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

1. An example
2. Causal affine models
3. Gaussian Quasi-Maximum Likelihood Estimation
4. Consistency of a penalized QML criterion and goodness-of-fit test
5. Numerical results
Example

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

→ 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 \(\forall k \in \mathbb{N}^*, \forall (t_1, \ldots, t_k) \in \mathbb{Z}^k,\)

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

- 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 \(\exists H : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}\) such as \(X_t = H((\xi_{t-k})_{k \geq 0}).\)
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 \(t \in \mathbb{Z}\)

\[ X_t + a_1 X_{t-1} + \cdots + a_p X_{t-p} = \xi_t + b_1 \xi_{t-1} + \cdots + b_q \xi_{t-q} \]

- Stationarity and causality: \(1 + a_1 z + \cdots + a_p z^p \neq 0\) for any \(|z| \leq 1\).
GARCH processes (Engel, 1982) (Bollerslev, 1986)

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

- **GARCH($p, q$) process**: with $c_0, c_p > 0$ and $d_q > 0$, for any $t \in \mathbb{Z}$
  \[
  \begin{align*}
  X_t &= \sigma_t \xi_t, \\
  \sigma_t^2 &= c_0 + c_1 X_{t-1}^2 + \cdots + c_p X_{t-p}^2 + d_1 \sigma_{t-1}^2 + \cdots + d_q \sigma_{t-q}^2
  \end{align*}
  \]

- $\sum_{j=1}^{q} d_j + \mathbb{E}(\xi_0^2) \sum_{i=1}^{p} c_i < 1 \implies \text{Stationarity and causality}$
Model selection and Goodness-of-fit test

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

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

We want to:

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

1. An example
2. Causal affine models
3. Gaussian Quasi-Maximum Likelihood Estimation
4. Consistency of a penalized QML criterion and goodness-of-fit test
5. Numerical results
Examples: Causal AR[$\infty$] and ARCH($\infty$) models

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

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

  $\Rightarrow$ **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($\infty$) processes**, (Robinson, 1991), with $b_0 > 0$ and $b_j \geq 0$

  \[
  \begin{cases}
  X_t = \sigma_t \xi_t, \\
  \sigma_t^2 = \phi_0 + \sum_{j=1}^{\infty} \phi_j X_{t-j}^2.
  \end{cases}
  \]

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

  \[
  \begin{cases}
  X_t = \sigma_t \xi_t, \\
  \sigma_t^2 = c_0 + \sum_{j=1}^{p} c_j X_{t-j}^2 + \sum_{j=1}^{q} d_j \sigma_{t-j}^2.
  \end{cases}
  \]
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 $\mathbb{R}^N$;
- $(\xi_t)_{t \in \mathbb{Z}}$ a white noise with $\mathbb{E}(\xi_0) = 0$ and $\mathbb{E}(|\xi_0|^r) < \infty$, $r \geq 1$. 

J.-M. Bardet, Paris 1 (MMMS2 CIRM Conference)
Extensions of univariate ARCH models

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

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

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

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

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

J.-M. Bardet, Paris 1 (MMMS2 CIRM Conference)
Combinations of models

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

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

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

  \[
  \begin{aligned}
  X_t &= \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} b_j \varepsilon_{t-j}, \\
  \varepsilon_t &= \sigma_t \zeta_t, \quad \text{with} \quad \sigma_t^\delta = \omega + \sum_{j=1}^{p'} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^{q'} \beta_j \sigma_{t-j}^\delta
  \end{aligned}
  \]
Existence and stationarity of causal affine models

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

We will assume that \( f \) and \( M \) satisfy Lipschitzian conditions:

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

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 \( \sum_{j=1}^{\infty} \alpha_j(f) + \left( \mathbb{E}(|\xi_0|^r) \right)^{1/r} \sum_{j=1}^{\infty} \alpha_j(M) < 1 \), 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$[\infty]$**:
  \[ X_t = \sum_{j=0}^{\infty} a_j \xi_{t-j} \implies \sum_{j=0}^{\infty} |a_j| < 1; \]

- **Causal ARCH$[\infty]$**:
  \[ X_t = \xi_t \sqrt{c_0 + \sum_{j=1}^{\infty} c_j X_{t-j}^2} \implies \left( \mathbb{E}[|\xi_0|^r] \right)^{1/r} \sum_{j=1}^{\infty} c_j < 1; \]

