

RAPPORT D'ETUDES

ETUDE TECHNIQUE
**EVALUATION DES OUTILS DE CARTOGRAPHIE
DE LA POLLUTION PAR L'OZONE DANS LES
BOUCHES-DU-RHONE, LE VAR, ET LE VAUCLUSE**



Surveillance de la qualité de l'air de l'Est des Bouches-du-Rhône, du Var et du Vaucluse

67-69, avenue du Prado ; 13 286 Marseille Cedex 6

Tel : 04 91 32 38 00 – Fax : 04 91 32 38 29 – Internet : www.airmaraix.com – Serveur téléphonique : 04 91 326 327

Date de publication : Juin 2004

Avec la participation de :



Intitulé de la demande :

AIRMARAIX développe depuis 2000 des techniques de cartographie, notamment pour la représentation des épisodes de pollution photochimique. Depuis 2003, les outils de cartographie permettent une certaine automatisation et la représentation de situations de pollution sous la forme de cartes peut être pratiquement envisagée en temps réel. L'objet de cette étude technique est de valider la méthodologie d'élaboration des cartographies d'épisodes de pollution photochimique en vue de systématiser notamment leur diffusion durant l'été 2004.

Traitement de la demande :

Une sélection d'épisodes de pollution représentatifs a été finement analysée à l'aide d'outils géostatistiques. Cette analyse s'appuie sur l'expertise et la validation par un organisme spécialisé dans le domaine de la géostatistique (GEOVARIANCES).

Rapport édité en Juin 2004.

Pour AIRMARAIX :

Laetitia Mary, Sylvain Fayet.

Pour GEOVARIANCES :

Nicolas Jeannée

SOMMAIRE

Introduction.....	3
1. Historique.....	4
2. Généralités	5
3. Cartographie des niveaux maximum d'ozone.....	6
3.1. Choix du meilleur modèle de variogramme	7
3.2. Ajustement par un modèle sphérique à grande portée + un effet de pépité :	9
3.3. Choix d'une portée pour le modèle sphérique et d'un effet de pépité :.....	13
3.4. Conclusion pour la cartographie des maximum d'ozone journaliers :.....	16
4. Cartographie des niveaux d'ozone en moyenne horaire glissante	17
4.1. Choix du meilleur modèle de variogramme :	18
4.2. Ajustement par un modèle sphérique :	24
4.3. Choix de la meilleure portée pour le modèle sphérique :.....	31
4.4. Conclusion pour la cartographie des moyennes horaires glissantes en ozone :.....	41

INTRODUCTION

Face à la demande accrue du public d'avoir une information visuelle de la pollution et face aux besoins des réseaux de mieux maîtriser les comportements des polluants et l'espace à surveiller, la cartographie devient une activité incontournable pour la surveillance de la qualité de l'air.

Au sein du réseau Airmaraix, des premiers travaux de cartographie ont déjà été réalisés en 2000 pour visualiser la pollution photochimique et tenter de mieux la comprendre. Les outils et l'expérience du réseau en la matière ne permettaient pas alors de maîtriser les algorithmes de cartographie à la base des études mais ont cependant apporté des éclaircissements sur certains phénomènes.

L'arrivée depuis de nouveaux outils de géostatistiques tels qu'ISATIS (Géovariances) ou GEOSTATISTICAL ANALYST (ESRI) nous aide à présent à valider ces premières études mais surtout à en réaliser de nouvelles plus performantes et plus robustes sur lesquelles pourront être basées de nombreuses réflexions.

Un audit réalisé par *Géovariances* et *Géovalor* en fin d'année 2003 sur ces travaux de cartographie nous a permis de mettre en évidence qu'une étude plus approfondie était nécessaire pour les finaliser.

Ce rapport présente donc, après un bref historique de ce qui a déjà été réalisé en terme de cartographie au sein du réseau, une étude détaillée des cartographies de l'ozone dans le but de réaliser des cartes de façon semi-automatique pour mettre rapidement à la disposition du public la visualisation des phénomènes photochimiques dans les Bouches-du-Rhône, le Var et le Vaucluse.

Ce travail a fait l'objet d'une relecture et d'une validation de la part de la société *Géovariances*.

1. HISTORIQUE

Depuis trois ans Airmaraix s'intéresse à l'interpolation spatiale des niveaux d'ozone.

Le but est d'interpoler les maximums d'ozone journaliers par station de mesures du réseau, mais aussi les moyennes horaires glissantes par station afin de représenter sur toute une journée l'évolution dans le temps et dans l'espace d'un épisode de pollution photochimique (animation). Ainsi, une information visuelle peut être fournie, apportant au grand public une aide à la compréhension du phénomène global (c'est-à-dire la visualisation du déplacement des masses d'air pollué en ozone en fonction de la direction du vent ainsi que la visualisation des zones les plus touchées par une pointe de pollution).

Diverses méthodes ont été testées en 2000 avec le logiciel Surfer : polygonale, moyenne mobile, inverse des distances au carré, krigeage ordinaire linéaire (paramètres choisis par défaut)... Le choix de la meilleure méthode à utiliser s'est basée sur des critères subjectifs visuels (meilleure continuité de l'interpolation, interpolation aux limites...) et sur une bibliographie sommaire : le krigeage linéaire est apparu être le plus performant.

La méthodologie utilisée n'était pas assez bien maîtrisée pour justifier de façon statistique la qualité de l'interpolation. C'est l'expertise du réseau en terme de photochimie et particulièrement les campagnes temporaires qui ont permis de valider l'interpolation réalisée en apportant des informations sur les mesures éloignées des stations fixes. Les premiers résultats ont alors été utilisés comme tels mais en tenant compte de tous ces artefacts.

Il fallait rendre de plus en plus robustes les réflexions dans ce domaine : l'évolution du réseau (plus grande couverture du territoire), l'arrivée de nouvelles compétences et de nouveaux outils (Isatis, ArcGIS Geostatistical Analyst...) sont allées dans ce sens (plus de facilité d'utilisation, moins de problèmes aux limites...).

Les résultats obtenus et la fonctionnalité de ces nouveaux outils nous ont permis d'envisager l'automatisation de certaines tâches pour la création rapide des cartographies d'ozone.

L'utilisation des Systèmes d'Information Géographiques (SIG avec Arcview d'ESRI) bien ancrée dans la structure a favorisé le choix de l'outil Geostatistical Analyst (module d'Arcview) pour la réalisation de cette automatisation.

Il fallait que l'utilisation de l'outil soit simple pour un utilisateur lambda, que le lien soit immédiat avec le SIG, que le travail soit facilement et rapidement reproductible sur n'importe quel domaine d'étude.

Un système de récupération automatique des données de qualité de l'air du réseau et de cartographie de ces données par krigeage a donc été mis en place sous Spatial Analyst (autre module d'Arcview - pas d'automatisation possible avec Geostatistical Analyst, plus puissant mais non accessible en langage de programmation).

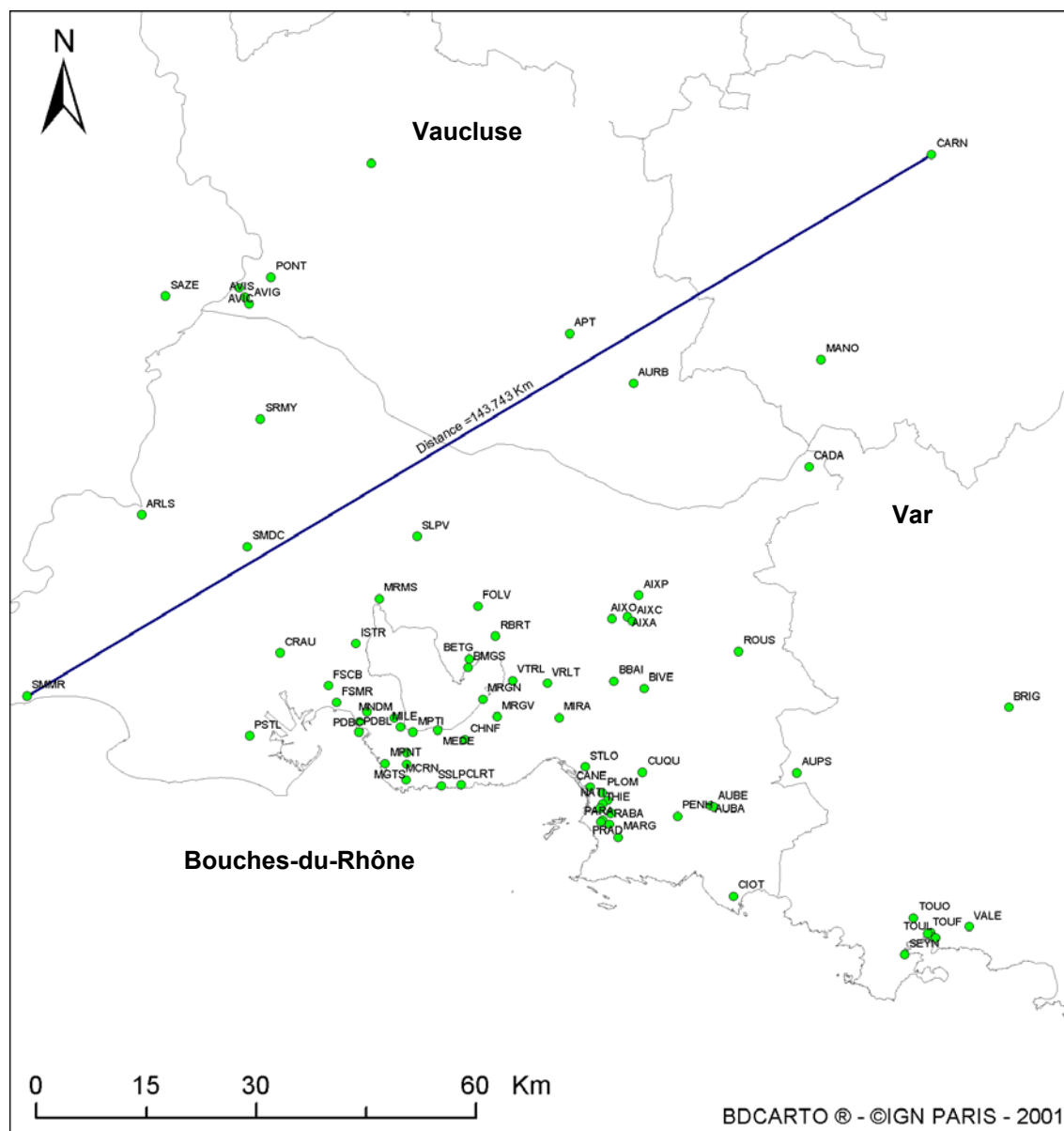
Pour finaliser ce travail, il est à présent nécessaire :

- de bien déterminer le ou les modèles de variogrammes adéquats pour représenter spatialement au plus juste les niveaux d'ozone maximum sur des journées polluées et les niveaux d'ozone aux différentes heures de la journée, pour des jours pollués ou non pour la période de mai à septembre.
- De trouver ensuite le modèle de variogramme le plus « universel » possible, dans chacun de ces deux cas, qui soit automatisable par Spatial Analyst.
- Enfin, d'évaluer l'erreur éventuelle entraînée par le choix de ce modèle.

Ce travail est réalisé avec le logiciel **Isatis** qui offre un plus grand éventail d'outils géostatistiques, automatisables, et qui est surtout mieux adapté que Geostatistical ou Spatial Analyst pour ce type d'étude prospective : il permet d'étudier en détails les variogrammes expérimentaux et ainsi de mieux déterminer les modèles de variogrammes à privilégier.

2. GENERALITES

Notre domaine d'étude comprend toutes les stations des Bouches-du-Rhône, du Var et du Vaucluse. Les deux points les plus distants de notre domaine d'étude sont situés à environ 140 km l'un de l'autre.



Pour l'étude sur les maximums journaliers comme pour celle sur les moyennes horaires glissantes, nous choisirons donc les paramètres du krigeage suivants :

- un **lag = 10000 m**
- un **nombre de lag = 7**

De façon générale, le nombre de lag * le lag doit être égal au maximum à la moitié de la dimension du domaine d'étude : en effet, le nombre de paires de points diminue sérieusement à partir de la moitié de la taille du champ et le calcul réalisé à partir de cette distance sur-représente les informations en bordure du champ ; dans notre cas, le nombre de points est suffisant pour permettre un calcul relativement robuste du variogramme jusqu'à cette distance.

L'échelle de couleur choisie pour les cartographies finales qui seront diffusées s'étale de 0 à $360\mu\text{g}/\text{m}^3$ avec un pas de $20\mu\text{g}/\text{m}^3$ afin de représenter au mieux les niveaux les plus élevés qui sont les plus intéressants pour le public.

Dans notre étude, l'échelle s'étend de valeurs $< 10\mu\text{g}/\text{m}^3$ à des valeurs $> 200\mu\text{g}/\text{m}^3$ avec un pas de $10\mu\text{g}/\text{m}^3$ dans le but d'avoir un plus grand degré de finesse dans notre analyse.

3. CARTOGRAPHIE DES NIVEAUX MAXIMUM D'OZONE

Pour les maximum d'ozone, nous nous attarderons sur les journées de pic de pollution par l'ozone qui seront les journées que nous choisirons de cartographier, les autres n'ayant pas d'intérêt particulier.

Plusieurs journées sont donc choisies de façon à étudier différentes situations types d'épisodes de pollution :

Descriptif des journées choisies pour l'étude :

04/08/03	Pic d'ozone généralisé sur les Bouches-du-Rhône, le Var et le Vaucluse
29/05/03	Pic d'ozone sur les zones Nord-Est et Sud-Est des BdR et le Var
10/03/03	Pic élevé durant 2h sur la zone NE des BdR
27/06/03	Pic sur la zone SE des BdR
09/06/03	Pic sur la zone Ouest des BdR
19/09/03	Pic sur la zone Ouest des BdR et le Vaucluse
16/09/03	Pic sur le Var
21/07/03	Pic sur le Vaucluse

Pic : dépassement du niveau de recommandations

Le but est d'observer sur toutes ces journées les variogrammes expérimentaux caractéristiques des maximum journaliers et de déterminer le meilleur modèle de variogramme que l'on peut leur associer.

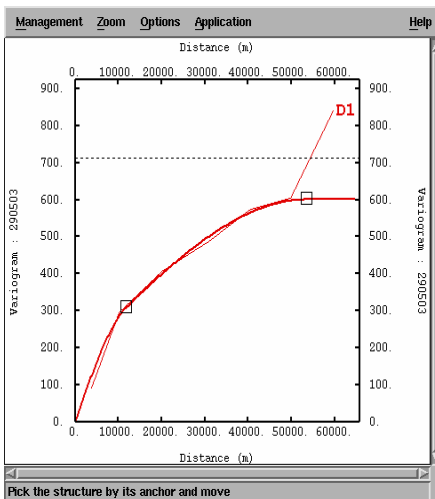
3.1. CHOIX DU MEILLEUR MODELE DE VARIOGRAMME

Après étude de ces différentes journées, il est apparu que deux types de modèles, qui ajustent le mieux le variogramme expérimental, se détachent :

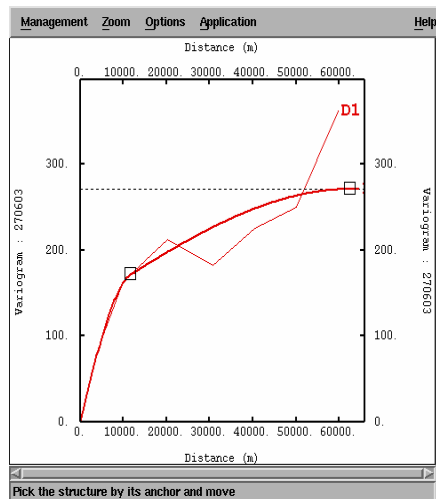
- un modèle double sphérique sans effet de pépite avec une portée à 10 km et une autre portée entre 50 et 60 km
- la combinaison d'un effet de pépite fixé à un certain pourcentage de la variance et d'un modèle linéaire

Remarque : la plupart des variogrammes expérimentaux sont créés à partir de toutes les mesures. Parfois un à trois points considérés comme des « outliers » sont supprimés pour faire le meilleur choix de variogramme mais ils sont réintégré par la suite dans les interpolations.

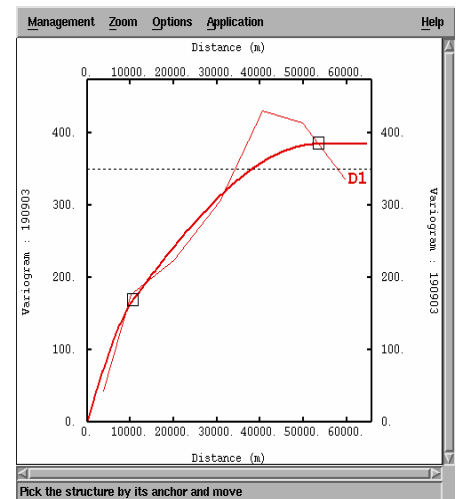
Exemples de modèle double sphérique pour le 29/05/03, le 27/06/03, et le 19/09/03 :



29/05/03



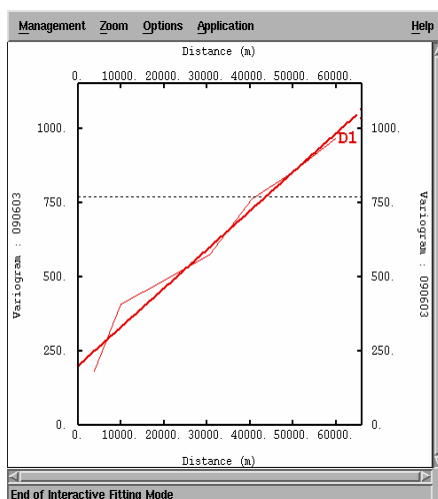
27/06/03



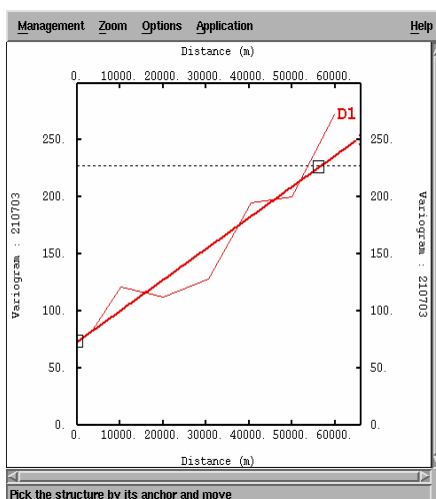
19/09/03

La petite portée correspond à la première cassure qui caractérise la variabilité des données dans les 10 premiers kilomètres. La deuxième portée correspond à la deuxième échelle de variabilité des données entre 10 et 60 km.

