Consistent model selection criteria and goodness-of-fit test for common time series models

Joint work with K. Kare (Paris 1) and W. Kengne (Cergy)

Jean-Marc Bardet, SAMM, Université Paris 1, France
bardet@univ-paris1.fr

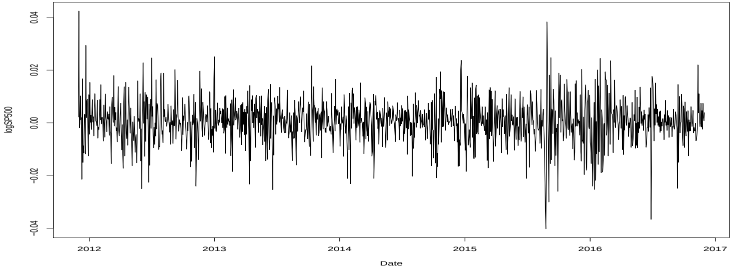
MMMS2 CIRM Conference

2 June 2020

Outline

Example

Observe the trajectory of the logarithmic returns of S&P 500 :



$\Longrightarrow$ **Two aims :**

- Chose an "optimal" model for these data ;
- Test its goodness-of-fit.

Two intuitive definitions

Let $(X_t)_{t\in\mathbb{Z}}$ be a time series (sequence of r.v. on $(\Omega, \mathcal{A}, \mathbb{P})$)

- $(X_t)_{t\in\mathbb{Z}}$ is a stationary process if $\forall k \in \mathbb{N}^*$, $\forall (t_1, \ldots, t_k) \in \mathbb{Z}^k$,

$$(X_{t_1}, \ldots, X_{t_k}) \stackrel{\mathcal{L}}{\sim} (X_{t_1+h}, \ldots, X_{t_k+h}) \quad \text{for all } h \in \mathbb{Z}.$$

- Assume that $(\xi_t)_{t\in\mathbb{Z}}$ is a white noise (centered i.i.d.r.v.)

  $(X_t)_{t\in\mathbb{Z}}$ causal process if $\exists H : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ such as $X_t = H\big((\xi_{t-k})_{k\geq 0}\big)$.

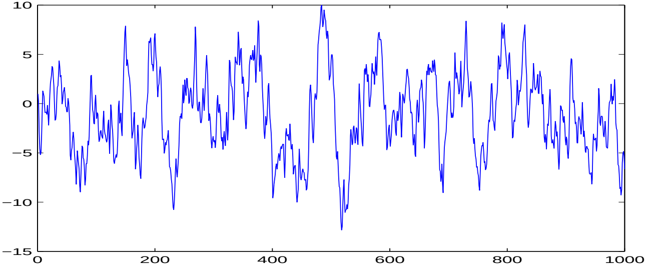
ARMA processes (Whittle, 1951)

With $(\xi_t)_{t\in\mathbb{Z}}$ a white noise, $(a_i)\in\mathbb{R}^p$, $(b_j)\in\mathbb{R}^q$

- ARMA(p,q) process : with $a_p\neq 0$ and $b_q\neq 0$, for any $t\in\mathbb{Z}$

$$X_t+a_1X_{t-1}+\cdots+a_pX_{t-p}=\xi_t+b_1\xi_{t-1}+\cdots+b_q\xi_{t-q}$$

- Stationarity and causality : $1+a_1z+\cdots+a_pz^p\neq 0$ for any $|z|\leq 1$.



ARMA$(1,1)$

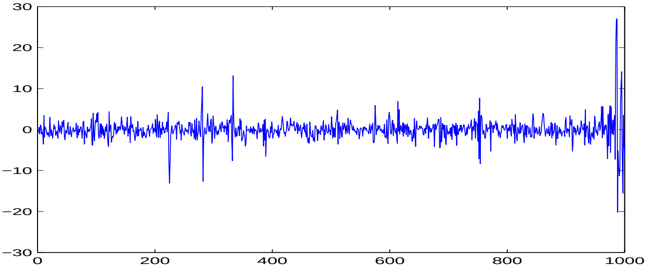
GARCH processes (Engel, 1982) (Bollersev, 1986)

