

L'Impact de la Crise Sanitaire sur le Trafic Portuaire de Béjaïa(EPB)

Ramdani A., Ghiles L., Saadi N. et Aïssani D.

Research Unit **LaMOS** (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria
lamos_bejaia@hotmail.com ,

Résumé L'application des méthodes de la recherche opérationnelle au niveau des entreprises est indispensable pour améliorer les différents processus de service et production, c'est à cet effet qu'il est plus nécessaire pour une entreprise d'établir une planification future dans toutes les filières et les secteurs, un préalable indispensable pour réduire les risques ou encore annexer plus de bénéfice, et cela se fait notamment avec les prévisions. Dans ce travail, nous avons étudié un phénomène clé de la prévision relevé durant notre stage au sein de l'EPB. L'objectif de notre étude est de traiter l'évolution de trafic des hydrocarbures et le trafic des conteneurs durant l'ère de la crise Covid- 19 et de prévoir les quantités de tonnage des hydrocarbures qui seront engendrées pour l'année 2021. Des modèles mathématiques sont formulés sur la base de la modélisation statistique et linéaire qui reflètent mieux les caractéristiques des systèmes étudiés. Dans un premier temps, nous avons modélisé les deux séries chronologiques en faisant appel à la méthode de Box et Jenkins et aux calculs des prévisions. Dans un deuxième temps, avec le principe de modélisation linéaire, nous avons établi un modèle qui sera ensuite appliqué pour le calcul des prévisions. La comparaison entre les résultats obtenus et les valeurs réelles permet d'apprécier l'impact de la crise sanitaire sur les deux terminaux.

Mots clés : Entreprise Portuaire de Bejaia ; Modélisation statistique ; Prévision ; Covid-19.

2.1 Introduction

La statistique est une science qui permet l'exploitation d'une masse d'informations provenant de l'observation d'un phénomène. Elle concerne un ensemble de méthodes mathématiques indispensables à l'analyse de grands ensembles de données et dont l'objectif est de fournir les éléments nécessaires à l'appréciation d'une situation et l'aide à la décision. Il existe, de nos jours, un ensemble de méthodes rigoureuses permettant d'analyser et de traiter des données concernant un phénomène et réaliser une étude qui permettra de comprendre son évolution et effectuer des prévisions, notamment sur les séries temporelles et régression linéaire. L'étude des séries temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Cette étude est appliquée dans plusieurs domaines, en météorologie, en économie, etc. Elle consiste à appliquer les méthodes de prévision. On distingue notamment, deux méthodes qui sont souvent utilisées, la méthode de Box et Jenkins (1976) et le lissage exponentiel. La première est très populaire grâce à sa simplicité et ses résultats fiables. La deuxième méthode a l'avantage de fournir une prévision "bon marché" et parfois très satisfaisante, mais les deux méthodes sont parfois loin d'être les mieux adaptées. L'analyse de régression peut être définie comme la recherche de la relation stochastique qui lie deux ou plusieurs variables. Son champ d'application recouvre de multiples domaines, parmi lesquels on peut citer la physique, l'astronomie, la biologie, la chimie, la médecine, la géographie, la sociologie, l'histoire, l'économie, la linguistique et le droit. La régression est

l'une des méthodes les plus connues et les plus appliquées en statistique pour l'analyse de données quantitatives. Elle est utilisée pour établir une liaison entre une variable quantitative et une ou plusieurs autres variables quantitatives, sous la forme d'un modèle. Si on s'intéresse à la relation entre deux variables, on parlera de régression simple en exprimant une variable en fonction de l'autre. Si la relation porte entre une variable et plusieurs autres variables, on parlera de régression multiple. La régression linéaire simple et multiple font une classe particulière de modèles de régression. Le but est d'expliquer une variable dépendante, appelée variable endogène par une ou plusieurs variables explicatives, dites exogènes, à travers une fonction affine. Afin d'avoir une idée précise et claire de la manière de la mise en œuvre des outils statistiques, nous avons effectué un stage pratique à la direction exploitation de L'EPB, au niveau du service statistique, où nous avons pu mettre en pratique nos connaissances théoriques acquises dans la matière. Durant notre stage pratique nous nous sommes intéressés aux données statistiques des deux terminaux : le terminal hydrocarbure et le terminal à conteneur, qui figurent parmi les principaux terminaux du port, pour voir l'impact de la pandémie Covid-19 sur ces deux potentiels, et cela sur la base de modélisations et des prévisions, en faisant appel aux deux méthodes ; la méthode de box et Jenkins et la méthode de la régression linéaire.

2.2 Position du problème

Le nombre des navires qui ont accosté au port de Béjaïa a atteint 1.573 navires contre 1.974 navires en 2019, cette baisse est induite, explique l'entreprise publique, "par la diminution du mouvement de la navigation durant cette crise sanitaire dont nous subissons les effets depuis mars 2020, notamment après la fermeture des frontières, la baisse de la production mondiale et des échanges commerciaux internationaux".

Le terminal à conteneur, comme le terminal hydrocarbure, sont les terminaux les plus actifs dans le mouvement de la navigation de port de Béjaïa, les quantités de tonnage qui s'engendrent par le trafic des hydrocarbures connaissent une variation très élevée avant la crise sanitaire, bien précis, avant mars 2020, mais depuis, ces quantités ont connues un contraste remarquable, ce qui a causé des perturbations sur la gestion de cette marchandise à l'heure actuelle, et des difficultés de planification future pour le terminal hydrocarbure, en raison du manque d'informations présentes, et aussi l'incertitude de l'information future.

Pour opter à une planification adéquate, la question qui se pose est de savoir ; sur la base des prévisions, quelles sont les quantités de tonnage qui seront engendrées par le trafic des hydrocarbures dans cette crise sanitaire pour l'année 2021 ?, et quel est l'impact de la crise sanitaire sur le trafic à conteneur et le trafic des hydrocarbures durant l'année 2020 ?

2.3 Modélisation et résolution

Cette section concerne en premier temps l'étude statistique descriptive des données recueillies au niveau de EPB. Ensuite nous réalisons une étude prévisionnelle par la méthode

de Box et Jenkins sur la quantité en tonne engendrée par le trafic des hydrocarbures .

La deuxième partie de cette section est consacrée à l'application numérique d'une régression linéaire multiple, avec les données qu'on a pu récupérer au sein de service statistique de l'EPB qui concerne le nombre des navires traités (les navires qui transportent les conteneurs) durant l'année 2019.

Le logiciel utilisé pour la réalisation de nos applications dans ce chapitre est le logiciel R.

2.3.1 Description des données

Lors de notre stage effectué au sein de l'entreprise EPB, nous avons recueilli les données concernant les quantités en tonne engendrées par le trafic des hydrocarbures de 2014 jusqu'à 2019 et le nombre de navires traités durant l'année 2019. Les données collectées sont représentées dans les tableaux ci-dessous.

Collecte de données

Les données dont nous disposons proviennent des annuaires statistiques établis par le service d'exploitation de la Direction de l'EPB, qui regroupe les données des tonnages engendré par le trafic des hydrocarbures de 2014 jusqu'à 2019. Après la collecte de ces données, nous avons regroupé ces données dans un tableau de 12 mois sur 6 ans. Dans le but de prévoir le tonnage qui sera engendré pour l'année 2020, nous appliquons la méthode de Box et Jenkins.

