

TABLES DES MATIERES

DEDICACES	i
REMERCIEMENTS	ii
TABLES DES MATIERES.....	iii
NOTATIONS ET UNITES	vi
ACRONYMES.....	x
LISTE DES FIGURES.....	xi
LISTE DES TABLEAUX.....	xiv
1. Introduction générale	1
Chapitre 1. Généralités sur la mécanique	3
1.1. Introduction	3
1.2. La mécanique	4
1.2.1. Les référentiels de travail.....	4
1.2.1.1. Référentiel [6].....	4
1.2.1.2. Référentiel de Copernic [6]	4
1.2.1.3. Référentiel géocentrique R_G [6].....	5
1.2.1.4. Référentiel terrestre R_T [6]	6
1.2.2. Loi fondamentale en mécanique terrestre [7].....	7
1.2.2.1. Bilan des forces appliquées [9].....	8
1.2.2.2. Loi d'attraction Universelle [1]	11
1.2.2.3. Loi fondamentale de la mécanique terrestre	12
1.2.3. Conclusion	12
1.3. Les théories des marées depuis Isaac Newton jusqu'à Doodson.....	14
1.3.1. Modèle de Newton [3] [9]	14
1.3.2. Interprétation du modèle de Newton.....	16
1.3.2.1. Cas où le mouvement de rotation de la Terre sur elle-même est prépondérante	17
1.3.2.2. Cas où le mouvement de translation circulaire de la Lune autour de la Terre est prépondérante	18
1.3.2.3. Cas où le mouvement de translation circulaire de la Terre autour du Soleil est prépondérante .	19
1.3.3. Cas où la déclinaison δ est différent de zéro [3]	20
1.3.4. Modèle harmonique.....	22
1.3.4.1. Potentiel générateur des marées [11].....	22
1.3.4.2. Développement en espèce de Laplace [3].....	23
1.3.5. Conclusion	29
Chapitre 2. Matériels et méthodes	31
2.1. Introduction	31
2.2. Matériels	32
2.3. Méthodes.....	33

2.3.1.	Méthode de Hurst [12].....	33
2.3.2.	Méthode statistique de préparation et d'analyse des données	34
2.3.2.1.	Courbe représentative des données initiales	34
2.3.2.2.	Préparation des données	34
2.3.3.	Méthode spectrale de détection et de calcul de la périodicité S	37
2.3.3.1.	Estimation de la DSP.....	38
2.3.3.2.	Formulation de la fréquence f et de la période S [20].....	41
2.3.3.3.	Réduction du domaine d'étude	42
2.3.3.4.	Saisonnalité S en terme journalier	42
2.3.3.5.	Choix de la périodicité S	42
2.3.4.	Méthode de détection de la tendance.....	45
2.3.4.1.	Les tests de corrélation de rang de Mann et de Kendall [32] [33]	46
2.3.4.2.	Le test de corrélation sur les rangs de Spearman	48
2.3.5.	Méthode d'identification du schéma de prévision [36]	49
2.3.5.1.	Méthode du profil	51
2.3.5.2.	Méthode de la bande.....	51
2.3.5.3.	Méthode analytique	52
2.3.6.	Conclusion	52
2.4.	Les différents types de modèles.....	53
2.4.1.	Modèle Census X11 [38]	53
2.4.1.1.	Estimation de la Tendance $t_{t,1}$ par moyenne mobile centrée d'ordre S [36].....	53
2.4.1.2.	Estimation de la composante saisonnière – irrégulière $s_{t,1} + r_{t,1}$ [38]	53
2.4.1.3.	Estimation de la série corrigée des variations saisonnières $scvs_{t,1}$ [38]	54
2.4.1.4.	Estimation de la tendance-cycle $t_{t,2}$ par moyenne de Henderson à 13 ou 9 termes [38]	54
2.4.1.5.	Nouvelle expression de la composante saisonnière-irrégulière $s_{t,2} + r_{t,2}$ [38]	55
2.4.1.6.	Nouvelle expression de la série corrigée des variations saisonnières $scvs_{t,2}$ [38].....	55
2.4.2.	Modèle ARIMA (p+1,d,q) - GARCH (P,Q)	55
2.4.2.1.	Identification de d et p	57
2.4.2.2.	Identification de $q_{initial}$	60
2.4.2.3.	Choix du modèle ARIMA(p+1,d,q).....	61
2.4.2.4.	Analyse des Résidus	61
2.4.2.5.	Analyse de la qualité de la prévision	65
2.4.2.6.	Analyse de la stabilité du modèle	66
2.4.3.	Conclusion	67
Chapitre 3. Résultats et discussions : applications numériques au cas des marées de Hell-Ville		69
3.1.	Introduction	69
3.2.	Résultats et discussions.....	70
3.2.1.	Sources des données	70
3.2.2.	Localisation su site	70
3.2.3.	Exposant de Hurst H_I	71
3.2.4.	Représentation graphique des données initiales.....	72
3.2.5.	Histogramme	73
3.2.6.	Préparation des données homogénéisées.....	75
3.2.7.	Estimation de la saisonnalité S	77
3.2.7.1.	Estimation de la saisonnalité S par corrélogramme ou la méthode de BLACKMAN-TUCKEY	77

3.2.7.2.	Estimation de la saisonnalité S par périodogramme moyenné	78
3.2.7.3.	Estimation de la saisonnalité par périodogramme par la Méthode de l'Entropie Maximale ..	80
3.2.8.	Saisonnalité S en terme journalier	92
3.2.9.	Choix de la Saisonnalité S	94
3.2.9.1.	Test sur les Résidus	94
3.2.9.2.	Analyse de la Variance à un Facteur	95
3.2.10.	Analyse de la tendance sur les données stabilisées	98
3.2.10.1.	Test de corrélation de rang de Kendall.....	98
3.2.10.2.	Test de corrélation de rang de Mann-Kendall	99
3.2.10.3.	Test de corrélation de rang de Spearman	99
3.2.11.	Schéma de prévision	100
3.2.11.1.	Méthode du profil.....	100
3.2.11.2.	Méthode de la bande	101
3.2.11.3.	Méthode analytique	102
3.2.12.	Méthode Census X11	103
3.2.13.	Méthode ARIMA(p+1,d,q)	104
3.2.13.1.	Identification des ordres d, q et p du modèle ARIMA(p+1,d,q).....	104
3.2.13.2.	Identification de q_{\max}	118
3.2.13.3.	Choix du modèle ARIMA(p+1,d,q)	120
3.2.13.4.	Analyse des résidus	122
3.2.13.5.	Choix du modèle ARIMA(p+1,1,q)-GARCH(P,Q)	123
3.2.13.6.	Analyse de la qualité de la prévision.....	131
3.2.14.	Conclusion	156
Conclusion générale		157
Avantages et perspectives		159
Références bibliographiques		160
LISTE DE PUBLICATIONS		163

NOTATIONS ET UNITES

a_{k_2}	[m]	: Amplitude des classes de la méthode de sturges en mètre
a_{k_1}	[-]	: Estimation de l'autocorrelation
$\vec{a}(P, t)_{/R_T}$	[m.s ⁻²]	: Accélération relative ou l'accélération du point P à l'instant t dans R_T
$\vec{a}(P, t)_e$	[m.s ⁻²]	: Accélération d'entraînement ou accélération de P à l'instant t en considérant P fixe dans R_T
$\vec{a}(P, t)_c$	[m.s ⁻²]	: Accélération de Coriolis du point P à l'instant t
$\vec{a}(P, t)_{/R_C}$	[m.s ⁻²]	: Accélération absolue ou accélération du point P à l'instant t dans R_C
$\vec{a}(G_T, t)_{/R_C}$	[m.s ⁻²]	: Accélération du centre d'inertie GT de la Terre dans le référentiel de Copernic
$\vec{a}(P, t)_{/R_T}$	[m.s ⁻²]	: Accélération du fluide de masse m_p dans le référentiel terrestre
A	[-]	: Terme Constante
a	[m]	: Borne inférieure de la boîte à moustache en mètre
ad	[m]	: Ecart moyen
a	[-]	: Constante du modèle de Dickey Fuller Augmenté
B	[-]	: Opérateur de retard
b	[m]	: Borne supérieure de la boîte à moustache en mètre
bt	[m]	: Tendance du modèle de Dickey Fuller Augmenté en mètre
C	[m]	: Moyenne en valeur absolue des croissances de la tendance t_i en mètre
Census X11	[-]	: Méthode de dessaisonnalisation par moyenne mobile développé dans les années 50-60 au US Bureau of Census
d	[-]	: Ordre de différenciation du modèle ARIMA
d_{final}	[-]	: Ordre final de différenciation du modèle ARIMA
d_{min}	[-]	: Valeur de « d » minimisant l'écart type des séries différenciées
d_{TS}	[m]	: Distance entre GT et GS égale à 149 000 000 000 mètres
d_{TL}	[m]	: Distance entre GT et GL égale à 384 400 000 mètres
d_{k_2}	[m]	: Représente la borne inférieure du k-ième classe de l'histogramme
d_{k_2+1}	[m]	: Représente la borne supérieure du k-ième classe
E(S)	[Lag]	: Forme entière de la Saisonnalité S en jour ou en Lag
$E_1(\tau_b)$	[-]	: Représente la moyenne de τ_b
F_e	[Hz]	: Fréquence d'échantillonnage en Hertz
fft	[m]	: Transformée de Fourier Rapide
$\vec{F}_{G_i}(P, t)$	[N]	: Force d'attraction de la Terre, de la Lune, du Soleil appliqués à P à l'instant t en Newton
$\vec{F}_{ap}(P, t)$	[N]	: Forces appliquées à P autres que celle de la gravitation à l'instant t en Newton
$\vec{F}(P, t)$	[N]	: Résultante des forces appliquées à P en Newton
m_p	[kg]	: Masse d'une particule de fluide P en kilogramme
$\vec{F}_{ie}(P, t)$	[N]	: Force d'inertie d'entraînement en Newton
$\vec{F}_{ic}(P, t)$	[N]	: Force d'inertie de Coriolis en Newton
$\vec{F}_{G_T}(P, t)$	[N]	: Force d'attraction terrestre en Newton
$\vec{F}_{G_{S,L}}$	[N]	: Force d'attraction exercée par la Lune et le Soleil en Newton
$\vec{F}'_{ap}(P, t)$	[N]	: Forces non gravitationnelles autre que le poids en Newton
$\vec{F}_{marée} \Big _{(\vec{X}, \vec{Z}_0)}$	[N]	: Force des marées dans le repère (\vec{X}, \vec{Z}_0) en Newton
\vec{F}	[N]	: Forces appliquées en Newton
Fisher (e; f)	[-]	: Statistique de Fisher pour « (e; f) » degré de liberté
g	[m.s ⁻²]	: Accélération de la pesanteur égale à 9,81
G_S	[-]	: Centre d'inertie du Soleil
G_L	[-]	: Centre d'inertie de la Lune
G_T	[-]	: Centre d'inertie de la Terre
$\vec{G}_{G_{L,S}}(G_T, t)$	[m]	: Champ gravitationnel crée par la Lune et le Soleil au centre d'inertie de la Terre
$\vec{G}_{G_T}(G_{S,L}, t)$	[m]	: Champ de gravitation exercée par la Terre aux centres d'inertie du Soleil et de la Lune

$\overrightarrow{G_T G_L}$	[m]	: Vecteur reliant le centre d'inertie de la Terre au centre d'inertie de la Lune
$\overrightarrow{G_T G_S}$	[m]	: Vecteur reliant le centre d'inertie de la Terre au centre d'inertie du Soleil
$\ \overrightarrow{G_T G_L}\ $	[m]	: Norme du vecteur $\overrightarrow{G_T G_L}$ égale à 384 400 000 mètres
$\ \overrightarrow{G_T G_S}\ $	[m]	: Norme du vecteur $\overrightarrow{G_T G_S}$ égale à 149 000 000 000 mètres
Go	[m]	: Représente un giga
Ho	[-]	: Hypothèse de référence
$H1$	[-]	: Hypothèse alternative
$H_{13 \text{ ou } 9}$	[m]	: Moyenne de Henderson à 13 ou 9 termes en mètre
H	[rad]	: Angle horaire en radian
$H1$	[-]	: Exposant de Hurst
$h(t)$	[m]	: Hauteur des marées de la formule harmonique en mètre
h_i	[m]	: Représente l'amplitude de la composante « i »
H_{k_3}	[m]	: Entropie maximale
$hist$	[-]	: Fonction histogramme sous Matlab
i^3	[-]	: Parallaxe commun aux trois termes
I	[m]	: Moyenne en valeur absolue des croissances des résidus r_i en mètre
I_{THEIL}	[-]	: Indicateur de THEIL
I_{k_2}	[-]	: Représente l'intervalle du k-ième classe
JB	[-]	: Test de Normalité de Jarque et Berra
k	[-]	: Ordre du test de Ljung-Box
k_1	[-]	: Coefficients liés au premier retard, à la tendance et à la constante du modèle de Dickey Fuller Augmenté
k_2	[-]	: Nombre de classes
k_3	[-]	: Variable fréquentielle variant de 1 à L
K_t	[-]	: Kurtosis
k_i	[-]	: Déphasage ou situation de l'argument $V_{i,0}$
L	[-]	: Le nombre de points fréquents
Lag	[-]	: Rang des temps de prélèvement des hauteurs d'eau
log	[-]	: Logarithme
log_{10}	[-]	: Logarithme de base 10
l_p	[-]	: Entier positif
l	[-]	: Représente le 1, 2 et 3ème terme de l'espèce de Laplace
ln	[-]	: Logarithme népérien
m_{G_T}	[kg]	: Masse de la Terre égale à $5,97.10^{24}$ kilogrammes
m_{G_L}	[kg]	: Masse de la Lune égale à $7,35.10^{22}$ kilogrammes
m_{G_S}	[kg]	: Masse du Soleil égale à $1,99.10^{30}$ kilogrammes
$m_{inertielle}$	[kg]	: Masse inertielle de la Terre en kilogrammes
m_{grave}	[kg]	: Masse grave de la Terre en kilogrammes
$M_{3 \times 3}$	[m]	: Moyenne mobile centrée à 5 termes de coefficient de pondération égal à $\left\{ \frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{1}{9} \right\}$
M_{2m+1}	[m]	: Moyenne mobile centrée d'ordre impair égal à $2m+1$
M_{2l_p}	[m]	: Moyenne mobile centrée d'ordre pair égal à $2m$
$M_{3 \times 5}$	[m]	: Moyenne mobile centrée à 7 termes de coefficient de pondération égal à $\left\{ \frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15} \right\}$
M	[m]	: Le nombre de décomposition en bloc
$m1$	[-]	: Nombre de groupes de rangs identiques sur les rangs R_x
$m2$	[-]	: Nombre de groupes de rangs identiques sur les rangs R_y
m_T	[kg]	: Masse moyenne de la Terre en kilogrammes
N	[-]	: Le nombre de points temporels
N_x	[-]	: Nombre de rangs identiques contenu dans le x- ième groupe
N_y	[-]	: Nombre de rangs identiques contenus dans le y- ième groupe
$M1$	[m]	: Forme matricielle du vecteur représentatif des données en mètres
$N(0,1)$	[-]	: Loi normale centrée réduite
$p_{initial}$	[-]	: Ordre de la partie autorégressive choisi pour démarrer la modélisation ARIMA
p	[-]	: Ordre de la partie autorégressive du modèle ARIMA
pd	[m]	: Ecart probable en mètre
$p(x)$	[-]	: Loi de probabilité pour x
$\overrightarrow{PG_L}$	[m]	: Vecteur reliant le point P au centre d'inertie de la Lune en mètres

\overrightarrow{PG}_S	[m]	: Vecteur reliant le point P au centre d'inertie du Soleil en mètres
$\ \overrightarrow{PG}_L\ $	[m]	: Norme du vecteur \overrightarrow{PG}_L en mètres
$\ \overrightarrow{PG}_S\ $	[m]	: Norme du vecteur \overrightarrow{PG}_S en mètres
\overline{P}	[m]	: Force de pression
Poids(P, t)	[m]	: Poids
$p_{optimal}$	[-]	: Ordre de la partie autorégressive optimisant le modèle ARIMA
Plot(x)	[-]	: Fonction graphique sous Matlab
Q1	[m]	: Premier quartile en mètre
Q3	[m]	: Troisième quartile en mètre
q_i	[rad.s ⁻¹]	: Vitesse angulaire
q	[-]	: Ordre de la partie moyenne mobile du modèle ARIMA
$q_{initial}$	[-]	: Ordre de la partie moyenne mobile choisi pour démarrer la modélisation ARIMA
\bar{R}_x	[-]	: Rang moyen correspondant aux rangs R_x
\bar{R}_y	[-]	: Rang moyen correspondant aux rangs R_y
R_x	[-]	: Rangs de la hauteur des pleines mers ou des basses
R_y	[-]	: Rangs des temps de prélèvement des pleines et basses mers
r	[-]	: Rayon moyenne de la Terre
R	[m]	: Représente l'étendue d'une série en mètre
R_C	[-]	: Référentiel de Copernic
R_G	[-]	: Référentiel géocentrique
R_T	[-]	: Référentiel Terrestre
\hat{R}_{xx}	[m]	: Estimateur de la fonction d'autocorrélation R_{xx}
rs	[-]	: Coefficient de corrélation de rang de Spearman
r_t	[m]	: Représente la composante résiduelle du schéma de prévision en mètre
r_t'	[m]	: Processus rendu stationnaire par différenciation en mètre
s_t	[m]	: Représente la composante saisonnière du schéma de prévision en mètre
S	[Lag ou j]	: Saisonnalité ou périodicité en Lag ou en jour
s	[m]	: Ecart type non biaisé en mètre
S''	[-]	: Signe de la différence entre deux composantes de la grandeur X égale 1, 0 ou -1
$scvst_{t,1}$	[m]	: Expression initiale de la série corrigée des variations saisonnières en mètre
$scvst_{t,2}$	[m]	: Nouvelle expression de la série corrigée des variations saisonnières en mètre
S_t	[-]	: Skewness
T_e	[m]	: Période d'échantillonnage
t_t	[m]	: Représente la composante tendancielle du schéma de prévision en mètre
$t_{1-\alpha/2}$	[-]	: Statistique de Student au seuil de signification de 5%
t	[s]	: Temps civil en seconde
$\vec{T}(P, t)$	[N]	: Tension qui permet de maintenir le fluide de masse m_p en équilibre dans le référentiel terrestre
$U_{0,1,0,05}$	[-]	: Statistique de la loi normale centrée réduite
\overrightarrow{V}_{RT}	[m.s ⁻¹]	: Représente le vecteur vitesse du fluide de masse m_p dans le référentiel terrestre
\overrightarrow{V}_d	[m]	: Vecteur représentatif des données en mètres
$V_{i,0}$	[m]	: Argument astronomique à l'instant t=0
x_{max}	[m]	: Valeur maximale d'une chronique en mètre
x_{min}	[m]	: Valeur minimale d'une chronique en mètre
xcorr	[-]	: Estimateur non biaisé de la fonction d'autocorrélation R_{xx}
$x(t)$	[m]	: Variable représentative des données
X	[m]	: Représentation d'une grandeur en mètre
X_i	[m]	: Composante i de la grandeur X
X	[-]	: Axe du référentiel terrestre parallèle à la direction qui suit le méridien passant au point P dans le sens Nord-Sud
Xo	[-]	: Axe du système solaire dirigé vers l'étoile fixe 1
Y	[-]	: Axe du référentiel terrestre parallèle à la direction qui suit le parallèle au point P dans le sens Ouest-Est
Yo	[m]	: Axe du système solaire dirigé vers l'étoile fixe 2
Z1	[-]	: Représente la statistique du test de Kendall
Z2	[-]	: Représente la statistique de Mann-Kendall
Zo	[-]	: Axe du système solaire dirigé vers l'étoile fixe 3
(x,y)	[m]	: Coordonnées du point P dans le repère cartésien (G_T, \vec{X}, \vec{Z}_0) en mètres

$\ \vec{\omega}\ $ ou ω	[rad.s ⁻¹]	: Norme du vecteur angulaire de rotation de la terre égale à $0,729 \cdot 10^{-4}$ rad.s ⁻¹
Θ	[rad]	: Angle zénithale en radian
σ	[m]	: Représente l'écart type en mètre
δ	[rad]	: Déclinaison en radian
$\overrightarrow{\text{grad}}$	[-]	: Vecteur gradient
ζ	[m ³ .s ⁻¹ .kg ⁻¹]	: Constante de gravitation égale à $6,67 \cdot 10^{-11}$ m ³ .s ⁻¹ .kg ⁻¹
φ	[m]	: Latitude
τ	[]	: Représente la taille de l'échantillon prise égale à N, N/2, N/4...
σ_t^2	[m]	: Représente la variance des résidus
φ_i	[m]	: i-ème coefficient du modèle autorégressif
τ_a	[m]	: Coefficient de corrélation des rang de Kendall sans considération des rangs identiques
τ_b	[m]	: Coefficient de corrélation des rang de Kendall avec considération des rangs identiques
μ et g_1	[m]	: Représente respectivement la pente et la constante de la régression de l'écart type par la moyenne de chaque période
θ_j	[m]	: Coefficient moyenne mobile du modèle ARIMA
Φ	[m]	: Coefficient du modèle de Dickey Fuller Augmenté sans retard
Φ_j	[-]	: Coefficient des termes retardés du modèle de Dickey Fuller Augmenté
K	[m]	: Constante du modèle de GARCH, EGARCH ou GJR
γ_i	[-]	: Coefficients de la partie GARCH du modèle GARCH, EGARCH ou GJR
α_j	[m]	: Coefficients de la partie ARCH du modèle GARCH, EGARCH ou GJR
ξ_j	[m]	: Leverage du modèle EGARCH ou GJR
ν	[-]	: Nombre de degré de liberté de la statistique de Student
π	[-]	: pi égal à 3,14
$\chi^2 (l - 1)$	[m]	: Khi-2 à « (l - 1) » degré de liberté

ACRONYMES

<i>ARIMA</i>	: <i>Autoregressive Integrated and Moving Average</i>
<i>ARCH</i>	: <i>Autoregressive Conditional Heteroscedastic</i>
<i>CIA</i>	: <i>Critère d'Information d'Akaïke</i>
<i>CS</i>	: <i>Critère de Schwarz</i>
<i>diff</i>	: <i>Différentielle</i>
<i>DyACO</i>	: <i>Dynamique de l'Atmosphère du climat et des Océans</i>
<i>DSP</i>	: <i>Densité Spectrale de Puissance</i>
<i>DFA</i>	: <i>Dickey Fuller Augmenté</i>
<i>EGARCH</i>	: <i>Exponential Generalized Autoregressive Conditional Heteroscedastic</i>
<i>FARIMA</i>	: <i>Fractional Autoregressive Integrated and Moving Average</i>
<i>GARCH</i>	: <i>Generalized Autoregressive Conditional Heteroscedastic</i>
<i>GJR</i>	: <i>Glosten Jagannathan et Runkle</i>
<i>JB</i>	: <i>Jarque et Berra</i>
$ K_u $ ou K_u	: <i>Kurtosis</i>
<i>Log</i>	: <i>Logarithme</i>
<i>LogL</i>	: <i>Log-Likelihood</i>
<i>Matlab 2013a</i>	: <i>Matrix Laboratory version première semestre 2013</i>
<i>MAD</i>	: <i>Mean Absolute Deviation</i>
<i>QQ-Plot</i>	: <i>Quantile-Quantile Plot</i>
<i>RMSE</i>	: <i>Root Mean Squared Error</i>
<i>SCVS</i>	: <i>Série Corrigée des Variations Saisonnières</i>
<i>S</i>	: <i>Saisonnalité</i>
<i>SHOM</i>	: <i>Service Hydrographique et Océanographique de la Marine</i>
<i>TS</i>	: <i>Tracking Signal</i>
<i>Var (x)</i>	: <i>Variance de x</i>

LISTE DES FIGURES

Figure 1 . 1 : Référentiel de Copernic	5
Figure 1 . 2 : Lien entre le référentiel de Copernic et le référentiel Géocentrique.....	6
Figure 1 . 3 : Lien entre le référentiel Géocentrique et le référentiel terrestre	7
Figure 1 . 4 : Approche théorique de newton : cas général pour le système Terre-Lune-Soleil.....	14
Figure 1 . 5 : Représentation es forces des marées pour le modèle théorique de Newton.....	17
Figure 1 . 6 : Pseudo-période liée au mouvement de rotation de la Terre (terme courte période ou semi-diurne)	18
Figure 1 . 7 : Pseudo-période liée au mouvement de translation circulaire de la Lune (terme longue période) .	19
Figure 1 . 8 : Pseudo-période liée au mouvement de rotation la Terre (terme longue période).....	20
Figure 1 . 9 : Force des marées pour l'angle de déclinaison δ différent de zéro.....	21
Figure 1 . 10 : Coordonnées géocentriques	24
Figure 1 . 11 : Répartition zonale du terme longue période.....	25
Figure 1 . 12 : Répartition tessérale du terme diurne des marées	26
Figure 1 . 13 : Répartition sectorielle semi-diurne des marées	27
Figure 2 . 1 : Test des valeurs aberrantes à l'aide de la Boite à Moustache	36
Figure 2 . 2 : Transformation des données vectorielles en matrice	51
Figure 2 . 3 : Organigramme de la méthode Census X11-ARIMA(p,d,q)-GARCH(P,Q).....	68
Figure 3 . 1 : Localisation de Hell-Ville	71
Figure 3 . 2 : Valeurs des exposants de Hurst pour les pleines mers et les basses mers	71
Figure 3 . 3 : Données initiales composées des hauteurs des pleines et des basses mers.....	72
Figure 3 . 4 : Histogramme des données initiales composées des hauteurs des pleines et des basses mers.....	73
Figure 3 . 5 : Histogrammes des pleines et des basses mers après homogénéisation.....	74
Figure 3 . 6 : Boîtes à moustache des hauteurs des pleines et basses mers avant et après correction des points aberrants.....	76
Figure 3 . 7 : Boîtes à moustache avant et après correction des points aberrants pour les données stabilisées des hauteurs des pleines et basses mers	77
Figure 3 . 8 : Corrélogramme des hauteurs des pleines et des basses mers stabilisées	78
Figure 3 . 9 : Périodogrammes moyennés des hauteurs des pleines mers stabilisées	79
Figure 3 . 10 : Périodogrammes moyennés des hauteurs des basses mers stabilisées	79
Figure 3 . 11 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 1 à 4.....	80
Figure 3 . 12 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 5 à 8.....	81
Figure 3 . 13 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 9 à 12.....	81
Figure 3 . 14 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 13 à 16.....	82
Figure 3 . 15 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 17 à 20.....	82
Figure 3 . 16 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 21 à 24.....	83
Figure 3 . 17 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 25 à 28.....	83
Figure 3 . 18 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 29 à 32.....	84
Figure 3 . 19 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 1 à 4.....	86
Figure 3 . 20 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 5 à 8.....	87
Figure 3 . 21 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 9 à 12.....	87
Figure 3 . 22 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 13 à 16.....	88
Figure 3 . 23 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 17 à 20.....	88
Figure 3 . 24 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 21 à 24.....	89
Figure 3 . 25 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 25 à 28.....	89
Figure 3 . 26 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 29 à 32.....	90
Figure 3 . 27 : Courbes des résidus pour $S = 28$ et $S = 29$	95
Figure 3 . 28 : Test de Stabilité.....	97

Figure 3 . 29 : Méthode du profil.....	101
Figure 3 . 30 : Méthode de la bande	102
Figure 3 . 31 : Courbes des tendances des données initiales.....	103
Figure 3 . 32 : Courbes des résidus du modèle 1 basses mers pour $d = 1$ et p variant de 11 à 14.....	106
Figure 3 . 33 : Courbes des résidus du modèle 1 basses mers pour $d = 1$ et p variant de 7 à 10.....	106
Figure 3 . 34 : Courbes des résidus du modèle 2 basses mers pour $d = 1$ et p variant de 11 à 14.....	107
Figure 3 . 35 : Courbes des résidus du modèle 2 basses mers $d = 1$ et p variant de 7 à 10	107
Figure 3 . 36 : Courbes des résidus du modèle 3 basses mers pour $d = 1$ et p variant de 11 à 14.....	108
Figure 3 . 37 : Courbes des résidus du modèle 3 basses mers $d = 1$ et p variant de 7 à 10	108
Figure 3 . 38 : Courbes des résidus du modèle 1 basses mers $d = 2$ et p variant de 11 à 14.....	109
Figure 3 . 39 : Courbes des résidus du modèle 2 basses mers pour $d = 2$ et p variant de 11 à 14.....	109
Figure 3 . 40 : Courbes des résidus du modèle 3 basses mers pour $d = 2$ et p variant de 11 à 14.....	110
Figure 3 . 41 : Courbes des résidus du modèle 1 pleine mers pour $d = 1$ et p variant de 11 à 14	111
Figure 3 . 42 : Courbes des résidus du modèle 1 pleine mers pour $d = 1$ et p variant de 7 à 10	112
Figure 3 . 43 : Courbes des résidus du modèle 2 pleine mers pour $d = 1$ et p variant de 11 à 14	112
Figure 3 . 44 : Courbes des résidus du modèle 2 pleine mers pour $d = 1$ et p variant de 7 à 10	113
Figure 3 . 45 : Courbes des résidus du modèle 3 pleine mers pour $d = 1$ et p variant de 11 à 14	113
Figure 3 . 46 : Courbes des résidus du modèle 3 pleine mers pour $d = 1$ et p variant de 7 à 10	114
Figure 3 . 47 : Courbes des résidus du modèle 4 pleine mers pour $d = 1$ et p variant de 11 à 14	114
Figure 3 . 48 : Courbes des résidus du modèle 4 pleine mers pour $d = 1$ et p variant de 7 à 10	115
Figure 3 . 49 : Courbes des résidus du modèle 1 pleine mers pour $d = 2$ et p variant de 11 à 14	115
Figure 3 . 50 : Courbes des résidus du modèle 2 pleine mers pour $d = 2$ et p variant de 11 à 14	116
Figure 3 . 51 : Fonctions d'autocorrélation des quatre modèles sur les pleines mers	119
Figure 3 . 52 : Fonctions d'autocorrélation des trois modèles sur les basses mers	119
Figure 3 . 53 : Courbe des résidus et des résidus au carré.....	122
Figure 3 . 54 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 1 Basse mer	124
Figure 3 . 55 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 2 Basse mer	125
Figure 3 . 56 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 3 Basse mer	126
Figure 3 . 57 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 1 Pleine mer	127
Figure 3 . 58 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 2 Pleine mer	128
Figure 3 . 59 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 3 Pleine mer	129
Figure 3 . 60 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale : Modèle 4 Pleine mer	130
Figure 3 . 61 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 1 Pleine mer.....	132
Figure 3 . 62 : Prévision série corrigée des variations saisonnières Modèle 1 Pleine mer	132
Figure 3 . 63 : Prévision des données initiales Modèle 1	133
Figure 3 . 64 : Tracking Signal du Modèle 1 Pleine mer.....	133
Figure 3 . 65 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 2 Pleine mer.....	135
Figure 3 . 66 : Prévision série corrigée des variations saisonnières Modèle 2 Pleine mer	135
Figure 3 . 67 : Prévision des données initiales Modèle 2 Pleine mer	136
Figure 3 . 68 : Tracking Signal du Modèle 2 Pleine mer.....	136
Figure 3 . 69 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 3 Pleine mer.....	137
Figure 3 . 70 : Prévision série corrigée des variations saisonnières Modèle 3 Pleine mer	138
Figure 3 . 71 : Prévision des données initiales Modèle 3 Pleine mer	138

Figure 3 . 72 : Tracking Signal du Modèle 3 Pleine mer.....	139
Figure 3 . 73 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 4 Pleine mer.....	140
Figure 3 . 74 : Prévision série corrigée des variations saisonnières Modèle 4 Pleine mer.....	140
Figure 3 . 75 : Prévision des données initiales Modèle 4 Pleine mer.....	141
Figure 3 . 76 : Tracking Signal du Modèle 4 Pleine mer.....	141
Figure 3 . 77 : Moyenne des valeurs avant et après rupture.....	143
Figure 3 . 78 : Prévision série corrigée des variations saisonnières avant et après homogénéisation pour le Modèle 4 Pleine mer ARIMA(12,1,1)-EGARCH(1,7)	144
Figure 3 . 79 : Prévision série corrigée des variations saisonnières avant et après homogénéisation pour le Modèle 4 Pleine mer ARIMA(12,1,1)-EGARCH(3,6)	145
Figure 3 . 80 : Prévision données initiales du Modèle 4 Pleine mer après homogénéisation	145
Figure 3 . 81 : Tracking Signal du Modèle 4 Pleine mer après homogénéisation.....	146
Figure 3 . 82 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 1 Basse mer.....	147
Figure 3 . 83 : Prévision série corrigée des variations saisonnières Modèle 1 Basse mer.....	147
Figure 3 . 84 : Prévision des données initiales Modèle 1 Basse mer.....	148
Figure 3 . 85 : Tracking Signal du Modèle 1 Basse mer.....	148
Figure 3 . 86 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 2 Basse mer.....	149
Figure 3 . 87 : Prévision série corrigée des variations saisonnières Modèle 2 Basse mer.....	150
Figure 3 . 88 : Prévision des données initiales Modèle 2 Basse mer.....	150
Figure 3 . 89 : Tracking Signal du Modèle 2 Basse mer.....	151
Figure 3 . 90 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 3 Basse mer.....	152
Figure 3 . 91 : Prévision série corrigée des variations saisonnières Modèle 3 Basse mer.....	152
Figure 3 . 92 : Prévision des données initiales Modèle 3 Basse mer.....	153
Figure 3 . 93 : Tracking Signal Modèle 3 Basse mer.....	153
Figure 3 . 94 : Prévision série corrigée des variations saisonnières Modèle 3 Basse mer avant et après homogénéisation	154
Figure 3 . 95 : Prévision des données initiales Modèle 3 Basse mer après homogénéisation.....	155
Figure 3 . 96 : Tracking Signal Modèle 3 Basse mer après homogénéisation.....	155

LISTE DES TABLEAUX

Tableau 1 . 1 : Tableau de variation du terme 2	26
Tableau 1 . 2 : Tableau de variation du terme 3	27
Tableau 3 . 1 : Paramètres de l'histogramme des données initiales.....	73
Tableau 3 . 2 : Paramètres des histogrammes des pleines et des basses mers après homogénéisation	74
Tableau 3 . 3 : Paramètres de la boîte à moustache des données homogénéisées.....	75
Tableau 3 . 4 : Paramètres de la boîte à moustache des données stabilisées	76
Tableau 3 . 5 : Entropie maximale de la première pseudo-période pour q variant de 19 à 32.....	85
Tableau 3 . 6 : Entropie maximale de la deuxième pseudo-période pour q variant de 28 à 32.....	85
Tableau 3 . 7 : Entropie maximale de la première pseudo-période pour q variant de 25 à 32.....	91
Tableau 3 . 8 : Entropie maximale de la deuxième pseudo-période pour q variant de 19 à 32.....	91
Tableau 3 . 9 : Test d'égalité de la saisonnalité obtenue par les méthodes traditionnelles et par nos méthodes .	93
Tableau 3 . 10 : Indépendance, homoscedasticité et normalité	96
Tableau 3 . 11 : Analyse de la stabilité pour $S = 29$	97
Tableau 3 . 12 : Test de corrélation de rang de Kendall	98
Tableau 3 . 13 : Test de corrélation de rang de Mann-Kendall.....	99
Tableau 3 . 14 : Test de corrélation de rang de Spearman.....	99
Tableau 3 . 15 : Régression entre la moyenne et l'écart type	102
Tableau 3 . 16 : Minimisation de l'écart type des séries différenciées	104
Tableau 3 . 17 : Stratégie de Dickey Fuller Augmenté avec constante et tendance pour un nombre de retard $p =$ 14 [40]	105
Tableau 3 . 18 : Test de Ljung-Box pour les trois modèles sur les basses mers	110
Tableau 3 . 19 : Test de Ljung-Box pour les quatre modèles sur les pleines mers	116
Tableau 3 . 20 : « Ordre d'intégration final d_{final} et optimal »	117
Tableau 3 . 21 : Test de significativité des coefficients : Modèles pleines mers	117
Tableau 3 . 22 : Test de significativité des coefficients : Modèles basses mers	118
Tableau 3 . 23 : Identification de q_{max} pour les trois modèles sur les basses mers	120
Tableau 3 . 24 : Identification de q_{max} pour les quatre modèles sur les pleines mers	121
Tableau 3 . 25 : Choix du modèle ARIMA-GARCH pour le Modèle 1 Basse mer	123
Tableau 3 . 26 : Choix du modèle ARIMA-GARCH pour le Modèle 2 Basse mer	125
Tableau 3 . 27 : Choix du modèle ARIMA-GARCH pour le Modèle 3 Basse mer	126
Tableau 3 . 28 : Choix du modèle ARIMA-GARCH pour le Modèle 1 Pleine mer	127
Tableau 3 . 29 : Choix du modèle ARIMA-GARCH pour le Modèle 2 Pleine mer	128
Tableau 3 . 30 : Choix du modèle ARIMA-GARCH pour le Modèle 3 Pleine mer	129
Tableau 3 . 31 : Choix du modèle ARIMA-GARCH pour le Modèle 4 Pleine mer	130
Tableau 3 . 32 : Caractéristique de la boîte à moustache : données initiales modèle 4 pleine mer + 355 points	142

1. Introduction générale

Plusieurs théories ont été avancées pour expliquer le phénomène des marées. En 1687, Isaac Newton proposait un modèle théorique des marées [1] basé sur l'application de la Loi Fondamentale de la Dynamique. La limite de cette méthode est de ne pas prévoir des marées de hauteur plus grande que 52 centimètres [2].

