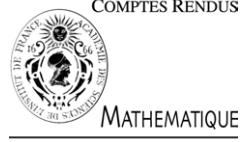




Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

C. R. Acad. Sci. Paris, Ser. I 339 (2004) 435–440



Statistics/Probability Theory

Estimating parameters of a k -factor GIGARCH process

Abdou Kâ Diongue ^{a,b}, Dominique Guégan ^a

^a ENS Cachan IDHE-MORA, UMR CNRS 8533, 61, avenue du président Wilson, 94231 Cachan cedex, France

^b EDF R&D, 1, avenue du général de Gaulle, 92141 Clamart cedex, France

Received 13 January 2004; accepted after revision 12 July 2004

Available online 26 August 2004

Presented by Paul Deheuvels

Abstract

Some crucial time series of market data, such as electricity spot prices, exhibit long-memory, in the sense of slowly-decaying correlations combined with heteroskedasticity. To be able to modelize such a behaviour, we consider in this Note the k -factor GIGARCH process and we propose two methods to address the related parameter estimation problem. For each method, we develop the asymptotic theory for the estimation. *To cite this article: A.K. Diongue, D. Guégan, C. R. Acad. Sci. Paris, Ser. I 339 (2004).*

© 2004 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Résumé

Estimation des paramètres d'un processus GIGARCH à k facteurs. Plusieurs données de marché, telles que les prix spot de l'électricité, présentent de la longue mémoire, au sens de la décroissance hyperbolique des autocorrélations combinée avec un phénomène d'hétéroskédasticité. Pour modéliser de tels comportements, nous considérons dans cette Note les processus GIGARCH à k facteurs et nous proposons deux méthodes d'estimation des paramètres de ce modèle. Enfin, nous développons les propriétés asymptotiques de ces estimateurs. *Pour citer cet article : A.K. Diongue, D. Guégan, C. R. Acad. Sci. Paris, Ser. I 339 (2004).*

© 2004 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Version française abrégée

Dans cette Note, nous nous intéressons à l'estimation des paramètres d'un processus GIGARCH à k facteurs par la méthode des moindres carrés conditionnels (CSS) et la méthode du maximum de vraisemblance de Whittle. Ce processus défini par les équations (1), (2) a été introduit et étudié dans les articles de Guégan [7,8]. Soit $\{X_t\}_{t=1}^T$ un processus GIGARCH à k facteurs stationnaire défini par les équations (1), (2). Posons $\gamma =$

E-mail addresses: abdou-ka.diongue@edf.fr (A.K. Diongue), guegan@ecogest.ens-cachan.fr (D. Guégan).

$(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, d_1, \dots, d_k)$, $\delta = (a_0, a_1, \dots, a_r, b_1, \dots, b_s)$ et $\omega = (\gamma, \delta)$. Supposons que $\omega_0 = (\gamma_0, \delta_0)$ soit la vraie valeur du paramètre ω et se trouve à l'intérieur du compact $\Theta \subseteq \mathbb{R}^{p+q+k+r+s+1}$. Nous supposons, dans toute la suite, que toutes les G-fréquences sont connues. Dans le Théorème 2.1, nous donnons les propriétés asymptotiques de l'estimateur des paramètres par la méthode CSS. Les propriétés asymptotiques des estimateurs des paramètres par la méthode de Whittle sont fournies dans le Théorème 2.2 pour les paramètres de mémoire longue et de mémoire courte homoscédastiques et dans le Théorème 2.3 pour les paramètres hétéroscléastiques.

