

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

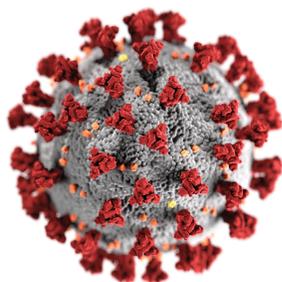


Université A. Mira de Béjaïa
Faculté des Sciences Exactes
Département Recherche Opérationnelle

Mémoire de Fin de Cycle
En vue de l'obtention du diplôme de Master
En Recherche Opérationnelle
Option : Modélisation Mathématique et Évaluation des Performances des Réseaux

Thème

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



Présenté par :
M^r. Ramdani Ali
M^r. Ghilas Lyamine

Soutenu le 17 octobre 2021 devant le jury composé de :

Président	D ^r S. KENDI	U.A/Mira Béjaïa
Encadreur	P ^r D. AISSANI	U.A/Mira Béjaïa
Co-Encadreur	D ^r N. SAADI	U.A/Mira Béjaïa
Examinatrice	D ^r L. DJERROUD	U.A/Mira Béjaïa
Examinatrice	M ^{me} D. MEDJANI	E.P.B

Béjaïa, Octobre 2021.

** Remerciements **

Nous remercions d'abord et avant tout le bon dieu tout puissant de nous avoir donné le courage et la force pour réaliser ce travail.

Nous désirons remercier notre encadreur Pr D.AISSANI et la co-encadreur Dr N.SAADI pour leurs disponibilité, leurs aide précieuse et leurs conseils qui nous ont accompagné tout au long de ce travail.

Nous remercions également l'ensemble des ingénieurs du département statistique pour leurs accueil et leurs aide affective ainsi tout le personnel des différents services au sein de l'entreprise.

Nous remercions ainsi la présidente Dr S. KENDI, l'ensemble des jurys Dr L. DJERROUD et M^{me} D. MEDJANI, qui ont accepté d'évaluer notre travail.

Nous remercions aussi nos parents et proches pour leurs aides et encouragements.

Sans oublier nos amis et camarades pour leurs soutiens qu'ils nous ont apportés.

✧ *Dédicaces* ✧

Nous dédions ce travail
A nos très chères parents
A nos frères et nos soeurs
A tous nos amis

Ali et Lyamine

Table des matières

Table des matières	i
Table des figures	v
Liste des tableaux	vi
Introduction générale	1
1 PRESENTATION DE L'ENTREPRISE	3
1.1 Généralités sur l'Entreprise Portuaire de Béjaia(EPB)	3
1.1.1 Présentation	3
1.1.2 Historique l'entreprise portuaire de Béjaia	4
1.1.3 Implantation géographique du port de Bejaia	4
1.1.4 Bassins du port	4
1.1.5 Surfaces d'entreposage	5
1.2 Missions et activités de l'entreprise	6
1.2.1 Missions de l'entreprise	6
1.2.2 Activités de l'entreprise	7
1.3 Organigramme de L'EPB	8
1.4 Organisation de l'entreprise	8
1.4.1 Direction Générale	9
1.4.2 Direction prospective et marketing	9
1.4.3 Direction Générale Adjointe Fonctionnelle	9
1.4.4 Direction des Ressources Humaines	9
1.4.5 Direction Digitalisation et Numérique	10
1.4.6 Direction des Finances et Comptabilité	10
1.4.7 Direction du Système de Contrôle Interne	10
1.4.8 Direction Achats	10
1.4.9 Direction Générale Adjointe Opérationnelle	10

1.4.10	Direction Domaine et Développement	11
1.4.11	Direction Maintenance	11
1.4.12	Direction Exploitation	11
1.5	Description du service Statistique	12
1.6	Processus de collection de données	13
1.6.1	Sources de la collecte de données	13
1.7	Type de données collectées	14
1.7.1	Type des navires	14
1.7.2	Type des marchandises	16
1.7.3	Les postes à quai	19
1.7.4	Ressources humaines et matériels	19
1.8	Organisation de données collectées	20
1.8.1	Données manipulées	21
1.8.2	Données manifestées	21
1.9	Position du problème	21
1.10	Conclusion	22
2	Généralités sur les séries chronologiques	23
2.1	Introduction	23
2.2	Les séries chronologiques	23
2.2.1	Les composantes d'une série chronologique	24
2.2.2	Schémas de décomposition d'une série chronologique	24
2.2.3	Processus stochastique	24
2.3	Série chronologique stationnaire	24
2.3.1	Processus bruit blanc	25
2.3.2	Fonction d'autocovariance	25
2.3.3	Fonction d'autocorrélation	25
2.3.4	Fonction d'autocorrélation partielle	26
2.4	Modélisation et prévision	26
2.4.1	Modèles de prévision	26
2.5	Méthode de Box et Jenkins	28
2.5.1	Les prévisions par la méthode de Box et Jenkins	28
2.6	Méthode du lissage exponentiel	31
2.6.1	Lissage exponentiel simple	32
2.6.2	Lissage exponentiel double	32
2.6.3	Lissage exponentiel de Holt	32

2.7	Conclusion	33
3	Régression linéaire et prévision	34
3.1	Introduction	34
3.2	Le modèle de la régression linéaire multiple	34
3.2.1	Hypothèses stochastiques	35
3.3	Forme matricielle du modèle	35
3.3.1	Hypothèses structurelle	36
3.4	Estimation des paramètres	36
3.4.1	Méthode des moindres carrés	37
3.4.2	Propriétés des estimateurs	37
3.4.3	La variance estimée des erreurs	38
3.5	Analyse de la variance et coefficient de détermination	38
3.5.1	Le coefficient de détermination R^2	39
3.5.2	Le coefficient de détermination ajusté R_{adj}^2	39
3.6	Les tests statistiques	39
3.6.1	Test de Fisher	40
3.6.2	Test de Student	40
3.7	Intervalles de confiances	41
3.8	Prévision	42
3.9	Méthode d'élimination pas à pas (Backward Elimination procedure)	42
3.10	Conclusion	43
4	Modélisation et résolution	44
4.1	Introduction	44
4.2	Présentation du logiciel R	44
4.3	Description des données	45
4.3.1	Données statistiques	45
4.3.2	Collecte de données	45
4.4	Prévision par de la méthode de Box-Jenkins(2020)	47
4.4.1	Présentation graphique des données	48
4.4.2	Analyse préliminaire	49
4.4.3	Stationnarité	49
4.4.4	Identification du modèle	50
4.4.5	Estimation des paramètres du modèle	51
4.4.6	Validation du modèle	52

4.4.7	Prévision	53
4.4.8	Interprétation des résultats	54
4.5	Prévision par la méthode de Box-Jenkins (2021)	55
4.5.1	Présentation graphique des données	55
4.5.2	Stationnarité	56
4.5.3	Identification du modèle	57
4.5.4	Prévision	58
4.5.5	Interprétation des résultats	59
4.6	Prévision par la méthode de régression linéaire	59
4.6.1	Modèle de régression	60
4.6.2	Les tests statistiques de Student et Fisher	62
4.6.3	Test global de Fisher	62
4.6.4	Test individuel de Student	62
4.6.5	Intervalle de confiance	64
4.6.6	Prévision	64
4.6.7	Interprétation des résultats	65
4.7	Conclusion	65
	Conclusion générale	66
	Bibliographie	68

Table des figures

1.1	Plan du Port de Béjaia	5
1.2	Les principaux terminaux du port de Béjaia	6
1.3	Organigramme de L'EPB	8
2.1	Organigramme de la démarche Box et Jenkins.	31
4.1	Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019	46
4.2	Tonnage engendré par le trafic des hydrocarbures (2015-2020)	47
4.3	Le graphe de la série originale	48
4.4	Grappe de la série et ses différentes composantes.	49
4.5	Corrélogramme et corrélogramme partiel de la série originale	49
4.6	Grappe de la série chronologique differencier	50
4.7	Corrélogramme et corrélogramme partiel de la série differencier	51
4.8	Corrélogramme et corrélogramme partiel des résidus.	53
4.9	Grappe de la série et la prévision	54
4.10	Grappe de la série originale	55
4.11	Corrélogramme et corrélogramme partiel de la série originale	56
4.12	Grappe de la série et la prévision	59

Liste des tableaux

1.1	Les informations collectées sur les navires	16
1.2	Les informations collectées de tonnage	17
1.3	Le nombre des Conteneurs au cours des dernières années	17
1.4	Tonnage (Import Export)	17
1.5	Tonnage des céréales	18
1.6	Tonnage de sucre (import export)	18
1.7	Bois (import)	18
1.8	Métaux ferreux (import)	18
1.9	Les postes à quai	19
1.10	Ressources matériels	20
2.1	Les critères de choix d'un modèle de prévision	28
4.1	Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019	45
4.2	Les quantités de tonnage engendrées par le trafic des hydrocarbures (2015-2020)	46
4.3	Les nombre de navires (porte-conteneurs) traités durant l'année 2019	47
4.4	Les résultats de la prévision de tonnage engendré par le trafic des hydrocarbures pour 2020	54
4.5	Les résultats de la prévision de tonnage des hydrocarbures pour l'année 2021	58
4.6	Nombre des navires réelles et prévus pour l'année 2020	65

Introduction générale

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 effecteur 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 que 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.

Afin d'atteindre nos objectifs, nous avons structuré notre travail en quatre chapitres :
Le premier chapitre est consacré à la présentation de l'entreprise d'accueil, où nous avons effectué notre stage, en retraçant son historique, sa localisation et ses principales caractéristiques ainsi que les différentes tâches des services de la direction.

Le deuxième chapitre portera sur le cadre général d'une série chronologique, ainsi que les éléments théoriques indispensables à l'élaboration des modèles de prévision, en exposant tous les outils statistiques qui vont nous permettre de réaliser notre étude empirique. Dans ce cadre, la méthode de Box et Jenkins sera présentée. Le troisième chapitre introduit la régression linéaire multiple.

Le quatrième chapitre concerne la partie pratique sur le traitement statistique des données du stage, en utilisant les méthodes développées dans les chapitres précédents.
Enfin, nous terminons ce mémoire par une conclusion générale.

PRESENTATION DE L'ENTREPRISE

1.1 Généralités sur l'Entreprise Portuaire de Béjaïa(EPB)

1.1.1 Présentation

De par sa position stratégique, les qualités nautiques remarquables et les infrastructures performantes dont il dispose, le port de Béjaïa reste un moteur de développement économique pour la région et le pays. Principale plaque tournante du commerce, il constitue l'accès privilégié aux différentes industries, parce qu'il offre à ses clients des terminaux propices et compétitifs ainsi que des équipements modernes et performants, tous dédiés pour l'accueil et le traitement de tous types de marchandises. L'entreprise portuaire de Béjaïa(EPB) est née le 14 Aout 1982 suite à la restriction de l'Office national des Ports(ONP), de la Société Nationale de Manutention (SONAMA) et de la Compagnie Nationale Algérienne de Navigation (CNAN). Le 15 Février 1989, elle devient une entreprise publique autonomie, son capital social fut fixé à dix millions (10.000.000) de dinars algériens par décision du conseil de la planification n° 191/SP/DP du 09 Novembre 1988. Actuellement, le capital social de l'entreprise a été ramené à 3.500.000.000 Da.

Le port de Béjaïa aujourd'hui est réputé mixte ; hydrocarbures et marchandises générales y sont traités. L'aménagement moderne des superstructures, le développement des infrastructures, l'utilisation de moyens de manutention et de techniques adaptés à l'évolution de la technologie des navires et enfin ses outils de gestion moderne, a fait évoluer le port de Béjaïa depuis le milieu des années 1990. Le port de Béjaïa est classé comme premier port en Algérie concernant le transit et la manutention des marchandises générales. Il est également le premier port du bassin sud méditerranéen certifié ISO 9001.2000 pour l'ensemble de ses prestations, et à avoir ainsi installé un système de management de la qualité. Cela constitue une étape dans le processus d'amélioration continue de ses prestations au grand bénéfice de ses clients. L'Entreprise Portuaire a connu d'autres succès depuis, elle est notamment certifiée à la Norme ISO 14001 :2004 et au référentiel OHSAS 18001 :2007, respectivement pour l'environnement, l'hygiène.

1.1.2 Historique l'entreprise portuaire de Béjaia

L'histoire du port de Béjaïa, c'est bien plus que l'accueil des navires en provenance des quatre coins du monde. Depuis son existence, qui remonte à l'époque Phénicienne où fût installé le comptoir punique, le port de Béjaïa a su s'adapter aux mutations internationales et au développement. Derrière tous les succès, on retrouve un nombre innombrable de personnes qui, grâce à leur compétence et leur dévouement, ont su hisser cette entreprise qui constitue un important moteur économique pour la région. C'est ainsi que s'est écrite son histoire et c'est de cette façon que s'exprimera son futur.

1.1.3 Implantation géographique du port de Bejaia

Le port est situé dans la baie de la ville de Béjaia, le domaine public artificiel maritime et portuaire est délimité :

- ◇ Au nord par la route nationale n°9.
- ◇ Au sud par les jetées de fermeture et du large sur une longueur de 2.750 m.
- ◇ A l'est par la jetée Est.
- ◇ A l'ouest par la zone industrielle de Béjaia.

Le Port de Béjaïa dessert un hinterland important et très vaste. La ville et le port de Béjaia disposent de ce fait de voies ferroviaires et d'un aéroport international.

D'une desserte routières reliant l'ensemble des villes du pays jusqu'au porte du Sahara.