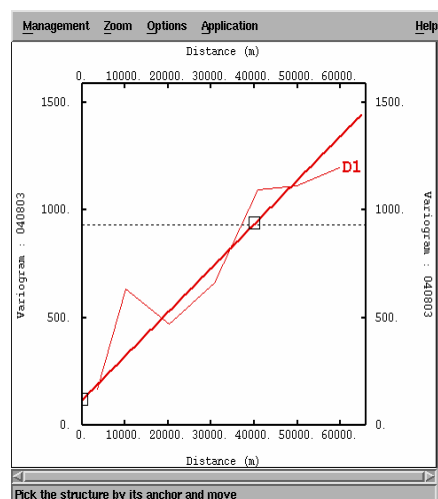
Exemples de modèle alliant effet de pépite + modèle linéaire pour le 09/06/03, le 21/07/03 et le 04/08/03 :



09/06/03



21/07/03



04/08/03

Dans ces cas le variogramme expérimental ne converge pas vers un seuil mais s'étend au-delà de 60 km, ce qui justifie le choix d'un modèle linéaire.

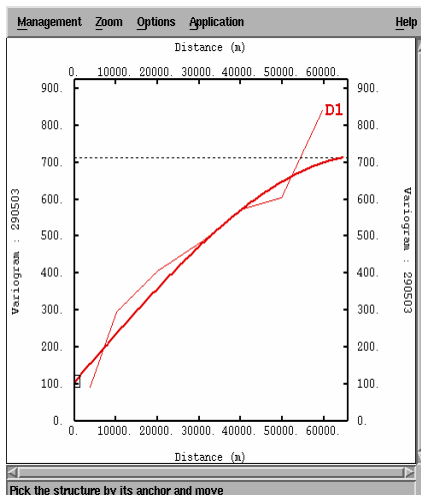
L'effet de pépite peut être expliqué par le fait que les maximums journaliers ne se produisent pas aux mêmes heures de la journée pour l'ensemble des stations, il se peut donc que des stations assez proches n'aient pas des valeurs maximales tout à fait semblables. Cependant, cet effet de pépite doit être minime si l'on veut que les valeurs aux points de mesures soient bien représentées : un écart d'environ $15 \mu\text{g}/\text{m}^3$ entre la cartographie et la réalité est acceptable. On se retrouve dans ce cas en fixant un effet de pépite entre 10 et 20% de la variance qui reste inférieure à 1000.

Avec une échelle à un pas de $10 \mu\text{g}/\text{m}^3$, la différence ne sera pas visible sur la cartographie ; encore moins avec un pas de $20 \mu\text{g}/\text{m}^3$.

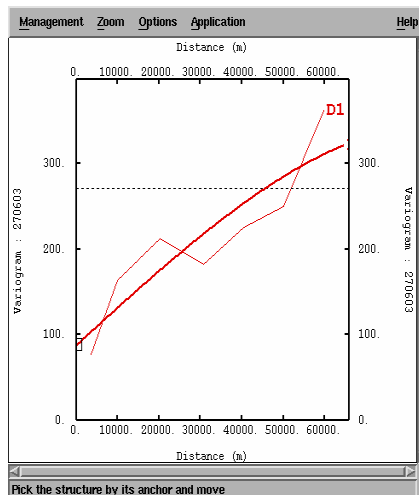
Pendant, compte tenu du choix (contraintes techniques) d'utiliser Spatial Analyst pour automatiser la réalisation de cartographies, il n'est pas possible d'utiliser un modèle double sphérique, ni un modèle linéaire. En revanche, il est tout à fait possible de combiner un effet de pépite à un modèle sphérique de grande portée qui se rapproche alors d'un modèle linéaire.

3.2. AJUSTEMENT PAR UN MODELE SPHERIQUE A GRANDE PORTEE + UN EFFET DE PEPITE :

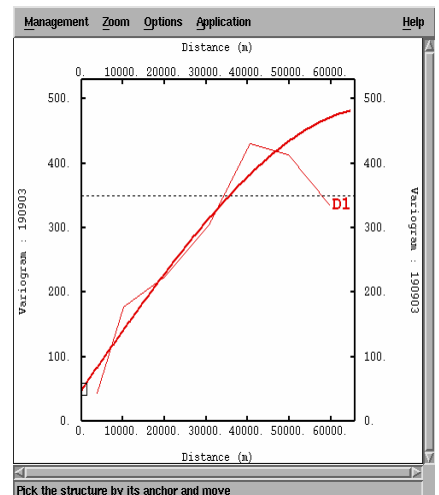
Exemples sur les journées testées précédemment :



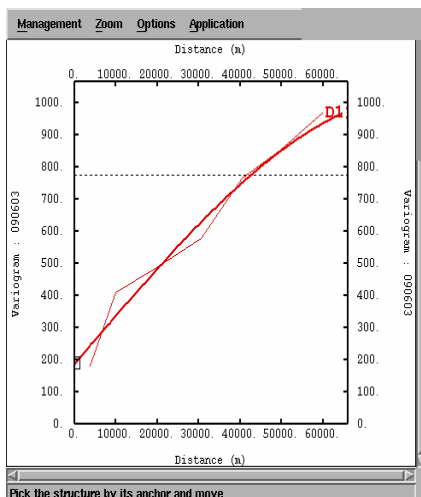
29/05/03



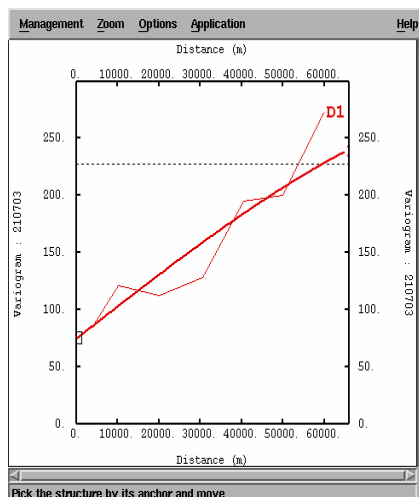
27/06/03



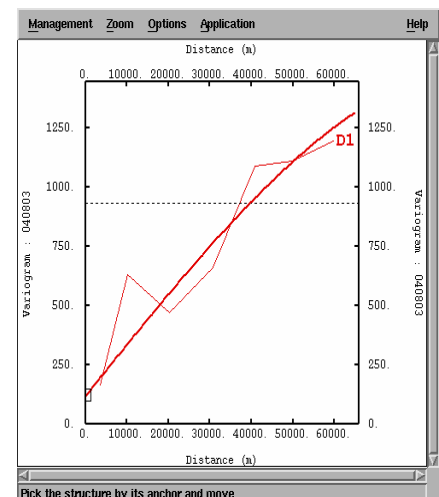
19/09/03



09/06/03



21/07/03



04/08/03

Dans ces cas on se rapproche tout à fait d'une combinaison d'un effet de pépite et d'un modèle linéaire.

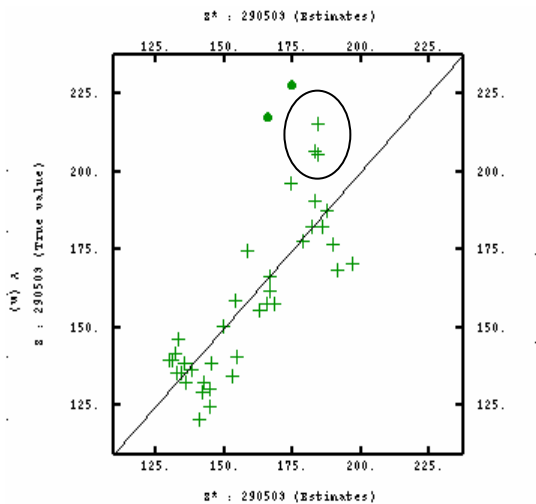
L'effet de pépite peut être situé entre 10 et 20% de la variance.

La portée du modèle sphérique s'étend au-delà de 60km, et se situe entre 80 et 100 km dans la plupart des cas.

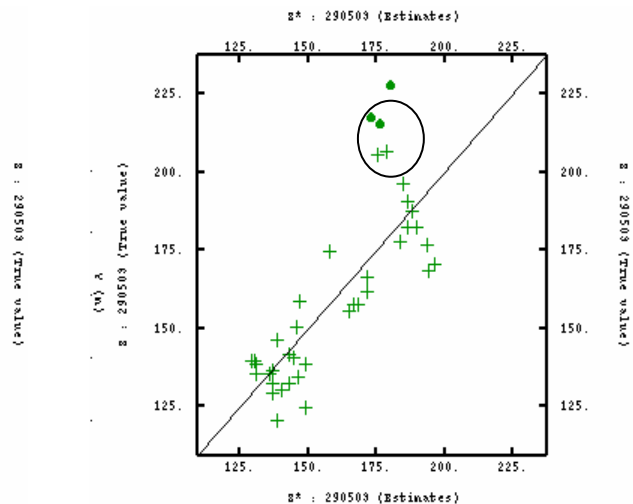
On peut s'attarder particulièrement sur le 29/05/03, le 27/06/03 et le 19/09/03 où l'utilisation de cette combinaison de modèles semble être malgré tout moins intéressante que l'utilisation d'un modèle double sphérique.

L'analyse des validations croisées et cartes finales obtenues nous permettra d'évaluer si la différence engendrée par l'utilisation des deux modèles est importante.

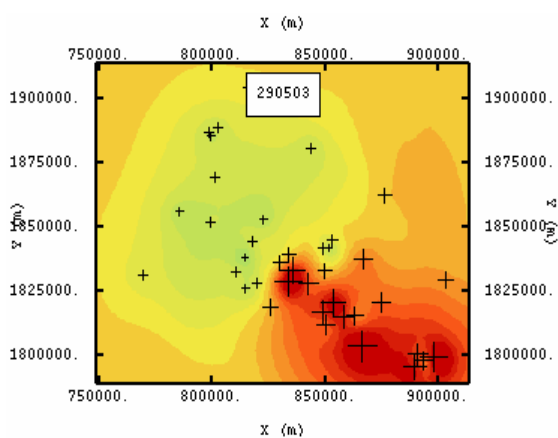
Pour le 29/05/03 :



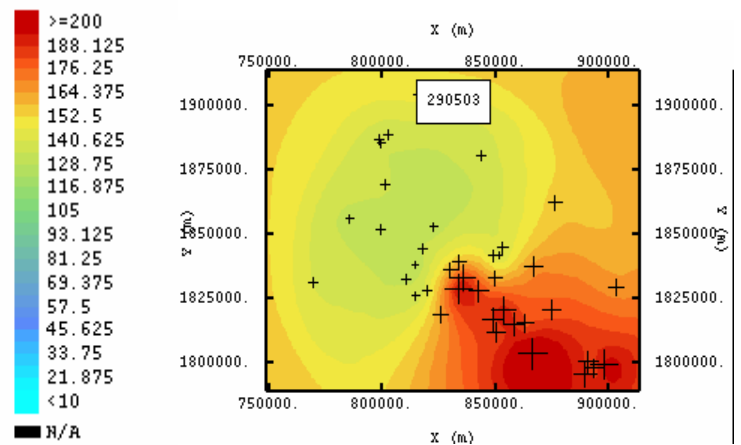
Validation croisée avec modèle double sphérique



Validation croisée avec modèle sphérique + effet de pépite



Cartographie par krigeage avec modèle double sphérique



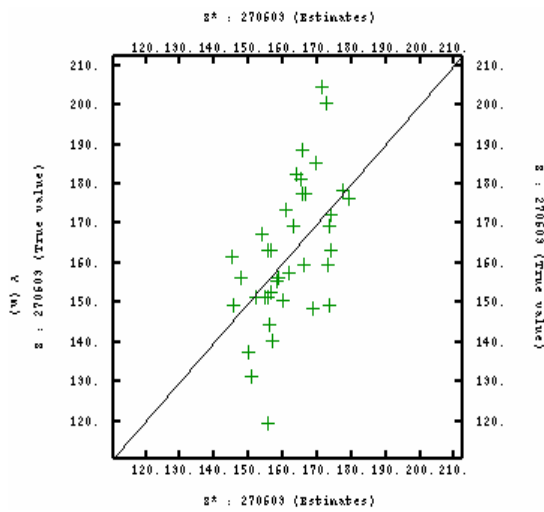
Cartographie par krigeage avec modèle sphérique + effet de pépite

Les deux résultats sont assez similaires autant en terme de validation croisée que de cartographie.

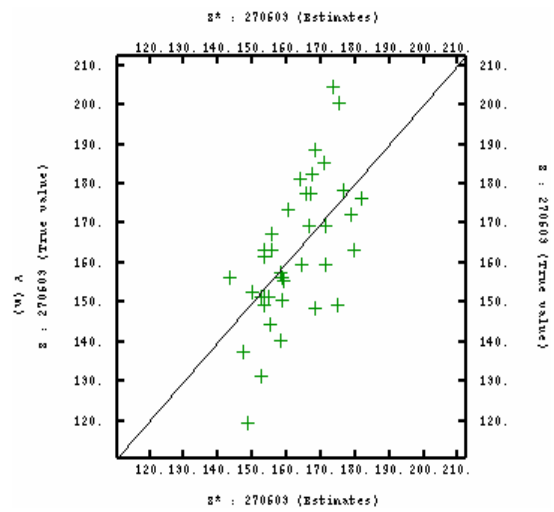
On constate juste une sous-estimation plus importante des valeurs fortes avec le modèle comprenant un effet de pépite (points entourés sur les validations croisées), d'environ 10 à 15 $\mu\text{g}/\text{m}^3$. Cette différence n'est pas très visible sur les cartes avec l'échelle que nous avons choisie pour notre étude (pas de $10\mu\text{g}/\text{m}^3$) et le sera encore moins avec l'échelle définie pour les cartographies qui seront diffusées (pas de $20\mu\text{g}/\text{m}^3$).

En terme d'interprétation du phénomène, on ne perd pas l'information principale indiquant le pic d'ozone et son étendue.

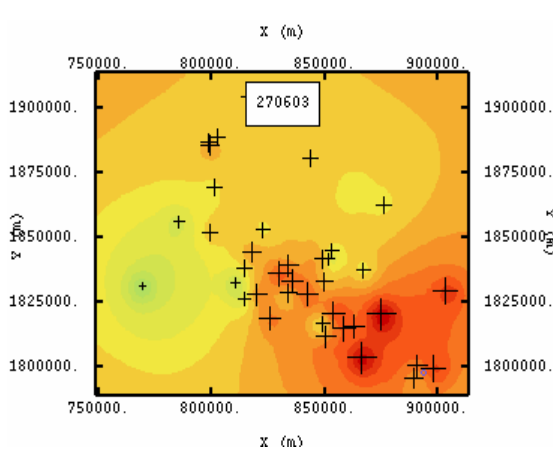
Pour le 27/06/03 :



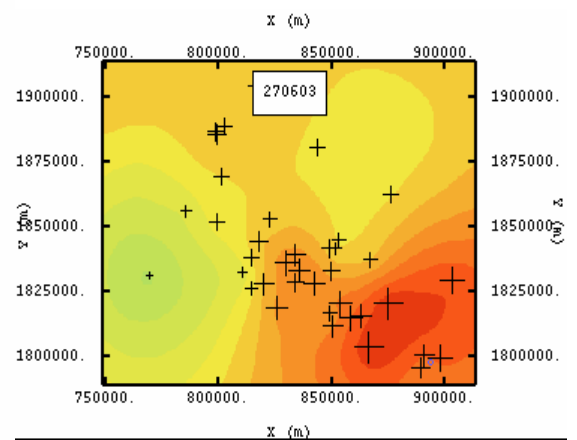
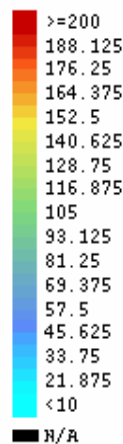
Validation croisée avec modèle double sphérique



Validation croisée avec modèle sphérique + effet de pépité



Cartographie par krigeage avec modèle double sphérique



Cartographie par krigeage avec modèle sphérique + effet de pépité

Sur les cartographies, on note des tâches plus rondes autour des points avec le modèle double sphérique : en effet, ce modèle suit mieux le variogramme expérimental sur les petites distances et respecte ainsi mieux les valeurs mesurées aux points.

