LINEAR TIME SERIES: ARIMA MODELLING OF A TIME SERIES
French arms and ammunition production index

Authors:
Jérôme ALLIER
Axel PINCON

Course leader:
Joseph NGATCHOU WANDJI

April 17, 2025

Contents

Chapter 1

THE DATA

1.1 The French arms and ammunition production index

The data chosen for this study is the industrial production index for arms and ammunition. It is produced by the INSEE. The index corresponds to the level of arms and ammunition French production for a given date. The base year is 2021 (basis 100) and the data is corrected from seasonal variations and working days (CVS-CJO index).

The data is available and can be download here : Arms and Ammunition Production - Insee

This series represents the index, calculated monthly, from January 1990 to January 2024. The dataset contains 409 observations. To simplify the study and to avoid some extraordinary periods (two falls during the 90's, 2008 crisis), we will focus our work on the observations between January 2008 to January 2024. Thus, our dataset contains 193 observations.

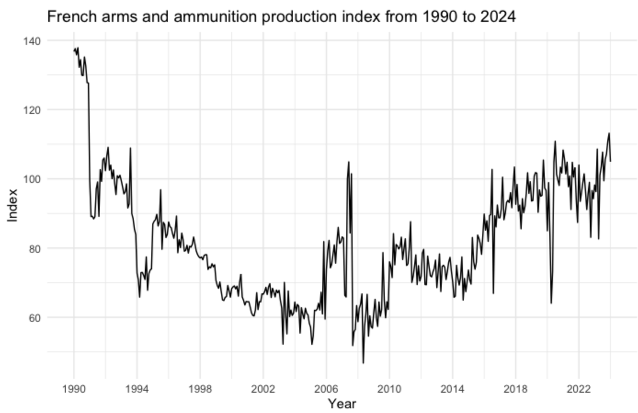

Figure 1.1: French arms and ammunition production index from 1990 to 2024

1.2 Transformation for a stationary serie

We can state that the data is not stationary.
When the series is plotted, one can easily note that the series follows an overall positive trend across time between 2008 and 2024. This hypothesis can be verified by doing a linear regression : the coefficient associated to the regression is positive (2.675) and statistically significant at the 1 % level. In order to eliminate this trend, we differentiated the series in the first order. The new plotted series seems now stationary. The Augmented Dickey-Fuller test validates this hypothesis for the differentiated series : we obtain a p-value smaller than 0.01 and we reject the null hypothesis of non-stationary condition at the 1% level. A similar result is obtained with the Phillip-Perron test.

1.3 The serie before and after transformation

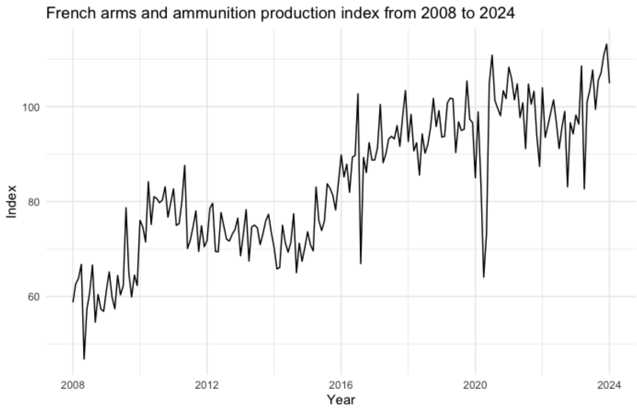

Figure 1.2: Series truncated before differentiation

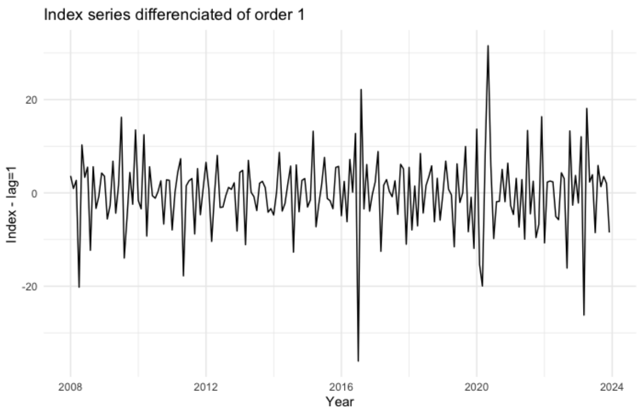

Figure 1.3: Series differentiated of order 1

Chapter 2

The ARMA Modelisation

1.1 Fitting an ARMA(p,q)

In order to fit an ARMA model to our stationary serie, we followed these 3 steps :

1. We first define the maximal degrees of the ARMA model that explain the stationary data by analyzing the ACF and PACF plot of the data.

