**Projet de Prévision des séries temporelles Ouatiti Youssef Esseddiq**

Remarque à propos des scripts Matlab de lecture des données à partir des fichiers Excel : Prière de changer l’adresse de l’emplacement du fichier source des données sur les fichiers ‘lecture.m‘ de celle actuelle à celle choisi sur votre ordinateur.

***1ère partie :***

***1-***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Critère\Modèle** | **Tendance linéaire** | **Tendance quadratique** | **Exponentiel** | **Exponentiel modifié** | **Gompertz** | **Logistique** |
| RMSE | 486.9672865 | 347.7901821 | 679.463996 | 406.5144199 | 402.309548 | 400.0346061 |
| MAE | 373.6491795 | 265.2527971 | 569.044517 | 257.5435638 | 260.5523879 | 265.1120853 |
| MAPE | 9.285707623 | 6.120432326 | 13.86812 | 5.368057135 | 5.705923909 | 6.147886713 |

***2-*** Il y a plusieurs modèles ayant meilleur résultat pour un critère mais non pour les autres critères. On choisit le plus consistant qui est : le modèle exponentiel modifié.

***3-*** Puisque la série est annuelle on a le choix entre 3 type de modèles selon Gardner : le lissage simple, DA-N ou le lissage de Holt. On calcule la variance des 3 séries : Xt, (1-B)Xt et (1-B)²Xt.

La variance minimale est celle de la série : (1-B)Xt (Voir le fichier Excel A1 feuille 3). Ainsi on ajuste la série à l’aide d’un modèle Damped Additive None.

On obtient un MAPE de ***21.1913*** pour l’initialisation suivante du modèle DA-N :

|  |  |
| --- | --- |
| *α* | 0.1 |
| *δ* | 0.2 |
| *[Greek phi Didot.svg](https://en.wikipedia.org/wiki/File:Greek_phi_Didot.svg)* | 0.3 |
| S1 | y1 |
| T1 | 0 |

Maintenant, a l’aide de Matlab nous allons écrire un script permettant de calculer le MAPE, puis on va essayer de trouver à l’aide d’une fonction d’optimisation les paramètres d’initialisation minimisant le MAPE.

Nous obtenons alors :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *α* | *δ* | *[Greek phi Didot.svg](https://en.wikipedia.org/wiki/File:Greek_phi_Didot.svg)* | *S1* | *T1* | *MAPE* |
| ***1*** | ***0.4*** | ***0.9*** | ***1620.1*** | ***90.6*** | ***3.7845*** |

***4-*** Le modélisateur expert de SPSS a choisi le modèle de lissage de Holt avec les critères et paramètres suivants :

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | | Ljung-Box Q(18) | | | Number of Outliers |
| Stationary R-squared | RMSE | MAPE | MAE | Statistics | DF | Sig. |
| série 208-Model\_1 | 0 | .259 | ***228.794*** | ***3.692*** | ***150.931*** | 13.855 | 16 | .610 | 0 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Exponential Smoothing Model Parameters** | | | | | | |
| Model | | | Estimate | SE | t | Sig. |
| série 208-Model\_1 | No Transformation | Alpha (Level) | ***1.000*** | .161 | 6.199 | .000 |
| Gamma (Trend) | ***.001*** | .055 | .015 | .988 |

***5-***Nous obtenons les résultats suivants :

|  |  |  |  |
| --- | --- | --- | --- |
| Exponentiel modifié | DA-N | SPSS | Naïve |
| *4.62395989* | ***5.8724099*** | ***8.822699*** | ***2.8548896*** |

La méthode naïve donne des résultats mieux que les autres modèles pour les prévisions.

***2ème partie :***

***1-*** On effectue le test de Buys-Ballot sur la série trimestrielle, on obtient une valeur de : **0.236513**. Ainsi, la série des moyennes annuelles est plutôt corrélée avec la série des écart-types annuelles et on choisit le schéma multiplicatif.

