

DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 20 au 24 MAI 1985

STATISTIQUE DE LA PROPAGATION HERTZIENNE
A 900 MHZ EN ZONE URBAINE

Gérald MAZZIOTTO et Ammar SEDRATI

Centre National d'Etudes des Télécommunications - 38-40, Rue du Général Leclerc -
92 131 - ISSY LES MOULINEAUX - FRANCE

RESUME

Les phénomènes de propagation intervenant dans une communication hertzienne à 900 MHz en milieu urbain entre un émetteur fixe et un récepteur mobile, conduisent à modéliser la puissance reçue en un point par une variable aléatoire P vérifiant:

$$P = K P_0 d^{-\alpha} R^2 M^2$$

Dans cette formule, P_0 est une puissance de référence, d la distance entre l'émetteur et le récepteur, α un coefficient de l'ordre de 3.4, R une variable aléatoire de loi de Rayleigh normalisée représentant les fluctuations rapides dues aux réflexions multiples des ondes sur les obstacles, M une variable aléatoire de loi Log-Normale représentant les fluctuations plus lentes dues à l'environnement, à l'effet de "masques", indépendante de R , et K est une constante de proportionnalité.

Quand le mobile se déplace dans une région D en suivant un trajet déterminé, les mesures successives de la puissance reçue sont modélisées par un processus aléatoire $P = (P_t; t \in \mathbb{Z})$. On précise ici la statistique de ce processus en exploitant les résultats d'une campagne de mesures effectuée par les Télécoms en région parisienne en 1983.

Exprimé en décibels, le processus des puissances reçues s'écrit:

$$\begin{aligned} A_t &= 10 \log P_t = \bar{Y} - 10\alpha \log d_t + Y_t \\ Y_t &= X_t + W_t, \quad \forall t \in \mathbb{Z}; \end{aligned}$$

où \bar{Y} est déterministe, d_t est la distance émetteur-récepteur lors de la t -ème mesure et où Y est la superposition des processus X et W indépendants, représentant respectivement, les fluctuations lentes dues à l'effet de "masques" et les fluctuations rapides dues aux réflexions multiples de Rayleigh.

Dans cet article, on montre que si la distance entre deux mesures consécutives est supérieure à 1 mètre, alors le processus W est un bruit blanc de variance de l'ordre de 30. De plus, on identifie la loi du processus X comme solution de l'équation:

$$X_{t+1} = a X_t + B_t$$

où $B = (B_t; t \in \mathbb{Z})$ est un bruit blanc de variance de l'ordre de 2, indépendant de W , et a est un coefficient évalué à environ 0.98.

SUMMARY

The mode of propagation of the electromagnetic energy between mobile stations in urban environment leads to the following statistical model for the received power:

$$P = K P_0 d^{-\alpha} R^2 M^2$$

In this formula, P_0 is a reference level, d is the distance between the fixed station and the mobile receptor, α is approximately 3.4, R is a normalized Rayleigh random variable which modelizes the fast fluctuations due to scattering, M is a Log-Normal random variable independent of R , which represents the slow-varying fluctuations due to the environment, and K is a constant factor.

As the mobile receptor travels in the region D , the successive measures of the received power can be modeled by a stochastic process $P = (P_t; t \in \mathbb{Z})$. By studying the results of a set of propagation measures made by the French Telecommunication Administration in Paris in 1983, we identify the statistical law of this process.

Expressed in decibels, the received power process can be written as follows:

$$\begin{aligned} A_t &= 10 \log P_t = \bar{Y} - 10\alpha \log d_t + Y_t \\ Y_t &= X_t + W_t, \quad \forall t \in \mathbb{Z}. \end{aligned}$$

In this formula, \bar{Y} is deterministic, d_t is the distance between the emitter and the receptor at the t -th measure, and Y is a process which is the superposition of the independent processes X and W . These processes describe respectively the slow varying fluctuations and the fast Rayleigh fading.

In this paper, we show that, if the distance between two successive measures is more than 1 meter, then W is a white noise having a variance of order 30. Moreover, X can be identified as the solution of the following equation:

$$X_{t+1} = a X_t + B_t$$

where $B = (B_t; t \in \mathbb{Z})$ is a white noise independent of W with variance of order 2, and the coefficient a can be evaluated as 0.98.



1 - LA PROPAGATION RADIO EN ZONE URBAINE :

Dans ce paragraphe, on propose un modèle statistique qui décrit les variations de la puissance radioélectrique reçue par un mobile se déplaçant dans une zone urbanisée, où une antenne fixe rayonne de façon constante à une fréquence déterminée, voisine de 900 MHz. Les phénomènes de propagation intervenant dans ce type de transmission, dite radiomobile, sont assez bien connus du point de vue physique (cf. /7/). Ils mettent en jeu les trois facteurs suivant.

i) Un affaiblissement dû à la distance, d , entre l'émetteur et le récepteur. En espace libre, la puissance reçue varie en d^{-2} ; par contre en zone urbaine, compte tenu des hauteurs d'antennes et des obstacles, cette variation se fait en $d^{-\alpha}$, où α est généralement compris entre 3 et 4 (cf. /2/, /8/). Celle-ci est donc très lente et parfaitement déterministe une fois que l'on a évalué α .