1.1.4 Bassins du port

Le port est composé de trois bassins :

- Bassin de l'Avant-Port : sa superficie est de 75 hectares et ses profondeurs varient entre 10,5 m et 13,5 m. Disposant d'installations spécialisées, l'avant-port est destiné à traiter les navires pétroliers.
- Bassin du Vieux Port : sa superficie est de 25 hectares et ses profondeurs de quai varient entre 8 et 9 m.
- Bassin de l'Arrière Port : Sa superficie est de 55 hectares et ses profondeurs varient entre 10,5 m et 12 m.



FIGURE 1.1 – Plan du Port de Béjaïa

1.1.5 Surfaces d’entreposage

Le port de Béjaïa dispose de 400.000 m² de surface non couverte et de 17500m² de surface couverte, ces surfaces sont réparties comme suit :

1. Terminal à conteneurs :le terminal à conteneurs est exploité depuis juillet 2005 par une nouvelle entité BMT (Béjaïa méditerranéen Terminal). C’est le seul terminal algérien à disposer de portiques de quai pouvant traiter les navires postpanamax. Le terminal dispose également de sa propre structure douanière (guichet unique).
2. Centre de transit des marchandises dangereuses ”CTMD” :unique en Algérie, le CTMD est dédié au stockage, dans les meilleures conditions de sécurité, de toutes les marchandises classées dangereuses sur une surface totale de 7000m².
3. Terminal céréalier : l’EPB, occupant la première position pour ce qui est du trafic céréalier, l’activité est dotée d’équipements les plus performants sur le marché ; qui ont fait que les rendements de la manipulation des céréales sont passés de 50T/heure à 750T/heure. L’EPB peut ainsi traiter de gros céréaliers de 60.000 tonnes.
4. Terminal roulier :il permet la réception des matériels roulants (Véhicules touristiques, utilitaires et de transport public, engins de Travaux Publics, engins d’exploitation ferroviaire).
5. Abri papier :il est d’une superficie de 1.200 m², spécialement aménagé pour accueillir des bobines de papier de grand et petit formats,
6. Terminal à bois :terminal dispose de sa propre surface d’entreposage ainsi que d’un guichet unique. Il est doté d’équipements performants et parfaitement adaptés à ce type de marchan-

dise. Les nouvelles grues à palonniers, permettent de traiter dix fardeaux en même temps et de réduire la durée de séjour à quai des navires, tout en réduisant le risque d'endommagement de la marchandise. L'EPB de cette manière, a consolidé sa part de marché et sa position de leader sur le produit Bois et offre un service de meilleure qualité à sa clientèle, qui compte d'importantes entreprises du Bâtiment et Travaux Publics, très exigeantes et pointilleuses en matière de qualité de service.

Caractéristiques :

- 300 ml de quai avec 02 postes à quai favorisés
- Emplacement : Nouveau Quai et quai de la gare
- Tirant d'eau maximum autorisé : 8 m à 11,50 m
- Superficie du terminal : 7 hectares
- Capacité totale : 19.300 fardeaux gerbés sur 04 niveaux



FIGURE 1.2 – Les principaux terminaux du port de Béjaïa

1.2 Missions et activités de l'entreprise

1.2.1 Missions de l'entreprise

Elle a pour missions :

- Organisation de l'accueil des navires.
- Aide à la navigation (Pilotage des navires).
- Activité d'acconage (entreposage et livraison des marchandises à l'import et l'export)..
- Transit des passagers et de leurs véhicules.
- Gestion et développement du domaine portuaire.
- Prise en charge des cargaisons à l'embarquement / débarquement et pré- évacuation.
- Assurer une disponibilité permanente des moyens humains et matériels.
- Améliorer en continu les performances (humaines, matérielles et budgétaires).
- Rentabiliser au maximum les infrastructures et superstructures portuaires.
- Gérer les systèmes de management de la qualité, de l'environnement, de la santé et sécurité.

1.2.2 Activités de l'entreprise

Les principales activités de l'entreprise sont :

- **L'acheminement des navires de la rade vers le quai** : Dans certains cas exceptionnels, d'arrivée massive en rade, les navires restent en attente dans la zone de mouillage (rade) jusqu'à obtention de l'autorisation de rejoindre un poste à quai. Cette dernière est délivrée après une conférence de placement qui se tient quotidiennement au niveau de la Direction Capitainerie.
- **Le remorquage** : il consiste à tirer ou à pousser le navire, pour effectuer les manoeuvres d'accostage, d'appareillage du navire. Il consiste également à effectuer les opérations de convoyage et d'aide dans l'exécution d'autres manoeuvres.
- **Le pilotage** : il est assuré de jour comme de nuit par la Direction Capitainerie et est obligatoire à l'entrée et à la sortie du navire. Il consiste à assister le commandant dans la conduite de son navire à l'intérieur du port.
- **Le lamanage** : il consiste à amarrer ou désamarrer le navire de son poste d'accostage.
- Les opérations de manutention et d'acconage pour les marchandises consistent en :
 - ◇ Les opérations d'embarquement et de débarquement des marchandises.
 - ◇ La réception des marchandises.
 - ◇ Le transfert vers les aires d'entreposage, hangars et terre-pleins, ports secs.
 - ◇ La préservation ou la garde des marchandises sur terre-pleins ou hangar et hors port.
 - ◇ Pointage des marchandises.
- D'autres prestations sont également fournies aux navires et aux clients tels enlèvement des déchets des navires et assainissement des postes à quai pesage des marchandises (ponts

bascules). Location de remorqueurs ou vedettes (pour avitaillement des navires, transport de l'assistance médicale, assistance et sauvetage en haute mer).

1.3 Organigramme de L'EPB

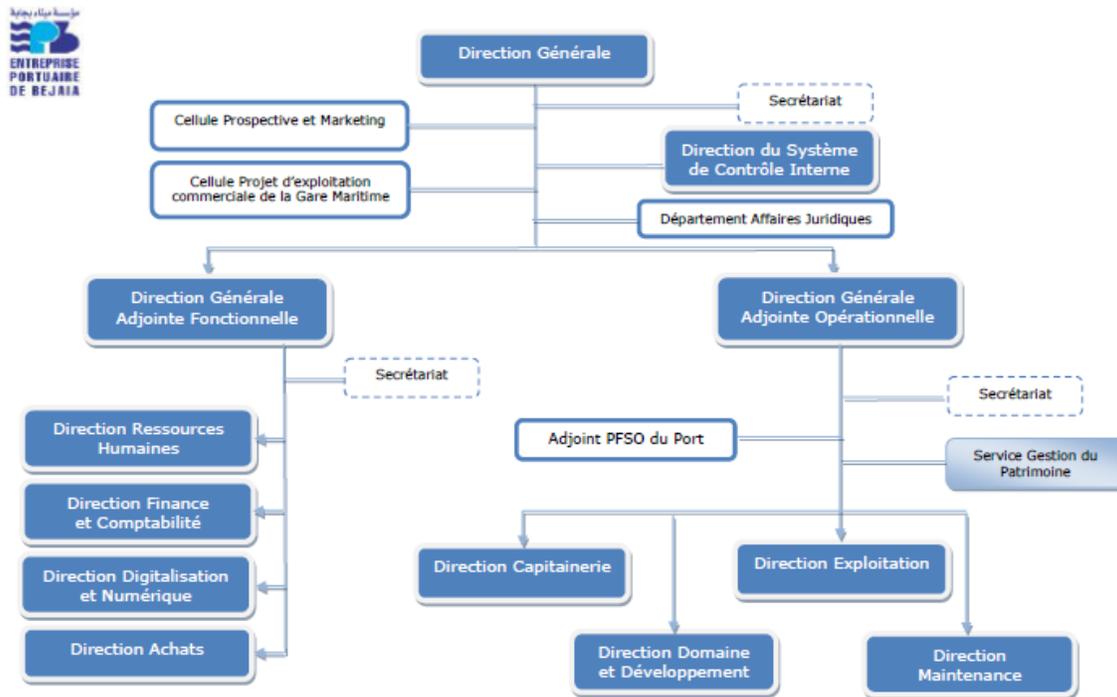


FIGURE 1.3 – Organigramme de L'EPB

1.4 Organisation de l'entreprise

L'entreprise portuaire est structurée en directions au nombre de neuf, elles-mêmes subdivisées en département et services. Les activités de l'entreprise sont organisées suivant un schéma d'excéntration des pouvoirs de décision au profit des directeurs centraux et de leur chef de départements. Un projet de réorganisation de l'entreprise a été présenté au Conseil d'administration en date du 17 février 2020, qui s'est soldé par une validation (résolution n°04).

Des ajustements ont été apportés à l'organisation actuelle par la fusion, la suppression ainsi que la création de certaines directions. Vu la mise en place de cette nouvelle organisation, des décisions de mobilité ont été établies au profit du personnel concerné ; d'où la nécessité d'actualiser les fiches emplois avant de procéder à leur diffusion.

1.4.1 Direction Générale

Elle est chargée de concevoir, coordonner et contrôler les actions liées à la gestion et au développement de l'Entreprise.

- ◇ Le transfert du Département Affaires Juridiques à la Direction Générale ;
- ◇ Le rehaussement du Département Audit interne et Contrôle de Gestion en " Direction du Système de Contrôle Interne " ;

1.4.2 Direction prospective et marketing

Elle est chargée de tous les travaux d'étude et de planification concernant l'entreprise. Cette structure élabore les opérations de marketing visant à augmenter les parts de marché de l'EPB.

1.4.3 Direction Générale Adjointe Fonctionnelle

- ◇ La dissolution de la Direction du Management Intégré et le transfert du Département Mangement Intégré à la Direction des Ressources Humaines ;
- ◇ Le changement de dénomination de la Direction des Systèmes d'information en " Direction Digitalisation et Numérique " .

1.4.4 Direction des Ressources Humaines

La Direction des Ressources Humaines est rattachée à la Direction Générale Adjointe Fonctionnelle. Cette direction est chargée d'exécuter toutes les taches liées à la gestion et au développement des structures, et d'appliquer rigoureusement les lois de gestions et les règlements intérieurs de l'entreprise. Elle a comme principal département celui du personnel et de la formation qui s'occupe de la gestion administrative du personnel, ainsi que l'établissement de la paie, le suivi des affaires sociales, le suivi de la situation de la carrière de l'effectif et l'élaboration des actions de formations afin d'atteindre une grande maîtrise des métiers pour adapter le personnel à l'exigence de la modernisation elle comporte trois trois départements, à savoir :

1. Département Ressources Humaines et Développement,
2. Département Administration du Personnel et Moyens,
3. Département Qualité de Vie au Travail.

1.4.5 Direction Digitalisation et Numérique

La Direction des Systèmes d'information est, désormais, dénommée Direction Digitalisation et Numérique, rattachée à la Direction Générale Adjointe Fonctionnelle. La Direction comportera trois départements :

1. Département Numérique,
2. Département Génie Logiciel,
3. Département Infrastructures et Systèmes.

1.4.6 Direction des Finances et Comptabilité

Cette direction gère les finances et les dépenses de l'entreprise.

La Direction des Finances et comptabilité comprend deux départements :

1. Département Finances,
2. Département Comptabilité générale.

1.4.7 Direction du Système de Contrôle Interne

La Direction du Système de Contrôle Interne comprend trois (03) services :

1. Service Audit Interne,
2. Service Management des Risques,
3. Service Contrôle de Gestion.

1.4.8 Direction Achats

La Direction Achats comprend trois départements :

1. Département Passations de Commandes,
2. Département Passation de Marchés,
3. Département Approvisionnements,

1.4.9 Direction Générale Adjointe Opérationnelle

- ◇ La création du " Service Gestion du Patrimoine ", rattaché à la Direction Générale Adjointe Opérationnelle,
- ◇ Le rattachement du Bureau de Sûreté Portuaire B.S.P au DGAO,

- ◇ La création de la " Direction Exploitation " par la fusion des directions de la Manutention et Acconage, les Zones Logistiques Extra portuaires, le Département Exploitation des Engins de la Direction Logistique,
- ◇ La fusion des Directions : Remorquage et Capitainerie en "Direction Capitainerie",
- ◇ La fusion de la Direction du Domaine et Développement et la Direction de la Gare Maritime en " Direction Domaine et Développement ",
- ◇ La centralisation des achats et l'approvisionnement et maintenances des différentes structures en une seule "Direction Achats et Maintenances" : un groupe de travail a été mis en place ayant pour missions la proposition d'une organisation opérationnelle.

1.4.10 Direction Domaine et Développement

Elle a pour mission essentielle la mise en oeuvre des programmes d'entretien et d'investissement de l'entreprise (travaux et acquisition du matériel) . Elle est subdivisée en trois départements cette Direction comporte quatre départements et un service ; à savoir :

1. Département Domaine,
2. Département Entretien et Assainissement,
3. Département Projets,
4. Département Gestion du Transit des Passagers,
5. Service Commercial.

1.4.11 Direction Maintenance

Il s'occupe de la maintenance curative et préventive des engins, la sécurité, la fiabilité pendant leur utilisation, il est chargé aussi du développement et l'extension des moyens et équipements.

Cette Direction comporte deux départements à savoir :

1. Département Maintenance Engins.
2. Département Maintenance Navale.

1.4.12 Direction Exploitation

La direction est composée de cinq départements :

1.4.12.1 Organisation de la direction exploitation

1. **Département Manutention et Acconage** : L'acconage est une opération qui consiste à assurer la réception, la reconnaissance à terre des marchandises ainsi que de leurs gardes jusqu'à leur embarquement ou livraison au destinataire.
2. **Département Affectation de Moyens** : l'affectation des moyens est une opération qui consiste à assurer des affectations engins et grues optimale.
3. **Département Logistique d'IGHIL OUBEROUAK et TIXTER** : elle est érigée en deux départements : Département de la zone logistique de TIXTER et de la zone logistique d'IGHIL OUBEROUAK. Ses principales missions sont :
 - Rapprocher la marchandise du client final.
 - Décongestionner les surfaces dans l'enceinte portuaire.
 - Réduire les temps d'attente en rade des navires dus au manque d'espace d'entreposage.
 - Développer le transfert de masse des marchandises par voie ferroviaire.
 - Réduire les congestions sur les routes et réduire l'émission de gaz polluants.
4. **Département Statistiques et Commercial** :

Il se compose de trois services :

 - (a) **Service commercial** :

Il est chargé de l'établissement des factures, des prestations fournies aux clients, ainsi que la détermination du chiffre d'affaires.
 - (b) **Service intermodal.**
 - (c) **Service statistique.**

1.5 Description du service Statistique

Ce service est parmi les principaux services de l'EPB, vu son rôle et les besoins qui lui sont affectés. Certaines tâches de ce service complètent les tâches des autres départements de l'entreprise, et certaines autres sont confiées seulement à ce dernier.