Théorème 2.1. Soit $\{X_t\}_{t=1}^T$ un processus générée par les équations (1), (2). Supposons que $a_0 > 0$, $a_1, \dots, a_r, b_1, \dots, b_s \geq 0$, $\sum_{i=1}^r a_i + \sum_{i=1}^s b_i < 1$, $E(\varepsilon_t^4) < \infty$, $0 < d_i < \frac{1}{2}$ si $|v_i| < 1$, ou $0 < d_i < \frac{1}{4}$ si $|v_i| = 1$ pour $i = 1, \dots, k$ et toutes les racines de $\phi(B)$ et $\theta(B)$ soient en dehors du cercle unité. Si les fréquences v_i sont connues alors

- (i) Il existe un estimateur CSS $\hat{\omega}_T$ satisfaisant $\frac{\partial L(\omega)}{\partial \omega} = 0$ et $\hat{\omega}_T \xrightarrow{P} \omega_0$ quand $T \rightarrow \infty$.
- (ii) $\sqrt{T}(\hat{\omega}_T - \omega_0) \xrightarrow{D} N(0, \Omega_0^{-1})$ quand $T \rightarrow \infty$, avec $\Omega_0 = \text{diag}(\Omega_{\gamma_0}, \Omega_{\delta_0})$.
- (iii) De plus, les estimateurs consistants des matrices d'information Ω_γ et Ω_δ sont donnés en (5).

Théorème 2.2. Soit $\{X_t\}_{t=1}^T$ un processus défini par les équations (1), (2). Supposons vérifiées les hypothèses du Théorème 2.1. Alors, les estimateurs de Whittle des paramètres γ sont tels que :

- (i) $\hat{\gamma}_T \xrightarrow{P.s} \gamma_0$ quand $T \rightarrow \infty$.
- (ii) $\sqrt{T}(\hat{\alpha}_T - \alpha_0) \xrightarrow{D} N(0, 4\pi V(\alpha_0)^{-1})$, quand $T \rightarrow \infty$, où $\alpha = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$, et où $V(\alpha)_{ij}$ est défini par l'équation (6).
- (iii) $\sqrt{T}(\hat{d}_T - d) \xrightarrow{D} N(0, 4\pi V(d)^{-1})$, où $V(d)_{ij}$ est donnée en (7).

Théorème 2.3. Soit $\{X_t\}_{t=1}^T$ un processus défini par les équations (1), (2). Supposons vérifiées les hypothèses du Théorème 2.1. Alors

- (i) Sous $(H_0)(J=4)$ et (H_1) , on a $\hat{d}_T \xrightarrow{P} d_0$, quand $T \rightarrow \infty$.
- (ii) Sous $(H_0)(J=8), (H_1)$ et (H_2) , on a $\sqrt{T}(\hat{d}_T - d_0) \xrightarrow{D} N(0, 2W^{-1} + W^{-1}VW^{-1})$, quand $T \rightarrow \infty$, où V est donné en (9).

1. Introduction

Assume that $(\xi_t)_{t \in \mathbb{Z}}$ is a white noise process with unit variance and let the polynomials $\phi(B)$ and $\theta(B)$ denote the ARMA operators. Let B denote the backshift operator and $0 < d_i < \frac{1}{2}$ if $|v_i| < 1$ or $0 < d_i < \frac{1}{4}$ if $|v_i| = 1$ for $i = 1, \dots, k$. We define a centered k -factor GIGARCH process $(X_t)_{t \in \mathbb{Z}}$ by, $\forall t$

$$\phi(B) \prod_{i=1}^k (I - 2v_i B + B^2)^{d_i} X_t = \theta(B) \varepsilon_t, \quad (1)$$

where

$$\varepsilon_t = \sqrt{h_t} \xi_t \quad \text{with } h_t = a_0 + \sum_{i=1}^r a_i \varepsilon_{t-i} + \sum_{i=1}^s b_i h_{t-i}. \quad (2)$$

For $i = 1, \dots, k$, the frequencies $\lambda_i = \arccos(v_i)$ are called the Gegenbauer frequencies (or G-frequencies). The process defined by Eqs. (1), (2) was introduced by Guégan (see [7,8]). In the following section, we provide some

results related to the asymptotic properties of the k -factor GIGARCH process estimators, obtained by two methods: the conditional sum of squares and the Whittle approach.

2. Asymptotic theory for estimation

Given a stationary k -factor GIGARCH process $\{X_t\}_{t=1}^T$ defined by Eqs. (1), (2). We denote $\gamma = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, d_1, \dots, d_k)$, $\delta = (a_0, a_1, \dots, a_r, b_1, \dots, b_s)$ and $\omega = (\gamma, \delta)$ its parameters. We assume that $\omega_0 = (\gamma_0, \delta_0)$ is the true value of ω and is in the interior of the compact set $\Theta \subseteq \mathbb{R}^{p+q+k+r+s+1}$. Let us assume that all the G-frequencies are known.