***2-*** On calcule MMC(4) pour enlever la composante saisonnière (Voir feuille T2 dans le fichier Excel Partie2):

***3-*** Voir T2 du fichier partie2.

***4-*** A partir du tableau ci-dessous :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **T/Année** | **79** | **80** | **81** | **82** | **83** | **84** | **85** | **86** |
| T1 |  | 0.95653622 | 0.95391334 | 0.96016887 | 0.98974195 | 0.98117841 | 0.94995641 | 0.93678054 |
| T2 |  | 1.00054727 | 1.01407824 | 0.97989866 | 1.00512662 | 1.04504442 | 1.02589447 | 1.01007246 |
| T3 | 1.05737212 | 1.04347826 | 1.05977732 | 1.05209941 | 1.03702875 | 1.03687801 | 1.05752958 | 1.04915757 |
| T4 | 0.97847251 | 0.99189375 | 0.95879444 | 0.98550725 | 0.974174 | 0.968468 | 0.95795739 | 0.9805344 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **87** | **88** | **89** | **90** | **91** | **92** | **93** | **94** | **95** |
| 0.966180187 | 0.966916295 | 0.982221208 | 0.959648138 | 0.961642089 | 0.989033276 | 0.981875227 | 0.947625 | 1.002843 |
| 1.039699647 | 1.036120428 | 1.011182918 | 1.021748078 | 1.028803932 | 1.035996813 | 1.033988986 | 1.014877 |  |
| 1.027461299 | 1.024928634 | 1.038839291 | 1.036902444 | 1.019267406 | 1.006156745 | 1.026313513 | 1.015049 |  |
| 0.973642216 | 0.959656453 | 0.974581467 | 0.984288783 | 0.975439872 | 0.968350035 | 1.008457876 | 0.984372 |  |

* On remarque que la composante saisonnière pour le trimestre 1 est toujours inférieur à 1, sauf pour l’année 1995. Donc, **le trimestre 1 est assez stable.**
* Pour le 2eme trimestre, pour toutes les années la composante saisonnière est supérieure à 1, sauf pour l’année 1982 ainsi **T2 est aussi assez stable**.
* Pour T3 toutes les années présentent une composante saisonnière supérieur à 1 donc **ce trimestre (T3) est stable.**
* Pour T4, toutes les années présentent une composante saisonnière inferieur à 1, sauf l’année 1993, **ainsi ce trimestre (T4) est assez stable.**

***5-*** On choisit le modèle linéaire, quadratique et exponentiel. (Voir T3 fichier Partie2).

***6-*** On calcule les modèles de décomposition saisonnière, et les critères de comparaison sur la feuille T3 du fichier partie2.

|  |  |  |  |
| --- | --- | --- | --- |
| **Critères/Modèle** | **Modèle linéaire** | **Modèle Quadratique** | **Modèle exponentiel** |
| RMSE | 232.2523081 | 217.2017608 | 347.2570621 |
| MAE | 195.3307248 | 178.8244057 | 251.0040231 |
| MAPE | 4.536026985 | 4.028308866 | 6.170395905 |

***7-*** Le meilleur modèle est le modèle quadratique.

**8-** L’analyse de Gardner pour le choix du modèle convenable est faite sur la feuille T4 du fichier Excel Partie2.

Puisque (1-B)(1-B4)Xt a la variance minimale, on va ajuster la série en utilisant le modèle DA-M.

L’ajustement est fait sur la feuille T5 du fichier Partie2.

**9-** Les résultats du modélisateur expert de SPSS sont :

* Le modèle choisit est le modèle de Holt-Winters additive.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description** | | | |
|  | | | Model Type |
| Model ID | série trim 880 (Xt) | Model\_1 | Winters' Additive |

* Les critères de comparaison :

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | Ljung-Box Q(18) | | | Number of Outliers |
| RMSE | MAPE | MAE | Statistics | DF | Sig. |
| série trim 880 (Xt)-Model\_1 | 0 | 136.188 | 2.427 | 99.928 | 14.097 | 15 | .518 | 0 |

* Les paramètres estimés :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Exponential Smoothing Model Parameters** | | | | | | |
| Model | | | Estimate | SE | t | Sig. |
| série trim 880 (Xt)-Model\_1 | No Transformation | Alpha (Level) | .922 | .127 | 7.261 | .000 |
| Gamma (Trend) | 1.556E-7 | .019 | 8.028E-6 | 1.000 |
| Delta (Season) | 1.641E-5 | .394 | 4.163E-5 | 1.000 |

