

STATISTIQUE ET ANALYSE DES DONNÉES

BERNARD CHALMOND

Régression avec résidus spatialement autocorrélés et recherche de la tendance spatiale

Statistique et analyse des données, tome 11, n° 2 (1986), p. 1-25

http://www.numdam.org/item?id=SAD_1986__11_2_1_0

© Association pour la statistique et ses utilisations, 1986, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

REGRESSION AVEC RESIDUS SPATIALEMENT AUTOCORRELÉS
ET RECHERCHE DE LA TENDANCE SPATIALE

Bernard CHALMOND

Université de Paris-Sud
Statistique appliquée
UA CNRS 743
91405 Orsay Cédex

Résumé: Reprenant les travaux [9] [19] [2], nous présentons les distributions asymptotiques des estimateurs du maximum de vraisemblance approchée pour un modèle de régression dont les résidus constituent un processus spatial stationnaire coloré. Comme illustration de ce modèle, nous considérons le problème classique de l'ajustement d'une surface spline bicubique, mais avec des écarts d'ajustement spatialement autocorrélés; une méthode numérique est proposée.

Abstract: Following the papers [9] [19] [20], we present the asymptotic distribution of the approximated maximum likelihood estimates for a regression model whose the residuals are a colored stationary spatial process. To illustrate this model, we consider the classical problem of the fit of a bicubic spline surface but with colored spatial deviations; a numerical method is proposed.

Mots clés: régression spatiale, processus spatial, estimateur du maximum de vraisemblance approchée, spline bicubique.

Classification STMA: 07-020, 07-160, 12-050.

Manuscrit reçu le Janvier 1986
révisé le 31 octobre 1986

1. INTRODUCTION

Les modèles de régression avec résidus non indépendants ont été largement étudiés. Le plus souvent, les résidus sont assujettis à un modèle ARMA et l'estimation des paramètres est pratiquée en maximisant une approximation de la vraisemblance, Pierce [16]. Ces dernières années, des travaux ont étendus ces résultats à des modèles avec résidus spatialement autocorrélés Ord [15], Mardia [13]. Pour situer les notes qui suivent, citons deux domaines d'applications qui prennent en compte l'autocorrélation spatiale: les expériences sur champ, Martin [14], l'analyse de texture et la restauration d'image, Kashyap [11], Chalmond [7].

Le modèle de régression spatial s'écrit:

$$z_{r,s} = \sum_k \sum_l X_{(r,s),(k,l)} \beta_{k,l} + y_{r,s} \quad (1)$$

où $(r,s) \in \mathbb{Z}^2$. En posant $t=(r,s)$ et $q=(k,l)$, on notera:

$$z_t = \sum_q X_{t,q} \beta_q + y_t = X_t \beta + y_t$$

Le processus z a été observé aux sommets d'une grille de taille $n=n_1 n_2$ suivant: $\{z_{r,s}; 1 \leq r \leq n_1, 1 \leq s \leq n_2\}$. Les $X_{t,q}$ sont des variables certaines et les β_q sont les coefficients de régression inconnus. Le processus des résidus y est un processus stationnaire au second ordre, centré et régulier.

Dans l'application [7], les résidus constituent un bruit ou une texture et la composante $X_t \beta$ de la régression représente l'image débruitée ou la surface portant la texture. Les objectifs poursuivis dans [7] sont: le débruitage, la caractérisation de texture, la détection de bord et la compression

d'image. Dans ce but, il faut alors définir les résidus et la surface, se donner une méthode d'estimation ainsi que les propriétés de cette estimation.

Au paragraphe II, nous rappelons des résultats maintenant classiques sur la définition des processus spatiaux, Guyon [10]. Dans cette esprit, y n'est pas défini par un modèle paramétré particulier, mais par une densité spectrale, ce qui permet de travailler dans un cadre général et ainsi d'énoncer des propriétés statistiques valides pour plusieurs classes de modèles. L'étude du comportement asymptotique des estimateurs basés sur une approximation de la vraisemblance est menée directement à partir des travaux de Whittle [20], Walker [19] et Guyon [9]. Au paragraphe III, un modèle de régression particulier est entrepris, permettant l'estimation de la tendance spatiale: les $X_{i,q}$ sont des fonctions splines bicubiques, DeBoor [2]. La mise en oeuvre numérique est faite pour des résidus régis par un processus spatial autorégressif unilatéral (ARU).

II. REGRESSION AVEC RESIDUS SPATIALEMENT AUTOCORRELÉS.

II.1. Définition des résidus.

Le processus des résidus y du modèle (1) est une suite de variables aléatoires indexées par Z^2 , telle que $E[y_{r,s}^2] < \infty$. $y_{r,s}$ appartient donc à l'espace L^2 des v.a. possédant un moment d'ordre 2; cet espace est hilbertien pour le produit scalaire $\langle x, y \rangle = E[xy]$. Le processus y est stationnaire au second ordre et centré: $E[y_{r,s}] = 0$, $\text{Cov}(y_{r,s}, y_{r+k, s+l}) = \gamma_{k,l} = \gamma_{-k, -l}$, $\forall r, s, k, l$. Le processus y a alors une mesure spectrale que l'on suppose absolument continue de densité spectrale $f(\omega_1, \omega_2)$. Le processus étant l'échantillonnage $\{\dots, y(r-n\Delta, s-m\Delta), \dots, y(r, s), \dots\}$ avec $\Delta=1$ d'un processus continue, la fréquence spatiale (ω_1, ω_2)

varie dans $T =]-\pi, \pi]^2$. On montre alors que:

$$\gamma_{k,l} = \frac{1}{4\pi^2} \int_T f(\omega_1, \omega_2) e^{I(\omega_1 k + \omega_2 l)} d\omega \quad (2)$$

$$f(\omega_1, \omega_2) = \sum_{(k,l) \in \mathbb{Z}^2} \gamma_{k,l} e^{-I(\omega_1 k + \omega_2 l)} \quad (3)$$

L'ordre lexicographique sur $\mathbb{Z}^2$ est l'ordre total suivant: $(r_1, s_1) < (r_2, s_2)$ si et seulement si $r_1 < r_2$ ou $r_1 = r_2$ et $s_1 < s_2$. Ainsi pour tout (r, s) , on fait apparaître une notion de "passé" et de "futur". Les résultats de la théorie de la prédiction linéaire dans L^2 , Anderson [1], s'étendent à $\mathbb{Z}^2$. Le processus y est dit régulier lexicographique si:

$$\sigma^2 = \text{Exp} \left[\frac{1}{4\pi^2} \int_T \text{Log } f(\omega_1, \omega_2) d\omega \right] > 0 \quad (4)$$