- **Causal TARCH$[\infty]$**:
  \[ X_t = \xi_t \left( b_0 + \sum_{j=1}^{\infty} \left[ b_j^+ \max(X_{t-j}, 0) - b_j^- \min(X_{t-j}, 0) \right] \right) \implies \left( \mathbb{E}[|\xi_0|^r] \right)^{1/r} \sum_{j=1}^{\infty} \max(b_j^-, b_j^+) < 1; \]
Outline

1. An example
2. Causal affine models
3. Gaussian Quasi-Maximum Likelihood Estimation
4. Consistency of a penalized QML criterion and goodness-of-fit test
5. Numerical results
Gaussian QMLE of causal affine model

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

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

- 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: \(q_t(\theta) = -\frac{1}{2} \left[ \frac{(X_t - f_{\theta}^t)^2}{(M_{\theta}^t)^2} + \log((M_{\theta}^t)^2) \right] \)

- Let \(\hat{f}_{\theta}^t = f_{\theta}(X_{t-1}, \ldots, X_1, 0, \cdots)\) and \(\hat{M}_{\theta}^t = M_{\theta}(X_{t-1}, \ldots, X_1, 0, \cdots)\)

\[
\hat{q}_t(\theta) = -\frac{1}{2} \left[ \frac{(X_t - \hat{f}_{\theta}^t)^2}{(\hat{M}_{\theta}^t)^2} + \log ((\hat{M}_{\theta}^t)^2) \right].
\]

\(\implies\) Gaussian QMLE: \(\hat{\theta}_n = \arg\max_{\theta \in \Theta} \hat{L}_n(\theta)\) with \(\hat{L}_n(\theta) = \sum_{t=1}^{n} \hat{q}_t(\theta)\).
Assumptions and strong consistency

We assume:

- **C0**: \( r \geq 2 \) and \( \mathbb{E}(\xi_0^2) = 1 \);
- **C1**: \( \Theta \) is a compact set included in
  \[
  \Theta(r) = \left\{ \theta \in \mathbb{R}^d \mid \sum_{j=1}^{\infty} \alpha_j^{(0)}(f_{\theta}) + (\mathbb{E}(|\xi_0|^r))^{1/r} \sum_{j=1}^{\infty} \alpha_j^{(0)}(M_{\theta}) < 1 \right\}.
  \]
- **C2**: \( \exists M > 0 \) such that \( M_{\theta}(x) \geq M \) for all \( \theta \in \Theta, \ x \in \mathbb{R}^n \).
- **C3**: \( M_{\theta} \) and \( f_{\theta} \) are such that for all \( \theta_1, \theta_2 \in \Theta \), then:
  \[
  (M_{\theta_1} = M_{\theta_2} \text{ and } f_{\theta_1} = f_{\theta_2}) \implies \theta_1 = \theta_2.
  \]
- **A(\(K_{\theta}, \Theta\))**: There exists \((\alpha_j(K_{\theta}, \Theta))_j\) such that \( \forall x, y \in \mathbb{R}^\infty 
  \]
  \[
  \sup_{\theta \in \Theta} |K_{\theta}(x) - K_{\theta}(y)| \leq \sum_{j=1}^{\infty} \alpha_j(K_{\theta}, \Theta)|x_j - y_j|,
  \]
  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

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

Then the QMLE $\hat{\theta}_n$ is strongly consistent, i.e. $\hat{\theta}_n \xrightarrow{a.s.} \Theta^*$. 

J.-M. Bardet, Paris 1 (MMMS2 CIRM Conference)
Asymptotic normality

Théorème (Bardet and Wintenberger, 2009)

Under conditions of SLLN, and if \( r \geq 4 \), if \( \Theta^* \in \Theta \cap \Theta(4) \) and if \( A(K_\theta, \Theta), A(\partial_\theta K_\theta, \Theta) \) and \( A(\partial_\theta^2 K_\theta, \Theta) \) hold for \( K_\theta = f_\theta \) or \( M_\theta \), and if

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

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

\[
\sqrt{n}(\hat{\Theta}_n - \Theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}_d(0, F(\Theta^*)^{-1} G(\Theta^*) F(\Theta^*)^{-1}). \tag{2}
\]

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

1. An example
2. Causal affine models
3. Gaussian Quasi-Maximum Likelihood Estimation
4. Consistency of a penalized QML criterion and goodness-of-fit test
5. Numerical results
Additivity of causal affine models

