



Faculté des Sciences Exacte et Informatique

Département de Mathématique

**Mémoire de fin d'études**

Présenté pour l'obtention du diplôme de

**Master**

**Spécialité** : Mathématiques.

**Option** : Probabilités et Statistique.

**Thème**

**Sur Les modèles vectoriels autorégressifs**  
**VAR(p)**

**Présenté par :**

\* **Mezhoud Ibtissam**

**Devant le jury composé de :**

Cheraitia Hassen M.C.A Université de Jijel

Encadreur

Sellami Nawel M.A.A Université de Jijel

Président

Ghouil Djoweyda M.A.A Université de Jijel

Examineur

Promotion **2021/2022**

## ♡ Remerciements ♡

Nous tenons à remercier ALLAH, le tout puissant de nous avoir donné courage et volonté pour accomplir ce modeste travail.

Nous tenons à exprimer tous nos reconnaissances à tous ceux qui ont contribué de prêt ou de loin à la réalisation de ce travail

Nous tenons à présenter notre profonde gratitude à mon encadreur *Cheraitia Hassen* pour son disponibilité, ses remarques et conseils.

Nos remerciements vont également à l'ensemble de nos enseignants du département de mathématiques

## ♡ Dédicace ♡

J'ai le grand plaisir de dédier ce modeste mémoire :

♡ A mes très chers parents

♡ A nos frères et sœurs

♡ A tous mes proches sans exception

♡ A tous mes amis sans exception

♡ A tous ceux qui je connaissent

※ *Ibtissam* ※

---

# RÉSUMÉ

Une série temporelle multivariée (STM) est constituée de plus d'une variable temporelle et chaque variable dépend non seulement de ses valeurs passées mais aussi des valeurs passées des autres variables. Pour traiter les STM, l'une des méthodes les plus populaires est le modèle Vector Auto Regressive (VAR) qui est une forme vectorielle de processus auto-régressive (AR) qui peut être utilisée pour examiner les relations entre plusieurs variables dans l'analyse des séries chronologiques multivariées.

Dans ce mémoire, trois séries météorologiques mensuelles (Température, Précipitations et Humidité) ont été modélisées au premier lieu séparément en appliquant la méthode de Box et Jenkins, et au deuxième lieu simultanément par estimation d'un processus VAR(3).

A la fin, des prévisions à court terme ont été calculées par les deux approches. En termes de RMSE et MAE, le modèle VAR(3) a fourni des meilleures prévisions que de Box et Jenkins.

**Mots clés :** Série temporelles univariées, Série temporelles multivariées, Processus stochastique, Dicky-Fuller, Box-Jenkins, Modèles VAR, Prévion.

---

# ABSTRACT

A multivariate time series consists of more than one time variable and each variable depends not only on its past values but also on the past values of other variables. To deal with MTS, one of the most popular methods is Vector Auto Regressive models (VAR) that is a vector form of autoregressive process (AR) that can be used to examine the relationships among several variables in multivariate time series analysis.

In this thesis, three monthly weather series (Temperature, Precipitation and Humidity) were modeled first separately by applying the Box and Jenkins method, and secondly simultaneously by estimating a VAR(3) process.

In the end, short-term forecasts were calculated by both approaches. In terms of RMSE and MAE, the VAR(3) model provided better forecasts than Box Jenkins.

**Key Words** : univariate time series ,multivariate time series,stochastic process ,Dicky-Fuller,Box-Jenkins,Models VAR, Forecast.

---

# TABLE DES MATIÈRES

<b>Résumé</b>	<b>v</b>
<b>Introduction</b>	<b>1</b>
<b>1 Les Séries Temporelles Univariées</b>	<b>3</b>
1.1 Introduction aux séries temporelle . . . . .	3
1.1.1 Les composantes principales d'une série temporelle . . . . .	3
1.1.2 Caractéristiques d'une série temporelle [8] . . . . .	4
1.2 Le processus stochastique . . . . .	5
1.2.1 Généralisation et notation . . . . .	5
1.2.2 Processus linéaire . . . . .	7
1.2.3 Processus linéaires stationnaire . . . . .	7
1.2.4 Processus $AR(p)$ (autoregressive) . . . . .	7
1.2.5 Processus $MA(q)$ (moyenne mobile) . . . . .	9
1.2.6 Le processus $ARMA(p,q)$ . . . . .	12
1.2.7 processus $ARIMA(p, d, q)$ [5] . . . . .	14
1.2.8 Processus linéaire non stationnaire(TS,DS) . . . . .	15
1.3 Tests de stationnarité . . . . .	17
1.3.1 Test de Dickey Fuller . . . . .	18
1.3.2 Test de Dickey Fuller augmenté . . . . .	21
1.4 Procédure de Box et Jenkins . . . . .	22

1.4.1	Identification . . . . .	22
1.4.2	Estimation . . . . .	24
1.4.3	Validation . . . . .	25
1.4.4	Prévision . . . . .	29
<b>2</b>	<b>Les Séries Temporelles Multivariées et Modèle <math>VAR(p)</math></b>	<b>32</b>
2.1	Présentation du Modèle . . . . .	32
2.1.1	Ecriture $VAR(1)$ d'un processus $VAR(p)$ . . . . .	34
2.1.2	Représentation canonique et processus des innovations : . . . . .	34
2.2	Stationnarité . . . . .	35
2.3	Les Moments de processus $VAR(p)$ . . . . .	36
2.3.1	Esperance . . . . .	36
2.3.2	Fonction d'auto-covariance . . . . .	37
2.3.3	Fonction d'autocorrélation . . . . .	38
2.3.4	Fonction d'autocorrélation partielle(identification) . . . . .	39
2.4	Estimation des paramètres d'un $VAR(p)$ . . . . .	39
2.4.1	la méthode des moindres carrés ordinaires . . . . .	39
2.4.2	La méthode de maximum de vraisemblance . . . . .	43
2.5	Validation de modèle $VAR(p)$ . . . . .	43
2.5.1	la statistique du rapport de vraisemblance . . . . .	44
2.5.2	Tests de bruit blanc des erreurs [2] . . . . .	45
2.5.3	Critère d'information . . . . .	45
2.6	Prévision . . . . .	46
2.7	Causalité . . . . .	47
2.7.1	La notion de causalité . . . . .	47
2.7.2	Causalité au sens de Granger(1969) . . . . .	48
2.7.3	Causalité au sens de Pierce et Haugh(1977) . . . . .	49
2.7.4	causalité au sens Sims (1980) . . . . .	49
2.8	Analyse de réponse impulsionnelle . . . . .	50
2.8.1	Représentation VMA d'un processus $VAR$ . . . . .	50
2.8.2	$VAR$ structurel . . . . .	50
2.8.3	Orthogonalisation des chocs . . . . .	51
2.8.4	Décomposition de Cholesky . . . . .	52

---

2.8.5	Méthode d'identification des chocs . . . . .	54
2.9	Décomposition de la variance . . . . .	54
<b>3</b>	<b>Application de Modèle VAR(<math>p</math>) en Météorologie</b>	<b>56</b>
3.1	Introduction . . . . .	56
3.2	Etude Univarié . . . . .	57
3.2.1	Etude de stationnarité . . . . .	57
3.2.2	Box-Jenkins . . . . .	60
3.3	Etude multivarié . . . . .	63
3.3.1	Choix de retard $p$ . . . . .	64
3.3.2	Estimation de modèle VAR . . . . .	64
3.3.3	Étude de stationnaire du VAR . . . . .	64
3.3.4	Causalité . . . . .	65
3.3.5	Fonction de réponse impulsionnelle et analyse de choc . . . . .	66
3.3.6	Décomposition de Variance . . . . .	69
3.3.7	Validation . . . . .	71
3.3.8	Prévision . . . . .	72
3.4	comparaison entre le prévision de Box-Jenkins et modèle VAR . . . . .	73
	<b>Conclusion</b>	<b>75</b>
	<b>Annexes</b>	<b>77</b>
	<b>Bibliographie</b>	<b>85</b>

---

# TABLE DES FIGURES

1.1	schéma 1 de stratégie de dicky-fuller . . . . .	19
1.2	schéma 2 de stratigie de Dicky-Fuller . . . . .	21
3.1	Représentations graphiques des trois séries . . . . .	58
3.2	ACF et PACF des trois séries . . . . .	58
3.3	ACF et PACF des séries stationnaires . . . . .	60
3.4	Cercle des inverses des racines . . . . .	65
3.5	Reponse de prec à un choc dans hum . . . . .	66
3.6	Réponse de Temp à un choc dans hum . . . . .	66
3.7	Reponse de Prec à un choc dans Temp . . . . .	67
3.8	Réponse de Hum à un choc dans Temp . . . . .	67
3.9	Réponse de Temp à un choc dans Prec . . . . .	68
3.10	Réponse de Hum à un choc dans Prec . . . . .	68
3.11	D.V de l'erreur de prévision des Séries . . . . .	71
3.12	graphe de prévision De modèle VAR(3) . . . . .	72

---

# LISTE DES TABLEAUX

3.1	Statistiques descriptives . . . . .	57
3.2	coefficients saisonniers . . . . .	59
3.3	résultats de test ADF . . . . .	60
3.4	Modèles candidats . . . . .	61
3.5	Estimation et validation des processus estimés.() représente p-value . . . . .	61
3.6	Estimation et validation des processus estimés.() représente p-value . . . . .	62
3.7	Estimation et validation des processus estimés.() représente p-value . . . . .	62
3.8	Prévisions des trois séries par la méthode Box et Jenkins . . . . .	63
3.9	Critères de choix de retard optimal . . . . .	64
3.10	Résultats du test de causalité . . . . .	65
3.11	D.V de l'erreur de prévision de Temp . . . . .	69
3.12	D.V de l'erreur de prévision de Hum . . . . .	70
3.13	D.V de l'erreur de prévision de Prec . . . . .	70
3.14	comparaison des données réelles et des résultats prévisionnels de (temp),(Hum)et(prec) . . . . .	72
3.15	Série Résiduels de Prec . . . . .	73
3.16	Série Résiduels de Temp . . . . .	73
3.17	RMSE et MAE de prévision des séries . . . . .	74



---

## NOTATION

$T_t$ :	<i>Latendance</i>
$S_t$ :	<i>Lasaisonnalité</i>
$E$ :	<i>Esperance.</i>
$var$ :	<i>Lavariance.</i>
$Cov$ :	<i>Lacovariance.</i>
$\gamma_h$ :	<i>La fonction d'auto – covariance</i>
$\rho_h$ :	<i>La fonction d'auto – corrélation.</i>
$ACF$ :	<i>La fonction d'auto – corrélation.</i>
$PACF$ :	<i>La fonction d'auto – corrélation Partielle.</i>
$AR$ :	<i>Auto – Régressif.</i>
$MA$ :	<i>Moyennemobile</i>
$ARMA$ :	<i>Auto – Régressif Moyennemobile.</i>
$ARCH$ :	<i>Auto – Régressif Conditionnellement Hétéroscédastique.</i>
$VAR$ :	<i>vecteur Auto – Régressif.</i>
$\varepsilon_t$ :	<i>Le résidus.</i>
$TS$ :	<i>Trend Stationary.</i>
$DS$ :	<i>Difference Stationary.</i>
$bb$ :	<i>bruit blanc.</i>
$SK$ :	<i>Skewness.</i>
$Kr$ :	<i>Kurtosis.</i>
$SCR$ :	<i>Somme des carrés des résidus.</i>
$SSi$ :	<i>Si et Selement Si.</i>
$i.i.d$ :	<i>indépendant identiquement distribué</i>

---

# INTRODUCTION

L'une des méthodes les plus populaires dans la théorie des séries temporelles est les modèles autorégressifs vectoriels (VAR), qui est une forme vectorielle de processus autorégressive (AR) qui peut être utilisée pour examiner les relations entre plusieurs variables dans l'analyse de séries temporelles multivariées.

Dans un modèle VAR, chaque série temporelle est une fonction linéaire des valeurs passées d'elle-même et des valeurs passées de toutes les autres séries. Le VAR est capable de comprendre et d'utiliser la relation entre plusieurs séries ce qui permet de décrire la dynamique des comportements des données et pourra fournir de meilleurs résultats de prévision. Le modèle VAR est caractérisé par les points suivants :

- Les séries à modéliser sont toutes stationnaires
- Les séries à modéliser sont toutes potentiellement dépendantes
- Le nombre de retards associé à chaque série dans chaque équation est identique

Notre partie pratique a porté sur l'analyse des séries chronologiques représentant la température, les précipitations et l'humidité. L'analyse nous permettra de construire un modèle mathématique qui explique ces trois variables, nous allons proposer dans un premier temps d'étudier individuellement les séries. Chaque série sera étudiée séparément, négligeant l'effet des autres séries suivant la méthodologie de Box Jenkins qui permet en plusieurs étapes de trouver un modèle adéquat avec lequel nous calculons les prévisions.

Cependant la méthodologie de Box Jenkins ne prend pas en compte l'interdépendance des séries. Dans notre cas pratique et à titre d'exemple plusieurs variables doivent être prise en

considération pour prédire de manière optimale des précipitations et pour cette raison que nous allons proposer par la suite une approche multivariée pour pallier aux insuffisances de l'approche univariée. Étant donné qu'à l'issue de la méthodologie de Box Jenkins les séries sont rendues stationnaires, l'application de la modélisation VAR(p) à la base de ces séries est possible.

Afin d'atteindre nos objectifs, nous articulons notre mémoire autour de trois chapitres :

Dans le premier chapitre intitulé « les séries temporelles univariées », nous présentons quelques rappels sur les processus stochastiques et la représentation des séries temporelles linéaires stationnaires et non stationnaires, la stratégie de Dickey Fuller et la méthodologie de Box et Jenkins.

Le deuxième chapitre intitulé « les séries temporelles multivariées et modèle VAR » est consacré à la présentation du modèle VAR, les méthodes d'estimation à savoir la méthode de maximum de vraisemblance et la méthode des moindres carrés ordinaires, les tests de validation. Ce chapitre est clôturé par une exposition de la dynamique de modèle VAR via la fonction de réponse impulsionnel et la décomposition de la variance .

Enfin, dans le troisième chapitre intitulé « applications de modèle VAR en météorologie », on fait compléter notre travail par la mise en place de deux approches ; la première est univariée qui consiste à appliquer la stratégie de Dickey Fuller et la méthode de Box et Jenkins séparément sur les trois séries météorologiques (température, précipitations, et humidité). La deuxième est multivariée dans laquelle nous allons appliquer le modèle VAR sur les trois séries simultanément. Les prévisions résultants de ces deux modélisations sont calculées et comparées.

A la fin, une conclusion générale englobe les principaux résultats obtenus au cours de notre travail.

---

---

# CHAPITRE 1

---

## LES SÉRIES TEMPORELLES UNIVARIÉES

### 1.1 Introduction aux séries temporelle

**Définition 1.1.** Une série temporelle est une suite d'observations indicées par le temp, notée  $\{X_t, t = 1 \dots T\}$ . Elle représente généralement l'évolution d'un phénomène au cours du temps

**Remarque 1.1.1.** •  $T \subset \mathbb{Z}$  (cas discret)

•  $T \subset \mathbb{R}$  (cas continue).

#### 1.1.1 Les composantes principales d'une série temporelle

Il est classique de décomposer une série temporelle  $\{X_t, t = 1, \dots, T\}$  en trois termes :

1-La tendance (trend) : notée  $T_t$  est une fonction représente le mouvement à long terme de la série étudiée.

2-Saisonnalité : notée  $S_t$  est une fonction périodique de période  $d$ .

3 - Résidus : notée  $R_t$  la partie aléatoire est dit bruit blanc correspond à des fluctuations irrégulières, en général faible amplitude.

On s'intéresse habituellement à deux modèles de décompositions d'une série chronologique : [9]

- Le modèle additif : les trois composantes précédentes indépendantes

$$X_t = T_t + S_t + R_t$$

où  $E(R_t) = 0$

- Le modèle multiplicatif : les composantes sont dépendantes

$$X_t = T_t \times S_t \times R_t$$

où  $E(R_t) = 1$

### 1.1.2 Caractéristiques d'une série temporelle [8]

Soit une série stationnaire  $\{X_t, t = 1, \dots, T\}$  :

a-La moyenne :  $E(X_t) = \frac{1}{T} \sum_{t=1}^T X_t$

b-La variance :  $var(X_t) = \frac{1}{T} \sum_{t=1}^T [(X_t - E(X_t))^2]$

c-La fonction d'auto covariance :

On appelle fonction d'auto covariance  $\gamma_h$  de  $X_t$  la fonction :

$$\gamma(h) = cov(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)E(X_{t+h})$$

d-Fonction d'auto corrélation :

On appelle fonction d'auto corrélation  $\rho_h$  la fonction :

$$\rho_h = \frac{\gamma_h}{\gamma_0}$$

**Remarque 1.1.2.** Le graphe de la fonction d'auto corrélation est appelé "**corrélogramme**"

**propriété 1.1.1.** soit  $(X_t)_t$  un processus stationnaire, la fonction d'auto covariance vérifie les propriétés suivantes :

$$1 - \gamma_0 = cov(X_t, X_t) = E((X_t - E(X_t))^2) = var(X_t) = \sigma_X^2 \geq 0$$

$$2 - |\gamma_h| \leq \gamma_0$$

3 -  $\gamma_h = \gamma_{-h}$ ,  $\gamma$  est une fonction paire.

**propriété 1.1.2.** soit  $(X_t)_t$  un processus stationnaire, la fonction d'auto corrélation vérifie les propriétés suivantes :

$$1 - \rho_0 = cov(X_t, X_t) = E((X_t - E(X_t))^2) = var(X_t) = \sigma_X^2 \geq 0$$

$$2 - |\rho_h| \leq \rho_0$$

3 -  $\rho_h = \rho_{-h}$ ,  $\rho$  est une fonction paire.

e-Fonction d'auto corrélation partielle :

Soit la fonction d'auto corrélation notée  $\rho_h$  et la fonction d'auto corrélation partielle noté  $\phi_{hh}$

La matrice symétrique formée des  $h - 1$  premier auto corrélation de  $X_t$  :

$$\rho_h = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{h-1} \\ \rho_1 & \ddots & \cdots & \rho_{h-2} \\ \vdots & \cdots & \ddots & \\ \rho_{h-1} & \rho_{h-2} & \cdots & 1 \end{bmatrix}$$

et soit  $\rho_h^*$  la matrice  $\rho_h$  dans laquelle on a remplacé la dernière colonne par le vecteur  $[\rho_1, \dots, \rho_h]$

$$\rho_h^* = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_2 \\ \vdots & \cdots & \ddots & \\ \rho_{h-1} & \rho_{h-2} & \cdots & \rho_h \end{bmatrix}$$

La fonction d'auto corrélation partielle est donnée par :

$$\phi_{hh} = \frac{|\rho_h^*|}{|\rho_h|}$$

où  $|\cdot|$  est déterminant .

## 1.2 Le processus stochastique

### 1.2.1 Généralisation et notation

#### L'opérateur de retard

L'opérateur de retard noté  $\mathbf{B}$  (back wards), l'opérateur qui fait passer de  $X_t$  à  $X_{t-1}$  :

$$BX_t = X_{t-1}$$

plus généralement, on a :

$$B^n X_t = X_{t-n}$$

#### Stationnarité des processus

**Définition 1.2.** (La stationnarité forte)

On dit qu'un processus  $X_t$  est stationnaire au sens strict (ou fortement stationnaire) si : pour tout  $(t_1, t_2, \dots, t_n), t_i \in T$  et si pour tout  $h \in T : \{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$  a la même distribution de probabilité jointe que  $\{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}\}$

**Définition 1.3.** (La stationnarité faible)

Soit un processus  $X_t$  à valeur réelle et ent temps discret, il est dit stationnaire au sens faible (ou "second ordre")

- $E(X_t^2) < \infty, \forall t \in \mathbb{Z}$  .
- $E(X_t) = m, \forall t \in \mathbb{Z}$  ( $m$  est une constante indépendant du temps)
- $cov(X_t, X_{t+h}) = \gamma_h, \forall t \in \mathbb{Z}, \forall h \in \mathbb{Z}$  ( fonction indépendant de temps)

**Exemple 1.2.1.** Un bruit blanc est stationnaire.

### processus bruit blanc

**Définition 1.4.** (Bruit blanc)

Un bruit blanc est une suite de variables aléatoire, notée  $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$ , vérifiée les propriété suivantes :

- $E(\varepsilon_t) = 0 \quad \forall t$
- $var(\varepsilon_t) = \sigma_\varepsilon^2 \quad \forall t$
- $cov(\varepsilon_t, \varepsilon_{t'}) = 0 \quad \forall t \neq t'$

### Le théorème de wold (1954)

Le théorème de wold (1954) est un théorème fondamental dans la modélisation des processus stationnaires.

**Théorème 1.2.1.** Considérons un processus stationnaire  $X_t$  . Il est toujours possible de décomposer  $X_t$  en une composante déterministe  $d_t$  et une composante stochastique  $u_t$  tel que :

$$X_t = d_t + u_t$$

où :

$$u_t = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i}$$

avec  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$

Le théorème de wold montre ainsi que tout processus stationnaire peut s'écrire sous la forme d'une somme de deux composantes, un composante déterministe et une composante stochastique.

### 1.2.2 Processus linéaire

**Définition 1.5.** Un processus linéaire est un processus stochastique  $(X_t)_{t \in \mathbb{T}}$  formé par une combinaison linéaire (non nécessairement finie) de bruit blanc  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$

**Définition 1.6.**  $(X_t)_{t \in \mathbb{Z}}$  est un processus linéaire de moyenne  $m$  s'il peut être écrit sous la forme :

$$X_t = m + \sum_{k=-\infty}^{+\infty} b_k \varepsilon_{t-k}$$

Avec,  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$  et telle que :  $\sum_{k=-\infty}^{+\infty} b_k^2 < +\infty$

L'écriture polynomial :

$$X_t = m + b(B)\varepsilon_t$$

### 1.2.3 Processus linéaires stationnaire

#### Les processus ARMA(p,q)

Les processus ARMA (autoregressive moving average) sont des processus linéaires très important dans la modélisation des séries chronologiques.