σ^2 est la variance du processus d'innovation ε , bruit blanc, dans la décomposition de Wold:

$$y_{r,s} = \varepsilon_{r,s} + \sum_{(i,j) > 0} \theta_{i,j} \varepsilon_{r-1,s-j} \quad (5)$$

qui est une représentation spatiale moyenne mobile unilatérale d'ordre infini MAU(∞). Sous la condition $\int f^{-1}(\omega_1, \omega_2) d\omega < \infty$, on a de plus la représentation autorégressive unilatérale d'ordre infini ARU(∞):

$$y_{r,s} + \sum_{(i,j) > 0} \phi_{i,j} y_{r-1,s-j} = \varepsilon_{r,s} \quad (6)$$

et aussi la représentation markovienne:

$$y_{r,s} + \sum_{(i,j) \neq 0} \lambda_{i,j} y_{r-1,s-1} = \varepsilon_{r,s} \quad (7)$$

où $e_{r,s}$ est un bruit stationnaire coloré, orthogonal à $y_{t,u}$ si $(t,u) \neq (r,s)$. Le processus est dit alors régulier non ordonné. Pour chacune de ces représentations la densité spectrale d'un tel processus est:

$$f(\omega_1, \omega_2) = \sigma^2 |\Theta(e^{-I\omega_1}, e^{-I\omega_2})|^2 \quad (5.1)$$

$$= \sigma^2 |\varphi(e^{-I\omega_1}, e^{-I\omega_2})|^{-2} \quad (6.1)$$

$$= \sigma_e^2 / \Lambda(e^{-I\omega_1}, e^{-I\omega_2}) \quad (7.1)$$

$$\text{avec:} \quad \Theta(B_1, B_2) = 1 + \sum_{(i,j)>0} \theta_{i,j} B_1^i B_2^j \quad (6.2)$$

$$\varphi(B_1, B_2) = 1 + \sum_{(i,j)>0} \phi_{i,j} B_1^i B_2^j \quad (6.2)$$

$$\Lambda(B_1, B_2) = 1 + \sum_{(i,j) \neq 0} \lambda_{i,j} B_1^i B_2^j \quad (7.2)$$

(6.1) peut encore s'écrire:

$$f(\omega_1, \omega_2) = \left| \sum_{(i,j) \geq 0} d_{i,j} e^{-I(\omega_1 i + \omega_2 j)} \right|^{-2}$$

où $d_{i,j} = \phi_{i,j} / \sigma$ avec $\phi_{0,0} = 1$. En développant cette expression, f^{-1} est une série de Fourier de coefficient $\alpha_{k,l}$:

$$f(\omega_1, \omega_2) = \left[\sum_{(k,l) \in \mathbb{Z}^2} \alpha_{k,l} e^{-I(\omega_1 k + \omega_2 l)} \right]^{-1} \quad (8)$$

$$\text{où:} \quad \alpha_{k,l} = \sum_{\substack{(i,j) \geq 0 \\ (i+k, j+l) \geq 0}} d_{i,j} d_{i+k, j+l}$$

$$= \frac{1}{4\pi} \int_T f^{-1}(\omega_1, \omega_2) e^{I(\omega_1 k + \omega_2 l)} d\omega \quad (9)$$

Par la suite, nous supposons que la densité spectrale est définie par un nombre fini de paramètres η .

II.2. Approximation de Whittle de la vraisemblance.

Si on suppose pour l'instant que le processus des résidus est gaussien, alors sa densité de probabilité s'écrit:

$$p(y) = (2\pi\sigma^2)^{-n/2} |V(\zeta)|^{-1/2} \text{Exp}\left[-\frac{1}{2\sigma^2} y'V(\zeta)^{-1}y\right]$$

pour la paramétrisation $\eta=(\zeta, \sigma^2)$ et où $\sigma^2 V(\zeta)$ est la matrice de variance covariante de y . Le logarithme de la vraisemblance est:

$$L(\zeta, \sigma^2 | y) = -\frac{1}{2} \text{Log}|V(\zeta)| - \frac{n}{2} \text{Log}(2\pi\sigma^2) - \frac{1}{2\sigma^2} y'V(\zeta)^{-1}y$$

Une première transformation de cette expression consiste à travailler conditionnellement au bord, c'est à dire à reprendre la transformation (6) en posant $y_{r,s} = 0$ si (r,s) n'appartient pas à la grille d'observation D . Ainsi, le passage de ε à y est univoque et:

$$L(\zeta, \sigma^2 | y) \approx -\frac{n}{2} \text{Log}(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{(r,s) \in D} (\varepsilon_{r,s} | Y_0)^2$$

où Y_0 rappelle le conditionnement. La somme des carrés résiduels conditionnels notée $S(\zeta)$, se développe comme suit:

$$S(\zeta) = \sum_{(r,s) \in D} \left[\sum_{\substack{(i,j) \geq 0 \\ (r-1, s-j) \in D}} \varphi_{i,j} y_{r-1, s-j} \right]^2$$

et approximativement:

$$S(\zeta) \approx n \sum_{k=-n_1+1}^{n_1-1} \sum_{l=-n_2+1}^{n_2-1} c_{k,l} \left[\sum_{\substack{(i,j) \geq 0 \\ (i+k, j+l) \geq 0}} \varphi_{i,j} \varphi_{i+k, j+l} \right] \quad (10)$$

$$C_{k,l} = \frac{1}{n} \sum_{\substack{(r,s) \in \mathcal{D} \\ (r+k,s+l) \in \mathcal{D}}} y_{r,s} y_{r+k,s+l}$$

En tenant compte de (9), il vient l'approximation:

$$\frac{S(\zeta)}{\sigma^2} \approx n \sum_{(k,l) \in \mathcal{D}} C_{k,l} \alpha_{k,l}(\eta)$$

où $\mathcal{D}$ désigne le domaine $]-n_1, n_1[\times]-n_2, n_2[$. Pour des problèmes de biais qui seront précisés plus loin, il est proposé dans [9] de remplacer $C_{k,l}$ par l'estimateur $G_{k,l} = n_1 n_2 C_{k,l} / ((n_1 - k)(n_2 - l))$ estimateur non biaisé de $\gamma_{k,l}$:

$$\frac{S(\zeta)}{\sigma^2} \approx U(\eta) = n \sum_{(k,l) \in \mathcal{D}} G_{k,l} \alpha_{k,l}(\eta) \quad (11)$$

L'approximation du logarithme de la vraisemblance est à une constante et un facteur près:

$$\begin{aligned} L(\eta|y) &= -n \text{Log} \sigma^2 - U(\eta) \\ &= -\frac{n}{4\pi^2} \int_T \text{Log} f(\omega, \eta) d\omega - U(\eta) \end{aligned} \quad (12)$$

$U(\eta)$ peut revêtir une expression approchée si on introduit l'estimateur de f appelé spectrogramme:

$$F(\omega) = \sum_{(k,l) \in \mathcal{D}} G_{k,l} e^{-I(\omega_1 k + \omega_2 l)} \quad (13)$$

et à partir duquel:

$$G_{k,l} = \frac{1}{4\pi^2} \int_T F(\omega) e^{I(\omega_1 k + \omega_2 l)} d\omega$$

D'où en tenant compte de (8):

$$U(\eta) = \frac{n}{4\pi} \int_T \frac{F(\omega)}{f(\omega, \eta)} d\omega \quad (14)$$

II.3. Distribution asymptotique des estimateurs.

