

# STATISTIQUE ET ANALYSE DES DONNÉES

RÉGIS KALAJDJIAN

**Efficacités comparées de certaines méthodes de prédiction  
pour un ARMA perturbé**

*Statistique et analyse des données*, tome 6, n° 3 (1981), p. 47-64

[http://www.numdam.org/item?id=SAD\\_1981\\_\\_6\\_3\\_47\\_0](http://www.numdam.org/item?id=SAD_1981__6_3_47_0)

© Association pour la statistique et ses utilisations, 1981, 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/>

EFFICACITES COMPAREES  
DE CERTAINES METHODES DE PREDICTION  
POUR UN ARMA PERTURBE

Régis KALAJDZIAN

Ingénieur de l'Ecole Polytechnique

Elève de l'E.N.S.A.E.

Résumé : Pour un  $AR(1)$  perturbé additivement, on calcule selon la méthode "de Box et Jenkins" et selon d'autres procédés non paramétriques, l'estimateur du prédicteur théorique. On compare alors les erreurs quadratiques des différentes méthodes sur le prédicteur. D'autres cas plus complexes d'ARMA perturbés sont envisagés par des simulations.

Abstract : We perturb an  $AR(1)$  process by another additional one to estimate the best forecast by the "Box & Jenkins" method and by other non parametric ones. Then we compare the different mean square errors. By simulations, we make a study of other more complex perturbed ARMA.

Mots clés : ARMA, Prédicteur, Estimateur.

Définitions préliminaires :

Tous les processus envisagés dans cet article seront réels, centrés.

a) Un processus  $(u_t, t \in \mathbb{Z})$  est un bruit blanc s'il est de variance constante lorsque  $t$  décrit  $\mathbb{Z}$ , et si  $\text{cov}(u_t, u_s)$  est nulle dès que  $t$  et  $s$  sont des dates différentes.

b) Un processus  $(x_t, t \in \mathbb{Z})$  est auto-régressif d'ordre  $p$ , noté  $AR(p)$ , s'il existe un entier positif  $p$  et une famille de réels  $(\pi_j, j \in \mathbb{Z})$  tels que

$$x_t = \sum_{j=1}^p \Pi_j x_{t-j} + u_t \text{ pour tout } t \text{ de } \mathbb{Z}, \text{ où } \Pi_p \text{ est non nul et } (u_t, t \in \mathbb{Z})$$

est un bruit blanc.

c)  $(x_t, t \in \mathbb{Z})$  est une moyenne mobile d'ordre  $q$ , notée  $MA(q)$ , s'il existe un entier positif  $q$ , une famille de réels  $(\psi_j, j \in \mathbb{Z})$  telle que  $\psi_0=1$  et  $\psi_q$  soit non nul, et un bruit blanc  $(u_t, t \in \mathbb{Z})$ , vérifiant pour tout  $t$  de  $\mathbb{Z}$  :

$$x_t = \sum_{j=0}^q \psi_j u_{t-j} .$$

d)  $(x_t, t \in \mathbb{Z})$  est un  $ARMA(p,q)$  s'il existe deux entiers positifs  $p$  et  $q$ , et une famille de réels  $(\Pi_j, j \in \mathbb{Z})$  telle que  $\Pi_p$  soit non nul et que le processus  $(y_t, t \in \mathbb{Z})$  défini par  $y_t = x_t - \sum_{j=1}^p \Pi_j x_{t-j}$ , pour tout  $t$ , soit une  $MA(q)$ .

## 0 - INTRODUCTION

La méthode de prédiction pour un  $ARMA(p,q)$ , dite "de Box et Jenkins", se compliquant rapidement lorsque le nombre de paramètres à estimer augmente, se voit concurrencée par des méthodes non paramétriques, surtout si le processus considéré n'est pas rigoureusement un  $ARMA$ . Les résultats ci-après font état des robustesses relatives des différentes méthodes appliquées à un  $ARMA$  perturbé. L'étude est générale pour le cas d'un  $AR(1)$  (§ 1, 2, 3, et 4), et se fait sur des cas particuliers pour des  $ARMA$  moins simples (§ 5).

Remarque : - Les densités apparaissant dans les calculs sont toutes majorées par  $\theta_1$  et minorées par  $\theta_2$  ( $\theta_1$  et  $\theta_2$  étant des réels positifs).

- Dans la phase calculatrice de l'étude, comparer les robustesses signifiera comparer les erreurs de l'estimateur sur le prédicteur ; les simulations montreront ce que valent les estimateurs par rapport au processus réel.

## 1 - PREDICTION ET ESTIMATION PAR LA METHODE DE BOX ET JENKINS

1.1 - Données et hypothèses : Soit  $(x_t, t \in \mathbb{Z})$  un processus (centré à valeurs dans un compact  $[-X, X]$  de  $\mathbb{R}$ . Il vérifie  $x_{t+1} = \rho x_t + u_{t+1} + \varepsilon_{t+1}$  pour

tout  $t$  de  $\mathbb{Z}$  ;  $\rho$  est une constante de  $]0,1[$  ,  $(u_t, t \in \mathbb{Z})$  est un bruit blanc de variance  $\sigma^2$ , à valeurs dans  $[-U,U]$  ,  $(\varepsilon_t, t \in \mathbb{Z})$  est un processus additionnel, sans lequel  $(x_t)$  serait markovien stationnaire, et prenant ses valeurs dans  $[-\varepsilon, \varepsilon]$  où  $\varepsilon$  est un réel positif.

$(\varepsilon_t)$  vérifie en outre pour tout couple  $(s,t)$  de  $\mathbb{Z}^2$  :

1)  $\text{cov}(\varepsilon_s, u_t) = 0$ . Ce qui se produit dès que les deux processus  $(u_t)$  et  $(\varepsilon_t)$  sont indépendants.

2)  $|\text{corr}(\varepsilon_s, \varepsilon_t)| < K \rho^{\alpha|s-t|}$  où  $\alpha$  et  $K$  sont des constantes.  $\alpha > 1$  et  $K > 1$ . En d'autres termes, lorsque  $s$  et  $t$  sont des instants assez éloignés l'un de l'autre  $\varepsilon_s$  et  $\varepsilon_t$  sont peu corrélés.