Quelques années plus tard, les travaux de Laplace (1776) [3], de Kelvin (1868) [3], de Darwin (1883) [3] et de Doodson (1921) [4] ont permis d'améliorer le modèle théorique de Newton en décomposant le potentiel générateur des marées en trois espèces périodiques développables en séries de Fourier. Mais, ces nouvelles approches harmoniques sont parfois très couteuses, car l'estimation de certains paramètres nécessite des matériels et outillages spéciaux. Ceci est donc la problématique de la Thèse.

C'est pourquoi, nous avons décidé de rechercher les caractéristiques des marées en utilisant pour :

- **les codes de programmation**, des fonctions sous Matlab ou sous Excel 2007 qui sont simples et faciles à manipuler ;
- **l'analyse des données**, des outils d'analyses statistiques (statistique descriptive, statistique inférentielle) et d'analyses spectrales (Transformé de Fourier Rapide, Méthode de l'Entropie Maximale) ;
- **la prévision**, la méthode de dessaisonalisation par Census X11 combiné avec la modélisation ARIMA - GARCH.

De plus, avec les importants travaux déjà réalisés en climatologie au sein du Laboratoire Dynamique de l'Atmosphère du Climat et des Océans (DyACO), nous avons préféré orienter nos travaux de recherche dans le domaine de l'océanographie.

Ce travail constitue donc le fer de lance des travaux futurs sur l'océanographie et va expliquer les grandes lignes sur les théories des marées et les techniques pour aboutir à des modèles de prévision fiables, simples et traduisant au mieux les phénomènes étudiés.

Le cas de Hell-ville sera détaillé afin de permettre aux lecteurs de comprendre les différentes étapes à entreprendre pour arriver aux validations des différents modèles.

Les difficultés rencontrées dans cette étude concernent le choix de la méthode de dessaisonalisation et la recherche d'un modèle vérifiant le critère de parcimonie [5], c'est-à-dire, un modèle fiable permettant d'effectuer la prédiction des phénomènes en utilisant un nombre minimal de paramètres.

Les objectifs que nous nous sommes fixés dans ce travail sont donc de :

- contribuer à promouvoir la recherche dans le domaine de l'océanographie,
- reconnaître les caractéristiques des pleines et basses mers,
- doter les Chercheurs d'outils de modélisation fiables et valables en toutes circonstances,
- prévoir les niveaux de la mer à l'horizon moyen terme,
- prévoir les crues exceptionnelles.

Finalement, le présent rapport comprend :

- une introduction,
- trois chapitres correspondant au développement,
- et une conclusion, avantages et perspectives.

Dans l'introduction, nous avons défini les limites des anciens modèles de prévision des phénomènes des marées proposés par Newton, Laplace, Kelvin, Darwin et Doodson. Nous avons également parlé dans cette partie des difficultés que nous avons rencontrées durant ces trois années de thèse et des objectifs que nous nous sommes fixés pour réussir nos travaux de recherche.

Le premier chapitre renferme les généralités sur la mécanique terrestre. Elle décrit le raisonnement qui a permis à Newton et aux autres chercheurs cités précédemment d'établir le modèle théorique des marées, puis d'arriver à la forme harmonique par décomposition du potentiel générateur des marées en trois espèces périodiques développables chacun en série de Fourier.

Le deuxième chapitre montre les matériels et méthodes utilisés. Comme ce travail concerne la modélisation des marées, nous avons utilisés les logiciels Matlab et xlstat. Dans ce chapitre sera également développé les différentes méthodes. Ce chapitre est également consacrée aux parties théoriques de la modélisation des pleines mers et des basses mers par la méthode Census X11-ARIMA-GARCH.

Le troisième chapitre contient les résultats et discussions. Nous avons commencé par calculer les coefficients de Hurst pour montrer que nos données sur la hauteur des pleines mers et des basses mers proviennent d'un processus à mémoire longue. Nous avons terminé ce chapitre par la modélisation des pleines mers et des basses mers à Hell-Ville.

Nous avons terminé cette Thèse par la conclusion de nos travaux de recherche en spécifiant les avantages, les limites et les perspectives.

Chapitre 1. Généralités sur la mécanique

1.1.Introduction

Les modèles de calculs antérieurs des phénomènes des marées sont basés sur la Loi Fondamentale de la Dynamique.

Pour établir cette Loi, nous avons besoin de plusieurs référentiels de travail.

Dans ce premier chapitre, nous allons définir en premier lieu les trois référentiels dont nous avons besoin pour établir l'équation de la Loi Fondamentale de la Dynamique en mécanique terrestre.

Ensuite, nous présenterons les démarches qui ont permis à Newton, Laplace, Kelvin, Darwin et Doodson, d'établir leurs modèles de calcul des marées.

Nous terminerons ce chapitre par l'interprétation des phénomènes des marées à partir du modèle de Newton et du modèle harmonique de Laplace en spécifiant les fréquences d'apparition des marées en fonction des astres perturbateurs (lune et ou soleil) et en présentant les avantages et les limites de chaque modèle.

1.2. La mécanique

1.2.1. Les référentiels de travail

1.2.1.1. *Référentiel [6]*

Un référentiel est un système de coordonnées de l'espace à trois dimensions qui possède une origine et qui permet de repérer la position d'un point à un instant donné.

Un référentiel absolu est un référentiel que nous pouvons supposer immobile et un référentiel relatif est un référentiel en mouvement par rapport à un référentiel absolu.

Un référentiel relatif animé d'un mouvement de translation rectiligne uniforme par rapport à un référentiel absolu est « Galiléens ».

Pour pouvoir appliquer les Lois Fondamentales, nous devons travailler dans un référentiel absolu ou Galiléen.

Pour confirmer qu'un référentiel est Galiléen, il nous faut montrer que le système matériel est mécaniquement isolé c'est à dire que son centre d'inertie est animé d'un mouvement de translation rectiligne uniforme par rapport à un référentiel absolu ou Galiléen.

1.2.1.2. *Référentiel de Copernic [6]*

Le référentiel de Copernic R_C a pour origine le centre de masse G_S du système solaire. Ses axes (X_o , Y_o , Z_o) sont dirigés vers trois étoiles suffisamment éloignées pour pouvoir être considérées comme fixes.

En tenant compte du fait que les forces d'attraction newtoniennes exercées par les étoiles sur le corps du système solaire sont extrêmement faibles, nous pouvons assimiler le système solaire comme un système mécaniquement isolé.

Par conséquent, le référentiel de Copernic peut être considéré comme un référentiel Galiléen.

Nous pouvons représenter le référentiel de Copernic comme à la figure 1.1

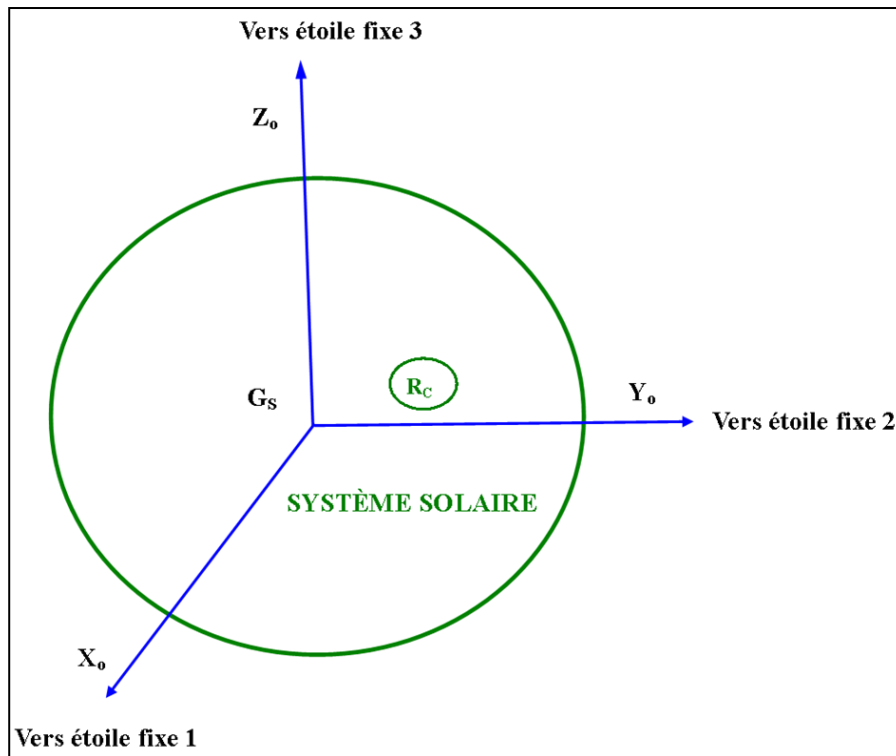


Figure 1 . 1 : Référentiel de Copernic

Le référentiel de Copernic nous permet d'établir la relation fondamentale de la dynamique pour le système Terre, Soleil et Lune.

1.2.1.3. *Référentiel géocentrique R_G [6]*

Le référentiel géocentrique R_G défini par (G_T, X_o, Y_o, Z_o) a pour origine le centre de masse G_T de la terre et ses axes ont des directions parallèles à ceux de Copernic.

En négligeant l'inclinaison de l'axe de rotation de la terre par rapport à la perpendiculaire à l'écliptique, nous pouvons admettre que l'axe Z_o est confondu avec l'axe de rotation de la Terre.

Dans ce cas, pour un vecteur rotation $\vec{\omega}$ orienté du pôle Sud vers le pôle Nord, la rotation de la terre évaluée en temps absolu peut être considérée comme uniforme.

Par conséquent, le référentiel géocentrique R_G n'est pas Galiléen son mouvement n'est pas en translation uniforme par rapport au référentiel R_c de Copernic.

Nous pouvons observer dans la figure 1.2 la liaison entre le référentiel de Copernic, et le référentiel géocentrique.

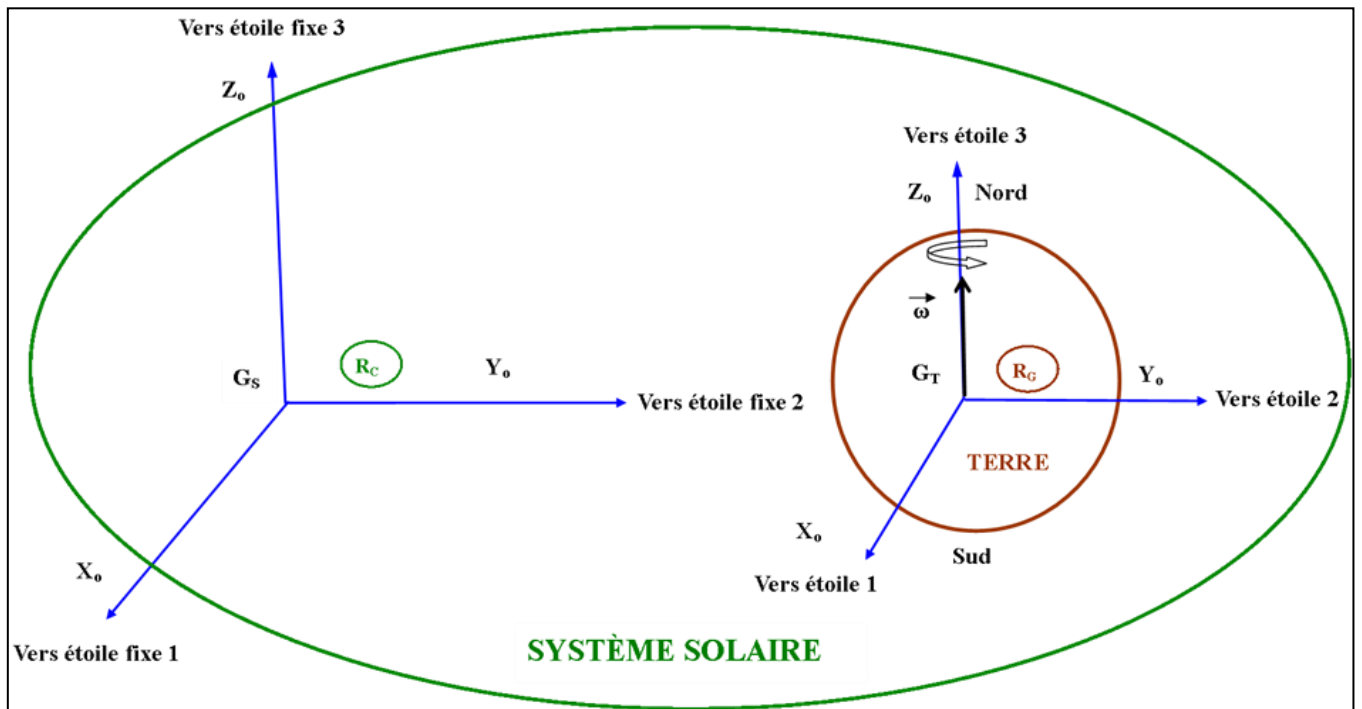


Figure 1.2 : Lien entre le référentiel de Copernic et le référentiel Géocentrique

Le référentiel géocentrique nous permet d'établir les forces agissant entre la lune, la terre et le soleil comme : les forces d'attractions, le champ de gravitation, l'accélération d'inertie d'entraînement

1.2.1.4. Référentiel terrestre R_T [6]

Le référentiel terrestre R_T défini par (G_T, X, Y, Z_0) est animé d'un mouvement de rotation uniforme ω autour de l'axe Z_0 . Pour un point P à la surface de la Terre, l'axe X sera défini dans ce référentiel terrestre R_T par la parallèle à la direction qui suit le méridien passant au point P dans le sens Nord-Sud et l'axe Y par le parallèle à la direction qui suit le parallèle au point P dans le sens Ouest-Est.

Le référentiel terrestre nous permet d'établir les forces agissant sur la particule de fluide sur la terre.

Les trois référentiels ont permis d'établir en premier lieu la loi fondamentale de la dynamique d'une particule de fluide en mouvement sur la terre. Ensuite, ils ont également permis de réduire le nombre d'inconnus à partir des conditions d'équilibre dans chaque référentiel.

La figure 1.3 montre le lien entre le Référentiel géocentrique et le référentiel terrestre.

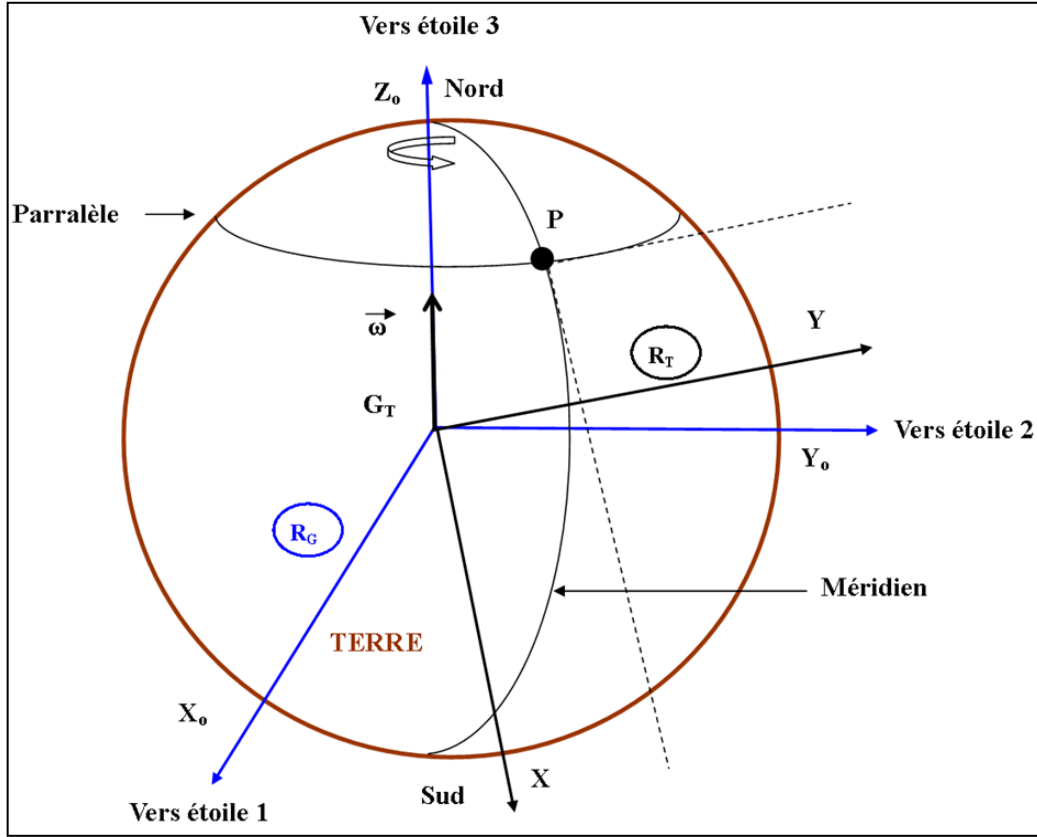


Figure 1.3 : Lien entre le référentiel Géocentrique et le référentiel terrestre

1.2.2. Loi fondamentale en mécanique terrestre [7]

En appliquant la seconde loi de Newton ou le Principe Fondamental de la Dynamique à une particule de fluide P sur terre de masse m_p dans son mouvement par rapport au référentiel de Copernic R_C , nous pouvons écrire :

$$m_p \vec{a}(P, t)_{R_C} = \sum_{Terre} \vec{F}(P, t) = \vec{F}_{G_i}(P, t) + \vec{F}_{ap}(P, t) \quad (1.1)$$

où $\vec{a}(P, t)_{R_C}$ est l'accélération absolue ou accélération du point P à l'instant t dans R_C ,
 $\vec{F}_{G_i}(P, t)$ la force d'attraction de la Terre, de la Lune, du Soleil appliqués à P à l'instant t,
 $G_{i=S,L,T}$ le centre d'inertie du Soleil, de la Lune et de la Terre,
 $\vec{F}_{ap}(P, t)$ les forces appliquées à P autres que celle de la gravitation à l'instant t,
 $\vec{F}(P, t)$ la résultante des forces appliquées à P,
et m_p la masse d'une particule de fluide P.

Dans le référentiel de Copernic R_C , la loi de composition de l'accélération appliquée au point P se traduit par la relation [8] :

$$\vec{a}(P, t)_{/R_C} = \vec{a}(P, t)_{/R_T} + \vec{a}(P, t)_e + \vec{a}(P, t)_c \quad (1.2)$$

où $\vec{a}(P, t)_{/R_T}$ est l'accélération relative ou l'accélération du point P à l'instant t dans R_T ,
 $\vec{a}(P, t)_e$, l'accélération d'entraînement ou accélération de P à l'instant t en considérant P fixe dans R_T ,
et $\vec{a}(P, t)_c$ l'accélération de Coriolis du point P à l'instant t.

En portant les valeurs de l'équation 1.2 dans l'équation 1.1, nous trouvons :

$$m_p \vec{a}(P, t)_{/R_T} = \vec{F}_{G_i}(P, t) + \vec{F}_{ap}(P, t) - m_p \vec{a}(P, t)_e + m_p \vec{a}(P, t)_c \quad (1.3)$$

C'est la loi fondamentale en mécanique terrestre pour un référentiel terrestre R_T supposé non Galiléen.

1.2.2.1. Bilan des forces appliquées [9]

A l'instant t, le point matériel P est soumis aux forces d'attraction de la terre, de la lune et du soleil défini par :

$$\vec{F}_{G_i}(P, t) = \zeta \frac{m_p * m_{G_T}}{\|PG_T\|^3} \overrightarrow{PG_T} + \zeta \frac{m_p * m_{G_L}}{\|PG_L\|^3} \overrightarrow{PG_L} + \tau \frac{m_p * m_{G_S}}{\|PG_S\|^3} \overrightarrow{PG_S} \quad (1.4)$$

Dans l'équation 1.4, la Terre, la Lune et le Soleil ont été assimilés respectivement à leurs centres d'inerties G_T , G_L et G_S

Au point P s'applique également les forces autres que la gravitation, la force d'inertie d'entraînement $\vec{F}_{ie}(P, t)$ et la force d'inertie de Coriolis $\vec{F}_{ic}(P, t)$ définies par les relations :

$$\vec{F}_{ap}(P, t) = \text{Force non gravitationnelle} \quad (1.5)$$

$$\vec{F}_{ie}(P, t) = m_p \vec{a}(P, t)_e \quad (1.6)$$

$$\vec{F}_{ic}(P, t) = m_p \vec{a}(P, t)_c = 2 m_p \vec{\omega} \wedge \vec{V}_{R_T}(P, t) \quad (1.7)$$

où \vec{V}_{R_T} représente le vecteur vitesse du fluide de masse m_p dans le référentiel terrestre.

Comme la masse de fluide m_p est en équilibre relatif dans le repère Terrestre, nous pouvons écrire :

$$\overrightarrow{V_{R_T}} = 0 \quad \text{ce qui implique} \quad \vec{a}(P, t)_{/R_T} = 0 \quad (1.8)$$

Et la relation d'équilibre s'exprime par la relation :

$$\text{Poids}(P, t) + \vec{F}_{ap}(P, t) = 0 \quad (1.9)$$

$\vec{F}_{ap}(P, t)$, ici représente l'équivalent de la force opposée au poids qui permet de maintenir en équilibre le fluide de masse m_p dans le référentiel terrestre [9] .

Dans le référentiel terrestre les forces appliquées au fluide de masse m_p sont :

- la force gravitationnelle terrestre $\vec{F}_{G_T}(P, t)$,
- la force $\vec{F}_{ap}(P, t)$ égale à la tension $\vec{T}(P, t)$ qui permet de maintenir le fluide de masse m_p en équilibre,
- la force d'inertie d'entraînement $\vec{F}_{ie}(P, t)$,
- et la force d'inertie de Coriolis $\vec{F}_{ic}(P, t)$.

A l'équilibre, en tenant compte des forces terrestres appliquées au point P du paragraphe précédent et des équations 1.7 et 1.8, l'expression de l'équation 1.3 devient :

$$0 = \vec{F}_{G_T}(P, t) + \vec{T}(P, t) - m_p \vec{a}(P, t)_e \quad (1.10)$$

Nous pouvons alors tirer des équations 1.9 et 1.10 la valeur de la tension $\vec{T}(P, t)$ ou du poids $\text{Poids}(P, t)$ qui est donnée par la relation :

$$\vec{T}(P, t) = -\text{Poids}(P, t) = m_p \vec{a}(P, t)_e - \vec{F}_{G_T}(P, t) \quad (1.11)$$

Nous constatons que la force d'inertie d'entraînement $\vec{F}_{ie}(P, t) = m_p \vec{a}(P, t)_e$ et la force d'attraction gravitationnelle de la terre sont implicitement comprises dans le poids.

En remarquant que la norme de la force d'inertie de Coriolis est en général très petite devant le poids pour les mouvements dans le référentiel terrestre [6], alors nous pouvons négliger sa valeur dans l'équation 1.3.

En effet, en prenant l'accélération de la pesanteur g égale à $9,81 \text{ m.s}^{-2}$ et l'angle de rotation de la terre égale à

$$\omega = \frac{2\pi}{86\,164} = 0,729.10^{-4} \text{ rad.s}^{-1} \quad (1.12)$$

le calcul donne :

$$\frac{F_{ic}(P,t)}{m_p g} \cong \frac{2 \omega V_{RT}(P,t)}{g} = 0,148.10^{-4} V_{RT}(P,t) \quad (1.13)$$

Nous voyons que même pour une vitesse relative $V_{RT}(P,t)$ égal à 1000 m.s^{-1} , $F_{ic}(P,t)$ ne peut dépasser $\frac{1}{70}$ du poids. Il semblerait donc qu'à l'échelle terrestre, nous pouvons négliger la force de Coriolis par rapport au poids.

Le référentiel terrestre permet donc de montrer que la force d'inertie d'entraînement est compris dans le poids et que la force d'inertie de Coriolis peut être négligé devant ce dernier. Ce qui revient à dire que le référentiel terrestre peut être considéré comme galiléen.

En négligeant la force de Coriolis par rapport au poids et en tenant compte des équations 1.3, 1.4 et 1.11, nous trouvons une autre formulation de la loi fondamentale en mécanique terrestre qui est définie par :

$$m_p \vec{a}(P,t)_{/RT} = \vec{F}'_{ap}(P,t) + \zeta \frac{m_p \times m_{GT}}{\|\vec{PG}_T\|^3} \vec{PG}_T + \zeta \frac{m_p \times m_{GL}}{\|\vec{PG}_L\|^3} \vec{PG}_L + \zeta \frac{m_p \times m_{GS}}{\|\vec{PG}_S\|^3} \vec{PG}_S - m_p \vec{a}(P,t)_e \quad (1.14)$$

où $\vec{F}'_{ap}(P,t)$ représente les forces non gravitationnelles autre que le poids.

En appliquant de nouveau la seconde loi de Newton au mouvement d'ensemble de la Terre que l'on pourrait assimiler au mouvement de son centre d'inertie G_T , nous pouvons écrire :

$$m_{inertielle} \times \vec{a}(G_T,t)_{/RC} = m_{grave} \times \vec{G}_{GL,S}(G_T,t) \quad (1.15)$$

ou $\vec{a}(G_T,t)_{/RC}$ est l'accélération du centre d'inertie G_T de la Terre dans le référentiel de Copernic,

$\vec{G}_{GL,S}(G_T,t)$ le Champ gravitationnel crée par la Lune et le Soleil au centre d'inertie G_T de la Terre,

$m_{inertielle}$ la masse inertielle de la Terre et m_{grave} sa masse gravitationnelle.

Pour l'établissement de cette Loi, nous avons considéré comme référentiel absolu, le référentiel de Copernic et comme référentiel relatif, le référentiel géocentrique.

En pratique, la masse inertielle et la masse grave peuvent être confondues [9] pour une Terre supposée parfaitement sphérique.

Nous pouvons donc réécrire l'équation 1.15 sous la forme :

$$\vec{a}(G_T, t)_{/R_C} = \vec{G}_{G_{L,S}}(G_T, t) \quad (1.16)$$

En utilisant la définition de l'accélération d'entraînement dans le référentiel géocentrique, nous pouvons assimiler l'accélération au point « P » à celle du centre d'inertie de la Terre en G_T [10].

Nous pouvons alors écrire la relation :

$$\vec{a}(P, t)_e = \vec{a}(G_T, t)_{/R_C} \quad (1.17)$$

Les équations 1.16 et 1.17 nous permettent de définir l'accélération d'entraînement au point P comme la force de gravitation créée par la lune et le soleil au centre d'inertie de la terre G_T .

Finalement l'accélération d'entraînement au point P peut donc être définie par la relation suivante :

$$\vec{a}(P, t)_e = \vec{G}_{G_{L,S}}(G_T, t) \quad (1.18)$$

1.2.2.2. Loi d'attraction Universelle [1]

Nous avons vu au paragraphe 1.2.1.2 que le système solaire est mécaniquement isolé. En supposant de plus que dans le système solaire seul le Soleil, la Lune et la Terre de masse $m_{i=S,L,T}$ sont en interaction et que leurs masses sont réparties uniformément et symétriquement par rapport à leurs centres d'inertie $G_{i=S,L,T}$, nous pouvons appliquer à la terre la troisième loi de Newton ou principe des actions réciproques [7].

Ce principe définit l'existence de deux forces d'attraction opposées \vec{F}_{G_T} et $\vec{F}_{G_{S,L}}$ de même intensité, la première liée à la Terre et la seconde liée à la Lune et au Soleil.

D'après Newton, les deux forces sont reliées par l'expression suivante :

$$\vec{F}_{G_T}(G_{S,L}, t) = -\zeta \frac{m_{G_T} * m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} \vec{G}_T \vec{G}_L(t) - \zeta \frac{m_{G_T} * m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \vec{G}_T \vec{G}_S(t) = -\vec{F}_{G_{S,L}}(G_T, t) \quad (1.19)$$

Dans l'expression de $\vec{F}_{G_T}(G_{S,L}, t)$, ζ correspond à la constante de gravitation universelle et le terme

$$-\zeta \frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} \vec{G}_T \vec{G}_L(t) - \zeta \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \vec{G}_T \vec{G}_S(t) \quad (1.20)$$

correspond au champ de gravitation $\vec{G}_{G_T}(G_{S,L}, t) = -\vec{G}_{G_{L,S}}(G_T, t)$.

1.2.2.3. Loi fondamentale de la mécanique terrestre

En utilisant les équations 1.14 et 1.18 et l'expression de $\vec{G}_{G_{L,S}}(G_T, t)$ dans l'équation 1.19, nous arrivons à l'expression suivante de la loi fondamentale en mécanique terrestre :

$$m_p \vec{a}(P, t)_{/R_T} = \vec{F}'_{ap}(P, t) + \zeta \cdot m_p \left[\frac{m_{G_T}}{\|\vec{P} \vec{G}_T\|^3} \vec{P} \vec{G}_T + \frac{m_{G_L}}{\|\vec{P} \vec{G}_L\|^3} \vec{P} \vec{G}_L + \frac{m_{G_S}}{\|\vec{P} \vec{G}_S\|^3} \vec{P} \vec{G}_S - \frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} \vec{G}_T \vec{G}_L - \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \vec{G}_T \vec{G}_S \right] \quad (1.21)$$

Le terme de l'équation 2.20 défini par [8] l'équation suivante :

$$\vec{F}_{marée} = \zeta \cdot m_p \times \left[\frac{m_{G_L}}{\|\vec{P} \vec{G}_L\|^3} \vec{P} \vec{G}_L + \frac{m_{G_S}}{\|\vec{P} \vec{G}_S\|^3} \vec{P} \vec{G}_S - \frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} \vec{G}_T \vec{G}_L - \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \vec{G}_T \vec{G}_S \right] \quad (1.22)$$

correspond à la force génératrice des marées.

En utilisant l'équation 1.19, nous pouvons définir la force génératrice des marées comme la différence entre la force d'attraction gravitationnelle exercée par la Lune et le Soleil à la particule de fluide P de la Terre de masse m_p et la force d'attraction gravitationnelle exercée par la Lune et le Soleil au centre d'inertie G_T de la Terre.

1.2.3. Conclusion

Nous avons vu, dans ce deuxième paragraphe du chapitre 1, que les modèles de calcul traditionnels des marées sont basés sur la Loi Fondamentale de la Dynamique.

Dans ce même chapitre, trois référentiels de travail ont été définis pour établir la Loi Fondamentale en mécanique terrestre : le référentiel de Copernic, le référentiel géocentrique et le référentiel terrestre.