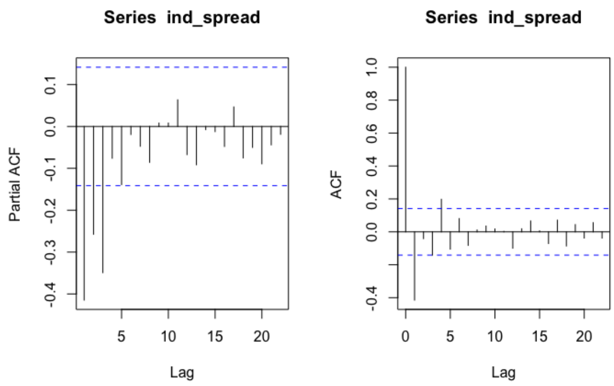

Figure 2.1: ACF-PACF of the diffrenciated series

The ACF graph shows that the first 4 past values have a significant power to explain the serie. Thus, let's define qmax = 4. Similarly, the PACF graph shows that the first 5 past values have a significant power to explain the serie. Thus, let's define pmax = 5.

2. We then find the set of valid ARMA model for the data. Within all the ARMA(p,q) models (associated to all the combinations of (p,q) with p < pmax and q < qmax), let's consider the subset of valid models. To check the validity of a model, we tested the significance of the coefficients for each model and the non-existence of autocorrelation between the residuals.

After these tests, only 8 models are conserved :

p	3	0	3	4	3	1	0	1
q	0	1	1	1	2	3	4	4

3. Finally, we choose within the previous set of models, retain the models that minimize the information criteria (AIC, BIC) (cf. Information criteria values 4) Finally, the ARMA model chosen to fit the corrected series is an ARMA(4,1). In fact, the ARMA(4,1) minimizes the Akaike information criterion (AIC) whereas the ARMA(0,1) minimizes the Bayesian information criterion (BIC). An ARMA model that minimizes the Akaike Information Criterion (AIC) is typically preferred when the focus is on achieving the best predictive performance, as it tends to balance model complexity and goodness-of-fit. In contrast, models that minimize the Bayesian Information Criterion (BIC) prioritize simplicity, which might be suitable when overfitting is a concern or when the dataset is large. Here we decided to continue the study with the ARMA(4,1), a more complex model than ARMA(0,1).

We can check of the validity and accuracy of this ARMA model.

The model is valid: significance of the coefficients, no correlation between the residuals. The adjusted R-squared is such that $R^2 > 0.99$. While the model appears to be precise, the relatively high R^2 value raises concerns about potential overfitting. As the roots of the AR polynomial lie outside the unit circle, it indicates that the model is causal. Estimation of the parameters:

Order	AR1	AR2	AR3	AR4	MA1
Coeff	0.3116	0.1151	0.0366	0.2779	-1

Based on the results of the Shapiro-Wilk test and the information provided in Figure 6, it appears that the residuals exhibit a normal distribution.

1.2 Corresponding ARIMA(p,d,q)

The differentiated series ΔX_t can be expressed as an ARMA(4,1):

$$\Delta X_t = \epsilon + \epsilon_{t-1} + 0.3116\Delta X_{t-1} + 0.1151\Delta X_{t-2} + 0.0366\Delta X_{t-3} + 0.2779\Delta X_{t-4}$$

Since we differentiated the original series once to make it stationary, the additional parameter in the ARIMA model (in comparison to the previous ARMA(4,1) model for the differentiated series) is equal to 1. Therefore, the ARIMA model for the series is an ARIMA(4,1,1).

$$\phi(L)(1-L)X_t = \psi(L)\epsilon_t \tag{2.1}$$

where $\phi(.)$ and $\psi(.)$ are the polynomials of degrees 4 and 1 respectively whose coefficients are given in the 3 ($\phi(.)$ for the AR part and $\psi(.)$ for the MA part) and L is the lag operator.

Chapter 3

Forecasting

We have the following model,

$$Y_t = \Delta X_t = \sum_{i=1}^{4} \phi_i \Delta X_{t-i} + \epsilon_t + \psi_1 \epsilon_{t-1}$$

1.1 Confidence Region

We have $E[\epsilon_{t+h}|Y_t, Y_{t-1}, Y_{t-2}, \ldots] = 0$, for all $h > 0$. Therefore, the optimal prediction is given by :

$$\begin{cases} \hat{Y}_{t+1|t} = \sum_{i=1}^{4} \phi_i Y_{t-i+1} + \psi_1 \epsilon_t \\ \hat{Y}_{t+2|t} = \phi_1 \hat{Y}_{t+1|t} + \sum_{i=2}^{4} \phi_i X_{t+2-i} \end{cases}$$

Let Y and $\hat{Y}$ such that:

$$Y = \begin{pmatrix} Y_{t+1} \\ Y_{t+2} \end{pmatrix}$$

and

$$\hat{Y} = \begin{pmatrix} \hat{Y}_{t+1|t} \\ \hat{Y}_{t+2|t} \end{pmatrix}$$

Thus, we have :

$$Y - \hat{Y} = \begin{pmatrix} \epsilon_{t+1} \\ \epsilon_{t+1}(\phi_1 + \psi_1) + \epsilon_{t+2} \end{pmatrix}$$