***10-*** Nous obtenons les résultats suivants : (Les calculs sont dans la feuille Evaluation du fichier Partie2).

|  |  |  |  |
| --- | --- | --- | --- |
| Modèle Quadratique | DA-M | SPSS | Naïve |
| *6.024588116* | ***8.243182*** | ***3.0469003*** | ***6.2333616*** |

La méthode Winters-Holt additive proposée par SPSS donne des résultats mieux que les autres modèles pour les prévisions.

***3ème partie :***

***1-*** On effectue le test de Buys-Ballot sur la série mensuelle, on obtient une valeur de : ***0.673551***. Ainsi, la série des moyennes annuelles est plutôt corrélée avec la série des écart-types annuelles et on choisit le schéma multiplicatif.

***2-*** On calcule MMC(12) pour enlever la composante saisonnière (Voir feuille M2 dans le fichier Excel Partie3):

***3-*** Voir M2 du fichier partie3.

***4-*** A partir du tableau ci-dessous :

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Mois/Année** | **85** | **86** | **87** | **88** | **89** | **90** | **91** | **92** | **93** |
| Jan |  | 0.79230769 | 0.94448349 | 0.67605634 | 0.72693383 | 1.45513339 | 0.63348416 | 0.95565611 | 1.313432836 |
| Fev |  | 0.65837321 | 0.5734072 | 0.91610738 | 0.77005348 | 0.70694864 | 1.14191961 | 0.51382694 | 0.536443149 |
| Mar |  | 0.82487725 | 1.34018499 | 0.74054759 | 1.37846154 | 0.44192635 | 1.20528634 | 0.87234043 | 0.92917847 |
| Avr | 0.82107175 | 0.76433121 | 0.6557377 | 0.72593951 | 1.35729387 | 0.71035941 | 1.1152815 | 1.09955947 | 0.741573034 |
| Mai | 1.24266999 | 1.19001664 | 0.77553857 | 1.43931461 | 1.05997819 | 1.39179632 | 1.25531915 | 1.45325022 | 1.224489796 |
| Jui | 1.26905132 | 1.50452565 | 1.20923723 | 0.92352655 | 1.87755102 | 0.98488121 | 1.25342466 | 0.83146067 | 0.99911426 |
| Jul | 0.77158604 | 1.01660454 | 1.32956951 | 0.97473997 | 0.99769053 | 1.50453172 | 0.85359801 | 1.39130435 | 1.372031662 |
| Auo | 0.99628253 | 1.05198777 | 1.18550107 | 1.36349925 | 0.47580175 | 2.02609363 | 0.87162162 | 0.9645933 | 1.308310992 |
| Sep | 1.32533922 | 0.80428954 | 1.05830904 | 0.76608479 | 0.53932584 | 0.85144124 | 0.65608466 | 0.89655172 | 1.015228426 |
| Oct | 1.11610006 | 1.17343905 | 1.23730099 | 0.85098227 | 0.72777395 | 0.67144907 | 0.85867621 | 0.91317073 |  |
| Nov | 0.86409342 | 1.05548038 | 0.83101242 | 0.84109458 | 0.77865898 | 0.782042 | 1.12866016 | 1.01510574 |  |
| Dec | 1.20661693 | 0.95764788 | 1.20685714 | 1.1788269 | 1.12012242 | 0.63436123 | 1.4676259 | 0.58181818 |  |

* On remarque que la composante saisonnière pour les mois : Février et Mai est toujours inférieur à 1 ou supérieure à 1 sauf pour une valeur. Donc, **ces mois sont assez stables.**
* Pour les autres mois, la composante saisonnière prend plusieurs valeurs supérieurs et inferieur a 1, **donc ces mois sont instables.**

***5-*** On choisit le modèle linéaire, quadratique et exponentiel. (Voir M2 fichier Partie3).

***6-*** On calcule les modèles de décomposition saisonnière et les critères de comparaison sur la feuille M3 du fichier partie3.