With $(\xi_t)_{t\in\mathbb{Z}}$ a white noise, $(c_i) \in \mathbb{R}_+^p$, $(d_j) \in \mathbb{R}_+^q$

- GARCH(p,q) process : with $c_0,\ c_p > 0$ and $d_q > 0$, for any $t \in \mathbb{Z}$

$$\begin{cases} X_t &= \sigma_t\,\xi_t, \\ \sigma_t^2 &= c_0 + c_1 X_{t-1}^2 + \cdots + c_p X_{t-p}^2 + d_1\sigma_{t-1}^2 + \cdots + d_q\sigma_{t-q}^2 \end{cases}$$

- $\sum_{j=1}^q d_j + \mathbb{E}(\xi_0^2)\sum_{i=1}^p c_i < 1 \implies$ Stationarity and causality



GARCH(1, 1)

Model selection and Goodness-of-fit test

Consider a family $\mathcal{M}$ of models. For instance,

$$\mathcal{M} = \Big\{ \text{ARMA}(p,q) \text{ or } \text{GARCH}(p',q'), \qquad \text{with } 0 \le p,p' \le p_{\max},\ 0 \le q,q' \le q_{\max} \Big\}$$

We want to :

- Chose an "optimal" model in $\mathcal{M}$ for $(X_1,\dots,X_n)$;
- Estimate its parameters ;
- Test its goodness-of-fit.

Outline

Examples : Causal AR[∞] and ARCH(∞) models

With $(\xi_t)_{t\in\mathbb{Z}}$ a white noise,

- AR(∞) processes $X_t = \sum_{i=1}^{\infty} \theta_i X_{t-i} + \xi_t$

  $\Longrightarrow$ Causal ARMA(p, q) processes $X_t + \sum_{i=1}^{p} a_i X_{t-i} = \xi_t + \sum_{i=1}^{q} b_i \xi_{t-i}$.

- ARCH(∞) processes, (Robinson, 1991), with $b_0 > 0$ and $b_j \geq 0$

$$\left\{ \begin{array}{lcl} X_t & = & \sigma_t \xi_t, \\ \sigma_t^2 & = & \phi_0 + \sum_{j=1}^{\infty} \phi_j X_{t-j}^2. \end{array} \right.$$

  $\Longrightarrow$ GARCH(p, q) processes, with $c_0 > 0$, c_j, $d_j \geq 0$, c_p, $d_q > 0$

$$\left\{ \begin{array}{lcl} X_t & = & \sigma_t \xi_t, \\ \sigma_t^2 & = & c_0 + \sum_{j=1}^{p} c_j X_{t-j}^2 + \sum_{j=1}^{q} d_j \sigma_{t-j}^2 \end{array} \right.$$

A common frame for studying time series

A common class of models for AR, ARMA, ARCH and GARCH processes :

Causal affine models : class $\mathcal{CA}(M, f)$

$$X_t = M(X_{t-1}, X_{t-2}, \ldots)\,\xi_t + f(X_{t-1}, X_{t-2}, \ldots), \quad \forall\, t \in \mathbb{Z}, \text{ a.s..}$$

- $M(\cdot)$ and $f(\cdot)$ are real valued function on $\mathbb{R}^{\mathbb{N}}$;
- $(\xi_t)_{t\in\mathbb{Z}}$ a white noise with $\mathbb{E}(\xi_0) = 0$ and $\mathbb{E}\big(|\xi_0|^r\big) < \infty$, $r \geq 1$.

Extensions of univariate ARCH models

- TGARCH(∞) processes, (Zakoïan, 1994), with $b_0,\ b_j^+,\ b_j^- \geq 0$

$$\begin{cases} X_t &= \sigma_t\, \xi_t, \\ \sigma_t &= b_0 + \sum_{j=1}^{\infty} \left[b_j^+ \max(X_{t-j}, 0) - b_j^- \min(X_{t-j}, 0) \right] \end{cases} .$$

- APARCH(δ, p, q) processes, (Ding *et al.*, 1993)

$$\begin{cases} X_t &= \sigma_t\, \zeta_t, \\ \sigma_t^\delta &= \omega + \sum_{j=1}^{p} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta, \end{cases}$$

with $\delta \geq 1$, $\omega > 0$, $-1 < \gamma_i < 1$ and $\alpha_i,\ \beta_j \geq 0$.