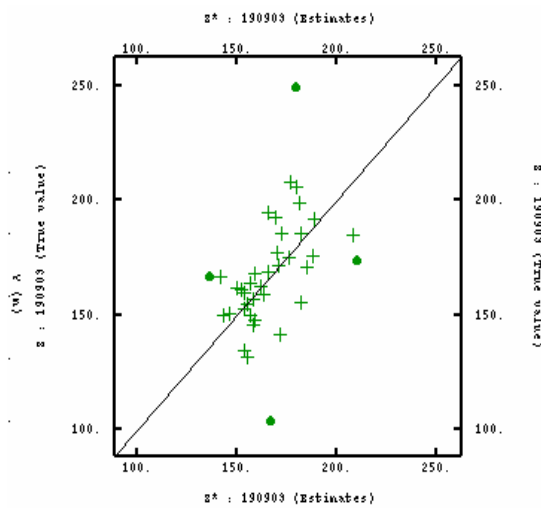
Cependant, en terme de masses d'air polluée (phénomène global), l'information est retranscrite de manière assez semblable.

Ceci est confirmé avec la validation croisée pour laquelle on ne distingue pas ou très peu de différence.

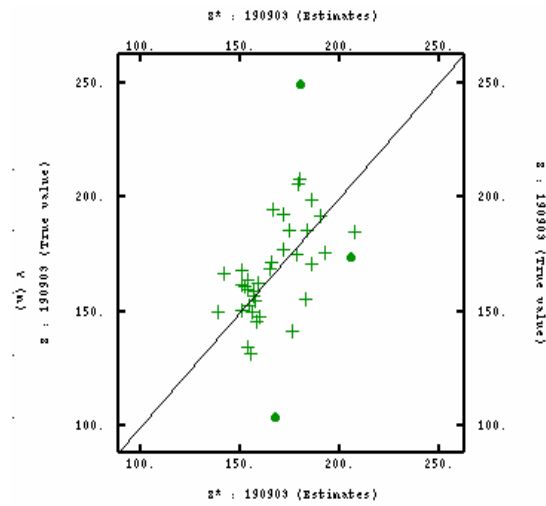
De plus, il faut noter que le variogramme ne dit pas tout car nous avons malgré tout peu de points de mesures, la plupart groupés, et un échantillonnage assez hétérogène ; ce qui favorise une variabilité plus grande à petite distance. Si nous avons plus de points aux extrémités, cette variabilité serait vraisemblablement atténuée. On peut par conséquent lui donner un poids moins fort et se retrouver alors avec une cartographie plus linéaire sur les petites distances.

Au regard de ce point de vue, la cartographie réalisée avec le modèle sphérique associé à un effet de pépité est bien valable.

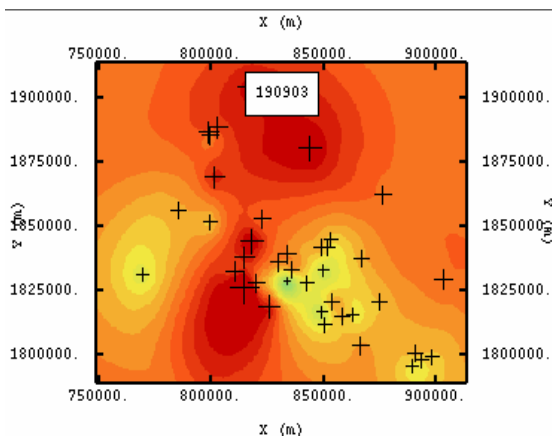
Pour le 19/09/03 :



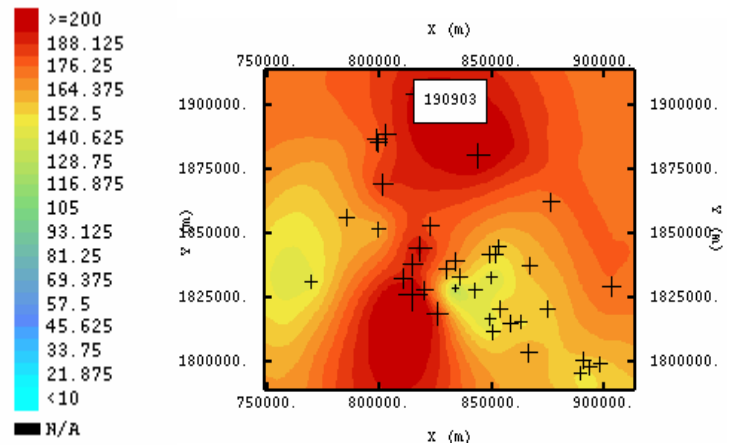
Validation croisée avec modèle double sphérique



Validation croisée avec modèle sphérique + effet de pépité



Cartographie par krigeage avec modèle double sphérique



Cartographie par krigeage avec modèle sphérique + effet de pépité

Dans ce cas les cartographies et la validation croisée sont très similaires.

Dans tous nos cas d'étude, le modèle double sphérique ou le modèle linéaire associé à un effet de pépite sont les modèles les mieux adaptés.

Un modèle sphérique avec une grande portée associé à un effet de pépite reste assez proche des variogrammes expérimentaux dans la plupart des cas et fournit des cartographies et validations croisées similaires. Dans les autres cas où le modèle s'éloigne plus du variogramme, on ne perd pas en qualité au niveau de l'information sur le phénomène global de photochimie que l'on souhaite représenter.

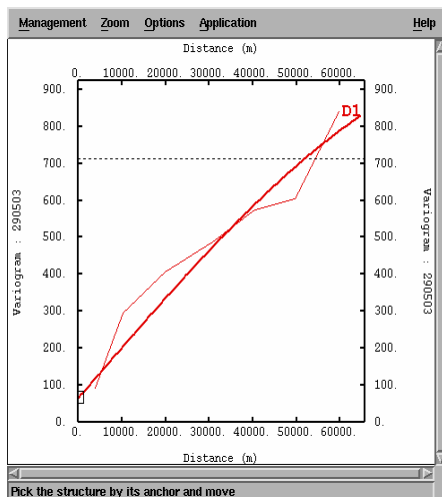
Ces cas étant représentatifs de plusieurs situations types d'épisodes de pollution, on peut donc généraliser l'utilisation d'un modèle de variogramme sphérique avec une grande portée supérieure à 60 km et un effet de pépite entre 10 et 20% de la variance pour cartographier les maximum d'ozone journaliers lors de pointe de pollution.

Il reste à déterminer où fixer cette portée et cet effet de pépite dans notre souci d'automatisation.

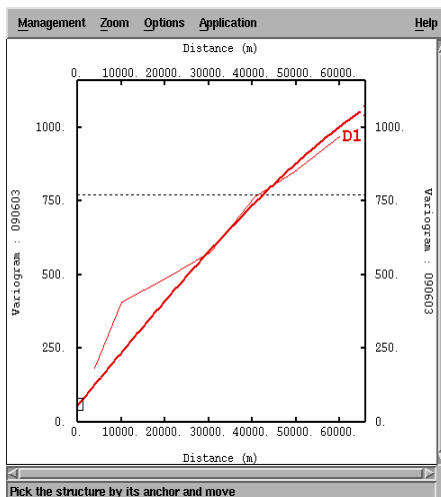
3.3. CHOIX D'UNE PORTEE POUR LE MODELE SPHERIQUE ET D'UN EFFET DE PEPITE :

Au regard des éléments ci-dessus, on pourrait fixer l'**effet de pépite à 10% de la variance et la portée à 90 km** : cela semblerait être un bon compromis (cf. p.7).

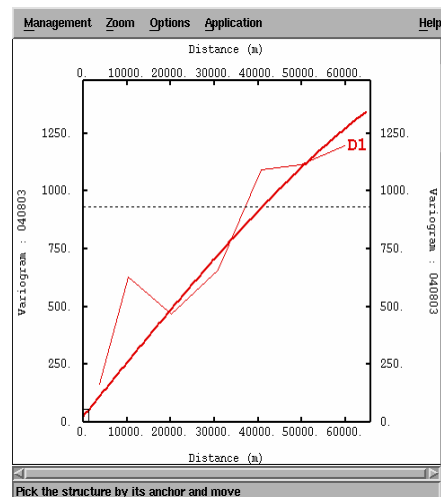
Pour les 29/05, 09/06 et 04/08, si l'on fixe ces paramètres, on obtient des variogrammes semblables à ceux déjà validés :



29/05/03



09/06/03

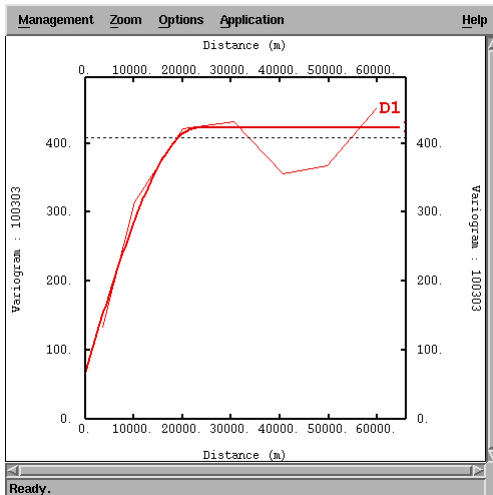


04/08/03

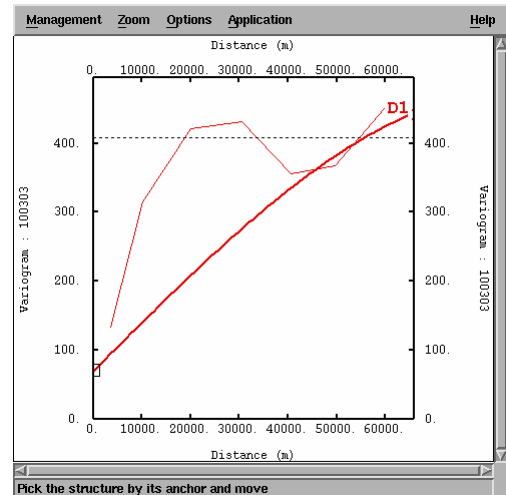
Cas particulier :

Le 10/03/03 est un cas particulier : le pic observé a été très localisé et très court par rapport aux autres journées où les pics sont généralement de grande ampleur et durent plusieurs heures...

Pour cette journée, il semble plus judicieux d'associer une combinaison d'un effet de pépité à environ 10% de la variance et d'un modèle sphérique à petite portée (autour de 20 km), qu'un modèle sphérique à grande portée (90 km) associé au même effet de pépité (cf. variogrammes ci-dessous).

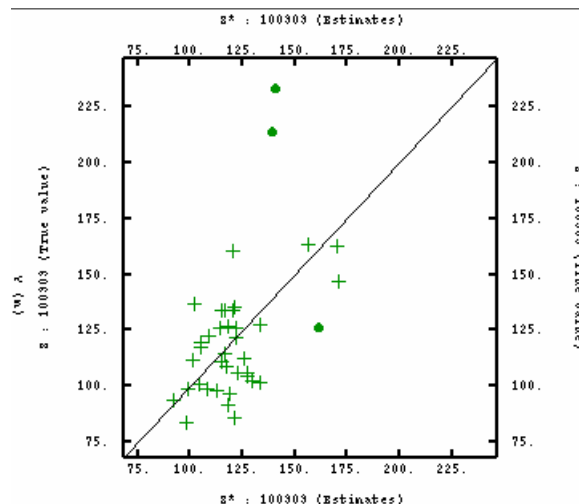


Modèle sphérique – portée : 20 km +
effet de pépité à 10% de la variance

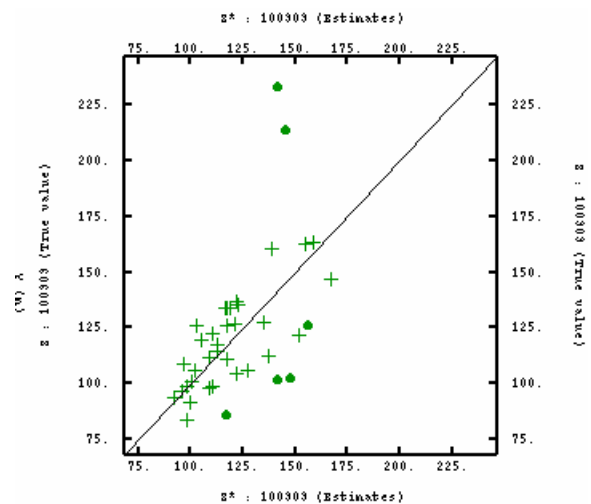


Modèle sphérique – portée : 90 km +
effet de pépité à 10% de la variance

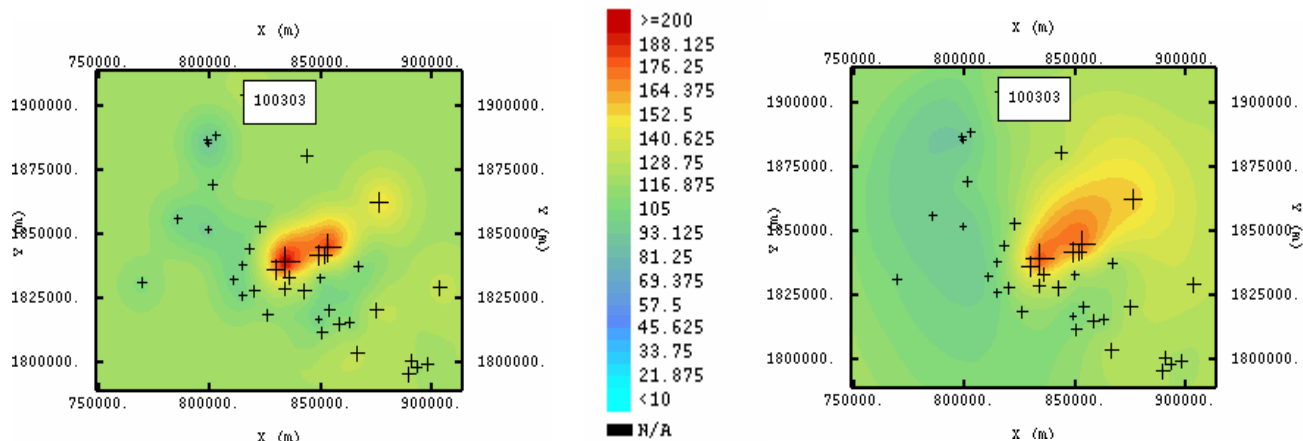
La validation croisée et les cartographies correspondantes ci-dessous, apportent des compléments d'information.



Validation croisée pour modèle
sphérique à petite portée



Validation croisée pour modèle
sphérique à grande portée



Cartographie pour modèle sphérique à petite portée

Cartographie pour modèle sphérique à grande portée

On note qu'avec un modèle sphérique à grande portée, on a plus « d'outliers », la cartographie est plus linéaire et les valeurs aux points sont moins proches de la réalité (les 2 valeurs les plus élevées sont à $232 \mu\text{g}/\text{m}^3$ et $213 \mu\text{g}/\text{m}^3$, ce qui est mieux représenté sur la 1^{ère} carte que sur la 2^{ème}).

Cependant, la différence des cartes n'est pas aussi flagrante que celle des modèles de variogrammes. De plus, même si le pic de pollution de cette journée était très localisé, on a pu observer un transfert des masses d'air pollué de l'Etang de Berre jusqu'à Cadarache. La 2^{ème} carte représente mieux cette logique de déplacement de la masse d'air.

Là encore, on retombe sur le fait que le variogramme ne montre pas tout. Il est perturbé par la variabilité à petite distance due au manque de points aux extrémités et accentué ici par le côté très local du phénomène alors que généralement, la pointe de pollution est plus homogène. Il semble donc plus pertinent de donner moins de poids à cette variabilité aux petites distances comme le fait le modèle de variogramme à grande portée. Ce qui explique que la deuxième carte représente mieux la continuité du phénomène.

Il semble finalement que l'on puisse aussi utiliser un modèle sphérique à grande portée dans ce cas.

Quoiqu'il en soit, ce type de journée méritera d'être étudié à chaque fois plus en profondeur : la généralité obtenue sur le type de modèle à choisir pour cartographier les maximum d'ozone journaliers qui peut être bien utilisée pour les cas habituels d'épisodes de pollution généralisés peut ne pas s'adapter à ces cas particuliers, ou du moins mérite d'être observée de plus près....

3.4. CONCLUSION POUR LA CARTOGRAPHIE DES MAXIMUM D'OZONE JOURNALIERS :

Cette étude de différents cas représentatifs de différents types de journées de l'été 2003 permet de tirer des conclusions générales sur la façon de cartographier les maximums d'ozone journaliers par krigeage.

On a pu observer que le meilleur type de modèle à adapter pour cette cartographie était :

- soit un modèle double sphérique avec une portée à 10 km et une autre entre 50 et 60 km
- soit la combinaison d'un effet de pépite à 10% de la variance et d'un modèle linéaire

Ces types de modèles n'étant pas utilisables avec Spatial Analyst que nous avons choisi d'utiliser au préalable pour réaliser ces cartographies automatiquement, notre étude a permis de montrer **qu'une combinaison d'un effet de pépite à 10% de la variance et d'un modèle sphérique avec une grande portée fixée à 90 km** pouvait convenir dans l'ensemble des journées traitées, avec un impact acceptable sur le résultat final au vu de l'objectif fixé.

Pour certains cas particuliers comme la journée du 10/03/03, la cartographie devrait être étudiée spécifiquement : aucun algorithme par défaut ne peut gérer ce genre d'épisode.

Quoiqu'il en soit, toute automatisation de cartographie devrait être validée avant d'être communiquée pour être en mesure d'écarter ce type d'épisode très particulier.