La tâche initiale de ce service est la récolte des données. Elle se fait en communication avec les autres secteurs du port. Une mise à jour quotidienne de chaque secteur sera envoyée à ce service. Ce dernier est chargé de saisir, classifier et vérifier la disponibilité et la sûreté des données envoyées.

La récolte des données est la tâche confiée seulement à ce département : mettre à jour les données des différents secteurs, élaborer les bilans statistiques et les rapports quotidiens, mensuels et encore annuels. Par la suite ces données seront envoyées au service marketing de l'entreprise.

1.6 Processus de collection de données

Pour collecter les données, il est donc nécessaire de noter l'ensemble des éléments d'information suivant les quatre dimensions et de spécifier les conditions de la collecte, lorsqu'une dimension doit être négligée. Même dans ce cas, il est indispensable de noter, ne serait-ce qu'à titre de vérification des conditions, les données correspondantes, au moins à chaque fois que ces conditions évoluent. La collecte de données consiste donc à recueillir, si possible, tous les éléments quantitatifs, qualitatifs, environnementaux et temporels nécessaires permettant d'évaluer l'intensité d'un phénomène, d'apprécier la qualité de cette valeur, de localiser le phénomène dans l'environnement et dans le temps.

L'Entreprise Portuaire de Béjaia est une prestataire de service, elle s'appuie sur de petites données pour déterminer son revenu. Il y a donc plusieurs types de données qui peuvent être collectées à partir de diverses sources.

1.6.1 Sources de la collecte de données

Comme L'Entreprise Portuaire de Béjaia possède plusieurs types de données, elle se sont recueillies en basant sur trois sources de collecte de données.

1.6.1.1 Système d'information portuaire(SIP)

Le Système d'Information Portuaire est un outil de gestion permettant d'évaluer rapidement la productivité du port. Il met à la disposition des gestionnaires, des renseignements pertinents qui favorisent la maîtrise de l'activité portuaire. Les impératifs de sécurité du trafic maritime, et la nécessité de faciliter autant que possible le flux des marchandises au travers de leurs installations, ont conduit les ports maritimes vers des systèmes d'informations utilisant de plus en plus l'informatique et les technologies modernes de télécommunications. Pour la plupart des ports maritimes, les systèmes de gestion du trafic des navires et de leurs escales.

1.6.1.2 Déclarations des consignataires

L'Entreprise Portuaire de Béjaia assure des services pour les clients. Pour que, les marchandises arrivent en bon état, il faudrait que toute opération commerciale, que se soit à l'importation ou à l'exportation, fait appel à divers intervenants dans le but de faciliter le trafic maritime et d'éviter les pertes de temps. Et pour bien expliquer cette opération commerciale, nous allons évoquer les principaux intervenants dans le port. Afin de réaliser tout cela, le port est un espace voué au transit de marchandises et de passagers, qui nécessite une multiplicité d'acteurs. La communauté

portuaire est l'ensemble des acteurs qui donnent vie à l'espace portuaire. Les fonctions décrites succinctement ci-dessous existent dans tous les ports, mais ne sont pas toujours remplies par les mêmes acteurs.

1.6.1.3 Clients du port

◇ Les armateurs

L'armateur exploite des navires pour la navigation commerciale. Il peut être le propriétaire ou le locataire.

Dans les ports où il n'est pas installé directement, l'armateur est en liaison directe avec un agent maritime consignataire, qui défend ses intérêts et agit en son nom, auprès de toutes les professions.

◇ Les chargeurs

Ce sont les propriétaires et les expéditeurs de la marchandise dans le cadre d'un transport international. Le chargeur est le propriétaire de la cargaison d'un navire, ou d'une partie de cette cargaison. Le chargeur peut être l'importateur ou l'exportateur. Selon, la nature du contrat commerciale, il s'agit de propriétaire et/ou d'expéditeur de marchandises.

En général, le chargeur n'est pas en relation directe avec l'autorité portuaire, c'est une relation commerciale avec les différents opérateurs à savoir les agents maritimes, les consignataires, les transitaires ou éventuellement manutentionnaire, à qui ils confient leurs marchandises.

1.7 Type de données collectées

Les différentes opérations qui se déroulent au niveau du port de Béjaia permettent de fournir aux différents secteurs du port des données qui seront exploitées ultérieurement dans les services concernés, ce qui résulte une diversité dans les types de données recueillies dans le port. A fin d'avoir une gestion meilleure, les données sont organisées selon :

- données des navires ;
- données des marchandises ;
- données des postes à quai ;
- données ressources humaines et matériels.

1.7.1 Type des navires

On distingue différents types de navires de charge selon le type de marchandises qu'ils transportent, et leur façon de les transporter.

1.7.1.1 Cargaison sèche

◇ Cargos polyvalents

Aussi appelés (cargos de divers)ou (cargos mixtes) lorsqu'ils transportent des passagers, se sont des navires qui embarquent à la fois des marchandises et des passagers, en général une douzaine , mais les (cargos de divers) n'embarquent pas de passagers.

◇ Porte-conteneurs

Ces navires sont spécialisés dans le transporteur de conteneurs, boîtes au format prédéfini. Leur taille varie du caboteur transportant une centaine de conteneurs aux géants pouvant en transporter une dizaine de milliers.

◇ Vraquiers

ils transportent de marchandises solides en vrac comme des granulats, des céréales, du charbon, etc

◇ Navires frigorifiques (reefers)

Petits, esthétiques et rapides, souvent peints en blanc, ils transportent des denrées périssables : bananes, jus de fruits, viande,poisson... à basse température, grâce à des installations spécialisées.

◇ Les rouliers

Ces navires transportent des véhicules sur plusieurs ponts, chargés grâce à une ou plusieurs rampes d'accès. Ils transportent souvent aussi d'autres marchandises sur leur pont, ou des passagers sur certaines routes très fréquentées

1.7.1.2 Cargaison liquide

Les navires-citernes sont destinés au transport des marchandises liquides. On y trouve les types suivants :

◇ Les pétroliers

Navires citernes transportant du pétrole. les transporteurs de pétrole brut sont les plus grands, dépassant 100000 tonnes.

◇ Les Chimiquiers

Pouvant transporter une grande variété de produits, ils disposent de nombreuses citernes et de tuyautages séparés, ainsi que de systèmes de chauffage. Ils sont soumis à des normes de sécurité drastiques.

◇ Les Gaziers

Ces navires transportent soit du gaz naturel (les méthaniers), soit du gaz de pétrole liquéfié (GPL), à basse température ($-160C$) ou à haute pression, dans des citernes spécialement conçues. On trouve encore d'autres navires-citernes spécialisés dans d'autres produits comme les huiliers, les pinardiens pour le vin, les bitumiers, etc. . .

Le tableau suivant montre un exemple sur les informations collectées sur les navires :

Imo	9486180
Navire	CHEM HYDRA
Type	NAVIRE CARBURANT
Pavillon	LIBERIA
Armateur	CHEM HYDRA SA TRUST COMP.MAJURO
Longueur	144.8
Largeur	23
T.J.B	11948
T.J.N	5180
Port en lourd	17055
Tirant d'eau été	8.9
Tirant eau max	8.9
Call sign	V7RU2
MMSI	636017991

Tableau 1.1 – Les informations collectées sur les navires

1.7.2 Type des marchandises

L'activité de l'import (export) des marchandises dans le port de Béjaia est connue par sa vitalité, ce dernier enregistre quotidiennement une diversité de type des marchandises embarquées et encore débarquées, ainsi, les marchandises sont classifiées comme suit :

1.7.2.1 Hydrocarbures

Le port de Béjaia est parmi les points d'entrées maritimes les plus importants de l'Algérie pour l'exportation et l'importation des produits énergétiques (hydrocarbures liquides et gaz liquéfiés). Les exportations d'hydrocarbures avoisinent le million de tonnes grâce à :

- ◇ Un appontement pétrolier ;
- ◇ Un appontement gazier.

Le tableau suivant montre un exemple sur les informations de tonnage collecté :

Les années	2015	2016	2017	2018	2019	2020
Pétrole brut	7177	7081	8311	8012	7804	4452
Hydrocarbures raffinés	1088	1030	862	873	976	758

Tableau 1.2 – Les informations collectées de tonnage

1.7.2.2 Conteneurs

Les conteneurs sont des grandes boîtes métalliques standardisées, permettant le transfert rapide de marchandises d'un navire à un autre mode de transport. Les terminaux portuaires qui les gèrent sont appelés terminaux à conteneurs et ont des activités plus complexes que les activités des ports de passagers ou de vrac sec ou liquides.[3] Avec son terminal à conteneurs le port de Béjaia engendre une vitalité dans ce secteur, ce qui justifie la création de l'entreprise BMT et les deux zones d'entreposage Ighil Oubarouaq à Béjaia et Tixter à Borj Bouririj. Le tableau suivant montre un exemple sur le nombre des Conteneurs au cours des dernières années

Les années	2015	2016	2017	2018	2019	2020
EVP	248766	267338	236193	252530	243506	246811

Tableau 1.3 – Le nombre des Conteneurs au cours des dernières années

1.7.2.3 Oléagineux

Le tableau suivant montre un exemple sur les informations collectées sur le tonnage

Les années	2015	2016	2017	2018	2019	2020
Import	502	542	607	633	688	636
Export	2	19	23	10	13	9

Tableau 1.4 – Tonnage (Import Export)

1.7.2.4 Céréales

Le tableau suivant montre un exemple sur les informations collectées sur le tonnage

Les années	2015	2016	2017	2018	2019	2020
Blé	1144	1020	750	853	639	647
Autres céréales	2577	2459	2308	2420	2490	2933
Soja	985	909	731	892	968	740

Tableau 1.5 – Tonnage des céréales

1.7.2.5 Sucre

L'EPB possède également des silos portuaires ainsi qu'un terminal de déchargement portuaire d'une capacité de 2000 tonnes/heure ce qui en fait le premier terminal de déchargement portuaire en Méditerranée. Le terminal de sucre qui appartient à l'entreprise CIVITAL.

Le tableau suivant montre un exemple sur les informations collectées de tonnage

Les années	2015	2016	2017	2018	2019	2020
Import	1552	1546	1673	1628	1555	1550
Export	397	463	514	551	602	762

Tableau 1.6 – Tonnage de sucre (import export)

1.7.2.6 Bois

Le trafic de bois, qui représente l'un des trafics stratégiques du port de Bejaia, un terminal dispose de sa propre surface d'entreposage ainsi que d'un guichet unique. Il est doté d'équipements performants et parfaitement adaptés à ce type de marchandise.

le tableau suivant montre un exemple sur les informations de tonnage collecté

Les années	2015	2016	2017	2018	2019	2020
Bois	659	613	393	765	554	487

Tableau 1.7 – Bois (import)

1.7.2.7 Métaux ferreux

Le tableau suivant montre un exemple sur les informations collectées sur le tonnage

Les années	2015	2016	2017	2018	2019	2020
Métaux ferreux	917	820	458	296	559	301

Tableau 1.8 – Métaux ferreux (import)

1.7.3 Les postes à quai

Le port s'étale sur une superficie totale de 79 hectares. Sa surface d'entreposage s'étale sur $410.000m^2$ dont $17.500m^2$ couverts. Il dispose de $3575ml$ de quai, répartis sur 16 postes à quai pour les navires de marchandises générales, 03 postes à quai pour les navires pétroliers et 01 poste gazier.

Le tableau suivant montre les caractéristiques de chaque quai.

Nom du Quai	N° des postes à quais	Bassins	Profondeurs	Longueurs
Port pétrolier	01 à 03	Avant port	11.5 à 13.5	250,260 et 260 ml
Quai Nord	06 à 08	Vieux port	08	290 ml
Quai Nord Ouest	09 à 11	Vieux port	08	273 ml
Quai de la Casbah	12 et 13	Vieux port	08	257 ml
Quai de la passe	14	Passe Casbah	08.5	146 ml
Quai Sud Ouest	15 et 16	Arrière port	10	230 ml
Quai de la gare	17 à 19	Arrière port	10	530 ml
Nouveau Quai	21 à 24	Arrière port	12	750 ml
Grande jetée du large	26	Arrière port	12	78 ml

Tableau 1.9 – Les postes à quai

1.7.4 Ressources humaines et matériels

L'entreprise portuaire de Bejaia a un énorme potentiel. Pour réaliser l'ensemble des opérations, le chargement et le déchargement des marchandises à la bonne destination. L'opération effectuée par l'entreprises pour assurer l'interface entre la mer et la terre. Elle réalise d'autres opérations de réception, reconnaissance et gardiennage des marchandises dans les ports.

On peut citer dans le tableau suivant les ressources humaines et matériels de l'entreprise.