Combinations of models

- ARMA-GARCH processes, (Ding *et al.*, 1993, Ling and McAleer, 2003)

$$\begin{cases} X_t &= \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \\ \varepsilon_t &= \sigma_t \, \zeta_t, \quad \text{with} \quad \sigma_t^2 = c_0 + \sum_{i=1}^{p'} c_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q'} d_j \sigma_{t-j}^2 \end{cases}$$

- ARMA-APARCH processes, (Ding *et al.*, 1993)

$$\begin{cases} X_t &= \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \\ \varepsilon_t &= \sigma_t \, \zeta_t, \quad \text{with} \quad \sigma_t^\delta = \omega + \sum_{j=1}^{p'} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^{q'} \beta_j \sigma_{t-j}^\delta \end{cases}$$

Existence and stationarity of causal affine models

$$X_t = M(X_{t-1}, X_{t-2}, \dots)\,\xi_t + f(X_{t-1}, X_{t-2}, \dots), \quad \forall\ t \in \mathbb{Z},$$

We will assume that f and M satisfy Lipschitzian conditions :

$$\begin{cases} |f(x) - f(y)| & \leq & \sum_{j=1}^{\infty} \alpha_j(f)|x_j - y_j| \\ |M(x) - M(y)| & \leq & \sum_{j=1}^{\infty} \alpha_j(M)|x_j - y_j|. \end{cases}$$

for $x = (x_j)_{j\in\mathbb{N}}$ and $y = (y_j)_{j\in\mathbb{N}}$ two sequences of $\mathbb{R}^{\infty}$.

Proposition (from Doukhan and Wintenberger, 2007)

If $\sum_{j=1}^{\infty} \alpha_j(f) + \big(\mathbb{E}(|\xi_0|^r)\big)^{1/r} \sum_{j=1}^{\infty} \alpha_j(M) < 1$, there exists a unique causal solution $(X_t)_{t\in\mathbb{Z}}$ which is stationary, ergodic, such as $\mathbb{E}(|X_0|^r) < \infty$.

Examples

Conditions on stationarity become :

- Causal AR[∞] :
  $X_t = \sum_{j=0}^{\infty} a_j \xi_{t-j} \Longrightarrow \sum_{j=0}^{\infty} |a_j| < 1$;
- Causal ARCH[∞] :
  $X_t = \xi_t \sqrt{c_0 + \sum_{j=1}^{\infty} c_j X_{t-j}^2} \Longrightarrow \left(\mathbb{E}\left[|\xi_0|^r\right]\right)^{1/r} \sum_{j=1}^{\infty} c_j < 1$;
- Causal TARCH[∞] :
  $$X_t = \xi_t \left(b_0 + \sum_{j=1}^{\infty} \left[b_j^+ \max(X_{t-j}, 0) - b_j^- \min(X_{t-j}, 0)\right]\right)$$
  $$\Longrightarrow \left(\mathbb{E}\left[|\xi_0|^r\right]\right)^{1/r} \sum_{j=1}^{\infty} \max\left(b_j^-, b_j^+\right) < 1 ;$$

Outline

Gaussian QMLE of causal affine model

Let $(X_1, \ldots, X_n)$ an observed trajectory of an $\mathcal{CA}(M_{\theta^*}, f_{\theta^*})$

$$X_t = M_{\theta^*}(X_{t-1}, X_{t-2}, \ldots)\,\xi_t + f_{\theta^*}(X_{t-1}, X_{t-2}, \ldots), \quad \forall\, t \in \mathbb{Z}$$

- With $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \ldots)$, $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \ldots)$,