Revenons au modèle de régression (1). Les paramètres inconnus $\xi = (\eta, \beta)$ sont estimés en maximisant la fonction:

$$L_n(\xi|z) = - \frac{n}{4\pi} \int_T \text{Log} f(\omega, \eta) d\omega - U_n(\xi) \quad (15)$$

On note: $q=(k, l)$, $t=(r, s)$, $\omega.q = \omega_1 k + \omega_2 l$, $n_q = (n_1 - k)(n_2 - l)$.

$$U_n(\xi) = n \sum_{q \in D} \alpha_q(\eta) G_q(\beta) = \frac{n}{4\pi} \int_T \frac{F(\omega, \beta)}{f(\omega, \beta)} d\omega \quad (16)$$

$$G_q(\beta) = \frac{1}{n_q} \sum_{\substack{t \in D \\ t+q \in D}} (z_t - X_t \beta) (z_{t+q} - X_{t+q} \beta)$$

$$= \frac{1}{4\pi} \int_T F(\omega, \beta) e^{i\omega.q} d\omega$$

$$F(\omega, \beta) = \sum_{q \in D} G_q(\beta) e^{-i\omega.q}$$

$\xi_0 = (\eta_0, \beta_0)$ désigne les vraies valeurs des paramètres et $\hat{\xi}_n$ leurs estimations.

Les hypothèses sur le modèle (1) sont:

(1) le processus des résidus y_t a une représentation stationnaire du type ARMA unilatérale ou du type markovien de densités spectrales respectives $f(\omega, \eta) = \sigma^2 |\theta(e^{-i\omega})|^2 |\varphi(e^{-i\omega})|^{-2}$ et $f(\omega, \eta) = \sigma^2 / \Lambda(e^{-i\omega})$ où θ , φ , Λ définis par (5.2), (6.2), (7.2) sont d'ordre fini, c'est à dire qu'ils n'ont qu'un nombre fini de paramètres non nuls. (D'autres formes sont possibles [9]):

(11) les variables de régression $X_{t,j}$ sont telles que:

$\lim_{n \rightarrow \infty} n^{-1/2} \sum_{t \in D} X_{t,j} e^{-I_0 \cdot t}$ est définie $\forall j \in T$ et la matrice de terme

$\lim_{n \rightarrow \infty} n^{-1} \sum_{t, u \in D} \alpha_{u-t}(\eta_0) X_{t,j} X_{u,j}$ est inversible.

Proposition: Pour ce modèle et la vraisemblance (15) les propriétés asymptotiques classiques s'étendent:

- . $\hat{\xi}_n$ converge en probabilité vers ξ_0 .
- . $n^{1/2}(\hat{\xi}_n - \xi_0)$ converge en loi vers une gaussienne centrée de matrice de variance covariance W (donnée à la fin de ce paragraphe).

La démonstration de ces propriétés suit celles données dans [18], [19], [9]. Ici, nous donnons seulement les éléments de la démonstration de la seconde propriété. $\hat{\xi}_n$ est solution du système $L_n^{(j)}(\hat{\xi}_n) = 0$, l'indice supérieur j désignant la dérivée par rapport au j ième paramètre de ξ . La distribution asymptotique des estimateurs du maximum de vraisemblance s'obtient en partant du théorème de la moyenne appliqué à $L_n^{(j)}(\hat{\xi}_n)$: il existe un vecteur $\xi_* \in [\hat{\xi}_0, \hat{\xi}_n]$ tel que:

$$L_n^{(j)}(\hat{\xi}_n) = L_n^{(j)}(\xi_0) + \sum_1 (\hat{\xi}_{n1} - \xi_{01}) L_n^{(ij)}(\xi_*) \quad (17)$$

où $\lim \xi_* = \xi_0$. Sous forme matricielle, il vient:

$$H_n(\xi_*) n^{1/2}(\hat{\xi}_n - \xi_0) = B_n(\xi_0)$$

avec: $H_n(\xi_*)_{ij} = -n^{-1} L_n^{(ij)}(\xi_*)$ et $B_n(\xi_0)_j = n^{-1/2} L_n^{(j)}(\xi_0)$.

La convergence en distribution de $n^{1/2}(\hat{\xi}_n - \xi_0)$ découle de l'étude la convergence en probabilité de $H_n(\xi_*)$ et de la convergence en distribution de $B_n(\xi_0)$. Pour β_0 connu, la conclusion est donnée par la proposition 4 de [9] qui étend le théorème 2 de [19] dont la démonstration dans le cas de Z est très détaillée. Pour notre modèle de régression cette démonstration est peu sensiblement modifiée:

a) plim $H_n(\xi_*) = H_n(\eta_0)$. Pour cela, on considère:

$$n^{-1} L_n^{(ij)}(\xi_*) = n^{-1} [L_n^{(ij)}(\xi_*) - L_n^{(ij)}(\xi_0)] + \\ n^{-1} [L_n^{(ij)}(\xi_0) - E[L_n^{(ij)}(\xi_0)]] + n^{-1} E[L_n^{(ij)}(\xi_0)]$$

• Le premier terme de cette décomposition converge en probabilité vers 0. Il s'écrit:

$$- \frac{1}{4\pi^2} \int_T \text{Log} \frac{f(\omega, \eta_*)}{F(\omega, \beta_0)} d\omega - [U_n^{(ij)}(\xi_*) - U_n^{(ij)}(\xi_0)]$$

L'intégrale converge en probabilité vers 0 car elle est continue en η_* . Quand les dérivations portent η et β , la différence des $U_n^{(ij)}$ est: $\sum_{q \in \mathcal{D}} \alpha_q^{(i)}(\eta_*) G_q^{(j)}(\beta_*) - \alpha_q^{(i)}(\eta_0) G_q^{(j)}(\beta_0)$

En appliquant le théorème de la moyenne à $G_q^{(j)}$:

$$G_q^{(j)}(\beta_*) = G_q^{(j)}(\beta_0) + \sum_k (\beta_* - \beta_0)_k G_q^{(jk)}(\beta_*)$$

où $\beta_* \in [\beta_0, \beta_*]$, cette différence devient:

$$\sum_k \sum_{q \in \mathcal{D}} \alpha_q^{(i)}(\eta_*) (\beta_* - \beta_0)_k G_q^{(jk)}(\beta_*) + \sum_{q \in \mathcal{D}} [\alpha_q^{(i)}(\eta_*) - \alpha_q^{(i)}(\eta_0)] G_q^{(j)}(\beta_0).$$

$\forall k$, en notant $h=f^{-1}$, le premier terme de cette somme s'écrit comme (16):

$$(\beta_* - \beta_0)_k \frac{1}{4\pi^2} \int_T F^{(jk)}(\omega, \beta_*) h^{(i)}(\omega, \eta_*) d\omega$$

dont la majoration:

$$\leq |\beta_* - \beta_0|_k \frac{M}{4\pi^2} \int_T |F^{(jk)}(\omega, \beta_*)| d\omega$$

converge vers 0 en probabilité car pour n suffisamment grand: $|\eta_* - \eta_0| < \delta$ et M est la borne supérieure de $|h^{(i)}(\omega, \eta)|$ en ω et η avec $|\eta - \eta_0| < \delta$ et par ailleurs:

$$F^{(1k)}(\omega, \beta_+) = \sum_{q \in \mathcal{D}} G_q^{(jk)}(\beta_+) e^{-i\omega \cdot q}$$

$$\# \frac{2}{n} \sum_{q \in \mathcal{D}} \sum_{\substack{t \in \mathcal{D} \\ t+q \in \mathcal{D}}} X_{t+q, j} X_{t, k} e^{-i\omega \cdot q} = \frac{2}{n} \sum_{t \in \mathcal{D}} X_{t, j} e^{-i\omega \cdot t} \sum_{t \in \mathcal{D}} X_{t, k} e^{i\omega \cdot t}$$

est définie quand $n \rightarrow \infty$ de par (11). Le deuxième terme de la somme s'écrit aussi:

$$\frac{1}{4\pi^2} \int_T F^{(j)}(\omega, \beta_0) [h^{(1)}(\omega, \eta_+) - h^{(1)}(\omega, \eta_0)] d\omega$$

qui converge en probabilité vers 0 suivant les mêmes considérations que précédemment par majoration de la différence des $h^{(1)}$ en utilisant une nouvelle fois le théorème de la moyenne. Quand les dérivations portent sur β seul ou sur η seul, la convergence est établie de manière semblable.

