



INSTITUT NATIONAL DE L'ENVIRONNEMENT INDUSTRIEL ET DES RISQUES

Méthodes de représentation de la qualité de l'air dans les zones peu/pas couvertes

Partie I : Stratégies et préconisations

***Laboratoire Central de Surveillance
de la Qualité de l'Air***

Convention 31/2001

*Laurence ROUÏL, Laure MALHERBE
Unité Modélisation et Analyse Economique
pour la gestion des Risques (MECO)*

Direction des Risques Chroniques (DRC)

Décembre 2002

Méthodes de représentation de la qualité de l'air dans les zones peu/pas couvertes

Partie I : Stratégies et préconisations

*Laboratoire Central de Surveillance
de la Qualité de l'Air*

Convention 31/2001

Décembre 2002

**LAURE MALHERBE
LAURENCE ROUÏL**

Ce document comporte 54 pages (hors couverture et annexes).

	Rédaction	Vérification	Approbation
NOM	Laurence ROUÏL Laure MALHERBE	Michel NOMINE	Martine RAMEL
Qualité	Ingénieurs Etudes et Recherches Direction des Risques Chroniques	Ingénieurs Etudes et Recherches Direction des Risques Chroniques	Coordination LCSQA Direction des Risques Chroniques
Visa			

TABLE DES MATIERES

1. RÉSUMÉ	3
2. INTRODUCTION	4
2.1 CONTEXTE ET OBJECTIFS.....	4
2.2 OBJET DE L'ETUDE	5
3. LES OUTILS DE CARTOGRAPHIE	7
3.1 L'INTERPOLATION	7
3.1.1 <i>Les méthodes d'interpolation déterministes</i>	7
3.1.2 <i>Les méthodes d'estimation géostatistiques</i>	9
3.1.3 <i>Données de concentration requises - Echelle spatiale et temporelle</i> Erreur! Signet non défini.	
3.1.4 <i>Amélioration de l'estimation au moyen de données auxiliaires</i>	21
3.1.5 <i>Incertitude d'une cartographie</i>	23
3.2 LES MODÈLES DÉTERMINISTES	24
3.2.1 <i>Description des méthodes</i>	24
3.2.2 <i>L'échelle spatiale</i>	25
3.2.3 <i>L'échelle temporelle</i>	25
3.2.4 <i>Précision des modèles déterministes et données d'entrée</i>	26
3.2.5 <i>Evaluation et validation</i>	27
3.3 LES MÉTHODES HYBRIDES.....	27
4. LES DONNÉES	30
4.1 DONNÉES DE CONCENTRATION	30
4.1.1 <i>Types de données</i>	30
4.1.2 <i>Collecte des données</i>	31
4.2 DONNÉES D'ÉMISSION.....	36
4.2.1 <i>Nature de l'information</i>	36
4.2.2 <i>Acquisition des données</i>	36
4.3 DONNÉES DE SITE ET MÉTÉOROLOGIQUES	39
5. CONCLUSION ET RECOMMANDATIONS	40
5.1 MÉTHODES D'INTERPOLATION DÉTERMINISTE	40
5.2 MÉTHODES DE GEOSTATISTIQUE	40
5.3 MÉTHODES DÉTERMINISTES	41
6. REFERENCES	43

1. RESUME

L'élaboration de cartes de répartition des concentrations des polluants réglementés sur des bases temporelles appropriées est un bon moyen de répondre aux objectifs de surveillance de la qualité de l'air définis dans les textes réglementaires nationaux et européens. Celles-ci sont relativement aisées à construire dans les zones de petite surface, où la densité de points de mesure est suffisante (sites urbains et péri-urbains), en appliquant par exemple des règles simples d'interpolation. En revanche, là où peu de stations de mesure sont implantées, cartographier la qualité de l'air nécessite l'usage de techniques de traitement complémentaires et plus sophistiquées, souvent regroupées sous le terme de « modélisation » ou d'analyse objective. Elles regroupent :

- Des méthodes d'interpolation,
- Des approches statistiques ou géostatistiques
- Des approches purement déterministes

Les deux premières catégories de méthodes reposent plus naturellement sur l'usage des données de concentrations mesurées. Elles correspondent à une vision stochastique des phénomènes qui induit des simplifications parfois abusives (mauvaise représentation des effets locaux ou d'une météorologie très complexe par exemple). Les méthodes déterministes incorporent les données « annexes » comme données d'entrée et calculent les concentrations par la résolution numérique d'équations aux dérivées partielles. Cette approche entraîne également des simplifications dues au formalisme mathématique des équations choisies et des processus de calcul de moyennes dans l'espace et le temps engendrés par la discrétisation spatiale et temporelle du problème.

Opérer une distinction stricte entre ces catégories de méthodes n'est donc pas judicieux par rapport à la problématique de l'évaluation de la qualité de l'air qui loin d'être une discipline complètement déterminée, répond tout de même à certaines règles.

Les approches les plus pertinentes résident donc dans les développements mixtes incluant toutes les informations disponibles (concentrations mesurées et données annexes). Celles-ci sont en plein essor actuellement.

L'objet principal de cette étude est de fournir aux AASQA un bilan sur les méthodes de modélisation disponibles pour réaliser des cartographies de la qualité de l'air dans les zones peu concernées par le réseau de mesure. Une analyse de ces techniques, des données nécessaires et de leur bon usage, de leurs performances et limites sont en particulier fournies.

Le présent document ne constitue qu'une version préliminaire d'un guide méthodologique dont la rédaction aboutira en 2003. En effet les études menées dans le cadre du LCSQA en 2003 alimenteront largement cette réflexion qui mérite d'être approfondie.

2. INTRODUCTION

2.1 CONTEXTE ET OBJECTIFS

La Directive Cadre 96/62/EC sur la qualité de l'air ambiant, et les Directives filles qui en découlent (1999/30/EC, 2000/69/EC et 2002/3/CE) fixent un cadre réglementaire de surveillance et de gestion des polluants atmosphériques. Outre la définition de valeurs cibles de concentrations, de seuils d'alerte et éventuellement de seuils d'information, ces textes prévoient l'évaluation de la qualité de l'air dans les états membres. La description de la situation observée dans chaque pays découpé en zones géographiques fait partie de cette mission. De même une information quotidienne, voire horaire (pour les oxydes d'azote, le dioxyde de soufre et l'ozone) doit être fournie au public et aux autorités compétentes.

L'élaboration de cartes de répartition des concentrations des polluants réglementés sur des bases temporelles appropriées est un bon moyen de répondre à ces objectifs. Celles-ci sont relativement aisées à construire dans les zones de petite surface, où la densité de points de mesure est suffisante (sites urbains et péri-urbains), en appliquant par exemple des règles simples d'interpolation. En revanche, là où peu de stations de mesure sont implantées, cartographier la qualité de l'air nécessite l'usage de techniques de traitement complémentaires et plus sophistiquées, souvent regroupées sous le terme de « modélisation » ou d'analyse objective.

L'application de méthodes numériques est précisément considérée dans les textes réglementaires :

- Elles peuvent être utilisées en complément des mesures disponibles sur le réseau de stations fixes dans les zones où les concentrations varient entre une limite inférieure et une limite supérieure fixées pour chaque polluant.
- La qualité de l'air peut être évaluée à l'aide des seules techniques de modélisation, dans les zones où les concentrations sont toujours plus faibles que la limite inférieure.

Suivant les polluants et suivant les indicateurs recherchés (moyenne horaire, sur 8 heures ou journalière) l'erreur commise en simulant les concentrations plutôt qu'en les mesurant doit varier entre 30 et 50%.

Cependant aucune recommandation sur le type de démarche et d'outils applicables afin de fournir l'information recherchée sur la qualité de l'air n'est disponible dans les directives.

En fait les méthodes de modélisation s'appuient nécessairement sur les sources d'information disponibles, qui se répartissent en trois grandes classes :

- Les données de mesure fournies par un réseau permanent de stations fixes dont le nombre est évidemment limité par des considérations de coût et de maintenance,
- Les données issues de campagnes de mesure spécifiques (tubes à diffusion, camions mobiles), limitées dans le temps et dans l'espace. Le nombre de points de prélèvement avec les tubes passifs peut être important afin de mieux cerner l'information obtenue par les stations fixes,

- L'ensemble des données qui ne sont pas des mesures de concentrations mais qui contribuent à expliquer la présence de tel ou tel polluant. Il s'agit des données d'activité et d'émission, des caractéristiques de site (bâti, couverture végétale, relief), et des données météorologiques.

Ces informations permettent de reconstituer des cartes de concentrations à l'aide de techniques mathématiques qui relèvent :

- Des méthodes d'interpolation,
- Des approches statistiques ou géostatistiques
- Des approches purement déterministes

Les deux premières catégories de méthodes reposent plus naturellement sur l'usage des données de concentrations mesurées. Elles correspondent à une vision stochastique des phénomènes qui induit des simplifications parfois abusives (mauvaise représentation des effets locaux ou d'une météorologie très complexe par exemple). Les méthodes déterministes incorporent les données « annexes » comme données d'entrée et calculent les concentrations par la résolution numérique d'équations aux dérivées partielles. Cette approche entraîne également des simplifications dues au formalisme mathématique des équations choisies et des processus de calcul de moyenne dans l'espace et le temps engendrés par la discrétisation spatiale et temporelle du problème.

Opérer une distinction stricte entre ces catégories de méthodes n'est donc pas judicieux par rapport à la problématique de l'évaluation de la qualité de l'air qui loin d'être une discipline complètement déterminée, répond tout de même à certaines règles.

Les approches les plus pertinentes résident donc dans les développements mixtes incluant toutes les informations disponibles (concentrations mesurées et données annexes). Celles-ci sont en plein essor actuellement.

2.2 OBJET DE L'ETUDE

L'objet principal de cette étude est de fournir aux AASQA un bilan sur les méthodes de modélisation disponibles pour réaliser des cartographies de la qualité de l'air dans les zones peu concernées par le réseau de mesure. Une analyse de ces techniques, des données nécessaires, des performances et limites, illustrées par une synthèse bibliographique des expériences recensées sont fournies en première partie. La deuxième partie du document est plus largement consacrée aux données d'entrée (mesures et autres) et propose un certain nombre de recommandations sur la nature de l'information qu'il est nécessaire de recueillir pour utiliser efficacement les modèles. En conclusion des préconisations générales sur l'usage des modèles pour la cartographie de la qualité de l'air sont fournies.

Le présent document ne constitue qu'une version préliminaire d'un guide méthodologique dont la rédaction aboutira en 2003. En effet les études menées dans le cadre du LCSQA en 2003 alimenteront largement cette réflexion. En particulier :

- Une analyse sur l'évaluation des incertitudes associées à l'usage des modèles géostatistiques et des modèles déterministes (études INERIS)
- Une évaluation de l'usage des modèles photochimiques en zone rurale (Ecole des Mines de Douai)
- Une synthèse bibliographique des méthodes d'estimation objective (INERIS+EMD)

- La définition d'une méthodologie, basée sur l'exploitation de mesures de concentrations, permettant d'élaborer des cartographies régionales de la qualité de l'air à partir des résultats issus d'un modèle déterministe continental du type CHIMERE (IPSL/LMD et Laboratoire de Statistiques de l'Université d'Orsay). Le descriptif de cette étude est donné en annexe A.
- L'extrapolation, sur une large base temporelle (annuelle) des cartographies établies à partir de campagnes de mesures (tubes passifs et camions mobiles), analyse menée avec l'outil de géostatistique ISATIS (Centre de Géostatistiques de l'Ecole des Mines de Paris). Le descriptif de cette étude est donné en annexe B.

3. LES OUTILS DE CARTOGRAPHIE

3.1 L'INTERPOLATION

L'interpolation a pour but d'obtenir une information spatiale continue sur la qualité de l'air à partir de mesures ponctuelles de concentrations. Elle consiste à mettre en œuvre des algorithmes mathématiques ou probabilistes afin d'estimer les concentrations de polluants entre les points d'échantillonnage.

Il existe de nombreuses méthodes d'interpolation, parmi lesquelles il peut sembler délicat de faire un choix.

Nous distinguerons deux catégories de méthodes :

- Les méthodes d'interpolation classique, qui utilisent des algorithmes purement mathématiques. Nous n'incluons pas dans cette catégorie les modèles physico-chimiques de la dispersion des polluants qui ne sont pas des méthodes d'interpolation.
- Les méthodes d'estimation **géostatistiques**, fondées sur une modélisation probabiliste du phénomène étudié. Ainsi elles sont basées sur la reconstitution du phénomène à l'aide d'une fonction relationnelle établie sur des analyses statistiques.

Elles s'appliquent à des *variables régionalisées*, c'est-à-dire à des fonctions numériques qui prennent leurs valeurs dans des régions délimitées de l'espace appelées *champs*. Dans notre cas, les variables régionalisées étudiées sont les concentrations de polluants.

3.1.1 Les méthodes d'interpolation classiques

Ces méthodes se divisent en deux groupes, selon les données qu'elles requièrent.

3.1.1.1 Les méthodes qui traitent uniquement les données de la variable étudiée

Cette classe regroupe un grand nombre de techniques. *La plupart définissent la valeur recherchée en un point comme une combinaison linéaire pondérée des mesures disponibles.* Ce sont souvent des méthodes utilisées dans de nombreux domaines de la physique, accessibles et implantées dans les logiciels de représentation graphique les plus connus.

Parmi les approches les plus usitées il faut noter :

- Les méthodes par partitionnement de l'espace, basées sur la définition de zones d'influence susceptibles de participer au calcul de la concentration au point cible (méthode des polygones de Thiessen par exemple). Seules les mesures les plus influentes sont donc incorporées.
- Les méthodes barycentriques qui considèrent un nombre plus important de données en affectant un poids plus représentatif à celles qui sont les plus proches du point cible, dans la formule d'interpolation : méthode d'interpolation par l'inverse des distances ou par l'inverse des carrés des distances par exemple.

- Les méthodes de régression polynomiale qui recherchent une surface définie par une équation polynomiale, s'ajustant au mieux, au sens des moindres carrés, aux points de mesure disponibles.

La mise en œuvre de ces méthodes ne nécessite pas d'effort de modélisation et permet de réaliser sans difficulté une cartographie de la variable étudiée.

De tels exemples se trouvent en Belgique (www.irceline.be/~celinair), aux Pays-Bas (www.lml.rivm.nl) ou encore en Suède (www.ivl.se), pays qui présentent des réseaux de mesure relativement denses.

Varns et al. (2001) interpolent par la méthode de l'inverse des distances les concentrations d'ozone entre 30 points de mesure situés dans l'agglomération de Dallas (mesures par tubes passifs, superficie de la zone d'étude : 24 000 km²).

Ces techniques présentent néanmoins deux défauts majeurs :

- Elles ignorent la structure spatiale de la variable, en offrant des contours très lisses d'interpolation (effet « œil de bœuf » par exemple) et peuvent omettre de représenter des situations locales très spécifiques, d'où le risque d'aboutir à des cartes peu réalistes,
- Aucun critère permettant de juger de la précision de ces cartes n'est formulé.

Les fonctions *splines* peuvent être également employées pour l'estimation locale. L'interpolation par spline consiste à ajuster aux données une surface de courbure minimale. Cette méthode produit également des surfaces interpolées lisses. C'est pourquoi elle s'applique de préférence aux variables qui évoluent lentement dans l'espace. Si en revanche les données présentent de brusques variations, elle est inappropriée.

Coyle et al. (2002) utilisent ainsi un algorithme d'interpolation qui minimise la courbure pour estimer les concentrations journalières ([O₃] entre 12h et 18h) moyennes d'ozone entre vingt stations de mesure réparties dans tout le Royaume-Uni.