  Gaussian conditional log-density : $q_t(\theta) = -\dfrac{1}{2}\Big[\dfrac{(X_t - f_\theta^t)^2}{(M_\theta^t)^2} + \log(M_\theta^t)^2\Big]$
- Let $\widehat{f}_\theta^t = f_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$ and $\widehat{M}_\theta^t = M_\theta(X_{t-1}, \ldots, X_1, 0, \cdots)$

$$\widehat{q}_t(\theta) = -\frac{1}{2}\Big[\frac{(X_t - \widehat{f}_\theta^t)^2}{(\widehat{M}_\theta^t)^2} + \log\big((\widehat{M}_\theta^t)^2\big)\Big].$$

$\Longrightarrow$ Gaussian QMLE : $\widehat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \widehat{L}_n(\theta)$ with $\widehat{L}_n(\theta) = \sum_{t=1}^n \widehat{q}_t(\theta)$.

Assumptions and strong consistency

We assume :

- **C0** : $r \geq 2$ and $\mathbb{E}(\xi_0^2) = 1$;
- **C1** : Θ is a compact set included in

$$\Theta(r) = \Big\{ \theta \in \mathbb{R}^d \ \Big/ \ \sum_{j=1}^{\infty} \alpha_j^{(0)}(f_\theta) + \big(\mathbb{E}(|\xi_0|^r)\big)^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M_\theta) < 1 \Big\}.$$

- **C2** : $\exists \underline{M} > 0$ such that $M_\theta(x) \geq \underline{M}$ for all $\theta \in \Theta$, $x \in \mathbb{R}^{\mathbb{N}}$.
- **C3** : M_θ and f_θ are such that for all $\theta_1, \ \theta_2 \in \Theta$, then :

$$\big(M_{\theta_1} = M_{\theta_2} \quad \text{and} \quad f_{\theta_1} = f_{\theta_2}\big) \quad \Longrightarrow \quad \theta_1 = \theta_2$$

- **A(K_θ, Θ)** : There exists $\big(\alpha_j(K_\theta, \Theta)\big)_j$ such that $\forall x, \ y \in \mathbb{R}^\infty$

$$\sup_{\theta \in \Theta} \big| K_\theta(x) - K_\theta(y) \big| \leq \sum_{j=1}^{\infty} \alpha_j(K_\theta, \Theta) |x_j - y_j|,$$

$$\text{with } \sum_{j=1}^{\infty} \alpha_j(K_\theta, \Theta) < \infty.$$

Strong consistency

Théorème (Bardet and Wintenberger, 2009)

Assume $r \geq 2$, $\Theta \subset \Theta(2)$, *Conditions C0-3 and* $A(f_\theta, \Theta)$ *and* $A(M_\theta, \Theta)$ *with*

$$\alpha_j(f_\theta, \Theta) + \alpha_j(M_\theta, \Theta) = O\big(j^{-\ell}\big) \quad \text{for some} \quad \ell > \min(1, 3/r).$$

Then the QMLE $\widehat{\theta}_n$ *is strongly consistent, i.e.* $\widehat{\theta}_n \xrightarrow[n\to\infty]{a.s.} \theta^*$.

Asymptotic normality

Théorème (Bardet and Wintenberger, 2009)

Under conditions of SLLN, and if $r \geq 4$, *if* $\theta^* \in \overset{\circ}{\Theta} \cap \Theta(4)$ *and if* $\mathbf{A}(K_\theta, \Theta)$, $\mathbf{A}(\partial_\theta K_\theta, \Theta)$ *and* $\mathbf{A}(\partial^2_{\theta^2} K_\theta, \Theta)$ *hold for* $K_\theta = f_\theta$ *or* M_θ, *and if*

$$\alpha_j(\partial_\theta f_\theta, \Theta) + \alpha_j(\partial_\theta M_\theta, \Theta) = O\big(j^{-\ell'}\big) \quad \text{for some} \quad \ell' > 1, \tag{1}$$

then the QMLE $\widehat{\theta}_n$ *is asymptotically normal, i.e., there exists matrix* $F(\theta^*)^{-1}$ *and* $G(\theta^*)$ *such that*

$$\sqrt{n}\big(\widehat{\theta}_n - \theta^*\big) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\big(0\ ,\ F(\theta^*)^{-1} G(\theta^*) F(\theta^*)^{-1}\big). \tag{2}$$

- Could be applied to all cited processes ARMA, ARCH, APARCH,...
- But requires $r \geq 4$ and not very robust.

