

Séries temporelles – Modèles ARIMA.
Didier Delignières
Séminaire EA "Sport – Performance – Santé"
Mars 2000

Il existe deux catégories de modèles pour rendre compte d'une série temporelle. Les premiers considèrent que les données sont une fonction du temps ($y = f(t)$). Cette catégorie de modèle peut être ajustée par la méthode des moindres carrés, ou d'autres méthodes itératives. L'analyse des modèles par transformée de Fourier est une version sophistiquée de ce type de modèle.

Une seconde catégorie de modèles cherche à déterminer chaque valeur de la série en fonction des valeurs qui la précède ($y_t = f(y_{t-1}, y_{t-2}, \dots)$). C'est le cas des modèles ARIMA ("Auto-Regressive – Integrated – Moving Average"). Cette catégorie de modèles a été popularisée et formalisée par Box et Jenkins (1976).

A noter que le choix de l'un ou l'autre type de modèle est surtout théorique: est-il raisonnable de penser que dans un phénomène donné, les points sont fondamentalement fonction des points précédents et de leurs erreurs, plutôt qu'un signal, périodique ou non, entaché de bruit. On peut noter cependant que souvent, on a recours à l'analyse de variance pour traiter les séries temporelles. Or une des assumptions majeures de l'ANOVA est que les résidus des différentes mesures ne sont pas auto-corrélés. Ce n'est évidemment pas le cas si la performance à l'essai t est liée à la performance réalisée à l'essai $t-1$.

Les processus autorégressifs supposent que chaque point peut être prédit par la somme pondérée d'un ensemble de points précédents, plus un terme aléatoire d'erreur.

Le processus d'intégration suppose que chaque point présente une différence constante avec le point précédent.

Les processus de moyenne mobile supposent que chaque point est fonction des erreurs entachant les points précédant, plus sa propre erreur.

Un modèle ARIMA est étiqueté comme modèle ARIMA (p,d,q), dans lequel:

p est le nombre de termes auto-régressifs

d est le nombre de différences

q est le nombre de moyennes mobiles.

1. Différenciation.

L'estimation des modèles ARIMA suppose que l'on travaille sur une série stationnaire. Ceci signifie que la moyenne de la série est constante dans le temps, ainsi que la variance. La meilleure méthode pour éliminer toute tendance est de différencier, c'est-à-dire de remplacer la série originale par la série des différences adjacentes. Une série temporelle qui a besoin d'être différenciée pour atteindre la stationnarité est considérée comme une version intégrée d'une série stationnaire (d'où le terme *Integrated*).

La correction d'une non-stationnarité en termes de variance peut être réalisée par des transformation de type logarithmique (si la variance croît avec le temps) ou à l'inverse exponentielle. Ces transformations doivent être réalisées avant la différenciation.

Une différenciation d'ordre 1 suppose que la différence entre deux valeurs successives de y est constante.

$$y_t - y_{t-1} = \mu + \varepsilon_t$$

μ est la constante du modèle, et représente la différence moyenne en y . Un tel modèle est un ARIMA(0,1,0). Il peut être représenté comme un accroissement linéaire en fonction du temps. Si μ est égal à 0, la série est stationnaire.

Les modèles d'ordre 2 travaillent non plus sur les différences brutes, mais sur les différences de différence. La seconde différence de y au moment t est égale à $(y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$, c'est-à-dire à $y_t - 2y_{t-1} + y_{t-2}$.

Un modèle ARIMA(0,2,0) obéira à l'équation de prédiction suivante :

$$y_t - 2y_{t-1} + y_{t-2} = \mu + \varepsilon_t$$

ou encore:

$$y_t = \mu + 2y_{t-1} - y_{t-2} + \varepsilon_t$$

2. Auto-régression

Les modèles auto-régressifs supposent que y_t est une fonction linéaire des valeurs précédentes.

$$y_t = \mu + \phi_1 y_{(t-1)} + \phi_2 y_{(t-2)} + \phi_3 y_{(t-3)} + \varepsilon_t$$

Littérairement, chaque observation est constituée d'une composante aléatoire (choc aléatoire, ε) et d'une combinaison linéaire des observations précédentes. ϕ_1 , ϕ_2 et ϕ_3 dans cette équation sont les coefficients d'auto-régression

A noter que cette équation porte soit sur les données brutes, soit sur les données différenciées si une différenciation a été nécessaire. Pour un modèle ARIMA(1,1,0) on aura :

$$y_t - y_{t-1} = \mu + \phi(y_{t-1} - y_{t-2}) + \varepsilon_t$$

Ce qui peut également être écrit:

$$y_t = \mu + y_{t-1} + \phi(y_{t-1} - y_{t-2}) + \varepsilon_t$$

Notez qu'un processus auto-régressif ne sera stable que si les paramètres sont compris dans un certain intervalle ; par exemple, s'il n'y a qu'un paramètre auto-régressif, il doit se trouver dans l'intervalle $-1 < \phi_1 < +1$. Dans les autres cas, les effets passés s'accumuleraient et les valeurs

successives des x_t se déplaceraient infiniment vers l'avant, ce qui signifie que la série ne serait pas stationnaire. S'il y a plus d'un paramètre auto-régressif, des restrictions similaires (générales) sur les valeurs des paramètres peuvent être posées (par exemple, voir Box et Jenkins, 1976 ; Montgomery, 1990).

3. Moyenne mobile

Les modèles à moyenne mobile suggèrent que la série présente des fluctuations autour d'une valeur moyenne. On considère alors que la meilleure estimation est représentée par la moyenne pondérée d'un certain nombre de valeurs antérieures (ce qui est le principe des procédures de moyennes mobiles utilisées pour le lissage des données). Ceci revient en fait à considérer que l'estimation est égal à la moyenne vraie, auquel on ajoute une somme pondérée des erreurs ayant entaché les valeurs précédentes :

$$y_t = \mu - \theta_1 \varepsilon_{(t-1)} - \theta_2 \varepsilon_{(t-2)} - \theta_3 \varepsilon_{(t-3)} + \varepsilon_t.$$

Littérairement, chaque observation est composée d'une composante d'erreur aléatoire (choc aléatoire, ε) et d'une combinaison linéaire des erreurs aléatoires passées. θ_1 , θ_2 et θ_3 sont les coefficients de moyenne mobile du modèle.

Comme précédemment cette équation porte soit sur les données brutes, soit sur les données différenciées si une différenciation a été nécessaire. Pour un modèle ARIMA(0,1,1) on aura :

$$y_t - y_{t-1} = \mu - \theta \varepsilon_{t-1} + \varepsilon_t$$

Ce qui peut également être écrit:

$$y_t = \mu + y_{t-1} - \theta \varepsilon_{t-1} + \varepsilon_t$$

Un modèle de moyenne mobile correspond à des séries exhibant des fluctuations aléatoires autour d'une moyenne variant lentement. Plutôt que de prendre comme précédemment la valeur précédente comme prédicteur, on utilise une moyenne de quelques observations précédentes, de manière à éliminer le bruit, et estimer plus précisément la moyenne locale.