La méthode des *splines* n'est pas plus détaillée, son formalisme équivalant celui du krigeage ; ce dernier a l'avantage de permettre le choix d'un modèle cohérent avec la structure spatiale du phénomène étudié (Arnaud et Emery, 2000).

3.1.1.2 Les méthodes qui nécessitent l'usage de variables secondaires

La concentration d'un polluant en un point est déduite d'une relation mathématique avec une ou plusieurs variables explicatives déterminées, relation établie généralement par régression multilinéaire.

Deletraz et Dabos (1999) ont recours à cette technique afin de cartographier les dépôts secs de NO₂ à proximité d'un axe routier. Cette méthode a servi aussi à exploiter les données du programme INTERREG II mené sur le Fossé Rhénan sous la coordination française de l'ASPA (2000) et à réaliser des cartographies thématiques. Une relation est établie entre la variable de concentration et chaque variable auxiliaire et permet de dresser une carte de concentration pour chacune de ces variables. Les variables auxiliaires sont classiquement puisées parmi les données d'émission, les modèles numériques de terrain ou l'occupation des sols. L'intégration de données satellitaires dans cette démarche tend à se développer (Ung et al, 2001, programme GMES du 6ième PCRD).

La principale difficulté de l'interpolation par régression est la recherche de variables explicatives corrélées linéairement avec la variable étudiée. Il faut par ailleurs disposer de données auxiliaires avec une résolution spatiale suffisante. Si ces difficultés et ces contraintes peuvent être surmontées, l'emploi de méthodes de régression constitue une manière simple et efficace de réaliser une cartographie. En effet, l'étape délicate qu'est la modélisation du variogramme, étape sur laquelle repose toute l'estimation géostatistique, n'est pas requise. Dans une étude comparative consistant à interpoler les concentrations moyennes saisonnières d'ozone par la méthode du plus proche voisin, des moindres carrés, du krigeage et par régression, Nikiforov et al. (1998) montrent la bonne performance de cette dernière technique.

La cartographie par régression peut être en fait considérée comme un cas particulier du krigeage avec dérive externe, dans l'hypothèse où les résidus de la régression ne sont pas spatialement corrélés, et à ceci près qu'il ne s'agit pas d'un interpolateur exact. (Pour une description du krigeage avec dérive externe, voir le chapitre 3.1.2.). Si toutefois cette hypothèse n'est pas vérifiée, les résidus contiennent alors une information sur la variabilité spatiale du polluant. Se limiter à la régression conduit à négliger cette information. Afin d'en tenir compte, il est souhaitable de faire appel à la géostatistique, en complétant la régression par un krigeage des résidus ou en effectuant directement un krigeage avec dérive externe.

Les méthodologies officielles de cartographie de la qualité de l'air adoptées au Royaume Uni font largement appel à cette approche, notamment pour les polluants principalement influencés par des sources locales (trafic routier) tels que le benzène ou les oxydes d'azote (Stedman 1998). La méthode de régression adoptée dans ce contexte sera plus largement étudiée par l'Ecole des Mines de Douai dans son programme LCSQA de 2003.

3.1.2 Les méthodes d'estimation géostatistiques

3.1.2.1 Introduction et définitions

Les principaux inconvénients des méthodes d'interpolation déterministes (structure spatiale du phénomène ignorée, pas de quantification de l'incertitude) disparaissent lorsque l'on opte pour des méthodes d'estimation géostatistiques.

Celles-ci permettent de décrire via *une analyse probabiliste le comportement des données, prenant ainsi en compte la part d'aléatoire propre à l'évaluation de la qualité de l'air, et d'intégrer une information relative à la position géographique de ces données*, d'où une amélioration des estimations dans le contexte spatial.

Les techniques de la géostatistique introduisent la notion de fonction aléatoire afin de traduire d'une part l'aspect erratique de la variable régionalisée étudiée (ici la concentration) qui empêche de prédire avec certitude les valeurs prises en différents points, et d'autre part l'existence d'une certaine structure spatiale.

a) Fonction aléatoire et loi spatiale

Dans la suite nous noterons pour les points s du domaine D étudié $z(s)$ la valeur observée. Cette dernière est considérée comme la réalisation d'une **variable aléatoire** $Z(s)$. La **fonction aléatoire** est la famille de toutes les variables aléatoires $\{Z(s); s \in D\}$. Cette fonction est également définie là où il n'y a pas de mesures disponibles.

Les données des lois de probabilité qui régissent le vecteur aléatoire $\{Z(s_1) \dots Z(s_n)\}$ que l'on peut extraire de $\{Z(s)\}$ sont régies par **la loi spatiale**. En géostatistique linéaire, celle-ci est décrite :

- par son moment du premier ordre, *l'espérance* (ou encore moyenne)

$$E[Z(s)] = m(s)$$

- par son moment du deuxième ordre, *la covariance entre deux points*. Celle-ci mesure la ressemblance entre les valeurs prises en ces deux points.

$$\text{cov}[Z(s_1), Z(s_2)] = E\{[Z(s_1) - m(s_1)][Z(s_2) - m(s_2)]\} = C(s_1, s_2)$$

L'interprétation de ces deux premiers moments impose de considérer les propriétés de stationnarité de la fonction aléatoire.

- Une fonction aléatoire est **stationnaire d'ordre deux** si l'espérance et la covariance existent et sont stationnaires :

$$\forall s, E[Z(s)] = m$$

$$\forall s, \forall h, E[Z(s)Z(s+h)] - m^2 = C(h)$$

- Une fonction aléatoire est **intrinsèque** si ses accroissements $Z(s+h) - Z(s)$ sont stationnaires d'ordre deux.

$$\forall s, E[Z(s+h) - Z(s)] = m$$

$$\forall s, \forall h, E\{[Z(s) - Z(s+h)]^2\} = 2\gamma(h)$$

$\gamma(h)$ est appelé *semi-variogramme* ou plus couramment **variogramme**.

Dans l'hypothèse de stationnarité d'ordre deux, covariance et variogramme existent et sont liés par la relation $\gamma(h) = C(0) - C(h)$. Dans l'hypothèse intrinsèque, seul le variogramme existe. C'est pourquoi il est généralement préféré à la covariance pour décrire et interpréter la structure spatiale du phénomène étudié.

b) Variogramme

C'est l'outil fondamental en géostatistique pour analyser et modéliser la structure spatiale de la variable régionalisée.

Comme il a été établi ci-dessus, le **variogramme théorique** d'une fonction aléatoire est défini par l'équation :

$$\gamma(h) = \frac{1}{2} E\{[Z(s+h) - Z(s)]^2\}$$
, où $Z(s)$ est la concentration du polluant au point de l'espace s .

$\gamma(h)$ quantifie la variabilité spatiale des concentrations entre deux sites en fonction du vecteur h qui sépare ces sites. Ses principales caractéristiques, qui nous renseignent sur celles du phénomène étudié, méritent ici d'être définies :

1. Portée (zone d'influence d'un échantillon). Pour un variogramme $\gamma(h)$ borné, la **portée** est la distance au delà de laquelle $\gamma(h)$ est constant, c'est à dire au-delà de laquelle les variables $Z(s)$ et $Z(s+h)$ ne sont plus corrélées. On appelle alors **palier** la valeur de stabilisation du variogramme.
2. Isotropie et anisotropie. Si $\gamma(h)$ ne dépend pas de la direction, il est dit **isotrope**. S'il révèle des différences selon les directions de l'espace, il est dit **anisotrope**. On distingue couramment deux types d'anisotropie :
 - les *anisotropies géométriques* (portées différentes selon les directions mais un palier identique). Dans ce cas, $\gamma(h)$ est constant lorsque le vecteur h décrit une ellipse. On peut donc retrouver une situation isotrope par un simple changement de coordonnées.
 - les *anisotropies zonales* (paliers différents selon les directions). Ces situations plus complexes sont souvent observées en présence d'une stratification de l'espace : la fonction aléatoire est constante dans le cas d'une anisotropie zonale pure ou faiblement variable à l'intérieur des strates, mais elle présente de fortes variations perpendiculairement à ces strates.
3. Comportement aux grandes distances et stationnarité. La notion de stationnarité a déjà été introduite. Le variogramme d'une fonction aléatoire stationnaire d'ordre deux est borné, celui d'une fonction aléatoire intrinsèque croît moins vite que la fonction h^2 . Une croissance plus rapide que la fonction h^2 est révélatrice de la présence d'une *dérive*, i.e. de l'absence de stationnarité. Remarquons que la stationnarité dépend de l'échelle à laquelle le phénomène est considéré.
4. Régularité. Le degré de régularité de la variable régionalisée est donné par la continuité et la dérivabilité en moyenne quadratique du variogramme $\gamma(h)$ lorsque h tend vers 0.
 - Un comportement à l'origine parabolique de $\gamma(h)$ ($\gamma(h) \sim h^2$) indique une grande régularité de la variable régionalisée qui est continue et différentiable.
 - Un comportement à l'origine linéaire ($\gamma(h) \sim h$) montre que la variable régionalisée est moins régulière (elle est continue mais non différentiable).
 - Un comportement discontinu, qui se manifeste par un saut à l'origine appelé **effet de pépité**, signale une grande irrégularité de la variable régionalisée et l'absence de corrélation entre les valeurs prises en deux points très proches. Il y a effet de pépité pur si $\gamma(h)$ est constant pour tout h strictement positif.
5. Structures gigognes. Un variogramme est constitué de structures gigognes s'il est la somme de plusieurs variogrammes de portées différentes. Il traduit un phénomène agissant à diverses échelles spatiales.

Le variogramme réel d'une fonction aléatoire est inconnu mais il peut être évalué à partir des données d'échantillonnage. On obtient ainsi le **variogramme expérimental** :

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} [z(s_i) - z(s_j)]^2$$

$N(h)$: nombre de couples de points de mesure distants de h

$z(s_i)$: concentration au point de mesure s_i .

Le variogramme expérimental n'est pas défini aux distances h pour lesquelles il n'existe aucune paire de points de mesure. En vue de l'estimation, il convient donc de lui ajuster un **modèle de variogramme** qui est la somme d'une ou plusieurs fonctions mathématiques respectant certaines contraintes et d'un éventuel effet de pépité. Parmi les principaux modèles mathématiques admissibles, on compte :

- le modèle pépitique (de palier C)	$\gamma(h) = 0$ si $h=0$ $\gamma(h) = C$ si $ h >0$
- le modèle sphérique (de palier C et de portée a)	$\gamma(h) = C[(3/2).(h /a)-(1/2).(h ^3/a^3)]$ si $ h <a$ $\gamma(h) = C$ si $ h >a$
- le modèle exponentiel (de palier C et de portée a)	$\gamma(h) = C[1-\exp(- h /a)]$
- le modèle cubique (de palier C et de portée a)	$\gamma(h) = C[7(h ^2/a^2)-(35/4).(h ^3/a^3)+(7/2).(h ^5/a^5)-(3/4).(h ^7/a^7)]$ si $ h <a$ $\gamma(h) = C$ si $ h >a$
- le modèle gaussien (de palier C et de portée a)	$\gamma(h) = C[1-\exp(- h ^2/a^2)]$
- le modèle puissance d'exposant α et de facteur d'échelle ω (hypothèse intrinsèque)	$\gamma(h) = \omega. h ^\alpha$ avec $0 \leq \alpha < 2$

c) Krigeage linéaire

Le krigeage a pour but de fournir une estimation locale **non biaisée la plus précise possible** de la variable régionalisée à l'aide d'une **combinaison linéaire pondérée des données expérimentales**.

Dans tout ce qui suit, Z désigne la variable régionalisée étudiée et Z^* l'estimateur par krigeage de cette variable. On peut rechercher une estimation ponctuelle de Z ou une estimation moyenne sur de petits volumes (ou surfaces) contigus appelés blocs. Nous nous plaçons désormais dans le cas le plus général d'une estimation par blocs. Les notations Z_v et Z_v^* correspondent respectivement à la valeur moyenne de la variable Z dans le bloc V et à la valeur moyenne estimée dans ce bloc.

Soit $Z_v^* = a + \sum_{i=1}^n w_i Z(s_i)$ un estimateur linéaire de la variable.

Pour que cet estimateur réponde aux exigences du krigeage, le paramètre a et les pondérateurs w_i doivent respecter les contraintes suivantes :

- L'erreur d'estimation est une combinaison linéaire autorisée, c'est-à-dire que son espérance et sa variance existent. Cela implique que $a=0$ et que $\sum_{i=1}^n w_i = 1$.

Cette dernière relation assure également l'absence de biais.

- L'estimateur est le plus précis possible au sens où la variance de l'erreur d'estimation est minimale, soit $\frac{\partial}{\partial w_i} \text{Var}(Z_V - Z_V^*) = 0 \quad \forall i$

Les poids w_i qui remplissent ces conditions sont solutions d'un système matriciel faisant intervenir le modèle de variogramme :

$$\begin{bmatrix} \gamma_{11} & \dots & \gamma_{1N} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{N1} & \dots & \gamma_{NN} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \mu \end{bmatrix} = \begin{bmatrix} \bar{\gamma}(s_1, V) \\ \vdots \\ \bar{\gamma}(s_N, V) \\ 1 \end{bmatrix}$$

γ_{ij} est la valeur du variogramme entre les points s_i et s_j .

$\bar{\gamma}(s_i, V)$ est la moyenne du variogramme entre s_i et le volume V .

Le paramètre μ est un multiplicateur de Lagrange qui intervient dans la résolution du système de krigeage et dans le calcul de la variance de l'erreur.

Le krigeage possède deux propriétés importantes :

- C'est un **interpolateur exact**, c'est-à-dire que les valeurs estimées aux points de mesure sont égales aux valeurs expérimentales.
- La valeur moyenne estimée dans un bloc est égale à la moyenne des estimations à l'intérieur de ce bloc.

D'autre part, il effectue généralement un **lissage**. Cela signifie que les estimations sont moins variables que les concentrations que l'on cherche à estimer.

La variance de l'erreur d'estimation, appelée **variance de krigeage**, est donnée par la relation :

$$\sigma_K^2 = \sum_i w_i \bar{\gamma}(s_i, V) - \bar{\gamma}(V, V) + \mu$$

Elle ne dépend que du modèle de variogramme et de la configuration relative du point ou du bloc à estimer et des données expérimentales.

Nous avons défini le principe du krigeage. Le terme de krigeage couvre en réalité différentes méthodes d'estimation locale qui sont choisies selon que la moyenne de la variable régionalisée est ou non connue dans le champ, selon qu'une non stationnarité a été ou non mise en évidence lors de l'étude du variogramme, et selon que des variables auxiliaires sont ou non disponibles (Tableau 1).

Tableau 1 – Les méthodes d'estimation géostatistiques

	Méthode	Variables requises	Cadre d'utilisation	Objectif
Géostatistique linéaire Cadre stationnaire ou intrinsèque	Krigeage simple	Concentration du polluant étudié	Fonction aléatoire stationnaire ou intrinsèque à moyenne connue	Cartographier les concentrations du polluant
	Krigeage ordinaire (cas le plus fréquent en pollution de l'air)	Concentration du polluant étudié	Fonction aléatoire stationnaire ou intrinsèque à moyenne inconnue	
	Cokrigeage ordinaire	Concentration du polluant étudié Une ou plusieurs variables auxiliaires connues aux points de mesure (cas <i>isotopique</i>) ou sur un ensemble de points différents (cas <i>hétérotopique</i>)	Fonctions aléatoires stationnaires ou intrinsèque Corrélation entre la variable d'intérêt et la ou les variables auxiliaires. Existence d'un lien physique entre ces variables.	
	Cokrigeage multicolocalisé	Concentration du polluant étudié Une ou plusieurs variables auxiliaires connues en tout point d'un maillage dense (cas <i>hétérotopique</i> dense)	Fonctions aléatoires stationnaires ou intrinsèque. L'hypothèse de stationnarité n'est toutefois pas réellement requise sur les variables denses. Corrélation entre la variable d'intérêt et les variables auxiliaires Le variogramme croisé de la variable d'intérêt et d'une variable auxiliaire est proportionnel au variogramme simple de cette variable auxiliaire.	

