18
4	2671934	40	22,8	0,27	31	22,8	0,27
5	2672033	40	21,9	0,57	37	21,9	0,57
6	2672101	40	5,0	2,14	40	5,0	2,14
7	2672150	40	12,4	0,41	30	5,4	1,32
8	2672213	40	15,8	2,22	30	13,5	0,39
9	2672235	40	7,7	1,37	31	5,6	1,53
10	2672303	40	8,8	1,08	40	8,8	1,08
11	2681829	40	19,8	1,67	30	19,8	1,67
12	2681849	40	7,4	0,95	27	7,4	0,94
13	2682256	40	7,7	0,97	33	7,7	0,97
14	2682320	40	2,2	1,42	29	2,2	1,42
15	2682353	40	1,1	1,88	25	2,7	1,91
16	2692054	40	26,7	1,70	26	26,7	1,72
17	2692131	40	7,0	0,90	31	7,0	0,90
18	2692157	40	14,3	1,15	33	17,5	1,01
19	2692223	40	22,3	1,76	35	22,3	1,76
20	2692250	40	32,5	0,83	31	32,1	0,81

Tableau 4.7. Résultats obtenus dans le cadre des travaux de (Kumar et al, 2015) : Reconstruction de la source pour les 20 expériences de l'expérience de MUST.

#### 4.3.2.1. Vérification de la simulation inverse

L'objectif de cette partie est de vérifier que notre code d'inversion pour la caractérisation de la source (décrit dans la partie 3.5.2.3.2) fonctionne convenablement dans le cadre d'une approche CFD. Ainsi, les éléments de volume sont intégrés dans le calcul de la matrice de Gram pondérée  $H_w$ . Pour ce faire, nous avons récupéré les données du maillage (tailles des mailles) et les 40 rétropanaches pour les 4 expériences choisies parmi les 20 essais (tableau 4.4). La vérification consiste donc à reconstituer les caractéristiques de chaque source et de comparer nos résultats à ceux obtenus par le code d'inversion de (Kumar et al, 2015).

No. Essai	Nom Essai (JJJhhmm)	« All Measurements »			« Résultats de nos simulations »		
		$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$
6	2672101	40	5,0	2,14	40	5,0	2,14
8	2672213	40	15,8	2,22	40	15,8	2,22
14	2682320	40	2,2	1,42	40	2,2	1,42
19	2692223	40	22,3	1,76	40	22,3	1,76

Tableau 4.8. « All Measurements » (Kumar et al, 2015), VS « Résultats de nos simulations »

Nous pouvons constater à partir du tableau 4.8 que les résultats de nos estimations sont strictement identiques à ceux obtenus par (Kumar et al, 2015) pour le réseau «All Measurements ». En conséquence, nous considérons avoir validé notre algorithme d'inversion dans le cadre de cette approche CFD.



### 4.3.3. Définition du problème d'optimisation

Dans cette partie, la région d'intérêt  $\Omega$  contient 40 positions potentielles. L'ensemble  $G_I$  de ces positions a été défini par l'expérimentateur.

$$G_I = \{x_1, x_2, \dots, x_{40}\} \subset \Omega$$

Parmi l'ensemble  $P$ , nous sommes amenés à sélectionner les meilleures  $m$  positions pour implémenter les capteurs. Le meilleur sous ensemble est le réseau optimal  $R_o$  défini par

$$R_o = \{x_1^o, x_2^o, \dots, x_m^o\} \subset G_I$$

#### 4.3.3.1. Description des calculs

##### 4.3.3.1.1. Echelle de complexité

Dans ce cas, nous cherchons parmi 40 positions, les réseaux optimaux de taille  $m$  variant de 4 à 16 capteurs. Ainsi, le nombre de combinaisons possibles varie de  $C_{27}^4$  à  $C_{27}^{16}$ . La figure 4.16 représente l'évolution du nombre de combinaisons en fonction du nombre de capteurs. Dans ce cas l'espace de recherche à une taille de départ de  $91.10^3$  et atteint  $6.10^{10}$ .

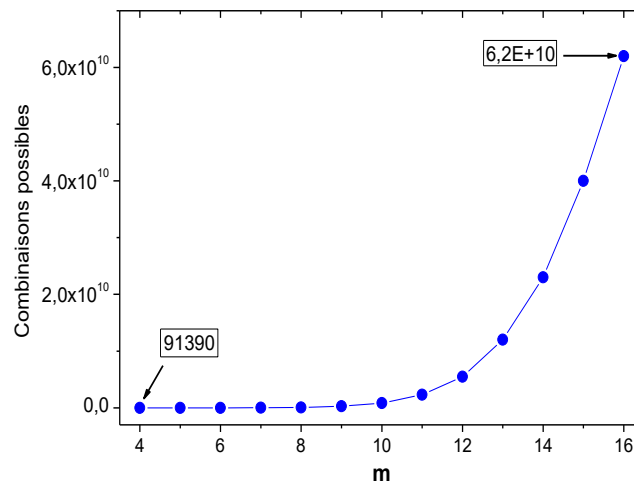


Figure 4.16. Evolution du nombre des combinaisons en fonction du nombre de capteurs

##### 4.3.3.1.2. Couplage entre l'algorithme d'optimisation l'algorithme de renormalisation et les « champs adjoints CFD »

Nous expliquons ici le principe de couplage entre l'algorithme du recuit simulé, l'algorithme de renormalisation et les « champs adjoint CFD ». Dans ce cas, le calcul des 40 champs adjoints (retropanaches) ne fait pas partie de la boucle d'optimisation, mais il est fait plutôt en « offline » par (Kumar et al, 2015). Le principe de couplage est représenté sur la figure 4.17.

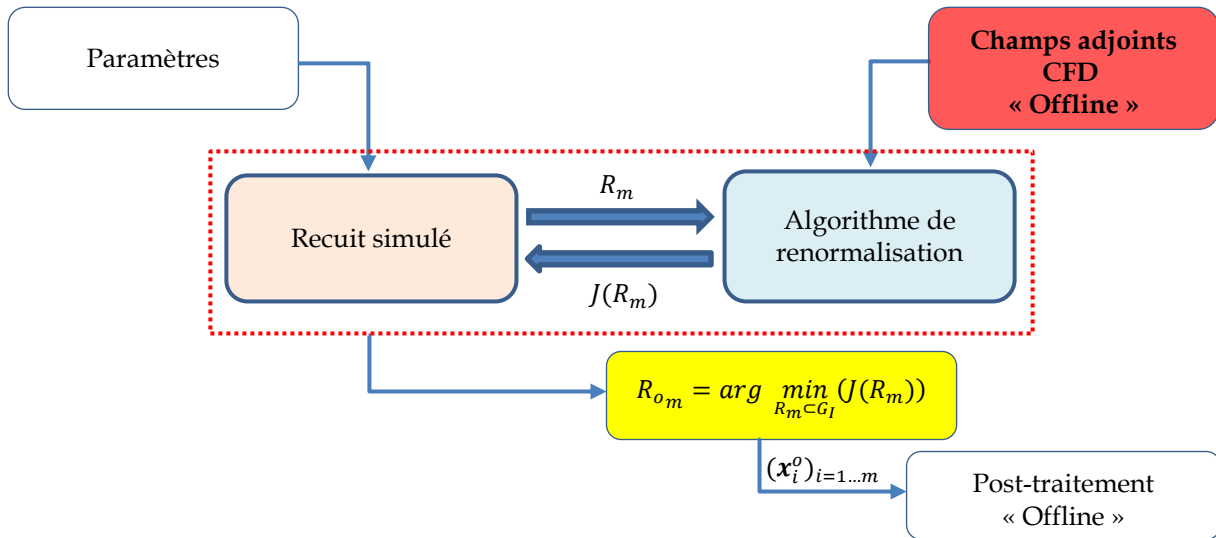


Figure 4.17. Couplage entre le recuit simulé, l’algorithme de renormalisation et les champs adjoints CFD

### 4.3.4. Problème 1 « Détermination du nombre optimal de capteurs pour la caractérisation de source »

Pour des conditions météorologiques fixes (direction et vitesse de vent), la source est supposée inconnue (position et débit). L’objectif est, dans ce cas, de déterminer le nombre minimal de capteurs qui permet de caractériser la source. Pour obtenir le réseau optimal  $R_{om}$ , nous utilisons la fonction objectif  $Jp_1$  (résidus normalisés) dont nous rappelons l’expression

$$Jp_1(R_m) = 1 - \frac{s_{//}^2(\mathbf{x}_0)}{\|\boldsymbol{\mu}\|_{H_w^{-1}}^2}$$

#### 4.3.4.1. Paramétrage

L’objectif de cette évaluation est de choisir  $m = 4, \dots, 16$  capteurs parmi 40 positions potentielles. Le paramétrage retenu pour la détermination des réseaux optimaux est regroupé dans le tableau 4.7.

$m$	$T_0$	$T_f$	$L$	$\theta$
4 à 10	10	$1.10^{-13}$	100	0,9
11 à 16	10	$1.10^{-13}$	200	0,9

Tableau 4.9. Paramétrage recuit simulé « problème 1 »

#### 4.3.4.2. Résultats et discussions

Pour le « problème 1 » nous examinons qualitativement les solutions obtenues par le recuit simulé, ensuite nous évaluons quantitativement leurs performances à partir des erreurs dans la caractérisation des sources.

- **Expérience 1, « Essai 6 » : atmosphère très stable.**

Nous présentons sur la figure 4.18 deux réseaux optimaux ( $m=10$  et  $m=13$ ). La totalité des solutions (de 4 à 16 capteurs) est regroupée en annexe IV.

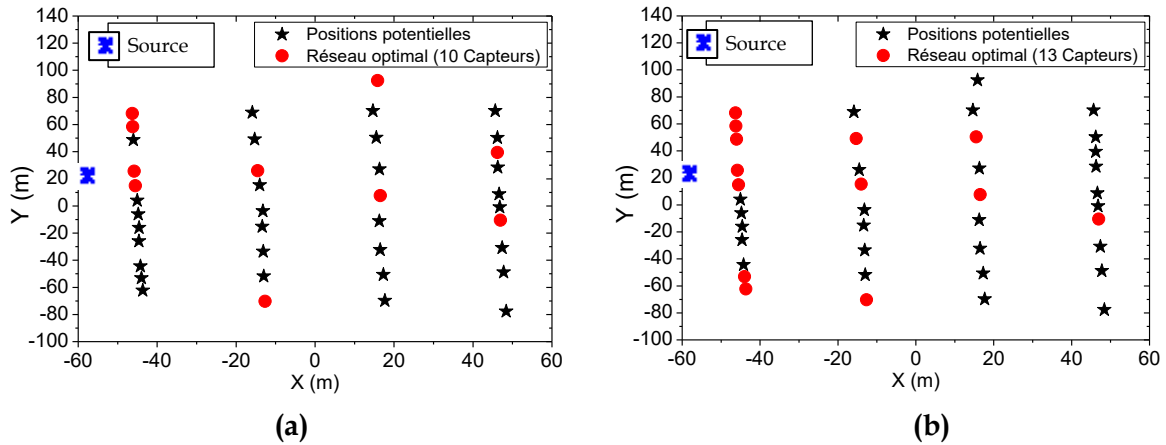


Figure 4.18. Configurations optimales (a)  $m=10$  et (b)  $m=13$  « Essai 6 »

Sur la figure 4.19 les cartes représentant la visibilité  $\varphi(x)$  et l'estimation de la source  $s_{//\varphi}(x)$  (disques noir et blanc: position de la vraie source et sa position estimée). Les champs de visibilité délimitent, en pratique, les zones où seront opérées les recherches de source.

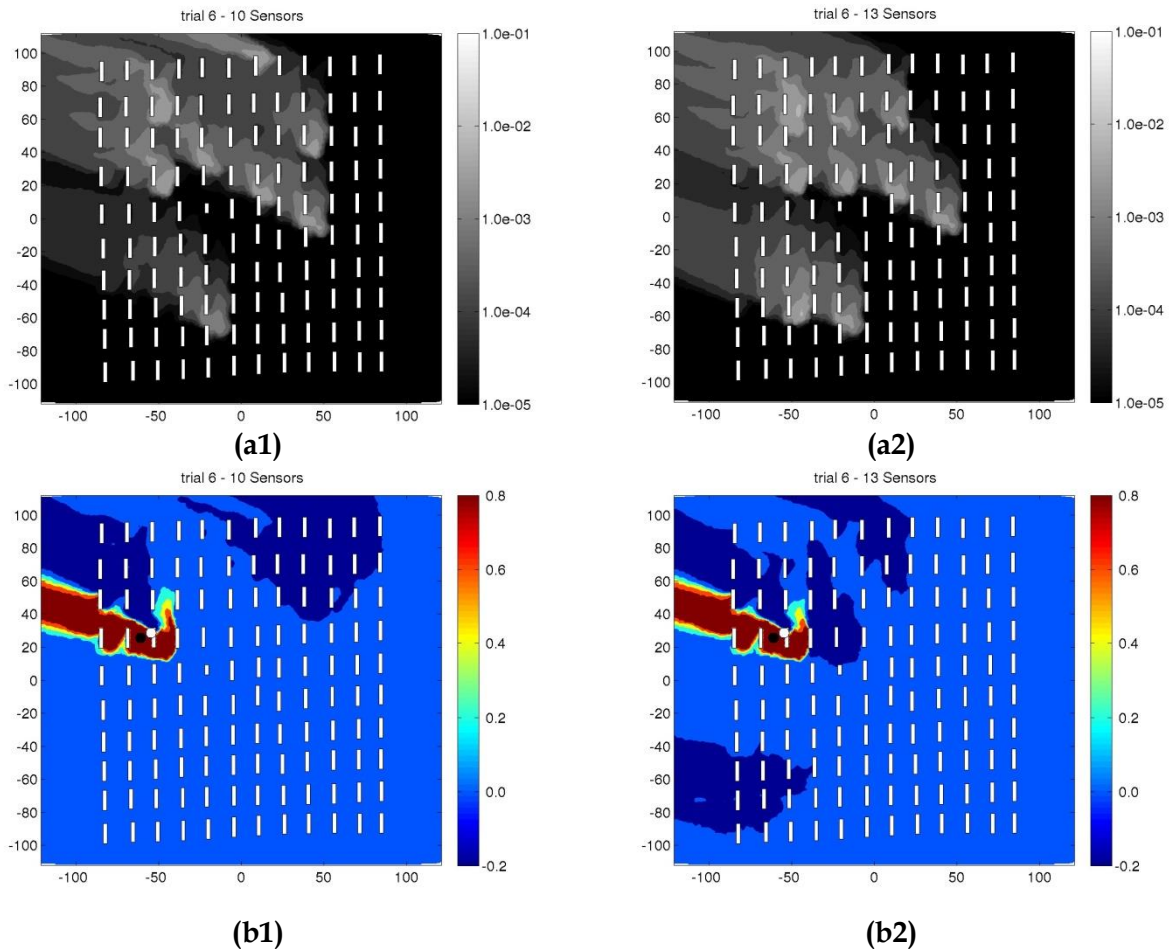


Figure 4.19. Fonctions de visibilité  $\varphi(x)$  des réseaux optimaux (a1)  $m=10$  à (a2)  $m=13$  et les sources  $s_{//\varphi}(x)$  (b1)  $m=10$  et (b2)  $m=13$

On rappelle que l'on peut retrouver la direction du vent à partir des champs de visibilité, car elle est opposée la direction des rétro-panache (direction des champs de visibilité).

Pour analyser la performance des configurations optimales, nous examinons l'évolution de  $El = \sqrt{(x_0 - y_s)^2 + (y_0 - y_s)^2}$  et de  $Eq = \frac{q_0}{q_s}$  en échelle logarithmique en fonction de  $m$ . Les résultats obtenus sont regroupés dans le tableau 4.8 et sont présentés sur la figure 4.20. Nos résultats sont ainsi comparés à ceux obtenus par (Kumar et al, 2015).

**Expérience 1 : « Essai 6 »**

Résultat de (Kumar et al, 2015)					
« All Measurements »			« Nonzero Measurements »		
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$
40	5,0	2,14	40	5,0	2,14

Résultats de l'optimisation		
$m$	$El(m)$	$Eq$
4	57,722	6.4256e+15
5	57,257	25,65
6	20,138	0,19432
7	20,138	0,19432
8	28,818	4,5579
9	20,138	0,19431
10	6,9635	1,0404
11	22,405	2,6331
12	6,9635	1,0406
13	6,9635	1,0406
14	7,2406	3,2955
15	16,475	4,1071
16	7,0226	4,0831

Tableau 4.10. Essai 6 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015)

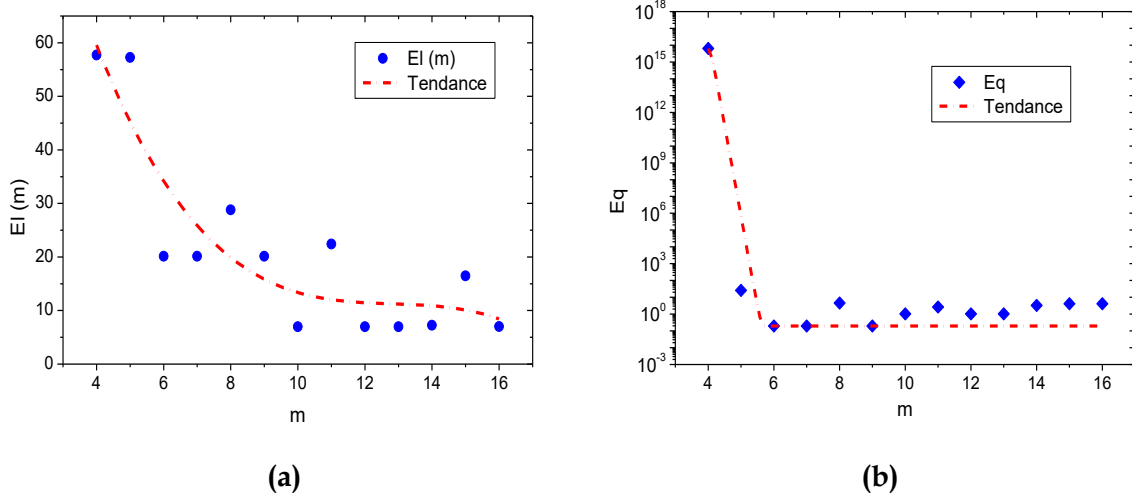


Figure 4.20. Erreurs d'estimation (a) de la position et (b) du débit de la source « Essai 6 »

Lorsque l'on observe l'évolution de  $Eq$  en fonction du nombre de capteurs  $m$ , on note une forte variation pour  $m < 10$ . Pour cet intervalle le débit estimé  $q_0$  est soit trop élevé ( $> 4 \times q_s$ ) soit trop faible (proche de zéro). Parallèlement, pour le même intervalle ( $m < 10$ ) les erreurs de

localisation  $El$  sont systématiquement supérieures à  $20\text{ m}$  et elles varient de  $20$  à  $57\text{ m}$ . Lorsque le nombre de capteurs est supérieur à  $10$  ( $m > 10$ ), l'erreur de localisation  $El$  se situe entre  $6,96\text{ m}$  et  $22\text{ m}$  tandis que  $Eq$  varie de  $1,04$  à  $4,1$ .

Les résultats montrent que les performances des réseaux optimaux formés de  $m = 10, 12$  et  $13$ , sont semblables à ceux obtenues par (Kumar et al, 2015) voir meilleurs dans l'estimation du débit, pour ces configurations les erreurs sont respectivement  $El < 7\text{ m}$  et  $Eq = 1.04$ .

Pour une analyse plus fine, nous réalisons la « matrice des structures communes ». Nous rappelons que cette matrice montre le nombre de capteurs commun entre les différents réseaux optimaux.

$m$	4	5	6	7	8	9	10	11	12	13	14	15	16
4	4	3	2	3	0	3	3	4	3	4	2	4	4
5	-	5	4	4	0	3	3	4	5	4	4	3	4
6	-	-	6	4	2	5	4	4	5	4	6	4	4
7	-	-	-	7	1	4	3	3	5	5	5	4	5
8	-	-	-	-	8	3	4	4	4	3	6	6	7
9	-	-	-	-	-	9	6	6	4	5	8	7	6
10	-	-	-	-	-	-	10	6	5	7	7	7	8
11	-	-	-	-	-	-	-	11	7	6	7	9	7
12	-	-	-	-	-	-	-	-	12	8	9	8	9
13	-	-	-	-	-	-	-	-	-	13	7	8	9
14	-	-	-	-	-	-	-	-	-	-	14	9	10
15	-	-	-	-	-	-	-	-	-	-	-	15	11
16	-	-	-	-	-	-	-	-	-	-	-	-	16

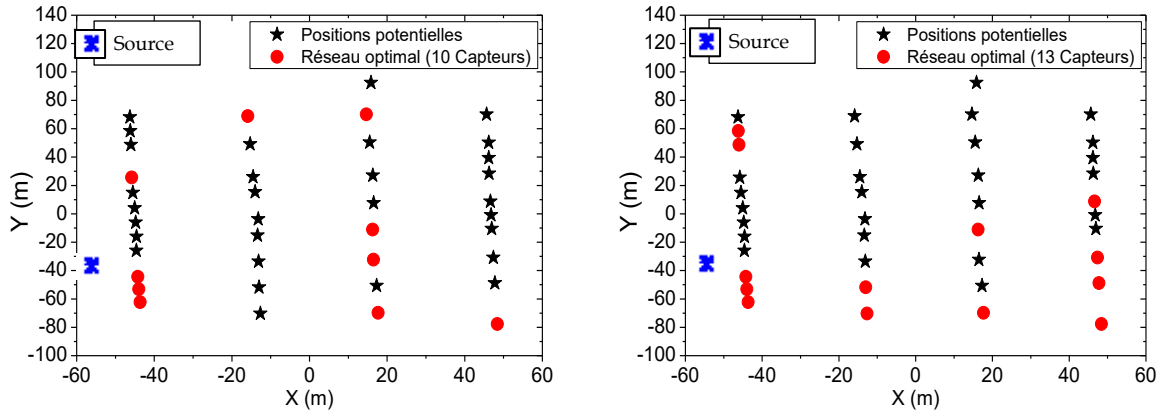
Tableau 4.11. Matrice des structures communes « Essai 6 »

L'analyse des meilleurs réseaux ( $m = 10, 12$  et  $13$ ) montre que seules les structures pour  $m = 12$  et  $m = 13$  ont une forte base commune (8 capteurs en commun). Ceci est moins vrai pour le réseau de  $m = 10$  qui contient seulement 5 capteurs en commun avec les réseaux de  $m = 12$  et  $m = 13$ . A l'issue de cette première analyse, on peut déduire que des réseaux dont la structure est plus au moins commune peuvent conduire à un minimum global.

Les structures des réseaux  $m = 4$  sont incluses dans le réseau dont le nombre de capteurs est  $m = 13$ . De même, la structure  $m = 5$  est incluse dans le réseau de  $m = 12$  donnant de bons résultats. Ceci montre que  $m = 4$  et  $m = 5$  capteurs sont des nombres insuffisants pour former des réseaux optimaux.

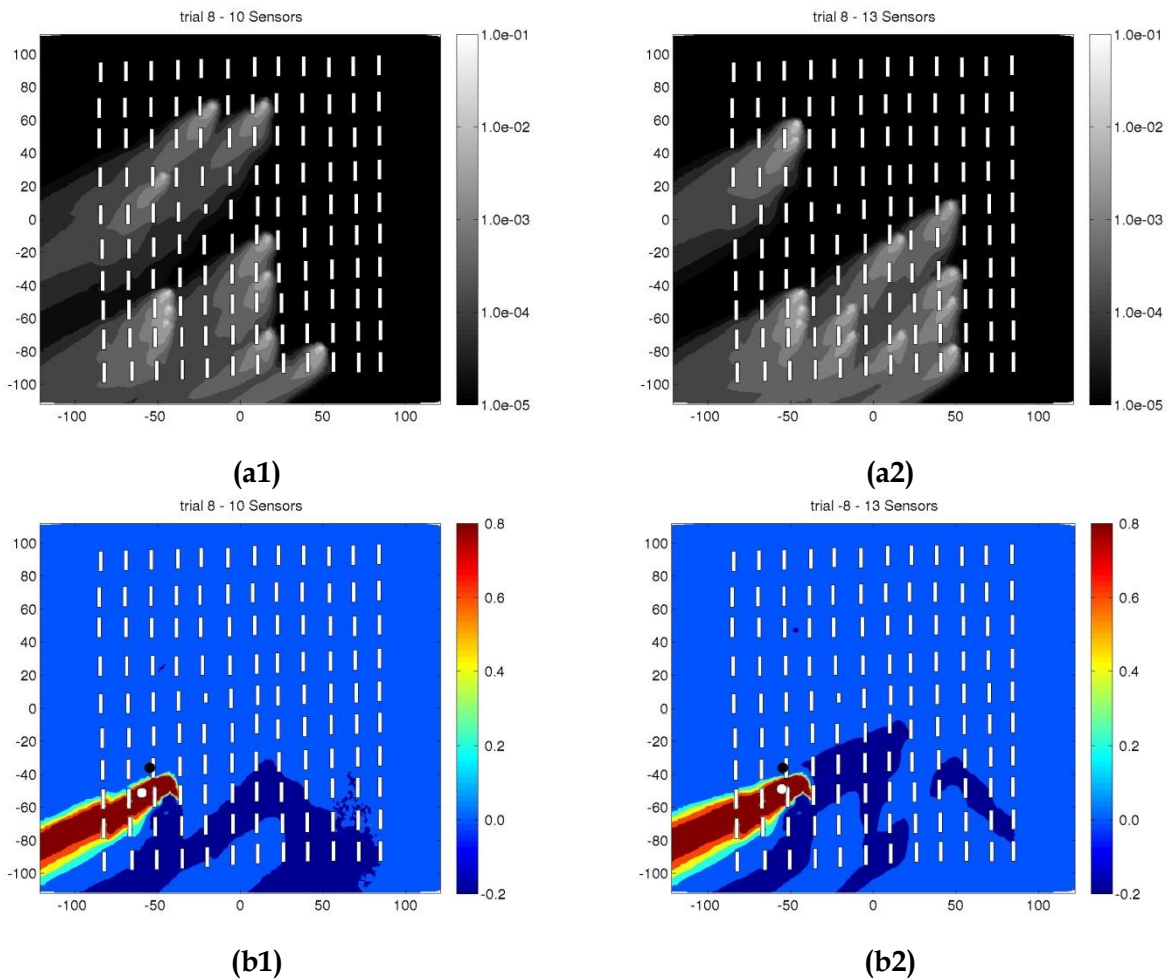
- **Expérience 2, « Essai 8 » : atmosphère modérément stable**

Pour l'expérience 2, nous présentons sur la figure 4.21 un échantillon des réseaux optimaux ( $m = 10$  et  $m = 13$ ). La totalité des solutions (de 4 à 16 capteurs) sont regroupées en annexe IV.



(a) (b)  
Figure 4.21. Configurations optimales (a)  $m=10$  et (b)  $m=13$  « Essai 8 »

Pour ces réseaux les cartes de  $\varphi(x)$  et de  $s_{//\varphi}(x)$  sont présentées sur la figure 4.22.



(a1) (a2)  
(b1) (b2)  
Figure 4.22. Fonctions de visibilité  $\varphi(x)$  des réseaux optimaux (a1)  $m=10$  à (a2)  $m=13$  et les sources  $s_{//\varphi}(x)$  (b1)  $m=10$  et (b2)  $m=13$

Qualitativement, nous constatons que des positions proches (3 capteurs) de la source sont sélectionnées. Cette tendance est générale pour tous les réseaux optimaux (de tailles  $m=4$  à  $m=16$ ) (voir annexe IV). Pour l'analyse quantitative, nous examinons l'évolution de l'erreur

de localisation  $El$  et de l'estimation du débit  $Eq$  en échelle logarithmique sur la figure 4.23. Les résultats obtenus sont aussi regroupés dans le tableau 4.9.

**Expérience 2 : « Essai 8 »**

Résultats de (Kumar et al, 2015)					
« All Measurements »			« Nonzero Measurements »		
$m$	$El (m)$	$Eq$	$m$	$El (m)$	$Eq$
40	15,8	2,22	30	13,5	0,39

Résultats de l'optimisation			
$m$	$El (m)$	$Eq$	
4	31,031	2.3e+08	
5	16,02	82,48	
6	8,6624	1,524	
7	7,6951	22,06	
8	16,02	82,48	
9	10,877	1,846	
10	15,785	1,762	
11	17,031	1,445	
12	14,382	1,630	
13	12,861	1,317	
14	13,013	3,045	
15	14,382	1,63	
16	15,785	2,105	

Tableau 4.12. Essai 8 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015).

Les résultats obtenus montrent que si l'on exclut les réseaux optimaux formés de  $m = 4, 5, 7$  et  $8$ , les performances sont semblables à ceux obtenus par (Kumar et al, 2015) et sont meilleures pour l'ensemble des réseaux de  $m = 6, 9, 10, 12, 13$  et  $15$ . Pour cet ensemble  $Eq$  varie de  $1,31$  à  $1,84$  et  $El$  varie de  $8,66m$  à  $17,03 m$ .

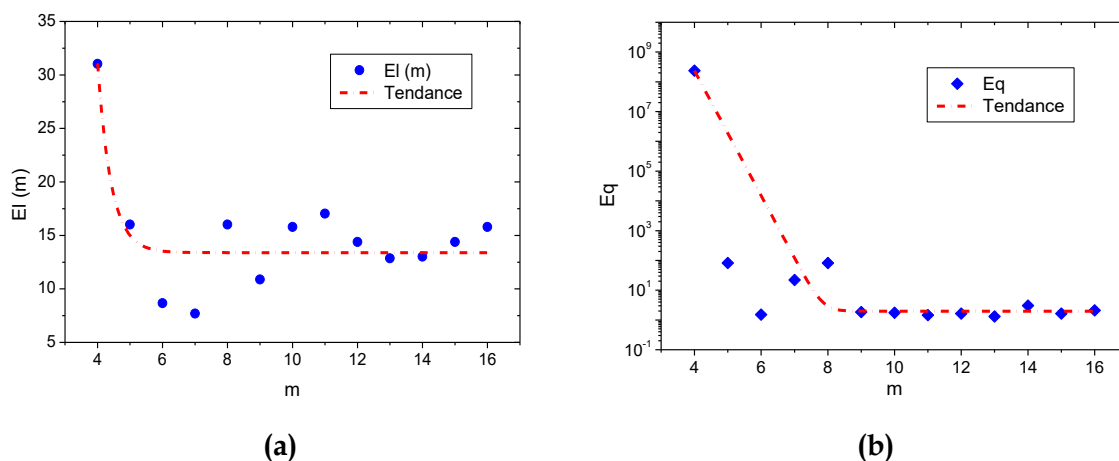


Figure 4.23 : Erreurs d'estimation (a) de la position et (b) du débit de la source « Essai 8 »

Pour une analyse plus fine nous réalisons la MSC. Après examen de la nature des structures communes nous constatons que les capteurs du réseau  $m=4$  sont inclus dans les réseaux de taille  $m=11, 12, 13$ , et  $15$  pour lesquels on obtient de bons résultats. Ceci prouve que

l'information délivrée pour ces 4 capteurs est insuffisante pour caractériser la source. On note également que pour  $m=5$ , on ne peut pas caractériser la source, il en est de même pour  $m=8$  dans laquelle  $m=5$  est incluse. On note également que 5 des 6 capteurs du réseau de taille  $m=6$  sont incluses dans les réseaux de taille  $m=12,13$  et 15. Il en résulte des performances en retrait comparativement à celles obtenues avec  $m=6$ .

$m$	4	5	6	7	8	9	10	11	12	13	14	15	16
4	4	4	3	3	4	3	3	4	4	4	4	4	2
5	-	5	3	3	5	3	4	4	4	4	4	4	3
6	-	-	6	4	3	3	4	3	5	5	4	5	3
7	-	-	-	7	4	3	3	5	5	6	7	6	6
8	-	-	-	-	8	4	5	6	5	6	6	6	4
9	-	-	-	-	-	9	4	5	6	7	7	5	4
10	-	-	-	-	-	-	10	4	3	5	4	5	3
11	-	-	-	-	-	-	-	11	5	8	8	8	5
12	-	-	-	-	-	-	-	-	12	8	10	8	7
13	-	-	-	-	-	-	-	-	-	13	10	7	6
14	-	-	-	-	-	-	-	-	-	-	14	11	9
15	-	-	-	-	-	-	-	-	-	-	-	15	8
16	-	-	-	-	-	-	-	-	-	-	-	-	16

Tableau 4.13. Matrice des structures communes « Essai 8 »

Les résultats obtenus pour le réseau de taille  $m=7$  montrent que le processus d'optimisation n'est pas satisfaisant. L'estimation du débit est quasiment divergente avec  $Eq=22,06$ . Ceci prouve que le recuit simulé a convergé vers un minimum local. Ainsi, le nombre de capteurs dans cette configuration ne peut être remis en question car pour  $m=6 < 7$  le recuit simulé est parvenu à proposer un bon réseau. Une des solutions dans ce cas est de réeffectuer des calculs probabilistes jusqu'à converger vers un minimum global, mais cette procédure nécessite un temps de calcul important. Nous proposons ici une alternative qui permet contourner ce problème.