ii) Une variation rapide, dite de Rayleigh, due aux trajets multiples. Du fait que le récepteur se trouve au milieu d'obstacles réfléchissants divers (immeubles, autres véhicules, etc ..), l'onde reçue en un point est la superposition d'une onde directe et d'un très grand nombre d'ondes réfléchies déphasées, plus ou moins atténuées. Cela a comme conséquence que l'amplitude du signal résultant pris en divers points voisins varie beaucoup, et de façon quasi-imprévisible. En général, on la modélise par une variable aléatoire suivant une loi de Rice ou de Rayleigh en première approximation (cf. /7/, /13/, /15/). Des travaux récents faits au Centre National d'Etudes des Télécommunications (cf. /6/, /9/, /11/) ont montré que pour les mesures effectuées en région parisienne, on pouvait accepter l'hypothèse que cette loi est de Rayleigh (physiquement, cela signifie que les trajets multiples des ondes réfléchies sont prépondérants).

iii) Une variation lente due aux "masques" qui viennent s'interposer sur les trajets radioélectriques. En effet, si on élimine les variations rapides précédentes en faisant la moyenne du champ reçu dans des zones suffisamment petites, on obtient un champ moyen local qui fluctue encore d'une zone à l'autre de façon aléatoire. Un des facteurs possibles de cette variabilité est le degré d'urbanisation de la zone considérée, on peut en imaginer facilement d'autres. Diverses études effectuées à l'étranger (cf. /10/, /14/, /15/) suggèrent que ces fluctuations peuvent être modélisées par une variable aléatoire de loi Log-Normale.

L'idée développée dans la suite de cet article consiste à modéliser la puissance reçue en divers points d'une zone donnée par un processus aléatoire indicé sur l'ensemble des points de mesure: $P = (P; z \in D)$. D'après ce qui précède, la loi de chaque variable aléatoire P_z , en un point z fixé, est assez bien connue. On étudie dans la suite la loi du processus P dans son ensemble.

On considère une antenne fixe qui émet une puissance constante P_0 à une fréquence fixée, voisine de 900 MHz, et on s'intéresse à la puissance P reçue en un point situé à la distance d de l'émetteur. Du fait de la conjonction des trois facteurs précédents, P est une variable aléatoire représentée par le modèle suivant (cf. /14/, /15/) :

$$P = K P_0 d^{-\alpha} |R_M|^2$$

où K est une constante (qui dépend d'autres facteurs fixés ici: par exemple les hauteurs d'antennes, etc..). R_M est une variable aléatoire représentant un champ électrique résultant des deux derniers facteurs. On représente le champ moyen local par une variable aléatoire M qui suit une loi Log-Normale de paramètres (μ, σ^2) (i.e. $\text{Log } M$ suit une loi gaussienne $N(\mu, \sigma^2)$); et on suppose que conditionnellement à M , R_M suit une loi de Rayleigh de paramètre M (i.e., avec un léger abus de langage, sachant que l'événement $\{M=m\}$ est réalisé, R_M suit une loi de densité définie par

$$p_M(x) = e^{-x^2/2m^2} x/m^2 \text{ pour } x \text{ positif.} \text{ Finalement, la loi de } R_M \text{ a pour densité la fonction } f \text{ définie par :}$$

$$f(x) = \int_0^\infty (e^{-x^2/2m^2} x/m^2) \frac{e^{-(\text{Log } m - \mu)^2/2\sigma^2}}{\sigma m \sqrt{2\pi}} dm .$$

Tenant compte du fait que la classe des lois de Rayleigh est stable par multiplication, un simple calcul montre que l'on peut représenter R_M comme le produit de deux variables aléatoires indépendantes R et M , où M est la variable précédente, et R une nouvelle variable de Rayleigh, normalisée (i.e. $R=R_1$). On suppose donc dans la suite que la puissance P reçue en un point est donnée par la formule suivante :

$$P = K P_0 d^{-\alpha} R^2 M^2 .$$

Dans la pratique, les mesures de puissance sont exprimées en décibel par rapport à un certain niveau de référence. Ceci conduit à l'expression suivante en logarithme décimal :

$$10 \log P = 10 \log K P_0 - 10 \alpha \log d + 20 \log M + 20 \log R .$$

Par hypothèse, la variable aléatoire $20 \log M$ est gaussienne. Quant à la variable aléatoire $20 \log R$, on peut calculer explicitement sa densité de probabilité: elle est unimodale mais pas tout à fait symétrique, dès que l'on travaille sur un échantillon important, il n'est pas possible de l'assimiler à une variable gaussienne: cf. /9/ . Néanmoins, c'est une variable de carré intégrable. On se ramène à des variables centrées en posant:

$$X = 20 (\log M - E(\log M)) , W = 20 (\log R - E(\log R))$$

$$Y = X + W \text{ et } 10 \log P = \bar{Y} - 10 \alpha \log d + Y ,$$

où \bar{Y} est une constante.

Finalement, la variable Y représente les fluctuations de la puissance corrigée de l'affaiblissement en distance. Dans la décomposition $Y = X + W$, la variable aléatoire non-gaussienne W représente les fluctuations rapides dues au phénomène de Rayleigh, et X celles dues aux "masques". Rappelons aussi que X est gaussienne et indépendante de W , et qu'elles sont centrées par construction.