Le référentiel galiléen de Copernic a permis en premier lieu d'appliquer la Relation Fondamentale de la Dynamique au système Terre-Lune-Soleil et d'obtenir une première équation fondamentale.

Le référentiel géocentrique va nous permettre d'obtenir les informations nécessaires pour démontrer que le référentiel terrestre est galiléen. En particulier, il permet d'identifier la masse inertielle à la masse grave et l'accélération d'inertie d'entraînement au point P au champ gravitationnelle créée par la lune et le soleil au centre d'inertie de la Terre.

En travaillant dans le référentiel Terrestre, ces égalités permettent d'inclure la force d'inertie d'entraînement dans le poids, de négliger la force d'inertie de Coriolis devant ce dernier et de prouver l'égalité entre l'accélération dans le référentiel absolue (référentiel de Copernic) et le référentiel relatif (référentiel terrestre).

Par conséquent, nous pouvons directement travailler dans le référentiel Terrestre au lieu d'étudier dans le référentiel de Copernic, car les lois restent invariantes dans les référentiels galiléens.

Finalement, les trois référentiels nous sont utiles pour ramener un problème de dynamique au niveau du système solaire à un problème de dynamique au niveau terrestre par changement de base.

1.3. Les théories des marées depuis Isaac Newton jusqu'à Doodson

1.3.1. Modèle de Newton [3] [9]

Le modèle de calcul de Newton décrit le phénomène des marées observée sur la Terre comme le résultat de l'attraction exercée par la Lune et le Soleil sur les océans. Sur la figure 1.4, nous avons représenté dans le repère $(G_T, \vec{X}, \vec{Z}_0)$ un graphique simplifié des paramètres agissant à une particule de fluide de masse m_p sur terre en supposant que le point P, G_T , G_S et G_L appartiennent au même plan méridien et que G_T , G_S et G_L appartiennent également au plan de l'équateur.

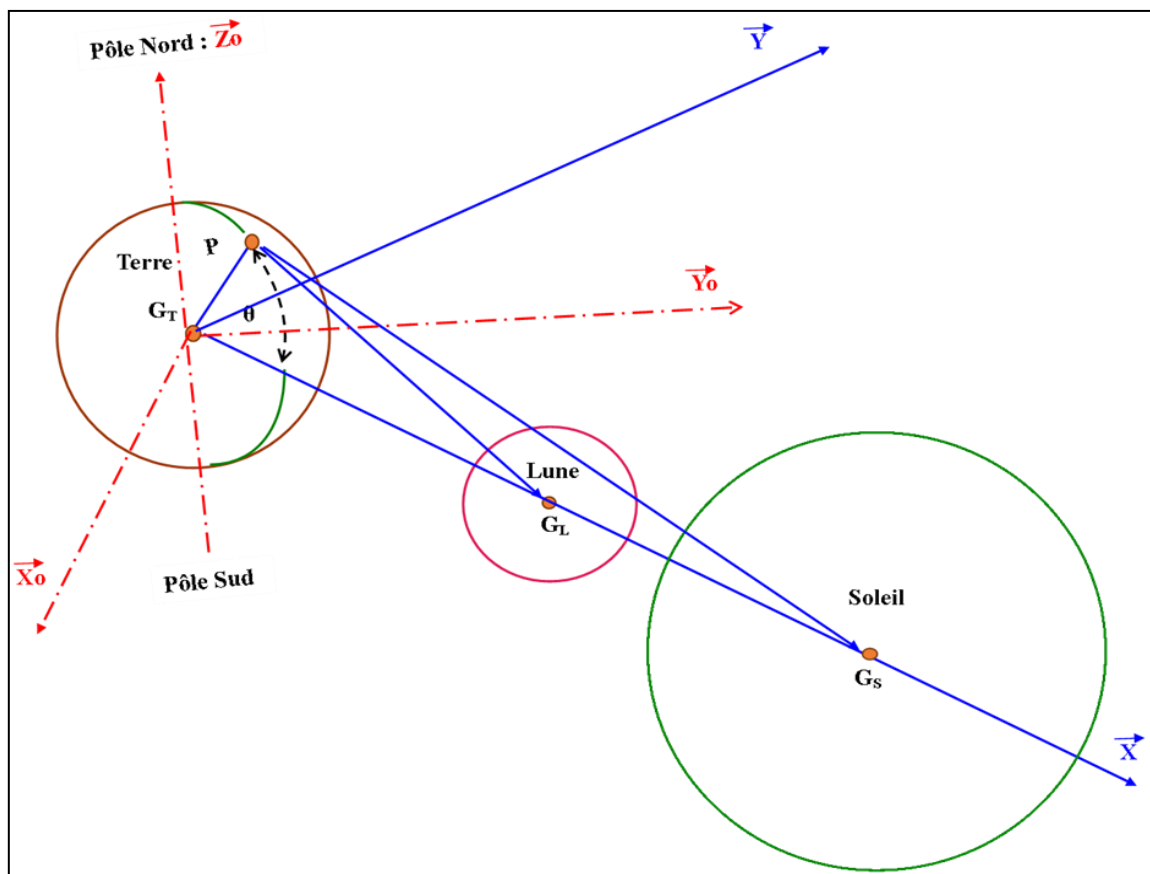


Figure 1.4 : Approche théorique de Newton : cas général pour le système Terre-Lune-Soleil

G_T , G_S , G_L , représentent respectivement les centres d'inertie de la Terre, du Soleil et de la Lune,

Θ , la distance zénithale,

et (x, z) , les coordonnées du point P dans le repère cartésien $(G_T, \vec{X}, \vec{Z}_0)$.

Dans le même repère (G_T, \vec{X}, \vec{ZO}) , nous pouvons écrire les coordonnées (x, z) du point P sous la forme :

$$x = r \cos \theta \quad (1.23)$$

$$z = r \sin \theta \quad (1.24)$$

avec

$$G_T P = r \quad (1.25)$$

$$G_T G_S = d_{TS} \quad (1.26)$$

$$G_T G_L = d_{TL} \quad (1.27)$$

où r représente le rayon moyen de la Terre,

θ , la distance zénithale,

d_{TS} , la distance entre G_T et G_S ,

et d_{TL} la distance entre G_T et G_L .

Comme ces valeurs sont connues, alors nous pouvons obtenir \vec{PG}_L et $\|\vec{PG}_L\|^3$ en fonction de r et de d_{TL} et \vec{PG}_S , $\|\vec{PG}_S\|^3$ en fonction de r et de d_{TS} .

En effet, nous pouvons écrire les relations suivantes

$$\vec{PG}_L = \vec{PG}_T + \vec{G_T G_L} = -\vec{G_T P} + \vec{G_T G_L} \quad (1.28)$$

$$\vec{PG}_S = \vec{PG}_T + \vec{G_T G_S} = -\vec{G_T P} + \vec{G_T G_S} \quad (1.29)$$

$$\vec{G_T P} = r \cos \theta \vec{X} + r \sin \theta \vec{ZO} \quad (1.30)$$

$$\vec{G_T G_L} = d_{TL} \vec{X} \quad (1.31)$$

$$\vec{G_T G_S} = d_{TS} \vec{X} \quad (1.32)$$

et déduire les valeurs

$$\vec{PG}_L = [(d_{TL} - r \cos \theta) \vec{X} + r \sin \theta \vec{ZO}] \quad (1.33)$$

$$\vec{PG}_S = [(d_{TS} - r \cos \theta) \vec{X} + r \sin \theta \vec{ZO}] \quad (1.34)$$

$$\|\vec{PG}_L\|^3 = [(d_{TL} - r \cos\theta)^2 + (r \sin\theta)^2]^{3/2} = [d_{TL}^2 - 2 d_{TL} r \cos\theta + r^2]^{3/2} = d_{TL}^3 \left[1 - \frac{2 r \cos\theta}{d_{TL}} + \left(\frac{r}{d_{TL}} \right)^2 \right]^{3/2} \quad (1.35)$$

$$\|\vec{PG}_S\|^3 = [(d_{TS} - r \cos\theta)^2 + (r \sin\theta)^2]^{3/2} = [d_{TS}^2 - 2 d_{TS} r \cos\theta + r^2]^{3/2} = d_{TS}^3 \left[1 - \frac{2 r \cos\theta}{d_{TS}} + \left(\frac{r}{d_{TS}} \right)^2 \right]^{3/2} \quad (1.36)$$

En supposant d_{TS} et d_{TL} très grand devant r , nous pouvons appliquer un développement limité à l'ordre 1 en $\left(\frac{r}{d_{TS}}\right)$ pour les transformations de \vec{PG}_L et de $\|\vec{PG}_L\|^3$ et un autre développement limité à l'ordre 1 en $\left(\frac{r}{d_{TL}}\right)$ pour les transformations de \vec{PG}_S et $\|\vec{PG}_S\|^3$ et nous trouvons finalement une autre formulation de l'équation 1.21.

En particulier, l'équation 1.21 devient :

$$\vec{F}_{\text{marée}} \Big|_{(\vec{X}, \vec{ZO})} = \zeta \cdot m_p \times \begin{bmatrix} 2 r \cos \theta \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) \\ - r \sin \theta \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) \end{bmatrix} \quad (1.37)$$

Ou encore :

$$\vec{F}_{\text{marée}} \Big|_{(\vec{X}, \vec{ZO})} = \zeta \cdot m_p \times \begin{bmatrix} 2 x \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) \\ - y \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) \end{bmatrix} \quad (1.38)$$

D'après l'équation 1.29 ou 1.30, l'interprétation du phénomène des marées à partir du modèle de Newton dépend donc du rayon de la Terre r et de l'angle Zénithal θ .

1.3.2. Interprétation du modèle de Newton

La représentation graphique de la force génératrice des marées dans le repère (G_T, \vec{X}, \vec{ZO}) sur la figure 1.5 montre une répartition symétrique par rapport au centre de la Terre G_T .

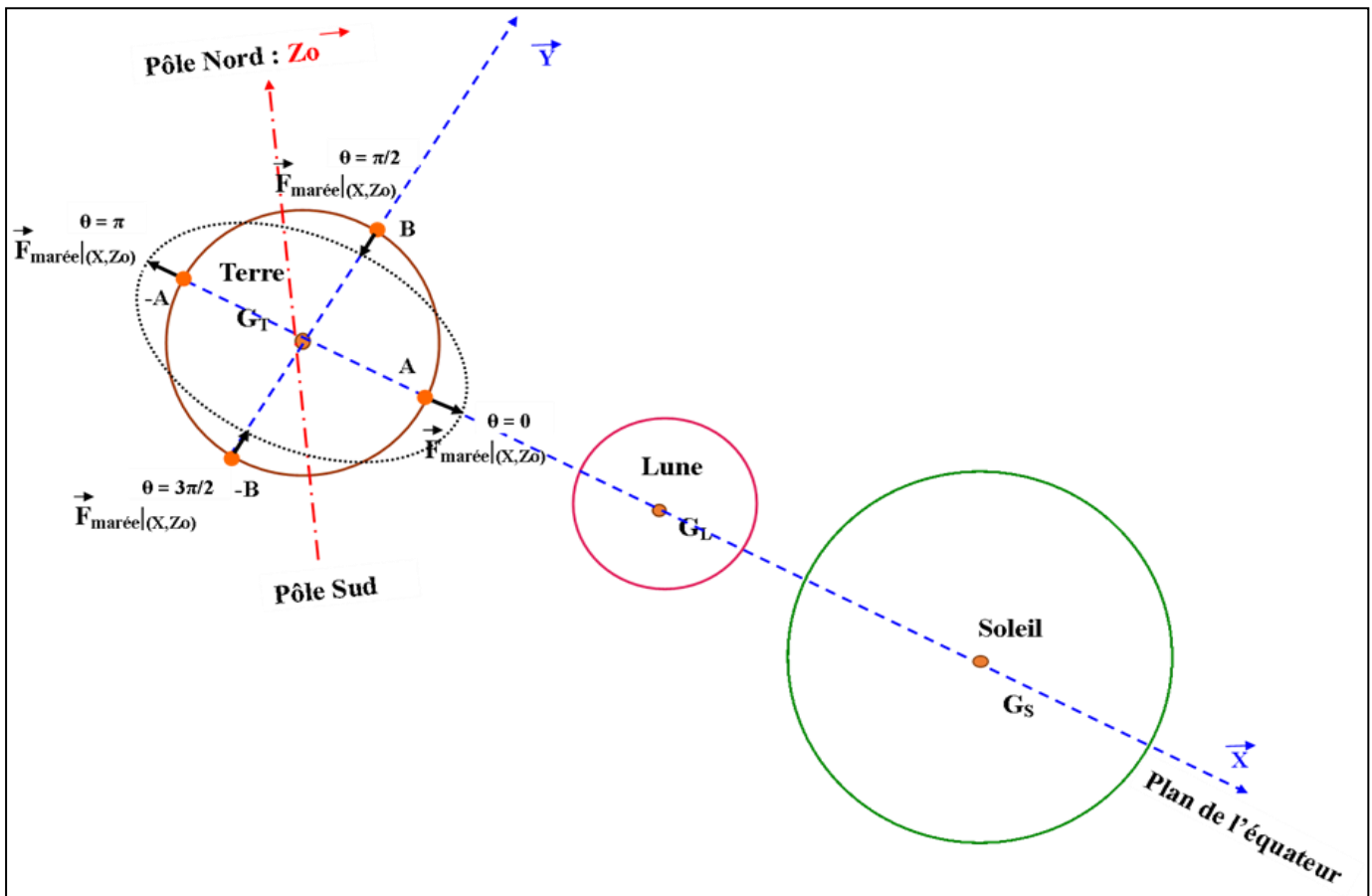


Figure 1.5 : Représentation es forces des marées pour le modèle théorique de Newton

D'après la figure 1.5, deux points situés en deux endroits opposés du globe sont donc soumis à la même force des marées. Théoriquement, l'amplitude des marées en ces deux points est la même.

Comme la position d'un point à la surface de la Terre à un instant donné dépend essentiellement du mouvement de rotation de la Terre sur elle-même, du mouvement de translation circulaire de la lune autour de la terre et du mouvement de translation circulaire de la Terre autour du Soleil, alors pour trouver les fréquences d'apparition des marées, il faut prendre en compte ces trois facteurs.

1.3.2.1. Cas ou le mouvement de rotation de la Terre sur elle-même est prépondérante

La Terre effectue un tour complet en une journée solaire et la Lune se déplace de $1/29,53$ jour au cours de la même journée. Le point P de la figure 1.6 retrouve donc sa position initiale en $30,53/29,53$ jours ($1,034$ jours ≈ 24 heures 49 minutes) et celle de son opposé en $30,53/(2 \times 29,53)$ jour ($0,517$ jour ≈ 12 heures 25 minutes) [3].

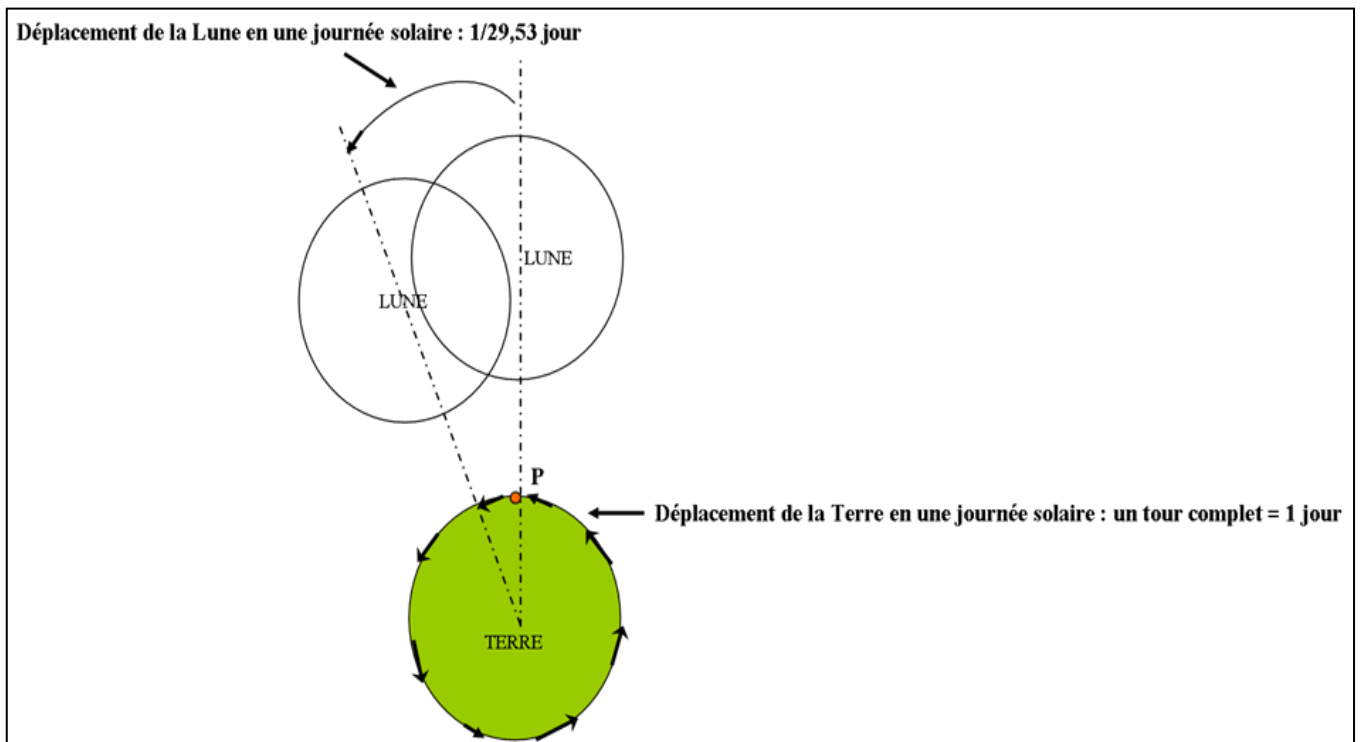


Figure 1 . 6 : Pseudo-période liée au mouvement de rotation de la Terre (terme courte période ou semi-diurne)

En tenant compte seulement du mouvement de rotation de la terre sur elle-même, le modèle de calcul des marées de Newton permet de trouver une pseudo-période de 12 heures 25 minutes pour les pleines et les basses mers : soit la moitié d'une demi-journée lunaire d'où l'appellation marées semi-diurne.

1.3.2.2. Cas ou le mouvement de translation circulaire de la Lune autour de la Terre est prépondérante

De la même façon que dans le paragraphe 1.3.2.1, en tenant compte seulement du mouvement de translation circulaire de la Lune autour de la Terre dans le modèle de Newton, nous trouvons, à la figure 1.7, pour les marées une autre pseudo-période de 14,765 jours (1/2 tour de la lune autour de la terre) qui correspond à la moitié d'un mois Lunaire [3].

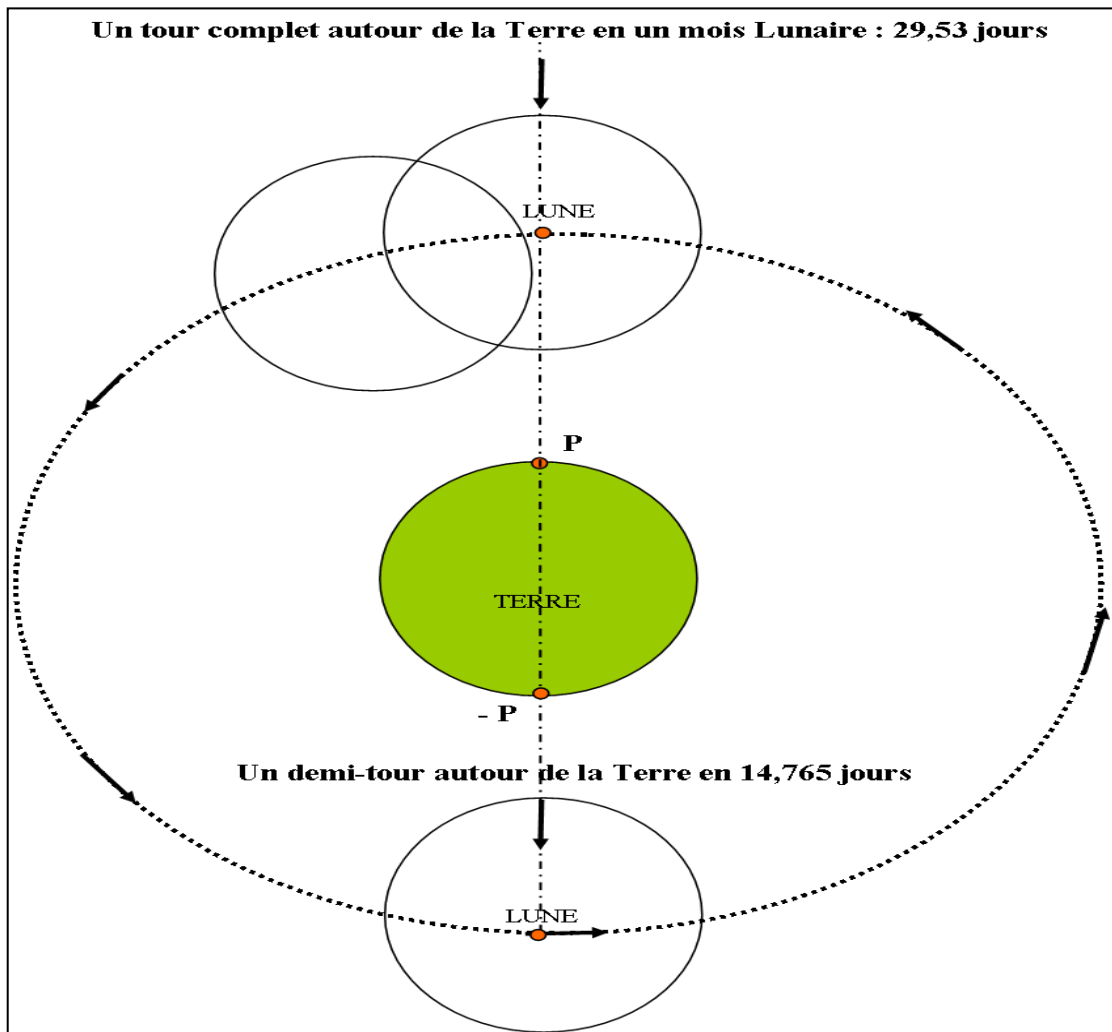


Figure 1 . 7 : Pseudo-période liée au mouvement de translation circulaire de la Lune (terme longue période)

Ces marées liés au cycle Lunaire sont appelés marées de vives ou mortes eaux

1.3.2.3. Cas ou le mouvement de translation circulaire de la Terre autour du Soleil est prépondérante

Si nous tenons compte seulement du mouvement de translation circulaire de la Terre autour du Soleil dans le modèle de Newton, nous nous apercevons à la figure 1.8 qu'il nous faut attendre 6 mois ou la moitié d'une année tropique pour observer le même phénomène des marées [3].

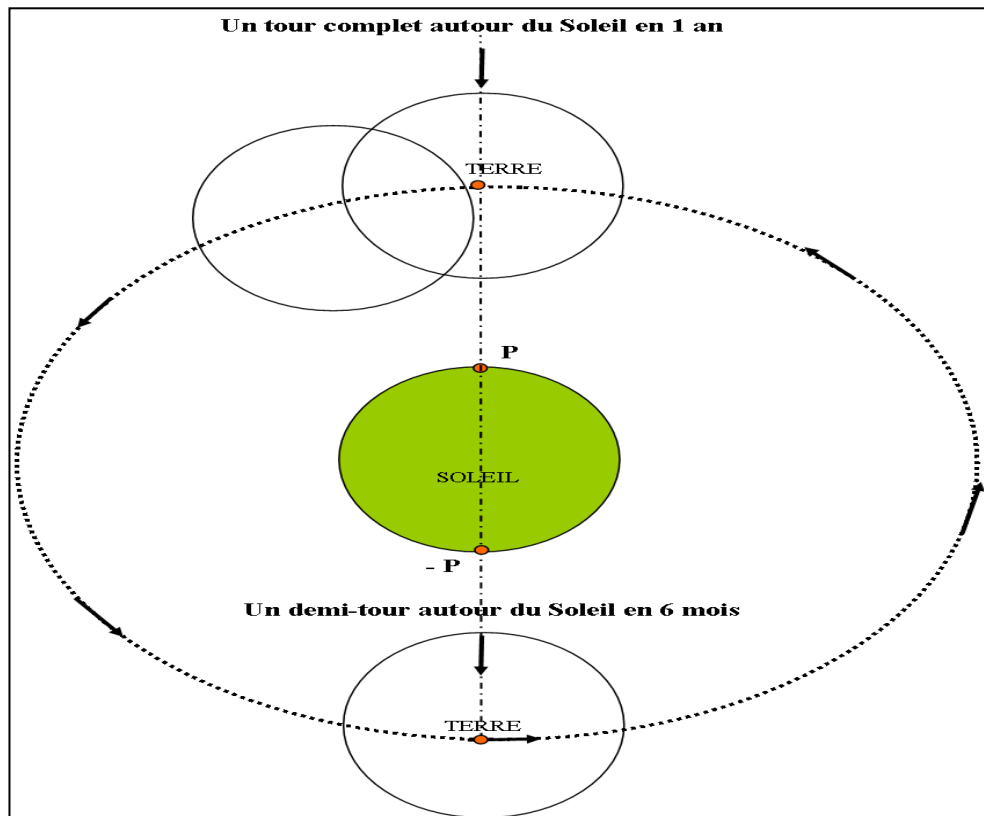


Figure 1 . 8 : Pseudo-période liée au mouvement de rotation la Terre (terme longue période)

1.3.3. Cas ou la déclinaison δ est différent de zéro [3]

Si nous introduisons l'angle de déclinaison δ dans le modèle de Newton c'est-à-dire que nous ne suggérons plus les astres perturbateurs (Lune ou Soleil) dans le même plan que l'équateur, alors nous obtenons le graphique de la figure 1.9 qui montre que les forces des marées en deux points opposés du globe terrestre ne sont plus de sens contraire et provoquent en ces points deux marées d'amplitudes différentes : on parle alors de marées à inégalité diurne car le point P ne retrouve sa position initiale qu'après un tour complet correspondant à une journée lunaire.

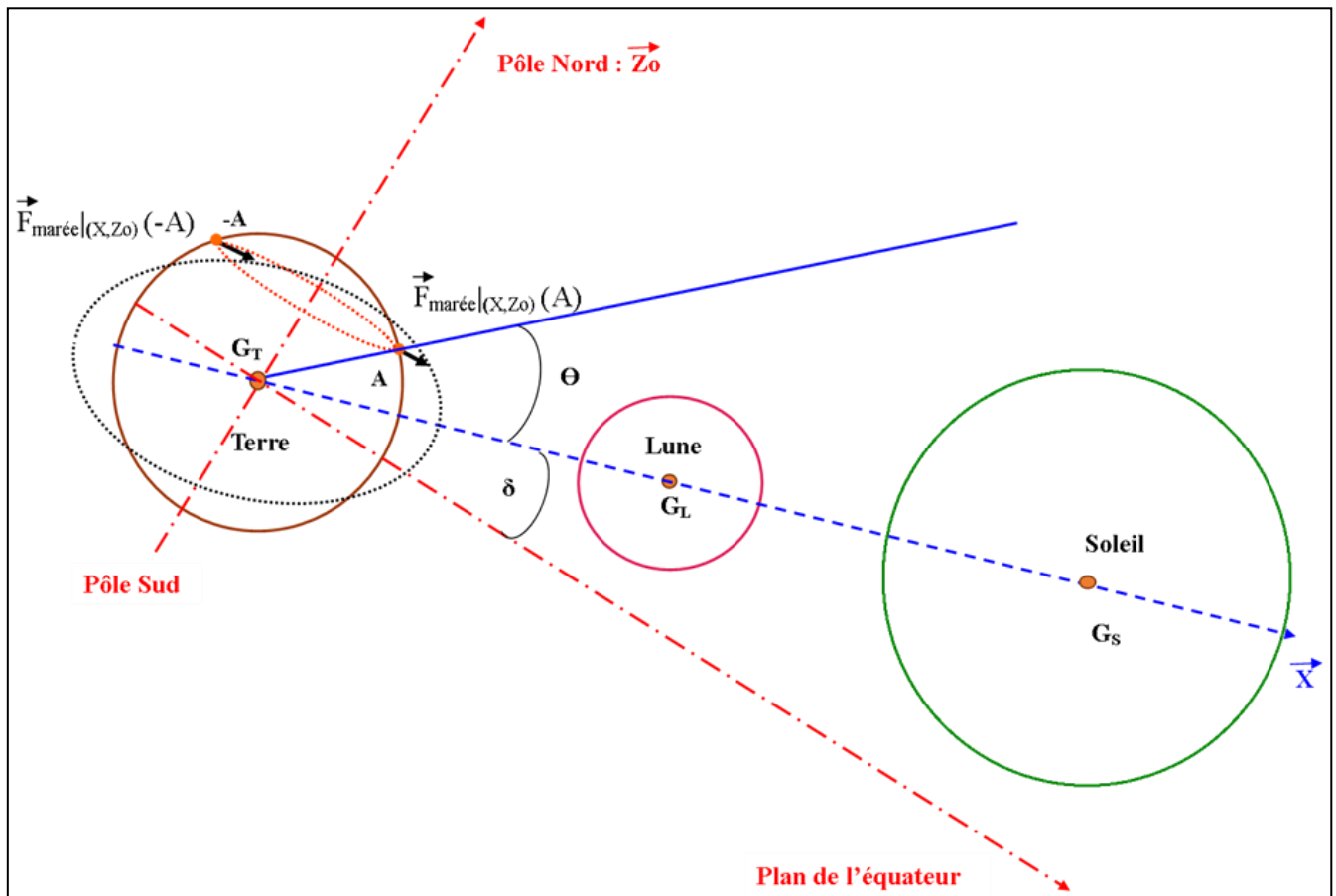


Figure 1 . 9 : Force des marées pour l'angle de déclinaison δ différent de zéro

En résumé, nous pouvons donc dire que l'approche théorique de Newton a permis de mettre en évidence :

1. la périodicité diurne du niveau de la mer,
2. l'inégalité diurne du niveau de la mer,
3. l'interdépendance entre le niveau de la mer et les périodes de rotation de la terre et de la lune,
4. l'intervention d'autres paramètres comme l'angle de déclinaison « δ » dans les phénomènes des marées.

Nous avons également trouvé que le modèle de Newton est trop restrictif car il n'arrive pas à prévoir les marées d'amplitudes supérieures à 52 centimètres [2]. De plus, il semblerait que d'autres facteurs en plus de la distance zénithale interviennent dans le phénomène des marées comme la déclinaison δ : d'où le recours aux modèles harmoniques, des versions améliorées du modèle de Newton.

1.3.4. Modèle harmonique

1.3.4.1. Potentiel générateur des marées [11]

L'équation 1.38 montre que la force génératrice des marées dépend seulement de x et de y puisque $\zeta, m_p, m_{G_L}, m_{G_S}, \|\vec{G_T G_L}\|, \|\vec{G_T G_S}\|$ peuvent être considérés comme constantes.

En remarquant que la force génératrice des marées par unité de masse possède le même ordre de grandeur que l'accélération, si cette force dérive d'un potentiel de vitesse $V_{\text{marée}}$ alors nous avons la relation

$$\frac{\vec{F}}{m_p} = \vec{\nabla} V_{\text{marée}} \quad (1.39)$$

En intégrant la force génératrice des mares définie par l'équation 1.38, nous trouvons, à une constante près, la valeur du potentiel générateur des marées $V_{\text{marée}}$.

$$V_{\text{marée}} \Big|_{(\vec{x}, \vec{z}_0)} = \zeta \times \left[\begin{aligned} & x^2 \left(\frac{m_{G_L}}{\|\vec{G_T G_L}\|^3} + \frac{m_{G_S}}{\|\vec{G_T G_S}\|^3} \right) \\ & - \frac{y^2}{2} \left(\frac{m_{G_L}}{\|\vec{G_T G_L}\|^3} + \frac{m_{G_S}}{\|\vec{G_T G_S}\|^3} \right) \end{aligned} \right] \quad (1.40)$$

Finalement avec l'hypothèse d'une terre complètement sphérique, tout point P de coordonnées (x, y) à la surface de la terre vérifie la relation :

$$x^2 + y^2 = r^2 \quad (1.41)$$

Ce qui nous permet d'écrire :

$$\begin{aligned} x^2 - \frac{y^2}{2} &= \frac{1}{2}(2x^2 - y^2) = \frac{1}{2}(2r^2 \cos^2 \theta - r^2 \sin^2 \theta) \\ \frac{1}{2}(2x^2 - y^2) &= \frac{1}{2}[2r^2 \cos^2 \theta - r^2(1 - \cos^2 \theta)] \\ \frac{1}{2}(2x^2 - y^2) &= \frac{r^2}{2}(3\cos^2 \theta - 1) \end{aligned} \quad (1.42)$$

En passant à la norme du potentiel des marées $V_{\text{marée}}$ défini par l'équation 1.40, nous trouvons

$$V_{\text{marée}} = \frac{\zeta}{2} \times \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) (3 \cos^2 \theta - 1) \times r^2 = A \times g \times (3 \cos^2 \theta - 1) \quad (1.43)$$

avec
$$A = \frac{1}{2} \times \frac{r^4}{m_T} \times \left(\frac{m_{G_L}}{\|\vec{G}_T \vec{G}_L\|^3} + \frac{m_{G_S}}{\|\vec{G}_T \vec{G}_S\|^3} \right) \quad (1.44)$$

et
$$g = \frac{\zeta \times m_{G_T}}{r^2} \quad (1.45)$$

1.3.4.2. Développement en espèce de Laplace [3]

En partant de l'équation 1.43 du potentiel générateur des marées, Laplace en utilisant la relation « $\cos \theta = \cos \delta \cos \varphi \cos H + \sin \delta \sin \varphi$ » et les propriétés du triangle sphérique aboutissait au développement du potentiel générateur des marées $V_{\text{marée}}$ en trois espèces harmoniques définis par les terme 1, 2 et 3 dans l'équation 1.46.

$$V_{\text{marée}} = \frac{3}{4} * A \times g \times i^3 \left[\underbrace{\frac{(1-3 \sin^2 \delta)(1-3 \sin^2 \varphi)}{3}}_{\text{Terme 1}} + \underbrace{\sin 2\delta \cdot \sin 2\varphi \cdot \cos H}_{\text{Terme 2}} + \underbrace{\cos^2 \delta \cdot \cos^2 \varphi \cdot \cos 2H}_{\text{Terme 3}} \right] \quad (1.46)$$

A et g correspondent aux mêmes notations que dans le paragraphe précédent et i^3 est le parallaxe commun aux trois termes. Il reste de l'ordre de l'unité, avec un maximum de 1,067 pour la lune et de 1,017 pour le soleil. En pratique, nous pouvons considérer A , g et i^3 comme constantes [3].