➤ **Optimisation hybride « probabiliste & déterministe »**

La technique d'optimisation « probabiliste & déterministe » est une méthode d'optimisation hybride qui permet de déterminer le nombre optimal de capteurs  $m$ . La première étape consiste à déterminer un réseau optimal de  $k$  capteurs (avec  $k < m$ ) par un algorithme probabiliste de type recuit simulé. Il faut s'assurer que le résultat correspond bien à un minimum global. La seconde étape consiste à déterminer les  $m-k$  capteurs manquants. Pour compléter la structure de base de  $k$  capteurs, un processus itératif est appliqué de façon déterministe. Ainsi, si on considère  $p$  positions potentielles,  $k$  le nombre de capteurs obtenu par technique probabiliste, le processus itératif déterministe consiste à chercher à chaque fois le meilleur capteur dans l'espace des  $p-k$  positions restantes. Le meilleur capteur ajouté à la



structure de base ( $k$  capteurs) est celui pour lequel une fonction coût est minimale ( $J(R_{k+1})_{min}$ ). On applique ce processus  $m-k-1$  fois et on obtient finalement un réseau optimal formé de  $m$  capteurs. Dans notre cas, on choisira  $k=6$  (structure obtenue par le recuit simulé) et on générera de façon déterministe les réseaux correspondant à  $m=7,8,\dots,16$  capteurs. On applique les différentes étapes pour le réseau de taille  $m=7$ . Comme on a 40 positions potentielles, on commence par générer les 34 configurations possibles ( $40-6$ ) formées de 7 capteurs puis on évalue leurs coûts et on choisit le meilleur réseau (coût minimal). Ce processus est répété pour les réseaux de  $m=8,9,\dots,16$  capteurs. Les résultats de cette technique sont regroupés dans le tableau 4.10.

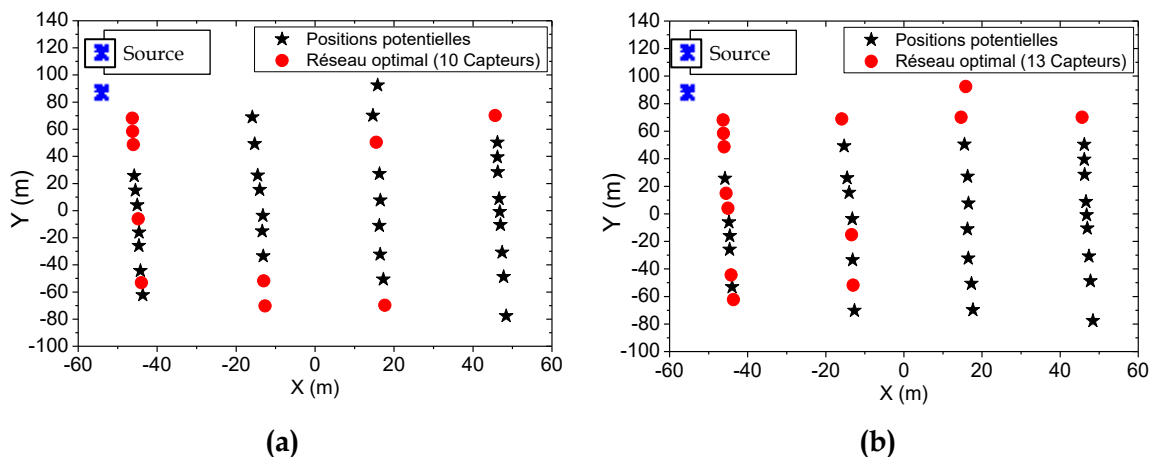
Nombre de capteurs	$El(m)$	$Eq$
7	8,6624	1,5236
8	8,6624	1,5236
9	8,6624	1,5237
10	8,6624	1,5237
11	8,6624	1,5237
12	8,6624	1,5237
13	8,6624	1,5238
14	8,6624	1,5238
15	8,6624	1,5238
16	8,6624	1,5238

Tableau 4.14. Essai 8 : Erreurs obtenues par technique d'optimisation hybride

Les résultats obtenus par la technique d'optimisation hybride sont quasiment identiques à ceux obtenus par le réseau de taille  $m=6$ . Il est à noter que toutes les nouvelles configurations incluent les six capteurs du réseau  $m=6$ . On constate également que les performances des anciens réseaux de taille  $m=7,\dots,16$  sont améliorées (voir tableau 4.9).

- **Expérience 3, « essai 14 » : atmosphère faiblement stable**

Pour l'essai 3, nous présentons sur la figure 4.24 un échantillon des réseaux optimaux ( $m=10$  et  $m=13$ ). La totalité des solutions (de 4 à 16 capteurs) est regroupée en annexe IV.



Nous dressons sur la figure 4.25 pour cet échantillon ( $m=10$  et  $m=13$ ) les cartes de visibilité  $\varphi(\mathbf{x})$  et des sources estimées  $s_{//\varphi}(\mathbf{x})$ . Nous rappelons que la position de la vraie source est représentée par un disque noir et la position estimée est représentée par un disque blanc.

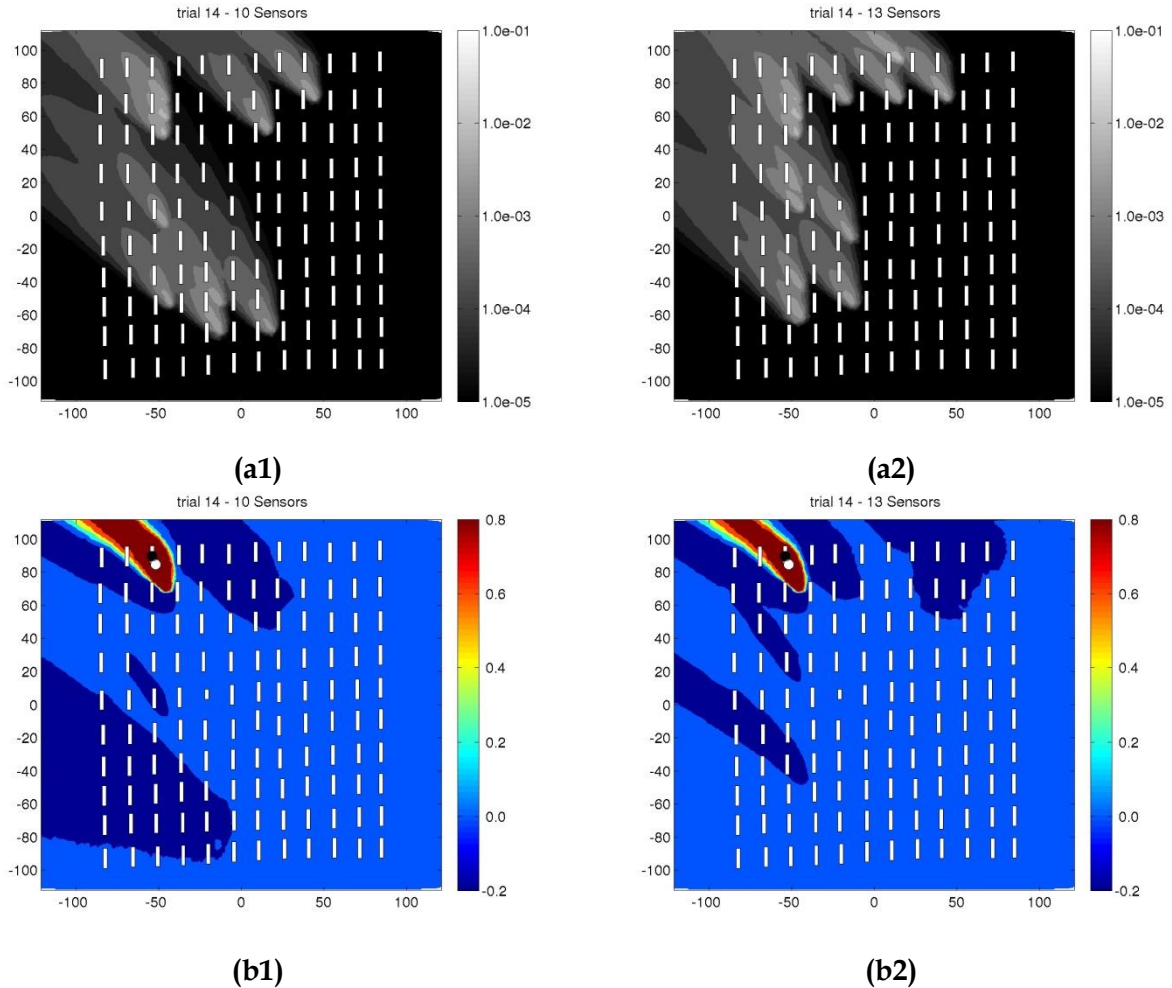


Figure 4.25. Fonctions de visibilité  $\varphi(\mathbf{x})$  des réseaux optimaux (a1)  $m=10$  à (a2)  $m=13$  et les sources  $s_{//\varphi}(\mathbf{x})$  (b1)  $m=10$  et (b2)  $m=13$

Qualitativement, nous constatons que des positions les plus proches (3 capteurs) de la source sont sélectionnées. Cette tendance est générale pour tous les réseaux optimaux (de tailles  $m=4$  à  $m=16$ ) (voir annexe IV).

Pour analyser la performance des configurations optimales, nous examinons les évolutions des erreurs de localisation  $El$  et de l'estimation du débit  $Eq$  en échelle logarithmique représentées sur la figure 4.26. Les résultats sont également regroupés dans le tableau 4.11.

Pour les réseaux de taille  $m=7$  à  $m=16$  capteurs (excepté  $m=15$ ), on note une sous-estimation du débit ; en effet,  $Eq$  varie de 0,8 à 1,07. On note également une erreur de localisation  $El$  qui varie de 2,06 m à 10,14 m. Pour cette expérience, les débits estimés pour les réseaux de tailles

$m=4, 5$  et  $6$  sont issus de calculs qui divergent ( $E_q \rightarrow 10^{17}$ ). On remarque aussi que les résultats obtenus à partir des réseaux formés de  $12$  et  $14$  capteurs sont identiques, avec  $E_q=0,9$  et  $El=4,8m$ .

**Expérience 3 : « Essai 14 »**

Résultats de (Kumar et al, 2015)					
« All Measurements »			« Nonzero Measurements »		
$m$	$El(m)$	$E_q$	$m$	$El(m)$	$E_q$
40	2,2	1,42	29	2,2	1,42

Résultats de l'optimisation			
$m$	$El(m)$	$E_q$	
4	156,43	1,4651e+17	
5	103,07	1,4651e+17	
6	137,2	1,4648e+17	
7	5,8642	0,80469	
8	5,8642	0,8047	
9	4,6407	0,86307	
10	5,4966	0,87744	
11	5,8642	0,80468	
12	4,8088	0,90286	
13	5,4966	0,87739	
14	4,8088	0,90263	
15	2,6096	1,0778	
16	10,147	0,97103	

Tableau 4.15. Essai 14 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015).

Les résultats obtenus montrent que si l'on exclut les réseaux optimaux formés de  $m = 4, 5$  et  $6$  les performances sont semblables à ceux obtenus par (Kumar et al, 2015) et sont meilleures pour le réseau de taille  $m=15$ . Ainsi, on peut noter pour cette expérience que le nombre optimal de capteurs est  $15$ .

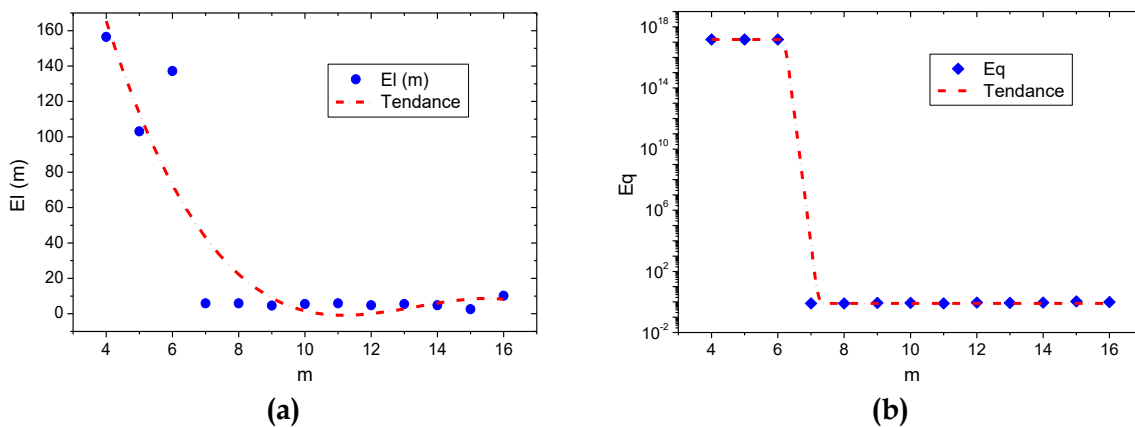


Figure 4.26. Erreurs d'estimation (a) de la position et (b) du débit de la source « Essai 14 »

Pour une analyse plus fine nous examinons la MSC. On constate que les réseaux formés de 12 et 14 capteurs ont une forte structure commune (10 capteurs). On note aussi que les résultats pour des réseaux de tailles  $m=4, 5$  et  $6$  ne sont pas satisfaisants (les calculs divergent), malgré plusieurs calculs d'optimisations successifs. Ceci prouve que pour cette configuration le nombre minimal de capteurs permettant de caractériser la source est  $m=7$ .

$m$	4	5	6	7	8	9	10	11	12	13	14	15	16
4	4	4	2	1	1	1	1	2	3	2	3	2	2
5	-	5	3	1	1	1	1	3	3	3	3	3	2
6	-	-	6	1	1	1	1	3	4	5	5	3	3
7	-	-	-	7	4	3	5	4	4	5	4	4	4
8	-	-	-	-	8	4	6	5	5	4	4	5	6
9	-	-	-	-	-	9	4	5	4	5	5	5	7
10	-	-	-	-	-	-	10	6	5	5	6	5	6
11	-	-	-	-	-	-	-	11	6	7	8	6	6
12	-	-	-	-	-	-	-	-	12	8	10	8	9
13	-	-	-	-	-	-	-	-	-	13	9	9	7
14	-	-	-	-	-	-	-	-	-	-	14	8	9
15	-	-	-	-	-	-	-	-	-	-	-	15	10
16	-	-	-	-	-	-	-	-	-	-	-	-	16

Tableau 4.16. Matrice des structures communes « Essai 14 »

- **Expérience 4, « essai 19 » : atmosphère modérément stable**

Pour l'expérience 4, nous présentons de la même façon que pour les autres expériences un échantillon des réseaux optimaux ( $m=10$  et  $m=13$ ) (voir figure 4.27). La totalité des solutions (de 4 à 16 capteurs) ainsi que les détails des calculs sont également regroupés en annexe IV.

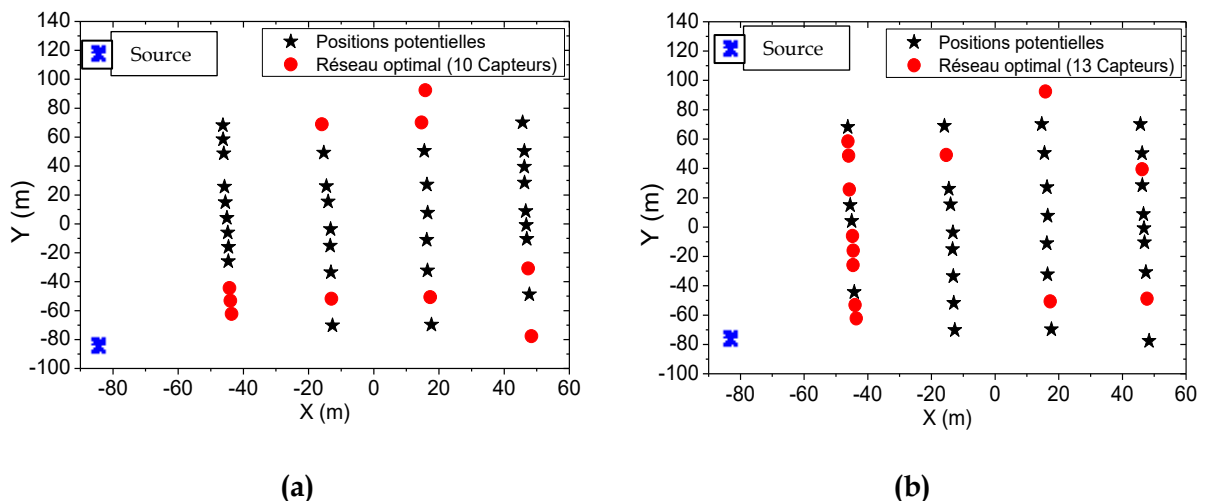


Figure 4.27. Configurations optimales (a)  $m=10$  et (b)  $m=13$  « Essai 19 »

Nous présentons sur la figure 4.28 pour les réseaux sélectionnés de tailles  $m=10$  et  $m=13$  les cartes de visibilité  $\varphi(\mathbf{x})$  et des sources estimées  $s_{//\varphi}(\mathbf{x})$ .

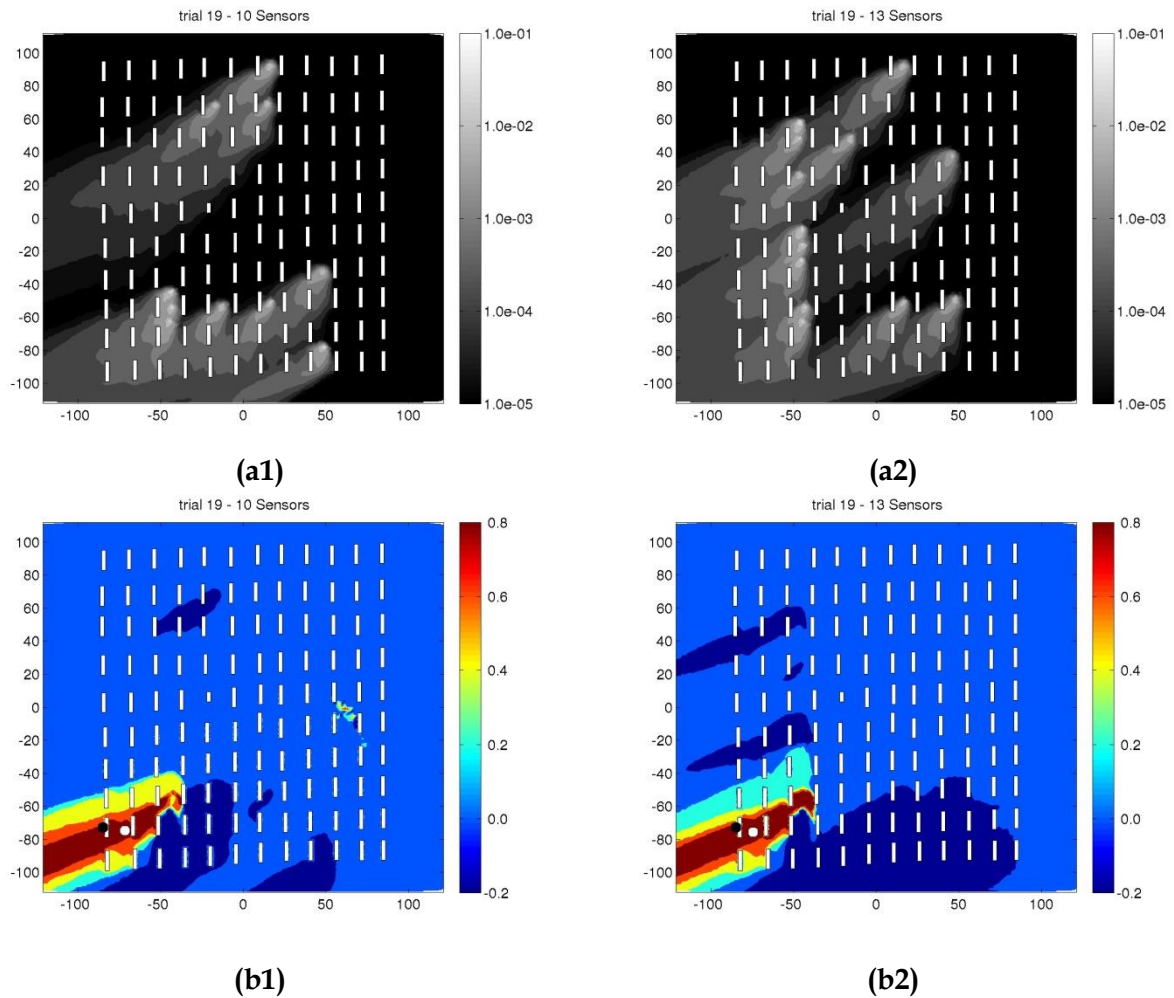


Figure 4.28. Fonctions de visibilité  $\varphi(\mathbf{x})$  des réseaux optimaux (a1)  $m=10$  à (a2)  $m=13$  et les sources  $s_{//\varphi}(\mathbf{x})$  (b1)  $m=10$  et (b2)  $m=13$

Qualitativement, nous constatons que des positions les plus proches de la source sont souvent sélectionnées. Cette tendance est générale pour tous les réseaux optimaux (de tailles  $m=4$  à  $m=16$ ) (voir annexe IV).

Pour analyser la performance des configurations optimales, nous examinons l'évolution de l'erreur de localisation  $El$  ainsi que l'évolution de l'erreur de l'estimation du débit  $Eq$  (en échelle logarithmique) sur la figure 4.29. Les résultats sont aussi regroupés dans le tableau 4.12.

Les résultats obtenus montrent que l'ensemble des configurations testées donnent des estimations meilleures comparativement à celles obtenues par (Kumar et al, 2015). L'erreur maximale de localisation est de  $13,45\text{ m} < 22,3\text{ m}$  (pour 40 capteurs) ; l'ensemble des réseaux optimaux sous-estime le débit de la source exceptés les réseaux de tailles  $m=12$  et  $m=14$ .

**Expérience 4 : « Essai 19 »**

Résultats de (Kumar et al. 2015)					
« All Measurements »			« Nonzero Measurements »		
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$
40	22,3	1,76	35	22,3	1,76

Résultats de l'optimisation			
$m$	$El(m)$	$Eq$	
4	13,459	0,83288	
5	13,459	0,8329	
6	13,459	0,83289	
7	13,459	0,83279	
8	13,459	0,83286	
9	13,459	0,83281	
10	13,459	0,83282	
11	13,459	0,83281	
12	9,2805	1,472	
13	10,769	0,9923	
14	5,6547	1,0519	
15	10,769	0,99237	
16	13,459	0,83294	

Tableau 4.17. Essai 19 : Caractérisation de la source : nos résultats VS les résultats de (Kumar et al, 2015).

On note que le meilleur réseau est obtenu pour une taille  $m=14$  avec  $El=5,65 m$  et  $Eq=1,05$ . Ainsi le nombre optimal dans ce cas est de 14 capteurs. Contrairement à ce qu'on a pu constater dans les autres « Essais » des expériences MUST, dans cette dernière expérience, 4 capteurs sont suffisants pour caractériser la source.

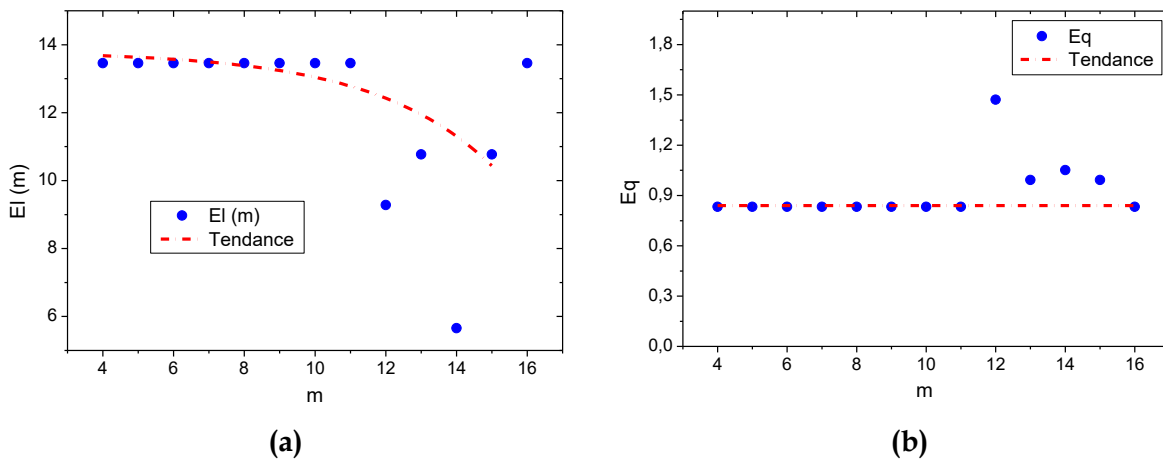


Figure 4.29. Erreurs d'estimation de débit (a) et position de la source (b) « Essai 19 »

De la même façon que pour les autres « Essais » nous construisons la MSC pour effectuer une analyse des structures des réseaux optimaux.

On note que les réseaux, formés de 4 à 10 capteurs, ont des performances quasi-identiques malgré qu'ils ont des structures relativement différentes. Ceci montre que plusieurs

configurations présentant des structures différentes peuvent conduire à une caractérisation précise de la source.

<i>m</i>	4	5	6	7	8	9	10	11	12	13	14	15	16
4	4	4	3	4	3	3	3	3	3	2	2	2	3
5	-	5	4	4	3	4	3	3	4	2	3	2	3
6	-	-	6	4	5	5	4	5	4	3	4	4	5
7	-	-	-	7	4	4	5	4	3	2	4	5	5
8	-	-	-	-	8	4	6	7	5	6	6	4	6
9	-	-	-	-	-	9	5	5	7	3	5	5	6
10	-	-	-	-	-	-	10	7	7	4	8	6	8
11	-	-	-	-	-	-	-	11	6	7	7	7	9
12	-	-	-	-	-	-	-	-	12	5	7	5	6
13	-	-	-	-	-	-	-	-	-	13	8	9	9
14	-	-	-	-	-	-	-	-	-	-	14	8	10
15	-	-	-	-	-	-	-	-	-	-	-	15	11
16	-	-	-	-	-	-	-	-	-	-	-	-	16

Tableau 4.18. Matrice des structures communes « Essai 19 »

#### 4.3.4.3. Conclusion à l'issue du problème 1

Les différents résultats obtenus montrent que l'on peut réduire la taille d'un réseau tout en conservant ses capacités de caractérisation de la source. Cette démarche a été appliquée avec succès pour différentes conditions de stabilité atmosphérique et pour différentes positions de la source.

Il ressort de cette étude qu'on ne peut pas définir a priori un nombre de capteurs minimal permettant de caractériser la source. Ce nombre est variable et dépend du contexte expérimental.

Les analyses des meilleurs résultats montrent que les réseaux présentant des structures relativement différentes peuvent conduire aux meilleurs résultats. Ceci prouve que pour ce problème la solution n'est pas unique.

L'analyse des différentes « Matrices des Structures Communes » montre que si une structure de base d'un réseau optimal (qui produit de bons résultats) est incluse dans un réseau de taille plus importante, cela ne conduit pas nécessairement à des réseaux ayant de bonnes performances.

Pour les expériences considérées, on note que le minimum de la fonction coût (résidus normalisés) ne correspond pas systématiquement à une meilleure caractérisation de la source. On note également que pour un réseau optimal, l'estimation précise de la position et du débit de la source n'est pas toujours obtenue simultanément.

### 4.3.5. Problème 2 « Déploiement en urgence d'un réseau mobile sans information a priori sur la source »

Nous rappelons que le *problème 2* consiste à déterminer un réseau optimal capable de caractériser une source inconnue localisée n'importe où dans un domaine large. Les conditions météorologiques, dans ce cas, sont fixes et la fonction objectif adéquate à ce problème est basée sur l'entropie de réseau (définie dans la partie 3.7.2) et est définie par

$$Jp_2 = \frac{1}{2} \log(\det(\mathbf{H}_\varphi))$$

#### 4.3.5.1. Paramétrage

Suite à l'étude réalisée dans le cadre du *problème 1* nous avons noté que généralement le nombre optimal de capteurs n'est pas trivial à déterminer. En effet, pour cette étude nous considérons un nombre minimal de 8 capteurs. Nous avons remarqué que des réseaux optimaux à partir de cette taille ( $m=8$ ), garantissent une caractérisation raisonnable de la source pour toutes les expériences testées. Au-delà de cette taille nous considérons deux autres réseaux de tailles arbitraires  $m=10$  et  $m=12$ . Ainsi, les paramètres retenus pour la recherche des configurations optimales sont regroupés dans le tableau 4.13.

$m$	$T_0$	$T_f$	$L$	$\theta$
8	15	$1.10^{-11}$	100	0,9
10	15	$1.10^{-11}$	100	0,9
12	15	$1.10^{-11}$	200	0,9

Tableau 4.19. Paramétrage recuit simulé « Problème 2 »

#### 4.3.5.2. Résultats et discussions

Pour chaque expérience nous analysons les structures des configurations optimales obtenues, ensuite nous examinons la performance de ces réseaux pour la caractérisation des sources disponibles.

- **Expérience 1 : « Essai 6 » : atmosphère très stable**

Nous présentons tout d'abord un échantillon des réseaux optimaux (8 et 12 capteurs) basés sur le concept entropique (voir la figure 4.30). La totalité des réseaux optimaux ainsi que l'évolution des fonctions objectifs sont regroupées dans l'annexe IV.



Du point de vue qualitatif, nous remarquons que les différentes structures optimales obtenues ont tendance à se disperser sur l'ensemble du domaine. Cette tendance est plutôt cohérente avec le concept de surveillance d'un espace large.

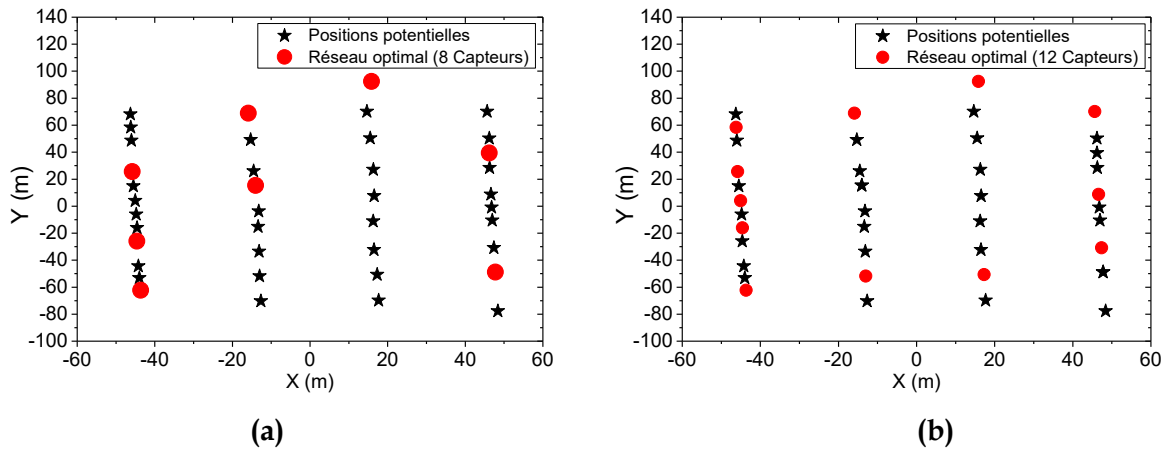


Figure 4.30. Configurations optimales (a)  $m=10$  et (b)  $m=12$  « Essai 6 »

Du point de vue quantitatif, les résultats obtenus (voir tableau 4.14) sont assez homogènes pour ce qui concerne l'estimation du débit et la position de la source. Ainsi,  $E_q$  varie de 1,91 à 2,31 et  $E_l$  varie de 4,64 m à 16,47 m.

On note que la performance de ces réseaux est équivalente à celle obtenue par (Kumar et al, 2015) (voir tableau 4.14) voir meilleure pour la configuration optimale formée de 12 capteurs avec des erreurs respectivement de  $E_q=1,91$  et de  $E_l=4,64m$ .

$m$	$E_l(m)$	$E_q$
8	16,4748	2,3162
10	5,3891	1,9827
12	4,6436	1,9124

Tableau 4.20. Essai 6 : Erreurs obtenues pour le problème 2

- **Expérience 2 « Essai 8 » : atmosphère modérément stable**

A l'image de l'expérience 1, nous présentons un échantillon des réseaux optimaux (8 et 12 capteurs) (voir figure 4.30). Egalement, la totalité des réseaux optimaux ainsi que l'évolution des fonctions objectifs pour cette expérience sont regroupées dans l'annexe IV.

Qualitativement, on note la même tendance observée pour l'expérience 1. Ainsi les capteurs formant les réseaux optimaux sont dispersés sur l'ensemble du domaine.