	Méthode	Variables requises	Caractéristiques du phénomène	Objectif
Géostatistique linéaire Cadre non stationnaire	Krigeage universel	Concentration du polluant étudié	Fonction aléatoire non stationnaire La variable régionalisée est considérée comme la somme d'une dérive déterministe et d'un résidu aléatoire.	
	Krigeage avec dérive externe (prolonge la théorie du krigeage universel)	Concentration du polluant étudié Une ou plusieurs variables auxiliaires connues en tout point d'un maillage dense	On suppose qu'il existe une fonction de forme qui modèle la structure générale de la variable régionalisée, et qui peut être définie au moyen des variables auxiliaires.	
	Krigeage intrinsèque généralisé (généralisation des krigeages simple, ordinaire, et universel)	Concentration du polluant étudié	Fonction aléatoire non stationnaire	
Géostatistique non linéaire	Krigeage disjonctif Espérance conditionnelle	Concentration du polluant étudié		Etablir une carte de probabilité de dépassement de seuil

A ce stade de l'étude, nous nous intéressons uniquement **aux méthodes de la géostatistique linéaire**.

3.1.2.2 Mise en œuvre de la méthodologie

La mise en œuvre des techniques d'estimation géostatistique et l'établissement de cartes de concentration nécessitent au préalable une analyse approfondie des données expérimentales afin de décrire et de modéliser la structure spatiale de la variable de pollution étudiée. Elle comporte généralement les étapes suivantes :

- L'analyse exploratoire des données

Cette étape a pour fin de dégager les principales caractéristiques de la variable de pollution et de construire un modèle géostatistique.

1. L'examen de l'histogramme permet de détecter la présence d'éventuelles observations atypiques qui risquent d'influencer les résultats de l'analyse. Il n'existe pas de règle générale sur la façon de prendre en compte ces données. Celle-ci relève plutôt de l'expérience de l'utilisateur et du jugement qu'il porte sur la validité et la représentativité de ces données.

L'examen de l'histogramme est souvent conduit en parallèle avec l'étude du variogramme expérimental $\hat{\gamma}(h)$. Ce dernier est calculé à partir des données d'échantillonnage, ce qui implique qu'on définisse différents paramètres de calcul :

- les directions de calcul. Ce paramètre est indissociable de **l'analyse des anisotropies**. En calculant le variogramme dans différentes directions de l'espace, on cherchera à voir s'il se différencie ou non selon ces directions. S'il est identique quelle que soit la direction considérée, on est dans le cas isotrope et le variogramme peut être calculé simultanément dans toutes les directions. Dans le cas contraire, on est en présence d'une anisotropie dont il convient d'identifier les directions principales. On peut s'aider à cette fin de la carte variographique, qui est la représentation en deux dimensions du variogramme.
- le pas de calcul. Il est pris égal à la maille d'échantillonnage dans le cas d'un échantillonnage régulier.
- la tolérance angulaire et la tolérance sur les distances. Ces deux paramètres permettent d'atténuer le caractère erratique du variogramme et de le rendre plus robuste lorsque les données expérimentales employées dans son calcul sont réparties irrégulièrement dans l'espace.

La définition de ces paramètres détermine l'aspect du variogramme expérimental que l'on doit pouvoir aisément interpréter et modéliser.

Notons que le variogramme expérimental est calculé d'ordinaire jusqu'à une distance voisine du tiers ou de la moitié du champ.

2. Un modèle composé d'une ou plusieurs fonctions mathématiques élémentaires choisies dans la liste des fonctions autorisées lui est alors ajusté (modèle de régionalisation).

On accordera un soin particulier à l'ajustement de l'**effet de pépité** que l'on pourrait observer. Celui-ci est dû à la variabilité des concentrations à très courte distance et/ou à la présence d'une erreur d'échantillonnage ou de mesure. L'effet de pépité peut influencer de façon significative les résultats de l'estimation. Il a la propriété d'atténuer les contrastes dans une carte et d'accroître la variance de krigeage.

3. L'analyse variographique exige que l'on prête également attention à la question de la **stationnarité**. Les algorithmes de la géostatistique stationnaire (krigeage simple, krigeage ordinaire, cokrigeage) ne s'appliquent en effet que si la fonction aléatoire est stationnaire d'ordre deux ou intrinsèque. Si l'analyse du variogramme met en évidence une dérive, il faut recourir aux techniques géostatistiques plus complexes développées dans le cadre non stationnaire. Un modèle de variogramme adapté à cette situation doit être alors défini.

4. Une façon de contrôler l'ajustement du modèle de variogramme est la **validation croisée**. Celle-ci consiste à ignorer provisoirement une donnée de mesure et à l'estimer par krigeage en utilisant les données restantes. La comparaison entre valeurs mesurées et valeurs estimées et l'étude statistique des erreurs permettent d'apprécier la qualité du modèle. La validation croisée sert aussi à déterminer le **voisinage de krigeage** le plus approprié. Le voisinage de krigeage est dit *unique* si tous les sites de mesure interviennent dans l'estimation en un point. Il est dit *glissant* s'il se réduit à une portion de l'espace autour de chaque point à estimer. Dans ce dernier cas seuls les sites de mesure contenus à l'intérieur du voisinage sont pris en compte dans l'estimation. Un voisinage glissant a généralement une forme circulaire ou elliptique. Il doit contenir un nombre suffisant de points de mesure et être cohérent avec les caractéristiques de la variable et le domaine de validité du modèle variographique.

- Estimation

Une fois que le modèle géostatistique a été défini, l'estimation par krigeage est mise en œuvre.

La résolution des équations de krigeage permet d'estimer les concentrations en tout nœud d'une grille dans le cas d'un krigeage ponctuel, ou à l'intérieur de surfaces carrées contiguës dans le cas d'un krigeage de blocs. Le pas de la grille d'estimation est fixé en fonction de l'échelle du phénomène étudié, des dimensions du domaine d'étude et du détail que l'on souhaite avoir.

Remarque : Lorsque le phénomène étudié ne présente pas de structure de corrélation spatiale (phénomène purement erratique), le meilleur estimateur linéaire que l'on puisse avoir est la moyenne mobile. L'absence de corrélation spatiale peut être en effet modélisée par un effet de pépite pur, lequel attribue des poids égaux à tous les points.

- Etude multivariable

En plus de prendre en compte la structure spatiale du polluant, les techniques géostatistiques offrent la possibilité d'améliorer les représentations cartographiques au moyen d'informations secondaires. C'est pourquoi l'analyse des variables auxiliaires disponibles et de leurs corrélations avec les concentrations de polluants exige un soin tout particulier. Ce point est détaillé au paragraphe 3.1.2.6.

Le rapport intitulé *Estimation de la qualité de l'air dans les zones peu ou pas couvertes : application à la problématique d'une association* illustre la démarche qui vient d'être décrite.

3.1.2.3 Les données de concentration

Les cartographies par interpolation permettent d'exploiter des données de concentration issues de mesures. Ces données peuvent provenir de stations de mesure fixes, de moyens mobiles ou de tubes à échantillonnage passif. Elles sont décrites plus précisément au chapitre 4. Les AASQA réalisent presque exclusivement leurs cartographies à partir des données de tubes passifs. En effet, seul ce mode d'échantillonnage permet de collecter à des coûts acceptables un grand nombre de données dans l'espace, afin d'évaluer la qualité de l'air d'une agglomération, d'un département ou d'une région entière. L'information fournies par ces cartographies reste cependant limitée à la période de la campagne, soit deux à quinze semaines environ.

Les données d'analyseurs (stations fixes et de moyens mobiles) n'en constituent pas moins une aide précieuse pour exploiter les données de tubes et quantifier l'incertitude de ces dernières. Des techniques de calcul de l'incertitude, étudiées notamment par AIR NORMAND et fondées sur la norme NF/ISO 13752 ou en ce qui concerne l'ozone sur la norme NF ENV 13005, ont été appliquées par des AASQA à l'occasion de campagnes de mesure (ATMO Picardie, 2000, ASCOPARG, 2001). Cette incertitude peut être ensuite incorporée dans la modélisation géostatistique en tant que variance de l'erreur de mesure.

Des travaux sont en cours pour tirer parti des différents types de mesures (tubes passifs, camions mobiles et stations fixes) et extraire de celles-ci une information plus riche et plus précise sur les caractéristiques spatiales et temporelles de la qualité de l'air. L'étude conduite en 2002 en partenariat avec le Centre de Géostatistique de l'Ecole des Mines de Paris devrait fournir des éléments de réponse préliminaires.

3.1.2.4 L'échelle spatiale

Les techniques de la géostatistique sont communément employées par les AASQA pour estimer la qualité de l'air dans une agglomération (ex : ASPA, 2002). L'étude de la qualité de l'air dans les zones peu couvertes par les réseaux de mesure impose cependant de considérer des domaines relativement vastes, larges de plusieurs dizaines, voire plusieurs centaines de kilomètres. On peut ainsi chercher à cartographier la pollution atmosphérique sur des départements ou des régions entières, ce qui soulève quelques questions. A cause de l'échantillonnage nécessairement limité, l'imprécision des estimations ne risque-t-elle pas d'être importante et de rendre peu fiables et peu significatives les cartographies produites ? Une même carte peut-elle englober des zones très contrastées (du fait du relief, de l'occupation des sols, de la nature des émissions...) ?

L'expérience de plusieurs AASQA est riche d'enseignement sur ce sujet. L'ASCOPARG a étudié en profondeur l'élaboration d'une cartographie de l'ozone dans le département de l'Isère (125 km x 125 km), qui se caractérise par une occupation du sol et un relief très variés. Une cartographie à si grande échelle se révèle faisable et pertinente, pourvu que le schéma d'échantillonnage soit judicieusement défini et que la ou les variables auxiliaires soient soigneusement choisies. Ces deux points font l'objet de paragraphes spécifiques. Notons que Casado et al. (1994) cartographient les concentrations horaires moyennes d'ozone au sud-ouest des Etats-Unis dans un domaine de superficie comparable (160 km x 160 km).

Pendant l'été 2000, six associations ont coopéré pour mesurer l'ozone et le dioxyde d'azote dans le Nord de la France. Le domaine d'étude couvre une superficie de 56000 km². Les représentations cartographiques de ces polluants ont été obtenues par la géostatistique. Cette entreprise de grande ampleur confirme la possibilité d'utiliser les méthodes géostatistiques à l'échelle régionale. Mais comme il a été mentionné, **il convient d'être vigilant sur l'échantillonnage spatial qui doit s'adapter aux gradients présumés de concentration, et sur le choix des sites de mesure utilisés dans l'étude variographique.** Dans cette étude seuls les sites ruraux, considérés comme les plus représentatifs du phénomène global, interviennent dans le calcul du variogramme. Les autres sites (urbains, périurbains) ne sont employés qu'au moment de l'estimation.

La coexistence dans un même domaine d'étude de zones où le polluant agit différemment constitue l'une des principales difficultés de la cartographie à l'échelle régionale.

Remarquons qu'Anglais et Américains n'hésitent pas à élaborer des cartographies sur de très vastes territoires, uniquement à partir des données de stations fixes (Coyle et al., 2002, Purushothaman et al., 1999, Nikiforov et al., 1998, Lefohn et al., 1988). La recherche de la précision ne constitue pas nécessairement une exigence. Leur objectif est de fournir à grande échelle une image possible des concentrations auxquelles les populations ou les milieux naturels sont en moyenne exposés. Les conclusions que l'on peut tirer localement de telles représentations sont de nature qualitative et peuvent guider la réalisation de campagnes de mesure complémentaires.

3.1.2.5 L'échelle temporelle

Plusieurs critères relatifs au comportement du polluant peuvent intervenir dans le choix de l'échelle de temps. La littérature témoigne de la variété des types de carte, notamment en ce qui concerne l'ozone. Selon les caractéristiques de ce polluant que l'on cherche à mettre en évidence, et suivant l'objectif de l'étude (évaluation journalière, saisonnière ou sur le long terme de l'exposition à l'ozone, compréhension des épisodes de pollution...), le pas de temps considéré est différent (Tableau 2).

Tableau 2 - Cartographie par interpolation de l'ozone: exemples d'échelles temporelles

Auteur	Variable étudiée	Période considérée	Objectif
Georgopoulos et al., 1997	[O ₃]h à chaque heure du jour	Jours de pointe d'ozone	Compréhension de l'évolution spatio-temporelle de l'ozone pendant les épisodes de pollution
Casado et al., 1994	[O ₃]h à chaque heure du jour	Moyenne sur un mois	Etude d'un profil moyen journalier
Lefohn et al., 2	[O ₃] sur 7 h (entre 9h et 16h) ----- Pourcentage de [O ₃]h supérieures à 0,07 ppm	Moyenne sur un mois ou sur une saison (mai-sept par ex) -----	Etude d'une situation moyenne pour évaluer l'exposition de la végétation à l'ozone
Liu et al., 1996	[O ₃] sur 12h (entre 8h et 20h pour le jour, 20h-8h ? pour la nuit)	Etude jour par jour durant l'été	Etude de la répartition spatiale de l'ozone en milieu urbain et de son évolution au cours de l'été
Nikirov et al., 1998	[O ₃] h max du jour [O ₃] 8h max du jour [O ₃] 8h (10-18h) [O ₃] 12h (10-22 h) Σ([O ₃] > 60 ppm)	Moyenne sur la saison estivale (Juin-Août) et sur 10 années consécutives (1981-1990)	Etude de la pollution sur le long terme (à des fins épidémiologiques)

La réalisation de telles cartographies exige dans la plupart des cas que l'on dispose de données spatio-temporelles denses.

Jusqu'à présent, les cartes de pollution établies par les AASQA représentent la qualité de l'air sur des durées moyennes d'une à quelques semaines environ qui correspondent aux durées des campagnes par échantillonnage passif. **Les stations fixes sont en effet trop peu nombreuses dans une région pour que l'on puisse établir des cartes annuelles ou pluriannuelles à partir de leurs seules données.**

On pourrait toutefois envisager d'exploiter les données de stations fixes et de moyens mobiles afin d'extrapoler à l'échelle d'une saison entière les cartes qui résultent des mesures par tubes passifs ou de restituer les fluctuations horaires ou journalières de ces cartes. Actuellement cette approche ne peut être aisément mise en œuvre à l'aide des outils de géostatistique disponibles. Ce point est à l'étude.

3.1.2.6 Amélioration de l'estimation au moyen de données auxiliaires

Intérêt des variables auxiliaires

Les coûts, le manque de moyens humains et les contraintes techniques peuvent limiter grandement la taille de l'échantillonnage. Afin de pallier l'insuffisance des données et accroître la précision de l'estimation dans les zones peu échantillonnées, il est recommandé d'utiliser des variables auxiliaires, connues en de plus nombreux points de l'espace et corrélées avec la variable étudiée. De plus les variables auxiliaires peuvent avoir un effet déterminant sur l'estimation des concentrations et la précision de cette estimation, lorsque le variogramme du polluant est peu structuré.

L'amélioration des résultats du krigeage par l'introduction de variables auxiliaires est attestée par de nombreuses études, que ce soit en pollution de l'air (Bobbia et al., 2001, AIRPARIF, 2000, Philips et al., 1997), en météorologie (Kyriakidis et al., 2001) ou en pollution des sols (Bishop et McBratney, 2001, Bourennane et al., 2000).