|  |  |  |  |
| --- | --- | --- | --- |
| **Critères/Modèle** | **Modèle linéaire** | **Modèle Quadratique** | **Modèle exponentiel** |
| RMSE | 1109.932495 | 1069.349021 | 1079.1804 |
| MAE | 871.8978942 | 835.6012033 | 839.4524967 |
| MAPE | 26.84805304 | 24.14932567 | 24.27772538 |

***7-*** Le meilleur modèle est le modèle quadratique.

**8-** L’analyse de Gardner pour le choix du modèle convenable est faite sur la feuille M4 du fichier Excel Partie3.

Puisque (1-B)Xt a la variance minimale, on va ajuster la série en utilisant le modèle DA-N.

L’ajustement est fait sur la feuille M5 du fichier Partie2.

**9-** Les résultats du modélisateur expert de SPSS sont :

* Le modèle choisit est le modèle de Holt-Winters multiplicatif.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description** | | | |
|  | | | Model Type |
| Model ID | V3 | Model\_1 | Winters' Multiplicative |

* Les critères de comparaison :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | Ljung-Box Q(18) | | | Number of Outliers |
| RMSE | MAPE | MAE | Statistics | DF | Sig. |
| V3-Model\_1 | 0 | 1113.432 | 24.512 | 846.390 | 30.543 | 15 | .010 | 0 |

* Les paramètres estimés :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Exponential Smoothing Model Parameters** | | | | | | |
| Model | | | Estimate | SE | t | Sig. |
| V3-Model\_1 | No Transformation | Alpha (Level) | .050 | .028 | 1.754 | .082 |
| Gamma (Trend) | .207 | .133 | 1.554 | .123 |
| Delta (Season) | .028 | .026 | 1.074 | .285 |

***10-*** Nous obtenons les résultats suivants : (Les calculs sont dans la feuille Evaluation du fichier Partie3).

|  |  |  |  |
| --- | --- | --- | --- |
| Modèle Quadratique | DA-N | SPSS | Naïve |
| *43.82256206* | ***40.99823*** | ***37.8814*** | ***43.8948*** |

La méthode Winters-Holt multiplicative proposée par SPSS donne des résultats mieux que les autres modèles pour les prévisions.

***4ème partie :***

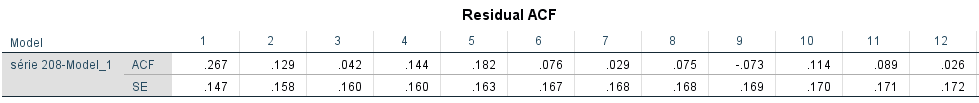
1. ***Série annuelle :***

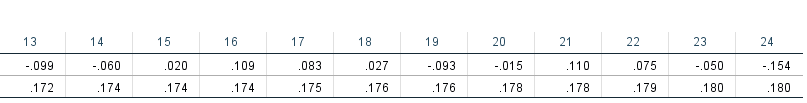
|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description** | | | |
|  | | | Model Type |
| Model ID | série 208 | Model\_1 | ARIMA(0,1,0) |

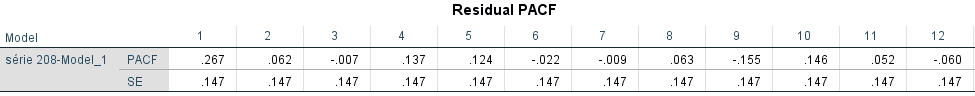
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | | | Ljung-Box Q(18) | | | Number of Outliers |
| Stationary R-squared | RMSE | MAPE | MAE | Normalized BIC | Statistics | DF | Sig. |
| série 208-Model\_1 | 0 | -4.441E-16 | 245.580 | 4.002 | 166.836 | 11.090 | 12.035 | 18 | .845 | 0 |