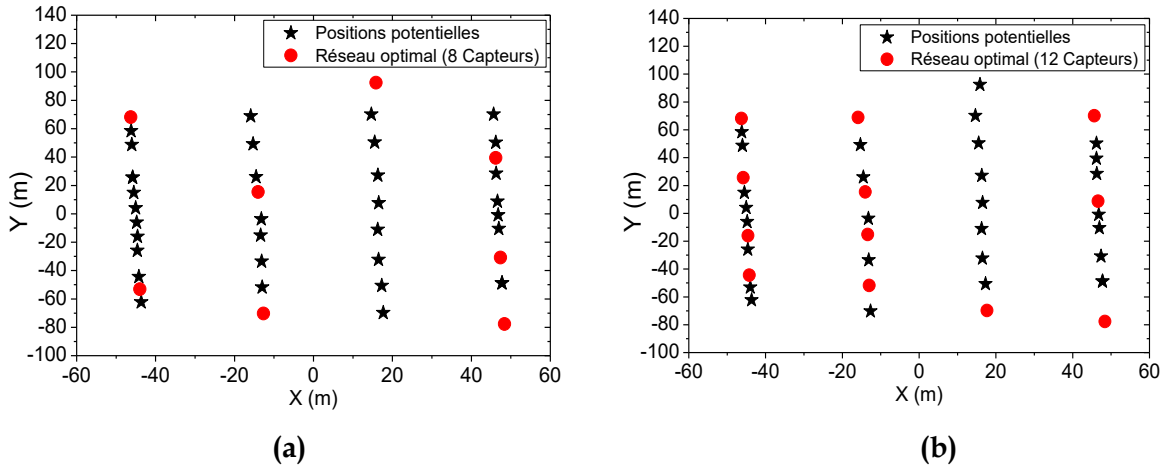


Figure 4.31. Configurations optimales (a)  $m=8$  et (b)  $m=12$  « Essai 8 »

Les erreurs  $E_q$  et  $E_l$  sont regroupées dans le tableau 4.15. Dans ce cas précis, les résultats obtenus sont assez homogènes comme pour la première expérience et sont de bien meilleures qualités que ceux obtenus par (Kumar et al, 2015) (voir tableau 4.9), excepté pour le réseau de taille  $m=8$ . Pour ce dernier, l'erreur de localisation de la source est de  $16,84\text{ m} > 15,8\text{ m}$ . On note de façon générale que  $E_q$  varie de 0,66 à 1,48 et  $E_l$  varie de 6,3 à 16,84. Le meilleur réseau dans ce cas est celui formé de 10 capteurs avec respectivement  $E_q=1,46$  et  $E_l=6,37\text{ m}$ .

$m$	$E_l (m)$	$E_q$
8	16,8400	1,4866
10	6,3795	1,4634
12	12,3481	0,6663

Tableau 4.21. Essai 8 : Erreurs obtenues pour le problème 2

- **Expérience 3 « Essai 14 » : atmosphère faiblement stable**

Pour l'expérience 3, nous présentons sur la figure 4.32 les réseaux optimaux formés de 8 et 12 capteurs. La totalité des réseaux optimaux ainsi que l'évolution des fonctions objectives pour cette expérience sont regroupées dans l'annexe IV.

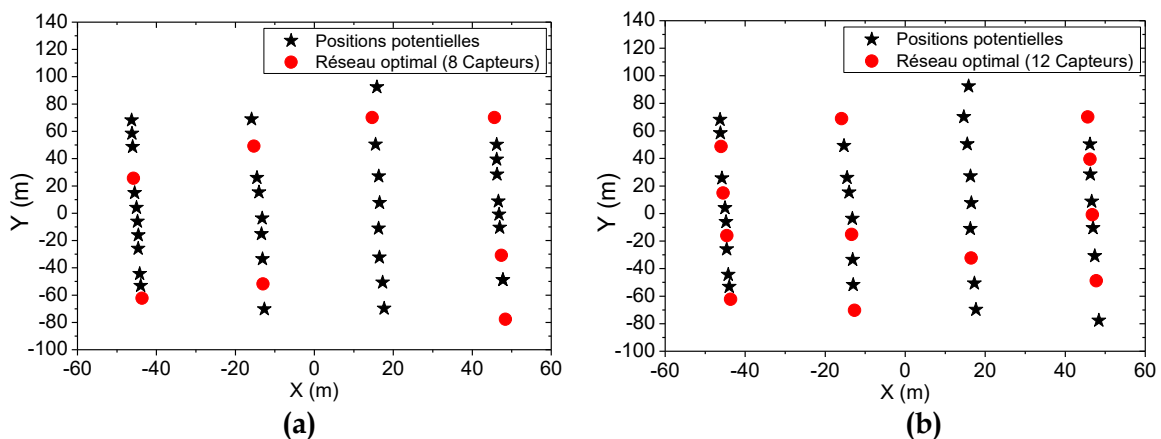


Figure 4.32. Configurations optimales (a)  $m=8$  et (b)  $m=12$  « Essai 14 »

Pour cette expérience, les erreurs d'estimation du débit et de la position de la source sont regroupées dans le tableau 4.16.

$m$	$El (m)$	$Eq$
8	10,991	1,7018
10	5,8465	1,9175
12	8,4088	1,9402

Tableau 4.22. Essai 14 : Erreurs obtenues pour le problème 2

Les résultats obtenus sont assez homogènes comme pour la première et la deuxième expérience. Nous constatons malgré tout que ces réseaux optimaux de tailles  $m=8, 10$  et  $12$  sont moins performants que ceux testés par (Kumar et al, 2015) (voir tableau 4.10). On note que  $Eq$  varie de 1,70 à 1,94 et  $El$  varie de 5,84 m à 10,99 m.

Les cas testés ne permettent pas d'obtenir pour un nombre de capteur  $m$  considéré, une meilleure estimation simultanée du débit et de la position de la source. Si l'on considère  $El$ , alors le meilleur réseau est obtenu pour  $m=10$ , tandis que si l'on considère  $Eq$ , alors le meilleur réseau sera celui de taille  $m=8$ .

- **Expérience 4 « Essai 19 » : atmosphère modérément stable**

Pour la dernière expérience, nous présentons un échantillon des réseaux optimaux de taille  $m=8$  et  $m=12$  capteurs (voir figure 4.32). Egalement, la totalité des réseaux optimaux ainsi que l'évolution des fonctions objectives sont regroupées dans l'annexe IV.

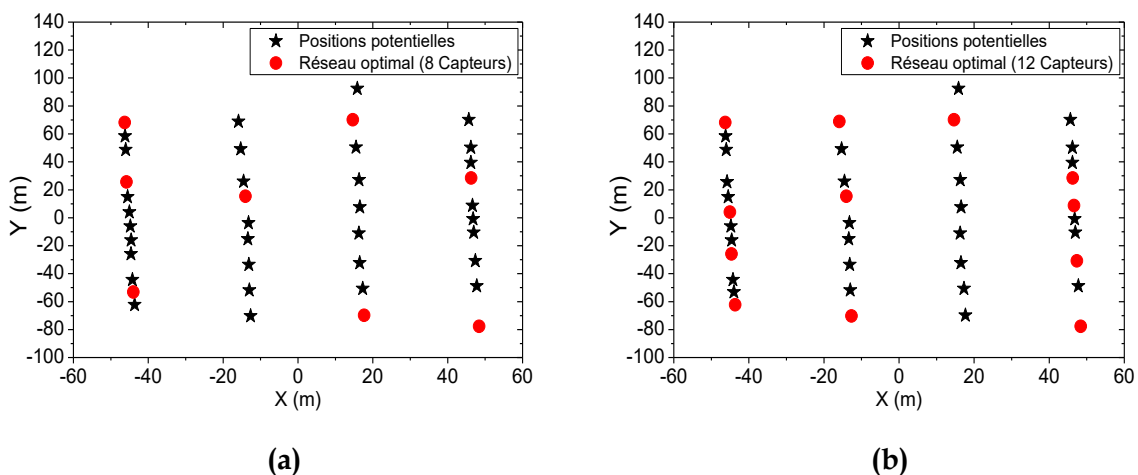


Figure 4.33. Configurations optimales (a)  $m=10$  et (b)  $m=12$  « Essai 19 »

Pour cette expérience, la tendance de dispersion des capteurs sur tout le domaine est également notée. Pour effectuer l'étude quantitative, les erreurs  $Eq$  et  $El$  sont regroupées dans le tableau 4.17.

Nombre de capteurs	$El (m)$	$Eq$
8	7.0958	1.2341
10	13.5779	1.3992
12	4.1333	1.6365

Tableau 4.23. Essai 19 : Erreurs obtenues pour le problème 2

Dans ce cas, nous observons de meilleurs résultats comparés à ceux de (Kumar et al, 2015) (voir tableau 4.12). Ainsi,  $Eq$  varie de 1,23 à 1,63 et  $El$  varie de 4,13 m à 13,57 m. Les cas testés ne permettent pas d'obtenir pour un nombre de capteur  $m$  considéré, une meilleure estimation du débit et de la position de la source. Si l'on considère  $El$ , alors le meilleur réseau est obtenu pour  $m=12$ , tandis que si l'on considère  $Eq$ , alors le meilleur réseau sera celui de taille  $m=8$ .

### 4.3.5.3. Conclusion à l'issue du problème 2

Pour le problème 2 dans le cadre de l'approche CFD, les différents résultats obtenus montrent que l'on peut déterminer un réseau optimal sur la base d'une fonction coût entropique.

Il ressort de cette étude qu'on ne peut pas définir a priori un nombre de capteurs minimal permettant de caractériser la source sachant qu'elle peut être localisée n'importe où dans un large domaine.

Pour les expériences considérées dans cette étude, on note que la fonction coût (entropie) est efficace et conduit à des réseaux qui caractérisent la source avec succès. On note également que pour un réseau optimal, une estimation précise de la position et du débit de la source n'est pas toujours obtenue simultanément. Enfin, on note qu'aucune divergence des calculs n'a eu lieu et que par l'entropie le processus de recherche probabiliste est assez performant.

### 4.3.6. Problème 3 « Conception optimale d'un réseau statique sans information a priori sur la source »

Dans ce cas les conditions météorologiques varient et correspondent à plusieurs directions et vitesses de vent. Le réseau est considéré statique et la source est supposée inconnue. Nous rappelons que la fonction objectif adéquate à ce problème (définie dans le paragraphe 3.7.2) est une entropie généralisée sous la forme :

$$Jp_3 = - \sum_{i=1}^k P_i \log(P_i) + \sum_{i=1}^k P_i S_i(R)$$

où  $k$  est le nombre de directions de vent considérées et  $P_i$  est la probabilité d'occurrence de chaque vent dans la direction  $i$ . Cette fonction objectif est simplifiée car dans l'expérience MUST, nous considérons que les vents sont équiprobables. Ceci revient à considérer chaque situation météorologique avec la même importance.

Ainsi la fonction objectif pour l'optimisation du réseau statique est donnée par

$$Jp_3 = \sum_{i=1}^k \log(\det(\mathbf{H}_{\varphi_i}))$$

Physiquement, cette fonction coût estime la quantité d'information fournie par le réseau pour caractériser la source, et ce pour différentes directions et vitesses de vent ayant la même probabilité d'occurrence.

#### 4.3.6.1. Paramétrage

Pour la détermination des tailles de réseaux à tester pour le *problème 3*, nous nous basons sur les constatations du *problème 1*. Aussi, nous considérons trois réseaux de tailles  $m=8,10$  et  $13$  capteurs. Il est à noter que pour des tailles de réseaux à partir de  $m=8$ , nous garantissons des estimations raisonnables de la source. Autrement dit, pour des réseaux de petites tailles ( $m<8$ ) les calculs sont susceptibles de diverger.

Les paramètres retenus pour chercher les configurations optimales sont regroupés dans le tableau 4.18.

$m$	$T_0$	$T_f$	$L$	$\theta$
8	10	$1.10^{-11}$	50	0,9
10	10	$1.10^{-11}$	100	0,9
12	10	$1.10^{-11}$	100	0,9

Tableau 4.24. Paramétrage recuit simulé « problème 3 »

#### 4.3.6.2. Résultats et discussions

Les résultats du *problème 3* concernent des champs météo multiples. Les conditions météorologiques des 4 expériences ont été prises en compte avec la même probabilité d'occurrence ( $P_i=0,25$ ).

Sur la figure 4.33 on observe comme pour le cas du *problème 2* (météo unique) que les réseaux optimaux sont constitués de capteurs dispersés à l'échelle de l'ensemble du domaine.

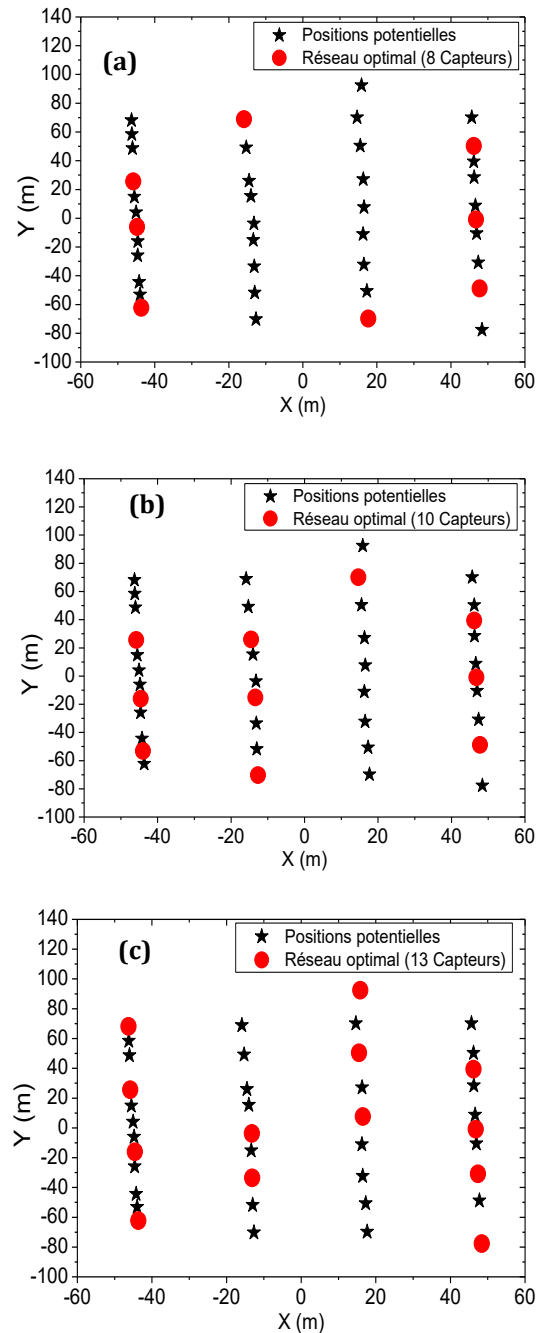


Figure 4.34. Configurations optimales (a)  $m=8$ , (b)  $m=10$  et (c)  $m=13$

Afin de vérifier les performances des réseaux correspondant à ce problème, nous les avons évaluées pour chacune des sources disponibles (composante des 4 expériences considérées).

Les résultats de la caractérisation des 4 sources sont regroupés dans le tableau 4.19. Pour l'expérience 1 les réseaux optimaux surestiment le débit,  $E_q$  varie de 2.16 à 2.32, tandis que l'erreur de localisation de la source  $E_l$  varie de 4.11 m à 6.78 m. Ainsi, on obtient des résultats comparables à ceux obtenus pour 40 capteurs (Kumar et al, 2015).

Résultats de (Kumar et al, 2015)									
Expérience 1 : « Essai 6 »									
« All Measurements »					« Nonzero Measurements »				
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$
40	5,0	2,14	40	5,0	2,14				
Expérience 2 : « Trial 8 »									
« All Measurements »					« Nonzero Measurements »				
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$
40	15,8	2,22	30	13,5	0,39				
Expérience 3 : « Trial 14 »									
« All Measurements »					« Nonzero Measurements »				
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$
40	2,2	1,42	29	2,2	1,42				
Expérience 4 : « Trial 19 »									
« All Measurements »					« Nonzero Measurements »				
$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$	$El(m)$	$Eq$	$m$
40	22,3	1,76	35	22,3	1,76				
Résultats de l'optimisation									
« Réseau optimal »		Expérience 1		Expérience 2		Expérience 3		Expérience 4	
$m$	$El(m)$	$Eq$	$El(m)$	$Eq$	$El(m)$	$Eq$	$El(m)$	$Eq$	$Eq$
8	4,11	2,32	19,66	1,63	26,69	1,63	4,13	1,56	
10	6,78	2,16	16,84	1,49	3,98	1,78	4,20	1,52	
13	6,78	2,22	22,09	1,76	1,90	1,79	4,13	1,51	

Tableau 4.25. Caractérisation des 4 sources : Nos résultats VS les résultats de (Kumar et al, 2015)

Les résultats obtenus pour l'expérience 2 montrent que les réseaux optimaux estiment mieux le débit que le réseau de taille  $m=40$ . Dans ce cas,  $Eq$  varie de 1,49 à 1,76 (à comparer avec 2,22). Toutefois les erreurs de localisation  $El$  varient de 16,84 m à 19,66 m. Pour l'ensemble des configurations optimales, ces estimations sont supérieures à 15,84 m obtenue pour  $m=40$ .

Pour l'expérience 3, le débit de la source est surestimé et  $Eq$  varie de 1,63 à 1,79 tandis que l'erreur de localisation  $El$  varie de 1,9 m à 26,69 m. Pour cette expérience, le réseau formé de 13 capteurs a quasiment les mêmes performances que celui formé de 40 capteurs pour l'estimation de la position de la source ( $1,90\text{ m} < 2,2\text{ m}$ ). Il est toutefois moins performant dans l'estimation du débit ( $1,79 > 1,42$ ).

Pour l'expérience 4, les réseaux optimaux de taille  $m=8,10$  et 13 capteurs parviennent à bien caractériser les sources. L'erreur de localisation  $El$  varie de 4,13 m à 4,20 m tandis que l'erreur de l'estimation du débit de la source  $Eq$  varie de 1,51 à 1,56. Pour cette expérience, l'ensemble des réseaux optimaux estime la source avec davantage de précision que le réseau de taille  $m=40$ .

Globalement, si l'on considère les 4 expériences simultanément, les réseaux optimaux de tailles  $m=8,10$  et  $13$  capteurs caractérisent convenablement les sources. Les performances de ces réseaux optimaux sont semblables, voire meilleures que celles obtenues par (Kumar et al, 2015). Pour plus de clarté, nous présentons tous les résultats du problème 3 sous forme d'histogramme (bar plot). Sur la figure 4.34, nous illustrons les erreurs  $E_q$  et  $E_l$  uniquement pour le réseau de taille  $m=10$  capteurs. La même représentation est disponible en annexe IV pour les réseaux de tailles  $m=8$  et  $m=13$ .

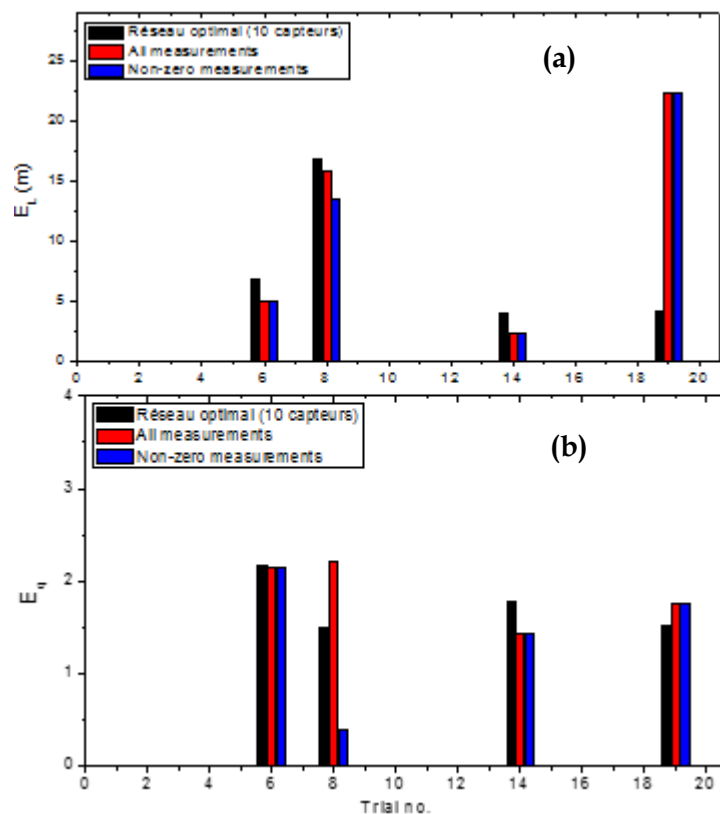


Figure 4.35. Réseaux optimaux  $m=10$  VS réseaux (Kumar et al, 2015) : (a) Erreurs de localisation  $E_l$  (m) et (b) erreurs d'estimation du débit de la source  $E_q$ .

#### 4.3.6.3. Conclusion à l'issue du problème 3

Les différents résultats obtenus dans ce cadre, montrent qu'il n'y a pas de relation évidente entre la taille du réseau et sa capacité à caractériser les sources. Ainsi, pour chaque expérience des résultats plus au moins satisfaisant sont notés, mais globalement nous pouvons conclure que les performances des réseaux optimaux sont acceptables.

Les fonctions objectifs dérivées du concept entropique et de la méthode de renormalisation s'avèrent robustes, car pour toutes les configurations testées aucune divergence des calculs n'est notée. De plus, elles permettent d'obtenir des réseaux dont la performance est



homogène (estimation précise du débit et de la position de la source). Nous avons noté à l'issue de ce problème que les réseaux obtenus pour une entropie maximale permettent de bien caractériser simultanément les paramètres des sources à savoir le débit et la position.

#### 4.4. CONCLUSION DU CHAPITRE

Dans ce chapitre, nous avons testé trois problèmes. Le premier problème vise à déterminer le nombre optimal de capteurs pour *caractériser* une source dans une situation météorologique particulière. Il a été évalué aussi bien dans le cadre expérimental de DYCE (terrain plat et sans obstacles) que celui de MUST (milieu urbain idéalisé). La fonction coût à minimiser (résidus normalisés qui examinent l'écart entre les observations et les données modélisées) s'est avérée particulièrement efficace lorsque l'on est dans un environnement sans obstacles. Pour un milieu urbain, dont les fonctions adjointes sont simulées par une approche CFD, cette même fonction coût s'est trouvée également performante. Néanmoins, pour ce cas précis, nous avons noté des divergences de calculs (en termes de position et de débit estimés), en particulier pour les réseaux de faibles tailles. Nous avons noté également qu'un minimum de la fonction coût ne correspond pas toujours à la caractérisation « précise » et simultanée du débit et de la position de la source. Ainsi, une fonction coût minimale peut correspondre à une « bonne » localisation de la source et à une « mauvaise » estimation du débit (et vice versa).

Le deuxième problème vise à déterminer un réseau optimal capable de caractériser une source inconnue localisée n'importe où dans un domaine large pour une situation météorologique particulière. Il a été testé aussi bien dans le cadre expérimental de DYCE que celui de MUST. Les résultats obtenus valident le concept d'entropie pour la caractérisation d'une source inconnue. Dans les deux cas expérimentaux, nous avons constaté la robustesse de la fonction coût entropique à maximiser (qui évalue l'information fournie par un réseau pour caractériser une source d'émission). Pour le milieu urbain, contrairement au *problème 1*, aucune divergence de calculs n'a été notée avec cette même fonction coût. Enfin, quelle que soit l'expérience considérée, les réseaux optimaux générés ont tendance à se disperser sur l'ensemble du domaine.

Le troisième problème concerne la détermination d'un réseau optimal capable de caractériser une source inconnue localisée n'importe où dans un domaine large pour des conditions météorologiques multiples. Il a été uniquement évalué dans le cadre expérimental

de MUST, et ce avec succès. La fonction coût que nous avons définie spécifiquement pour ce problème est une fonction entropique généralisée. Cette dernière s'est avérée robuste dans le cas de situations météorologiques multiples. Pour chacune des configurations, nous avons obtenu des résultats en accord avec les données expérimentales. Nous avons noté que les réseaux optimaux générés ont tendance à se disperser sur l'ensemble du domaine.

Enfin, pour les trois problèmes considérés, les analyses des Matrices des Structures Communes MSC (précisant les capteurs communs entre les réseaux optimaux de tailles différentes), ont permis de mettre en évidence la complexité à déterminer une solution optimale. En effet, elles ont montré que des réseaux présentant des structures semblables ne conduisent pas systématiquement aux mêmes performances. Nous avons également noté que des réseaux de structures différentes peuvent converger vers le même optimum (approché ou global), ceci montre que pour ces expériences la solution n'est pas unique.

« BIBLIOGRAPHIE CHAPITRE 4 »

Beljaars, A., et Holtslag, A., (1991). Flux parameterization over land surfaces for atmospheric models, *J. Appl. Meteorol.*, 30(3), 327-341, doi:10.1175/1520-0450(1991)030<0327:FPOLSF>2.0.CO;2.

Biltoft, C. A. (2001). Customer report for Mock Urban Setting Test, Tech. Rep., West Desert Test Center, U.S Army Dugway Proving Ground. Dugway, Utah.

DISCARD, (2014). REGIME D'APPUI AUX PME POUR L'INNOVATION DUALE (RAPID). Lot N°6. Application aux jeux de données expérimentales MUST (MockUrban Setting Test) pour l'étude de la performance.

Fluidyn-PANACHE, (2010). User manual. Version 4.0.7 ed., FLUIDYN France/TRANSOFT Int., Saint-Denis, France.

Gryning, S.-E., Batchvarova, E., Brümmner, E., Jørgensen, H., Larsen S., (2007). On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer, *Boundary Layer Meteorol.*, 124(2), 251-268, doi:10.1007/s10546-007-9166-9.

Kumar, P., Feiz, A.-A., Singh, S. K., Ngae, P., Turbelin, G., (2015). Reconstruction of an atmospheric tracer source in an urban like environment, *J. Geophys. Res. Atmos.*, 120, 12,589-12,604, doi:10.1002/2015JD024110.

Lepley, J. J., Lloyd, D., Robins, A., Wilks, A., Rudd, A., Belcher, S., (2011). Dynamic sensor deployment for the monitoring of chemical releases in urban environments (DYCE), *Proc. of SPIE*, 8018(12), 1-11, doi:10.1117/12.883373.

Rodgers, C.D, (2000). Inverse methods for atmospheric sounding: theory and practice. World Scientific Publishing Co. Pte. Ltd., Singapore, ISBN 978-981-02-2740-1, 85 pp.

Yee, E, et Biltoft, C., (2004). Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles, *Boundary Layer Meteorol.*, 111(3), 363-415, doi:10.1023/B:BOUN.0000016496.83909.ee.

**CONCLUSION GENERALE  
ET PERSPECTIVES**

L'objectif principal de cette thèse était de définir les méthodes nécessaires pour optimiser un réseau de surveillance conçu pour la *caractérisation* de source de rejets atmosphériques. Ainsi, nous avons défini d'une part le cadre général du problème de conception des réseaux de capteurs et d'autre part, une méthodologie à adopter afin d'en assurer la conception optimale. Nous avons également donné un aperçu détaillé des besoins spécifiques, en termes de configuration de réseaux de capteurs pour plusieurs domaines, ainsi que les moyens à adopter pour répondre à ces différents besoins. Ceci nous a permis de mettre en évidence la problématique qui relève de cette étude, à savoir la *caractérisation* de source.

L'optimisation des réseaux utilisés pour la *caractérisation* de source de contamination de l'air a été traitée, pour la première fois, par un couplage entre la méthode d'inversion de données dite de « renormalisation » et des algorithmes d'optimisation métaheuristiques (recuit simulé et algorithme génétique). Dans un premier temps, nous avons montré que la méthode d'inversion « basée récepteurs » était efficace et permettait d'estimer rapidement les caractéristiques de sources ponctuelles (position et débit). Pour sa mise en œuvre, nous avons proposé un algorithme performant. À l'aide de ce dernier, nous avons évalué la méthode pour des expériences réalisées en soufflerie (DYCE). Dans un second temps, nous avons mis en évidence que des notions issues de la « renormalisation » (résidus normalisés, entropie, visibilité) permettaient de définir des critères d'optimalité utiles pour la conception de réseaux.

Dans cette étude, le processus d'optimisation a été utilisé pour trois problèmes et dans deux cadres expérimentaux différents. Ces derniers concernent (i) un domaine plat et sans obstacles (DYCE); (ii) un domaine contenant plusieurs bâtis/obstacles (milieu urbain idéalisé, MUST). Ceux-ci ont donc permis d'analyser finement la performance du processus dans des configurations de terrains diverses.

L'optimisation a été évaluée pour un *premier problème* dont l'objectif est de déterminer la taille optimale d'un réseau permettant de *caractériser* la source de pollution à la fois en terrain plat et en milieu urbain. Pour ce faire, une fonction coût adéquate traduisant l'écart entre les observations et les données modélisées, a été identifiée. Cette fonction, dite « résidus normalisés », a été minimisée dans le cadre d'une approche d'optimisation combinatoire. Nous avons ainsi pu former des réseaux optimaux de tailles différentes où dans certains cas, nous avons montré qu'il existait une taille minimale permettant de *caractériser* la source avec précision. Cette taille n'est donc pas unique et varie selon le contexte expérimental. Nous

avons également noté que des architectures différentes de réseaux peuvent conduire au même optimum (approché ou global). Ceci confirme que pour ce problème d'optimisation, la solution n'est pas unique. Nous avons aussi montré que l'augmentation de la taille d'un réseau optimal performant (i.e. *caractérisation* précise de la source) ne conduisait pas nécessairement à des réseaux efficaces et ceci a été constaté dans différentes configurations.

Le *deuxième problème* a consisté à concevoir un réseau optimal permettant de *caractériser* une source ponctuelle inconnue pour une condition météorologique particulière. A l'aide d'un algorithme métaheuristique, nous avons pu maximiser une fonction coût entropique afin d'augmenter la quantité d'information fournie par le réseau. Cette information est utile uniquement pour la *caractérisation* de la source. Nous avons ainsi montré l'efficacité de cette démarche pour la *caractérisation* d'une source inconnue. Les performances de ces réseaux ont permis de mettre en évidence la robustesse de la notion d'information contenue dans le réseau.

Le *troisième problème* a eu pour objectif de déterminer un réseau optimal statique permettant de *caractériser* une source ponctuelle inconnue pour des configurations météorologiques multiples. Pour ce problème, nous avons défini une fonction coût entropique généralisée. Une fois maximisée, cette fonction a permis d'obtenir des réseaux optimaux performants capables de *caractériser* séparément différentes sources pour diverses conditions météorologiques. Chacune des configurations a été considérée avec la même probabilité d'occurrence.

Dans cette étude, nous étions confrontés à un problème réputé complexe dans la littérature académique et dans le monde industriel. En effet, d'après le retour d'expériences, la conception optimale des réseaux de surveillance n'est pas un défi trivial. Nous avons examiné de près cette difficulté, surtout pour la détermination de solutions optimales. Nous avons approché ce problème dans le cadre d'une optimisation combinatoire, qualifiée de *NP-difficile*. Ainsi, nous avons pu identifier les moyens algorithmiques adéquats. Ces algorithmes une fois mise en œuvre, ont fait apparaître des complexités que nous avons dû gérer. Outre le problème de paramétrage très empirique des algorithmes métaheuristiques, nous avons pu stabiliser les divergences des calculs constatés pour certaines configurations. Ces divergences ont été uniquement remarquées pour des environnements urbanisés. Aussi, pour remédier à ce problème, nous avons proposé une technique d'optimisation hybride

« probabiliste & déterministe » capable de produire des solutions performantes à partir d'une structure de base obtenue par une technique de recherche probabiliste.

En guise de perspectives, plusieurs questions restent toutefois posées :

- Une question relative à la détermination du *degré de confiance* associé à chaque réseau optimal pour l'ensemble des problématiques abordées.
- Dans cette étude, les *positions potentielles* allouées pour l'implantation d'un réseau étaient préalablement prédéfinies. Ceci a conduit à optimiser des réseaux sous cette contrainte. En vue d'améliorer la performance des réseaux, est-il possible d'adopter une méthodologie permettant de prédéterminer les positions potentielles au sein d'une zone d'implantation ?
- Il serait enfin intéressant de prolonger cette étude dans l'objectif d'étendre l'opérationnalité des réseaux conçu pour la *caractérisation* de la source. Autrement dit, est-il possible de concevoir des réseaux optimaux capables de répondre simultanément à des besoins différents, par exemple en associant l'objectif de la *mise en alerte précoce* à l'objectif de la *caractérisation* de la source ?

## Annexe I

### « Approches de stabilité et modèles des écarts types de dispersion »

#### I. Approches de Stabilité

##### 1. L'approche de Doury

À partir du gradient thermique vertical de l'atmosphère  $\Gamma$  deux classes de stabilité atmosphérique sont établies (Doury, 1980)

- Une première classe à laquelle la diffusion est considérée comme "diffusion normale"

$$\Gamma \leq -0,5^\circ \frac{C}{100} m$$

- La seconde classe indique la "diffusion faible" est conditionnée pour un gradient thermique

$$\Gamma > -0,5^\circ \frac{C}{100} m$$

D'après Doury, la situation de diffusion, normale ou faible, dépend d'une vitesse seuil du vent. La diffusion faible s'établit la nuit pour des vitesses du vent inférieures à 3 m/s tandis que la diffusion normale apparaît le jour, ou la nuit avec des vitesses de vent supérieures à 3 m/s. Pour indication, le mode de diffusion classé normal par Doury correspond aux classes de stabilité A, B et C de Pasquill et la diffusion classée faible correspond aux classes D et F.