Mois \ Année	2019	2018	2017	2016	2015	2014
jan	1.2	1.6	1.9	1.9	2	2.8
Fév	2.5	2.5	2.6	2	1.7	2.5
Mar	1.3	1.3	1.3	1.5	2.8	1.8
Avr	2	1.8	1.5	1.3	1	1.115
Mai	1	1.5	1.3	1.4	1.3	1.4
Jun	1.2	1.6	1.2	1.16	1.21	1.2
Jui	1.3	1.5	1.2	1.2	1.3	1.3
Aou	1.24	1.32	1.5	1.3	1.5	1.5
Sept	1.5	1.2	1.6	1.7	1.2	1.7
Oct	2	1.3	1.2	1.5	1.7	1.9
Nov	1.9	2.5	1.5	2.5	1.8	1.9
Déc	2	2	2	2	2.5	1

TABLE 2.1: Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

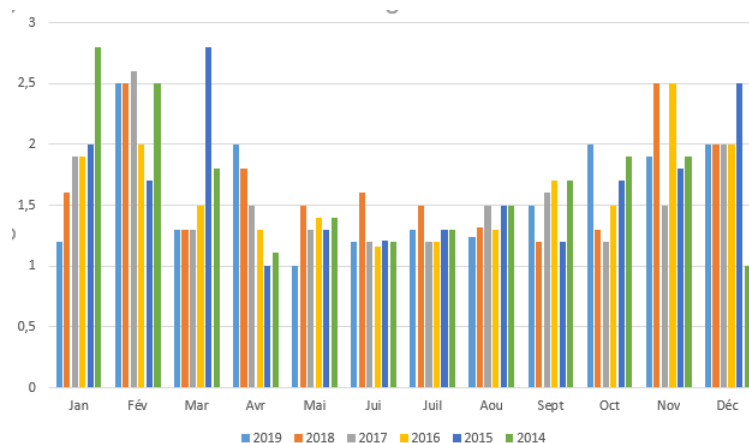


FIGURE 2.1: Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

D'après la figure (1), les valeurs de tonnage mensuels de 2014 jusqu'à 2019, c'est-à-dire avant la crise de Covid- 19, se varient dans un intervalle de 1 million de tonnes jusqu'à 2 million de tonnes en moyenne par mois, on remarque qu'elle atteint la valeur 2.8 million de tonnes en mars 2019, dans l'ensemble, si on varie la quantité totale qui s'engendre par année on trouve 18 million de tonnes en moyenne chaque ans.

Mois,Année	2020	2019	2018	2017	2016	2015
Jan	2	1.2	1.6	1.9	1.9	2
Févr	1.7	2.5	2.5	2.6	2	1.7
Mar	1.2	1.3	1.3	1.3	1.5	2.8
Avr	1.15	2	1.8	1.5	1.3	1
Mai	1.2	1	1.5	1.3	1.4	1.3
Jui	1.1	1.2	1.6	1.2	1.16	1.214
Juil	1.2	1.3	1.5	1.2	1.2	1.3
Juil	1.2	1.3	1.5	1.2	1.2	1.3
Aou	1.2	1.244	1.32	1.5	1.3	1.5
Sept	1.1	1.5	1.2	1.6	1.7	1.2
Oct	1.4	2	1.3	1.2	1.5	1.7
Nov	2	1.9	2.5	1.5	2.5	1.8
Déc	1	2	2	2	2	2.5

TABLE 2.2: Les quantités de tonnage engendrées par le trafic des hydrocarbures (2015-2020)

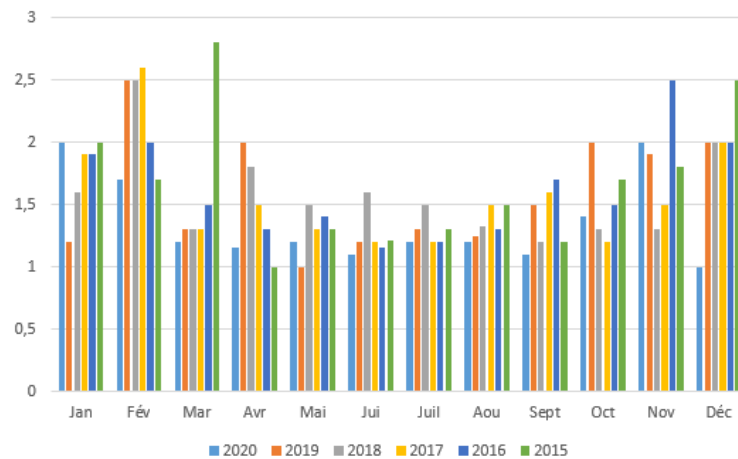


FIGURE 2.2: Tonnage engendré par le trafic des hydrocarbures (2015-2020)

D'après la figure (2), on constate que les valeurs de tonnage mensuels de 2015 jusqu'à 2019 se variaient dans un intervalle de 1 million de tonnes jusqu'à 2 millions de tonnes en moyenne par mois, bien avant la crise sanitaire, en revanche, pour l'année 2020, on remarque une baisse dans les quantités de tonnage. Cependant, on a enregistré une quantité de 15.7 millions de tonnes pour l'année 2020.

Temps(mois)	Importation(EVP)	Exportation(EVP)	Nombrenavires	Livraison
Jan	7080	950	25	6225
Fév	6200	1500	18	6030
Mar	7400	800	25	6100
Avr	6500	800	20	6500
Mai	7200	500	22	6420
Jui	5600	600	23	6325
Juil	6500	1000	26	4090
Aou	7100	750	27	6380
Sep	5300	650	21	6020
Oct	6500	756	19	5400
Nov	6060	850	25	5990
Déc	6550	990	22	7080

TABLE 2.3: Les nombre de navires (porte-conteneurs) traités durant l'année 2019

2.3.2 Prévision par de la méthode de Box-Jenkins(2020)

Dans cet exemple, on s'intéresse à effectuer une prévision sur le tonnage qui sera engendré par le trafic des hydrocarbures pour l'années 2020 à partir des données sur la période 2014-2019 (tableau 1). Pour cela on applique la méthode de Box-Jenkins dont le principe est d'étudier systématiquement une série chronologique à partir de leurs caractéristiques. L'objectif est de chercher dans la famille ARIMA et SARIMA le modèle le plus adapté pour représenter le phénomène étudié.