4. CARTOGRAPHIE DES NIVEAUX D'OZONE EN MOYENNE

HORAIRE GLISSANTE

Nous avons observé ici le comportement variographique de 10 journées entre les mois de mai et septembre 2003 représentant chacune une typologie particulière (journées sans pointe d'ozone, journées avec pic généralisé sur la zone d'étude, journées avec pic d'ozone sur une seule zone) pour 6 heures différentes de la journée (6h, 8h, 12h, 14h, 16h, et 20h TU) de façon à réaliser le nombre de tests le plus représentatif possible d'un été et tenter de pouvoir en extraire des conclusions générales pour réaliser des cartographies automatiques quotidiennes.

Descriptif des journées choisies pour l'étude :

04/08/03	Pic d'ozone généralisé sur les Bouches-du-Rhône, le Var et le Vaucluse
29/05/03	Pic d'ozone sur les zones Nord-Est et Sud-Est des BdR et le Var
10/03/03	Pic élevé durant 2h sur la zone NE des BdR
27/06/03	Pic sur la zone SE des BdR
09/06/03	Pic sur la zone Ouest des BdR
19/09/03	Pic sur la zone Ouest des BdR et le Vaucluse
16/09/03	Pic sur le Var
21/07/03	Pic sur le Vaucluse
01/05/03	Pas de pointes d'ozone
25/09/03	Pas de pointes d'ozone

Pic : dépassement du niveau de recommandations

Le but est d'observer sur toutes ces journées les variogrammes expérimentaux caractéristiques de chacune des heures et de déterminer le meilleur modèle de variogramme que l'on peut leur associer.

4.1. CHOIX DU MEILLEUR MODELE DE VARIOGRAMME :

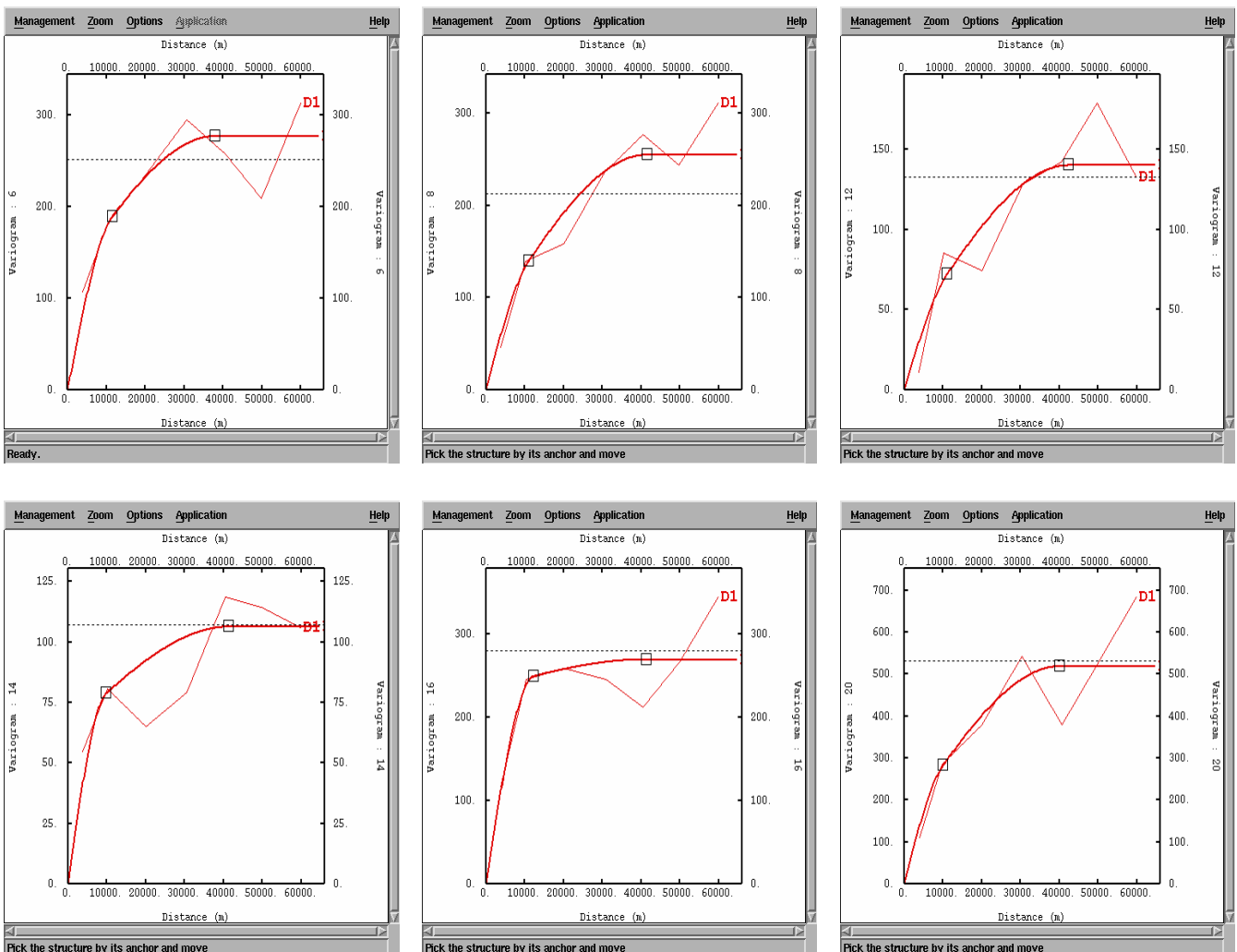
Après étude des différentes heures de toutes ces journées, il est apparu que dans le plus grand nombre de cas, le modèle le mieux adapté aux variogrammes expérimentaux observés est **un modèle double sphérique sans effet de pépite** avec :

- une portée à 10 km, représentant la variabilité des mesures dans les 10 premiers km
- et une portée autour de 40 km, représentant la variabilité des mesures entre 10 et 40 km.

Les exemples ci-dessous illustrent l'application du modèle double sphérique aux épisodes étudiés.

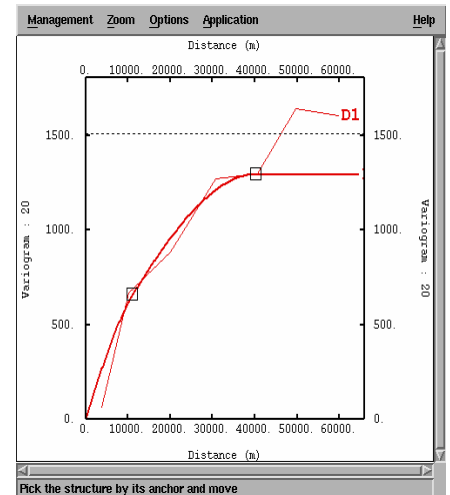
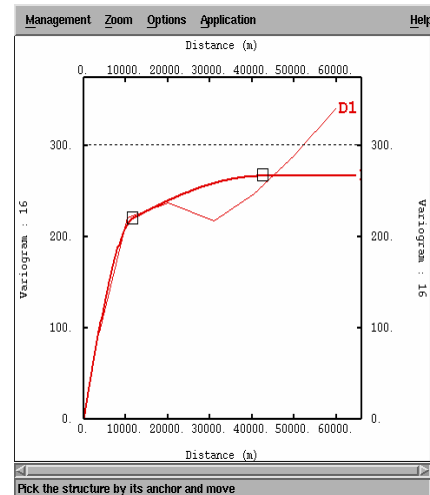
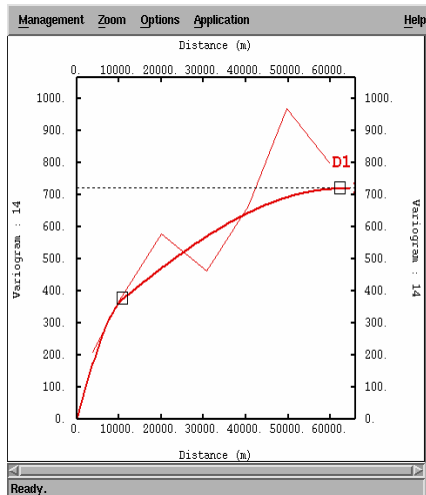
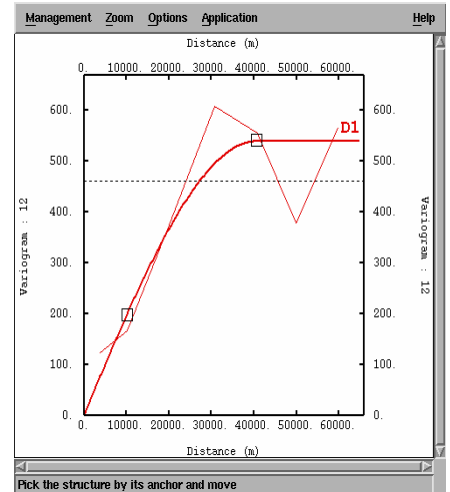
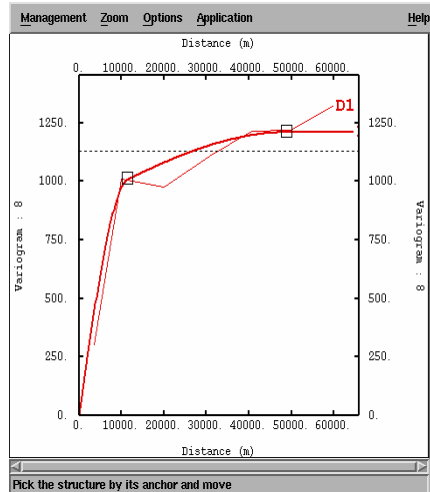
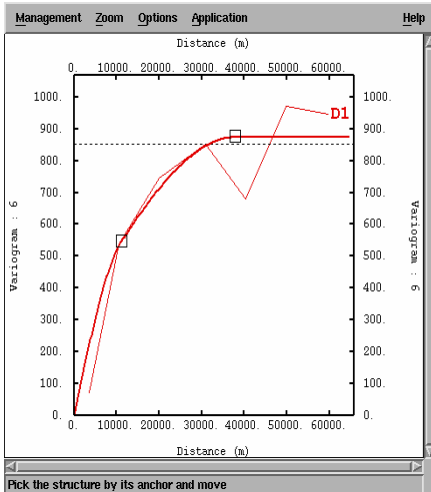
Remarque : la plupart des variogrammes expérimentaux sont créés à partir de toutes les mesures. Parfois un à trois points considérés comme des « outliers » sont supprimés.

01/05/03 : pas de pointe ozone



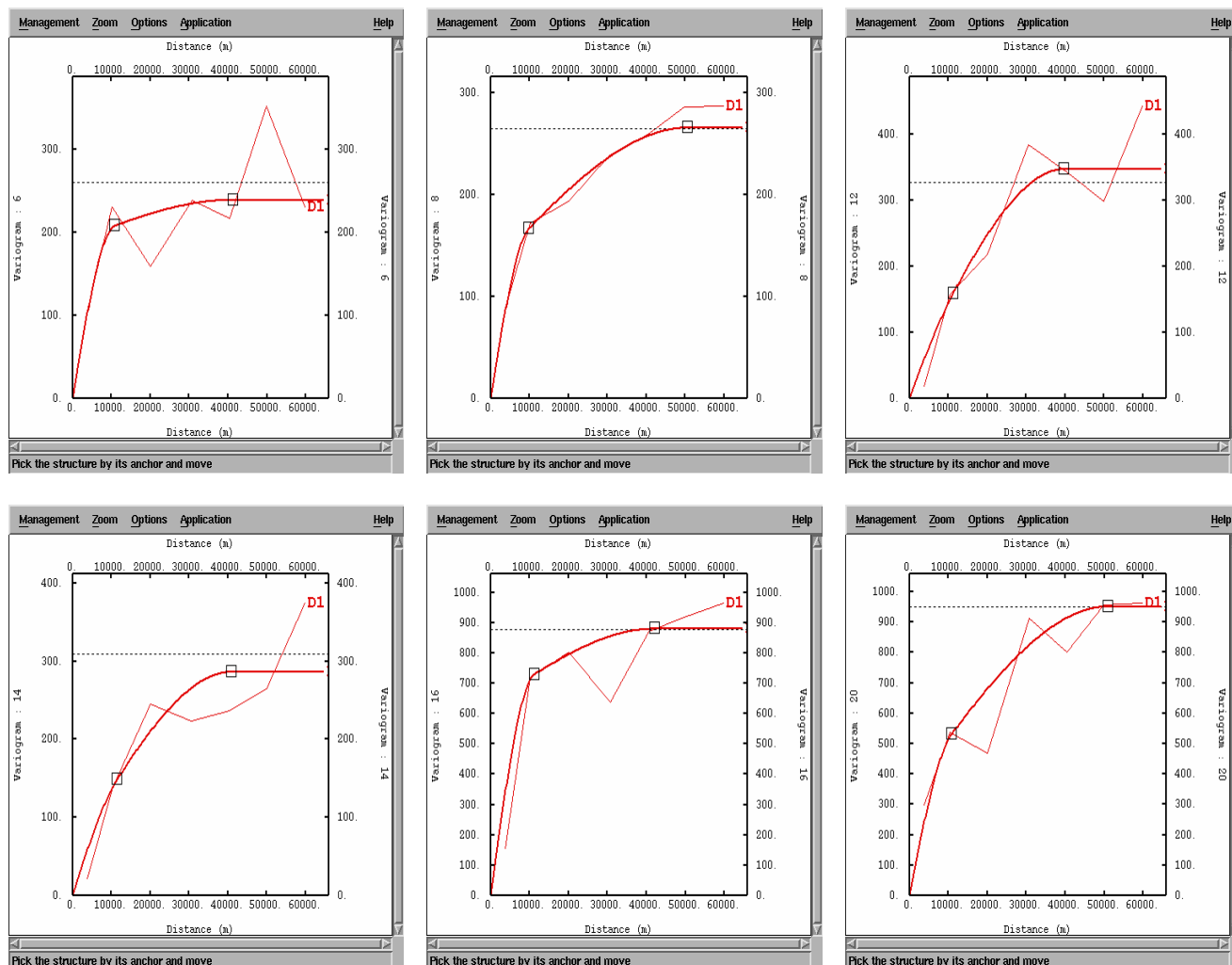
Variogrammes expérimentaux et modèles de variogramme doubles sphériques associés, à différentes heures de la journée

04/08/03 : pic d'ozone généralisé sur la zone d'étude



Variogrammes expérimentaux et modèles de variogramme doubles sphériques associés, à différentes heures de la journée

27/06/03 : pic sur la zone SE des BdR



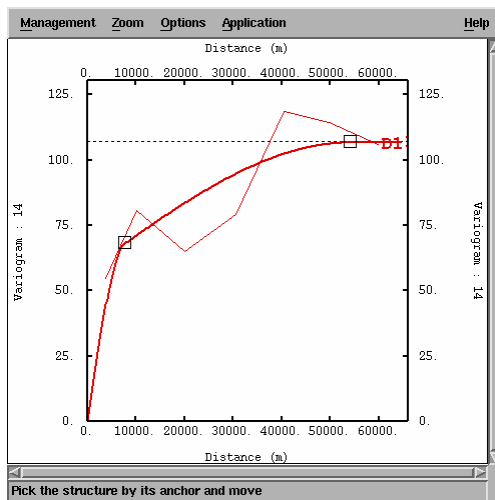
Variogrammes expérimentaux et modèles de variogramme doubles sphériques associés, à différentes heures de la journée

On note dans la plupart des cas que les variogrammes **tendent vers un effet de pépité nul** : ce qui en terme d'ozone paraît plutôt cohérent car pour des points séparés d'une distance nulle, la valeur en ozone est sensé être la même au même moment. On constate que pour des points proches, les valeurs d'ozone sont similaires.

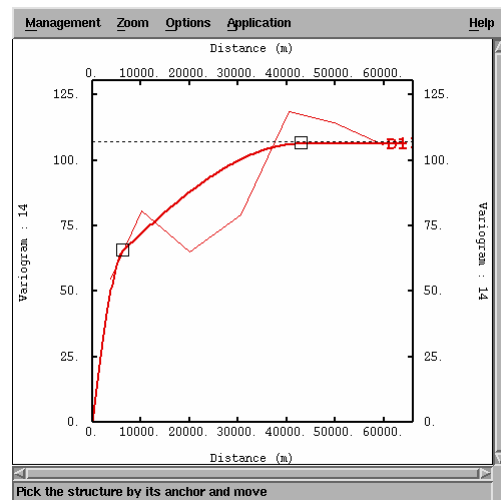
Pour l'ensemble des journées testées (avec ou sans pic d'ozone), on voit très nettement la première cassure du variogramme à 10 km. Elle correspond à la première portée du modèle double sphérique, et signifie une plus grande variabilité des mesures dans les 10 premiers kilomètres. Ensuite, on note clairement aussi, soit une deuxième cassure autour de 40 km, soit une deuxième portée qui peut se situer autour de 40 km pour que le modèle de variogramme s'ajuste au mieux au variogramme expérimental. Cela correspond à une seconde échelle de variabilité des mesures entre 10 et 40 km.

Dans l'ensemble des cas testés, on peut fixer la deuxième portée autour de 40 km mais parfois autour de 50 km, et parfois même à 60 km. Mais dans tous ces derniers cas on peut facilement et sans grande modification du modèle de variogramme fixer cette deuxième portée à 40 km, cas le plus courant, si l'on veut en déterminer une que l'on puisse utiliser dans le cas d'une automatisation.