Cette logique correspond au **lissage exponentiel simple**, qui considère chaque observation comme la résultante d'une constante (b) et d'un terme d'erreur ε , soit :

$$y_t = b + \varepsilon_t$$

La constante b est relativement stable sur chaque segment de la série, mais peut se modifier lentement au cours du temps. Si ce modèle est approprié, l'une des manières d'isoler la réelle valeur de b, et donc la partie systématique ou prévisible de la série, consiste à calculer une sorte de moyenne mobile, où les observations courantes et immédiatement précédentes ("les plus récentes") ont une pondération plus forte que les observations plus anciennes. C'est exactement ce que fait un lissage exponentiel simple, où les pondérations les

plus faibles sont affectées exponentiellement aux observations les plus anciennes. La formule spécifique de lissage exponentiel simple est :

$$y_t = \alpha \hat{y}_t - (1-\alpha) y_{t-1}$$

Lorsqu'on l'applique de façon récurrente à chaque observation successive de la série, chaque nouvelle valeur prédite est calculée comme la moyenne pondérée de l'observation courante et de l'observation précédente prédite ; la précédente observation prédite était elle-même calculée à partir de la valeur (précédente) observée et de la valeur prédite avant cette valeur (précédente), et ainsi de suite. Par conséquent, chaque valeur prédite est une moyenne pondérée des observations précédentes, où les poids décroissent exponentiellement selon la valeur des paramètres α . Si α est égal à 1 les observations précédentes sont complètement ignorées ; si α est égal à 0, l'observation courante est totalement ignorée, et la valeur prédite ne porte que sur les valeurs prédites précédentes (qui est calculée à partir de l'observation lissée qui lui précède, et ainsi de suite ; c'est pourquoi toutes les valeurs prédites auront la même valeur que la valeur initiale \hat{y}_0). Les valeurs intermédiaires de α produiront des résultats intermédiaires (noter que la valeur $1-\alpha$ correspond au θ des équations précédentes).

On peut également envisager des modèles mixtes: par exemple un modèle ARIMA(1,1,1) aura l'équation de prédiction suivante:

$$y_t = \mu + y_{t-1} + \phi(y_{t-1} - y_{t-2}) - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Néanmoins on préfère généralement utiliser de manière exclusive les termes AR ou MA.

4. Signification des paramètres des modèles ARIMA

L'objectif essentiel des modèles ARIMA est de permettre une prédiction de l'évolution future d'un phénomène. Son développement dans le domaine de l'économétrie est basé sur ce principe. On en verra plus loin une illustration.

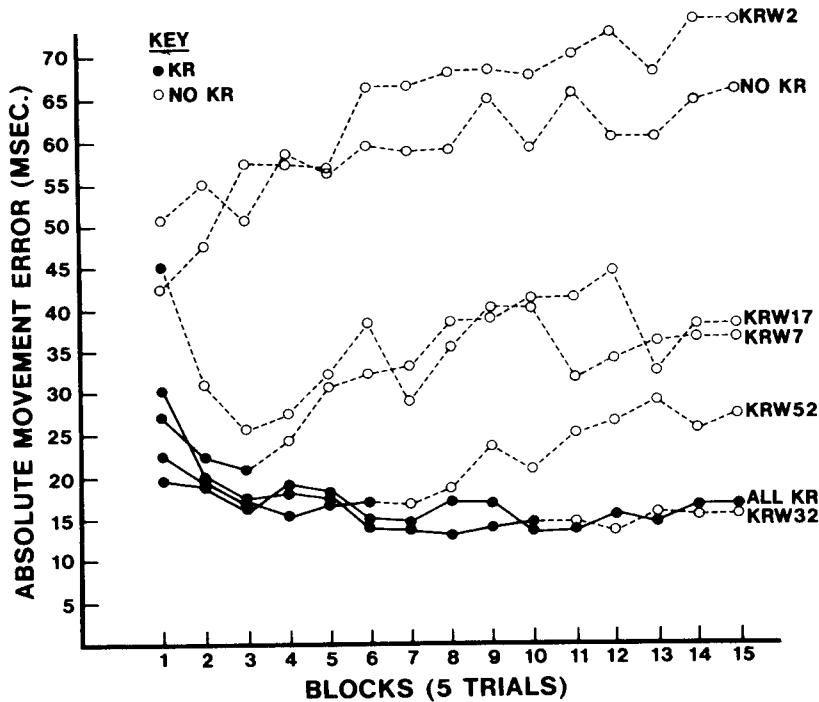
Un autre intérêt, peut-être plus essentiel en ce qui concerne la recherche scientifique, est de comprendre la signification théorique de ces différents processus. Il est clair cependant que cette interprétation dépend de la nature du phénomène étudié, et des modèles dont le chercheur dispose pour en rendre compte.

- Un processus non différencié à bruit blanc (ARIMA(0,0,0)) suggère des fluctuations aléatoires autour d'une valeur de référence. Cette valeur de référence peut être considérée comme une caractéristique stable du système étudié (trait de personnalité, mémoire, capacité stabilisée, etc..)

- Un processus de moyenne mobile suggère que la valeur de référence évolue d'une mesure à l'autre. Plus précisément, la valeur de référence est fonction de la valeur de référence précédente et de l'erreur ayant entaché la mesure précédente.

- Un processus auto-régressif suggère que le phénomène étudié n'est pas déterminé par une valeur de référence. C'est la performance précédente (ou les performances précédentes) qui déterminent entièrement la performance présente.

Par exemple, Spray et Newell (1986) analysent des données tirées d'une expérimentation portant sur le rôle de la connaissance des résultats dans l'apprentissage. Les sujets réalisent 77 essais dans une tâche manuelle. Le protocole comprenait plusieurs groupes, différenciés par des combinaisons spécifiques d'essais avec ou sans connaissance des résultats. Notamment, certains sujets disposaient de connaissance des résultats tout au long des 77 essais, pour d'autre la CR était supprimée au-delà de 17, 32 ou 52 essais. Un groupe n'avait pas de CR du tout.



Les résultats de la modélisation montrent que les séries avec CR (ou les portions de séries avec CR) peuvent être représentée par des processus à bruit blanc du type:

$$y_t = \mu + \varepsilon_t$$

C'est-à-dire un modèle ARIMA (0,0,0). Cette équation suggère donc que les performances successives oscillent de manière aléatoire autour d'une valeur moyenne, sorte de référence interne construite par la connaissance des résultats.

Les séries sans CR (ou les portions de série sans CR) sont quant à elles modélisées selon un ARIMA(0,1,1) selon la formule:

$$y_t = \mu - \theta_1 \varepsilon_{(t-1)} + \varepsilon_t$$

ou

$$y_t = r_t + \varepsilon_t$$

r_t représentant la valeur de référence, qui cette fois change à chaque essai. On peut dériver du modèle que

$$r_t = r_{t-1} - \theta_1 \varepsilon_{(t-1)}$$

C'est-à-dire que la référence est une combinaison de la référence précédente et de l'erreur ayant entaché l'essai précédent. Ce modèle indique clairement que l'essai en cours est influencé par l'essai précédent, ce qui n'était pas le cas dans les essais avec CR.

Ce modèle peut également être écrit sous la forme d'une interpolation pondérée entre la performance au temps t et la référence au temps $t-1$:

$$r_t = -\theta_1 y_t + (1+\theta_1)r_{t-1}$$

L'analyse des données de Diggles (1977) suggère que la référence précédente est plus importante que la performance actuelle.

