A SARIMAX coupled modelling applied to individual load curves intraday forecasting

Frédéric Proïa

Workshop EDF
Institut Henri Poincaré - Paris
05 avril 2012

INRIA Bordeaux Sud-Ouest
Institut de Mathématiques de Bordeaux
1 Introduction

2 On a SARIMAX coupled modelling
   - Stationary ARMA processes
   - Identification for stationary AR\((p)\) and MA\((q)\) processes
   - Linear relationship between consumption and temperature
   - The SARIMAX modelling
   - Application to forecasting

3 Application to forecasting on a load curve
   - Procedure
   - Seasonality and stationarity
   - ACF and PACF
   - Selection criteria
   - Selection on bayesian criteria
   - Selection on bayesian criteria

4 Application on a huge dataset
   - Technical issues
   - Procedure
   - Modelling
   - Forecasting

5 Conclusion
• **Motivations.**
  - New energy meters to gather individual consumption with high frequency.
  - Economic issue for EDF: anticipate to optimize.

• **Objectives.**
  - Distinguish nonthermosensitive from thermosensitive customers.
  - Intraday daily forecasting.
  - Introduce temperature as an exogenous contribution.
• Detecting thermosensitivity.
  - Deterministic criterion.
  - Thermosensitivity if
    \[
      \max_{\Delta \leq t \leq N} M_t - \min_{\Delta \leq t \leq N} M_t > \delta
    \]
    where \( M_t \) is the empirical median of \((C_t \ldots C_{t-\Delta+1})\).
  - \( \Delta = 1440 \) et \( \delta = 1000 \).

- More pertinent than SR/DR.
  - 70% of nonthermosensitive customers, only including 75% of SR.
1. Introduction

2. On a SARIMAX coupled modelling
   - Stationary ARMA processes
   - Identification for stationary AR($p$) and MA($q$) processes
   - Linear relationship between consumption and temperature
   - The SARIMAX modelling
   - Application to forecasting

3. Application to forecasting on a load curve
   - Procedure
   - Seasonality and stationarity
   - ACF and PACF
   - Selection criteria
     - Selection on bayesian criteria
     - Selection on bayesian criteria

4. Application on a huge dataset
   - Technical issues
   - Procedure
   - Modelling
   - Forecasting

5. Conclusion
• **Stationary ARMA processes.**

**Definition (Stationarity)**

A time series \((Y_t)\) is said to be weakly stationary if, for all \(t \in \mathbb{Z}\), \(E[Y_t^2] < \infty\), \(E[Y_t] = m\) and, for all \(s, t \in \mathbb{Z}\), \(\text{Cov}(Y_t, Y_s) = \text{Cov}(Y_{t-s}, Y_0)\).

**Definition (ARMA)**

Let \((Y_t)\) be a stationary time series with zero mean. It is said to be an ARMA\((p, q)\) process if, for every \(t \in \mathbb{Z}\),

\[
Y_t - \sum_{k=1}^{p} a_k Y_{t-k} = \varepsilon_t + \sum_{k=1}^{q} b_k \varepsilon_{t-k}
\]

where \((\varepsilon_t)\) is a white noise of variance \(\sigma^2 > 0\), \(a \in \mathbb{R}^p\) and \(b \in \mathbb{R}^q\).
• **Causality of ARMA processes.**
  
  - Compact expression, for all \(1 \leq t \leq T\),
    
    \[
    \mathcal{A}(B)Y_t = \mathcal{B}(B)\varepsilon_t
    \]
    
    where the polynomials
    
    \[
    \mathcal{A}(z) = 1 - a_1z - \ldots - a_pz^p \quad \text{and} \quad \mathcal{B}(z) = 1 + b_1z + \ldots + b_qz^q.
    \]

**Definition (Causality)**

Let \((Y_t)\) be an ARMA\((p, q)\) process for which the polynomials \(\mathcal{A}\) and \(\mathcal{B}\) have no common zeroes. Then, \((Y_t)\) is causal if and only if \(\mathcal{A}(z) \neq 0\) for all \(z \in \mathbb{C}\) such that \(|z| \leq 1\).