Cas du 01/05/03 à 14h TU :



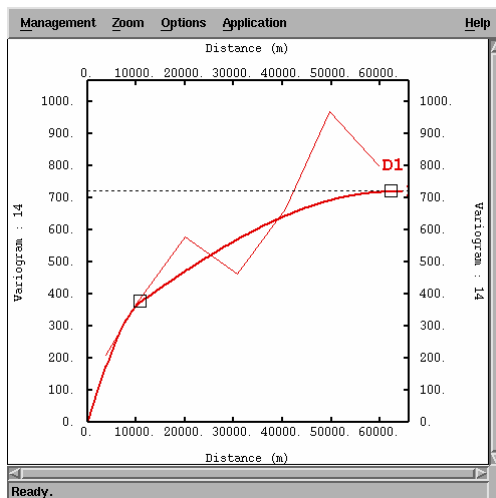
Deuxième portée autour de 50 km



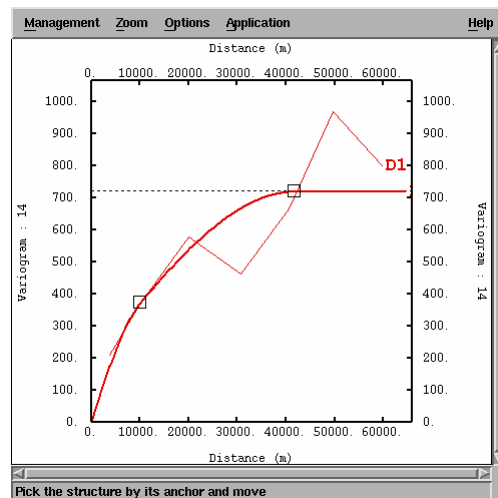
Deuxième portée autour de 40 km

Avec une deuxième portée à 50 km, le variogramme semble optimisé. En fixant cette 2^{ème} portée à 40 km, le modèle de variogramme est très légèrement modifié ; ce qui ne devrait pas entraîner de différences sur la cartographie finale.

Cas du 04/08/03 à 14h et à 20h TU :

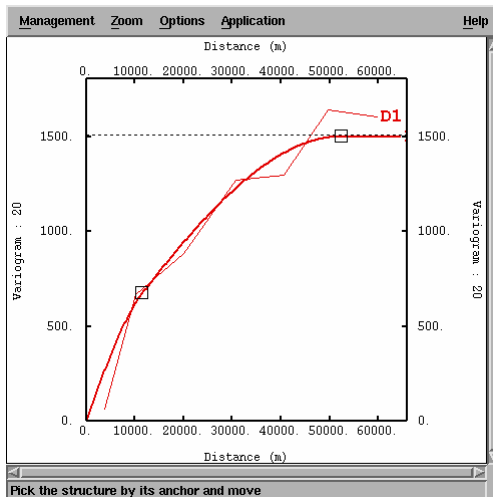


Deuxième portée autour de 60 km

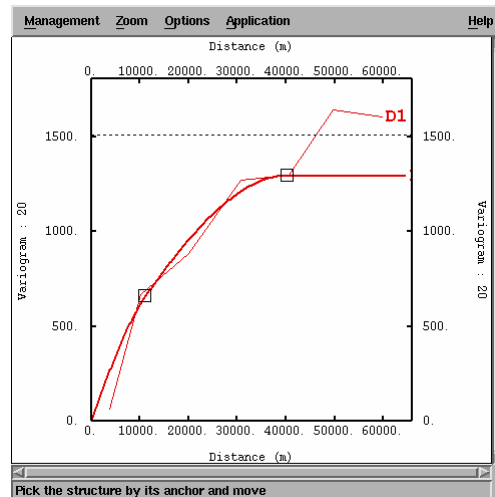


Deuxième portée autour de 40 km

Le modèle de variogramme est ici légèrement modifié mais garde une forme assez similaire. La cartographie finale est peu modifiée (cf. cartographies en exemple ci-dessous).



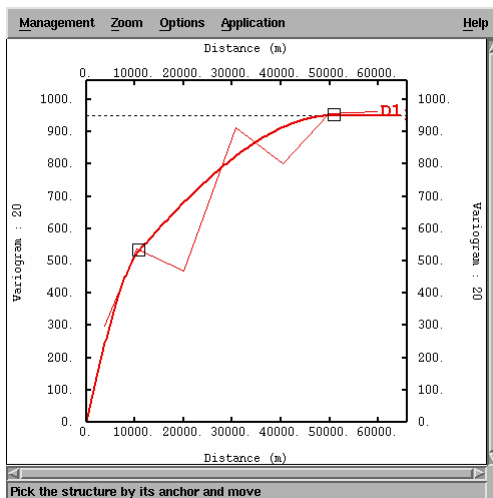
Deuxième portée autour de 50 km



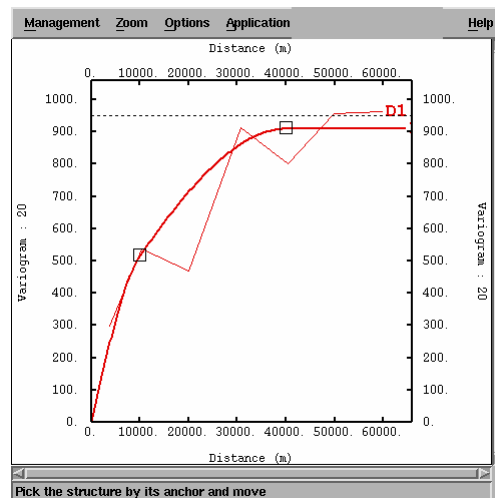
Deuxième portée autour de 40 km

Après modification de la deuxième portée, le modèle est presque identique dans les trois premiers quarts les plus importants du variogramme expérimental.

Cas du 27/06/03 à 20h TU :



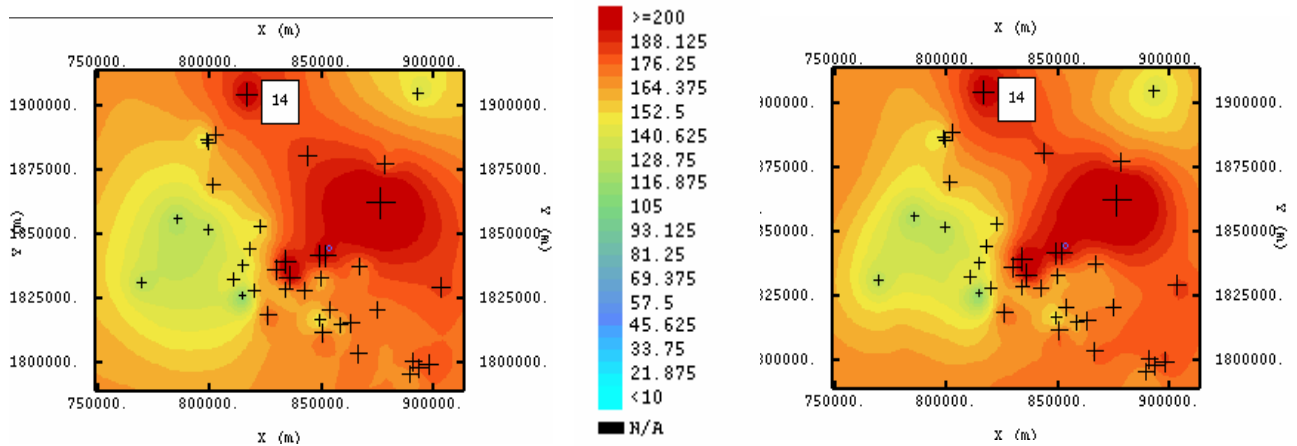
Deuxième portée autour de 50 km



Deuxième portée autour de 40 km

Même conclusion là encore.

Exemples de cartographies le 04/08 à 14h illustrant le fait que ces modifications de portée n'engendrent pas de différences sensibles au niveau des cartes finales :



Cartographie avec un modèle double sphérique – 2^{ème} portée à 60 km

Cartographie avec un modèle double sphérique – 2^{ème} portée à 40 km

Le meilleur modèle de variogramme pour représenter les niveaux horaires glissants d’ozone est donc un modèle double sphérique avec une portée à 10 km et une autre portée autour de 40 km.

Ce modèle semble être le mieux adapté pour représenter les différentes variabilités des mesures d’ozone sur notre zone d’étude.

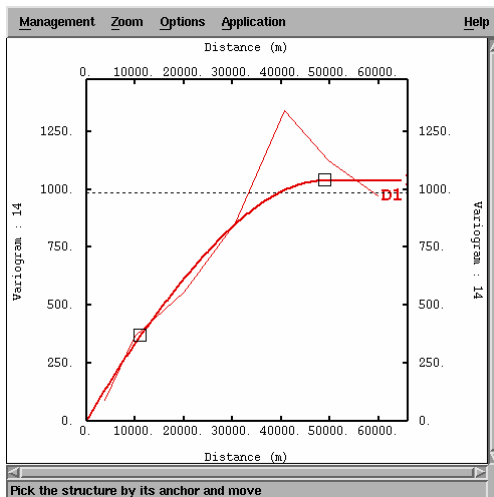
Cependant, le logiciel Spatial Analyst, avec lequel nous avons choisi de travailler initialement, n’offre pas la possibilité d’utiliser ce genre de modèle. La combinaison de deux modèles sphériques ne peut pas être effectuée : seul un modèle sphérique avec une seule portée peut être utilisé.

Si l’on observe bien l’ensemble des cas testés pour notre étude, on note que dans tous les cas, un modèle sphérique peut aussi être ajusté.

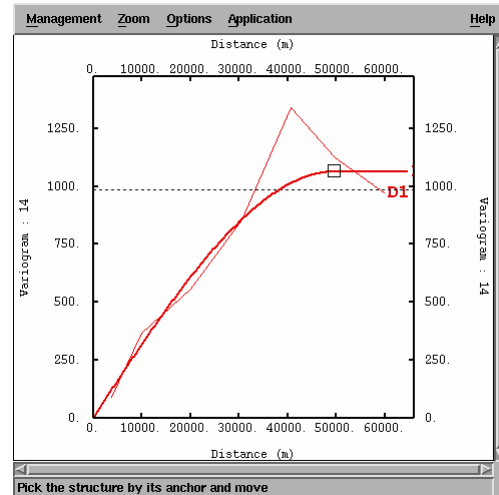
4.2. AJUSTEMENT PAR UN MODELE SPHERIQUE :

Dans de nombreux cas, le modèle sphérique semble pouvoir être aussi bien adapté qu'un modèle double sphérique. Dans ces cas, les deux modèles sont assez similaires pour choisir aussi bien l'un que l'autre (cf. exemples suivants) sans engendrer de modifications sur le résultat final.

Cas du 09/06/03 à 14h et à 16h TU : pic sur la zone Ouest des BdR

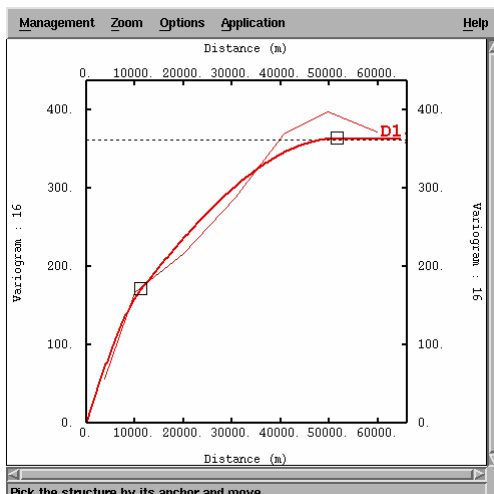


Modèle double sphérique

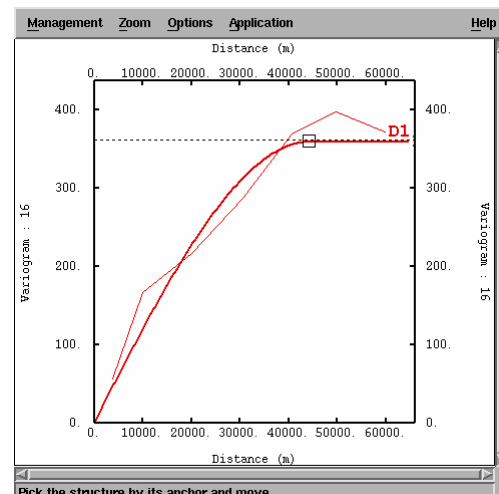


Modèle sphérique

Les deux modèles sont ici très similaires : le modèle sphérique convient tout aussi bien que le double sphérique.

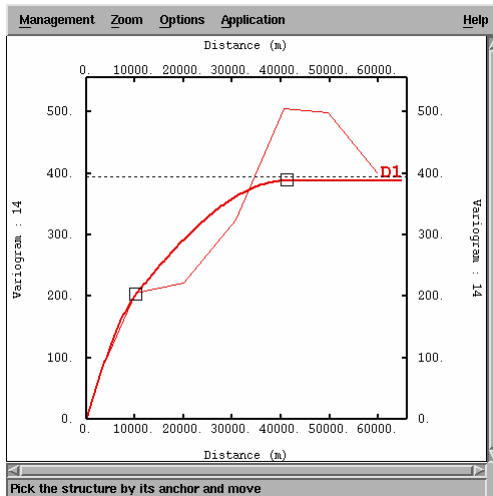


Modèle double sphérique

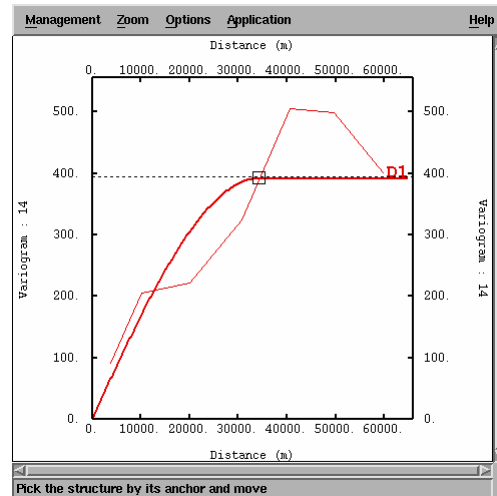


Modèle sphérique

Cas du 19/09/03 à 14h TU : pic sur la zone Ouest des BdR et sur le Vaucluse



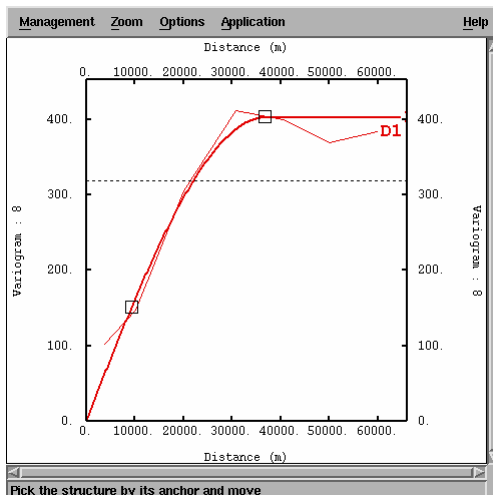
Modèle double sphérique



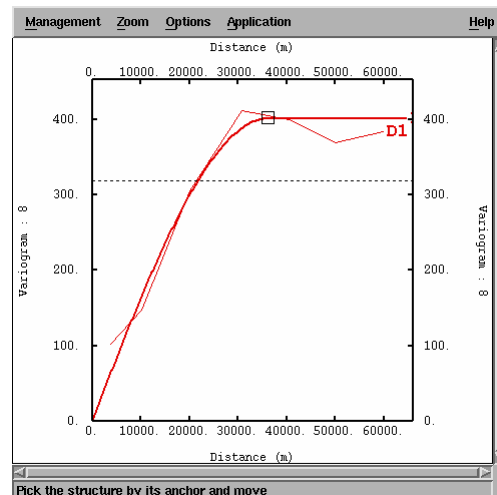
Modèle sphérique

Dans les deux derniers cas, le modèle double sphérique est plus optimisé dans les 10 premiers km, ensuite les modèles sont très similaires. La différence est légère : au final les résultats ne devraient pas différer.

Cas du 10/03/03 à 8h TU : pic sur la zone Nord-Est des BdR



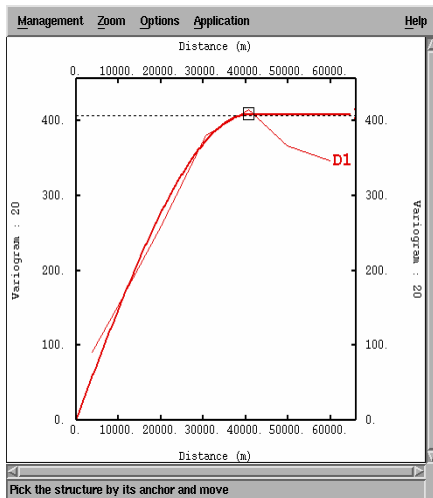
Modèle double sphérique



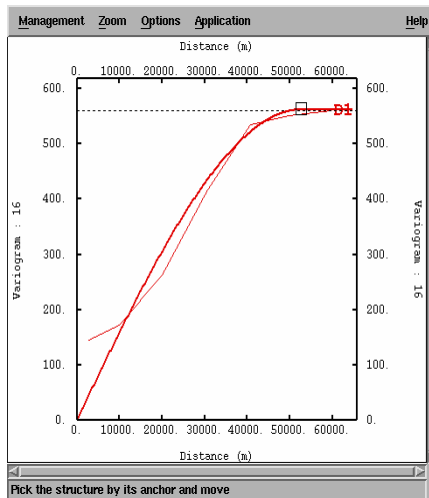
Modèle sphérique

Les deux modèles sont ici presque semblables.