Les paramètres δ , φ , et H de la figure 1.10 représentent respectivement la déclinaison, la latitude et l'angle horaire relatifs au point P.

En supposant donc A , g et i^3 constantes, nous pouvons dire que l'analyse du potentiel générateur des marées $V_{\text{marée}}$ s'effectue à travers les trois termes de l'équation 1.46.

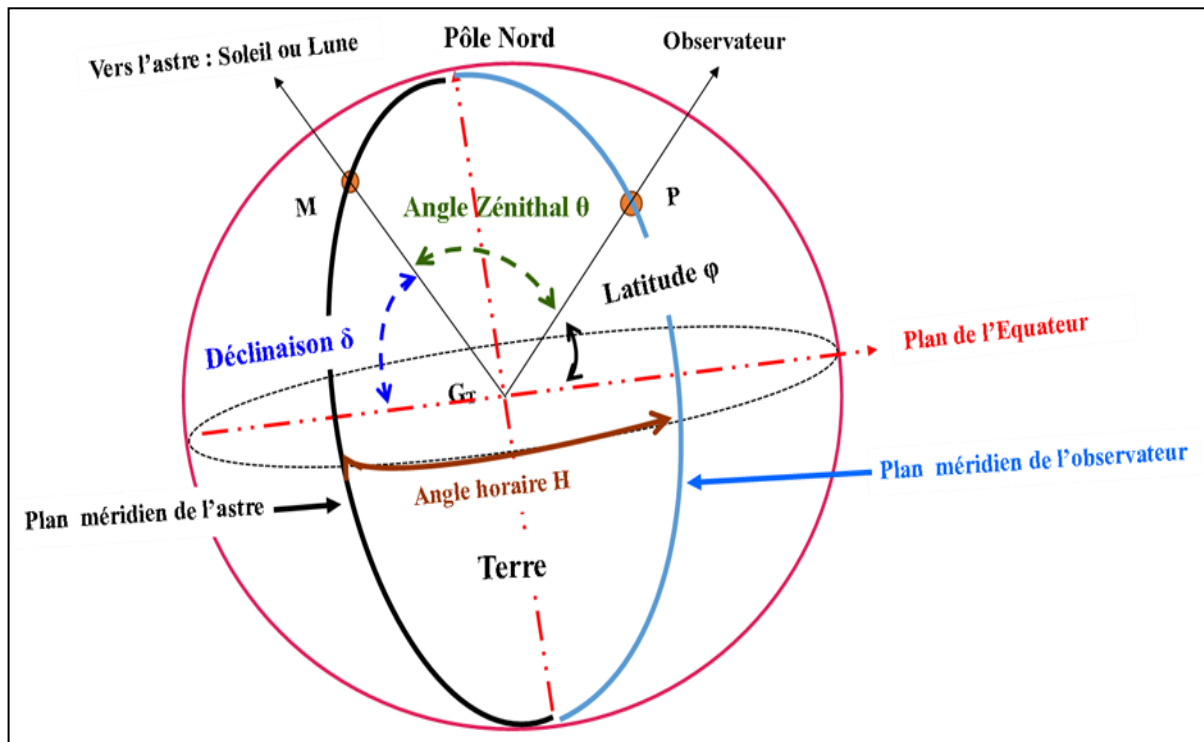


Figure 1 . 10 : Coordonnées géocentriques

Explication du terme 1 [3]

Il est recommandé de travailler sous la forme linéarisée du terme 1 en utilisant la relation :

$$(\sin \delta)^2 = \frac{1 - \cos 2\delta}{2} \quad (1.47)$$

$$\text{et de } (\sin \varphi)^2 = \frac{1 - \cos 2\varphi}{2} \quad (1.48)$$

On peut ainsi démontrer facilement que :

- L'expression $(1 - 3 \sin^2 \delta) = \frac{-1 + 3 \cos 2\delta}{2}$ est toujours positive puisque la valeur maximale trouvée dans la littérature pour la déclinaison est de $28^\circ 30'$ pour la Lune et $23^\circ 27'$ pour le Soleil [3] et que
- L'expression $(1 - 3 \sin^2 \varphi) = \frac{-1 + 3 \cos 2\varphi}{2}$ est toujours positive entre les deux nœuds correspondant à l'ensemble des points de latitudes $\varphi = \pm 35^\circ 27'$ [3]

Finalement, nous pouvons affirmer que le terme 1 restera toujours positif à l'intérieur des nodales et négatif ailleurs. La variable qui conditionne le signe du terme 1 est donc la latitude φ dont la période de variation est de 14,765 jours pour la lune et 6 mois pour le soleil. Comme ces valeurs sont assez élevées devant une journée lunaire, alors Laplace qualifiait le terme 1 de longue période.

Nous pouvons observer à la figure 1.11 la répartition zonale du terme longue période du potentiel générateur des marées sur la Terre.

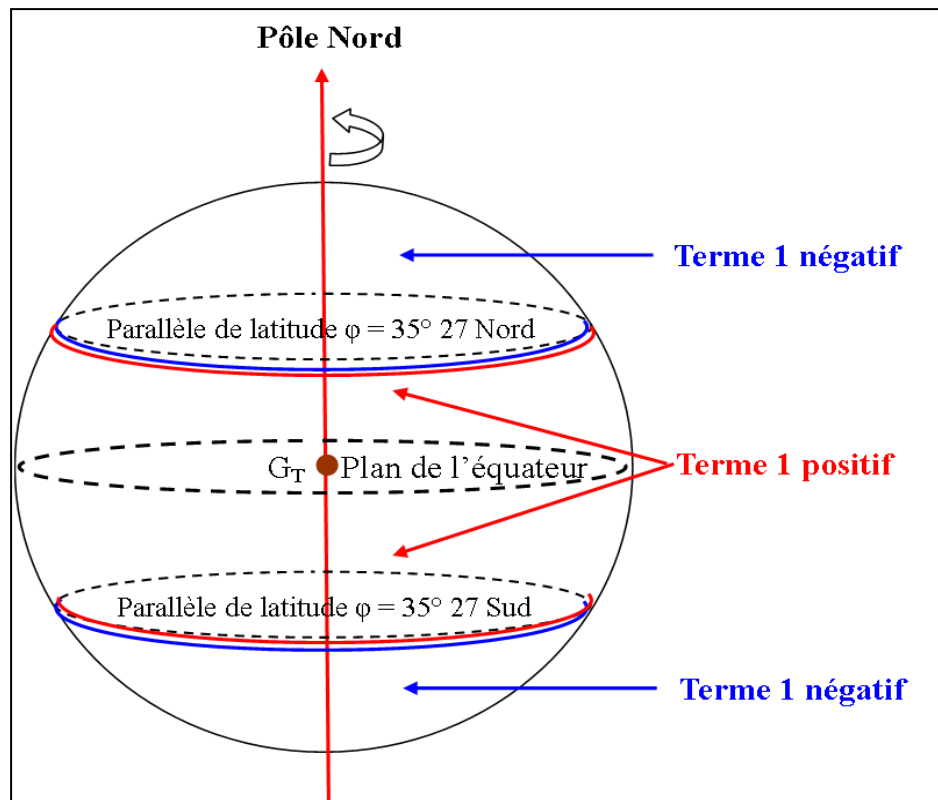


Figure 1 . 11 : Répartition zonale du terme longue période

Explication du terme 2 [3]

Il suffit d'établir le tableau de variation de la fonction « $\sin 2\delta \cdot \sin 2\varphi \cdot \cos H$ » en un point P pour comprendre la participation de ce terme 2 au phénomène des marées. En remarquant que cette fonction est paire, nous pouvons réduire le domaine par exemple à $(\varphi, \delta, H) \in [0, -\pi/2] \times [0, -\pi/6] \times [0, 2\pi]$ et compléter le résultat obtenu au tableau 1.1 par symétrie par rapport à l'origine G_T de la Terre. Sur ce même tableau sont montrés les signes des différentes fonctions.

Tableau 1 . 1 : Tableau de variation du terme 2

Fonction	$\forall \varphi \in \left[0, -\frac{\pi}{2}\right]$ $\forall H \in \left[0, \frac{\pi}{2}\right]$ $\forall \delta \in \left[0, -\frac{\pi}{6}\right]$	$\forall \varphi \in \left[0, -\frac{\pi}{2}\right]$ $\forall H \in \left[\frac{\pi}{2}, \pi\right]$ $\forall \delta \in \left[0, -\frac{\pi}{6}\right]$	$\forall \varphi \in \left[0, -\frac{\pi}{2}\right]$ $\forall H \in \left[\pi, \frac{3\pi}{2}\right]$ $\forall \delta \in \left[0, -\frac{\pi}{6}\right]$	$\forall \varphi \in \left[0, -\frac{\pi}{2}\right]$ $\forall H \in \left[\frac{3\pi}{2}, 2\pi\right]$ $\forall \delta \in \left[0, -\frac{\pi}{6}\right]$
$\sin 2\delta$	-	-	-	-
$\sin 2\varphi$	-	-	-	-
$\cos H$	+	+	-	-
$\sin 2\delta \cdot \sin 2\varphi \cdot \cos H$	+	+	-	-

En construisant la figure 1.12 sur la base du tableau 1.1, nous obtenons une répartition téssérale du terme diurne des marées. Diurne puisque le paramètre qui détermine le signe du terme 2 est l'angle horaire H qui possède une période de variation journalière.

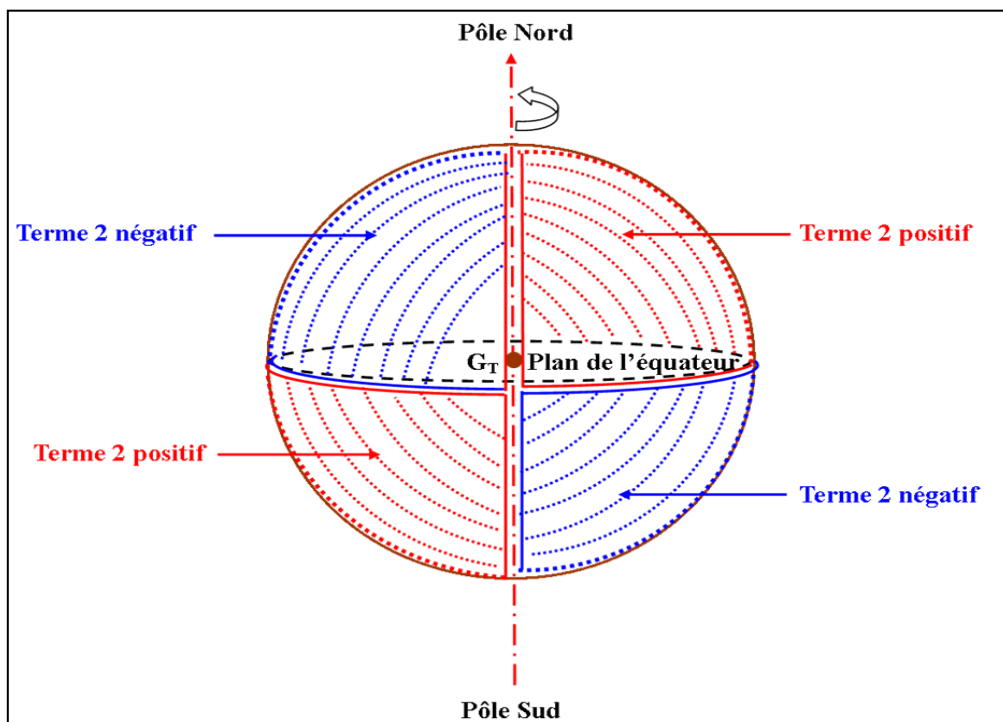


Figure 1 . 12 : Répartition téssérale du terme diurne des marées

Explication du terme 3 [3]

En remarquant que la fonction « $\cos^2 \delta \cdot \cos^2 \varphi \cdot \cos 2H$ » est de même signe que « $\cos 2H$ », nous pouvons étudier seulement la variation de cette deuxième fonction. La parité de cette dernière application nous permet de réduire alors le domaine d'étude à $[0, \pi]$, puis de compléter le résultat obtenu au tableau 1.2 par symétrie par rapport à G_T .

Tableau 1 . 2 : Tableau de variation du terme 3

Fonction	$\forall \varphi, \forall \delta H \in \left[0, \frac{\pi}{4}\right]$	$\forall \varphi, \forall \delta H \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$	$\varphi, \forall \delta H \in \left[\frac{\pi}{2}, \frac{3\pi}{4}\right]$	$\forall \varphi, \forall \delta H \in \left[\frac{3\pi}{4}, \pi\right]$
$\cos 2H$	+	-	-	+

La représentation de cette variation du tableau 1.2 sur la figure 1.13 montre une répartition sectorielle du terme semi-diurne des marées. Nous parlons de courte période car c'est l'angle horaire H possédant une période de 12 heures 25 minutes (inférieure à une journée lunaire) qui conditionne le signe du terme 3.

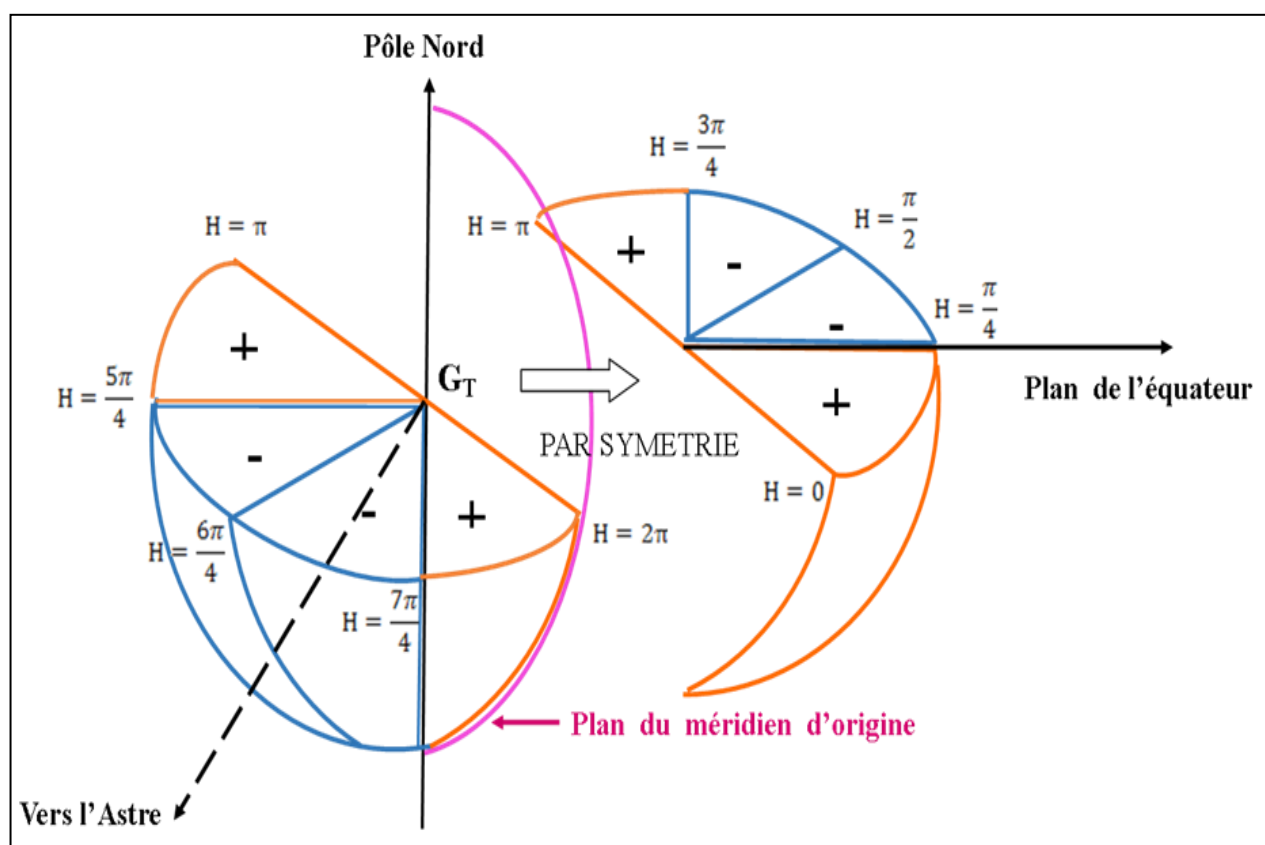


Figure 1 . 13 : Répartition sectorielle semi-diurne des marées

La **formulation dynamique** de Laplace est basée sur [3] :

- le principe de superposition des petits mouvements (linéarité de la réponse) ;
- et le principe des oscillations forcées (dissipation des conditions initiales et périodicité des réponses).

Elle est basée sur la décomposition du potentiel générateur des marées en trois espèces développable chacun en séries de Fourier.

Après quelques transformations, Laplace aboutit à l'expression suivante de la hauteur d'eau [3]

$$h_{A,l}(P) = b_{A,l}(P) \sum_n U_{P,l,n} \cos[(l\omega T + n\Omega_A)(t - T_l) - \varphi_{A,l,n} - l\lambda_l] \quad (1.49)$$

où l représente le 1, 2 et 3^{ème} terme de l'espèce de Laplace

et $t - T_l$ le temps de prélèvement dans la forme harmonique de Laplace.

Le modèle de calcul de Laplace permet également de retrouver les différentes pseudo-périodicités obtenues à partir de l'approche de Newton.

Laplace a aussi prouvé à travers la décomposition en espèce du potentiel générateur des marées que le phénomène des marées dépend au moins de trois paramètres, la déclinaison δ , la latitude ϕ et l'angle horaire H et non d'un seul paramètre comme l'a décrit Newton.

Cependant, la condition de linéarité avancée par Laplace limite son application au cas des marées à dominance semi-diurne. En effet le modèle reste satisfaisant seulement dans le cas où le terme 2 de la décomposition en espèce du potentiel générateur des marées est faible devant le terme 3 [3].

De même, le modèle de Laplace n'est plus valable pour les marées de faible profondeur car les termes non linéaires des équations de l'hydrodynamique ne peuvent plus être négligés [3].

Sur la base des travaux de Laplace qui a montré la possibilité de décomposition du potentiel générateur des marées en plusieurs composantes harmoniques, les chercheurs qui lui ont succédé, en particulier Kelvin, Darwin et Doodson ont essayé de trouver un développement harmonique plus complet du potentiel générateur des marées. En plus des trois paramètres trouvés par Laplace, ils ont ajouté un autre facteur qui intervient dans le phénomène des marées : la perturbation de l'orbite lunaire.

La réponse de ces modèles de calcul prend en général la forme suivante.

$$h(t) = \sum_i^{\infty} h_i \cos[V_{i,0} + q_i t - k_i] \quad (1.50)$$

ou h_i représente l'amplitude de la composante « i » ;

$V_{i,0}$ l'argument astronomique à l'instant $t=0$;

q_i la vitesse angulaire ;

t le temps civil du lieu ;

et k_i le déphasage ou situation de l'argument $V_{i,0}$.

1.3.5. Conclusion

Dans ce troisième paragraphe du chapitre 1, nous avons commencé à définir l'approche de Newton pour décrire les phénomènes des marées.

Basé sur la Loi Fondamentale de la Dynamique en mécanique terrestre, le modèle de calcul des marées de Newton dépend d'un seul paramètre : l'angle zénithal Θ .

En second lieu, nous avons vu le modèle dynamique de Laplace pour le calcul des marées. C'est une version améliorée du modèle de Newton qui consiste à décomposer le potentiel générateur des marées du modèle de Newton en trois espèces harmoniques dépendant : de la latitude ϕ , de la déclinaison δ et de l'angle horaire H .

En troisième lieu, l'interprétation des phénomènes des marées à partir des modèles de Newton et de Laplace nous a permis de mettre en évidence l'interaction entre les différents cycles des marées et les périodes de rotation des astres perturbateurs : le Soleil et la Lune.

En particulier, dans cette analyse, nous avons pu déterminer plusieurs cycles des marées :

- les marées de période 0,517 jour qui correspondent aux marées semi-diurne,
- les marées de période 1,034 jours qui correspondent aux marées diurne,
- les marées de période 14,765 jours qui correspondent aux marées de vive ou morte eau,
- les marées de période 181 jours qui correspondent aux marées longues périodes.

Nous avons constaté également que l'amélioration de ces modèles de calcul est reliée à l'accroissement du nombre de composantes du potentiel générateur des marées, un pour Newton en 1687, trois pour Laplace en 1776, une vingtaine pour Kelvin en 1868, de l'ordre de 40 pour Darwin en 1883 et plus de 200 pour Doodson en 1921 [3].

Actuellement, dans le cadre où la théorie de Newton se limitait à la prévision des marées d'amplitude inférieure ou égale à 52 centimètres et que l'approche dynamique de Laplace n'arrivait à prévoir que les marées semi-diurnes, les prévisionnistes utilisaient généralement des modèles harmoniques reliant, l'angle zénithal Θ , la déclinaison δ , l'angle horaire H et la perturbation de l'orbite Lunaire, dans leur décomposition du potentiel générateur des marées en sommes de composantes harmoniques.

Bien qu'universellement reconnues, les méthodes harmoniques faisant intervenir ces paramètres sont compliquées à mettre en œuvre (difficulté dans la séparation des composantes harmoniques, recours à des matériels spéciaux,...) et exigent une longue période d'observation : une année au minimum [3].

Ces limites quant à leurs applications nous amènent à rechercher des modèles mathématiques, fiables et valables en toute circonstance permettant d'obtenir des réponses en cohésion avec les phénomènes des marées.

Nous suggérons alors d'appliquer la méthode Census X11-ARIMA-GARCH comme modèle de prévision des marées.

Sa mise en œuvre nécessite l'utilisation de matériels simples et le recours à des méthodes statistiques et mathématiques de prévision que nous verrons dans la deuxième partie de cette thèse.

Chapitre 2. Matériels et méthodes

2.1.Introduction

Dans cette étude, nous avons assimilés les données issues des prélèvements à l'instant t à des séries chronologiques.

Par conséquent, les matériels et méthodes que nous allons décrire dans ce deuxième chapitre concernent les outillages et méthodes classiques de préparation, d'analyse et de traitement des séries chronologiques.

Après une brève présentation des matériels, nous verrons en premier lieu la méthode de Hurst pour tester la possibilité d'une modélisation ARIMA d'une série chronologique.

Ensuite nous allons montrer les méthodes statistiques de préparations des données.

Des méthodes de calcul et de détection de la saisonnalité et de la tendance seront également traitées dans ce deuxième chapitre ainsi que des techniques permettant de choisir le schéma de prévision le plus adapté à la distribution des données étudiées.

Avant de conclure ce deuxième chapitre, nous expliquerons les principes généraux de la méthode Census X11 et ceux de la méthode ARIMA-GARCH.

2.2.Matériels

Nous avons décidé d'utiliser des matériels informatiques simples dans cette thèse.

En particulier, nous avons utilisé un ordinateur portable 4Go de RAM interne avec un processeur pentium ® Dual-Core CPU T7400 pour l'exécution des programmes.

Les unités centrales et ordinateurs du Laboratoire DyACO ont surtout servi pour le stockage des données et des informations téléchargées sur Internet.

Pour l'exécution des codes et programmes, nous avons eu recours au logiciel Matlab 2013a, à xlstat 2014 et à l'utilitaire d'analyse d'Excel 2007.

Nous avons combiné les trois logiciels car ils ont tous les trois leurs avantages.

En effet, pour transformer par exemple une matrice 539 lignes et 29 colonnes, Excel 2007 met 1 seconde alors que Matlab a besoin d'environ une heure pour effectuer la même opération et pour transformer la même matrice en vecteur composé de 15631 variables, nous devons avoir recours obligatoirement à Matlab car cette fonction n'existe pas sous Excel.

De même, pour les tests statistiques, il vaut mieux utiliser xlstat pour gagner du temps car le logiciel fournit plusieurs réponses en une seule exécution alors que sous Matlab, il faut programmer une à une les tests pour obtenir plusieurs réponses.

Pour le test de Dickey Fuller Augmenté également, pour tenir compte du nombre de retard dans le modèle autorégressif de Dickey Fuller, il faut travailler par régression linéaire sous l'utilitaire d'analyse Excel car Matlab ne fournit que le test à l'ordre 1, c'est à dire avec zéro retard.

Ainsi, nous avons combiné dans cette thèse les trois logiciels pour optimiser le temps de calculs des programmes que nous avons exécutés pour obtenir nos résultats.

2.3.Méthodes

Nous avons utilisé dans ce travail pour analyser et modéliser nos chroniques :

- des techniques d'analyse statistique,
- des outils d'analyse spectrale,
- la technique de dessaisonalisation par Census X11,
- la méthodologie ARIMA –GARCH.

2.3.1. Méthode de Hurst [12]

Hurst a proposé le modèle suivant liant l'étendue R , l'écart type σ et l'exposant HI pour caractériser le mémoire d'un processus.

$$\frac{R(t)}{\sigma(t)} = t^{H1} \quad (2.1)$$

Dans cette relation, la pente de la régression linéaire de $\log\left(\frac{R}{\sigma}\right)$ sur $\log \tau$ constitue une bonne approximation de HI avec :

$$\bar{x} = \frac{1}{\tau} \sum_{i=1}^{\tau} x_i \quad (2.2)$$

$$y_j = x_j - \bar{x} \quad j \text{ variant de } 1 \text{ à } \tau \quad (2.3)$$

$$z_j = \sum_{i=1}^j y_i \quad j \text{ variant de } 1 \text{ à } \tau \quad (2.4)$$

$$R(\tau) = \max_{1 \leq j \leq \tau} (z_j) - \min_{1 \leq j \leq \tau} (z_j) \quad (2.5)$$

$$\sigma(\tau) = \sqrt{\frac{1}{\tau} \sum_{i=1}^{\tau} (x_i - \bar{x})^2} \quad (2.6)$$

ou τ représente la taille de l'échantillon prise égale à N , $N/2$, $N/4$...
et N la taille de l'échantillon total.

Dans cette étude, pour trouver HI , nous avons conçus un petit programme sous Matlab.

Selon Hurst, le mémoire d'un processus est fonction de la valeur de l'exposant HI .

Ainsi :

- si la valeur de HI est comprise entre 0,5 et 1 au sens strict, on dit que le processus est persistant ou à mémoire longue ;
- si la valeur de HI est comprise entre 0 et 0,5 au sens strict, on dit que le processus est anti-persistant ou à mémoire courte;
- si $HI = 0,5$ le processus n'a pas de mémoire.

La valeur du paramètre HI donne une indication sur l'interdépendance des données. Par conséquent, elle nous informe sur la possibilité de modélisation du processus par la méthode ARIMA.

Par exemple, on modélise généralement les processus à mémoire longue par des modèles FARIMA (ARMA intégré avec un ordre fractionnaire) [13], mais cela n'exclut pas la possibilité de modélisation par des modèles ARIMA

2.3.2. Méthode statistique de préparation et d'analyse des données

2.3.2.1. Courbe représentative des données initiales

Elle permet de visualiser l'allure générale de la courbe et de détecter s'il existe une tendance, une saisonnalité ou une rupture au niveau des observations [14]. Pour faire apparaître cette courbe sous Matlab 2013a, nous avons utilisé la fonction « Plot (x) ».

2.3.2.2. Préparation des données

Avant tout analyse, il est important de s'assurer de la fiabilité des données, d'une part, il faut confirmer l'homogénéité des données et d'autre part, il faut détecter les points atypiques, les corriger et si nécessaire les supprimer.

a) Vérification de l'homogénéité des données

On peut vérifier l'homogénéité des données en construisant l'histogramme, si les données sont homogènes, nous observerons un seul mode et si la chronique est hétérogène, nous obtenons un histogramme multimodal [15].

Pour construire l'histogramme, nous avons utilisé la règle de **STURGES** [16].

Elle permet en premier lieu de calculer le nombre de classes k_2 en utilisant la relation

$$k_2 = 1 + 3,322 \log_{10}(N) \quad (2.7)$$

puis de calculer l'amplitude des classes en appliquant

$$a_{k_2} = \frac{x_{max} - x_{min}}{k_2} = \text{constante} \quad (2.8)$$

et de déduire l'intervalle des classes en se servant de la formule

$$I_{k_2} = [d_{k_2}, d_{k_2+1}[\quad (2.9)$$

dans laquelle

$$d_{k_2+1} = d_{k_2} + a_{k_2} \text{ avec } d_1 = x_{min} \text{ et } d_{k_2+1} = x_{max} \text{ pour } k_2 \text{ variant de } 1 \text{ à } k_2 \quad (2.10)$$

ou k_2 représente le nombre de classes

et x_{max} et x_{min} , les valeurs maximales et minimale de la chronique

Pour afficher l'histogramme sous Matlab, il suffit d'exécuter la fonction « hist (x , k_2) »

b) Détection, correction et élimination des points aberrants

Pour détecter les points aberrants, nous avons utilisé la « **Méthode de la boîte à moustache** ou **Box-plot** » [17].

Le principe du Box-plot considère comme atypiques les valeurs situées en dehors de l'intervalle $[a, b]$ défini par

$$[a, b] = [Q1 - 1,5 \times (Q3 - Q1), Q3 + 1,5 \times (Q3 - Q1)] \quad (2.11)$$

$Q1$ et $Q3$ représentent respectivement le premier et le troisième quartile de la série étudiée.

Graphiquement, le test de détection de point aberrant par la méthode de la boîte à moustache se traduit comme à la figure 2.1.

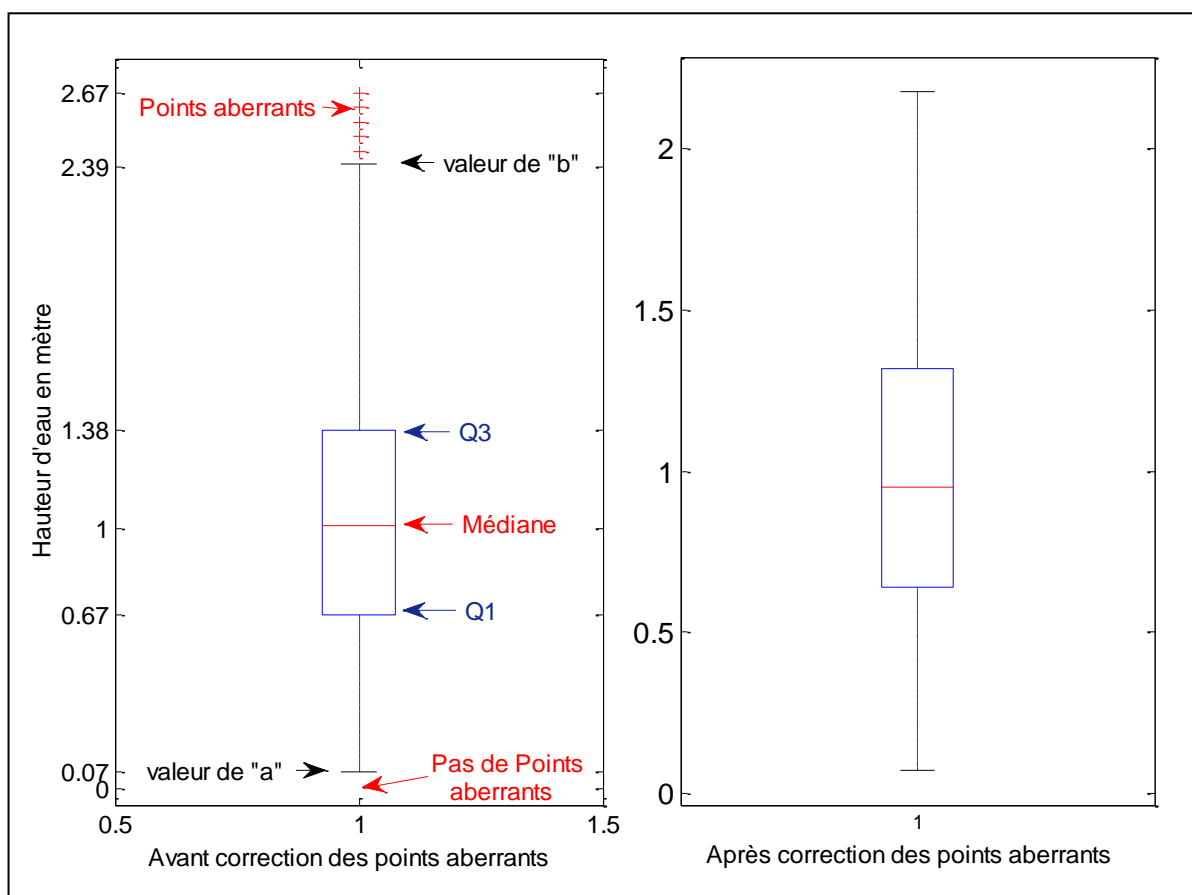


Figure 2.1 : Test des valeurs aberrantes à l'aide de la Boîte à Moustache

La boîte à moustache de la figure 2.1 permet d'obtenir graphiquement les valeurs des paramètres « a, b, Q1 et Q3 de l'équation 2.11 et de repérer les domaines dans lesquels se situent les points aberrants, c'est-à-dire inférieur à « a » ou supérieur à « b » (partie en rouge sur la figure 2.1).

Une fois détecté, les points aberrants peuvent être corrigés ou éliminés de la série temporelle.