2.1. Conditional sum of squares estimation

The conditional sum of squares estimator $\hat{\omega}_T$ of ω in Θ maximizes the conditional logarithmic likelihood $L(\omega)$ on F_0 , where F_t is the σ -algebra generated by $(X_s, s \leq t)$. If we assume that the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ have a conditional Gaussian distribution then the conditional log-likelihood is defined by

$$L(\omega) = \frac{1}{T} \sum_{t=1}^T \ell_t, \quad \ell_t = -\frac{1}{2} \log(h_t) - \frac{\varepsilon_t^2}{2h_t}. \quad (3)$$

Now, if we assume that the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ have a conditional Student distribution with l degrees of freedom, then the CSS estimator $\hat{\omega}_T$ maximizes the likelihood function $L(\omega)$ defined by

$$\begin{aligned} L(\omega) = T & \left[\log \Gamma \left\{ \frac{(l+1)}{2} \right\} - \log \Gamma \left(\frac{l}{2} \right) - \frac{1}{2} \log(l-2) \right] \\ & - \frac{1}{2} \sum_{t=1}^T \left\{ \log(h_t) + (l+1) \left[\log \left(1 + \frac{\varepsilon_t^2}{h_t(l-2)} \right) \right] \right\}. \end{aligned} \quad (4)$$

In the following theorem, $L(\omega)$ represents the log likelihood introduced in (3) or in (4).

Theorem 2.1. Suppose that the process $(X_t)_{t \in \mathbb{Z}}$ is generated by Eqs. (1), (2). Assume that $a_0 > 0, a_1, \dots, a_r, b_1, \dots, b_s \geq 0, \sum_{i=1}^r a_i + \sum_{i=1}^s b_i < 1, E(\varepsilon_t^4) < \infty, 0 < d_i < \frac{1}{2}$ if $|v_i| < 1$ or $0 < d_i < \frac{1}{4}$ if $|v_i| = 1$ for $i = 1, \dots, k$ and all roots of the polynomials $\phi(B)$ and $\theta(B)$ lie outside the unit circle. Then

- (i) There exists a CSS estimator $\hat{\omega}_T$ that satisfies $\frac{\partial L(\omega)}{\partial \omega} = 0$ and $\hat{\omega}_T \xrightarrow{P} \omega_0$ as $T \rightarrow \infty$.
- (ii) $\sqrt{T}(\hat{\omega}_T - \omega_0) \xrightarrow{D} N(0, \Omega_0^{-1})$ as $T \rightarrow \infty$, where \xrightarrow{P} and \xrightarrow{D} denotes respectively the convergence in probability and in distribution. Furthermore, $\Omega_0 = \text{diag}(\Omega_{\gamma_0}, \Omega_{\delta_0})$ and Ω_{γ_0} and Ω_{δ_0} are values of Ω_γ and Ω_δ at $\omega = \omega_0$, with $\Omega_\gamma = E([\frac{1}{h_t} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma^T} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \gamma} \frac{\partial h_t}{\partial \gamma^T}])$ and $\Omega_\delta = E([\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \delta} \frac{\partial h_t}{\partial \delta^T}])$.
- (iii) The information matrices Ω_γ and Ω_δ can be estimated consistently by

$$\widehat{\Omega}_\gamma = \frac{1}{T} \sum_{t=1}^T \left[\frac{1}{h_t} \frac{\partial \varepsilon_t}{\partial \gamma} \frac{\partial \varepsilon_t}{\partial \gamma^T} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \gamma} \frac{\partial h_t}{\partial \gamma^T} \right] \quad \text{and} \quad \widehat{\Omega}_\delta = \frac{1}{T} \sum_{t=1}^T \left(\left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \delta} \frac{\partial h_t}{\partial \delta^T} \right] \right). \quad (5)$$

The proof is given in Section 3 for the Gaussian case and details can be found in Diongue, Guégan and Vignal [4]. Note that, if the innovations $(\varepsilon_t)_{t \in \mathbb{Z}}$ have a conditional Student distribution with l degrees of freedom, then the proof can easily be performed using the same steps as in Section 3.