Il s'agit maintenant de définir un modèle qui représente conjointement les puissances reçues en divers points d'une zone de réception arbitraire. Si on repère chaque point de la zone D où on effectue une mesure par ses coordonnées cartésiennes (x,y) dans un système topographique donné, on désigne la puissance reçue en ce point par:

$$A_{x,y} = 10 \log P_{x,y} = Y_{x,y} - 10 \alpha \log d_{x,y} + Y_{x,y}$$

où $d_{x,y}$ représente la distance entre le point (x,y) et l'antenne fixe de l'émetteur. On définit ainsi un processus aléatoire $A = (A_{x,y}; (x,y) \in D)$ indexé sur les points de la zone D où sont effectuées les mesures.

Le problème qui est posé est d'élaborer un modèle statistique décrivant la loi de ce processus qui soit en accord avec l'expérience. On supposera que le coefficient d'affaiblissement α est constant dans la région étudiée: à condition de ne pas s'approcher trop de l'antenne, cette hypothèse semble réaliste (cf. /11/). De plus si la zone D a des dimensions pas trop importantes (un rectangle d'une centaine de mètres de côté), on peut raisonnablement admettre que la quantité non aléatoire $Y_{x,y}$ ne dépend pas du point (x,y) . Cette dernière hypothèse est en fait assez difficile à vérifier: elle est certainement inexacte dans une zone trop étendue, ou dans certains cas exceptionnels, mais inversement, il n'y a pas de raisons de la refuser dès que la zone est suffisamment petite.

On adopte donc le modèle suivant:

$$A_{x,y} = \bar{Y}_D - 10 \alpha \log d_{x,y} + Y_{x,y} , \forall (x,y) \in D .$$

où \bar{Y}_D est évalué pour chaque zone D en faisant la moyenne de la puissance reçue, corrigée de l'affaiblissement dû à la distance :

$$Y_D = \sum_{(x,y) \in D} (A_{x,y} + 10\alpha \log d_{x,y}) / \text{cardinal } D$$

On choisit une valeur moyenne du coefficient α qui est obtenue par une méthode de régression sur l'ensemble des zones considérées. On a trouvé $\alpha = 3.4$ avec les fichiers de mesures exploités pour cet article.

On définit aussi les processus X et W comme précédemment les variables X et W, on a

$$\forall (x,y) \in D : Y_{x,y} = X_{x,y} + W_{x,y}$$

En chaque point $(x,y) \in D$, on sait que $W_{x,y}$ est une variable centrée de carré intégrable, et de loi connue, et que $X_{x,y}$ suit une loi gaussienne de moyenne nulle et d'écart type $\sigma_{x,y}$, indépendante de $W_{x,y}$. Si la zone D n'est pas trop grande, on peut supposer en première approximation, que l'écart type $\sigma_{x,y}$ ne dépend pas du point (x,y) . Tout cela entraîne que la loi du couple $(X_{x,y}, W_{x,y})$ est invariante par translation. Il est naturel de renforcer cette hypothèse en supposant qu'il en est de même pour la loi de tout 2n-uple associé à un n-uple de points pris dans D. Finalement, on exprime cela en disant que le processus (X,W) est stationnaire dans la zone D. On supposera de plus que les processus X et W sont indépendants.

Le processus X représente les fluctuations lentes dues aux "masques" et si on s'intéresse plus particulièrement à celui-ci, le processus W, qui représente les fluctuations rapides de Rayleigh, peut s'interpréter comme un "bruit". Il est intuitif que si la distance qui sépare deux points de mesure adjacents est suffisamment grande, alors toutes les variables $(W_{x,y}; (x,y) \in D)$ sont mutuellement indépendantes, et W est un bruit blanc (non-gaussien). Par contre, si les points de mesure sont trop rapprochés, on s'attend à ce qu'il existe des corrélations non-nulles des variations rapides. On précisera cette question dans la suite de l'article (cf. aussi /11/).

2 - CHOIX DU MODELE STATISTIQUE :

Le problème posé est celui de l'étude statistique des processus X et W à partir de celle de Y, avec la mise au point de modèles descriptifs pas trop compliqués, susceptibles d'applications à la simulation.

Les données dont on dispose pour identifier la loi du processus Y ont été relevées lors d'une campagne de mesures effectuées par le Centre National d'Etudes des Télécommunications et la Direction des Télécommunications du Réseau Extérieur en 1983. Chaque fichier de données est relatif à la mesure de la puissance reçue en des points d'une zone carrée de 500 mètres de côté. L'émetteur est situé en dehors de la zone, à quelques kilomètres, et il émet une puissance fixe à une fréquence voisine de 900 MHz. Le récepteur est monté sur un camion se déplaçant dans les rues de la zone étudiée selon un trajet choisi pour remplir au mieux tout le carré. Pour plus de détails sur les conditions techniques des mesures, on se reportera à /6/. Chaque fichier contient une série de quadruplets (x_n, y_n, A_n, d_n) pour n allant de 1 à N, où N est le nombre total de mesures effectuées sur le trajet, et n le numéro d'ordre de la mesure. (x_n, y_n) définit les coordonnées cartésiennes du point où a été effectué la mesure, A_n la puissance reçue mesurée en décibel par rapport au niveau de référence et d_n est la distance calculée séparant l'émetteur du récepteur. Parmi les 56 fichiers étudiés, il y en avait 19 où les données étaient relevées tous les mètres au cours du trajet, avec N de l'ordre de 3000, et 37 où les mesures étaient relevées tous les 5 mètres, avec N de l'ordre de 800. L'émetteur, dénommé "Pastourelle" était situé dans le quartier du Marais à Paris. Il faut souligner que l'on est forcé d'admettre que la puissance reçue en divers

points de la zone D ne fluctue qu'en fonction de la position de ces points, et ne dépend pas de l'instant auquel la mesure a été effectuée. En effet il est probable qu'il s'écoule un laps de temps non-négligeable entre le relevé de la première mesure, correspondant à $n=1$, et de la dernière, pour $n=N$, au cours de la constitution du fichier. Pour étudier cette dépendance en temps, il faudrait disposer de mesures effectuées en des points fixes à des instants différents. Néanmoins, si on suppose que la dépendance en temps des différentes variables est stationnaire et que l'intervalle de temps qui sépare deux mesures consécutives est bien supérieur au temps de corrélation moyen du champ, on peut effectivement négliger cette dépendance, en première approche.