Désignation	Capacités	Nombre d'engins
Chariots élévateurs à pinces	0.95 a 1.5 T	06
	4T	03
Chariots élévateurs à fourches	3T	12
	5T	09
	6T	05
	7T	11
	8T	03
	10T	13
	12T	03
	18T	05
	28T	03
	32T	03
	42T	01
	52T	01
Grue portuaire GOTTWALD	63T	01
	80T	01
Grue portuaire LIEBHERR	64T	06
	124T	01
Grue télescopique GROVE	90T	01
	60T	01
Grue camion	50T	02
	59T	01
	83T	01
Portique à grains	550T/H	01
Chargeurs sur pneus	2m3	02
	4m3	09
Pelles Chargeuses excavatrices	De 01-1.1T	03
Minis pelles	De 0.6 -1T	09
Stackers	42T	01
Tracteurs semi-remorques	20.5T	18
	38T	06
	50T	04
Tracteurs Ro-Ro	36T	01
	46T	01
Bennes	sup 14T	11
Spreader automatique	41T	01
Grappin motorisé	9.4T	02

Tableau 1.10 – Ressources matériels

1.8 Organisation de données collectées

Les données sont organisées en trois types

- Les données manipulées :
Les données saisies directement dans les applications.
- Les données manifestées :
Les données reçues par les e-mails et les déclarations des clients.

1.8.1 Données manipulées

Les données manipulées sont les données saisies dans les différents secteurs du port, la saisie se fait avec des applications, cependant chaque service dispose d'une ou plusieurs applications destinées à saisir et classifier et encore manipuler les données qui leur sont envoyées directement du terrain. la première donnée provient directement du navire avant qu'il soit à quai, une déclaration du commandant du navire sera envoyé à la tour marine, cette déclaration contient des informations sur le navire, type de marchandises, le tonnage . . . , une fois le navire est à quai, les agents du terrain s'occupent de remplir leurs documents à la main, classifier toutes les données sur le navire(nom du navire,le types de marchandises,le poste à quai, le tonnage débarqué (embarqué), les shifts...), ces données seront envoyées aux services concernés de l'entreprise puis seront saisies dans les applications, ensuite elles s'organisent dans des tableaux (SIP).

1.8.2 Données manifestées

Les données manifestées sont les données recueillies dans les formulaires et les déclarations des clients au niveau de la douane.

L'agent doit déclarer tous les informations concernant le navire, cela par le remplissage d'un formulaire au niveau de la douane, ce formulaire sera envoyé au service statistique de l'entreprise, puis le contenu de ce dernier sera enregistré manuellement dans un registre spécial.

1.9 Position du problème

Le nombre des navires qui ont accosté au port de Béjaia 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 Béjaia, les quantités de tonnage qui s'engendre par

le trafic des hydrocarbures connaissent une variation très élevées avant la crise sanitaire, bien précis, avant mars 2020, mais depuis, ces quantités ont connues un contraste remarquable, ce qui a causé des perturbation 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 hydrocarbure 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 ?

1.10 Conclusion

Dans ce chapitre, nous avons présenté les différents moyens, structure de l'Entreprise Portuaire de Béjaia ainsi que la position de problème,le chapitre suivant nous allons présenter les notions de base sur les séries chronologiques.

Généralités sur les séries chronologiques

2.1 Introduction

La théorie des séries chronologiques (ou temporelles) abordée est appliquée de nos jours dans des domaines aussi variés que l'économétrie, la médecine ou la démographie, pour n'en citer qu'une petite partie. On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prévoir ce phénomène dans le futur. On dispose ainsi d'observations à des dates différentes, c'est-à-dire d'une suite de valeurs numériques indicées par le temps. Cette suite d'observations d'une famille de variables aléatoires réelles notées $(Y_t)_{t \in T}$ est appelée Série chronologique où T espace des temps qui peut être discret ou continu [1].

Dans ce chapitre, nous donnons quelques notions sur les séries chronologiques, afin d'avoir une idée sur le choix des modèles appropriés pour nos séries chronologiques et calculer ainsi les prévisions.

2.2 Les séries chronologiques

Définition 2.2.1. On appelle une série chronologique (série temporelle), une série statistique à deux variables (t, Y_t) , avec $t \in T$, $T = (t_1, t_2, \dots, t_n)$ où la première composante de l'ensemble T est le premier instant et la deuxième composante est une variable numérique prenant ses valeurs aux instants t , suivant la nature du problème étudié la chronique peut être journalière, mensuelle, annuelle [2].

Il y a deux types de séries chronologiques :

1. **Série continue** : C'est une série où l'observation se fait d'une manière continue.
2. **Série discrète** : C'est une série où l'observation se fait sur des intervalles du temps fixés à priori (heure, jour, mois, années), la série discrète est notée $(Y_t, t = 1 \dots T)$, où T désigne le nombre d'observations.

2.2.1 Les composantes d'une série chronologique

1. **La tendance**(C_t) : cette composante représente l'évolution de la série à long terme, elle se présente comme une fonction du temps (linéaire ou non linéaire) et décrit une structure d'évolution déterministe.
2. **Les variations saisonnières** (C_s) : la composante saisonnière où le mouvement saisonnier, représente des effets périodiques, de période p , qui se reproduisent d'une façon plus ou moins identique, d'une période sur l'autre, la chronique correspondante est notée $S_t, t = 1, \dots, T$.
3. **la composante résiduelle** : Elle regroupe tout ce qui n'est pas pris en charge par les autres composantes saisonnières et tendancielle, elle contient donc de nombreuses fluctuations, en particulier accidentelles, dont le caractère est exceptionnel et imprévisible [3].

2.2.2 Schémas de décomposition d'une série chronologique

La technique de décomposition d'une série chronologique repose sur un modèle qui l'autorise, ce modèle porte le nom de schéma de décomposition, il en existe essentiellement trois types [4] :

1. **Schéma additif** : $Y_t = C_t + S_t + \epsilon_t$.
2. **Schéma multiplicatif** : $Y_t = C_t \times S_t \times \epsilon_t$.
3. **Schéma mixte** : $C_t \times S_t + \epsilon_t$.

2.2.3 Processus stochastique

Un processus est un ensemble de phénomènes, conçus comme une chaîne causale progressive à travers le temps, de point de vue mathématique, c'est la solution d'une équation composée de variables indicées par le temps : si l'évolution temporelle d'une ou plusieurs variables n'est pas totalement déterministe, nous parlons alors d'un processus stochastique[4]. Un processus stochastique est dit stationnaire si et seulement si :

1. $\forall t \in T E[Y_t]$ est finie et indépendante du temps.
2. $\forall t \in T E[Y_t^2]$ est finie et indépendante du temps.
3. $\forall t \in T Cov[Y_t, Y_{t+k}]$ dépend uniquement de k (indépendante de t).

2.3 Série chronologique stationnaire

Définition 2.3.1. Une série chronologique est stationnaire si elle est la réalisation d'un processus stationnaire, ceci implique que la série ne possède ni tendance ni saisonnalité et plus généralement

aucun facteur n'évolue avec le temps [10].

2.3.1 Processus bruit blanc

Un processus bruit blanc est une suite de variables aléatoires (ϵ_t) de même distribution mutuellement indépendantes.

Dans le cas de l'analyse des séries chronologiques, un bruit blanc $\{\epsilon_t, t \in T\}$ est tel que [11] :

$$E[\epsilon_t] = 0, \forall t \in T; \text{Var}(\epsilon_t) = E[\epsilon_t^2] = \sigma^2, \forall t \in T; E[\epsilon_t, \epsilon_k] = 0, \forall t \in T, t \neq k. \quad (2.1)$$

2.3.2 Fonction d'autocovariance

Soit $(Y_t, \forall t \in T)$ un processus stochastique, on appelle fonction d'autocovariance de Y_t la fonction γ [10] :

$$\gamma(k) = \text{Cov}(Y_t, Y_{t-k}), \forall k \in T. \quad (2.2)$$

Propriété 2.3.1. *La fonction d'autocovariance vérifie les propriétés suivantes :*

1. $\text{Cov}(Y_t, Y_{t-k}) = \gamma(|t - s|)$.
2. $\gamma(0) = \text{Var}(Y_t) = \sigma^2 < \infty$.
3. $|\gamma(k)| \leq \gamma(0)$, pour tout $k \in T$, la fonction d'autocovariance est donc bornée.
4. $\gamma(k) = \gamma(-k)$, pour tout $k \in T$, autrement dit la fonction d'autocovariance est paire.

2.3.3 Fonction d'autocorrélation

La fonction d'autocorrélation d'une série chronologique est définie par [10] :

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)} = \frac{\gamma(k)}{\gamma(0)}, k \in T. \quad (2.3)$$

Propriété 2.3.2. *La fonction d'autocorrélation vérifie les propriétés suivantes :*

1. $\rho(0) = 1$.
2. $|\rho(k)| < 1$, pour tout $k \in T$.
3. $\rho(k) = \rho(-k)$, pour tout $k \in T$.

2.3.4 Fonction d'autocorrélation partielle

L'autocorrélation partielle d'ordre k est un coefficient de corrélation partielle entre les deux processus $Y_t, t \in T$ et $Y_{t-K}, K \in T$ donnée par l'expression suivante [10] :

$$\pi_K = \frac{Cov(Y_t - Y_t^*, Y_{t-K} - Y_{t-K}^*)}{Var(Y_t, Y_t^*)}. \quad (2.4)$$

avec Y_t^* et Y_{t-K}^* sont les variables aléatoires obtenues par régression linéaire sur $Y_{t-1}, \dots, Y_{t-K+1}$.

2.4 Modélisation et prévision

La modélisation d'une série chronologique est une représentation de la série par un modèle qui décrit son évolution dans le temps. Plusieurs méthodes de modélisation des séries chronologiques ont été établit, citons en particulier, les méthodes du lissage exponentiel et la méthode de Box et Jenkins. La prévision d'une série consiste à déterminer son évolution sur les périodes non observées.

On distingue généralement trois types de prévisions [12] :

- La prévision à court terme : elle a un horizon de 6 à 18 mois ;
- La prévision à moyen terme : elle a un horizon de 2 à 5 ans ;
- La prévision à long terme : elle a un horizon de 6 à 20 ans.

2.4.1 Modèles de prévision

2.4.1.1 Modèles stationnaires

a) Processus moyenne mobile d'ordre q , $MA(q)$

Le processus utilisant plus d'une erreur passée est appelé un processus moyenne mobile d'ordre q , où q est le plus grand retard utilisé. Il est défini par l'équation suivante [18] :

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} = (1 - \theta_1 \beta - \theta_2 \beta^2 - \dots - \theta_q \beta^q) \epsilon_t = \phi(\beta) \epsilon_t.$$

où $\theta_1, \theta_2, \dots, \theta_q, q$ paramètres réels indépendants de $t, \phi(\beta)$ l'opérateur moyenne mobile d'ordre q et ϵ_t : Processus bruit blanc.

b) Processus autorégressif d'ordre p , $AR(p)$

Il s'agit d'un processus utilisant plusieurs variables du passé. p est le plus grand retard utilisé dans l'autorégression. Il s'écrit :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t.$$

ϕ est appelé paramètre du processus.

c) Processus mixte ARMA(p,q)

Il s'agit de processus utilisant plus d'une variable du passé ou plus d'une innovation, tel que p est le retard maximum des variables du passé et q le retard maximum des innovations du passé. Il s'écrit :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}.$$

Equivalent à :

$$\phi_\beta X_t = \Theta(\beta) \epsilon_t.$$

où les innovations ϵ_t constituent un processus bruit blanc de variance σ^2 .

$\phi(\beta)$: s'appelle le polynôme autorégressif de degré p ;

$\Theta(\beta)$: s'appelle le polynôme moyenne mobile de degré q .

2.4.1.2 Modèles non stationnaires

Les processus AR, MA et ARMA sont introduit pour des processus stationnaires. Or, en général les séries chronologiques ont une tendance et aussi une saisonnalité. Par conséquent l'utilisation de ces modèles est limité. Une différence première d'une série non stationnaire peut être stationnaire. Il est donc possible de rendre stationnaire une série non stationnaire [18].

a) Processus ARIMA(p,d,q) : Une série suit un processus ARIMA(p,d,q) si elle suit un processus ARMA(p + d,q) défini par :

$$\phi(\beta)(1 - \beta)^d X_t = \Theta(\beta) \epsilon_t.$$

$$\phi(\beta) \nabla^d X_t = \Theta(\beta) \epsilon_t.$$

où : ∇^d l'opérateur de différenciation, $\phi(\beta)$ est un polynôme de degré p et $\Theta(\beta)$ est un polynôme de degré q .

b) Processus SARIMA(p, d, q)(P, D, Q)_s : Une série (Y_t) suit un processus SARIMA (p, d, q)(P, D, Q)_s, si cette série a une saisonnalité de période s . On peut l'écrire de la manière suivante :

$$\phi_1(\beta) \phi_2(\beta^s) (1 - \beta)^d (1 - \beta^s)^D Y_t = \Theta_1(\beta) \Theta_2(\beta^s) \epsilon_t.$$

où ϕ_1 est un polynôme de degré p ;

ϕ_2 est un polynôme de degré P ;

Θ_1 est un polynôme de degré q ;

Θ_2 est un polynôme de degré Q .

Remarque 2.4.1. Les modèles ARIMA et SARIMA permettent de représenter des séries non stationnaires puisqu'il faut leurs appliquer les opérations de différentiation pour avoir un bruit blanc.

Modèle	Autocorrélation	Autocorrélation partielle	Condition de stationnarité	Condition de réversibilité
Bruit blanc	nulle pour $k > 0$	nulle pour $k > 0$	non	non
AR(p)	fonction dans l'enveloppe	nulle pour $k > 0$	oui	non
MA(q)	nulle pour $k > p$	fonction dans l'enveloppe	non	oui
ARMA(p,q) dans l'enveloppe	fonction amortie $k > q - p$	comportement général $k > q - p$	oui	oui

Tableau 2.1 – Les critères de choix d'un modèle de prévision

2.5 Méthode de Box et Jenkins

Dans la méthodologie d'analyse des séries chronologiques synthétisée par Box et Jenkins, on utilise trois types de processus ARMA, ARIMA ou SARIMA pour construire un modèle restituant le mieux possible le comportement d'une série temporelle.