##### 2. L'approche de Détrie, (1969), Triplet et Roche (1986)

Cette approche, se base sur les hypothèses que l'air saturé d'humidité se refroidit de 0,55°C/100 m, cette quantité représente le gradient adiabatique saturé d'humidité, alors que l'air sec se refroidit de 0,98°C/100 m (gradient adiabatique sec) et définit la stabilité atmosphérique à partir du gradient thermique, sous la dépendance des variations thermiques avec l'altitude. Trois classes de stabilité atmosphérique sont proposées par cette approche : l'état neutre, l'état stable et l'état instable.

- Atmosphère stable : lorsque la situation de l'atmosphère influe considérablement la dispersion des nuages des espèces et cela quelle que soit leur densité. Dans cette situation, une particule déplacée de sa position initiale y retourne spontanément. Sur la figure 1 est présentée une atmosphère stable.



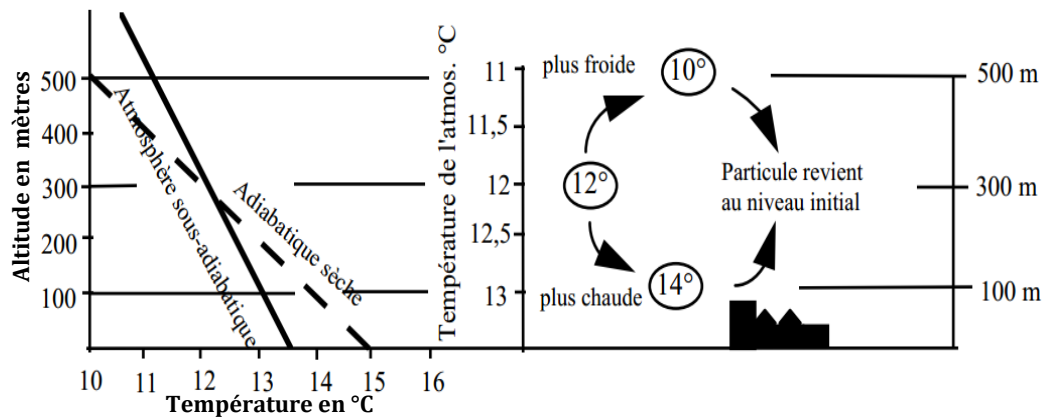


Figure 1. Atmosphère stable (Détrie, 1969).

- Atmosphère instable : dans ce genre de situation, la décroissance de température est supérieure au gradient adiabatique sec, les conditions réelles de température sont dites sur-adiabatiques, de ce fait, une particule déplacée de sa position initiale tend à s'éloigner spontanément. Sur la figure 2 est présentée une atmosphère instable.

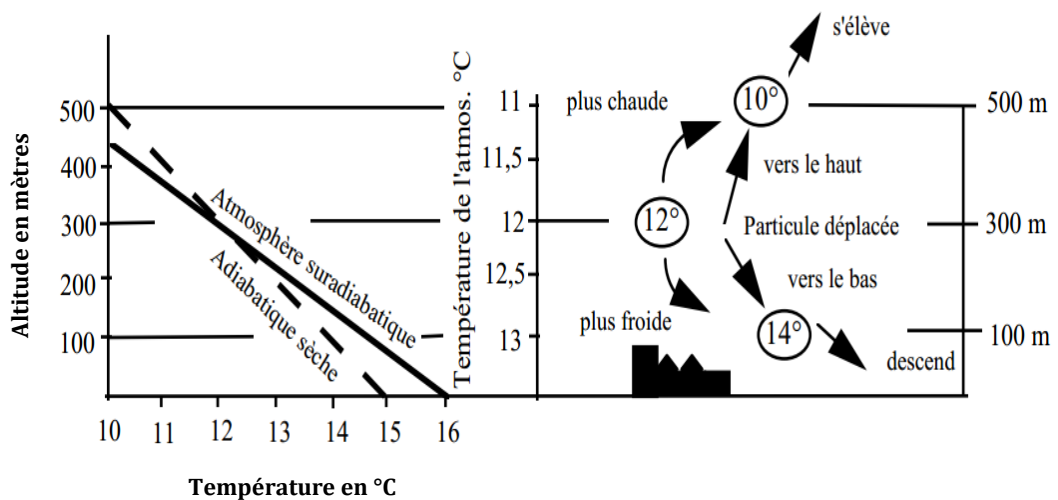


Figure 2. Atmosphère instable (Détrie, 1969).

- Atmosphère neutre : dans cette atmosphère, quelle que soit la position d'une particule dans la masse d'air, elle est à la même température que l'atmosphère donc à la même densité vu que la décroissance verticale de la température est égale au gradient adiabatique sec. Sur la figure 3 est présentée une atmosphère neutre.

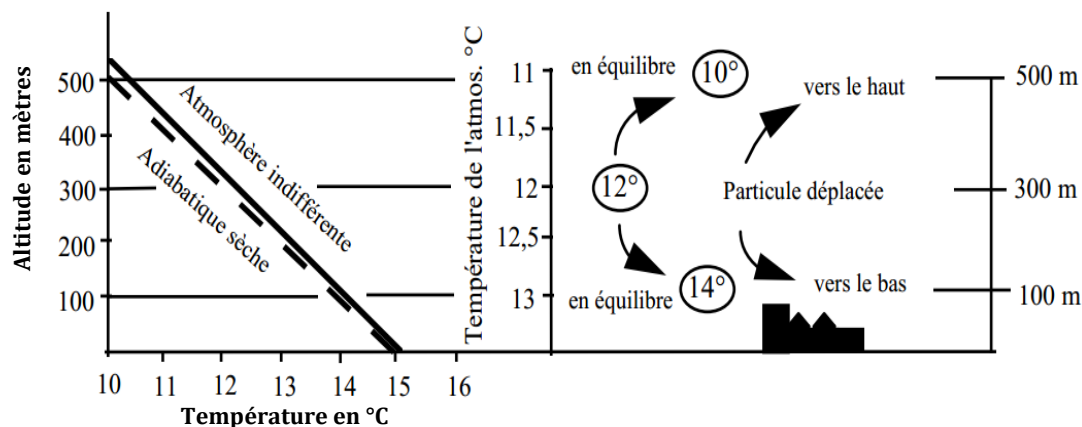


Figure 3. Atmosphère neutre (Détrie, 1969)

### 3. L'approche de US-EPA

Cette approche est proposée par l'agence américaine de protection de l'environnement l'US-EPA, et met en relation les classes de stabilité de Pasquill avec les écarts types de la fluctuation de la direction du vent horizontal qui leurs sont associés. Le Tableau 1 ci-dessous regroupe les classes de stabilité de Pasquill et les écarts types correspondants.

Classes de Stabilité	Ecart-type de la fluctuation de la direction du vent horizontal
A	> 22,5 °
B	17,5 ° - 22,5°
C	12,5° - 17,5°
D	7,5 ° - 12,5°
E	3,8 ° - 7,5°
F	< 3,8°

Tableau 1 Classes de stabilité de Pasquill selon la fluctuation de la direction du vent (Zannetti, 1986).

## II. Écarts types

### 1. Modèle ADMS

ADMS (Atmospheric Dispersion Modelling System) est un modèle de dispersion atmosphérique avancé. Il a été développé par Cambridge Environmental Research Consultants (CERC) du Royaume-Uni en collaboration avec l'Office météorologique du Royaume-Uni, National Power plc (maintenant Innogy Holdings plc) et de l'Université de Surrey. La première version d'ADMS a été délivrée en 1993. Le modèle utilise un préprocesseur météorologique permettant de calculer les paramètres de la couche limite atmosphérique : hauteur (h) et longueur de Monin-Obukhov (L) (CERC, 2009). Ces

paramètres sont utilisés pour définir les profils verticaux de vent et de turbulence selon la stabilité atmosphérique. ADMS utilise une distribution gaussienne de concentration pour calculer la dispersion des rejets en condition stable et neutre (Hunt, 1985 ; Weil, 1985) et une distribution asymétrique dans le cas de conditions convectives (Carruthers et al. 2009a ; Lamb, 1982). Par ailleurs, il possède plusieurs modules permettant de prendre en considération l'effet de rabattement de bâtiments à proximité dans la trajectoire du panache de pollution en dispersion (Carruthers et al, 2009b ; Jackson et al, 1975). Ces modules prennent également en considération, les effets de relief complexe, les effets de bord de mer, les dépôts humides, la sédimentation gravitationnelle, les dépôts secs, la fluctuation à court terme de concentration des polluants, les réactions chimiques, la désintégration radioactive et les gamma-dose...

Les écarts types de dispersion dépendent de la hauteur de rejet  $z_0$ , de la turbulence atmosphérique, et du temps de transfert,  $t$ , depuis la source :

$$\sigma_z = \sigma_w t \left[ \frac{1}{b^2} + \frac{N^2 t^2}{1 + 2Nt} \right]^{-1/2}$$

où :

$\sigma_w$  écart type de la vitesse verticale,

$b$  un paramètre fonction du rapport  $z_0/h$  et de la vitesse de friction  $u$ ,

$N$  fréquence de Brunt Väisala donnée par

$$N = \left[ \frac{g}{\theta} \frac{\partial \theta}{\partial z} \right]^{-1/2}$$

$\theta$  température potentielle

$g$  accélération de la pesanteur.

La fréquence de Brunt Väisala correspond à l'oscillation verticale d'une portion d'air dans l'atmosphère stratifiée. Le paramètre de dispersion horizontale est calculé par

$$\sigma_y^2 = \sigma_{yt}^2 + \sigma_{yw}^2$$

Avec

$$\sigma_{yt} = \sigma_v t \left[ 1 + 2.5 \frac{U_* t}{h} \right]^{-1/2}$$

$$\sigma_{yw} = \sigma_{\theta x}$$

où :

$\sigma_v$  écart type horizontal de la direction moyenne du vent

$\sigma_\theta$  écart type vertical de la direction moyenne du vent qui est soit défini en entrée parmi les paramètres météorologiques, soit calculée par le préprocesseur météorologique comme suite

$$\sigma_\theta = \frac{\sigma_w}{u_{10}} = 0.065 \left( \frac{7 T}{u_{10}} \right)$$

où :

$T$  temps d'intégration en heure

$u_{10}$  vitesse moyenne du vent à 10 m (*Weil, 1985 ; Hunt 1985*).

En condition d'instabilité, ADMS adopte une distribution de la concentration asymétrique et non gaussienne, (*Briggs, 1985*) et de (*Lamb, 1982*).

Dans une couche limite convective, la hauteur à laquelle la concentration est maximale dans le panache diminue pendant le transport du panache alors que l'épaisseur moyenne du panache augmente. Une fois que la concentration maximale a atteint le sol, sa hauteur augmente à nouveau. Les travaux de (*Willis et al, 1976, 1978, 1981*) menés en chambre de simulation illustrent la distribution de panaches de polluants atmosphériques dans des conditions convectives. La figure 4 présente les contours de la concentration intégrée dans la direction perpendiculaire au vent en fonction des distances sans dimension  $X^*$  et  $Z$  pour des rejets non turbulents.

La dispersion horizontale est la somme des contributions des processus de convection thermique  $\sigma_{yc}$  de turbulence mécanique  $\sigma_{yn}$  et de variation de la direction du vent  $\sigma_{yw} = \sigma_\theta \times$

$$\sigma_y^2 = \sigma_{yc}^2 + \sigma_{yn}^2 + \sigma_{yw}^2$$

avec

$$\sigma_{yc} = \sigma_{vc} t \left[ 1 + \frac{t}{h} \sqrt[3]{0.75 w_*} \right]^{-1/2}$$

$$\sigma_{yn} = \sigma_{vn} t \left[ 1 + \frac{t}{h} \sqrt[3]{15.6 w_*} \right]^{-1/2}$$

où :

$\sigma_{vc}$  et  $\sigma_{vn}$  écarts types de la vitesse horizontale générés par la turbulence thermique et mécanique.  $w_*$  échelle de vitesse convective, donnée par

$$w_* = \sqrt[3]{\frac{hu_*^3}{K|L|}}$$

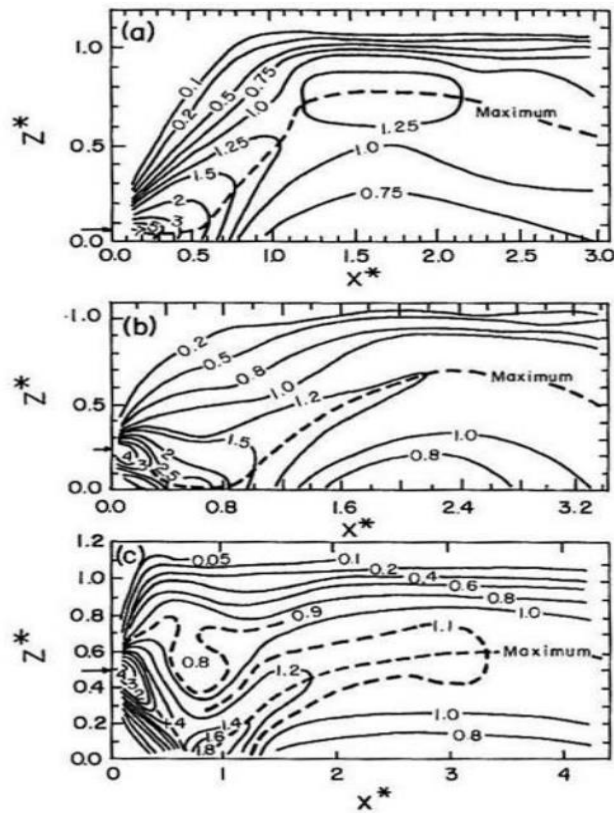


Figure 4. Contours de la concentration intégrée dans la direction perpendiculaire au vent en fonction des distances sans dimension  $X^*$  et  $Z$  pour des rejets non turbulents (*Willis et Deardoff, 1976, 1978, 1981*).

Pour le paramètre de dispersion verticale  $\sigma_z$ , la trajectoire du panache la nature oblique conduit à définir des écarts-types de vitesse verticale turbulente d'ascendance et de subsidence :  $\sigma_{w^+}$  et  $\sigma_{w^-}$ . Deux écarts type de dispersion verticale, un premier positif et un deuxième négatif (*Carruthers et al. 2009a*) sont définis et donnés par

$$\sigma_{z^+} = \sigma_{w^+} t \left[ 1 + \frac{t}{2T_L} \right]^{-1/2}$$

$$\sigma_{z^-} = \sigma_{w^-} t \left[ 1 + \frac{t}{2T_L} \right]^{-1/2}$$

où :

$T_L$  échelle de temps intégrale lagrangienne.

## 2. Modèle de Weill et Brower

Ce modèle constitue une actualisation du modèle de diffusion de panache gaussien pour les grandes cheminées et se base sur les mêmes conditions de développement de l'approche Briggs. Pour les classes de stabilité variant de A à D le modèle de Weill et Brower redéfinit les coefficients de dispersion en fonction du rapport de la vitesse moyenne sur l'échelle de vitesse convective  $\bar{u}/w_*$ . Les nouveaux écarts types sont donnés par

$$\sigma_y(x) = \sqrt{0.0064 + \left(0.56 \frac{w_*}{\bar{U}}\right)^2} \cdot x$$

$$\sigma_z(x) = \sqrt{0.0036 + \left(0.56 \frac{w_*}{\bar{U}}\right)^2} \cdot x$$

où :

$w_*$  échelle de vitesse convective.

## 3. Modèle split sigma theta Cirillio et Poli (1992) (SST)

Cette approche a la réputation de prendre en compte la sinuosité du panache. La formulation introduit des valeurs observées de  $\sigma_\theta$  et les coefficients de dispersion sont donnés par (Cirillio et al, 1992)

$$\sigma_x = x \cdot \sqrt{\cosh(\sigma_\theta^2) - 1}$$

$$\sigma_y = x \cdot \sqrt{\sinh(\sigma_\theta^2)}$$

où :

$\sigma_\theta$  déterminée sur une moyenne horaire.

## 4. Modèle de Doury

Historiquement pour les besoins des autorités de sûreté nucléaire françaises, les formulations de (Doury, 1976) ont été conclues de plusieurs campagnes de mesure.

Deux classes de stabilité sont définies suivant le gradient vertical de température (GVT),

- Diffusion normale pour  $GVT \leq -0,5 \text{ °C}/100 \text{ m}$
- Diffusion faible pour  $GVT > -0,5 \text{ °C}/100 \text{ m}$

Les écarts type de Doury sont donnés par

$$\sigma_y = (A_h t)^{k_h}$$

$$\sigma_z = (A_h t)^{k_z}$$

où :

$K_h$  et  $K_z$  exposants de diffusion horizontale et verticale

$A_h, A_z$  paramètres de diffusion horizontale et verticale

Les valeurs d' $A_h, A_z, k_h$  et  $k_z$  sont données en fonction du temps de transfert de la masse d'air.

Selon (Doury, 1981), seule la distribution verticale est sensible à la stabilité atmosphérique

Temps de transfert (s)	$A_h$ [m <sup>1</sup> /k/s] k = $K_h$ ou $K_z$	$A_z$	$k_h$	$k_z$
0 < t < 240	0,405	0,42	0,859	0,814
240 < t < 3 280	0,135	1,00	1,130	0,685
3 280 < t < 97 000	0,135	20,0	1,130	0,500
97 000 < t < 508 000	0,463	20,0	1,000	0,500
508 000 < t < 1 300 000	6,50	20,0	0,824	0,500
t > 1 300 000	2 × 10 <sup>5</sup>	20,0	0,500	0,500

Tableau 2 Coefficients relatifs aux écarts-types pour une « Diffusion Normale »  
(Doury 1980)

Temps de transfert (s)	$A_h$ [m <sup>1</sup> /k/s] k = $K_h$ ou $K_z$	$A_z$	$k_h$	$k_z$
0 < t < 240	0,405	20,0	0,859	0,500
240 < t < 3 280	0,135	20,0	1,130	0,500
97 000 < t < 508 000	0,463	20,0	1,000	0,500
508 000 < t < 1 300 000	6,50	20,0	0,824	0,500
t > 1 300 000	2 × 10 <sup>5</sup>	20,0	0,500	0,500

Tableau 3 Coefficients relatifs aux écarts-types pour une « Diffusion Faible »  
(Doury 1980)

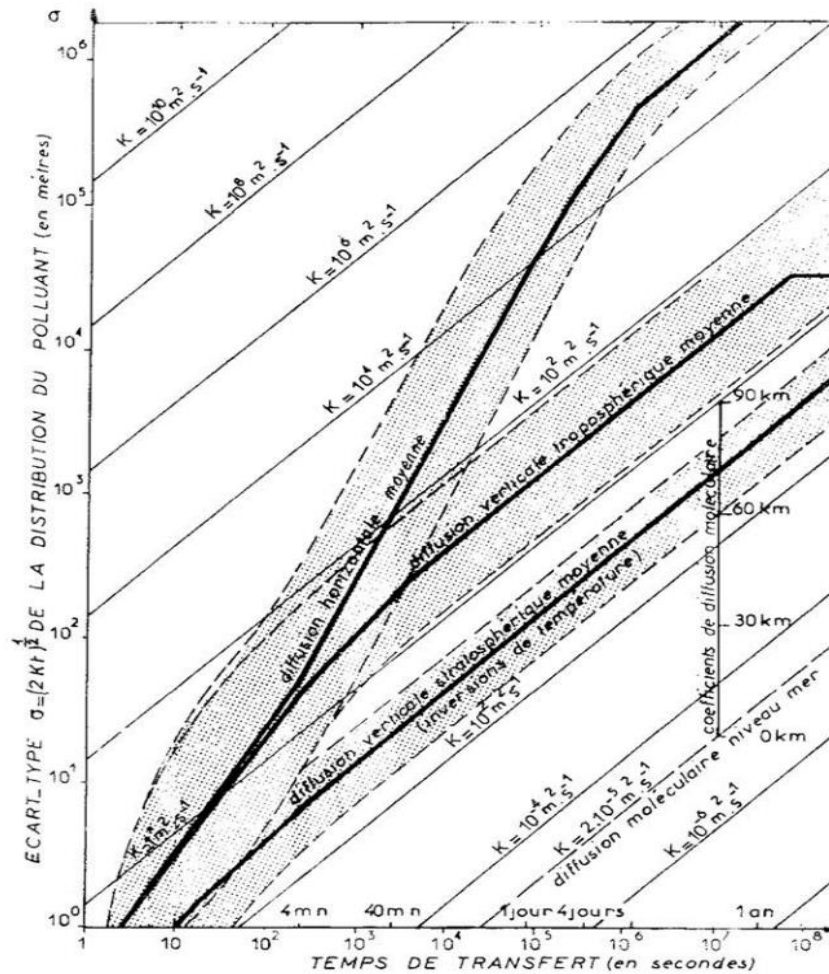


Figure 5. Écarts types de dispersion horizontale et verticale établis par (Doury 1976)

## 5. Coefficients de diffusivité verticale $K_z$

### 5.a. Modèle de Businger et Arya (1974)

Ce modèle est applicable pour n'importe quelle condition atmosphérique stable, instable ou neutre, le coefficient de diffusivité verticale est donné par

$$K_z = \frac{k U_* z}{0.74 + 4.7(z/L_{MO})} \exp\left(-8 \frac{fz}{U_*}\right)$$

Où :

$K$  constante de von Karman

$L_{MO}$  longueur de Monin-Obukhov

$F$  force de Coriolis

$Z$  hauteur de la CLA

$u_*$  vitesse de friction.



5.b. *Modèle de Similitude de Panofsky et Dutton (1984)*

Cette approche est adéquate pour les conditions neutres et stables, le coefficient de diffusivité verticale est de la forme

$$K_z = \frac{kU_*z}{\phi_h(z/L_{MO})}$$

La fonction  $\phi_h$  est déduite de la corrélation de (Dyer, 1965)

$$\phi_h = 1 + 5 \frac{z}{L_{MO}}$$

5.c. *Modèle de Troen et Mahrt*

Ce modèle est applicable pour les conditions stables et instables.

$$K_z = \frac{kU_*z (1 - \frac{z}{h})^2}{\phi_h(z/L_{MO})}$$

avec

$\phi_h$  Fonction de (Dyer, 1965)

« BIBLIOGRAPHIE ANNEXE I »

- Doury A, Gerard R, Picol M. (1980). Abaques d'évaluation directe des transferts atmosphériques d'effluents gazeux. Rapport CEA/ DSN 1980; n884 1980 [57 p].
- Détrie J.-P. et Jarrault Mme P. (1969). La pollution atmosphérique. Revue de géographie alpine, Année 1969, Volume 57, Numéro 4. p. 865 – 866.
- Triplet, J.P. and Roche, G. (1986). The influence of free stream turbulence on the local heat transfer from cylinders. *J. Heat Transfer*, 82, 101-107.
- CERC. 2009, ADMS 4.0 4 User Guide. Cambridge Environmental Research Consultants (CERC).
- Hunt J.C.R. (1985). Turbulent diffusion from sources in complex flows. *Ann. Rev. fluid. Mech.* 17, 447-458.
- Weil J.C. (1985). Updating applied diffusion models. *J. Clim. Appl. Met.*, 24, 1111-1130.
- Carruthers D.J., Weng W.S., Dyster S.J., Singles R. and Higson H. (2009a). Complex terrain module. Tech. rept. ADMS 4.0, P14/010/09. Cambridge Environmental Research Consultants (CERC).
- Lamb R.G. 1982, Diffusion in the convective boundary layer; atmospheric turbulence and air pollution modelling, 159-230. Eds F.T.M. Nieuwstadt and H. Van Dop, D. Reidel, Dordrecht.
- Carruthers D.J., Weng W.S., Hunt J.C.R., Holroyd R.J., McHugh C.A. and Dyster S.J. (2009b). Plume/Puff spread and mean concentration module specifications. Tech.
- Jackson P.S. and Hunt J.C.R. (1975). Turbulent wind flow over a low hill. *Q. J. R. Meteorol. Soc.* Vol101, 929-955.
- Briggs G. A. (1985). Analytical parametrizations of diffusion—the convective boundary layer, *J. Clim. Appl. Met.* 14, 1167-1186.
- Willis G.E. and Deardoff J.W. (1976). A laboratory model of diffusion from into the convective boundary layer. *Quart. J. Roy. Meteorol. Soc.* 102, 427-445.
- Willis G.E. and Deardoff J.W. (1978). A laboratory study of dispersion from an elevated source in a modified convective planetary boundary layer. *Atmos. Environ.* 12, 1305- 1311.
- Willis G.E. and Deardoff J.W. (1981). A laboratory study of dispersion from a source in the middle of the convective planetary boundary layer. *Atmos. Environ.* 17, 2435- 2447.
- Cirillio, M.C., Poli, A.A., (1992). An inter-comparison of semiempirical diffusion models under low wind speed, stable conditions. *Atmospheric Environment* 26A, 765- 774.
- Doury A. (1981). Le vademécum des transferts atmosphériques, rapport n°CEADSN- 440, CEA.
- BUSINGER. J. et ARYA .S. (1974). Height of the mixed layer in the stably stratified planetary boundary layer. *Advan. Geophys.*, 18A:73–92.

Panofsky, H.A., Dutton J.A. (1984). *Atmospheric Turbulence: Models and Methods for Engineering Applications*. John Wiley and Sons, NY, 417 pp.

DYER A. (1965). The flux-gradient relation for turbulent heat transfer in the lower atmosphere. *Quart. J. Roy. Meteorol. Soc.*, 91:151-157.

