

Chapitre 11, exercice 2

Instructions pour employer Time Series Expert et Demetra pour analyser la série CHAMPC du répertoire CH011EX02

Le répertoire CH11EX02 comporte un exercice de base destiné à tous les apprenants et un exercice avancé réservé aux seuls apprenants de la version avancée du cours.

Exercice de base (Pour tous les utilisateurs du cours)

Préalable



Le chapitre 11 du cours de base doit avoir été suivi jusqu'à la page 63 pour la partie 1, jusqu'à la page 99 pour la partie 2, jusqu'à la page 100 pour la partie 3, jusqu'à la page 106 pour la partie 4, jusqu'à la page 107 pour la partie 5.

Objectif



Le but est de modéliser la série au moyen d'un modèle ARIMA complété par une analyse d'intervention, et d'en déduire des prévisions. La méthode de Box et Jenkins est employée à cette fin.

Le but est de se préparer à utiliser la méthode RegARIMA dans le module X-12-ARIMA de Demetra qui servira au chapitre 12 à effectuer un ajustement saisonnier, et la méthode TRAMO dans le module TRAMO/SEATS de Demetra qui servira au chapitre 13 à effectuer un ajustement saisonnier.

Données



Il s'agit des ventes mensuelles de champagne en France entre janvier 1962 et septembre 1970, en millions de bouteilles (Wheelwright et Makridakis, 1977), déjà employée dans le chapitre 4, exercice 3, dans le chapitre 5, exercices 6 à 10, dans le chapitre 6, exercice 5, dans le chapitre 8, exercice 3, dans le chapitre 9, exercice 2 et dans le chapitre 10, exercice 6. La série sera encore utilisée dans les chapitres suivants du cours.

Structure de l'exercice

L'exercice comporte cinq parties :

- Dans la partie 1, le but de l'exercice est de modéliser la série, dans Time Series Expert for Windows, au moyen d'un modèle ARIMA complété par une analyse d'interventions.
- Dans la partie 2, le but de l'exercice est de modéliser la série, par la méthode RegARIMA dans le module X-12-ARIMA de Demetra, au moyen d'un modèle ARIMA complété par une analyse d'interventions.
- Dans la partie 3, le but de l'exercice est de modéliser la série, par la méthode TRAMO dans le module TRAMO/SEATS de Demetra, au moyen d'un modèle ARIMA complété par une analyse d'interventions.
- Dans la partie 4, le but de l'exercice est de modéliser la série, par la méthode RegARIMA dans le module X-12-ARIMA de Demetra, au moyen d'un modèle ARIMA complété par une analyse d'interventions, ainsi qu'un traitement des jours ouvrables et de l'effet de Pâques.
- Dans la partie 5, le but de l'exercice est de modéliser la série, par la méthode TRAMO dans le module TRAMO/SEATS de Demetra, au moyen d'un modèle ARIMA complété par une analyse d'interventions, ainsi qu'un traitement des jours ouvrables et de l'effet de Pâques.

Partie 1 Dans la partie 1, le but de l'exercice est d'introduire des interventions dans le modèle ARIMA obtenu dans l'exercice 6 du chapitre 10.

1.1 INTRODUCTION

- ⇒ Afin de lancer le logiciel, suivez les instructions données en annexe de l'introduction du cours.
- ⇒ Choisissez le répertoire de données approprié sur votre disque (pas sur le CD-ROM) : menu File ⇒ Open. Choisissez DATA puis CHAP11 puis CH10EX02.



Remarque

Des instructions plus détaillées ont été données sous forme de remarques lors de la partie 1 de l'exercice 1 du chapitre 10. Prière de s'y référer.

- ⇒ Chargez le problème déjà préparé : CHAMPC2. La variable dépendante est CHAMPC, l'échantillon d'estimation est 1962.01 – 1969.12 et les prévisions seront calculées jusqu'en 1970.09.
- ⇒ Pour visualiser graphiquement la série : menu Graphics ⇒ Series et sélectionnez CHAMPC. Cliquez OK pour obtenir le graphique.

La spécification qui a été obtenue est la suivante :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t) - m] = (1 - \Theta_1 B^{12}) e_t.$$

Nous avons sauvegardé le fichier de sortie, sous le nom CHAMPC2.

- ⇒ Pour accéder au fichier de résultats : menu Reports ⇒ Statistic report. Sélectionnez CHAMPC2 et cliquez Open.

Le fichier s'ouvre.

- ⇒ Descendez sur la ligne “=== Residual analysis”.



L'analyse des résidus est-elle favorable au modèle ? Quels sont les aspects qui sont les plus défavorables ? Comment pourrait-on y remédier ?



1.1.1 Votre réponse

1.2 AJOUT D'UNE INTERVENTION



Pour visualiser graphiquement la série : menu Graphics \Rightarrow Series et sélectionnez RCHAMPC2. Cliquez OK.

On décide d'ajouter une intervention du type intervention sur la variable, à la date de janvier 1968. Elle a été justifiée par l'introduction de la TVA en France qui a apparemment eu un effet sur les ventes de champagne.



Justifiez l'intérêt de cette intervention ou effet de type AO ou *additive outlier*.



1.2.1 Votre réponse

La démarche est la même que dans l'exercice 1. Elle consiste à employer une variable binaire égale à 1 au temps où l'événement atypique s'est produit et égale à 0 ailleurs. Dans ce cas-ci, on définit une variable — qu'on appelle I_{6801} — comme suit : $I_{6801}_t = 1$ si $t = \text{janvier } 1968$, et $I_{6801}_t = 0$ si $t \neq \text{janvier } 1968$. L'impact de l'événement sera quantifié par un paramètre qu'il faudra estimer. Notons b_{6801} ce paramètre. En négligeant pour le moment l'autocorrélation, un modèle de régression pour la variable $\nabla_{12} \log(\text{CHAMPC}_t)$ serait le suivant :

$$\nabla_{12} \log(\text{CHAMPC}_t - b_{6801} I_{6801}_t) = m + e_t.$$

On remarque que l'intervention agit sur la variable elle-même et dans ce cas-ci n'apparaît pas devant le terme d'erreur. Le paramètre b_{6801} est exprimé dans les mêmes unités que les données, donc des millions de bouteilles.

En tenant compte de l'autocorrélation, on aura le modèle suivant :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t - b_{6801} I_{6801}_t) - m] = (1 - \Theta_1 B^{12}) e_t.$$

L'intervention agit sur la variable tandis que la constante m est soustraite de la différence saisonnière des logarithmes et le polynôme autorégressif

est appliqué sur l'expression entre crochets,

$$Y_t = \nabla_{12} \log(\text{CHAMPC}_t - b_{6801} \text{I6801}_t) - m.$$

Il faut noter que l'expression de Y_t fait intervenir deux paramètres. Dans ce cas-ci, l'écriture du modèle sans le recours à l'opérateur de retard — et en profitant pour mettre le plus de termes possibles dans le second membre — est relativement complexe :

$$\begin{aligned} \nabla_{12} \log(\text{CHAMPC}_t - b_{6801} \text{I6801}_t) \\ = m + \phi_1 [\nabla_{12} \log(\text{CHAMPC}_{t-1} - b_{6801} \text{I6801}_{t-1}) - m] + e_t - \Theta_1 e_{t-12}. \end{aligned}$$

Remarque



Pour ceux que les opérateurs de retard ne dérangent pas, on rencontre parfois une autre écriture du modèle, dans laquelle l'opérateur de différence est 'inversé' et qui permet alors de mettre plus de termes dans le second membre :

$$\nabla_{12} \log(\text{CHAMPC}_t - b_{6801} \text{I6801}_t) = m + \frac{(1 - \Theta_1 B^{12})}{(1 - \phi_1 B)} e_t$$

1.3 ESTIMATION DU MODÈLE AVEC UNE INTERVENTION

Il y a quatre paramètres à estimer : m , b_{6801} , ϕ_1 et Θ_1 . Nous avons vu le principe de la méthode d'estimation dans la partie A de l'exercice 1. Nous nous contentons ici d'observer les résultats, qui sont obtenus par des méthodes numériques itératives.

Nous employons à cette fin Time Series Expert. Dans les deux autres parties de l'exercice, nous emploierons aussi le logiciel Demetra qui sera utilisé pour la désaisonnalisation dans les chapitres 12 et 13.

- ⇒ Pour accéder à la boîte de dialogue pour l'estimation : menu Methods ⇒ Arima model ⇒ Estimation.
- ⇒ Dans la fenêtre de dialogue Arima model – Estimation, vous devez avoir 0 en face de Differences et 1 en face de Seasonal differences.
- ⇒ Cliquez Add Parameters pour ouvrir la fenêtre de dialogue. La fenêtre doit montrer les deux types de paramètres sélectionnés, 'Ordinary autoregressive' d'ordre 1 et 'Seasonal moving average' d'ordre 1. Les valeurs initiales sont 0.0000. Cliquez OK pour valider le choix des paramètres.
- ⇒ Cliquez Add Interventions pour ouvrir la fenêtre de dialogue. La

zone Type apparaît.

- ⇒ Cliquez sur la ligne On the variable : Box-Tiao (I). Cliquez Add. Cliquez sur la zone date où 0 est affiché. Tapez 1968.01, avec un point décimal, en aucun cas une virgule. Cliquez dans la zone Coefficient et sélectionnez Parameter.
- ⇒ Cliquez OK pour revenir à la fenêtre de dialogue Arima model – Estimation.
- ⇒ Sur la ligne Save residuals, vous devez voir “RCHAMPC2”. Cliquez afin d’éditer la zone. Pressez la touche de correction pour enlever 2, et tapez 3 pour que le nom devienne RCHAMPC3.
- ⇒ Sur la ligne Save forecasts or predictions, vous devez voir “PCHAMPC2”. Cliquez afin d’éditer la zone. Pressez la touche de correction pour enlever 2, et tapez 3 pour que le nom devienne PCHAMPC3.
- ⇒ Sur la ligne Save fitted values, vous devez voir “FCHAMPC2”. Cliquez et effacez le nom.
- ⇒ Cliquez pour lancer le programme. Nous allons consulter le fichier de sortie.

Après l’estimation, qui dure une fraction de seconde sur un ordinateur moderne, le fichier de sortie apparaît.



Remarque

Le format est identique à celui discuté dans l’exercice 1 du chapitre 11.

Vous pouvez sauvegarder le fichier de sortie, par exemple sous le nom CHAMPC3.

- ⇒ Dans la fenêtre de sortie, clic droit et choisissez Save as. Tapez CHAMPC3. Cliquez Save.



Ecrivez l’équation du modèle estimé accompagné de l’écart-type résiduel, AIC et BIC. Interprétez les résultats de l’estimation.



1.3.1 Votre réponse

?

L'analyse des résidus est-elle favorable au modèle ? Notez le critère MAPE pour les prévisions.



1.3.2 Votre réponse



Pour visualiser la série résiduelle : menu Graphics \Rightarrow Residuals.

?

Remarquez-vous quelque chose d'anormal ?



1.3.3 Votre réponse

?

Que pouvez-vous conclure ?



1.3.4 Votre réponse

Nous allons poursuivre le problème. Il est donc recommandé de sauvegarder le problème dans l'état actuel.



Pour sauvegarder le problème en cours : menu File \Rightarrow Save as. Tapez le nom CHAMPC3 et cliquez Save pour sauvegarder le problème sous le nom CHAMPC3.

1.4 ESTIMATION DU MODÈLE AVEC DEUX INTERVENTIONS

Le programme signale encore une valeur aberrante fortement significative, cette fois à la date de mai 1968. Ceci suggère l'ajout d'une intervention à cette date. Elle est justifiée par les grèves quasi insurrectionnelles qui ont suivi l'agitation étudiante à cette date. Le modèle s'écrit alors :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t - b_{6801} \text{I6801}_t - b_{6805} \text{I6805}_t) - m] \\ = (1 - \Theta_1 B^{12}) e_t.$$

⇒ Procéder comme dans le paragraphe 1.3 avec les modifications suivantes :

- ajoutez une intervention de type On the variable : Box-Tiao (I), mais avec la date 1968.05.
- changez les noms des séries des rubriques suivantes
Save residuals en "RCHAMPC4",
Save forecasts or predictions en "PCHAMPC4".
- sauvegardez la sortie sous le nom CHAMPC4, par exemple.
- sauvegardez le problème sous le nom CHAMPC4, par exemple, quand les réponses aux questions ci-dessous auront été trouvées.



Écrivez l'équation du modèle estimé accompagné de l'écart-type résiduel, AIC et BIC. Interprétez les résultats de l'estimation.



1.4.1 Votre réponse



L'analyse des résidus est-elle favorable au modèle ? Notez le critère MAPE pour les prévisions.



1.4.2 Votre réponse



Pour visualiser la série résiduelle : menu Graphics ⇒ Residuals.



Remarquez-vous quelque chose d'anormal ?



1.4.3 Votre réponse



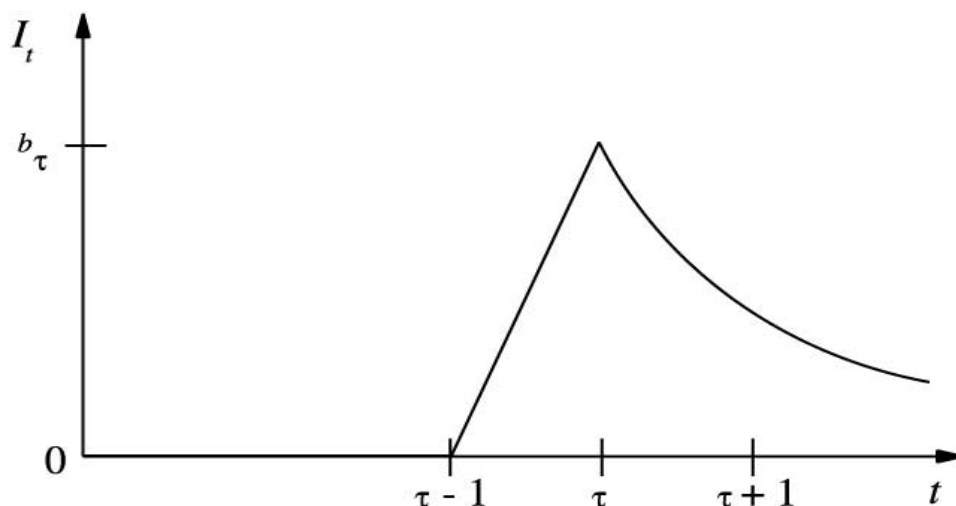
Que pouvez-vous conclure ?