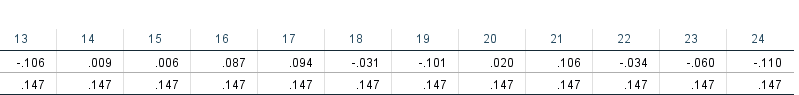
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | |
|  | | | | Estimate | SE | t | Sig. |
| série 208-Model\_1 | série 208 | Natural Logarithm | Constant | .029 | .008 | 3.541 | .001 |
| Difference | 1 |  |  |  |

***3-***

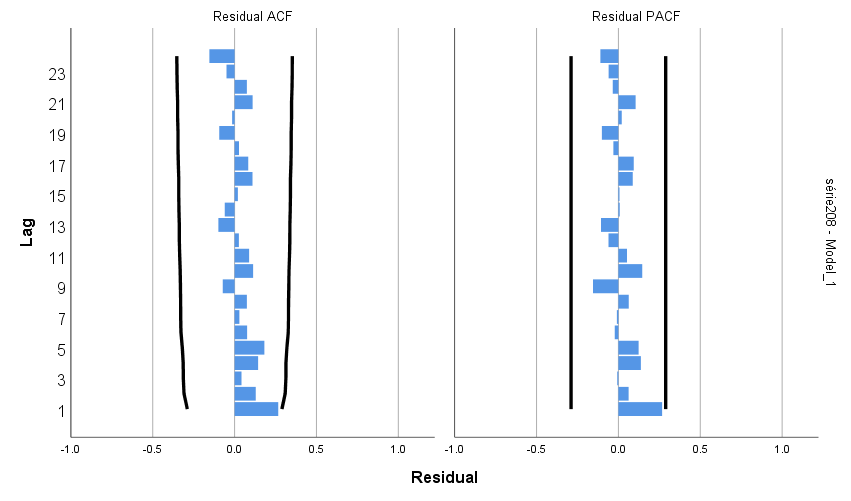






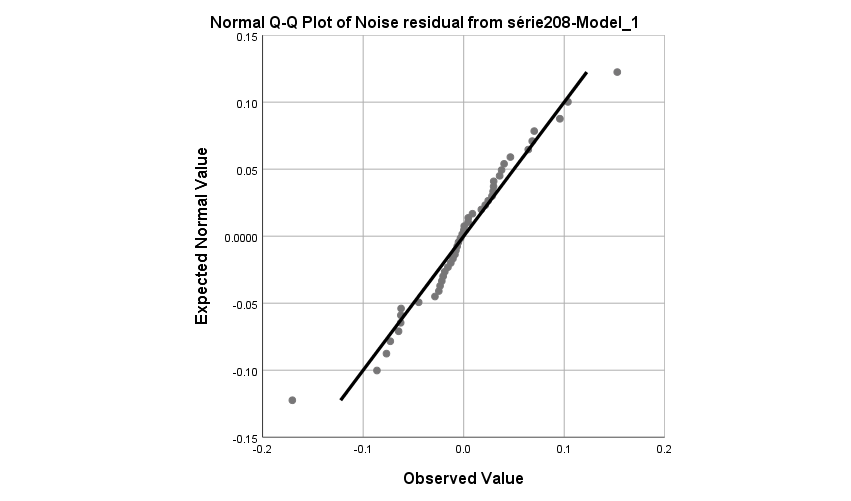


***4-***



***5-*** Le modèle choisi est ARIMA(0,1,0), on peut rejeter le test de bruit blanc, puis le paramètre estimée (la constante) est statistiquement significatif puisque sa p-valeur est inferieur à 0.05. On a absence d’autocorrélations et d’autocorrélations partielles au niveau du résidu. Ainsi on peut valider ce modèle.

***6- Test de normalité :***



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tests of Normality** | | | | | | |
|  | Kolmogorov-Smirnova | | | Shapiro-Wilk | | |
| Statistic | df | Sig. | Statistic | df | Sig. |
| Noise residual from série208-Model\_1 | .110 | 46 | .200\* | .968 | 46 | .232 |
| \*. This is a lower bound of the true significance. | | | | | | |
| a. Lilliefors Significance Correction | | | | | | |

**Selon les tests les résidus suivent une loi normale.**

1. ***Série trimestrielle :***

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description** | | | |
|  | | | Model Type |
| Model ID | série trim 880 | Model\_1 | ARIMA(0,1,0)(0,1,1) |