Pour corriger les points aberrants, nous avons utilisé des moyennes mobiles pondérées :

- deux points en avant pour les points situés à l'extrémité de la chronique (t-2 et t-1) ;
- deux points en arrière (t+1 et t+2) pour les points situés au début de la chronique ;
- un point en avant et un point en arrière pour les points à l'intérieur de la chronique (t-1 et t+1).

Suivant l'emplacement de la variable, les moyennes mobiles pondérées utilisées sont définies par

$$X(t) = \sum_{i=t-1}^{i=t+1} c(i) \times x(i) \quad (2.12)$$

pour les valeurs se trouvant à l'intérieur de la chronique

$$X(t) = \sum_{i=t}^{i=t+2} c(i) \times x(i) \quad (2.13)$$

pour les valeurs se trouvant au début de la chronique et

$$X(t) = \sum_{i=t-2}^{i=t} c(i) \times x(i) \quad (2.14)$$

pour les valeurs se trouvant à la fin de la chronique,

où $x(i)$ représente la hauteur d'eau à l'instant « i » ;

$c(i)$, le coefficient de pondération de $x(i)$ prise égal à $\frac{|b-x(i)|}{b-a}$ si $x(i)$ est un point aberrant et prise égal à 1 dans le cas contraire,

et a et b représentent respectivement les bornes inférieures et supérieures de l'intervalle de Box-plot.

On peut **éliminer** les points aberrants se trouvant à l'extrémité d'une chronique si l'on ne disposait pas d'informations nécessaires pour les corriger.

2.3.3. Méthode spectrale de détection et de calcul de la périodicité S

Plusieurs méthodes existent pour déterminer la valeur de la saisonnalité S. L'approche retenue consiste à construire pour chaque chronique, la courbe de la densité spectrale en fonction de la période (ou fréquence).

Notre problème se ramène donc :

- à l'estimation de la Densité Spectrale de Puissance (DSP),
- et au calcul de la fréquence ou période correspondante.

2.3.3.1. Estimation de la DSP

La densité spectrale contient la même information que la fonction d'autocovariance dans le domaine fréquentiel. En particulier, elle permet d'identifier la tendance, les saisonnalité, les cycles

Pour trouver la DSP, nous avons décidé d'appliquer avec une fonction d'apodisation rectangulaire :

- la méthode du corrélogramme ou méthode de BLACKMAN-TUCKEY ;
- et la méthode du périodogramme qui se subdivise en deux catégories :
 - ⇒ la méthode du périodogramme moyenné ;
 - ⇒ la méthode du périodogramme par maximisation de l'entropie.

a) *Méthode indirecte, méthode du corrélogramme ou méthode de BLACKMAN-TUCKEY [18]*

Cette méthode consiste à estimer la densité spectrale comme la Transformée de Fourier de la fonction de corrélation R_{xx} :

$$DSP(f) = \text{fft} [\hat{R}_{xx}(k_3)] \quad (2.15)$$

$$\hat{R}_{xx}(k_3) = \frac{1}{N - k_3} \sum_{t=0}^{N-k_3-1} x_{t+k_3} x_t^* \quad (2.16)$$

$$\text{fft}(k_3) = T_e \sum_{t=1}^N x_t e^{-2i\pi(t-1)\times\frac{(k_3-1)}{L}} \quad [19] [20] \quad (2.17)$$

où on a fft la Transformé de Fourier Rapide ;

DSP la Densité Spectrale de Puissance ;

N le nombre de points temporels ;

t la variable temporelle variant de 1 à N ;

x_t les données périodique de période t ;

L le nombre de points fréquentiels ;

k_3 la variable fréquentielle variant de 1 à L ;

\hat{R}_{xx} l'estimateur de la fonction d'autocorrélation R_{xx} ;

T_e la période d'échantillonnage ;

et F_e la fréquence d'échantillonnage.

R_{xx} est obtenu sous Matlab en utilisant la forme non biaisée de « xcorr » et la fonction « fft ».

b) Méthode du périodogramme moyenné ou méthode directe [20]

Dans cette méthode, la densité spectrale devrait être estimée à l'aide de la relation :

$$DSP(f) = \frac{1}{M} \sum_{i=1}^M |\text{fft}(k_3)|^2 \quad (2.18)$$

où M est le nombre de décomposition en bloc de l'échantillon initial égal à L/Te

c) Périodogramme par la Méthode de l'Entropie Maximale

C'est une variante de la méthode du corrélogramme. Elle consiste à maximiser l'entropie ou la probabilité d'obtenir le maximum d'information sur les données.

Pour le décrire nous nous sommes limités à la représentation de BURG [21]. Il considérait tout d'abord un processus aléatoire centré et stationnaire fini $X(n)$, dont la fonction d'autocorrélation est définie par :

$$R_{xx}(k_3) = E[X(n)X(n + k_3)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} DSP(f) e^{jmf} df \text{ avec } k_3 \text{ varie de } 0 \text{ à } K \quad (2.19)$$

La densité spectrale de puissance $DSP(f)$ peut donc s'écrire sous la forme

$$DSP(f) = \sum_{m=-\infty}^{\infty} R_{xx}(k_3) e^{-jmf} \quad (2.20)$$

Ensuite, il attribuait une loi de probabilité $p(x)$ au vecteur

$$\vec{X} = \begin{bmatrix} X(0) \\ \vdots \\ X(N-1) \end{bmatrix} \quad (2.21)$$

Puis, en appliquant le principe du maximum d'entropie [22] et en tenant compte du fait que les contraintes de l'équation 2.19 sont quadratiques en X , on peut facilement établir que X suit une loi gaussienne.

En particulier, pour un processus centré, stationnaire et gaussien, lorsque le nombre d'échantillons N est assez élevé, nous pouvons démontrer l'expression de l'entropie suivante :

$$H = \int_{-\pi}^{\pi} \ln DSP(f) df \quad (2.22)$$

Finalement, en maximisant H sous les contraintes de l'équation 2.19, nous arrivons facilement à la solution définie par la relation

$$DSP(f) = \frac{1}{\left| \sum_{k_3=-K}^K \lambda_{k_3} e^{jk_3 f} \right|^2} \quad (2.23)$$

dans laquelle λ_m représente les multiplicateurs de Lagrange [23] associés aux contraintes de l'équation 2.19 et qui sont équivalents aux coefficients φ_i d'une modélisation autorégressive du processus $X(n)$ défini par

$$X(n) = \sum_{i=1}^p \varphi_i X(n-i) \quad (2.24)$$

Cette représentation paramétrique autorégressive de l'équation 2.24 permet de réécrire l'équation 2.23 sous la forme [24] suivante

$$DSP(f) = \frac{\sigma_p^2}{\left| 1 + \sum_{i=1}^p a_i e^{jif} \right|^2} \quad (2.25)$$

où σ_q^2 représente la variance du bruit blanc issu du modèle autoregressif d'ordre q et a_i les coefficients du modèle autorégressif d'ordre q .

La méthode de l'entropie maximale consiste en premier lieu à tracer au seuil de significativité α égal à 5% la courbe (DSP, S) , en faisant varier l'ordre q du modèle autorégressif de BURG.

Ensuite, pour chaque valeur de q , nous cherchons la période la plus significative.

Enfin en attribuant une probabilité $p_{k_3 \leq q}$, à l'apparition de la période S_{k_3} correspondant à une valeur donnée de q , nous trouvons la bonne saisonnalité S en choisissant la période possédant la plus grande entropie [22] définie par

$$H_{k_3} = - \sum_{i=1}^N p_{k_3 \leq q} \ln p_{k_3 \leq q} \quad (2.26)$$

En pratique, pour chercher la valeur de l'entropie maximale, il faut fixer un ordre q assez élevé de manière à obtenir 25 à 30 valeurs de q rendant significative la saisonnalité S .

Pour obtenir l'ordre minimal du modèle autorégressif de BURG, il faut diminuer progressivement q jusqu'à ce que le H_{k_3} trouvé précédemment ne soit plus maximal. Le q correspondant à cette limite est l'ordre maximal du modèle autorégressif de BURG.

Nous pouvons nous servir de cette valeur de q pour obtenir le q_{initial} du modèle ARIMA(p+1,d,q) que nous traiterons au paragraphe 2.4.2. 2.

2.3.3.2. Formulation de la fréquence f et de la période S [20]

Pour la formulation de la fréquence, il suffit de remarquer l'équivalence entre les deux termes :

$$\Rightarrow -2i\pi \times (t-1) \times \frac{(k_3-1)}{L} = -2i\pi \times (t-1)T_e \times \frac{(k_3-1)}{L}F_e \text{ de la}$$

Transformé de Fourier Discrète

$$\Rightarrow \text{et } -2i\pi.f.n \text{ de la Transformé de Fourier Continue.}$$

L'identification entre les deux termes nous permet d'établir les relations suivantes :

$$n = (t-1)T_e \quad (2.27)$$

$$\text{et } f = \frac{(k_3-1)}{L}F_e \quad (2.28)$$

Finalement, en tenant compte du fait que k variant de 1 à L , la relation reliant la fréquence f à la fréquence d'échantillonnage et au nombre de points est donc donnée par :

$$f = \frac{F_e}{L} \times [0, L-1] \quad (2.29)$$

Si on utilise la période et non la fréquence, elle se déduit en utilisant la relation :

$$S = \frac{1}{f} \quad (2.30)$$

2.3.3.3. Réduction du domaine d'étude

La parité de la fonction $DSP(f)$ (les x_t sont des réels positifs), le théorème de Nyquist-Shannon qui stipule que la fréquence F_e d'échantillonnage doit être au moins égale au double de la fréquence maximale du signal à reproduire [25] et le mode d'itération sous Matlab (itérations à partir de 1 et non de 0), nous ont conduit à adopter le domaine fréquentiel

$$f = \frac{F_{Nyquist}}{\frac{L}{2}} \times \left[1, \frac{L}{2}\right] = \frac{F_e}{\frac{L}{2}} \times \left[1, \frac{L}{2}\right] \quad (2.31)$$

2.3.3.4. Saisonnalité S en terme journalier

Notons que la saisonnalité S trouvée par les trois méthodes précédentes correspond à la saisonnalité liée aux rangs des temps de prélèvement.

Pour convertir cette saisonnalité en terme journalier, nous devons calculer la moyenne arithmétique en considérant comme variable la différence en jour entre deux temps de prélèvement successif sur les pleines mers ou sur les basses mers.

Pour ramener 1 Lag, sur les figures, en terme journalier, il suffit d'appliquer la relation d'équivalence

$$1 \text{ Lag} = \frac{\sum_{i=1}^N D_i}{N} \text{ (jours)} \quad (2.32)$$

où D_i est la différence en jour entre deux temps de prélèvement consécutifs

Après conversion, nous pourrions comparer la valeur des saisonnalités journalière S obtenus par nos méthodes de calcul avec celles fournies par les modèles de Newton ou de Laplace en utilisant le « **test t de Student** » (test d'égalité de la moyenne d'une série à une valeur donnée) défini au point c) du paragraphe 2.4.2.4.

2.3.3.5. Choix de la périodicité S

Comme la méthode du corrélogramme et celle du périodogramme donnent en général des valeurs non entières, il nous faut chercher leurs meilleures estimations sous la forme entière $E(S)$ ou $E(S+1)$ pour pouvoir mettre en œuvre la méthode Census X-11.

Pour reconnaître la meilleure estimation de la saisonnalité S , nous pouvons effectuer:

- le test sur les résidus ;
- l'analyse de la variance à un facteur.

a) Test sur les résidus [26]

Cette méthode consiste à choisir l'estimation de la saisonnalité S créant moins de fluctuation résiduelle (propriété des moyennes mobiles).

b) Analyse de la Variance à un facteur ou test de Fisher

Cette technique est utilisée pour des données indépendantes, homoscedastes et possédant une distribution normale ou approximativement normale [27][28].

Comme au paragraphe précédent, on choisira toujours la période qui offre une meilleure transformation des données, c'est-à-dire, celle qui assure une meilleure stabilité saisonnière.

L'indépendance des données découle de l'indépendance des moments de prélèvement des échantillons.

Pour tester **la normalité**, nous avons utilisé le test de Jarque et Bera et deux autres tests basés sur le coefficient d'aplatissement et le coefficient d'asymétrie [29]. La distribution de la variable $x(t)$ est normale

$$\text{si} \quad JB = \frac{N}{6} S_t^2 + \frac{N}{24} K_t^2 \leq 5,99 \quad (2.33)$$

$$\text{ou si} \quad |S_k| = \sqrt{\frac{N}{6}} \cdot S_t \leq 1,96 \quad (2.34)$$

$$\text{ou si} \quad |K_u| = \sqrt{\frac{N}{24}} \cdot K_t \leq 1,96 \quad (2.35)$$

$$S_t = \frac{\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^3}{\left[\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^2 \right]^{3/2}} \quad (2.36)$$

$$K_t = \frac{\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^3}{\left[\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^2 \right]^2} - 3 \quad (2.37)$$

Si la normalité totale n'est pas confirmée, on teste *la normalité grossière* en utilisant les relations entre l'écart type, l'écart moyen et l'écart probable d'une distribution normale [30] définie par :

$$\mathbf{ad} = \frac{\sum_{i=1}^N |x(i) - \bar{x}|}{N} = 0,7979 \times \sigma \quad (2.38)$$

$$\mathbf{pd} = \frac{Q3-Q1}{2} = 0,6745 \times \sigma \quad (2.39)$$

où σ est l'écart type,

ad l'écart moyen,

et **pd** l'écart probable.

Si les écarts en valeur absolue entre le premier et le second terme des équations 2.38 et 2.39 restent inférieur à 5% , la vraisemblance avec la loi normale sera acceptée.

L'**homoscedasticité** a été testée en utilisant à la fois le test de Bartlett et le test de Levene [31].

Principe du test de Bartlett :

Nous concluons que les données sont homoscedastes

$$\text{Si } \frac{B}{c} < \chi^2 (l - 1) \quad (2.40)$$

avec

$$B = (N - l) \ln(s_p^2) - \sum_{i=1}^l (N_i - 1) \ln(s_i^2) \quad (2.41)$$

$$s_p^2 = \frac{\sum_{i=1}^l (N_i - 1) s_i^2}{n - l} \quad (2.42)$$

$$C = 1 + \frac{1}{3(l - 1)} \left(\sum_{i=1}^l \frac{1}{N_i - 1} - \frac{1}{N - l} \right) \quad (2.43)$$

Principe du test de Levene

Nous concluons que les données sont homoscedastes

$$\text{Si } D < \text{Fisher } (l - 1; n - l) \quad (2.44)$$

$$D = \frac{(n - l) \sum_{i=1}^l N_i (\bar{E}_i - \bar{E})^2}{(l - 1) \sum_{i=1}^l \sum_{j=1}^{N_i} (E_{i,j} - \bar{E}_i)^2} \quad (2.45)$$

$$\bar{X}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} E_{i,j} \quad (2.46)$$

$$\bar{E}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} E_{i,j} \quad (2.47)$$

$$\bar{E} = \sum_{i=1}^l N_i \bar{E}_i \quad (2.48)$$

où N est la taille totale égale à $N_i \times l$;

l , le nombre de groupes égal à S ;

N_i , la taille de chaque groupe ;

s_i^2 , la variance de chaque groupe qui correspond à la colonne « i » ;

$x_{i,j}$, la variable qui correspond à la colonne « i », ligne « j » ;

\bar{X}_i , la moyenne du groupe « i » ;

$\text{Fisher } (l - 1; n - l)$ la valeur de F sur la table de Fisher pour « $(l - 1; n - l)$ » degré de liberté ;

et $\chi^2 (l - 1)$ la valeur de Khi-2 sur table avec « $l-1$ » degré de liberté.

2.3.4. Méthode de détection de la tendance

En pratique, les conditions d'application des tests paramétriques (indépendance, normalité, homoscedasticité) sont souvent difficiles à respecter.

Nous avons donc préféré nous orienter vers l'utilisation des tests non paramétriques pour déterminer l'existence de tendance au niveau de nos chroniques.

Pour étudier la corrélation entre les rangs de la hauteur des pleines mers ou des basses mers noté R_x et les rangs de leurs temps de prélèvement noté R_y , nous avons mis en œuvre les tests de corrélation de rang de Mann et de Kendall et le test de corrélation de rang de Spearman

L'objectif de ces tests est de constater la significativité de la composante tendancielle dans le schéma de prévision.

2.3.4.1. Les tests de corrélation de rang de Mann et de Kendall [32] [33]

Pour étudier la corrélation entre deux grandeurs X et Y , Mann (1945) et Kendall (1970) [32] [33] proposaient d'effectuer des tests basés sur l'utilisation du :

- coefficient de corrélation de rang τ ;
- terme S'' .

Test basé sur l'utilisation du coefficient de corrélation de rang τ (Test de Kendall)

Dans son article « *Rank Correlation Methods* » en 1975, Kendall définissait le coefficient τ_a par la relation

$$\tau_a = \frac{2S''}{N(N-1)} \quad (2.49)$$

$$\text{Avec } S'' = \sum_{i < j}^N \text{signe}(X_i - X_j) \text{signe}(Y_i - Y_j) \text{ et } \text{signe}(G) = \begin{cases} +1 & \text{si } G > 0 \\ 0 & \text{si } G = 0 \\ -1 & \text{si } G < 0 \end{cases} \quad (2.50)$$

Pour tenir compte des valeurs identiques, il modifia τ sous la forme τ_b suivante

$$\tau_b = \frac{S''}{\sqrt{\frac{N(N-1)}{2} - E_x} \sqrt{\frac{N(N-1)}{2} - E_y}} \quad (2.51)$$

$$\text{où } E_x = \sum_{x=1}^{m1} \frac{N_x (N_x - 1)}{2} \quad (2.52)$$

$$\text{et } E_y = \sum_{y=1}^{m2} \frac{N_y (N_y - 1)}{2} \quad (2.53)$$

$m1$ étant le nombre de groupes de rangs identiques sur les rangs R_x de la grandeur X ,
 N_x le nombre de rangs identiques contenu dans le x -ième groupe,
 $m2$ le nombre de groupes de rangs identiques sur les rangs R_y de la grandeur Y ,
et N_y le nombre de rangs identiques contenus dans le y -ième groupe.

Sous l'hypothèse **H₀** « il n'y a pas de corrélation entre X et Y », la moyenne et la variance de τ_b ont été estimés par M.G Kendall sous les formes suivantes que nous avons retenues

$$E_1(\tau_b) = 0 \quad (2.54)$$

et

$$Var(\tau_b) = \frac{2(2N+5)}{9N(N-1)} \quad (2.55)$$

Pour un grand échantillon, la règle de décision est, si la valeur absolue de $Z_1 = \frac{\tau_b - E(\tau_b)}{\sqrt{Var(\tau_b)}}$ est supérieure à la valeur critique $U_{0,1,0,05}$ de la loi normale centrée- réduite $N(0,1)$, nous rejetons l'hypothèse **H₀** en faveur de l'hypothèse **H₁** « il existe une relation monotone entre les deux grandeurs X et Y ».

Test basé sur l'utilisation du terme S'' (Test de Mann-Kendall)

Dans cette méthode, nous n'avons pas besoin de calculer τ_b pour effectuer le test de significativité sur l'existence de tendance. Il suffit de connaître la valeur de S'' pour pouvoir passer au test de significativité en utilisant la fonction Z_2 définie par :

$$Z_2 = \begin{cases} \frac{S''-1}{\sqrt{Var(S'')}} & \text{pour } S'' > 0 \\ 0 & \text{pour } S'' = 0 \\ \frac{S''+1}{\sqrt{Var(S'')}} & \text{pour } S'' < 0 \end{cases} \quad (2.56)$$

Sous l'hypothèse **H₀**, la moyenne et la variance de S'' pourront être estimées en utilisant respectivement les relations suivantes :

$$E_1(S'') = 0 \quad (2.57)$$

$$Var(S'') = \frac{1}{18} [N(N-1)(2N+5) - \sum_{g=1}^{m3} N_g (N_g - 1)(2N_g + 5)] \quad (2.58)$$

où $m3$ est le nombre de groupes de rangs identiques,
et N_g , le nombre de rangs identiques du g -ième groupe.

Pour un grand échantillon, la règle de décision est de rejeter l'hypothèse H_0 dès que la valeur absolue de Z_2 est supérieure à la valeur critique $U_{0,1,0,05}$ de la loi normale centrée-réduite $N(0,1)$.

Nous dirons que la tendance est :

- croissante en fonction du temps pour Z_1 positif,
- décroissante en fonction du temps pour Z_1 négatif.

Le même raisonnement reste valable pour Z_2 .

2.3.4.2. *Le test de corrélation sur les rangs de Spearman*

Le calcul du coefficient de corrélation de rang de Spearman r_s permet de mesurer la dépendance monotonique entre deux grandeurs X et Y .

Dans cette étude, nous avons retenu la formulation de r_s dans la publication de MATIAS Cathérine [34] qui permet de tenir compte des rangs identiques et qui est définie par la relation :

$$r_s = \frac{\sum_{i=1}^N (R_x - \bar{R}_x)(R_y - \bar{R}_y)}{\sqrt{[\sum_{i=1}^N (R_x - \bar{R}_x)^2] * [\sum_{i=1}^N (R_y - \bar{R}_y)^2]}} \quad (2.59)$$

où \bar{R}_x est le rang moyen correspondant aux rangs R_x ,

\bar{R}_y , le rang moyen correspondant aux rangs R_y ,

R_x , le rang correspondant à la hauteur des pleines mers ou des basses mers,

et R_y , le rang correspondant au temps de prélèvement.

Pour chaque groupe de rangs identiques, les rangs de chaque élément seront remplacés par la moyenne des rangs du groupe.

Le test de significativité [35] utilisé est la suivante :

En prenant comme Hypothèses :

- **H_0** il n'a pas de corrélation des rangs ou $a_{k_1} = 0$,
- et **H_1** il existe une corrélation monotone ou $a_{k_1} \neq 0$.

La règle de décision pour un grand échantillon est :

- si $\left| r_s \sqrt{\frac{N-2}{1-r_s^2}} \right|$ est inférieure à $t_{1-\alpha/2}(N-1)$, alors on accepte **H₀**,
- si $\left| r_s \sqrt{\frac{N-2}{1-r_s^2}} \right|$ est supérieure à $t_{1-\alpha/2}(N-1)$, alors on rejette **H₀**.

où $t_{1-\alpha}(N-1)$ est la valeur dans la table de Student à $N-1$ degré de liberté.

Si l'hypothèse **H₁** est acceptée, l'interprétation du résultat se fait :

- par rapport au signe de r_s ;
- ou par rapport à la valeur de r_s .

Par rapport au signe de r_s , nous disons que :

- la corrélation est négative et la tendance est décroissante si $r_s < 0$,
- la corrélation est positive et la tendance est croissante si $r_s > 0$.

Par rapport à la valeur de r_s , nous disons que la liaison entre les rangs R_x et R_y :

- est absente si $|r_s|$ est proche de 0 ;
- est très faible si $|r_s|$ est comprise entre 0,025 et 0,5 ;
- est faible si $|r_s|$ est comprise entre 0,50 et 0,70 ;
- est modérée si $|r_s|$ est comprise entre 0,70 et 0,80 ;
- est forte si $|r_s|$ est comprise entre 0,80 et 0,90 ;
- est très forte si $|r_s|$ est comprise entre 0,90 et 1.

2.3.5. Méthode d'identification du schéma de prévision [36]

Une fois la périodicité et la tendance confirmées ou rejetées, nous pouvons procéder au choix du schéma de prévision.

Avant d'appliquer les différentes méthodes, nous avons dû stabiliser notre série en appliquant une transformation logarithmique de base 10. Cela a aussi permis de transformer notre schéma multiplicatif en schéma additif.

Nous avons aussi transformé le vecteur issu des données stabilisées en matrice N/S lignes et S colonnes en tenant compte qu'il faudrait laisser environ 5% des observations pour vérifier la qualité de la projection à partir des modèles de calcul.

En général, nous choisissons un des deux schémas de prévision suivants.

- le schéma additif qui est défini par la relation

$$x_t = t_t + s_t + r_t \quad (2.60)$$

- ou le schéma multiplicatif défini par

$$x_t = t_t \times s_t \times r_t \quad (2.61)$$

où t_t représente la composante tendancielle du schéma de prévision,

s_t la composante saisonnière du même schéma,

et r_t les résidus ou innovation.

Des variantes mixtes de ces deux schémas sont également possibles.

Pour reconnaître le schéma en cohérence avec nos données [37], nous avons testé trois méthodes :

- ⇒ la méthode du profil,
- ⇒ la méthode de la bande,
- ⇒ et la méthode analytique.

Avant d'appliquer ces trois techniques, nous devons transformer le vecteurs $\overrightarrow{V_d}$ de taille N représentant nos données initiales en matrice M1 composé de S colonnes et N/S lignes. S étant la saisonnalité en valeur entière que nous avons trouvée au paragraphe 2.3.2.6.

Sur la figure 2.2, nous avons représenté la procédure de transformation du vecteur $\overrightarrow{V_d}$ en matrice M1.

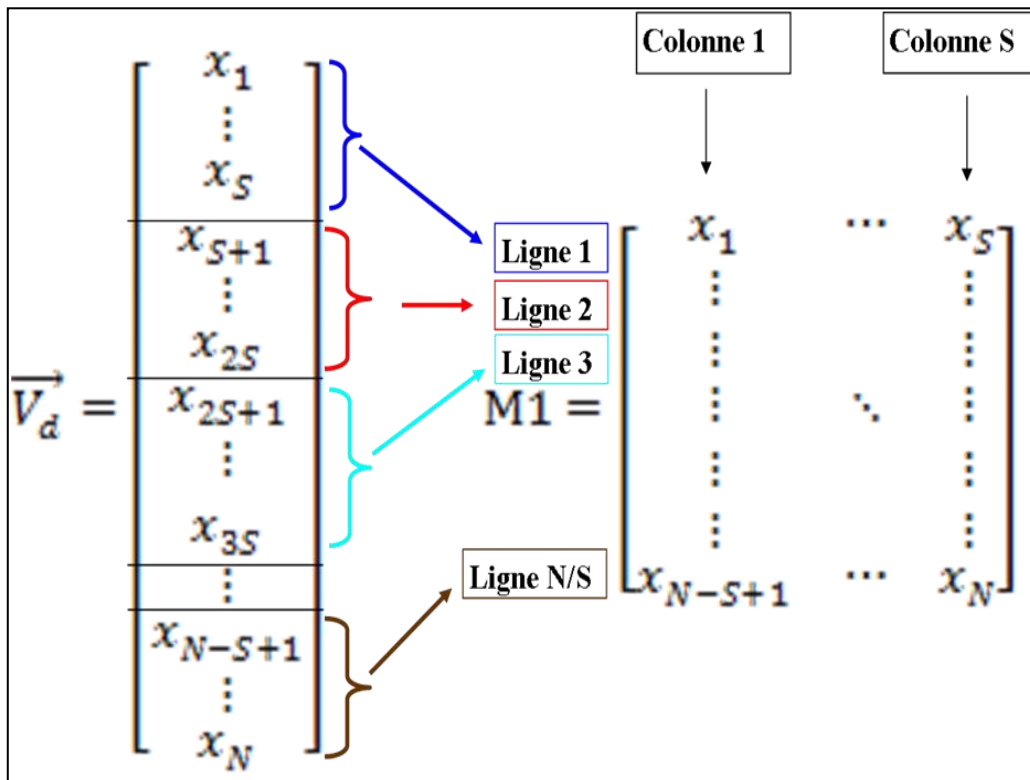


Figure 2 . 2 : Transformation des données vectorielles en matrice

Nous obtenons donc la matrice M1 en mettant :

- à sa ligne 1 les S premières valeurs du vecteur $\vec{V_d}$,
- à sa ligne 2 les S+1 à 2S valeurs du vecteur $\vec{V_d}$,
- à la ligne 3 les 2S+1 à 3S valeurs du vecteur $\vec{V_d}$,
- et ainsi de suite jusqu'à la ligne N/S.

N correspond à la taille de l'échantillon et S à la valeur entière de la saisonnalité.

2.3.5.1. Méthode du profil

Elle consiste à superposer sur un même graphique, les courbes représentatives du profil dans chaque période. Si les courbes sont parallèles entre elles, nous choisissons un schéma additif et si elles s'entrecroisent alors nous optons pour un schéma multiplicatif.

2.3.5.2. Méthode de la bande

La méthode de la bande consiste à construire pour chaque période, d'une part la courbe des minima, et d'autre part la courbe des maxima. On adopte les deux règles de décisions suivantes :

- Si les deux courbes sont parallèles, on opte pour le schéma multiplicatif,
- Si les deux courbes s'entrecroisent, on choisit un schéma additif.

2.3.5.3. *Méthode analytique*

La méthode analytique consiste à régresser pour chaque période l'écart type σ_i par rapport à la moyenne \bar{x}_i .

Si la pente de la droite de régression $\sigma_i = \mu\bar{x}_i + g_1$ est significativement nulle alors nous adoptons un modèle additif avec saisonnalité rigide en amplitude et en période et si la pente de cette même droite est significativement non nulle, alors nous optons pour le choix d'un modèle multiplicatif. La saisonnalité est dans ce cas souple avec variation de l'amplitude au cours du temps

Règles de décision

Sous l'hypothèse H_0 défini par « le coefficient n'est pas significativement différent de 0 » :

Si t-statistic < t-student de la table, alors on accepte **H_0** ,

Si t-statistic > t-student de la table, alors on rejette **H_0** .

2.3.6. Conclusion

Dans ce troisième paragraphe du chapitre 2, nous avons rappelé que :

- si la valeur de l'exposant de Hurst appliqué à la chronique sur les hauteurs des marées est supérieure à 0,5 alors nous avons la possibilité de modéliser ces hauteurs par ARIMA ;
- si les différents tests confirment la présence de saisonnalité et de tendance alors nous devons inclure ses 2 composantes dans le schéma de prévision ;
- et enfin, nous avons constaté que pour trouver le schéma de prévision en cohérence avec la réalité, nous devons mettre en œuvre les trois méthodes et choisir le schéma, multiplicatif ou additif, qui est le plus confirmé après mise en œuvre.

2.4. Les différents types de modèles

2.4.1. Modèle Census X11 [38]

La méthode Census X11 est une technique de dessaisonnalisation des séries chronologiques par moyenne mobile centrée et pondérée.

Cette méthode utilise les moyennes mobiles de Henderson pour l'estimation de la tendance.

Pour sa mise en œuvre, nous devons utiliser l'un des schémas de prévision que nous avons défini au paragraphe 2.3.5.

Quelque soit le schéma choisi, additif ou multiplicatif, cette méthode utilise la forme matricielle M1 des données que nous avons définie à la figure 2.2 et permet de séparer la composante saisonnière, des composantes tendancielle et résiduelle.

Pour trouver ces composantes, nous devons exécuter successivement les opérations suivantes.

2.4.1.1. Estimation de la Tendance $t_{t,1}$ par moyenne mobile centrée d'ordre S [36]

La Tendance extra-saisonnière s'obtient en appliquant aux données brutes stabilisées une moyenne mobile centrée d'ordre S .

Cette moyenne mobile se calcule en utilisant pour une saisonnalité paire S la relation

$$t_{t,1} = \frac{1}{2l_p} \left(\frac{1}{2}x_{t-l_p} + x_{t-l_p+1} + \dots + x_{t+l_p-1} + \frac{1}{2}x_{t+l_p} \right) = M_{2l_p}(x_t) \quad (2.62)$$

et pour une saisonnalité S impaire la formule

$$t_{t,1} = \frac{1}{2l_p + 1} (x_{t-l_p} + x_{t-l_p+1} + \dots + x_{t+l_p-1} + x_{t+l_p}) = M_{2l_p+1}(x_t) \quad (2.63)$$

où l_p est un entier positif

2.4.1.2. Estimation de la composante saisonnière – irrégulière $s_{t,1} + r_{t,1}$ [38]

La composante Saisonnière-Irrégulier est obtenu par différence entre les données initiales et la tendance

$$s_{t,1} + r_{t,1} = x_{t,1} - t_{t,1} \quad (2.64)$$

La composante saisonnière $s_{t,1}$ se déduit du résultat précédent en appliquant une moyenne mobile 5 termes « 3x3 » puis en soustrayant à $M_{3 \times 3}(x_{t,1} - t_{t,1})$ la valeur correspondant à sa moyenne mobile centrée d'ordre S.

$s_{t,1}$ s'obtient alors en utilisant la formule :

$$s_{t,1} = [M_{3 \times 3}(x_{t,1} - t_{t,1})] - M_{2m+1}[M_{3 \times 3}(x_{t,1} - t_{t,1})] \quad (2.65)$$

où $M_{3 \times 3}$ est la moyenne mobile centrée à 5 termes de coefficient de pondération égal à $\left\{ \frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{1}{9} \right\}$

2.4.1.3. Estimation de la série corrigée des variations saisonnières $scvs_{t,1}$ [38]

Pour obtenir la série corrigée des variations saisonnières $scvs_{t,1}$, on utilise la relation :

$$scvs_{t,1} = x_{t,1} - s_{t,1} \quad (2.66)$$

2.4.1.4. Estimation de la tendance-cycle $t_{t,2}$ par moyenne de Henderson à 13 ou 9 termes [38]

Il suffit d'utiliser la relation

$$t_{t,2} = H_{13 \text{ ou } 9}(scvs_{t,1}) \quad (2.67)$$

Le **principe** est simple, si le rapport $\frac{I}{C}$ est supérieur à 1, nous appliquons une moyenne mobile de Henderson à 13 termes et dans le cas contraire nous appliquons une moyenne mobile de Henderson à 9 termes

C est donné par la relation :

$$C = \frac{1}{n-1} \sum_{i=2}^n (t_{i,2} - t_{i-1,2}) \quad (2.68)$$

et I est obtenu en appliquant la formule :

$$I = \frac{1}{n-1} \sum_{i=2}^n (scvs_{i,1} - t_{i,2}) - (scvs_{i-1,1} - t_{i-1,2}) \quad (2.69)$$

2.4.1.5. Nouvelle expression de la composante saisonnière-irrégulière $s_{t,2} + r_{t,2}$ [38]

Après connaissance de l'estimation de la tendance $t_{t,2}$, nous calculons la nouvelle expression de la composante Saisonnière-Irrégulière définie par

$$s_{t,2} + r_{t,2} = x_{t,1} - t_{t,2} \quad (2.70)$$

Le terme $s_{t,2}$ s'obtient en utilisant la relation :

$$s_{t,2} = [M_{3 \times 5}(x_{t,1} - t_{t,2})] - M_{2m+1}[M_{3 \times 5}(x_{t,1} - t_{t,2})] \quad (2.71)$$

où $M_{3 \times 5}$ est la moyenne mobile centrée à 7 termes de coefficient de pondération égal à $\left\{ \frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15} \right\}$

2.4.1.6. Nouvelle expression de la série corrigée des variations saisonnière $scvs_{t,2}$ [38]

La nouvelle formulation de la série corrigée des variations saisonnière $scvs_{t,2}$ est obtenue en utilisant

$$scvs_{t,2} = x_{t,1} - s_{t,2} \quad (2.72)$$

Dans notre étude, l'intervalle de la boîte à moustache a été estimé en utilisant toutes les données alors que pour le logiciel Census X11, les points aberrants sont détectés en utilisant :

- les données appartenant à 2 saisonnalités antérieures au point considéré ;
- les données appartenant à la saisonnalité du point considéré ;
- et les données appartenant à 2 saisonnalités futures au point considéré.