On peut noter que pour les sujets ayant bénéficié de la CR durant 52 essais sur 77, la série demeure stationnaire et à bruit blanc jusqu'à la fin de l'expérimentation.

5. Détermination de l'ordre de différenciation.

Une série stationnaire fluctue autour d'une valeur moyenne et sa fonction d'autocorrélation décline rapidement vers zéro.

Si une série présente des auto-corrélations positives pour un grand nombre de décalages (par exemple 10 ou plus), alors elle nécessite d'être différenciée.

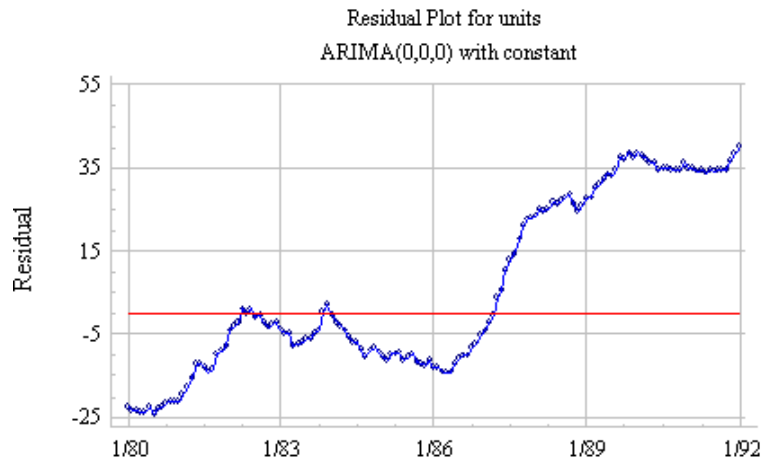
La différenciation tend à introduire des auto-corrélations négatives.

Si l'auto-corrélation de décalage 1 est égale à 0 ou négative, la série n'a pas besoin d'être différenciée. Si l'auto-corrélation de décalage 1 est inférieure à -0.5 , la série est sur-différenciée.

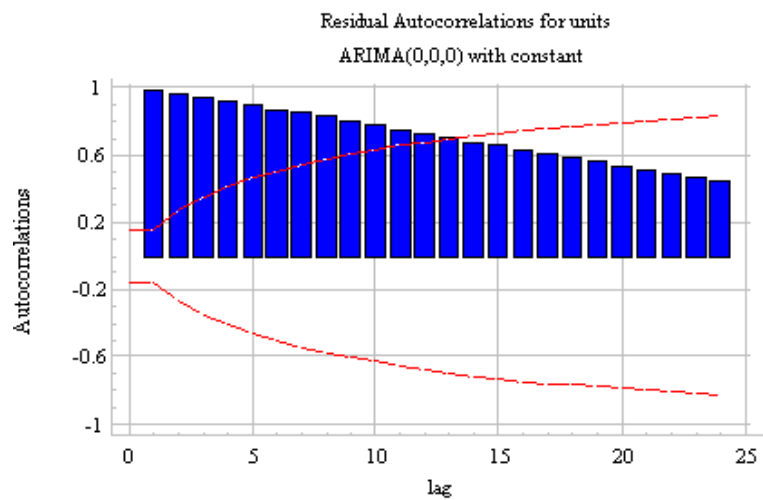
L'ordre optimal de différenciation est souvent celui pour lequel l'écart-type est minimal. Un accroissement de l'écart-type doit donc être considéré comme un symptôme de sur-différenciation.

Un troisième symptôme de sur-différenciation est un changement systématique de signe d'une observation à l'autre.

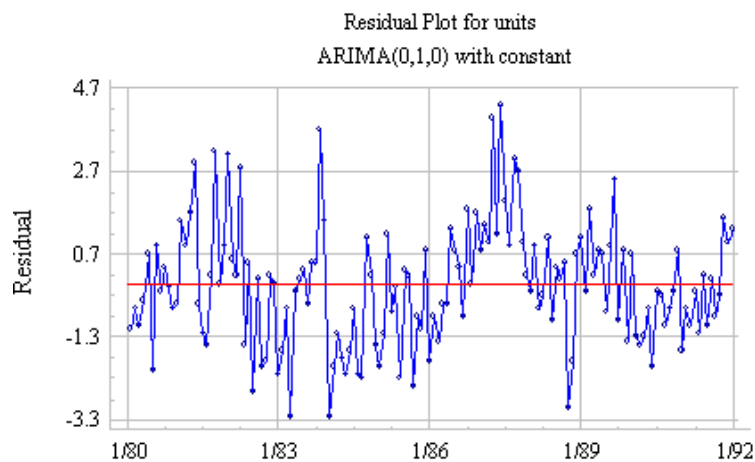
Par exemple, la figure suivante représente une série temporelle originale (il s'agit de l'évolution temporelle d'unités de vente).



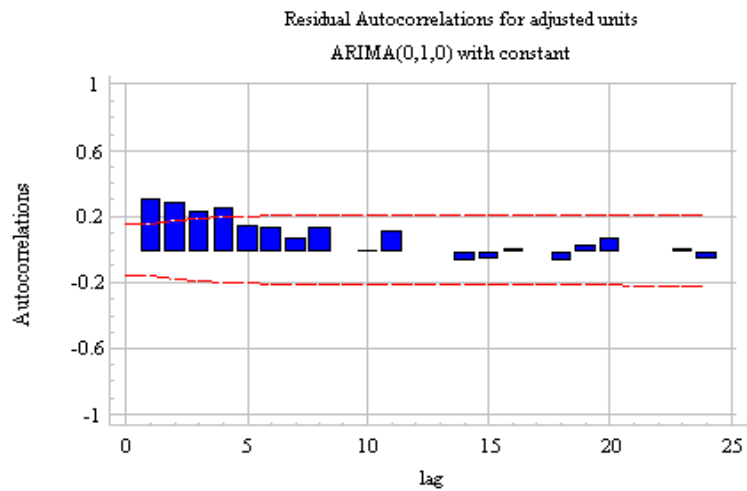
Le graphique de la fonction d'auto-corrélation présente une régression lente et linéaire, typique des séries non stationnaires:



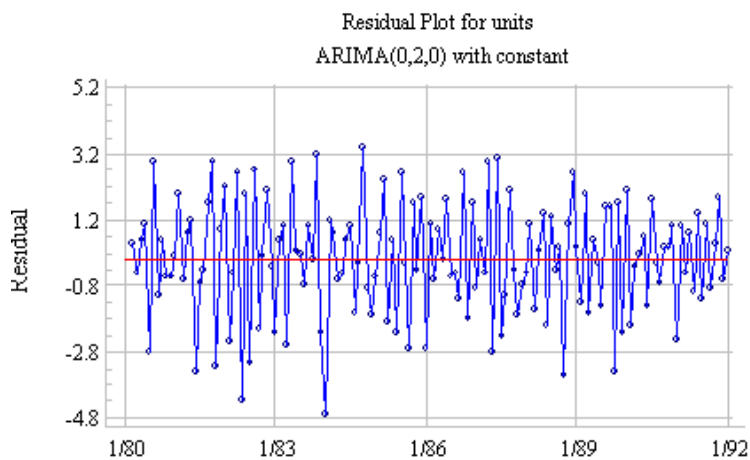
Enfin l'écart-type de la série est important: 17.56. Cette série nécessite de toute évidence d'être différenciée. Une première différenciation (c'est-à-dire l'application d'un ajustement ARIMA (0,1,0) avec constante) donne les résidus suivants:



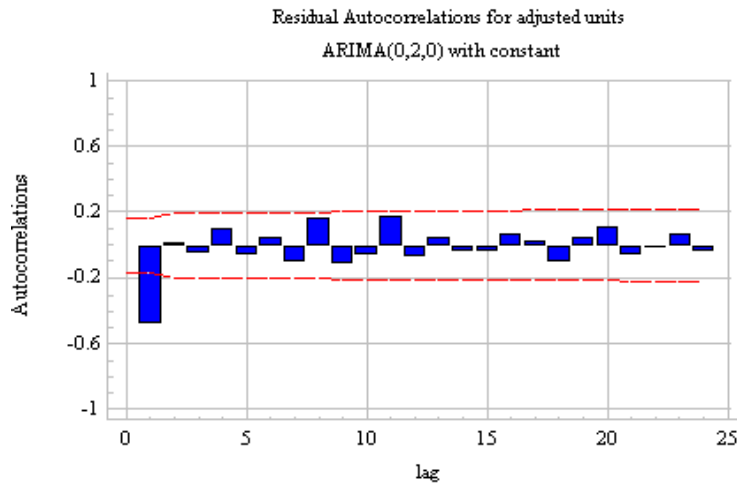
La série semble a peu près stationnaire, sans tendance marquée à long terme: la série semble toujours revenir à sa valeur moyenne, bien que cette tendance reste pour le moins "paresseuse". La fonction d'auto-corrélation confirme la persistance d'auto-corrélation positive:



L'écart-type a cependant été réduit de manière importante, de 17.56 à 1.54. Si l'on essaie une seconde différenciation (c'est-à-dire l'application d'un modèle ARIMA(0,2,0) avec constante, on obtient les résultats suivants:



Cette série montre de clairs signes de sur-différenciation. Les valeurs passe de manière quasi-systématique du signe positif au signe négatif. La fonction d'auto-corrélation montre un pic négatif au premier décalage; proche de .5.



Enfin, l'écart-type a augmenté de 1.54 à 1.81. Ceci semble indiquer que l'ordre optimal de différenciation, pour cette série, est de 1. Le modèle devra cependant être optimisé (pour faire disparaître toute trace d'auto-corrélation dans les résidus), par l'ajout de termes AR ou MA.

La différenciation dont on a parlé est dite non-saisonnnière, dans le sens où elle porte sur des données adjacentes. On verra plus loin qu'une différenciation saisonnière peut être nécessaire pour rendre compte de variations systématiques, circadiennes, hebdomadaires, mensuelles ou annuelles.

Un modèle sans différenciation suppose que la série originale est stationnaire. Un modèle avec une différenciation d'ordre 1 suppose que la série originale présente une tendance constante. Un modèle avec une différenciation d'ordre 2 suppose que la série originale présente une tendance variant dans le temps.

Les modèles ARIMA peuvent inclure une constante ou non (sans constante signifie que la constante est égale à 0). L'interprétation d'une constante (significativité statistique) dépend du modèle.

- Un modèle sans différenciation possède généralement une constante (qui représente dans ce cas la moyenne de la série).

- Si la série est différenciée, la constante représente la moyenne ou l'ordonnée à l'origine de la série différenciée ; par exemple, si la série est différenciée une fois, et qu'il n'y a pas de paramètre auto-régressif dans le modèle, la constante représentera la moyenne de la série différenciée, et donc la pente du trend linéaire de la série non différenciée..

- Dans le cas des modèles avec un ordre de différenciation de 2, la constante représente la *tendance moyenne de la tendance*. Dans la mesure où en général on ne suppose pas l'existence de telles tendances, la constante est généralement omise.

- S'il n'y a pas de paramètre auto-régressif dans le modèle, l'espérance mathématique de la constante est m , la moyenne de la série ;

- S'il y a des paramètres auto-régressifs dans la série, la constante représente l'ordonnée à l'origine.

A noter que la moyenne, dans les modèles ARIMA, renvoie à la moyenne des séries différenciées, alors que la constante est un facteur qui apparaît dans la partie droite des équations de prédiction. Moyenne et constante sont liées par l'équation suivante:

$$\mu = \text{moyenne} \times (1 - \sum \text{AR}(p))$$

La constante est égale à la moyenne, multipliée par 1 moins la somme des coefficients des termes auto-régressifs.

6. Identification des termes AR.

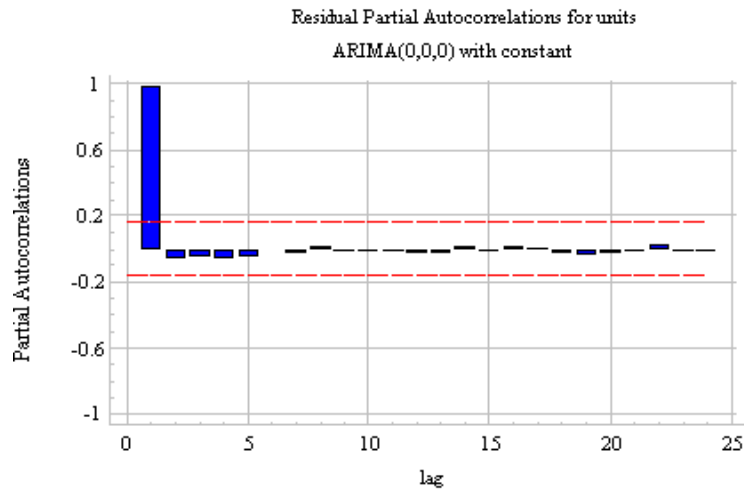
Après que la série ait été stationnarisée, l'étape suivante consiste à identifier les termes AR et MA nécessaires pour corriger les auto-corrélations résiduelles. Cette analyse est basée sur l'examen des fonctions d'auto-corrélation et d'auto-corrélation partielle. Rappelons que l'auto-corrélation est la corrélation d'une série avec elle-même, selon un décalage (*lag*) défini. L'auto-corrélation de décalage 0 est par définition égale à 1. La fonction d'auto-corrélation fait correspondre à chaque décalage l'auto-corrélation correspondante.

D'une manière générale, une corrélation partielle entre deux variables est la quantité de corrélations qui n'est pas expliquée par les relations de ces variables avec un ensemble spécifié d'autres variables. Supposons par exemple que l'on réalise la régression de Y sur trois variables X1, X2 et X3. La corrélation partielle entre Y et X3 contrôlant X1 et X2 est la quantité de corrélation entre Y et X3 qui n'est pas expliqué par leurs relations communes avec X1 et X2. Elle peut être calculée comme la racine carrée du gain de variance expliquée obtenu en ajoutant X3 à la régression de Y sur X1 et X2.

Dans le cas des séries temporelles, la corrélation partielle de décalage k est la corrélation entre y_t et y_{t-k} , contrôlant l'influence des k-1 valeurs interposées.