#### 1.2.4 Processus AR(p)(autoregressive)

**Définition 1.7.** On appelle processus auto régressif d'ordre  $p$ , noté  $AR(p)$ , un processus  $(X_t)$  vérifier une relation du type :

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \varepsilon_t$$

ù les  $\phi_i (i = 1, \dots, p)$  sont des réels et  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$

En utilisant l'opérateur de retard, la relation peut encore s'écrire :

$$(1 - \phi_1 B - \dots - \phi_p B^p) X_t = \varepsilon_t$$

Soit

$$\Phi(B) X_t = \varepsilon_t$$

où

$$\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

**Auto corrélation et équation de yule-walker**

Soit  $(X_t)_{t \in T}$  un processus autoregressive  $AR(p)$  on écrit :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

On a :

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = E(X_t X_{t-h})$$

Car :  $E(X_t) = E(X_{t-h}) = 0$  ( $AR(p)$  possède une représentation  $MA(\infty)$ , on suppose que  $X_t$  stationnaire) Alors,

$$\begin{aligned} \gamma(h) &= E(X_t X_{t-h}) \\ &= E(X_{t-h}(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \varepsilon_t)) \\ &= E(\phi_1 X_{t-h} X_{t-1}) + E(\phi_2 X_{t-h} X_{t-2}) + \dots + E(X_{t-h} \varepsilon_t) \\ &= \phi_1 E(X_{t-1} X_{t-k}) + \dots + \phi_p E(X_{t-p} X_{t-k}) + E(\varepsilon_t X_{t-k}) \end{aligned}$$

On a :

$$E(X_{t-h} \varepsilon_t) = \begin{cases} 0 & h \neq 0 \\ \sigma_\varepsilon^2 & h = 0 \end{cases}$$

et  $E(X_i X_j) = \gamma_{i-j}$

donc ;

$$\gamma_h = \begin{cases} \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} & \text{si } 0 < h \leq p \\ \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} + \sigma_\varepsilon^2 & \text{si } h=0 \end{cases}$$

On notons :  $\rho_h = \frac{\gamma_h}{\gamma_0}$

donc :

$$\rho_h = \sum_{i=1}^p \phi_i \rho_{h-i}, \forall h > 0$$

Les autocorrélations d'un processus  $AR(p)$  sont ainsi décrites par une équation de récurrence

linéaire d'ordre  $p$ . En écrivant cette relation pour différentes valeurs de  $k$  ( $k=1,2,\dots,p$ ), on obtient les équations de Yule-walker :

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_p \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & & \dots & \rho_{p-2} \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \rho_{p-1} & \rho_{p-2} & & \dots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \phi_p \end{pmatrix}$$

### Auto corrélation partielle

Il est possible de calculer les autocorrélations partielles du processus AR à partir des équations de Yulle -Walker et des auto corrélations.

**propriété 1.2.1.** *Pour un processus AR( $p$ ),  $\phi_{kk} = 0, \forall k > p$ . En d'autres termes, pour un processus AR( $p$ ) les auto corrélations partielles s'annulent à partir du rang  $p + 1$ .*

### Les propriétés d'un processus AR( $p$ )

- La stationnarité et la causalité :

Le processus autorégressif AR( $p$ ) est causal et stationnaire ssi le polynôme  $\Phi_z = 1 -$

$\sum_{k=1}^p \phi_k z^k$  est tel que :

$\Phi_z \neq 0, \forall z \in (\mathbb{C}),$  telle que :  $|z| \leq 1$  ou  $\Phi_z = 0, \forall z \in (\mathbb{C}),$  telle que :  $|z| > 1$

- L'inversibilité : AR( $p$ ) toujours est un processus inversible.

### 1.2.5 Processus MA( $q$ ) (moyenne mobile)

**Définition 1.8.** *On appelle processus moyenne mobile d'ordre  $q$  noté MA( $q$ ), un processus  $X_t$  stationnaire vérifiant une relation de type :*

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

où les  $\theta_i$  ( $i = 1, \dots, q$ ) sont des réels et  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$ .

En utilisant l'opérateur de retard, la relation peut encore s'écrire :

$$X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t$$

Soit encore :

$$X_t = \Theta(B) \varepsilon_t$$

Avec,

$$\Theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)$$

### Auto covariance et l'auto corrélation

La fonction d'auto covariance d'un processus MA(q) est donné par :

$$\gamma_h = cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = E(X_t X_{t-h})$$

(car  $E(X_t) = E(X_{t-k}) = 0$ )

$$\gamma_h = E(X_t X_{t-h}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-h} - \dots - \theta_q \varepsilon_{t-h-q})]$$

Quelque calculs simples conduisant alors à l'expression suivante :

$$\gamma_h = \begin{cases} \sigma_\varepsilon^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & \text{si } 0 \leq h \leq q \\ 0 & \text{si } h > q \end{cases}$$

où,  $\theta_0 = 1$

Donc la fonction d'auto corrélation  $\rho_h = \frac{\gamma_h}{\gamma_0}$  est :

$$\rho_h = \begin{cases} \frac{-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h > q \end{cases}$$

où  $\gamma_0 = 1 + \theta_1^2 + \dots + \theta_q^2$

**propriété 1.2.2.** Pour un processus MA(q),  $\rho_h = 0$  pour  $h > q$ . En d'autre termes les auto corrélations s'annulent à partir du rang  $q+1$ .

### Auto corrélation partielle

La fonction d'auto corrélation partielle d'un MA(q) n'a pas de propriété particulière et son expression est relativement compliquée.

### Les propriétés d'un processus $MA(q)$

- La causalité : le processus moyenne mobile  $MA(q)$  est toujours causale. ( $MA(q)$  est linéaire donc  $\sum_{i=0}^q \theta_i^2 < \infty$ )
- La stationnarité : le processus  $MA(q)$  est toujours stationnaire.

### Démonstration. (Stationnaire $MA(q)$ )

1.  $E(X_t) = E(\sum_{j=0}^q \theta_j \varepsilon_{t-j}) = \sum_{j=0}^q \theta_j E(\varepsilon_{t-j}) = 0 < \infty (\varepsilon_t \sim bb(0, \sigma_\varepsilon^2))$ .
2.  $\gamma_h = cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = E(X_t X_{t-h})$

$$\begin{aligned} \gamma_h &= E(X_t X_{t+h}) = E\left[\left(\sum_{j=0}^q \theta_j \varepsilon_{t-j}\right)\left(\sum_{k=0}^q \theta_k \varepsilon_{t+h-k}\right)\right] \\ &= E\left(\sum_{j=0}^q \sum_{k=0}^q \theta_j \theta_k \varepsilon_{t-j} \varepsilon_{t+h-k}\right) \\ &= \sum_{j=0}^q \sum_{k=0}^q \theta_j \theta_k E(\varepsilon_{t-j} \varepsilon_{t-(k-h)}) \\ &= \begin{cases} \sigma_\varepsilon^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+h} & \text{si } j = k - h \\ 0 & \text{sinon} \end{cases} \\ \gamma_h &= \begin{cases} \sigma_\varepsilon^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+h} & \text{si } 0 < h < q \\ 0 & \text{si } h > q \end{cases} \end{aligned}$$

donc  $\gamma_h$  indépendant de  $t$ .

3.  $E(X_t^2) = var(X_t) = \gamma_0 = \sigma_\varepsilon^2 \sum_{j=0}^q \theta_j^2 < \infty$

- L'inversibilité : un processus  $MA(q)$  est inversible ssi le polynome  $\Theta(z) = \sum_{j=0}^q \theta_j z^j$  où  $\theta_0 = 1$  est tel que :

$\Theta(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| \leq 1$

ou,  $\Theta(z) = 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| > 1$

### 1.2.6 Le processus ARMA(p,q)

Les modèles ARMA sont un mélange des modèles AR(auto régressif) et MA(moyenne mobile).

**Définition 1.9.** *Un processus linéaire stationnaire  $(X_t)_{t \in \mathbb{Z}}$  est dit ARMA(p,q),  $(p,q) \in (\mathbb{N}^*)^2$ , s'il existe des constantes :  $\phi_1, \phi_2, \dots, \phi_p$  ( $\phi_p \neq 0$ ), et  $\theta_1, \dots, \theta_q$  ( $\theta_q \neq 0$ ) et un processus  $(\varepsilon_t) \sim bb(0, \sigma_\varepsilon^2)$  tel que :*

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

En utilisant l'opérateur de retard  $\beta$ , l'équation devient :

$$X_t - \phi_1 B X_t - \phi_2 B^2 X_t - \dots - \phi_p B^p X_t = \varepsilon_t + \theta_1 \beta \varepsilon_t + \dots + \theta_q B^q \varepsilon_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t$$

On ajoute :

$$\Phi(\beta) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

et

$$\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

donc,

$$\Phi(B) X_t = \Theta(B) \varepsilon_t$$

#### Auto covariance et l'auto corrélation (ACF d'un ARMA(p,q))

Pour calculer les auto corrélations d'un processus ARMA. On procède comme dans le cas des processus AR. A partir de l'équation :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

On a

$$\gamma(h) = cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t) E(X_{t-h}) = E(X_t X_{t-h})$$

(car  $E(X_t) = E(X_{t-k}) = 0$ , on suppose  $X_t$  stationnaire)

Donc,

$$\begin{aligned}
\gamma(h) &= \text{cov}(X_t, X_{t-h}) \\
&= E(X_t X_{t-h}) \\
&= \phi_1 E(X_{t-1} X_{t-h}) + \dots + \phi_p E(X_{t-p} X_{t-h}) + E(\varepsilon_t X_{t-h}) + \theta_1 E(\varepsilon_{t-1} X_{t-h}) + \dots + \theta_q E(\varepsilon_{t-q} X_{t-h}).
\end{aligned}$$

Puisque,  $\varepsilon_t \sim \text{bb}$  donc  $E(\varepsilon_{t-j} X_{t-h}) = 0, h > q$  On en déduit que :  $\gamma_h = \phi_1 E(X_{t-1} X_{t-h}) + \dots + \phi_p E(X_{t-p} X_{t-h})$  On a :

$$E(X_i X_j) = \gamma_{i-j}$$

Donc :

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p}$$

Et on a :

$$\rho_h = \frac{\gamma_h}{\gamma_0}$$

Alors,

$$\rho_h = \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2} + \dots + \phi_p \rho_{h-p}$$

Ou bien :

$$\rho_h = \sum_{i=1}^p \phi_i \rho_{h-i}$$

On constate donc que la fonction d'auto corrélation des processus ARMA répond à la même équation aux différences que celle des processus AR.

### Auto corrélation partielle (PACF d'un ARMA(p, q))

La fonction d'Auto corrélation partielle d'un processus ARMA n'a pas d'expression simple. Elle se caractérise le plus fréquemment soit par une forme exponentielle décroissante, soit par une forme aléatoire.

### Les propriétés d'un processus ARMA(p, q)

- **Stationnarité et causalité :**

Un processus ARMA(p, q) est stationnaire et causal ssi  $\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \neq 0, \forall z \in \mathbb{C}$  tel que  $|z| \leq 1$

• **Inversibilité :**

Un processus ARMA(p,q) est inversible ssi  $\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \neq 0, \forall z \in \mathbb{C}$  tel que  $|z| \leq 1$

### 1.2.7 processus ARIMA(p, d, q)[5]

**Définition 1.10.** on dit que  $X_t$  est un modèle ARMA intégré ("Auto-Regressive Moving Average") abrégé ARIMA(p, d, q), si  $\Delta^d X_t$  est un ARMA(p, q).

on d'autre terme  $X_t \sim ARIMA(p, d, q)$  s'il vérifie une équation du type :

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\varepsilon_t, \forall t, \text{ et } \varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$$

où :

$$\begin{cases} \Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p, \text{ où } \phi_p \neq 0 \\ \Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q, \text{ où } \theta_q \neq 0 \end{cases}$$

sont des polynôme dont les racines sont de module supérieur strictement à 1.

**propriété 1.2.3.** 1-un modèle de type ARIMA(p, d, q) est non stationnaire :il contient des tendances de degré d.

2-Le paramètre d indique le nombre de différences prises pour obtenir la stationnarité.

### L'écriture AR( $\infty$ ) d'un processus MA(q)

Soit un processus MA(q) :  $X_t = \Theta(B)\varepsilon_t$ , si  $\Theta(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| \leq 1$  (inversible), alors on peut écrire  $X_t$  sous la forme AR( $\infty$ ).

$$\varepsilon_t = \Theta^{-1}(B)X_t = \sum_{i=0}^{\infty} \beta_i X_{t-i}$$

Avec,  $\beta_0 = 1$  et  $\sum_{i=0}^{\infty} |\beta_i| < \infty$

### L'écriture MA( $\infty$ ) d'un processus AR(p)

AR(p) s'écrit :  $\Phi(B)X_t = \varepsilon_t$

Si  $\Phi(z) = 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| > 1$  (causale), alors on peut écrire  $X_t$  sous la forme MA( $\infty$ ) :

$$X_t = \Phi^{-1}(B)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

Avec,  $\psi_0 = 1$  et  $\sum_{j=0}^{\infty} |\psi_j| < \infty$

### L'écriture d'un processus ARMA(p,q) sous la forme MA( $\infty$ )

$$\text{On a : } \Phi(B)X_t = \Theta(B)\varepsilon_t$$

$$\text{d'ou : } X_t = \frac{\Theta(B)}{\Phi(B)}\varepsilon_t = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i}$$

Avec :  $\sum_{i=0}^{\infty} |\alpha_i| < \infty$  et  $\alpha_0 = 1$

### Ecriture ARMA(p,q) sous la forme AR( $\infty$ )

$$\text{On a : } \Theta(B)\varepsilon_t = \Phi(B)X_t$$

$$\text{Donc : } \varepsilon_t = \frac{\Phi(B)}{\Theta(B)}X_t = \sum_{i=0}^{\infty} \beta_i X_{t-i}$$

Avec,  $\sum_{i=0}^{\infty} |\beta_i| < \infty$  et  $B_0 = 1$

## 1.2.8 Processus linéaire non stationnaire(TS,DS)

Nelson et Plossen ont retenu en 1982 deux classes de processus non stationnaire peuvent être distinguées :

- Le processus TS(trend stationary) : correspondent à une non stationnarité de type déterministe.
- Le processus DS(difference stationary) : correspondent à une non stationnarité de type stochastique.

### Les processus TS(trend stationary)

$X_t$  est Un processus non stationnaire TS. S'il peut s'écrire sous la forme :[1]

$$X_t = f_t + \varepsilon_t$$

où,

$f_t$  : est une fonction (déterministe) du temps

$\varepsilon_t$  : est un processus stationnaire.

L'exemple le plus simple est celui de la tendance linéaire :

$$X_t = \alpha + \beta t + \varepsilon_t$$

( $f_t$  fonction polynomiale d'ordre 1)

### Les caractéristique des processus TS

Supposant que  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$ , on a alors les propriété suivant :

- $E(X_t) = E(\alpha + \beta t + \varepsilon_t) = \alpha + \beta t$ .
- $var(X_t) = E(X_t - E(X_t))^2 = E(\varepsilon_t)^2 = var(\varepsilon_t) = \sigma_\varepsilon^2$
- $cov(X_t, X_s) = E[(X_t - E(X_t))(X_s - E(X_s))] = E(\varepsilon_t \varepsilon_s) = 0 \quad \forall t \neq s$ .

Ce processus est en effet non stationnaire puisque son espérance vaut  $\alpha + \beta t$  a la date  $t$ , et donc dépend de  $t$ .

Une des propriétés importantes de ce type de processus est qu'il n'ya pas persistance des chocs : l'influence d'un chocs subit à un instant  $T$  aura tendance à s'estomper au cours du temps, et la variable rejoint alors sa dynamique de long terme, déterminé par  $f(t)$ .

### Le processus DS(difference stationary)

Un processus DS s'écrit sous la forme :

$$X_t = \rho X_{t-1} + \beta + \varepsilon_t$$

$\varepsilon_t$  : processus stationnaire.

Si l'on suppose  $\rho = 1$  et  $\varepsilon_t$  : bruit blanc

Donc on écrire :

$$X_t = X_{t-1} + \beta + \varepsilon_t$$

Si  $\beta = 0$  le processus est dit marche aléatoire sans dérivée.

Si  $\beta \neq 0$  le processus est dit marche aléatoire avec dérivée.

### Les caractéristiques des processus DS

Premièrement on démontre (par récurrence) que :

$$X_t = X_{t-1} + \beta + \varepsilon_t = X_0 + \beta t + \sum_{i=1}^t \varepsilon_i$$

On a :

$$X_t = X_{t-1} + \beta + \varepsilon_t$$

$$X_1 = X_0 + \beta + \varepsilon_1$$

$$X_2 = X_1 + \beta + \varepsilon_2$$

$$X_3 = X_2 + \beta + \varepsilon_3 = (X_1 + \beta + \varepsilon_2) + \beta + \varepsilon_3 = X_0 + \beta + \varepsilon_1 + \beta + \varepsilon_2 + \beta + \varepsilon_3 = X_0 + 3\beta + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

Par récurrence :

$$X_t = X_0 + t\beta + \sum_{i=1}^t \varepsilon_i$$

Alors,

- $E(X_t) = E(X_0 + t\beta + \sum_{i=1}^t \varepsilon_i) = X_0 + t\beta$
- $var(X_t) = var(X_0 + t\beta + \sum_{i=1}^t \varepsilon_i) = t\sigma_\varepsilon^2$
- $cov(X_t, X_{t'}) = E[(X_t - E(X_t))(X_{t'} - E(X_{t'}))] = E[(\sum_{i=1}^t \varepsilon_i)(\sum_{i=1}^{t'} \varepsilon_i)] = \sigma_\varepsilon^2 \min(t, t') \quad t \neq t'$

L'espérance et la variance dépendent du temps, donc le processus non stationnaire (marche aléatoire avec dérivé).

Si le marche aléatoire sans dérivé ( $\beta = 0$ )

$$E(X_t) = X_0$$

$$var(X_t) = t\sigma_\varepsilon^2$$

$$cov(X_t, X_{t'}) = \sigma_\varepsilon^2 \min(t, t') \quad t \neq t'.$$

Donc le marche aléatoire non stationnaire en variance.

**Remarque 1.2.1.** *Un processus (DS) est un processus que l'on peut rendre stationnaire par l'utilisation d'un filtre aux différences :*

$$(1 - B)^d X_t = \beta + \varepsilon_t$$

$\varepsilon_t$  : processus stationnaire

$d$  : ordre de différenciation

Le processus  $X_t$  est intégrer d'ordre  $d$  noté  $(I(d))$ , si  $d=1$  le processus est dit alors le processus du premier ordre.

## 1.3 Tests de stationnarité

Ce sont des tests qui permettent de détecter la présence ou l'absence d'une racine unitaire. Dickey Fuller suggère un test qui révèle la présence ou l'absence de racine unitaire.

### 1.3.1 Test de Dickey Fuller

Dickey et Fuller ont construit leur test à partir des modèles de bases suivants :

\* **Modèle 1** :  $X_t = \rho X_{t-1} + \varepsilon_t$  (sans constant et sans trend)

\* **Modèle 2** :  $X_t = c + \rho X_{t-1} + \varepsilon_t$  ( avec constante)

\* **Modèle 3** :  $X_t = c + \beta t + \rho X_{t-1} + \varepsilon_t$  ( avec constante et tendance)

où  $\varepsilon_t$  : bruit blanc

$\beta, c$  : constantes réelles

Le principe du test consiste à tester :

$$\begin{cases} H_0 : \rho = 1 & \text{(présence du racine unitaire)} \\ H_1 : |\rho| < 1 & \text{(absence du racine unitaire).} \end{cases}$$

Les modèles de base du test étant théorique, l'application du test requiert l'estimation en pratique du modèle :

\* **Modèle 1'** :  $\Delta X_t = \phi X_{t-1} + \varepsilon_t$

\* **Modèle 2'** :  $\Delta X_t = \phi X_{t-1} + c + \varepsilon_t$

\* **Modèle 3'** :  $\Delta X_t = \phi X_{t-1} + c + \beta t + \varepsilon_t$

Où :  $\phi = \rho - 1$  dans les trois modèles.

La t-statistique  $t_{\hat{\phi}}$  qui est donnée par :

$$t_{\hat{\phi}} = \frac{\hat{\phi} - 1}{\sigma^2}$$

$t_{\hat{\phi}}$  sera comparée à la valeur critique tabulé noté  $t_{tab}$  et on applique la règle suivante :

— Si  $t_{\hat{\phi}} < t_{tab}$  : on rejette  $H_0$

— si  $t_{\hat{\phi}} \geq t_{tab}$  : on accepte  $H_0$

En pratique, on n'effectue pas ce test sur les trois modèles mais on procède par une stratégie séquentielle en trois étapes :

#### ⊗ **Étape 1**

On estime le modèle (3) et on test la significativité de la tendance déterministe (test de Student sur le paramètre ).

— Si cette tendance estimée n'est pas significativement différente de zéro (t-statistique de la tendance inférieur aux valeur critiques de la tendance tabulé par Dickey Fuller). alors on passe à l'étape 2.

— Si la tendance est différente de zéro. On teste l'hypothèse nulle unitaire :

- si on accepte  $H_0$ ,  $X_t$  est non stationnaire de type DS.
- Sinon,  $X_t$  est non stationnaire de type TS.

⊗ **Etape 2**

On aura à appliquer cette étape1 on a rejeté l'idée d'une tendance significative. On estime le modèle (2') et on test la significativité de la constante  $c$ .

- Si  $H_0$  est acceptée, alors  $X_t$  est non stationnaire de type DS.
- Sinon,  $X_t$  est stationnaire.