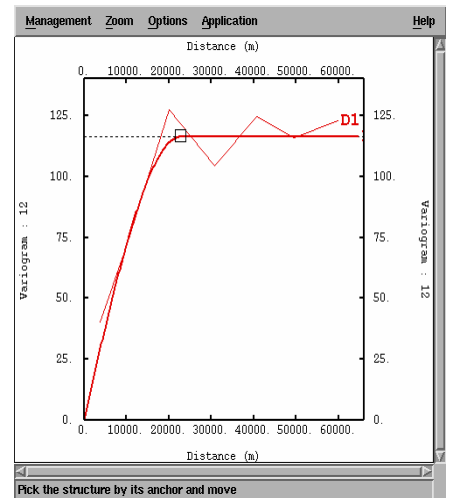
Dans certains cas, le modèle sphérique convient aussi bien qu'un modèle double sphérique qui ne serait pas plus intéressant : on choisirait deux portées pour finalement s'ajuster sur un modèle sphérique à une seule portée (cf. les 3 cas suivants).



Cas du 09/06/03 à 20h TU



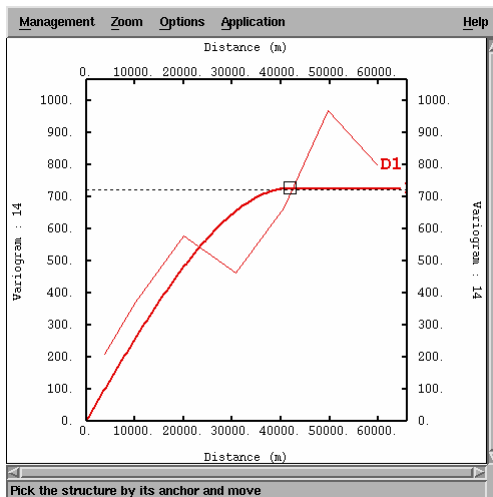
Cas du 16/09/03 à 16h TU



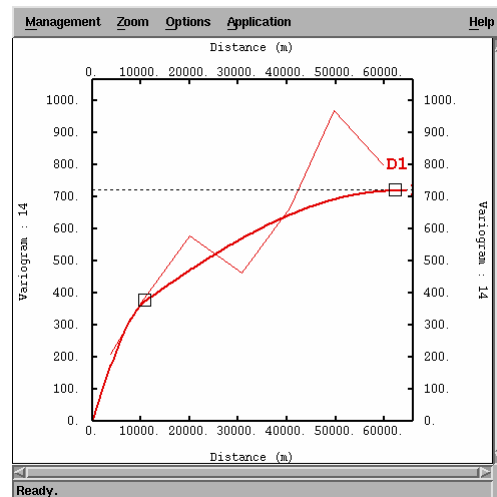
Cas du 25/09/03 à 12h TU

Enfin, dans les autres cas, le modèle double sphérique est nettement mieux adapté qu'un modèle sphérique.

Cas du 04/08/03 à 14h TU :

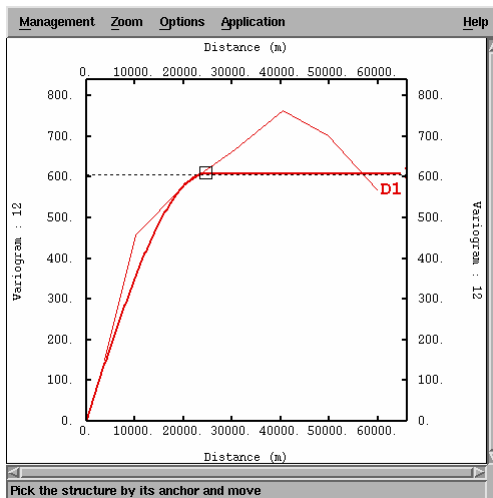


Modèle sphérique

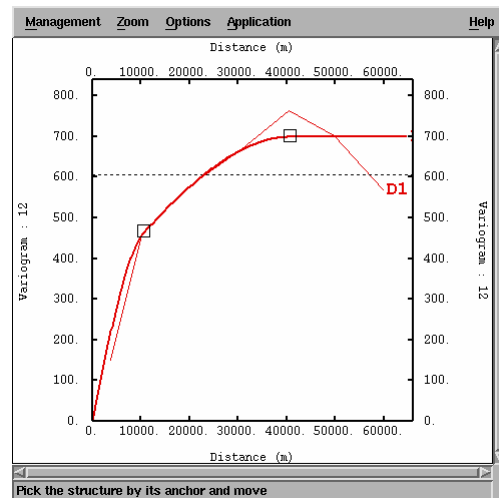


Modèle double sphérique

Cas du 29/05/03 à 12h TU :

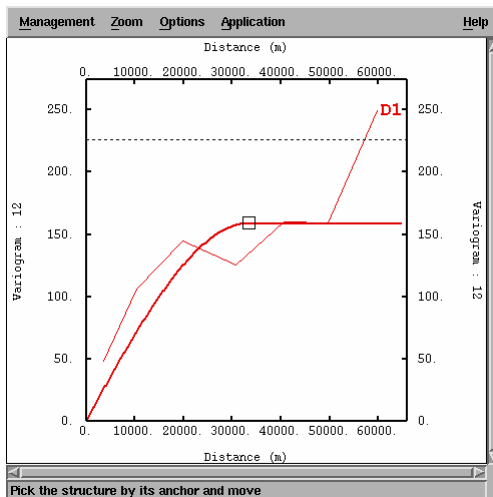


Modèle sphérique

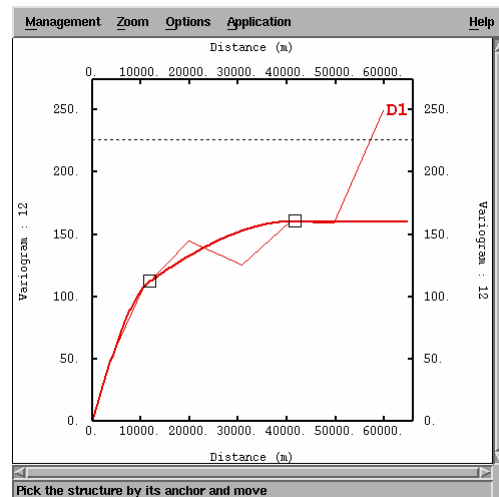


Modèle double sphérique

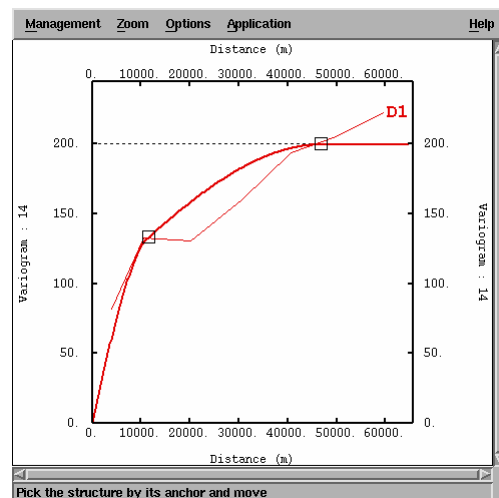
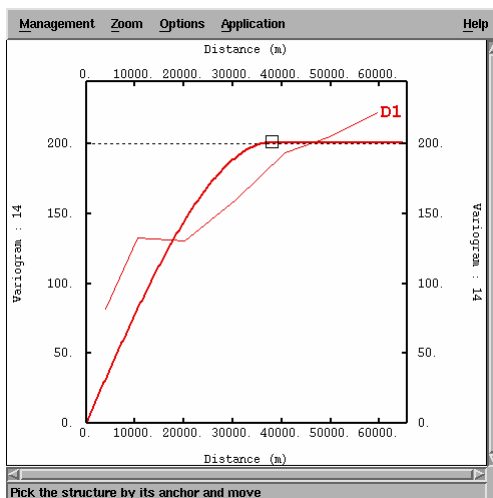
Cas du 21/07/03 à 12h et 14h TU :



Modèle sphérique



Modèle double sphérique

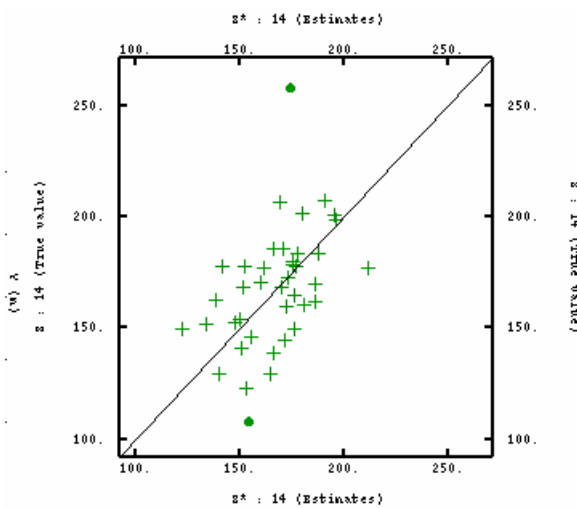


Dans ces quelques cas, il est difficile d'ajuster un modèle de variogramme sphérique qui prenne bien en compte les différentes « cassures » du variogramme expérimental. Le modèle double sphérique répond mieux à cette problématique.

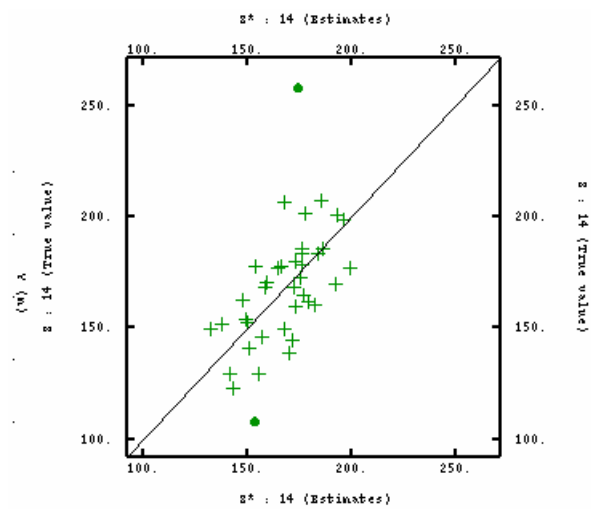
La question est de quantifier dans ces cas précis, la perte de qualité que peut engendrer l'utilisation, malgré tout, d'un modèle de variogramme sphérique, de façon à savoir si l'on peut utiliser ce type de modèle dans notre optique d'automatisation.

La comparaison des validations croisées et des cartographies finales obtenues pour ces différents cas est un moyen de répondre à cette question.

Si l'on observe le cas du 04/08/03 à 14h :



Validation croisée pour le modèle sphérique



Validation croisée pour le modèle double sphérique

On note une légère amélioration de la validation croisée avec un modèle double sphérique : la dispersion autour de la première bissectrice est plus faible ; ce qui confirme que ce type de modèle correspond mieux.

On peut le noter aussi au travers des **statistiques** :

Statistics based on 40 test data

	Mean	Variance
Error	-0.55755	591.13248
Std. Error	-0.01923	1.90158

Statistiques avec modèle sphérique

Statistics based on 40 test data

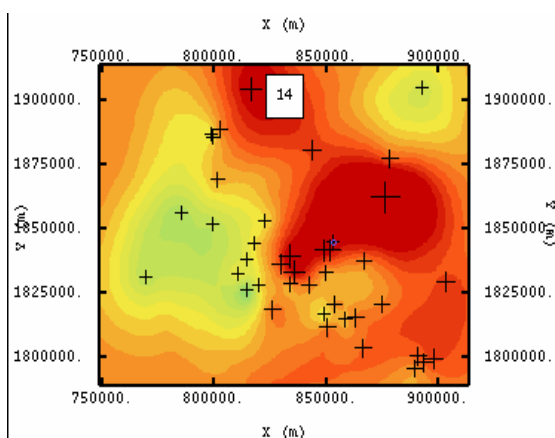
	Mean	Variance
Error	-0.12349	497.00637
Std. Error	-0.00496	1.14142

Statistiques avec modèle double sphérique

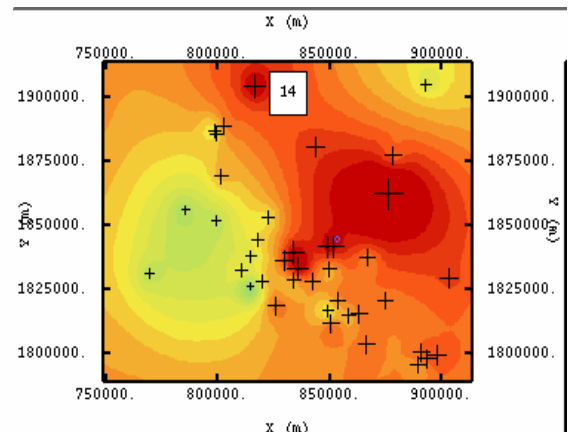
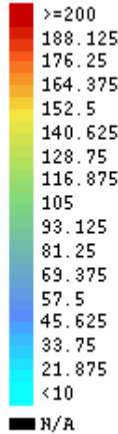
Avec le modèle double sphérique, l'erreur moyenne est divisée par 4 et la variance de l'erreur standardisée est plus proche de 1.

Cependant, les résultats obtenus avec le modèle sphérique restent très raisonnables. La perte en qualité est acceptable.

En ce qui concerne les deux cartographies correspondantes de la valeur horaire glissante d'ozone pour le 04/08/03 à 14h, elles diffèrent aussi légèrement mais en terme de représentation d'un phénomène global de mouvement des masses d'air pollué en ozone, on a deux cartes représentant assez bien le même mouvement et ainsi le même phénomène. On ne perd pas en qualité d'information.

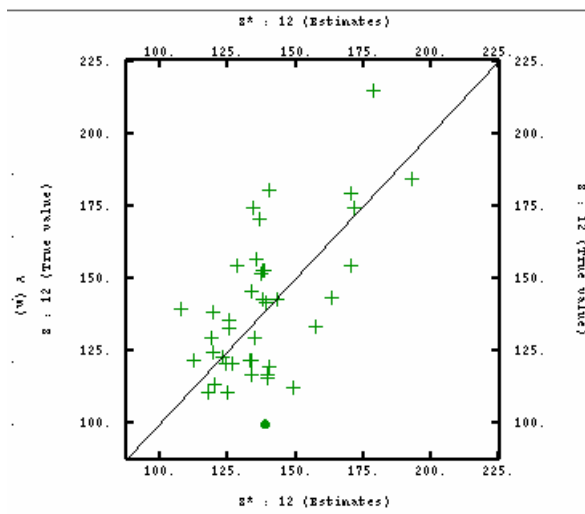


Cartographie par krigeage avec modèle sphérique

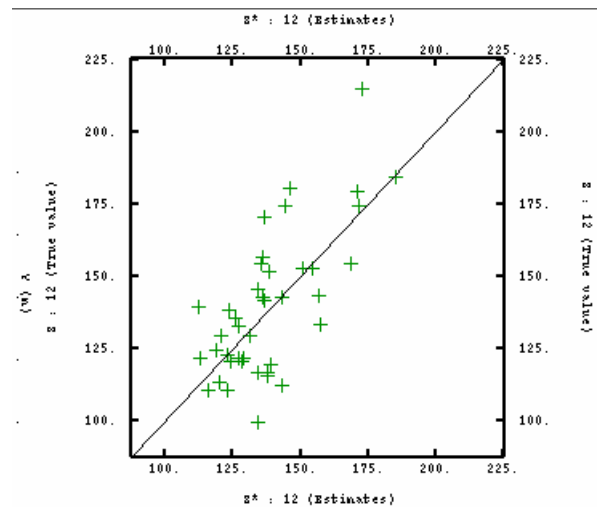


Cartographie par krigeage avec modèle double sphérique

Pour le 29/05/03 à 12h :



Validation croisée pour le modèle sphérique



Validation croisée pour le modèle double sphérique

Même remarque que précédemment.

Au niveau statistiques :

Statistics based on 41 test data

	Mean	Variance
Error	-1.06309	390.60868
Std. Error	-0.04003	1.35528

Statistiques avec modèle sphérique

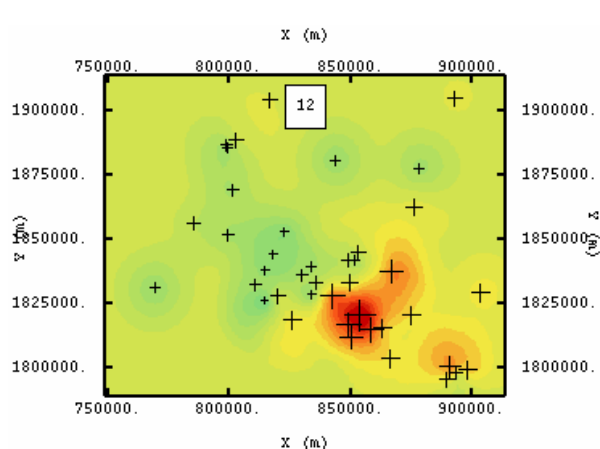
Statistics based on 41 test data

	Mean	Variance
Error	-0.8305	305.08974
Std. Error	-0.02865	0.81932

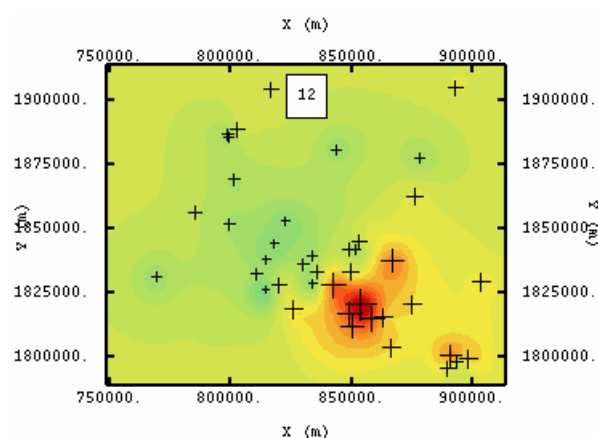
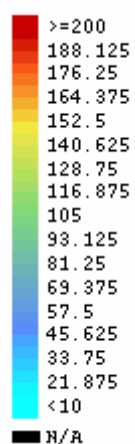
Statistiques avec modèle double sphérique

Même si avec le modèle double sphérique on a de meilleurs résultats, on reste quand même dans un domaine largement acceptable et assez similaire avec le modèle sphérique.