1.4.4 Votre réponse

1.5 ESTIMATION DU MODÈLE AVEC DEUX AUTRES INTERVENTIONS

L'examen graphique des résidus montre un changement de niveau à partir du début de 1968. Il reste notamment une donnée aberrante en février 1968. On peut supposer que l'introduction de la TVA en France a eu un effet plus long qu'en un seul mois. Une représentation adéquate est probablement de la forme suivante.



Dans un premier temps, nous allons retirer l'intervention sur mai 1968 au profit d'un changement temporaire '*temporary change*' ou '*transient change*' qu'on peut représenter par une impulsion, donc une intervention de type AO (en janvier 1968, déjà choisie) suivi d'une décroissance exponentielle jusqu'à l'infini ou jusqu'à une date déterminée, ici septembre 1970. L'intervention devient plus complexe et il devient préférable de la noter, par exemple I_t , mais de la décrire de manière séparée. Le modèle s'écrit alors :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t - I_t) - m] = (1 - \Theta_1 B^{12}) e_t,$$

$$\text{où } \begin{aligned} I_t &= b_{6801} && \text{pour } t = 73 \\ I_t &= b_{6801} (1 + \delta_{6801})^{t-73} && \text{pour } t > 73, \end{aligned}$$

où b_{6801} et δ_{6801} sont deux paramètres et sachant que $t = 73$ correspond à la date de janvier 1968. D'un temps à l'autre, l'effet est multiplié par le facteur $(1 + \delta_{6801})$ qu'on suppose être inférieur à 1, de manière à représenter une décroissance. Par exemple, pour $t = 74, 75$ et 76 , on a respectivement $I_{74} = b_{6801} (1 + \delta_{6801})$, $I_{75} = b_{6801} (1 + \delta_{6801})^2$ et $I_{76} = b_{6801} (1 + \delta_{6801})^3$.

Remarques



1. Pour le point suivant, rappelons que pour supprimer une intervention, il faut choisir le type Delete current intervention.

2. L'intervention de type TC (temporary change) est représentée par la paire de paramètres b_{6801} et δ_{6801} , notés dans le programme respectivement KI6801 et KE7009.

⇒ Procéder comme dans le paragraphe 1.4 avec les modifications suivantes :

- supprimez l'intervention de type On the variable : Box-Tiao (I),

à la date 1968.05.

- ajoutez une intervention de type Continue : exponential line (E), avec la date de fin 1970.09 (en n'obéissant pas à l'instruction qui demande d'entrer une date entre 1962.01 et 1969.12).
- changez les noms des séries des rubriques suivantes
Save residuals en "RCHAMPC5",
Save forecasts or predictions en "PCHAMPC5",
Save fitted values, en "FCHAMPC5".
- sauvegardez la sortie sous le nom CHAMPC5, par exemple.
- sauvegardez le problème sous le nom CHAMPC5, par exemple, quand les réponses aux questions ci-dessous auront été trouvées.



Ecrivez l'équation du modèle estimé accompagné de l'écart-type résiduel, AIC et BIC. Interprétez les résultats de l'estimation.



1.5.1 Votre réponse



L'analyse des résidus est-elle favorable au modèle ? Notez le critère MAPE pour les prévisions.



1.5.2 Votre réponse



Pour visualiser la série résiduelle : menu Graphics \Rightarrow Residuals.



Remarquez-vous quelque chose d'anormal ?



1.5.3 Votre réponse



Que pouvez-vous conclure ?



1.5.4 Votre réponse

1.6 ESTIMATION DU MODÈLE AVEC TROIS INTERVENTIONS

Il subsiste une donnée aberrante extrêmement significative à la date de janvier 1965. Nous n'avons pas d'information sur sa cause. Nous allons toutefois la corriger. Nous utilisons une impulsion, donc une intervention de type AO. Si l'on note encore l'intervention globale I_t , on peut la décrire de la manière suivante :

$$\begin{aligned} I_t &= b_{6501} && \text{pour } t = 37 \\ I_t &= b_{6801} && \text{pour } t = 73 \\ I_t &= b_{6801} (1 + \delta_{6801})^{t-73} && \text{pour } t > 73, \end{aligned}$$

où b_{6501} , b_{6801} et δ_{6801} sont trois paramètres et sachant que $t = 37$ et $t = 73$ correspondent respectivement aux dates de janvier 1965 et de janvier 1968. Le modèle s'écrit alors, comme au paragraphe 1.5 :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t - I_t) - m] = (1 - \Theta_1 B^{12}) e_t.$$



Procéder comme dans le paragraphe 1.5 avec les modifications suivantes :

- ajoutez une intervention de type On the variable : Box-Tiao (I), à la date 1965.01.
- changez les noms des séries des rubriques suivantes
Save residuals en "RCHAMPC6",
Save forecasts or predictions en "PCHAMPC6", Save fitted values, en "FCHAMPC6".
- sauvegardez la sortie sous le nom CHAMPC6, par exemple.
- sauvegardez le problème sous le nom CHAMPC6, par exemple,

quand les réponses aux questions ci-dessous auront été trouvées.



? Ecrivez l'équation du modèle estimé accompagné de l'écart-type résiduel, AIC et BIC. Interprétez les résultats de l'estimation.

1.6.1 Votre réponse



? L'analyse des résidus est-elle favorable au modèle ? Notez le critère MAPE pour les prévisions.

1.6.2 Votre réponse



⇒ Pour visualiser la série résiduelle : menu Graphics ⇒ Residuals.

? Remarquez-vous quelque chose d'anormal ?

1.6.3 Votre réponse

? Que pouvez-vous conclure ?



1.6.4 Votre réponse

1.7 ESTIMATION DU MODÈLE AVEC QUATRE INTERVENTIONS

Une inspection du graphique des résidus en fonction du temps montre que plusieurs des premières années sont affectées par un mauvais ajustement au mois d'août. Il n'est pas très utile de corriger un seul de ces mois d'août. En revanche, en mai 1968 subsiste une donnée aberrante extrêmement significative. Nous allons donc corriger cette donnée grâce à une impulsion, donc une intervention de type AO. Si l'on note encore l'intervention globale I_t , on peut la décrire de la manière suivante :

$$\begin{aligned} I_t &= b_{6501} && \text{pour } t = 37 \\ I_t &= b_{6801} && \text{pour } t = 73 \\ I_t &= b_{6801} (1 + \delta_{6801})^{t-73} && \text{pour } t > 73, t \neq 77 \\ I_t &= b_{6805} + b_{6801} (1 + \delta_{6801})^{t-73} && \text{pour } t = 77 \end{aligned}$$

où b_{6501} , b_{6801} , b_{6805} et δ_{6801} sont quatre paramètres et sachant que $t = 37$, $t = 73$ et $t = 77$ correspondent respectivement aux dates de janvier 1965, de janvier 1968 et de mai 1968. Le modèle s'écrit alors, comme au paragraphe 1.6 :

$$(1 - \phi_1 B)[\nabla_{12} \log(\text{CHAMPC}_t - I_t) - m] = (1 - \Theta_1 B^{12}) e_t.$$

⇒ Procéder comme dans le paragraphe 1.6 avec les modifications suivantes :

- ajoutez une intervention de type On the variable : Box-Tiao (I), à la date 1968.05.
- changez les noms des séries des rubriques suivantes
Save residuals en "RCHAMPC7",
Save forecasts or predictions en "PCHAMPC7",
Save fitted values, en "FCHAMPC7".
- sauvegardez la sortie sous le nom CHAMPC7, par exemple.
- sauvegardez le problème sous le nom CHAMPC7, par exemple, quand les réponses aux questions ci-dessous auront été trouvées.

**?**

Ecrivez l'équation du modèle estimé accompagné de l'écart-type résiduel, AIC et BIC. Interprétez les résultats de l'estimation.

1.7.1 Votre réponse

**?**

L'analyse des résidus est-elle favorable au modèle ? Notez le critère MAPE pour les prévisions.

1.7.2 Votre réponse

**?**

Remarquez-vous quelque chose d'anormal ?

1.7.3 Votre réponse

?

Que pouvez-vous conclure ?



Pour visualiser la série résiduelle : menu Data ⇒ Plot ⇒ Residuals.



1.7.4 Votre réponse

SYNTHÈSE

Nous avons introduit des interventions dans le modèle ARIMA obtenu au chapitre 10 pour les ventes de champagne en France. Nous avons utilisé plusieurs interventions ponctuelles ou impulsions pour corriger des données aberrantes. On les appelle aussi interventions de type AO, *additive outlier*.

L'intervention de type TC, *temporary change*, que nous avons considérée en janvier 1968 est plus spécifique que ce qui est habituellement pratiqué en régression multiple à l'aide de variables binaires ou muettes, *dummy variables*. Ceci explique le changement de terminologie. Les portées des ces interventions de type TC est souvent infinie mais elle peut être limitée dans Time Series Expert. Ces interventions ont des cas particuliers importants : (a) les interventions en niveau ou de type LS, *level shift*, pour lesquelles le paramètre δ vaut 0, de portée finie ou infinie, selon le cas;

(b) les interventions de rattrapage ou de compensation pour lesquelles le paramètre δ vaut -1 et la portée est d'une seule période.

Nous en verrons des applications dans d'autres exemples.

Partie 2 **2.1 INTRODUCTION**

Dans le chapitre 5 du cours, nous avons utilisé la méthode Census II X-11 pour effectuer la décomposition saisonnière de la série. Dans le chapitre 12, nous allons traiter une version plus récente de cette méthode, appelée X-12-ARIMA. Pour cela, nous devons aborder la modélisation ARIMA complétée par diverses corrections pour données aberrantes. Dans la partie 4, nous ajouterons les corrections pour jours ouvrables et pour l'effet de la fête de Pâques.

C'est le module RegARIMA de la méthode X-12-ARIMA qui est chargé de cette tâche. Il peut réaliser ceci de manière **automatique mais nous ne considérons pas ceci dans cet exercice**. La première étape va consister à reprendre le modèle du chapitre 10, exercice 6. Dans une seconde étape, la spécification du modèle RegARIMA ne tient pas compte des effets des jours ouvrables ni de Pâques, seulement des corrections pour données aberrantes. C'est dans la partie 4 de l'exercice que nous ajouterons des variables de régression pour tenir compte des jours ouvrables et de l'effet de Pâques.

2.2 DÉMARRAGE D'UN PROJET

- ⇒ Afin de lancer le logiciel Demetra, suivez les instructions données en annexe de l'introduction du cours. Demetra est un logiciel sous Windows. Il requiert néanmoins quelques précautions.
- ⇒ En particulier, relisez
 - comment utiliser les menus
 - comment manipuler les projets et les modèles.

Nous devons d'abord récupérer un projet préparé pour Demetra.

- ⇒ Pour récupérer un projet préparé pour Demetra, menu Project ⇒ Open. Dans la liste déroulante intitulée Files of type, choisissez Demetra detailed analysis projects *.dm2. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPX12.

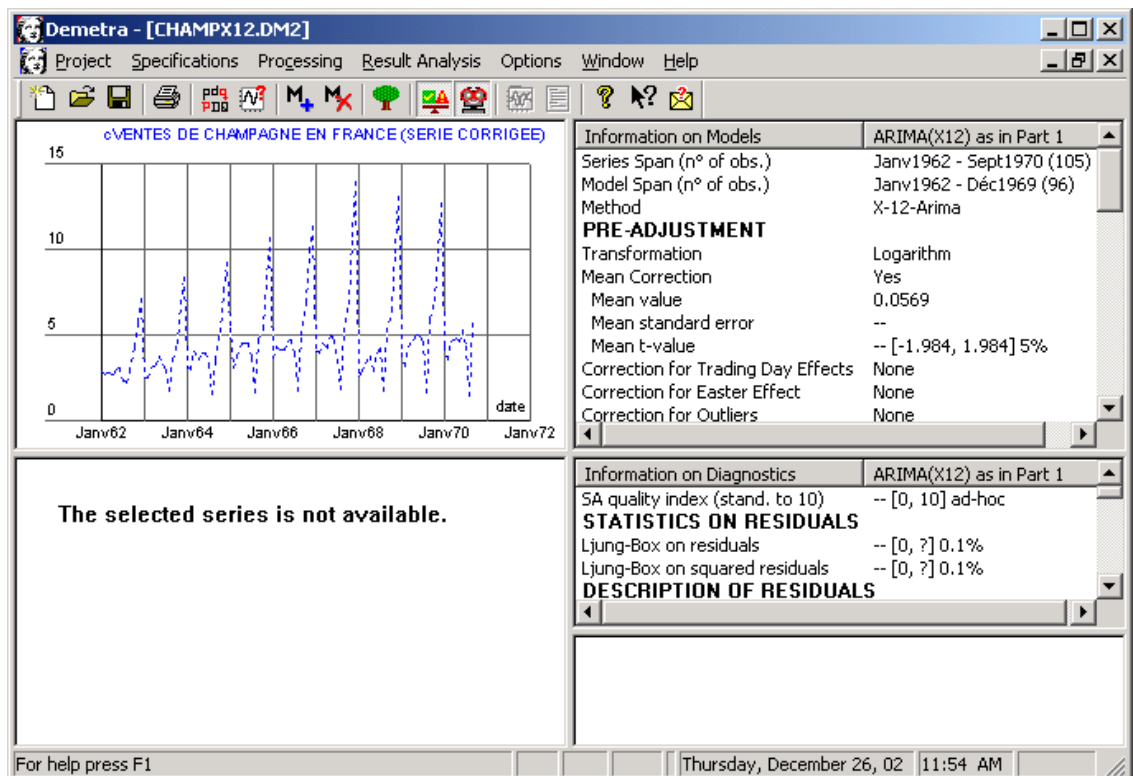
Une fenêtre de Demetra apparaît avec les menus : Projet, Specifications, Processing, Result analysis, Options, Window, Help.