2.5.1 Les prévisions par la méthode de Box et Jenkins

2.5.1.1 Familiarisation avec les données :

L'utilisateur s'informe d'abord sur le domaine dont relèvent les données (précision, incertitude,...). L'utilisateur examine les représentations graphiques des données qui peuvent révéler des changements de structures dans la série, des erreurs grossières.

2.5.1.2 Analyse préliminaire :

L'analyse préliminaire nous permet de détecter la non stationnarité et voir ainsi les transformations à faire sur la série brute. Cette étape est indispensable avant toute modélisation, on commence l'analyse de nos données par enlever les valeurs aberrantes, les erreurs de saisie, etc. Après avoir nettoyé la série, on trace son graphe ainsi que l'ACF et le PACF. A partir de ces graphes nous pouvons déduire la stationnarité ou la non stationnarité de notre série, et on peut aussi appliquer des tests de stationnarité tel que le test de Dickey-Fuller (simple ou augmenté), le test de Phillips-Perron [17][5].

2.5.1.3 Identification du modèle :

Cette étape consiste à identifier le modèle ARMA susceptible de représenter la série, nous allons définir l'ordre d si la série originale a une tendance, et l'ordre s si celle-ci présente une saisonnalité, le corrélogramme et le corrélogramme partiel nous permettent de déterminer le p et le q grâce aux pics significatifs, tels que le nombre de pics significatifs dans l'ACF représente Q_{max} et le nombre de pics significatifs dans le PACF représente P_{max} . Il existe trois modèles stationnaire : AR(p), MA(q) et ARMA(p, q); et si une différenciation est appliquée sur AR(p) et MA(q), on obtient le modèle ARIMA (p, d, q). Si en plus d'une différenciation, on applique un filtre saisonnier, on aura le modèle SARIMA (p, d, q).

2.5.1.4 Estimation des paramètres :

Après avoir terminé l'identification, il convient d'estimer les paramètres qui sont les coefficients des polynômes AR et MA. La méthode d'estimation la plus utilisée est celle du maximum de vraisemblance ou la méthode des moindres carrés.

2.5.1.5 Validation du modèle :

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

- a) **Test sur les paramètres :** On utilise le test de l'intervalle de confiance. Si le zéro n'appartient pas à l'intervalle de confiance du paramètre estimé, alors ce paramètre est significativement différent de zéro et vice-versa.
- b) **Test sur les résidus :** Lorsque le modèle est bien estimé, les résidus entre les valeurs observées et les valeurs estimées doivent se comporter comme un bruit blanc. Le test de Box et Pierce et le test de Box-Ljung nous permettent de le vérifier.

1. Test de Box et Pierce :

Il s'agit de tester

$$\begin{cases} H_0 : p_k = 0 \forall k; \\ H_1 : \exists k; p_k \neq 0. \end{cases}$$

Où $p_k = cov(y_t, y_{t-k})$ Le test repose sur la stat.

$$Q_k = n \sum_{t=1}^k \hat{p}_t^2.$$

$$\hat{p}_t^2 = \frac{\sum_{t=k+1}^n e_t e_{t-k}}{\sum_{t=k+1}^n e^{2t}}, t = 1 \dots nk = 1 \dots M.$$

Où $M = [\min(n/3), 3\sqrt{n}]$

Sous $H_0, Q_k \rightsquigarrow X_{k-l}^2$,

l =nombre de paramètres estimées. On rejette H_0 si $H_0 > X_{k-1}^2(1 - \alpha)$.

2. **Test de Ljung-Box** : Ce test repose sur la statistique suivante :

$$\phi^* = n(n - 1) \sum_{i=1}^n \frac{p_i^2}{n^{-i}} \forall i = 1 \dots M.$$

Où n désigne la taille de l'échantillon et p_t^2 désigne les coefficients d'autocorrélation théoriques.

Cette statistique est aussi distribuée selon une Khi-deux à degrés de liberté.

Si $\phi^* > X_{M,\alpha}^2$ nous rejetons H_0 , le modèle estimé n'est pas acceptable.

Si $\phi^* < X_{M,\alpha}^2$ nous acceptons l'hypothèse H_0 , donc le modèle estimé est acceptable.

Le critère d'information d'Akaike (AIC) : En anglais Akaike information criterion est une mesure de la qualité d'un modèle statistique proposée par Hirotugu Akaike en 1973. Lorsque l'on estime un modèle statistique, il est possible d'augmenter la vraisemblance du modèle en ajoutant un paramètre. Le critère d'information d'Akaike, tout comme le critère d'information bayésien, permet de pénaliser les modèles en fonction du nombre de paramètres afin de satisfaire le critère de parcimonie. Akaike a démontré que le meilleur des modèles ARMA validés est celui qui minimise la valeur AIC suivante :

$$AIC(p, q) = n \log(\hat{\sigma}_\epsilon^2 + 2(p + q)).$$

Où n est la taille de l'échantillon.

p est l'ordre de la partie AR.

q est l'ordre de la partie MA.

$\hat{\sigma}_\epsilon^2$ est la variance de l'erreur de prévision.

2.5.1.6 La prévision :

C'est la dernière étape de la méthodologie de Box et Jenkins. Etant donné une série stationnaire, observée entre 1 et T , on cherche à faire de la prévision à horizon h , et donc de prévoir Y_{T+1}, \dots, Y_{T+h} .

Il s'agit de calculer les prévisions optimales du modèle ARIMA estimé, à savoir $Y_T(h)$ la prévision de Y_{T+h} sachant l'ensemble d'information disponible en T , noté :

$$\hat{Y}_{T+1} = E(Y_{T+h} / Y_T, Y_{T-1}, \dots, Y_1).$$

Les étapes de la méthode de Box-Jenkins peuvent être résumées par le schéma donné dans la figure (2.1) :

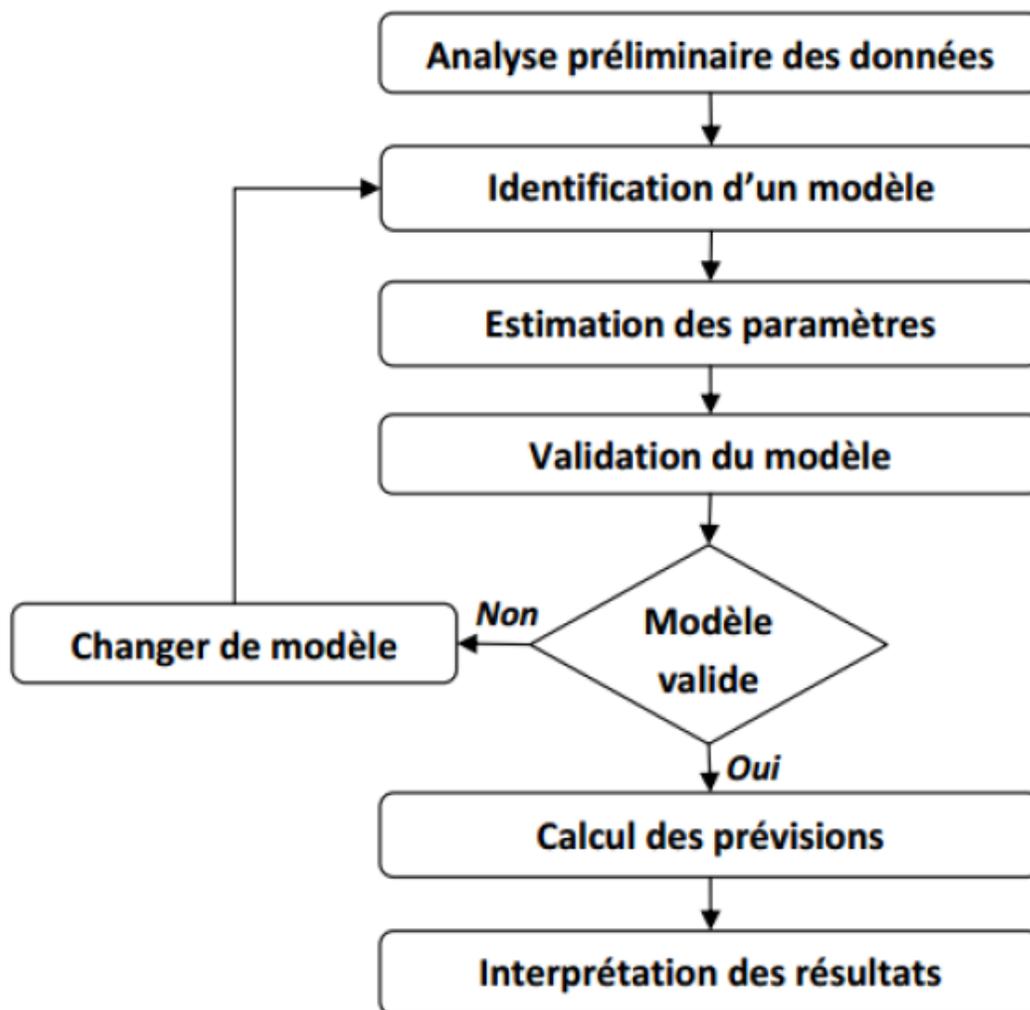


FIGURE 2.1 – Organigramme de la démarche Box et Jenkins.

2.6 Méthode du lissage exponentiel

Les méthodes de lissage exponentiel ont été proposées à la fin des années 60 . Leur principe se base sur le fait que pour le calcul des prévisions, on accorde plus de poids (plus d'importance) aux valeurs de l'historique par ordre décroissant d'une façon exponentielle au fur et à mesure qu'on remonte dans le passé. En d'autres termes, les informations récentes sont jugées plus indispensables pour le calcul des prévisions [1].

2.6.1 Lissage exponentiel simple

Cette méthode de lissage exponentiel a été proposée par Brown en 1959. Elle s'adapte au cas où la série peut être ajustée par une droite au voisinage de T , c'est-à-dire que la série originale ne présente pas de saisonnalité et sa tendance est constante.

Si on dispose d'une série chronologique Y_1, Y_2, \dots, Y_T , le modèle s'écrira sous la forme suivante :

$$Y_t = a_t + \epsilon_t, t = 1 \dots T.$$

où a_t : est le niveau de la série, ϵ_t : est la composante irrégulière.

2.6.2 Lissage exponentiel double

Quand la courbe de la série présente une tendance, le lissage simple n'est plus adapté. Alors, on utilise le lissage double de Brown pour modéliser cette série. On suppose que la série peut être ajustée localement par une droite au voisinage de T .

L'équation de la série sera alors :

$$Y_t = a_0 + (t - T)a_1 + \epsilon_t; t \in 1 \dots T;$$

2.6.3 Lissage exponentiel de Holt

La méthode de lissage de Holt est applicable à une série qui admet une tendance linéaire mais pas une saisonnalité.

2.6.3.1 Lissage exponentiel de Winters

La méthode de lissage de Winters est une généralisation de la méthode de Holt. Cette méthode est adaptée aux séries présentant une tendance et une saisonnalité. Il existe deux types de modèles de Winters :

2.6.3.2 Le modèle additif

est obtenu en ajoutant une composante saisonnière au modèle de Holt. La série peut être ajustée au voisinage de T par :

$$Y_t = a_0 + a_1(t - T) + S_t + \epsilon_t; t \in 1 \dots T;$$

2.6.3.3 Modèle multiplicatif

est obtenu en multipliant la composante saisonnière au modèle de Holt. On aura :

$$Y_t = [a_0 + a_1(t - T)]S_t \times \epsilon_{t+}; t \in 1 \dots T.$$

2.7 Conclusion

Dans ce chapitre, nous avons donné un bref aperçu sur quelques méthodes de prévision. La méthode de Box et Jenkins est adaptée au cas des prévisions à long, moyen et court terme tandis que les méthodes du lissage exponentiel se limitent aux prévisions à court et moyens terme. En effet, la méthode de Box et Jenkins possède une base mathématique plus solide que les méthodes de lissage exponentiel. C'est pour cela que pour nos prévisions, nous allons faire appel à la méthode de Box et Jenkins.

Régression linéaire et prévision

3.1 Introduction

La régression linéaire est l'une des méthodes les plus fréquemment utilisée pour faire de la prévision. Elle a pour but d'expliquer les fluctuations d'une variable dépendante notée Y en fonction d'une combinaison linéaire de variables $X_1; \dots; X_p$ dites explicatives ou indépendantes [6].

Dans ce chapitre, nous allons introduire le modèle de régression linéaire. On parle de régression multiple.

3.2 Le modèle de la régression linéaire multiple

Étant donné un échantillon $(Y_i; X_i^1, \dots, X_i^2, \dots, X_i^p); i = 1 \dots n$. Nous cherchons à expliquer, avec plus de précision possible, les valeurs prises par la variable endogène Y_i à partir d'une série de variables explicatives $X_i^1, \dots, X_i^2, \dots, X_i^p$. Le modèle théorique prend la forme suivante [7] :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_p X_{ip} + \epsilon_i; i = 1, \dots, n. \quad (3.1)$$

Où,

- Y est la variable à expliquer ;
- $(X^1 \dots X^p)$ sont les variables explicatives ;
- ϵ est le terme d'erreur aléatoire du modèle ;
- $\beta_0, \beta_1, \dots, \beta_p$ sont les paramètres du modèle que l'on veut estimer à l'aide des données.

Remarque 3.2.1. La désignation “multiple” fait référence au fait qu'il y a plusieurs variables explicatives X_i pour expliquer Y . La désignation “linéaire” correspond au fait que le modèle (3.1) est linéaire.