La plupart de ces études montrent une diminution de l'écart-type de krigeage, qui quantifie la dispersion possible des valeurs réelles non mesurées autour des valeurs estimées. Les modèles sur lesquels s'appuie le krigeage semblent de surcroît reproduire plus fidèlement la réalité observée : des procédures de validation comme la validation croisée, le jackknife ou l'utilisation de jeux de données tests, conduisent en effet à une erreur d'estimation (différence entre observation et estimation) plus faible.

La possibilité d'améliorer la précision de l'estimation à l'aide de variables auxiliaires constitue l'un des avantages majeurs des méthodes géostatistiques de cartographie. L'introduction de variables auxiliaires permet aussi de produire des cartes plus détaillées, moins lisses que celles que l'on obtient par krigeage ordinaire, et plus représentatives de la réalité (Bobbia et al., 2000)

Quel type de variables ?

Ces variables peuvent expliquer directement les concentrations du polluant étudié ou la dispersion de celui-ci dans l'atmosphère (ex : émissions, topographie). Mais elles peuvent être aussi liées à ces concentrations de manière indirecte (ex : les émissions de composés issus du même type de source, le comptage automobile, indicateur des émissions du trafic, la densité de population ou de bâti, indicative de l'importance des émissions anthropiques).

Les qualités requises de ces variables sont :

- d'être stables :
Cette propriété peut être évaluée sur les statistiques et sur le variogramme de ces variables. Une forte instabilité se traduit par une modification notable de ceux-ci lorsqu'un petit nombre de points sont retirés du calcul.
- d'être explicatives :
On recherche des variables corrélées linéairement à la variable d'étude.

Or ces deux propriétés ne sont pas toujours vérifiées, ce qui conduit parfois à transformer les variables auxiliaires initiales (par des fonctions polynomiales ou logarithmiques par exemple). En particulier la transformation logarithmique peut avoir un effet stabilisateur et améliorer la corrélation. Plusieurs variables auxiliaires peuvent être aussi combinées dans une unique fonction (ASCOPARG, 2000).

En présence de nombreuses variables auxiliaires, une **analyse en composantes principales** est un moyen de sélectionner les variables explicatives les plus pertinentes. Elle permet de distinguer des groupes de variables qui présentent des comportements similaires, et de repérer ceux qui sont les plus proches de la variable de pollution.

Comment incorporer ces variables ?

Nous considérons désormais les variables auxiliaires après une éventuelle transformation (cf. ce qui précède). Différents cas se présentent :

- a) Les données sont disponibles aux mêmes points de mesure que le polluant étudié (cas homotopique)
- b) Les données sont disponibles en un plus grand nombre de points (cas hétérotopique), sans être toutefois connues de manière très dense : sortie d'un modèle d'émission selon un maillage relativement lâche (AIR Pays de la Loire, *in* Table Ronde des Utilisateurs d'Isatis, 2002), données d'autres polluants issues d'un plus grand nombre de stations...

Dans ces deux situations, la prise en compte des données auxiliaires peut se faire par un **cokrigage** classique, ce qui nécessite une modélisation multivariée : les variogrammes simples de la variable de pollution et de la variable auxiliaire et le variogramme croisé ($\hat{\gamma}_{12}(h) = \frac{1}{2N(h)} \sum_{N(h)} [z_1(s_i) - z_1(s_j)][z_2(s_i) - z_2(s_j)]$) sont modélisés (modèle dit de *corégionalisation*) et introduits dans le système d'équations du cokrigage.

Si en général le cokrigage homotopique améliore peu les résultats du krigage, voire pas du tout en cas de parfaite corrélation entre les variables, l'amélioration apportée par le cokrigage hétérotopique est souvent significative. **Une variable auxiliaire, même lorsqu'elle n'est pas connue de manière très dense, mérite d'être prise en compte.** L'inférence d'un modèle de corégionalisation n'est toutefois possible que s'il existe un nombre suffisant de points pour lesquels les valeurs des deux variables sont connues. Par ailleurs, on choisira préférentiellement une variable auxiliaire liée physiquement à la variable d'étude.

Le cokrigage peut être étendu à plusieurs variables.

- c) Les données auxiliaires sont disponibles en tout point d'une grille relativement fine à l'échelle du domaine.

Deux techniques permettent l'introduction de données hétérotopiques denses dans le krigage (Wackernagel, 2001).

- L'une, celle du cokrigeage prend en considération la structure de covariance croisée. Toutefois les difficultés numériques induites par la grande densité de données secondaires imposent de réduire la taille de voisinage. On effectue alors un **cokrigeage colocalisé ou multicolocalisé** (Wackernagel et al., 2002, Rivoirard, 2001). :
 - Dans un cokrigeage colocalisé au sens strict, l'estimation en un point tient compte de la valeur de la variable auxiliaire en ce seul point. Hormis le cas où la variable de concentration se décompose en une fonction linéaire de la variable auxiliaire et en un résidu sans corrélation spatiale, le cokrigeage colocalisé ne constitue jamais qu'une approximation du cokrigeage.
 - Dans un cokrigeage multicolocalisé, l'estimation en un point tient compte des données de la variable auxiliaire en ce point et aux points expérimentaux. L'utilisation la plus rigoureuse et la plus efficace du cokrigeage multicolocalisé impose que les modèles de covariance simple et croisée respectent certaines hypothèses (Le variogramme croisé des deux variables doit être proportionnel au variogramme simple de la variable auxiliaire). Si tel n'est pas le cas, le cokrigeage multicolocalisé n'est là encore qu'une approximation du cokrigeage classique.
- L'autre technique consiste à considérer la variable auxiliaire comme une dérive externe : il s'agit du *krigeage avec dérive externe* (KDE). Le KDE s'applique aux situations où la variable auxiliaire $s(x)$ peut s'interpréter comme la tendance générale du comportement du polluant. On suppose ainsi que $Z(x)$ est liée à $s(x)$ par la relation :

$$Z(x) = a_0 + a_1 \cdot s(x) + R(x)$$

$s(x)$ est la dérive à une constante additive et multiplicative près.

$R(x)$ est le résidu (aléatoire).

Il faut noter que le KDE ignore la structure de covariance croisée avec la variable principale. Aussi est-il intéressant de s'orienter vers cette approche dans le cas où les données de la variable principale, en nombre insuffisant, ne permettent pas d'inférer un modèle de corégionalisation.

Et en l'absence de variable auxiliaire ?

Même en l'absence de variable auxiliaire, l'estimation par krigeage demeure préférable à toute autre méthode d'interpolation rapide, pourvu qu'il existe une structure de corrélation spatiale.

Signalons à cette occasion que l'utilisation du krigeage ordinaire pour cartographier l'ozone est une pratique courante (Purushothaman et Georgopoulos (1999), Liu et al., 1996, Casado et al., 1994, Lefohn et al., 1987, 1988).

3.1.2.7 Incertitudes d'une cartographie

Les méthodes géostatistiques présentent l'avantage d'associer aux cartes d'estimation une carte de la variance de l'erreur d'estimation appelée *variance de krigeage*. Celle-ci quantifie la dispersion possible de la valeur vraie autour de la valeur estimée. Comme il a été mentionné, elle ne dépend que du modèle de variogramme et de l'implantation relative du point ou du bloc à estimer et des données expérimentales.

Moins le phénomène est structuré (variogramme à forte composante pépitique, courte portée), plus la variance de krigeage augmente. La variance de krigeage traduit ainsi la plus grande difficulté du krigeage à fournir des estimations locales précises des concentrations lorsque la variable régionalisée est très irrégulière. Par ailleurs la variance de krigeage est plus importante dans les endroits plus pauvres en données de mesure. La carte de variance de krigeage permet donc d'évaluer la précision de l'estimation selon l'endroit considéré dans le domaine d'étude et de repérer les zones où l'échantillonnage est insuffisant.

Il est important de remarquer que la variance de krigeage n'intègre pas l'incertitude de modélisation et qu'elle est sensible au modèle de variogramme. C'est la raison pour laquelle une modification du variogramme, si elle altère peu les résultats du krigeage, peut en revanche conduire à des valeurs de variance de krigeage sensiblement différentes. Si la variance de krigeage constitue un bon indicateur de l'efficacité de l'échantillonnage et de la qualité de l'estimation, il faut se garder d'en tirer un jugement trop hâtif sur l'incertitude réelle de la carte obtenue.

La question de l'incertitude des cartes sera plus largement développée en 2003.

3.2 LES MODELES DETERMINISTES

3.2.1 Description des méthodes

Les méthodes d'interpolation précédemment décrites sont basées sur le traitement mathématique de données de concentrations mesurées dans l'air ambiant. *Les méthodes déterministes s'appuient sur l'analyse et la simulation mathématique, en fonction du site et de la météorologie, de la transformation chimique et du transport d'émissions polluantes d'origine anthropogénique et naturelle.*

Ces modèles reposent sur la résolution numérique des équations aux dérivées partielles qui régissent le comportement physico-chimique des polluants. On distingue différentes approches pour aborder cette question (Rouil et Wroblewski, 2001). Parmi les plus courantes :

- les méthodes analytiques qui proposent une solution exacte d'un problème simplifié. Dans le domaine de la dispersion atmosphérique les approches basées sur la représentation gaussienne du panache de polluant sont parmi les plus utilisées.
- Les méthodes lagrangiennes considèrent le panache de polluants comme une infinité de particules, l'évolution de chacune d'entre elle étant suivie dans un référentiel qui lui est propre. Les concentrations sont déduites par sommation du nombre de particules à un instant donné, dans un volume donné.
- Les méthodes eulériennes considèrent un référentiel fixe sur lequel est bâti un maillage tridimensionnel. Les concentrations en chaque nœud de cette grille sont alors calculées par intégration numérique des équations, via des processus itératifs souvent très élaborés.

Quelle que soit la démarche retenue, lorsque les composants étudiés sont considérés comme inertes chimiquement, seule la partie transport est traitée (équation de transport-diffusion). En revanche le traitement des polluants secondaires issus de la transformation chimique de composés primaires (ozone, aérosols), est assuré par un module spécifique, qui gère un nombre souvent réduit de réactions chimiques (par souci de simplicité et de performance numérique) agissant sur des espèces modèles. En effet la chimie de l'ozone met en jeu des centaines de Composés Organiques Volatils (COV) qu'il est illusoire de prendre en compte de manière individuelle. Aussi il est d'usage de « réduire » les mécanismes chimiques en raisonnant sur un nombre limité de classes de composés jugées représentatives (espèces modèles) et de réactions. Les phénomènes de dépôt sec et humide, déterminants dans le cas des particules, et de certains polluants gazeux (ozone, dioxyde de soufre) sont également modélisés par des modules spécifiques.

3.2.2 L'échelle spatiale

Ces méthodes sont par nature bien adaptées à la représentation cartographique de la qualité de l'air. En effet elles permettent de calculer les concentrations en chaque point d'une grille tridimensionnelle entièrement définie par l'utilisateur.

Différentes hypothèses de modélisation sont applicables suivant l'échelle considérée, mais il n'y a pas de contraintes strictes sur l'étendue du domaine de calcul : il pourra aussi bien se limiter à la rue (à condition que le modèle puisse reproduire les phénomènes de turbulence caractérisant cette échelle), que s'étendre au niveau du continent (ce qui autorise quelques simplifications).

Il s'agit certainement de l'un des principaux atouts de la modélisation déterministe. Cette technique reposant sur une conceptualisation mathématique des phénomènes physiques, elle est théoriquement adaptée là où peu de données de mesures sont disponibles, et typiquement en zone rurale.

De plus si cela est cohérent avec la précision des données d'entrée et s'il n'existe pas de contraintes de temps de calcul ou de stockage informatique (ce qui est parfaitement illusoire en pratique) il est théoriquement possible de représenter les phénomènes sur des mailles de petite taille et de les cartographier avec une bonne précision.

Cependant la part d'aléatoire propre aux phénomènes de dispersion atmosphérique ne peut être ainsi représentée, ce qui induit par défaut une incertitude naturelle dans les résultats.

3.2.3 L'échelle temporelle

La cartographie de la qualité de l'air comme outil d'aide à la surveillance implique de disposer d'informations sur des bases temporelles relativement longues, de l'ordre de la saison ou de l'année. Cela pose un certain nombre de difficultés par rapport à l'usage de certains modèles déterministes. En effet, le temps nécessaire à la réalisation des calculs constitue indéniablement une limitation à leur usage en conditions opérationnelles.

Cependant, cette constatation ne vaut pas pour les modèles analytiques de type gaussien qui reposent sur le calcul en tous points et pour toute échéance temporelle des concentrations par une simple formule mathématique. Le domaine de validité de ces méthodes est cependant assez limité (de quelques centaines de mètres à une dizaine de kilomètres de la source), et leur mise en œuvre se cantonne à des problèmes spécifiques tels que l'évaluation de l'impact des rejets d'une installation industrielle ou d'un noyau urbain. Les modèles eulériens adaptables à tous types de situation spatiale, sont en revanche nettement plus coûteux à utiliser. D'immenses efforts se concentrent aujourd'hui afin de réduire notablement les coûts en temps de calcul de ces méthodes.

Le modèle CHIMERE, développé par l'IPSL a été conçu sur la base de simplifications importantes (Schmidt et al 2001) qui permettent de réduire considérablement les coûts en temps de calcul par rapport à d'autres modèles de la même gamme. Par ailleurs le développement de versions parallélisées de codes de chimie transport des polluants atmosphériques tend à se développer (Tremback et al, 2000, Wolke et al, 2001). Le principe est de permettre l'usage de ces modèles en répartissant différentes tâches sur des machines multi-processeurs, afin de diminuer les temps de calculs. Cependant il est encore admis que les coûts prohibitifs liés à l'intégration numérique des équations qui régissent ces phénomènes reste un facteur limitant pour la réalisation de simulations sur de longues périodes.

L'usage des modèles déterministes pour cartographier la qualité de l'air étant donc suspendu à ces contraintes, différentes solutions sont envisagées par les modélisateurs pour réaliser à l'aide de ces outils des simulations annuelles. Si le modèle est suffisamment optimisé et supporté par un système informatique performant, il est possible de tirer des indicateurs annuels par le traitement de résultats de simulations obtenus pour chaque jour et chaque heure de l'année. Il est par exemple possible de réaliser une année de simulation avec le modèle CHIMERE Continental en 2 jours de calcul.

D'autres méthodes s'appuient sur des analyses climatologiques et l'exploitation d'un nombre restreint de situations météorologiques, représentatives du site, auxquelles sont affectées des fréquences d'apparition (Moussiopoulos, 1998).

La question relative à la réalisation de simulations déterministes (à l'aide de modèles de type eulérien ou lagrangien) de la qualité de l'air sur de longues périodes reste néanmoins complètement d'actualité.

3.2.4 Précision des modèles déterministes et données d'entrée

La qualité et l'incertitude des résultats fournis par les modèles déterministes sont dépendantes de plusieurs facteurs (Rouil, 2001):

- La résolution du modèle, dans les directions verticale et horizontale. Cela concerne concrètement le nombre de couches utilisées pour stratifier la troposphère libre (zone dans laquelle évolue le panache de polluants). Des couches de faible épaisseur près du sol sont nécessaires pour appréhender correctement les mécanismes thermique et mécanique qui conditionnent le transport des polluants. Le pas de maillage dans les directions horizontales détermine la finesse avec laquelle l'impact des sources de pollution pourra être modélisé. Généralement plus le maillage est fin et plus les résultats sont précis à condition que cette précision soit cohérente avec celle des données d'entrée.

- La qualité des schémas et approximations numériques utilisés dans le modèle. Sur ce point, l'utilisateur du code non spécialiste des développements numériques, n'a que très peu de marge de manœuvre.
- La précision des données d'entrée.