• **Implications.**
  
  - Causality implies the existence of a MA\((\infty)\) structure for \((Y_t)\).
  - Causality implies stationarity of the process.
  - On \(\mathbb{N}^*\), causality often coincides with stationarity.
On a SARIMAX coupled modelling
Stationary ARMA processes

• **Existence and unicity of a stationary solution.**

**Proposition**

If \( A(z) \neq 0 \) for all \( z \in \mathbb{C} \) such that \( |z| \leq 1 \), then the ARMA equation \( A(B) Y_t = B(B) \varepsilon_t \) have the unique stationary solution

\[
Y_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k},
\]

and the coefficients \( (\psi_k)_{k \in \mathbb{N}} \) are determined by the relation

\[
A^{-1}(z)B(z) = \sum_{k=0}^{\infty} \psi_k z^k \quad \text{with} \quad \sum_{k=0}^{\infty} \psi_k^2 < \infty.
\]

• **Explosive cases.**
  - On \( \mathbb{Z} \), no zeroes on the unit circle is a sufficient condition.
  - Irrelevant for practical purposes.
On a SARIMAX coupled modelling
Identification for stationary AR($p$) and MA($q$) processes

- **Autocorrelation function.**
  - To identify $q$.

**Definition (ACF)**

Let $(Y_t)$ be a stationary time series. The autocorrelation function $\rho$ associated with $(Y_t)$ is defined, for all $t \in \mathbb{Z}$, as

$$\rho(t) = \frac{\gamma(t)}{\gamma(0)}$$

where the autocovariance function $\gamma(t) = \text{Cov}(Y_t, Y_0)$.

**Proposition**

The stationary time series $(Y_t)$ with zero mean is a MA($q$) process such that $b_q \neq 0$ if and only if $\rho(q) \neq 0$ and $\rho(t) = 0$ for all $|t| > q$. 
• **Examples.**
  - MA(1) : only $\rho(1)$ nonzero, exponential decay of $\alpha(t)$.
  - MA(2) : only $\rho(1)$ and $\rho(2)$ nonzero, damped exponential and sine wave for $\alpha(t)$. 
- **Partial autocorrelation function.**
  - To identify $p$.

**Definition (PACF)**

Let $(Y_t)$ be a stationary time series with zero mean. The partial autocorrelation function $\alpha$ is defined as $\alpha(0) = 1$, and, for all $t \in \mathbb{N}^*$, as

$$\alpha(t) = \phi_t, t$$

where $(\phi_t, t)_{t\in\mathbb{N}^*}$ are computed via the Durbin-Levinson recursion.

**Proposition**

If there exists a square-integrable sequence $(\psi_k)_{k \in \mathbb{N}}$ such that $(Y_t)$ has a $MA(\infty)$ expression with $\psi_0 = 1$, then the stationary time series $(Y_t)$ with zero mean is an $AR(p)$ process such that $a_p \neq 0$ if and only if $\alpha(p) \neq 0$ and $\alpha(t) = 0$ for all $t > p$. 

---

On a SARIMAX coupled modelling

Identification for stationary AR($p$) and MA($q$) processes
On a SARIMAX coupled modelling
Identification for stationary AR($p$) and MA($q$) processes

- **Examples.**
  - AR(1) : only $\alpha(1)$ nonzero, exponential decay of $\rho(t)$.
  - AR(2) : only $\alpha(1)$ and $\alpha(2)$ nonzero, damped exponential and sine wave for $\rho(t)$. 
• **Examples.**
  • ARMA(1,1) : exponential decay of $\rho(t)$ and $\alpha(t)$ from first lag.
  • ARMA(2,2) : exponential decay of $\rho(t)$ and $\alpha(t)$ from second lag.
On a SARIMAX coupled modelling
Linear relationship between consumption and temperature

- **Variance-stabilizing Box-Cox transformation.**
  - Logarithmic transform given, for all $1 \leq t \leq T$, by
    $$Y_t = \log (C_t + e^m)$$
  where $m$ ensures that $Y_t = m$ when $C_t = 0$. 

![Graph 1](image1.png)  
![Graph 2](image2.png)
On a SARIMAX coupled modelling
Linear relationship between consumption and temperature

• **Linear relationship between consumption and temperature.**
  - On a thermosensitive load curve, for all \( 1 \leq t \leq T \),
    \[
    Y_t = c_0 + C(B)U_t + \epsilon_t.
    \]
  - For all \( z \in \mathbb{C} \),
    \[
    C(z) = \sum_{k=0}^{r-1} c_k+1 z^k.
    \]
  - Unknown vector \( c \in \mathbb{R}^{r+1} \) estimated by OLS.