On a vu au paragraphe précédent que la puissance reçue en un point de coordonnées (x,y) se représente par la formule :

$$A_{x,y} = 10 \log P_{x,y} = \bar{Y} - 10\alpha \log d_{x,y} + Y_{x,y}$$

Le problème posé est l'estimation de la loi du processus Y. Celui-ci est par définition indexé sur des points d'un rectangle du plan : c'est ce qu'on appelle un processus à deux indices. La branche correspondante de la théorie des processus aléatoires est actuellement en plein développement, notamment en raison des applications possibles au traitement des images. Cependant, pour pouvoir identifier la loi du processus Y, il serait bon de disposer d'une mesure en chaque sommet d'un quadrillage suffisamment serré de la zone D. Pour un carré de 500 mètres de côté avec un quadrillage de pas 5mètres, cela fait environ 10 000 points. Ceci est à comparer aux 800 données dont on dispose, le long d'un trajet arbitraire dans la zone. Cette remarque donne une idée des difficultés de mise en oeuvre de cette méthode; celle-ci sera développée dans une étude ultérieure.

Dans cet article, on adopte une autre approche, fondée sur le fait suivant. L'ordre dans lequel sont effectuées les mesures constituant le fichier définit une série chronologique. Pour t variant de 1 à N, on pose :

$$A_t = \bar{Y}_D - 10\alpha \log d_{x_t, y_t} + Y_{x_t, y_t}$$

pour (x_t, y_t) les coordonnées du point où la t^{ième} mesure est effectuée. On note aussi d_t la distance séparant ce point de l'émetteur, et on définit finalement un processus indexé sur une partie de \mathbb{N} en posant :

$$A_t = \bar{Y}_D - 10\alpha \log d_t + Y_t, \quad \forall t = 1, \dots, N.$$

On définit de même les processus X et W tels que :

$$Y_t = X_t + W_t, \quad \forall t = 1, \dots, N$$

Le processus $W = (W_t; t = 1, \dots, N)$ reste un bruit (blanc si la distance entre deux mesures est grande) centré non-gaussien, indépendant du processus $X = (X_t; t = 1, \dots, N)$. Par contre, le fait que les mesures soient relevées le long d'un chemin arbitraire dans la zone D, et qui plus est, se reboucle plusieurs fois, rend l'hypothèse de stationnarité difficile à admettre en toute rigueur. Néanmoins, si on se restreint à des fichiers où le trajet est formé de segments rectilignes relativement longs par rapport au pas d'acquisition avec des boucles assez larges, on peut penser que la loi du processus Y reste approximativement invariante par des translations pas trop grandes.

D'un point de vue pratique, cette approche est tout aussi intéressante que celle des processus à deux indices. En effet, le processus $A = (A_t; t = 1, \dots, N)$ représente la puissance effectivement reçue par un usager qui se déplace, en suivant un certain trajet dans la zone D. La connaissance de sa statistique sera particulièrement utile si on veut, par exemple, mettre au point un système régulateur de gain fondé sur un algorithme de prédiction récursif. On verra dans l'exploitation des résultats qu'en fait les lois statistiques estimées dans chaque fichier sont assez semblables, et donc assez indépendantes du chemin, en moyenne.



On se propose donc d'analyser les enregistrements de mesures de chaque fichier comme une série chronologique. A partir de la série $A = (A_t; t = 1, \dots, N)$, on construit la série $Y = (Y_t; t = 1, \dots, N)$ en posant:

$$Y_t = A_t + 10 \alpha \log d_t - \bar{Y}_D, \quad \forall t = 1, \dots, N$$

avec $\alpha = 3.4$ et \bar{Y}_D évalué par un moyennage sur la zone D, couvrant un carré d'environ 500 mètres de côté. Les représentations graphiques de Y (cf. figure 1) suggèrent que la série n'est pas parfaitement stationnaire. La technique classique en théorie des séries chronologiques (cf. /1/, /3/) permettant de se ramener à des séries plus stationnaires est la différenciation. On remplace ainsi la série Y par la série Y' définie par

$$Y'_t = Y_t - Y_{t-1}, \quad \forall t = 1, \dots, N \quad (\text{avec } Y_{-1} = 0).$$

L'opération peut être répétée jusqu'à l'obtention d'une série stationnaire; pour le cas présent, la série Y' semble suffisamment stationnaire (cf. figure 2). Ce procédé fait, en particulier, disparaître toute tendance déterministe de variation négligeable entre deux points consécutifs. Dans ce qui précède, on a écrit Y comme la somme des processus X et W, où W représente des fluctuations rapides, et où X a des raisons physiques de varier peu entre deux mesures consécutives. Cette dernière propriété s'exprime aussi en disant que le processus dérivé X' a une énergie très faible. On en déduit que, d'une part, l'observation de la série dérivée Y' peut apporter des informations sur le processus W. Mais que d'autre part, il ne paraît pas judicieux de dériver Y pour étudier X.