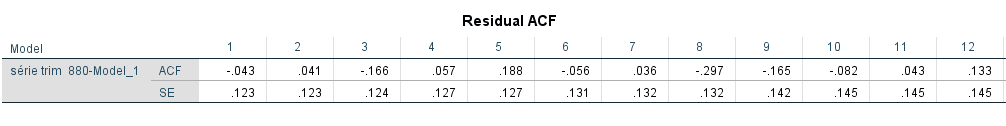
***1-***

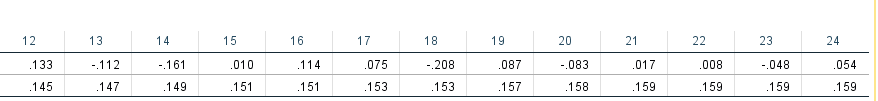
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | | | Ljung-Box Q(18) | | | Number of Outliers | |
| Stationary R-squared | RMSE | MAPE | MAE | Normalized BIC | Statistics | DF | Sig. |
| série trim 880-Model\_1 | 0 | .352 | 148.601 | 2.248 | 105.019 | 10.066 | 25.555 | 17 | .083 | 0 |

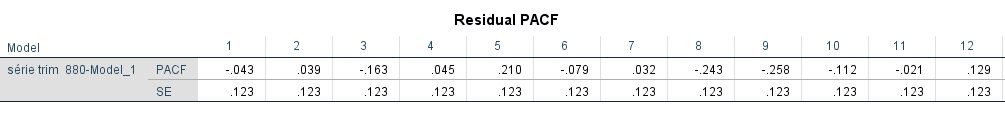
***2-***

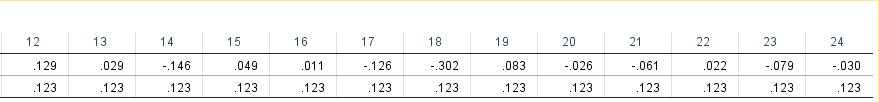
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| série trim 880-Model\_1 | série trim 880 | No Transformation | Difference | | 1 |  |  |  |
| Seasonal Difference | | 1 |  |  |  |
| MA, Seasonal | Lag 1 | .859 | .097 | 8.872 | .000 |

***3-***

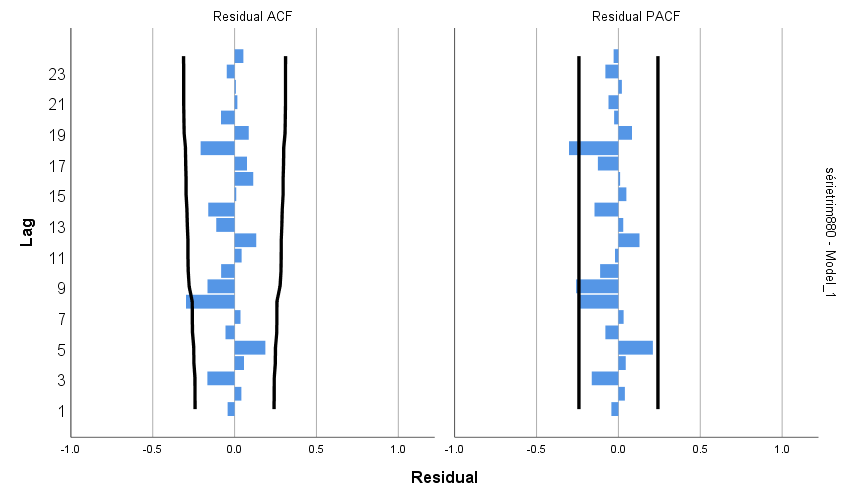








***4-***



***5-*** Le modèle choisi est ARIMA(0,1,1), on peut rejeter le test de bruit blanc, puis les paramètres estimées (la constante, et le paramètre q de MA) sont statistiquement significatif puisqu’il leur p-valeur est inférieure à 0.05. Mais on a présence d’autocorrélations et d’autocorrélations partielles au niveau du résidu. Ainsi on ce modèle est invalide.