En effet les modèles déterministes sont exigeants en terme de données d'entrée : caractéristiques de site, données météorologiques, données d'émission et conditions de concentration et de flux aux limites du domaine considéré sont indispensables pour mettre en œuvre ces méthodes. Si la résolution spatiale et temporelle de ces informations est l'un des facteurs conditionnant la qualité des résultats, il faut néanmoins en apprécier l'impact véritable, en regard du coût que représente leur acquisition, et de l'incertitude qui accompagne le traitement d'une information de plus en plus fine et donc dépendante d'un plus grand nombre de facteurs (localisation, clefs de répartition temporelle, régimes de fonctionnement des usines...). Ces considérations amènent à repositionner les exigences des modèles déterministes sur les données d'entrée par rapport aux objectifs visés (voir paragraphe 2).

3.2.5 Evaluation et validation

Les modèles déterministes s'appuient sur la résolution approchée d'équations mathématiques sensées décrire la physique du problème. Comme cela est évoqué dans le paragraphe précédent, disposer d'une formulation satisfaisante et de données d'entrée adaptées ne suffit pourtant pas à assurer la qualité parfaite des résultats. En effet, le nombre de paramètres susceptibles d'influencer le résultat final est tel qu'il est indispensable de procéder à une phase de « calage » afin d'optimiser les valeurs qui leur sont affectées. Généralement cette étape s'effectue sur des épisodes de courte durée (quelques jours) pour lesquels l'on dispose de suffisamment de données de mesure validées. Elle est incontournable, et ne dispense en aucun cas le modélisateur de remettre régulièrement au cause les résultats qu'il obtient par rapport aux informations dont il dispose.

Cette constatation conduit petit à petit à privilégier le développement d'approche « hybrides » liant modélisation déterministe et exploitation de mesures ponctuelles afin d'affiner les diagnostics.

3.3 LES METHODES HYBRIDES

La caractéristique majeure des questions liées à la qualité de l'air, et qui la distingue d'autres disciplines est que les phénomènes mis en jeu ne sont pas ni complètement déterminés ni totalement stochastiques. Une part d'aléatoire, représentée souvent par la notion de turbulence, régit l'ensemble de ces comportements.

Ainsi une vision déterministe du problème induit des simplifications dues au formalisme mathématique et à la manipulation de moyennes en temps et en espace. De même une vision purement statistique s'avère également restrictive par le manque de physique prise en compte dans la modélisation.

Ainsi l'on s'accorde de plus en plus à considérer que la solution réside certainement dans l'exploitation conjointe de ces méthodes, afin d'affiner les résultats issus de la modélisation de la qualité de l'air. Elles supposent principalement des traitements numériques complémentaires, appliqués aux méthodes déterministes, afin d'introduire dans la simulation physique et mathématique des phénomènes, les données mesurées (et supposées donc réelles), généralement des concentrations. L'idée est donc de mieux coller à la réalité du terrain qui comprend une part stochastique non considérée par les modèles déterministes.

Quelques exemples d'approches hybrides sont données ci-dessous :

- ***L'assimilation de données*** qualifie le principe de base qui vise à assurer une certaine cohérence entre calcul déterministe et mesures. Très développées dans le domaine de la météorologie ou de l'océanographie, ces approches tendent désormais à largement se répandre pour la qualité de l'air. Plusieurs méthodologies sont proposées dans ce cadre. Globalement elles consistent à réduire, par une procédure spécifique mise en œuvre au cours du processus itératif propre au modèle déterministe, l'écart entre la valeur numérique calculée, et la valeur exacte de la concentration. Cette procédure donne lieu au calcul d'une valeur « analysée » supposée plus « juste », qui est introduite dans le calcul déterministe à la place de la valeur initialement calculée. Plusieurs philosophies sont développées :
 - L'analyse peut être « séquentielle ». dans ce cas, à chaque fois qu'une mesure est disponible, une nouvelle valeur du champ analysée est calculée. Les méthodes d'interpolation optimale sont parmi les plus anciennes (1960). Elles se basent sur la décomposition de la variable analysée en combinaison linéaire de la valeur initialement calculée, et des erreurs, « correctement » pondérées. Ces poids sont déterminés afin de minimiser une mesure de la différence entre la valeur réelle et la valeur analysée. Les méthodes basées sur les filtres de Kalman entrent également dans cette catégorie (Van Loon et al, 1997). Très complexes à mettre en œuvre, elles permettent d'introduire une composante d'évolution temporelle au système. Ces méthodes sont généralement bien appropriées pour ajuster des prévisions effectuées par un modèle.
 - L'autre méthode repose sur une formulation « variationnelle » du système (Le Dimet et al, 2002, Daescu et al, 2003). Dans ce cas, les résultats du modèle sont calculés sur une période donnée (par exemple 24h), durant laquelle les mesures disponibles sont également stockées. La fonction mesurant la différence entre mesure et calcul est construite pour cette période. Le problème d'assimilation revient à rechercher les valeurs de la variable analysée qui minimisent cette fonction de coût. Pour se faire, suivant les techniques classiques d'optimisation sans contraintes, un modèle adjoint est construit, dont est déduit le gradient de la fonction de coût, qui doit être nul en situation optimale. Les méthodes dites « 3D-VAR » ou plus récemment « 4D-VVAR » sont des exemples de ces techniques. Elles sont généralement coûteuses et difficiles à mettre en œuvre de manière opérationnelle.

- ***L'adaptation statistique*** (Blond et al, 2002): les résultats issus de modèles déterministes peuvent être analysés en regard de ceux déduits d'analyses statistiques des observations aux stations disponibles. Généralement le but de cette démarche est d'améliorer la précision de la simulation déterministe localement, dans les zones où ces mesures sont localisées. En effet le modèle déterministe calcule les concentrations en chaque nœud d'un maillage tridimensionnel plus ou moins raffiné. Aussi l'idée de les corriger en fonction des données obtenues sur un réseau de stations de mesures, conduit à des résultats optimisés sur une surface relativement réduite. Ceci vaut en particulier lorsque le modèle est utilisé en mode prédictif.
- ***L'exploitation de cartographies d'erreurs*** établies à partir de l'interpolation (par exemple par la géostatistique) de la différence entre mesure et calcul aux stations du domaine permet de développer des méthodologies de réduction de l'incertitude sur les cartographies.
- ***L'utilisation des résultats issus d'un modèle déterministe comme dérive externe dans une procédure d'interpolation géostatistique*** (Lajaunie et al, 2001, Wackernagel, 2002) : des expériences de ce type ont été menées, par AIRPARIF, de manière très récente sur l'Ile de France afin d'affiner les résultats issus du modèle CHIMERE.

Même si elles font encore l'objet de nombreuses recherches menées actuellement, les méthodes hybrides se développent largement dans le domaine de la qualité de l'air avec des résultats prometteurs. Il est cependant important de rappeler qu'il s'agit de développements purement mathématiques qui peuvent s'avérer très complexes, et qui restent donc réservés aux spécialistes.

4. LES DONNEES

4.1 DONNEES DE CONCENTRATION

Elles constituent la donnée d'entrée des méthodes d'interpolation et de géostatistique, mais peuvent être utilisées pour l'analyse des résultats issus de modèles déterministes afin d'en améliorer la qualité. L'objet de ce paragraphe est d'émettre un certain nombre de recommandations pour le recueil et l'utilisation de ces informations pour cartographier la qualité de l'air.

4.1.1 Types de données

L'évaluation de la qualité de l'air d'une région peut s'appuyer sur deux catégories de mesures :

- des **mesures en continu** faites aux stations fixes. Le nombre minimal de stations est défini par la réglementation en fonction de l'environnement (urbain, périurbain, rural) et de l'importance de la population.
- des campagnes de mesures réalisées pendant des périodes de temps limitées. Les directives européennes les désignent sous l'expression de **mesures indicatives**. Ces campagnes peuvent être conduites, soit à l'aide de moyens mobiles, c'est-à-dire d'analyseurs installés dans un camion laboratoire ou une remorque, soit à l'aide de tubes à échantillonnage passif.

La nature de l'information obtenue dépend du mode d'échantillonnage choisi :

Source de données	Types de données	Information apportée par ces données pour la représentation de la qualité de l'air	Remarques
Stations fixes	Concentrations quart-horaires fournies en continu	<p>Les données de stations fixes peuvent servir :</p> <ul style="list-style-type: none"> - à établir des cartographies si elles sont suffisamment nombreuses dans le domaine étudié - à extrapoler à de plus longues périodes les informations qui résultent des campagnes de mesure, sous réserve d'un traitement statistique ou géostatistique adapté - à évaluer l'incertitude de mesure des échantillonneurs passifs 	

Tubes échantillonnage passif	à Concentrations moyennes sur une à plusieurs semaines en plusieurs points de l'espace	<ul style="list-style-type: none"> - Méthode actuellement la plus employée pour disposer à moindre coût et avec une résolution spatiale suffisante de données moyennes sur la qualité de l'air - Possibilité d'interpoler entre les données mesurées afin de cartographier la pollution sur de vastes domaines - Estimation de concentrations moyennes saisonnières ou annuelles après un traitement statistique approprié¹ 	
Moyens mobiles	Concentrations quart-horaires fournies en continu pendant quelques jours à quelques semaines, en un point donné de l'espace	<ul style="list-style-type: none"> - Moyen de disposer d'une information sur la qualité de l'air dans des communes non équipées de stations fixes - Même rôle que les stations fixes : couplage avec les échantillonneurs passifs afin de quantifier leur incertitude de mesure - Estimation de moyennes saisonnière et annuelles et de centiles, reconstitution possible de séries chronologiques, après un traitement statistique approprié¹ 	Nous ne considérons pas ici les études locales (étude d'impact, validation d'une station fixe, choix d'une future station fixe) qui ne concernent pas la réalisation d'une cartographie régionale

4.1.2 Collecte des données

4.1.2.1 Echantillonnage dans l'espace

Les données d'un analyseur peuvent être jugées représentatives de toute une zone de caractéristiques relativement homogènes (émissions, occupation du sol, relief, météorologie). La recherche d'une information spatiale plus détaillée au moyen de campagnes de mesure est néanmoins indispensable à la réalisation d'une cartographie.

Plusieurs considérations interviennent dans la définition d'un schéma spatial d'échantillonnage, en particulier :

- la précision recherchée
- la densité de l'information auxiliaire disponible

¹ Ce point est actuellement étudié au sein des groupes de travail *Moyens Mobiles* et *Tubes Passifs*.

- la taille du domaine que l'on souhaite cartographier et les moyens de mesure que l'on peut mettre en œuvre
- l'échelle spatiale du phénomène étudié (polluant transporté ou non sur de grandes distances)

Schéma d'échantillonnage

La répartition des échantillons dépend des possibilités matérielles d'installer un capteur ou un analyseur, et du type de site (urbain de fond, périurbain, rural) auquel l'on s'intéresse. On cherchera cependant, dans la mesure du possible, à disposer les échantillons selon un **maillage régulier**. Outre sa simplicité, une telle configuration a l'avantage de l'efficacité. A nombre d'échantillon égal, la régularité de la maille facilite l'inférence de la structure spatiale et garantit une meilleure précision. Il est en effet prouvé que la variance de l'erreur d'estimation d'un échantillonnage aléatoire stratifié est inférieure à celle d'un échantillonnage aléatoire uniforme. Dans le cas d'un échantillonnage régulier, cette variance est généralement plus faible encore (Arnaud et Emery, 2000, Fouquet C. (de), 1997).

Ces recommandations, qui ont été également émises par le Groupe de Travail *Echantillonneurs Passifs* s'appliquent en premier lieu aux campagnes par tubes passifs.

Densité spatiale d'échantillonnage

Peu de recommandations sur la densité d'échantillonnage apparaissent dans la littérature.

Dans son étude sur l'ozone dans le département de l'Isère, l'ASCOPARG analyse *a posteriori* l'efficacité du maillage choisi pour sa campagne par tubes (double maille de 20 km et 5km de large) et tire quelques préconisations de cette analyse.

Pour cartographier les concentrations de NO₂ dans une agglomération, le Groupe de Travail *Echantillonneur Passifs* suggère de subdiviser préalablement le domaine étudié en fonction de l'occupation du sol, et d'adapter la densité de l'échantillonnage à chaque subdivision. Ces conseils peuvent tout aussi bien s'appliquer à la cartographie de vastes régions, dans lesquelles le degré d'urbanisation est très variable.

Le Tableau 3 et le Tableau 4 synthétisent ces diverses recommandations.

Tableau 3 - Recommandations sur le maillage émises par les AASQA

Recommandations sur le maillage	Remarques
La taille maximale de la maille d'échantillonnage est celle au-delà de laquelle l'inférence d'une structure spatiale devient peu aisée, la corrélation spatiale étant très faible ou nulle.	Cette valeur est inconnue <i>a priori</i> . Toutefois, elle peut être évaluée à partir de campagnes de mesure antérieures ou d'études réalisées dans des régions de caractéristiques comparables (météorologie, topographie, émissions). Elle est généralement estimée à 20-25 km pour l'ozone.
Là où la corrélation avec la variable auxiliaire est plus faible, il est conseillé de diminuer la taille de la maille.	La corrélation avec les variables auxiliaires peut être estimée d'après le comportement dispersif des polluants. Ainsi, lorsque la variable auxiliaire <i>topographie</i> a vraisemblablement peu d'influence sur les concentrations (par exemple dans les régions de plaine où la migration des polluants est importante), il est suggéré de resserrer le maillage.
Adapter le maillage en fonction de la population présente. Dans les zones étendues, un maillage très resserré induit des coûts d'échantillonnage élevés qui ne sont pas nécessairement justifiés.	<p>Dans les zones peuplées et hautement fréquentées, il convient d'échantillonner selon un maillage suffisamment fin (5 km par exemple).</p> <p>Dans les zones peu peuplées, la perte de précision due à l'élargissement de la maille (10 à 15 km par exemple) peut être jugée acceptable.</p>
Ajuster le maillage en fonction de l'occupation du sol et des sources de pollution présentes	Resserrer le maillage dans les zones où les gradients de concentration risquent d'être élevés (ex : dans l'étude de la pollution de NO ₂ et d'ozone dans le Nord de la France, la maille de 25 km est divisée par deux et par quatre dans ces zones)

Dans tous les cas, il est préconisé **de resserrer la maille d'échantillonnage en quelques endroits**, afin de faciliter la modélisation du variogramme aux petites distances (Fouquet, 1997). Lorsqu'une région présente des zones de pollution contrastées, il convient d'effectuer ce resserré local dans un secteur de fortes concentrations et dans un secteur de plus faibles concentrations, afin d'éviter l'introduction de biais.

Tableau 4 - Quelques tailles de maille recommandées (ordres de grandeur)

Dimension de la maille	Cas dans lesquels cette recommandation peut s'appliquer	Type de cartographie
20 à 25 km	Zone étendue peu peuplée. Faibles gradients de concentration. Concentration fortement corrélée avec une variable auxiliaire	Cartographie régionale
10 à 15 km	Gradients de concentration plus élevés. Concentration moins corrélée avec les variables auxiliaires	Cartographie régionale
5 km	Forts gradients de concentration. Zone densément peuplée	Cartographie régionale
1 à 2 km	Zone rurale	Cartographie d'une agglomération
250 m à 1 km	Zone urbaine ou industrielle, zone à points chauds	Cartographie d'une agglomération

Il s'agit de valeurs indicatives à ajuster selon les moyens disponibles et les caractéristiques du domaine et du polluant étudiés.

4.1.2.2 Echantillonnage dans le temps

La question de l'échantillonnage temporel est particulièrement délicate. A quels moments de l'année, à quelle fréquence et pendant combien de temps faut-il échantillonner en un point pour disposer d'une information suffisamment représentative de la qualité de l'air en ce point ?