• **Seasonal residuals.**
  - Residuals \( (\epsilon_t) \) regarded as a seasonal time series.
On a SARIMAX coupled modelling

The SARIMAX modelling

- **Residuals** \( (\varepsilon_t) \) as a seasonal time series.
  - SARIMA\((p, d, q) \times (P, D, Q)\) modelling, for all \(1 \leq t \leq T\),
    \[
    (1 - B)^d (1 - B^s)^D \mathcal{A}(B) \mathcal{A}_s(B) \varepsilon_t = \mathcal{B}(B) \mathcal{B}_s(B) V_t,
    \]
    where \((V_t)\) is a white noise of variance \(\sigma^2 > 0\).
  - Polynomials defined, for all \(z \in \mathbb{C}\), as
    \[
    \mathcal{A}(z) = 1 - \sum_{k=1}^{p} a_k z^k, \quad \mathcal{A}_s(z) = 1 - \sum_{k=1}^{P} \alpha_k z^{sk},
    \]
    \[
    \mathcal{B}(z) = 1 - \sum_{k=1}^{q} b_k z^k, \quad \mathcal{B}_s(z) = 1 - \sum_{k=1}^{Q} \beta_k z^{sk},
    \]
  - Parameters \(a \in \mathbb{R}^p, b \in \mathbb{R}^q, \alpha \in \mathbb{R}^P\) and \(\beta \in \mathbb{R}^Q\) estimated by GLS.
  - \(\mathcal{A}\) and \(\mathcal{A}_s\) are causal.
On a SARIMAX coupled modelling

The SARIMAX modelling

- The dynamic coupled modelling.

**Definition (SARIMAX)**

In the particular framework of the study, a random process \((Y_t)\) will be said to follow a SARIMAX\((p, d, q, r) \times (P, D, Q)_s\) coupled modelling if, for all \(1 \leq t \leq T\), it satisfies

\[
\begin{align*}
Y_t &= c_0 + C(B)U_t + \varepsilon_t, \\
(1 - B)^d(1 - B^s)^D A(B)A_s(B)\varepsilon_t &= B(B)B_s(B)V_t.
\end{align*}
\]

- As soon as \(d + D > 0\),

\[
(1 - B)^d(1 - B^s)^D A(B)A_s(B) (Y_t - C(B)U_t) = B(B)B_s(B)V_t.
\]
• **Existence of a stationary solution.**
  
  Let \( I \) be the identity matrix of order \( T \) and

  \[
  U = \begin{pmatrix}
  1 & U_T & U_{T-1} & \ldots & U_{T-r+1} \\
  1 & U_{T-1} & U_{T-2} & \ldots & U_{T-r} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & U_1 & U_0 & \ldots & U_{-r+2}
  \end{pmatrix}, \quad Y = \begin{pmatrix}
  Y_1 \\
  Y_2 \\
  \vdots \\
  Y_T
  \end{pmatrix}.
  \]

**Theorem**

Assume that \( U'U \) is invertible. Then, the differenced process \( (\nabla^d \nabla^D \varepsilon_t) \) where \( \varepsilon_t \) is given, for all \( 1 \leq t \leq T \), by the vector form

\[
\varepsilon = \left( I - U(U'U)^{-1}U' \right) Y
\]

is a stationary solution of the coupled model suitably specified.

• ADF and KPSS tests: \( d \) and \( D \).
• Box and Jenkins methodology: \( p, q, r, P, Q, s \).
• **Forecasting using time series analysis.**
  
  - Let $\tilde{\epsilon}_{T+1}$ be the predictor of $(\epsilon_t)$ at stage $T + 1$.
  - Let $\hat{c}_T$ be the OLS estimate of $c$.
  - Assume that $\hat{r}$ has been evaluated.
  - The predictor at horizon 1 is given by
    \[
    \tilde{Y}_{T+1} = \hat{c}_{0,T} + \sum_{k=1}^{\hat{r}} \hat{c}_{k,T} U_{T-k+2} + \tilde{\epsilon}_{T+1}.
    \]
  - The predictor at horizon $H$ is given by
    \[
    \tilde{Y}_{T+H} = \hat{c}_{0,T} + \sum_{k=1}^{\hat{r}} \hat{c}_{k,T} U_{T-k+H+1} + \tilde{\epsilon}_{T+H}.
    \]
1. Introduction