## Annexe II

### « Algorithme de calcul de la fonction de pondération optimale »

---

```
function [IHw,w] = renorm(A,pre)


---


%           " Function RENORM "
% Compute the inverse of the weighted Gram matrix
% and the optimal weights
% Input parameters:
%   A   - the [mxN] sensibility matrix
%   pre - the precision parameter (for defining the
%         stopping criteria)
% Output:
%   IHw - the [mxm] inverse weighted Gram matrix
%   w   - the vector of optimal weights (dimension N)


---


tA=A';
m=size(A,1);
N=size(A,2);
w=(m/N)*ones(N,1);

while ~(p_min>= (1-pre) & p_max <=(1+pre))
    tAw=bsxfun(@rdivide,tA,w);
    IHw=inv(tAw'*tA);
    p=sum(tAw*IHw.*tAw,2);
    p_min=min(p);
    p_max=max(p);
    w=w.*p.^0.5
end


---

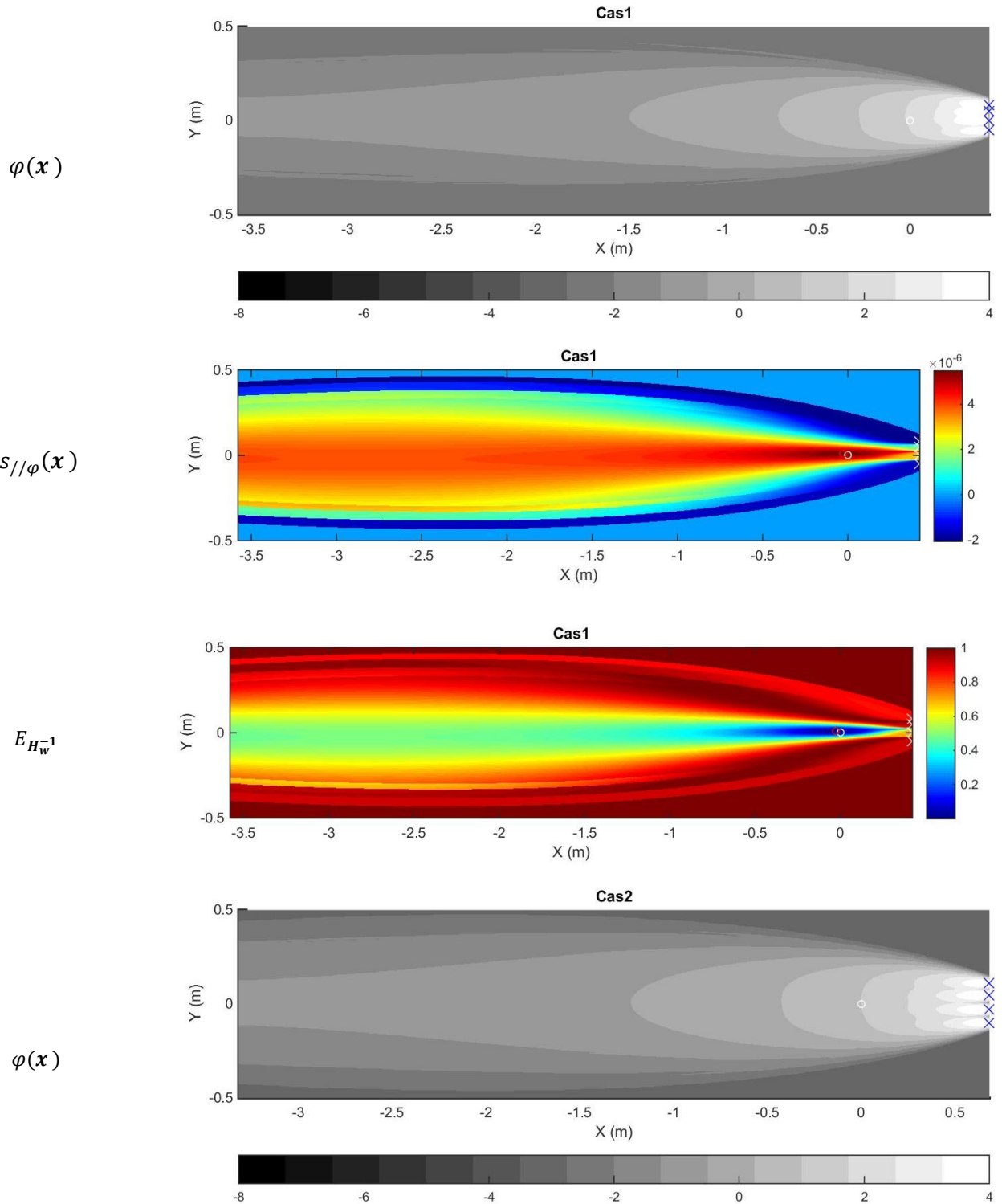

end

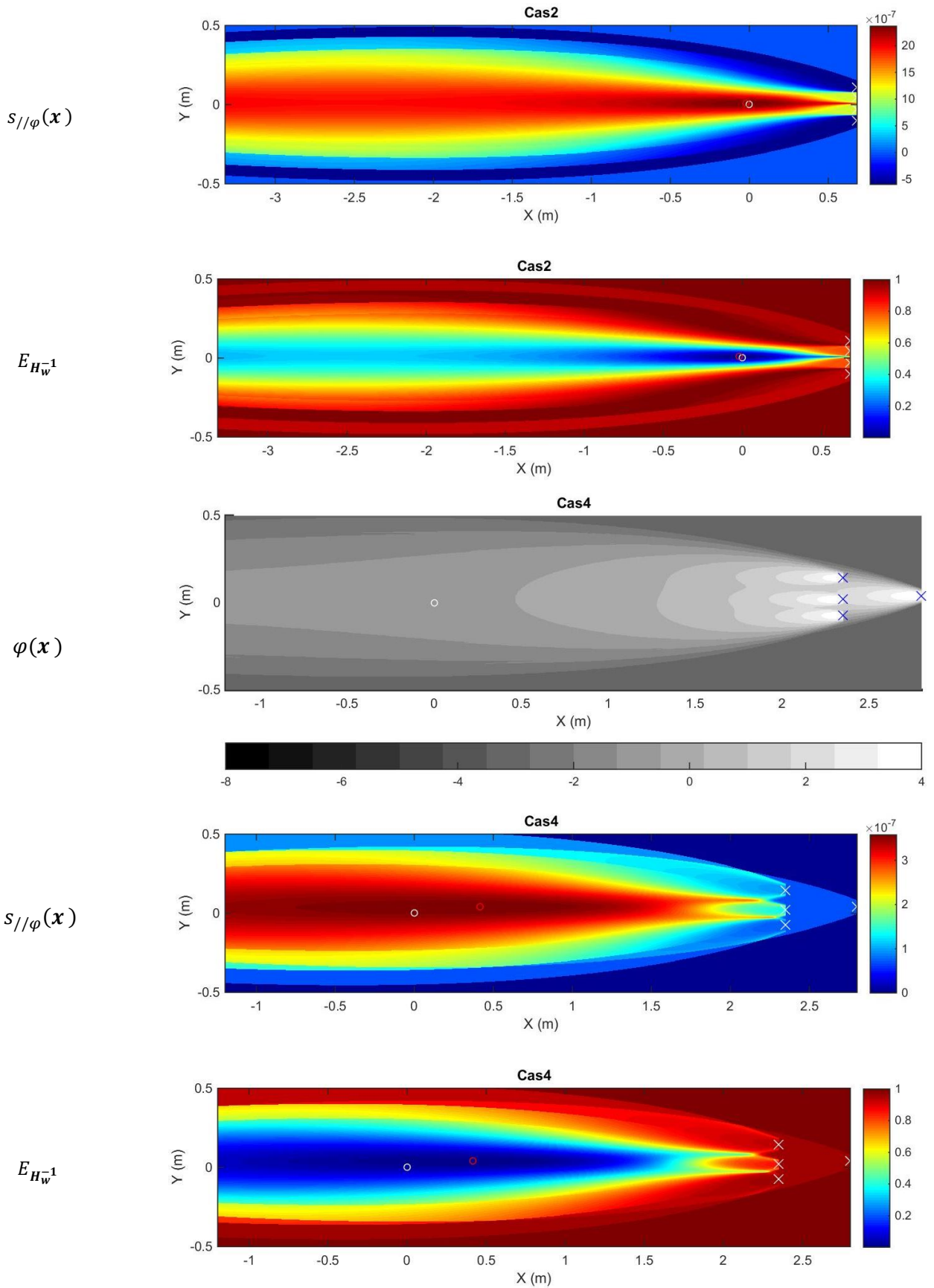

---

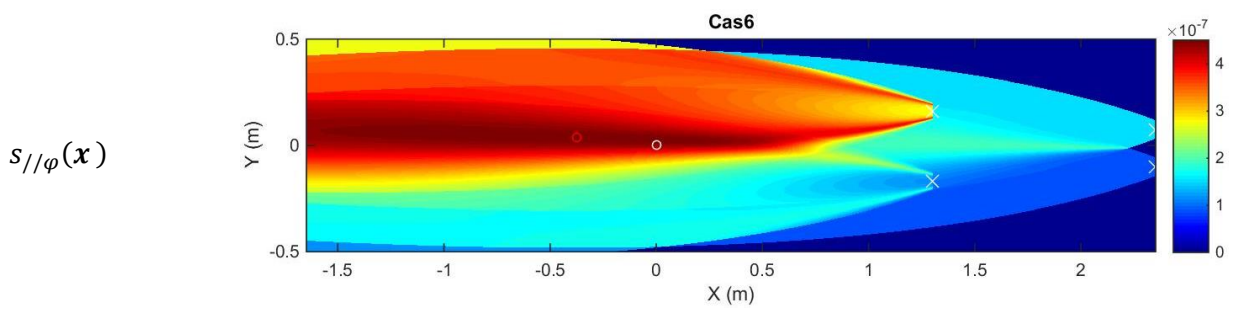
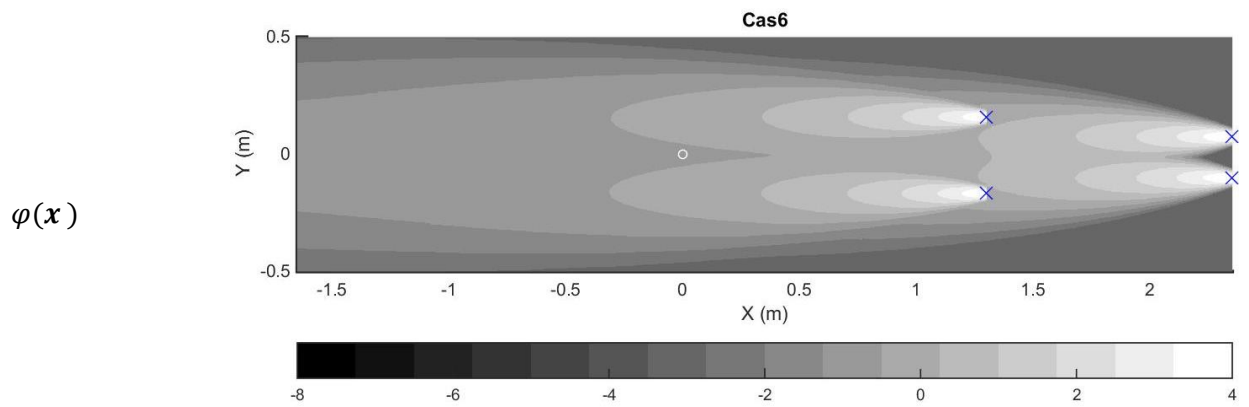
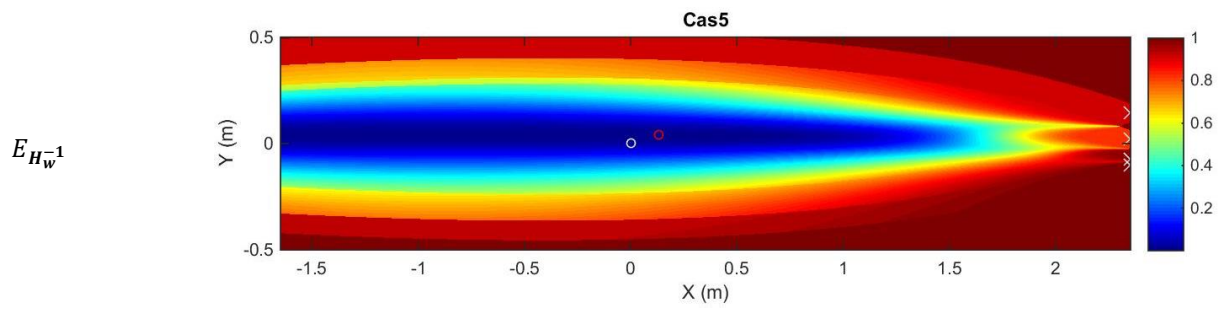
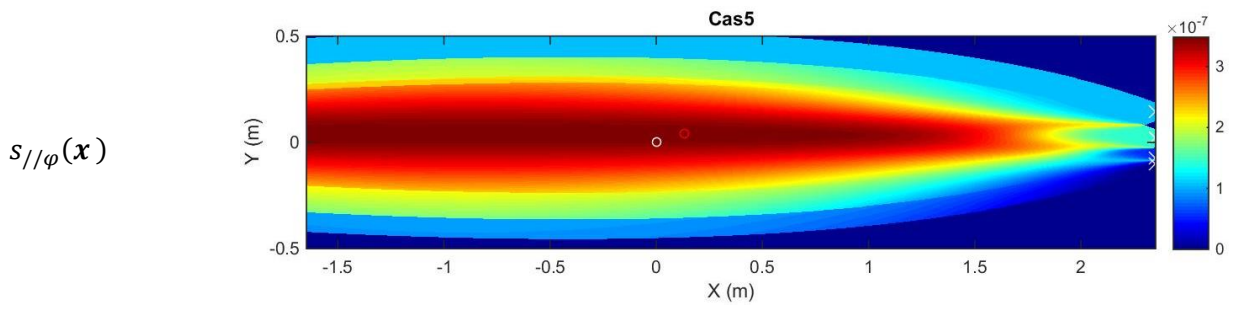
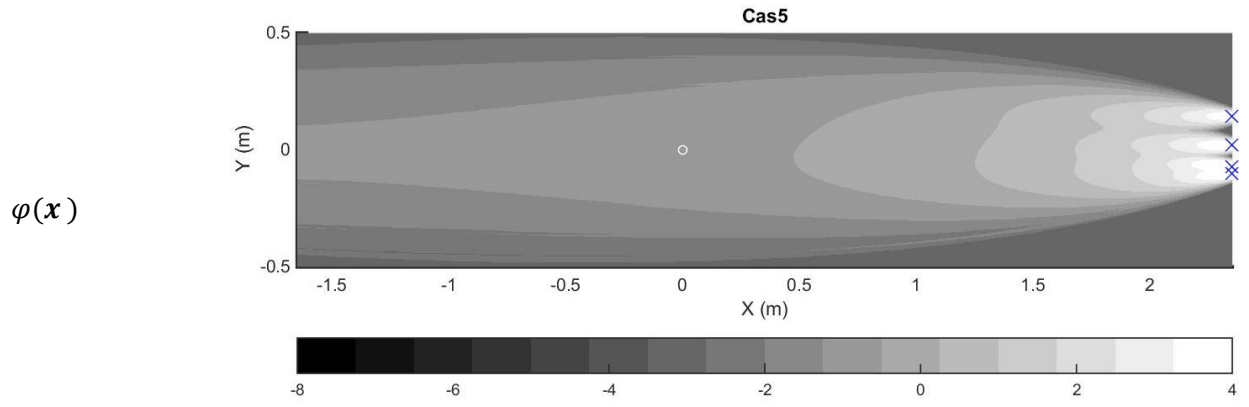

```

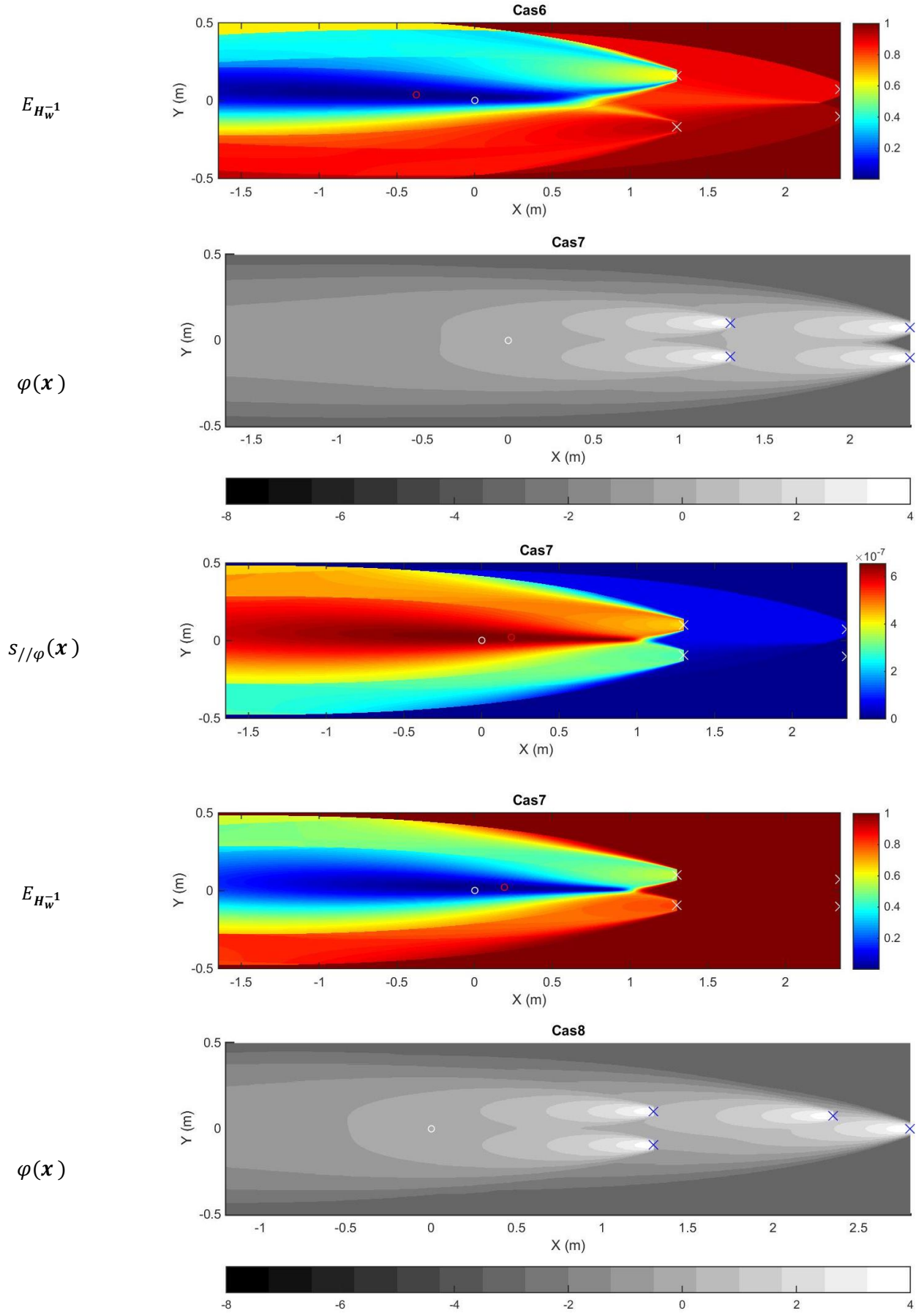
**Fonctions de visibilité  $\varphi(x)$ , (échelle logarithmique),  $s_{//\varphi}(x)$  et résidus normalisés  $E_{H_w^{-1}}$  des configurations 1, 2, 4, 5, 6, 7, 8, 9 et 11 du projet DYCE.**

La position réelle de la source correspond au rond blanc, la position estimée au rond rouge.  
Les capteurs sont symbolisés par des croix

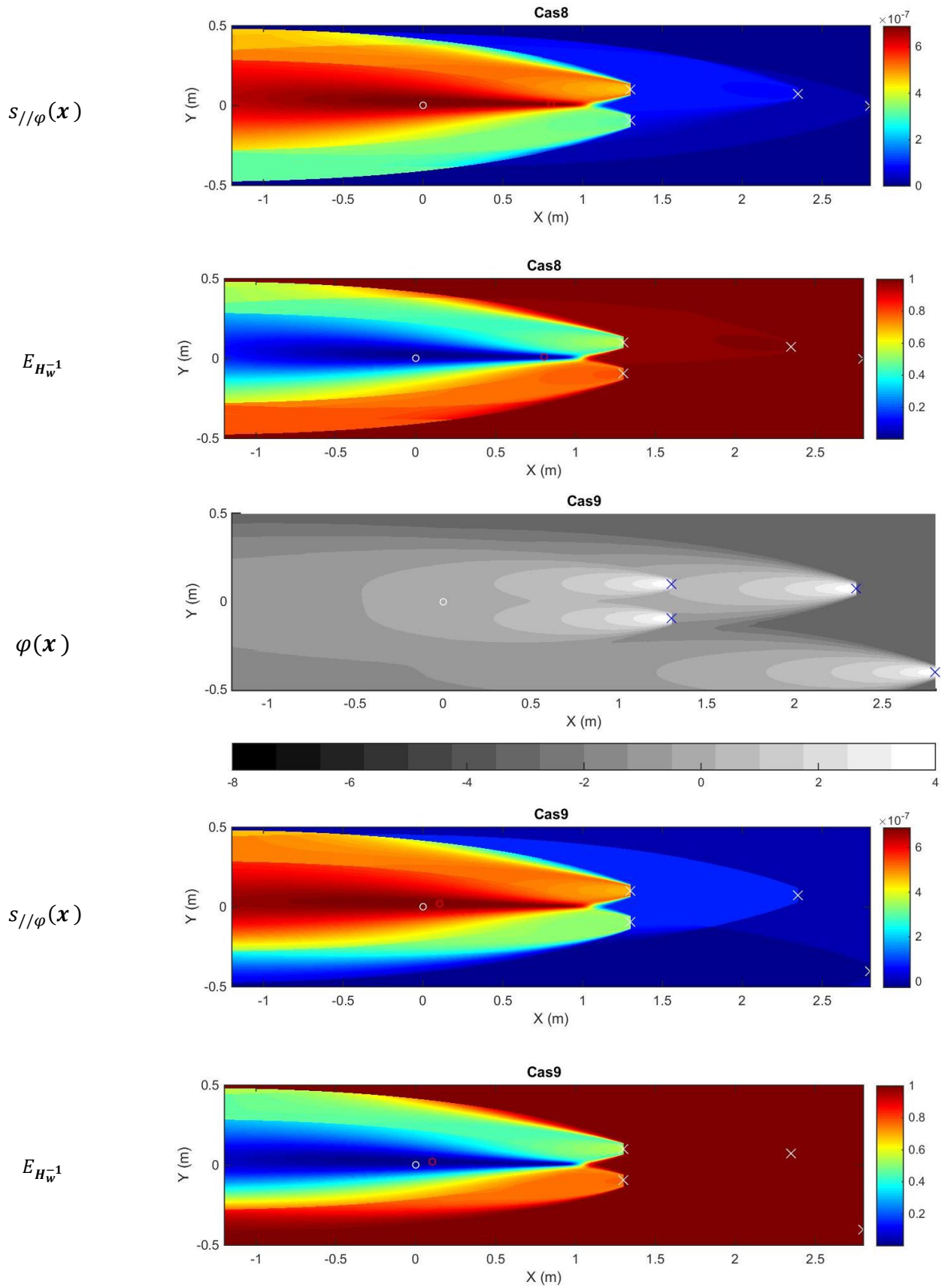


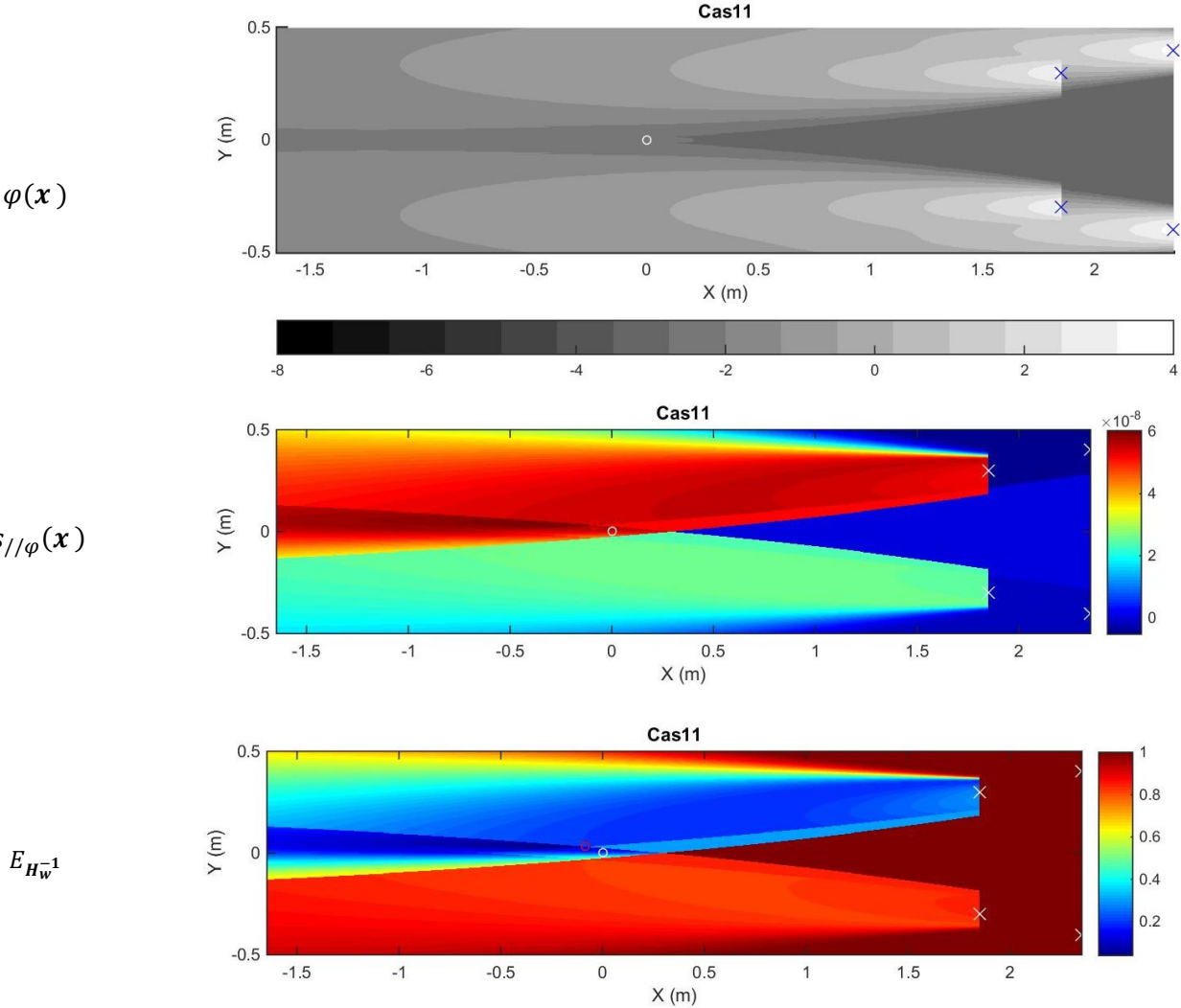












## Annexe III

### « Modèle de turbulence RANS et conditions d'entrée »

#### I. Modélisation des grandeurs turbulentes de la CLA

Pour rappel, la représentation d'une couche limite atmosphérique passe par la résolution des équations de Navier-Stokes moyennées, mais aussi par la détermination d'un modèle de turbulence, qui dans notre cas sera un modèle de type  $k$ - $\varepsilon$ . Par conséquent, le modèle de fermeture doit faire intervenir deux autres grandeurs que sont l'énergie cinétique turbulente ( $k$ ) et son taux de dissipation ( $\varepsilon$ ). Les profils de turbulence (vitesse et température) sont obtenus à partir d'une équation pronostique (utilisant une fermeture 1D) pour l'énergie cinétique turbulente :

$$\begin{aligned} \frac{\partial}{\partial z} \left( v_l + \frac{v_t}{\sigma_k} \right) \frac{\partial k}{\partial z} + P_k + P_b - \varepsilon &= 0 \\ \frac{\partial}{\partial z} \left( v_l + \frac{v_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial z} + \frac{\varepsilon}{k} [C_{\varepsilon 1} (P_k + C_b P_b) - C_{\varepsilon 2} \varepsilon] &= 0 \end{aligned}$$

où  $P_k = v_t \left( \frac{\partial U}{\partial z} \right)^2$  correspond au terme de production d'énergie cinétique turbulente due aux forces de cisaillement. Et  $P_b = v_t \frac{\beta}{Pr_t} g \frac{\partial \theta}{\partial z}$  correspond à la production due à la stratification thermique. Dans ce modèle, la viscosité turbulente,  $v_t$ , est obtenue à partir de l'énergie cinétique turbulente,  $k$ , et de son taux de dissipation,  $\varepsilon$ . On a :

$$v_t = C_\mu \frac{k^2}{\varepsilon}$$

où  $C_\mu$  est déterminé expérimentalement. Dans le modèle original de [\(Launder and Spalding, 1974; Kumar et al, 2015\)](#), les valeurs des coefficients de fermeture sont ajustées expérimentalement afin de représenter la physique de l'écoulement. Les valeurs originales de ces coefficients sont :  $C_{\varepsilon 1}=1.44$ ,  $C_{\varepsilon 2}=1.92$ ,  $C_b=1$ ,  $C_\mu=0.09$ ,  $\sigma_k=1$  et  $\sigma_\varepsilon=1.3$ . Dans cette étude, la valeur du coefficient  $C_b$  a été modifiée, afin de s'assurer que les distributions empiriques (basées sur la théorie de Monin-Obukhov) soient solutions du modèle de turbulence. Pour ce faire, des profils empiriques du modèle de turbulence  $k$ - $\varepsilon$  pour des conditions de stratification thermique stable ont été obtenus par [\(Han et al, 2000\)](#). Le problème étudié étant

un écoulement de couche limite turbulente complètement développée, les distributions verticales des champs de vitesse, de température ( $\theta$ ), de  $k$  et  $\varepsilon$ , ne dépendent que de la position verticale,  $z$ , et de la longueur de Monin-Obukhov,  $L$ . Les solutions obtenues montrent que  $C_b$  n'est pas constant mais varie avec la hauteur  $z$ . La valeur moyenne de  $C_b$ , obtenue en condition de stabilité thermique stable, est de -0.4 (*Peters and Baumert, 2007*).

Une caractérisation de la turbulence pour un écoulement est l'existence de mouvements (visualisés par des "tourbillons") balayant un large spectre d'échelles spatiales et temporelles. Ceci conduit à augmenter les propriétés de brassage de l'écoulement en transférant de l'énergie cinétique des grandes échelles vers les petites échelles, où cette énergie est dissipée en chaleur (via la diffusion moléculaire). Les effets de couche limite induisent la forte turbulence de la Couche Limite Atmosphérique (CLA). On distingue deux effets :

- la couche limite dynamique liée aux effets de frottement sur le sol induite par la condition d'adhérence pour le vent ( $V = V_{\text{sol}}$  au sol) ;
- la couche limite thermique liée aux cycles de température du sol : réchauffement et refroidissement du sol lors d'un cycle diurne (gradient vertical de température).

La présence des obstacles, constituant un groupe de bâtis interagissant entre eux, peut sensiblement affecter la dispersion. Ce phénomène est ainsi associé à un effet de ralentissement du champ horizontal de vitesse au voisinage du sol (du fait de la propagation de la condition d'adhérence  $V = 0$ ) : il est donc essentiellement lié à la stratification verticale. La prise en compte implicite des containers doit ainsi introduire des paramétrisations appropriées dans le calcul du panache.

Habituellement trois états de la CLA sont distingués :

- CLA neutre, peu fréquente ;
- CLA instable, pour laquelle les polluants sont fortement brassés par la turbulence ;
- CLA stable (souvent nocturne) où les polluants restent accumulés au sol.

Dans le cas instable, les deux effets (dynamique et thermique) génèrent de la turbulence. Dans le cas stable, elle est uniquement due aux effets dynamiques. Dans ce cas, la turbulence est détruite par les effets de flottabilité (conjonction de la force de gravité et de la poussée d'Archimède).

La théorie de Monin-Obukhov permet de décrire (dans les premières dizaines de mètres de l'atmosphère) ces gradients verticaux de vitesse et de température en conditions neutre et non-neutre. Son principe est d'inclure les effets de la stratification dans la description du transport turbulent. Pour des conditions stable et instable, cette théorie reprend le profil logarithmique, bien connu en condition neutre (situation de neutralité thermique : isothermie), en y rajoutant un coefficient de correction fonction de la stabilité de l'atmosphère.

En effet, l'amplitude des mouvements turbulents et la forme des profils dépendent des conditions de stratification thermique. En condition instable, la vitesse du vent croît plus vite avec l'altitude que dans les conditions de stabilité thermique. Le profil de vent est alors représenté par une courbe dont la concavité est tournée vers le haut. Au contraire, dans les conditions de stratification thermique stable, la vitesse du vent est plus réduite au niveau du sol et le profil de vitesse est représenté par une courbe dont la concavité est orientée vers le bas (en échelle semi-logarithmique).

Il apparaît donc important dans le cadre d'une simulation fine de l'évolution dynamique d'une CLA, de pouvoir bien paramétriser les différents processus qui prennent naissance à l'interface entre le sol et l'air. Aussi, ce travail s'est focalisé principalement sur la modélisation d'une couche limite atmosphérique par une approche RANS et un modèle de fermeture de type  $k-\varepsilon$ . D'après l'expérience (*De Visscher, 2014 ; Wyngaard, 2010*), la prise en compte des profils de la CLA est totalement dépendante de l'état de stabilité atmosphérique. Dans le cas du modèle  $k-\varepsilon$ , cet état induit une incompatibilité entre les profils analytiques (notamment en condition très stable) et le système d'équations dans l'approche RANS. Pour lever cette indétermination, une attention particulière est apportée à la manière dont sont représentés ces effets de stabilité, influençant directement les différents paramètres qui décrivent la turbulence (profils turbulents découlant de la théorie de similitude).

Plusieurs profils de vitesse de vent à l'entrée du domaine sont ainsi évalués. Deux cas d'étude sont plus spécifiquement utilisés dans ce rapport : le profil de vitesse est évalué par deux modèles différents de longueur de mélange : (*Gryning et al, 2007*), et (*Beljaars and Holtslag 1991*). Nous rappelons que la notion de longueur de mélange suppose que le transport de quantité de mouvement se fait sur une distance de l'ordre de la distance à la paroi (au sol). Un des problèmes essentiels des modèles de longueur de mélange est le fait

que l'on est loin de pouvoir faire des hypothèses aussi simples pour les écoulements dans le CLA (influence de la rugosité, de la stabilité, etc.), surtout quand interviennent des échelles de longueur caractéristique multiple. Cela souligne ainsi la grande difficulté de déterminer une loi de comportement valable dans l'ensemble de la CLA.

Dans ce qui suit, nous exposons succinctement les profils théoriques développés par nos deux auteurs où différentes paramétrisations influençant l'épaisseur de la CLA sont proposées, où des modèles à longueur de mélange ont été étendu à l'ensemble de la CLA plus spécifiquement dans les cas de conditions stables, voire très stables. Pour ce faire, ces paramétrisations sont proposées pour résoudre la distribution verticale du vent, qui combine ainsi la théorie de Monin-Obukhov avec le modèle de turbulence  $k-\varepsilon$ .

## II. Profils de vitesse dans la couche limite atmosphérique

### 1. Profil de vent développé pour l'ensemble de la CLA en condition stable ou modérément stable (*Gryning et al, 2007*)

Une première formulation de longueur de mélange a été proposée par (*Gryning et al, 2007*). Celle-ci tient compte de la hauteur de la couche limite ; ils ont notamment comparé avec succès les solutions analytiques (développées pour le profil de vent sur terrain plat et homogène) à des mesures de vitesse de vent jusqu'à 160m au Høvsøre (Danemark), et jusqu'à 250m à Hambourg (Allemagne).

Pour ce faire, Ces auteurs ont subdivisé la CLA en plusieurs sous-couches (au nombre de trois), où l'échelle de longueur totale,  $l$ , représente la sommation de trois termes :

$$\frac{1}{l} = \frac{1}{L_{SL}} + \frac{1}{L_{MBL}} + \frac{1}{L_{UBL}} = \frac{1}{z} + \frac{1}{L_{MBL}} + \frac{1}{(h-z)}$$

avec  $L_{SL}$ , l'échelle de longueur dans la couche de surface,  $L_{MBL}$ , l'échelle de longueur située au milieu de la couche limite et  $L_{UBL}$ , l'échelle de longueur dans la partie supérieure de la couche limite. En condition stable, (*Gryning et al 2007*) montre que la variation de la vitesse du vent avec l'altitude s'exprime de cette manière :

$$U(z) = \frac{u_*}{\kappa} \left[ \ln \left( \frac{z}{z_0} \right) + \frac{bz}{L} \left( 1 - \frac{z}{2h} \right) + \frac{z}{L_{MBL}} - \frac{z}{h} \left( \frac{z}{2L_{MBL}} \right) \right]$$

où  $h$  représente la hauteur de la couche limite atmosphérique,  $L$  la longueur de Monin-Obukhov,  $b = 4.7$ , la constante de Von Karman ( $\kappa$ ) est prise à 0.41 et enfin  $z_0$  représente la hauteur de rugosité (dans le cadre de cette étude  $z_0 = 0.045$  m). La paramétrisation de  $L_{MBL}$  pour le cas stable est donnée par la relation suivante :

$$L_{MBL} = \frac{h}{2} \left[ \left\{ \left( \ln \left( \frac{u_*}{f_c z_0} \right) - B(\mu) \right)^2 + A^2(\mu) \right\}^{1/2} - \ln \left( \frac{h}{z_0} \right) - \frac{bh}{2L} \right]^{-1}$$

où  $f_c$  est le paramètre de Coriolis ( $f_c = 1.10^{-4} \text{ s}^{-1}$ ) et  $u_*$ ,  $A(\mu)$  et  $B(\mu)$  sont des fonctions similitude établies par (Arya, 1975) pour la CLA :

$$A(\mu) = 1.01 - 0.105 \times (\kappa \mu) - 0.00099 \times (\kappa \mu)^2 + 0.00000081 \times (\kappa \mu)^3$$

et

$$B(\mu) = 5.14 + 0.142 \times (\kappa \mu) + 0.00117 \times (\kappa \mu)^2 - 0.0000033 \times (\kappa \mu)^3$$

avec

$$\mu = \frac{u_*}{f_c L}$$

Les valeurs de la vitesse de frottement  $u_*$  et de la longueur de Monin-Obukhov  $L$  sont celles des quatre essais de l'expérience MUST renseignées dans le tableau 4.4 du chapitre 4.

## 2. Profil de vent développé pour l'ensemble de la CLA en condition très stable (Beljaars and Holtslag, 1991)

Dans les cas précédents, le profil a été développé pour des situations modérément ou faiblement stables. Sous des conditions très stables, la longueur de Monin-Obukhov ( $L$ ) tend vers zéro et les limites imposées par les fonctions universelles sont atteintes. (Beljaars and Holtslag, 1991) modélisent un profil particulièrement adapté pour une couche atmosphérique très stable. Ce dernier a été reformulé pour tenir compte explicitement de l'importance des processus laminaires dans la limite d'une atmosphère très stable. Le profil de vitesse suivant sera donc utilisé uniquement pour l'essai 6 (expérience 1) de l'expérience MUST :

$$U(z) = \frac{u_*}{\kappa} \left[ \ln \left( \frac{z}{z_0} \right) - \Psi_m \left( \frac{z}{L} \right) \right]$$

où

$$\Psi_m = -a \left( \frac{z - z_0}{L} \right) - b \left( \frac{z}{L} - \frac{c}{d} \right) e^{-d \frac{z}{L}} + b \left( \frac{z_0}{L} - \frac{c}{d} \right) e^{-d \frac{z_0}{L}}$$

avec  $a=1$ ,  $b=0.667$ ,  $c=5$  et  $d=0.35$ .  $L$  est la longueur de Monin-Obukhov,  $\kappa = 0.41$ , la constante de Von Karman et  $z_0$  représente la hauteur de rugosité ( $z_0 = 0.045$  m).



« BIBLIOGRAPHIE ANNEXE III »

Arya, S. P. S. (1975). Geostrophic drag and heat transfer relations for the atmospheric boundary layer. *Quarterly Journal of the Royal Meteorological Society*, 201, pp. 147-161.