Dans le cas de mesures indicatives, les directives européennes fournissent à titre d'orientation des objectifs de qualité en ce qui concerne la période minimale à prendre en compte. Cette période, exprimée en % de temps, devrait être :

- supérieure à 10% en été pour l'ozone, le NO et le NO₂ (directive 2002/3/CE)
- de 14% pour le SO₂, le NO₂, les NO_x, les particules, le plomb, le benzène et le monoxyde de carbone (une mesure par semaine, au hasard, également répartie sur l'année ou huit semaines, également réparties sur l'année) (directives 2000/69/CE et 1999/30/CE)

Il n'existe pas de plan d'échantillonnage unique, applicable en toute situation. Un plan qui se révèle efficace pour un polluant et un site donné ne l'est pas nécessairement dans d'autres circonstances. Toutefois, les méthodes statistiques étudiées au sein du groupe de travail Etudes Mobiles, en particulier celle des **plans de sondage** (développée par ATMO Poitou-Charentes et en cours d'évaluation à l'INERIS) et celle de l'échantillonnage stratifié (Ecole des Mines de Douai) devraient aider à formuler des recommandations générales.

Traitement temporel des données

Quelle information tirer des données d'une campagne de mesures ? L'objectif est de disposer de concentrations moyennes saisonnières ou annuelles, que l'on puisse non seulement comparer avec des valeurs seuils mais aussi intégrer dans une cartographie. Le calcul d'un intervalle de confiance autour de la moyenne estimée est également de première importance.

L'estimation de concentrations moyennes annuelles par simple corrélation avec une station fixe semble être la méthode la plus répandue (LCSQA, Rapport *Représentativité des mesures et méthodes statistiques*, 2001), mais elle ne fournit pas d'intervalle de confiance autour de la moyenne estimée. Or les directives européennes définissent des objectifs de qualité en terme d'exactitude, ce qui suppose que l'on puisse évaluer les incertitudes sur les concentrations estimées.

Des méthodes plus poussées, que l'INERIS s'est chargé de recenser dans sa mission d'assistance au groupe de travail Moyens Mobiles, pourraient être employées pour estimer une moyenne et l'incertitude sur cette moyenne.

- Dans une étude sur le benzène en Île de France, AIRPARIF a développé une approche par modélisation, afin de réaliser des cartographies moyennes annuelles de ce polluant à partir de campagnes passives (Roth et Dégardin, 2001). Un modèle a été préalablement construit grâce à des données de stations fixes. Il exprime les concentrations journalières de benzène comme le produit de trois facteurs explicatifs : l'accumulation journalière (représentant principalement l'effet du trafic automobile), la dispersion journalière (effet du vent), et le facteur saisonnier (effet des conditions météorologiques). Ce modèle peut être calé localement par régression sur les concentrations mesurées aux points d'échantillonnage. Selon les tests effectués par AIRPARIF sur le benzène, il permettrait d'estimer la moyenne annuelle avec une incertitude de 5%. Rappelons qu'il s'agit là encore, d'une zone dense en nombre de stations de mesures, et que l'extrapolation de ce résultat en zone rurale n'est pas trivial.

- La méthode étudiée à l'Ecole des Mines de Douai est fondée sur la norme « ISO 9359 – Qualité de l'air – Echantillonnage aléatoire stratifié pour l'évaluation de la qualité de l'air ambiant ». Elle consiste à regrouper les mesures par classe météorologique, et à utiliser cette stratification pour estimer une concentration sur le long terme et l'incertitude sur cette moyenne.

- La théorie des sondages précédemment citée permet d'estimer une moyenne avec un intervalle de confiance, en s'appuyant si possible sur les données d'un site fixe auxiliaire afin d'améliorer l'estimation. L'avantage de cette méthode est de s'appliquer à tous les types de sites et de polluants, alors que ceux-ci ne se prêtent pas tous nécessairement à une modélisation, et de fournir un intervalle de confiance estimé selon une théorie rigoureuse (Tillé, 2001). En outre, elle ne requiert pas de variables externes telles que le trafic ou la météorologie. L'évaluation de cette méthode est en cours à l'INERIS dans le cadre de la mission d'assistance au GT *Moyens Mobiles* qui lui a été confiée.

4.2 DONNEES D'EMISSION

4.2.1 Nature de l'information

La description des rejets émis par les sources considérées pour évaluer la qualité de l'air dans une zone donnée, est une étape fondamentale de la modélisation déterministe. En effet, il serait illusoire d'espérer obtenir des résultats exploitables en alimentant les modèles de dispersion, aussi sophistiqués soient ils avec des données d'émission trop approximatives ou erronées. L'étape d'inventaire et de quantification des émissions conditionne la qualité des résultats des modèles de dispersion.

De plus, comme indiqué en 3.1.2, les données d'émission font partie des variables auxiliaires, qui expliquent directement ou indirectement les concentrations de polluant observées, et qui, peuvent être employées dans l'interpolation par cokrigeage (multi)colocalisé, krigage avec dérive externe ou régression.

Les difficultés liées à l'élaboration d'un inventaire d'émission résident essentiellement dans le recensement des sources de toute nature (industrie, trafic, chauffage, végétaux...) sur l'ensemble du domaine qui peut s'étendre sur plusieurs centaines de kilomètres. Ainsi le travail minutieux d'inventaire peut-être très coûteux en temps et en moyens humains. Cela est dû à l'étendue du domaine de calcul, à la variété des sources à considérer, aux lourdes incertitudes qui entachent certaines données (chauffage, végétaux), mais aussi à la bonne volonté et à la disponibilité, des organismes qui détiennent ces informations.

Le terme de « cadastre d'émissions » implique une notion de répartitions spatiale et temporelle des données de l'inventaire. Le pas de maillage du cadastre et son échelle temporelle sont déterminants suivant les applications visées et doivent être choisis comme étant le meilleur compromis entre le coût d'élaboration et la précision recherchée.

La précision spatiale du cadastre doit demeurer cohérente avec celle adoptée pour la modélisation. Ainsi pour réaliser des cartographies de veille sur de vastes étendues en zone rurale il est rarement opportun de procéder à des modélisations très fines, au km² par exemple. La maille unitaire du cadastre doit être définie de manière cohérente avec ce choix. En situation rurale des mailles de l'ordre de 5 à 10 km² sont des choix complètement acceptables.

Le problème est le même pour la précision temporelle. Elaborer un cadastre au pas horaire est extrêmement coûteux. De plus un nombre important de sources, les grandes sources industrielles par exemple peuvent enregistrer des fluctuations significatives à très court terme. Afin de lisser les résultats et d'établir des cartographies représentatives d'une situation moyenne, la solution réside plutôt dans l'identification de quelques situations types (heures de pointe du soir et du matin, effet week-end, périodes de vacances) et leur répartition convenable dans le temps suivant les périodes que l'on entend simuler. L'élaboration d'un cadastre au pas horaire est plutôt dédiée à la réalisation de simulations déterministes pour l'étude de scénarios.

4.2.2 Acquisition des données

Des inventaires d'émission à une échelle globale, existent et sont disponibles:

- inventaire EMEP (www.emep.int dans le cadre de UN-ECE²)

² UN-ECE : United-Nations, Commission Economique Européenne

- inventaire PRQA 1994 (France, inventaire CITEPA)
- inventaire GENEMIS (IER Stuttgart, www.uni-stuttgart.de/genemis programme EUROTRAC2)
- inventaire GEIA (Global Emission Inventory Activities, <http://weather.engin.umich.edu/geia>)

Néanmoins suivant les situations il peut être utile d'accéder à une information plus ciblée. Elle est établie en plusieurs étapes :

- **Définition du problème** : Cela inclut la définition des polluants que l'on souhaite incorporer dans l'inventaire et est très dépendante de l'application visée. Ainsi en cas de modélisation photochimique, il sera indispensable de disposer d'un inventaire précis des Composés Organiques Volatils émis (COV).

Mais il s'agit également d'inventorier les sources, leurs caractéristiques géométriques, et leur localisation. Les sources sont habituellement réparties suivant de grandes catégories :

- Les sources ponctuelles : elles se réfèrent typiquement aux rejets canalisés d'une installation fixe telle qu'une usine. Une enquête auprès des industriels concernés ou des DRIRE doit permettre à l'inventoriste d'accéder au type de polluants rejetés, par quelles voies et sous quelles formes, afin de juger de la nécessité d'incorporer la source ou non.
- Les sources linéiques : elles sont imputables en grande partie au trafic routier, et plus modestement à d'autres moyens de communication (train, fluvial). Dans le premier cas leur recensement nécessite une modélisation du réseau routier, et des données chiffrées sur les caractéristiques des véhicules circulant (nature du parc, style de conduite),
- Les sources surfaciques de nature anthropogénique : ce terme désigne typiquement les sources dues à l'activité humaine. Leur origine peut être liée à l'habitat (chauffage), mais aussi au transport routier (agrégation de sources linéiques sur le noyau urbain), assimilées à des surfaces d'émission sur le domaine d'étude,
- Les sources surfaciques d'origine biogénique : il s'agit des émissions de COV liées aux surfaces végétales. Leur prise en compte est en particulier déterminante dans la modélisation photochimique.

Il est nécessaire de définir ces sources par rapport à un référentiel cohérent. Plusieurs types de formats sont à la disposition de l'utilisateur, établis suivant des règles économiques (SECTEN, NAMEA) ou géographiques (PRQA, DEPARTEMENT) pour ce qui est des formats produits par le CITEPA.

Les activités émettrices sont représentées à l'aide de nomenclatures adaptées. En Europe la classification SNAP (Selected Nomenclature for Air Pollution) associée à la méthodologie CORINAIR (CORE Inventory of AIR pollution in Europe) est largement utilisée et retenue par la Commission Economique pour l'Europe des Nations Unies (www.aeat.co.uk/netcen/corinair/corinair.html) .

- **Choix de la méthode d'estimation des sources** : Il s'agit de définir une stratégie qui permette de quantifier les émissions que l'on associera à chaque source recensée. Deux techniques sont usuellement rencontrées :

- Le Bottom-up : les émetteurs sont individuellement recensés et un travail d’investigation permet d’affecter à chacun d’eux une donnée chiffrée d’émission. Par sommation, des entités plus importantes sont ainsi reconstituées automatiquement.
- Le top-down : Des sources d’informations globales, telles que les enquêtes régionales, des données INSEE, des statistiques de consommation, qui font référence à des états annuels sont exploitées et désagrégées de manière à en tirer une estimation des émissions horaires.

Au cours de l’élaboration de l’inventaire, il sera certainement nécessaire de coupler ces deux approches, suivant la nature des sources considérées (grandes sources ponctuelles ou source surfacique), les informations disponibles, le degré de complexité que l’on s’autorise. Ainsi la précision requise, mais également le budget disponible sont des critères de choix à prendre en considération.

- **Collecte des données et des informations** : Elle constitue l’étape la plus fastidieuse de l’élaboration effective du cadastre des émissions. Il s’agit du recensement de toutes les données chiffrées nécessaires au calcul des émissions. Tout le problème est d’extraire de ces éléments, ceux qui, associés à des facteurs d’émission qu’il reste à choisir judicieusement, permettront d’aboutir aux émissions. La méthodologie européenne CORINAIR, présentée par le CITEPA résume cette démarche par les équations suivantes :

$$E_{s,t} = \sum_{a=1}^{a=n} E_{s,a,t}$$

où $E_{s,t}$ est l’émission totale de la substance s pendant le temps t , et $E_{s,a,t}$, l’émission liée à l’activité a . Celle-ci se définit comme suit :

$$E_{s,a,t} = A_{a,t} \times F_{s,a}$$

$A_{a,t}$ est la quantité d’activité relative à l’activité a pendant le temps t (une consommation annuelle par exemple), et $F_{s,a}$ est le facteur d’émission relatif à la substance s pendant le temps t .

- **Traitement des données et allocation spatio-temporelle** : Les données récoltées et calculées dans les étapes précédentes doivent être formatées de manière à pouvoir être utilisées pour les applications visées (répartition spatio-temporelle des émissions).
- **Validation et formatage des résultats finaux** : Cette dernière étape concerne l’évaluation de la qualité de l’inventaire, sa vérification et sa mise en forme pour une utilisation en conditions opérationnelles. Elle peut être menée en recoupant les résultats obtenus avec des données publiées, en calculant des indicateurs de qualité ou en évaluant l’incertitude associée aux émissions calculées. Sur ce dernier point, les travaux de la communauté scientifique sont encore au stade de la recherche.

4.3 DONNEES DE SITE ET METEOROLOGIQUES

Cette catégorie recoupe les données annexes (hors émissions) susceptibles de fournir une information complémentaire lors de la mise en œuvre de modèles d'interpolation et de géostatistique (variable auxiliaire, dérive externe). Certaines d'entre elles (conditions météorologiques, modèle numérique de terrain, rugosité) constituent également des données d'entrée incontournables pour la mise en oeuvre des modèles déterministes. Les variables concernées sont :

- ***Des variables météorologiques***

- température
- vitesse de vent
- direction de vent
- humidité
- couverture nuageuse
- ...

Elles sont généralement fournies par des stations de mesure au sol représentatives à une altitude de dix mètres afin de se dégager des perturbations du sol.

Ces stations fournissent généralement des informations suffisantes pour les études locales, de type impact industriel, et sont les supports de constitution des roses de vents. De plus les données sont disponibles avec un pas de temps adapté à l'application (tri-horaire, horaire ou quart-horaire).

Il est également possible de disposer de données traitées statistiquement de manière à disposer d'informations moyennes. Il s'agit des roses des vents qui fournissent suivant chaque direction (par classe de 20°) la fréquence et l'intensité du vent. Elles sont généralement établies mensuellement, par saisons ou annuellement.

Parfois une information mesurée par radiosondage ou par des instruments tels que les radars et sodars permet de disposer d'une information complémentaire sur la structure verticale de la variable météorologique.

Enfin les sorties numériques issues des modèles de prévision météorologique (ARPEGE ou ALADIN pour la France) offrent également une discrétisation sur le domaine d'étude des variables météorologiques.

- ***La topographie*** représentée par un modèle numérique de terrain fourni par l'IGN par exemple avec des pas de discrétisation pouvant être très faibles
- ***L'occupation du sol, la densité de bâti***
- ***La densité de population*** (information INSEE)

Pour les méthodes d'interpolation géostatistique peu de références font état de l'usage des variables météorologiques. L'étude appliquée menée en 2002 par l'INERIS/LCSQA sur le département de l'Allier a montré que l'occupation du sol et la densité de population étaient en revanche des variables pertinentes pour ajuster les cartographies de qualité de l'air. Ces variables peuvent être transformées (§ 3.1.2) de manière à posséder les caractéristiques nécessaires à leur emploi dans les techniques géostatistiques.

5. CONCLUSION ET RECOMMANDATIONS

L'objet de ce document est de fournir aux AASQA une synthèse et une analyse critique des méthodes numériques disponibles pour élaborer des cartographies pertinentes de la qualité de l'air à l'aide des données disponibles (concentrations atmosphériques, émissions, météorologie, données de site). Ces techniques de modélisation ont été décrites dans la première partie du document. La seconde partie est dédiée à la description fine des données d'entrée nécessaires suivant l'application visée.

Bien que ce rapport soit la version intermédiaire d'un guide méthodologique dédié aux Associations, nous pouvons déjà émettre un certain nombre de recommandations sur l'usage de la modélisation appliquée à la représentation cartographique de la qualité de l'air.