- ⇒ Maximisez la fenêtre CHAMPX12.dm2.

**Remarque**

Il faut souvent maximiser les fenêtres de Demetra. Nous n'insisterons pas chaque fois sur ce détail.

On voit le graphique des données dans la partie supérieure gauche de la fenêtre. Ce projet comporte un modèle que nous avons appelé 'ARIMA(X12) as in Part 1' et qu'on voit en partie dans la partie supérieure droite de la fenêtre.



Nous pouvons parcourir les fenêtres des spécifications détaillées, sans rien modifier. À cette fin, procédez comme suit. Il faut distinguer les spécifications du modèle et le traitement statistique.

Le modèle comporte la même spécification ARIMA qu'au début de la partie 1, qui est aussi la même que dans l'exercice 3 du chapitre 10.

Avant de lancer l'estimation et d'interpréter les résultats, nous pouvons parcourir les fenêtres des spécifications détaillées, sans rien modifier. À cette fin, procédez comme suit. Il faut distinguer les spécifications du modèle et le traitement statistique.



Pour examiner les spécifications du modèle, menu Specifications ⇒ Modelling. Cliquez sur le nom du modèle, ici ARIMA(X12) as in Part 1.

La fenêtre qui s'ouvre possède cinq onglets.



Cliquez tour à tour sur chacun d'eux pour vérifier ce qui suit.

1. "Data Handling" décrit les transformations, les interpolations pour données manquantes, l'utilisation d'une constante dans le modèle.
2. "Regression Variables" décrit les corrections pour jours ouvrables (trading day), les effets de la fête de Pâques (Easter effect) et le traitement des données aberrantes.
3. "ARIMA Model Specification" permet de spécifier un modèle ARIMA ou SARIMA : différences ordinaire et saisonnière, nombres de termes autorégressifs ordinaire et saisonnier et nombres de termes moyenne mobile ordinaire et saisonnière. On doit ici indiquer qu'on veut estimer les paramètres et fournir éventuellement des valeurs initiales.
4. "Automatic Model Identification/Selection" permet de décrire les paramètres d'une méthode de spécification automatique.
5. "Model Estimation" permet de choisir la méthode d'estimation des paramètres et la période d'estimation.

Le modèle est spécifié dans la fenêtre ARIMA Model Specification qui se présente comme suit :

On voit bien que nous avons choisi un coefficient moyenne mobile saisonnier d'ordre 1, donc de retard 12, et un coefficient autorégressif ordinaire, d'ordre 1.

Pour le reste, nous avons choisi une transformation logarithmique et

l'utilisation d'une constante dans le modèle, mais nous avons supprimé les interpolations pour données manquantes, les corrections pour jours ouvrables, les effets de la fête de Pâques, le traitement des données aberrantes et la méthode de spécification automatique. En outre la méthode d'estimation choisie est Exact maximum likelihood.



Vérifiez tout ce qui a été indiqué ci-dessus.

2.2.1 Votre réponse



Cliquez sur le bouton Cancel pour annuler les modifications éventuelles faites par erreur.



Pour examiner les spécifications du traitement statistique, menu Specifications \Rightarrow Statistical Treatment. Cliquez sur le nom du modèle, ici ARIMA(X12) as in Part 1.

La fenêtre qui s'ouvre possède quatre onglets.



Cliquez tour à tour sur chacun d'eux pour vérifier ce qui suit.

1. "Decomposition" avec une partie 'ARIMA-Model Based Method' pour le module TRAMO/SEATS et une partie 'Moving Averages Based Method "X11"' pour la méthode X-12-ARIMA.

2. "Forecasting" pour la détermination des prévisions.

3. "Sliding Spans Analysis".

4. "Revision History Analysis".

Nous avons choisi de cocher sur le bouton Produce and use forecasts, choisi 9 comme horizon, spécifié Out-of-sample forecasts within the time series length, et coché le bouton Produce confidence intervals, avec la probabilité par défaut de 95 %.

Nous avons supprimé la décomposition saisonnière en décochant la case Perform Decomposition. Nous ne discuterons pas ces deux dernières fenêtres avant le chapitre 12.



Vérifiez tout ce qui a été indiqué ci-dessus.

2.2.2 Votre réponse

- ⇒ Cliquez sur le bouton Cancel pour annuler les modifications éventuelles faites par erreur.
- ⇒ Pour choisir le niveau de détail de la partie de fenêtre principale consacrée aux modèles qui correspond à ce que verrons dans la suite, dans le menu Result Analysis ⇒ Information on Models ⇒ Complete List.
- ⇒ Pour choisir le niveau de détail de la partie de fenêtre principale consacrée aux diagnostics qui correspond à ce que verrons dans la suite, dans le menu Result Analysis ⇒ Information on Diagnostics ⇒ Complete List.

2.3 ESTIMATION DES PARAMÈTRES DU MODÈLE

Pour estimer les paramètres du modèle et calculer les prévisions, il faut procéder comme suit.

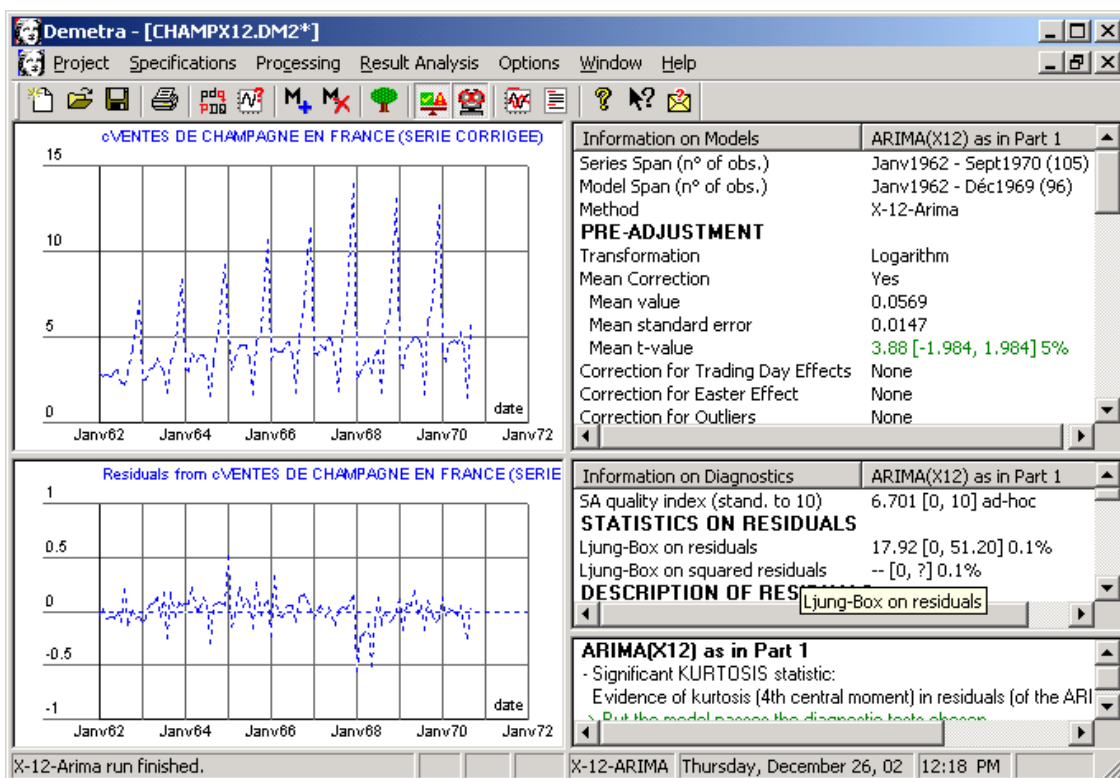
- ⇒ Pour estimer les paramètres du modèle ARIMA, menu Processing ⇒ Execute SA. Une question relative à la mise à jour des données s'affiche. Cliquez sur No. Cliquez sur le nom du modèle, ici ARIMA(X12) as in Part 1.



Indiquez ce qui a changé en première apparence.

2.3.1 Votre réponse





Voici les sorties détaillées.

Information on Models	ARIMA(X12) as in Part 1
Series Span (n° of obs.)	Janv1962 - Sept1970 (105)
Model Span (n° of obs.)	Janv1962 - Déc1969 (96)
Method	X-12-Arima
PRE-ADJUSTMENT	
Transformation	Logarithm
Mean Correction	Yes
Mean value	0.0569
Mean standard error	0.0147
Mean t-value	3.88 [-1.984, 1.984] 5%
Correction for Trading Day Effects	None
Correction for Easter Effect	None
Correction for Outliers	None
Critical t-value	Automatic
Critical t-value (LS)	Automatic
Critical t-value (TC)	Automatic
Corr. for Missing Obs.	None
Corr. for Other Regr. Effects	None
Specif. of the ARIMA model	(1 0 0)(0 1 1) (fixed)
Non-seas. AR (lag 1) value	-0.2809
Non-seas. AR (lag 1) standard error	0.1030
Non-seas. AR (lag 1) t-value	2.73 [-1.984, 1.984] 5%
Seasonal MA (lag 12) value	-0.4692
Seasonal MA (lag 12) standard error	0.0882
Seasonal MA (lag 12) t-value	5.32 [-1.984, 1.984] 5%
Method of Estimation	Exact Maximum Likelihood
DECOMPOSITION	
X-11 Decomposition	None

**Remarques**

1. Pour visualiser ces résultats et les imprimer, plaçant le pointeur de la souris dans la fenêtre en haut à droite, cliquez sur le bouton droit de la souris et choisissez Print Preview

2. On peut aussi obtenir ces sorties détaillées comme suit. Dans la fenêtre en haut à droite, cliquez sur le bouton droit de la souris et choisissez Copy Table Image. Vous pouvez coller l'image dans un logiciel de traitement de texte. Alternativement Copy Table Text permet de copier le texte sans le formatage.

On peut faire de même pour les informations sur les diagnostics. Nous ne les reprenons pas ici.



Comparez les résultats avec ceux qui avaient été obtenus dans l'exercice 6 du chapitre 10.

2.3.2 Votre réponse**Remarque**

1. On peut obtenir des sorties encore plus détaillées en choisissant Result Analysis puis Show Log File mais ce n'est pas intéressant à ce stade parce qu'il ne contient pas beaucoup plus d'information intéressante.

2. Nous n'examinons pas les prévisions ici parce qu'elles ne sont pas relatives à la période qui nous intéresse, allant de janvier à septembre 1970, mais à la période commençant en octobre 1970. Il ne semble pas possible de les obtenir au moyen du programme X-12-ARIMA utilisé dans Demetra.

2.4 AJOUT AUTOMATIQUE D'INTERVENTIONS

Nous allons compléter la modélisation ARIMA par une analyse d'interventions, comme dans la partie 1 de l'exercice. Nous allons laisser le soin au module RegARIMA de Demetra de choisir les interventions les plus appropriées.

Nous repartons du même modèle qu'au paragraphe 2.1.

- ⇒ Si vous avez interrompu la session, et seulement dans ce cas, reprenez la récupération du projet préparé pour Demetra, menu Project ⇒ Open. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPX12.

Nous allons créer un nouveau modèle à partir de celui-ci et nous allons le nommer Interventions(X12). La procédure est la suivante.

- ⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici ARIMA(X12) as in Part 1.
- ⇒ Dans la fenêtre Save as qui s'ouvre, naviguez jusqu'au dossier DATA du cours (pas celui de Demetra!) puis CHAP11 puis CH11EX02 et changez le nom de fichier dataARIMA(X12) as in Part 1 en Intervention(X12). Cliquez sur le bouton Save. Le modèle est sauvegardé sous un nouveau nom.
- ⇒ Pour relire le modèle dans Demetra à partir du fichier, menu Specifications ⇒ Add New Model. Dans la zone Seasonal adjustment method, sélectionnez X-12-ARIMA. et cliquez sur le bouton Model from File. Cliquez sur OK.
- ⇒ Dans la fenêtre Open qui s'ouvre, sélectionnez le fichier Intervention(X12).mdl. Cliquez sur le bouton Open. Le modèle est chargé et apparaît dans une nouvelle colonne dans la partie en haut à droite de la fenêtre (il y avait la colonne "ARIMA(X12) as in Part 1, la nouvelle colonne porte le titre du nouveau modèle "Intervention(X12). Pour le voir, modifiez la taille de la fenêtre, éventuellement les largeurs des colonnes ou déplacez l'ascenseur vers la droite.

Remarque



Il est aussi possible de copier un modèle vers un autre mais c'est plus hasardeux car le modèle sélectionné n'est pas contrasté et la copie ne se fait pas où on l'attendrait. Ce procédé ne permet pas non plus d'interrompre un exercice. Pour copier, placez le pointeur dans la colonne du modèle à copier, employez le bouton droit de la souris et sélectionnez Copy. Pour coller, employez le bouton droit de la souris et sélectionnez Paste.

Nous allons maintenant compléter le modèle avant de relancer l'estimation des paramètres.

⇒ Pour modifier les spécifications du modèle, menu Specifications
⇒ Modelling. Cliquez sur le nom du modèle, ici Intervention(X12).

Il n'est pas nécessaire de modifier les spécifications des autres fenêtres que celle d'onglet "Regression Variables". Nous ne changeons pas l'absence de sélection des corrections pour jours ouvrables (trading day) ni des effets de la fête de Pâques (Easter effect) mais bien le traitement des données aberrantes.

⇒ Cliquez sur l'onglet Regression Variables. Dans la liste déroulante Automatic Outlier Detection and Correction, sélectionnez Additive outliers, transitory changes.