• Le deuxième terme de la décomposition converge en probabilité vers 0. Quand les dérivations portent sur η , la variance de ce terme tends vers 0 par application de la proposition 2 de [9] entraînant la convergence en moyenne quadratique et donc en probabilité. Quand les dérivations portent sur β , ce terme est nul et quand les dérivations portent sur η et β , la convergence s'obtient par un développement comme ci-dessus.

• On a donc: $\text{plim } -n^{-1} L_n^{(1j)}(\xi_+) = -\lim E[-n^{-1} L_n^{(1j)}(\xi_0)] = H_{1j}(\xi_0)$. Quand les dérivations portent sur η , alors d'après [19], [9]:

$$H_{1j}(\xi_0) \stackrel{\Delta}{=} H_{11j}(\xi_0) = \frac{1}{4\pi^2} \int_T \frac{\partial \text{Log} f(\omega, \eta_0)}{\partial \eta_1} \frac{\partial \text{Log} f(\omega, \eta_0)}{\partial \eta_j} d\omega \quad (18)$$

Quand les dérivations portent sur η et β : $E[-n^{-1} L_n^{(1j)}(\xi_0)] = 0$. Quand les dérivations portent sur β :

$$H_{1j}(\xi_0) = \lim_{n \rightarrow \infty} \sum_{q \in \mathcal{D}} \alpha_q(\eta_0) G_q^{(1j)}(\beta_0)$$

$$\# \lim_{n \rightarrow \infty} \frac{2}{n} \sum_{q \in \mathcal{D}} \alpha_q(\eta_0) \sum_{\substack{t \in \mathcal{D} \\ t+q \in \mathcal{D}}} X_{t,1} X_{t+q,j} = \lim_{n \rightarrow \infty} \frac{2}{n} \sum_{t \in \mathcal{D}} \sum_{u \in \mathcal{D}} \alpha_{u-t}(\eta_0) X_{t,1} X_{u,j}$$

et d'après (9) et l'hypothèse (11):

$$H_{1,j}(\xi_0) = \lim_{n \rightarrow \infty} \frac{2}{n 4\pi^2} \int_T \sum_{t \in \mathcal{D}} X_{t,1} e^{-i\omega \cdot t} \sum_{u \in \mathcal{D}} X_{u,j} e^{i\omega \cdot u} f^{-1}(\omega, \eta_0) d\omega \quad (19)$$

que l'on note $H_{1,j}(\xi_0)$. $H(\xi_0)$ est une matrice bloc diagonale qui ne dépend pas de β_0 ; on la note:

$$H(\eta_0) = \begin{bmatrix} H1(\eta_0) & 0 \\ 0 & H2(\eta_0) \end{bmatrix}$$

$$b) B_n(\xi_0) \xrightarrow{L} B: LG(0, W(\eta_0)).$$

$\lim E[B_n(\xi_0)] = 0$. Quand la dérivation porte sur η_0 , la nullité de cette limite tient au fait que G_q a été considéré au lieu de C_q ; sinon, il existerait un biais asymptotique [9].
Exprimons alors:

$$\lim_{n \rightarrow \infty} E[n^{-1} L_n^{(i)}(\xi_0) L_n^{(j)}(\xi_0)] = W(\xi_0)$$

Quand les dérivations portent sur η , alors la sous matrice correspondante dans W , notée $W1$, s'écrit d'après [19], [9]:

$$W1(\xi_0) = 2H1(\eta_0) + F(\eta_0) \stackrel{\Delta}{=} W1(\eta_0)$$

$$F_{1,j}(\eta_0) = \frac{K_4}{(2\pi)^4} \int_T \frac{\partial \text{Log} f(\omega, \eta_0)}{\partial \eta_1} d\omega \int_T \frac{\partial \text{Log} f(\omega, \eta_0)}{\partial \eta_j} d\omega \quad (20)$$

où $K_4 = E[\varepsilon_t^4]$. Quand les dérivations portent sur η et β :
 $W_{1,j}(\xi_0) = 0$. Quand les dérivations portent sur β seul:

$$n^{-1/2} L_n^{(j)}(\xi_0) = n^{1/2} \sum_{q \in \mathcal{D}} \alpha_q(\eta_0) G_q^{(j)}(\beta_0)$$

$$\# n^{-1/2} \sum_{t \in \mathcal{D}} \sum_{u \in \mathcal{D}} \alpha_{t-u}(\eta_0) (-X_{t,j} y_u - X_{u,j} y_t)$$

$$\begin{aligned}
E[n^{-1}L_n^{(i)}(\xi_0)L_n^{(j)}(\xi_0)] &= \frac{4}{n} \sum_{t,u} \sum_{r,s} \alpha_{t-u} \alpha_{r-s} E[X_{t,i} y_u X_{r,j} y_s] \\
&= \frac{4}{n} \sum_{t,r} X_{t,i} X_{r,j} \sum_{u,s} \gamma_{u-s} \alpha_{t-u} \alpha_{r-s}
\end{aligned}$$

Or avec (9), on montre que $\sum_u \gamma_{u-s} \alpha_{t-s} = 1$ si $t=s$ et 0 sinon. Par conséquent :

$$W_{ij}(\xi_0) = \lim_{n \rightarrow \infty} \frac{4}{n} \sum_{t \in D} \sum_{r \in D} X_{t,i} X_{r,j} \alpha_{t-r}(\eta_0)$$

qui n'est autre que l'expression $2H_{2ij}(\xi_0)$ à un facteur près :

$$W_{ij}(\xi_0) = 2H_{2ij}(\eta_0) \stackrel{\Delta}{=} W_{2ij}(\eta_0) \quad (21)$$

Enfin, la limite gaussienne se démontre comme dans [19].

Epilogue: $n^{1/2}(\hat{\xi}_n - \xi_0) \xrightarrow{L} H(\eta_0)^{-1} B$. La matrice de variance covariance asymptotique est :

$$W(\eta_0) = H(\eta_0)^{-1} W(\eta_0) H(\eta_0)^{-1} = \begin{bmatrix} W1(\eta_0) & 0 \\ 0 & W2(\eta_0) \end{bmatrix}$$

où : $W1(\eta_0) = H1(\eta_0)^{-1} [2H1(\eta_0) + F(\eta_0)] H1(\eta_0)^{-1}$, $W2(\eta_0) = 2H2(\eta_0)^{-1}$ sont définis par (18), (19) et (20). $\hat{\eta}_n$ et $\hat{\beta}_n$ sont asymptotiquement non corrélés.