2.3.3 Présentation graphique des données

Afin de tracer le graphe de la série originale, il faut introduire les données sous forme d'un vecteur dans le logiciel R puis faire appel aux commandes suivantes qui permettent de le rendre sous forme d'une série chronologique.

```
> y = c(1.2, 2.5, 1.3, 2, 1, 1.2, 1.3, 1.244, 1.5, 2, 1.9, 2, 1.6, 2.5, 1.3, 1.8, 1.5, 1.6, 1.5, 1.32, 1.2, 1.3, 2.5, 2, 1.9, 2, 1.5, 1.3, 1.4, 1.116, 1.2, 1.3, 1.7, 1.5, 2.5, 2, 2, 1.7, 2.8, 1, 1.3, 1.214, 1.3, 1.5, 1.2, 1.7, 1.8, 2.5, 2.8, 2.5, 1.8, 1.115, 1.4, 1.2, 1.3, 1.5, 1.7, 1.9, 1.9, 1)
> serie=ts(y,start=c(2014,1),end=c(2019,12),frequency=12)
```

La commande qui nous permet de visualiser le graphe de la série originale est :

```
> plot(serie)
```

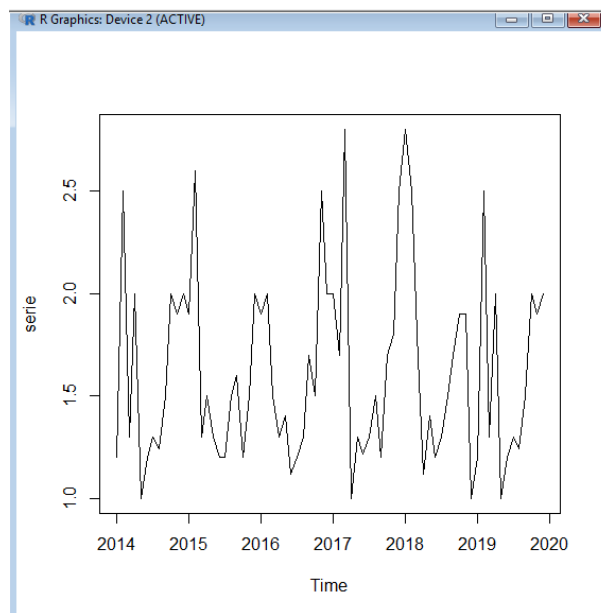


FIGURE 2.3: Le graphe de la série originale

Le graphe de la série originale figure(3) montre que celle-ci comporte une tendance et une composante saisonnière, pour illustrer cela nous allons décomposer la série originale en ses différentes composantes. Nous utilisons les commandes :

```
> domp=decompose(serie,type=c("additive","multiplicative"),filter=NULL)
> plot(domp)
```

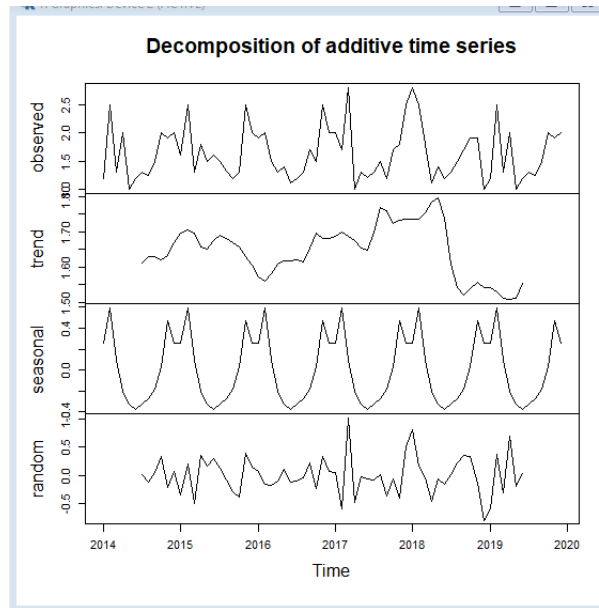


FIGURE 2.4: Graphe de la série et ses différentes composantes.

2.3.4 Analyse préliminaire

Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série originale sont :

- > acf(serie)
- > pacf(serie)

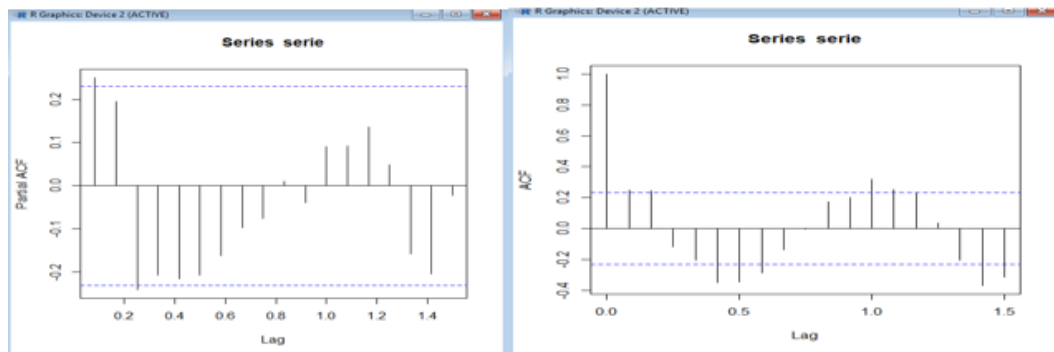


FIGURE 2.5: Corrélogramme et corrélogramme partiel de la série originale

2.3.5 Stationnarité

L'examen des deux graphes ACF et PACF nous suggèrent l'existence d'une tendance donc il sera nécessaire de la différencier, ainsi qu'une composante saisonnière, donc une différenciation d'ordre 12 sera nécessaire pour enlever cette saisonnalité.

L'instruction suivante nous permet d'affaiblir la variance de la série. On utilise la transformation en logarithme pour rendre la série stationnaire, en appliquant une différence première ordinaire et une différence saisonnière d'ordre 12. Les commandes qui nous permettent d'obtenir la série

chronologique differencier sont les suivantes :

```
> dift=diff(serie)
> difs=diff(dift,frequency=12)
> plot(difs)
```

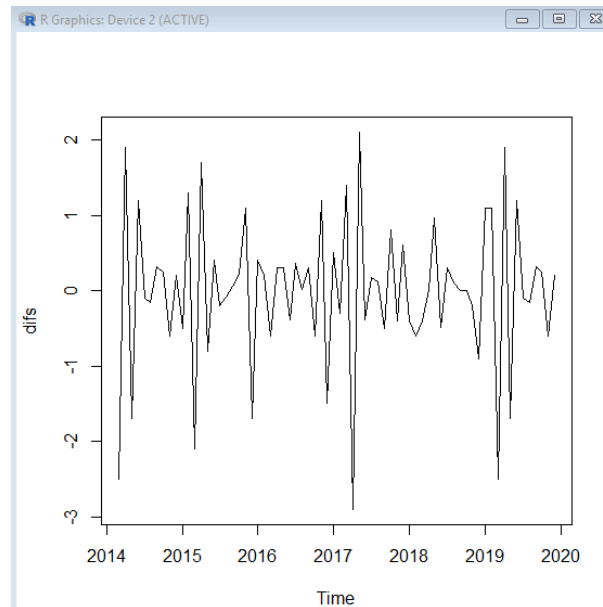


FIGURE 2.6: Graphe de la série chronologique differencier

2.3.6 Identification du modèle

D'après le graphe de la série originale, on constate que cette série est sujet a une saisonalite d'ordre 12 (12 mois) et une tendance. Donc le modèle adéquat est $SARIMA(p, d, q)(P, D, Q)_{(s=12)}$. En appliquant à la série originale un filtre ordinaire d'ordre 1 ($d = 1$), et un filtre saisonnier d'ordre 1 ($D = 1$), on a obtenu une série sans tendance ni saisonnalité.