1.2 - Notations : Dans ce paragraphe, on suppose connue et indépendante de  $t$  la quantité notée  $\tau^2 = \text{Var}(x_t)$ . On peut appliquer la méthode de Box et Jenkins à condition de négliger la perturbation, donc de l'annuler. Alors le prédicteur est  $E(x_{t+1}/x_t)$ , noté  $\hat{x}_t$ . L'estimateur  $x_t^*$  du prédicteur est, comme dans le cas markovien stationnaire,  $x_t^* = \hat{\rho} x_t$  où  $\hat{\rho}$  est un estimateur de  $\rho$ . Le problème consiste à majorer  $E((\hat{x}_t - x_t^*)^2)$ , erreur quadratique moyenne de l'estimation.  $\rho$  est estimé à l'aide de  $n$  valeurs du processus.

1.3 - Résultat : Un calcul donne :

$$E((\hat{x}_t - x_t^*)^2) \leq 2\varepsilon^2 + \frac{4X^2}{\tau^4} K^2 \varepsilon^4 C_1 + \frac{2X^2(1-\rho^2)}{n}$$

$$\text{où } C_1 = A_1 + B_1 ; A_1 = \frac{\rho}{1-\rho^{\alpha-1}} \left( \frac{1}{1-\rho^2} - \frac{\rho^{2(\alpha-1)}}{1-\rho^{\alpha+1}} \right) ; B_1 = \frac{\rho^{2+\alpha}}{(1-\rho^2)(1-\rho^{1+\alpha})}$$

La majoration trouvée est donc la somme d'une grandeur croissant avec l'amplitude de la perturbation, et faisant intervenir  $\varepsilon^2$  et  $\varepsilon^4$ , et d'un terme fonction de  $\frac{1}{n}$ , petit si l'estimation de  $\rho$  est précise. Le facteur  $C_1$  pouvant devenir très grand lorsque  $\rho$  s'approche de 1, c'est peut-être par le terme en  $\varepsilon^4$

que la majoration est trop lâche.

## 2 - PREDICTION ET ESTIMATION PAR LA METHODE DE LA SUBDIVISION

2.1 - Données et hypothèses : Nous considérons toujours le même processus,  $(x_t, t \in \mathbb{Z})$ . La loi de  $x_t$  est supposée indépendante de  $t$ . Il existe un processus markovien stationnaire,  $(y_t, t \in \mathbb{Z})$ , naturellement associé à  $(x_t)$  :  $y_t = \rho y_{t-1} + u_t$  pour tout  $t$  de  $\mathbb{Z}$ .

Sur ce nouveau processus, nous faisons une hypothèse de "mélangeance" tirée de [2] :

Il existe une suite  $(\phi_n, n \in \mathbb{N})$  réelle, telle que  $\phi_0 = 1$ , et vérifiant pour tout couple  $(i, h)$  de  $\mathbb{N}^2$  :

$|\text{cov}(A_i, B_{i+h})| \leq 2E|A_i| b \phi_h$ , où  $b$  majore la v.a.r.  $|B_{i+h}|$ , dès que

$A_i = \alpha(y_0, \dots, y_i)$  et  $B_{i+h} = \beta(y_{i+h}, y_{i+1+h}, \dots)$

Cette hypothèse qui fait jouer à la date  $t=0$  un rôle privilégié, suppose que le "début" du processus a un sens.

On réalise alors une subdivision  $\mathcal{S}_k$  de  $[-X, X]$ , réunion de  $k$  segments disjoints, tels que  $P(x_t \in J) = 1/k$  dès que  $t$  appartient à  $\mathbb{Z}$  et  $J$  à  $\mathcal{S}_k$ ,  $P$  désignant la loi de  $x_t$ .

Notre prédicteur restant  $\hat{x}_t = E(x_{t+1}/x_t)$ , l'estimateur se définit à l'aide de la subdivision :

$$\tilde{x}_t = \frac{\sum_{i=0}^{t-1} \sum_{j \in \mathcal{S}_k} x_{i+1} l_j(x_i) l_j(x_t)}{\sum_{i=0}^{t-1} \sum_{j \in \mathcal{S}_k} l_j(x_i) l_j(x_t)} \quad \text{si le dénominateur est non nul}$$

$$\tilde{x}_t = 0 \quad \text{sinon}$$

2.2 - Résultat : La majoration s'effectue comme suit

$$\tilde{x}_t - \hat{x}_t = \frac{A_t}{D_t} + \frac{B_t}{D_t}$$

$$\text{où } A_t = \frac{1}{t} \sum_{i=0}^{t-1} \sum_{j \in \mathcal{S}_k} (x_{i+1} - E(x_{i+1}/x_i)) 1_j(x_i) 1_j(x_t)$$

$$B_t = \frac{1}{t} \sum_t \sum_i \sum_j (E(x_{i+1}/x_i) - E(x_{t+1}/x_t)) 1_j(x_i) 1_j(x_t)$$

$$D_t = \frac{1}{t} \sum_i \sum_j 1_j(x_i) 1_j(x_t)$$

$$\text{Il vient : } \frac{1}{2} E((\tilde{x}_t - \hat{x}_t)^2) \leq E \frac{A_t^2}{D_t^2} + E \frac{B_t^2}{D_t^2}$$

le processus  $(y_t)$  sert à "approcher"  $(x_t)$  car  $|x_t - y_t| < \varepsilon$  ; certains termes rectangulaires intervenant dans les expressions de  $E \frac{A_t^2}{D_t^2}$  ou  $E \frac{B_t^2}{D_t^2}$  sont alors majorables grâce à l'hypothèse de mélangeance vérifiée par  $(y_t)$ .

En notant  $\omega = \frac{\varepsilon}{1-\rho}$  et  $\Omega_k = \inf(\frac{1}{\theta_2^k}, \omega)$ , on trouve :

$$\begin{aligned} E((\tilde{x}_t - \hat{x}_t)^2) &\leq 2\left(\frac{\rho}{k} + 2\varepsilon\right)^2 + 2M \\ M &= 8X^2 \left(1 + \frac{\theta_1}{\theta_2} + \theta_1\right) k \Omega_k + 16X^2 \theta_1^2 (k \Omega_k)^2 + 16X^2 \theta_1 k^2 \Omega_k \omega \\ &+ 4Xk\varepsilon + 16X^2 k\varepsilon\omega \\ &+ 32X^2 \frac{\theta_1}{1-\rho} (k \Omega_k^2 + k\omega\Omega_k) + 16X^2 / (k(1-\rho) \theta_2) \\ &+ 2X \sigma \frac{k}{\sqrt{t}} + 8X^2 (1+2k \theta_1 \Omega_k) \left(\frac{k}{t}\right) \sum_{h=0}^t \phi_h \end{aligned}$$