Beljaars, A., Holtslag, A., (1991). Flux parameterization over land surfaces for atmospheric models, *J. Appl. Meteorol.*, 30 (3), pp. 327-341, doi:10.1175/1520-0450(1991)030<0327:FPOLSF>2.0.CO;2.

Dyer A. (1965). The flux-gradient relation for turbulent heat transfer in the lower atmosphere. *Quart. J. Roy. Meteorol. Soc.*, 91, pp. 151-157.

De Visscher A. (2014). *Air dispersion modeling*. Ed. Wiley.

Gryning, S.-E., Batchvarova, E., Brümmner, E., Jørgensen, H., Larsen S., (2007). On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer, *Boundary Layer Meteorol.*, 124 (2), pp. 251-268, doi:10.1007/s10546-007-9166-9.

Han J., Arya S.P., Shen S., and Lin Y-L. (2000). An Estimation of Turbulent Kinetic Energy and Energy Dissipation Rate Based on Atmospheric Boundary Layer Similarity Theory, NASA/CR-2000- 210298.

Kumar, P., Feiz, A.-A., Ngae, P., Singh, S. K., Issartel, J.P., (2015). CFD simulation of short range plume dispersion from a point release in an urban like environment, *Atmospheric Environment*, 122, pp. 645-656.

Launder, B.E.; Spalding, D.B. (1974). The numerical computation of turbulent flows. *Computer Methods in Applied Mechanics and Engineering*. 3 (2): 269-289. doi:10.1016/0045-7825(74)90029-2

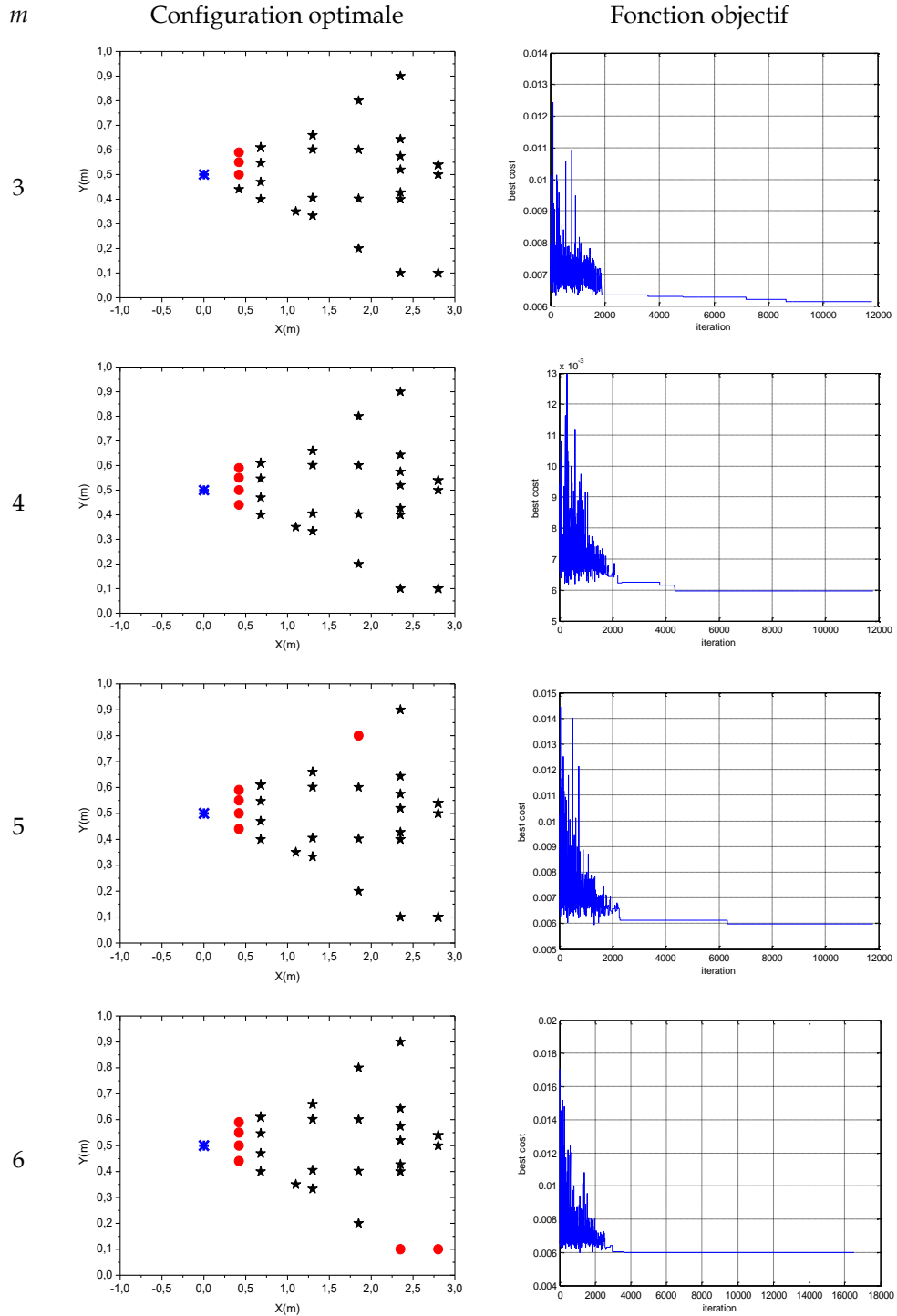
Peters H.; Baumert H.Z. (2007). Validating a turbulence closure against estuarine microstructure measurements. *Ocean. Model.* 19 (34), pp. 183-203.

Wyngaard J.C. (2010). *Turbulence in the atmosphere*. Ed. Cambridge University Press.

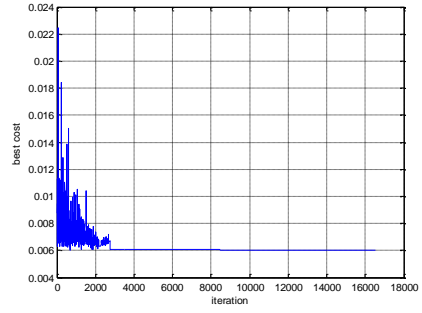
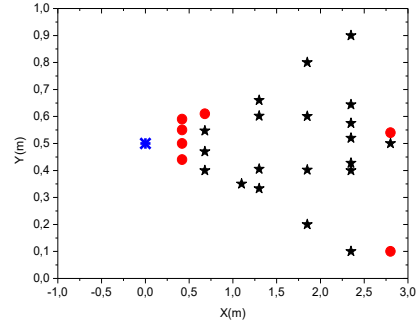
# Annexe IV :

## Approche analytique « problème 1 »

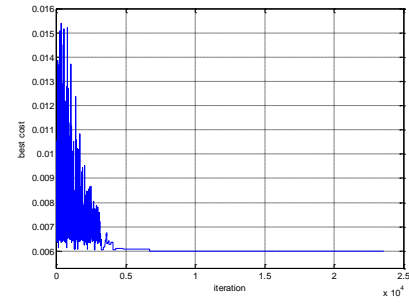
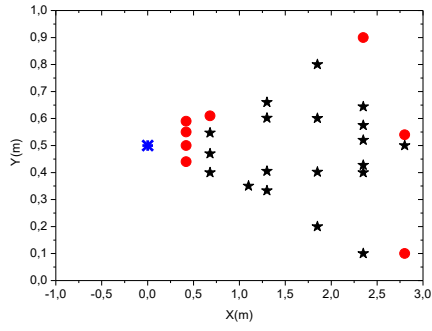
### Configurations optimales « Recuit Simulé »



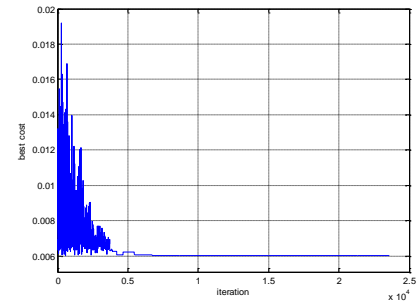
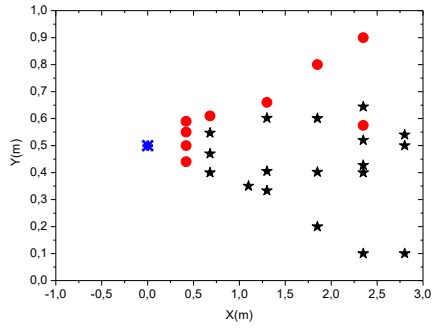
7



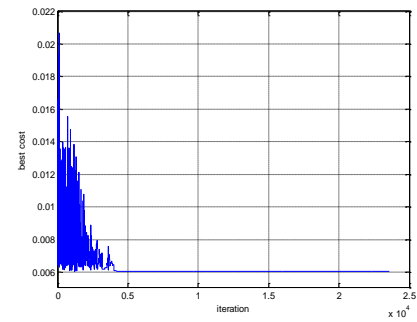
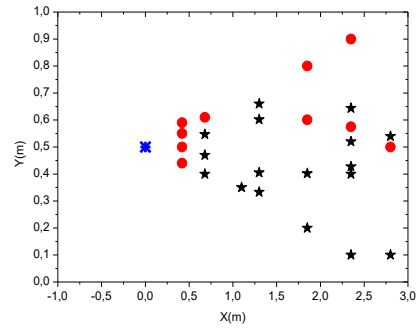
8



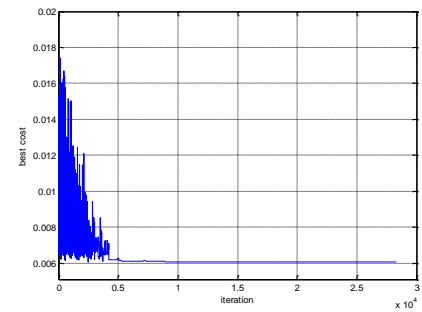
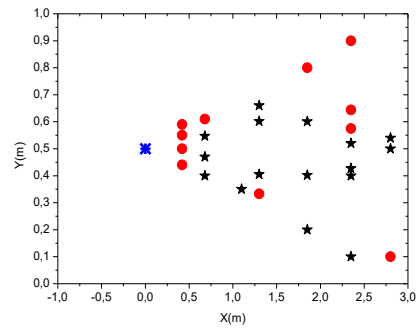
9



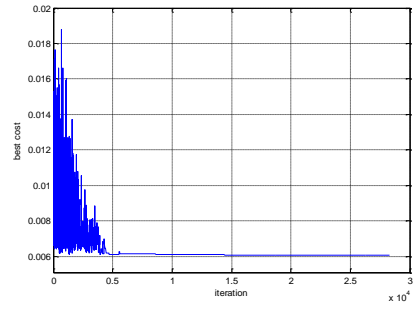
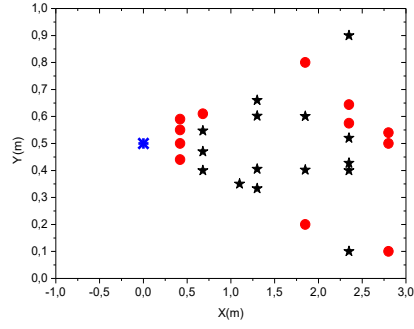
10



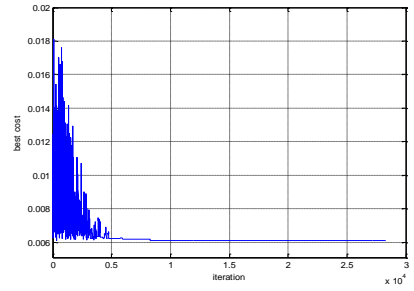
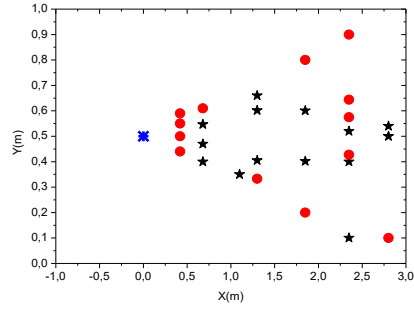
11



12



13



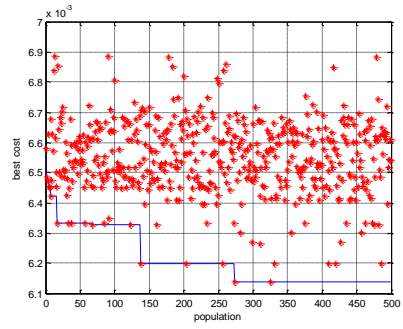
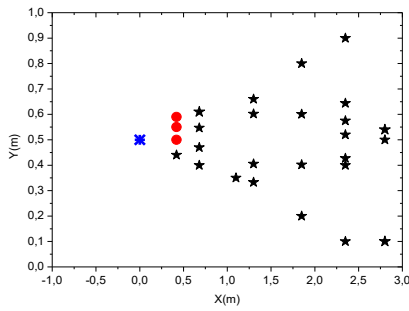
Configurations optimales « Algorithme génétique »

*m*

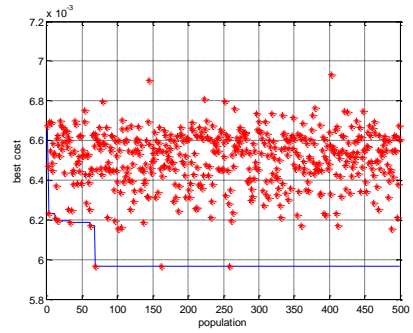
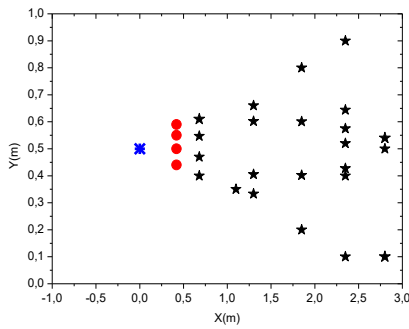
Configuration optimale

Fonction objectif

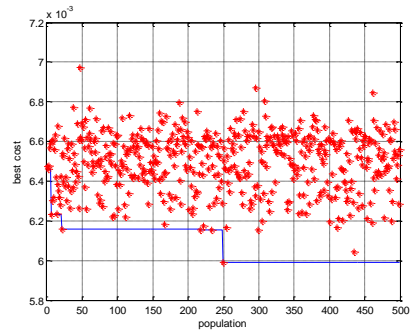
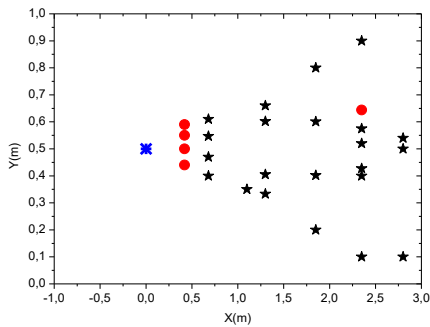
3



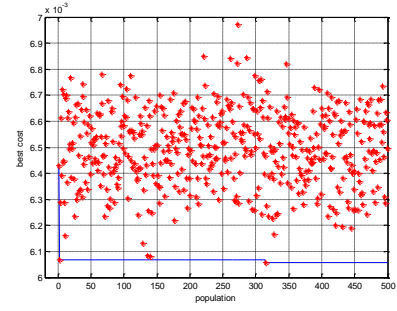
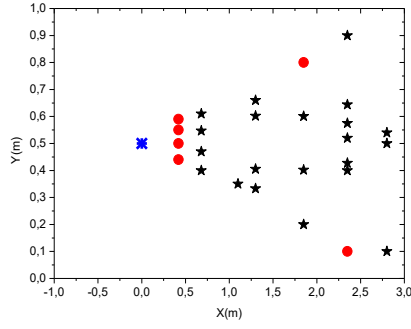
4



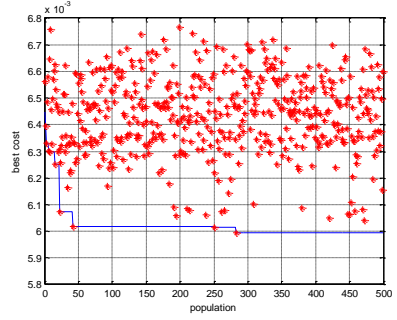
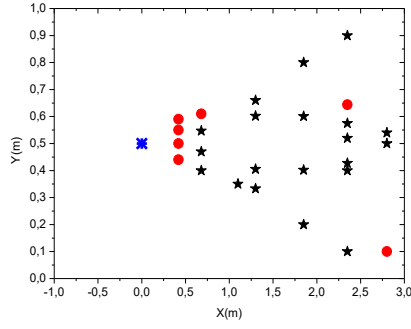
5



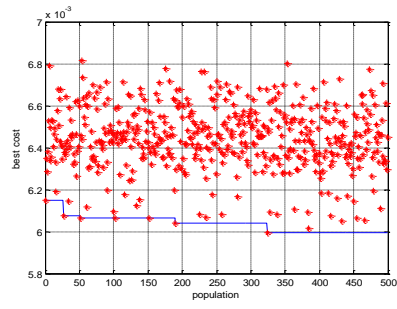
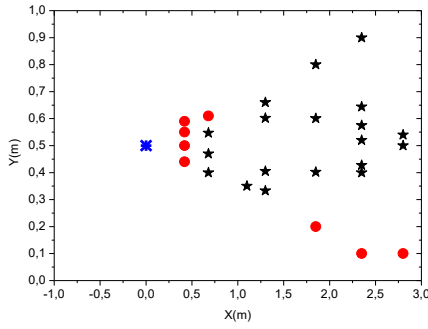
6



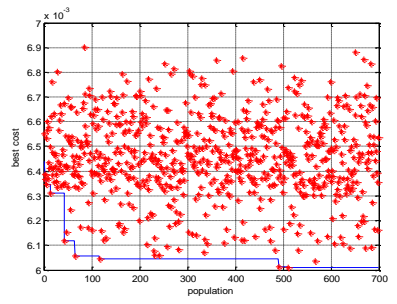
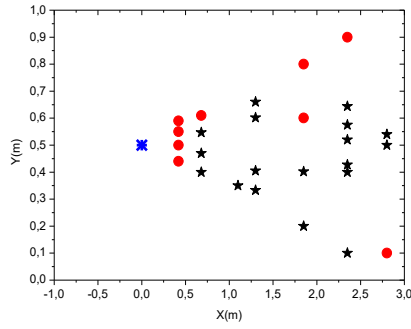
7



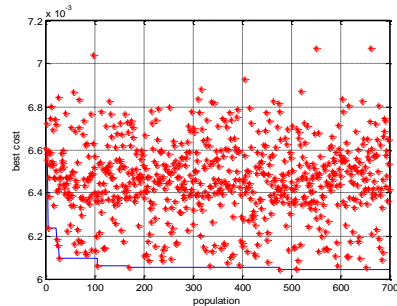
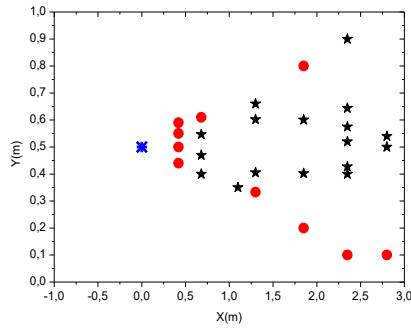
8



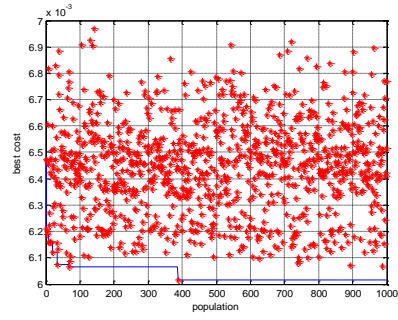
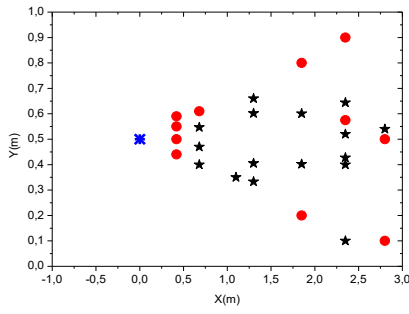
9



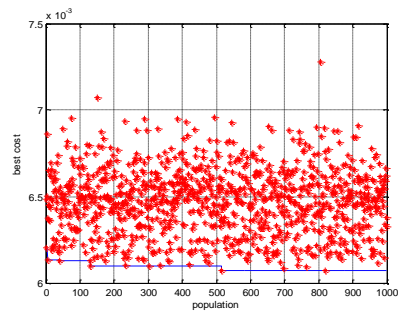
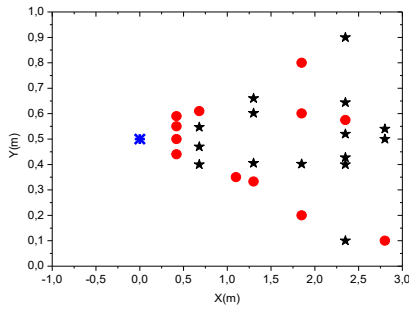
10



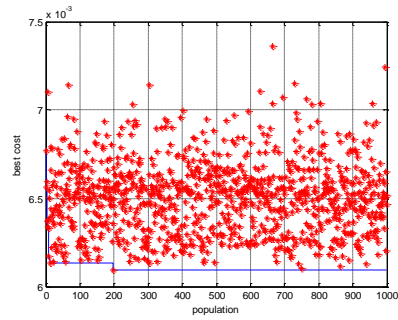
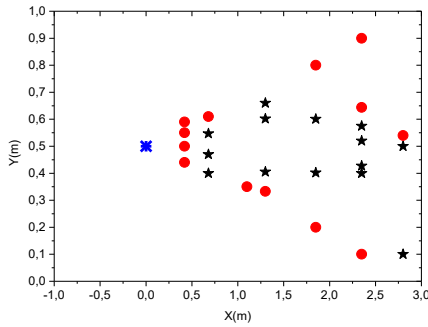
11



12



13



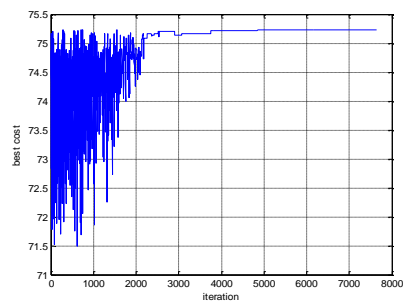
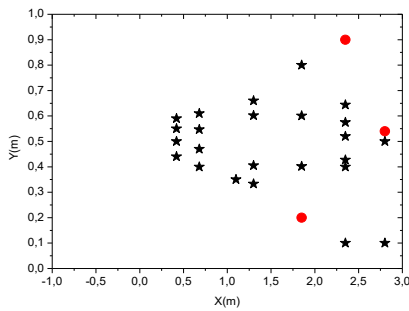
### Approche analytique « problème 2 »

*m*

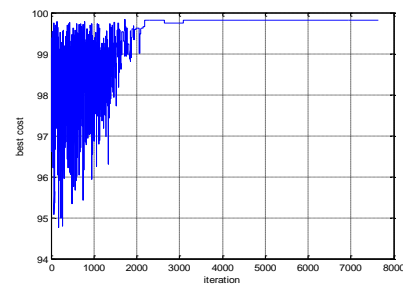
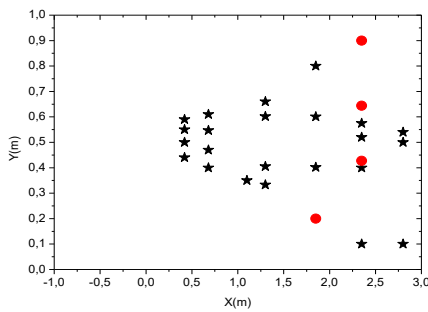
Configuration optimale