L'autocorrélation de décalage 1 est la corrélation entre y_t et y_{t-1} . On suppose que c'est également la corrélation entre y_{t-1} et y_{t-2} . Si y_t et y_{t-1} sont corrélés, et que y_{t-1} et y_{t-2} le sont également, on peut supposer qu'une corrélation sera présente entre y et y_{t-2} . C'est-à-dire que la corrélation de décalage 1 se propage au décalage 2 et sans doute aux décalages d'ordre supérieurs. Plus précisément, la corrélation attendue au décalage 2 est la carré de la corrélation observée au décalage 1. L'auto-corrélation *partielle* de décalage 2 est donc la différence entre l'auto-corrélation de décalage 2 et la corrélation attendue due à la propagation de la corrélation de décalage 1.

Si l'on revient à la fonction d'auto-corrélation de l'exemple précédent (avant différenciation), on peut supposer que la présence d'auto-corrélations fortes pour un grand nombre de décalages successifs est lié à ce phénomène de propagation. Ceci est confirmé par l'examen de la fonction d'auto-corrélation partielle, qui n'a qu'une valeur significative au décalage 1 (notons que l'auto-corrélation partielle de décalage 1 est égale à l'auto-corrélation correspondante, aucune valeur n'étant intercalée).



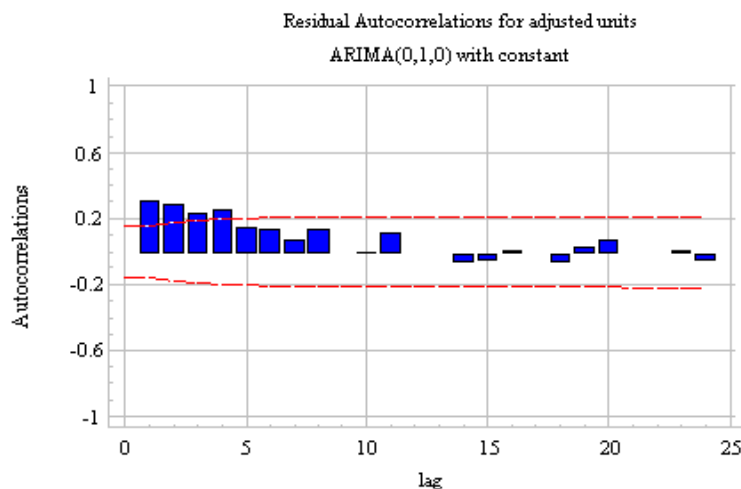
Ceci indique clairement que les auto-corrélations de décalage supérieur à 1 ne sont dues qu'à la propagation de l'auto-corrélation de décalage 1.

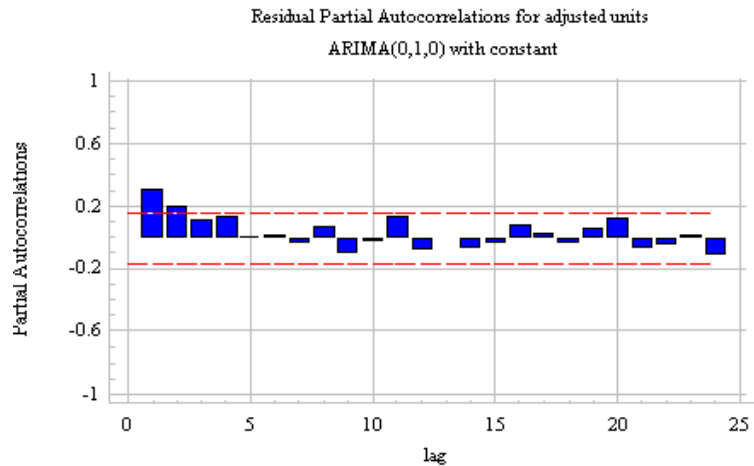
Une extinction brutale de l'auto-corrélation partielle associée à un déclin plus progressif de l'auto-corrélation constitue la signature caractéristique d'un processus auto-régressif. Plus particulièrement, l'auto-corrélation partielle de décalage k est égale au coefficient $AR(k)$ estimé dans un modèle contenant k termes AR. On pourrait d'ailleurs déterminer les coefficients AR par régression multiple, en prédisant $(y_t - y_{t-1})$ à partir de k échantillons représentant les k décalages.

Le décalage auquel l'auto-corrélation partielle disparaît indique le nombre de termes auto-régressifs à inclure.

Généralement, ce pattern est associé à une auto-corrélation de décalage 1 positive, signe que la série demeure sous-différenciée. Une légère sous-différenciation peut donc être compensée par l'ajout d'un terme auto-régressif.

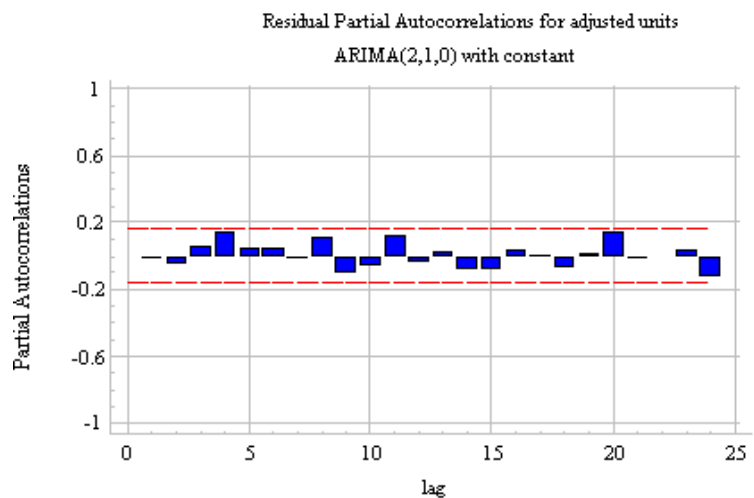
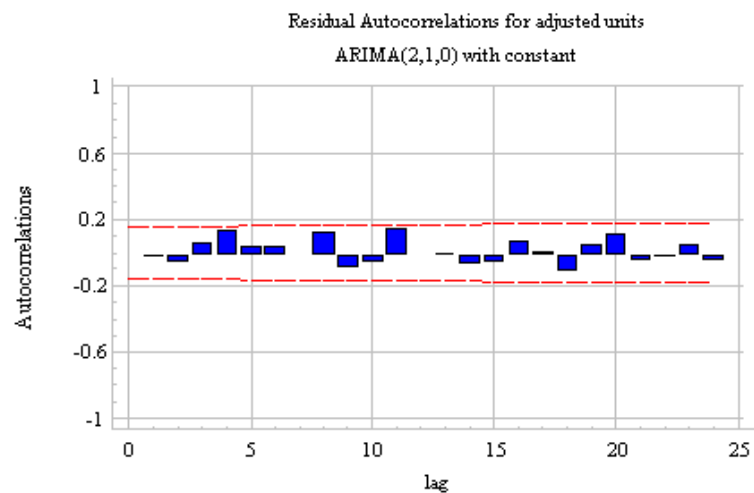
Si l'on reprend la série temporelle qui a précédemment servi d'exemple, on trouvait donc avec un ordre de différence (ajustement par ARIMA(0,1,0)) les fonctions d'auto-corrélation et d'auto-corrélation partielle suivantes:





On peut noter que l'auto-corrélation de décalage 1 est significative et positive, et que l'auto-corrélation partielle présente une extinction plus brutale que l'auto-corrélation. L'auto-corrélation partielle ne présente en fait que deux pics significatifs, alors que l'auto-corrélation en présente quatre. La série différenciée présente donc une signature de processus auto-régressif d'ordre 2.