II.4. Cas particulier: modèle ARMAU.

Le processus des résidus y est régi par le modèle ARMA spatial unilatéral d'ordre fini :

$$\sum_{(i,j) \in \mathcal{D}_\varphi} \varphi_{i,j} y_{r-i, s-j} = \sum_{(i,j) \in \mathcal{D}_\theta} \theta_{i,j} \varepsilon_{r-i, s-j}$$

où $\mathcal{D}_\varphi$ et $\mathcal{D}_\theta$ sont inclus dans le demi-plan positif $\{(i,j) \geq 0\}$, avec $\varphi_{0,0}=1$, $\theta_{0,0}=1$. Sa densité spectrale est:

$$f(\omega, \eta) = \sigma^2 |\theta(e^{-i\omega})|^2 |\varphi(e^{-i\omega})|^{-2} \stackrel{\Delta}{=} \sigma^2 g(\omega, \zeta)$$

ε_t étant un processus d'innovation, il vient à partir de (4):

$$\int_T \text{Log } g(\omega, \zeta) d\omega = 0, \quad \int_T \frac{\partial}{\partial \zeta_j} \text{Log } g(\omega, \zeta) d\omega = 0 \quad (22)$$

Ces égalités permettent d'exprimer W . Par exemple, pour la covariance asymptotique W_{ij} entre $\hat{\sigma}^2$ et $\hat{\theta}_j$, on calcule: $F_{ij}(\eta)=0$ et

$$H1_{ij}(\eta) = \frac{2}{4\pi^2 \sigma^2} \int_T \frac{e^{-i\omega \cdot j}}{\theta(e^{-i\omega})} d\omega = 0.$$

Cette dernière est nulle car elle correspond à (22) avec $g(\omega, \zeta) = \theta(e^{-i\omega})$, θ étant unilatéral. Les autres termes F_{ij} et $H1_{ij}$ se calculent de façon semblable; ainsi, pour la covariance asymptotique entre θ_i et θ_j , on calcule: $F_{ij}(\eta)=0$ et

$$H1_{ij}(\eta) = \frac{2}{4\pi^2} \int_T \frac{e^{-i\omega(i-j)}}{|\theta(e^{-i\omega})|^2} d\omega = 2\gamma_{\theta_{ij}} \quad (23)$$

Le résultat est:

$$W1(\sigma^2, \varphi, \theta) = \begin{bmatrix} \sigma^4 (2 + K_4) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \dots & 0 \\ \gamma_\varphi & & \gamma_{\varphi\theta} \\ \gamma_{\varphi\theta} & & \gamma_\theta \end{bmatrix}^{-1}$$

$\hat{\sigma}^2$ et $(\hat{\varphi}, \hat{\theta})$ sont asymptotiquement non corrélés. γ_θ définie en (23) n'est autre que la fonction d'autocovariance associée au spectre $|\theta(e^{-i\omega})|^{-2}$. Similairement, γ_φ est la fonction d'

autocovariance associée au spectre $|\varphi(e^{-i\omega})|^{-2}$ et $\gamma_{\theta\varphi}$ est associée au cospectre $[\theta(e^{-i\omega})\varphi^*(e^{-i\omega})]^{-1}$. Enfin, en développant $\varphi(e^{-i\omega})/\theta(e^{-i\omega})$ suivant $\sum_{q \geq 0} \psi_q e^{-i\omega \cdot q}$, de (19) il vient:

$$[W_2(\eta)^{-1}]_{ij} = \lim_{n \rightarrow \infty} \frac{1}{n\sigma^2} \sum_{t \in D} \left[\frac{\varphi(B)}{\theta(B)} X_{t,i} \right] \left[\frac{\varphi(B)}{\theta(B)} X_{t,j} \right]$$

où $B=(B_1, B_2)$ est ici vu comme l'opérateur retard ligne-colonne sur $t=(r, s)$: $B_1^k B_2^l X_{(r,s),i} = X_{(r-k, s-l),i}$. On retrouve ainsi, les expressions données par Pierce [16] dans le cas de régression avec résidus ARMA sur Z .

III. RECHERCHE DE LA TENDANCE SPATIALE.

Le processus z observé $\{z_{r,s}; 1 \leq r \leq n_1; 1 \leq s \leq n_2\}$ est maintenant régi par un modèle ARU non stationnaire au premier ordre:

$$z_{r,s} = \mu_{r,s} + y_{r,s} \quad (24)$$

$\mu_{r,s} = E[z_{r,s}]$ est la tendance spatiale qu'il s'agit d'estimer.

III.1. Régression sur fonctions splines bicubiques.

$\mu_{r,s}$ est supposé être l'échantillon d'une surface bicubique par morceaux de la manière suivante, DeBoor [2], Prenter [17]. Soient $\{\tau_0, \dots, \tau_{K+1}\}$ et $\{\lambda_1, \dots, \lambda_{L+1}\}$ deux sous ensembles de $\{1, \dots, n_1\}$ et $\{1, \dots, n_2\}$. Ils constituent chacun une partition $P\tau$ sur l'intervalle $[1, n_1]$ et $P\lambda$ sur l'intervalle $[1, n_2]$ telles que $\tau_0 = 1 \leq \tau_1 \leq \dots \leq \tau_{K+1} = n_1$ et $\lambda_0 = 1 \leq \lambda_1 \leq \dots \leq \lambda_{L+1} = n_2$ et de là une partition P sur $[1, n_1] \times [1, n_2]$ suivant des rectangles $R_{kl} = [\tau_{k-1}, \tau_k] \times [\lambda_{l-1}, \lambda_l]$. On suppose que les pas $p_\tau = \tau_k - \tau_{k-1}$ et $p_\lambda = \lambda_l - \lambda_{l-1}$ sont constants.

On définit les fonctions splines bicubiques sur P comme:
 $S_P = \{f: [1, n_1] \times [1, n_2] \rightarrow R \text{ telle que: } 1^\circ) f \text{ est un polynome de degré } 3 \text{ en } r \text{ et } s, \forall R_{k,l}; 2^\circ) f \in C_A^2: \partial^{p+q} f / \partial r^p \partial s^q \text{ est continue pour } 0 \leq p, q \leq 2\}$. S_P est de dimension $(K+4)(L+4)$. Donnons une de cet espace à partir des fonctions B-splines sur Pr et $P\lambda$. On définit les B-splines de rang k sur Pr au moyen des différences divisées d'ordre 4:

$$N_k^\tau(r) = \sum_{i=k}^{k-4} (r - \tau_i)_+^3 / \prod_{\substack{j=k \\ j \neq i}}^{k-4} (\tau_j - \tau_i)$$