Au niveau des cartographies on observe deux représentations du phénomène très similaires.



Cartographie par krigeage avec modèle sphérique



Cartographie par krigeage avec modèle double sphérique

Pour le 21/07/03 à 12h et à 14h on pourra arriver aux mêmes conclusions.

Un modèle sphérique peut donc être aussi bien utilisé qu'un modèle double sphérique dans notre étude. La perte de qualité que peut engendrer ce choix est acceptable au regard de l'utilisation finale (cartographie à des fins de communication auprès du grand public) même dans les cas où le modèle double sphérique semble être plus approprié.

La question à présent est de savoir si l'on peut fixer une portée spécifique à ce modèle sphérique et si oui, laquelle ?

4.3. CHOIX DE LA MEILLEURE PORTEE POUR LE MODELE SPHERIQUE :

Suite à tous les tests réalisés, on a pu noter que :

- sur une même journée, il est difficile d'associer une même portée aux modèles sphériques correspondants aux différentes heures
- idem sur différentes journées aux mêmes heures.

En effet, le phénomène de pollution par l'ozone est très variable d'une journée sur l'autre (ce ne sont pas toujours exactement les mêmes heures qui sont atteintes par un pic d'ozone) et sur une même journée (évolution spatiale du phénomène).

La réponse à notre question précédente serait donc qu'il n'est pas possible de fixer une portée unique pour toutes les heures et toutes les journées.

Cependant, notre objectif consiste à bien représenter le phénomène global de pollution par l'ozone aux heures où les niveaux en ozone sont les plus importants ($> 180 \mu\text{g}/\text{m}^3$).

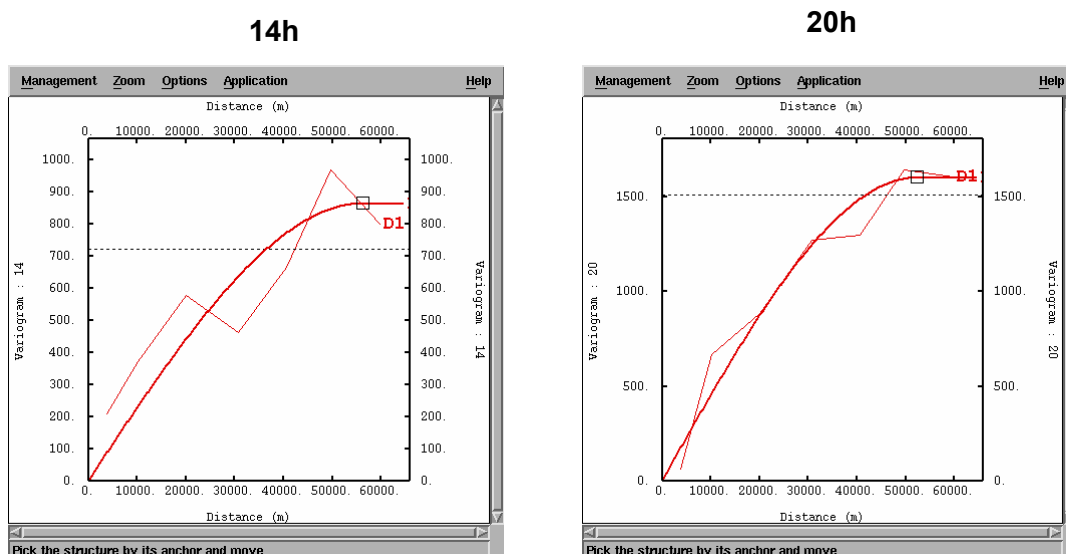
Ce sont à ces heures que l'on souhaite définir le modèle de variogramme le plus juste possible pour avoir la représentation du phénomène la plus juste possible.

4.3.1. POUR LE CAS DES JOURNEES ET HEURES AVEC DES EPISODES D'OZONE

STRUCTURES :

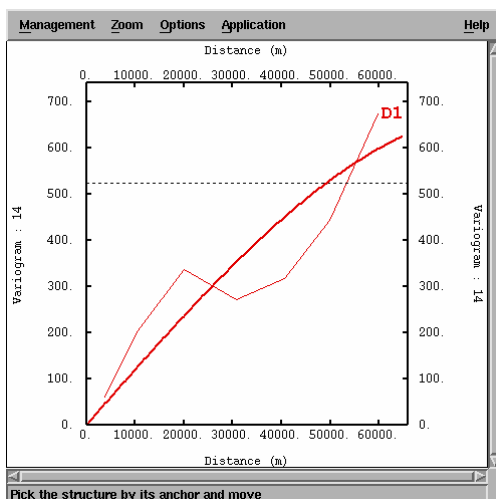
Pour les journées de test fixées initialement, si l'on observe particulièrement **les heures de pointes en ozone** (entre 10h et 20h TU, la plupart du temps), on note que la plupart des **modèles ont une portée comprise entre 50 et 80 km** (cf. exemples suivants).

Le 04/08/03 à 14h et 20h (le pic a duré de 10h à 20h TU) :

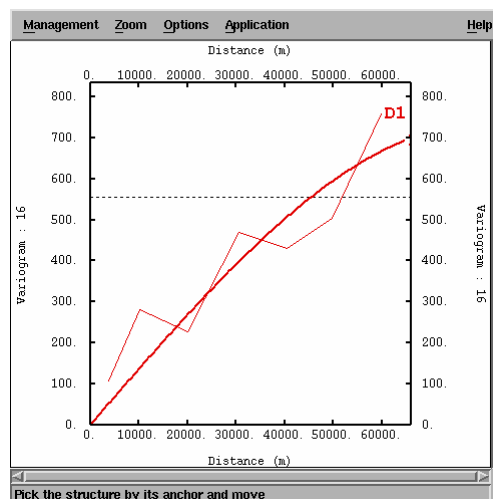


Le 29/05/03 à 14h et 16h (le pic a duré de 10h à 16h TU) :

14h



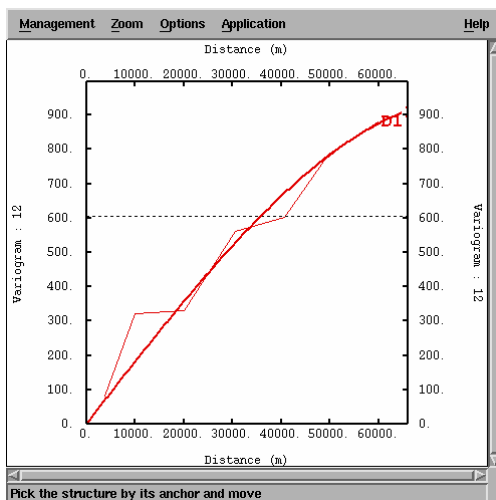
16h



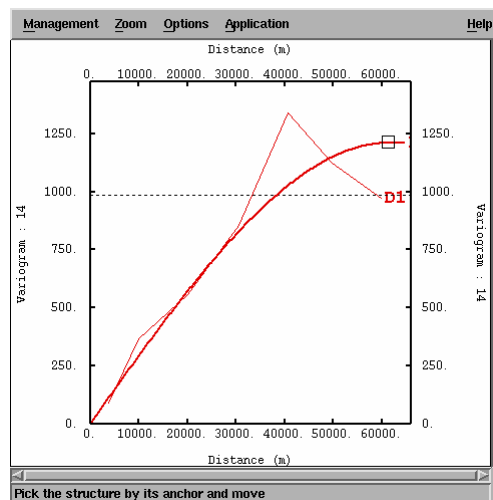
Portée autour de 80 km

Le 09/06/03 à 12h et 14h (le pic a duré de 10h à 14h) :

12h



14h

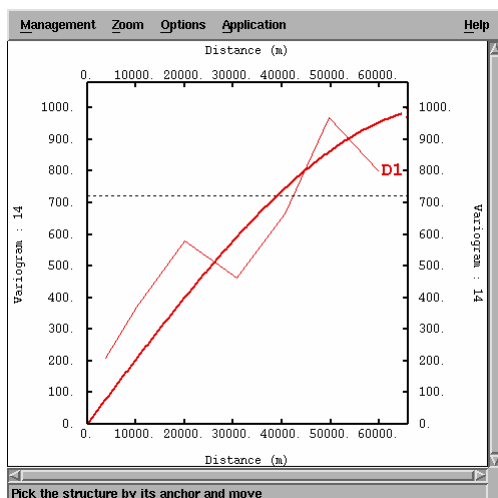


Portée entre 60 et 80 km

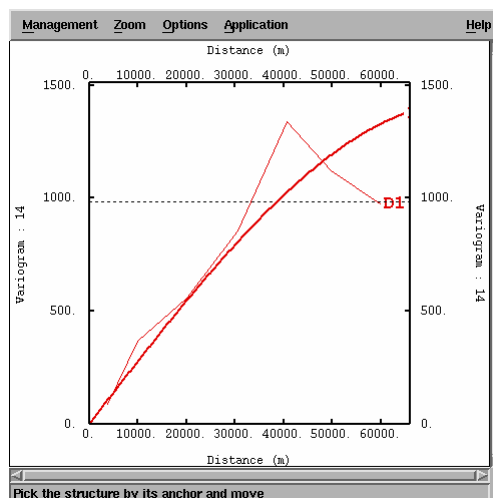
Dans tous ces cas, il est facile d'étirer légèrement le modèle de variogramme pour l'amener à une portée de 80 km sans grande modification du modèle et donc de la qualité du résultat final.

Etirement de la portée à 80 km pour le 04/08/03 à 14h et le 09/06/03 à 14h :

04/08/03 14h



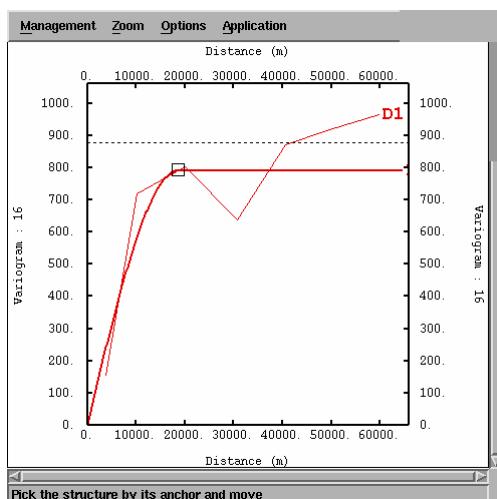
09/06/03 14h



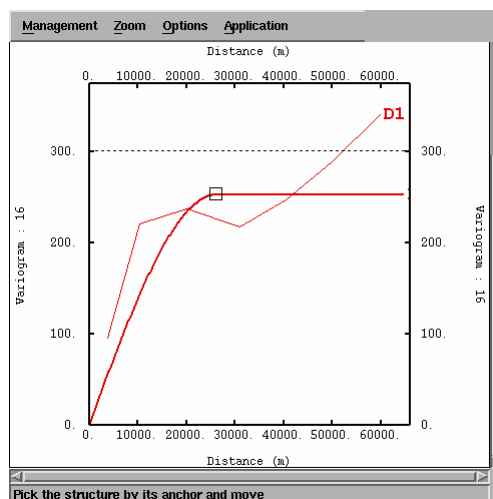
Ajustement de la portée à 80 km

On a pu aussi observer des cas, plus rares, où cette portée pouvait être plus petite, entre 20 et 40 km.

Le 04/08/03 à 16h et le 27/06/03 à 16h :



04/08/03 à 16h : portée à 20 km



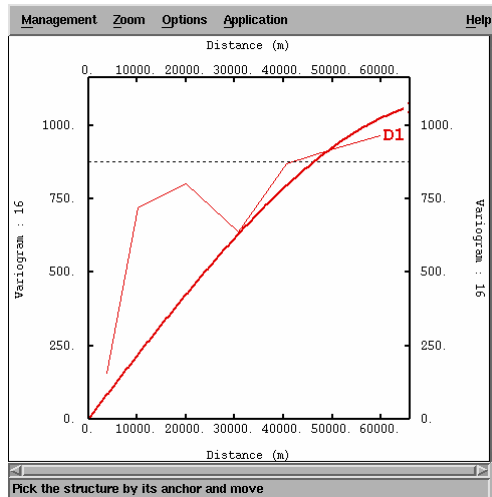
27/06/03 à 16h : portée entre 20 et 30km

Dans ces cas, on constate que le variogramme expérimental ne se stabilise pas mais se linéarise aux grandes distances : cela peut favoriser la possibilité d'une augmentation de la portée.

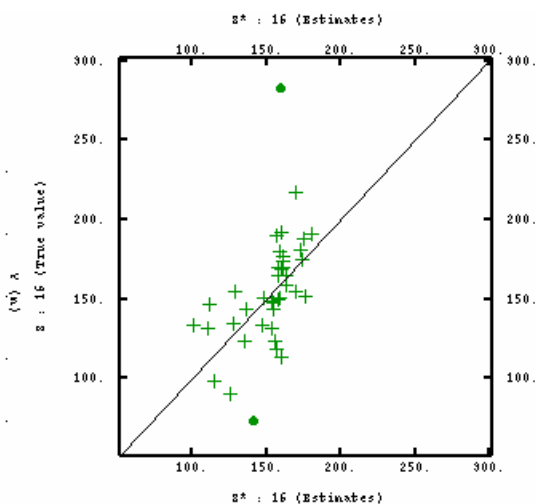
Dans tous les cas, il est nécessaire de fixer une portée dont on maîtrise les conséquences sur le résultat final.

Quelle est donc ici la conséquence d'utiliser aussi une portée à 80 km ?

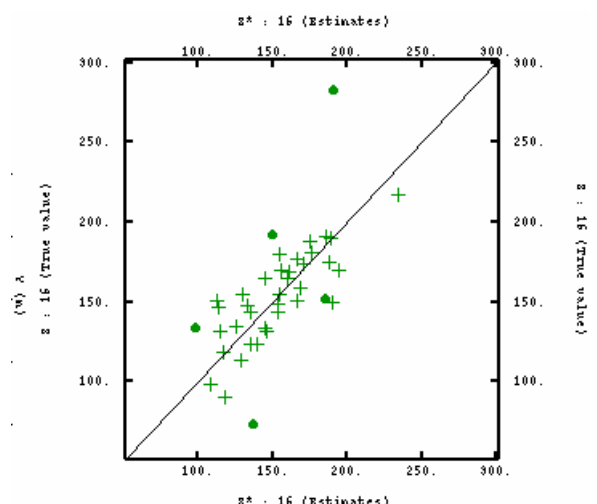
Etirement de la portée à 80 kms pour le 04/08/03 à 16h :



Le modèle de variogramme change, mais finalement « rattrape » une autre partie du modèle pour les grandes distances. On ne s'éloigne pas complètement de la réalité.



Validation croisée pour le modèle sphérique avec une portée à 20 km



Validation croisée pour le modèle sphérique avec une portée à 80 km

On note plus « d'outliers » avec une grande portée, mais les résultats sont largement acceptables dans les deux cas.

En terme de statistiques :

Statistics based on 40 test data

	Mean	Variance
Error	-1.88556	925.28248
Std. Error	-0.04204	1.56655

Statistiques pour une portée à 20 km

Statistics based on 40 test data

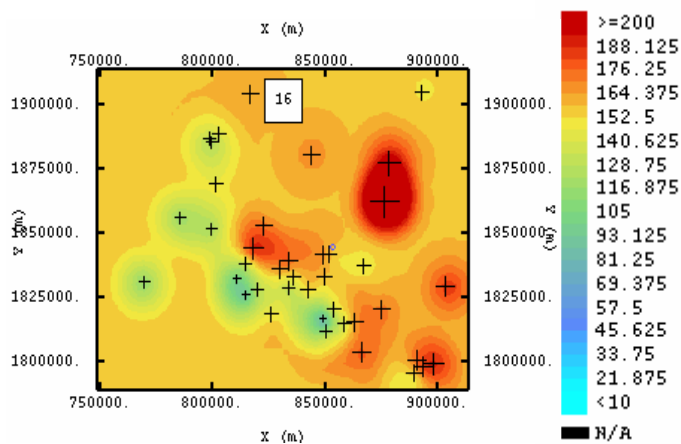
	Mean	Variance
Error	-0.56361	693.90491
Std. Error	-0.03441	3.16032

Statistiques pour une portée à 80 km

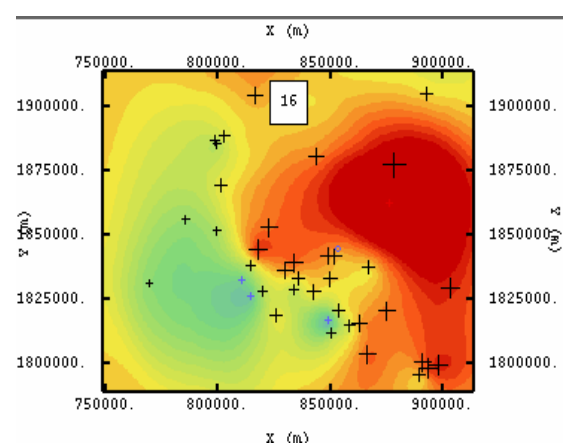
La moyenne de l'erreur est légèrement meilleure avec une grande portée, tandis que la variance autour de l'erreur est divisée par 2 avec une petite portée.