2. On a SARIMAX coupled modelling
   - Stationary ARMA processes
   - Identification for stationary AR(\(p\)) and MA(\(q\)) processes
   - Linear relationship between consumption and temperature
   - The SARIMAX modelling
   - Application to forecasting

3. Application to forecasting on a load curve
   - Procedure
   - Seasonality and stationarity
   - ACF and PACF
   - Selection criteria
   - Selection on bayesian criteria
   - Selection on bayesian criteria

4. Application on a huge dataset
   - Technical issues
   - Procedure
   - Modelling
   - Forecasting

5. Conclusion
Application to forecasting on a load curve

Procedure
Application to forecasting on a load curve

Procedure

- **Box and Jenkins methodology.**
  - For a given \( \hat{r} \), estimation of the residual set \((\hat{\varepsilon}_t)\), for all \( 1 \leq t \leq T \),
  
  \[
  \hat{\varepsilon}_t = Y_t - \hat{c}_0, T - \sum_{k=1}^{\hat{r}} \hat{c}_k, T U_{t-k+1}.
  \]

  - Select \( \hat{s} \) by investigating the seasonality of \((\hat{\varepsilon}_t)\).
  - Select \( \hat{d} \) and \( \hat{D} \) by investigating the stationarity of \((\nabla^d \nabla^D_s \hat{\varepsilon}_t)\).
  - Select \( \hat{p}, \hat{q}, \hat{P} \) and \( \hat{Q} \) by looking at ACF and PACF on \((\nabla^d \nabla^D_s \hat{\varepsilon}_t)\).
  - Adjust \( \hat{p}, \hat{q}, \hat{r}, \hat{P} \) and \( \hat{Q} \) by minimizing bayesian or prediction criteria.
  - Test of white noise on the fitted innovations.
• **Seasonality.**
  - Fourier spectrogram on $(\hat{\varepsilon}_t)$, $(\nabla_{12}\hat{\varepsilon}_t)$ and $(\nabla_{24}\hat{\varepsilon}_t)$, for $T = 730 \times 24$ and $\hat{r} = 1$.

• **Stationarity.**
  - $(\hat{\varepsilon}_t)$ is not stationary.
  - $(\nabla\hat{\varepsilon}_t)$, $(\nabla_{24}\hat{\varepsilon}_t)$ and $(\nabla\nabla_{24}\hat{\varepsilon}_t)$ are stationary around a deterministic trend.
• **Sample autocorrelations.**
  - On the estimated residuals ($\hat{\varepsilon}_t$).

![Sample autocorrelations](image1)

• On the seasonally differenced residuals ($\nabla_{24}\hat{\varepsilon}_t$).

![Seasonally differenced residuals](image2)
• **Sample autocorrelations.**
  - On the doubly differenced residuals ($\nabla \nabla_{24} \hat{e}_t$).

![Graph showing autocorrelation](image)

• **Identified models.**
  - SARIMAX$(p, 0, 0, r) \times (P, 1, Q)_{24}$ with $p \leq 5$, $r \leq 2$, $P \leq 1$ and $Q = 1$.
  - SARIMAX$(p, 1, q, r) \times (P, 1, Q)_{24}$ with $p \leq 1$, $q = 2$, $r \leq 2$, $P \leq 1$ and $Q = 1$. 
Application to forecasting on a load curve

Selection criteria

- **Bayesian criteria.**
  - Akaike information criterion and Schwarz bayesian criterion,
    \[
    \text{AIC} = -2 \log \mathcal{L} + 2k \quad \text{and} \quad \text{SBC} = -2 \log \mathcal{L} + k \log T
    \]
    where \( \mathcal{L} \) is the model likelihood and \( k \) the number of parameters.
  - Log-likelihood.
  - Overall randomness of successive innovations.

- **Prediction criteria.**
  - We define \( C_A \) and \( C_R \) as follows,
    \[
    C_A = \frac{1}{NH} \sum_{k=1}^{NH} \left| \tilde{C}_{T+k} - C_{T+k} \right| \quad \text{and} \quad C_R = \left( \sum_{k=1}^{NH} C_{T+k} \right)^{-1} \left( \sum_{k=1}^{NH} \left| \tilde{C}_{T+k} - C_{T+k} \right| \right)
    \]
    where \((\tilde{C}_{T+1}, \ldots, \tilde{C}_{T+NH})\) are \( N \) consecutive predictions at horizon \( H \) from time \( T \).
Application to forecasting on a load curve

Selection on bayesian criteria.