Il est alors utile de supposer la série  $(\phi_h)$  convergente. Les termes proportionnels à  $\varepsilon$  ou  $\varepsilon^2$  prennent en compte l'intensité de la perturbation, comme dans la majoration paramétrique ; ceux tendant vers 0 avec  $\frac{1}{t}$  montrent qu'il est plus précis de calculer la moyenne qu'est  $\hat{x}_t^y$  avec un grand nombre de points  $x_i$ . Mais  $\hat{x}_t^y$  dépend aussi du critère par lequel ces points sont sélectionnés. Avec une subdivision trop fine, peu de points servent à construire  $\hat{x}_t^y$  ; si, au contraire, la subdivision est trop large, on perd évidemment en précision. Cela porte à croire à l'existence d'une ou plusieurs subdivisions optimales, suggérée par la forme même de notre majorant dont certains termes sont proportionnels à  $k$  ou  $1/k$ .

Lorsque  $k$  est très grand, la majoration trouvée ne veut plus rien dire, car, toute chose égale par ailleurs, elle tend vers l'infini, alors que simultanément  $\hat{x}_t^y$  tend vers 0, et l'erreur vers la variance du prédicteur théorique.

### 3 - METHODE DU NOYAU REGULARISANT

#### 3.1 - Données

Les hypothèses sont les mêmes que dans le paragraphe précédent ; la subdivision nous est maintenant inutile.  $\hat{x}_t^y$  est exprimé comme barycentre des  $x_{i+1}$  affectés de poids d'autant plus faibles que  $x_i$  est éloigné de  $x_t$ .

$$x_k = \frac{\sum_{i=0}^{t-1} x_{i+1} F_\lambda(x_i - x_t)}{\sum_{i=0}^{t-1} F_\lambda(x_i - x_t)} \quad \text{où } F_\lambda(z) = \exp(-z^2/2\lambda^2)$$

$\lambda$  est censée tendre vers 0. Ce noyau est dit gaussien.

#### 3.2 - Résultat : La méthode de majoration ressemble à la précédente.

$$E((\tilde{x}_t - \hat{x}_t)^2) \leq 2M + 2N \quad \text{où}$$

$$M = 8\epsilon^2 + 4(f(\lambda))^2 (M' + U^2 M'')$$

$$M' = U^2 \omega_1 \sqrt{2\pi} \lambda/t + 2\omega U^2 \theta_1 \sqrt{2} + 2U^3 \theta_1 \sqrt{2} / (t(1-\rho))$$

$$M'' = 2X \theta_1 \sqrt{2\pi} \lambda/t + 32 \theta_1 \omega + 4 \theta_1 \sqrt{2\pi} (\lambda/t) \sum_{h=0}^t \phi_h$$

$$N = 4X \epsilon + 2X f(\lambda) (2\rho \theta_1 \lambda^2 + 2X\rho (2X\theta_1 \sqrt{2\pi} \frac{\lambda}{t} + 32\omega \theta_1 + 4\theta_1 \sqrt{2\pi} \frac{\lambda}{t} \sum_{h=0}^t \phi_h)^{1/2})$$

$f(\lambda)$  est une quantité dépendant de  $\lambda$  et des paramètres supposés fixes,  $X$  et  $\theta_2$ . Lorsque  $\lambda$  tend vers 0,  $f(\lambda)$  est équivalente à  $1/(2X \lambda \theta_2^2 \sqrt{2\pi})$ .

Les deux facteurs  $\lambda$  et  $1/k$  jouant le même rôle, la majoration obtenue ressemble beaucoup à la précédente. Les termes prépondérants sont ici en  $\epsilon$  ou  $\epsilon^2$ , en  $1/\lambda$  ou  $1/\lambda^2$ , en  $1/t\lambda$  ou  $1/t\lambda^2$ . Là encore, nous sommes conduits à soupçonner l'existence d'un  $\lambda$  optimal, minimisant le majorant.

#### 4 - AUTRE TYPE DE PERTURBATION - ERREURS PARAMETRIQUE ET NON PARAMETRIQUE

4.1 - Données et hypothèses : Envisageons le cas d'un AR(2) que l'on traite comme s'il s'agissait d'un AR(1) :

$$x_{t+1} = \alpha x_t + \beta x_{t-1} + u_{t+1} \quad \text{pour tout } t \text{ de } \mathbb{Z}.$$

$(u_t, t \in \mathbb{Z})$  est un bruit blanc à valeurs dans  $[-U, U]$  ;  $\alpha$  et  $\beta$  sont des constantes réelles,  $(x_t)$  est à valeurs dans  $[-X, X]$ , et  $\text{Var } x_t = \tau^2$  est indépendante de  $t$ .

4.2 - Erreur paramétrique : Théoriquement,  $\hat{x}_t = \alpha x_t + \beta x_{t-1}$ , alors qu'on utilise  $x_t^* = \alpha x_t$ . Compte tenu des approximations faites dans [1] (Ch. 5), l'erreur peut se majorer par :



$$E((\hat{x}_t - x_t^*)^2) \leq 2\chi^2 \frac{1-\alpha^2}{n} + 2\beta^2 \tau^2$$

Ce résultat est d'une forme analogue à celle de la précédente majoration paramétrique. On retrouve un terme en  $1/n$  représentant la précision de l'estimation, et un terme dû à la perturbation.

#### 4.3 - Erreur non paramétrique : méthode de la subdivision

Nous appliquons les résultats exposés dans [3]. Nous supposons le processus  $(x_t)$  strictement stationnaire et, comme cela nous est possible avec ce processus linéaire, lui imposons de satisfaire l'hypothèse de mélangeance vérifiée par  $(y_t)$  tout à l'heure. L'estimateur est défini à l'aide de la subdivision  $\mathcal{D}_k$  de manière inchangée. Nous obtenons :

$$E((\hat{x}_t - \hat{x}_t^*)^2) \leq 16 \frac{\theta_1^2}{\theta_2} U^2 k \left(1 + \delta(k + \sum_{i=0}^{t-k} \phi_i)\right) / (t-k)$$

$$+ 2C k \theta_1 \left(k + \sum_{i=0}^{t-k} \phi_i\right) / (t-k) + 2\left(\frac{\alpha}{\theta_2 k} + 2\chi\beta\right)$$

Lorsque  $k$  augmente en restant petit devant  $t$ , les deux premiers termes sont de l'ordre de  $k^2/t$ .