⊗ **Etape 3**

Si l'étape2 detecté une constante nul, alors on estime le modèle (1') et on effectue le test de racine unitaire tel que :

- si  $H_0$  est acceptée,  $X_t$  est non stationnaire de type DS.
- si  $H_0$  est rejetée  $X_t$  stationnaire . [6]

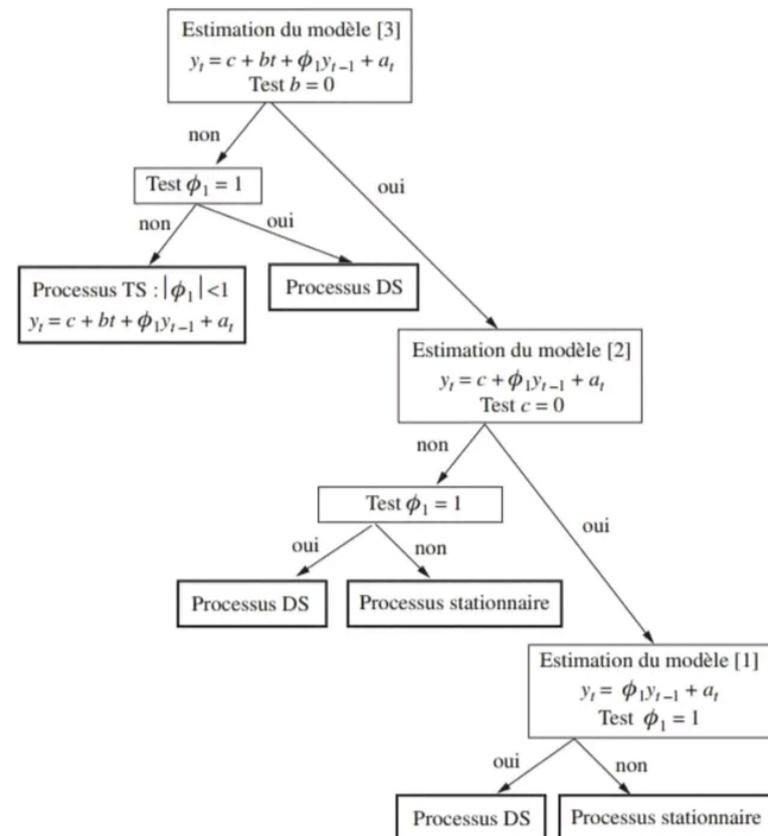


FIGURE 1.1 – schéma 1 de stratégie de dicky-fuller

**Remarque 1.3.1.** *Il est également possible de procéder à des tests l'hypothèses jointes dans le cas des modèles(2) et (3).*

★ Pour le modèle (2) :

$$X_t = \rho X_{t-1} + c + \varepsilon_t$$

-On peut tester l'hypothèse nulle( $c=0, \rho = 1$ ) :

On calcule alors :

$$F_1 = \frac{(SCR_c - SCR_{nc})/2}{SCR_{nc}/(T-2)} \sim F_{(2, T-2)}$$

$SCR_c$  : La somme des carré des résidus du modèle ( $X_t = X_{t-1} + \varepsilon_t$ )

$SCR_{nc}$  : La somme des carré des résidus du modèle (2')

★ Pour le modèle (3) :

$$X_t = \rho X_{t-1} + c + \beta t + \varepsilon_t$$

On peut mener deux test :

tester( $c = 0, \beta = 0, \rho = 1$ )

$$F_2 = \frac{(SCR_c - SCR_{nc})/3}{SCR_n/(T-3)} \sim F_{(3, T-3)}$$

$SCR_c$  : La somme des carrée des résidus du modèle ( $c=0, \beta = 0, \rho = 1$ )

$SCR_{nc}$  : La somme des carré des résidus du modèle(3).

Tester( $c = c, \beta = 0, \rho = 1$ ), on calcule :

$$F_3 = \frac{(SCR_c - SCR_{nc})/2}{SCR_{nc}/(T-3)} \sim F_{(2, T-3)}$$

$SCR_c$  : La somme des carré des résidus du modèle( $\beta = 0, \rho = 1$ )

$SCR_{nc}$  : La somme des carré des résidus du modèle(3).

La stratigie séquentielle :

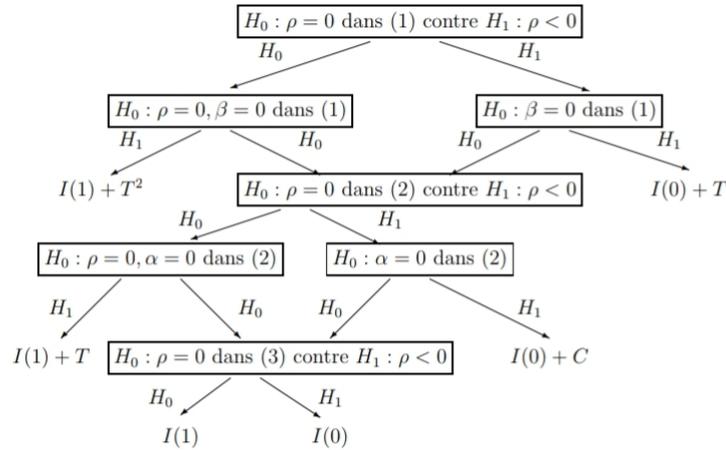


FIGURE 1.2 – schéma 2 de stratigie de Dicky-Fuller

### 1.3.2 Test de Dickey Fuller augmenté

Ce test se déroule exactement comme dans le cas de la version simple, à la seule différence que les modèles de la base à la construction de ce test sont les suivant :

- **Modèle(1)** :  $\Delta X_t = \phi X_{t-1} + \sum_{j=1}^p \phi X_{t-j} + \varepsilon_t$
- **Modèle(2)** :  $\Delta X_t = \phi X_{t-1} + c + \sum_{j=1}^p \phi X_{t-j} + \varepsilon_t$
- **Modèle (3)** :  $\Delta X_t = \phi X_{t-1} + c + \beta t + \sum_{j=1}^p \phi X_{t-j} + \varepsilon_t$

(On adopte la même stratégie séquentielle, aussi les statistiques test sont les même que dans le cas du test DF-simple).

Les tests joints au moyen de statistiques de Fisher dans les modèles (2) et (3).

#### M2

On peut tester  $(c, \phi) = (0, 0)$

La statistique du test :

$$F_4 = \frac{(SCR_c - SCR_{nc})/2}{SCR_{nc}/(T - p - 1)} \sim F_{(2, T-p-1)}$$

$SCR_c$  : La somme des carré résidus du modèle( $c=0, \phi = 0$ )

$SCR_{nc}$  : La somme des carré résidus du modèle (2)

#### M3

Deux types du test :

- Tester  $\{(c, \beta, \phi) = (0, 0, 0)\}$

On calcule :

$$F_5 = \frac{(SCR_c - SCR_{nc})/3}{SCR_{nc}/(T - p - 2)} \sim F_{(3, T-p-2)}$$

$SCR_c$  : La somme des carré résidus du modèle  $(c, \beta, \phi) = (0, 0, 0)$

$SCR_{nc}$  : La somme des carré résidus du modèle (3).

- Tester  $\{(c, \beta, \phi) = (c, 0, 0)\}$

On calcule :

$$F_6 = \frac{(SCR_c - SCR_{nc})/2}{SCR_{nc}/(T - p - 2)} \sim F_{(2, T-p-2)}$$

$SCR_c$  : La somme des carré résidus du modèle  $(c, \beta, \phi) = (c, 0, 0)$

$SCR_{nc}$  : La somme des carré résidus du modèle (3).

## 1.4 Procédure de Box et Jenkins

Les processus ARMA(Autoregressive moving average) ont été introduits par Box et Jenkins (1970). L'objet est de modéliser une série temporelle en fonction de ses valeurs passées, mais aussi en fonction des valeurs présentes et passées d'un bruit. Afin de déterminer le processus ARMA adéquat, Box et Jenkins ont suggéré une procédure en quatre étapes :[8]

- Identification du modèle.
- Estimation des paramètres.
- Validation du modèle.
- Prévision à l'aide du modèle validé.

### 1.4.1 Identification

**Identification** : détermination des ordres  $p, q$  des processus ARMA l'étape suivante consiste à analyser l'auto corrélation et l'auto corrélation partielle.

**Fonction d'auto corrélation** :

La fonction d'auto corrélation estimée d'ordre  $k$  d'un processus stationnaire  $X_t$  s'écrit :

$$\hat{\rho}_k = \frac{\frac{1}{T} \sum_{t=1}^{T-K} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^2}, k = 1, \dots, K; t = 1, \dots, T$$

$T$  : nbr d'observation de  $X_t$ ,  $\bar{X}$  : moyenne de  $X_t$ .

Box et Jenkins suggèrent de retenir un nombre maximal de retard  $K = T/4$

Après avoir évalué la fonction  $\hat{\rho}_k$ , on peut tester la significativité statistique de chaque coefficient d'auto corrélation, pour cela, on utilise la formule de Bartelett donnant l'écart type de la distribution des  $\hat{\rho}_k$  :

$$\hat{\sigma}_{\hat{\rho}_k} = \sqrt{\frac{1}{T}(1 + 2 \sum_{i=1}^{k-1} \hat{\rho}_i^2)}$$

$T$  : nombre d'observation.

Pour un nombre d'observation  $T$  suffisamment important, Bartelett a montré que  $\hat{\rho}_k$  suit une loi normale.

Pour tester :

$$\begin{cases} H_0 : \rho_k & \text{ne sont pas significatif} \\ H_1 : \rho_k & \text{significatif} \end{cases}$$

-La valeur de la statistique de student :

$$t_{\hat{\rho}_k} = \frac{\hat{\rho}_k}{\hat{\sigma}(\hat{\rho}_k)}$$

$$\begin{cases} \text{si } |t_{\hat{\rho}_k}| < t_{1-\alpha/2} & \text{on accepte } H_0 \\ \text{sinon on rejette } & H_0 \end{cases}$$

$t_{(1-\alpha/2)}$  : valeur critique lue dans table de student :

$$\begin{cases} h : \text{nbr de paramètres estimée} \\ (t-h) \text{ degré de liberté} \end{cases}$$

$\Rightarrow$  Ce test permet d'identifier l'ordre  $q$  des processus  $MA$ , car les auto corrélations d'un processus  $MA(q)$  s'annulent à partir du rang  $q + 1$ .

### Fonction d'auto-corrélation partielle

On tester :

$$\begin{cases} H_0 : \phi_{kk} \text{ non significatif} \\ H_1 : \phi_{kk} \text{ significatif} \end{cases}$$

Pour de grands échantillons,  $\phi_{kk} \sim N(0, 1/\Gamma)$ .

• Statistique de test :

$$t_{\hat{\phi}_{kk}} = \frac{\hat{\phi}_{kk}}{\sqrt{1/T}}$$

- La règle de décision :

$$\begin{cases} \text{si } t_{\hat{\phi}_{kk}} < t_{1-\alpha/2} \text{ on accepte } H_0 \\ \text{sinon on rejette } H_0 \end{cases}$$

- où  $t_{1-\alpha/2}$  : valeur critique lue dans la table de loi de student d'ordre  $(1 - \alpha/2)$ , degré de liberté  $(T - h)$ , avec  $h$  : nombre des paramètres estimés.

⇒ Ce test permet d'identifier l'ordre  $p$  des processus AR( $p$ ), car les auto corrélations partielles d'un processus AR( $p$ ) s'annulent à partir du rang  $p + 1$ .

A l'issue de cette étape d'identification, on a sélectionné un ou plusieurs modèles, il convient à présent d'estimer chaque modèle sélectionné.

On passe alors à la deuxième étape.

## 1.4.2 Estimation

A cette étape les ordres  $p, q$  ont été fixés, il reste donc d'estimer les paramètres  $\phi_i, \theta_j$  et aussi  $\sigma_\varepsilon^2$ , nous serons intéressés par les deux méthodes :

- $\Omega$  : matrice  $(T \times T)$  de variance-covariance de vecteur  $X = (X_1, \dots, X_T)^t$  elle dépend de  $\phi_i (i = 1, \dots, p)$  et  $\theta_j (j = 1, \dots, q)$  du processus ARMA.
- $L$  : fonction de vraisemblance.

Cette vraisemblance est difficile à calculer à cause du déterminant  $\Omega$  et de son inverse  $\Omega^{-1}$  quand  $T$  très grand.

On ajoute :

$$S(\phi, \theta) = X^t \Omega^{-1} X = \sum_{t=-\infty}^T (E(X_t, \phi_i, \theta_j, \sigma_\varepsilon^2))^2, i = 1, \dots, p; j = 1, \dots, q$$

La log-vraisemblance d'un processus ARMA( $p, q$ ), est donnée par :

$$\log L_T = \log L(X, \phi, \theta, \sigma_\varepsilon^2) = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log \sigma_\varepsilon^2 - \frac{1}{2} \log(\det \Omega) - \frac{S(\phi, \theta)}{2\sigma_\varepsilon^2}$$

$T$  : nombre d'observation.

On maximise  $\log L_T$  par rapport aux paramètres  $\phi_i, \theta_j$  et  $\sigma_\varepsilon^2$  avec  $i = 1, \dots, p; j = 1, \dots, q$

En pratique, on commence par estimer  $\sigma_\varepsilon^2$  en calculant :

$$\frac{\partial \log L_T}{\partial \sigma_\varepsilon^2} = 0 \Leftrightarrow -\frac{T}{2} \sigma_\varepsilon^2 + \frac{S(\phi, \theta)}{2\sigma_\varepsilon^4} = 0$$

donc :

$$\sigma_\varepsilon^2 = \frac{S(\phi, \theta)}{T}$$

On ajoute  $\sigma_\varepsilon^2$  dans  $\log L_T$  on obtient :

$$\log L_T^* = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log \frac{S(\phi, \theta)}{T} - \frac{1}{2} \log(\det \Omega) - \frac{T}{2}$$

On a ainsi le nombre de paramètres par rapport aux quels la log vraisemblance doit être maximiser. Maximisé la log vraisemblance par rapport aux paramètre autoregressive et moyenne mobile revient à minimiser l'expression suivant :

$$l_T^* = T \log \frac{S(\phi, \theta)}{T} + \log(\det \Omega)$$

La minimisation de cette expression nous permet d'obtenir les estimateurs du maximum de vraisemblance des paramètres  $\phi_i (i = 1, \dots, p)$  et  $\theta_j (j = 1, \dots, q)$  des processus ARMA(p,q). **Estimation par la méthode des moindres carrés :**

Soit  $X_t \sim AR(p)$ , s'écrit sous la forme :

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t = Z^t \beta + \varepsilon_t.$$

où  $\varepsilon_t \sim bb(0, \sigma_\varepsilon^2)$ ,  $Z^t = (1, X_{t-1}, \dots, X_{t-p})$ ,  $\beta^t = (c, \phi_1, \dots, \phi_p)$

L'estimation des paramètres de ce modèle par la méthode des moindres carrés est donné par [1]

$$\hat{\beta} = (ZZ^t)^{-1} ZX_t$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{T - (p + 1)} \sum_{t=1}^T (X_t - Z^t \hat{\beta})^2$$

Si les racines du polynôme caractéristique  $\Phi(z)$  sont à l'extérieure du disque unité alors :

$$\hat{\beta} \xrightarrow{P} \beta \text{ et } \hat{\sigma}_\varepsilon^2 \xrightarrow{P} \sigma_\varepsilon^2$$

Et de plus :  $\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{D} N(0, \sigma_\varepsilon^2 V)$  où  $V = \rho \lim_{T \rightarrow +\infty} \frac{1}{T} ZZ'$

### 1.4.3 Validation

Au début de cette étape on dispose de plusieurs processus ARMA dont on a estimé les paramètres. Il faut maintenant valider ces modèles afin de choisir les meilleur modèles adéquats. Pour cela, on applique des tests sur les paramètres et sur les résidus. Si plusieurs modèles sont validés, l'étape de validation doit se poursuivre par une comparaison des qualités de ces derniers.

**A) Tests sur les paramètres :**

Après avoir estimé les paramètres d'un modèle on peut se poser la question de savoir si ces paramètres sont significative.

L'hypothèse :

$$\begin{cases} H_0 : \text{le modèle ARMA}(p-1, q) (\phi_p = 0) \\ H_1 : \text{le modèle ARMA}(p, q) (\phi_p \neq 0) \end{cases}$$

- Pour  $p' = p - 1$  et  $q' = q$

La statistique du test :  $t_c = \frac{\hat{\phi}_p}{\hat{\sigma}(\hat{\phi}_p)}$

La valeur obtenue est à comparer à la valeur critique lue dans la table de loi de student, la règle de décision est alors la suivant :

$\Rightarrow$  Si  $|t_c| < t_{tab}$ , on accepte  $H_0(\phi_p = 0)$

$\Rightarrow$  Si  $|t_c| \geq t_{tab}$ , on rejette  $H_0(\phi_p \neq 0)$

Bien entendu, on peut appliquer un raisonnement similaire au test de l'hypothèse nulle  $p' = p$  et  $q' = q - 1$

- Pour  $p' = p + 1$  et  $q' = q$  de même façon, on tester la significativité du coefficient  $\phi_{p+1}$ .

L'hypothèse :

$$\begin{cases} H_0 : \text{le modèle ARMA}(p, q) (\phi_p = 0) \\ H_1 : \text{le modèle ARMA}(p + 1, q) (\phi_p \neq 0) \end{cases}$$

la règle de décision est alors la suivant :

$\Rightarrow$  Si  $|t_{\hat{\phi}_{p+1}}| < t_{tab}$ , on accepte  $H_0(\phi_p = 0)$

$\Rightarrow$  Si  $|t_{\hat{\phi}_{p+1}}| \geq t_{tab}$ , on rejette  $H_0(\phi_p \neq 0)$

Bien entendu, on peut appliquer un raisonnement similaire au test de l'hypothèse nulle  $p' = p$  et  $q' = q + 1$

**B) Tests sur les résidus :**

Ces tests ont pour objet de vérifier que les résidus estimés  $\hat{\varepsilon}_t = \frac{\hat{\phi}(\beta)}{\hat{\theta}(\beta)} X_t$ . Suivant bien un processus de bruit blanc.

Il existe plusieurs des tests d'absence d'auto corrélation le test le plus usuel est le test proposé par Box-Piére(1970).

- **Le test de Box et Pière :**

Ce test appelé "test portmanteau". Pour tester l'hypothèse :

$$\begin{cases} H_0 : \hat{\rho}_{\hat{\varepsilon}}(1) = \hat{\rho}_{\hat{\varepsilon}}(2) = \dots = \hat{\rho}_{\hat{\varepsilon}}(k) = 0 \\ H_1 : \exists ! : \hat{\rho}_{\hat{\varepsilon}}(i) \neq 0 \end{cases}$$

La statistique du test :

$$Q = T \sum_{k=1}^K \hat{\rho}_{\hat{\varepsilon}}(k).$$

où  $\hat{\rho}_{\hat{\varepsilon}}(k)$  : est le coefficient d'auto corrélation d'ordre  $k$  des résidus estimés.

$K$  : nombre maximale de retard.

$T$  : nombre d'observation.

La statistique  $Q$  suit une loi de Khi deux de degré de liberté ( $K-p-q$ ), la région critique :(la règle de décision) :

$$\begin{cases} \text{si } Q > \chi_{1-\alpha}^2(K-p-q) \text{ on rejette } H_0 \\ \text{Sinon} & \text{on accepte } H_0 \end{cases}$$

- **Le test de Ljung Box**

Ce test est d'une part comme une amélioration du test de Box et Pièce. La distribution de la statistique du test de Ljung-Box est en effet plus proche de celle du Khi-deux en petit échantillon que ne l'est celle du test de Box-Pièce. La statistique du test s'écrit :

$$Q_{L\beta} = T(T+2) \sum_{k=1}^K \frac{\hat{\rho}_{\hat{\varepsilon}}(k)}{T-k}$$

- **Le test de normalité des résidus**

Le but de ce test est de vérifier la normalité des résidus.

Un test qui permet de tester la normalité du résidus est celui de Jarque-Bera(1980).

Les hypothèses du test :

$$\begin{cases} H_0 : \varepsilon_t \sim N(0,1) \text{(normalité)} \\ H_1 : \varepsilon_t \not\sim N(0,1) \end{cases}$$

La statistique du test est :

$$JB = \frac{T}{6}\beta_1^2 + \frac{T}{24}(\beta_2 - 3)^2$$

où :

$$\beta_1 = \frac{1}{T} \sum_{i=1}^T \left( \frac{X_i - \bar{X}}{S} \right)^3$$

$$\beta_2 = \frac{1}{T} \sum_{i=1}^T \left( \frac{X_i - \bar{X}}{S} \right)^4$$

où :  $\bar{X}$ ,  $S$  : sont respectivement la moyenne et l'écart type empiriques.

La loi asymptotique de JB est la distribution de Khi deux à deux degrés de liberté tel que :

$$\frac{n}{6} \beta_1 \sim \chi_{(1)}^2 \text{ et } \frac{n}{24} (\beta_2 - 3)^2 \sim \chi_{(1)}^2$$

La règle de décision :

$$\begin{cases} \text{Si } Q_{JB} > \chi_{(\alpha,2)}^2 & \text{on rejette } H_0 \\ \text{Sinon } Q_{JB} \leq \chi_{(\alpha,2)}^2 & \text{on accepte } H_0 \end{cases}$$

ou,  $\alpha$  : le risque.

**Remarque 1.4.1.** *En général, les logiciels fournissent la valeur de la statistique de test, ainsi que la p.valeur. On rejette  $H_0$  au niveau  $\alpha$  si (p.valeur)  $< \alpha$*

### C) Les critères de choix de modèles

- **Les critères standard** Ils sont fondés sur le calcul de l'erreur de prévision que l'on cherche à minimiser. On rappelle ici l'expression des trois critères les plus fréquemment utilisés.

— **Erreur absolue moyenne :**

$$MAE = \frac{1}{T} \sum_t |\hat{\varepsilon}_t|$$

— **Racine de l'erreur quadratique moyenne :**

$$RMSE = \sqrt{\frac{1}{T} \sum_t \hat{\varepsilon}_t^2}$$

— **Écart absolu moyen en pourcentage :**

$$MAPE = 100 \frac{1}{T} \sum_t \left| \frac{\hat{\varepsilon}_t}{X_t} \right|$$

$T$  : nombre d'observation.

$\hat{\varepsilon}_t$  : résidus estimés.

Plus la valeur de ces critères est faible. Plus le modèle estimé est proche des observations.

D'autre critère est basé sur la théorie de l'information.

- **Les critères d'information** L'idée sous-jacente consiste à choisir un modèle sur la base d'une mesure de l'écart entre la vraie loi inconnue et le modèle estimé. Cette mesure peut être fournie par la quantité d'information de Kullback. Les différents critères ont alors pour objet d'estimer cette quantité d'information. Il existe plusieurs, nous présentons ici les trois critères les plus fréquemment employés.

— **Le critère AIC (AKaike information criterion 1969)**

$$AIC = \ln \hat{\sigma}_\varepsilon^2 + \frac{2(p+q)}{T}$$

— **Le critère BIC (Shwartz Bayesian Information Criterion 1977)**

$$BIC = \ln \hat{\sigma}_\varepsilon^2 + \frac{(p+q) \ln T}{T}$$

— **Le critère de Hannan-Quinn(1979)**

$$\ln \hat{\sigma}_\varepsilon^2 + \alpha(p+q) \ln\left(\frac{\ln T}{T}\right)$$

$\alpha$  : ( $\alpha > 2$ ) est une constante.

$p$  : l'ordre de la partie AR.

$q$  : l'ordre de la partie MA.

**Remarque 1.4.2.** - *Le critère le plus utilisé est le critère AIC.*

- *L'objet de (AIC) est de faire la prévision.*

- *L'objet de (BIC) et (QH) est de s'ajuster à la série observée.*

#### 1.4.4 Prévision

Étant donné une série stationnaire  $(X_t)_{t \in \mathbb{Z}}$ . Observé entre 1 et T, on cherche à faire la prévision à l'horizon h. Et donc à prévoir :  $X_{t+1}, X_{t+2}, \dots, X_{t+h}$ .