- Selection on bayesian criteria.

<table>
<thead>
<tr>
<th>Model</th>
<th>$p$</th>
<th>$d$</th>
<th>$q$</th>
<th>$r$</th>
<th>$P$</th>
<th>$D$</th>
<th>$Q$</th>
<th>$s$</th>
<th>AIC</th>
<th>SBC</th>
<th>LL</th>
<th>VAR</th>
<th>WN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMAX</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-362.4</td>
<td>-330.4</td>
<td>186.2</td>
<td>0.053</td>
<td>✓</td>
</tr>
<tr>
<td>SARIMAX</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-393.6</td>
<td>-348.8</td>
<td>203.8</td>
<td>0.053</td>
<td>✓</td>
</tr>
<tr>
<td>SARIMAX</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-424.7</td>
<td>-367.2</td>
<td>221.4</td>
<td>0.053</td>
<td>✓</td>
</tr>
<tr>
<td>SARIMAX</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-446.9</td>
<td>-389.4</td>
<td>232.5</td>
<td>0.052</td>
<td>✓</td>
</tr>
<tr>
<td>SARIMAX</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-457.2</td>
<td>-393.3</td>
<td>238.6</td>
<td>0.052</td>
<td>✓</td>
</tr>
<tr>
<td>SARIMAX</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>100.4</td>
<td>132.3</td>
<td>-45.2</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>SARIMAX</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-238.7</td>
<td>-200.4</td>
<td>125.4</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>SARIMAX</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-389.5</td>
<td>-351.2</td>
<td>200.8</td>
<td>0.053</td>
<td></td>
</tr>
<tr>
<td>SARIMAX</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>-435.1</td>
<td>-384.0</td>
<td>225.6</td>
<td>0.052</td>
<td>✓</td>
</tr>
</tbody>
</table>

- Modelling with SARIMAX($3, 0, 2, 2) \times (0, 1, 1)_{24}$ with $T = 730 \times 24$. 

![Graph showing load curve and SARIMAX model comparison]
Application to forecasting on a load curve
Selection on bayesian criteria

- **Selection on bayesian criteria.**
  - Modelling with SARIMAX$(3, 0, 2, 2) \times (0, 1, 1)_{24}$ with $T = 730 \times 24$.

- Least squares estimation, for all $28 \leq t \leq T$,

\[
\begin{align*}
C_t &= \exp \left( \hat{c}_0 + \hat{c}_1 U_t + \hat{c}_2 U_{t-1} + \varepsilon_t \right) - \exp(5), \\
\varepsilon_t &= \varepsilon_{t-24} + \hat{a}_1 (\varepsilon_{t-1} - \varepsilon_{t-25}) + \hat{a}_2 (\varepsilon_{t-2} - \varepsilon_{t-26}) + \hat{a}_3 (\varepsilon_{t-3} - \varepsilon_{t-27}) \\
&\quad + (V_t - \hat{b}_1 V_{t-1} - \hat{b}_2 V_{t-2}) - \beta_1 (V_{t-24} - \hat{b}_1 V_{t-25} - \hat{b}_2 V_{t-26}),
\end{align*}
\]

in which $\hat{c}_0 = 7.9871, \hat{c}_1 = 0.0166, \hat{c}_2 = -0.0420, \hat{a}_1 = 0.4776, \hat{a}_2 = 0.9030, \hat{a}_3 = -0.4305, \hat{b}_1 = 0.0801, \hat{b}_2 = -0.8524, \beta_1 = -0.8125.$
Application to forecasting on a load curve

Selection on prediction criteria.

- Forecasting with SARIMAX(1, 0, 1, 2) × (0, 1, 1)_{24} with \( T = 730 \times 24 \).
Selection on prediction criteria.

- Forecasting with SARIMAX(1, 0, 1, 2) × (0, 1, 1)_{24} with \( T = 730 \times 24 \).

- Parsimony is a central issue in time series analysis.
- Results slightly improved with a sliding window of 2 months.
- Around 2% of relative error between \( (\tilde{Y}_{T+1}, \ldots, Y_{T+NH}) \) and \( (Y_{T+1}, \ldots, \tilde{Y}_{T+NH}) \).
• **Refining...**
  - Influence of $r$ and the size of the sliding windows $M$ on $C_R$.
  