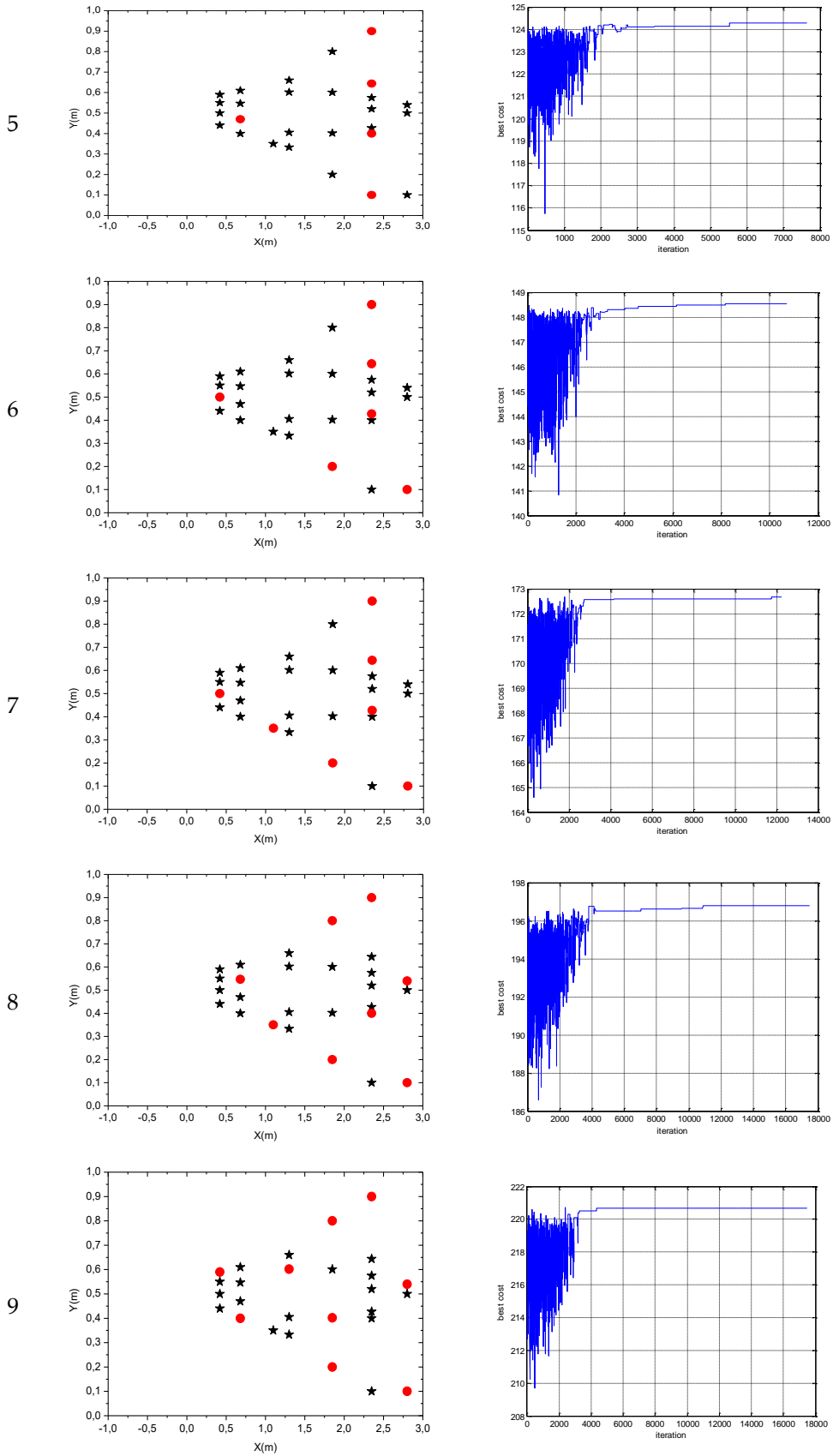
Fonction objectif

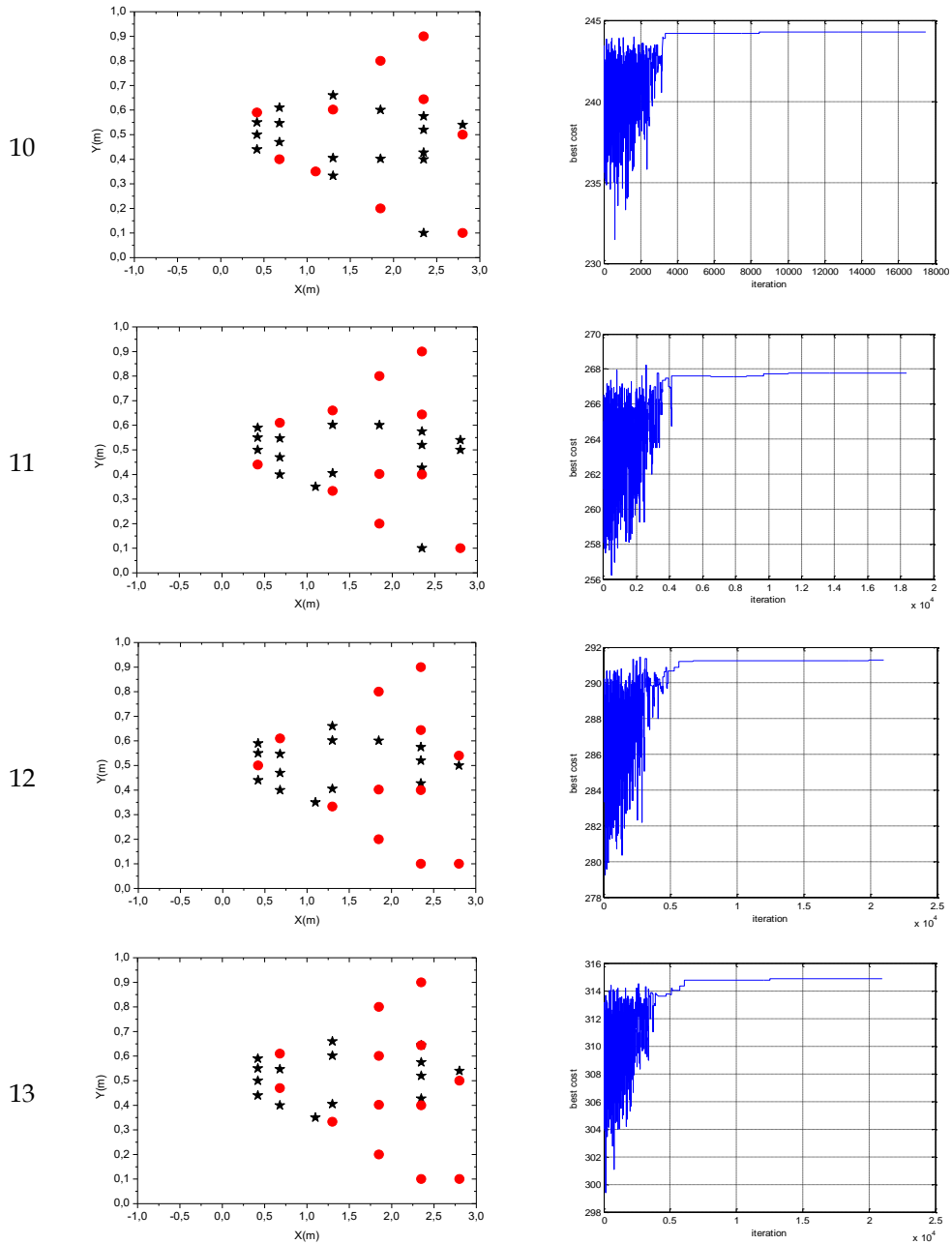
3



4

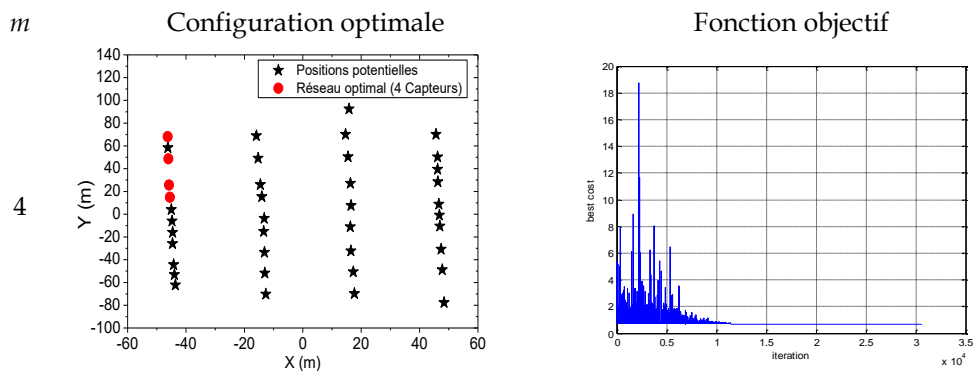




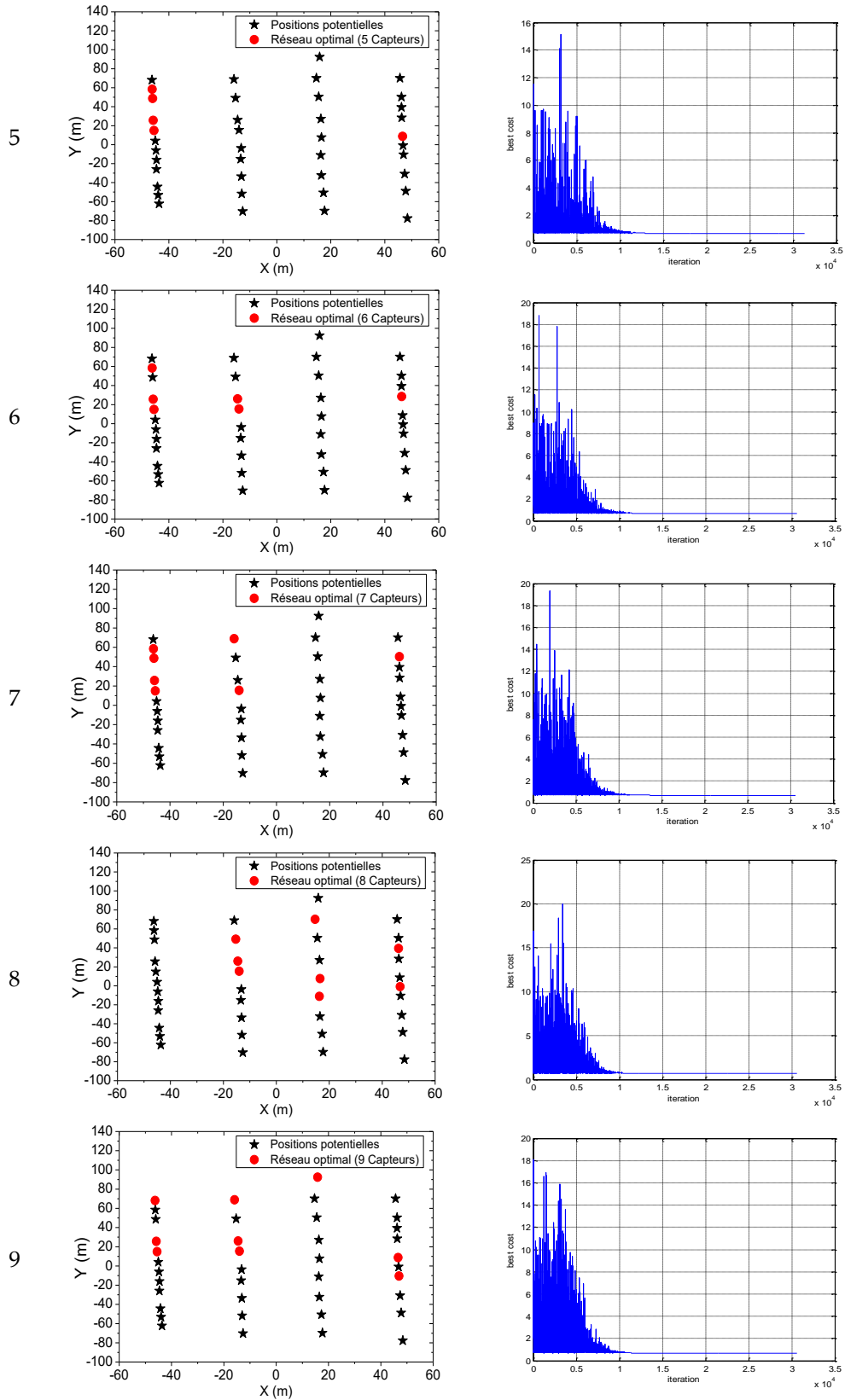


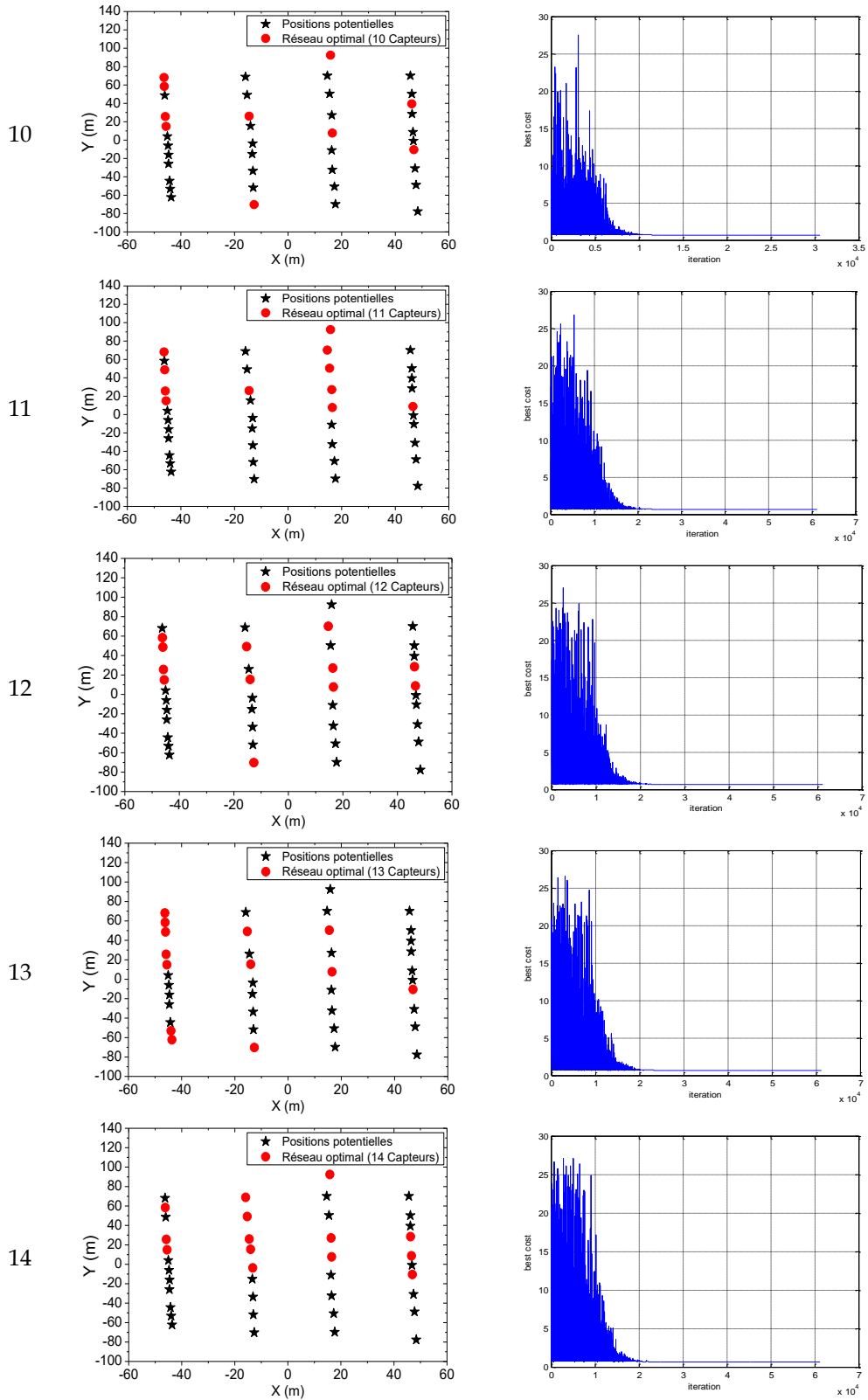
### Approche CFD « problème 1 »

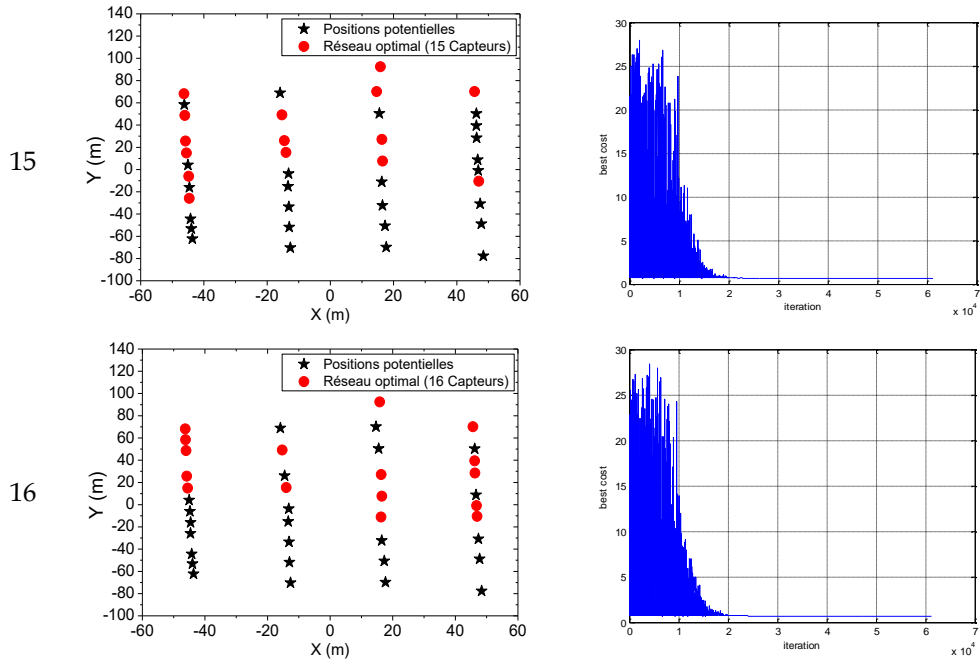
#### Expérience 1 « Essai 6 »











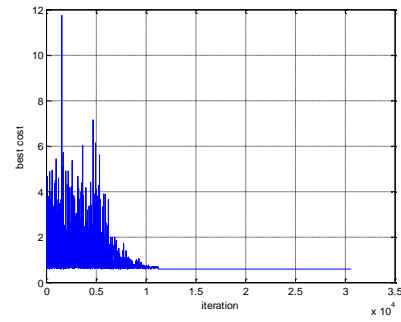
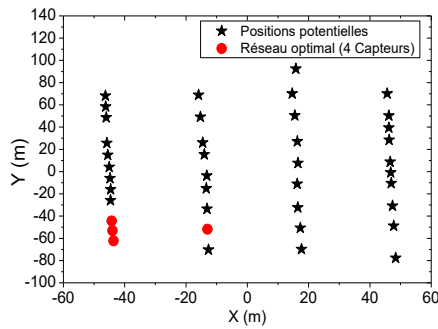
Expérience 2 « Essai 8 »

*m*

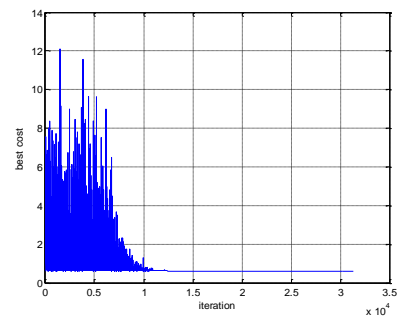
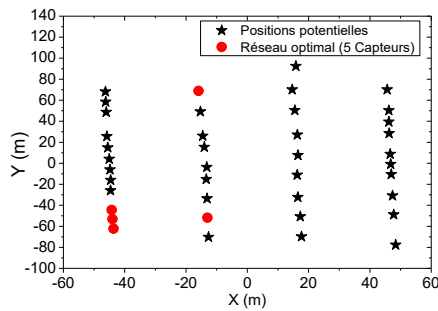
Configuration optimale

Fonction objectif

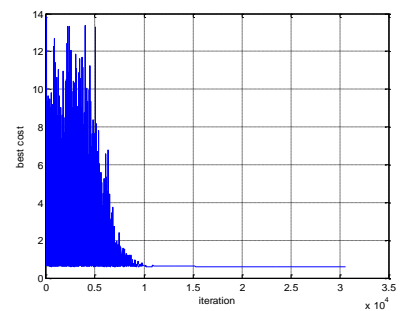
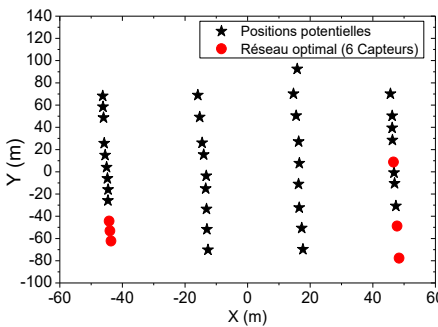
4

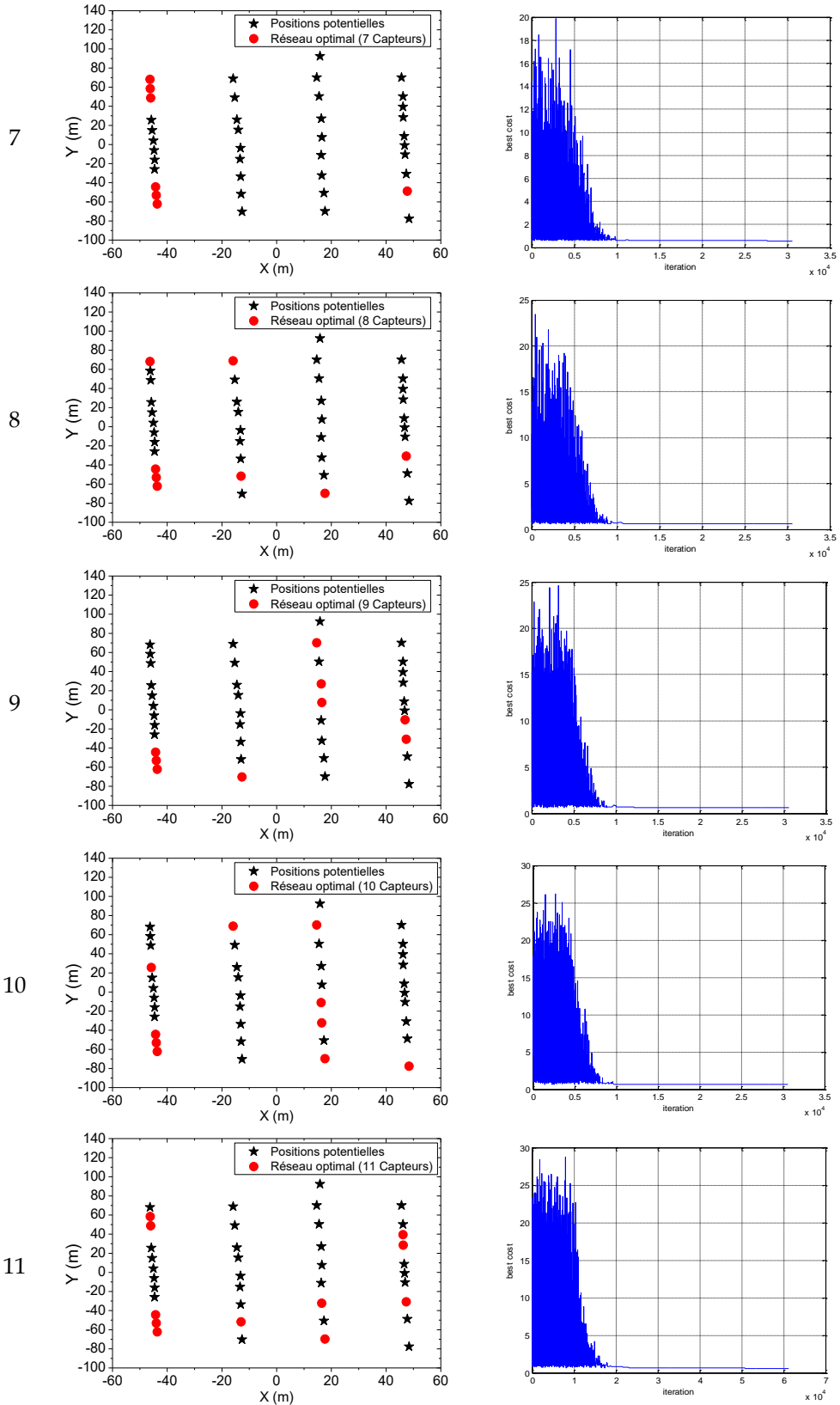


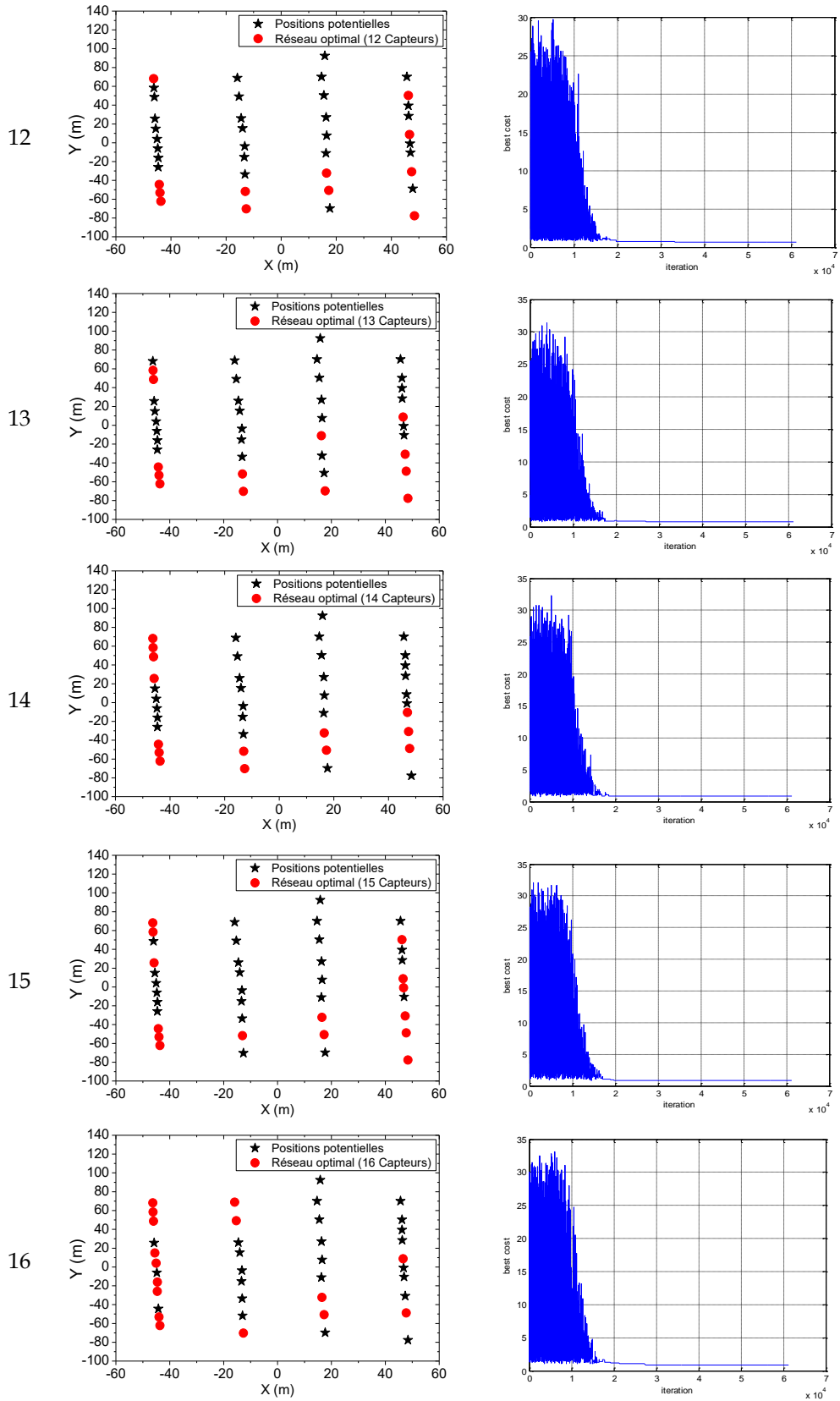
5



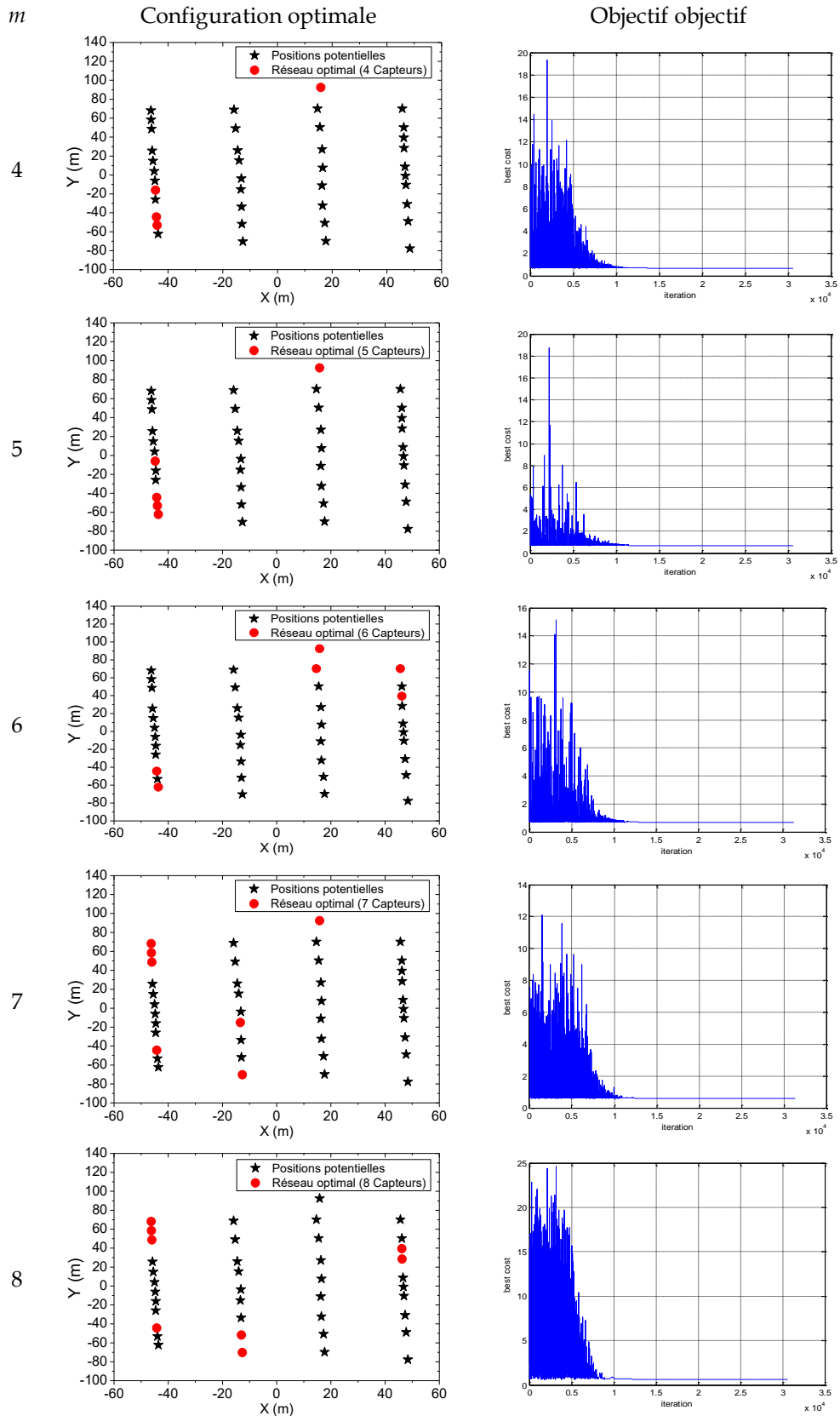
6

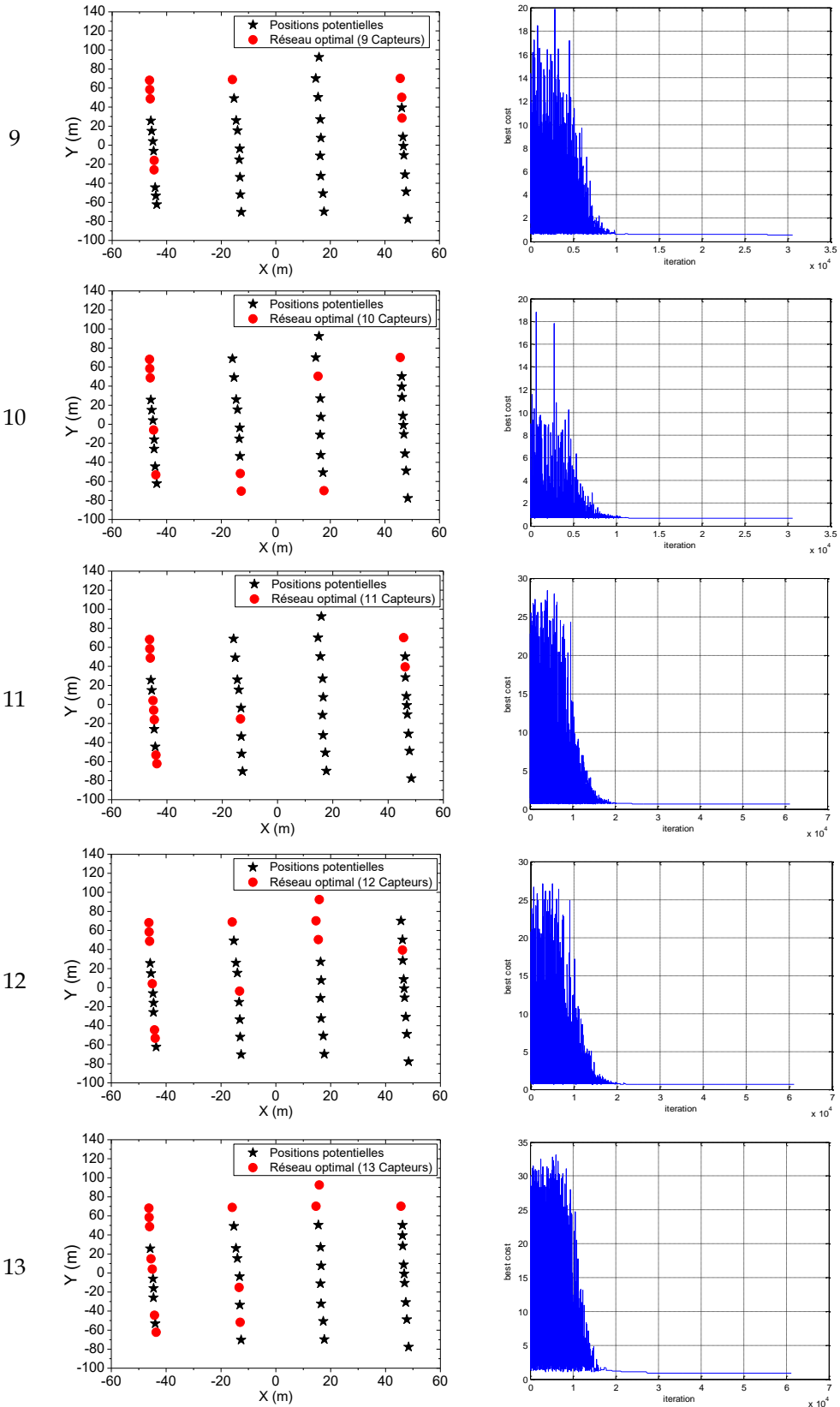


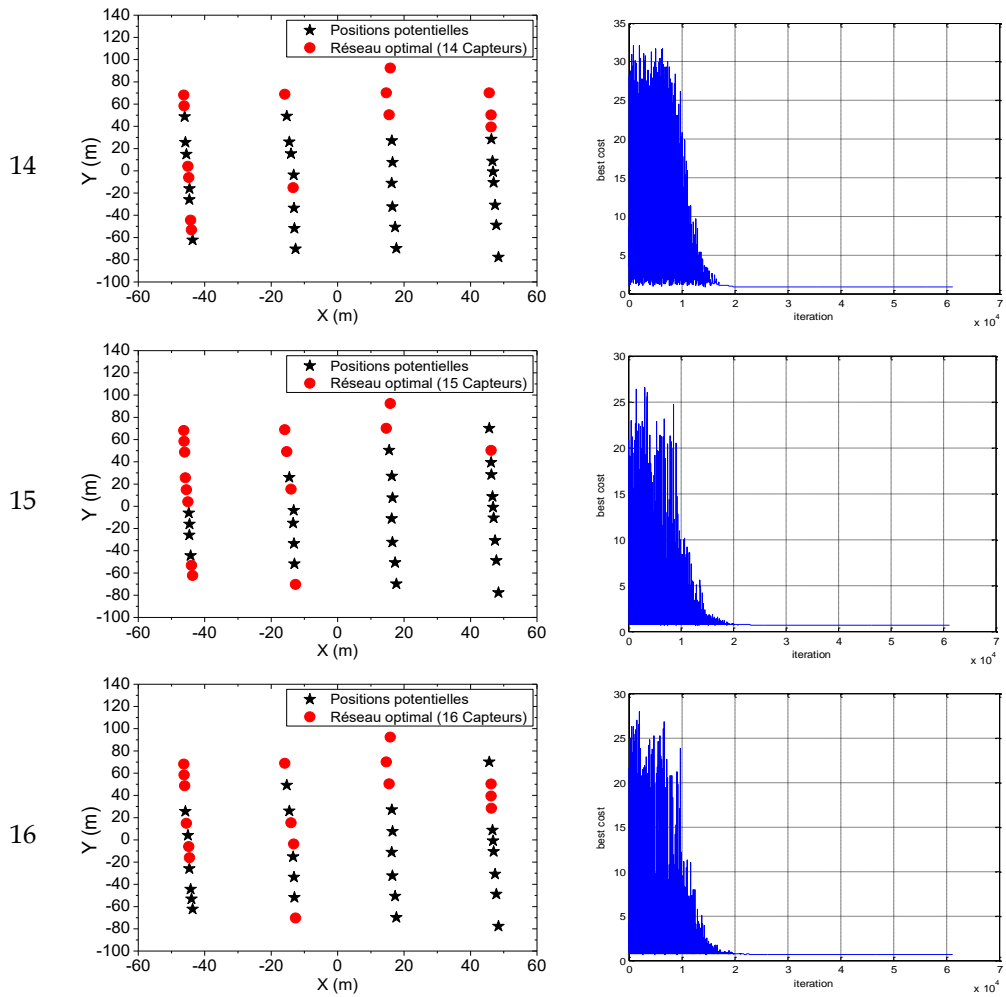




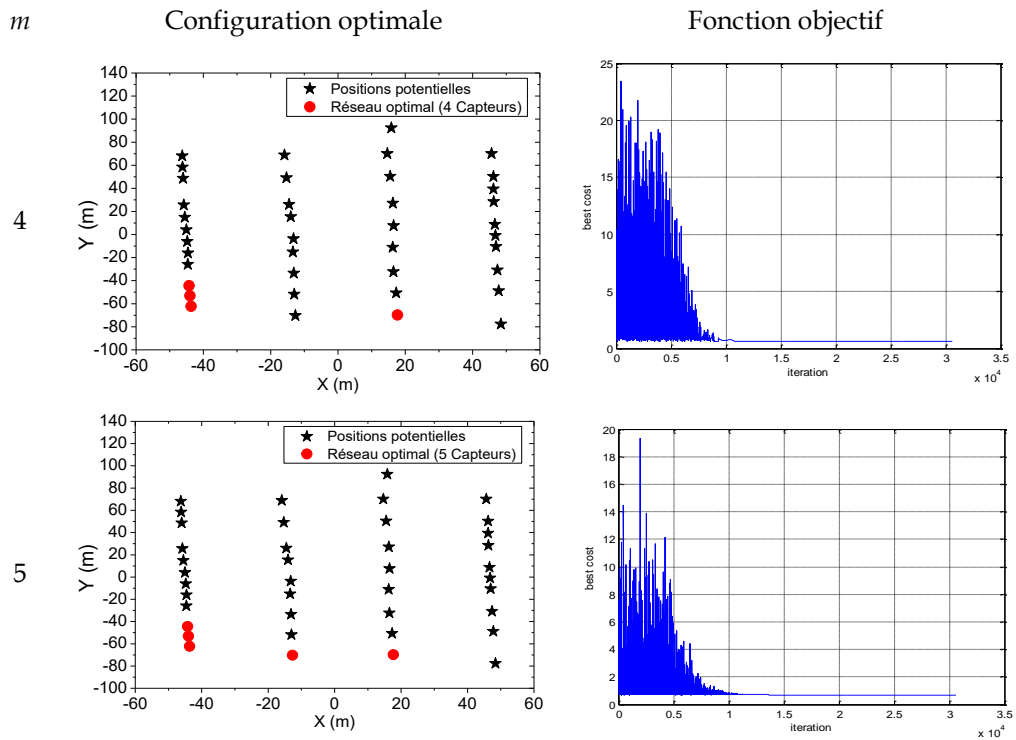
Expérience 3 « Essai 14 »