Après avoir tester différents modèles, nous avons opté pour le modèle $SARIMA(1, 1, 0)(0, 1, 1)_{12}$. Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série transformée sont :

```
> acf(difs)
> pacf(difs)
```

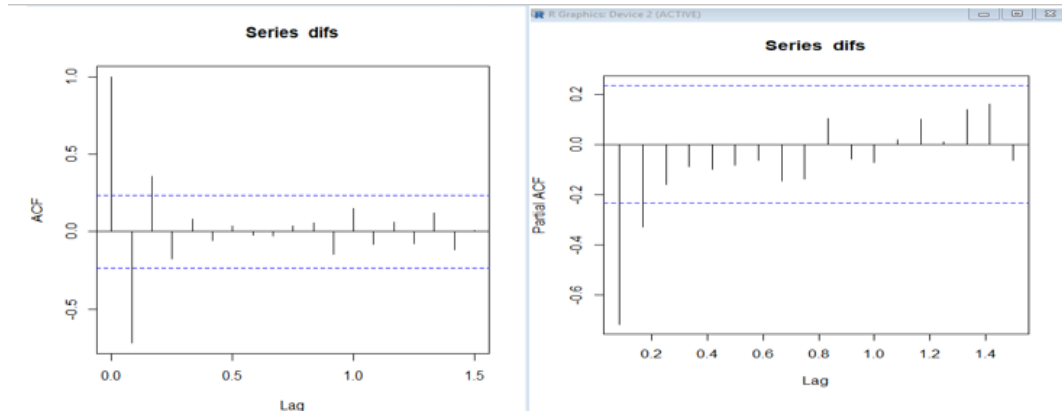



FIGURE 2.7: Corrélogramme et corrélogramme partiel de la série differencier

2.3.7 Estimation des paramètres du modèle

Après avoir opté pour un modèle SARIMA(1, 1, 0)(0, 1, 1)₁₂ la commande :
`> sarima=arima(s, order = c(1, 1, 0), seasonal = list(order = c(0, 1, 1), period = 12)),`
 permet d'effectuer l'ajustement de la série et d'obtenir les résultats suivants : Les estimateurs des paramètres du modèle sont :

$$\hat{\phi}_1 = -0.6213.$$

$$\hat{\theta}_s = -0.5808.$$

Les écarts type des paramètres estimés sont :

$$\hat{\sigma}(\hat{\phi}_1) = 0.1273.$$

$$\hat{\sigma}(\hat{\theta}_s) = 0.2269.$$

Nous remarquons que tous les estimateurs vérifient les conditions de stationnarité :

$|\hat{\phi}| < 1, |\hat{\theta}_s| < 1$ et d'inversibilité.

Le modèle s'écrit :

$$(1 - \hat{\phi}_1 B)(1 - B)(1 - B^{12})y_t = (1 - \hat{\theta}_s B^{12})\epsilon_t$$

.

Alors

$$(1 + 0.6213B)(1 + B^{12})y_t = (1 + 0.5808B^{12})\epsilon_t.$$

2.3.8 Validation du modèle

Tests sur les coefficients du modèle

La valeur critique lue sur la table de student est $T(72, 0.025) = 2,39$

1. Test de $H_0 : \phi_1 = 0^n$ contre $H_1 : \phi_1 \neq 0^n$ Ce test est basé sur la statistique $T_s = \frac{|\hat{\phi}_1|}{\hat{\sigma}_{\phi_1}} \longrightarrow t(n-2, \alpha/2)$
sa réalisation $t_1 = \frac{0.6213}{0.1273} = 4.88 > 2.39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_1 est significativement différent de zéro.
2. $H_0 : \theta_s = 0$ contre $H_1 : \theta_s \neq 0$ Ce test est basé sur la statistique :
 $T_s = \frac{|\hat{\theta}_s|}{\hat{\sigma}_{\theta_s}} \longrightarrow t(n-2, \alpha/2)$
sa réalisation $t_1 = \frac{0.5808}{0.2269} = 2.63 > 2.39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre θ_s est significativement différent de zéro

Test sur les résidus

Ce test est basé sur la statistique suivante :

$$Q = n(n+2) \sum_{i=1}^n \frac{p_i^2}{n-i} \longrightarrow \chi_{n-k}^2$$

Test : $H_0 : \rho_1 = \rho_2 = \dots = 0, i = 1, \dots, n$ contre $H_1 : \exists i \in 1 \dots, n; \rho_i \neq 0^n$,

où ρ_i désignent les coefficients d'autocorrélations théoriques,

M est la partie entière de : $\text{Min}(n/2, 3\sqrt{n})$ donc $n = 25$.

L'instruction :

> Box.test(residu1, lag=23, type="Ljung") nous donne la valeur empirique de la statistique de Box et Ljung, $Q^* = 13,6550$. Au niveau de signification $\alpha = 0.05$ la valeur critique lue sur la table de khi-deux est :

$$\chi_{M-k}^2(\alpha) = \chi_{25-2-1}^2(0,05) = \chi_{22}^2(0,05) = 26,245$$

Comme $Q^* = 13,6550 < \chi_{22}^2(0,05) = 26,245$, alors on accepte l'hypothèse H_0 , par conséquent les résidus forment un bruit blanc, le corrélogramme et le corrélogramme partiel des résidus nous montrent clairement que tous les pics sont à l'intérieure de l'enveloppe. Les commandes sous R qui nous permettent de les représenter sont :

```
> residu=residuals(sarima)
> residu1=ts(residu,start=2010,frequency=12)
> acf(residu1)
> pacf(residu1)
```

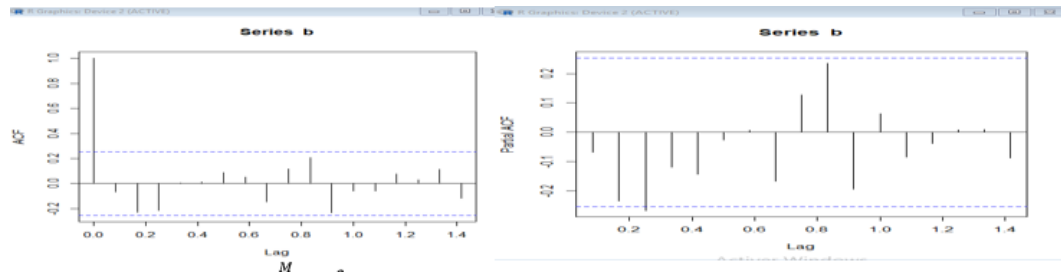


FIGURE 2.8: Corrélogramme et corrélogramme partiel des résidus.