Malheureusement Demetra part du principe qu'un modèle estimé ne doit plus être modifié. Il faut donc *chaque fois* effectuer les modifications suivantes.

⇒ Cliquez sur l'onglet ARIMA Model Specification. Décochez les deux coefficients de modèles et ensuite, dans la zone Estimation of unfixed coefficients, choisissez With initial values computed by SA method (au lieu de No estimation, values given by the user or by defaults) SANS QUOI LES VALEURS DES COEFFICIENTS RESTERONT FIXES. Cliquez sur OK.

Remarque



Une manière plus rapide consiste à placer le pointeur de la souris dans le corps de la colonne 3 intitulée Intervention(X12) et de cliquer sur le bouton droit de la souris. Choisissez Model parameters puis Unfix coefficients.

Avant d'estimer les paramètres du modèle, nous allons sauvegarder celui-ci sous son nom d'origine.

⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici Intervention(X12). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA du cours, puis CHAP11 puis CH11EX02 et sélectionnez le fichier Intervention(X12).mdl. Cliquez sur le bouton Save.

⇒ Pour estimer les paramètres du modèle ARIMA et les interventions, menu Processing ⇒ Execute SA.



Cliquez sur le nom du modèle, ici Intervention(X12) et refusez la mise à jour des données.

La fenêtre du haut à droite, colonne intitulée Intervention(X12), est maintenant mise à jour et certaines lignes sont complétées, notamment celles relatives à Correction for outliers. Pour la voir, il faut éventuellement déplacer la barre de défilement vers la droite. Pour faciliter la lecture des résultats, signalons que AO signifie Additive Outlier, ce qui équivaut aux interventions de type impulsion, ou Pulse, de la partie 1 et TC correspond à Transitory Change. Notons aussi que ces impulsions sont appliquées sur la série transformée en logarithmes, mais avant l'application des différences.



Comparez les résultats avec ceux de la colonne précédente. Qu'est-ce qui est nouveau ? Comparez les interventions trouvées par rapport à celles introduites manuellement dans la partie 1.



2.4.1 Votre réponse

Demetra fournit la représentation des données et des résidus. Il permet de représenter des graphiques pour toutes les séries intermédiaires et finales produites par le programme en représentant une ou plusieurs de ces séries sur le même graphique. Parmi celles-ci, il est possible d'obtenir l'effet des interventions. Cela peut se faire comme suit.



Pour représenter une série, menu Result Analysis ⇒ Graphical Comparison Tool. Pour créer un graphique dans une des 4 cases de la nouvelle fenêtre qui s'ouvre alors, cliquez deux fois dans la case. Pour développer partiellement l'arborescence de séries, cliquez sur le signe '+' de Preliminary Series puis sur le signe '+' de Factors puis sur le signe '+' de Pre-adjustment Factors puis sur le signe '+' de Total Outlier Effect et enfin sur le nom du modèle, "2: Intervention(X12)". Cliquez sur le bouton Add to Plot. Pour achever le graphique, cliquez sur le bouton OK.

**Remarques**

1. On peut créer un graphique dans chacune des 4 cases de la nouvelle fenêtre. On peut ajouter d'autres fenêtres, si nécessaire. Il est recommandé d'agrandir la fenêtre pour éviter que les graphiques soient trop petits. On peut déplacer les séparations entre les cases.
2. Chaque graphique peut être traité séparément. En cliquant sur le bouton droit de la souris, on peut notamment copier le graphique dans le presse-papiers de Windows, imprimer le graphique ou en exporter les valeurs. On peut aussi changer l'aspect des courbes et réaliser un agrandissement sur une partie de l'intervalle de temps.
3. Pour développer partiellement l'arborescence de séries, cliquez sur les signes '+' — et ensuite, éventuellement, sur '-' pour refermer les branches non désirées — jusqu'à ce que les modèles apparaissent précédés d'une lettre M en rouge et d'un numéro. Cliquez sur le nom désiré puis sur le bouton Add to Plot. Pour rajouter une autre série, procédez de même, sans oublier de cliquer sur Add to Plot. On peut aussi retirer une série. Pour achever le graphique, cliquez sur le bouton OK.

Le graphique réalisé reprend les effets des quatre interventions introduites par le module RegARIMA.

Dans l'exercice 6 du chapitre 5, partie 3, nous disions ce qui suit :

« Deux dates particulières sont indiquées. Rappelez-vous que les données concernent la France. Sachez que le mois de mai 1968 a connu des grèves d'étudiants gauchistes (11 mai, nuit des barricades, 400 blessés, 188 véhicules endommagés) suivies de grèves générales qui commençaient à dégénérer en insurrection, au point que le Président de la République, le Général Charles de Gaulle, a dû menacer de ramener l'armée française stationnée à l'époque en Allemagne (à cause de la guerre froide) pour calmer la situation. Il a déclenché des élections législatives qui ont donné à la fin de juin 1968 une majorité écrasante au parti gaulliste.

Le mois de janvier 1968 correspond à l'introduction de la taxe à la valeur ajoutée en France. Auparavant, il existait onze taxes dont la taxe locale qui frappait les produits chaque fois qu'ils changeaient de main. »



Voyez-vous les effets de mai et de janvier 1968 ?



2.4.2 Votre opinion

Si vous arrêtez ici, il faut sauvegarder le projet sous un nom quelconque, *mais pas CHAMPX12, ni CHAMPTS*



Pour sauvegarder le projet préparé, menu Project ⇒ Save As.
Donnez un nom au projet, par exemple PARTIE2.



Remarque

Quand on sauve le projet, il n'est plus possible d'accéder aux résultats détaillés ni de réaliser des graphiques, à moins d'exécuter de nouveau le programme. Quand Demetra s'arrête, il est souvent possible de récupérer le projet traité. Ceci réduit l'intérêt de sauver fréquemment le projet en cours.

SYNTHÈSE

Dans cette partie, nous avons repris l'étude de la série des ventes de champagne à l'aide d'un modèle ARIMA et avec une analyse d'interventions. Nous avons cette fois employé le module RegARIMA de Demetra qui permet une détection automatique des interventions appropriées dans la méthode du logiciel. Les résultats sont conformes à l'attente puisque la méthode a permis de détecter les observations aberrantes liées aux grèves de mai 1968 et à l'introduction de la TVA en janvier 1968. Dans la partie 4, nous introduirons l'effet des jours ouvrables et l'effet de Pâques.

Annexe

Résultats de l'estimation du modèle final de l'exercice 6 du chapitre 10

```

=== MODEL DESCRIPTION
- SEASONAL PERIOD
- NORMALIZED BOX COX TRANSFORMATION
- DIFFERENCE
- ADDITIVE CONSTANT
- ARMA MODEL
  AUTOREGRESSIVE POLYNOMIAL
  MOVING AVERAGE POLYNOMIAL
NON LINEAR ESTIMATION:
ITER SUM OF SQ AR 1 SMA 1
0 48.79 .000 .000
1 40.02 .219 .356
2 39.56 .275 .439
3 39.54 .281 .460
4 39.54 .282 .464
5 39.54 .282 .466
6 39.54 .282 .466
=ITERATION STOPS - RELATIVE CHANGE IN EACH COEFFICIENT LESS THAN 1.00000E-03
FORM DEGREE/ORD PARAMETERS NUMBER
12
LOGARITHMS BOXC 1 0
SEASONAL 1
AUTOMATIC
REGULAR 1 AR nn 1
SEASONAL 1 SMA nn 1

```

```

CORRELATION MATRIX
  AR 1 SMA 1
AR 1 1.00
SMA 1 .00 1.00
FINAL VALUES OF THE PARAMETERS
      NAME      VALUE      STD ERROR      T-VALUE      95% CONFIDENCE LIMITS
      1      AR 1      .28190      .10656      2.6      6.89E-02      .49
      2      SMA 1      .46576      .11412      4.1      .24      .69
THE FOLLOWING PARAMETERS WERE ESTIMATED SEPARATELY
      MEAN      .26290
THE FOLLOWING CONSTANTS ARE DEPENDENT PARAMETERS OF THE MODEL
      BOXCOX      4.2430
THE FOLLOWING CONSTANTS WERE INVOLVED IN THE LEAST SQUARES ESTIMATION METHOD
      ARMA      .98220
=== SUMMARY MEASURES <V>
SUM OF SQUARES :      COMPUTED = 39.5379      ADJUSTED = 38.1429
VARIANCE ESTIMATES :      BIASED = .454082      UNBIASED = .470900
TOTAL NUMBER OF PARAMETERS = 3      STANDARD DEVIATION = .686222
INFORMATION CRITERIA :      AIC = 209.237      SBIC = 220.960
=== RESIDUAL ANALYSIS WITH 84 RESIDUALS, BEGINNING AT TIME JAN1963===
MEAN = -.433058E-02 , T-STATISTIC = -.06      (FOR TESTING ZERO MEAN)
=OUTLIERS <R(OR)S>
.01-.2 % JAN1968: -2.366      MAY1968: -2.225
.2 - 1 % JAN1965: 2.100
1 - 5 % AUG1963: -1.360      FEB1966: 1.391
=SIGNIFICANT AUTOCORRELATIONS (USING BARTLETT LIMITS) <A(OR)S>
=SIGNIFICANT PARTIAL AUTOCORRELATIONS <P(OR)S>
=LJUNG-BOX PORTMANTEAU TEST STATISTICS ON RESIDUAL AUTOCORRELATIONS <L>
      ORDER D.F. STATISTIC SIGNIFICANCE
      6      4      5.44      .245
      12     10     8.57      .573
      18     16    10.86      .818
      24     22    13.56      .916
      26     24    14.07      .945
----> WRITTEN TO FILE : RCHAMPC2.DB , LENGTH = 84
=== FITTING INTERVALS AT THE 95% LEVEL, WITH LEAD TIME 1
      3 POINTS BELOW THE LOWER LIMIT (TOTAL: 72 POINTS)
      2 POINTS ABOVE THE UPPER LIMIT (TOTAL: 72 POINTS)
----> WRITTEN TO FILE : FCHAMPC2.DB , LENGTH = 72
----> WRITTEN TO FILE : FCHAMPC2.DBM , LENGTH = 72
----> WRITTEN TO FILE : FCHAMPC2.DBP , LENGTH = 72
=== FORECASTING FROM DEC1969 WITH FRESH DATA <F>
      DATE      OBSERVATION      FORECAST      ERROR      % ERROR      95% FORECAST INTERVAL
JAN1970      4.3480      3.8684      .4796      11.0      2.8175      5.3112
FEB1970      3.5640      3.5797      -.0157      .4      2.5752      4.9759
MAR1970      4.5770      4.5031      .0739      1.6      3.2364      6.2655
APR1970      4.7880      4.8361      -.0481      1.0      3.4754      6.7293
MAY1970      4.6180      4.8496      -.2316      5.0      3.4851      6.7482
JUN1970      5.3120      5.1286      .1834      3.5      3.6857      7.1365
JUL1970      4.2980      4.8061      -.5081      11.8      3.4539      6.6877
AUG1970      1.4310      1.9094      -.4784      33.4      1.3722      2.6569
SEP1970      5.8770      6.2020      -.3250      5.5      4.4571      8.6301
CUMULATED ERROR :      -.8699 (= -2.2%); MEAN ERROR:      -.0967
MEAN ABSOLUTE ERROR (MAE):      .2604 (= 6.0%);
ROOT MEAN SQUARE ERROR :      .3194 (= 7.4%); MEAN SQUARE ERROR:      .102
MEAN ABSOLUTE PERCENTAGE ERROR (MAPE):      8.1%
      0 POINTS BELOW THE LOWER LIMIT (TOTAL: 9 POINTS)
      0 POINTS ABOVE THE UPPER LIMIT (TOTAL: 9 POINTS)
----> WRITTEN TO FILE : PCHAMPC2.DB , LENGTH = 9
----> WRITTEN TO FILE : PCHAMPC2.DBM , LENGTH = 9
----> WRITTEN TO FILE : PCHAMPC2.DBP , LENGTH = 9

```

Partie 3 **3.1 INTRODUCTION**

Dans le chapitre 13, nous allons utiliser une autre méthode de décomposition saisonnière que X-12-ARIMA appelée TRAMO/SEATS. Le module TRAMO de TRAMO/SEATS est l'équivalent du module RegARIMA. Le sigle TRAMO signifie Time series Regression with Arima noise, Missing observations and Outliers. TRAMO possède des caractéristiques assez semblables mais aussi quelques différences d'avec RegARIMA. Heureusement, le logiciel Demetra qui sert d'interface commune à X-12-ARIMA et à TRAMO/SEATS atténue les différences. Nous allons donc employer TRAMO pour la modélisation ARIMA complétée par diverses corrections pour données aberrantes. Dans la partie 5, nous ajouterons la correction pour les jours ouvrables et pour l'effet de la fête de Pâques, qui s'emploient presque comme RegARIMA.

Demetra peut réaliser ceci de manière **automatique mais nous ne considérons pas ceci dans cet exercice**. La première étape va consister à reprendre le modèle du chapitre 10, exercice 6. Dans une seconde étape, la spécification du modèle RegARIMA ne tient pas compte des effets des jours ouvrables ni de Pâques, seulement des corrections pour données aberrantes. Dans la partie 5, nous ajouterons des variables de régression pour tenir compte des jours ouvrables et de l'effet de Pâques.

3.2 DÉMARRAGE D'UN PROJET

- ⇒ Afin de lancer le logiciel Demetra, suivez les instructions données en annexe de l'introduction du cours. Demetra est un logiciel sous Windows. Il requiert néanmoins quelques précautions.
- ⇒ En particulier, relisez
 - comment utiliser les menus
 - comment manipuler les projets et les modèles.

Nous devons d'abord récupérer un projet préparé pour Demetra.

- ⇒ Pour récupérer un projet préparé pour Demetra, menu Project ⇒ Open. Dans la liste déroulante intitulée Files of type, choisissez Demetra detailed analysis projects *.dm2. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPTS.