  - $M = 2$ months is the optimal sliding window.
  - No more influence of $r$ as soon as $r \geq 2$. 
Introduction

2 On a SARIMAX coupled modelling
   - Stationary ARMA processes
   - Identification for stationary AR($p$) and MA($q$) processes
   - Linear relationship between consumption and temperature
   - The SARIMAX modelling
   - Application to forecasting

3 Application to forecasting on a load curve
   - Procedure
   - Seasonality and stationarity
   - ACF and PACF
   - Selection criteria
   - Selection on bayesian criteria
   - Selection on bayesian criteria

4 Application on a huge dataset
   - Technical issues
   - Procedure
   - Modelling
   - Forecasting

5 Conclusion
Application on a huge dataset

Technical issues

- **Huge dataset.**
  - More than 2000 load curves.
  - Around 70% nonthermosensitive.
  - High quality: more than 9 months of data per curve, very little missing values.

- **Technical problems.**
  - Exponential growth of computing time with parsimony.
Application on a huge dataset

Procedure

- **On a representative panel.**
  - Selection of 200 heterogeneous thermosensitive curves (size, peaks intensity, etc.)
  - Massive statistical KPSS procedures and visual first conclusions.
  - Bayesian criteria to select the best models **on average**.
  - Application to forecasting.
  - Prediction criteria to select the best models **on average**.
  - All parameters vary in their neighborhood.
  - Consider technical issues: a night of computation for some models.

- **2 more bayesian criteria.**
  - Reliability index.
  - Percentage of significance of the first exogenous coefficient.
  - Assess the relevance on large-scale.
  - Caution: main assumptions for t-test not satisfied!
Application on a huge dataset
Modelling

- **Stationarity, for \( r = 1 \) and \( M = 3 \) months.**
  - Less than 30% of \((\hat{\varepsilon}_t)\) stationary.
  - 100% of \((\nabla_{24}\hat{\varepsilon}_t)\) and \((\nabla \nabla_{24}\hat{\varepsilon}_t)\) stationary.
  - SARIMAX\((p, 0, q, r) \times (P, 1, Q)_{24}\) and SARIMAX\((p, 1, q, r) \times (P, 1, Q)_{24}\).
  - Making \( r \) increase.

- **Relative evolution of AIC and SBC with \( r \).**

![Graph showing relative evolution of AIC and SBC with \( r \).]

- Clearly, \( r = 1 \).
- SARIMAX\((3, 0, 2, 1) \times (0, 1, 1)_{24}\), SARIMAX\((3, 1, 2, 1) \times (0, 1, 1)_{24}\).
- Gain over the naive model: \( \approx 50\% \).
- Gain over the SARIMA model: \( \approx 2\% \).
• Significance of $c_1$ for $M = 6$ months.

• Substantial on thermosensitive curves.
Application on a huge dataset
Forecasting

- **Evolution of $C_R$.**

- $\text{SARIMAX}(1, 1, 1, 1) \times (0, 1, 1)_{24}$ and $\text{SARIMAX}(2, 0, 1, 1) \times (0, 1, 1)_{24}$.
- $M = 2$ months.
- Gain over the naive model: $\approx 65\%$.
- Gain over the SARIMA model: $\approx 3\%$. 
• Examples.
1 Introduction

2 On a SARIMAX coupled modelling
   - Stationary ARMA processes
   - Identification for stationary AR($p$) and MA($q$) processes
   - Linear relationship between consumption and temperature
   - The SARIMAX modelling
   - Application to forecasting

3 Application to forecasting on a load curve
   - Procedure
   - Seasonality and stationarity
   - ACF and PACF
   - Selection criteria
   - Selection on bayesian criteria
   - Selection on bayesian criteria

4 Application on a huge dataset
   - Technical issues
   - Procedure
   - Modelling
   - Forecasting

5 Conclusion
• **Nonthermosensitive curves.**
  - SARIMA(1, 0, 1) × (0, 1, 1)\textsubscript{24}

• **Thermosensitive curves.**
  - SARIMAX(1, 1, 1, 1) × (0, 1, 1)\textsubscript{24} and SARIMAX(2, 0, 1, 1) × (0, 1, 1)\textsubscript{24}.

• **A careful study curve by curve would provide better results.**
  - Of course...

• **Caution : technical issues for huge datasets.**
  - Is a computation time ×1000 for a gain of 0.1% relevant?
  - Engineering approach, corporate vision.

• **Thank you for your attention.**

• **Comments or questions ? 🤓**