5.1 METHODES D'INTERPOLATION CLASSIQUES

- ✓ Elles reposent sur des concepts simples d'interpolation, accessibles à tous, et sont disponibles dans bon nombre de logiciels de représentation graphique,
- ✓ Pour limiter les incertitudes, elles nécessitent, par définition, de disposer d'un réseau assez dense de mesures. Elles peuvent également intégrer par des procédures de régression des variables externes telles que les émissions, les caractéristiques de site. Cette approche est pertinente lorsque les concentrations observées sont très corrélées aux émissions (polluants passifs émis au niveau du sol où les phénomènes de dispersion sont limités, par exemple).
- ✓ Ces méthodes ne peuvent généralement reproduire la véritable structure spatiale des concentrations ce qui peut aboutir à des résultats peu réalistes.
- ✓ L'élaboration de cartographies sur de longues périodes ne peut se faire qu'à partir des données issues du réseau de mesure fixe ce qui limite l'accès à ces méthodes pour les zones rurales.

5.2 METHODES DE GEOSTATISTIQUE

- ✓ Parmi les techniques d'interpolation, les méthodes géostatistiques présentent un intérêt certain pour la réalisation de cartographies de la qualité de l'air dans les zones peu couvertes par les réseaux de mesure. La géostatistique suppose que les valeurs de concentration mesurées sont la réalisation d'un processus aléatoire dont est modélisée la fonction de covariance ou le variogramme. A la différence des méthodes d'interpolation déterministes, elle permet de prendre en compte la structure spatiale du phénomène étudié par l'intermédiaire du variogramme et de joindre à la carte d'estimation la carte de la variance de l'erreur d'estimation.

- ✓ La géostatistique propose différents algorithmes selon les caractéristiques du phénomène mises en évidence dans l'étude du variogramme et selon les données disponibles. Lorsque l'on dispose de données secondaires liées directement ou indirectement à la variable de pollution étudiée (ex : cadastre des émissions, occupation du sol, altimétrie...), il est intéressant de les prendre en compte dans l'estimation. Des variables auxiliaires, suffisamment corrélées aux concentrations servent à compenser le manque d'échantillonnage en certains endroits, à améliorer la précision de l'estimation et à rendre la carte plus réaliste.
- ✓ Les méthodes de krigeage ordinaire, de cokrigeage et de krigeage avec dérive externe ont fait leur preuve dans la cartographie de la pollution d'agglomérations urbaines pour des polluants tels que le dioxyde d'azote, le benzène ou l'ozone. Elles tendent aujourd'hui à s'élargir à des domaines de grande superficie, larges de plusieurs dizaines voire centaines de kilomètres. Or c'est précisément à de vastes territoires que l'on est conduit à s'intéresser dans l'étude des zones peu couvertes par les réseaux de mesure. L'expérience montre néanmoins qu'un minimum de 30 points de mesures est nécessaire pour réaliser une étude significative.
- ✓ Le risque de rencontrer des processus non stationnaires s'accroît avec la dimension du domaine, ce qui peut conduire à envisager de nouvelles techniques.
- ✓ La réalisation de cartes à grande échelle exige la rigueur dans la stratégie de collecte des données. L'échantillonnage spatial doit ainsi s'adapter au type d'environnement, aux caractéristiques du polluant et aux gradients supposés de concentration. L'échantillonnage passif à l'aide de tubes est la technique de mesure la plus employée pour dresser des cartes de concentrations moyennes sur plusieurs semaines ou plusieurs mois. L'exploitation des données d'analyseurs (stations fixes, moyens mobiles) offre une information plus riche sur l'évolution temporelle des concentrations. A l'heure actuelle, les outils de la géostatistique ne permettent pas d'intégrer directement cette information dans les cartographies, et donc de généraliser des cartographies sur de longues périodes. Ce point est à l'étude.
- ✓ L'évaluation des incertitudes d'une carte fait à elle seule un sujet d'étude qui sera abordé en 2003.

5.3 METHODES DETERMINISTES

- ✓ Les méthodes déterministes permettent de reconstruire des champs de concentrations en tous points d'une grille de discrétisation par la résolution numérique des équations régissant les phénomènes physico-chimiques mis en jeu. Elles s'appuient pour cela sur les caractéristiques du site étudié, les données d'émission et les données météorologiques qui constituent les points d'entrée du modèle.
- ✓ Ces méthodes fournissent toujours une cartographie quel que soit le nombre de points de mesure disponible. Pour cela elles sont une alternative idéale dans les zones peu/pas couvertes par le réseau de mesure. Il n'y a pas de limitation dans l'extension du domaine d'étude. Par nature plus le pas de discrétisation est fin, meilleure est la résolution du modèle, à condition de disposer du même niveau de qualité pour les données d'entrée. Cela induit une réflexion sur la recherche du meilleur compromis entre précision et coût de mise en œuvre. Cette question n'est pas encore résolue.

- ✓ La principale limitation de ces méthodes est l'aspect temporel. En effet l'usage des modèles les plus élaborés (eulériens ou lagrangiens), induit un coût en temps de calcul encore prohibitif. Ainsi ces méthodes s'adaptent bien à la cartographie d'épisodes de pollution (quelques jours) mais demeurent moins appropriées pour le traitement de longues périodes. Cette application, outre un matériel informatique performant, nécessite de formuler un certain nombre d'hypothèses simplificatrices pour limiter les calculs ou d'avoir accès à des techniques de parallélisation. Des progrès récents sont néanmoins constatés dans ces voies.
- ✓ L'évaluation de l'incertitude des résultats fournis par ces modèles demeure également une question en suspens. Les sources d'incertitude sont nombreuses et les processus complexes mis en jeu induisent parfois des compensations d'erreur qu'il est délicat d'évaluer (pour l'ozone en particulier). Ce point fait également l'objet d'études en cours et sera approfondi dans la prochaine version du document.
- ✓ L'incorporation de données de mesures pour analyser et retraiter, a posteriori ou durant les processus itératifs les résultats fournis par les modèles déterministes de qualité de l'air est une voie d'avenir pour améliorer la qualité des cartes produites par les modèles déterministes. Des méthodologies sont élaborées, mais elles demeurent encore délicates à utiliser en conditions opérationnelles. En revanche il faut s'attendre dans le futur à l'adoption de cette démarche de manière quasi-systématique.

6. REFERENCES

- ARNAUD M. et EMERY X. , 2000. Estimation et interpolation spatiale - *Méthodes déterministes et méthodes géostatistiques*. Hermès Science Publication.
- ASCOPARG, 2001. Mise en place d'une méthodologie pour la cartographie de l'ozone à l'échelle du département de l'Isère.
- ASPA, UMEG, 2000. Analyse transfrontière de la qualité de l'air dans le Rhin supérieur. Programme INTERREG II.
- ASPA, 2002. Diagnostic de la qualité de l'air sur l'agglomération de Mulhouse. Annexe au rapport final sur la répartition spatiale de la pollution atmosphérique (ASPA 02031901-I-D)- Méthodes d'interpolation spatiale.
- ATMO PICARDIE, AIRPARIF, AIR NORMAND, OPAL'AIR, AREMA LM, AREMARTOIS, 2000. Campagne interrégionale d'étude de l'ozone et du dioxyde d'azote par tubes à diffusion passive, 26 juin-04 septembre 2000.
- BISHOP T.F.A., McBRATNEY A.B., 2001. *A comparison of prediction methods for the creation of field-extent soil property maps*. Geoderma, 103. 149-160.
- BLOND N., BEL L., VAUTARD R., 2002, Three dimensional ozone data analysis with an air quality model over paris area, soumis à Atmospheric Environment
- BOBBIA M., PERNELET V., ROTH C., 2001. L'intégration des informations indirectes à la cartographie géostatistique des polluants. Pollution Atmosphérique n° 170 - Avril-Juin.
- BOURENNANE H., KING D., COUTURIER A., 2000. *Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities*. Geoderma, 97, 255-271.
- CASADO L.S., ROUHANI S., CARDELINO C.A., FERRIER A.J., 1994. Geostatistical analysis and visualization of hourly ozone data. Atmospheric Environment 28, n°12, 2105-2118.
- CHRISTAKOS G., VYAS V.M., 1998. A composite space/time approach to studying ozone distribution over Eastern United States. Atmospheric Environment, Vol.32, n°16, 2845-2857.
- COYLE M., SMITH R.I., STEDMAN J.R., WESTON K.J., FOWLER D., 2002. Quantifying the spatial distribution of surface ozone concentration in the UK. Atmospheric Environment, vol. 36, 1013-1024.
- DAESCU D.N., CHARMICHAEL G.R., An adjoint sensitivity method for the adaptive location of the observations in air quality modeling, J. Of the Atmospheric Sciences, Vol 60, n°2, pp 434-450, 2003
- DELETRAZ G., DARBOS P., 2001. Modélisation statistique de la pollution azotée à proximité d'un axe routier et évaluation des incidences sur l'environnement. Application au site de Buriatou. Colloque Risques, oct 2001, Besançon.

FOUQUET C. (de), 1997. Influence de la méthode d'estimation et de la maille de reconnaissance sur la quantification des pollutions : étude méthodologique à 2D. In Nicolas (ed.), Echantillonnage et environnement. Liège : Cebedoc, pp. 39-63.

GEORGOPOULOS P.G., PURUSHOTAMAN V., CHIOU R., 1997. Comparative evaluation of methods for estimating potential human exposure to ozone : photochemical modeling and ambient monitoring. *Journal of Exposure and Environmental Epidemiology*, Vol. 7, N°2

Groupe de Travail des AASQA "Echantillonneurs passifs pour le dioxyde d'azote". Guide version 1, oct. 2001

KYRIAKIDIS P.C., KIM J., MILLER N.L., 2001. Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics. *J. American Meteorological Society*, 40, 1855-1877.

LAJAUNIE C, WACKERNAGEL H, BERTINO L, 2001. Geostatistical Normalization: Case Studies, Technical Report N-31/01/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.

LE DIMET F.X., NAVON I.M., DAESCU D.N., *Second order information in data assimilation*, *Monthly Weather Review*, Vol 130, n°3, pp 629-648, 2002

LEFOHN A.S., KNUDSEN H.P., LOGAN J.A., SIMPSON J., BHUMRALKAR C., 1987. An evaluation of the kriging method to predict 7-h seasonal mean ozone concentrations for estimating crop losses. *J. Air Pollut. Control. Assn.*, 37(5), 595-602.

LEFOHN A.S., KNUDSEN H.P., McEVOY L.R., 1988. The use of kriging to estimate monthly ozone exposure parameters for the Southern United States. *Environmental pollution*, 53, 27-42.

LIU S. L.-J., ROSSINI A.J., 1996. Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan Toronto - A pilot study. *Environment International*, N° 6, 667-692.

MOUSSIOPOULOS N., 1998, Air quality in Athens : long terme trend and expected evolution in 2004, Air Pollution VI, WIT Press, Computational Mechanics Publications, pp 425-434

NF ENV 13005. Normes fondamentales. Guide pour l'expression de l'incertitude de mesure, AFNOR, 1999.

NF ISO 13752 (novembre 1998). Qualité de l'air - Évaluation de l'incertitude d'une méthode de mesurage sur site en utilisant une seconde méthode comme référence.

NIKIFOROV S.V., AGGARWAL M., NADAS A., KINNEY P.L., 1998. *Methods for spatial interpolation of long-term ozone concentrations*. *Journal of Exposure Analysis and Environmental Epidemiology*, vol. 8, n°4, 465-481.

PHILLIPS D.L., LEE E. H., HERSTOM A. A., HOGSETT W. E., TINGEY D.T., 1997. *Use of auxiliary data for spatial interpolation of ozone exposure in southern forests*. *Environmetrics*, 8, 43-61.

PURUSHOTHAMAN V., GEORGOPOULOS P.G., 1999. Integrating physico-chemical modeling, geostatistical techniques et geographical information systems for ozone exposure assessment. Ozone Research Center, Technic Report, ORC-TR99-01, May 99.

- RIVOIRARD J., 2001. Which models for collocated cokriging ? *Mathematical Geology*, vol 33, n°2, pp 117-131
- ROTH C., DÉGARDIN D., 2001. Interpreting the results of diffusive sampling campaigns with respect to yearly standards. International Conference Measuring Air Pollutants by Diffusive Sampling, Montpellier, France, 26-28 septembre 2001.
- ROUÏL L. *Evaluation des modèles et des incertitudes*, Rapport LCSQA 2001
- ROUÏL L., WROBLEWSKI A. , Guide méthodologique en modélisation, rapport LCSQA 2001
- SCHAUG J., IVERSEN T., PEDERSEN U., 1993. Comparision of measurements and model results for airborne sulfur and nitrogen components with kriging. *Atmospheric Environment*, 27A, N°6, 831-844.
- Table Ronde des Utilisateurs Isatis, 11 et 12 mars 2002. Présentations techniques, Géovariances
- STEDMAN J. Revised high resolution maps of background levels of air pollutants, report AEA Technologies, 1998, www.aeat.co.uk/netcen/airqual/reports/jsmaps/mphead.html
- TAYANÇ M., 2000. *An assessment of spatial and temporal variation of sulfur dioxide levels over Istanbul, Turkey*. *Environmental pollution*, 107, 61-69.
- TILLÉ Y., 2001. *Théorie des Sondages - Echantillonnage et estimation en populations finies*. Dunod
- TREMBACK C.J, WALKO R.L. 2000, The regional Atmospheric Modeling System (RAMS): development for parallel Processing computer architectures, <http://atmet.com/html/papers/parallel.pdf>
- UNG A., WEBER C., PERRON G., HIRSCH J., KLEINPETER J., WALD L., MANDIN T., *Air pollution mapping over a city-virtual stations and morphological indicators*, 10th Int. Symposium "Transport and Air Pollution", Septembre 2001, Boulder, Colorado
- VAN LOON M., HEEMICK A.W, 1997, kalman filtering for non linear atmospheric chemistry models : first experiences, CWI report
- VARNS J.L., MULIK J.D., SATHER M.E., GLEN G., SMITH L., STALLINGS C., 2001. Passive sampling of ambient, gaseous air pollutants : an assessment from an ecological perspective. *Environmental Pollution*, Vol.107, 31-45.
- WACKERNAGEL H., 1992. Cours de géostatistique multivariable. Ecole des Mines de Paris, 62 p.

WACKERNAGEL H., BERTINO L., SIERRA J.P., GONZÁLEZ DEL RÍO, 2002. *Multivariate kriging for interpolating with data from different sources*. In Quantitative Methods for Current Environmental Issues, Anderson et al. (eds), pp 57-75, Springer-Verlag, Londres

WACKERNAGEL H., 2002, Geostatistical normalization of air pollution transport model output and station data using ISATIS, IMPACT project report n°4 (ref N-20/02/G), <http://cg.ensmp.fr>

WOLKE R., KNOTH O., HELLMUTH O. , SHRODER W., WEICKERT J., 2000, load balancing in the parallel model system LM-MUSCAT for Multiscale chemistry Transport simulations, contribution au projet GLOREAM (EUROTRAC 2)

Liste des annexes

Repère	Désignation précise	Nb/N° pages
1	Présentation de l'étude IPSL/université d'Orsay sur l'adaptation statistique des sortie d'un modèle continental à l'échelle régionale	3
2	Présentation de l'étude de l'Ecole des Mines de Paris sur l'estimation de la qualité de l'air dans les zones peu/pas couvertes.	3
3		
4		
5		
6		

Annexe Technique de la convention CNRS-LMD/INERIS

Cartographie et prévision des champs de pollution à l'échelle locale, à partir des résultats de simulation d'un modèle continental

- Introduction

La prévision et la représentation cartographiée des champs de polluants sont aujourd'hui devenues des priorités pour la surveillance de la qualité de l'air. Le problème est particulièrement aigu sur les zones fortement polluées, autour des grandes agglomérations et pôles industriels, mais la loi impose la surveillance sur l'ensemble du territoire, comprenant aussi de grandes régions peu couvertes par les observations de routine des réseaux.