2.3.9 Prédiction

Les commandes qui nous permettent d'obtenir les prévisions pour l'année 2020 sont :

```
> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
```

```
|
> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
Time Series:
Start = 61
End = 72
Frequency = 1
[1] 1.479052 1.357856 1.677086 1.565544 1.667962 1.573922 1.660269 1.580985 1.653783 1.586940 1.648316 1.591961
```

Le tableau ci-dessus nous montre les résultats de la prédiction de tonnage engendré par le trafic des hydrocarbures (unité : million tonnes) obtenus avec le logiciel R est les valeurs réelles de l'année 2020 :

mois (année 2020) \ année	Valeurs prévues	Valeurs réelles
jan	1.479052	2
Fév	1.357856	1.7
Mar	1.677086	1.2
Avr	1.565544	1.15
Mai	1.667962	1.2
Jun	1.573922	1.1
Jui	1.660269	1.2
Aou	1.653783	1.2
Sept	1.580985	1.1
Oct	1.653783	1.4
Nov	1.648316	2
Déc	1.591961	1

TABLE 2.4: Les résultats de la prévision de tonnage engendré par le trafic des hydrocarbures pour 2020

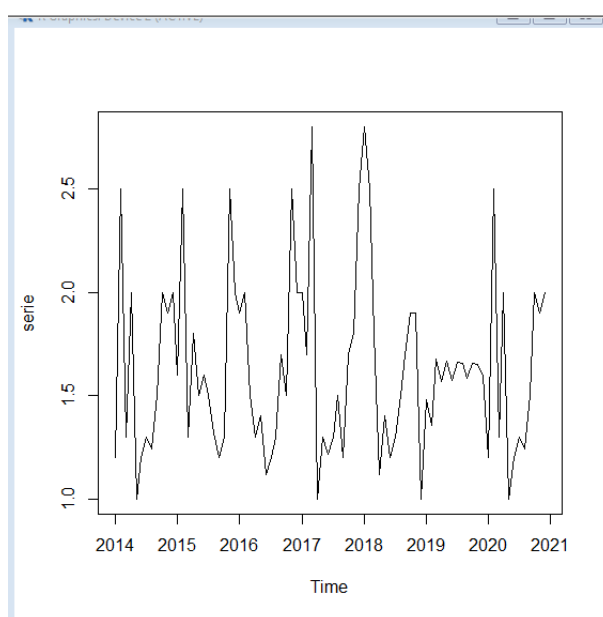


FIGURE 2.9: Graphe de la série et la prévision

2.3.10 Interprétation des résultats

D'après le tableau, les prévisions données par Box-Jenkins concernant les quantités de tonnage, montrent qu'il y a un grand écart avec les valeurs réelles, en revanche, si on fait la somme des quantités prévues pour l'année 2020 on obtient une quantité de 19.11 millions de tonnes, sachant que qu'à cette année le port a enregistré au total 15.7 millions de tonnes, cela s'explique qu'il y a plusieurs facteurs influents, parmi ces facteurs la crise de Corona virus qui a causé un grand recul dans le secteur de l'import et l'export.

cette prévision effectué nous éclaircit sur l'effet de cette crise sur trafic des hydrocarbures, on remarque qu'une perte de 4 million de tonnes induite de ce recul, sinon une quantité de 19.11 million de tonnes sera recueillie si les opérations import-expot n'ont été pas touchées par la crise sanitaire.

2.4 Prédiction par la méthode de Box-Jenkins (2021)

Dans cet exemple, nous allons prédire les quantités de tonnage engendrées par le trafic des hydrocarbures pour 2021. Nous utilisons les données réelles de 2020.

2.4.1 Présentation graphique des données

Afin de tracer le graphe de la série originale, il faut introduire les données sous forme d'un vecteur dans le logiciel R puis appliquer les commandes suivantes qui permettent de le rendre sous forme d'une série chronologique.

```
> m = c(2, 1.7, 1.2, 1.15, 1.2, 1.1, 1.2, 1.2, 1.1, 1.4, 2, 1, 1.9, 2.6, 1.3, 1.5, 1.3, 1.2, 1.2, 1.5, 1.6, 1.2, 1.5, 2, 1.9, 2, 1.3, 1.4, 1.116, 1.2, 1.3, 1.7, 1.5, 2.5, 2, 2, 1.7, 2.8, 1, 1.3, 1.214, 1.3, 1.5, 1.2, 1.7, 1.8, 2.5, 2.8, 2.5, 1.8, 1.115, 1, 1.2, 1.3, 1.5, 1.7, 1.9, 1.9, 1)
```

```
> serie = ts(m, start = c(2015, 1), end = c(2020, 12), frequency = 12)
```

La commande qui nous permet de visualiser le graphe de la série originale est :

```
> plot(serie)
```

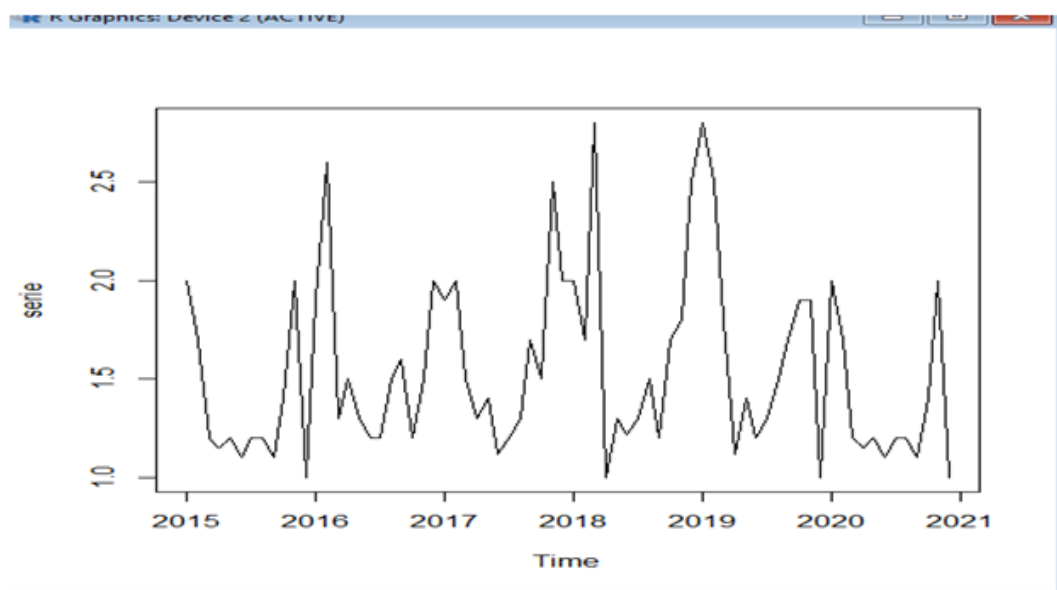


FIGURE 2.10: Graphe de la série originale

Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série originale sont :

```
> acf(serie)
```

```
> pacf(serie)
```