2.2. Whittle estimation

In this paragraph, we investigate Whittle's method of estimating all parameters of the k -factor GIGARCH process defined by Eqs. (1), (2). The first step consists of estimating the long-memory parameters $d = (d_1, \dots, d_k)$ and the ARMA(p, q) parameters $\alpha = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ using the Whittle's approach (for more details, see Chung [2,3] and Ferrara and Guégan, Chapter 8 of [5]). In the second step, the GARCH(r, s) parameters $\delta = (a_0, a_1, \dots, a_r, b_1, \dots, b_s)$ are estimated using Whittle's method applied to the residuals of the long-memory process (see Giraitis and Robinson [6] for more details).

Theorem 2.2. Let $\{X_t\}_{t=1}^T$ be a k -factor GIGARCH process defined by Eqs. (1), (2). Let us assume that the same hypothesis given in Theorem 2.1 are verified. Then

- (i) $\hat{\gamma}_T \xrightarrow{a.s} \gamma_0$ as $T \rightarrow \infty$.
- (ii) Furthermore: $\sqrt{T}(\hat{\alpha}_T - \alpha_0) \xrightarrow{D} N(0, 4\pi V(\alpha_0)^{-1})$, as $T \rightarrow \infty$, where

$$V(\alpha)_{ij} = \int_{-\pi}^{\pi} g^2(\lambda, \omega) \frac{\partial g^{-1}(\lambda, \omega)}{\partial \alpha_i} \frac{\partial g^{-1}(\lambda, \omega)}{\partial \alpha_j} d\lambda. \quad (6)$$

Here $g(\lambda, \omega)$ denotes the spectral density of the process $(X_t)_{t \in \mathbb{Z}}$.

- (iii) Moreover $\sqrt{T}(\hat{d}_T - d) \xrightarrow{D} N(0, 4\pi V(d)^{-1})$, with

$$V(d)_{ij} = \int_{-\pi}^{\pi} \log \left| 4 \sin \left[\frac{(\lambda - \lambda_i)}{2} \right] \sin \left[\frac{(\lambda + \lambda_j)}{2} \right] \right| \log \left| 4 \sin \left[\frac{(\lambda - \lambda_j)}{2} \right] \sin \left[\frac{(\lambda + \lambda_i)}{2} \right] \right| d\lambda. \quad (7)$$

Theorem 2.2 follows from the proof of Hosoya's Theorem 2.3 [9].

To estimate the GARCH(r, s) parameters δ , we consider the process $(\varepsilon_t^2)_{t \in \mathbb{Z}}$ in its ARMA representation. This means that we can rewrite (2) as: $\varepsilon_t^2 - \sum_{i=1}^{\max(r,s)} (a_i + b_i) \varepsilon_{t-i}^2 = a_0 + v_t - \sum_{j=1}^s b_j v_{t-j}$, where $b_i = 0$ if $i \in (s, r]$ and $a_i = 0$ if $i \in (r, s]$. The process $(v_t)_{t \in \mathbb{Z}}$ defined by $v_t = \varepsilon_t^2 - h_t$ constitutes a white noise sequence with mean zero and variance σ^2 . We introduce now some complementary assumptions to get the consistency and asymptotic normality of $\hat{\delta}_T$:

(H_0) For $t = 0, \pm 1, \dots$, the process $(\xi_t)_{t \in \mathbb{Z}}$ introduced in Eq. (2), is strictly stationary, ergodic with finite J th moment and $E(\xi_t | F_{t-1}) = 0$, $E(\xi_t^2 | F_{t-1}) = 1$, and $E(\xi_t^{2j} | F_{t-1}) = \nu_{2j}$ almost-surely, with $j = 2, \dots, J/2$, where ν_{2j} are constants such that $|\nu_J|^{2/J} (\sum_{i=1}^r a_i + \sum_{i=1}^s b_i) < 1$.