où $r \in [1, n_1]$, $k=1, \dots, K+4$ et $(r - \tau_i)_+ = r - \tau_i$ si $r > \tau_i$, 0 sinon. N_k^τ est nulle en dehors de l'intervalle $[\tau_{k-4}, \tau_k]$. Cette définition nécessite d'étendre l'ensemble $\{\tau_k\}$ suivant le pas p : $\tau_{-3} < \tau_{-2} < \tau_{-1} < \tau_0 < \dots < \tau_{K+1} < \tau_{K+2} < \tau_{K+3} < \tau_{K+4}$. $\{N_k^\tau; k=1, \dots, K+4\}$ est une base des fonctions splines cubiques sur Pr :
 $S_{P\tau} = \{f: [1, n_1] \rightarrow R \text{ telle que: } 1^\circ) f \text{ est un polynome de degré } 3 \text{ en } r, \forall [\tau_{k-1}, \tau_k]; 2^\circ) \partial^p f / \partial r^p \text{ est continue pour } 0 \leq p \leq 2\}$. On définit de même sur $P\lambda$ la base $\{N_\ell^\lambda; \ell=1, \dots, L+4\}$ de $S_{P\lambda}$. La base de S_P est alors $\{N_k^\tau N_\ell^\lambda; k=1, \dots, K+4; \ell=1, \dots, L+4\}$, (voir figure 1): $\forall f \in S_P, \forall r, s \in [1, n_1] \times [1, n_2], f(r, s) = \sum_k \sum_\ell \beta_{k\ell} N_k^\tau(r) N_\ell^\lambda(s)$ et en particulier:

$$\mu_{rs} = \sum_{k=1}^{K+4} \sum_{\ell=1}^{L+4} N_k^\tau(r) N_\ell^\lambda(s) \beta_{k\ell} \quad (25)$$

III.2. Mise en oeuvre numérique.

Dans ce paragraphe, un tableau de nombres, disons t , a deux représentations possibles: la forme matricielle notée $[t]$ et la forme vectorielle notée $\underline{t}$. $\underline{t}$ est remplie suivant l'ordre lexicographique ligne à ligne sur $[t]$, c'est à dire la première ligne de $[t]$ est placée au début de $\underline{t}$, puis la deuxième, etc... Avec ces conventions, le système (25) peut

s'écrire $\underline{\mu} = X\underline{\beta}$, où X est une matrice $(n_1 n_2, (K+4)(L+4))$ dont la ligne $(r-1)n_2 + s$ est le vecteur $((N_k^\tau(r) N_\ell^\lambda(s); \ell=1, \dots, L+4); (k=1, \dots, K+4))$. Même pour des champs de petite dimension un tel système est gigantesque et donc impraticable. Il faudra donc essayer d'utiliser la propriété de séparabilité de la base des B-splines bicubiques qui permet de réécrire (25) suivant:

$$[\mu] = N^\tau [\beta] N^{\lambda'} \quad (26)$$

où $N_{rk}^\tau = N_k^\tau(r)$ et $N_{s\ell}^\lambda = N_\ell^\lambda(s)$ sont des matrices bandes blocs de largeur 4. Le modèle (24) est:

$$[z] = N^\tau [\beta] N^{\lambda'} + [y]$$

$$\sum_{(i,j) \in \mathcal{D}} \sum_{\varphi} \phi_{i,j} y_{r-1,s-j} = e_{r,s} \quad (27)$$

où $\mathcal{D} \subset \{(i,j) \geq 0\}$ avec $\phi_{0,0}=1$. L'estimation des ϕ et β s'obtient en maximisant l'approximation du logarithme de la vraisemblance décrite au paragraphe II:

$$L(\phi, \beta^2, \sigma^2 | z) = -n \log \sigma^2 - \sigma^{-2} S(\phi, \beta)$$

$$S(\phi, \beta) = \sum_{(r,s) \in \mathcal{D}} \sum_{\varphi} \left[\sum_{(i,j) \in \mathcal{D}} \phi_{i,j} (z_{r-1,s-j} - (X\beta)_{r-1,s-j}) \right]^2 \quad (28)$$

où maintenant on travaille conditionnellement au $\mathcal{D}$ -bord observé. Dans ce cas, l'expression qui suit (10):

$$S(\phi, \beta) = n \sum_{(r,s) \in \mathcal{D}} \sum_{\varphi} G_{k,i} \left(\sum_{(i,j) \in \mathcal{D}} \phi_{i,j} \phi_{i+k,j+i} \right)$$

est une écriture exacte. Sous forme matricielle, il vient:

$$S(\phi, \beta) = [1, \underline{\phi}'] \begin{bmatrix} \sum \sum y_{r,s}^2 & d' \\ d & D \end{bmatrix} \begin{bmatrix} 1 \\ \underline{\phi} \end{bmatrix}$$

où les éléments de d et D sont proportionnels à l'expression:
 $\sum_{r,s} (z - X\beta)_{r-1, s-j} (z - X\beta)_{r-k, s-l}$.

La minimisation de $S(\phi, \beta)$ en (32) est un problème d'optimisation non linéaire. Cependant à ϕ fixé, l'estimation de β est la solution des moindres carrés généralisés (MCG) et à β fixé, l'estimation de ϕ est $-D^{-1}d$. Ainsi, est-on conduit à une méthode de relaxation, Céa [5] [6], déjà proposée par Ord [15] dans le cadre de certains modèles de régression spatiale. La procédure est itérative: à l'étape (1), $\hat{\phi}^{(1)} = 0$ et $\hat{\beta}^{(1)}$ est l'estimateur des moindres carrés ordinaires (MCO); à l'étape (p), on minimise $S(\phi, \hat{\beta}^{(p-1)})$, puis $S(\hat{\phi}^{(p)}, \beta)$. Ici, cette procédure converge vers un minimum local car à chaque étape on minimise une fonction quadratique et par conséquent la suite décroissante et bornée résultante converge vers une limite [15]; cependant nous ne sommes pas assurés d'un minimum global.

Il reste à résoudre efficacement la minimisation de $S(\hat{\phi}^{(p)}, \beta)$ à l'étape (p) dont la solution est celle des MCG. Si on réécrit (27) sous forme vectorielle:

$$\underline{z} = \underline{\mu} + \underline{y} \quad \hat{\mathcal{P}}(B) y_{r,s} = \varepsilon_{r,s} \quad (29)$$

alors cette solution est aussi celle des MCO relative au "système blanchi": $\hat{\mathcal{P}}(B) z_{r,s} = \hat{\mathcal{P}}(B) \mu_{r,s} + \varepsilon_{r,s}$. L'opérateur $\mathcal{P}(B)$ qui transforme le bruit coloré y en un bruit blanc ε , est défini par (6.2). Plus précisément:

$$\hat{\mathcal{P}}(B) = \sum_{(I,J) \in \mathcal{D}} \sum_{\varphi} \hat{\phi}_{I,J}^{(p)} B_1^I B_2^J, \quad \hat{\phi}_{0,0} = 0.$$

La "moyenne blanchie" $\mu_{b_{r,s}} = \hat{\mathcal{P}}(B) \mu_{r,s}$ est définie par:

$$\mu_{b_{r,s}} = \sum_{(I,J) \in \mathcal{D}} \sum_{\varphi} \hat{\phi}_{I,J}^{(p)} \mu_{r-1, s-j} = \sum_{(K,I)} \sum_{(L,J)} \left(\sum_{(I,J)} \hat{\phi}_{I,J}^{(p)} X_{(r-1, s-j), (k, l)} \right) \beta_{k, l}$$

$$= \sum_{(k,l)} \sum_{(r,s)} (\hat{\varphi}(B)X_{(r,s),(k,l)}) \beta_{k,l} = (\hat{\varphi}(B)X)\underline{\beta}.$$

On note donc abusivement $\underline{\mu}_b = (\hat{\varphi}(B)X)\underline{\beta}$. Malheureusement, la propriété de séparabilité (26) ne s'applique plus à $\underline{\mu}_b$ et on est donc tenu de travailler vectoriellement. Or, des essais numériques menés sur des champs simulés ont montré l'inconsistance de cette démarche, même pour des champs de petite taille. Ceci est dû au mauvais conditionnement de la matrice explicative $\hat{\varphi}(B)X$ quand la structure d'autocovariance devient forte. (Pour ces essais, la transformation de Householder et les programmes de Lawson [12] ont été utilisés).