Remarquons aussi que les deux majorations du § 4 ont des comportements comparables devant la perturbation.

## 5 - SIMULATIONS

### 5.1 - AR(1) perturbé

Soit  $I = \{0, 1, \dots, 200\}$  et  $(x_t, t \in I)$  tel que  $x_{t+1} = \rho x_t + u_{t+1} + \varphi_{t+1}$  et  $x_0 = u_0$ , où  $(u_t, t \in I)$  et  $(\varphi_t, t \in I)$  sont des processus donnés, non corrélés.

$(u_t)$  est un bruit blanc,  $(p_t)$  ne l'est pas.

$\rho$  et  $\epsilon$  varient dans  $\{0.1, 0.5, 0.9\}$  et  $k$  dans  $\{1, 2, 5, 10, 20, 30, 50\}$  ;  $\lambda$  est inversement proportionnel à  $k$ . On peut ainsi comparer dans de nombreux cas l'efficacité des différentes méthodes en calculant la moyenne empirique des  $(\hat{x}_t - \hat{x}_t^v)^2$ ,  $(x_{t+1} - \hat{x}_t^v)^2$ ,  $(\hat{x}_t - x_t^*)^2$ ,  $(x_{t+1} - x_t^*)^2$  pour  $t$  supérieur à 100 (on choisit  $n = 100$ ).

Résultats :

Pour une perturbation faible devant  $\rho$ , la méthode paramétrique est la meilleure, mais celle du noyau n'est pas ridicule. La méthode de la subdivision est nettement devancée. Quand la perturbation est forte, tous les résultats sont du même ordre. Si les méthodes non paramétriques sont stables quand la perturbation s'amplifie, l'erreur paramétrique, elle peut varier pour  $\rho$  donné dans un rapport de 1 à 50. Enfin, le  $k$  optimal croit avec  $\rho$  et  $\epsilon$ .

On présente le cas particulier où  $\rho = 0.9$  et  $\epsilon = 0.5$  (perturbation non négligeable) ;  $k = 2$  donne le  $\lambda$  optimal.

Graphe pour  $t > 150$ .

Légende : $x_{t+1}$	_____	
$x_t^*$	— — — — —	erreur sur prédicteur = $1,8 \cdot 10^{-2}$ erreur sur processus = 0,27 ; $\hat{\rho} = 0,93$
$\hat{x}_t^v$ (gaussien)	-----	erreur sur prédicteur = $4,9 \cdot 10^{-2}$ erreur su processus = 0,3

5.2 - AR(4) perturbé

$x_{t+1} = 0.8 x_t - 0.4 x_{t-1} - 0.2 x_{t-2} - 0.1 x_{t-3} + u_{t+1} + \epsilon p_{t+1}$ , pour tout  $t$  de  $I$  ;  $\epsilon$  varie dans  $\{0, 0.03, 0.1, 0.5, 0.9\}$  et  $k$  dans  $\{1, 2, 4, 7, 10, 20, 30\}$ .

Dans la méthode de Box et Jenkins les paramètres sont déduits des équations de Yule-Walker ([1] Ch. 7), ce qui nécessite la résolution lourde d'un

systeme de Cramer 4 x 4.

Par la subdivision, l'estimateur devient :

$$\tilde{x}_t = \frac{\sum_{i=0}^{t-1} x_{i+1} \prod_{r=0}^3 \sum_j l_j(x_{i-r}) l_j(x_{t-r})}{\sum_i \prod_r \sum_j l_j(x_{i-r}) l_j(x_{t-r})}$$

tandis que le "noyau gaussien" donne :

$$\tilde{x}_t = \frac{\sum_i x_{i+1} \prod_r F_\lambda(x_{i-r} - x_{t-r})}{\sum_i \prod_r F_\lambda(x_{i-r} - x_{t-r})}$$

Donnons-nous aussi  $T_\mu(x) = \cos(\pi x/\mu) \quad 1_{|x| \leq \mu/2}$

Il suffit de remplacer  $F_\lambda$  par  $T_\mu$  où  $\mu$  est proportionnel à  $\lambda$  pour construire ce qu'on nommera la méthode du "noyau trigonométrique".

Résultats :

A nouveau la méthode paramétrique, très instable, donc peu "robuste", est concurrencée par le noyau gaussien (meilleur que le noyau trigonométrique), surtout si la perturbation est importante ; elle estime les paramètres d'autant mieux qu'ils sont de module grand.

Cas particulier : On choisit  $\epsilon = 0.5$ ,  $t$  supérieur à 100 et  $n = 100$ .

$x_{t+1}$  \_\_\_\_\_

$x_t^*$  \_\_\_\_\_

Paramètres estimés : 0.86, -0.37, -0.30  
 0.06  
 Erreur sur prédicteur :  $2 \cdot 10^{-2}$   
 Erreur sur processus : 0.27

$x_t$  (subd.) \_\_\_\_\_

Erreur sur prédicteur : 0.26 avec  $k = 2$   
 Erreur sur processus : 0.56

$\hat{x}_t^2$  (gaussien) ----- Erreur sur prédicteur :  $8.7 \cdot 10^{-2}$   
 Erreur sur processus : 0.38

### 5.3 - ARMA(2,2) perturbé

$x_{t+1} = 0.7 x_t - 0.5 x_{t-1} + 0.2 u_t - 0.3 u_{t-1} + u_{t+1} + \varepsilon_{t+1}$ , pour  $t$  dans  $I$ ,  
 $\varepsilon$  dans  $\{0, 0.03, 0.1, 0.5, 0.9\}$ ,  $k$  dans  $\{1, 2, 4, 7, 10, 20, 30\}$ .

Théoriquement,  $x_t^* = Ax_t + Bx_{t-1} + Du_t + Eu_{t-1}$ ; on peut alors supposer le bruit blanc connu et construire un estimateur paramétrique théorique, ou reconstituer  $(u_t)$  par  $u_{t+1} = x_{t+1} - x_t^*$  et le réintroduire dans l'expression de  $x_{t+1}^*$ : estimateur paramétrique "récurrent".

Les estimateurs non paramétriques s'expriment comme auparavant,  $r$  prenant les valeurs 0 et 1.