- (H_1)
- (i) $\int_{-\pi}^{\pi} \log f(\lambda, \delta) d\lambda = 0$, for all δ , with $f(\lambda, \delta)$ the spectral density of the process $(\varepsilon_t)_{t \in \mathbb{Z}}$.
 - (ii) $f(\lambda, \delta)^{-1}$ is continuous in $(\lambda, \delta) \in [-\pi, \pi] \times \Lambda$, where $\Lambda \subset \mathbb{R}^{r+s+1}$ is a compact.
 - (iii) $\mu_L(\{\lambda; f(\lambda, \delta) \neq f(\lambda, \delta_0)\}) \geq 0$, for $\delta \in \Lambda$ with μ_L the Lebesgue measure.

- (H_2)
- (i) δ_0 is an interior point of Λ and in a neighborhood of δ_0 , $\frac{\partial f(\lambda, \delta)^{-1}}{\partial \delta}$ and $\frac{\partial^2 f(\lambda, \delta)^{-1}}{\partial \delta \partial \delta^T}$ exist and are continuous in λ and δ .
 - (ii) $\frac{\partial f(\lambda, \delta_0)^{-1}}{\partial \delta}$ is K -Lipchitzienne with $K > \frac{1}{2}$.
 - (iii) The matrix W given by

$$W = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \log f(\lambda, \delta_0)}{\partial \delta} \frac{\partial \log f(\lambda, \delta_0)}{\partial \delta^T} d\lambda \quad (8)$$

is nonsingular.

Theorem 2.3. Let $(X_t)_{t \in \mathbb{Z}}$ be a stationary, causal and invertible process defined by Eqs. (1), (2).

- (i) Under (H_0) with $J = 4$ and (H_1) , $\hat{\delta}_T \xrightarrow{P} \delta_0$, as $T \rightarrow \infty$.
- (ii) Under (H_0) with $J = 8$, (H_1) and (H_2) , $\sqrt{T}(\hat{\delta}_T - \delta_0) \xrightarrow{D} N(0, 2W^{-1} + W^{-1}VW^{-1})$, as $T \rightarrow \infty$. Here V is given by

$$V = \frac{2\pi}{\sigma^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\partial f(\lambda, \delta_0)^{-1}}{\partial \delta} \frac{\partial f(\omega, \delta_0)^{-1}}{\partial \delta^T} h(\lambda, -\omega, \omega) d\lambda d\omega, \quad (9)$$

with $h(\lambda, \omega, \nu) = \frac{1}{8\pi^3} \sum_{j,k,l=-\infty}^{+\infty} e^{ij\lambda - ik\omega - il\nu} \text{Cum}(\varepsilon_0, \varepsilon_j, \varepsilon_k, \varepsilon_l)$, and Cum is the order four's cumulant for the process $(\varepsilon_t)_{t \in \mathbb{Z}}$.

The proof of Theorem 2.3 is similar to the proofs of Theorems 2.1 and 2.2 given in Giraitis and Robinson's [6].

3. Proof of Theorem 2.1

Here, we assume that the process $(\varepsilon_t)_{t \in \mathbb{Z}}$ have a conditional Gaussian distribution. We will first show (iii). The strict stationarity and ergodicity of the process $(X_t)_{t \in \mathbb{Z}}$ and $(\varepsilon_t)_{t \in \mathbb{Z}}$ imply the consistency of the information matrices $\widehat{\Omega}_{\gamma}$ and $\widehat{\Omega}_{\delta}$.

In order to proof (ii), we need to check the following Basawa's conditions (see Basawa, Feign and Heyde [1]):

- $\frac{1}{T} \sum_{t=1}^T \frac{\partial \ell_t(\omega_0)}{\partial \omega} \xrightarrow{P} 0$,
 - there exists a nonrandom positive definite matrix $M(\omega_0)$ such that for all $\epsilon > 0$,
- $$\mathbb{P}\left(-\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \ell_t}{\partial \omega \partial \omega^T} \geq M(\omega_0)\right) > 1 - \epsilon \quad \text{for all } T > T_1(\epsilon),$$
- there exists a constant $M < \infty$ such that $E|\frac{\partial^3 \ell_t(\omega)}{\partial \omega_i \partial \omega_j \partial \omega_k}| < M$ for all $\omega \in \Theta$, where ω_i is the i th component of ω .