Une raison possible de la non-stationnarité de la série Y est l'existence d'un terme de dérive ou tendance déterministe. Pour diminuer l'influence d'un tel facteur, on divise la zone D de 500 mètres de côté en 9 sous-zones: D1, ..., D9, et on calcule des moyennes locales: Y_{D1}, \dots, Y_{D9} . On considère maintenant la série Y définie par:

$$Y_t = A_t + 10 \alpha \log d_t - \sum_{i=1}^9 \bar{Y}_{Di} \mathbb{1}\{(x_t, y_t) \in Di\}$$

Les représentations graphiques de Y (cf. figure 3) ont en général un aspect plus stationnaire que les précédentes (cf. figure 1). On peut alors espérer obtenir des renseignements sur le terme X à partir de Y et W.

3 - ANALYSE SPECTRALE DES SERIES CHRONOLOGIQUES :

Dans ce paragraphe, on étudie les séries Y et Y' par des méthodes fondées sur l'analyse spectrale des processus aléatoires du second ordre. Celle-ci n'est valide que si les processus sous-jacents Y et Y' sont stationnaires au sens large (cf. /1/, /4/). L'estimation des spectres se fait au moyen des périodogrammes, ce qui suppose que diverses moyennes temporelles convergent vers des espérances mathématiques, donc que les processus considérés sont ergodiques (cf. /1/) ou de cumulants suffisamment intégrables (cf. /4/).

Etant donné un processus aléatoire réel indéxé sur tous les entiers relatifs Z, centré stationnaire du second ordre, $Z = (Z_t; t \in \mathbb{Z})$, on définit la fonction d'autocovariance R_Z sur Z par:

$$\forall t \in \mathbb{Z} : R_Z(t) = E(Z_s Z_{s+t}), \quad \forall s \in \mathbb{Z}.$$

Dans le cas où la fonction d'autocovariance R_Z décroît suffisamment vite à l'infini pour que la série de terme général $R_Z(t)$ converge, il existe une fonction f_Z , appelée densité spectrale de Z, qui est paire et de période 2π , et qui vérifie:

$$\forall t \in \mathbb{Z} : R_Z(t) = \int_{-\pi}^{+\pi} e^{it\lambda} f(\lambda) d\lambda$$

$$\forall \lambda \in \mathbb{R} : f_Z(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} R_Z(t) e^{-it\lambda}$$

Pour chaque λ , $f_Z(\lambda)$ représente la puissance spectrale

du processus Z à la fréquence λ .

Ayant observé le processus Z sur un intervalle de longueur T : $(Z_t; t = 0, \dots, T-1)$, il est naturel de chercher à estimer la covariance R_Z au point t , pris entre $-T+1$ et $T-1$, par la variable aléatoire $C_Z^T(t)$ suivante:

$$C_Z^T(t) = \frac{1}{T} \sum_{s=0}^{T-t-1} Z_s Z_{s+t},$$

et la densité spectrale $f_Z(\lambda)$ au point $\lambda \in [0, \pi]$ par la variable aléatoire $I_Z^T(\lambda)$ suivante:

$$I_Z^T(\lambda) = \frac{1}{2T} \left| \sum_{t=0}^{T-1} e^{-it\lambda} Z_t \right|^2.$$

On appelle périodogramme d'ordre T la fonction aléatoire I_Z^T . Pour $\lambda \in [-\pi, \pi]$, on note $\lambda(T)$ le nombre pris dans l'ensemble $\{k\pi/T; k = -T, \dots, T\}$ le plus proche de λ . Il est établi (cf. /4/) que $I_Z^T(\lambda(T))$ est un estimateur asymptotiquement sans biais de la quantité $f_Z(\lambda)$. Cependant, la variance de cet estimateur reste de l'ordre de $f_Z(\lambda)^2$, c'est à dire que ses fluctuations sont relativement importantes, et qu'elles n'ont même pas tendance à diminuer quand T croît. Un troisième résultat est que deux estimations, à des fréquences différentes même très proches, sont asymptotiquement non corrélées. Tout ceci donne un aspect "bruit blanc" au périodogramme (cf. figure 4). Pour obtenir une estimation plus régulière et plus précise de la densité spectrale f_Z , on a recours au périodogramme lissé. Le périodogramme lissé d'ordre (T, m) est défini en posant, pour $m \in \mathbb{N}$:

$$I_Z^{T,m}(\lambda) = \frac{1}{2m+1} \sum_{j=-m}^m I_Z^T(\lambda(T) + j\pi/T), \quad \lambda \neq 0 \quad (\pi)$$

Il est montré (cf. /4/) que l'estimateur $I_Z^{T,m}(\lambda)$ est asymptotiquement sans biais pour $f_Z(\lambda)$ et de variance $f_Z(\lambda)^2 / (2m+1)$, pour $\lambda \in]0, \pi[$. Ce procédé diminue donc les fluctuations de l'estimation quand m croît, en revanche il est bien connu qu'il a tendance à augmenter le biais; il y a un compromis à réaliser.