#### Résultats :

Les deux estimateurs paramétriques se valent. La stabilité de la méthode du noyau devant  $\varepsilon$  est encore plus nette que pour l'AR(4); et si  $\varepsilon$  est grand, le noyau atteint la précision paramétrique. Les  $k$  et  $\lambda$  optimaux sont peu sensibles à la perturbation.

#### Cas particulier :

On choisit  $\varepsilon = 0.5$ ; la légende est la même que précédemment; on fera figurer l'estimateur paramétrique récurrent.

$x_t^*$  paramètres estimés 0.58 -0.64 0.18 0.11  
 Erreur sur prédicteur :  $3.6 \cdot 10^{-2}$   
 Erreur sur processus : 0.28

$\hat{x}_t^2$  (subd.) Erreur sur prédicteur : 0.23  
 Erreur sur processus : 0.5

$\hat{x}_t^2$  (gaussien) Erreur sur prédicteur :  $8 \cdot 10^{-2}$   
 Erreur sur processus : 0.35

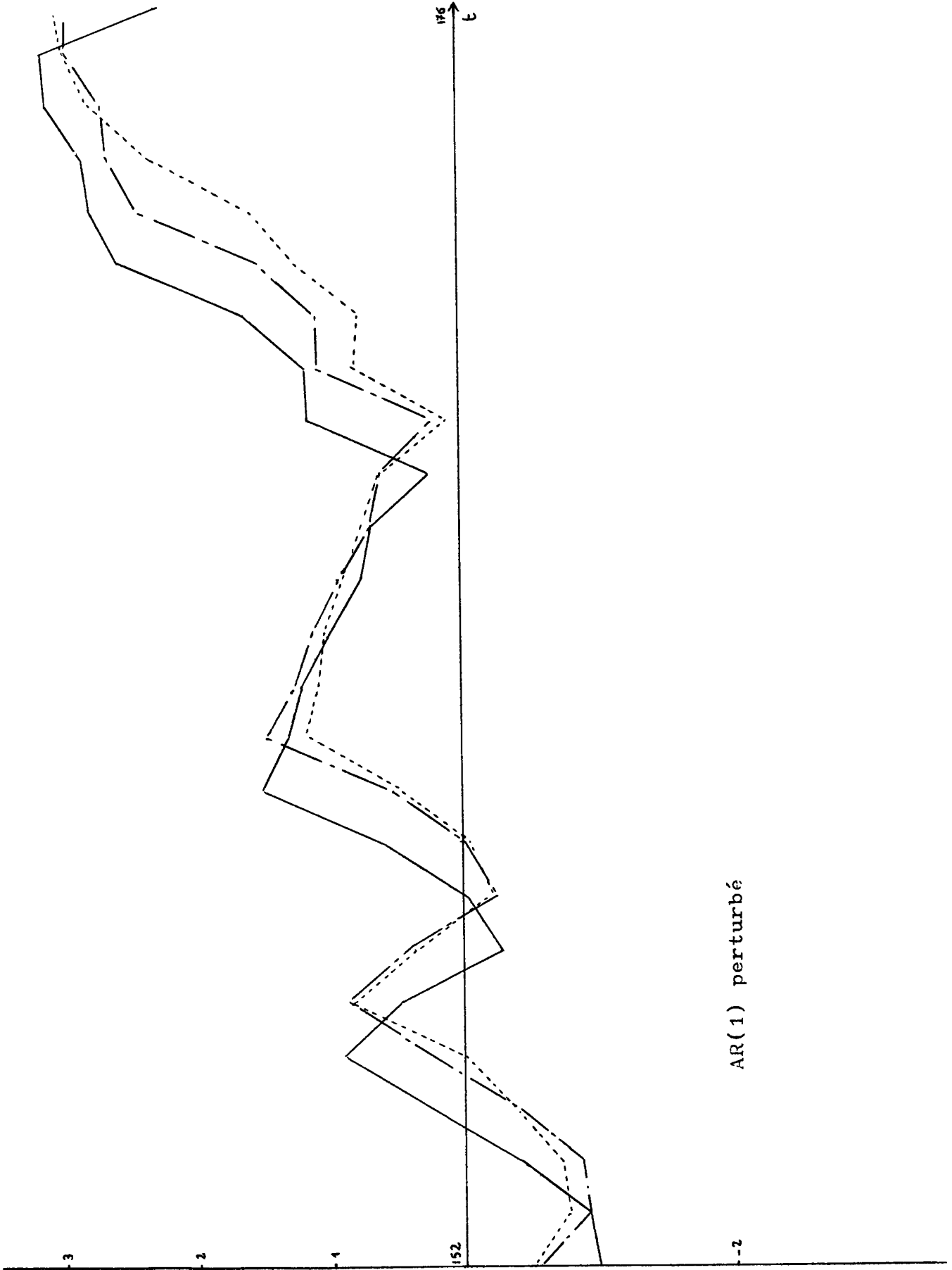
## 6 - CONCLUSION

Si, par une combinaison adéquate des paramètres manipulables ( $t$ ,  $k$ ,  $n$ ,  $\lambda$ ), le prédicteur peut être aisément approché, en revanche un bruit blanc de variance non négligeable empêche d'avoisiner le processus d'une erreur inférieure à celle-ci. Les simulations le montrent, l'erreur sur le processus reste dans tous les cas favorables, très supérieure à celle commise sur le prédicteur.

Voilà qui limite la portée de notre étude ; celle-ci aura prouvé toutefois que, dans les cas différant fortement de ceux étudiés par Box et Jenkins, la méthode du noyau s'avère efficace. Ses résultats sont très corrects dans les situations les plus favorables à la méthode paramétrique, et sa mise en oeuvre informatique toujours fort simple.