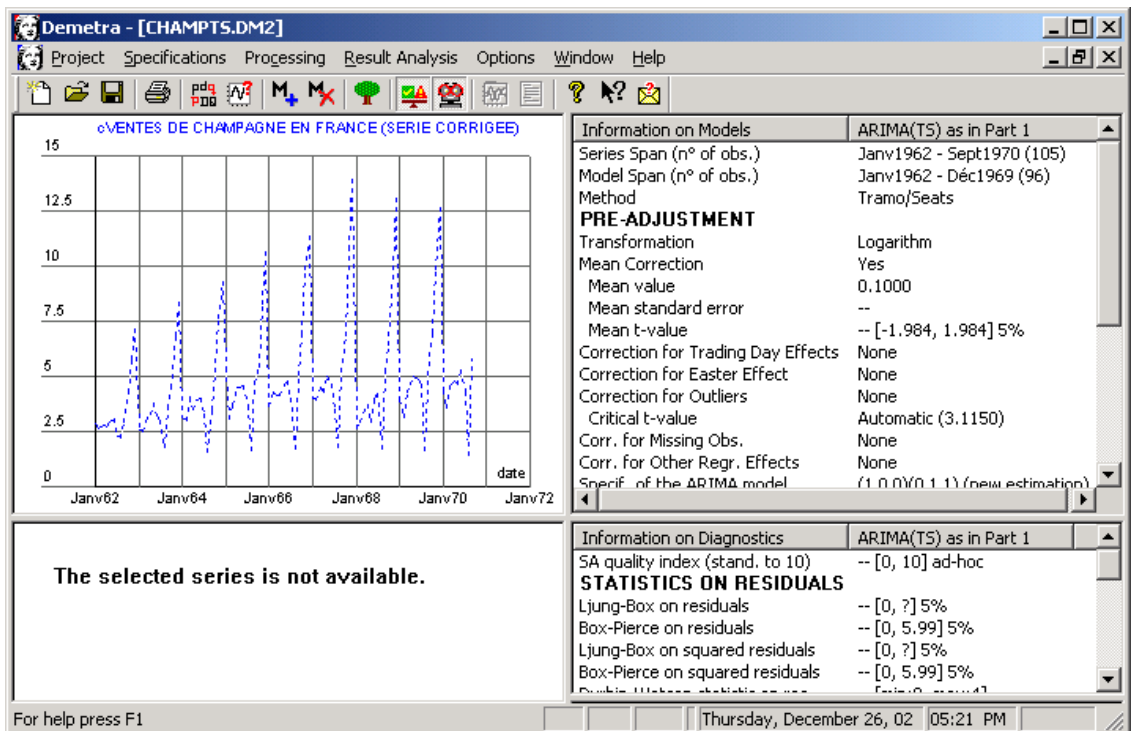
Nous verrons que CHAMPTS est pratiquement identique au projet CHAMPX12 de la partie 2. Une fenêtre de Demetra apparaît avec les menus : Projet, Specifications, Processing, Result analysis, Options, Window,

Help.



Maximisez la fenêtre CHAMPTS.dm2.

On voit le graphique des données dans la partie supérieure gauche de la fenêtre. Ce projet comporte un modèle que nous avons appelé 'ARIMA(TS) as in Part 1' et qu'on voit en partie dans la partie supérieure droite de la fenêtre. On peut le comparer avec le modèle obtenu dans la partie 2, 'ARIMA(X12) as in Part 1'.



Observez les différences.

3.2.1 Votre réponse



Remarque



Il est en principe possible de convertir le modèle 'ARIMA(X12) as in Part 1' pour TRAMO/SEATS. Il suffirait de placer le pointeur de la souris dans la colonne 'ARIMA(X12) as in Part 1' du projet CHAMPX12, de cliquer sur le bouton droit et de sélectionner Model parameters puis Convert to TRAMO/SEATS. Notons que le nombre d'observations utilisées dans la modélisation (Model Span) n'est apparemment pas égal à 96 mais à 105, alors que seules les 96 premières observations sont employées.

Le modèle comporte la même spécification ARIMA qu'au début des parties 1 et 2, qui est aussi la même que dans l'exercice 3 du chapitre 10.

Avant de lancer l'estimation et d'interpréter les résultats, nous pouvons parcourir les fenêtres des spécifications détaillées, sans rien modifier. À cette fin, procédez comme suit. Il faut distinguer les spécifications du modèle et le traitement statistique.

⇒ Pour examiner les spécifications du modèle, menu Specifications ⇒ Modelling. Cliquez sur le nom du modèle, ici ARIMA(TS) as in Part 1.

La fenêtre qui s'ouvre possède cinq onglets.

⇒ Cliquez tour à tour sur chacun d'eux pour vérifier ce qui suit.

1. "Data Handling" décrit les transformations, les interpolations pour données manquantes, l'utilisation d'une constante dans le modèle.
2. "Regression Variables" décrit les corrections pour jours ouvrables (trading day), les effets de la fête de Pâques (Easter effect) et le traitement des données aberrantes.
3. "ARIMA Model Specification" permet de spécifier un modèle ARIMA ou SARIMA : différences ordinaire et saisonnière, nombres de termes autorégressifs ordinaire et saisonnier et nombres de termes moyenne mobile ordinaire et saisonnière. On doit ici indiquer qu'on veut estimer les paramètres et fournir éventuellement des valeurs initiales.
4. "Automatic Model Identification/Selection" permet de décrire les paramètres d'une méthode de spécification automatique.
5. "Model Estimation" permet de choisir la méthode d'estimation des paramètres et la période d'estimation.

Pour le reste, nous avons choisi une transformation logarithmique et l'utilisation d'une constante dans le modèle, mais nous avons supprimé les interpolations pour données manquantes, les corrections pour jours ouvrables, les effets de la fête de Pâques, le traitement des données aberrantes et la méthode de spécification automatique. En outre la méthode d'estimation choisie est Exact maximum likelihood.



Vérifiez ce qui a été indiqué ci-dessus.



3.2.2 Votre réponse

Le modèle est spécifié dans la fenêtre ARIMA Model Specification qui se présente comme suit :

On voit bien que nous avons choisi un coefficient moyenne mobile saisonnier d'ordre 1, donc de retard 12, et un coefficient autorégressif ordinaire, d'ordre 1.

- ⇒ Cliquez sur le bouton Cancel pour annuler les modifications éventuelles faites par erreur.
- ⇒ Pour examiner les spécifications du traitement statistique, menu Specifications ⇒ Statistical Treatment. Cliquez sur le nom du modèle, ici ARIMA(X12) as in Part 1.

La fenêtre qui s'ouvre possède quatre onglets.

- ⇒ Cliquez tour à tour sur chacun d'eux pour vérifier ce qui suit.

1. "Decomposition" avec une partie 'ARIMA-Model Based Method' pour le module TRAMO/SEATS et une partie 'Moving Averages Based Method "X11"' pour la méthode X-12-ARIMA.
2. "Forecasting" pour la détermination des prévisions.
3. "Sliding Spans Analysis".

4. "Revision History Analysis".

Sur l'onglet "Forecasting", nous avons choisi de cocher l'option "Produce (and use) forecasts", choisi 9 comme horizon de prévision, spécifié "Out-of-sample forecasts within the time series length", et coché le bouton "Produce confidence intervals", avec la probabilité par défaut de 95 %.

**Remarque**

Contrairement au module RegARIMA de X-12-ARIMA, qui calculait des prévisions au delà de la fin de la série, les prévisions seront ici calculées sur les 9 premiers mois de l'année 1970 avec décembre 1969 comme origine des prévisions.

Nous avons supprimé la décomposition saisonnière en décochant la case Perform Decomposition. Nous ne discuterons pas ces deux dernières fenêtres avant le chapitre 12.



Vérifiez ce qui a été indiqué ci-dessus.

3.2.3 Votre réponse



Cliquez sur le bouton Cancel pour annuler les modifications éventuelles faites par erreur.

3.3 ESTIMATION DES PARAMÈTRES DU MODÈLE

Pour estimer les paramètres du modèle et calculer les prévisions, il faut procéder comme suit.



Pour estimer les paramètres du modèle ARIMA, menu Processing ⇒ Execute SA. Cliquez "No" sur le message intermédiaire vous demandant de mettre la série à jour puis cliquez sur le nom du modèle, ici ARIMA(TS) as in Part 1.



Indiquez ce qui a changé en première apparence.



3.3.1 Votre réponse

Voici les sorties détaillées qui avaient été obtenues pour le modèle ARIMA(X12) as in Part 1.

Information on Models	ARIMA(X12) as in Part 1
Series Span (n° of obs.)	Janv1962 - Sept1970 (105)
Model Span (n° of obs.)	Janv1962 - Déc1969 (96)
Method	X-12-Arima
PRE-ADJUSTMENT	
Transformation	Logarithm
Mean Correction	Yes
Mean value	0.0569
Mean standard error	0.0147
Mean t-value	3.88 [-1.984, 1.984] 5%
Correction for Trading Day Effects	None
Correction for Easter Effect	None
Correction for Outliers	None
Critical t-value	Automatic
Critical t-value (LS)	Automatic
Critical t-value (TC)	Automatic
Corr. for Missing Obs.	None
Corr. for Other Regr. Effects	None
Specif. of the ARIMA model	(1 0 0)(0 1 1) (fixed)
Non-seas. AR (lag 1) value	-0.2809
Non-seas. AR (lag 1) standard error	0.1030
Non-seas. AR (lag 1) t-value	2.73 [-1.984, 1.984] 5%
Seasonal MA (lag 12) value	-0.4692
Seasonal MA (lag 12) standard error	0.0882
Seasonal MA (lag 12) t-value	5.32 [-1.984, 1.984] 5%
Method of Estimation	Exact Maximum Likelihood
DECOMPOSITION	
X-11 Decomposition	None



Comparez les résultats du module TRAMO avec ceux qui ont été obtenus avec le module RegARIMA dans la partie 2.



3.3.2 Votre réponse



Comparez les résultats du module TRAMO avec ceux qui ont été obtenus avec Time Series Expert dans la partie 1.



3.3.3 Votre réponse



Pour avoir des sorties plus détaillées, menu Result Analysis ⇒ Show Log File. Cliquez sur le nom du modèle, ici ARIMA(TS) as in Part 1. Descendez dans la rubrique Forecasts.

Les données et les prévisions qui sont fournies sont transformées sauf les prévisions de la dernière colonne. Comparez avec les prévisions que voici fournies par Time Series Expert, dans le chapitre 10, exercice 3 :

```

=== FORECASTING FROM DEC1969 WITH FRESH DATA <F>
  DATE      OBSERVATION      FORECAST      ERROR      % ERROR      95% FORECAST INTERVAL
JAN1970      4. 3480      3. 8684      . 4796      11. 0      2. 8175      5. 3112
FEB1970      3. 5640      3. 5797      -. 0157      . 4      2. 5752      4. 9759
MAR1970      4. 5770      4. 5031      . 0739      1. 6      3. 2364      6. 2655
APR1970      4. 7880      4. 8361      -. 0481      1. 0      3. 4754      6. 7293
MAY1970      4. 6180      4. 8496      -. 2316      5. 0      3. 4851      6. 7482
JUN1970      5. 3120      5. 1286      . 1834      3. 5      3. 6857      7. 1365
JUL1970      4. 2980      4. 8061      -. 5081      11. 8      3. 4539      6. 6877
AUG1970      1. 4310      1. 9094      -. 4784      33. 4      1. 3722      2. 6569
SEP1970      5. 8770      6. 2020      -. 3250      5. 5      4. 4571      8. 6301
CUMULATED ERROR      :      -. 8699 (= -2. 2%); MEAN ERROR:      -. 0967
MEAN ABSOLUTE ERROR (MAE):      . 2604 (= 6. 0%);
ROOT MEAN SQUARE ERROR      :      . 3194 (= 7. 4%); MEAN SQUARE ERROR:      . 102
MEAN ABSOLUTE PERCENTAGE ERROR (MAPE):      8. 1%
      0 POINTS BELOW THE LOWER LIMIT (TOTAL:      9 POINTS)
      0 POINTS ABOVE THE UPPER LIMIT (TOTAL:      9 POINTS)

```



Comparez les prévisions obtenues par les deux programmes.



3.3.4 Votre réponse

3.4 AJOUT AUTOMATIQUE D'INTERVENTIONS

Nous allons compléter la modélisation ARIMA par une analyse d'interventions, comme dans la partie 1 de l'exercice. Nous allons laisser le soin au module TRAMO/SEATS de Demetra de choisir les interventions les plus appropriées.

Nous repartons du même modèle qu'au paragraphe 2.1.

⇒ Si vous avez interrompu la session, et seulement dans ce cas, reprenez la récupération du projet préparé pour Demetra, menu Project ⇒ Open. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPTS.

Nous allons créer un nouveau modèle à partir de celui-ci et nous allons le nommer Interventions(TS). La procédure est la suivante.

- ⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici ARIMA(TS) as in Part 1.
- ⇒ Dans la fenêtre Save as qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et tapez le nom de fichier Intervention(TS) à la place de ARIMA(TS) as in Part 1. Cliquez sur le bouton Save. Le modèle est sauvegardé sous un nouveau nom.
- ⇒ Pour relire le modèle dans Demetra à partir du fichier, menu Specifications ⇒ Add New Model. Dans la zone Seasonal adjustment method, sélectionnez TRAMO/SEATS et cliquez sur le bouton Model from File. Cliquez sur OK.
- ⇒ Dans la fenêtre Open qui s'ouvre, sélectionnez le fichier Intervention(TS).mdl. Cliquez sur le bouton Open. Le modèle est chargé et apparaît dans une nouvelle colonne dans la partie en haut à droite de la fenêtre. Pour le voir, modifiez la taille de la fenêtre, éventuellement les largeurs des colonnes ou déplacez l'ascenseur vers la droite.

Nous allons maintenant compléter le modèle avant de relancer l'estimation des paramètres.

