

Modélisation et prédiction de processus à mémoire longue par décomposition multirésolution

Skander SOLTANI¹, Patrice SIMARD¹, Daniel BOICHU²

¹Université de Technologie de Compiègne.
HEUDIASYC UMR CNRS 6599
BP 20529, 60205 Compiègne Cedex, France

²Université de Technologie de Compiègne.
Division des Mathématiques Appliquées
BP 20529, 60205 Compiègne Cedex, France

soltani@hds.utc.fr, simard@hds.utc.fr, boichu@utc.fr

Résumé – Nous nous intéressons ici à la modélisation et la prédiction de séries temporelles à mémoire longue. Pour la famille de modèles couramment utilisée, notée *ARfIMA*, c'est une intégration fractionnaire du processus qui suit un modèle ARMA. L'estimation du paramètre d'intégration fractionnaire est un problème déterminant et difficile. Nous proposons une méthode de modélisation et de prédiction qui évite cette estimation par une décomposition multirésolution. Ceci est possible à cause des propriétés statistiques des séries obtenues. Un exemple d'application est donné pour illustrer la méthode.

Abstract – This paper outlines a new method for long memory processes modeling and prediction, it uses a multiscale filtering. This method avoids the estimation of the fractional integration parameter used when dealing with the classical *ARfIMA* modeling. The method was applied to an ultra low temperature measurement problem.

1 Introduction

Le fait que certaines séries temporelles ont une propriété de mémoire longue est connu depuis longtemps. Par exemple: le trafic dans les réseaux de télécommunications, les séries financières, ou certaines séries biologiques.

Une série temporelle stationnaire est dite à mémoire longue si sa fonction d'autocorrélation décroît lentement¹. Sa densité spectrale a alors un comportement en $1/f^\alpha$, $\alpha > 0$ près de l'origine. Ceci implique que les basses fréquences ou les variations lentes sont très importantes. Si la série est non-stationnaire, c'est le spectre de Wigner-Ville qui, en moyenne, a ce comportement.

La méthode la plus courante pour la modélisation de ce type de série introduit une intégration fractionnaire: c'est la famille des modèles *ARfIMA* (p, d, q). Dans ce cas, c'est une dérivation fractionnaire, d'ordre d , du processus qui est supposée suivre un modèle ARMA:

$$\Phi(B)(1-B)^d X_k = \Theta(B)\varepsilon_k, \quad (1)$$

où B est l'opérateur de retard unitaire, $\Phi(B)$ et $\Theta(B)$ sont deux polynômes de degrés p et q respectivement. Les échantillons ε_k , $k \geq 0$ sont Gaussiens et indépendants. L'introduction du terme $(1-B)^d$ induit sous certaines conditions la longue dépendance à cause du fait que les coefficients de son développement de Taylor

$$(1-B)^d = \sum_{m=0}^{\infty} h_m B^m, \quad h_m = \frac{\Gamma(d)}{m! \Gamma(d-m)} \quad (2)$$

décroissent lentement.

1. La fonction d'autocorrélation n'est pas intégrable.

On peut ainsi démontrer qu'un processus X_k satisfaisant (1) est stationnaire si et seulement si $d < \frac{1}{2}$ et $\Phi(z) \neq 0$ pour $|z| \leq 1$, alors qu'il est inversible si et seulement si $d > -\frac{1}{2}$ et $\Theta(z) \neq 0$ pour $|z| \leq 1$ [Hos80].

Notons à la fin que les processus *ARfIMA* ($0, d, 0$) sont les similaires discrets des bruits Gaussiens fractionnaires qui sont des processus continus obtenus par dérivation des mouvements Browniens fractionnaires [Ber94]. La fonction d'autocorrélation d'un processus *ARfIMA* ($0, d, 0$), $d < \frac{1}{2}$ satisfait

$$\gamma_X(\tau) \sim C_d |\tau|^{2d-1} \quad \text{quand } |\tau| \rightarrow \infty, \quad (3)$$

alors que sa densité spectrale satisfait

$$S_X(\omega) = \frac{1}{2\pi} (2 \sin \frac{\omega}{2})^{-2d} \simeq \frac{1}{2\pi} |\omega|^{-2d} \quad \text{quand } \omega \rightarrow 0. \quad (4)$$

Nous remarquons bien à partir de ces deux équations que la longue dépendance apparait quand $d > 0$.