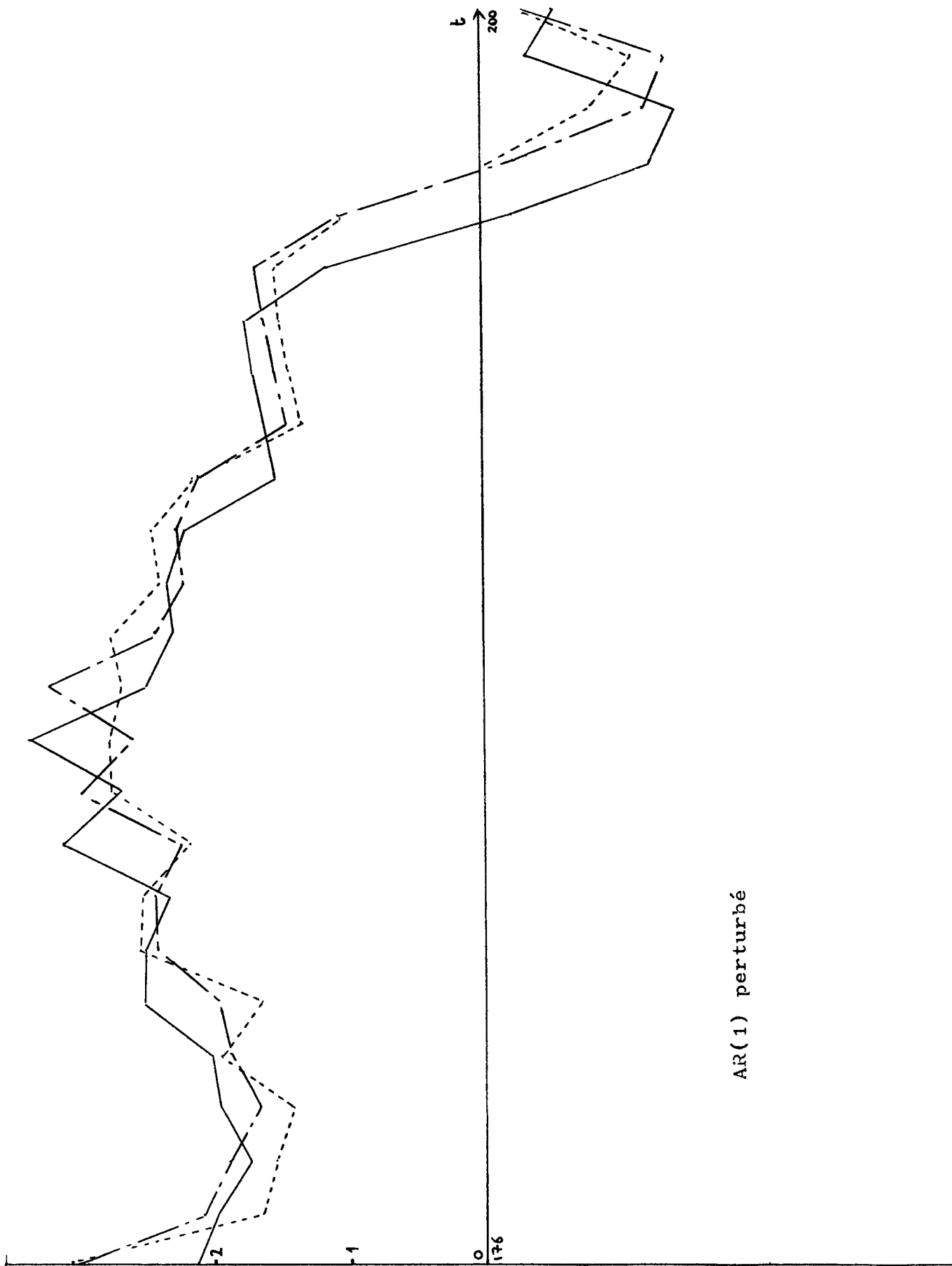
## B I B L I O G R A P H I E

- [ 1 ] Box & Jenkins : "Time Series Analysis" (Holden Day - 1979)
- [ 2 ] Billingsley : "Convergence of Probability Measures" (Wiley - 1968)
- [ 3 ] G. Collomb : "Prédiction non paramétrique : étude de l'erreur quadratique du Prédicogramme" (Université Paul Sabatier - Toulouse).