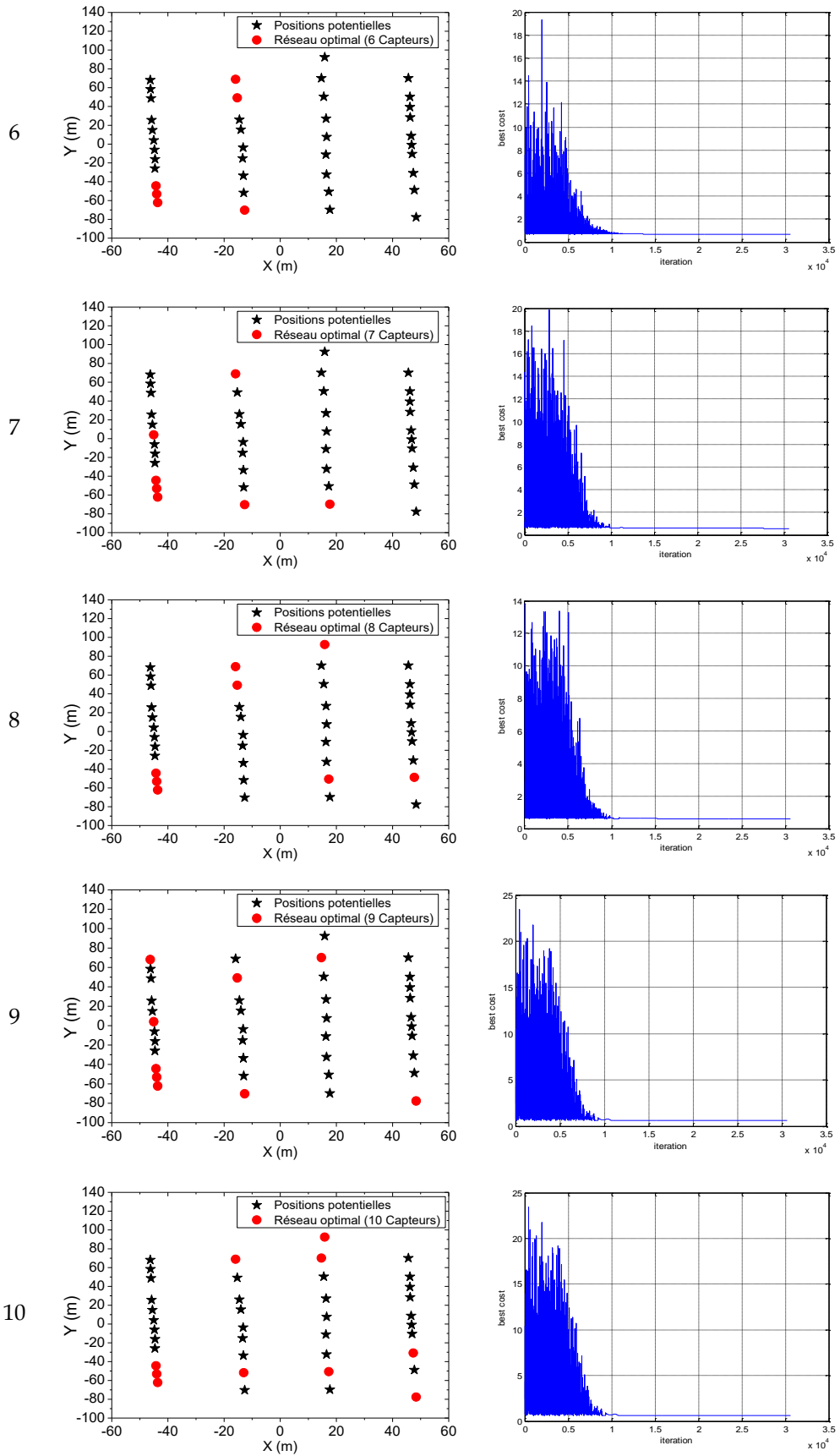


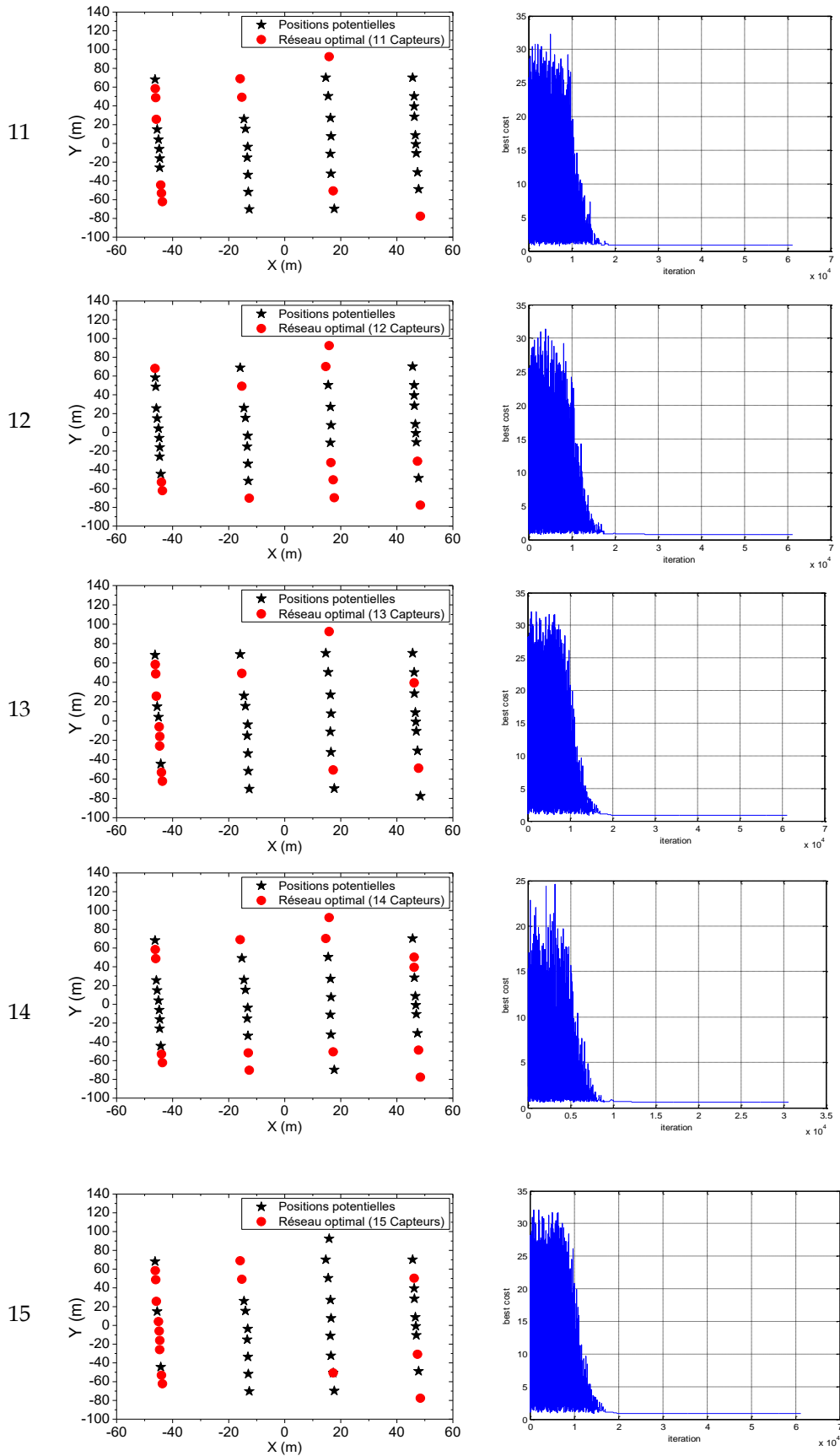


Expérience 4 « Essai 19 »

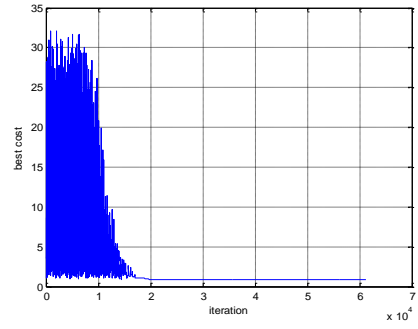
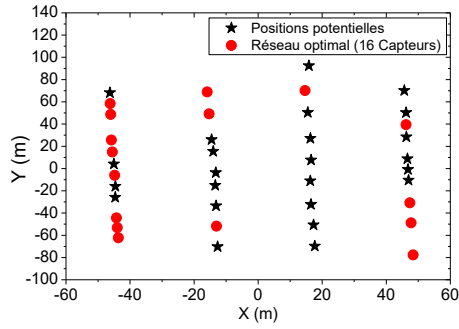








16



## Approche CFD « Problème 2 »

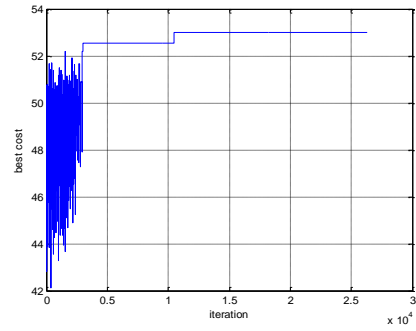
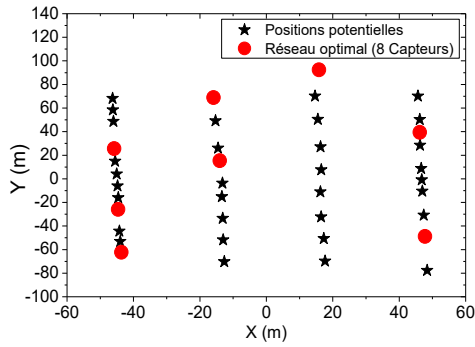
### Expérience 1 « Essai 6 »

*m*

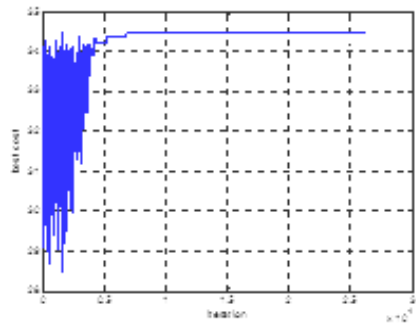
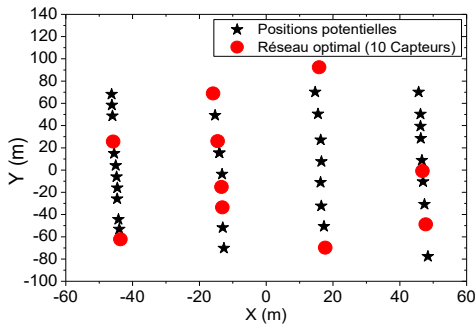
Configuration optimale

Fonction objectif

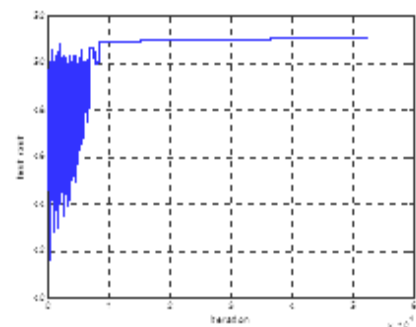
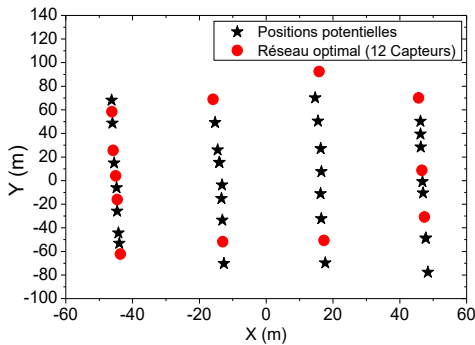
8



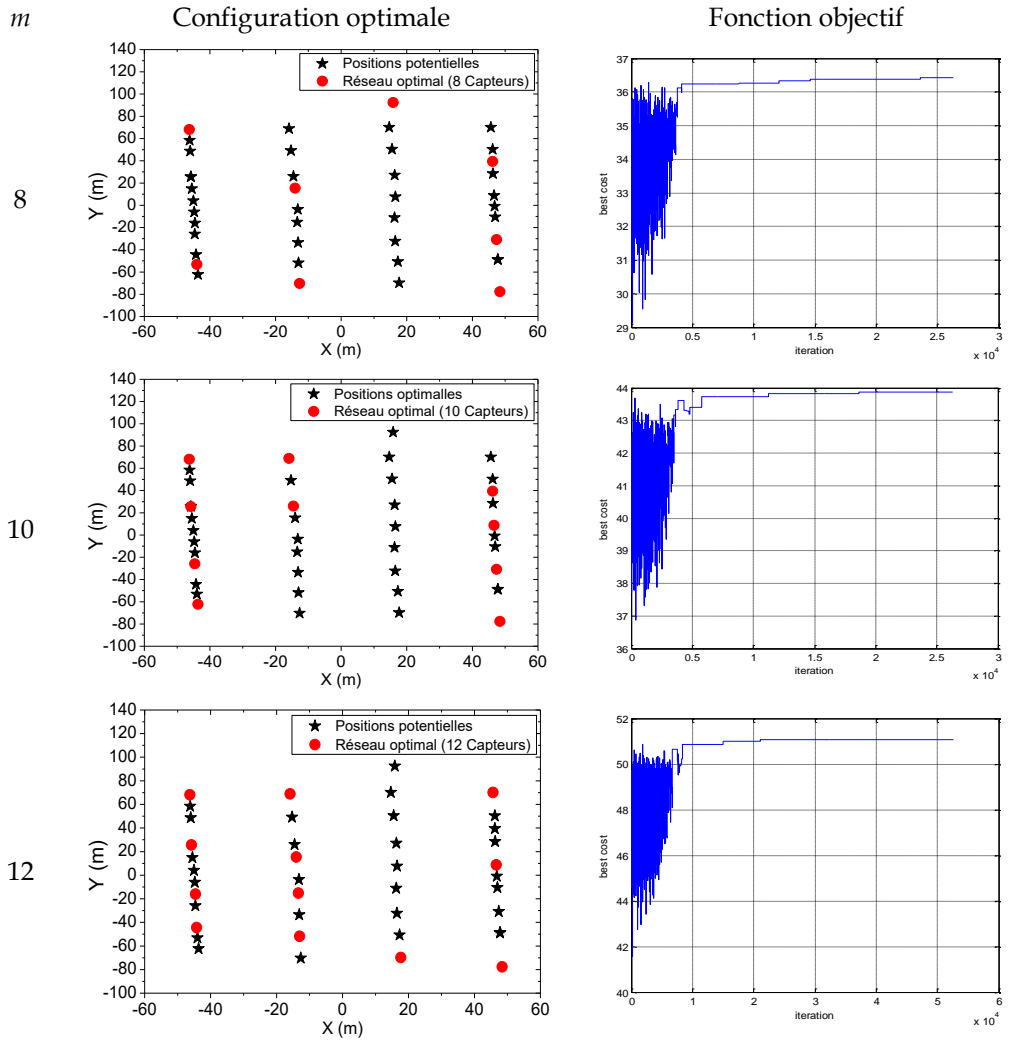
10



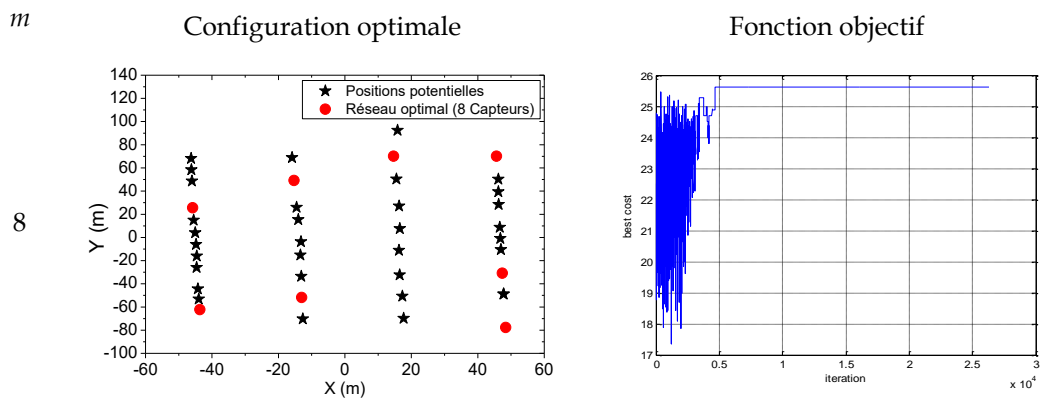
12

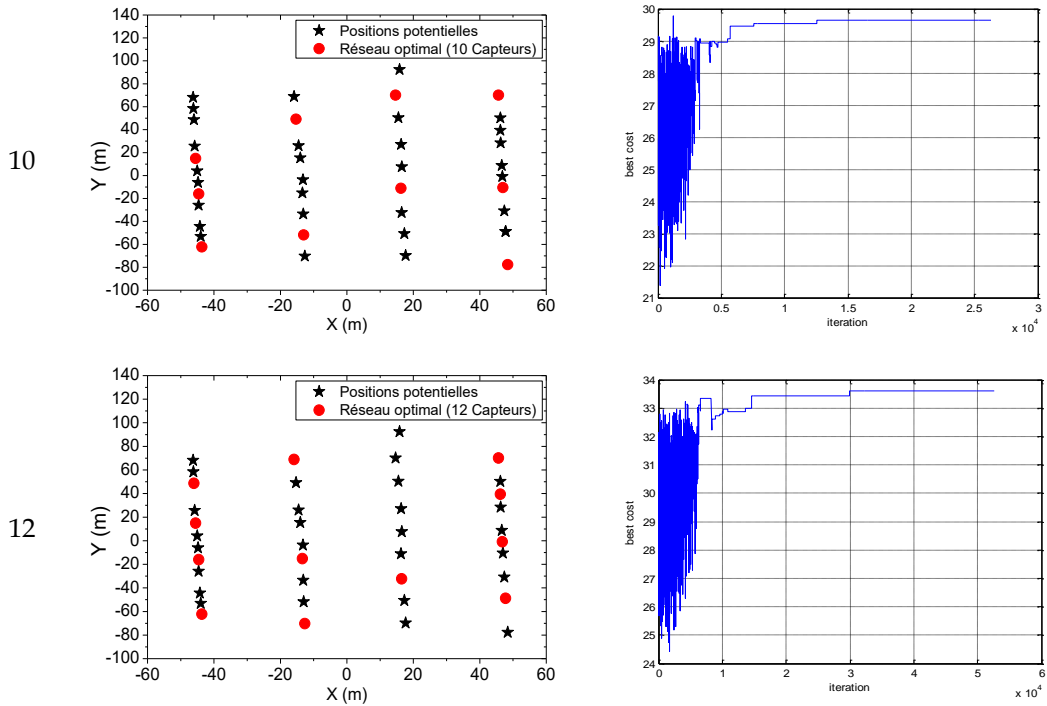


Expérience 2 « Essai 8 »



Expérience 3 « Essai 14 »



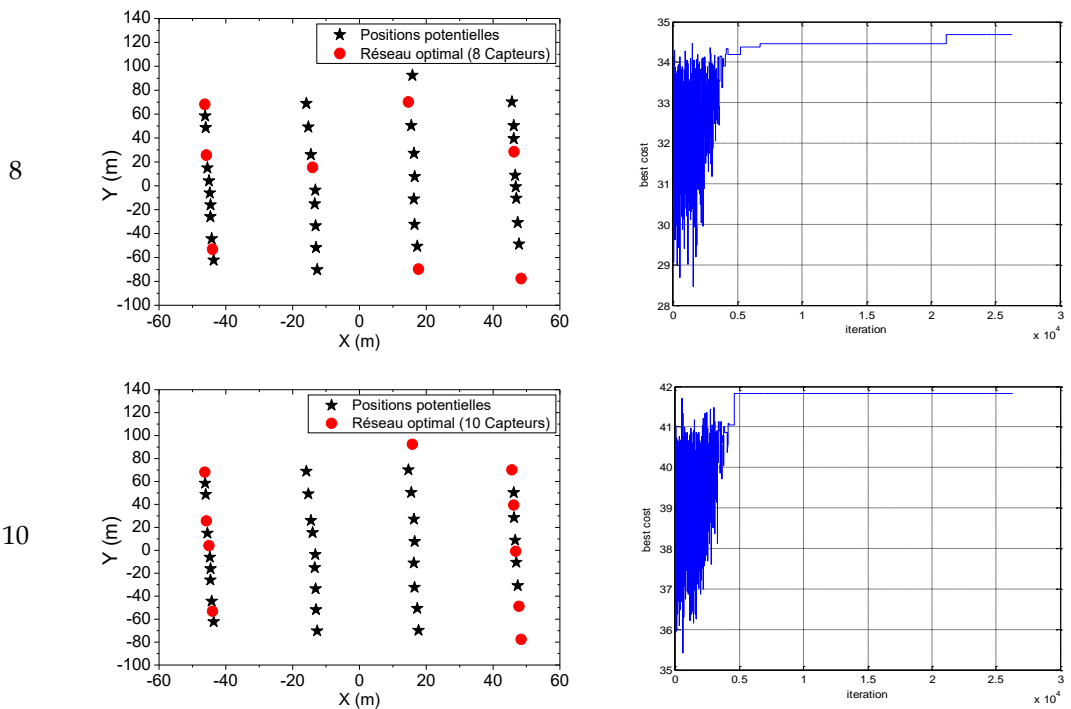


Expérience 4 « Essai 19 »

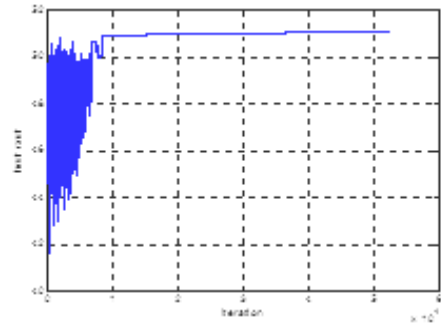
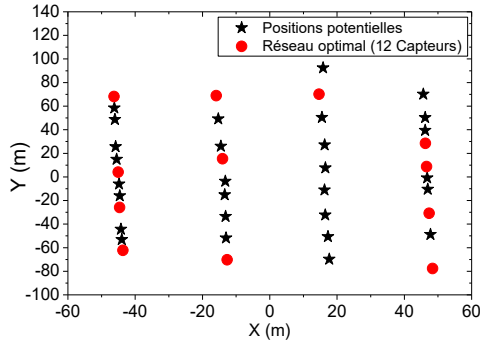
$m$

Configuration optimale

Fonction objectif



12



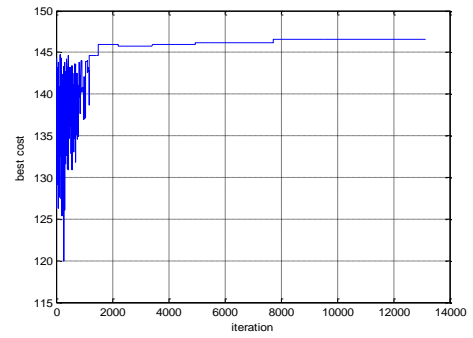
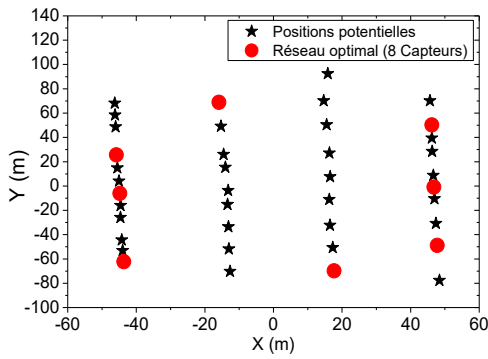
### Approche CFD « Problème 3 »

*m*

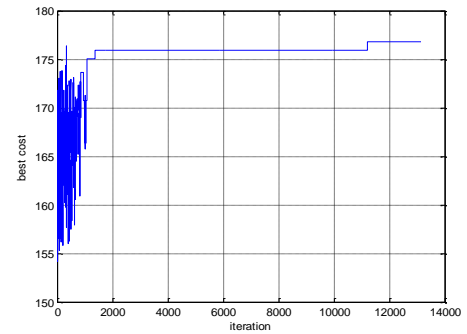
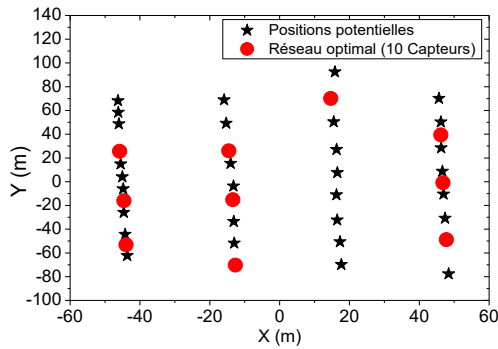
configuration optimale

objectif

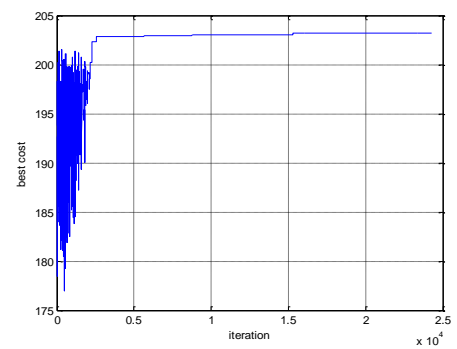
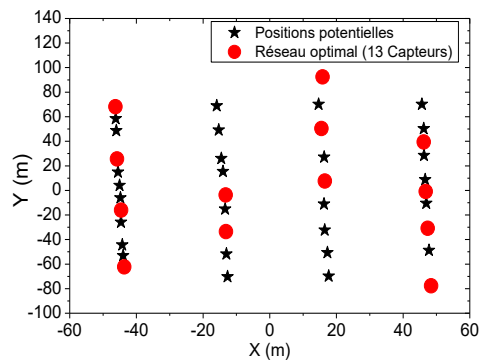
8

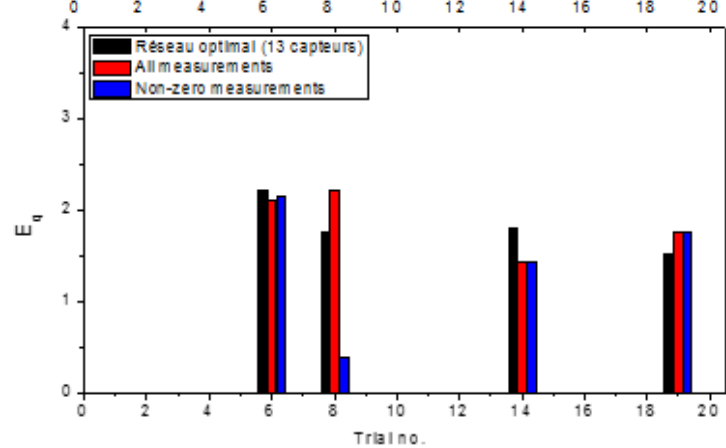
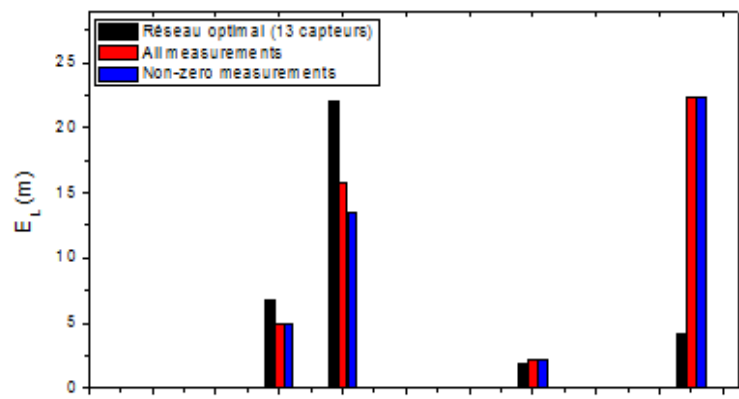
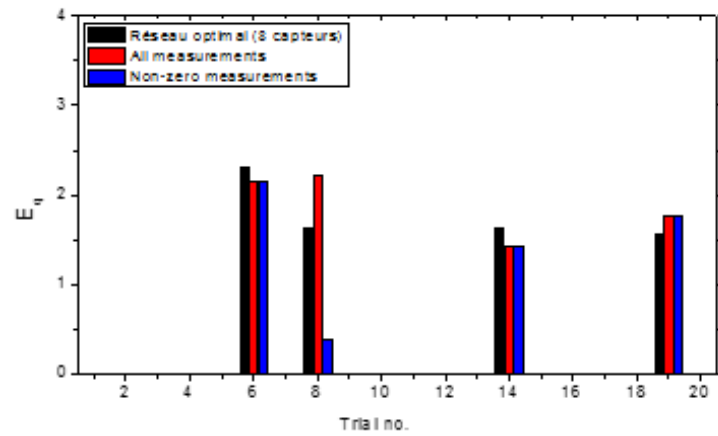
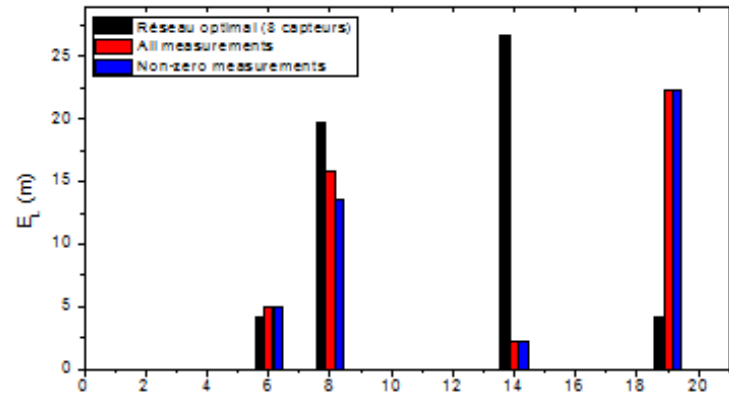


10



13



Resau optimal  $m=8$  et  $m=13$  VS reseaux (kumar et al, 2015)

**Titre :** Optimisation de réseaux de capteurs pour la caractérisation de source de rejets atmosphériques

**Mots clés :** *Optimisation, renormalisation, champs adjoints, modèle gaussien, CFD, caractérisation de source, réseau de capteurs, fonction objectif, erreurs normalisées, entropie de réseau, information, entropie généralisée.*

**Résumé :** L'objectif principal de cette étude est de définir les méthodes nécessaires pour optimiser un réseau de surveillance conçu pour la caractérisation de source de rejets atmosphériques. L'optimisation consiste ici à déterminer le nombre et les positions optimales de capteurs à déployer afin de répondre à ce type de besoin. Dans ce contexte, l'optimisation est réalisée pour la première fois par un couplage entre la technique d'inversion de données dite de « renormalisation » et des algorithmes d'optimisation métaheuristiques. La méthode d'inversion a été en premier lieu évaluée pour la caractérisation de source ponctuelle, et a permis ensuite, de définir des critères d'optimalité pour la conception des réseaux. Dans cette étude, le processus d'optimisation a été évalué dans le cadre d'expériences réalisées en terrain plat sans obstacles (DYCE) et en milieu urbain idéalisé (MUST). Trois problématiques ont été définies et testées sur ces expériences. Elles concernent (i) la détermination de la taille optimale d'un réseau permettant de caractériser une source de pollution, où une fonction coût (erreurs normalisées), traduisant l'écart entre les observations et les données modélisées, a été minimisée; (ii) la conception optimale d'un réseau permettant de caractériser une source ponctuelle inconnue, pour une condition météorologique particulière.

Dans ce contexte, une fonction coût entropique a été maximisée afin d'augmenter la quantité d'information fournie par le réseau; (iii) la détermination d'un réseau optimal permettant de caractériser une source ponctuelle inconnue pour des configurations météorologiques multiples. Pour ce faire, une fonction coût entropique généralisée, que nous avons définie, a été maximisée. Pour ces trois problématiques, l'optimisation est assurée dans le cadre d'une approche d'optimisation combinatoire. La détermination de la taille optimale d'un réseau (problématique 1) s'est révélée particulièrement sensible aux différentes conditions expérimentales (hauteur et débit de la source, conditions de stabilité, vitesse et direction du vent, etc.). Nous avons noté pour ces expériences, que les performances des réseaux sont meilleures dans le cadre d'une dispersion sur terrain plat comparativement aux milieux urbains. Nous avons également montré que différentes architectures de réseaux pouvaient converger vers le même optimum (approché ou global). Pour la caractérisation de sources inconnues (problématiques 2 et 3), les fonctions coûts entropiques se sont avérées robustes et ont permis d'obtenir des réseaux optimaux performants (de tailles raisonnables) capables de caractériser différentes sources pour une ou plusieurs conditions météorologiques.

**Title :** Sensors networks optimization for the characterization of atmospheric releases source

**Keywords :** *Optimization, renormalization, adjoint fields, Gaussian model, CFD, source characterization, sensor network, objective function, normalized errors, network entropy, information, generalized entropy.*

**Abstract :** The main objective of this study is to define the methods required to optimize a monitoring network designed for atmospheric source characterization. The optimization consists in determining the optimal number and locations of sensors to be deployed in order to respond to such needs. In this context, the optimization is performed for the first time by a coupling between the data inversion technique named "renormalization" and the metaheuristic optimization algorithms. At first, the inversion method was evaluated for a point source, and then have allowed to define optimality criteria for networks design. In this study, the optimization process was evaluated in experiments carried out in flat terrain without obstacles (DYCE) and in an idealized urban environment (MUST). Three problems were defined and tested based on these experiments. These problems concern (i) the determination of the optimal network size for source characterization, for which a cost function (standard errors) estimating the gap between observations and modeled data, has been minimized; (ii) the optimal design of a network to retrieve an unknown point source for a particular meteorological condition.

In this context, an entropy cost function has been maximized in order to increase the information's amount provided by the network; (iii) the determination of an optimal network to reconstruct an unknown point source for multiple meteorological configurations. For this purpose, a generalized entropic cost function that we have defined, has been maximized. For these all problems, optimization is ensured within the framework of a combinatorial optimization approach. The determination of the optimal network size (problem 1) was highly sensitive to experimental conditions (source height and intensity, stability conditions, wind speed and direction, etc.). We have noted that the networks performance is better for a dispersion on flat terrain compared to the urban environments. We have also shown that different networks architectures can converge towards the same optimum (approximate or global). For unknown sources reconstruction (problems 2 and 3), the entropic cost functions have proven to be robust and allowed to obtain optimal networks (for reasonable sizes) capable of characterizing different sources for one or multiple meteorological conditions.