Tout l'enjeu de ce type de modélisation réside dans l'estimation, à partir des données, du paramètre d'intégration fractionnaire. Plusieurs méthodes ont été proposées, parmi les quelles on peut citer: la méthode de Higuchi, l'agrégation de la variance de la série, le périodogramme, décomposition en ondelettes, la statistique R/S, ... [AGF95, Ber94, Let98, Rei94, TTW95]. Toutes ces méthodes ne donnent pas de résultats satisfaisants quand le modèle contient un partie autorégressive ou à moyenne mobile.

Nous introduisons ici une méthode de modélisation et de prédiction de ce type de séries qui évite justement ce problème d'estimation en utilisant des décompositions multirésolution. Nous obtenons ainsi une série trame qui

contient les basses fréquences, et une hiérarchie de séries de détail qui chacune d'elles contient une certaine bande de fréquences. En effet, nous pouvons démontrer que les séries de détail ne contiennent pas de longue dépendance et sont stationnaires. Ce qui implique qu'elles peuvent être modélisées par la famille ARMA. Cette méthode est exposée en détail dans §2. Les résultats sur une application de mesure de très basses températures sont donnés dans §3

2 La modélisation multirésolution

Le problème avec la famille ARfIMA (p, d, q) réside dans l'estimation du paramètre d'intégration d surtout en présence de d'une partie autorégressive ou moyenne mobile. Afin d'éviter ce problème, nous présentons une méthode de modélisation et de prédiction fondée sur une approche de filtrage multirésolution.

2.1 Le filtrage multirésolution

Le filtrage multirésolution utilise un couple de filtres d'analyse: le premier de réponse impulsionnelle h_n , $n \in \mathbb{Z}$, est un filtre passe-bas, alors que le second de réponse impulsionnelle g_n , $n \in \mathbb{Z}$ est passe-bande. Le filtrage d'une série, en une itération, donne une série (dite trame) contenant les basses fréquences et une autre série (dite détail) contenant les hautes fréquences. La synthèse ou la reconstruction de la série original est possible grâce à deux autres filtres, qui doivent satisfaire certaines conditions.

Dans la perspective d'une application de prédiction, les échantillons futurs sont supposés être inconnus, nous imposons donc la causalité des filtres d'analyse et de synthèse. Ceci implique que

$$h_n = g_n = 0, \quad n \leq -1. \quad (5)$$

On impose également que la reconstruction se fait de manière simple en additionnant directement la trame et le détail, ce qui implique

$$\begin{cases} h_0 + g_0 = 1, \\ h_n = -g_n, \quad n \geq 1. \end{cases} \quad (6)$$

Des contraintes sur le nombre de moments nuls des filtres peuvent être rajoutées. Les résultats obtenus jusqu'à maintenant ne sont guère intéressants, c'est pour cette raison que nous n'allons utiliser que les filtres de Haar [Dau92] donnés par

$$\begin{cases} h_0 = h_1 = \frac{1}{2}, \\ g_0 = -g_1 = \frac{1}{2}. \end{cases} \quad (7)$$

Ce schéma de décomposition peut être appliqué sur plusieurs itérations, en filtrant à chaque itération la trame obtenue dans l'itération précédente.