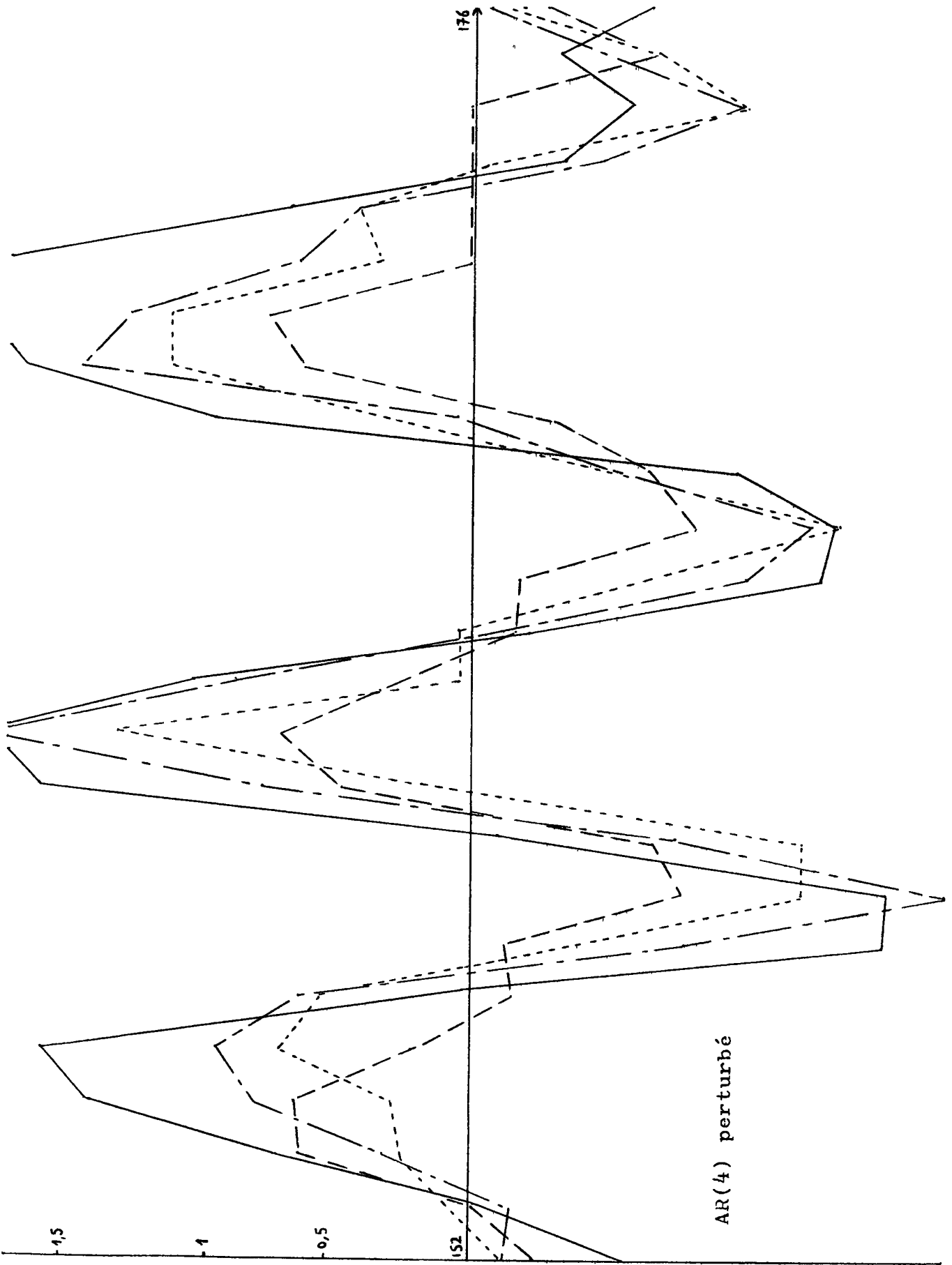


AR(1) perturbé

.60.