Ces statistiques sont malgré tout acceptables avec une grande portée.

Si l'on regarde les cartographies obtenues avec les différentes portées, les résultats diffèrent assez fortement.



Cartographie pour une portée à 20 km



Cartographie pour une portée à 80 km

En terme de représentation du phénomène global de pollution par l'ozone, on retrouve dans les deux cas une pollution visible au Nord-Est de la zone et pas de pollution au Sud-Ouest (en ne tenant pas compte de l'extrapolation).

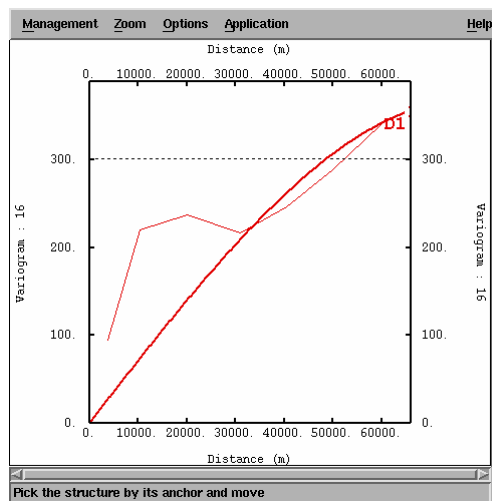
En revanche, dans le premier cas, on observe des bulles autour des points, caractéristiques d'une petite portée, et dans le deuxième cas, une image plus lissée liée à la plus grande portée et surtout plus représentative du phénomène de grande ampleur qu'est la pollution par l'ozone donc plus juste et plus acceptable pour notre étude.

Le modèle de variogramme à grande portée permet d'informer sur la variabilité à grande distance du phénomène, sans perdre l'information à petite distance même si le modèle suit moins bien le variogramme expérimental sur ces petites distances.

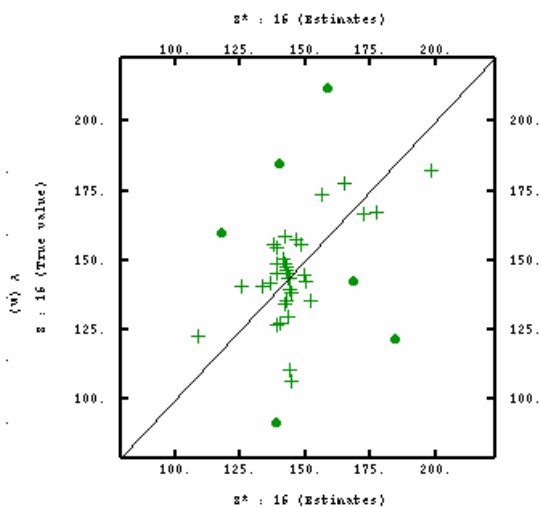
Le modèle de variogramme à petite portée suit bien le variogramme sur les petites distances, mais l'information à grande distance est totalement perdue : l'étendue du phénomène photochimique n'est pas représentée.

La variabilité à petite distance est moins importante par rapport à l'étendue du phénomène et du champ. Il faut donc lui donner moins de poids.

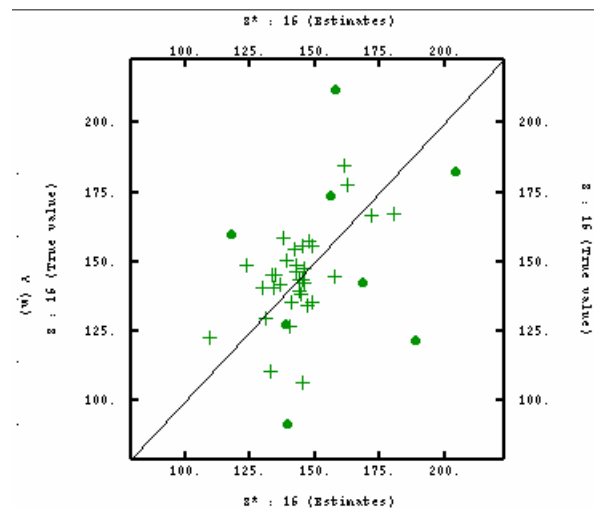
Etirement de la portée à 80 km pour le 27/06/03 à 16h :



Là aussi on note une différence au niveau du modèle entre une portée à 80 km et une portée entre 20 et 30 km.



Validation croisée pour une portée entre 20 et 30 km



Validation croisée pour une portée à 80 km

On note plus « d'ouliers » là aussi avec la grande portée mais les deux résultats sont très acceptables.

En terme de statistiques :

Statistics based on 41 test data

	Mean	Variance
Error	0.80528	479.34009
Std. Error	0.04218	4.47867

Statistiques avec une portée entre 20 et 30 km

Statistics based on 41 test data

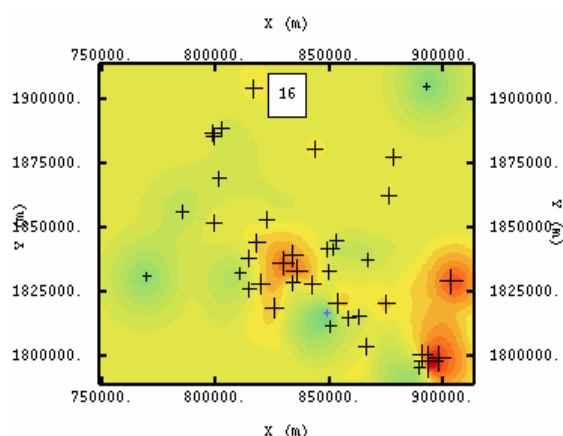
	Mean	Variance
Error	1.04101	459.91326
Std. Error	0.06638	9.200083

Statistiques avec une portée à 80km

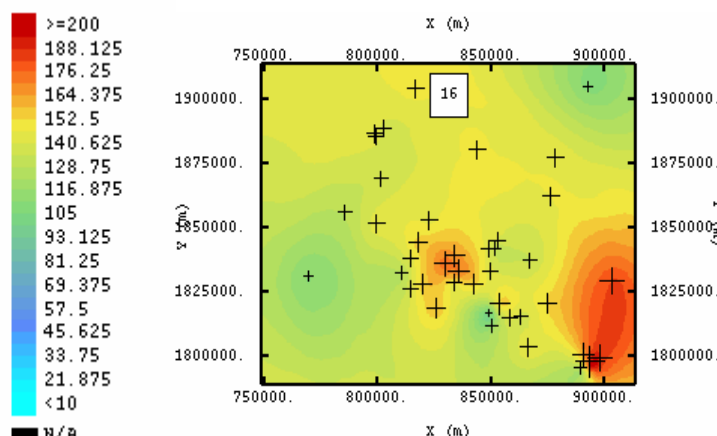
Les variances d'erreur standardisées sont très fortes ; cela vient du fait que l'on a masqué des points pour l'analyse variographique et donc réduit la variance expérimentale. Du coup, on divise des écarts quadratiques (calculés à partir de toutes les données) par une variance de krigeage dérivée d'un modèle dont la variance est inférieure à celle des données. Au final on obtient donc une variance d'erreur standardisée bien plus grande que 1. Celle du modèle à petite portée est deux fois plus acceptable.

La moyenne des erreurs est faible dans les deux cas.

En terme de cartographie, cette fois les résultats sont assez similaires.



Cartographie pour une portée entre 20 et 30 km.



Cartographie pour une portée à 80 km.

On note là aussi une image plus lissée au niveau du pic d'ozone (Sud-Est) avec une grande portée.

Là aussi, on constate que statistiquement, on a une interpolation plus juste aux points avec une petite portée, mais moins représentative du phénomène de grande ampleur de l'ozone.

Ce qui rejoint ce que l'on a pu observer tout au long de cette étude : lors des heures de pointes en pollution les résultats sont plus acceptables en terme d'interprétation avec une portée supérieure à 40 km : l'étendue du phénomène photochimique est mieux représentée qu'avec une petite portée.

En conclusion, on peut dans le cas présent généraliser l'utilisation d'un modèle sphérique avec une portée à 80 km et sans effet de pépité pour cartographier les épisodes de pollution par l'ozone.

4.3.2. POUR LES CAS DES JOURNEES ET HEURES SANS EPISODE D'OZONE

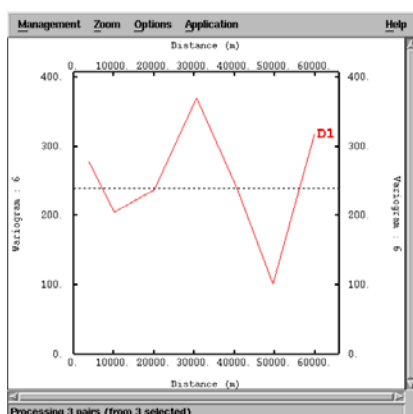
STRUCTURE :

Si l'on s'attarde à présent sur les **journées ou sur les heures sans pointe de pollution par l'ozone**, on peut noter des cas où les portées sont entre 40 et 80 km et d'autres cas où elles sont entre 10 et 40 km. Dans les deux cas, comme vu précédemment, on peut utiliser sans problème une portée à 80 km.

Si ce choix entraîne parfois une modification plus importante du modèle de variogramme, et que le résultat final est de moins bonne qualité car basé sur un modèle moins juste, cela n'importe pas devant notre objectif de bien représenter les heures de pointes. A ces heures, il n'y a pas d'ozone, et quelle que soit la qualité de la carte, ce sera bien pris en compte et on pourra l'observer...

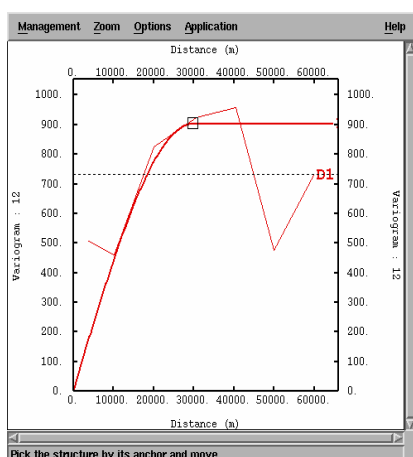
Il reste certaines heures où les variogrammes expérimentaux sont mal structurés.

Le 25/09/03 à 06h TU : (pas de pic ce jour-là)



On peut associer à ce variogramme expérimental un effet de pépite à 250.

Le 10/03/03 à 12h TU : pic très localisé et court sur la zone Nord-Est des BdR :



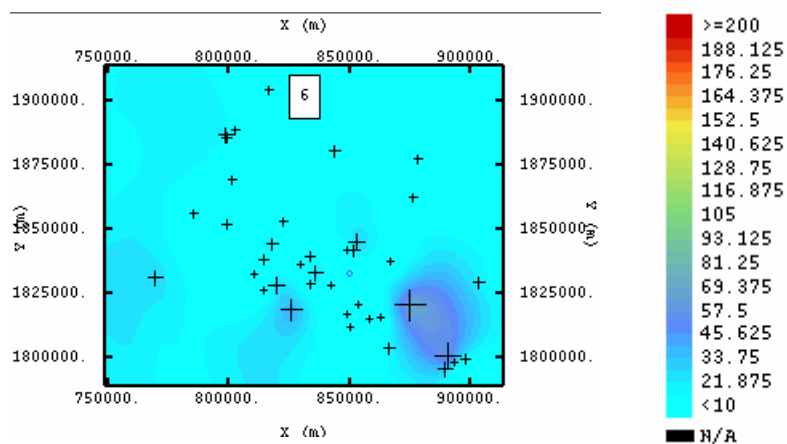
Portée autour de 30 km

On peut associer à ce variogramme un modèle sphérique avec une portée autour de 30 km sans prendre en compte alors la partie descendante du variogramme.

Quelle que soit la portée et le modèle que l'on va associer à ces variogrammes expérimentaux, le résultat final en terme de cartographie et de validation croisée restera le même : il ne devrait pas être de très bonne qualité donc à étudier avant d'en tirer des conclusions. Ce phénomène se produit en général pour des heures où les niveaux sont bas (tôt le matin) ou à des heures précises où les pics sont très localisés et où il peut ne pas y avoir de corrélation en fonction de la distance.

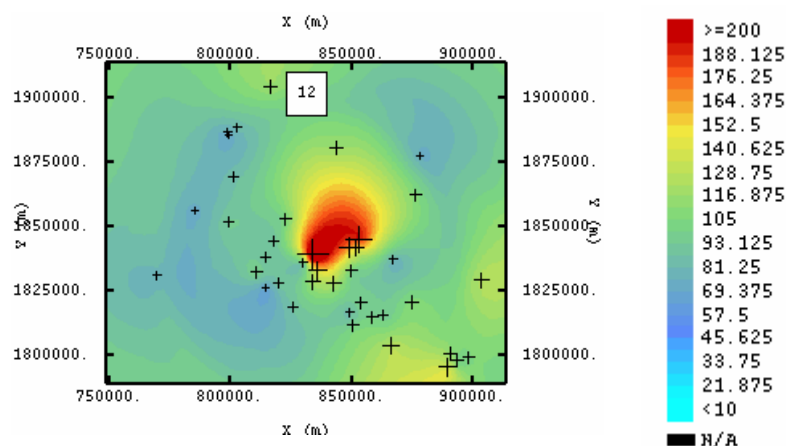
Dans le premier cas (exemple du 25/09/03 à 6h, cf. cartographie suivante), si l'on applique malgré tout un modèle sphérique à grande portée, les valeurs faibles sont quand même prises en compte et représentées correctement au niveau des stations, ce qui est suffisant comme information à ce stade de la journée.

Le choix du modèle ici et la cartographie finale n'ont donc pas trop d'importance si l'on considère l'objectif que l'on s'est fixé pour cette étude.



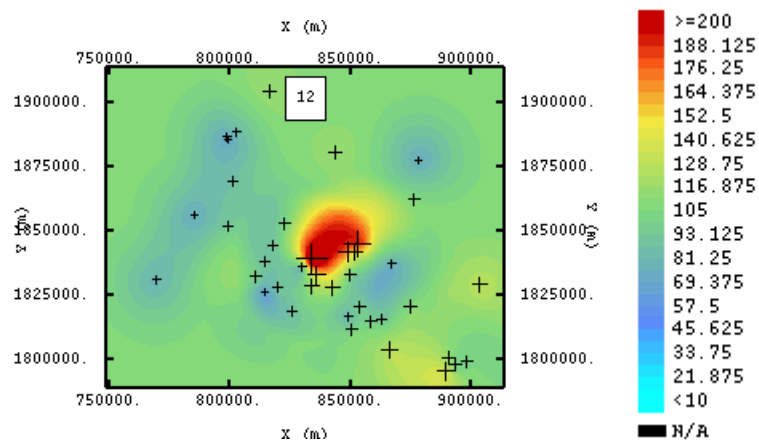
Cartographie du 25/09/03 à 6hTU avec un modèle sphérique et une portée à 80 km

Dans le deuxième cas (10/03/03 à 12h, cf. cartographie suivante), quelque soit le modèle choisi, on peut en revenir à la constatation faite pour des journées polluées avec une petite portée : il vaut mieux utiliser une grande portée qui lissera plus l'image et sera ainsi plus représentative du phénomène global de grande ampleur de l'ozone.



Cartographie du 10/03/03 à 12h TU avec un modèle sphérique et une portée à 80 km

Avec une portée autour de 30 km le résultat est assez similaire.



Cartographie du 10/03/03 à 12h TU avec un modèle sphérique et une portée à 30 km

Dans ce cas, le pic d'ozone est très localisé et très court. C'est un cas particulier où le phénomène n'a pas été de grande ampleur comme on peut le constater habituellement.

Ce cas fait partie de cas particuliers pour lesquels il faudra faire une étude plus détaillée avant de diffuser la cartographie obtenue.

4.4. CONCLUSION POUR LA CARTOGRAPHIE DES MOYENNES HORAIRES GLISSANTES EN OZONE :

Cette étude de différents types d'épisodes rencontrés durant l'été 2003 permet de tirer des conclusions générales sur la façon de cartographier les moyennes horaires glissantes d'ozone par krigeage de manière automatisée.

On a pu observer que le meilleur type de modèle à adapter pour cette cartographie était un **modèle double sphérique avec une portée à 10 km et une autre à 40 km**.

Ce type de modèle n'étant pas utilisable avec Spatial Analyst que nous souhaitons utiliser pour réaliser ces cartographies automatiquement, notre étude a permis de montrer **qu'un modèle de variogramme sphérique avec une grande portée fixée à 80 km** pouvait convenir dans l'ensemble des cas aux heures de pointe en ozone.

En dehors de ces heures et pour des journées sans pic d'ozone, ce type de modèle peut aussi convenir car il n'induit pas d'erreurs sensibles sur les résultats cartographiques.

Dans les autres cas où les données sont très déstructurées, le modèle sphérique avec une portée à 80 km sera malgré tout utilisé dans notre souci d'automatisation sachant que ces heures ont un intérêt moindre pour le public car non impactées par des niveaux élevés en ozone. Les valeurs faibles seront néanmoins prises en compte et représentées quelque soit le modèle choisi.

Pour certains cas particuliers comme la journée du 10/03/03 (pointe d'ozone très localisée), la cartographie doit être étudiée spécifiquement pour obtenir une carte fiable.

Quoiqu'il en soit, la validation humaine reste indispensable pour éviter la diffusion de cartes erronées (données manquantes, épisodes très particuliers...).