2.4.2. Modèle ARIMA (p+1,d,q) - GARCH (P,Q)

Le modèle ARIMA(p+1,d,q) consiste à rendre stationnaire un processus non stationnaire en différenciant d fois les données initiales, puis à appliquer à la série différenciée un modèle autorégressif d'ordre p+1 et un modèle moyenne mobile d'ordre q.

De façon analogue, pour tout entier relatif t un modèle ARIMA(p+1,d,q) est représenté par la formulation suivante.

$$r'_t = (I - B)^d x_t \quad (2.73)$$

$$r'_t = \sum_{i=1}^{p+1} \varphi_i r'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (2.74)$$

où x_t représente les données initiales à l'instant t qui sont non stationnaires,

B est l'opérateur de retard,

r'_t , le processus stationnaire,

et ε_t représente les résidus qui constitue un bruit blanc.

Notons que les résidus sont les termes aléatoires obtenus en appliquant le modèle de calcul à un échantillon d'étude tandis que les erreurs concernent le même phénomène pour toute la population.

Les résidus constituent un bruit blanc faible si sa moyenne est constante, sa variance est un nombre fini et si le processus est non auto corrélés.

Si de plus le processus est indépendant et identiquement distribué alors le bruit blanc est fort.

Pour déterminer les paramètres de notre modèle ARIMA(p+1,d,q), nous avons appliqué la méthodologie suivante :

- estimer par régression linéaire les coefficients du modèle AutoRegressif de Dickey Fuller Augmenté (DFA(p,d,0)) ;
- estimer les coefficients en utilisant la modélisation ARIMA(p+1,d,0) sous Matlab et garder seulement les coefficient significatifs ;
- Inclure dans le modèle ARIMA(p+1,d,0) une partie moyenne mobile d'ordre « q » et chercher le q_{\max} ;
- Estimer simultanément par maximum de vraisemblance les coefficients du modèle ARIMA(p+1,d,q) - GARCH(P,Q) en vérifiant l'adéquation des modèles au niveau de leurs résidus et de leurs résidus standardisés.

Dans ces modèles « q » correspond à l'ordre de la partie à **Moyenne Mobile**, « d » à l'ordre d'**Intégration** et « p » à l'ordre de la partie **AutoRégressive**.

2.4.2.1. Identification de d et p

Nous avons adopté la formulation de l'équation 2.3.72 proposée par Box et Jenkins [39] pour éliminer la tendance et stationnariser notre série en remplaçant les valeurs de x_t par celles de la série corrigée des variations saisonnière $scvs_{t,2}$

En pratique, nous pouvons retrouver l'ordre d'intégration « d » et l'ordre de la partie Autorégressive « p » qui optimise le modèle DFA en effectuant les trois étapes suivantes :

- Minimisation de l'écart type en fonction de « d »
- Application de la stratégie de Dickey Fuller Augmenté
- Recherche de l'ordre d'intégration final « d_{final} » et de l'ordre optimal « p_{optimal} » du modèle DFA

a) Minimisation de l'écart type de la série différenciée [39]

Cette méthode consiste à rechercher la valeur de « d » minimisant l'écart type de la série différenciée en faisant varier l'ordre d'intégration « d » entre 0 et 2 et rarement entre 0 et 3 ».

b) Stratégie de Dickey Fuller Augmenté

Cette approche permet de retrouver, d'une part, l'ordre d'intégration « d » et d'autre part, de fixer le nombre de retard « p_{initial} » du modèle DFA.

Le principe est simple, on fixe tout d'abord **un nombre de retard initial** « $p = p_{\text{initial}}$ », puis on effectue le test de stationnarité au sens de Dickey Fuller Augmenté en faisant varier l'ordre d'intégration « d » entre « d_{min} et $d_{\text{min}} + 1$ ».

« d_{min} » représente la valeur de « d » minimisant l'écart type des séries différenciées.

- Si le test de Dickey Fuller Augmenté confirme la stationnarité de la série différenciée et que les résidus sont non corrélés au test de Ljung-Box, alors :
⇒ le « d » choisi (d_{min} ou $d_{\text{min}} + 1$) correspond au bon ordre d'intégration
et

⇒ le nombre de retard « p_{initial} » choisi correspond à un ordre maximal acceptable pour notre modèle DFA.

- Si le test de stationnarité de Dickey Fuller Augmenté est rejeté ou le test de Ljung Box sur les résidus n'est pas significative, alors le nombre de retard choisi n'est pas suffisant ; nous augmentons « p_{initial} » d'une ou de plusieurs unités et on refait les tests qui cette fois-ci devront confirmer la stationnarité au sens de Dickey Fuller Augmenté et au sens de Ljung-Box. Cette nouvelle valeur de « p_{initial} » représentera alors l'ordre maximal acceptable pour notre modèle DFA.

Le test de **Dickey Fuller Augmenté** que nous avons utilisé consiste à appliquer une stratégie [40] à travers le modèle 1, 2 et 3 définis respectivement par

$$r'_t = \Phi r'_{t-1} + \sum_{j=1}^p \Phi_j \Delta r'_{t-j} + \varepsilon_t \quad (2.75)$$

$$r'_t = \Phi r'_{t-1} + a + \sum_{j=1}^p \Phi_j \Delta r'_{t-j} + \varepsilon_t \quad (2.76)$$

$$r'_t = \Phi r'_{t-1} + bt + a + \sum_{j=1}^p \Phi_j \Delta r'_{t-j} + \varepsilon_t \quad (2.77)$$

La règle de Dickey-Fuller Augmenté suggère de toujours commencer l'analyse à partir du modèle le plus général ou le modèle 3 défini par l'équation 2.3.76.

Hypothèses 1 :

H₀ : $\Phi = 1$, il existe une racine unitaire,

H₁ : $|\Phi| < 1$, il n'y a pas de racine unitaire, la série est stationnaire.

Décisions :

- Si la statistique de Student pour Φ est inférieur à la statistique de Dickey-Fuller sur table, on rejette H_0 .
- Si la statistique de Student pour Φ est supérieur à la statistique de Dickey-Fuller sur table, on accepte H_0 .

Si la série est stationnaire au sens de Dickey Fuller Augmenté, on teste la significativité de la tendance « b » et de la constante « a » en utilisant la loi t de Student défini par :

Hypothèses 2 :

H₀ : « b » = 0 , Il n'y a pas de tendance,

H₁ : « b » $\neq 0$, Il existe une tendance.

Décisions :

- Si la statistique de Student pour « b » est supérieure à 1,96, on rejette H_0 .
- Si la statistique de Student pour « b » est inférieure à 1,96, on accepte H_0 .

La significativité de la constante « a » se confirme de la même manière que pour la tendance avec comme hypothèse :

H₀ : « a » = 0 , Il n'y a pas de constante,

H₁ : « a » $\neq 0$, Il existe un terme constant dans le modèle.

c) Recherche de l'ordre d'intégration « d_{final} » et de « $p_{optimal}$ »

Si les valeurs de « d » obtenues par les deux méthodes coïncident alors elle correspond à l'ordre d'intégration finale « d_{final} ».

Nous avons déterminé « $p_{optimal}$ » en diminuant progressivement « $p_{initial}$ ».

« $p_{optimal}$ » représente la valeur minimale de « p » permettant d'assurer à la fois la stationnarité de la série différenciée au sens de Dickey Fuller Augmenté et la stationnarité des résidus au sens de Ljung-Box.

Dans le cas où les valeurs de « d » obtenues par les deux méthodes diffèrent, il faut mettre en œuvre le test de Ljung-Box en diminuant progressivement la valeur de « $p_{initial}$ » pour chaque valeur de « d égal à : d_{min} ou $d_{min} + 1$ ».

« d_{final} et $p_{optimal}$ » correspondent aux valeurs de « d et de p » pour lesquels l'ordre « k » au test de Ljung-box est maximal.

Si on obtient pour différentes valeurs de « d et de p » le même ordre « k » au test de Ljung-Box, alors le « p_{optimal} » correspond à la plus petite valeur de « p » car augmenté « p » d'une unité n'est plus significative sur la non autocorrélation des résidus.

Dans ce travail, nous avons choisi la méthode « **contrôle ex-post de l'absence d'autocorrélation des Résidus** » basé sur le choix de « d et de p » permettant de **mieux blanchir les résidus** avec le **minimum de paramètres à estimer** [41].

En cas de valeurs identiques de l'ordre « k », nous choisissons donc le modèle avec le minimum de paramètres à estimer incluant les coefficients du premier retard, de la tendance et de la constante.

Pour déterminer si nos modèles DFA possèdent ou non des tendances et des constantes, il faut comparer les résultats du test de Ljung-Box avec et sans les constantes et tendances, puis choisir toujours le cas pour lequel les résidus sont mieux blanchis.

Pour déterminer le nombre de paramètres, nous avons besoin d'estimer par exemple pour le modèle 3 DFA :

- P_{optimal} coefficients autorégressifs
- et k_I coefficients liés au premier retard, à la tendance et à la constante du modèle.

2.4.2.2. Identification de q_{initial}

Nous avons pris « q_{initial} » égal au dernier Lag significatif de la fonction d'autocorrélation des résidus du modèle ARIMA($p_{\text{optimal}}+1, d_{\text{final}}, 0$) [39] en utilisant :

- pour l'estimation de l'autocorrélation ρ_{k_1} la fonction a_{k_1} définie par

$$a_{k_1} = \frac{\sum_{t=k_1+1}^N (r_t - \bar{r})(r_{t-k_1} - \bar{r})}{\sum_{t=1}^N (r_t - \bar{r})^2} \quad (2.78)$$

- pour le test de significativité de la fonction d'autocorrélation, l'intervalle de Bartlett I défini par :

$$I = \pm 2\sigma_{a_{k_1}} \quad (2.79)$$

$$\sigma_{a_{k_1}} = \sqrt{\frac{(1+2\sum_{t=1}^{h-1} r_t^2)}{N}} \quad (2.80)$$

Les modèles $ARIMA(p_{\text{optimal}}+1, d_{\text{final}}, 0)$ que nous avons construit ici ont été blanchis des coefficients non significatifs, c'est-à-dire des coefficients avec des statistiques t de Student inférieures à 1,96 en valeur absolue. Dans certains cas, pour éviter des modèles mal conditionnés, nous pouvons garder certains coefficients dont la statistique t de Student reste au voisinage des limites de significativité.

2.4.2.3. *Choix du modèle ARIMA(p+1,d,q)*

Pour le choix du modèle en cohésion avec le phénomène étudié, nous avons :

- Tout d'abord, insérer progressivement dans nos modèles des termes moyennes mobiles en variant q entre 0 et q_{initial}
- Ensuite, récupérer les modèles pour lesquels l'ordre « k » au test de Ljung-box est maximal.
- Enfin, en cas de valeurs identiques pour l'ordre « k », maximiser le Critère d'Information d'Akaike (CIA) ou le Critère de Schwarz (SC) des modèles récupérés [40] définis par

$$CIA(p + q + k_1) = N \times \log(\sigma_t^2) + 2(p + q + k_1) \quad (2.81)$$

et $SC(p + q + k_1) = N \times \log(\sigma_t^2) + (p + q + k_1) \times \log(N) \quad (2.82)$

où σ_t^2 est la variance des résidus ;

et k_1 , le coefficient lié au premier retard, à la tendance et à la constante du modèle.

Nous avons décidé de maximiser les critères d'informations pour augmenter la variance au niveau de nos résidus car nous avons constaté que la méthode du maximum de vraisemblance arrive à une meilleure estimation des coefficients de nos modèles au fur et à mesure que la variance au niveau des résidus augmente.

2.4.2.4. *Analyse des Résidus*

a) *Test ARCH*

L'**hétéroscedasticité** a été testé sur le modèle $ARIMA(p+1,d,q)$ en utilisant le test ARCH sous Matlab 2013a défini par [19] :

$$r_t^2 = \pi_o + \sum_{k_1=1}^W \pi_k r_{t-k_1}^2 + e(t) \quad (2.83)$$

où r_t représente le vecteur des résidus

L'**hétéroscédasticité** est confirmée si la p-value issue du test ARCH est inférieure à 0,05.

En pratique, nous pouvons aussi visualiser la fonction d'autocorrélation des résidus au carré. En cas d'hétéroscédasticités, certaines valeurs de d'autocorrélation dépassent les limites de significativité.

En présence d'hétéroscédasticité des résidus, ces derniers seront modélisés [19]:

Soit par un modèle GARCH(P,Q) défini par :

$$r_t = \eta_t \sigma_t \quad (2.84)$$

$$\sigma_t^2 = K + \sum_{j=1}^Q \alpha_j r_{t-j}^2 + \sum_{i=1}^p \gamma_i \sigma_{t-i}^2 \quad (2.85)$$

Pour assurer la stationnarité et la positivité de la variance, nous avons les contraintes suivantes :

$$K > 0 \ ; \ \alpha_j \geq 0 \ ; \ \gamma_i \geq 0 \ \text{et} \ \sum_{i=1}^P \gamma_i + \sum_{j=1}^Q \alpha_j < 1 \quad (2.86)$$

Soit par un modèle EGARCH(P,Q) défini par :

$$r_t = \eta_t \sigma_t \quad (2.87)$$

$$\log \sigma_t^2 = K + \sum_{i=1}^P \gamma_i \sigma_{t-i}^2 + \sum_{j=1}^Q \alpha_j \left[\frac{|r_{t-j}|}{\sigma_{t-j}} - E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} \right] + \sum_{j=1}^Q \xi_j \left(\frac{r_{t-j}}{\sigma_{t-j}} \right) \quad (2.88)$$

$$E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} = E \{ |\eta_{t-j}| \} = \sqrt{\frac{2}{\pi}} \quad (2.89)$$

pour une distribution gaussienne des résidus

$$E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} = E \{ |\eta_{t-j}| \} = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \quad (2.90)$$

pour une distribution de Student avec $\nu \geq 2$ degré de liberté

où K est la constante du modèle de type GARCH ;

γ_i les coefficients GARCH ;

α_j les coefficients ARCH ;

et ξ_j le Leverage.

Soit par un modèle GJR(P,Q) ou modèle de Glosten défini par :

$$r_t = \eta_t \sigma_t \quad (2.91)$$

$$\sigma_t^2 = K + \sum_{j=1}^Q \alpha_j r_{t-j}^2 + \sum_{i=1}^P \gamma_i \sigma_{t-i}^2 + \sum_{j=1}^Q \xi_j \times I[r_{t-j} < 0] r_{t-j}^2 \quad (2.92)$$

$$I[r_{t-j} < 0] = 1 \quad \text{si } r_{t-j} < 0 \quad \text{et } I[r_{t-j} < 0] = 0 \text{ ailleurs} \quad (2.93)$$

Pour assurer la stationnarité et la positivité de la variance, nous avons les contraintes suivantes :

$$K > 0 ; \alpha_j \geq 0 ; \gamma_i \geq 0 ; \alpha_j + \xi_j \geq 0 \quad \text{et} \quad \sum_{i=1}^P \gamma_i + \sum_{j=1}^Q \alpha_j + \frac{1}{2} \sum_{j=1}^Q \xi_j < 1 \quad (2.94)$$

Avec les mêmes notations que pour le modèle EGARCH

b) Test de bruit blanc [39]

Le test de Ljung-Box consiste à comparer la p-Value de la statistique de Ljung-Box à la valeur 0,05 :

$$\chi_m^2 = N(N+2) \sum_{k=1}^m \left[\frac{a_{k_1}^2(r_t)}{N - k_1} \right] \quad (2.95)$$

où a_{k_1} est l'estimation de l'autocorrelation.

Si la p-Value ou $Prob [\chi^2(m-a) > \chi_m^2]$ est supérieure ou égale à 0,05 pour « k_1 » fixé, la série est stationnaire.

Ce test sera appliqué pour vérifier la stationnarité des résidus, des résidus standardisés et des résidus standardisés au carré.

c) Test de normalité

La **normalité** de la distribution des résidus standardisés sera confirmée de la même manière qu'au paragraphe 2.3.3.4.2 ou à l'aide du graphique QQ-plot sous Matlab. Cette variante de la droite de Henry permet de tester l'adéquation graphique avec la loi normale au seuil de signification α égale à 0,05 en comparant la correspondance entre les deux distribution entre l'intervalle $[-1,96, 1,96]$.

Pour vérifier que la **distribution** des résidus standardisés est **centrée et réduite**, il suffit de tester l'égalité entre les paramètres (\bar{r}_t, σ) et $(0, 1)$ ou (\bar{r}_t, σ^2) et $(0, 1)$.

- Pour vérifier l'égalité entre la **moyenne des résidus normalisés** \bar{r}_t et la **valeur zéro (0)**, nous avons appliqué le test d'égalité de la moyenne d'une série à une valeur donnée ou test de Student.

Cette technique permet de déterminer l'intervalle de confiance symétrique à 95 % pour zéro [30].

En notant $[c, d]$ cette intervalle, la **règle de décision** est la suivante :

- \Rightarrow si $0 \in [c, d] = [\bar{r}_t - t_{1-\alpha/2} \frac{s}{\sqrt{N}}, \bar{r}_t + t_{1-\alpha/2} \frac{s}{\sqrt{N}}]$, alors la moyenne \bar{r}_t est non significativement différent de zéro.
- \Rightarrow si $0 < \bar{r}_t - t \frac{s}{\sqrt{N}}$ ou $\bar{r}_t + t \frac{s}{\sqrt{N}} > 0$ alors la moyenne \bar{r} est significativement différente de zéro.

où \bar{r}_t représente la moyenne des résidus normalisés ;

σ l'écart type biaisé de l'échantillon ;

s l'écart type non biaisé égal à $\sigma \times \frac{N}{N-1}$;

N la taille de l'échantillon ;

et $t_{1-\alpha/2}$ la variable de Student à $N-1$ degré de liberté pour un risque de première espèce α .

- Avec la même technique, si on décide d'utiliser le test d'égalité entre **l'écart type des résidus normalisés et un (1)**, la règle de décision est la suivante [30] :

\Rightarrow Si $1 \in [s - t_{1-\alpha/2} \frac{s}{\sqrt{2N}}, s + t_{1-\alpha/2} \frac{s}{\sqrt{2N}}]$, alors l'écart type s est non significativement différent de un,

\Rightarrow Si $1 < s - t_{1-\alpha/2} \frac{s}{\sqrt{2N}}$ ou $s + t_{1-\alpha/2} \frac{s}{\sqrt{2N}} > 1$ alors l'écart type s est significativement différent de un.

- De la même manière que précédemment, si on décide d'utiliser le test d'égalité entre **la variance des résidus normalisés et un (1)**, la règle de décision est la suivante [30] :

\Rightarrow Si $1 \in [s^2 - t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}, s^2 + t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}]$, alors la variance s^2 est non significativement différent de un.

\Rightarrow Si $1 < s^2 - t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}$ ou $s^2 + t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}} > 1$ alors la variance s^2 est significativement différent de un.

2.4.2.5. Analyse de la qualité de la prévision

La prévision consiste à exploiter, sur la base d'analyse rationnelle, les données disponibles et ou l'expérience passée pour estimer des valeurs futures [41].

Pour arriver aux résultats de prévision, nous avons dû effectuer les transformations suivantes :

- à partir de l'estimation de la série corrigée des variations saisonnières différenciée une fois que nous avons notée $\text{Prévision}(\text{diff}(\text{SCVS}))$, pour revenir à l'expression des séries corrigées des variations saisonnières, nous avons appliqué la relation :

$$\text{SCVS}(t+1) = \text{SCVS}(t) + \text{somme cumulé} \left(\text{Prévision}(\text{diff}(\text{SCVS}))(t+1) \right) \quad (2.96)$$

- et pour revenir à l'estimation des données initiales, nous avons appliqué, aux séries corrigées des variations saisonnières $\text{SCVS}(t+1)$, la formule :

$$\text{Prévision (donnée initiale)}(t+1) = 10^{\text{SCVS}(t+1)} \quad (2.97)$$

Sous Matlab 2013a, la série (Prévision(diff(SCVS))(t+1)) est obtenue en récupérant la moyenne conditionnel obtenue en sortie après l'exécution de la commande « garchfit »

La qualité de prévision dans chaque cycle sera ensuite appréciée en utilisant l'indicateur de THEIL

Theil définissait l'indicateur suivant pour déterminer la précision d'un modèle [42] :

$$I_{THEIL} = \frac{RMSE_{modèle}}{RMSE_{méthode alternative}} \quad (2.98)$$

avec

$$RMSE = \sqrt{\frac{1}{N_t} \sum_{t=1}^{N_t} (x_t - x_{modèle})^2} \quad (2.99)$$

où x_t représente les données initiales à l'instant t ;

$x_{modèle}$, les valeurs prédites de x_t à l'instant t ;

et N_t , la taille de l'échantillon sur laquelle on effectue la projection.

Si le ratio de Theil est inférieur à 1, le modèle offre une meilleure précision que la méthode alternative pendant la durée de prévision.

Pour effectuer la prévision de l'année A, nous avons pris les coefficients saisonniers $s_{t,2}$ de l'année A-1. En particulier, l'estimation du coefficient saisonnier de l'année A-1 à l'instant « t » sera donc définie par la relation

$$s_{t,2}(A) = s_{t-713,2}(A-1) \quad (2.100)$$

2.4.2.6. Analyse de la stabilité du modèle

Pour analyser la stabilité des résidus à l'instant t , nous avons mis en œuvre la méthode tracking signal [43].

Cette méthode consiste à normaliser les résidus en les divisant par leurs écarts types, puis à vérifier que les variables normalisées restent à l'intérieur d'un intervalle de confiance choisi préalablement par le prévisionniste.

En pratique, l'écart type est estimé à 1,25 fois le MAD et les formules qui vont nous permettre de construire la courbe du Tracking signal sont défini par

$$TS_t = \frac{Biais_t}{MAD_t} \quad (2.101)$$

$$Biais_t = \sum_{i=1}^t x_{modèle(i)} - x_i \quad (2.102)$$

et

$$MAD_t = \frac{RMSE_t}{1,25} \quad (2.103)$$

Pour vérifier la stabilité des résidus, certains auteurs comme le Professeur Amar Ramudhin suggère de vérifier cette significativité dans l'intervalle ± 4 à ± 6 [43] et d'autres comme Philippe Marier propose un intervalle plus large entre ± 4 à ± 8 [44].

Dans cette étude, nous avons tout d'abord vérifié la stabilité à l'aide de ces trois intervalles. Si cela ne marche pas, alors nous définirons un intervalle de confiance à partir des prévisions issues de la méthode alternative. Ce nouvel intervalle prendra alors la forme

$$\left[-E \left(Max(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right), E \left(Max(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right) \right]$$

dans laquelle E représente la fonction partie entière, Max la valeur maximale, σ_{TS} l'écart type et N_{TS} la taille de l'échantillon considéré.

2.4.3. Conclusion

Dans ce quatrième paragraphe du chapitre 2, nous avons vu que la méthode Census X11 nous permet de décomposer la série des données stabilisées en trois composantes : saisonnières, tendancielle et résiduelles.

Elle nous permet également de récupérer les séries corrigées des variations saisonnières, variables assez stables dont nous aurons besoins pour procéder à la modélisation ARMA.

Nous avons aussi constaté que cette méthode s'applique aussi bien pour les séries avec ou sans tendance, de périodicités paires ou impaires.

Enfin, nous avons vu que notre modèle Census X11-ARIMA est une technique de double dessaisonalisation combinant la méthode empirique non paramétrique, Census X11 et la méthode paramétrique ARIMA.

Par conséquent, cette technique de dessaisonalisation devrait être valable en toute circonstance.

En particulier, nous avons représenté dans l'organigramme ci-après les démarches à suivre pour effectuer une modélisation Censu X11-ARIMA(p,d,q)-GARCH(P,Q).

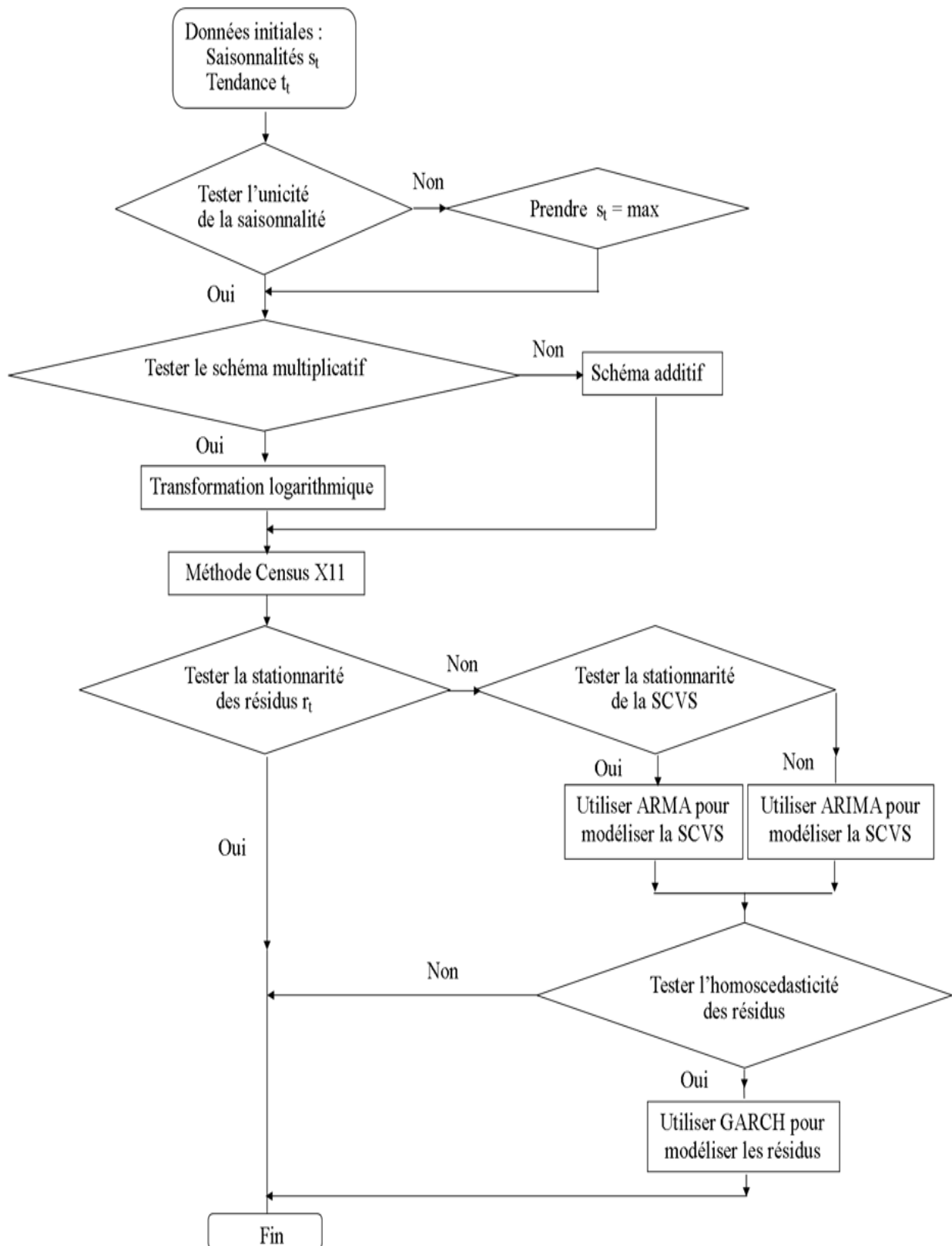


Figure 2 . 3 : Organigramme de la méthode Censu X11-ARIMA(p,d,q)-GARCH(P,Q)

Chapitre 3. Résultats et discussions : applications numériques au cas des marées de Hell-Ville

3.1.Introduction

Ce troisième chapitre concerne la présentation des résultats des travaux de recherche que nous avons effectués durant cette thèse.

Nous avons commencé par la localisation du site et par l'explication du mode de prélèvement des données sur les hauteurs des pleines et des basses mers à Hell-Ville qui est notre zone d'étude.

Ensuite, nous avons défini successivement les résultats conformément à la présentation de la deuxième partie que nous avons décrite au chapitre sur la méthodologie.

Nous allons donc calculer en premier l'exposant de Hurst H_1 de nos données stabilisées afin de confirmer la possibilité d'une modélisation ARIMA.

Nous passerons en second lieu à l'analyse statistique des hauteurs des marées pour nous assurer de l'homogénéité de nos données.

Enfin avant d'entamer la modélisation ARIMA-GARCH des séries corrigées des variations saisonnières, nous allons procéder aux différents tests de détection et d'estimation de la tendance, de la saisonnalité et du schéma de prévision.

3.2. Résultats et discussions

3.2.1. Sources des données

Les deux échantillons que nous avons constitués pour nos 25 années d'observation, entre 1990 et 2014, correspondent à :

- 35 286 données initiales sur les hauteurs des pleines et des basses mers en mètres,
- et à 35 286 données initiales sur les déphasages des temps de prélèvement successifs.

Ces données proviennent des tables des marées du Centre National de Recherches Océanographiques de Hell-Ville.

Les hauteurs des pleines et des basses mers en mètres ont été mesurées à partir du zéro hydrographique situé à 7,743 mètres en dessous d'un repère fondamental fixé dans le portique face au centre Nationale de Recherches Océanographiques de Hell-Ville [45].

Pour les déphasages, ils se déduisent des différences entre les temps de prélèvement successifs.

3.2.2. Localisation su site

Nosy Be est une ile située dans la partie Nord Ouest de Madagascar. Hell-Ville est repérée sur la carte géographique de Madagascar au point de coordonnées géographiques [45]

- de latitude $13^{\circ} 24'$ Sud,
- et de longitude $48^{\circ} 18'$ Est.

Sur la figure 3.1, la zone d'étude est représentée par la partie en couleur verte située au Nord Ouest de Madagascar.

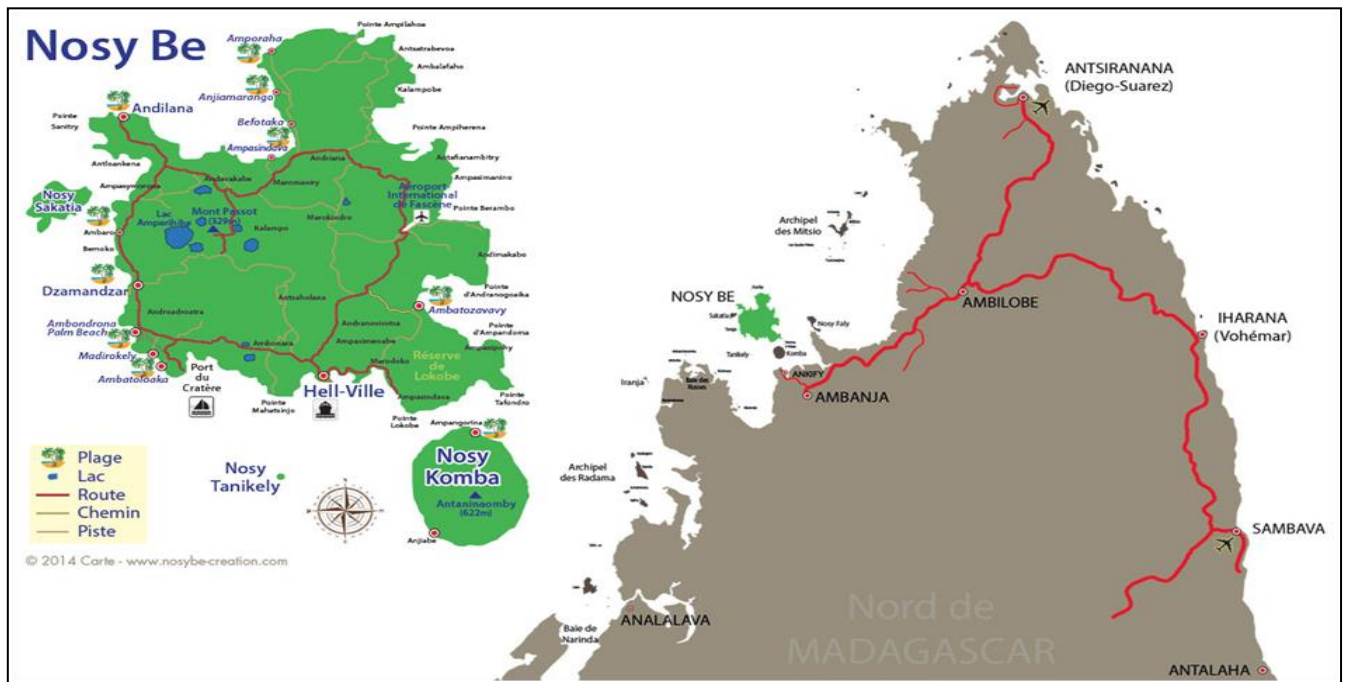


Figure 3 . 1 : Localisation de Hell-Ville

(source : <http://www.villas-orchidees-nosybe.com/images/Nosy-Be/ile-de-nosybe/carte-nosybe-Mada-Villa-Orchidees.jpg>)

3.2.3. Expositant de Hurst H_I

Sur la figure 3.2, nous avons obtenus les estimations des exposants de Hurst H_I pour les hauteurs des pleines et basses mers en régressant la fonction logarithme de $\left(\frac{R}{\sigma}\right)$ sur la fonction logarithme du temps τ .