AR(1) perturbé



AR(4) perturbé

-1,5

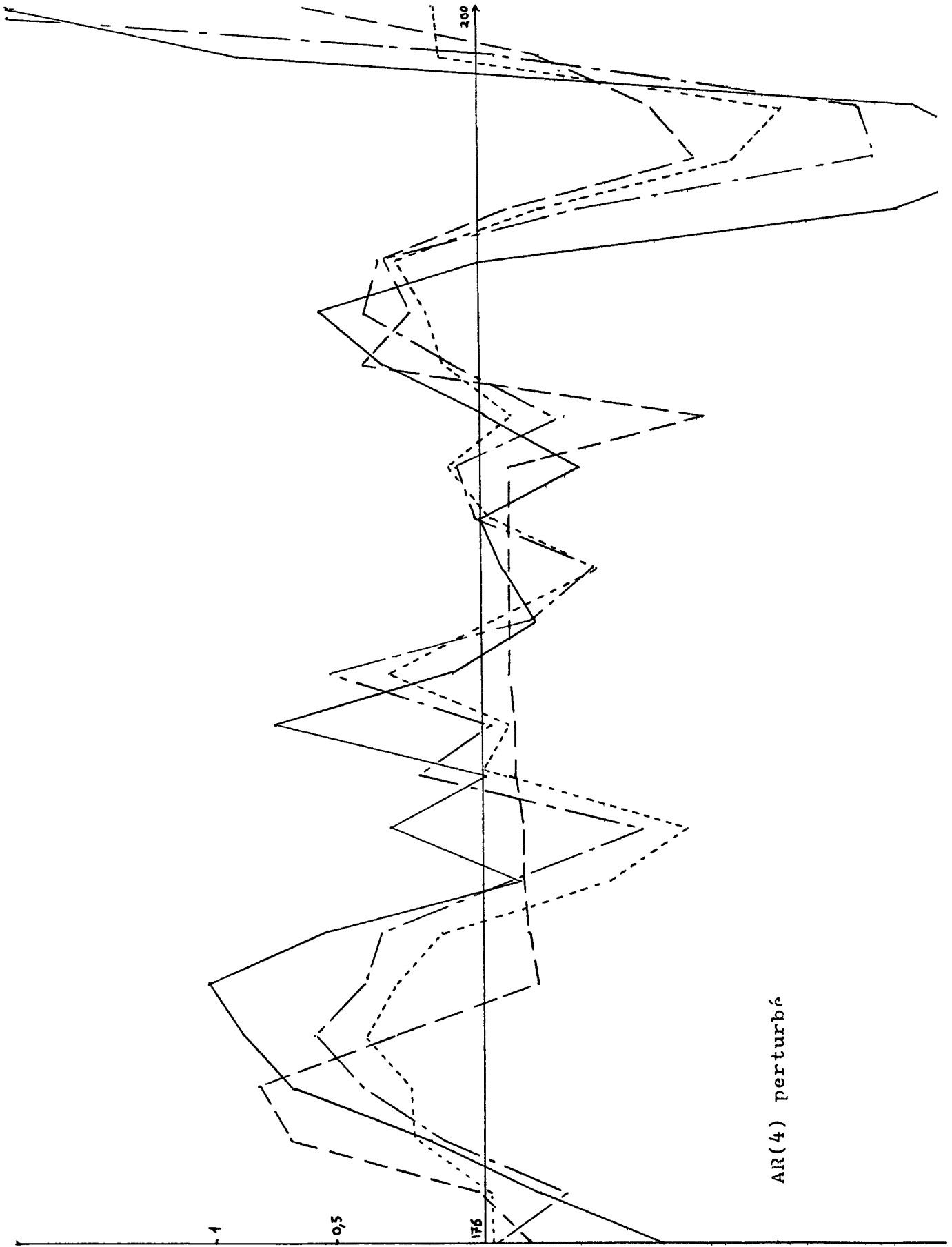
-1

-0,5

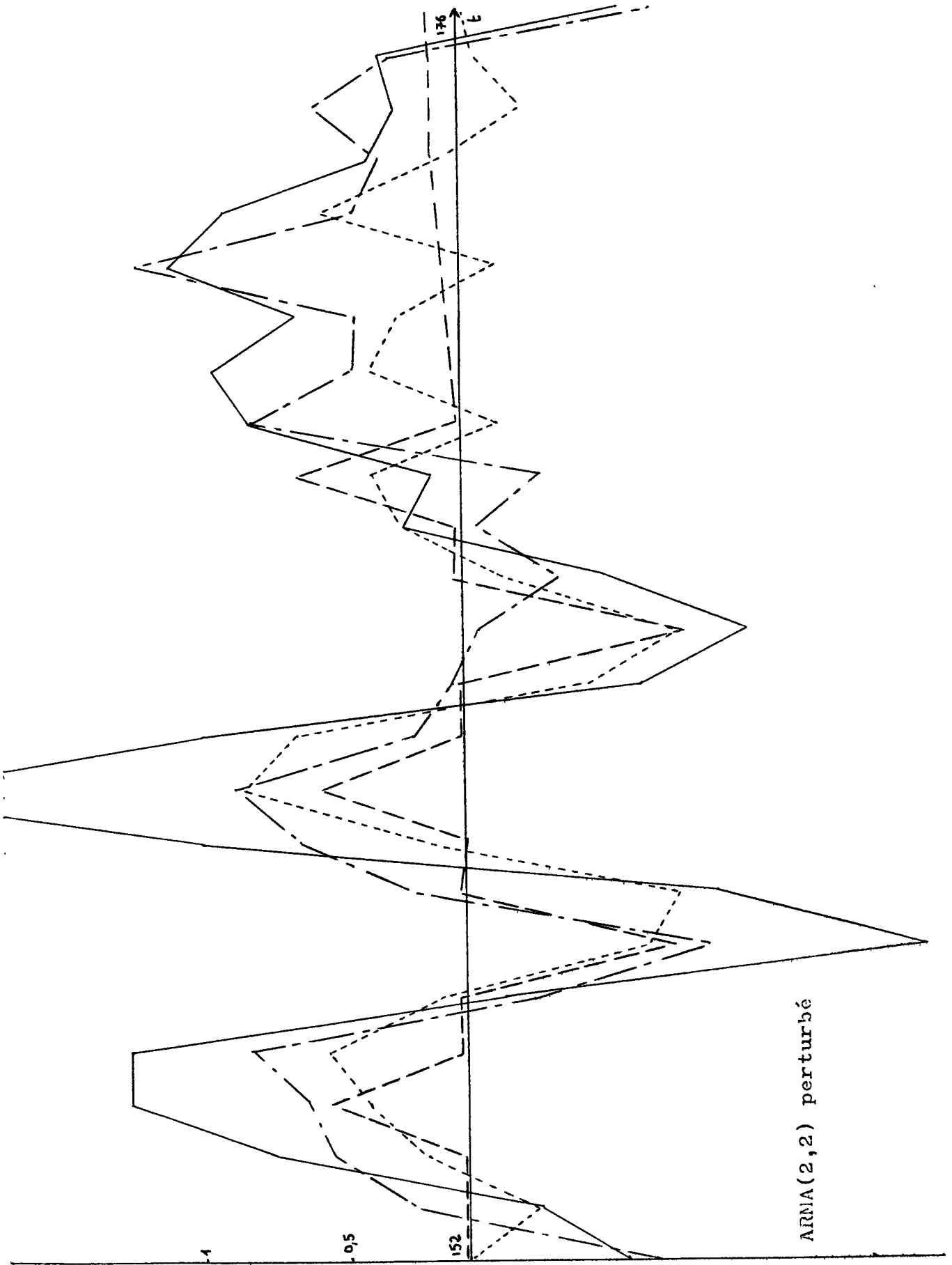
152

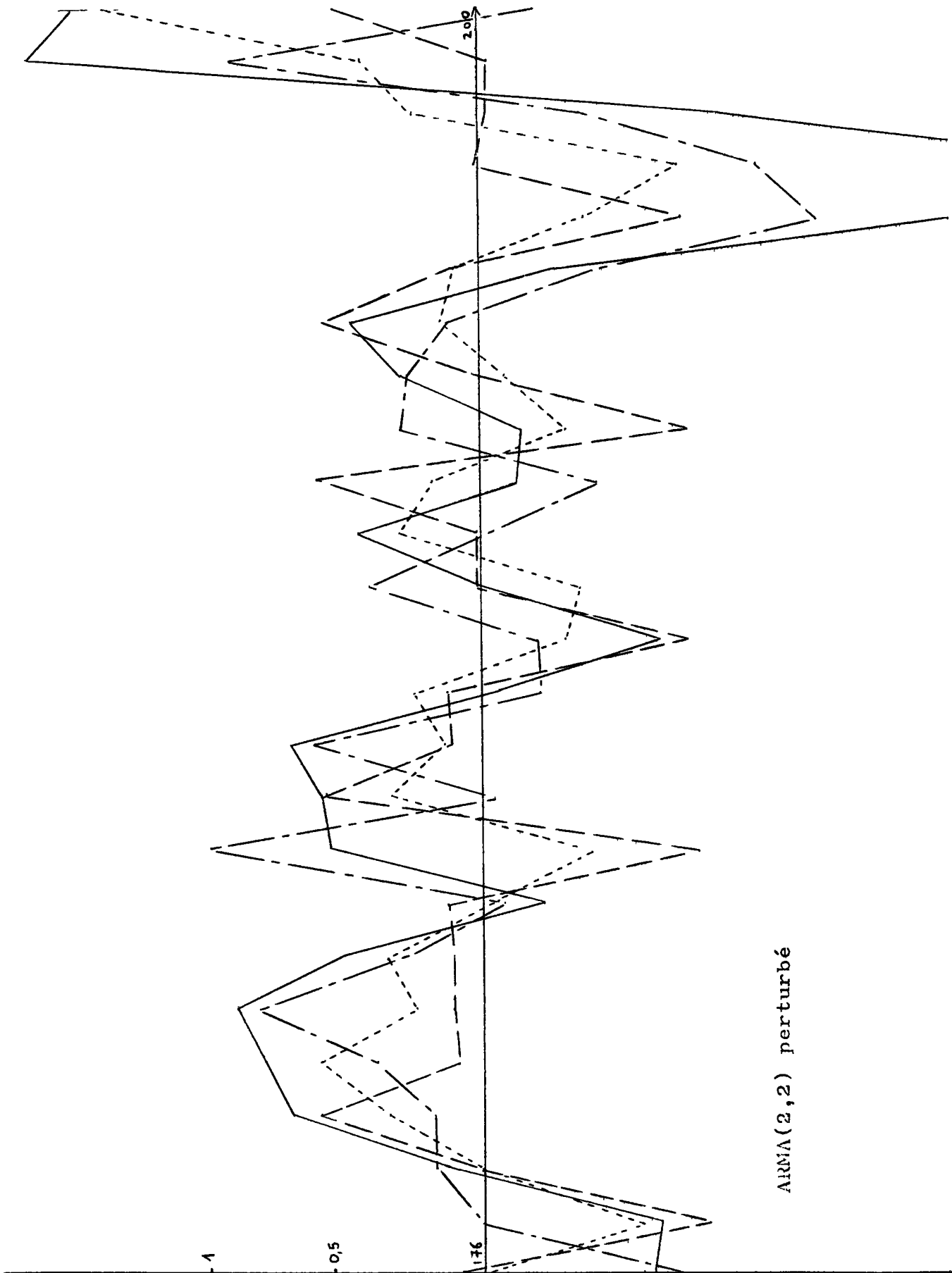
176





AR(4) perturbé





ARMA(2,2) perturbé