L'ajustement de la série avec un modèle ARIMA(2,1,0) donne les fonctions ACF et PACF suivantes:

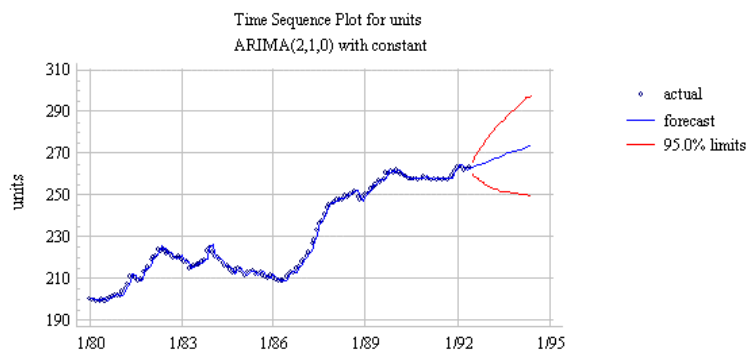


Les auto-corrélations aux décalages 1 et 2 ont été supprimées, et aucune auto-corrélation n'apparaît aux décalages supérieurs. L'analyse montre que les coefficients AR sont significativement différents de 0, et l'écart-type des résidus a été réduit de 10% (de 1.54 à 1.42 par l'ajout des termes AR L'équation de prédiction a la forme suivante:

$$y_t = \mu + y_{t-1} + \phi_1(y_{t-1} - y_{t-2}) + \phi_2(y_{t-2} - y_{t-3})$$

Dans ce cas, $\mu = 0.258178$
 $\phi_1 = 0.2524$
 $\phi_2 = 0.195572$

Cette équation permet la prédiction des données ultérieures à la période dite de validation du modèle. Le graphique suivant rend compte de cette prédiction:



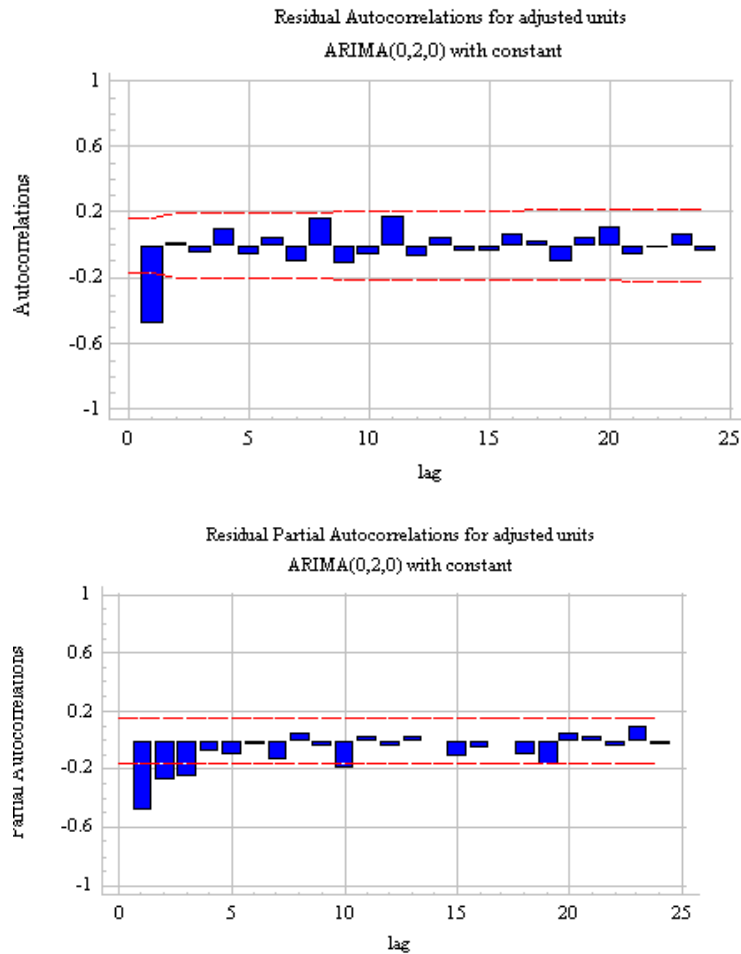
7. Identification des termes MA.

La fonction d'auto-corrélation joue pour les processus de moyenne mobile le même rôle que la fonction d'auto-corrélation partielle pour les processus auto-régressifs. Si l'auto-corrélation est significative au décalage k mais plus au décalage $k+1$, ceci indique que k termes de moyenne mobile doivent être ajoutés au modèle.

A noter que si les coefficients AR peuvent être estimés par une analyse en régression multiple, une telle démarche est impossible pour les coefficients MA. D'une part, parce que l'équation de prédiction est non-linéaire, et d'autre part les erreurs ne peuvent être spécifiées en tant que variable indépendantes. Les erreurs doivent être calculées pas à pas en fonction des estimations courantes des paramètres.

Une signature MA est généralement associée à une auto-corrélation négative au décalage 1, signe que la série est sur-différenciée. Une légère sur-différenciation peut donc être compensée par l'ajout d'un terme de moyenne mobile.

On a vu qu'avec un ordre de différenciation de 1, la série qui nous a servi d'exemple demeurait sous-différenciée. Si l'on travaille sur des différences d'ordre 2 (par ajustement d'un modèle ARIMA(0,2,0) on obtient les fonction ACF et PACF suivantes:



Ces fonctions sont typiques de la signature des processus de moyenne mobile. Le pic unique et négatif pour la fonction d'auto-corrélation indique un processus MA(1). Le modèle pertinent serait donc un ARIMA(0,2,1). L'équation de prédiction serait la suivante:

$$y_t = 2y_{t-1} - y_{t-2} - \theta_1 e_{t-1}$$

Ces deux modèles peuvent ajuster de manière alternative la série de départ. Sachant que les termes AR peuvent compenser une légère sous différenciation, et les termes MA une légère sous-différenciation, il est courant que deux modèles alternatifs soient possibles: un premier avec 0 ou 1 ordre de différenciation combiné avec des termes AR, et un autre avec le niveau de différenciation supérieur, combiné à des termes MA. Le choix d'un ou l'autre modèle peut reposer sur des présupposé théoriques liés au phénomène observé.

Les outils principaux utilisés lors de la phase d'identification sont donc les tracés de la série, les corrélogrammes d'autocorrélation (FAC), et d'autocorrélation partielle (FACP). La décision n'est pas simple et les cas les plus atypiques requièrent, outre l'expérience, de nombreuses expérimentations avec des modèles différents (avec divers paramètres ARIMA). Toutefois, les composantes des séries chronologiques empiriques peuvent souvent être assez bien approchées en utilisant l'un des 5 modèles de base suivants, identifiables par la forme de l'autocorrélogramme (FAC) et de l'autocorrélogramme partiel (FACP). Puisque le nombre de paramètres (à estimer) de chaque type dépasse rarement 2, il est souvent judicieux d'essayer des modèles alternatifs sur les mêmes données.