1. ***Série mensuelle:***

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Description** | | | |
|  | | | Model Type |
| Model ID | série mens 1707 | Model\_1 | ARIMA(0,1,1)(0,0,0) |

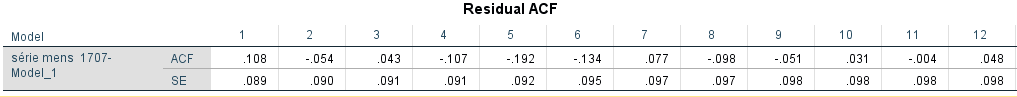
***1-***

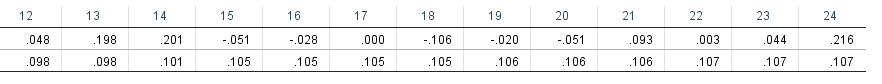
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model Statistics** | | | | | | | | | |
| Model | Number of Predictors | Model Fit statistics | | | | Ljung-Box Q(18) | | | Number of Outliers |
| Stationary R-squared | RMSE | MAPE | MAE | Statistics | DF | Sig. |
| série mens 1707-Model\_1 | 0 | .308 | 1309.314 | 33.093 | 1052.409 | 27.338 | 17 | .053 | 0 |

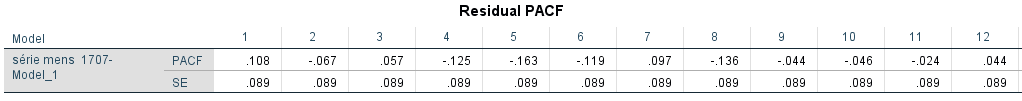
***2-***

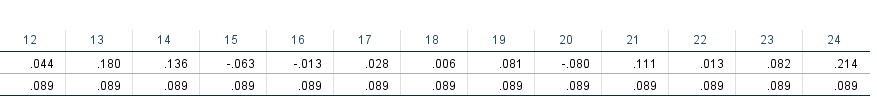
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ARIMA Model Parameters** | | | | | | | | |
|  | | | | | Estimate | SE | t | Sig. |
| série mens 1707-Model\_1 | série mens 1707 | No Transformation | Difference | | 1 |  |  |  |
| MA | Lag 1 | .818 | .054 | 15.188 | .000 |

***3-***

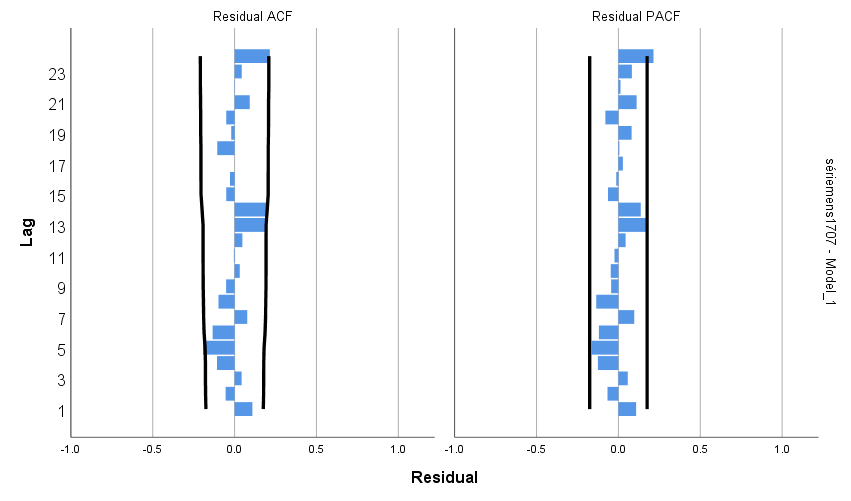








***4-***



***5-*** Le modèle choisi est ARIMA(0,1,1), on peut rejeter le test de bruit blanc, puis les paramètres estimées (la constante, et le paramètre q de MA) sont statistiquement significatif puisqu’il leur p-valeur est inférieure à 0.05. Mais on a présence d’autocorrélations et d’autocorrélations partielles au niveau du résidu. Ainsi on ce modèle est invalide.