On définit :  $\hat{X}_{t+h} = E(X_{t+h}/I_t)$

telle que :  $X_{t+h}$  : la prévision faite en t pour la date t + h.

$$I_t = (X_1, X_2, \dots, X_t, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$$

L'espérance est ici pris au sens d'espérance conditionnelle, elle représente la meilleure prévision de la série X conditionnellement à l'ensemble d'informations disponible. Donc on a :

$$\hat{X}_{t+h} = E(X_{t+h}/I_t) = \hat{X}_{t+h} \quad h = 1, 2, \dots$$

$$\hat{X}_{t-h} = E(X_{t-h}/I_t) = X_{t-h} \quad h = 0, 1, \dots$$

$$\hat{\varepsilon}_{t+h} = E(\varepsilon_{t+h}/I_t) = 0 \quad h = 1, 2, \dots$$

$$\hat{\varepsilon}_{t-h} = E(\varepsilon_{t-h}/I_t) = \varepsilon_{t-h} \quad h = 0, 1, \dots$$

Considérons le processus  $ARMA(p, q)$  :

$$\Phi(B)X_t = \Theta(B)\varepsilon_t$$

Prenons l'exemple  $ARMA(1, 1)$  :

$$X_t = \phi_1 X_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Avec,  $|\phi_1| < 1$  et  $|\theta_1| < 1$  (stationnarité et inversibilité).

$$\bullet X_{t+1} = \phi_1 X_t + \varepsilon_{t+1} - \theta_1 \varepsilon_t.$$

$$\hat{X}_{t+1} = E(X_{t+1}/I_t) = \phi_1 X_t - \theta_1 \varepsilon_{t-1}.$$

$$\bullet X_{t+2} = \phi_1 X_{t+1} + \varepsilon_{t+2} - \theta_1 \varepsilon_{t+1}.$$

$$\hat{X}_{t+2} = E(X_{t+2}/I_t) = E(\phi_1 X_{t+1} + \varepsilon_{t+2} - \theta_1 \varepsilon_{t+1}/I_t) = \phi_1 \hat{X}_{t+1}.$$

On en déduit :

$$\hat{X}_{t+h} = \phi_1 \hat{X}_{t+h-1} \quad \forall h > 1$$

Il est également possible d'utiliser la forme  $MA(\infty)$  d'un  $ARMA$  pour effectuer les prévisions.

On a :

$$\Phi(B)X_t = \Theta(B)\varepsilon_t \Rightarrow X_t = \frac{\Theta(B)}{\Phi(B)}\varepsilon_t = \psi(B)\varepsilon_t$$

On écrit  $(X_t)$  sous la forme d'une  $MA(\infty)$ .

$$X_t = \varepsilon_t + \sum_{j=1}^{+\infty} (\psi_j \varepsilon_{t-j})$$

$$\hat{X}_{t+h} = \sum_{i \geq 0} \psi_{h+i} \varepsilon_{t-i}$$

On a la prévision de l'erreur :

$$\hat{\varepsilon}_{t+h} = X_{t+h} - \hat{X}_{t+h}$$

**Par exemple :**

$$\hat{\varepsilon}_{t+1} = X_{t+1} - \hat{X}_{t+1} = \varepsilon_{t+1}$$

$$\hat{\varepsilon}_{t+2} = X_{t+2} - \hat{X}_{t+2} = \varepsilon_{t+2} + \psi_1 \varepsilon_{t+1}$$

En général :

$$\hat{\varepsilon}_{t+h} = X_{t+h} - \hat{X}_{t+h} = \sum_{i=0}^{h-1} \psi_i \varepsilon_{t+h-i}$$

avec  $\psi_0 = 1$

La variance de l'erreur de prévision : ( $\varepsilon_t$  bruit blanc gaussienne)

$$\text{var}(\hat{\varepsilon}_{t+h}) = E\left(\sum_{i=0}^{h-1} \psi_i \varepsilon_{t+h-i}\right)^2 = \left(\sum_{i=0}^{h-1} \psi_i E(\varepsilon_{t+h-i})\right)^2 = \sigma_\varepsilon^2 \sum_{i=0}^{h-1} \psi_i^2$$

L'intervalle de prévision :

$$X_t \in \left[ \hat{X}_{t+h} - 1.96\sigma_\varepsilon \left( \sum_{i=0}^{h-1} \psi_i^2 \right)^{\frac{1}{2}}, \hat{X}_{t+h} + 1.96\sigma_\varepsilon \left( \sum_{i=0}^{h-1} \psi_i^2 \right)^{\frac{1}{2}} \right]$$

Avec : ( $U_{1-\frac{\alpha}{2}} = 1.96$ ) est la quantile d'ordre  $(1 - \frac{\alpha}{2})$  de loi normale centré réduit, où ( $\alpha = 5\%$ )

---

---

# CHAPITRE 2

---

## LES SÉRIES TEMPORELLES MULTIVARIÉES ET MODÈLE $VAR(P)$

### 2.1 Présentation du Modèle

la représentation VAR à  $N$  variables et à  $p$  décalage  $VAR(p)$  s'écrit sous la forme matricielle :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon(t)$$

où :

$$Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{nt} \end{bmatrix}; \phi_0 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}; \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}; \phi_i = \begin{bmatrix} a_{11i} & a_{12i} & \dots & a_{1ni} \\ a_{21i} & a_{22i} & \dots & a_{2ni} \\ \vdots & \dots & \dots & \vdots \\ a_{n1i} & a_{n2i} & \dots & a_{nni} \end{bmatrix}; \quad (2.1)$$

- $\varepsilon(t) \sim bb(0, \Sigma_\varepsilon)$   
 $\Sigma_\varepsilon = E(\varepsilon'\varepsilon)$  : est la matrice de variance covariance d'erreur
- la forme (1) s'écrit encore :  $\Phi_B Y_t = \phi_0 + \varepsilon(t)$   
 $\Phi$  : polynôme matricielle ( $\mathbb{N} * \mathbb{N}$ ) telle que :

$$\Phi(B) = I - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p = I - \sum_{i=1}^p \phi_i B^i$$

- le nombre de paramètre :  $\mathbf{P} * \mathbf{N}^2$  telle que :
  - $\mathbf{p}$  : nombre de retard
  - $\mathbf{N}$  : nombre de variable

**Exemple 2.1.1.** considérons deux variable stationnaires  $y_{1t}$  et  $y_{2t}$  et suppose que :  $p = 3$   
le modèle  $var(4)$  décrivant :  $N = 2, p = 3$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} a_{111} & a_{121} \\ a_{211} & a_{221} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} a_{112} & a_{122} \\ a_{212} & a_{222} \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} a_{113} & a_{123} \\ a_{213} & a_{223} \end{bmatrix} \begin{bmatrix} y_{1t-3} \\ y_{2t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$\begin{cases} y_{1t} = a_1 + a_{111}y_{1t-1} + a_{121}y_{2t-1} + a_{112}y_{1t-2} + a_{122}y_{2t-2} + a_{113}y_{1t-3} + a_{123}y_{2t-3} + \varepsilon_{1t} \\ y_{2t} = a_2 + a_{211}y_{1t-1} + a_{221}y_{2t-1} + a_{212}y_{1t-2} + a_{222}y_{2t-2} + a_{213}y_{1t-3} + a_{223}y_{2t-3} + \varepsilon_{2t} \end{cases}$$

$$\begin{cases} y_{1t} = a_1 + \sum_{i=1}^3 a_{11i}y_{1t-i} + \sum_{i=1}^3 a_{12i}y_{2t-i} + \varepsilon_{1t} \\ y_{2t} = a_2 + \sum_{i=1}^3 a_{21i}y_{1t-i} + \sum_{i=1}^3 a_{22i}y_{2t-i} + \varepsilon_{2t} \end{cases}$$

**Remarque 2.1.1.** 1-un processus ARMA multivarié ,ce que l'on note processus VARMA qui s'écrit :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

et soit encore :

$$\Phi(B)Y_t = \Theta(B)\varepsilon(t) + \phi_0$$

$tq$  :  $\Phi(B) = I - \sum_{k=1}^p \phi_k B^k$  est une polynôme de degré  $p$  en  $B$

$\Theta(B) = I + \sum_{i=1}^q \theta_i B^i$  est une polynôme de degré  $q$  en  $B$

il est possible de distingés les processus VMA (obtenu quand  $p = 0$ ) :moyennes mobiles multi-variées.on a alors VMA( $q$ )s'écrit :

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

-VMA sont toujours stationnaires et sont inversibles si les racines du polynome caractéristique sont à l'extérieure du disque unité.

### 2.1.1 Ecriture VAR(1) d'un processus VAR(p)

soit  $Y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})' \in \mathbb{R}^N$  satisfait la représentation VAR(p) suivante :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon(t) \Leftrightarrow \Phi_B Y_t = \phi_0 + \varepsilon(t)$$

un processus VAR(p) peut s'écrire sous la forme d'un VAR(1) q'on notera  $Z_t$

$$Z_t = c + \phi Z_{t-1} + \mu_t$$

où :

$$\phi = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ \mathbb{I}_N & 0_N & \dots & 0_N \\ 0_N & \mathbb{I}_N & \dots & 0_N \\ & & \dots & \ddots \\ 0_N & \dots & \dots & \mathbb{I}_N \end{bmatrix} ; Z_t = \begin{bmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{bmatrix} ; c = \begin{bmatrix} \phi_0 \\ 0_N \\ \vdots \\ 0_N \end{bmatrix} ; \mu_t = \begin{bmatrix} \varepsilon_{1t} \\ 0_N \\ \vdots \\ 0_N \end{bmatrix} \quad (2.2)$$

;

**Exemple 2.1.2.** considérons le processus bivarié défini par :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.6 \\ 0.4 & 0.7 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} 0.8 & 0.1 \\ 0.9 & 0.2 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

on a alors :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{1t-1} \\ y_{2t-1} \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.6 & 0.8 & 0.1 \\ 0.4 & 0.7 & 0.9 & 0.2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \\ 0 \end{bmatrix}$$

### 2.1.2 Représentation canonique et processus des innovations :

considérons VAR(p) avec  $\phi_0 = 0$  :

$$\Phi_B Y_t = \varepsilon(t)$$

en peut écrire :

$$Y_t = (\Phi(B))^{-1}\varepsilon(t) = (\tilde{\Phi}(B)/\det \Phi(B))\varepsilon(t)$$

où :

- $y_t$  :représentation canonique

- $\varepsilon(t)$  :l'innovation du processus

## 2.2 Stationnarité

**Définition 2.1.** *le processus est stationnaire si :*

- $E(Y_t) = \mu$  vecteur constant,  $\forall t$
- $\text{var}(Y_t) = E([Y_t - \mu][Y_t - \mu]') = \Gamma$  :vecteur constant d'elements finis
- $\text{cov}(Y_t, Y_{t+h}) = E([Y_t - \mu][Y_{t+h} - \mu]) = \Gamma_h$  independant de  $t$  mais uniquement du retard.  $\forall t$

**propriété 2.2.1.** (Stationnarité)

- le processus VAR( $p$ )est stationnaire si :

$$\det \Phi(\mathbf{z}) \neq 0, |z| > 1$$

(tout les racines du  $\det \Phi(\mathbf{z}) = \det(I - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p)$  sont de module supérieure à 1 (à l'extérieur du cercle unité ))

- si :

$$\det \Phi(\mathbf{z}) \neq 0, |z| < 1$$

(non stationnaire), on put changer les racines en leur inverse et modifier le bruit blanc associé afin de se ramener à la représentation canonique.

- si au moins des racines de  $\det \Phi(\mathbf{z}) = 1$ , le processus n'est plus stationnaire et on ne peut pas se ramener à une représentation canonique.

**Exemple 2.2.1.** le processus bivarié défini par :[2]

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0.7 & 0.4 \\ 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

le polynome caractéristique :

$$\det \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.7 & 0.4 \\ 0.2 & 0.3 \end{bmatrix} \mathbf{z} \right) = \begin{bmatrix} 1 - 0.7\mathbf{z} & -0.4\mathbf{z} \\ -0.2\mathbf{z} & 1 - 0.3\mathbf{z} \end{bmatrix} = 1 - z + 0.13z^2$$

le polynôme admet pour racines : $\mathbf{z} = 0.84$  et  $\mathbf{z} = -0.15$  :donc le processus n'est pas stationnaire.

**Exemple 2.2.2.** processus bivarié défini par :

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.7 \\ 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

le polynôme caractéristique

$$\det \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.2 & 0.7 \\ 0.3 & 0.4 \end{bmatrix} \mathbf{z} \right) = \begin{bmatrix} 1 - 0.2\mathbf{z} & -0.7\mathbf{z} \\ -0.3\mathbf{z} & 1 - 0.4\mathbf{z} \end{bmatrix} = 1 - 0.6\mathbf{z} + 0.13\mathbf{z}^2$$

le polynôme admet pour racines  $\mathbf{z} = 1.3$  et  $\mathbf{z} = -5.91$  : donc le processus est stationnaire

**Remarque 2.2.1.** le processus VAR sont toujours inversibles.

## 2.3 Les Moments de processus VAR(p)

### 2.3.1 Esperance

on a VAR(p) s'écrit :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} + \varepsilon(t)$$

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \phi_0 + \varepsilon(t)$$

$$\mathbb{E}[Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p}] = \mathbb{E}[\phi_0 + \varepsilon(t)]$$

car  $\mathbb{E}[Y_t] = \mathbb{E}[Y_{t-1}] = \dots = \mathbb{E}[Y_{t-p}]$  et  $\mathbb{E}[\varepsilon(t)] = 0$  (on pose  $Y_t$  stationnaire)

$$\mathbb{E}[Y_t](I - \phi_1 - \phi_2 - \dots - \phi_p) = \phi_0$$

$$\mathbb{E}[Y_t] = (I - \phi_1 - \phi_2 - \dots - \phi_p)^{-1} \phi_0$$

par exemple :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon(t)$$

$$\mathbb{E}[Y_t] = \mathbb{E}[\phi_0 + \phi_1 Y_{t-1} + \varepsilon(t)] = \phi_0 + \phi_1 \mathbb{E}[Y_t] + \overbrace{\mathbb{E}[\varepsilon(t)]}^0$$

$$\mathbb{E}[Y_t] = (I - \phi_1)^{-1} \phi_0$$

car  $(\mathbb{E}[Y_t] = \mathbb{E}[Y_{t-1}]$ )(on pose  $Y_t$  stationnaire)

### 2.3.2 Fonction d'auto-covariance

$$\gamma(h) = \mathbb{E}[(Y_t - \mathbb{E}[Y_t])(Y_{t+h} - \mathbb{E}[Y_{t+h}])]$$

on pose :

$$X_t = Y_t - \mathbb{E}[Y_t]$$

$$X_t = (\phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon(t)) - \mathbb{E}[Y_t]$$

$$X_t = \overbrace{(\mathbb{I} - \phi_1 - \phi_2 - \dots - \phi_p)\mathbb{E}[Y_t]}^{\phi_0} + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon(t) - \mathbb{E}[Y_t]$$

$$X_t = \mathbb{E}[Y_t] + \phi_1 \overbrace{(Y_{t-1} - \mathbb{E}[Y_t])}^{X_{t-1}} + \phi_2 \overbrace{(Y_{t-2} - \mathbb{E}[Y_t])}^{X_{t-2}} + \dots + \phi_p \overbrace{(Y_{t-p} - \mathbb{E}[Y_t])}^{X_{t-p}} + \varepsilon(t) - \mathbb{E}[Y_t]$$

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon(t)$$

donc :

$$\begin{aligned} \gamma(0) &= \mathbb{E}[X_t X_t'] = \mathbb{E}[(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon(t)) X_t'] \\ &= \phi_1 \mathbb{E}[X_{t-1} X_t'] + \dots + \phi_p \mathbb{E}[X_{t-p} X_t'] + \mathbb{E}[\varepsilon(t) X_t'] \end{aligned}$$

(on a :  $\mathbb{E}[X_t X_{t+h}] = \gamma(t - (t+h))$ )

$$= \phi_1 \gamma(1)' + \phi_2 \gamma(2)' + \dots + \phi_p \gamma(p)' + \mathbb{E}[\varepsilon(t) X_t']$$

on a :

$$\begin{aligned} \mathbb{E}[\varepsilon(t) X_t'] &= \mathbb{E}[\varepsilon(t)(\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon(t))'] \\ &= \phi_1 \overbrace{\mathbb{E}[\varepsilon(t) X_{t-1}']}^0 + \dots + \phi_p \overbrace{\mathbb{E}[\varepsilon(t) X_{t-p}']}^0 + \mathbb{E}[\varepsilon(t) \varepsilon(t)'] \end{aligned}$$

et encore :

$$\mathbb{E}[\varepsilon(t) \varepsilon(t)'] = \begin{cases} 0 & \text{Si } t \neq t' \\ \Sigma_\varepsilon & \text{Si } t = t' \end{cases}$$

donc :

$$\gamma(0) = \phi_1 \gamma(1)' + \phi_2 \gamma(2)' + \dots + \phi_p \gamma(p)' + \Sigma_\varepsilon$$

donc la formule générale :

$$\gamma(h) = \begin{cases} \phi_1 \gamma(h-1)' + \phi_2 \gamma(h-2)' + \dots + \phi_p \gamma(h-p)' & \text{si } 0 < h \leq p \\ \phi_1 \gamma(h-1)' + \phi_2 \gamma(h-2)' + \dots + \phi_p \gamma(h-p)' + \Sigma_\varepsilon & \text{si } h = 0 \end{cases}$$

par exemple pour VAR(1) :

$$X_t = Y_t - \mathbb{E}[Y_t]$$

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon(t)$$

$$\gamma(0) = \mathbb{E}[X_t X_t'] = \mathbb{E}[(\phi_1 X_{t-1} + \varepsilon(t)) X_t'] = \phi_1 \mathbb{E}[X_{t-1} X_t'] + \mathbb{E}[\varepsilon(t) X_t']$$

or :

$$(\mathbb{E}[\varepsilon(t) X_t'] = \mathbb{E}[\varepsilon(t)(\phi_1 X_{t-1} + \varepsilon(t))']) = \phi_1 \overbrace{\mathbb{E}[\varepsilon(t) X_{t-1}']}^0 + \mathbb{E}[\varepsilon(t) \varepsilon(t)'] = \Sigma_\varepsilon)$$

donc :

$$\gamma(0) = \phi_1 \mathbb{E}[X_{t-1} X_t'] + \Sigma_\varepsilon$$

$$\gamma(0) = \phi_1 \gamma(0)' + \Sigma_\varepsilon$$

pour  $h = 1$  :

$$\gamma(1) = \mathbb{E}[X_t X_{t-1}'] = \mathbb{E}[(\phi_1 X_{t-1} + \varepsilon(t)) X_{t-1}'] = \phi_1 \mathbb{E}[X_{t-1} X_{t-1}']$$

$$= \phi_1 \gamma(0)'$$

**Remarque 2.3.1.** la propriété de système valable dans le cas univarié n'est plus vérifié ici :

pour  $h \neq 0$  :  $\gamma(h) \neq \gamma(-h)$  [2]

**propriété de fonction d'autocovariance** [4]

$$*\gamma(h) = \gamma(h)'$$

$$*|\gamma_{ij}(h)| \leq [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}, \forall h = 0, 1, \dots \text{ et } \forall i, j = 0, \dots, N$$

\* $\gamma_{ii}(h)$  : est la fonction d'autocovariance de la composante  $i, i = 1, \dots, N$

$\gamma_{ij}(h)$  : est la fonction de covariance croisée entre les composantes  $y_{i,t}$  et  $y_{j,t-h}$  du processus  $y_t, \forall i, j = 1, \dots, N$

$$* \sum_{j,N=1}^k a_j' \gamma_{j-h} a_N \geq 0, \forall k = 1, 2, \dots \text{ et } a_1, \dots, a_k \in \mathbb{R}^N$$

### 2.3.3 Fonction d'autocorrélation

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}}$$

aussi pour tout retard h la matrice d'autocorrélation :

$$\rho_h = \begin{bmatrix} \rho_{11}(h) & \rho_{12}(h) & \cdots & \rho_{1N}(h) \\ \rho_{21}(h) & \ddots & \cdots & \rho_{2N}(h) \\ \vdots & & & \\ \rho_{N1}(h) & \cdots & & \rho_{NN}(h) \end{bmatrix}$$

\* $\rho_h = \rho_h'$ , les éléments sur la diagonale correspondant aux autocorrélations usuelles.

### 2.3.4 Fonction d'autocorrélation partielle(identification)

Pour identifier le nombre de retard p, dans le cas multivarié on ne utilise pas la fonction d'autocorrélation partielle car elle est très difficile à calculer.

l'idée est de retenir des ordres de retards p suffisamment grands puis d'en réduire la taille à l'aide de tests.

on choisit ainsi en général un var(4) pour les données trimestrielles ,var(12) pour des données mensuelles...ect,on verra cependant qu'il existe des outils,notamment les critère d'information qui évitent de fixer arbitrairement la valeur de p afin que le nombre de paramètres à estimer  $pN^2$  ne soit pas trop grand .

## 2.4 Estimation des paramètres d'un VAR(p)

En pratique,les paramètres d'un VAR(p) ne sont pas connu.

ces paramètres sont les coefficients de régression, c-à-d : $\phi_0$ (le vecteur de constant)et les éléments des matrices  $\phi_0, \phi_1 \dots \phi_p$ ,il faut donc estimer aussi la matrice des variance covariance des bruit blanc vectoriels( $\Sigma$ )

Il existe différents méthodes :la méthode des moindres carrés(MCO) pour les processus VAR non contraints,et la technique du maximum de vraisemblance.