Proposition

Let $\Theta_1 \subset \mathbb{R}^{d_1}$, $\Theta_2 \subset \mathbb{R}^{d_2}$, $M_{\theta_1}^{(1)}$, $f_{\theta_1}^{(1)}$, $M_{\theta_2}^{(2)}$, $f_{\theta_2}^{(2)}$ 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 $M_{\theta}$, $f_{\theta}$ with $\theta \in \Theta$, such as for any $\theta_1 \in \Theta_1 \subset \mathbb{R}^{d_1}$ and $\theta_2 \in \Theta_2 \subset \mathbb{R}^{d_2}$,

$$\{CA(M_{\theta_1}^{(1)}, f_{\theta_1}^{(1)}) \cup CA(M_{\theta_2}^{(2)}, f_{\theta_2}^{(2)})\} \subset \{CA(M_{\theta}, f_{\theta})\}.$$ 

Consequence:

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

$$\Rightarrow \mathcal{M} = \bigcup_{i \in I} \left\{CA(M_{\theta}, f_{\theta})\right\}_{\theta \in \Theta_1 \subset \mathbb{R}^d}$$

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

$$\theta \in \Theta(m) \subset \{(x_1, \ldots, x_d) \in \mathbb{R}^d, x_i = 0 \text{ if } i \notin m\}$$
Penalized Quasi-Maximum Likelihood criterion

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

For \(m \in \mathcal{M}\), define:

\[
\begin{align*}
\hat{\theta}(m) &= \arg\max_{\theta \in \Theta(m)} \hat{L}_n(\theta) \\
\hat{m} &= \arg\min_{m \in \mathcal{M}} \hat{C}(m) \quad \text{with} \quad \hat{C}(m) = -2\hat{L}_n(\hat{\theta}(m)) + |m| \kappa_n,
\end{align*}
\]

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 \(CA(M_{\theta^*}, f_{\theta^*})\) where \(\theta^*\) unknown in \(\Theta \subset \Theta(r) \subset \mathbb{R}^d\) with \(r \geq 4\). Under previous assumptions and if

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

then

\[
\mathbb{P}(\hat{m} = m^*) \xrightarrow{n \to +\infty} 1.
\]

Consequence :

- If \(\alpha_j(f_{\theta}, \Theta) + \alpha_j(M_{\theta}, \Theta) + \alpha_j(\partial_{\theta} f_{\theta}, \Theta) + \alpha_j(\partial_{\theta} M_{\theta}, \Theta) = O(\rho^j), |\rho| < 1, \kappa_n \to \infty\) sufficient: BIC for ARMA, GARCH, APARCH, ..., processes.

- If \(\alpha_j(f_{\theta}, \Theta) + \alpha_j(M_{\theta}, \Theta) + \alpha_j(\partial_{\theta} f_{\theta}, \Theta) + \alpha_j(\partial_{\theta} M_{\theta}, \Theta) = O(j^{-\gamma}), \gamma > 1, \kappa_n = O(n^\delta)\) with \(\delta > 2 - \gamma\): not valid for BIC for \(AR(\infty), ARCH(\infty)\),...
Estimation of the parameters

**Théorème**

*Under the assumptions of the previous Theorem, then*

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

where $F$ and $G$ are defined in CLT.

\[\implies \text{Same convergence rate with or without the knowledge of the model}\]
Portmanteau goodness-of-fit test (1)

Define:

- Residuals: $\hat{\xi}_k = \frac{X_k - \hat{f}^t_{\hat{\theta}(\hat{m})}}{\hat{M}^t_{\hat{\theta}(\hat{m})}}$

- Covariogram of square residuals: $\hat{r}(k) = \frac{1}{n} \sum_{j=1}^{n-k} \xi_j^2 \xi_{j+k}^2 - 1$

- Correlogram of square residuals: $\hat{\rho}(k) = \frac{\hat{r}(k)}{\hat{r}(0)}$
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:
   \[
   \sqrt{n} (\hat{\rho}(1), \ldots, \hat{\rho}(K)) \xrightarrow{\mathcal{L}} N_K(0, V(\theta^*)).
   \]

2. If we define $\hat{Q}_K = n \left(\hat{\rho}(1), \ldots, \hat{\rho}(K)\right) (\hat{V}(\hat{\theta}(\hat{m})))^{-1} (\hat{\rho}(1), \ldots, \hat{\rho}(K))$,
then $\hat{Q}_K \xrightarrow{\mathcal{L}} \chi^2(K)$.