3.2.1 Hypothèses stochastiques

Dans cette partie nous donnons les hypothèses permettent de déterminer les propriétés des estimateurs (biais, convergence) et leurs lois de distributions pour les estimations par intervalle et les tests d'hypothèses [8][9][14].

Hypothèse 3.2.1. Les erreurs sont centrées, c.à.d $E[\epsilon_i] = 0, \forall i \in \{1, \dots, n\}$.

Hypothèse 3.2.2. La variance des erreurs est constante, on parle d'homogénéité des variances ou encore d'homoscédasticité :

$$V(\epsilon_i) = \sigma_\epsilon^2 \forall i \in \{1, \dots, n\}.$$

Hypothèse 3.2.3. Les ϵ_i sont indépendants et identiquement distribués de la loi normale :

$$\epsilon_i \rightsquigarrow \mathcal{N}(0, \sigma_\epsilon^2).$$

Hypothèse 3.2.4. Les termes X_i sont supposés fixés (facteurs contrôlés) ou bien l'erreur ϵ est indépendante de la distribution conjointe de $X_1 \dots X_p$. On écrit dans ce dernier cas que [19] :

$$E(Y/X_1 \dots X_p) = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1} + \epsilon.$$

$$Var(Y/X_1 \dots X_p) = \sigma_\epsilon^2.$$

Hypothèse 3.2.5. L'erreur est indépendants des variables explicatives $E(\epsilon_i, X_i) = 0$.

Hypothèse 3.2.6. Les paramètres inconnus sont supposés constants.

3.3 Forme matricielle du modèle

Ce modèle s'écrit sous la forme matricielle suivant :

$$Y = X\beta + \epsilon.$$

Où Y est le $(n \times 1)$ vecteur des observations relatives à la variable expliquée.

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{pmatrix}.$$

X est la $(n \times (p + 1))$ qui contient l'ensemble des observations, avec une première colonne formée par la valeur 1 indiquant que l'on intègre la constante β_0 dans l'équation où p étant le nombre de

variables explicatives réelles.

$$X = \begin{pmatrix} 1 & X_{11} & \dots & X_{1p} \\ \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{pmatrix}.$$

ϵ est le $(n \times 1)$ vecteurs des erreurs.

$$\epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}.$$

Remarque 3.3.1. Le coefficient β_0 est un paramètre appelé intercepte qui représente la moyenne des Y_i lorsque la valeur de chaque variable explicative est égale 0. Les coefficients $\beta_i (i = 1; \dots; p)$ représentent le changement subi par $E(Y_i)$ correspondant à un changement unitaire dans la valeur de la $i^{\text{ème}}$ variable explicative, lorsque les autres variables explicatives demeurent inchangées.

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

3.3.1 Hypothèses structurelle

Hypothèse 3.3.1. La matrice $(X^t X)$ (X^t est la matrice transposée de X), est non singulière de rang p , c.à.d $\det(X^t X) \neq 0$ et $(X^t X)^{-1}$ existe. Cette hypothèse implique que les différents vecteurs X_i sont linéairement indépendants.

Hypothèse 3.3.2. $\frac{X^t X}{n}$ tend vers une matrice finie non singulière lorsque $n \mapsto \infty$.

Hypothèse 3.3.3. $n > p + 1$ le nombre d'observations est supérieur au nombre des paramètres du modèle $\beta_i (i = 0, \dots, p)$.

3.4 Estimation des paramètres

L'estimation des paramètres β_i se fait avec la méthode des moindres carrés en minimisant la fonction des erreurs, on définit un critère qui mesure l'adéquation entre les paramètres du modèle et les données observées.

3.4.1 Méthode des moindres carrés

Le principe des moindres carrés consiste à rechercher les valeurs des paramètres qui minimisent la somme des carrés des résidus et qui proche mieux la dépendance des Y sur X .

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \hat{\beta}_{p-1} x_{i,p-1})^2. \quad (3.2)$$

où e_i est l'erreur observée (le résidu) est une évaluation de terme d'erreur ϵ .
c'est-à-dire à minimiser la fonction objectif :

$$z = f_{(\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1})} = \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \hat{\beta}_{p-1} x_{i,p-1})^2. \quad (3.3)$$

pour trouver le paramètre β qui minimisent z on doit résoudre :

$$\frac{\partial f(\beta)}{\partial \beta} = 0.$$

On a :

$$\begin{aligned} z &= f(\beta) = \sum_{i=1}^n [y_i - (\beta_0 + \sum_{i=1}^n \beta_i x_{ij})]^2 \\ &= (Y - X\beta)^t (Y - X\beta) \\ &= Y^t Y - 2Y^t X\beta + \beta^t X^t X\beta \end{aligned}$$

Le minimum est atteint pour avoir le resultat suivant :

$$\hat{\beta} = (X'X)^{-1} X'Y$$

3.4.2 Propriétés des estimateurs

L'étude des propriétés d'un estimateur se résume à l'étude de sa convergence et son biais.

3.4.2.1 Biais

Pour le biais de $\hat{\beta}$ on doit calculer son espérance :

On a

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1} X'Y \\ &= (X'X)^{-1} X'(X\beta + \epsilon) \\ &= \beta + (X'X)^{-1} X'\epsilon \\ E[\hat{\beta}] &= \beta + (X'X)^{-1} X'E[\epsilon] \\ &= \beta + (X'X)^{-1} X'E[\epsilon] = \beta \end{aligned}$$

L'estimateur $\hat{\beta}$ des moindres carrés est sans biais.

3.4.2.2 La matrice de variance-covariance

La matrice de variance-covariance des coefficients est importante, car elle renseigne sur la variance de chaque coefficient estimé, elle permet d'effectuer des tests d'hypothèses, notamment de voir si chaque coefficient est significativement différent de zéro, elle est définie comme suis [16] :

$$\begin{aligned} \text{Varcov}(\hat{\beta}) &= E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'] \\ &= E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] \\ &= \sigma_\epsilon^2 (X'X)^{-1} \end{aligned}$$

Et sa matrice de variance-covariance, notée par $\text{varcov}(\hat{\beta}) = \sigma_\epsilon^2 (X'X)^{-1}$ est de variance minimale parmi tous les estimateurs linéaires

$$\begin{pmatrix} v(\hat{\beta}_0) & \text{cov}(\hat{\beta}_0, \hat{\beta}_1) & \cdots & \text{cov}(\hat{\beta}_0, \hat{\beta}_p) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_0) & v(\hat{\beta}_1) & \cdots & \text{cov}(\hat{\beta}_1, \hat{\beta}_p) \\ \vdots & \cdots & \cdots & \vdots \\ \text{cov}(\hat{\beta}_p, \hat{\beta}_0) & \cdots & \cdots & v(\hat{\beta}_p) \end{pmatrix}$$

3.4.3 La variance estimée des erreurs

Le paramètre σ^2 est défini par

$$\sigma^2 = V(\epsilon_i)$$

(Hypothèse 3.2.1). soit $\hat{Y} = X\hat{\beta}$ comme estimateur de $E[Y]$, il apparaît naturel d'estimer σ^2 par

$$s^2 = \frac{1}{n - (p + 1)} \sum_{i=0}^n (y_i - \hat{y}_i)^2 = \frac{SCR}{n - p - 1}.$$

3.5 Analyse de la variance et coefficient de détermination

Un avantage du critère des moindres carrés est de fournir une estimation de qualité d'ajustement d'un modèle de régression fondé sur la décomposition de la variable dépendante Y , donc on va considérer que la quantité d'information totale contenue dans une variable Y est proportionnelle à sa variance, la qualité d'ajustement correspond au rapport entre information totale sur Y et l'information effectivement reconstituée à partir de la connaissance procurée par la variable X , est entre 0% (X n'apporte aucun élément de prévision sur Y) et 100% (la connaissance des valeurs de X permet de prédire intégralement les valeurs de Y) et dépend de l'intensité de la corrélation entre X et Y .

3.5.1 Le coefficient de détermination R^2

Le rapport entre SCE et SCT représente la proportion de variance expliquée et porte le nom de coefficient de détermination, noté par R^2 [16] :

$$R^2 = \frac{SCE}{SCT} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SCR}{SCT}$$

- SCT : représente la somme des carrés totale, elle signifie la quantité de variabilité dans les données originales ;
- SCE : représente la somme des carrés expliquée, elle signifie la quantité de variabilité dans le modèle ajusté aux données originales ;
- SCR : représente la somme des carrés résiduelle, elle signifie la quantité de variabilité non-expliquée par l'ajustement.

Ce coefficient R^2 est compris entre 0 et 1 : plus il est proche de 1 et plus grande est la part expliquée, autrement dit meilleure est la régression. Inversement, un coefficient R^2 proche de 0 indique que la quantité SCR est élevée.

3.5.2 Le coefficient de détermination ajusté R_{adj}^2

Le coefficient R^2 est un indicateur de la qualité de l'ajustement des valeurs observées par le modèle mais il a le défaut de ne pas tenir compte du nombre de variables explicatives utilisés dans le modèle. On ne peut pas l'utiliser pour comparer plusieurs modèles entre eux car, si on ajoute une variable explicative à un modèle, la part des erreurs diminue forcément et donc le coefficient R^2 augmente : cela signifie que plus il y a de variables explicatives et plus le R^2 est élevé. Or un modèle n'est pas nécessairement meilleur parce qu'il a plus de variables explicatives. On définit donc un coefficient R^2 ajusté qui tient compte des degrés de liberté. Ce coefficient, noté par R_{adj}^2 , est défini comme suit :

$$R_{adj}^2 = 1 - \frac{\frac{SCR}{n-p-1}}{\frac{SCT}{n-1}} = 1 - \frac{n-1}{n-p-1}(1 - R^2).$$

3.6 Les tests statistiques

Avant de juger que notre modèle est bon ou non, nous avons effectué des tests d'hypothèses qui suivent les démarches suivantes :

1. Choix d'hypothèses H_0 et H_1 ;
2. Détermination de la variable de décision ;

3. Détermination de la région critique en fonction de α ;
4. Décision : accepte ou rejette H_0 .

3.6.1 Test de Fisher

Ce test est utilisé pour montrer qu'un modèle de régression linéaire multiple est globalement significatif. Pour le construire nous utilisons l'hypothèse de normalité des erreurs :

$$\epsilon_i \rightsquigarrow \mathcal{N}(0, \sigma_\epsilon^2).$$

soit le test d'hypothèses :

$$\begin{aligned} H_0 : & \beta_1 = \beta_2 = \dots = \beta_p = 0; \\ H_1 : & \exists \beta_i \neq 0 \forall i \in 1 \dots p. \end{aligned}$$

On calcule F^* qui est un indicateur qui s'exprime comme un rapport de la variance expliquée sur la variance inexpliquée. On a :

$$\begin{aligned} F^* &= \frac{\frac{SCE}{p}}{\frac{SCR}{n-p-1}} = \frac{\frac{R^2}{1-R^2}}{\frac{R^2}{n-p-1}} \sim F_{(p, n-(p+1))} \\ &= \frac{\frac{p}{1-R^2}}{\frac{p}{n-p-1}} = \frac{p}{1-R^2} \cdot \frac{n-p-1}{p} \\ &= \frac{SCE}{SCR} = \frac{R^2}{1-R^2}. \end{aligned}$$

tel que :

- R^2 : le coefficient de détermination.
- P : est le degré de liberté de SCE.
- $(n - p - 1)$: est le degré de liberté de SCR.

On compare F^* calculée au F théorique, précalculée dans les tables de référence (table de Fisher) en fonction du risque de première espèce α et les degrés de liberté (p) et $(n - p - 1)$.

▷ Si on trouve $F^* > F$, on rejette H_0 donc le modèle est globalement bon.

▷ Si on trouve $F^* < F$, on accepte H_0 donc le modèle est rejeté [16].

3.6.2 Test de Student

Comme il est important de savoir la signification globale du modèle, il est aussi important de savoir la signification individuelle de chacune des variables explicatives (constantes incluses), en

effectuant le test de student, sous l'hypothèse de normalité des erreurs. En effectuant ce test, on peut savoir si chaque paramètre est significativement différent de zéro ou non [13][15].

$$\begin{aligned} H_0 : \beta &= 0 \forall i \in \overline{1..p}; \\ H_1 : \beta &\neq 0 \forall i \in \overline{1..p}. \end{aligned}$$

Sous H_0 on calcule le ratio de student :

$$t_{\hat{\beta}_i}^* = \left| \frac{\hat{\beta}_i}{\hat{\sigma}_\epsilon} \right| \rightsquigarrow t_{(n-(p+1))}.$$

Où $\hat{\sigma}$ est l'écart type estimé, obtenu à partir de la matrice des variances- covariances estimée :

$$\text{varcov}(\hat{\beta}) = \sigma_\epsilon^2 (X'X)^{-1}.$$

On compare $t_{\hat{\beta}_i}$ à $t_{(\alpha/2, n-p-1)}$, lu sur la table de student.

- ▷ Si $t_{\hat{\beta}_i} > t_{(\alpha/2, n-p-1)}$, alors on rejette H_0 .
- ▷ Si $t_{\hat{\beta}_i} < t_{(\alpha/2, n-p-1)}$. alors on accepte H_1

3.7 Intervalles de confiances

L'estimation par intervalle de confiance est une autre manière d'estimer les paramètres de régression mais cette fois-ci par intervalle. Au niveau de confiance α , l'intervalle de confiance pour chaque coefficient de régression β_i tel que ($i = 0, \dots, p$) est donnée par :

$$IC_{\beta_i} = [\hat{\beta}_i - \hat{\sigma} t_{(\alpha/2, n-p-1)}; \hat{\beta}_i + \hat{\sigma} t_{(\alpha/2, n-p-1)}].$$

Avec

- $t_{\alpha/2}$ est le quantile de la distribution t.
- $\hat{\sigma}$ est l'écart type estimé.