Outline

Additivity of causal affine models

Proposition

Let $\Theta_1 \subset \mathbb{R}^{d_1}$, $\Theta_2 \subset \mathbb{R}^{d_1}$, $M^{(1)}_{\theta_1}, f^{(1)}_{\theta_1}, M^{(2)}_{\theta_2}, f^{(2)}_{\theta_2}$ *for* $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$.
There exist $\max(d_1, d_2) \leq d \leq d_1 + d_2$, $\Theta \subset \mathbb{R}^d$, *and* M_θ, f_θ *with* $\theta \in \Theta$, *such as for any* $\theta_1 \in \Theta_1 \subset \mathbb{R}^{d_1}$ *and* $\theta_2 \in \Theta_2 \subset \mathbb{R}^{d_2}$,

$$\Big\{\mathcal{CA}\big(M^{(1)}_{\theta_1}, f^{(1)}_{\theta_1}\big) \bigcup \mathcal{CA}\big(M^{(2)}_{\theta_2}, f^{(2)}_{\theta_2}\big)\Big\} \subset \Big\{\mathcal{CA}\big(M_\theta, f_\theta\big)\Big\}.$$

Consequence :

- For any family $\mathcal{M} = \bigcup_{i\in I} \mathcal{CA}\big(M^{(i)}_{\theta_i}, f^{(i)}_{\theta_i}\big)$,

$$\Longrightarrow \quad \mathcal{M} = \bigcup_{i\in I} \Big\{\mathcal{CA}\big(M_\theta, f_\theta\big)\Big\}_{\theta\in\Theta_i\subset\mathbb{R}^d}$$

- $\mathcal{M}$ family of $\mathcal{CA}$ models $\Leftrightarrow$ $\mathcal{M} \sim \{m \subset \{1, \ldots, d\}\big)$,

$$\theta \in \Theta(m) \subset \Big\{(x_1, \ldots, x_d) \in \mathbb{R}^d,\ x_i = 0 \text{ if } i \notin m\Big\}$$

Penalized Quasi-Maximum Likelihood criterion

Let $(X_1, \ldots, X_n)$ an observed trajectory.

For $m \in \mathcal{M}$, define :

$$\begin{cases} \widehat{\theta}(m) & = & \underset{\theta \in \Theta(m)}{\text{argmax}} \, \widehat{L}_n(\theta) \\ \widehat{m} & = & \underset{m \in \mathcal{M}}{\text{argmin}} \, \widehat{C}(m) \quad \text{with} \quad \widehat{C}(m) = -2\widehat{L}_n(\widehat{\theta}(m)) + |m| \, \kappa_n, \end{cases}$$

using

- $(\kappa_n)_n$ an increasing sequence of positive real numbers ;
- $|m|$ denotes the cardinal of m, subset of $\{1, \ldots, d\}$.

Consistency

Théorème

Let $(X_1, \ldots, X_n)$ *be an observed trajectory of* $\mathcal{CA}(M_{\theta^*}, f_{\theta^*})$ *where* θ^* *unknown in* $\Theta \subset \Theta(r) \subset \mathbb{R}^d$ *with* $r \geq 4$. *Under previous assumptions and if*

$$\sum_{k\geq 1} \frac{1}{\kappa_k} \sum_{j\geq k} \alpha_j(f_\theta, \Theta) + \alpha_j(M_\theta, \Theta) + \alpha_j(\partial_\theta f_\theta, \Theta) + \alpha_j(\partial_\theta M_\theta, \Theta) < \infty,$$

$$\text{then} \qquad \mathbb{P}(\widehat{m} = m^*) \underset{n\to+\infty}{\longrightarrow} 1.$$

Consequence :

- If $\alpha_j(f_\theta, \Theta) + \alpha_j(M_\theta, \Theta) + \alpha_j(\partial_\theta f_\theta, \Theta) + \alpha_j(\partial_\theta M_\theta, \Theta) = \mathcal{O}(\rho^j)$, $|\rho| < 1$, $\kappa_n \to \infty$ sufficient : BIC for ARMA, GARCH, APARCH,..., processes.
- If $\alpha_j(f_\theta, \Theta) + \alpha_j(M_\theta, \Theta) + \alpha_j(\partial_\theta f_\theta, \Theta) + \alpha_j(\partial_\theta M_\theta, \Theta) = \mathcal{O}(j^{-\gamma})$, $\gamma > 1$, $\kappa_n = \mathcal{O}(n^\delta)$ with $\delta > 2 - \gamma$: not valid for BIC for AR(∞), ARCH(∞),...