Nous proposons donc l'artifice de calcul suivant qui permet d'utiliser la séparabilité. En posant $\underline{\mu}_b = X\underline{\beta}_b$, l'espérance blanchie est estimée par les MCO sur le modèle:

$[\hat{\varphi}(B)z] = N^{\tau} [\beta_b] N^{\lambda'} + [\varepsilon]$. La solution $\hat{\beta}_b$ est donnée par le système: $N^{\tau'} N^{\tau} [\beta_b] N^{\lambda'} N^{\lambda} = N^{\tau'} [\hat{\varphi}(B)z] N^{\lambda}$, d'où $[\hat{\mu}_b] = N^{\tau} [\hat{\beta}_b] N^{\lambda'}$. Pour obtenir l'estimation $\hat{\mu}$ de l'espérance μ , il suffit de "colorer" $\hat{\mu}_b$ suivant la récurrence:

$$\hat{\mu}_{r,s} = \sum_{\substack{(i,j) \neq 0 \\ (i,j) \in \mathcal{D}_\varphi}} \hat{\varphi}_{i,j}^{(p)} \hat{\mu}_{r-1,s-j} + \hat{\mu}_{b,r,s} \quad (30)$$

III.3. Expériences et commentaires.

Un exemple de traitement d'un champ obtenu par simulation est maintenant présenté. Les résidus sont régis par un ARU suivant: $y_{r,s} = 0.7y_{r-1,s} + 0.3y_{r-1,s-1} + \varepsilon_{r,s}$, avec $\sigma=1$ et $1 \leq r,s \leq 64$; la moyenne du champ est $\mu_{r,s}=0$, $\forall r,s$, (figures 2 et 3). μ est alors modélisé par une fonction spline bicubique suivant (25) avec $K=L=2$, les noeuds étant également espacés. L'estimation des MCO de μ , c'est à dire l'étape (1) de la procédure de relaxation, est donnée en figure 5. L'écart quadratique moyen relativement à $\mu=0$ est EQM=1.41. On devine

sur ces figures que cette estimation suit les basses fréquences des résidus. La relaxation a été poursuivie en supposant connu l'ordre de l'ARU soit $\varphi_{1,0}$, $\varphi_{0,1}$ et $\varphi_{1,1}$ inconnus. L'algorithme a convergé à l'étape (3); l'estimation des MCG estimés de μ est portée en figure 4 et $\hat{\varphi}_{1,0}=0.03$, $\hat{\varphi}_{0,1}=0.68$, $\hat{\varphi}_{1,1}=0.27$, EQM=1.04. L'estimation des MCG avec $\varphi_{1,0}$, $\varphi_{0,1}$, $\varphi_{1,1}$ connus a aussi été faite: l'estimation $\hat{\mu}$ est très peu différente de la précédente et EQM=0.93. Ceci laisse croire que le minimum local atteint par la relaxation est en fait un minimum global pour cette expérience.

Ces résultats seraient à juger en tenant compte des intervalles de confiance associés, intervalles que l'on pourrait approximer grace aux lois asymptotiques du paragraphe II, [7]. Cependant pour certaines applications, par exemple la restauration d'image [7], où le but est l'affichage de la surface $\hat{\mu}$ nettoyée du bruit y , on est amené à se demander ce qu'apportent les MCG pour le lissage d'un champ par régression spatiale polynomiale. Pour cela, rappelons un résultat sur l'estimation du paramètre β du modèle de régression $z=X\beta+y$ où les résidus y sont autocorrélés de matrice de variance covariance Σ : quand les colonnes de X sont des combinaisons linéaires des vecteurs propres de Σ alors les estimateurs des MCO et des MCG de β coïncident, Durbin [4]. Quand y est une série chronologique stationnaire de taille n , ce résultat est catastrophique pour la recherche d'une tendance polynomiale. En effet, dans ce cas Σ est approximativement une matrice de Toeplitz et donc ces vecteurs propres sont approximativement les vecteurs de Fourier de la (n,n) matrice de Fourier $\mathcal{F}$. Ainsi, lorsque les colonnes de X représentent des polynômes pouvant être approchés par des séries de Fourier, on se trouve dans la situation rappelée. Quand y est un processus spatial stationnaire de taille (n,n) , cette remarque ne s'étend pas sauf dans le cas particulier où y est un ARU factorisant:

$$y_{r,s} + \phi_1 y_{r-1,s} + \phi_2 y_{r,s-1} - \phi_1 \phi_2 y_{r-1,s-1} = \varepsilon_{r,s}$$

Pour ce processus, la matrice W des vecteurs propres de Σ s'écrit approximativement suivant un produit matriciel de Kroneker: $W = \mathcal{F} \otimes \mathcal{F}$, où $\mathcal{F}$ est la (n,n) matrice de Fourier. Cette propriété entraîne que si x est une combinaison linéaire des vecteurs propres: $\underline{x} = W \underline{\psi} \iff [x] = \mathcal{F}[\psi] \mathcal{F}'$, soit:

$$x_{r,s} = n^{-1} \sum_{k=0}^{n-1} \sum_{\ell=0}^{n-1} \psi_{k,\ell} e^{-I(r\omega_k + s\omega_\ell)} \quad (31)$$

où $\omega_k = 2\pi k/n$. Ainsi, lorsque les colonnes de X représentent des surfaces polynomiales pouvant être approchées par une série de Fourier suivant (31), on se trouve encore dans la situation rappelée.

Pour terminer ces notes, citons une intervention de Diggle [3] sur le problème du lissage d'une série chronologique par une spline non-paramétrique. A partir d'un processus AR(1) simulé, il présente une estimation de la tendance sensiblement meilleur que celle estimée sans prise en compte de l'autocorrélation. Cette démarche semble être une alternative à la notre tout au moins dans le cas des séries chronologiques.