(1) Un paramètre autorégressif (p) : FAC - décomposition exponentielle ; FACP - pic à la période 1, pas de corrélation pour les autres périodes.

(2) Deux paramètres autorégressifs (p) : FAC - une composante de forme sinusoïdale ou un ensemble de décompositions exponentielles ; FACP - pics aux périodes 1 et 2, aucune corrélation pour les autres périodes.

(3) Un paramètre de moyenne mobile (q) : FAC - pic à la période 1, aucune corrélation pour les autres périodes ; FACP - exponentielle amortie.

(4) Deux paramètres de moyenne mobile (q) : FAC - pics aux périodes 1 et 2, aucune corrélation pour les autres périodes ; FACP - une composante de forme sinusoïdale ou un ensemble de décompositions exponentielles.

(5) Un paramètre autorégressif (p) et un de moyenne mobile (q) : FAC - décomposition exponentielle commençant à la période 1 ; FACP - décomposition exponentielle commençant à la période 1.

8. Racines unitaires

Si une série est largement sur- ou sous-différenciée (c'est-à-dire qu'un ordre complet de différenciation doit être ôté ou ajouté), on obtient généralement une racine unitaire (*unit root*) au niveau des coefficients AR ou MA.

Dans le cas d'un modèle ARIMA(1,k,0), AR(1) est égal à 1. Dans le cas plus général d'un modèle ARIMA(p,k,0), la somme des coefficients AR est égale à 1 (c'est-à-dire non significativement différente de 1). On peut considérer qu'alors les termes AR simulent une différence (on a vu dans l'exemple précédent que le premier terme auto-régressif pouvait être substitué à la première différence). Le nombre de termes AR doit être réduit d'une unité et on doit accroître d'une unité l'ordre de différenciation.

Le même raisonnement peut être tenu pour les coefficients MA. Dans le cas d'un modèle ARIMA(0,k,1), MA(1) est égal à 1, et d'une manière générale la somme des coefficients MA est égale à 1. Les termes MA dans ce cas tendent à annuler exactement une différence. Le nombre de termes MA doit être réduit d'une unité et on doit réduire d'une unité l'ordre de différenciation.

9. Evaluation des modèles

L'objectif de la modélisation est de déterminer combien de paramètres auto-régressifs (p) et de moyennes mobiles (q) sont nécessaires pour obtenir un modèle effectif et parcimonieux du processus (parcimonieux signifie qu'il s'agit, parmi tous les modèles possibles, du modèle possédant le moins de paramètres et le plus grand nombre de degrés de liberté pour ajuster les données). En pratique, le nombre de paramètres p ou q dépasse très rarement 2.

Estimation des Paramètres: l'analyse produit les valeurs approchées de t, calculées à partir des erreurs-types des paramètres. Si cette valeur de t n'est pas significative, le paramètre respectif

peut être éliminé du modèle, sans affecter dans la plupart des cas l'ajustement global du modèle.

Un autre indicateur simple et courant pour mesurer la fiabilité du modèle est de comparer la précision des estimations basées sur des données partielles, afin que les prévisions puissent être comparées avec les dernières observations connues de la série (initiale).

Toutefois, un bon modèle ne doit pas seulement produire des prévisions suffisamment précises, il doit également être parcimonieux et produire des résidus statistiquement indépendants, ne contenant que du bruit, sans aucune composante régulière (par exemple, le corrélogramme des résidus ne doit pas révéler d'autocorrélations). Un bon test du modèle consiste (a) à tracer les résidus et à les examiner pour voir s'il existe des tendances systématiques, et (b) à étudier l'autocorrélogramme des résidus (il ne doit pas y avoir d'autocorrélations entre les résidus).

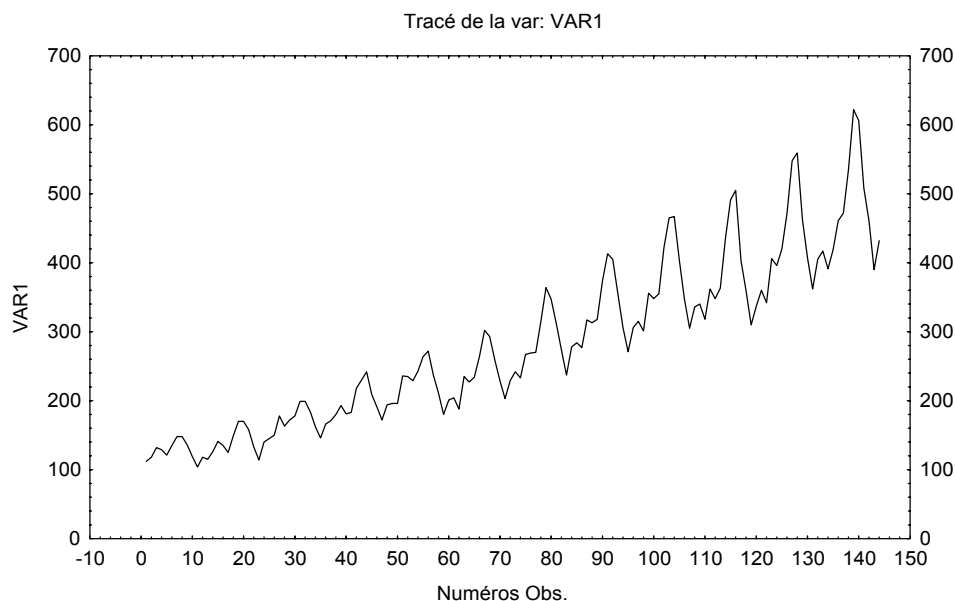
Ce que l'on cherche à savoir ici, c'est si les résidus ne sont pas distribués de façon systématique dans la série (par exemple, systématiquement négatifs dans une première partie de la série puis proches de zéro dans une seconde) ou s'ils ne sont pas autocorrélés ce qui tendrait à montrer que le modèle ARIMA est inadéquat. L'analyse des résidus de l'ARIMA constitue un test important du modèle. La procédure d'estimation postule que les résidus ne sont pas (auto-) corrélés et qu'ils sont distribués normalement.

La méthode ARIMA n'est appropriée que lorsque la série chronologique est stationnaire (c'est-à-dire que les moyennes, variances, et autocorrélations doivent être sensiblement constantes au cours du temps). Il est également recommandé d'avoir au moins 50 observations dans le fichier de données. Les paramètres estimés sont considérés constants dans toute la série.

10. Modèles saisonniers

L'ARIMA saisonnière multiplicative est une généralisation et une extension de la méthode présentée aux paragraphes précédents, pour des séries où le même phénomène se répète au cours du temps, à chaque saison. Outre les paramètres non saisonniers, des paramètres saisonniers pour un décalage spécifié (repéré lors de la phase d'identification) doivent être estimés. Comme pour les paramètres de l'ARIMA simple, ce sont : les paramètres auto-régressifs saisonniers (ps), de différenciation saisonnière (ds), et de moyenne mobile saisonnière (qs). Par exemple, le modèle $(0,1,2)(0,1,1)$ décrit un modèle sans paramètre auto-régressif, avec 2 paramètres de moyenne mobile normaux et 1 paramètre de moyenne mobile saisonnier. Le calcul de ces paramètres intervient après différenciation une fois de 1 période, et une fois saisonnièrement. Le décalage saisonnier utilisé pour les paramètres saisonniers est souvent déterminé lors de la phase d'identification et doit être explicitement spécifié.