Estimation of the parameters

Théorème

Under the assumptions of the previous Theorem, then

$$\sqrt{n}\left(\left(\widehat{\theta}_n(\widehat{m})\right)_i - (\theta^*)_i\right)_{i\in\widehat{m}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_d\left(0\ ,\ F(\theta^*,m^*)^{-1}G(\theta^*,m^*)F(\theta^*,m^*)^{-1}\right)$$

where F and G are defined in CLT.

$\Longrightarrow$ Same convergence rate with or without the knowledge of the model

Portmanteau goodness-of-fit test (1)

Define :

- Residuals : $\widehat{\xi}_k = \dfrac{X_k - \widehat{f}^t_{\widehat{\theta}(\widehat{m})}}{\widehat{M}^t_{\widehat{\theta}(\widehat{m})}}$
- Covariogram of square residuals : $\widehat{r}(k) = \dfrac{1}{n} \sum_{j=1}^{n-k} \widehat{\xi}_j^2 \widehat{\xi}_{j+k}^2 - 1$;
- Correlogram of square residuals : $\widehat{\rho}(k) = \dfrac{\widehat{r}(k)}{\widehat{r}(0)}$;

Portmanteau goodness-of-fit test (2)

Théorème

Under the assumptions of Theorem and if $\mathbb{E}(\xi_0^3) = 0$ *:*

1. *With* $V(\theta^*)$ *an explicit definite positive matrix, we have :*

$$\sqrt{n}\left(\widehat{\rho}(1), \ldots, \widehat{\rho}(K)\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_K\left(0\,,\; V(\theta^*)\right).$$

2. *If we define* $\widehat{Q}_K = n\ {}^t\!\left(\widehat{\rho}(1), \ldots, \widehat{\rho}(K)\right)\left(\widehat{V}(\widehat{\theta}(\widehat{m}))\right)^{-1}\left(\widehat{\rho}(1), \ldots, \widehat{\rho}(K)\right)$,

$$\text{then} \qquad \widehat{Q}_K \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2(K).$$

$$\Longrightarrow \text{Test} \left\{ \begin{array}{l} H_0 \,:\; X \in \mathcal{AC}(M_{\theta^*}, f_{\theta^*}) \\ H_1 \,:\; X \notin \mathcal{AC}(M_{\theta^*}, f_{\theta^*}) \end{array} \right.$$

Outline

Simulation results for classical models

1. Model 1, AR(2) : $X_t = 0.4X_{t-1} + 0.4X_{t-2} + \xi_t$
2. Model 2, ARMA(1,1) : $X_t = 0.3X_{t-1} + \xi_t + 0.5\xi_{t-1}$
3. Model 3, ARCH(2) : $X_t = \xi_t \sqrt{0.2 + 0.4X_{t-1}^2 + 0.2X_{t-2}^2}$