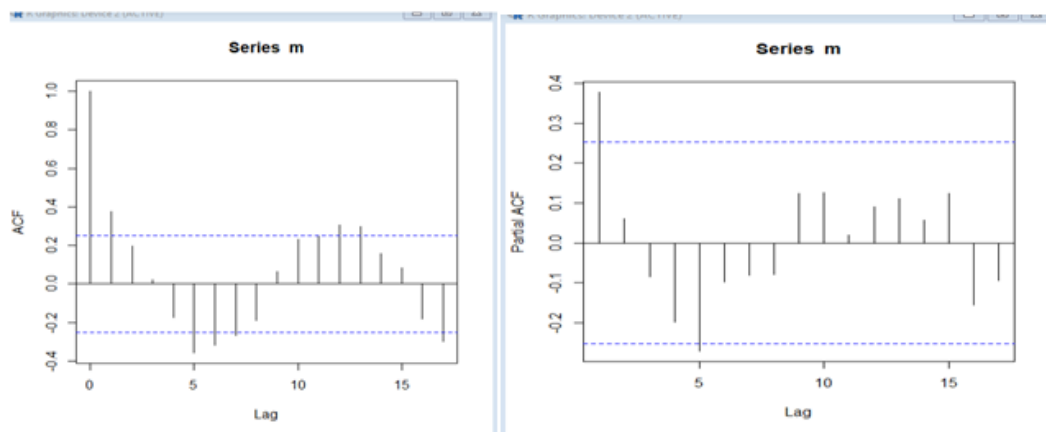


FIGURE 2.11: Corrélogramme et corrélogramme partiel de la série originale

2.4.2 Stationnarité

D'après les deux graphes de L'ACF et le PACF, la série que nous disposons est une série non stationnaire. Pour la rendre stationnaire, on doit éliminer sa saisonnalité et sa tendance, pour se faire nous utilisons la commande.

```
> adf.test(m, alternative="stationary")
```

```
> library("tseries")
Registered S3 method overwritten by 'quantmod':
  method      from
as.zoo.data.frame zoo

'tseries' version: 0.10-48

'tseries' is a package for time series analysis and computational
finance.

See 'library(help="tseries")' for details.

> adf.test(m, alternative="stationary")

Augmented Dickey-Fuller Test

data: m
Dickey-Fuller = -4.1215, Lag order = 3, p-value = 0.01038
alternative hypothesis: stationary
```

on remarque que la propriété de la stationnarité est vérifiée,

2.4.3 Identification du modèle

Nous avons une série chronologique stationnaire, les modèles qui représente ce type de série sont AR(p), MA(q) et ARMA(p,q), après avoir testé différents modèles sous R, nous avons opté pour le modèle ARIMA(1 ; 0 ; 1) qui s'écrit aussi ARMA(1,1). Les commandes qui nous permettent d'identifier le modèle sont :

```
> mydata.arima111=arima(y,order=c(1,0,1))
> mydata.arima111
```

```
> arima(m,order=c(1,0,0))

Call:
arima(x = m, order = c(1, 0, 0))

Coefficients:
      ar1  intercept
      0.3874      1.575
s.e.  0.1202      0.088

sigma^2 estimated as 0.178:  log likelihood = -33.44,  aic = 72.87
> arima(m,order=c(1,0,1))

Call:
arima(x = m, order = c(1, 0, 1))

Coefficients:
      ar1      ma1  intercept
      0.4543 -0.0790      1.5753
s.e.  0.2330  0.2492  0.0907

sigma^2 estimated as 0.1777:  log likelihood = -33.39,  aic = 74.78
> arima(m,order=c(1,0,2))

Call:
arima(x = m, order = c(1, 0, 2))

Coefficients:
      ar1      ma1      ma2  intercept
      0.3542  0.0017  0.1129      1.5757
s.e.  0.3214  0.3282  0.1581  0.0924

sigma^2 estimated as 0.1763:  log likelihood = -33.16,  aic = 76.33
> arima(m,order=c(1,0,3))

Call:
arima(x = m, order = c(1, 0, 3))
```

Cette technique est basée sur le critère d'AIC (Akaike Information Criterion),qui est calculée comme suit :

$$AIC = \ln(SCR/n) + (2p/n)$$

Tel que :

Ln : logarithme népérien.

SCR : la somme des carrés des résidus du modèle.

n : nombre d'observation.

P : le nombre de variable explicatives.

Premièrement on calcule l'AIC pour le modèle global et aussi pour chaque variable, puis nous éliminons les variables dont leurs AIC est petit, jusqu'à ce qu'on obtient le modèle approprié.

2.4.4 Prédiction

Les commandes qui nous permettent d'obtenir les prévisions pour l'année 2021 sont :

```
> mydata.pred1 = predict(mydata.arima111,n.ahead = 12)
> mydata.pred1pred
```

```

> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
Time Series:
Start = 61
End = 72
Frequency = 1
[1] 1.369764 1.481894 1.532839 1.555986 1.566503 1.571281 1.573452 1.574438
[9] 1.574886 1.575090 1.575182 1.575225

```

Le tableau ci-dessous représente les résultats de la prévision de tonnage des hydrocarbures pour l'année 2021 :

Mois(année2021)	Valeurs prévues
Jan	1.369764
Fév	1.481894
Mar	1.532839
Avr	1.555986
Mai	1.566503
Jui	1.571281
Juil	1.573452
Aou	1.574438
Sep	1.574886
Oct	1.575090
Nov	1.575182
Déc	1.575225

TABLE 2.5: Les résultats de la prévision de tonnage des hydrocarbures pour l'année 2021

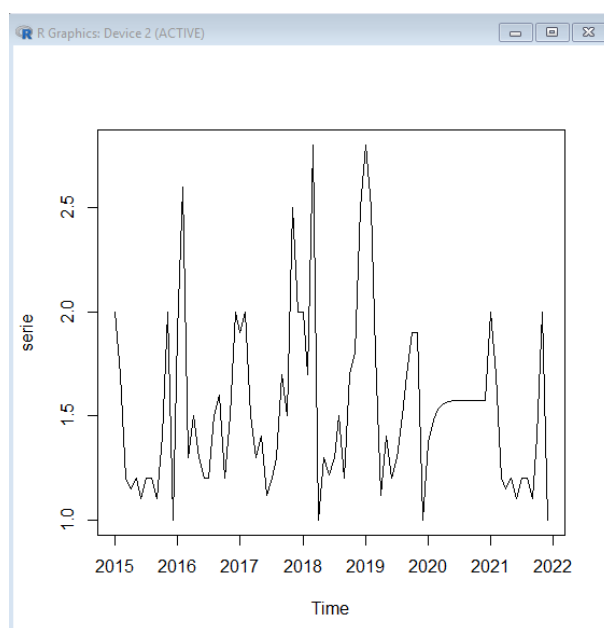


FIGURE 2.12: Graphe de la série et la prévision

2.4.5 Interprétation des résultats

D'après le tableau, les prévisions données par Box-Jenkins concernant les quantités de tonnage, nous donne une vision approximative pour l'année 2021.

Cependant, selon les résultats obtenus, le port de Béjaia enregistrera une quantité de 18.5 million de tonnes qui sera engendrée par le trafic des hydrocarbures en 2021, cela assure une hausse considérable de cette dernière par rapport à la quantité enregistrée en 2020.

Le premier exemples que nous avons traité nous fournis des informations importantes sur les quantités de tonnage des hydrocarbures pour les deux années 2020, cependant, les résultats obtenus nous confirme l'exactitude de la prévision en comparant les valeurs obtenues avec les valeurs réelles, cela nous confirme l'idée de la réalisation d'une prévision pour l'année 2021 dans le deuxième exemple.

L'étude réalisée dans le deuxième exemple nous donne une vision approximative sur les quantités de tonnages qui seront engendrées par le trafic des hydrocarbures en 2021, les résultats obtenus de cette prévision assurent une hausse des tonnages, ce qui permet aux responsables du port de Béjaia d'opter à une planification futur dans les secteurs concernés.