- ⇒ Pour modifier les spécifications du modèle, menu Specifications ⇒ Modelling. Cliquez "No" sur le message intermédiaire vous demandant de mettre la série à jour puis cliquez sur le nom du modèle, ici Intervention(TS).

Il n'est pas nécessaire de modifier les spécifications des autres fenêtres que celle d'onglet "Regression Variables". Nous ne changeons pas l'absence de sélection des corrections pour jours ouvrables (trading day) ni des effets de la fête de Pâques (Easter effect) mais bien le traitement des données aberrantes.



⇒ Cliquez sur l'onglet Regression Variables. Dans la liste déroulante Automatic Outlier Detection and Correction, sélectionnez "Additive outliers , trans. changes". La zone Interval for outlier correction doit contenir les dates de début et de fin, ici Jan 1962 to Dec 1969.

Remarque

A cause d'un bogue du programme, il faut fréquemment redéfinir cette zone.

Demetra part du principe qu'un modèle estimé ne doit plus être modifié. Il faut donc *chaque fois* effectuer les modifications suivantes.

⇒ Cliquez sur l'onglet ARIMA Model Specification. Décochez les deux coefficients de modèles et ensuite, dans la zone Estimation of unfixed coefficients, choisissez With initial values computed by SA method (au lieu de No estimation, values given by the user or by defaults) SANS QUOI LES VALEURS DES COEFFICIENTS RESTERONT FIXES. Cliquez sur OK.

Nous allons sauvegarder le modèle terminé sous son nom d'origine avant d'estimer ses paramètres.

- ⇒ Revenez à la fenêtre principale du projet (terminée par :1).
- ⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici Intervention(TS). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et sélectionnez le fichier Intervention(TS).mdl. Cliquez sur le bouton Save.
- ⇒ Pour estimer les paramètres du modèle ARIMA et les interventions, menu Processing ⇒ Execute SA. Cliquez sur le nom du modèle, ici Intervention(TS).

La fenêtre du haut à droite, colonne intitulée Intervention(TS), est maintenant mise à jour et certaines lignes sont complétées, notamment celles relatives à Correction for outliers. Pour la voir, il faut éventuellement déplacer la barre de défilement vers la droite. Pour faciliter la lecture des résultats, signalons que AO signifie Additive Outlier, ce qui équivaut aux interven-

tions de type ponctuelle (ou impulsion, *Pulse*) de la partie 1 et que TC signifie Transitory Change. Notons aussi que ces impulsions sont appliquées sur la série transformée en logarithmes, mais avant l'application des différences.



Comparez les résultats avec ceux de la colonne précédente. Qu'est-ce qui est nouveau ? Comparez les interventions trouvées par rapport à celles introduites manuellement dans la partie 1 ou celles de la partie 2.



3.4.1 Votre réponse

Demetra fournit la représentation des données et des résidus. Il permet de représenter des graphiques pour toutes les séries intermédiaires et finales produites par le programme en représentant une ou plusieurs de ces séries sur le même graphique. Parmi celles-ci, il est possible d'obtenir l'effet des interventions. Cela peut se faire comme suit.



Pour représenter une série, menu Result Analysis ⇒ Graphical Comparison Tool. Pour créer un graphique dans une des 4 cases de la nouvelle fenêtre qui s'ouvre alors, cliquez deux fois dans la case. Pour développer partiellement l'arborescence de séries, cliquez sur le signe '+' de Preliminary Series puis sur le signe '+' de Components puis sur le signe '+' de Pre-adjusted Components puis sur le signe '+' de Total Outlier Effect et enfin sur le nom du modèle, "2: Intervention(TS)". Cliquez sur le bouton Add to Plot. Pour achever le graphique, cliquez sur le bouton OK.

Remarques



1. On peut créer un graphique dans chacune des 4 cases de la nouvelle fenêtre. On peut ajouter d'autres fenêtres, si nécessaire. Il est recommandé d'agrandir la fenêtre pour éviter que les graphiques soient trop petits. On peut déplacer les séparations entre les cases.

2. Chaque graphique peut être traité séparément. En cliquant sur le bouton droit de la souris, on peut notamment copier le graphique dans le presse-

papiers de Windows, imprimer le graphique ou en exporter les valeurs. On peut aussi changer l'aspect des courbes et réaliser un agrandissement sur une partie de l'intervalle de temps.

3. Pour développer partiellement l'arborescence de séries, cliquez sur les signes '+' — et ensuite, éventuellement, sur '-' pour refermer les branches non désirées — jusqu'à ce que les modèles apparaissent précédés d'une lettre M en rouge et d'un numéro. Cliquez sur le nom désiré puis sur le bouton Add to Plot. Pour rajouter une autre série, procédez de même, sans oublier de cliquer sur Add to Plot. On peut aussi retirer une série. Pour achever le graphique, cliquez sur le bouton OK.

Le graphique réalisé reprend les effets des interventions introduites par le module TRAMO.

Dans l'exercice 6 du chapitre 5, partie 3, nous disions ce qui suit.

« Deux dates particulières sont indiquées. Rappelez-vous que les données concernent la France. Sachez que le mois de mai 1968 a connu des grèves d'étudiants gauchistes (11 mai, nuit des barricades, 400 blessés, 188 véhicules endommagés) suivies de grèves générales qui commençaient à dégénérer en insurrection, au point que le Président de la République, le Général Charles de Gaulle, a dû menacer de ramener l'armée française stationnée à l'époque en Allemagne (à cause de la guerre froide) pour calmer la situation. Il a déclenché des élections législatives qui ont donné à la fin de juin 1968 une majorité écrasante au parti gaulliste.

Le mois de janvier 1968 correspond à l'introduction de la taxe à la valeur ajoutée en France. Auparavant, il existait onze taxes dont la taxe locale qui frappait les produits chaque fois qu'ils changeaient de main. »



Voyez-vous les effets de mai et de janvier 1968 ?



3.4.2 Votre opinion

Si vous arrêtez ici, il faut sauvegarder le projet sous un nom quelconque, mais pas *CHAMPTS* ni *CHAMPX12*.



Pour sauvegarder le projet préparé, menu Project ⇒ Save As.



Donnez un nom au projet, par exemple PARTIE3.

Remarque

Quand on sauve le projet, il n'est plus possible d'accéder aux résultats détaillés ni de réaliser des graphiques, à moins d'exécuter de nouveau le programme. Quand Demetra s'arrête, il est souvent possible de récupérer le projet traité. Ceci réduit l'intérêt de sauver fréquemment le projet en cours.

SYNTHÈSE

Dans cette partie, nous avons repris l'étude de la série des ventes de champagne à l'aide d'un modèle ARIMA et avec une analyse d'interventions. Nous avons cette fois employé le module TRAMO/SEATS de Demetra qui permet une détection automatique des interventions appropriées. Les résultats sont conformes à l'attente puisque la méthode a permis de détecter les observations aberrantes liées aux grèves de mai 1968 et à l'introduction de la TVA en janvier 1968. Dans l'ensemble, les résultats sont très proches de ceux obtenus par le programme X-12-ARIMA. Dans la partie 5, nous introduirons l'effet des jours ouvrables et l'effet de la fête de Pâques.



Partie 4 **4.1 INTRODUCTION**

Nous allons poursuivre l'utilisation du module RegARIMA de la méthode X-12-ARIMA commencée dans la partie 2. Nous y avons abordé la modélisation ARIMA complétée par diverses corrections pour données aberrantes. Ici, nous ajouterons les corrections pour jours ouvrables et pour l'effet de la fête de Pâques.

Si vous avez laissé Demetra ouvert et si la fenêtre pour le projet CHAMPX12 est encore disponible, il n'est pas nécessaire d'effectuer le paragraphe suivant.

4.2 DÉMARRAGE D'UN PROJET

- ⇒ Afin de lancer le logiciel Demetra, suivez les instructions données en annexe de l'introduction du cours. Demetra est un logiciel sous Windows. Il requiert néanmoins quelques précautions.
- ⇒ En particulier, relisez
 - comment utiliser les menus
 - comment manipuler les projets et les modèles.

Nous devons d'abord récupérer un projet préparé pour Demetra.

- ⇒ Pour récupérer un projet préparé pour Demetra, menu Project ⇒ Open. Dans la liste déroulante intitulée Files of type, choisissez Demetra detailed analysis projects *.dm2. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPX12.

4.3 AJOUT DE L'EFFET DES JOURS OUVRABLES ET DE L'EFFET DE PÂQUES

Nous allons d'abord créer un nouveau modèle RegARIMA(X12) à partir du modèle Interventions(X12). L'objectif est ici d'examiner l'effet des jours ouvrables sur les données ainsi que l'effet de la fête de Pâques.

On récupère le modèle et on modifie son nom.

Remarque

Il est évidemment possible de commencer un nouveau modèle à partir de rien mais il faut alors effectuer un certain nombre de modifications dans les différentes fenêtres pour retrouver les mêmes résultats qu'ici.

- ⇒ Pour relire le modèle dans Demetra à partir du fichier, menu Specifications ⇒ Add New Model. Dans la zone Seasonal adjustment method, sélectionnez X-12-ARIMA et cliquez sur le bouton Model from File. Cliquez sur OK.
- ⇒ Dans la fenêtre Open qui s'ouvre, sélectionnez le fichier Intervention(X12).mdl. Cliquez sur le bouton Open. Le modèle est chargé et apparaît dans une nouvelle colonne dans la partie en haut à droite de la fenêtre. Pour le voir, déplacez l'ascenseur vers la droite.

Attention, le modèle apparaît en double si vous n'avez pas redémarré le projet !

- ⇒ Allez sur la deuxième version, cliquez avec le bouton droit de la souris et choisissez Edit Model Name et tapez le nom RegARIMA(X12).

Nous allons maintenant compléter le modèle avant de relancer l'estimation des paramètres.

- ⇒ Pour modifier les spécifications du modèle, menu Specifications ⇒ Modelling. Cliquez sur le nom du modèle, ici RegARIMA(X12).

Il faut modifier le contenu de deux fenêtres. Commençons par celle dont l'onglet est "Regression Variables". Nous allons demander en plus les corrections pour jours ouvrables (trading day) et les effets de la fête de Pâques (Easter effect) et gardons le traitement des données aberrantes.

- ⇒ Cliquez sur l'onglet Regression Variables.
- ⇒ Dans la zone Trading Day Effect, dans la liste déroulante Trading Day Correction, sélectionnez la troisième option Trading Day #Mon – #Sun, etc.
- ⇒ Dans la zone Special Effects, dans la liste déroulante Easter Effect, choisissez Yes.

Malheureusement Demetra part du principe qu'un modèle estimé ne doit plus être modifié. Or, on peut espérer que les effets de jours ouvrables et de la fête de Pâques affectent les données et donc le modèle qui en est déduit. Il faut donc *chaque fois* effectuer les modifications suivantes.

⇒ Cliquez sur l'onglet ARIMA Model Specification. Décochez les deux coefficients du modèles SANS QUOI LES VALEURS DES COEFFICIENTS RESTERONT FIXES, et DE PLUS, POUR LA MEME RAISON, dans la liste déroulante Estimation of Unfixed Coefficients, CLIQUEZ SUR With initial values given by the user or by default.

⇒ Cliquez sur OK.

Le module X-12-ARIMA emploie 6 variables binaires pour l'effet des jours ouvrables et 1 variable pour l'effet de Pâques. Pour plus de détails, voir la partie A de l'exercice. Les paramètres correspondants seront estimés ainsi que les paramètres du modèle ARIMA et les paramètres des interventions.

⇒ Pour estimer les paramètres du modèle ARIMA et les interventions, menu Processing ⇒ Execute SA. Cliquez "No" sur le message intermédiaire vous demandant de mettre la série à jour puis cliquez sur le nom du modèle, ici RegARIMA(X12).

La fenêtre du haut à droite a maintenant une colonne RegARIMA(X12) mise à jour. Pour la voir, il faut déplacer la barre de défilement vers la droite.

? Comparez les résultats avec ceux de la colonne précédente. Qu'est-ce qui est nouveau ? Peut-on dire qu'il y a un effet de jours ouvrables ? Un effet de Pâques ?



4.3.1 Votre réponse

Nous allons produire un graphique unique comparant les trois effets étudiés ici : l'effet des jours ouvrables, l'effet de Pâques et les interventions.

⇒ Pour représenter ces trois séries, menu Result Analysis ⇒ Gra-

phical Comparison Tool. Pour créer un graphique dans une des 4 cases de la nouvelle fenêtre qui s'ouvre alors, cliquez deux fois dans la case. Pour développer partiellement l'arborescence de séries, cliquez sur le signe '+' de Preliminary Series puis sur le signe '+' de Factors puis sur le signe '+' de Pre-adjustment Factors puis sur le signe '+' de Total Trading Day Effect et enfin sur le nom du modèle, "3: RegARIMA(X12)". Cliquez sur le bouton Add to Plot. Procédez de même pour Easter Effect et Total Outlier Effect. Pour achever le graphique, cliquez sur le bouton OK.

Remarque

Voir les remarques du paragraphe 2.4.



Comparez l'intensité des trois effets. Quel est l'effet le plus apparent ?

4.3.2 Votre réponse



Si vous souhaitez réutiliser au chapitre 12 le modèle que vous avez obtenu, il faut le sauvegarder sous son nom d'origine, comme suit.

- ⇒ Revenez à la fenêtre principale du projet (terminée par :1).
- ⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici RegARIMA(X12). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et tapez le nom de fichier RegARIMA(X12).mdl. Cliquez sur le bouton Save.

Nous allons terminer en examinant en partie le fichier de sortie de Demetra.

- ⇒ Pour examiner le fichier de sortie relatif au modèle en cours d'examen, menu Result analysis ⇒ Show log file. Choisissez RegARIMA(X12).

Après un titre et une longue liste de noms de fichiers intitulée File Save Requests, on peut voir le tableau des données, ici en logarithmes et avec

une seule décimale — donc illisible —, puis les rubriques Model defined, Model estimation/evaluation, Outlier detection, Regression model et ARIMA model. Les tableaux qui suivent sont relatifs à la méthode de décomposition saisonnière, bien qu'elle ne soit pas demandée.