Du point de vue pratique, on traite ici des fichiers où T varie de 800 à 4000 environ. Le périodogramme est calculé au moyen d'un algorithme de Transformation de Fourier Rapide. Le périodogramme lissé est évalué avec m de l'ordre de 5 à 20. Il faut noter que le lissage "gomme" les détails du spectre et élargit les "pics" (cf. figure 5). La fonction d'autocovariance est ensuite obtenue par une transformation de Fourier inverse sur le périodogramme.

Cette méthode a d'abord été appliquée à la série Y construite au paragraphe précédent. La densité spectrale f_Y estimée apparaît sur $[-\pi, \pi]$ comme la superposition d'un pic autour de zéro et d'une densité à peu près uniforme sur le reste de l'intervalle. La fonction d'autocovariance C_Y présente une discontinuité importante à l'origine, puis décroît ensuite régulièrement pour devenir négligeable au bout d'une distance correspondant à quelques centaines de mètres (cf. figure 5 et figure 6). Ces résultats suggèrent bien que Y peut être considéré comme la superposition d'un processus X à grande distance de corrélation, et d'un bruit blanc W. La variance de W correspond au saut de la fonction d'autocovariance à l'origine, et vaut environ 35. Tout ceci confirme le modèle proposé dans la Paragraphe 1.

La série dérivée Y' associée à Y est étudiée avec la même méthode. Au préalable, on remarque que, si Y est un processus stationnaire du second ordre, il en est a fortiori de même pour le processus Y', et la densité spectrale $f_{Y'}$ de Y' s'exprime à partir de la densité spectrale f_Y de Y par la formule suivante:

$$\forall \lambda \in \mathbb{R} : f_{Y'}(\lambda) = 2(1 - \cos \lambda) f_Y(\lambda).$$

Si on suppose que Y est la somme de deux processus stationnaires du second ordre X et W, alors on a aussi la décomposition $Y' = X' + W'$, où X' et W' sont les

processus dérivés de X et W, respectivement. Dire que X a une grande distance de corrélation ou qu'il fluctue peu entre deux mesures consécutives signifie que le processus dérivé X' a une énergie négligeable. Tout cela entraîne que f_Y , diffère peu de $f_{W'}$. On peut alors estimer la densité spectrale de W à partir de l'observation de la série dérivée Y'. Pour les fichiers étudiés, les périodogrammes lissés des séries Y' donnent des courbes homothétiques de celle de la fonction sur $[-\pi, \pi]$ $\lambda \rightarrow 2(1 - \cos \lambda)$. On en déduit donc que la densité spectrale $f_{W'}$ est constante sur tout l'intervalle $-\pi, \pi$, sauf peut-être au point zéro. On a ainsi vérifié que le processus W est un bruit blanc (pour les pas d'échantillonnage de 1 ou 5 mètres considérés ici). Un périodogramme lissé de Y' est représenté sur la figure 7.

4 - IDENTIFICATION PAR DES SERIES ARMA :

Une méthode particulièrement opératoire de spécifier la loi d'un processus aléatoire consiste à définir celui-ci comme la solution d'un système d'équations aux différences finies piloté par un bruit blanc. En d'autres termes, on identifie la loi du processus étudié avec celle de la sortie d'un filtre numérique dont l'entrée est un bruit blanc. La détermination des caractéristiques de ce filtre s'appelle le problème d'identification. Dans la pratique, on se fixe une classe de modèles i.e. de filtres, caractérisés par un nombre fini de coefficients, et on essaie d'approcher au mieux la loi du processus réel observé par la sortie d'un de ces systèmes. Evidemment, il y a un compromis à réaliser entre le nombre de coefficients et donc la complexité du problème d'identification, et la capacité du modèle à représenter convenablement la réalité.

Les classes les plus utilisées sont celles des processus ARMA d'ordre (p,q), où p et q sont des entiers fixés. Celles-ci ont vocation à représenter des processus stationnaires du second ordre dont les densités spectrales sont certaines fractions rationnelles de numérateur de degré q et de dénominateur de degré p. Ce type de problème est étudié dans de nombreux ouvrages, et on se réfère à /1/,/3/. On rappelle qu'un processus réel stationnaire centré du second ordre, $Z = (Z_t ; t \in \mathbb{Z})$ est un ARMA d'ordre (p,q) si il vérifie la relation suivante:

$$\forall t \in \mathbb{Z} : \sum_{j=0}^p a_j Z_{t-j} = \sum_{k=0}^q b_k B_{t-k}$$

où $B = (B_t ; t \in \mathbb{Z})$ est un bruit blanc centré de variance σ_B^2 , et où les coefficients $(a_j ; j=0, \dots, p)$ et $(b_k ; k=0, \dots, q)$ sont réels, avec $a_0 = b_0 = 1$. Le problème d'identification se ramène alors à la détermination des nombres $\sigma_B^2, a_1, \dots, a_p, b_1, \dots, b_q$, à partir de la série Z observée. Pour ce faire, on suit la méthode développée par Box et Jenkins /3/.