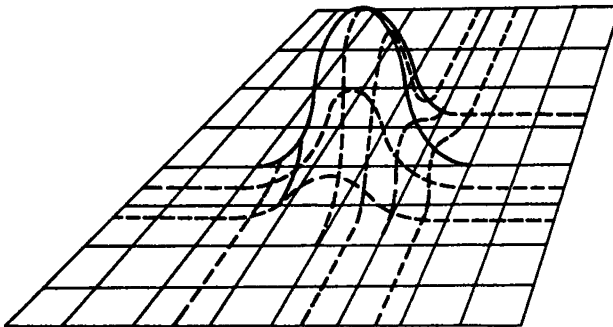
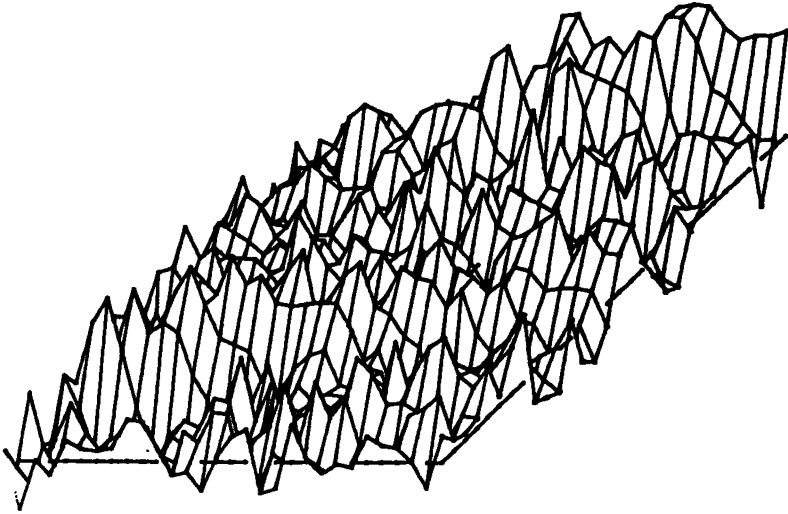
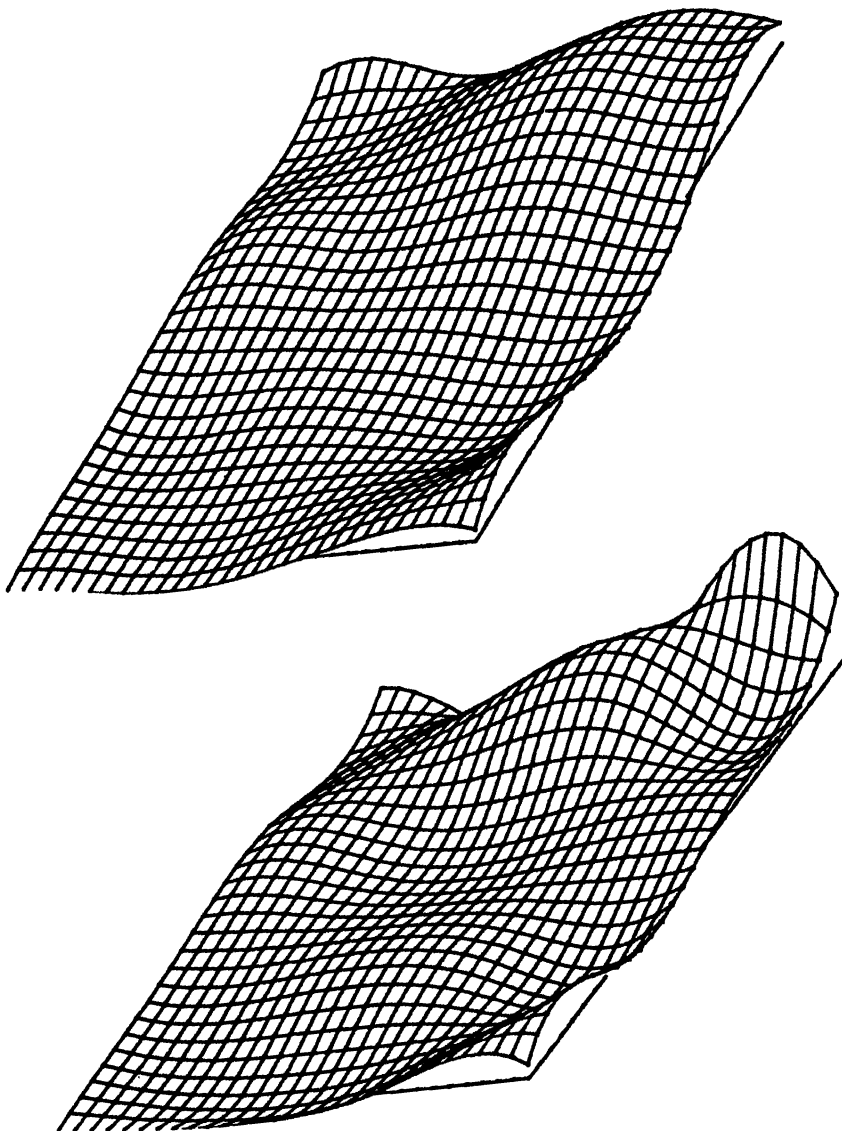


Figure 1.



Figures 2 et 3.



Figures 4 et 5.

BIBLIOGRAPHIE.

- [1] ANDERSON T. W., The statistical analysis of times series, John-Wiley, 1971.
- [2] DeBOOR C., Bicubic spline interpolation, J. Math. Phys., 41, p212-218, 1962.
- [3] DIGGLE P. J., Discussion on Silverman's paper: Some aspects of spline smoothing approach to non-parametric curve fitting, JRSS B, 47, 1, p28-29, 1985.
- [4] DURBIN J., Present position and potentiel developments: Some personal views. Time series analysis. JRSS A, 147, Part 2, p161-173, 1984.
- [5] CEA J., Optimisation, Dunod, 1971.
- [6] CEA J. et GLOWINSKI R., Sur des méthodes d'optimisation par relaxation, RAIRO, R-3, p5-32, 1973.
- [7] CHALMOND B., Restauration stochastique et analyse d'une surface recouverte par une texture, Actes du 4eme congrès de Reconnaissance des Formes et d'Intelligence Artificielle, Tome 1, p1-14, AFCET-INRIA, 1984.
- [8] CHALMOND B. Régression avec résidus spatialement autocorrélés: estimation par le maximum de vraisemblance, Journée de Statistique de l'ASU, 1984.
- [9] GUYON X., Parameter estimation for stationary proces on a d-dimensinal lattice, Biometrika, 69, 1, p95-105, 1982.
- [10] GUYON X., Champs stationnaires sur $\mathbb{Z}^2$: modèles, statistiques et simulation, Prépublication Orsay T21, 1985.
- [11] KASHYAP R.L. et CHELLAPA R., Estimation and choice of neighbors in spatial interaction models in image, IEEE Trans. Inf. Theory, IT-29, 1, p60-72, 1983.

- [12] LAWSON C. et HANSON R., Solving least squares problems, Prentice-Hall, 1974.
- [13] MARDIA K.V. et MARSHALL R.J., Maximum likelihood estimation of models for residual covariance in spatial regression, Biometrika, 71, 1, p136-146, 1984.
- [14] MARTIN R.J., Some aspects of experimental design and analysis when errors are correlated. Biometrika, 69, 3, p497-512, 1982.
- [15] ORD K., Estimation methods for models of spatial interaction. JASA, 70, p120-126, 1975.
- [16] PIERCE D., Least squares estimation in the regression model with autoregressive - moving average errors, Biometrika, 58, 2, p299-312., 1971.
- [17] PRENTER P., Splines and variational methods, Jonh-Wiley, 1971.
- [18] WALD A., Note on the consistency of the maximum likelihood estimates, Ann. Math. Stat., 20, p595-601, 1949.
- [19] WALKER A. M., Asymptotic properties of the least squares estimates of parameter of the spectrum of a stationary non-deterministic time series, J. Austr. Math. Soc., 35, 363-384, 1964.
- [20] WHITTLE P., On stationary processes in the plane, Biometrika, 41, p434-449, 1954.