La surveillance par la mesure est très coûteuse, en équipement et en maintenance. Aussi il s'avère important de développer, pour les régions faiblement couvertes, les techniques mathématiques et numériques permettant une bonne estimation, à la fois en temps réel et pour le passé, des champs de polluants. Dans la dernière décennie plusieurs modèles, statistiques ou physiques (aussi appelés déterministes), ont été développés en France, permettant aujourd'hui de répondre à ces besoins.

Plusieurs associations de surveillance de la qualité de l'air se sont dotées de modèles déterministes à maille relativement fine (quelques km), mais ces modèles sont relativement coûteux en développement et, dans un souci économique, leur exploitation pour des grandes régions faiblement ou modérément polluées peut ne pas apparaître comme une priorité.

Depuis 1998, l'Institut Pierre-Simon Laplace a développé le modèle d'échelle continentale CHIMERE de chimie-transport simplifié (Schmidt et al., 2001), qui a été testé dans un mode « prévision » durant les étés 2000 à 2002 (voir le serveur internet <http://euler.lmd.polytechnique.fr/pioneer>). Ce modèle présente les particularités d'être à la fois déterministe et de couvrir l'Europe de l'Ouest avec des mailles relativement larges (typiquement 30-40 km). Sa mise en œuvre quotidienne pour les besoins nationaux de surveillance de la qualité de l'air est en cours de réalisation à l'INERIS. Si un tel modèle suffisait pour donner une description et une prévision des champs de polluants pour les régions modérément polluées, cela contribuerait à satisfaire, à faible coût additionnel, les besoins mentionnés ci-dessus. Mais la taille de la maille de ce modèle, trop grande pour une représentation locale des champs de polluants, fait qu'un traitement statistique des données, en aval des sorties de ce modèle, est indispensable. Ce traitement a déjà été testé sur l'Ile-de-France par l'université Paris-Sud et l'IPSL, et donne des résultats satisfaisants (Blond et al., 2002).

- Objectifs

Nous proposons ici une étude dont l'objectif est de développer et valider les méthodes statistiques permettant cette descente d'échelle. Nous fixons ainsi deux objectifs:

- Donner une description relativement fine (typiquement quelques km) des champs d'ozone et de dioxyde d'azote par l'assimilation de données, en combinant par un traitement statistique les sorties de modèle et les observations de routine des réseaux.
- Donner une prévision des champs d'ozone et de dioxyde d'azote, avec la même échelle que précédemment.

La contrainte majeure est de n'utiliser qu'un modèle déterministe de grande échelle, le modèle CHIMERE-continental.

Afin de réaliser l'étude sur un cas concret, la région que nous avons sélectionnée couvre une partie de l'Ouest de la France, dont la Vendée, le Poitou, les Pays de la Loire et le Sud-Est de la Bretagne. Pour ces régions, aucun modèle déterministe à maille fine n'a encore été développé. Par ailleurs, la topographie est peu élevée, de sorte que nous pouvons nous affranchir des problèmes de circulation de brises de relief. En revanche la présence des côtes et des brises marines associées demandera un soin particulier afin d'obtenir des résultats d'une qualité suffisante.

- Méthodologie

La méthodologie générale pour réaliser les deux objectifs ci-dessus repose sur deux types de développement. Le premier concerne le modèle de chimie-transport lui-même, et le second concerne le développement de méthodes statistiques.

Tout d'abord, afin de mieux décrire les champs de polluants près de la surface, et notamment pour le dioxyde d'azote, nous devons améliorer la résolution verticale du modèle près du sol en introduisant, par exemple, un volume d'émissions dans la première couche du modèle. Une fraction de cette première couche reçoit donc directement les émissions, permettant de mieux décrire les fortes concentrations au voisinage des sources. Ensuite, le mélange turbulent, assez rudimentairement décrit dans la version actuelle du modèle, doit être amélioré, par exemple par l'introduction de schémas récents en flux de masse décrivant la convection par une approche spectrale des thermiques. Enfin, nous étudierons le réalisme des brises côtières fournies par le modèle de grande échelle du CEPMMT, et tenterons, si celles-ci ne sont pas réalistes, d'en donner, dans le modèle, une paramétrisation simplifiée, afin de mieux simuler les concentrations de polluants près des côtes.

Le deuxième développement concerne les méthodes statistiques permettant la descente d'échelle, à la fois pour la prévision et pour la représentation a posteriori (cartographie, ou encore assimilation de données) des champs de polluants, à une échelle d'environ 5 km.

Pour la *cartographie*, la méthode de base proposée, validée sur l'Ile de France seulement, est le *krigeage sur innovations*. Ces innovations sont la différence entre les sorties du modèle et les observations aux stations du réseau de surveillance. Le modèle CHIMERE-continental fournissant des sorties sur une grille de [0.5x0.25] degrés, une première étape pour la région Ouest sera d'interpoler ces simulations sur une grille 10 fois plus fine en tenant compte des observations des réseaux de surveillance. Cette interpolation ne pourra pas prendre en compte certaines particularités locales comme la présence de villes moyennes et de leurs émissions associées, qui vont, pour l'ozone, affecter les concentrations par titration avec le monoxyde d'azote. Le dépôt sec sur les zones boisées ou cultivées va également introduire des hétérogénéités dans les concentrations de surface. Ces particularités seront prises en compte, en introduisant, dans la méthode de krigeage, des données externes, par exemple issues de la base de données américaine GLCF (disponible sur internet) d'utilisation des sols, en testant par exemple la méthode de *dérive externe* récemment proposée par les travaux de Wackernagel et Lajaunie.

Pour la *prévision* régionalisée des champs de polluants, nous ne disposons pas, bien sur, des concentrations observées en certains points. Nous devons donc nous appuyer sur un traitement statistique des seules sorties de modèle, couramment appelé *adaptation statistique*. Là encore, le modèle statistique d'adaptation reposera sur la nature du terrain à petite échelle et dépendra de l'utilisation du sol. La validation de ces adaptations sera effectuée en utilisant toutes les données d'observations disponibles dans les réseaux.

Le problème des côtes nécessitera un soin particulier puisque les gradients de concentration peuvent y être particulièrement marqués. L'introduction de variables comme la distance à la côte peut par exemple permettre de mieux les prendre en compte.

La validation de la méthode demandera de nombreux tests, notamment sur la partie cotière sur une ou plusieurs années. L'année 2001, riche en événements, sera particulièrement utilisée pour ces tests.

Description de l'étude ARMINES/INERIS

L'extrapolation sur une large base spatiale et temporelle (saisonnaire ou annuelle) des cartographies établies à partir de campagnes de mesure (échantillonneurs passifs, moyens mobiles) soulève plusieurs questions :

- quel est le nombre optimal de campagnes de mesure à programmer ?
- combien de points de mesure sont nécessaires ?
- comment tirer profit des variables externes ?

Afin de prendre la mesure du problème et d'apporter des réponses, l'INERIS a confié une étude au Centre de Géostatistique de l'Ecole des Mines de Paris (interlocuteur : Mme de Fouquet). L'annexe technique de la convention entre l'INERIS et ARMINES est présentée à la page suivante.

Trois jeux de données ont été fournis à l'INERIS par l'ASPA et AIR Languedoc Roussillon à la suite d'une demande faite par l'INERIS auprès du groupe de travail *Moyens Mobiles*. Ils concernent la pollution par le dioxyde d'azote dans les régions de Mulhouse, de la Vallée de la Thur (données de l'ASPA) et de Montpellier (données d'AIR Languedoc Roussillon). Ils sont constitués de données de concentration (tubes à diffusion passive, moyens mobiles, stations fixes) et de données auxiliaires (cadastre des émissions, population, occupation du sol). Les campagnes de mesure de Mulhouse et Montpellier ont été conduites dans le but de cartographier la pollution à l'échelle de l'agglomération. La campagne de mesure de la Vallée de la Thur doit permettre d'évaluer l'impact atmosphérique du trafic automobile de part et d'autre d'un grand axe routier.

Les difficultés propres à chaque région étudiée montrent combien serait délicate l'élaboration d'une méthodologie générale de cartographie. Le travail accompli constitue néanmoins un guide précieux pour l'application des méthodes d'estimation géostatistiques au problème de la qualité de l'air. Compte tenu des données disponibles, la question de la stratégie d'échantillonnage et celle de l'extrapolation temporelle n'ont pu être abordées que d'un point de vue qualitatif. En revanche, l'utilisation de variables auxiliaires a donné lieu à une analyse approfondie.

Les résultats de cette étude seront consignés dans un rapport remis à l'INERIS par le Centre de Géostatistique au début de l'année 2003. Les conclusions de ce travail seront synthétisées par l'INERIS et intégrées dans la version finale (2003) du rapport *Représentation cartographique de la qualité de l'air dans les zones peu ou pas couvertes*.

Annexe technique de la convention INERIS/ARMINES (Centre de Géostatistique de l'Ecole des Mines de Paris)

Estimation de la qualité de l'air sur les zones peu ou pas couvertes par les stations fixes.

1. Problématique

La surveillance de la qualité de l'air repose actuellement sur trois types de mesures :

- les stations fixes, supposées les plus précises. Ces stations ne sont pas disponibles sur l'ensemble du territoire. Elles fournissent des mesures quart-horaires pour plusieurs composés simultanément.

- les tubes passifs. Ils fournissent des mesures indirectes de la concentration, intégrées sur une semaine ou plus fréquemment sur quinze jours. Des campagnes sont effectuées, durant lesquelles un assez grand nombre de tubes sont mis en place et relevés quasi-simultanément.

- les moyens mobiles, qui fournissent des mesures analogues à celles des stations fixes, y compris par la précision.

Une cartographie de la concentration moyenne annuelle ou saisonnière nécessite d'utiliser simultanément les différents ensembles de mesures disponibles, pour améliorer la précision. On s'intéresse ici à la pollution «de fond» en NO_x. Cette pollution dépend notamment des émissions. Pour cette raison, il peut être utile de tirer parti de variables explicatives comme le trafic routier, la densité de population, ainsi que l'altitude, variables que l'on peut supposer connues de façon dense, aux noeuds d'un maillage régulier. Parmi ces variables auxiliaires, le cadastre des émissions présente un intérêt particulier, en vue d'établir une méthodologie ou des recommandations pour l'estimation des concentrations, applicables de façon générale pour différents départements ou zones du territoire.

Toujours en vue d'établir une méthodologie applicable en différents contextes, l'INERIS s'intéresse également à la précision actuelle des estimations, ainsi qu'à la définition de stratégies de mesures appropriées, utilisant de façon conjointe les différents moyens de mesures.

2. Objectif de l'étude

En se basant sur deux sites différents par leur situation géographique, par les moyens de reconnaissance actuelle, ainsi que par leur environnement (topographie, activités industrielles...) on examinera les relations entre variables explicatives et concentration, afin de proposer une modélisation des concentrations à l'échelle saisonnière ou annuelle. Ce modèle servira ensuite pour l'estimation des moyennes saisonnières ou annuelles, ainsi que pour l'évaluation de la précision de ces estimations. On examinera ensuite des modifications de l'échantillonnage, et, dans le cadre du modèle précédent, on évaluera le gain ou la perte de précision associés. Ces études de sensibilité serviront à établir des recommandations pour l'échantillonnage.

Pour vérifier l'utilité et l'intérêt du cadastre des émissions, et examiner quel en est le maillage souhaitable, on comparera les estimations avec et sans prise en compte du cadastre, et pour les deux maillages disponibles (4*4 ou 1*1 km²).

L'utilisation d'un modèle déterministe (modèle Chimère, déduit des modèles météorologiques), ne rentre pas dans le cadre de la présente étude.

Pour qualifier l'impact des périodes de forte concentration, en particulier en terme d'exposition, il est d'usage de cumuler pour une période déterminée l'écart de la concentration à des valeurs de référence, cet écart étant nul lorsque la concentration est inférieure aux seuils fixés. Ces valeurs de référence ont été établies en particulier pour les teneurs en O₃. La présente étude ne contient pas le calcul de la probabilité de dépassement de seuils pour les moyennes annuelles.

3. Contenu de l'étude

Deux ensembles de mesures seront étudiés :

- ASPA : concentration en NO₂ et éventuellement en O₃.

L'ASPA dispose de plusieurs ensembles de mesures de la concentration, en zones urbaines, périurbaines ou rurales. Les trois types de mesures sont représentés, selon les caractéristiques des domaines mesurés. Ces mesures sont complétées par des bases spatialisées (mnt, population, densité de bâti), en particulier l'inventaire des émissions, disponible pour deux maillages (4*4 ou 1*1 km²).

Si les données fournies par l'INERIS le permettent, on intégrera dans la modélisation les mesures de nature météorologique (température, vitesse et orientation du vent).

Ces données seront exploitées de façon détaillée, selon leur disponibilité, pour une période couvrant une à deux années.

- Air Languedoc-Roussillon : concentration en NO₂.

L'inventaire des émissions n'est pas disponible. Les variables explicatives comme le mnt seront fournies si nécessaire par l'INERIS.

En agglomération, les trois types de mesures sont représentés.

Si les données fournies par l'INERIS le permettent, on intégrera dans la modélisation les mesures de nature météorologique (température, vitesse et orientation du vent).

Pour chacun des sites, pour des périodes couvrant une à deux années selon les données disponibles, les étapes seront les suivantes :

- étude exploratoire des données, pour rechercher les relations entre les différents ensembles de mesures, ainsi qu'entre les mesures et les variables explicatives.

Si nécessaire, des campagnes de mesures effectuées pour la comparaisons des stations fixes, des tubes et des moyens mobiles (Aveyron) seront utilisées, afin de préciser ces relations.

- modélisation variographique spatiale ou spatio-temporelle selon les besoins.

- estimation des moyennes saisonnières par krigeage ou cokrigeage, et examen de la précision associée, en particulier en terme de variance de l'erreur d'estimation. Les méthodes de validation croisée pourront être utilisées, en vue de visualiser les erreurs d'estimation.

- pour certaines variables auxiliaires (cadastres des émissions pour les données de l'ASPA, notamment), étude de sensibilité, suivant la prise en compte ou non du cadastre des émissions, et le maillage de ce dernier. Comparaison des estimations et des variances d'estimation associées.

- Si les données le permettent (ASPA) examen des évolutions interannuelles. Estimation des variations des moyennes et précision de cette estimation.

- Etude du schéma d'échantillonnage, et en particulier intérêt des différents moyens de mesure, examinés sur la base des variances d'estimation précédemment obtenues. Elaboration de différents schémas de reconnaissance, utilisant en particulier les moyens mobiles et calcul de la précision associée.

- Comparaison de la modélisation et des résultats relatifs aux deux sites, recommandations en vue de l'extension de la méthode d'estimation à d'autres sites, en particulier pour l'acquisition des données.

- Réflexion sur les questions en suspens, en particulier pour les questions de dépassement de seuil. Il s'agit d'évaluer les difficultés posées par ces questions.

Le rapport sera rédigé en portant attention à son aspect pédagogique : présentation des méthodes, évaluation critique des résultats. Le rapport ne se substitue pas à une présentation complète ou à un cours de géostatistique.

JP Chollet (LEGI)

LISTE DE DIFFUSION

Nom	Adresse/Service	Nb
BIRCK	Dossier maître	1
DOC		1
ROUÏL		1
MALHERBE		1
MATE		5
ADEME		2
EMD		1
LNE		1

TOTAL **13**

PERSONNES AYANT PARTICIPE A L'ETUDE

Travail	Nom	Qualité	Date	Visa

Fin du Complément non destiné au client