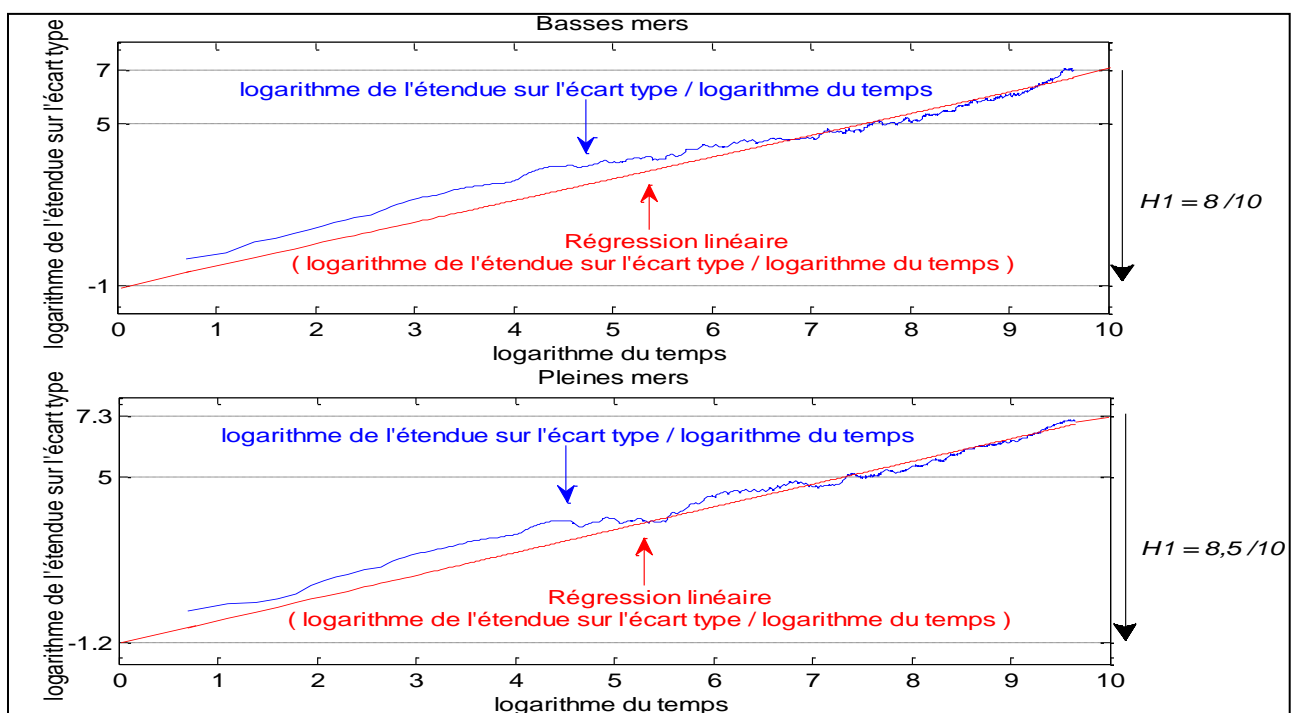


Figure 3 . 2 : Valeurs des exposants de Hurst pour les pleines mers et les basses mers

Pour les basses mers, l'estimation de HI vaut 0,8020. Le processus sur les basses mers est donc persistant ou à mémoire longue puisque HI est supérieur à 0,5.

Pour les pleines mers, l'estimation de HI vaut 0,8510. Le processus sur les pleines mers est également persistant ou à mémoire longue puisque HI reste supérieur à 0,5.

Avant d'appliquer une modélisation FARIMA propre aux séries à mémoire longue et qui est plus compliquée à mettre en œuvre, nous allons tester si une modélisation ARIMA (p,d,q) marche et suffit à prévoir les pleines et les basses mers dans notre zone d'étude et à expliquer le phénomène des marées à Hell-ville.

3.2.4. Représentation graphique des données initiales

Sur la figure 3.3, la représentation graphique des données initiales composées des hauteurs des pleines et des basses mers en mètres, montre une forte probabilité d'existence d'une rupture de pente aux alentours du 31 700 rang ou Lag.

Rappelons que le Lag correspond au rang de prélèvement des hauteurs d'eau consécutives.

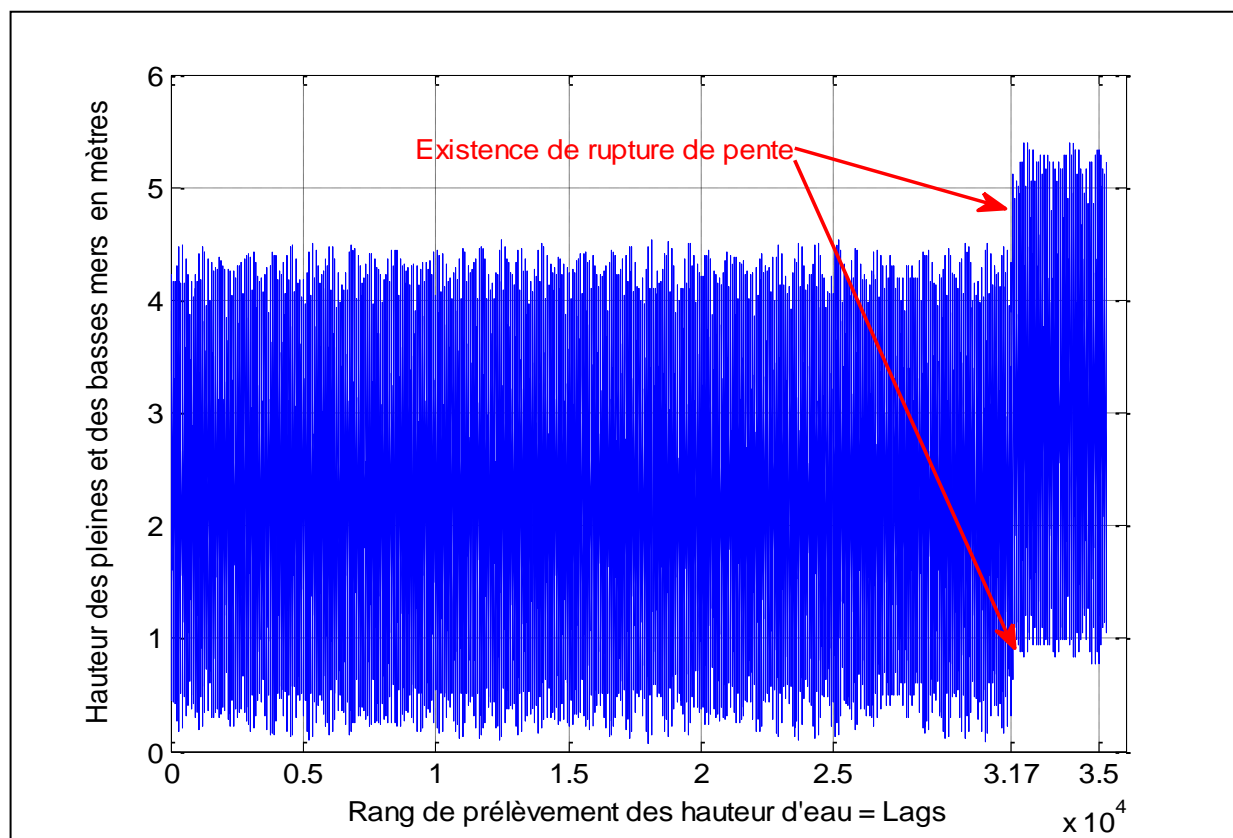


Figure 3.3 : Données initiales composées des hauteurs des pleines et des basses mers

3.2.5. Histogramme

Les résultats de la règle de STURGES a) du paragraphe 2.3.2.2, que nous avons appliquée aux données des hauteurs des pleines et des basses en mètres, sont regroupés dans le tableau 3.1. On y trouve les valeurs minimale (x_{min}) et maximale (x_{max}) de la série, le nombre de classes k_2 et l'amplitude a_{k_2} pour construire l'histogramme.

Tableau 3.1 : Paramètres de l'histogramme des données initiales

Taille de l'échantillon	x_{max}	x_{min}	k_2	a_{k_2}
35 286	5,39 mètres	0,07 mètres	16	0,3325 mètres

La figure 3.4 montre l'histogramme obtenu à partir des paramètres du tableau 3.1 en utilisant la fonction «hist(x,16)» sous Matlab 2013a. Elle montre une distribution bimodale. Il semble donc qu'il existe une hétérogénéité entre les données sur les pleines mers et les données sur les basses mers.

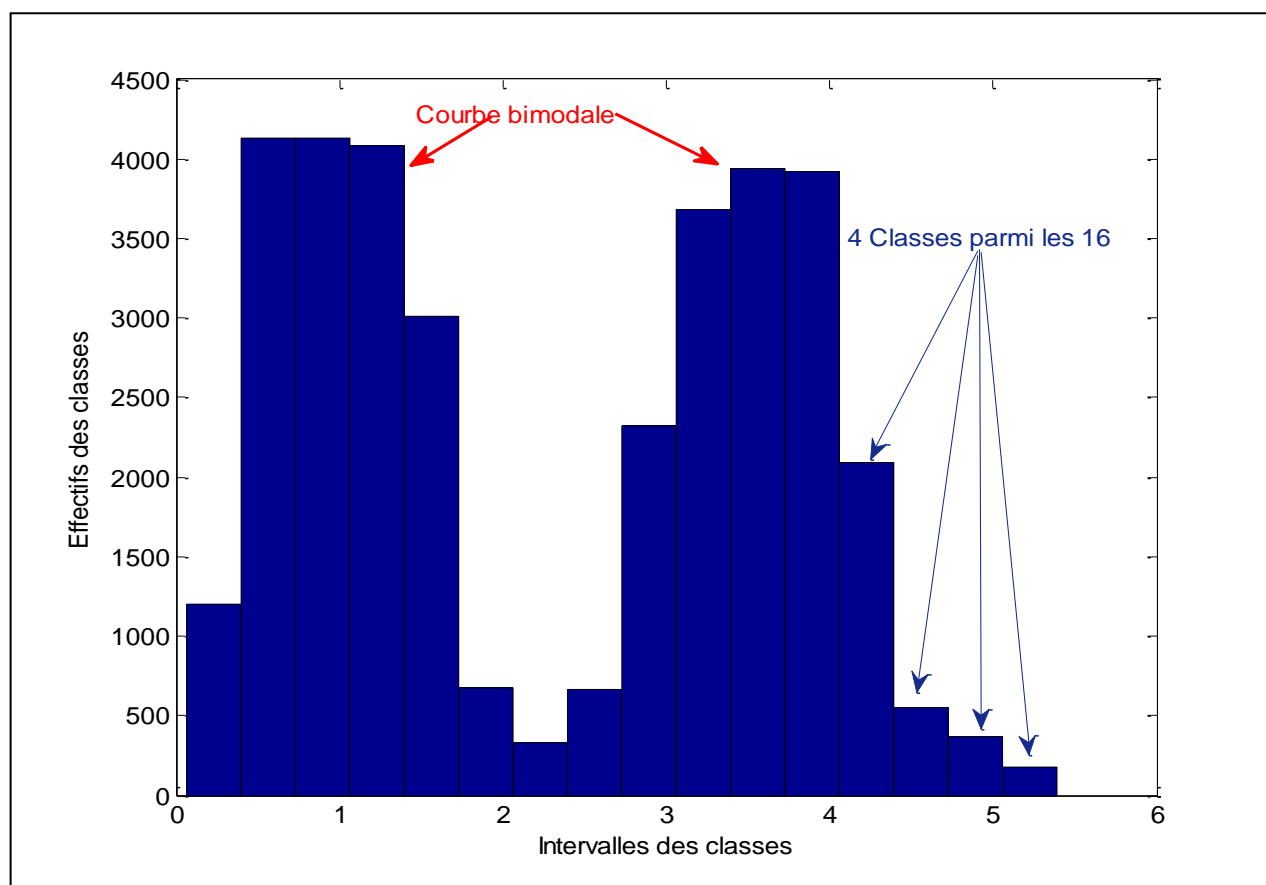


Figure 3.4 : Histogramme des données initiales composées des hauteurs des pleines et des basses mers

En gardant le même nombre de classes calculé dans le paragraphe précédent, nous avons construit deux différents échantillons en séparant les données sur les hauteurs des pleines mers et les données sur les hauteurs des basses mers : c'est l'opération d'homogénéisation.

Leurs caractéristiques sont reportées dans le tableau 3.2 :

Tableau 3.2 : Paramètres des histogrammes des pleines et des basses mers après homogénéisation

Nature des observations	Taille de l'échantillon	x_{max}	x_{min}	k_2	a_{k_2}
Pleines mers	17 643	5,39 mètres	2,30 mètres	16	0,1931 mètres
Basses mers	17 643	2,67 mètres	0,07 mètres	16	0,1625 mètres

Les histogrammes de ces deux échantillons à la figure 3.4 montrent deux distributions unimodales.

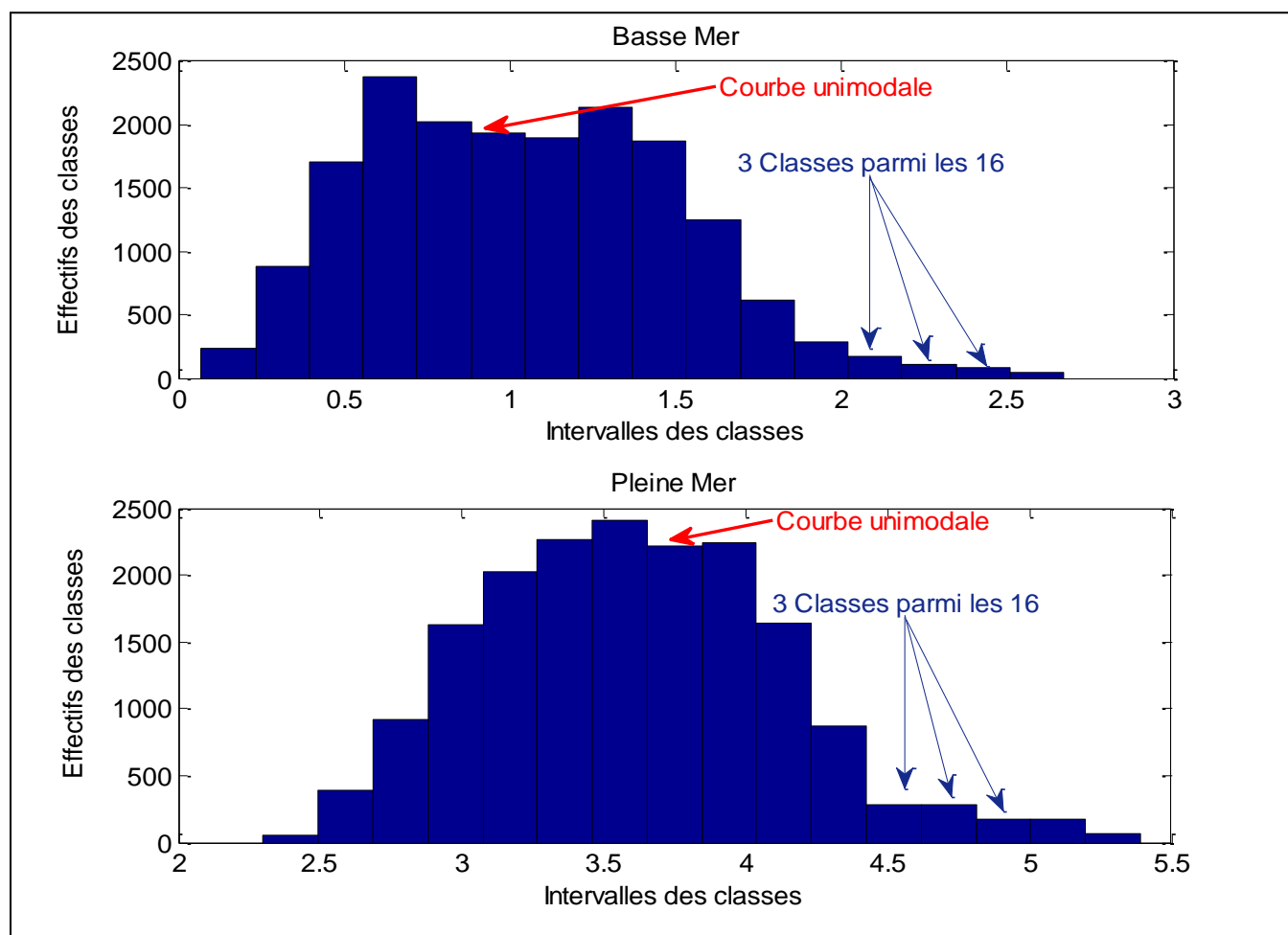


Figure 3.5 : Histogrammes des pleines et des basses mers après homogénéisation

Il semble donc que la séparation des données initiales, en chronique sur les pleines mers et sur les basses mers, permet d'homogénéiser les données issues de nos prélèvements initiaux.

Par conséquent, nous avons décidé d'effectuer deux modélisations : la modélisation des données sur les pleines mers et celle des données sur les basses mers.

3.2.6. Préparation des données homogénéisées

Le tableau 3.3 donne les caractéristiques des boîtes à moustache, des hauteurs des pleines et des basses mers, que nous devons appliquer pour définir les points aberrants.

Tableau 3 . 3 : Paramètres de la boîte à moustache des données homogénéisées

	Q1	Q3	a	b	Taille de l'échantillon
Basse mer	0,670	1,3800	-0,3950	2,4450	17 643
Pleine mer	3,2000	3,9600	2,0600	5,1000	17 643

Les résultats du tableau 3.3 montrent que les hauteurs d'eau normales pour les basses mers se situent entre -0,3950 mètres et 2,4450 mètres et pour les pleines mers entre 2,0600 mètres et 5,1000 mètres. Les points en dehors de ces deux intervalles sont donc atypiques.

Pour les basses mers, nous avons éliminé 1 744 points aberrants situés à l'extrémité de notre chronique pour corriger nos données.

Pour les pleines mers, nous avons éliminé 1 761 points pour blanchir la chronique des hauteurs des pleines mers des points aberrants.

Les boîtes à moustache de la figure 3.6 confirment que les données des hauteurs des pleines et basses mers après suppressions des points aberrants ont été parfaitement corrigées.

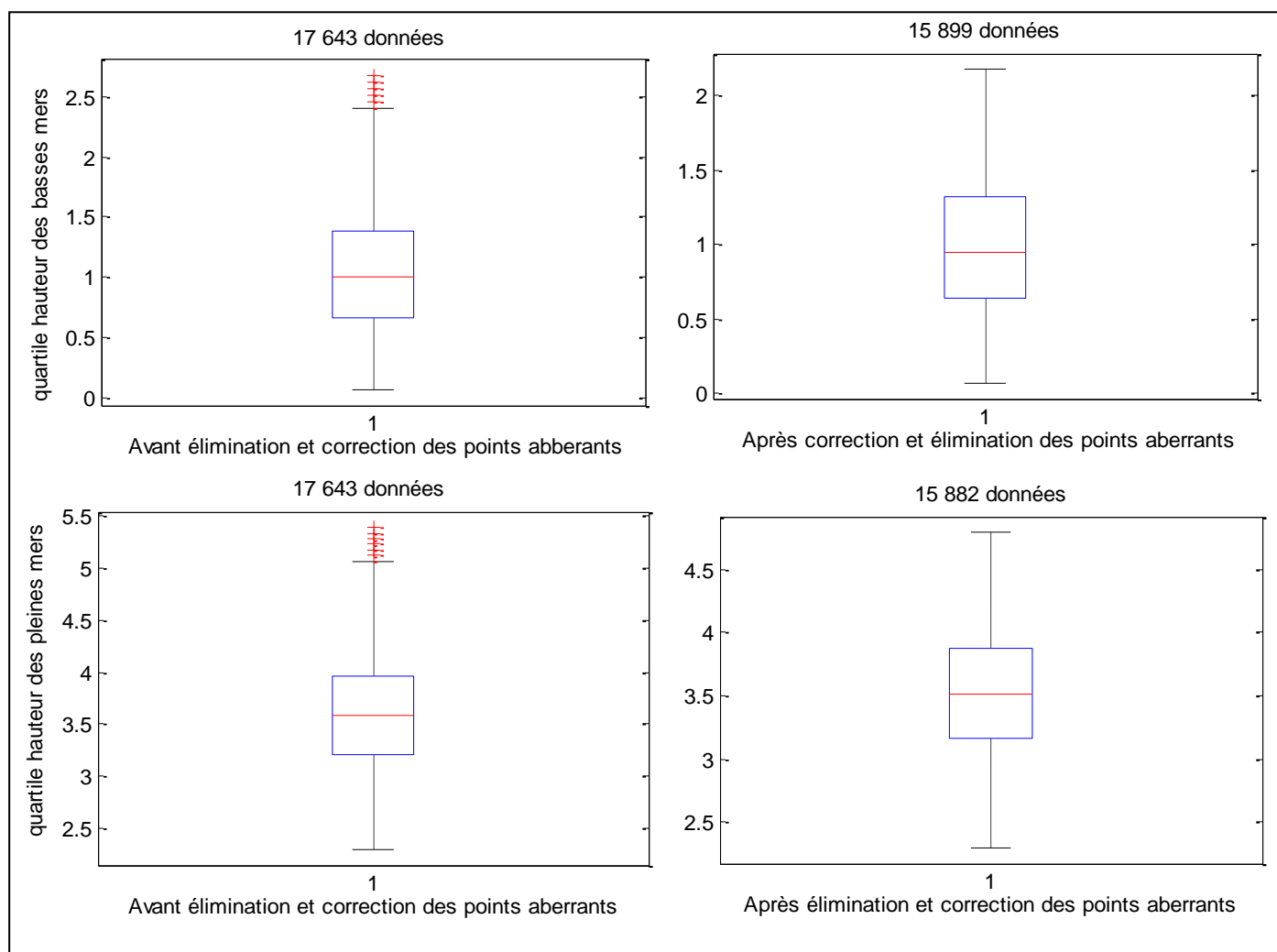


Figure 3.6 : Boîtes à moustache des hauteurs des pleines et basses mers avant et après correction des points aberrants

De la même manière que dans le paragraphe précédent, nous avons établi les caractéristiques et la représentation graphique des boîtes à moustache des données stabilisées, c'est-à-dire, les données initiales transformées par une fonction logarithmique de base 10.

Nous avons regroupé ces résultats dans le tableau 3.4 et sur la figure 3.7.

Tableau 3.4 : Paramètres de la boîte à moustache des données stabilisées

	Q1	Q3	a	b	Taille de l'échantillon
Basse mer	-0,1938	0,1206	-0,6654	0,5922	15 631
Pleine mer	0,4997	0,5877	0,367	0,7197	15 631

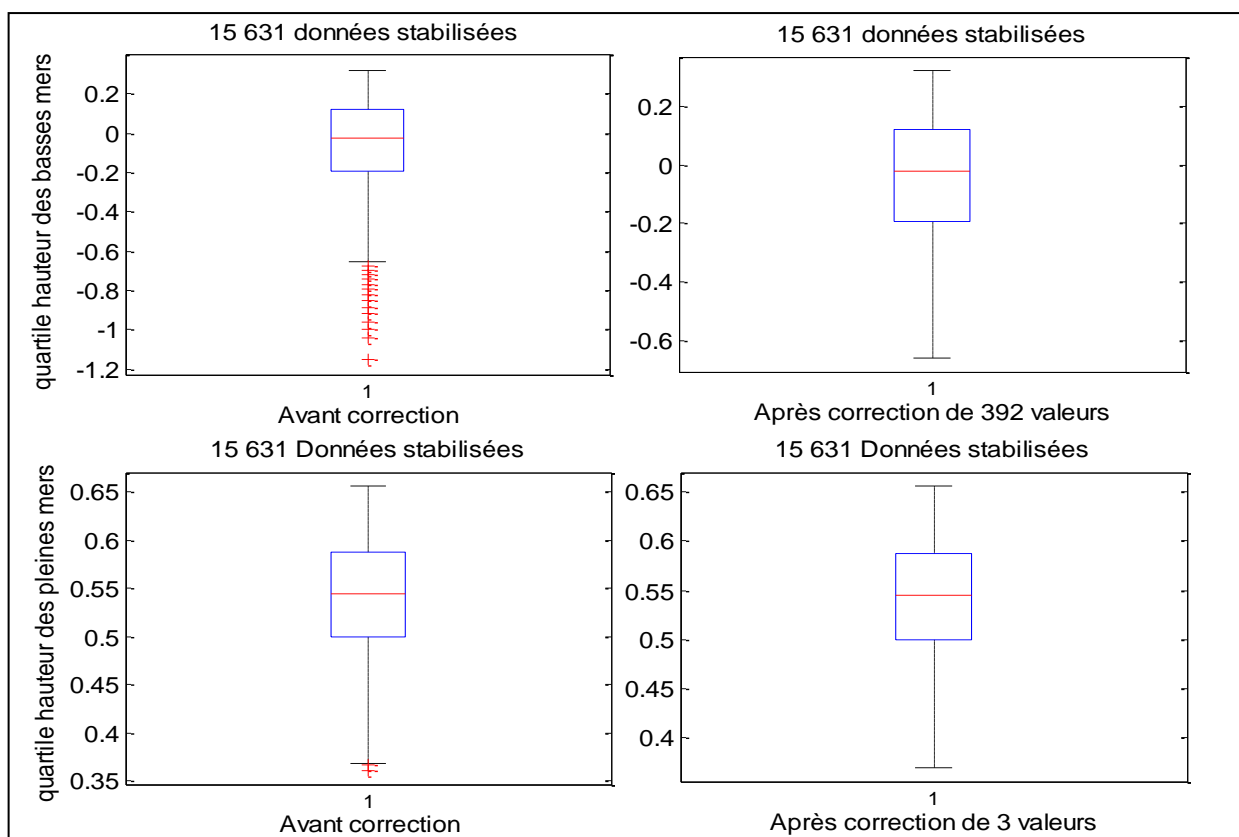


Figure 3.7 : Boîtes à moustache avant et après correction des points aberrants pour les données stabilisées des hauteurs des pleines et basses mers

Les résultats de la figure 3.7 montrent que nous avons dû corriger 3 points aberrants pour blanchir les données stabilisées sur les pleines mers, et 392 points aberrants pour blanchir les données stabilisées sur les basses mers.

Pour la mise en œuvre des modèles, nous avons ramené la taille de chaque échantillon (pleines mers, basses mers) à 15631.

Les données restantes vont servir à vérifier la qualité de la prévision.

En particulier, pour les pleines mers, nous avons gardé 1,58 % des données (soit 251 = 15882 - 15631 données), et pour les basses mers 1,68 % des données (soit 268 = 15899 - 15631 données).

3.2.7. Estimation de la saisonnalité S

3.2.7.1. Estimation de la saisonnalité S par corrélogramme ou la méthode de BLACKMAN-TUCKEY

La méthode du corélogramme aboutit à la construction de la figure 3.8.

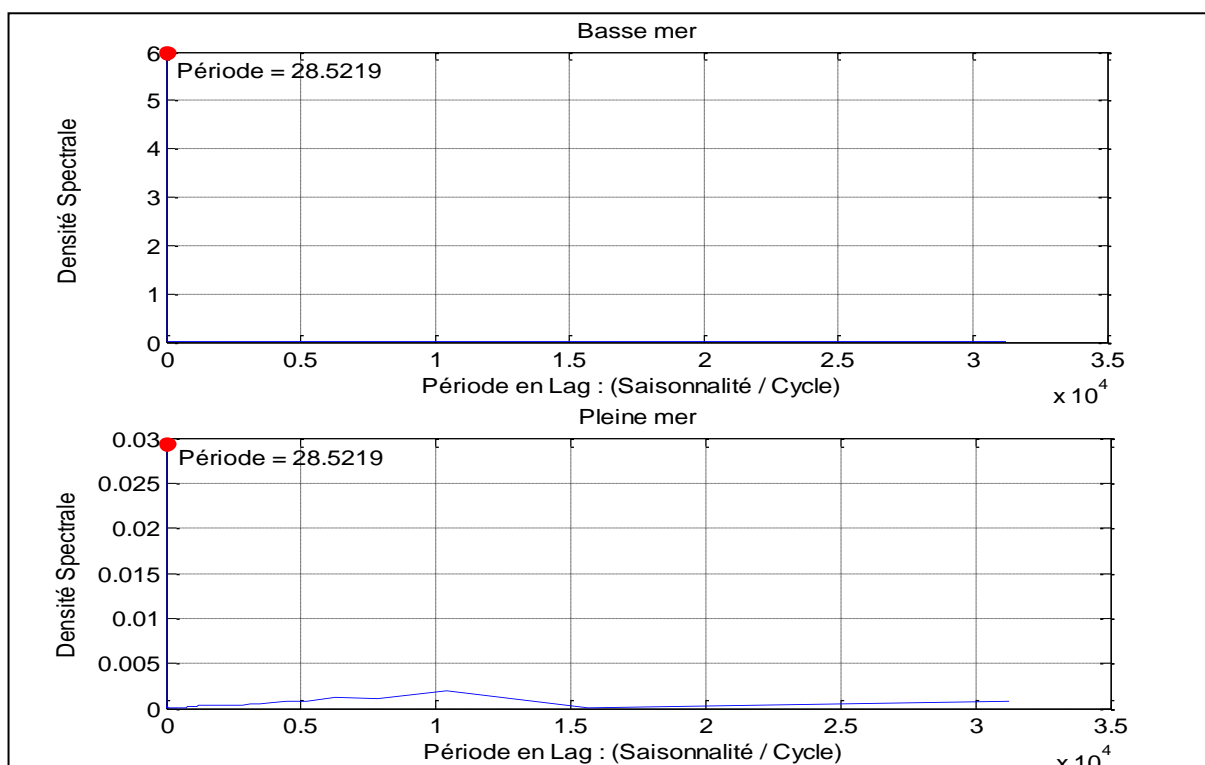


Figure 3.8 : Corrélogramme des hauteurs des pleines et des basses mers stabilisées

Les résultats de la de la figure 3.8 montrent que la réponse obtenue par la méthode du corrélogramme est unique et correspond à la saisonnalité S égale à 28,5219 Lags pour les pleines mers et pour les basses mers. Cette méthode nous suggère donc de prendre la saisonnalité S égale à 28,5219 Lags pour les séries des hauteurs des pleines et des basses mers à Hell-Ville.

3.2.7.2. Estimation de la saisonnalité S par périodogramme moyenné

Nous avons mis en œuvre les périodogrammes moyennés des hauteurs des pleines et basses mers en subdivisant successivement l'échantillon initial stabilisé en 1, 6 et 10 blocs de même taille et nous sommes arrivés aux résultats des figures 3.9 et 3.10.