Nous notons $c_{J-j_0, k}$ et $d_{J-j, k}$, $j = 1, \dots, j_0$, $k = 0, \dots, \ell - 1$ la trame et les détails obtenues après ℓ décomposition de l'ensemble de mesures disponibles $X_k = c_{J, k}$, $k = 0, \dots, \ell - 1$. La contrainte de reconstruction simple donne alors:

$$y_k = c_{J, k} = c_{J-j_0, k} + \sum_{j=1}^{j_0} d_{J-j, k}. \quad (8)$$

Le couple de filtres utilisés peuvent varier d'une itération à une autre à condition qu'ils satisfassent les contraintes de causalité et de reconstruction simple. Une application directe de cette remarque est l'insertion, à chaque itération, de zéros dans les réponses impulsionnelles des deux filtres. Ce procédé est couramment utilisé dans l'algorithme à trous [She92]. La figure 1 montre les réponses impulsionnelles des filtres utilisés pour une décomposition sur trois niveaux. Les filtres issus de la méthode avec insertion de zéros s'avèrent plus intéressants car les bandes de fréquences sont bien séparées, ce qui n'est pas le cas des filtres de Haar. La section suivante est dédiée à l'application de ce type de décomposition aux processus à mémoire longue.

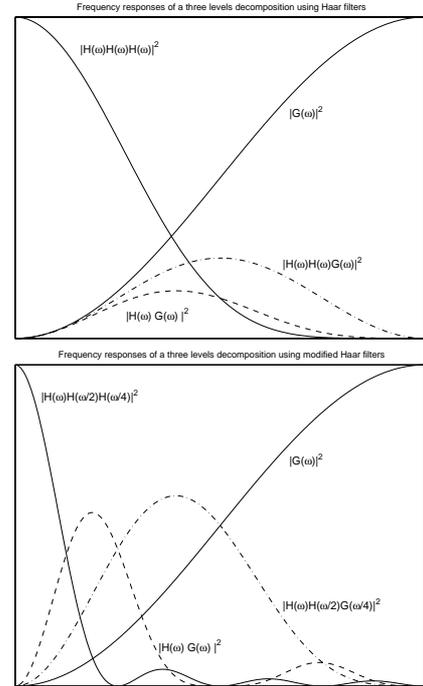


FIG. 1: Représentation des réponses fréquentielles des filtres utilisés pour une décomposition sur trois niveaux, en utilisant les filtres de Haar (en haut), et en utilisant les filtres de Haar avec insertion de zéros (en bas).

2.2 La modélisation multirésolution

Nous ne considérons ici que les modèles sans la partie moyenne mobile, *i.e.* les processus satisfaisant

$$\Phi(B)(1-B)^d X_k = \varepsilon_k, \quad \varepsilon_k \sim \text{i.i.d.} \quad (9)$$

L'introduction du terme d'intégration fractionnaire assure la prise en compte des valeurs lointaines dans le passé. En changeant de base dans l'équation (2), le développement de Taylor à l'ordre N de $(1-B)^d$ peut s'écrire en fonction des fonctions $\frac{(1+B)^N}{2^N}$ et $\frac{(1+B)^k}{2^k} \frac{(1-B)}{2}$, $k = 0, \dots, N-1$ au lieu de B^k , $k = 0, \dots, N$. Ainsi, on peut écrire

$$(1-B)^d \approx h'_N \left(\frac{1+B}{2} \right)^N + \sum_{k=0}^{N-1} h'_k \left(\frac{1+B}{2} \right)^k \left(\frac{1-B}{2} \right). \quad (10)$$

En injectant cette écriture dans l'équation 9, nous obtenons alors

$$\Phi(B)h'_N \left(\frac{1+B}{2} \right)^N + \sum_{k=0}^{N-1} h'_k \Phi(B) \left(\frac{1+B}{2} \right)^k \left(\frac{1-B}{2} \right) X_k = \varepsilon_k. \quad (11)$$

Les termes $\left(\frac{1+B}{2} \right)^N X_k$ et $\left(\frac{1+B}{2} \right)^m \left(\frac{1-B}{2} \right) X_k$, $m = 0, \dots, N-1$ correspondent aux séries trame et détails obtenus par une décomposition sur N échelles utilisant les filtres de Haar. Ainsi, on a

$$\Phi_N(B)c_{J-N,k} + \sum_{m=1}^N \Phi_k(B)d_{J-m,k} = \varepsilon_k. \quad (12)$$