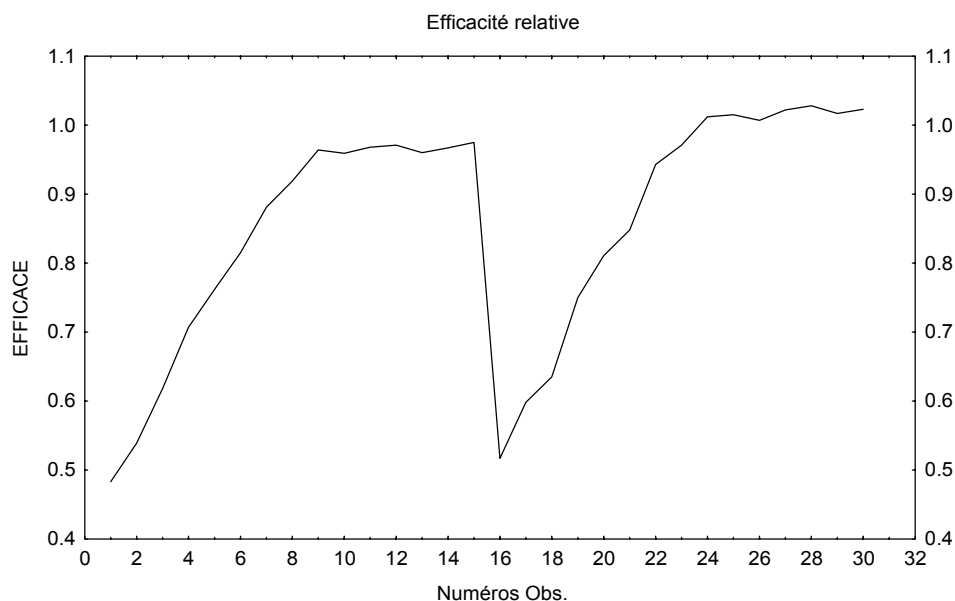
Les recommandations générales concernant la sélection des paramètres à estimer (basés sur les FAC et FACP) s'appliquent également aux modèles saisonniers (pour une illustration d'une ARIMA saisonnière, voir l'Exemple dans la section ARIMA du chapitre). La principale différence est que dans les séries saisonnières, les FAC et FACP auront des coefficients assez importants pour de nombreuses périodes saisonnières (en plus de leur composante globale reflétant la composante non saisonnière de la série).



11. ARIMA pour Séries Chronologiques Interrompues

Une question qui se pose fréquemment dans les analyses de séries chronologiques est de savoir si un événement extérieur (exogène) a perturbé les observations ultérieures. Par exemple, la mise en œuvre d'une nouvelle politique économique a-t-elle permis d'améliorer les performances économiques ; une nouvelle loi anti-crime a-t-elle permis d'infléchir les taux d'agression suivants ; et ainsi de suite. En général, nous souhaiterions pouvoir évaluer l'impact d'un ou plusieurs événements discrets sur les valeurs de la série chronologique. Ce type d'analyse de série chronologique interrompue est décrit en détail par McDowall, McCleary, Meidinger, et Hay (1980). McDowall distingue entre trois grands types d'impact possibles : (1) soudain et permanent, (2) permanent et graduel, et (3) soudain et temporaire.

Impact Soudain et Permanent



Un impact soudain et permanent implique simplement que la moyenne globale de la série chronologique se trouve modifiée après l'intervention ; le déplacement global est noté ω .

Impact Graduel et Permanent

La structure d'impact graduel et permanent implique que l'augmentation ou la diminution due à l'intervention soit progressive, et que l'impact permanent final ne devienne évident qu'après un certain temps :

$$\text{Impact}_t = \delta * \text{Impact}_{t-1} + \omega$$

(pour tout t période d'impact, sinon = 0).

Cette structure d'impact est définie par deux paramètres δ (delta) et ω (oméga). Si δ est proche de 0 (zéro), l'impact permanent final deviendra évident après quelques observations supplémentaires ; si δ est proche de 1, l'impact permanent final de l'impact deviendra évident après de nombreuses observations supplémentaires. Tant que le paramètre δ demeure supérieur à 0 et inférieur à 1 (les limites de stabilité du système), l'impact sera graduel et produira une modification (déplacement) asymptotique de la moyenne globale par la quantité :

$$\omega / (1 - \delta)$$

Les programmes d'analyse calculent automatiquement la modification asymptotique des impacts graduels et permanents. Notez que lors de l'évaluation d'un modèle, il est important que les deux paramètres soient statistiquement significatifs ; sinon, on pourrait aboutir à des conclusions paradoxales. Par exemple, supposons que le paramètre ω ne soit pas significativement différent de 0 (zéro) mais que le paramètre δ le soit ; cela voudrait dire qu'une intervention aurait causé une modification graduelle significative, sans que le résultat ne soit significativement différent de zéro.

Impact Soudain et Temporaire

Un impact soudain et temporaire implique que l'augmentation ou la diminution initiale soudaine due à l'intervention s'estompe lentement, sans modification permanente de la moyenne de la série. Ce type d'intervention peut être synthétisé par les expressions :

Avant l'intervention :	Impact_t = 0
Pendant l'intervention :	Impact_t = ω
Après l'intervention :	Impact_t = $\delta * \text{Impact}_{t-1}$

Notez que cette structure d'impact est à nouveau définie par les deux paramètres δ (delta) et ω (oméga). Tant que le paramètre δ demeurera supérieur à 0 et inférieur à 1 (les limites de stabilité du système), l'impact initial soudain se décomposera graduellement. Si δ est proche de 0 (zéro) la décomposition sera très rapide, et l'impact aura complètement disparu après quelques observations seulement. Si δ est proche de 1 la décomposition sera lente, et l'intervention continuera à affecter la série sur de nombreuses observations. Notez que lors de l'évaluation d'un modèle, il est encore une fois important que les deux paramètres soient statistiquement significatifs ; sinon on pourrait aboutir à des conclusions paradoxales. Par exemple, supposons que le paramètre ω ne soit pas significativement différent de 0 (zéro)

mais que le paramètre δ le soit ; cela voudrait dire qu'une intervention n'aurait pas causé de modification initiale soudaine, tout en montrant une décomposition significative.

Références essentielles :

Box, G.E.P. & Jenkins, G.M. (1976). Time series analysis : Forecasting and control. Oakland, CA : Holden-Day.

Spray, J.A. & Newell, K.M. (1986). Time series analysis of motor learning : KR versus no-KR. *Human Movement Science*, 5, 59-74.

Sites Internet :

<http://www.autobox.com/t1a13a.html>

<http://www.duke.edu/~rnau/arimest.htm>

Logiciels :

Systat (menu Time series)

Statistica (menu Séries chronologiques)

Minitab (menu Time series)