	n	100			500			1000			2000		
		$\log n$	$\sqrt{n}$	$\widehat{\kappa}_n$	$\log n$	$\sqrt{n}$	$\widehat{\kappa}_n$	$\log n$	$\sqrt{n}$	$\widehat{\kappa}_n$	$\log n$	$\sqrt{n}$	$\widehat{\kappa}_n$
	W	21.4	32.3	18.4	1.7	0.8	0.9	0.8	0.1	0.1	0.2	0	0
M1	T	74.2	67.6	79.7	97.2	99.2	99.1	98.2	99.9	99.9	99.2	100	100
	O	4.4	0.1	1.9	1.1	0	0	1.0	0	0	0.6	0	0
	W	30.4	57.7	28.0	4.8	4.2	4.0	0.7	0.3	0.3	0.4	0	0
M2	T	64.1	42.1	67.3	93.6	95.8	95.8	98.2	99.7	99.6	99.2	100	100
	O	5.5	0.2	4.7	1.6	0	0.2	1.1	0	0.1	0.4	0	0
	W	76.1	90.8	53.5	27.3	67.1	18.0	14.0	41.5	13.3	4.6	12.0	4.6
M3	T	23.8	9.2	39.8	72.7	32.9	79.9	85.9	58.5	86.7	95.4	88.0	95.4
	O	0.1	0	6.7	0	0	2.1	0.1	0	0	0	0	0
	W	83.8	94.3	73.4	22.1	61.5	20.4	5.8	31.3	5.7	1.8	6.2	0.7

Simulation results for classical models

1. Model 1, AR(2) : $X_t = 0.4X_{t-1} + 0.4X_{t-2} + \xi_t$
2. Model 2, ARMA(1,1) : $X_t = 0.3X_{t-1} + \xi_t + 0.5\xi_{t-1}$
3. Model 3, ARCH(2) : $X_t = \xi_t \sqrt{0.2 + 0.4X_{t-1}^2 + 0.2X_{t-2}^2}$

n		100		500		1000		2000	
		size	power	size	power	size	power	size	power
$K = 3$	Model 1	3.3	10.9	6.2	52.2	3.5	84.8	5.0	98.2
	Model 2	3.3	7.0	4.8	23.3	6.2	42.4	4.9	70.4
	Model 3	4.6	6.4	8.4	44.1	14.3	81.0	36.9	99.4
$K = 6$	Model 1	2.9	9.1	4.9	42.0	4.4	76.3	4.5	97.6
	Model 2	3.0	6.3	5.2	18.0	5.1	35.1	4.6	60.2
	Model 3	4.5	12.6	11.1	64.4	14.7	92.5	27.9	99.9

Simulation results for non hierarchical models

1. Model 4 : $X_t = 0.4X_{t-3} + 0.4X_{t-4} + \xi_t$.

n	100			500			1000			2000		
	$\log n$	$\sqrt{n}$	$\widehat{\kappa_n}$	$\log n$	$\sqrt{n}$	$\widehat{\kappa_n}$	$\log n$	$\sqrt{n}$	$\widehat{\kappa_n}$	$\log n$	$\sqrt{n}$	$\widehat{\kappa_n}$
T	70.4	67.3	71.0	90	100	100	93.2	100	100	95.3	100	100
O	25.0	1.6	28.8	10	0	0	6.8	0	0	4.7	0	0
W	4.6	31.1	0.2	0	0	0	0	0	0	0	0	0

Numerical results for SP500 log-returns

Log-return of SP500 closing values from 11/2011 → 11/2016

Table – Results of the model selection and goodness-of-fit analysis on FTSE index.

	$\kappa_n = \log(n)$	$\kappa_n = \sqrt{n}$	$\kappa_n = \widehat{\kappa}_n$
$\widehat{m}$	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)
$\widehat{Q}_{10}(\widehat{m})$	9.30	9.30	9.30
$p - value$	0.50	0.50	0.50

References

- Bardet, Kare and Kengne (2020). Consistent model selection criteria and goodness-of-fit test for common time series models. *Elec. Journ. Statist.*
- Bardet and Wintenberger (2009). Asymptotic normality of the Quasi-Maximum Likelihood Estimator for multidimensional causal processes. *Ann. Statist.*
- Doukhan and Wintenberger (2008). Weakly dependent chains with infinite memory. *Stoch. Proc. and their Appli.*
- Francq and Zakoïan (2010). *GARCH Models*. Wiley.
- Hsu, Ing and Tong. On model selection from a finite family of possibly misspecified time series models. *Ann. of Statist.*
- Li and Mak (1991). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis*
- Sin and White (1996). Information criteria for selecting possibly misspecified parametric models. *Journal of Econometrics*.
- Straumann (2005). *Estimation in conditionally heteroscedastic time series models*. Springer, New York.