### 2.4.1 la méthode des moindres carrés ordinaires

le processus VAR(p) :

$$\Phi(B)Y_t = \varepsilon(t) \text{ ou } : \varepsilon_t \sim bb(0, \Sigma_\varepsilon)$$

\* le nombre des paramètres à estimer pour un VAR(p) : $pN^2 + N(N + 1)/2$

$-pN^2$  : paramètre à estimer dans  $\phi$

$N(N + 1)/2$  : paramètre à estimer dans  $\Sigma_\varepsilon$

Décomposons l'écriture du VAR(p) la  $j^{ième}$  équation du VAR(p) s'écrit :

$$\underline{Y}_j = \begin{pmatrix} Y_{j1} \\ Y_{j2} \\ \vdots \\ Y_{jT} \end{pmatrix} = \begin{pmatrix} Y'_0 & \dots & Y'_{1-p} \\ Y'_1 & \dots & Y'_{2-p} \\ \vdots & & \\ Y'_{T-1} & \dots & Y'_{T-p} \end{pmatrix} \psi_j + \varepsilon_j$$

soit encore :

$$\underline{Y}_j = \underline{Y}\psi_j + \varepsilon_j$$

$$\text{où : } \underline{Y} = \begin{pmatrix} Y'_0 & \cdots & Y'_{1-p} \\ Y'_1 & \cdots & Y'_{2-p} \\ \vdots & & \\ Y'_{T-1} & \cdots & Y'_{T-p} \end{pmatrix} ; \varepsilon_j = \begin{pmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \vdots \\ \varepsilon_{jT} \end{pmatrix}$$

et soit :  $Y'_t = (Y_{1t-1}Y_{2t-1}\dots Y_{Nt-1}, Y_{1t-2}\dots Y_{Nt-2}, \dots, Y_{1t-p}\dots Y_{Nt-p})$  le modèle est un processus VAR(p) à N composantes indicées par le temps t :

$$\psi_j = \begin{pmatrix} \phi_{11j} \\ \vdots \\ \phi_{1Nj} \\ \phi_{21j} \\ \vdots \\ \vdots \\ \phi_{pNj} \end{pmatrix} ; \varepsilon_j = \begin{pmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \vdots \\ \varepsilon_{jT} \end{pmatrix}$$

la matrice  $\underline{Y}$  indépendant de j :

$$\underline{Y}_j = \underline{Y}\psi_j + \varepsilon_j$$

on empile les N équation pour retrouver le VAR :

$$\begin{pmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \vdots \\ \underline{Y}_N \end{pmatrix} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1T} \\ Y_{21} \\ \vdots \\ Y_{NT} \end{pmatrix} = \begin{pmatrix} \underline{Y} & 0 & \cdots & 0 \\ 0 & \underline{Y} & \cdots & 0 \\ \cdots & & \ddots & \\ 0 & \cdots & & \underline{Y} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1T} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{NT} \end{pmatrix}$$

on cherche à estimer  $(\psi_1, \psi_2, \dots, \psi_N)'$  la matrice de variance covariance des erreur devient un

peu plus compliquée et s'écrit :

$$\left( \begin{array}{c} \left( \begin{array}{cccc} \sigma_{11} & 0 & \dots & 0 \\ & \ddots & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \sigma_{11} \end{array} \right) & \left( \begin{array}{cccc} \sigma_{12} & 0 & \dots & 0 \\ 0 & \ddots & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \sigma_{12} \end{array} \right) & \dots & \cdot \\ \\ \left( \begin{array}{cccc} \sigma_{21} & 0 & \dots & 0 \\ 0 & \ddots & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \sigma_{21} \end{array} \right) & & \ddots & \dots & \cdot \\ \\ \vdots & & \dots & \ddots & \cdot \\ \\ \cdot & & \cdot & \dots & \left( \begin{array}{cccc} \sigma_{NN} & 0 & \dots & 0 \\ 0 & \ddots & & \\ \cdot & & & \\ \cdot & & & \\ 0 & & & \sigma_{NN} \end{array} \right) \end{array} \right)$$

l'observation de cette matrice indique la présence d'hétéroscédasticité (il n'y a en effet aucune raison pour que  $\sigma_{11} = \sigma_{22} = \dots = \sigma_{NN}$ ) et d'autocorrélation, il se pose en conséquence un problème pour l'application des MCO.

Rappelons en effet que les estimateurs sont sans biais, mais ne sont plus de variance minimale. Il convient dès lors d'utiliser la technique des MCG (moindres carrés généralisés), qui fournit un estimateur BLUE (best linear unbiased estimator).

on peut réécrire la matrice de variance comme :

$$var[\varepsilon] = \Sigma_\varepsilon \otimes \mathbf{I} = \Omega$$

où :  $\Sigma_\varepsilon = \sigma_{ij}$  et  $\otimes$  : désigne le produit de kroeker

Rappelons que  $A \otimes B = a_{ij}B$

par exemple :

$$\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \otimes \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_2b_1 & a_1b_2 & a_2b_2 \\ a_3b_1 & a_4b_1 & a_3b_2 & a_4b_2 \\ a_1b_3 & a_2b_3 & a_1b_4 & a_2b_4 \\ a_3b_3 & a_4b_3 & a_3b_4 & a_4b_4 \end{bmatrix}$$

Nous venons de voir que la matrice de variance covariance des résidus est telle que l'on devrait théoriquement appliquer les MCG cependant, puisque la matrice des variables explicatives est bloc diagonale, on peut appliquer les MCO bloc par bloc. le théorème de Zellner nous montre ainsi qu'estimer chacune des N équations par MCO et équivalent à estimer le modèle par les MCG. Afin de le prouver, considérons le modèle suivant :

$$X = aY + \varepsilon$$

$\varepsilon$  : est un bruit blanc

Rappelons que l'estimateur des MCO est donné par :

$$a_{MCO} = (Y'Y)^{-1}Y'X$$

et que l'estimateur des MCG s'écrit :

$$a_{MCO} = (Y'\Omega^{-1}Y)^{-1}Y'\Omega^{-1}X$$

$\Omega$  : matrice variance covariance de  $\varepsilon$

et on a :

$$Y = \begin{pmatrix} \underline{Y} & 0 & \cdots & 0 \\ 0 & \underline{Y} & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & \cdots & \underline{Y} \end{pmatrix} = \mathbf{I} \otimes \underline{Y}$$

$\mathbf{I}$  : matrice identité.

**Remarque 2.4.1.** Avant d'appliquer les MCG rappelons que l'on a les égalités suivantes concernant le produit de Kronecker :

$$*(A \otimes B)(C \otimes D) = AC \otimes BD$$

$$*(A \otimes B)' = A' \otimes B'$$

$$*(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

Afin de calculer l'estimateur des MCG, commençons par étudier la matrice  $(Y'\Omega^{-1}Y)$  :

$$Y'\Omega^{-1}Y = (\mathbf{I} \otimes Y')(\Sigma_\varepsilon^{-1} \otimes \mathbf{I})(\mathbf{I} \otimes \underline{Y}) = \Sigma_\varepsilon^{-1} \otimes \underline{Y}'\underline{Y}$$

$$\text{avec } \Omega^{-1} = \Sigma_\varepsilon^{-1} \otimes \mathbf{I}$$

on en déduit :

$$(Y'\Omega^{-1}Y)^{-1} = \Sigma_\varepsilon \otimes (\underline{Y}'\underline{Y})^{-1}$$

pour la matrice  $Y'\Omega^{-1}X$

$$Y'\Omega^{-1}X = (\mathbf{I} \otimes Y')(\Sigma_\varepsilon^{-1} \otimes \mathbf{I})X = (\Sigma_\varepsilon^{-1} \otimes \underline{Y'})X$$

on a donc :

$$\hat{a}_{MCG} = \Sigma_\varepsilon \otimes (\underline{Y'Y})^{-1}(\Sigma_\varepsilon^{-1} \otimes \underline{Y'}) = ((\mathbf{I} \otimes \underline{Y'Y})^{-1}\underline{Y'})X$$

D'où :

$$\hat{a}_{MCG} = \begin{pmatrix} (\underline{Y'Y})^{-1}Y' & 0 & \dots & 0 \\ 0 & (\underline{Y'Y})^{-1}Y' & \dots & 0 \\ \vdots & & \ddots & \\ 0 & \dots & & (\underline{Y'Y})^{-1}Y' \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} (\underline{Y'Y})^{-1}Y'X_1 \\ \vdots \\ \vdots \\ (\underline{Y'Y})^{-1}Y'X_N \end{pmatrix} \quad (2.3)$$

on retrouve l'estimateur des MCO équation par équation.

cependant, cette technique d'estimation des VAR n'est plus valable dès lors qu'il existe des contraintes sur les paramètres .

Il convient alors d'utiliser la technique du MV.

### 2.4.2 La méthode de maximum de vraisemblance

l'estimation par la méthode de maximum de vraisemblance(MV) nécessite la condition de bruit blanc soit gaussien ,donc que le processus VAR(p) soit gaussien.

donc : la fonction de densité de bruit blanc( $\varepsilon$ ) :

$$F_\varepsilon(\varepsilon) = \frac{1}{(2\pi)^{N/2}(\det\Sigma_\varepsilon)^{1/2}} \exp(-1/2\varepsilon'(\Sigma_\varepsilon)^{-1}\varepsilon)$$

$$\varepsilon = Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p}$$

$$F_\varepsilon(Y) = \frac{1}{(2\pi)^{N/2}(\det\Sigma_\varepsilon)^{1/2}} \exp(-1/2(Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p})'(\Sigma_\varepsilon)^{-1}(Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p}))$$

$$L(Y_1, \dots, Y_T) = \prod_{t=1}^T F_\varepsilon(Y) = \frac{1}{(2\pi)^{TN/2}(\det\Sigma_\varepsilon)^{T/2}} \exp(-1/2 \sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1}\varepsilon)$$

on peut en déduire la fonction de log-vraisemblance qui est :

$$\log L(Y_1, \dots, Y_T) = (-NT/2) \log 2\pi - (T/2) \log(\det \Sigma_\varepsilon) - (1/2) \sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1}\varepsilon$$

on maximise ensuite cette expression afin d'obtenir les estimations de  $\phi_1, \dots, \phi_p$  et de  $\Sigma_\varepsilon$

## 2.5 Validation de modèle VAR(p)

pour choisir le bon ordre du modèle ( $p$ ), on suppose différents tests :

### 2.5.1 la statistique du rapport de vraisemblance

l'objet de ce test est de comparer deux modèles, l'un qui est dit contraint et l'autre est le modèle non contraint(complet)[3]

- la statistique du rapport de vraisemblance :noté  $\lambda$  :

$$\lambda = 2(\ln(L^{Nc}) - \ln(L^c))$$

où :

- $\ln(L^{Nc})$  :la fonction de vraisemblance associée au modèle non contraint.
- $\ln(L^c)$  :la fonction de vraisemblance associée au modèle contraint.

on a le test :

$$\begin{cases} H_0 : \text{le modèle VAR}(p) \text{ est vrais} : \phi_{p+1} = 0 \\ \text{ou} \\ H_1 : \text{le modèle VAR}(p+1) \text{ est vrais} : \phi_{p+1} \neq 0 \end{cases}$$

la technique consiste à estimer le modèle contraint  $VAR(p)$  et un modèle non contraint  $VAR(p+1)$ .

la log-vraisemblance d'un processus VAR s'écrit :

$$\log L(X_1, \dots, X_T) = (-NT/2) \log 2\pi - (T/2) \log(\det \Sigma_\varepsilon) - (1/2) \sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1} \varepsilon$$

- $\sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1} \varepsilon$  est un scalaire.

on notant tr la trace donc :

$$\begin{aligned} \sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1} \varepsilon &= tr \left( \sum_{t=1}^T \varepsilon'(\Sigma_\varepsilon)^{-1} \varepsilon \right) = tr \left( \Sigma_\varepsilon^{-1} \sum_{t=1}^T \varepsilon \varepsilon' \right) = tr \left( T \Sigma_\varepsilon^{-1} \frac{1}{T} \sum_{t=1}^T \varepsilon \varepsilon' \right) \\ &= tr \left( T \Sigma_\varepsilon^{-1} \frac{1}{T} \sum_{t=1}^T T \right) = tr (T \Sigma_\varepsilon^{-1} \Sigma_\varepsilon) = tr(T I_N) = NT \end{aligned}$$

donc :La log-vraisemblance estimée du modèle contraint :

$$\bullet \ln(L^c) = (-NT/2) \log 2\pi - (T/2) \log(\det \hat{\Sigma}_\varepsilon^c) - (1/2)NT$$

donc :La log-vraisemblance estimée du modèle non contraint :

$$\bullet \ln(L^{Nc}) = (-NT/2) \log 2\pi - (T/2) \log(\det \hat{\Sigma}_\varepsilon^{Nc}) - (1/2)NT$$

où :  $\hat{\Sigma}_\varepsilon^c$  et  $\hat{\Sigma}_\varepsilon^{Nc}$  : désigne l'estimateur de la matrice de vraince covariance de résidus du modèle contraint respectivement non contraint.

donc :la statistique du test s'écrit encore :

$$\lambda = T \log\left(\frac{\det \hat{\Sigma}_\varepsilon^c}{\det \hat{\Sigma}_\varepsilon^{Nc}}\right)$$

cette statistique suit une loi du khi-deux à r degré de liberté , où : r nombre de constraints si l'on accepte l'hypothèse nulle ,on peut effectuer un deuxième test :

$$\left\{ \begin{array}{l} H_0 : \text{le modèle } VAR(p-1) \text{ est vrai : } \phi_p = 0 \\ \text{ou} \\ H_1 : \text{le modèle } VAR(p) \text{ est vrai : } \phi_p \neq 0 \end{array} \right.$$

ce test s'effectue de la même façon que précédemment.

on a ainsi une séquence de tests emboîtés dont le but est de déterminer l'ordre p du processus VAR.

### 2.5.2 Tests de bruit blanc des erreurs [2]

De la même façon que pour le modèle AR .Il convient de vérifier que l'erreurs correspondent à un bruit blanc.

on a  $\varepsilon_t$  :processus d'erreur et  $\rho(h)$  la fonction d'autocorrélation.

l'hypothèse nulle est :  $H_0 : \rho(1) = \rho(2) = \dots = \rho(h) = 0$

- la statistique  $\phi$  de Box-Piere(en multivarié) :

$$\Phi_h = n \sum_{i=1}^h \text{trace}(\hat{\rho}(i)\hat{\rho}(0)^{-1} \hat{\rho}(i)\hat{\rho}(0)^{-1})$$

- la statistique  $\phi'$  de Ljung-Box :

$$\Phi'_h = n^2 \sum_{i=1}^h \frac{1}{n-i} \text{trace}(\hat{\rho}(i)\hat{\rho}(0)^{-1} \hat{\rho}(i)\hat{\rho}(0)^{-1})$$

La distribution asymptotique suite une loi de khi-deux.

### 2.5.3 Critère d'information

Afin de déterminer l'ordre p du processus VAR. on peut également utiliser des critère d'information.

on estime un certain nombre de modèle VAR pour un ordre p .on retient le retard p qui minimise les critère AIC,SIC,HQ :

- $AIC = \log \det(\hat{\Sigma}_\varepsilon) + \frac{2N^2p}{T}$

$$\bullet SIC = \log \det(\hat{\Sigma}_\varepsilon) + \frac{N^2 p \log(T)}{T}$$

$$\bullet HQ = \log \det(\hat{\Sigma}_\varepsilon) + \frac{2N^2 p \log(\log(T))}{T}$$

où :

$N$  : nombre de variable

$T$  : nombre d'observation

$\hat{\Sigma}_\varepsilon$  : estimateur de la matrice variance covariance des résidus.

**Remarque 2.5.1.** • le critère AIC donnant des estimateurs efficaces de  $p$ .

• les critères SIC, HQ conduisent à des estimateurs convergents de  $p$ .

## 2.6 Prévision

la prévision à l'horizon  $h$  :

considérons un modèle VAR(1) :

$$Y_t = \phi_0 + \phi_1 Y_t + \varepsilon_t$$

on estime à partir de  $T$  observations :

$$\underline{h=1} : Y_T \hat{=} (1) = \hat{\phi}_0 + \hat{\phi}_1 Y_T$$

$$\underline{h=2} : Y_T \hat{=} (2) = \hat{\phi}_0 + \hat{\phi}_1 Y_T(1) = \hat{\phi}_0 + \hat{\phi}_1(\hat{\phi}_0 + \hat{\phi}_1 Y_T) = \hat{\phi}_0 + \hat{\phi}_1 \hat{\phi}_0 + (\hat{\phi}_1)^2 Y_T$$

En général :

$$Y_T \hat{=} (h) = \hat{\phi}_0 + \hat{\phi}_1 Y_T(h-1) = (I + \hat{\phi}_1 + \dots + \hat{\phi}_1^{h-1}) \hat{\phi}_0 + (\hat{\phi}_1)^h Y_T \quad E(\varepsilon_t) = 0$$

**la variance :**

$$\underline{h=1} : \Sigma_1 = \Sigma$$

$$\underline{h=2} : \Sigma_2 = \Sigma_1 + \hat{\phi}_1 \Sigma \hat{\phi}_1' = \Sigma + \hat{\phi}_1 \Sigma \hat{\phi}_1'$$

$$\underline{h=3} : \Sigma_3 = \Sigma_2 + \hat{\phi}_1^2 \Sigma \hat{\phi}_1^{2'} = \Sigma + \hat{\phi}_1 \Sigma \hat{\phi}_1' + \hat{\phi}_1^2 \Sigma \hat{\phi}_1^{2'}$$

En général :

$$\Sigma_h = \Sigma_{h-1} + \hat{\phi}_1^{h-1} \Sigma \hat{\phi}_1^{h-1'} = \Sigma + \hat{\phi}_1 \Sigma \hat{\phi}_1' + \hat{\phi}_1^2 \Sigma \hat{\phi}_1^{2'} + \hat{\phi}_1^{h-1} \Sigma \hat{\phi}_1^{h-1'}$$

Des formules analogues pour les processus VAR(p).

la variance de l'erreur de prévision de chacun des  $N$  variables est obtenue sur la diagonale des

$\Sigma_h$  : l'intervalle de confiance de la prévision au niveau  $(1 - \alpha/2)$  est donnée :

$$Y_{it}(h) \pm t_{\alpha/2} \sqrt{\Sigma_{ih}^i}$$

avec :

$t_{\alpha/2}$  : fractile de loi normal.

## 2.7 Causalité

### 2.7.1 La notion de causalité

Soit  $X_t, Y_t$  deux processus aléatoire.

La loi du processus couple  $(X_t, Y_t)$  s'écrit à la date  $t$  conditionnellement au passé (noté  $\underline{X}_{t-1}$  et  $\underline{Y}_{t-1}$ ) :  $l(x_t, y_t / \underline{x}_{t-1}, \underline{y}_{t-1})$

Si  $X_t, Y_t$  sont indépendants, alors :

$$l(x_t, y_t / \underline{x}_{t-1}, \underline{y}_{t-1}) = l(x_t / \underline{x}_{t-1}) \times l(y_t / \underline{y}_{t-1})$$

On pose  $I_t$  : un invers,  $\underline{I}_t$  : information relative au passé,  $\underline{\underline{I}}_t$  : information relative au passé et présent telle que :

$$\underline{I}_t = (I_s / s < t) \text{ et } \underline{\underline{I}}_t = (I_s / s \leq t)$$

On pose les processus  $X_t, Y_t$  sont stationnaire alors :

$X$  cause  $Y$  si l'erreur de prévision de  $Y$  est telle que :

$$\delta^2(Y_t / \underline{I}_t) < \delta^2(Y_t / \underline{I}_t - \underline{X}_t)$$

où  $\underline{I}_t - \underline{X}_t$  : l'information obtenue en retirant de  $\underline{I}_t$  les valeurs passé de  $X$ .

La condition signifie que, pour prévoir  $Y_t$ , le passé de  $X$  apporte une information supplémentaire par rapport à la seule prise en compte des autres figurant dans  $\underline{I}$ .

$X$  cause instantanément  $Y$  si :

$$\delta^2(Y_t / \underline{I}_t - \underline{X}_t) < \delta^2(Y_t / \underline{I}_t)$$

La condition signifie que la valeur présente de  $X$  apporte une information supplémentaire par rapport à la connaissance du passé des variables figurant dans  $\underline{I}$ .

Les inégalité sont toujours satisfaites. Elle se transforment en égalités si et seulement si :

- $X$  ne cause pas  $y$  si :  $E(Y_t / I_t) = E(Y_t / \underline{I}_t - \underline{X}_t)$ .
- $X$  ne cause pas instantanément  $Y$  si :  $E(Y_t / \underline{I}_t, \underline{X}_t) = E(Y_t / I_t)$

La définition de la causalité est ici relative à un univers.

On distingue trois grands types de causalité :

### 2.7.2 Causalité au sens de Granger(1969)

La causalité au sens de Granger(1969) est le plus connu on pose  $X_t, Y_t$  deux processus donc le but est de se demander si  $X_t$  cause  $Y_t$  et de voir dans quelle proportion la valeur courante de  $Y_t$  peut être expliquée par ses valeurs passées, et si en ajoutant des valeurs retardées de ( $X_t$ ), l'explication est meilleure plus formellement on a :

**Définition 2.2.** -  $X$  cause  $Y$  à la date  $t$  :

$$E(Y_t/\underline{Y}_{t-1}, \underline{X}_{t-1}) \neq E(Y_t/\underline{Y}_{t-1})$$

-  $X$  ne cause pas  $Y$  à la date  $t$  :

$$V_\epsilon(Y_t/\underline{Y}_{t-1}, \underline{X}_{t-1}) = V_\epsilon(Y_t/\underline{Y}_{t-1})$$

-  $X$  cause instantanément  $Y$  :

$$E(Y_t/\underline{Y}_{t-1}, X_t) \neq E(Y_t/\underline{Y}_{t-1}, X_{t-1})$$

où

$V_\epsilon(Y_t/\underline{Y}_{t-1}, \underline{X}_t)$  : la matrice de variance covariance de l'erreur de prévision associée à la régression linéaire de  $Y_t$  sur son passé ( $\underline{Y}_{t-1}$ ) et sur le passé de  $X_t$  jusqu'à la date  $t-1$  incluse ( $\underline{X}_{t-1}$ ).

#### Test de causalité [7]

Un modèle VAR(p) de dimension 2 expliquant  $X_t, Y_t$ , la première équation de ce modèle s'écrit :

$$y_t = \mu_1 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \epsilon_{1t}$$

Pour tester la non causalité au sens de Granger de X sur Y, il suffit de tester la nullité jointe de  $\beta_1, \beta_2$ .

Par un test en F. l'expression générale est :

$$S_1 = \frac{(SCR_0 - SCR)/P}{SCR/(T - 2P - 1)} \sim F(P, T - 2P - 1)$$

Ce test n'est pas un test exact puisque il ya des endogènes retardées dans la régression OLS.

On emploie donc un test asymptotique du  $\chi^2_{(P)}$  :

$$S_2 = \frac{T(SCR_0 - SCR)}{SCR} \sim \chi^2_{(P)}$$

**Par exemple :**

considérons le processus  $VAR(p)$  à deux variables  $X_t, Y_t$  :

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_{11} & b_{11} \\ a_{12} & b_{12} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} a_{1p} & b_{1p} \\ a_{2p} & b_{2p} \end{pmatrix} \begin{pmatrix} Y_{t-p} \\ X_{t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Pour tester l'absence de causalité de  $X_t$  vers  $Y_t$  :

**l'hypothèse :**

$$\begin{cases} H_0 : b_{11} = b_{12} = \dots = b_{1p} = 0 (X_t \text{ ne cause pas } Y_t) \\ \text{ou} \\ H_0 : a_{11} = a_{12} = \dots = a_{1p} = 0 (Y_t \text{ ne cause pas } X_t) \end{cases}$$

Les tests sont alors tests de Fisher classique.