Tout d'abord, on a appliqué cette méthode à la série $Y = (Y_1, \dots, Y_N)$ construite au Paragraphe 2. Pour la plupart des 56 fichiers, on a été conduit à accepter une représentation ARMA(1,1) du type suivant:

$$\forall t \in \mathbb{Z} : Y_t + a Y_{t-1} = \epsilon_t + b \epsilon_{t-1}$$

où ϵ est un bruit blanc de variance σ_ϵ^2 . Les valeurs estimées des coefficients a et b légèrement selon les fichiers. Des valeurs typiques sont:

$$a = -0.98, \quad b = -0.83, \quad \sigma_\epsilon^2 = 38.7$$

Dans cette première identification, on n'a pas tenu compte de toutes les hypothèses faites sur Y dans les paragraphes précédents. En particulier on a ignoré la décomposition $Y = X + W$. Pour la faire intervenir, il faut rappeler quelques résultats concernant l'identification des processus autorégressifs (i.e. ARMA(p,0)) bruités, développée notamment dans /5/ et /12/.

Supposons que le processus Y est la superposition de deux processus indépendants: un bruit blanc W de variance σ_W^2 et un processus ARMA(p,0) X.

On peut montrer qu'alors le processus Y satisfait une équation ARMA(p,p) (cf. /5/, /12/), dont les coefficients se déduisent des précédents. Détaillons le cas p=1 qui est utilisé ici. Supposons que Y vérifie le système:

$$\begin{aligned} X_t + a X_{t-1} &= B_t \\ Y_t &= X_t + W_t, \quad \forall t \in \mathbb{Z} \end{aligned}$$

où B et W sont des bruits blancs indépendants de variance σ_B^2 et σ_W^2 . On a alors aussi la relation suivante:

$$Y_t + a Y_{t-1} = \epsilon_t + b \epsilon_{t-1}$$

avec ϵ un bruit blanc de variance σ_ϵ^2 .

Les paramètres (a, σ_B^2, σ_W^2) sont liés aux paramètres (a, b, σ_ϵ^2) par les équations:

$$\sigma_W^2 = \sigma_\epsilon^2 b/a \quad \sigma_B^2 = \sigma_\epsilon^2 (1 + b^2 - b(a + 1/a))$$

Etant donnée une série (Y_1, \dots, Y_N) , on applique la méthode de Box et Jenkins pour déterminer les coefficients a, b et σ_ϵ^2 . On déduit alors des formules ci-dessus une estimation des coefficients a, σ_B^2 et σ_W^2 . On trouve ici les valeurs suivantes:

$$a = -0.98 \quad \sigma_W^2 = 32.6 \quad \text{et} \quad \sigma_B^2 = 1.2$$

pour un fichier typique.

Afin de comparer ces résultats avec ceux obtenus par analyse spectrale au Paragraphe 3, on peut calculer la densité spectrale du processus Y défini par le modèle ci-dessus. Celle-ci s'écrit:

$$f_Y(\lambda) = \frac{1}{2\pi} \left(\frac{\sigma_B^2}{1 + a^2 - 2a \cos \lambda} + \sigma_W^2 \right)$$

L'allure de f_Y est comparable à la figure 5. On en conclut que l'approche développée dans ce paragraphe et celle du Paragraphe 3 conduisent à des modèles similaires.

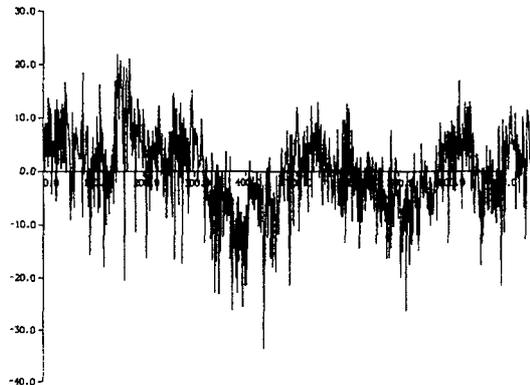


Figure 1 : Série Y, avec moyennage global.

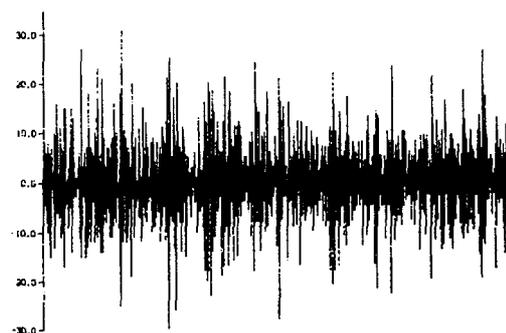


Figure 2 : Série dérivée Y'.

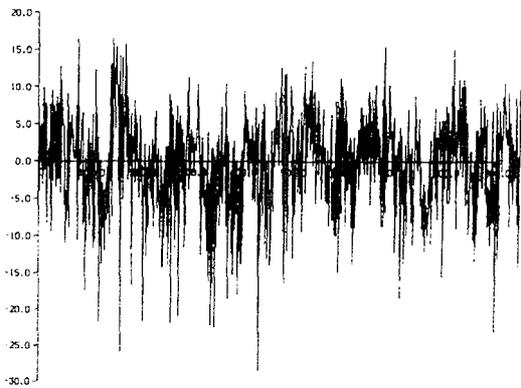


Figure 3 : Série Y avec moyennage localisé .

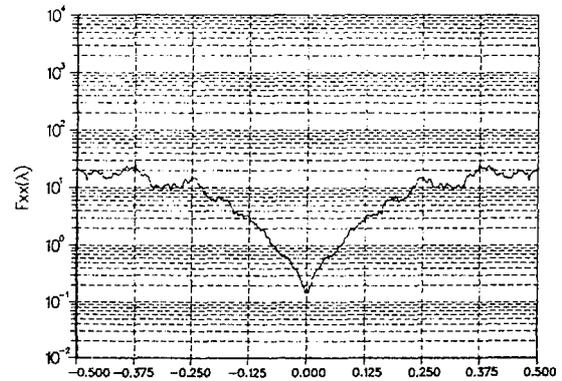


Figure 7 : Périodogramme lissé de Y' .