2.5 Prévision par la méthode de régression linéaire

Cette partie est consacrée à l'application numérique d'une régression linéaire multiple, avec les données qu'on a pu récupérer au sein de service statistique de l'EPB qui concerne le nombre des navires traités (les navires qui transportent les conteneurs) durant l'année 2019 (tableau 4.3). Nous avons essayé d'expliquer le nombre de navires traités en fonction de certain nombre de facteurs (Exportation, importation, temps, livraison) ensuite réaliser une prévision pour l'année 2020, tout en appliquant une régression linéaire multiple.

2.5.1 Modèle de régression

Pour formuler le modèle adéquat pour les choix des variables à expliquer, nous utilisons la méthode d'élimination (Back-Ward), cette méthode permet de sélectionner parmi les variables suggérées précédemment, celles les plus appropriées à prendre en considération pour notre modèle. Nous commençons par la création du tableau d'observation dans R..

```

> Importation=c(7080,6200,7400,6500,7200,5600,6500,7100,5300,650,6060,6550)
> Exportation=c(950,1500,800,800,500,600,1000,750,650,756,850,990)
> NbrNavire=c(25,18,25,20,22,23,26,27,21,19,25,22)
> Livraison=c(6025,6030,6100,6500,6420,6325,4090,6980,6020,5400,5990,7090)
> Temps=c(1:12)
> D=data.frame(Importation,Exportation,Livraison,NbrNavire,Temps)
> D
  Importation Exportation Livraison NbrNavire Temps
1         7080          950       6025         25     1
2         6200         1500       6030         18     2
3         7400          800       6100         25     3
4         6500          800       6500         20     4
5         7200          500       6420         22     5
6         5600          600       6325         23     6
7         6500         1000       4090         26     7
8         7100          750       6980         27     8
9         5300          650       6020         21     9
10          650          756       5400         19    10
11         6060          850       5990         25    11
12         6550          990       7090         22    12
~ |

```

Maintenant sous R nous allons faire appel la fonction `step()` qui sert à appliquer la méthode d'élimination comme suit :

```

> step(lm(NbrNavire~Importation+Exportation+Livraison+Temps,data=D),direction="backward")
Start: AIC=25.13
NbrNavire ~ Importation + Exportation + Livraison + Temps

      Df Sum of Sq  RSS   AIC
- Temps      1    7.359 49.689 25.050
<none>                 42.330 25.127
- Exportation 1    8.205 50.535 25.253
- Livraison   1   12.944 55.274 26.329
- Importation 1   42.826 85.156 31.515

Step: AIC=25.05
NbrNavire ~ Importation + Exportation + Livraison

      Df Sum of Sq  RSS   AIC
<none>                 49.689 25.050
- Livraison   1   10.188 59.877 25.289
- Exportation 1   12.191 61.880 25.683
- Importation 1   35.488 85.177 29.518

Call:
lm(formula = NbrNavire ~ Importation + Exportation + Livraison,
    data = D)

Coefficients:
(Intercept)  Importation  Exportation  Livraison
 28.107265      0.001059     -0.004241     -0.001339

```

Cette technique est basée sur le critère d'AIC(Akaike Information Criterion),utilisée dans l'exemple précédent. Premièrement on calcule l'AIC pour le modèle global et aussi pour chaque variable, puis nous éliminons les variables dont leurs AIC est petit, jusqu'à ce qu'on obtient le modèle approprié.

Dans notre cas l'AIC de modèle global= 25 :13, on compare l'AIC pour chaque variable, et on sélectionne l'AIC le plus petit. On remarque que l'AIC le plus petit est AIC = 25 :050, qui correspond à l'AIC de la variable temps, on le compare avec l'AIC global on trouve que $AIC(\text{temps}) < AIC \text{ global}$, donc on supprime la variable temps. D'où, les variables retenues sont : Importation, Exportation, Livraison.

Le modèles s'écrit :

$$NbrVessel = \beta_0 + \beta_1 Importation + \beta_2 Exportation + \beta_3 Livraison + \epsilon$$

D'après la fonction `step()` function, on obtient les estimateurs :

$$\hat{\beta}_0 = 28,1, \quad \hat{\beta}_1 = 0,001, \quad \hat{\beta}_2 = -0,004, \quad \hat{\beta}_3 = -0,001$$

alors

$$Nbr\hat{V}essel = 28.1 + (0.001)Import + (-0.004)Export + (-0.001)Livraison.$$

2.5.2 Tests statistiques de Student et Fisher

Pour effectuer les tests de Student et Fisher, nous utilisons la fonction `summary()` dans R.

```
> summary(lm(NbrNavire~Importation+Exportation+Livraison+Temps,data=D))

Call:
lm(formula = NbrNavire ~ Importation + Exportation + Livraison +
    Temps, data = D)

Residuals:
    Min       1Q   Median       3Q      Max
-2.6024 -1.7405  0.2027  0.9863  3.4892

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.5968373  7.2411876   3.535  0.00953 **
Importation  0.0012916  0.0004854   2.661  0.03241 *
Exportation -0.0035536  0.0030509  -1.165  0.28225
Livraison   -0.0015299  0.0010457  -1.463  0.18686
Temps       0.2609059  0.2365165   1.103  0.30645
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.459 on 7 degrees of freedom
Multiple R-squared:  0.5411,    Adjusted R-squared:  0.2789
F-statistic: 2.064 on 4 and 7 DF,  p-value: 0.1894
```

Le coefficient de détermination $R^2 = 0,54$, nous avons donc un bon ajustement linéaire.

2.5.3 Test global de Fisher

Nous testons les hypothèses suivantes :

$$H_0 : \beta_0 = 0; \hat{\epsilon}a_1 = 0; \hat{\epsilon}a_2 = 0; \beta_3 = 0;$$

contre

$$H_1 : \beta_0 \neq 0; \beta_1 \neq 0; \beta_2 \neq 0; \beta_3 \neq 0.$$

Ce test est basé sur la statistique de Fisher (statistique F) F^* :

Nous avons $F^* = 3.535$, au niveau de signification de $\alpha = 0.05$ la valeur sur la table de Fisher est $F(4, 9, 0.05) = 2.60$, comme $F^* > F(4, 9, 0.05)$ nous rejetons H_0 , donc le modèle est globalement significatif.

2.5.4 Test étudiant individuel

Au seuil de signification $\alpha = 0,05$ nous avons :

Pour $\hat{\beta}_0$

$$H_0 : \hat{\beta}_0 = 0;$$

Contre

$$H_1 : \hat{\beta}_0 \neq 0$$

$$t(0.05; 9) = 3.535 > T_{\hat{\beta}_0} = 0.00953.$$

Conclusion

Nous acceptons l'hypothèse H_1 .

Pour $\hat{\beta}_1$

$$H_0 : \hat{\beta}_1 = 0;$$

Contre

$$H_1 : \hat{\beta}_1 \neq 0.$$

$$t(0.05; 9) = 2.661 > T_{\hat{\beta}_1} = 0.03241$$

Conclusion

Nous acceptons l'hypothèse H_1 .

Pour $\hat{\beta}_2$

$$H_0 : \hat{\beta}_2 = 0;$$

Contre

$$H_1 : \hat{\beta}_2 \neq 0.$$

$$t(0.05; 9) = -1.165 > T_{\hat{\beta}_2} = 0.28225.$$

Conclusion

Nous rejetons l'hypothèse H_0 .