**Remarque 2.7.1.** Si l'on est amené à rejeter les deux hypothèses nulles. On a une causalité bi-directionnelle. On parle de boucle rétroactive (feedback effect).

### 2.7.3 Causalité au sens de Pierce et Haugh(1977)

Si  $I = (X, Y)$  alors la causalité de X vers Y peut être caractérisée par les corrélations des innovations de deux processus X et Y.

Soient  $Y_t$  le processus des innovations de  $Y_t$  et  $\eta_t$  le processus des innovations de  $X_t$ .

La fonction d'auto corrélation est donné par :

$$\rho_{\eta\mu}(h) = \frac{cov(\eta_t \mu_{t-h})}{\delta_\eta \delta_\mu}$$

Dans ce cas, Y ne cause pas X si  $\rho_{\eta\mu}(h) = 0, \forall h > 0$

(où  $h \geq 0$  Si la causalité instantanée est exclue).

L'innovation de  $X_t$  doit être non corrélée avec tous les innovations passées associées au processus  $Y_t$ .

### 2.7.4 causalité au sens Sims (1980)

selon, Sims(1980), propose de considérer les valeurs futures de  $Y_t$ , si les valeurs futures de Y peuvent permettre d'expliquer les valeurs présentes de X, alors X est la cause de Y.

De façon similaire, on dira que X cause Y si les innovations de X contribuent à la variance de

l'erreur de prévision de Y.

considérons un processus VAR(p) à deux variables :

$$\begin{cases} Y_t = a_1^0 + \sum_{i=1}^p a_{1i}^1 Y_{t-i} + \sum_{i=1}^p a_{1i}^2 X_{t-i} + \sum_{i=1}^p b_i^2 X_{t-i} + \varepsilon_{1t} \\ X_t = a_2^0 + \sum_{i=1}^p a_{2i}^1 X_{t-i} + \sum_{i=1}^p a_{2i}^2 Y_{t-i} + \sum_{i=1}^p b_i^1 Y_{t-i} + \varepsilon_{2t} \end{cases}$$

Dans ce cas :

- Y ne cause pas X si l'hypothèse nulle suivante est vérifiée :  $H_0 : b_1^2 = b_2^2 \dots = b_p^2 = 0$
- X ne cause pas Y si l'hypothèse nulle suivante est vérifiée :  $H_0 : b_1^1 = b_2^1 \dots = b_p^1 = 0$  IL s'agit encore de tests de Fisher de nullité des coefficients.

## 2.8 Analyse de réponse impulsionnelle

### 2.8.1 Représentation VMA d'un processus VAR

la représentation canonique de VAR :

$$\Phi(B)X_t = \varepsilon_t$$

où :  $\varepsilon_t \sim BB$  soit encore :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t$$

Dans ce cas, selon le théorème de Wold, ce processus peut être écrit sous la forme d'un processus VMA infini :

$$X_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} = \Theta(B)\varepsilon_t$$

où :  $\Theta(B) = \sum_{j \geq 0} \theta_j B^j$ ,  $\Theta_0 = I$

### 2.8.2 VAR structurel

Le modèle VAR structurel est un système d'équation simultanées dont la forme réduite est le modèle VAR canonique.

La représentation VAR structurel se déduit de la représentation VAR canonique.

On a la représentation de VAR canonique :

$$X_t = \sum_{i=1}^P \phi_i X_{t-i} + \varepsilon_t$$

$\varepsilon_t$  : Vecteur des innovations canoniques

$X_t$  : Vecteur des séries observables.

On a P : la matrice de passage (inversible et de dimension  $N \times N$ ) qui doit être estimée.

On pré multiplie les deux membres (1) par la matrice  $\hat{P}^{-1}$  ( $\hat{P}$  étant un estimateur de P) :

$$\hat{P}^{-1}X_t = \hat{P}^{-1} \sum_{i=1}^P \phi_i X_{t-i} + \hat{P}^{-1}\varepsilon_t$$

Soit encore :

$$\hat{P}^{-1}X_t + \hat{P}^{-1} \sum_{i=1}^P \phi_i X_{t-i} + \hat{P}^{-1}\varepsilon_t + X_t - X_t = 0$$

Donc,

$$X_t = \hat{P}^{-1} \sum_{i=1}^P \phi_i X_{t-i} + \hat{P}^{-1}\varepsilon_t + (I - \hat{P}^{-1})X_t$$

On ajoute :

$$\hat{P}^{-1}\varepsilon_t = w_t, I - \hat{P}^{-1} = \psi_0 \quad 1 \leq i \leq p, \hat{P}^{-1}\phi_i = \psi_i$$

Donc, l'expression du processus VAR structurel :

$$X_t = \sum_{i=0}^P \psi_i X_{t-i} + \psi_t \quad (2)$$

Les relations précédents montrent que l'estimation du VAR structurel est acquise dès que la matrice P a été estimée.

L'identification des chocs est réalisée puisqu'il est alors possible de passer des chocs estimés aux chocs structurels (interprétables économiquement) par :

$$\hat{\psi}_t = \hat{P}^{-1}\hat{\varepsilon}_t$$

### 2.8.3 Orthogonalisation des chocs

Par exemple le processus VAR(1) suivant :

$$\begin{cases} X_t = a_{11}X_{t-1} + a_{12}Y_{t-1} + \varepsilon_{1t} \\ Y_t = a_{21}X_{t-1} + a_{22}Y_{t-1} + \varepsilon_{2t} \end{cases}$$

On voit qu'un chocs sur  $\varepsilon_1$  affectera immédiatement la valeur présente de  $X_t$ , il affectera aussi les valeurs futures de y et x car les deux équations corrélé par la valeur passé de X.

•Si  $\varepsilon_{1t}$  et  $\varepsilon_{2t}$  ne sont pas corrélées, l'interprétation de la fonction de réponse implusionnelle est très simple.

- La fonction de réponse impulsionnelle pour  $\varepsilon_{2t}$  mesure l'effet chocs de  $y$  sur valeur passées et présentes de  $X$  et  $Y$ .
- On pratique généralement, les innovations sont corrélées, elles ont donc une composante commune qui ne peut pas être associée à une variable spécifique. il ya une méthode quelque peu arbitraire mais fréquemment utilisée consiste à attribuer la totalité de la composante commune à la variable qui intervient en premier dans le système VAR.

Techniquement, les erreurs peuvent être orthogonalisées en utilisant la décomposition de Cholesky (la matrice de variance covariance des innovations qui en résulte est diagonale).

### 2.8.4 Décomposition de Cholesky

Cette méthode permet de résoudre le système  $AX = b$ ,  $b \in \mathbb{R}^N$

$A$  : matrice symétrique définie positive d'ordre  $N$ .

**Théorème 2.8.1.** *Si  $A$  est une matrice symétrique définie positive, il existe une matrice triangulaire inférieure  $P$  telle que :*

$$A = P.P^t$$

De plus, si on pose aux coefficients diagonaux de  $P$  d'être strictement positifs, alors cette décomposition est unique.

**Algorithme :**

$A \in M_N(\mathbb{R})$  tq :

$$A = P.P^t$$

$$P = \begin{pmatrix} p_{11} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & \cdots & \cdots & p_{nn} \end{pmatrix}, P^t = \begin{pmatrix} p_{11} & \cdots & p_{n1} \\ 0 & \ddots & \vdots \\ 0 & \cdots & p_{nn} \end{pmatrix}$$

$$A = \begin{pmatrix} p_{11} & 0 & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & \cdots & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} p_{11} & \cdots & p_{n1} \\ 0 & \ddots & \vdots \\ 0 & \cdots & p_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} & \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

**Rappel**

Soit  $A \in M_n(\mathbb{R})$

A symétrique  $\Leftrightarrow A = A^t$ .

A définie positif si les n déterminants des sous matrices principales sont strictement positif.

**Exemple 2.8.1.**  $N = 3$

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 5 & -4 \\ 2 & -4 & 6 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

On écrit A sous la forme  $P.P^t$  par les décompositions de Cholesky :

On a  $A = A^t$  (A symétrique).

Pour démontrer que A positif, il faut que les 3 déterminant est positif :

$$\Delta_1 = 1 > 0$$

$$\Delta_2 = \begin{vmatrix} 1 & -1 \\ -1 & 5 \end{vmatrix} = 5 - 1 = 4 > 0$$

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 2 \\ -1 & 5 & -4 \\ 2 & -4 & 6 \end{vmatrix} = (30 - 16) + (-6 + 8) + 2(4 - 10) = 14 + 2 - 12 = 4 > 0$$

Donc A définie positif.

Donc  $\exists P$  triangulaire inférieure tq :

$$A = PP^t$$

$$\begin{pmatrix} 1 & -1 & 2 \\ -1 & 5 & -4 \\ 2 & -4 & 6 \end{pmatrix} = \begin{pmatrix} p_{11} & 0 & 0 \\ p_{21} & p_{22} & 0 \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \cdot \begin{pmatrix} p_{11} & p_{21} & p_{31} \\ 0 & p_{22} & p_{32} \\ 0 & 0 & p_{33} \end{pmatrix} = \begin{pmatrix} p_{11}^2 & p_{11}p_{21} & p_{11}p_{31} \\ p_{21}p_{11} & p_{21}^2 + p_{22}^2 & p_{21}p_{31} + p_{22}p_{32} \\ p_{31}p_{11} & p_{31}p_{21} + p_{32}p_{22} & p_{31}^2 + p_{32}^2 + p_{33}^2 \end{pmatrix}$$

Par suite :

$$\begin{cases} p_{11}^2 = 1 \Rightarrow p_{11} = 1 \\ p_{11}p_{21} = -1 \Rightarrow p_{21} = -1 \\ p_{11}p_{31} = 2 \Rightarrow p_{31} = 2 \\ p_{21}^2 p_{22}^2 = 5 \Rightarrow p_{22}^2 = 5 \Rightarrow p_{22} = \sqrt{5} \\ p_{21}p_{31} + p_{22}p_{32} = -4 \Rightarrow (-1 \times 2) + \sqrt{5}p_{32} = -4 \Rightarrow p_{32} = \frac{2}{\sqrt{5}} \\ p_{31}^2 + p_{32}^2 + p_{33}^2 = 6 \Rightarrow (2)^2 + \left(\frac{2}{\sqrt{5}}\right)^2 + p_{33}^2 = 6 \Rightarrow p_{33}^2 = \frac{30}{24} \Rightarrow p_{33} = \frac{\sqrt{30}}{\sqrt{24}} \end{cases}$$

### 2.8.5 Méthode d'identification des chocs

Pour facilité l'identification on pose que :

$$V(W_t) = I \quad (1)$$

Le différents chocs structurels  $W_t$  ne sont pas corrélés entre eux sur une même date, et ont une variance unitaire.

On a :

$$\epsilon_t = PW_t$$

$$V(\epsilon_t) = PV(W_t)P' = PP' = \Sigma_\epsilon$$

ou

$P$  : la matrice de passage ( $N^2$  paramètre inconnu).

$\Sigma_\epsilon$  : la matrice de variance covariance de  $\epsilon_t$  (innovation canonique).

• On pose  $N(N+1)/2$  contraintes sur les éléments de matrice  $P$  (puisque  $\Sigma_\epsilon$  est symétrique). Ces contraintes sont appelées **contraintes d'orthogonalisation**.

• Pour identifier les  $N^2$  éléments de la matrice  $P$ , il reste imposer  $N(N-1)/2$  contraintes supplémentaires sont appelées **contraintes identifiantes structurelles**.

## 2.9 Décomposition de la variance

Considérons le processus  $X_t$  admettant la représentation VAR(p) :

$$A(B)X_t = X_t - A_0 - A_1X_{t-1} - \dots - A_pX_{t-p} = M + \epsilon_t$$

$\epsilon_t$  : bruit blanc de matrice variance covariance  $\Sigma$ .

On suppose  $X_t$  est stationnaire

La forme VMA( $\infty$ ) donc :

$$X_t = \mu + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots = \mu + \Theta(\beta)\epsilon_t = \mu + \sum_{i=0}^{\infty} \theta_i\epsilon_{t-i}$$

L'erreur de prévision à l'horizon  $h$  se note :

$$X_{t+h} - E(X_{t+h}) = \sum_{i=0}^{h-1} \theta_i\epsilon_{t+h-i} = \epsilon_{t+h} + \theta_1\epsilon_{t+h-1} + \dots + \theta_{h-1}\epsilon_t$$

L'espérance vaut alors 0 (par définition, le bruit blanc des processus centré).

La matrice de variance covariance donné par :

$$\begin{aligned} E[(X_{t+h} - E(X_{t+h}))(X_{t+h} - E(X_{t+h}))'] \\ = E(X_{t+h} - E(X_{t+h}))^2 \\ = \Sigma + \theta_1 \Sigma \theta_1' + \dots + \theta_{h-1} \Sigma \theta_{h-1}' \dots (3) \end{aligned}$$

Cette erreur de prévision est alors exprimée en fonction de la matrice de variance covariance des résidus (non diagonale).

Considérons la transformation :

$$\begin{aligned} \mu_t = A^{-1} \varepsilon_t \text{ ou } \varepsilon_t = A \mu_t \text{ et } \Sigma = A.D.A' \\ \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{dt} \end{pmatrix} = (a_1, a_2, \dots, a_n) \begin{pmatrix} \mu_{1t} \\ \vdots \\ \mu_{dt} \end{pmatrix}, a_i \in \mathbb{R}^d \end{aligned}$$

$$\Sigma = E(\varepsilon \varepsilon_t') = a_1 a_1' V(\mu_{1t}) + a_2 a_2' V(\mu_{2t}) + \dots + a_d a_d' V(\mu_{dt})$$

on remplaçant cette expression dans (3) :

$$E[(X_{t+h} - E(X_{t+h}))(X_{t+h} - E(X_{t+h}))'] = \sum_{j=1}^d V(\mu_{jt}) [\theta_1 (a_1 a_1') \theta_1' + \theta_{h-1} (a_{h-1} a_{h-1}') \theta_{h-1}']$$

La quantité :

$$V(\mu_{jt}) [\theta_1 (a_1 a_1') \theta_1' + \theta_{h-1} (a_{h-1} a_{h-1}') \theta_{h-1}']$$

est contribution d'une innovation pure à la variance totale de la prévision à un horizon h.

---

---

## CHAPITRE 3

---

# APPLICATION DE MODÈLE VAR( $P$ ) EN MÉTÉOROLOGIE

### 3.1 Introduction

Après avoir présenter dans le chapitre précédent le cadre théorique de modèle VAR(p), nous poursuivons la démarche que nous avons adopté depuis le début de ce mémoire.

Dans ce contexte, nous entamons une phase pratique consiste à appliquer de toutes les techniques citées auparavant sur trois séries météorologiques mesurées dans la wilaya de Jijel, ces données sont fournit par la station d'Achouat et s'étalent du Janvier 1996 à Décembre 2005. notre base de données est composée de la température (notée Temp), les précipitations (notée Prec), et l'humidité (noté Hum).

Dans ce chapitre, nous allons essayer de mettre en place une modélisation univariée de chaque série séparément via la méthode de Box et Jenkins et une modélisation Multivariée de toutes les séries simultanément par le biais de modèle VAR(p). nous allons essayer par la suite de calculer les prévisions par les modèles résultants des deux approches et faire une comparaisons entre elles.

## 3.2 Etude Univarié

### 3.2.1 Etude de stationnarité

#### Statistiques descriptives

la table 3.1 montre les statistiques de base de chaque série :

	Prec	temp	hum
Mean	83.31	18.15	75.22
Med	64.20	17.10	75.65
Max	407.30	28.30	82.80
Min	0.00	9.00	66.60
1 st Q	15.97	13.60	72.97
3 st Q	121.95	23.43	77.42
Skewness	1.45	0.16	-0.29
Kurtosis	5.22	1.66	2.71
$V_1 = \left( \frac{ SK-0 }{\sqrt{6/120}} \right)$	6.48	0.71	1.29
$V_2 = \left( \frac{ Kr-3 }{\sqrt{24/120}} \right)$	4.96	2.99	0.64

TABLE 3.1 – Statistiques descriptives

nous comparons les valeurs de  $V_1$  et  $V_2$  pour chaque série avec la valeur (1.96 :qui représente la valeur de la loi normal au seuil  $\alpha = 5\%$ ).

D'après les résultats ,on rejette l'hypothèse de normalité pour la série (prec)et la série (temp)et on l'accepte pour la série(Hum).

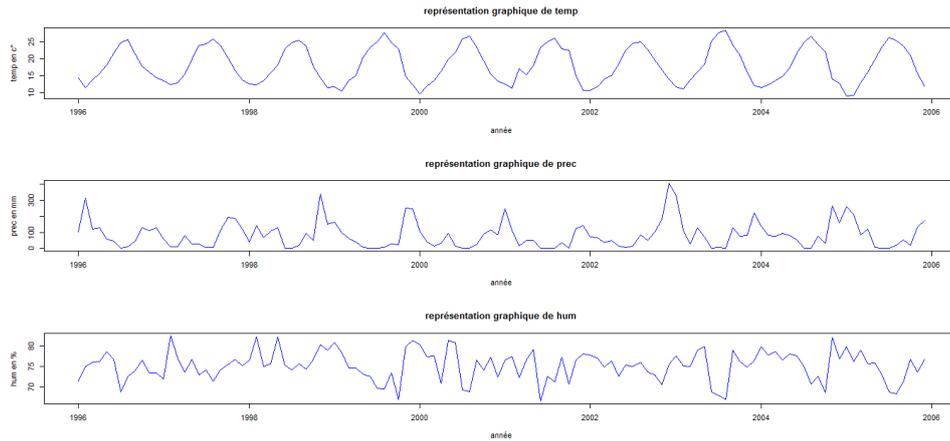


FIGURE 3.1 – Représentations graphiques des trois séries

on constate que les trois courbes sont parallèles à l'axe des abscisses, donc on peut les considérer comme des séries stationnaires.

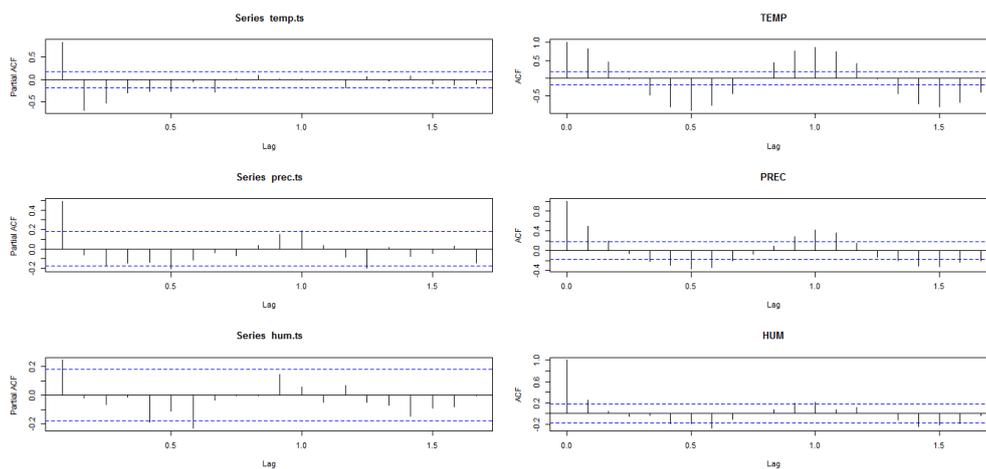


FIGURE 3.2 – ACF et PACF des trois séries

l'analyse de corrélogramme simple est partiel des trois séries nous indique une éventuelle présence de la saisonnalité.

### Etude de saisonnalité

la table suivante regroupe les coefficients saisonniers pour les trois séries, ces coefficients sont au nombre 12 car nos séries sont mensuelles :

	prec	temp	hum
01	63.73423	-6.4980827	2.1226087
02	32.036507	-6.7982657	2.7627454
03	-33.8911057	-4.2984491	0.8328811
04	-0.8844408	-2.6333782	0.3585463
05	-29.4677260	0.6216936	2.6242111
06	-67.4237856	4.8333828	-1.2305805
07	-79.0498108	7.1350719	-3.6653697
08	-66.4602717	8.1398829	-4.0830964
09	-16.5607505	5.3746934	-0.6108214
10	-7.7276270	2.3114726	-1.7190327
11	100.4155347	-2.6617494	1.2927625
12	105.2792275	-5.5262717	1.3151451

TABLE 3.2 – coefficients saisonniers

**test du racine unitaire**

on va appliquer le test de Dicky-Fuller sur les trois séries dépersonnalisées .a l'aide de logiciel R ,on estime par la méthode MV les paramètres des modelés [4],[5],[6] .  
les valeurs mentionnées dans ce tableau représentent les probabilités de signification (p-value).pour plus de détails sur les estimations de ces tests voir l'annexe

	M3 (M6)	M2 (M5)	M1 (M4)
Temp			
trend	0.919	/	/
const	$1.34 * 10^{-8}$	$2 * 10^{-16}$	/
R.U	$1.35 * 10^{-8}$	$2 * 10^{-16}$	/
Prec			
trend	0.29	/	/
const	0.0002	$1.6 * 10^{-7}$	/
R.U	$2.65 * 10^{-9}$	$3.96 * 10^{-9}$	/
Hum			
trend	0.55	/	/
const	$3.41 * 10^{-12}$	$2.99 * 10^{-12}$	/
R.U	$3.04 * 10^{-12}$	$2.89 * 10^{-12}$	/

TABLE 3.3 – résultats de test ADF .

l'application de la stratégie du Dicky-Fuller, sur les trois séries nos mène à conclure que nos séries sont stationnaires parce qu'elles ne sont pas affectées ni de tendance ni de racine unitaire.

### 3.2.2 Box-Jenkins

#### Identification

cette étape est effectuée par le biais de l'examination des fonctions ACF et PACF des séries désaisonnalisées stationnaires.