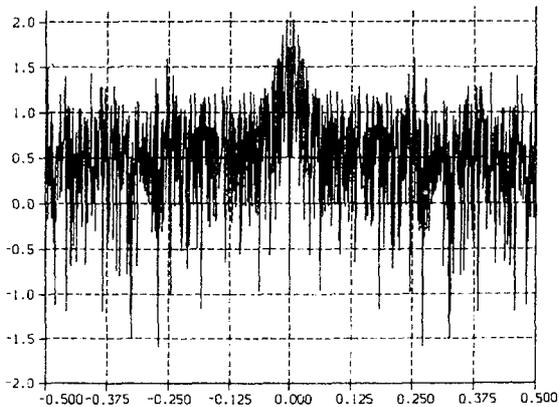


Figure 4 : Périodogramme de Y .

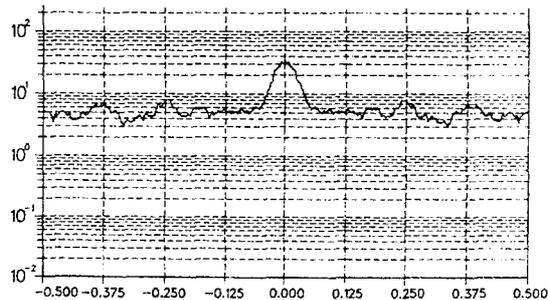


Figure 5 : Périodogramme lissé de Y .

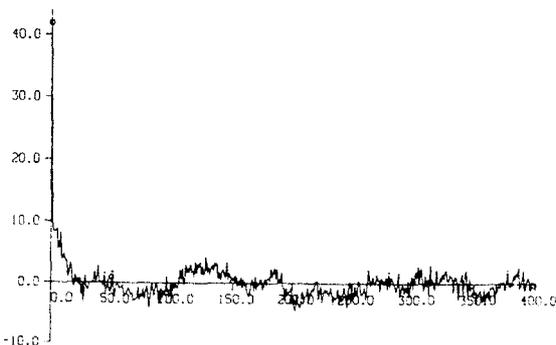


Figure 6 : Fonction d'autocovariance de Y .

REFERENCES :

- /1/ R. AZENCOTT et D. DACUNHA-CASTELLE : Séries d'observations irrégulières. Masson, Paris 1984.
- /2/ L. BOITHIAS : Propagation des ondes radio-électriques dans l'environnement terrestre. Dunod, Paris 1981.
- /3/ G.E.P. BOX et G.M. JENKINS : Time series analysis; forecasting and control. Holden-Day, San Francisco 1976.
- /4/ D.R. BRILLINGER : Time series; data analysis and theory. Holden-Day, San Francisco 1975.
- /5/ K. BRY : Etudes d'algorithmes d'identification de modèles autorégressifs lorsque les observations sont bruitées, et applications. Thèse Ecole Nationale Supérieure des Télécommunications 1982, Paris.
- /6/ C. HAVEL, A. AZOULAY, P. LE GAC et D. ROBERT : Mesures de propagation effectuées par le Département ETR/CEM pour le projet "MARATHON" au cours de l'année 1983 avec la collaboration de la D.T.R.E.. Note Technique CNET/PAB/ETR/680, Juin 1984.
- /7/ W.C. JAKES : Microwave mobile communications. John Wiley & Son, New-York 1974.
- /8/ A. MALOBERTI : Influence des caractéristiques de propagation sur un système de radiomobile cellulaire numérique. L'Onde Electrique, Juin 1984, Paris.
- /9/ G. MAZZIOTTO et A. SEDRATI : Modèle des fluctuations lentes de la propagation radioélectrique en milieu urbain à 900 MHz. Note Technique CNET/TIM/MTI/1479 et CNET/PAB/ETR/697, Janvier 1985.
- /10/ Y. OKUMURA, E. OHMORI, T. KAWANO et K. FUKUDA : Field strength and its variability in VHF and UHF land-mobile radio service. Rev. Elec. Comm. Lab. (Japan), Vol. 16 N° 9-10; pp. 825-873.
- /11/ P. OLIVIER et J. TIFFON : Propagation radioélectrique à 900 MHz en milieu urbain; modélisation à fréquence fixe et résultats expérimentaux. Note Technique C.R.P.E. 136, Septembre 1984.
- /12/ M. PAGANO : Estimation of models of autoregressive signal plus white noise. The Annals of Statistics Vol. 2-1 (1974) pp. 99-108.
- /13/ M. SCHWARTZ, W.R. BENNET et S. STEIN : Communication systems and techniques. Mac Graw Hill, New-York 1966.
- /14/ H. SUZUKI : A statistical model for urban multipath propagation. IEEE Trans. Communications, Vol. 25-7 (1977) pp. 673-680.
- /15/ G.L. TURIN, F.D. CLAPP, T.L. JOHNSTON, S.B. FINE et D. LAVRY : A statistical model for urban multipath propagation. IEEE Trans. Vehic. Techn. Vol. 21-1 (1972) pp. 1-9.