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

1. An example
2. Causal affine models
3. Gaussian Quasi-Maximum Likelihood Estimation
4. Consistency of a penalized QML criterion and goodness-of-fit test
5. Numerical results
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): \( X_t = \xi_t \sqrt{0.2 + 0.4X_{t-1}^2 + 0.2X_{t-2}^2} \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \log n )</th>
<th>( \sqrt{n} )</th>
<th>( \widehat{\kappa}_n )</th>
<th>( \log n )</th>
<th>( \sqrt{n} )</th>
<th>( \widehat{\kappa}_n )</th>
<th>( \log n )</th>
<th>( \sqrt{n} )</th>
<th>( \widehat{\kappa}_n )</th>
<th>( \log n )</th>
<th>( \sqrt{n} )</th>
<th>( \widehat{\kappa}_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>M1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.4</td>
<td>32.3</td>
<td>18.4</td>
<td>1.7</td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>74.2</td>
<td>67.6</td>
<td>79.7</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O</td>
<td>4.4</td>
<td>0.1</td>
<td>1.9</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>M2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.4</td>
<td>57.7</td>
<td>28.0</td>
<td>4.8</td>
<td>4.2</td>
<td>4.0</td>
<td>0.7</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>64.1</td>
<td>42.1</td>
<td>67.3</td>
<td>93.6</td>
<td>95.8</td>
<td>95.8</td>
<td>98.2</td>
<td>99.7</td>
<td>99.6</td>
<td>99.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>O</td>
<td>5.5</td>
<td>0.2</td>
<td>4.7</td>
<td>1.6</td>
<td>0</td>
<td>0.2</td>
<td>1.1</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>M3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76.1</td>
<td>90.8</td>
<td>53.5</td>
<td>27.3</td>
<td>67.1</td>
<td>18.0</td>
<td>14.0</td>
<td>41.5</td>
<td>13.3</td>
<td>4.6</td>
<td>12.0</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>23.8</td>
<td>9.2</td>
<td>39.8</td>
<td>72.7</td>
<td>32.9</td>
<td>79.9</td>
<td>85.9</td>
<td>58.5</td>
<td>86.7</td>
<td>95.4</td>
<td>88.0</td>
<td>95.4</td>
</tr>
<tr>
<td>O</td>
<td>0.1</td>
<td>0</td>
<td>6.7</td>
<td>0</td>
<td>0</td>
<td>2.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
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): \( X_t = \xi_t \sqrt{0.2 + 0.4X^2_{t-1} + 0.2X^2_{t-2}} \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>size</td>
<td>power</td>
<td>size</td>
<td>power</td>
</tr>
<tr>
<td>( K = 3 )</td>
<td>Model 1</td>
<td>3.3</td>
<td>10.9</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>3.3</td>
<td>7.0</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>4.6</td>
<td>6.4</td>
<td>8.4</td>
</tr>
<tr>
<td>( K = 6 )</td>
<td>Model 1</td>
<td>2.9</td>
<td>9.1</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>3.0</td>
<td>6.3</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>4.5</td>
<td>12.6</td>
<td>11.1</td>
</tr>
</tbody>
</table>
Simulation results for non hierarchical models

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

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\log n$</td>
<td>$\sqrt{n}$</td>
</tr>
<tr>
<td>T</td>
<td>70.4</td>
<td>67.3</td>
</tr>
<tr>
<td>O</td>
<td>25.0</td>
<td>1.6</td>
</tr>
<tr>
<td>W</td>
<td>4.6</td>
<td>31.1</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\log n$</td>
<td>$\sqrt{n}$</td>
<td>$k_n$</td>
</tr>
<tr>
<td>T</td>
<td>90</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>O</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\log n$</td>
<td>$\sqrt{n}$</td>
<td>$k_n$</td>
</tr>
<tr>
<td>T</td>
<td>93.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>O</td>
<td>6.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\log n$</td>
<td>$\sqrt{n}$</td>
<td>$k_n$</td>
</tr>
<tr>
<td>T</td>
<td>95.3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>O</td>
<td>4.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

J.-M. Bardet, Paris 1 (MMMS2 CIRM Conference)
Numerical results for SP500 log-returns


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

<table>
<thead>
<tr>
<th></th>
<th>$\kappa_n = \log(n)$</th>
<th>$\kappa_n = \sqrt{n}$</th>
<th>$\kappa_n = \hat{\kappa}_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{m}$</td>
<td>GARCH(1, 1)</td>
<td>GARCH(1, 1)</td>
<td>GARCH(1, 1)</td>
</tr>
<tr>
<td>$Q_{10}(\hat{m})$</td>
<td>9.30</td>
<td>9.30</td>
<td>9.30</td>
</tr>
<tr>
<td>$p - value$</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>
References

- Hsu, Ing and Tong. On model selection from a finite family of possibly misspecified time series models. *Ann. of Statist.*