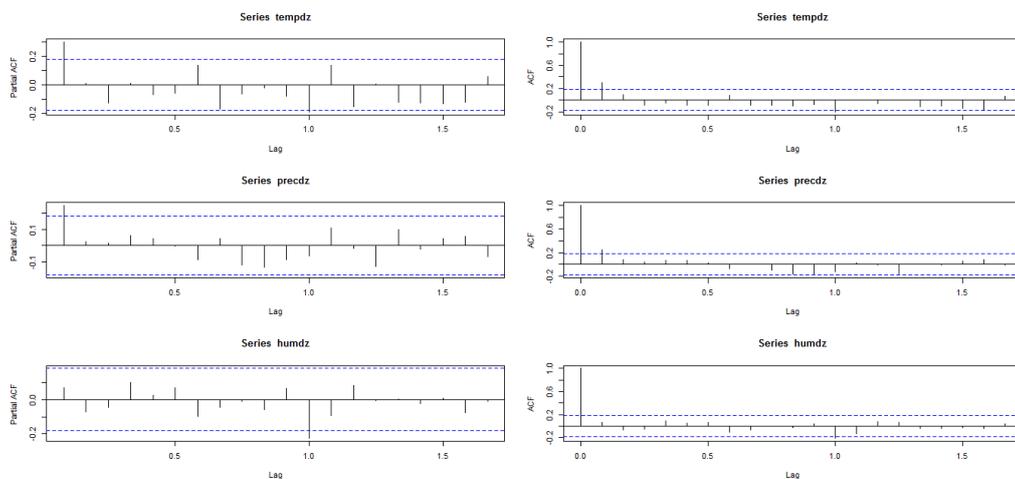


FIGURE 3.3 – ACF et PACF des séries stationnaires

pour la série des températures et la série précipitations, les deux fonctions ACF et PACF s'annulent à partir de deuxième pics

pour la série Humidité, les deux fonctions ACF et PACF s'annulent à partir de pic numéros 13. d'après l'analyse ci dessus, on peut identifier les processus dont les ordres mentionnés dans le tableau (3.4) pour chaque séries.

prec	temp	hum
AR(1)	AR(1)	AR(12)
MA(1)	MA(1)	MA(12)
ARMA(1,1)	ARMA(1,1)	ARMA(12,12)

TABLE 3.4 – Modèles candidats

### Estimation et Validation

après avoir identifier l'ordre des processus, il convient d'estimer les paramètres du modèle et de vérifier par la suite la validité de différents modèles estimés en se basant sur un ensemble de tests sur paramètres et sur résidus.

#### série prec

	AR(1)	MA(1)	ARMA(1,1)
cst	83.144 (0.000)	83.131(0.000)	83.174(0.000)
$\phi$	0.245(0.00638)		0.352(0.433)
$\theta$		0.229(0.00791)	-1.115(0.813)
AIC	1308.346	1308.931	1310.291
BIC	1313.921	1314.506	1318.654
Box-Piere	0.0022968( 0.9618)	0.020919(0.885)	0.00097522(0.9751)
J-Berra	1.5973( 0.4499)	2.0785( 0.3537)	83.517 (2.2 * 10 <sup>-16</sup> )

TABLE 3.5 – Estimation et validation des processus estimés.() représente p-value

#### série temp

	AR(1)	MA(1)	ARMA(1,1)
cst	18.161(0.000)	18.158(0.000)	18.1615(0.000)
$\phi$	0.312(0.000644)		0.3225(0.139)
$\theta$		0.0.263(0.000966)	-0.0117(0.958)
AIC	358.7815	360.5169	360.7788
BIC	364.3565	366.0918	369.1412
Box-Piere	0.031331(0.8595)	0.090848( 0.7631)	0.02558 (0.8729)
J-Berra	1.5973( 0.4499)	2.0785( 0.3537)	1.573 (0.4554)

TABLE 3.6 – Estimation et validation des processus estimés.() représente p-value

série hum

	AR(12)	MA(12)	ARMA(12,12)
cst	75.24805(0.000)	75.25352(0.000)	
$\phi$	-0.24835(0.00756)		
$\theta$		-0.28910(0.0235))	
AIC	606.7539	607.4742	
BIC	642.9913	643.7116	
Box-Piere	0.037895(0.8457)	0.075495(0.7835)	
J-Berra	1.5633( 0.4577)	1.2123(0.5454)	

TABLE 3.7 – Estimation et validation des processus estimés.() représente p-value

Au regard des résultats d'estimation, on peut valider :

les processus AR(1) et MA(1) pour les séries Prec et Temp dans la mesure ou les coefficients sont significatives et leurs résidus vérifient les hypothèses de l'homoscédasticité et de la normalité .

les processus AR(12) et MA(12) pour la série Hum dans la mesure ou les coefficients sont significatives et ses résidus vérifient les hypothèses de l'homoscédasticité et de la normalité.

d'après les valeurs de critères AIC et BIC, on peut choisir le processus AR(1) comme meilleur modèle (qui minimise les deux critères) pour Temp et Prec et le processus AR(12) pour Hum.

**Conclusion** : l'application de la méthodologie de Box et Jenkins nous a conduit à retenir un processus AR(1) pour la Température et Précipitation et un processus AR(12) pour l'humidité.

**prévision**

une fois les modèles adéquats sont choisis, on peut baser sur eux pour calculer les prévisions de 12 mois de l'année 2006 pour chaque série.

**Prec** :  $\hat{X}_t = 83.144 + 0.245X_{t-1}$

**Temp** :  $\hat{X}_t = 18.161 + 0.312X_{t-1}$

**hum** :  $\hat{X}_t = 75.24805 - 0.24835X_{t-12}$

	prec	temp	hum
01/2006	188.92023	15.3445173	57.5523287
02/2006	145.851077	18.1776255	59.0616904
03/2006	77.1374643	21.655029	56.4612811
04/2006	109.461559	23.625107	56.8065013
05/2006	80.711044	26.975341	59.0224961
06/2006	42.7137724	31.2167208	55.8879195
07/2006	31.0778962	33.5276734	54.4713653
08/2006	43.6650163	34.5353746	54.1778136
09/2006	93.5639455	31.7710868	56.9547086
10/2006	102.396924	28.7081473	54.4805723
11/2006	210.54005	23.7350131	58.2374173
12/2006	215.403734	20.8705181	57.4899151

TABLE 3.8 – Prévisions des trois séries par la méthode Box et Jenkins

### 3.3 Etude multivarié

Dans la section précédente nous avons modélisé séparément les trois séries temporelle. par contre dans cette section, le modèle VAR nous donne la possibilité de modéliser les trois séries simultanément afin de prendre en considération les interactions entre les variables.

### 3.3.1 Choix de retard p

	AIC	BIC	HQ
1	12.117	12.402	12.233
2	11.429	11.928	11.632
3	11.030	11.743	11.320
4	11.039	11.965	11.415

TABLE 3.9 – Critères de choix de retard optimal

on choisit le retard qui fournit le minimum des trois critères d'information (AIC,BIC,HQ), la valeur minimale correspond à un retard  $p=3$ .

### 3.3.2 Estimation de modèle VAR

l'estimation de modèle VAR(3) par la méthode MV(maximum de vraisemblance), nous permet d'écrire notre modèle sous les équations suivantes :

$$T = (-9.2576) + 1.0695T_{t-1} + 0.1736T_{t-2} - 0.6658T_{t-3} + 0.1718H_{t-1} + 0.0425H_{t-2} + 0.01422H_{t-3} - 0.00061P_{t-1} - 0.00528P_{t-2} - 0.00796P_{t-3}$$

$$H = 103.463 - 0.0774H_{t-1} - 0.172H_{t-2} - 0.049H_{t-3} - 0.274T_{t-1} - 0.349T_{t-2} + 0.3097T_{t-3} + 0.0032P_{t-1} - 0.0053P_{t-2} + 0.0026P_{t-3}$$

$$P = 447.688 + 0.304P_{t-1} - 0.0329P_{t-2} - 0.01415P_{t-3} - 9.0287T_{t-1} + 0.6788T_{t-2} + 8.568T_{t-3} - 6.624H_{t-1} + 0.0339H_{t-2} + 1.3985H_{t-3}$$

### 3.3.3 Étude de stationnaire du VAR

Il faut s'assurer que nous sommes en présence d'une VAR(3) stationnaire, ie, que tous les racines sont supérieurs à 1, pour cela en trace le cercle des inverses des racines du polynôme caractéristique.

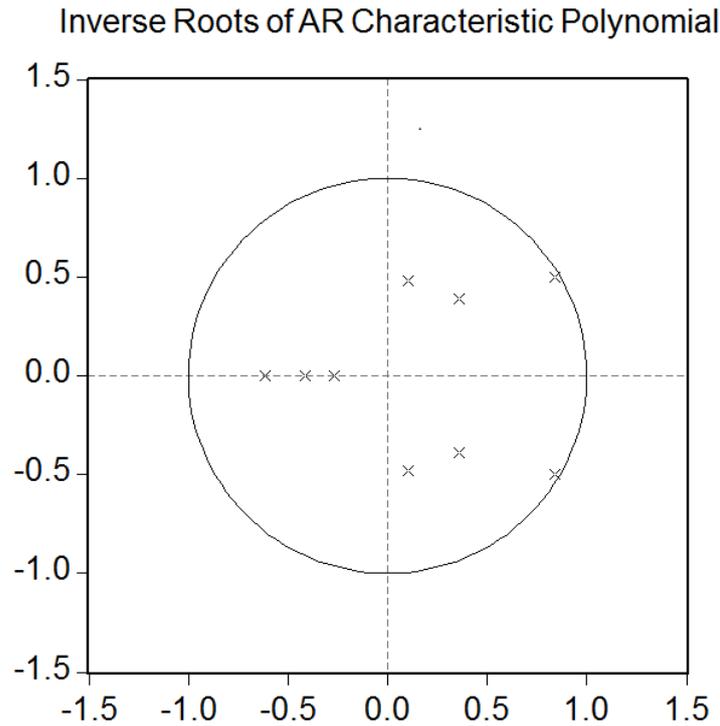


FIGURE 3.4 – Cercle des inverses des racines

D'après le cercle ,le modèle VAR(3) est stationnaire car tous les racines se situent à l'intérieure de cercle d'unité.

### 3.3.4 Causalité

l'hypothèse nul		obs	f.stat	P.value
T ne cause pas H-P	Granger	120	8.50	$1.401 * 10^{-8}$
	instant		23.51	$7.841 * 10^{-6}$
P ne cause pas T-H	Granger	120	2.18	0.04411
	instant		16.76	0.00022
H ne cause pas T-P	Granger	120	3.14	0.0052
	instant		24.24	$5.435 * 10^{-6}$

TABLE 3.10 – Résultats du test de causalité

D'après le tableau ,on peut conclue :qu'il existe des relations de causalité bidirectionnel entre les trois séries deux à deux au sens de Granger et instantanément

### 3.3.5 Fonction de réponse impulsionnelle et analyse de choc

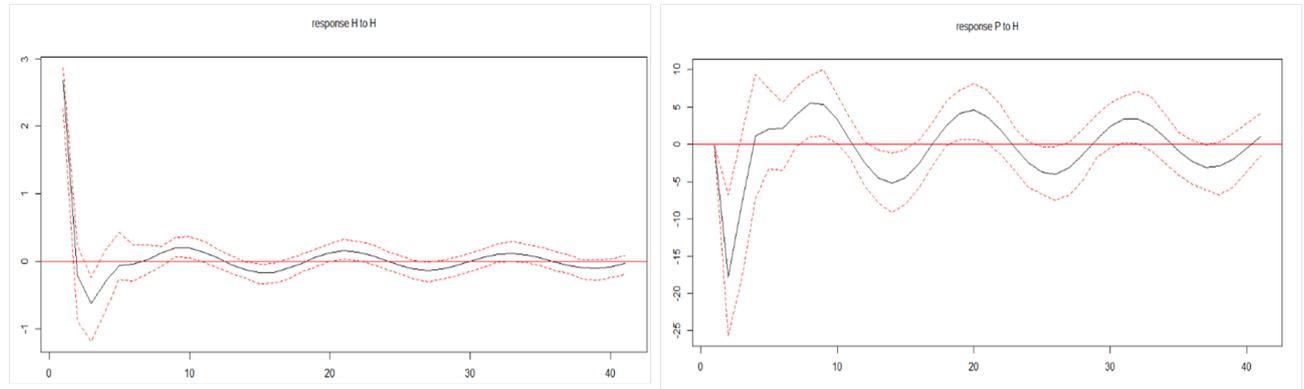


FIGURE 3.5 – Reponse de prec à un choc dans hum

Entre le premier et le deuxième mois après le choc on remarque une diminution de précipitations suite a la diminution de l'humidité; a partir du deuxième mois on remarque une évolution similaire des 2 graphes. Cela s'explique par l'influence de taux de l'humidité sur la quantité des pluies

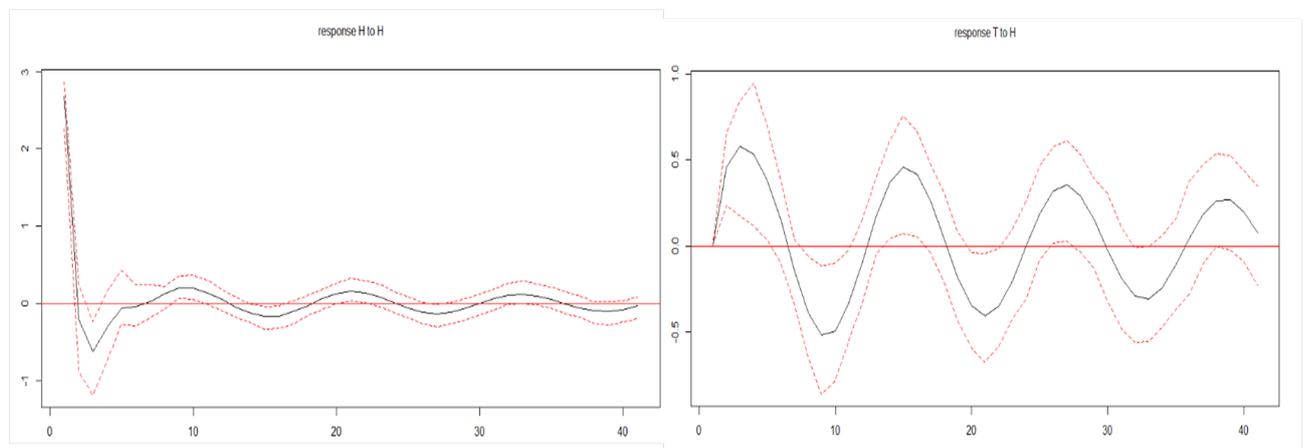


FIGURE 3.6 – Réponse de Temp à un choc dans hum

Entre le premier et le troisième mois après le choc on remarque une augmentation de températures suite à la diminution de l'humidité; a partir de la troisième mois on remarque une diminution de la température ou moment ou le H augmente. En générale on constate une évolution inverse des 2 graphes jusqu'à la disparition totale de ce choc .

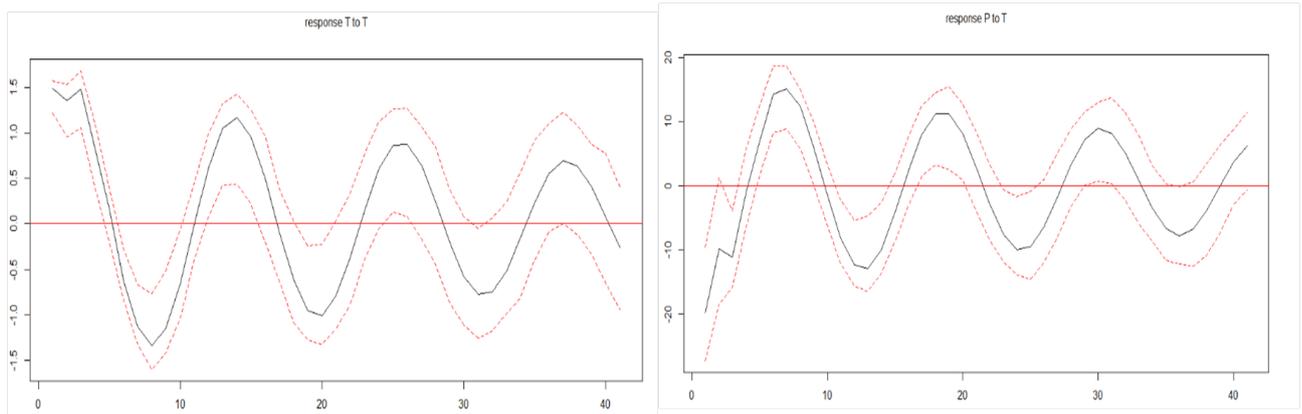


FIGURE 3.7 – Reponse de Prec à un choc dans Temp

Entre le premier et le sixième mois après le choc on remarque une augmentation de PREC suite à la diminution de TEMP ; puis à partir du sixième mois au deuxième mois on remarque une diminution du PREC suite à une augmentation de TEMP. En générale on constate une évolution inverse des 2 graphes alternant entre l'augmentation et la diminution chaque six mois jusqu'à la disparition totale de ce choc .

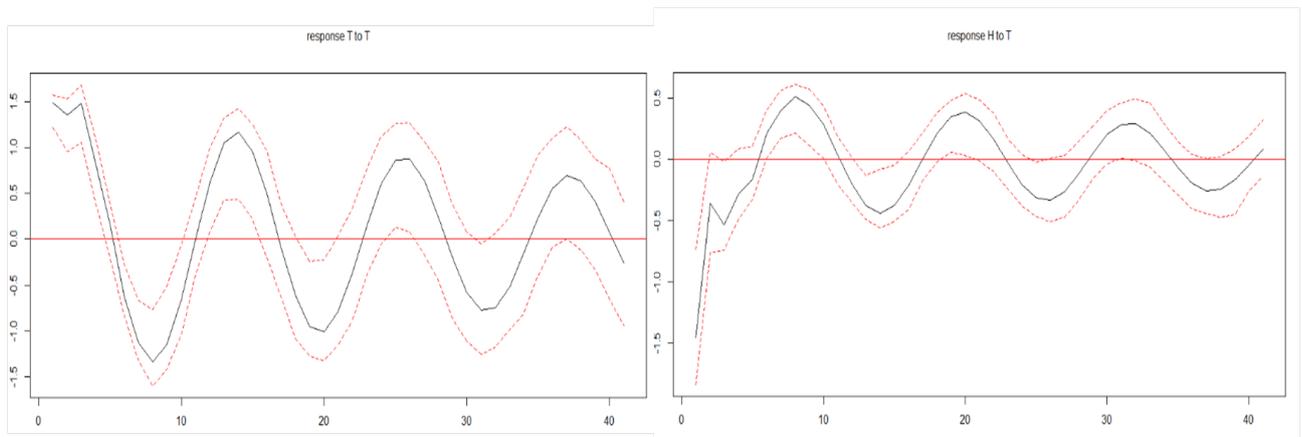


FIGURE 3.8 – Réponse de Hum à un choc dans Temp

Entre le premier et le sixième mois après le choc on remarque une augmentation de PREC suite à la diminution de TEMP ; puis à partir du sixième mois au deuxième mois on remarque une diminution du PREC suite à une augmentation de TEMP. En générale on constate une évolution inverse des 2 graphes alternant entre l'augmentation et la diminution chaque six mois jusqu'à la disparition totale de ce choc .

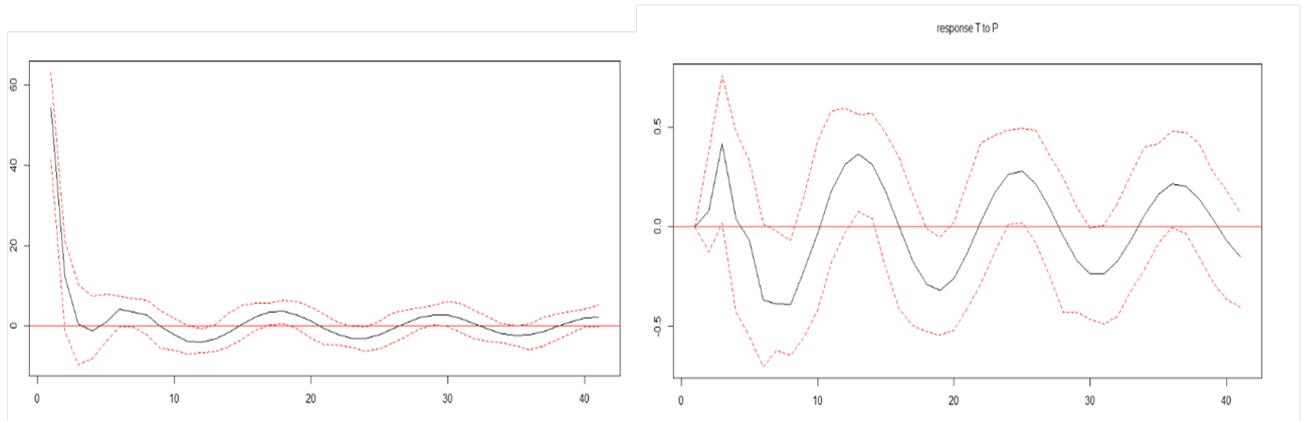


FIGURE 3.9 – Réponse de Temp à un choc dans Prec

Entre le premier et le troisième mois après le choc on remarque une augmentation de TEMP suite à la diminution de PREC ; puis à partir de la troisième mois au sixième mois on remarque une diminution du TEMP suite à une augmentation de PREC. En générale on constate une évolution inverse des 2 graphes alternant entre l'augmentation et la diminution chaque 03 mois jusqu'à la disparition totale de ce choc .

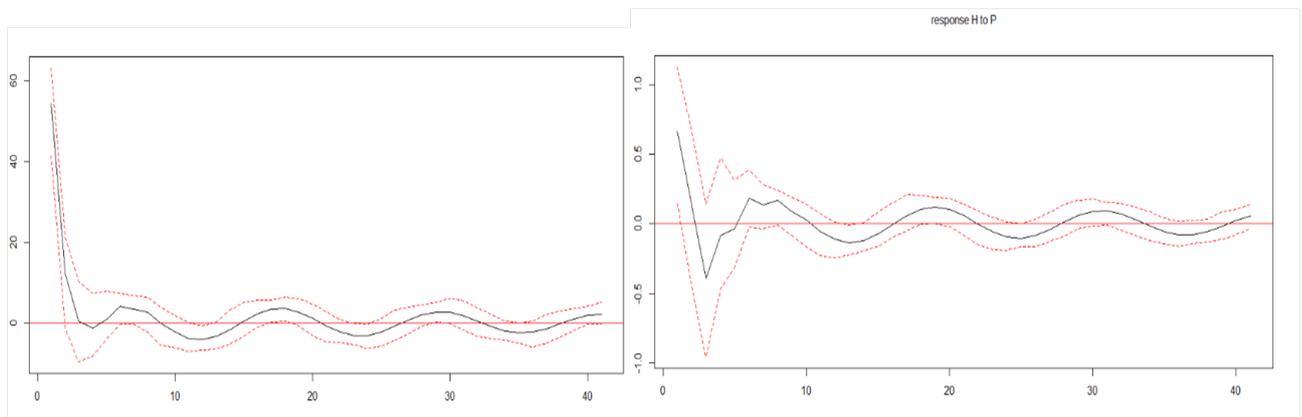


FIGURE 3.10 – Réponse de Hum à un choc dans Prec

Entre le premier et le troisième mois après le choc on remarque une diminution de HUM suite à la diminution de PREC ; puis à partir de la troisième mois au sixième mois on remarque une augmentation de la HUM suite à une augmentation de PREC. En générale on constate une évolution similaire des 2 graphes alternant entre l'augmentation et la diminution chaque 03 mois jusqu'à la disparition totale de ce choc .