From $\frac{\partial \ell_t}{\partial \gamma} = \frac{1}{2h_t} (\frac{\varepsilon_t^2}{h_t} - 1) \frac{\partial h_t}{\partial \gamma} - \frac{\varepsilon_t}{h_t} \frac{\partial \varepsilon_t}{\partial \gamma}$ and $\frac{\partial \ell_t}{\partial \delta} = \frac{1}{2h_t} (\frac{\varepsilon_t^2}{h_t} - 1) \frac{\partial h_t}{\partial \delta}$, we have $E(\frac{\partial \ell_t}{\partial \omega})_{\omega=\omega_0} = 0$ and using the ergodic theorem, Basawa's first condition follows.

The matrix Ω_0 is definite positive and hence the second Basawa's condition holds.

Now the last conditions of Basawa is obtained by differentiating $\frac{\partial^2 \ell_t}{\partial \omega \partial \omega^T}$ and using $E(\frac{\partial^3 \varepsilon_t}{\partial d_i \partial d_j \partial d_k})^2 < \infty$.

Condition (ii) follows.

Using (iii) of Theorem 2.1, we get

$$\frac{1}{T} \sum_{t=1}^T \left(\frac{\partial \ell_t}{\partial \omega} \frac{\partial \ell_t}{\partial \omega^T} \right)_{\omega=\omega_0} \xrightarrow{\text{a.s.}} \Omega_0.$$

Let S_T defined by $S_T = \sum_{t=1}^T (b_0 \frac{\partial \ell_t}{\partial \omega})_{\omega=\omega_0}$, where b_0 is an arbitrary constant vector and $b_0 b_0^T \neq 0$. Then, S_T is a martingal with $\frac{1}{T} E(S_T^2) = b_0^T \Omega_0 b_0 > 0$. Now, using the strict stationarity and ergodicity of the process $(X_t)_{t \in \mathbb{Z}}$ and $(\varepsilon_t)_{t \in \mathbb{Z}}$, we get $[\frac{1}{T} E(S_T)]^{-1} [\frac{1}{T} E(S_T^2 | F_{t-1})] \xrightarrow{\text{a.s.}} 1$. From the Central Limit Theorem of Stout [10], the asymptotic normality convergence of the CSS estimators is derived.

Remark. For applications and more details see Diongue, Guégan and Vignal [4].

References

- [1] I.V. Basawa, P.D. Feigin, C.C. Heyde, Asymptotic properties of maximum likelihood estimators for stochastic processes, *Sankhya Ser. A* 38 (1976) 259–270.
- [2] C.-F. Chung, Estimating a generalized long-memory process, *J. Econometrics* 73 (1996) 237–259.
- [3] C.-F. Chung, A generalized fractionally integrated ARMA process, *J. Time Series Anal.* 17 (1994) 111–140.
- [4] A.K. Diongue, D. Guégan, B. Vignal, Processus GIGARCH: Estimation et applications aux prix spot de l'électricité, Preprint MORA, 14, 2003.
- [5] L. Ferrara, D. Guégan, Comparison of parameter estimation methods in cyclical long-memory time series, in: C. Dunis, J. Timmermann (Eds.), *Development in Forecasts Combination and Portfolio Choice*, Wiley, 2001.
- [6] L. Giraitis, P.M. Robinson, Whittle estimation of ARCH models, *Econometrics Theory* 17 (2001) 608–631.
- [7] D. Guégan, A new model: The k -factor GIGARCH process, *J. Signal Process.* 4 (2000) 265–271.
- [8] D. Guégan, A prospective study of the k -factor Gegenbauer process with heteroscedastic errors and an application to inflation rates, *Finance India* 17 (2003) 1–20.
- [9] Y. Hosoya, A Limit theory for long-range dependence and statistical inference on related models, *Ann. Stat.* 25 (1997) 105–137.
- [10] W.F. Stout, *Almost Sure Convergence*, Academic Press, New York, 1974.