L'intervalle de confiance permet de savoir s'il y a une relation entre la variable explicative X_{ip} et la variable expliquée Y , en tenant compte des autres variables explicatives qui sont dans le modele.

- ▷ Si IC_{β_i} exclut la valeur zéro, alors on déduit qu'il y a une association linéaire significative entre Y et X_{ip}
- ▷ Si IC_{β_i} inclut la valeur zéro, alors on déduit que la variable X_{ip} n'apporte aucune information supplémentaire pour la prédiction de Y [16].

3.8 Prévision

En connaissant les valeurs des variables explicatives, on peut prédire la valeur que peut être attribuée à la variable expliquée. Le modèle estimé est donnée par :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_p x_{pi}.$$

La prévision pour la donnée $(i + h)$ est la suivante :

$$\hat{y}_{i+h} = \hat{\beta}_0 + \hat{\beta}_1 x_{1i+h} + \dots + \hat{\beta}_p x_{pi+h}.$$

L'erreur de la prévision est donnée par :

$$e_{i+h} = y_{i+h} - \hat{y}_{i+h}.$$

Et l'intervalle de confiance pour les prévisions est donnée par :

$$IC_{y_{i+h}} = [\hat{y}_{i+h} \sigma \hat{\epsilon}_{i+h} t_{(\alpha/2, n-p-1)}; y_{i+h} \sigma \hat{\epsilon}_{i+h} t_{(\alpha/2, n-p-1)}].$$

Avec

$$\hat{\sigma}_{\epsilon_{i+h}} = \sqrt{\hat{\sigma}_\epsilon (X'_{i+h} (X'X)^{-1}) X_{i+h} + 1}.$$

Et

$$Y_{i+h}^{\hat{}} = \begin{pmatrix} \hat{Y}_{11+h} \\ \vdots \\ \hat{Y}_{1p+h} \end{pmatrix}$$

Lorsque les variables explicatives sont dépendantes entre elles, le problème de multicollinéarité se pose, pour cela on propose la méthode d'élimination pas à pas pour régler ce problème.

3.9 Méthode d'élimination pas à pas (Backward Elimination procedure)

La méthode d'élimination "pas à pas" est une amélioration de "All possible régression" (méthode d'introduction progressive) dans la quelle on n'examine pas toutes les équations de régression mais seulement la meilleure régression contenant certains nombres de variable [6]. Soit l'équation de régression avec toutes les variables :

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p.$$

On calcule la valeur F-test (teste de Fisher) pour chaque variable traitée comme si elle était la dernière à rentrer dans le modèle : Soient $F_i, i = 1, \dots, p$ les valeurs respectives des F-tests calculées, Où :

$$F_i = \frac{\frac{\Delta(SCR)}{p-h}}{\frac{SCE}{n-(p+1)}}$$

Telque

$\Delta(SCR) = SCR_i - \widehat{(SCR)}$ et h étant le nombre de régressions qui reste. Puis on calcule F_L et F_0 tel que :

$$F_L = \min F_i, i = 1, \dots, p$$

F_0 La valeur théorique de F-test.

- ▷ Si $F_L < F_0$ on rejette la variable X_L et on considère l'équation de régression sans X_L , et on passe à la variable suivant.
- ▷ Si $F_L > F_0$ on adopte le modèle.

3.10 Conclusion

Dans ce chapitre, nous avons donné un bref aperçu sur les notion de base de la régression linéaire. Comme il s'agit d'un modèle avec une seule ou plusieurs variables explicative ; le chapitre suivant sera dédié à la modélisation et l'application numérique des deux méthodes que nous avons présenté.

Modélisation et résolution

4.1 Introduction

Ce chapitre 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 ce chapitre est consacré 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.

4.2 Présentation du logiciel R

R est un langage de programmation et un environnement mathématique utilisés pour le traitement de données et l'analyse statistique. Ce logiciel a été développé dans les années 90 par Robert Gentleman et Ross Ihaka , c'est un logiciel libre (avec code source) et peut être distribué librement. Il est le plus populaire dans la communauté des statisticiens-chercheurs car :

De nombreuses techniques statistiques modernes et classiques ont été implémentées.

L'utilisateur n'est pas limitée par les procédures préprogrammées par un package.

Il comporte de nombreuses fonctions pour les analyses statistiques et les graphiques, ceux-ci sont visualisé immédiatement dans une fenêtre propre et peuvent être exportés sous divers formats jpg, png, bmp, ps, pdf, emf, pictex et xfig. Les résultats des analyses statistiques sont affichés à l'écran, certains résultats partiels (valeurs de P, coefficients de régression, résidus,...) peuvent être sauvés à part, exportées dans un fichier utilisés dans des analyses ultérieures.

4.3 Description des données

4.3.1 Données statistiques

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 :

4.3.2 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

Tableau 4.1 – Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

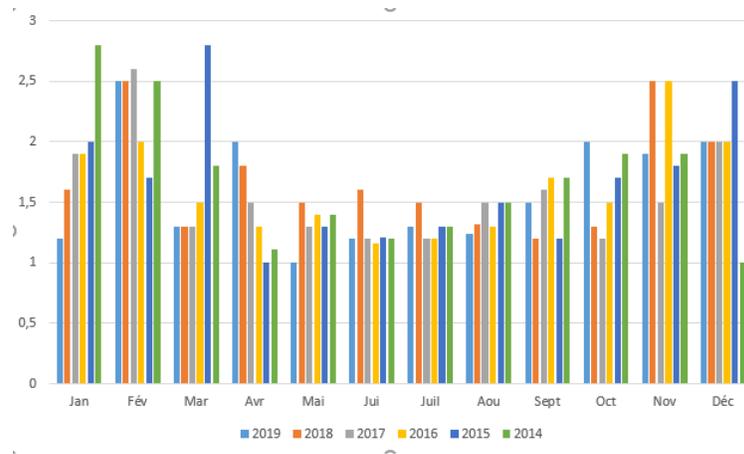


FIGURE 4.1 – Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

D'après la figure (4.1), 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
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

Tableau 4.2 – Les quantités de tonnage engendrées par le trafic des hydrocarbures (2015-2020)

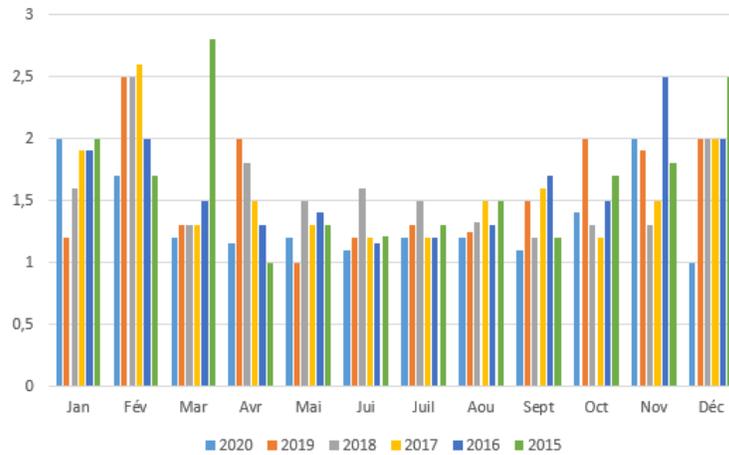


FIGURE 4.2 – Tonnage engendré par le trafic des hydrocarbures (2015-2020)

D'après la figure (4.2), 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

Tableau 4.3 – Les nombre de navires (porte-conteneurs) traités durant l'année 2019

4.4 Prédiction par de la méthode de Box-Jenkins(2020)

Dans cet exemple, on s'intéresse à effectuer une prédiction 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 4.1). 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.

4.4.1 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 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)
```

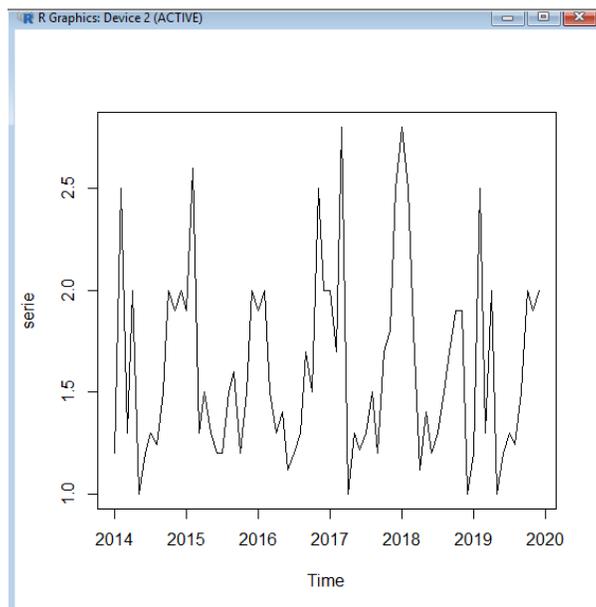


FIGURE 4.3 – Le graphe de la série originale

Le graphe de la série originale figure(4.3) 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)
```

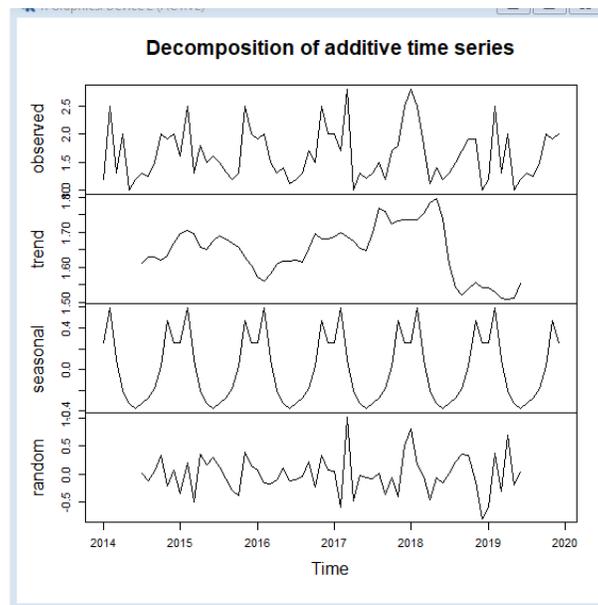


FIGURE 4.4 – Graphe de la série et ses différentes composantes.

4.4.2 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)
```

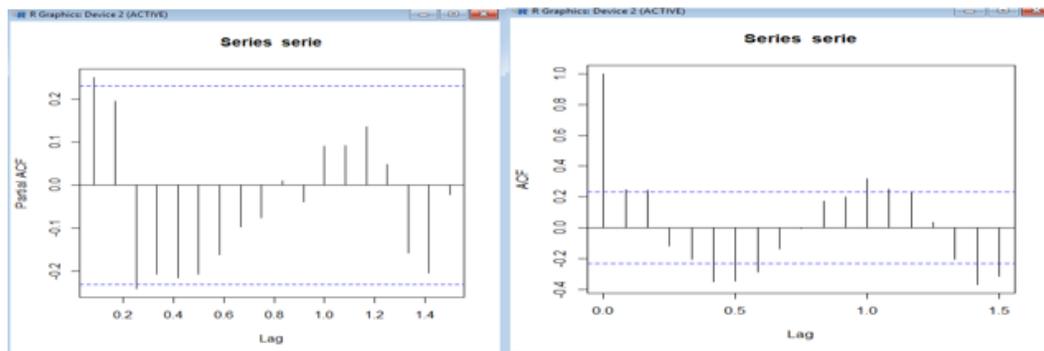


FIGURE 4.5 – Corrélogramme et corrélogramme partiel de la série originale

4.4.3 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)
```

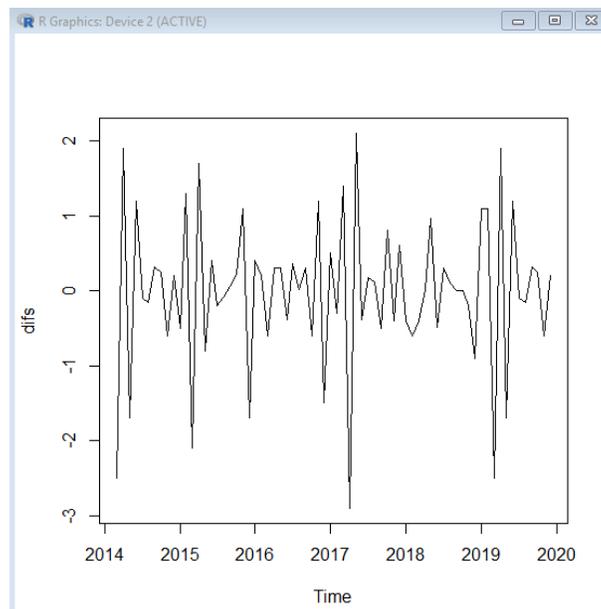


FIGURE 4.6 – Graphe de la série chronologique differencier

4.4.4 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)
```

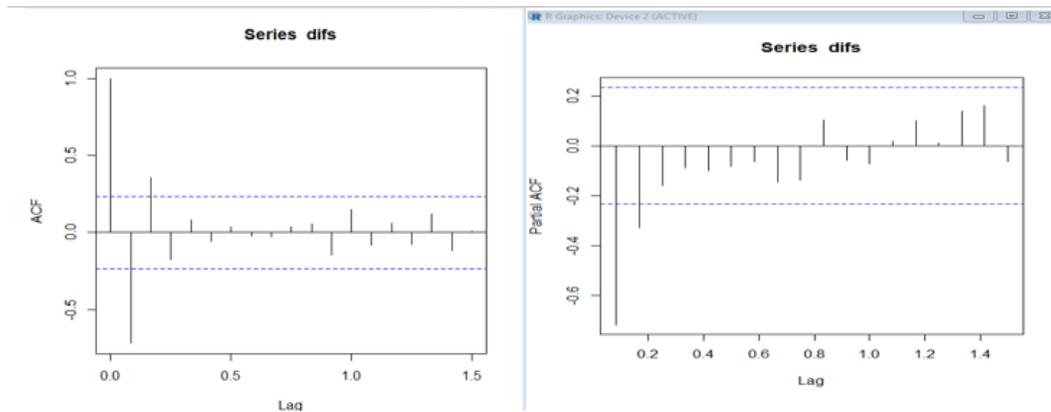


FIGURE 4.7 – Corrélogramme et corrélogramme partiel de la série differencier

4.4.5 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

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

4.4.6 Validation du modèle

4.4.6.1 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

4.4.6.2 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 : $Min(n/2, 3\sqrt{n})$ 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 $Q^* = 17.578 < \chi_{22}^2(0,05) = 26,245$, 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)
```

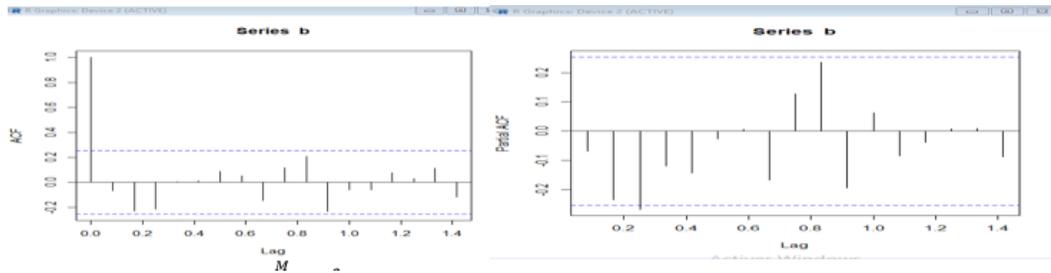


FIGURE 4.8 – Corrélogramme et corrélogramme partiel des résidus.