### 3.3.6 Décomposition de Variance

L'intérêt est de savoir quelle est contribution de chaque innovation a la variance totale de l'erreur de prévision .On présente les tableaux suivants qui donnent le pourcentage de contribution des résidus de chaque variable sur la variance de l'erreur de prévision de la variable considérée dont on peut tirer des conclusion sur la variable qui influence le plus au autres variables.

T	H	P
1.0000	0.0000	0.0000
0.9484358	0.05131760	0.0002466035
0.8948795	0.09495635	0.0101641171
0.8711016	0.11885561	0.0100428007
0.8552797	0.13188230	0.0128380047
0.8463566	0.12384127	0.0298021406
0.8512787	0.11152153	0.0371997669
0.8519694	0.11051644	0.0375141331
0.8463067	0.12000975	0.0336835647
0.8365733	0.13083397	0.0325927484
0.8284242	0.13502452	0.0365513017
0.8270006	0.13068663	0.0423128161

TABLE 3.11 – D.V de l'erreur de prévision de Temp

La variance de l'erreur de prévision de Temp est due pour 87%de T, de 10 % de H et de 2 %de P,donc l'impacte d'un choc affectant Temp sur Hum est plus important que l'impact d'un choc affectant TE sur Prec

T	H	P
0.2164448	0.7835552	0.00000
0.2253054	0.7717532	0.0002941457
0.2345184	0.7577082	0.007773406
0.2378942	0.7544511	0.007654716
0.2395469	0.7527911	0.007661974
0.2418800	0.7473774	0.010742604
0.2520922	0.7358735	0.012034299
0.2682078	0.7185811	0.013211100
0.2794221	0.7075320	0.013045913
0.2833756	0.7036824	0.012942001
0.2829221	0.7034777	0.013600192
0.2849602	0.7003392	0.014700636

TABLE 3.12 – D.V de l’erreur de prévision de Hum

variance de l’erreur de prévision de Hum est due de 74%de Hum, de 25% de Temp , et de 11% de Prec, donc l’impacte d’un choc affectant HUM sur Temp est plus important que l’impact d’un choc affectant Hum sur Prec

T	H	P
0.1165686	0.0506769	0.8327545
0.1244349	0.0960065	0.7795586
0.1492021	0.1062941	0.7445039
0.1493233	0.1063433	0.7443334
0.1585095	0.1061925	0.7352980
0.1966961	0.1029075	0.7003965
0.2350113	0.1019029	0.6630858
0.2574505	0.1053781	0.6371714
0.2612495	0.1095086	0.6292419
0.2607218	0.1106314	0.6286468
0.2700114	0.1088484	0.6211403
0.2904203	0.1073444	0.6022353

TABLE 3.13 – D.V de l’erreur de prévision de Prec

la variance de l'erreur de prévision de Prec est due de 69%de Prec, de 21%de Temp , et de 10% de Hum, donc l'impacte d'un choc affectant Prec sur Temp est plus important que l'impact d'un choc affectant Prec sur Hum

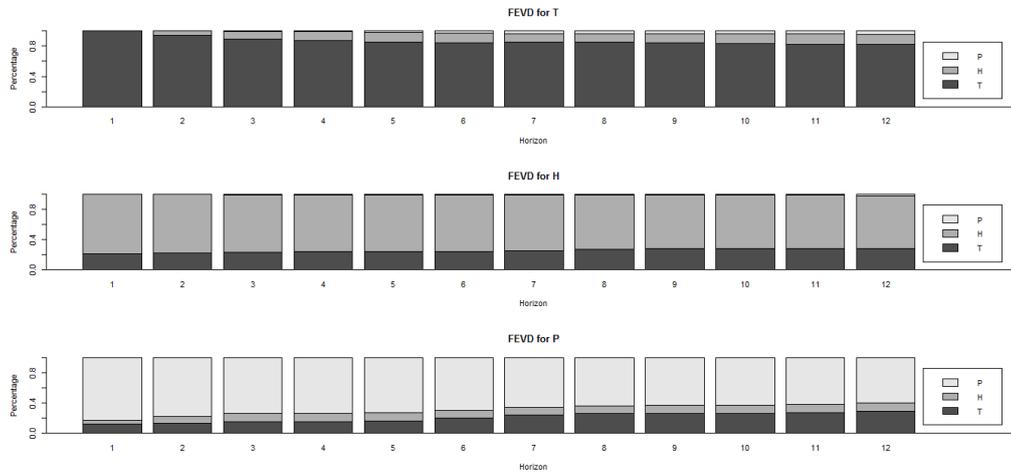


FIGURE 3.11 – D.V de l'erreur de prévision des Séries

### 3.3.7 Validation

#### test d'autocorrelation

L'application de test Ljung-Box sur la série résiduel du modèle VAR(3) nous a conduit à constater que les erreurs sont autocorrélées ( $p.value = 0.0107$ ).

#### test d'hétéroscédasticité

L'application de test ARCH sur la série résiduel du modèle VAR(3) ,nous a conduit à constater que les erreurs sont homoscedastique(variance homogène)( $p.value = 0.71$ ).

#### test normalité

L'application de test Jarque-Bera sur la série résiduel du modèle VAR(3) ,nous a conduit à constater que les erreurs ne sont pas gaussiens(normales) ( $p = 2.2 * 10^{-16}$ )

## 3.3.8 Prévision

	réel			Var			residual		
	Temp	Hum	Prec	Temp	Hum	Prec	Temp	Hum	prec
01/2006	10.555		162.1028	9.958	78.705	180.323	0.5911		-18.2205
02/2006	11.66		171.8056	10.548	78.534	132.009	1.1116		39.7957
03/2006	14.44		54.0004	13.332	77.236	82.20	1.1079		-28.0196
04/2006	17.77		24.1046	17.140	75.851	39.004	0.6296		-14.8999
05/2006	20.55		32.5882	21.155	74.353	9.674	-0.6057		22.9136
06/2006	23.33		2.9972	24.111	73.202	1.105	-0.7819		1.892
07/2006	26.11		0.000	25.346	72.597	14.367	0.7633		-14.3678
08/2006	25.55		34.6964	24.512	72.798	46.243	1.0371		-11.5467
09/2006	23.33		44.8056	21.978	73.666	86.372	1.3516		-41.5666
10/2006	22.22		38.3032	18.488	74.959	123.670	3.732		-85.3676
11/2006	17.77		39.1922	15.067	76.277	147.662	2.7029		-108.47
12/2006	13.33		205.5114	12.646	77.268	152.508	0.6835		53.0034

TABLE 3.14 – comparaison des données réelles et des résultats prévisionnels de (temp),(Hum)et(prec)

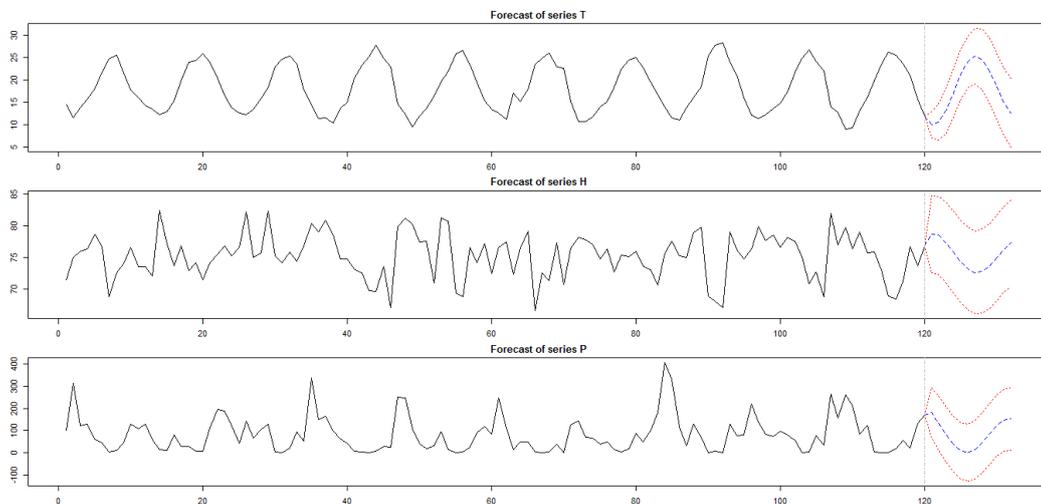


FIGURE 3.12 – graphe de prévision De modèle VAR(3)

### 3.4 comparaison entre le prévision de Box-Jenkins et modèle VAR

Prec					
	prvun	prvml	reel	res1	res2
01/2006	188.9202	180.3233	162.1028	-26.8174	-18.2205
02/2006	145.851	132.0099	171.8056	25.9546	39.7957
03/2006	77.1374	82.02	54.0004	-23.137	-28.0196
04/2006	109.4615	39.0045	24.1046	-85.3569	-14.8999
05/2006	80.711	9.6746	32.5882	-48.1228	22.9136
06/2006	42.7137	1.1052	2.9972	-39.7165	1.892
07/2006	31.0077	14.3678	0	-31.0077	-14.3678
08/2006	43.665	46.2431	34.6964	-8.9686	-11.5467
09/2006	93.5639	86.3722	44.8056	-48.7583	-41.5666
10/2006	102.3969	123.6708	38.3032	-64.0937	-85.3676
11/2006	210.54	147.6622	39.1922	-171.3478	-108.47
12/2006	215.4037	152.508	205.5114	-9.8923	53.0034

TABLE 3.15 – Série Résiduels de Prec

Temp					
	prvun	prvml	reel	res1	res2
01/2006	15.3445	9.9589	10.55	-4.7945	0.5911
02/2006	18.1776	10.5484	11.66	-6.5176	1.1116
03/2006	21.655	13.3321	14.44	-7.215	1.1079
04/2006	23.6251	17.1404	17.77	-5.8551	0.6296
05/2006	26.9753	21.1557	20.55	-6.4253	-0.657
06/2006	31.2167	24.1119	23.3	-7.8867	-0.7819
07/2006	33.5276	25.3467	26.11	-7.4176	0.7633
08/2006	34.5353	24.5129	25.55	-8.9853	1.0371
09/2006	31.771	21.9784	23.33	-8.441	1.3516
10/2006	28.7081	18.488	22.22	-6.4881	3.732
11/2006	23.735	15.0671	17.77	-5.965	2.7029
12/2006	20.8705	12.6465	13.33	-7.5405	0.6835

TABLE 3.16 – Série Résiduels de Temp

variable	RMSE(BJ)	RMSE(VAR)	MAE(BJ)	MAE(VAR)
prec	64.60	47.7682	148.59	36.67
temp	7.052	1.5645	6.96	1.25

TABLE 3.17 – RMSE et MAE de prévision des séries

En générale, les deux critères sont faibles pour la prévision par le modèle VAR que pour la prévision Box-Jenkins. Ce qui nous permet de conclure que la modélisation VAR est plus efficace en terme de la précision de prévision.

---

# CONCLUSION

Ce mémoire est juste une petite introduction aux modèles VAR et leur application sur des séries météorologiques, nous avons donné quelques outils de base de la théorie des séries temporelles multivariées permettant de prendre en compte la dépendance temporelle entre les séries et l'analyse de leur comportement dynamique via la fonction de réponse impulsionnelle (impact de choc) et la décomposition de la variance de l'erreur de prévision.

Notre étude empirique est partagée en deux parties :

–La première partie réservée à la modélisation univariée de trois séries météorologiques (température, précipitations, et humidité) . après avoir désaisonnaliser les trois série et confirmer leur stationnarité par le test de racine unitaire; nous avons suit la méthodologie de Box et Jenkins qui nous a conduit à obtenir un modèle AR(1) pour la série de température et la série de précipitations et un processus AR(12) pour la série de l'humidité.

–La seconde partie consacrée à la modélisation des trois séries météorologiques simultanément en appliquant la modélisation VAR. après avoir choisir le retard optimal ( $P=3$ ) et étudier la causalité , le modèle VAR(3) est estimé par la méthode de Maximum de vraisemblance et validé par les tests de résidus.

L'analyse de choc et la décomposition de la variance de l'erreur de prévision ont été discutées à la fin de cette partie.

L'étude de ces trois séries météorologiques était orientée essentiellement dans une optique prévisionnelle, les processus résultants de deux approche sont utilisées par la suite au calcul de prévisions pour une période de 12 mois du janvier 2006 à décembre 2006. En terme de RMSE et

MAE, le modèle VAR(3) avait donné des prévisions mieux que celles fournies par les processus AR(1) et AR(12) pour les trois séries météorologiques.

Nous espérons avoir répondu à la problématique posée et les résultats trouvés seront d'une utilité pertinente.

---

## ANNEXES

```
-----choi de p-----
> p<-VARselect(varm,lag.max=4,type="const")
> p
$selection
AIC(n)  HQ(n)  SC(n) FPE(n) |
      3      3      3      3

$criteria
              1          2          3          4
AIC(n)      12.11787    11.42993    11.03093    11.03945
HQ(n)       12.23350    11.63229    11.32002    11.41526
SC(n)       12.40272    11.92842    11.74307    11.96523
FPE(n)  183129.40223  92075.77655  61834.55410  62460.25885
```

Estimation results for equation T:

=====

T = T.11 + H.11 + P.11 + T.12 + H.12 + P.12 + T.13 + H.13 + P.13 + const

	Estimate	Std. Error	t value	Pr(> t )	
T.11	1.0695277	0.0874270	12.233	< 2e-16	***
H.11	0.1718098	0.0527298	3.258	0.00150	**
P.11	-0.0006129	0.0026932	-0.228	0.82042	
T.12	0.1736304	0.1364504	1.272	0.20596	
H.12	0.0425068	0.0553775	0.768	0.44443	
P.12	0.0052822	0.0026674	1.980	0.05024	.
T.13	-0.6658719	0.0882034	-7.549	1.54e-11	***
H.13	0.0142226	0.0540059	0.263	0.79279	
P.13	-0.0079627	0.0024661	-3.229	0.00165	**
const	-9.2576163	7.3494050	-1.260	0.21054	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.487 on 107 degrees of freedom

Multiple R-Squared: 0.9288, Adjusted R-squared: 0.9228

F-statistic: 155.1 on 9 and 107 DF, p-value: < 2.2e-16

Estimation results for equation H:

=====

H = T.11 + H.11 + P.11 + T.12 + H.12 + P.12 + T.13 + H.13 + P.13 + const

	Estimate	Std. Error	t value	Pr(> t )
T.11	-0.274187	0.183852	-1.491	0.1388
H.11	-0.077426	0.110886	-0.698	0.4865
P.11	0.003238	0.005664	0.572	0.5688
T.12	-0.349314	0.286944	-1.217	0.2261
H.12	-0.172102	0.116454	-1.478	0.1424
P.12	-0.005327	0.005609	-0.950	0.3444
T.13	0.309770	0.185485	1.670	0.0978 .
H.13	-0.049679	0.113570	-0.437	0.6627
P.13	0.002698	0.005186	0.520	0.6039
const	103.463839	15.455194	6.694	1.03e-09 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.128 on 107 degrees of freedom

Multiple R-Squared: 0.3153, Adjusted R-squared: 0.2577

F-statistic: 5.474 on 9 and 107 DF, p-value: 3.58e-06

```

|
Estimation results for equation P:
=====
P = T.11 + H.11 + P.11 + T.12 + H.12 + P.12 + T.13 + H.13 + P.13 + const

      Estimate Std. Error t value Pr(>|t|)
T.11   -9.02874    3.40731  -2.650  0.00927 **
H.11   -6.62431    2.05505  -3.223  0.00168 **
P.11    0.30472    0.10496   2.903  0.00449 **
T.12    0.67885    5.31791   0.128  0.89866
H.12    0.03391    2.15824   0.016  0.98749
P.12   -0.03291    0.10396  -0.317  0.75220
T.13    8.56817    3.43757   2.493  0.01422 *
H.13    1.39858    2.10479   0.664  0.50781
P.13   -0.01415    0.09611  -0.147  0.88323
const  447.68803   286.43006   1.563  0.12101
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 57.96 on 107 degrees of freedom
Multiple R-Squared: 0.5275,    Adjusted R-squared: 0.4877
F-statistic: 13.27 on 9 and 107 DF,  p-value: 4.794e-14

```

### Portmanteau Test (asymptotic)

```

|
data:  Residuals of VAR object var.model
Chi-squared = 113.09, df = 81, p-value = 0.01074

```

```
> bv.arch

      ARCH (multivariate)

data:  Residuals of VAR object var.model
Chi-squared = 414.57, df = 432, p-value = 0.7185

> bv.norm <- normality.test(var.model, multivariate.only = TRUE)#normalité
> bv.norm
$JB

      JB-Test (multivariate)

data:  Residuals of VAR object var.model
Chi-squared = 93.418, df = 6, p-value < 2.2e-16

$Skewness

      Skewness only (multivariate)

data:  Residuals of VAR object var.model
Chi-squared = 25.551, df = 3, p-value = 1.184e-05

$Kurtosis

      Kurtosis only (multivariate)

data:  Residuals of VAR object var.model
Chi-squared = 67.867, df = 3, p-value = 1.221e-14
```

`$Granger`

Granger causality H0: T do not Granger-cause H P

data: VAR object var.model

F-Test = 8.5012, df1 = 6, df2 = 321, p-value = 1.401e-08

`$Instant`

H0: No instantaneous causality between: T and H P

data: VAR object var.model

Chi-squared = 23.512, df = 2, p-value = 7.841e-06

`$Granger`

Granger causality H0: P do not Granger-cause T H

data: VAR object var.model

F-Test = 2.1857, df1 = 6, df2 = 321, p-value = 0.04411

`$Instant`

H0: No instantaneous causality between: P and T H

data: VAR object var.model

Chi-squared = 16.764, df = 2, p-value = 0.0002289

`$Granger`

Granger causality H0: H do not Granger-cause T P

data: VAR object var.model

F-Test = 3.1411, df1 = 6, df2 = 321, p-value = 0.005248

`$Instant`

H0: No instantaneous causality between: H and T P

data: VAR object var.model

Chi-squared = 24.245, df = 2, p-value = 5.435e-06

ARIMA(1,0,0) model is estimated for variable: precdz

Conditional-Sum-of-Squares & Maximum Likelihood Estimation

	Estimate	S.E	t.value	p.value	Lag
MU	83.144	6.6339	12.53	0.00000	1
AR 1	0.245	0.0883	2.78	0.00638	1

-----

n = 120; 'sigma' = 54.99504; AIC = 1308.346; SBC = 1313.921

Box-Pierce test

data: modar\$residuals

X-squared = 0.0022968, df = 1, p-value = 0.9618

Jarque Bera Test

data: modar\$residuals

X-squared = 1.5973, df = 2, p-value = 0.4499

ARIMA(1,0,0) model is estimated for variable: tempdz

Conditional-Sum-of-Squares & Maximum Likelihood Estimation

	Estimate	S.E	t.value	p.value	Lag
MU	18.161	0.139	130.61	0.000000	1
AR 1	0.312	0.089	3.51	0.000644	1

-----

n = 120; 'sigma' = 1.051858; AIC = 358.7815; SBC = 364.3565

Box-Pierce test

data: modar\$residuals

X-squared = 0.031331, df = 1, p-value = 0.8595

Jarque Bera Test

data: modtar\$residuals

X-squared = 1.5973, df = 2, p-value = 0.4499

ARIMA(12,0,0) model is estimated for variable: humdz

Conditional-Sum-of-Squares & Maximum Likelihood Estimation

	Estimate	S.E	t.value	p.value	Lag
MU	75.24805	0.2041	368.600	0.00000	1
AR 1	0.08902	0.0894	0.996	0.32135	1
AR 2	-0.08791	0.0896	-0.982	0.32848	2
AR 3	-0.04406	0.0896	-0.492	0.62400	3
AR 4	0.11384	0.0903	1.260	0.21037	4
AR 5	-0.01890	0.0913	-0.207	0.83632	5
AR 6	0.10090	0.0921	1.096	0.27567	6
AR 7	-0.10062	0.0907	-1.110	0.26952	7
AR 8	-0.02457	0.0910	-0.270	0.78762	8
AR 9	-0.00975	0.0904	-0.108	0.91430	9
AR 10	-0.09573	0.0910	-1.052	0.29506	10
AR 11	0.09122	0.0912	1.000	0.31960	11
AR 12	-0.24835	0.0912	-2.723	0.00756	12

-----

n = 120; 'sigma' = 2.68603; AIC = 606.7539; SBC = 642.9913

```
> Box.test(modhar$residuals,type="Box-Pierce")
```

Box-Pierce test

data: modhar\$residuals

X-squared = 0.037895, df = 1, p-value = 0.8457

```
> jarque.bera.test(modhar$residuals)
```

Jarque Bera Test

data: modhar\$residuals

X-squared = 1.5633, df = 2, p-value = 0.4577

---

# BIBLIOGRAPHIE

- [1] **Arthur Charpentier** : Cours de séries temporelle théorie et applications, Volume1, Paris Dauphine.
- [2] **Arthur Charpentier** : Cours de séries temporelle théorie et applications, Volume2, Paris Dauphine.
- [3] **Amenan Christiane Chukunyere** :Les modèles VAR(p),Maîtrise en statistique - avec mémoire Maître ès sciences (M. Sc.),Québec, Canada,© Amenan Christiane Chukunyere, 2019.
- [4] **Brockwell P.J** and **David R.A** :Introduction to time series and forecasting.Springer,2002
- [5] **Chaouchkouane Meriem** :Méthode de Box et Jenkins,MASTER en Mathématiques,DÉPARTEMENT DE MATHÉMATIQUES,FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de laVIE,UNIVERSITÉ MOHAMED KHIDER,BISKRA, Juin2021
- [6] **Kesraoui Mahrez** et **6** :Etude et modélisation d'une évolution d'une pathologie,Mémoire de fin d'étude,En vue de l'obtention du Diplôme de Master Professionnel en Mathématiques Appliquées à la Gestion,Département de Mathématiques,Faculté des Sciences, Promotion : 2016/2017
- [7] **Michel Lubrano** Modèles VAR, modèles VAR structurels,March 17, 2008
- [8] **Sandrine Lardic** and **Valérie Mignon** :Econométrie des séries temporelles macroéconomique et financières,ed.economica,49,rue Héricat,75015 Paris,2002.
- [9] **Yves Aragon** : Introduction aux séries temporelles,Septembre 2004 .