(ϵ_t) is an i.i.d. sequence of random variable following $N(0, \sigma^2)$. The variances of the prediction errors are given by :

$$V(Y_{t+1} - \hat{Y}_{t+1|t}) = V(\epsilon_{t+1}) = \sigma^2$$
$$V(Y_{t+2} - \hat{Y}_{t+2|t}) = \sigma^2 \cdot ((\phi_1 + \psi_1) + 1)$$

Therefore, $Y - \hat{Y} \sim N(0, \Sigma)$, where $\Sigma = \sigma^2 \begin{pmatrix} 1 & \phi_1 \\ \phi_1 & (\phi_1 + \psi_1) + 1 \end{pmatrix}$.
Since $\sigma^2 > 0$, we have $Det(\Sigma) = \sigma^2 > 0$. Then, based on the course, we have $(Y - \hat{Y})^t \Sigma^{-1} (Y - \hat{Y}) \sim \chi^2(2)$. The confidence region of level α of the future values (Y_{t+1}, Y_{t+2}) is given by:

$$\left\{ Y \in \mathbb{R}^2 \middle| (Y - \hat{Y})^t \Sigma^{-1} (Y - \hat{Y}) \leq q_{1-\alpha}^{\chi^2(2)} \right\}$$

1.2 Hypothesis for confidence region

The given confidence region holds under the following hypothesis :

1. The model is perfectly known and therefore the coefficients identified in part 1.1 are the reel coefficients of the model.

2. The white noise follows a centered normal distribution with a known variance $\sigma^2 > 0$.

In fact, if the variance of the residual is unknown, it must be estimated. Therefore, the confidence interval will be derived from the distribution of a Student's law, with wider tails and will be a less accurate interval. We also assumed that the true values of the parameters of the ARMA model are known. Indeed, if the true values of the ARMA parameters are unknown, then $\hat{\Sigma}$ would be doubly uncertain (due to the uncertainty associated with the estimation of the variance of the residual, and due to the uncertainty in the estimation of the model coefficients).

1.3 Visualisation of the Forecast

Here, we forecast the next values in the series using the ARIMA(4,1,1) model. The graph below shows our actual series in black, our forecast in blue and the 95% confidence interval in grey:

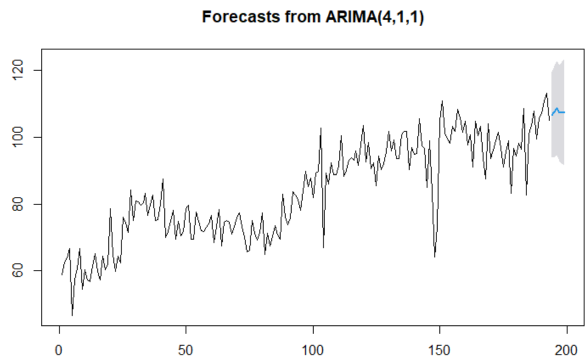

Figure 3.1: Confidence region of level $\alpha = 95\%$ of the future values

1.4 Open Question

The information given by Y_{T+1} may improve and make the prediction of X_{T+1} easier under specific assumptions: the process (Y_t) must instantly cause (X_t) in the sense given by Granger, that corresponds to the ability of Y_{T+1} to predict X_{T+1}. In other words: $E[X_t|X_{t-1}, Y_{t-1}, \ldots] \neq E[X_t|X_{t-1}, \ldots]$. To verify this condition, it is feasible to test causality with a Wald test.

Chapter 4

Complements

1.1 ACF PCF plots

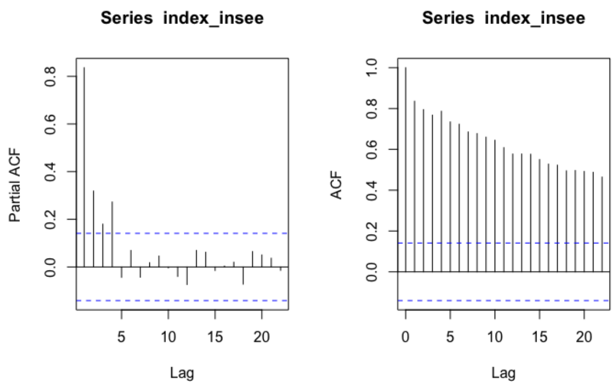

Figure 4.1: ACF - PACF of the original series

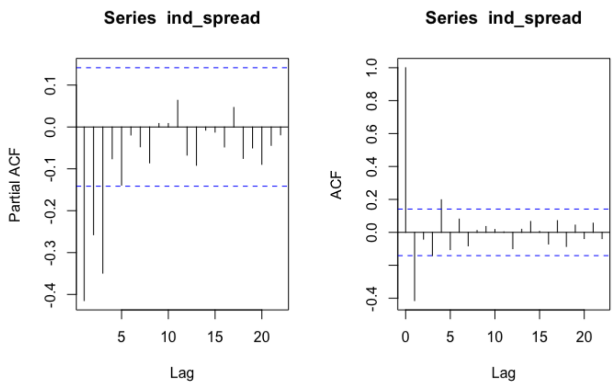

Figure 4.2: ACF-PACF of the diffrenciated series

1.2 Information criteria

1.3 Residuals