La rubrique Regression model présente les sorties de régression, selon la forme habituelle (variable, paramètre estimé, erreur-type, statistique t de Student). La rubrique ARIMA model reprend l'équivalent mais pour les paramètres du modèle ARIMA.



Retrouvez les éléments mentionnés ci-dessus. Commentez la significativité à 5 % des différents paramètres du modèle.



4.3.3 Votre réponse



Cliquez sur le nom du modèle, ici RegARIMA(X12). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et tapez le nom de fichier RegARIMA(X12).mdl. Cliquez sur le bouton Save.

Vous pouvez aussi sauvegarder le projet sous un nom quelconque, *mais pas CHAMPX12, ni CHAMPTS*.



Pour sauvegarder le projet préparé, menu Project \Rightarrow Save As. Donnez un nom au projet, par exemple PARTIE4.

SYNTHÈSE

Dans cette partie, nous avons repris l'étude de la série des ventes de champagne à l'aide d'un modèle ARIMA et avec une analyse d'interventions dans le module RegARIMA de Demetra. Nous avons également introduit l'effet des jours ouvrables et l'effet de la fête de Pâques mais avec moins de succès.

Partie 5 **5.1 INTRODUCTION**

Nous allons poursuivre l'utilisation du module TRAMO de TRAMO/SEATS commencée dans la partie 3. Nous y avons abordé la modélisation ARIMA complétée par diverses corrections pour données aberrantes. Ici, nous ajouterons les corrections pour jours ouvrables et pour l'effet de la fête de Pâques.

Si vous avez laissé Demetra ouvert et si la fenêtre pour le projet CHAMPTS est encore disponible, il n'est pas nécessaire d'effectuer le point suivant.

5.2 DÉMARRAGE D'UN PROJET

- ⇒ Afin de lancer le logiciel Demetra, suivez les instructions données en annexe de l'introduction du cours. Demetra est un logiciel sous Windows. Il requiert néanmoins quelques précautions.
- ⇒ En particulier, relisez
 - comment utiliser les menus
 - comment manipuler les projets et les modèles.

Nous devons d'abord récupérer un projet préparé pour Demetra.

- ⇒ Pour récupérer un projet préparé pour Demetra, menu Project ⇒ Open. Dans la liste déroulante intitulée Files of type, choisissez Demetra detailed analysis projects *.dm2. Choisissez DATA puis CHAP11 puis CH11EX02 puis CHAMPTS.

5.3 AJOUT DE L'EFFET DES JOURS OUVRABLES ET DE L'EFFET DE PÂQUES

Nous allons d'abord créer un nouveau modèle RegARIMA(TS) à partir du modèle Interventions(TS). L'objectif est ici d'examiner l'effet des jours ouvrables sur les données ainsi que l'effet de la fête de Pâques.

On récupère le modèle et on modifie son nom.

Remarques

Il est évidemment possible de commencer un nouveau modèle à partir de rien mais il faut alors effectuer un certain nombre de modifications dans les différentes fenêtres pour retrouver les mêmes résultats qu'ici.

- ⇒ Pour relire le modèle dans Demetra à partir du fichier, menu

Specifications \Rightarrow Add New Model. Dans la zone Seasonal adjustment method, sélectionnez TRAMO-SEATS. et cliquez sur le bouton Model from File. Cliquez sur OK.

\Rightarrow Dans la fenêtre Open qui s'ouvre, sélectionnez le fichier Intervention(TS).mdl. Cliquez sur le bouton Open. Le modèle est chargé et apparaît dans une nouvelle colonne dans la partie en haut à droite de la fenêtre. Pour le voir, déplacez la barre de défilement vers la droite.

Attention, le modèle apparaît en double !

\Rightarrow Allez sur la deuxième version, cliquez avec le bouton droit de la souris et choisissez Edit Model Name et tapez le nom RegARIMA(TS).

Nous allons maintenant compléter le modèle avant de relancer l'estimation des paramètres.

\Rightarrow Pour modifier les spécifications du modèle, menu Specifications \Rightarrow Modelling. Cliquez sur le nom du modèle, ici RegARIMA(TS).

Il faut modifier le contenu de deux fenêtres. Commençons par celle dont l'onglet est "Regression Variables". Nous allons demander en plus les corrections pour jours ouvrables (trading day) et les effets de la fête de Pâques (Easter effect) et gardons le traitement des données aberrantes.

\Rightarrow Cliquez sur l'onglet Regression Variables.

\Rightarrow Dans la zone Trading Day Effect, dans la liste déroulante Trading Day Correction, sélectionnez la troisième option Trading Day #Mon – #Sun, etc.

\Rightarrow Dans la zone Special Effects, dans la liste déroulante Easter Effect, choisissez Yes.

Malheureusement Demetra part du principe qu'un modèle estimé ne doit plus être modifié. Or, on peut espérer que les effets de jours ouvrables et de la fête de Pâques affectent les données et donc le modèle qui en est déduit. Il faut donc *chaque fois* effectuer les modifications suivantes.

\Rightarrow Cliquez sur l'onglet ARIMA Model Specification. Décochez les

deux coefficients du modèles SANS QUOI LES VALEURS DES COEFFICIENTS RESTERONT FIXES, et DE PLUS, POUR LA MEME RAISON, dans la liste déroulante Estimation of Unfixed Coefficients, CLIQUEZ SUR With initial values given by the user or by default.



Cliquez sur OK.

Le module TRAMO/SEATS emploie 6 variables binaires pour l'effet des jours ouvrables et 1 variable pour l'effet de Pâques. Pour plus de détails, voir la partie A de l'exercice. Les paramètres correspondants seront estimés ainsi que les paramètres du modèle ARIMA et les paramètres des interventions.



Pour estimer les paramètres du modèle ARIMA et les interventions, menu Processing \Rightarrow Execute SA. Cliquez "No" sur le message intermédiaire vous demandant de mettre la série à jour puis cliquez sur le nom du modèle, ici RegARIMA(TS).

La fenêtre du haut à droite a maintenant une colonne RegARIMA(TS) mise à jour. Pour la voir, il faut déplacer l'ascenseur vers la droite.



Comparez les résultats avec ceux de la colonne précédente. Qu'est-ce qui est nouveau ? Peut-on dire qu'il y a un effet de jours ouvrables ? Un effet de Pâques ?



5.3.1 Votre réponse

Nous allons produire un graphique unique comparant les trois effets étudiés ici : l'effet des jours ouvrables, l'effet de Pâques et les interventions.



Pour représenter ces trois séries, menu Result Analysis \Rightarrow Graphical Comparison Tool. Pour créer un graphique dans une des 4 cases de la nouvelle fenêtre qui s'ouvre alors, cliquez deux fois

dans la case. Pour développer partiellement l'arborescence de séries, cliquez sur le signe '+' de Preliminary Series puis sur le signe '+' de Components puis sur le signe '+' de Pre-adjusted Components puis sur le signe '+' de Total Trading Day Effect et enfin sur le nom du modèle, "3: RegARIMA(TS)". Cliquez sur le bouton Add to Plot. Procédez de même pour Easter Effect et Total Outlier Effect. Pour achever le graphique, cliquez sur le bouton OK.

Remarque



Voir les remarques du paragraphe 3.4.



Comparez l'intensité des trois effets. Quel est l'effet le plus apparent ?

5.3.2 Votre réponse



Si vous souhaitez réutiliser au chapitre 12 le modèle que vous avez obtenu, il faut le sauvegarder sous son nom d'origine, comme suit.

- ⇒ Revenez à la fenêtre principale du projet (terminée par :1).
- ⇒ Pour sauvegarder les spécifications du modèle dans un fichier, menu Specifications ⇒ Define User Model. Cliquez sur le nom du modèle, ici RegARIMA(TS). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et tapez le nom de fichier RegARIMA(TS).mdl. Cliquez sur le bouton Save.

Nous allons terminer en examinant en partie le fichier de sortie de Demetra.

- ⇒ Pour examiner le fichier de sortie relatif au modèle en cours d'examen, menu Result analysis ⇒ Show log file. Choisissez RegARIMA(TS).

Nous l'avons déjà fait dans le paragraphe 3.2 pour un modèle plus simple. Après un titre, on peut voir le tableau des données d'origine et celui des données transformées, puis transformées en différence, les autocorrélations et autocorrélations partielles, une rubrique ARIMA model estimation, une

rubrique White noise residuals, les tests statistiques sur les résidus et enfin les prévisions.

Ici, la rubrique ARIMA model estimation présente les sorties de régression, selon la forme habituelle (variable, paramètre estimé, erreur-type, statistique t de Student). Une rubrique Linear Series reprend la série corrigée.



Retrouvez les éléments mentionnés ci-dessus. Commentez la significativité à 5 % des différents paramètres du modèle.

5.3.3 Votre réponse



Cliquez sur le nom du modèle, ici RegARIMA(TS). Dans la fenêtre Save qui s'ouvre, naviguez jusqu'au dossier DATA puis CHAP11 puis CH11EX02 et tapez le nom de fichier RegARIMA(TS).mdl. Cliquez sur le bouton Save.

Vous pouvez aussi sauvegarder le projet sous un nom quelconque, *mais pas CHAMPX12, ni CHAMPTS.*



Pour sauvegarder le projet préparé, menu Project \Rightarrow Save As. Donnez un nom au projet, par exemple PARTIE5.

SYNTHÈSE

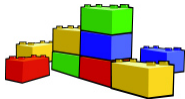
Dans cette partie, nous avons repris l'étude de la série des ventes de champagne à l'aide d'un modèle ARIMA et avec une analyse d'interventions dans le module TRAMO/SEATS de Demetra. Nous avons également introduit l'effet des jours ouvrables et l'effet de la fête de Pâques mais avec moins de succès. Dans l'ensemble, les résultats sont très proches de ceux obtenus par le programme X-12-ARIMA.



Exercice avancé

(Pour les utilisateurs de la version avancée du cours)

Préalable Le chapitre 11 du cours de base doit avoir été suivi jusqu'à la page 101.



Objectif Le but est de d'introduire les variables de régression utilisées dans Demetra de manière plus détaillée que dans la présentation du cours.



Données Néant.

Structure de l'exercice

L'exercice avancé comporte une seule partie :

- Dans la partie A, le but de l'exercice est d'expliquer, de manière plus détaillée que dans le cours, la manière de définir les différentes variables de régression employées dans les deux modules de Demetra que sont RegARIMA et TRAMO.

Partie A Le but de l'exercice est d'expliquer, de manière plus détaillée que dans le cours, la manière de définir les différentes variables de régression employées dans les deux modules de Demetra que sont RegARIMA et TRAMO. Nous traitons successivement des interventions, des effets des jours ouvrables ou de calendrier et de l'effet de la fête de Pâques.

A.a LES INTERVENTIONS

Nous allons ici définir la détection des données aberrantes et l'emploi des interventions à cette fin. N'importe quelle type d'intervention peut être définie par l'utilisateur. Nous allons donc nous limiter aux interventions définies à l'avance. On distingue les trois types suivants :

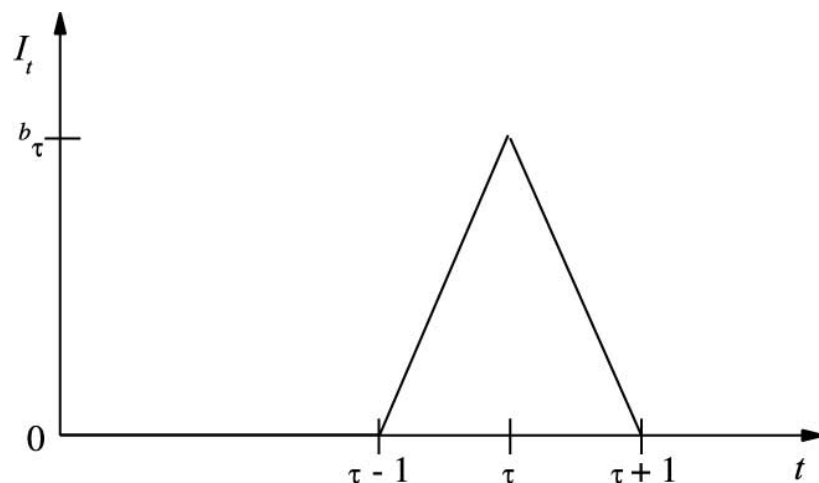
- 1) les interventions ponctuelles ou impulsions, pour traiter les valeurs aberrantes additives (*additive outlier*, AO) ;
- 2) les interventions temporaires, pour traiter les changements temporaires dont l'effet est dégressif (*temporary change*, TC) ;
- 3) les interventions de changement de niveau (*level shift*, LS).

Notons I_t la variable résumant une intervention. Nous allons distinguer les trois cas : I_t^{AO} , I_t^{TC} , I_t^{LS} . Bien entendu, I_t peut être une intervention combinant les trois types.

Représentons par $I_t(\tau)$ une variable binaire indicatrice pour le temps τ , telle que

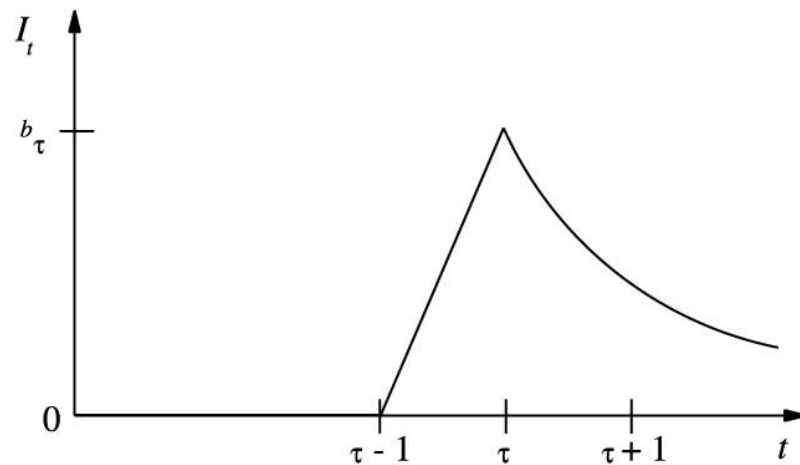
$$\begin{aligned} I_t(\tau) &= 1 \text{ pour } t = \tau, \\ I_t(\tau) &= 0 \text{ pour } t \neq \tau. \end{aligned}$$

L'intervention AO est de la forme suivante : $I_t^{\text{AO}} = b_\tau I_t(\tau)$. On peut la représenter de la manière suivante.