4.4.7 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 prévision 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

Tableau 4.4 – Les résultats de la prévision de tonnage engendré par le trafic des hydrocarbures pour 2020

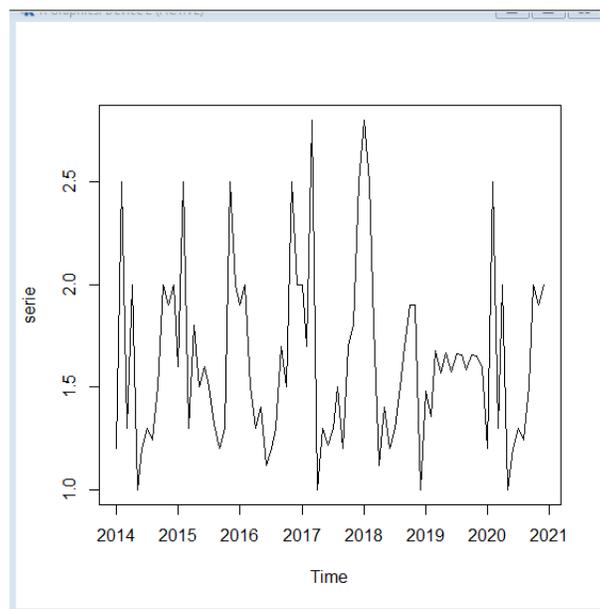


FIGURE 4.9 – Graphe de la série et la prévision

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

4.5 Prévision 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.

4.5.1 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, 1.5, 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(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)
```

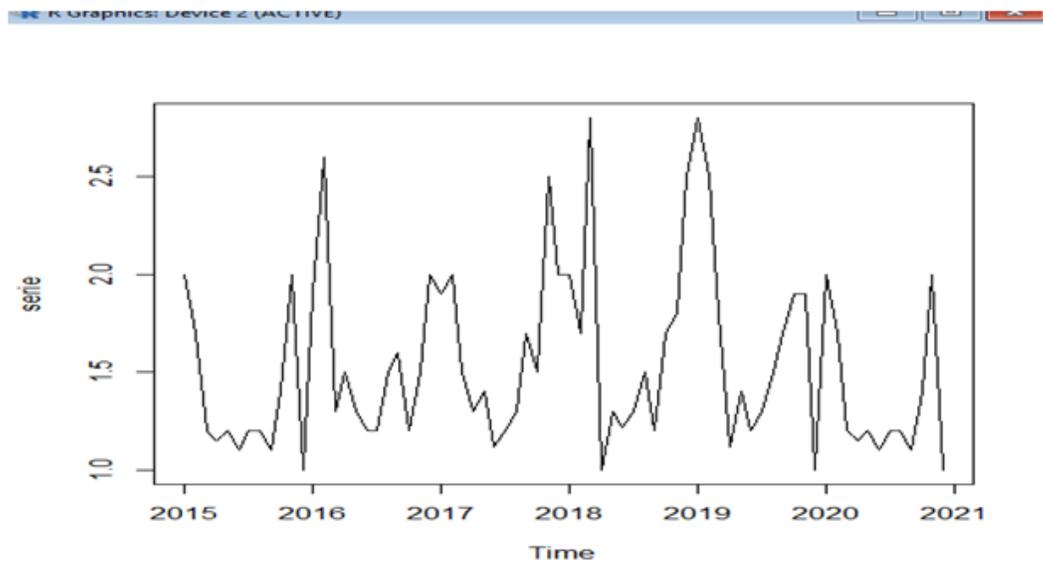


FIGURE 4.10 – Graphe de la série originale

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

```
> acf(serie)
> pacf(serie)
```

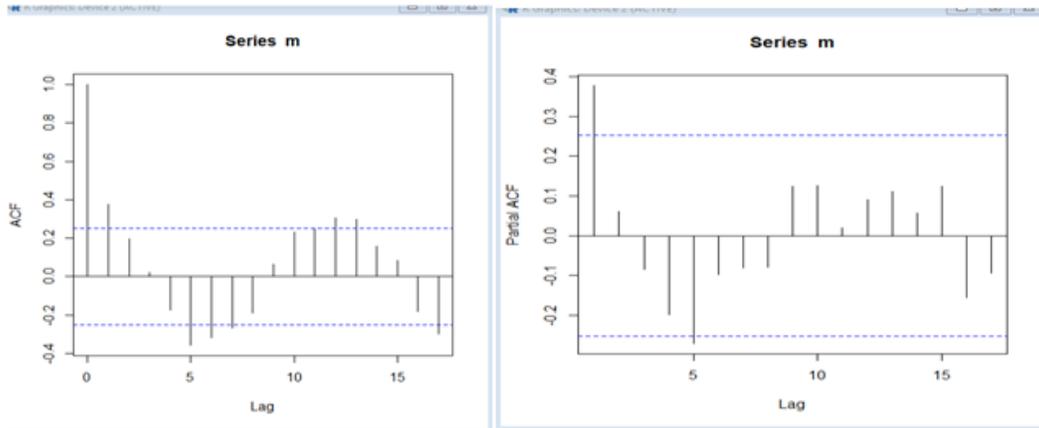


FIGURE 4.11 – Corrélogramme et corrélogramme partiel de la série originale

4.5.2 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,

4.5.3 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.

4.5.4 Prédiction

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

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

> 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 prédiction 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

Tableau 4.5 – Les résultats de la prédiction de tonnage des hydrocarbures pour l'année 2021

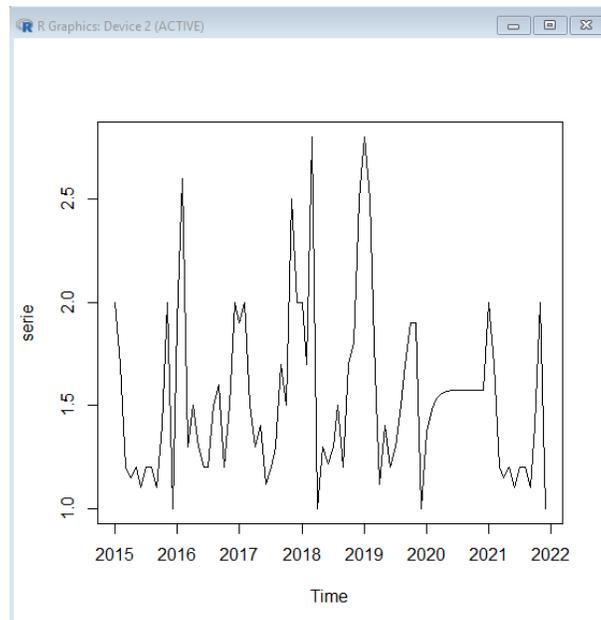


FIGURE 4.12 – Graphe de la série et la prévision

4.5.5 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.

4.6 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.

4.6.1 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(temps) < AIC global, donc on supprime la variable temps.

D'où les variables retenues sont : Importation, Exportation, Livraison.

Le modèles'écrit :

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

D'après la fonction step(), nous obtenons 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

$$Nbr\hat{N}avire = 28.1 + (0.001)Importation + (-0.004)Exportation + (-0.001)Livraison.$$

4.6.2 Les tests statistiques de Student et Fisher

Pour réaliser les tests de Student et Fisher, nous utilisons la fonction `summary()` sous 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 $R^2 = 0.54$, d'où nous avons un bon ajustement linéaire.

4.6.3 Test global de Fisher

Nous testons les hypothèses suivantes :

$$H_0 : \beta_0 = 0; \beta_1 = 0; \beta_2 = 0; \beta_3 = 0;$$

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 (F-statistic) F^* :

On a $F^* = 3.535$, au seuil de signification de $\alpha = 0.05$ la valeur sur la table de Fisher est $F(4, 9, 0.05) = 2.60$, comme $F^* > F(4, 9, 0.05)$ on rejette H_0 , donc le modèle est globalement significatif.

4.6.4 Test individuel de Student

Au seuil de signification $\alpha = 0.05$ on a :

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

On accepte 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

On accepte 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

On rejette 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

On accepte l'hypothèse H_1 .

4.6.5 Intervalle de confiance

On calcule l'intervalle de confiance pour les quatre coefficients avec la fonction `confint()` sous R.

```
> confint(lm(NbrNavire~Importation+Exportation+Livraison,data=D))
                2.5 %          97.5 %
(Intercept)  1.204171e+01 44.172817408
Importation  3.736097e-05  0.002080983
Exportation  -1.122047e-02  0.002739299
Livraison    -3.748486e-03  0.001071472
> |
```

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 $\hat{\beta}_3$: $IC_3 = [-3.748486 \exp(-3); 0.001071472]$.

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

Alors le modèle est validé.

4.6.6 Prévision

Dans le but de réaliser une prévision de nombre de navires qui seront traités pour l'année 2020, nous utilisons la fonction `predict()` sous R.

Ci-dessus les résultats approximatifs de la prévision :

```
> predict(lm(NbrNavire~Importation+Exportation+Livraison))
      1      2      3      4      5      6      7      8      9     10     11     12
23.51314 20.24205 24.38778 22.89912 25.01979 23.02822 25.27680 23.10417 22.90668 18.36191 22.90369 21.35665
> |
```

Le tableau suivant présente les résultats de la prévision et le nombre réel des 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

Tableau 4.6 – Nombre des navires réelles et prévus pour l'année 2020

4.6.7 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

4.7 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 tendance 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.

Conclusion générale

À la lumière de notre travail, nous avons présenté dans ce mémoire 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.

Les résultats obtenus dans ce travail permettent de dégager quelques perspectives de recherche :

- Modélisation statistique et prévision par les méthodes non paramétriques.
- Modélisation statistique en utilisant la régression non-paramétrique.

Bibliographie

- [1] S. Adjabi. "Cours des méthodes statistiques de prévisions". Département de Recherche Opérationnelle, Université de Bejaia, 2013-2014.
- [2] S. Adjabi. "Cours des méthodes statistiques de prévisions". Département de Recherche Opérationnelle, Université de Bejaia, 2003.
- [3] S. Adjabi, D. Aissani, and M.S. Radjef "Modèles statistiques de prévisions pour les demandes de branchement à un réseau électrique urbain." Lamos-Université de Bejaia, 1996.
- [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.
- [9] R. Bourbonnais,(1998). Econométrie, Manuel et exercices corrigés. Dunod, 2ème édition.
- [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.

-
- [15] A. M. Legendre, (1805) : Nouvelles méthodes pour la détermination des orbites des comètes, Paris, Firmin Didot.
- [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.
- [18] F. Teghbit et Y. Bouiche. "Modélisation et prévision des quantités totales de pétrole transportées par L. Oleoduc OB1 reliant Hassi Messaoud et Bejaia." . Mémoire de Master en STID, Université de Bejaia, 2013.
- [19] Y. Tillé, (2011). Résumé du Cours de Modèles de Régression. Institut de statistique, Université de Neuchâtel.

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 Béjaia ; Modélisation statistique ; Prévision ; Covid-19.

ABSTRACT

The application of operational research methods at company level is essential to improve the various service and production processes . . . , it's for this purpose that it's more necessary for a company to establish future planning in all areas and sectors, an essential prerequisite for reducing risks or even adding more profits, and this is done with forecasts. In this work, we have studied a key phenomenon of forecasting noted during our internship at Port Company of Bejaia (EPB). The objective of our study is to treat the evolution of hydrocarbon traffic and container traffic during the era of the Covid-19 crisis and to predict the quantities of hydrocarbon tonnage that will be generated for the year 2021, mathematically models formulated on the basis of statistical and linear modeling which better reflect the characteristics of the systems studied. First, we modeled the two times series using the Box and Jenkins method and forecast calculations. Secondly, with the principle of linear modeling, we established a model which will then be applied for the calculation of forecasts. the comparison between the results obtained and the actual values we were able to see the impact of the health crisis on the two terminals.

Key words : Port Company of Bejaia EPB ; Statical modeling ; Forecast ; Covid-19.