Cette écriture ressemble à un modèle AR classique, mais implique $N+1$ processus. Plutôt que d'estimer les $N+1$ polynômes, nous préférons écrire le problème sous forme vectorielle. Ainsi, on obtient :

$$\begin{pmatrix} c_{J-N,k} \\ d_{J-N,k} \\ \vdots \\ d_{J-1,k} \end{pmatrix} = A_1 \begin{pmatrix} c_{J-N,k-1} \\ d_{J-N,k-1} \\ \vdots \\ d_{J-1,k-1} \end{pmatrix} + \dots + A_p \begin{pmatrix} c_{J-N,t-p} \\ d_{J-N,t-p} \\ \vdots \\ d_{J-1,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{0,t} \\ \vdots \\ \varepsilon_{N,t} \end{pmatrix} \quad (13)$$

Le problème original d'estimation du paramètre d et du polynôme $\Phi(B)$ a été transformé en un problème, plus simple, d'estimation de plusieurs polynômes $\Phi_k(B)$, $k = 0, \dots, N$ dans un modèle autorégressif multivariable.

3 Application

3.1 Expérimentation

La méthode proposée a été testée sur les signaux obtenus durant l'expérience DIABOLO menée par des astrophysiciens de l'IAS, Paris. L'objectif de cette expérience est la construction d'un instrument capable de mesurer le rayonnement de l'univers aux ondes millimétriques. Cet instrument, dit bolomètre, est une résistance sensible à la température, de taille réduite ($\approx \mu M$) et de valeur très élevée ($\approx M\Omega$). La sensibilité de cet instrument est élevée et sa réponse est fortement non-linéaire.

La théorie de rayonnement du corps noir impose le refroidissement du bolomètre à des températures avoisinant les 100 mK pour pouvoir les ondes de longueur entre 1 mm et 2 mm. Le refroidissement est possible grâce à un système de dilution de ^3He dans ^4He . Zagury [Zag96] a montré que ce

système introduit de fortes fluctuations, et corrompt les mesures avec un bruit en $1/f$. Ce sont ces mesures que nous allons étudier.

3.2 Résultats

L'ensemble des mesures prises durant cette expérience contient 735.000 mesures prises à une fréquence de 74.4Hz. La figure 2 montre un sous-ensemble des ces mesures. Un ensemble 10000 mesures a été constitué dans l'objectif d'identifier les paramètres de notre modèle multirésolution. Pour la décomposition, nous avons préféré utiliser les filtres de Haar avec insertion de zéros afin de garantir la prise en compte d'échantillons lointains dans le passé. Le nombre d'itérations de la décomposition est égal à 4. Il est inutile de décomposer encore car la trame que nous avons obtenu à ce niveau ne contient pratiquement plus aucune information. L'ordre du modèle autorégressif multivariable a été choisi égal à 6. Afin de valider le résultat obtenu, nous l'avons testé sur un autre ensemble contenant 40000 mesures. Cet ensemble a été choisi le plus loin possible du premier afin de garantir le minimum de corrélation entre eux. La figure 3 montre la densité spectrale des mesures sur l'ensemble de validation, ainsi que celle des erreurs commises sur ce mme ensemble, les deux tracés ont été effectués dans des plans log-log. On voit bien le comportement en $1/f$ des mesures qui se traduit par une droite d'une certaine pente près de l'origine. Par contre, le spectre des résidus est pratiquement plat, ce qui implique que les résidus sont pratiquement blancs, et que la longue dépendance a été apprise par notre modèle.

4 Conclusion

Nous avons proposé une méthode de modélisation et de prédiction de séries temporelles à mémoire longue. Cette méthode est fondée sur un filtrage multirésolution, qui permet d'obtenir une suite de séries de détail stationnaires, et ne contenant que la courte dépendance. Elle est simple, rapide et de surcroît évite le problème d'estimation du paramètre d'intégration fractionnaire. Les résultats obtenus sur des séries simulées, non rapportés ici, et sur les données d'astrophysique attestent son efficacité.