L'intervention TC est de la forme suivante : $I_t^{\text{TC}} = b_\tau [1/(1 - \delta B)] I_t(\tau)$,

où B est l'opérateur de retard et δ est une constante ou un paramètre. On peut la représenter de la manière suivante.



On a donc un saut d'amplitude b_τ suivi d'une décroissance exponentielle de taux δ , compris entre -1 et 1 , généralement positif. En $t = \tau + 1$, l'effet est $b_\tau \delta$, en $t = \tau + 2$, l'effet est $b_\tau \delta^2$, en $t = \tau + 3$, l'effet est $b_\tau \delta^3$, et ainsi de suite.

Remarque



La justification est la suivante. En développant $b_\tau [1/(1 - \delta B)] I_t(\tau)$, on obtient $b_\tau [1 + \delta B + \delta^2 B^2 + \delta^3 B^3 + \delta^4 B^4 + \dots] I_t(\tau)$, donc

$b_\tau I_t(\tau) + b_\tau \delta I_{t-1}(\tau) + b_\tau \delta^2 I_{t-2}(\tau) + b_\tau \delta^3 I_{t-3}(\tau) + \dots$, ce qui vaut

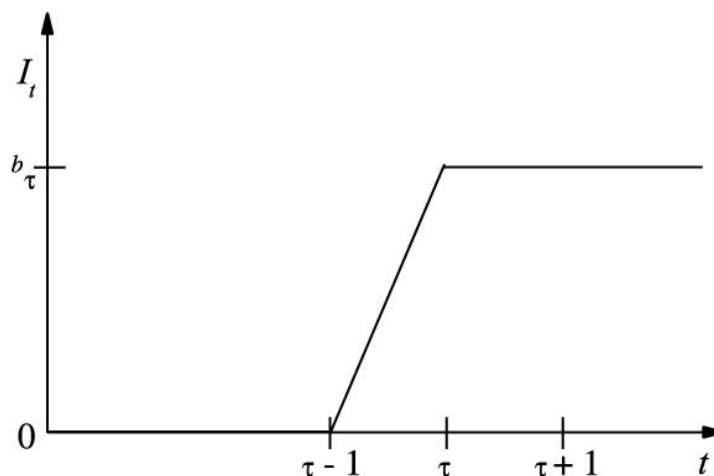
b_τ , si $t = \tau$,

$b_\tau \delta$, si $t = \tau + 1$,

$b_\tau \delta^2$, si $t = \tau + 2$,

...

L'intervention LS est de la forme suivante : $I_t^{LS} = b_\tau [1/(1 - B)] I_t(\tau)$, où B est l'opérateur de retard. On peut la représenter de la manière suivante.



On a donc un saut d'amplitude b_τ . En $t \geq \tau$, l'effet est b_τ .

**Remarque**

La justification est la suivante. En développant $b_\tau [1/(1-B)] I_t(\tau)$, on obtient $b_\tau [1 + B + B^2 + B^3 + B^4 + \dots] I_t(\tau)$, donc

$b_\tau I_t(\tau) + b_\tau I_{t-1}(\tau) + b_\tau I_{t-2}(\tau) + b_\tau I_{t-3}(\tau) + \dots$, ce qui vaut

b_τ , si $t = \tau$,
 b_τ , si $t = \tau + 1$,
 b_τ , si $t = \tau + 2$,
 b_τ , si $t = \tau + 3$,
 \dots

La méthodologie pour la détection des données aberrantes, l'identification et l'estimation, les références sont l'article de Chang, Chen et Tiao (1988) et le livre de Pankratz (1991, pp. 313-314) que nous suivons ici. Supposons un modèle ARMA avec intervention, par exemple un modèle ARMA(1, 1), d'équation

$$(1 - \phi B) (y_t - I_t) = (1 - \theta B) e_t.$$

Notons qu'on peut écrire

$$\begin{aligned} e_t &= (1 - \theta B)^{-1} (1 - \phi B) (y_t - I_t) \\ &= (1 - \theta B)^{-1} (1 - \phi B) y_t - (1 - \theta B)^{-1} (1 - \phi B) I_t \quad (*) \end{aligned}$$

Pour comprendre la procédure, il faut se souvenir que la statistique de Student d'un coefficient de régression b_j est de la forme $b_j/\sigma(b_j)$ où l'erreur-type $\sigma(b_j) = \sigma \sqrt{(\mathbf{X}'\mathbf{X})_{jj}^{-1}}$, utilise de l'information sur les valeurs des variables explicatives : $(\mathbf{X}'\mathbf{X})_{jj}^{-1}$. Ecrivons ceci sous la forme $b_j/\sigma K(X)$.

La procédure utilisée peut être résumée comme suit :

- 1) un modèle ARIMA est spécifié et estimé ;
- 2) les résidus e_t sont déterminés par $e_t = (1 - \theta B)^{-1} (1 - \phi B) y_t$ puis l'écart-type résiduel σ ;
- 3) pour chaque temps $\tau > 1$, on introduit tour à tour chacune des interventions AO, TC et LS, on estime les paramètres correspondants $b_{AO,\tau}$, $b_{TC,\tau}$, et $b_{LS,\tau}$
- 4) pour chaque temps $\tau > 1$, on détermine également les erreurs-types $\sigma K_{AO,\tau}(X)$, $\sigma K_{TC,\tau}(X)$, $\sigma K_{LS,\tau}(X)$ et les statistiques de Student $t_{AO,\tau} = b_{AO,\tau}/\sigma K_{AO,\tau}(X)$, $t_{TC,\tau} = b_{TC,\tau}/\sigma K_{TC,\tau}(X)$ et $t_{LS,\tau} = b_{LS,\tau}/\sigma K_{LS,\tau}(X)$;

- 5) on détermine ensuite, pour chaque temps τ ,
 $t_\tau = \max(|t_{AO, \tau}|, |t_{TC, \tau}|, |t_{LS, \tau}|)$, puis $t = \max_{\tau=2, \dots, T} (t_\tau)$
- 6) on compare t à une valeur critique appropriée C_d — voir ci-dessous; si $t \leq C_d$, on arrête la procédure ; dans le cas contraire, on décide d'ajouter une intervention au temps τ pour lequel $t_\tau = t$; le type d'intervention retenue, AO, TC ou LS, dépend de la statistique parmi $|t_{AO, \tau}|, |t_{TC, \tau}|, |t_{LS, \tau}|$, qui est égale à t_τ ;
- 7) on estime à nouveau les résidus e_t et l'écart-type résiduel σ de la manière suivante :
- a) dans le cas d'une intervention AO au temps τ , de paramètre estimé égal à b_τ , on applique (*) pour calculer les nouveaux résidus $e_t - b_\tau (1 - \theta B)^{-1} (1 - \phi B) I_t(\tau)$ puis l'écart-type résiduel σ ;
- b) dans le cas d'une intervention TC au temps τ , de paramètre estimé égal à b_τ , on applique (*) pour calculer les nouveaux résidus $e_t - b_\tau (1 - \theta B)^{-1} (1 - \phi B) [1/(1 - \delta B)] I_t(\tau)$ puis l'écart-type résiduel σ ;
- c) dans le cas d'une intervention LS au temps τ , de paramètre estimé égal à b_τ , on applique (*) pour calculer les nouveaux résidus $e_t - b_\tau (1 - \theta B)^{-1} (1 - \phi B) [1/(1 - B)] I_t(\tau)$ puis l'écart-type résiduel σ ;
- 8) utiliser le nouvel écart-type résiduel σ et la nouvelle série résiduelle pour répéter les étapes 4) à 7) jusqu'à ce qu'une intervention ne soit plus détectée.

On peut alors estimer de nouveau les paramètres du modèle avec toutes les interventions détectées, donc ϕ , θ , et les différents b_τ .

Les valeurs critiques C_d à employer ont été étudiées par simulation, dans des circonstances déterminées qui ne sont pas nécessairement toujours applicables. D'après ces résultats, $C_d = 3,5$ correspond approximativement à un niveau de test de 5 %. On acceptera un peu plus d'interventions en prenant $C_d = 3,3$.

A.b LES EFFETS DE JOURS OUVRABLES

Les corrections pour jours ouvrables tiennent compte des nombres de chaque jour de la semaine, par mois.

Notons X_{1t} le nombre de lundis dans le mois t
 X_{2t} le nombre de mardis dans le mois t
 ...

X_{6t} le nombre de samedis dans le mois t
 X_{7t} le nombre de dimanches dans le mois t
 X_t le nombre total de jours dans le mois t

Le nombre de jours du mois est la somme du nombre de lundis, de mardis, de mercredis, de jeudis, de vendredis, de samedis et de dimanches. On a donc $X_t = X_{1t} + X_{2t} + \dots + X_{6t} + X_{7t}$.

La première idée pour tenir compte du nombre de jours de la semaine est d'effectuer une régression des données y_t sur les variables $X_{1t}, X_{2t}, \dots, X_{6t}, X_{7t}$. Notons les coefficients respectivement $b_1, b_2, \dots, b_6, b_7$. Les valeurs ajustées se calculeront comme $b_0 + b_1 X_{1t} + b_2 X_{2t} + \dots + b_6 X_{6t} + b_7 X_{7t}$, et les résidus représenteront les données corrigées pour les nombres de jours ouvrables.

Cependant cette régression se présente sous de mauvais auspices. En effet, il y a de la quasi-colinéarité avec la constante parce que les $X_t = X_{1t} + X_{2t} + \dots + X_{6t} + X_{7t}$ sont presque égales d'un mois à l'autre. On propose alors plusieurs alternatives.

Alternative 1

Elle consiste à utiliser comme variables explicatives dans la régression: $X_{1t} - X_{7t}, X_{2t} - X_{7t}, \dots, X_{6t} - X_{7t}$ et X_t . Les avantages sont :

- a) ce modèle est équivalent au précédent mais souffre moins de la quasi-colinéarité;
- b) on pourrait omettre la variable X_t .

Remarque



Le test de Fisher simultané que les coefficients des variables 1 à 6 sont nuls est un test global de l'effet des jours ouvrables. Il est disponible dans les sorties détaillées de X-12-ARIMA et de TRAMO/SEATS mais n'apparaît malheureusement pas dans les sorties principales de Demetra.

Alternative 2

Elle consiste à utiliser deux variables explicatives dans la régression : $(X_{1t} + X_{2t} + \dots + X_{5t}) - 5(X_{6t} + X_{7t})/2$ et X_t . Les avantages sont :

- a) ce modèle est plus simple que le précédent;
- b) le test sur la présence d'un effet de jours ouvrables se limite à un test de Student sur un seul coefficient;
- c) le test examine si la moyenne des 5 jours ouvrables est égale à la

moyenne du samedi et du dimanche;

d) on pourrait omettre la variable X_t .

Il y a toutefois quelques inconvénients :

a) ce modèle ne convient pas pour les secteurs où il y a de l'activité le samedi mais pas le dimanche ;

b) il ne convient pas si l'activité n'est pas répartie régulièrement dans la semaine.

Quelle est la situation par rapport à Census X-11, qui comportait aussi des corrections pour les variations de calendrier, notamment des régressions pour jours ouvrables ?

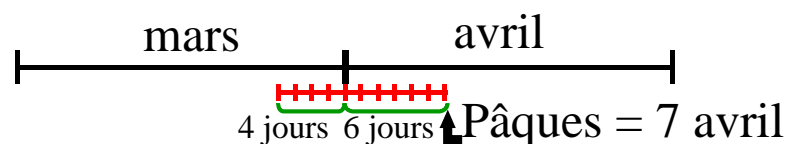
La grande différence est : ici, les régressions sont appliquées sur les données en tenant compte de ce que les erreurs du modèle sont autocorrélées et modélisées par un modèle ARIMA

Dans la méthode Census X-11 ces corrections étaient faites sur la composante d'erreur (corrigées pour valeurs aberrantes) avec le risque que les variations pour jours ouvrables soient pris dans les autres composantes

A.C LES EFFETS DE LA FÊTE DE PÂQUES

Pâques est une fête mobile importante dans les pays occidentaux, mais elle est moins importante que Noël et le Nouvel An. Elle a la particularité de se déplacer entre mars et avril, donc entre les trimestres I et II. Il n'est donc pas surprenant que sa position dans une année ait un effet sur la saisonnalité.

Les corrections généralement envisagées se basent sur une période de n jours avant Pâques, où n doit être choisi, et la détermination du nombre de jours de cette période en mars et en avril, par exemple $n = 10$.



Plus concrètement, on procède comme suit :

soit $n_{\text{mars},t}$, le nombre parmi ces n jours qui tombent dans le mois de mars, dans l'exemple 4;

soit $n_{\text{avril},t}$, le nombre parmi ces n jours qui tombent dans le mois d'avril,

dans l'exemple 6.

On définit alors une variable X_P à valeurs X_{P_t} :

pour t en mars, on définit $X_{P_t} = n_{\text{mars}}/n$, ici 0,4;

pour t en avril, on définit $X_{P_t} = n_{\text{avril}}/n$, ici 0,6;

pour t dans un autre mois, on définit $X_{P_t} = 0$.

Le test de l'effet de Pâques se limite à un test de Student sur le coefficient de la variable X_P .

SYNTHÈSE

Nous avons décrit ici de manière succincte la manière de définir les variables de régression pour les principales interventions, pour l'effet des jours ouvrables et pour l'effet de la fête de Pâques.

[Retour au chapitre 11](#)