Pour $\hat{\beta}_3$

$$H_0 : \hat{\beta}_3 = 0;$$

Contre

$$H_1 : \hat{\beta}_3 \neq 0.$$

$$t(0.05; 9) = -1.463 > T_{\hat{\beta}_3} = 0.18686.$$

Conclusion

Nous acceptons l'hypothèse H_1 .

2.5.5 Intervalle de confiance

Nous calculons l'intervalle de confiance pour les quatre coefficients avec la fonction confint() sous R. On obtient :

L'intervalle de confiance pour $\hat{\beta}_0$: $IC_0 = [1.204171e + 01; 44.172817408]$;

L'intervalle de confiance pour $\hat{\beta}_1$: $IC_1 = [3, 736097 \exp(-5); 0.002080983]$;

L'intervalle de confiance pour $\hat{\beta}_2$: $IC_2 = [-1.122047 \exp(-2); 0.002739299]$;

L'intervalle de confiance pour $\hat{\beta}_3$: $IC_3 = [-3, 748486 \exp(-3); 0.001071472]$.

Nous avons : $\hat{\beta}_0 \in IC_0$, $\hat{\beta}_1 \in IC_1$, $\hat{\beta}_2 \in IC_2$, $\hat{\beta}_3 \in IC_3$.

Ensuite, le modèle est validé.

2.5.6 Prévision

Afin de faire une prévision du nombre de navires qui seront traités pour l'année 2020, nous utilisons la fonction predict() sous R.

Ci-dessus, les résultats approximatifs des prévisions :

Le tableau suivant présente les résultats de la prévision et le nombre réel de navires traités en 2020 :

Mois	Nombre de navires prévus en 2020	Nombre de navires réel en 2020
Jan	23	22
Fév	20	22
Mar	24	25
Avr	23	23
Mai	25	27
Jui	23	20
Juil	25	19
Aou	23	29
Sep	23	25
Oct	18	21
Nov	23	24
Déc	21	23

TABLE 2.6: Nombre des navires réels et prévus pour l'année 2020

2.5.7 Interprétation des résultats

D'après les résultats obtenus, nous constatons que le nombres de navires prévus et réels sont approximativement analogues durant tous les mois, alors nous déduisons que dans l'ensemble, si nous comparons le nombre de navires traités dans les deux années, nous déduisons que le trafic des conteneurs au niveau du port de Béjaia face a la crise sanitaire est stable par rapport à l'année précédente.

2.5.8 Conclusion

Dans cette partie, nous avons procédé dans un premier temps à l'application de la méthode Box et Jenkins, étant donné que la première série n'est pas stationnaire (elle comporte une ten-

dance et une saisonnalité), il a été question d'une différenciation d'ordre 1 et saisonnière d'ordre 12. Après la stationnarisation de la série, et à partir des graphiques des fonctions d'autocorrélation et d'autocorrélation partielle, nous avons sélectionné le modèle SARIM A(1 ; 1 ; 0)(0 ; 1 ; 1)₁₂. La validation du modèle s'est faite grâce au test de student sur les coefficients et le test de Ljung-Box sur les résidus. Ainsi, nous avons pu obtenir des prévisions pour l'année 2020. Puis nous avons élaboré un modèle de série chronologique avec la même méthode pour effectuer une prévision pour l'année 2021. Dans un deuxième temps, nous avons modélisé le nombre de navires traités en 2019 avec la méthode de la régression linéaire multiple pour effectuer une prévision pour 2020.

2.6 Conclusion générale

À la lumière de notre travail, nous avons présenté d'une part, une revue statistique d'une série chronologique qui représente, dans notre cas, les quantités de tonnage engendrées par le trafic des hydrocarbures par mois sur six ans, d'une autre part, un modèle de régression linéaire multiple. Comme indiqué dans l'introduction, le but de cette étude sur cette série chronologique, et de ce modèle de régression linéaire est la prévision, qui consiste à utiliser le meilleur des modèles estimés pour donner les valeurs estimées à des instants futurs. En premier lieu, nous avons présenté les méthodes statistiques de la prévision telle que la méthode de Box et Jenkins et de la régression linéaire multiple. En deuxième lieu, grâce au logiciel R, nous avons appliqué la méthode de Box et Jenkins pour prévoir les quantités de tonnage de l'année 2020. Dans le but d'effectuer une comparaison avec les valeurs réelles et voir l'impact de la pandémie Covid-19 sur le trafic hydrocarbure, nous avons élaboré un modèle de série chronologique avec la même méthode pour effectuer une prévision pour l'année 2021. Ensuite, nous avons adopté la méthode de la régression linéaire multiple pour modéliser le nombre de navires traités en 2019 avec certains facteurs (livraison, importation, exportation, temps), et réaliser une prévision pour l'année 2020. Les résultats obtenus montrent que pour le trafic hydrocarbure, la quantité annuelle enregistrée en 2020 semble être inférieure à la quantité prévue, cela éclaire l'influence de la pandémie Covid-19 sur les opérations import-export de ce terminal. La prévision pour l'année 2021 permet de nous enrichir sur la vision future et envisager une planification assez adéquate pour les années à venir. L'étude réalisée sur le trafic des conteneurs montre que la crise sanitaire n'impacte pas sur les mesures de performance, au vu de la comparaison du nombre réel des navires traités en 2020 avec le nombre de navires obtenu par les prévisions.

2.7 Bibliographie

- [1] S. Adjabi. "Cours des méthodes statistiques de prévisions". Département de Recherche Opérationnelle, Université de Bejaia, 2013-2014.
- [4] F. Avram. "Séries temporelles : Régression et modélisation ARIMA (p, d, q)". Université de Pau, 2012.

- [5] A. Ahmed, Contribution à l'économétrie dans des séries temporelles à valeurs entières, Université Charle de Gaulls Lille 3 (2016).
- [6] G. Arnauad (2012/2013). Régression linéaire. Université Rennes 2 Master de statistique.
- [7] A. Antoniadis, J. Berruyer ,R. Carmona (1992). Régression non linéaire et applications, Economica.
- [8] J. Bergonzini et C. Duby (1995). Analyse et planification des expériences. Les dispositifs en blocs, ed. Masson, Paris, Milan, Barcelone, p 353.
- [10] G. Box and G.M. Jenkins. "Time series analysis : forecasting and control". 2nd edition, Holden- Day, San Francisco, 1976.
- [11] M. David et J.S. Michaud. "La prévision approche empirique d'une méthode statistique". Université de Lyon, Messon, 1983.
- [12] J. Daudin, J. c. duby ; s. robin et p. trécourt. "Analyse des série chronologiques". Université de Bordeaux, 1996.
- [13] R. Giraud et N. Chaix (1989). Econométrie. Presses Universitaires de France (PUF).
- [14] D. C. Hoaglin, and R. E. Welsch, (1978) : The hat matrix in regression and anova. The American Statistician.
- [16] V. Fourcassié et C. Jost, (2012). Introduction aux modeles linéaires généraux (General linear model - GLM). Cours Modules Statistiques Master 2 NCC.
- [17] G. Melard. "Méthodes de prévision à cours terme." Université de Bruxelles, 1990.
- [19] Y. Tillé, (2011). Résumé du Cours de Modèles de Régression. Institut de statistique, Université de Neuchâtel.