Nos travaux futurs concerneront la construction de nouveaux filtres, plus réguliers qui permettent une meilleure séparation des bandes de fréquences. D'autres domaines (finance, bruit dans les transistors à effet de champ, ...) représentent des applications intéressantes de la méthode. Il serait également intéressant d'étudier la possibilité d'utiliser la méthode pour estimer le paramètre d'intégration fractionnaire et l'utilisation des paramètres du modèle à des fins de diagnostic. Enfin, le problème des processus ayant une partie moyenne mobile reste encore ouvert.

Références

- [AGF95] P. Abry, P. Goncalvès, and P. Flandrin. Wavelets, spectrum analysis and $1/f$ processes. volume 103 of *Lectures in Statistics*, pages 15–30. Springer Verlag, 1995.
- [Ber94] Jan Beran. *Statistics of long-memory processes*. Number 61 in Monographs on Statistics and Applied Probability. Chapman and Hall, New York, 1994.
- [Dau92] Ingrid Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series on Applied Mathematics, No 61, SIAM, 1992.
- [Hos80] J. R. M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–176, 1980.
- [Let98] Emmanuel Letavernier. *Méthodes d'analyse des propriétés d'auto-affinité des processus comportant des composantes spectrales en $1/f$* . PhD thesis, Université de Technologie de Compiègne, France, 1998.
- [Rei94] Valderio A. Reisen. Estimation of the fractional difference parameter in the arima (p, d, q) using the smoothed periodogram. *Journal of time series Analysis*, 15(3):335–350, 1994.
- [She92] M. J. Shensa. Discrete wavelet transform: wedding the “à trous” and mallat algorithms. *IEEE Trans. signal processing*, 40:2464–2484, 1992.
- [TTW95] M. Taqqu, V. Teverovsky, and W. Wilinger. Estimators for long-range dependence: an empirical study. *Fractals*, 3(4):785–788, 1995.
- [Zag96] Frédéric Zagury. *Exemples de l'importance de l'interaction champs de rayonnement-matière pour la compréhension de la structure du milieu intersellaire proche*. PhD thesis, Université Paris XI, 1996.

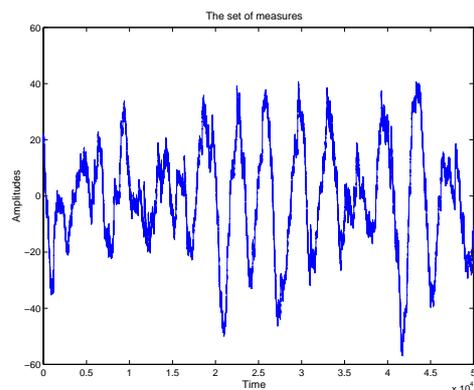


FIG. 2: Illustration d'un sous-ensemble de mesures du bolomètre.

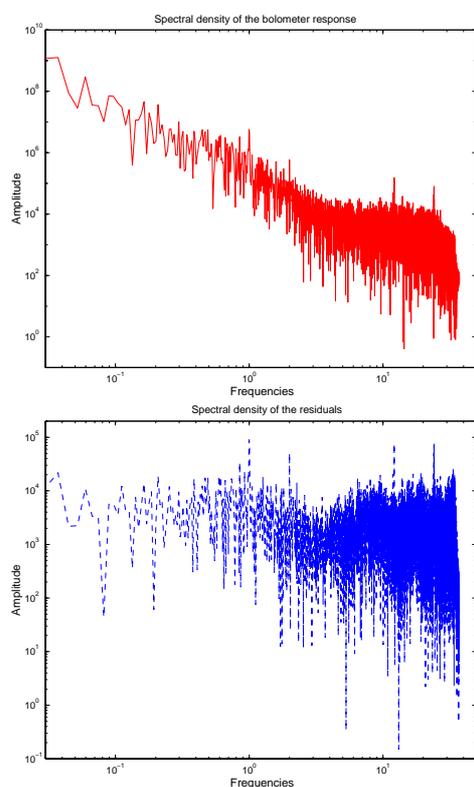


FIG. 3: Densité spectrale des mesures (en haut) et celle des résidus sur l'ensemble de validation (en bas).