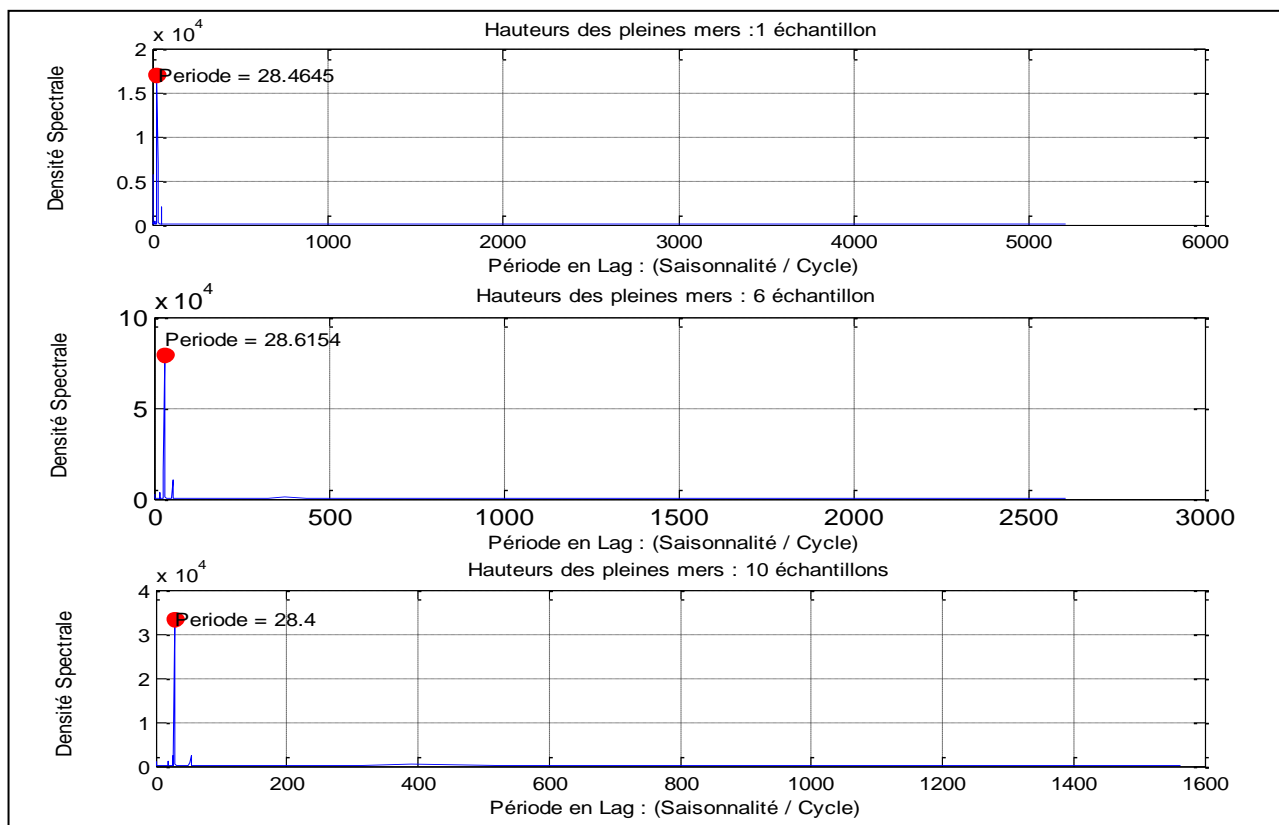


Figure 3.9 : Périodogrammes moyennés des hauteurs des pleines mers stabilisées

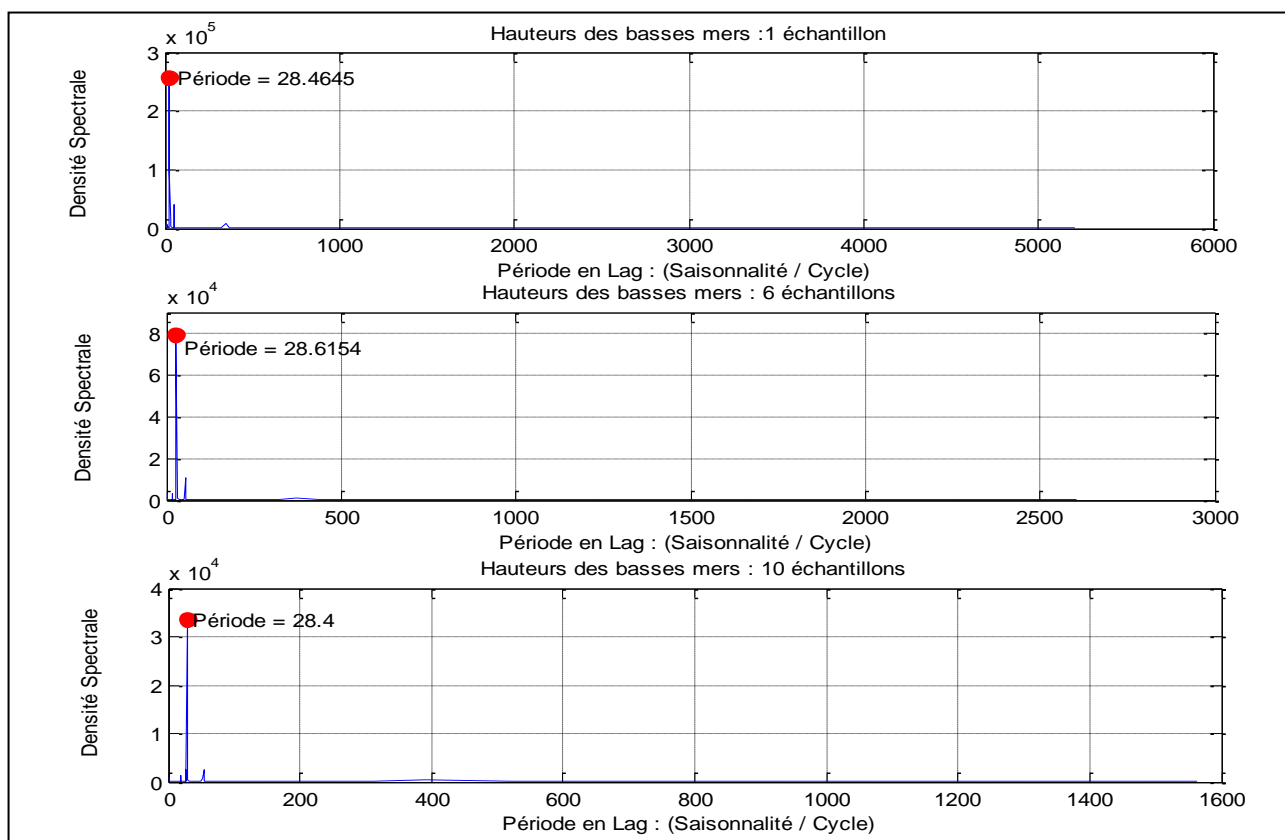


Figure 3.10 : Périodogrammes moyennés des hauteurs des basses mers stabilisées

Les résultats des figures 3.9 et 3.10 offrent trois choix identiques pour la saisonnalité S des hauteurs des pleines et des basses mers. Pour les deux chroniques, nous avons trouvé une saisonnalité S , décrite par les trois valeurs proches et cohérentes suivantes : 28,4 ; 28,4645 et 28,6154 Lags.

3.2.7.3. Estimation de la saisonnalité par périodogramme par la Méthode de l'Entropie Maximale

a) Cas des hauteurs des marées sur les basses mers

Nous avons tracé successivement les périodogrammes des hauteurs des basses mers en variant l'ordre q du modèle autorégressif de BURG de 1 à 32.

D'après la figure 3.11, les modèles autorégressifs de BURG d'ordre inférieur à 5 ne permettent pas de rechercher la saisonnalité S par la méthode de l'Entropie Maximale, car nous voyons que la densité spectrale en bleu sur cette figure ne dépasse pas le seuil de signification $\alpha = 5\%$ en rouge. Nous avons donc remis en œuvre la méthode de l'entropie maximale pour trouver la saisonnalité S pour des modèles de BURG d'ordre supérieur à 4.

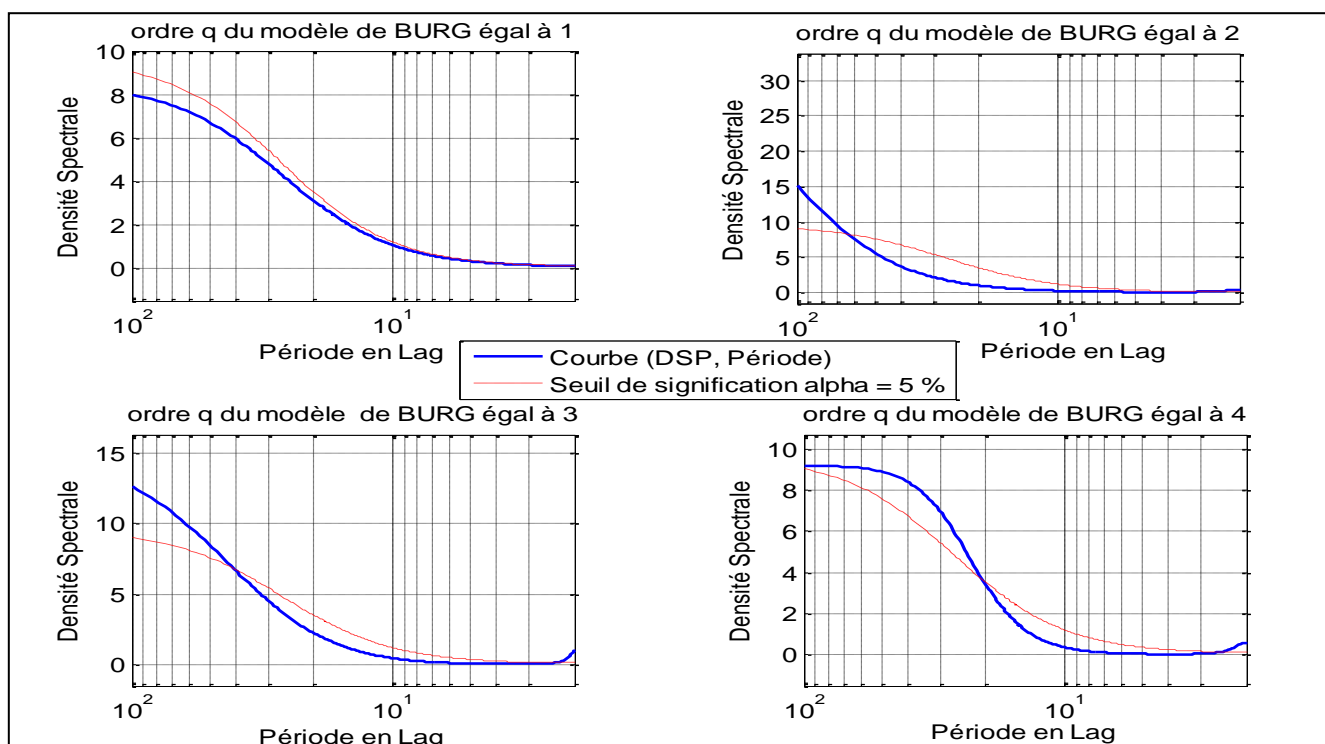


Figure 3.11 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 1 à 4

En variant l'ordre q du modèle autorégressif de BURG de 5 à 32, (28 valeurs de q), nous avons constaté sur les figures 3.12 à 3.18 que les saisonnalités S calculées par la

méthode de l'entropie maximale sont toutes significatives car les densités spectrales en bleues sur ces figures ont des pics situés en dessus du seuil de signification $\alpha = 5\%$ en rouge.

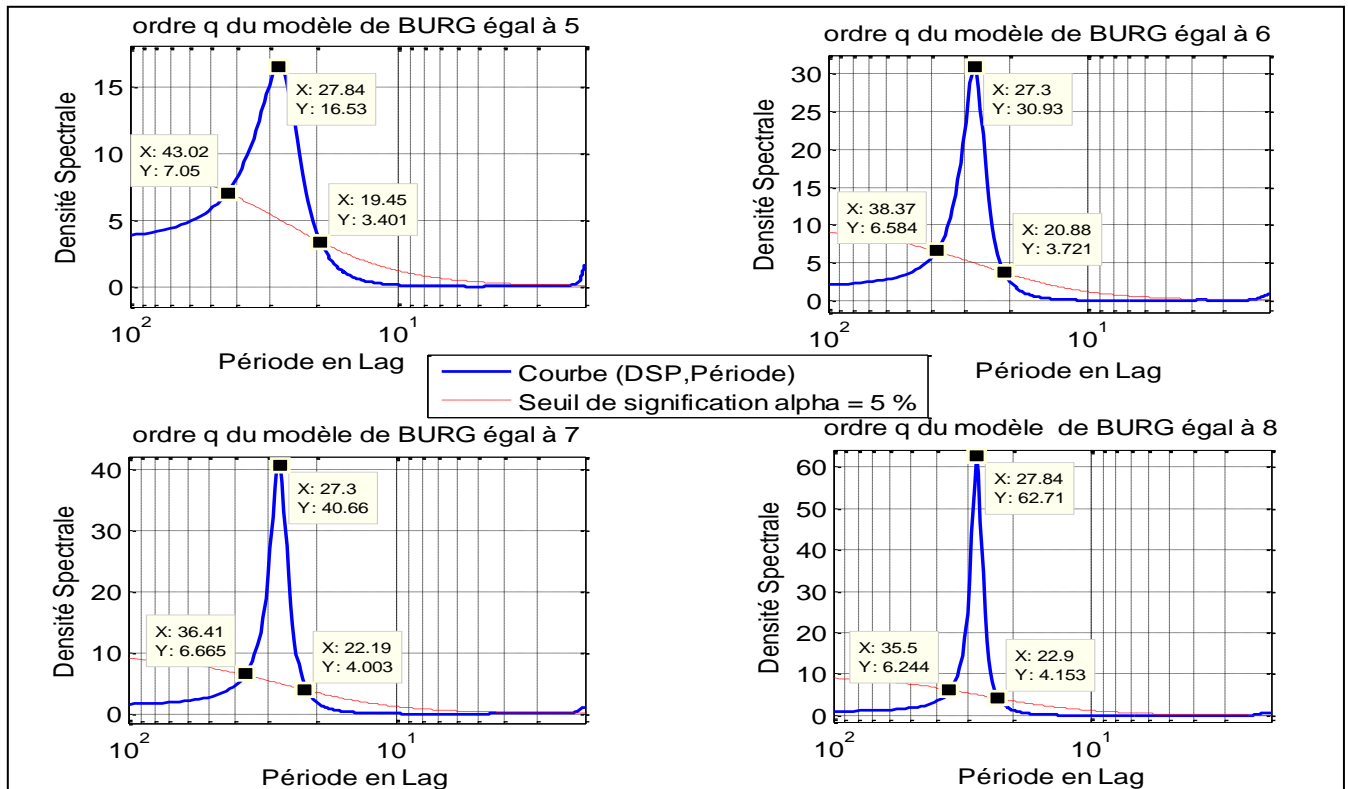


Figure 3.12 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 5 à 8

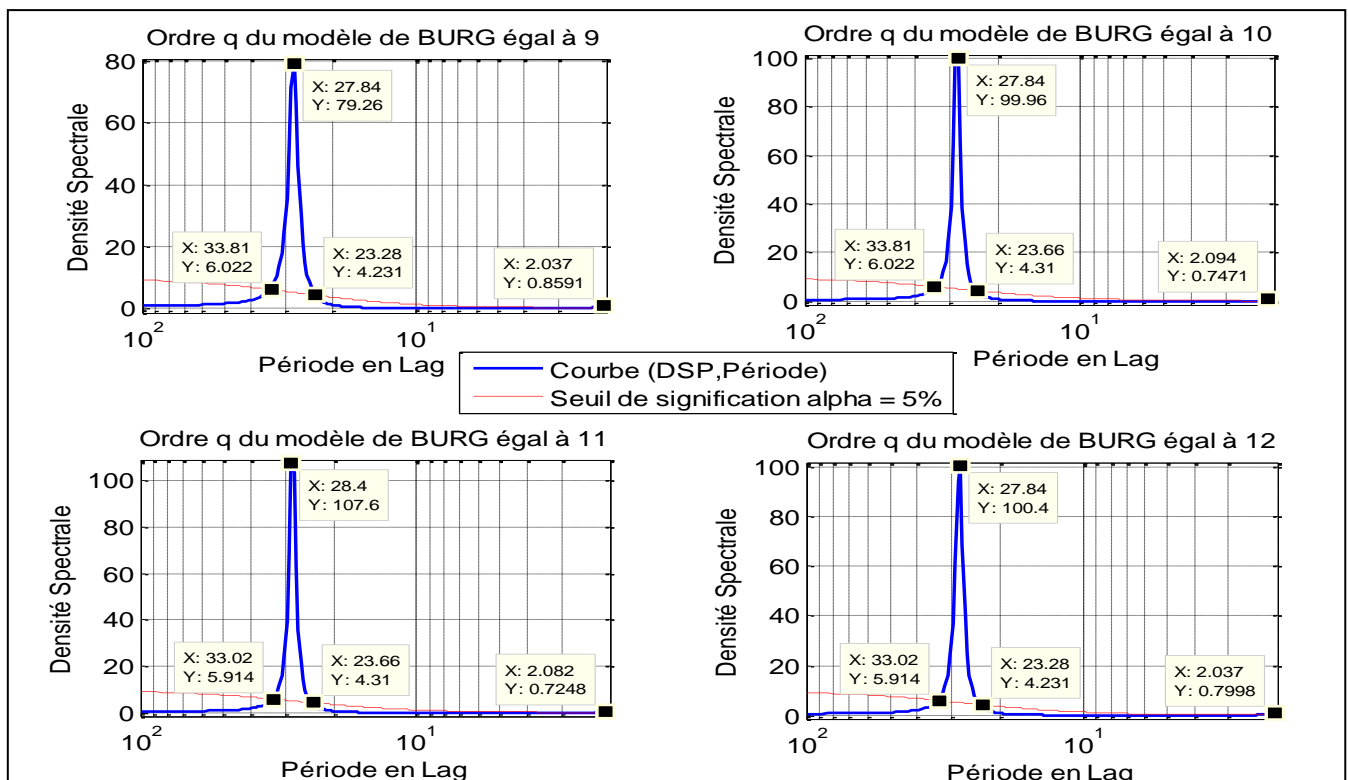


Figure 3.13 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 9 à 12

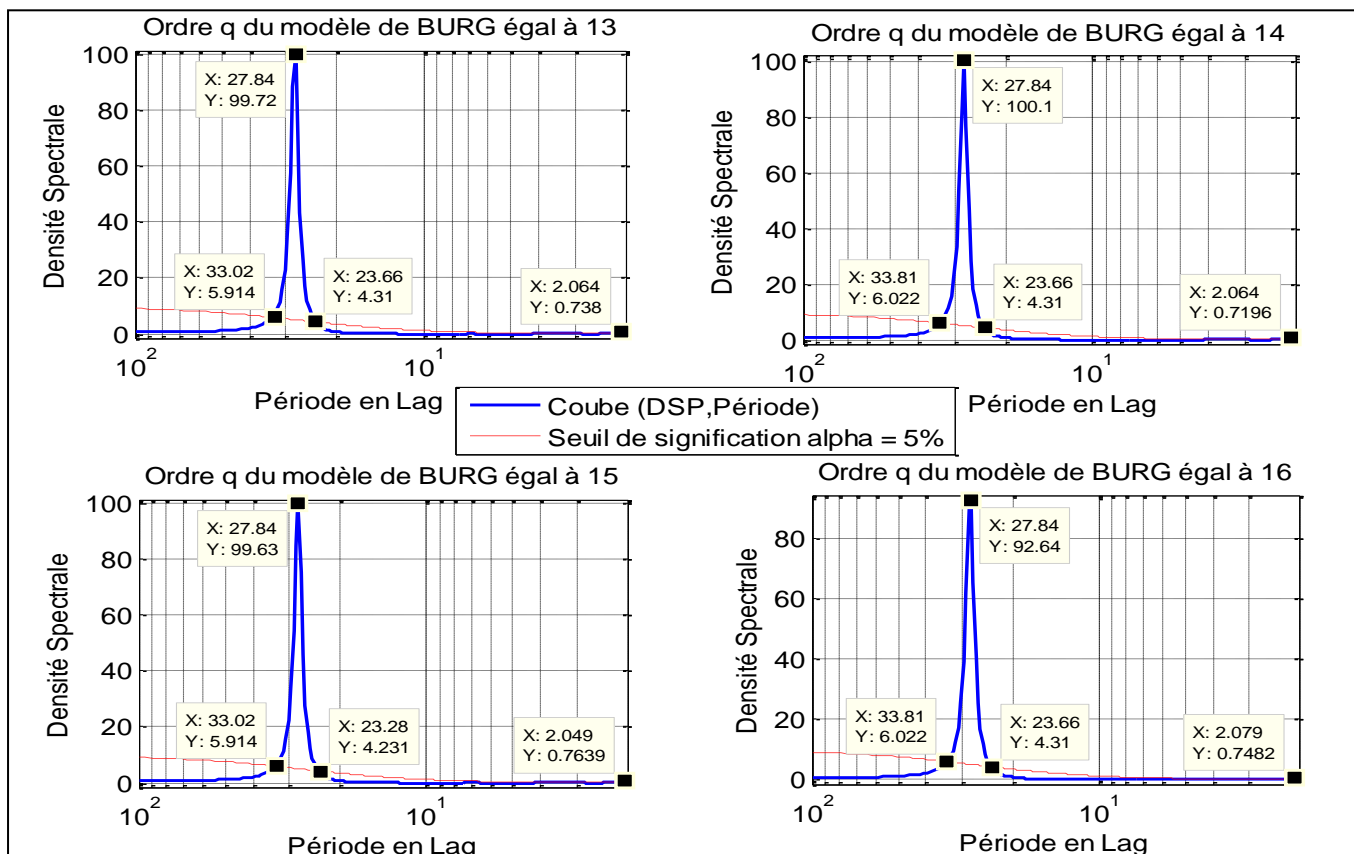


Figure 3. 14 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 13 à 16

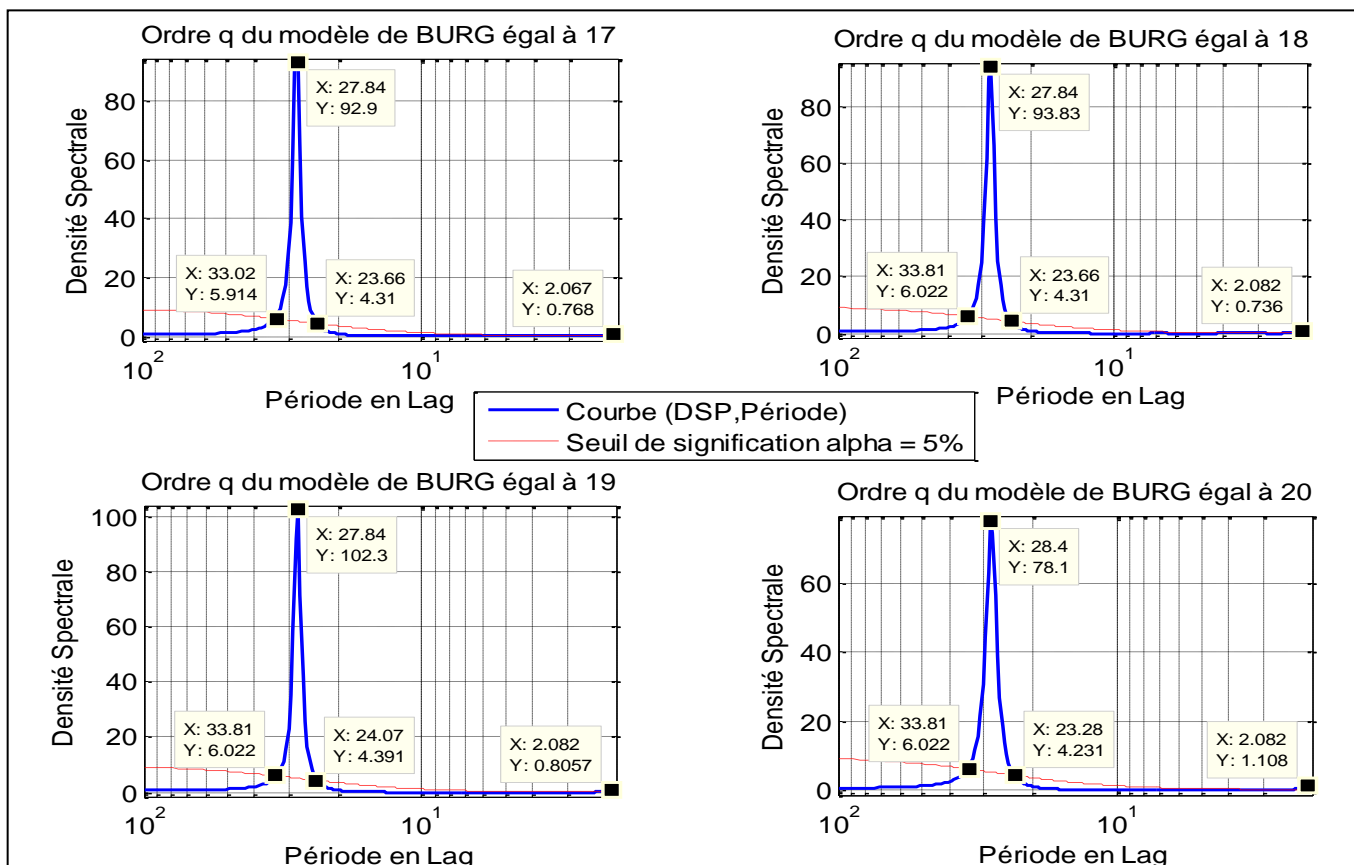


Figure 3. 15 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 17 à 20

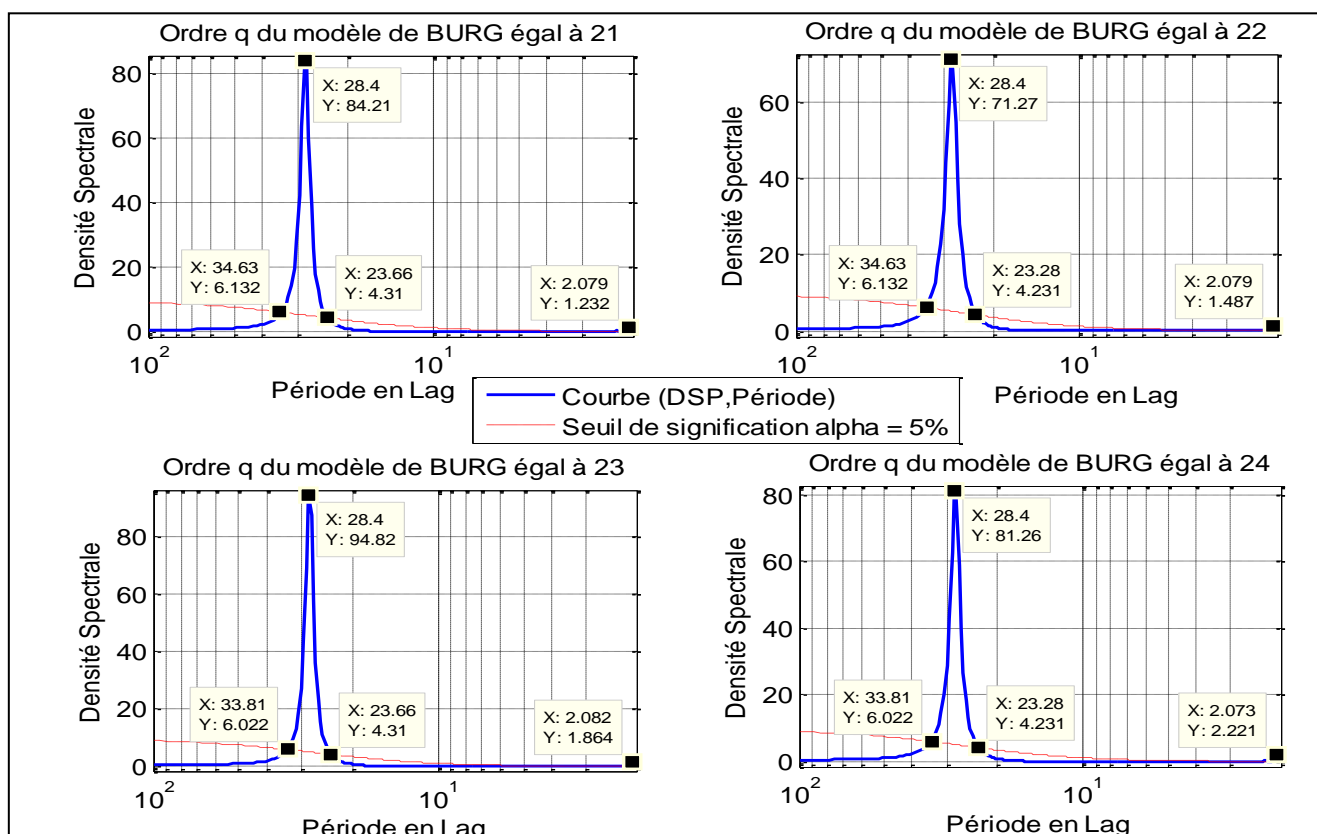


Figure 3.16 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 21 à 24

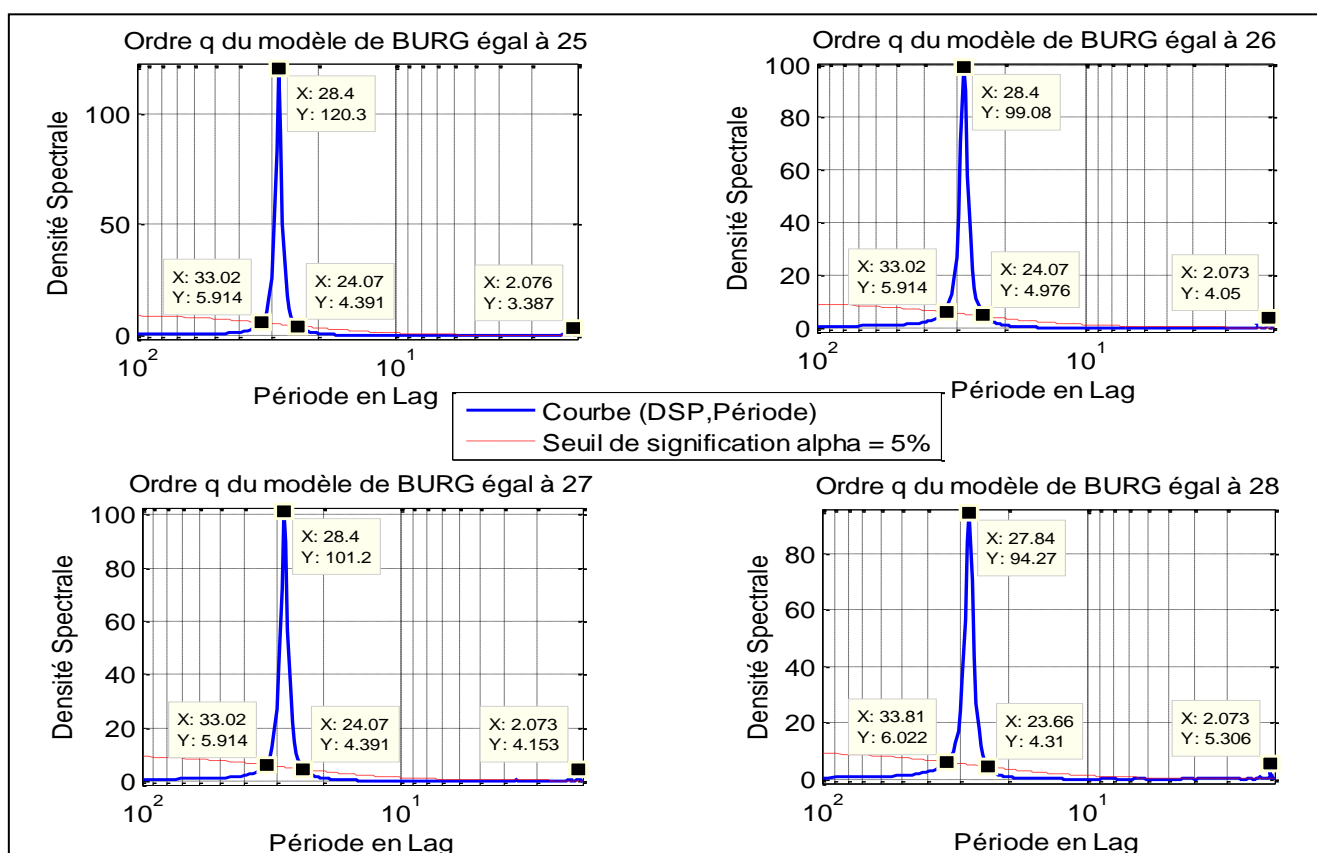


Figure 3.17 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 25 à 28

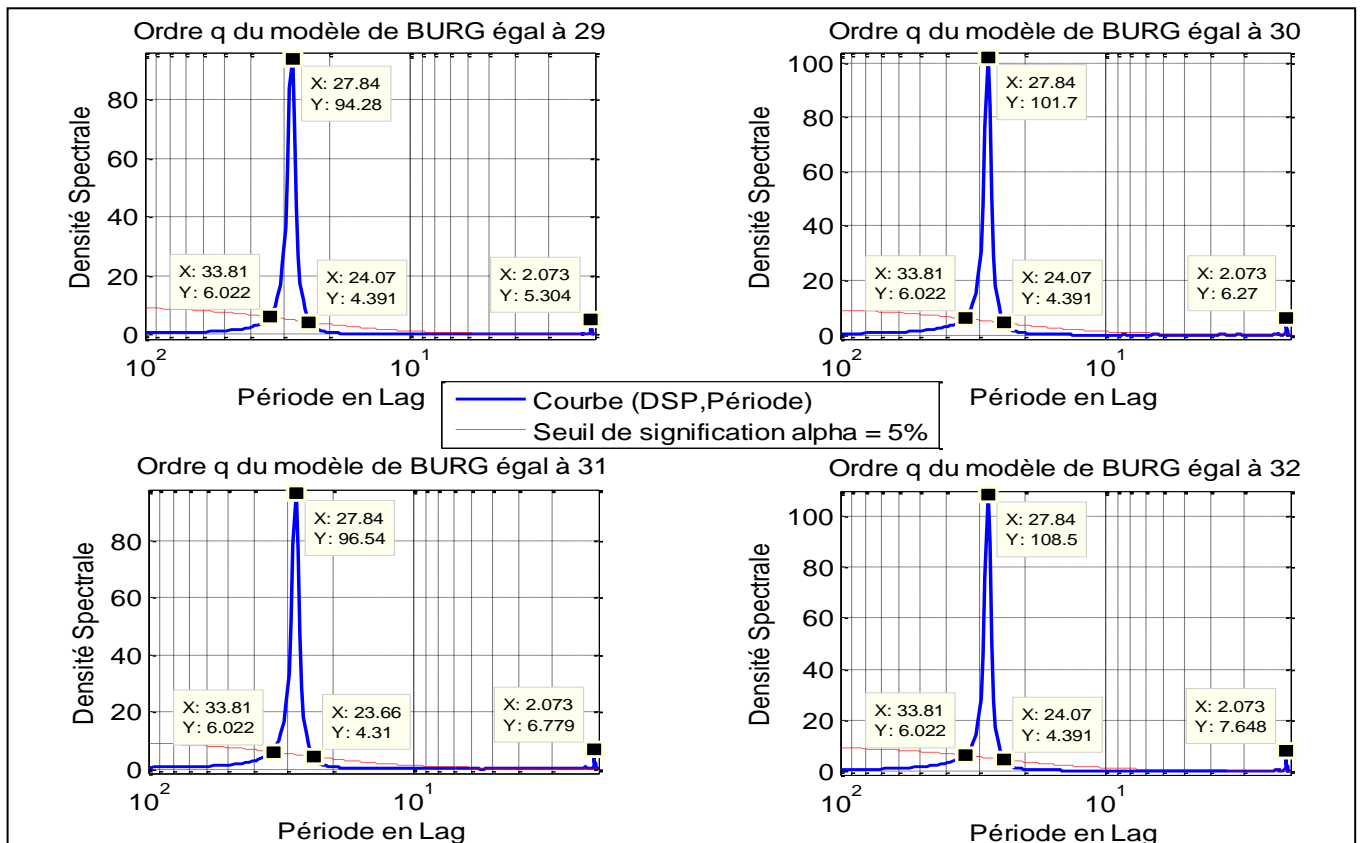


Figure 3.18 : Méthode de l'entropie maximale des hauteurs des basses mers pour q variant de 29 à 32

D'après les figures 3.12 à 3.18, nous avons trouvé deux pseudo-périodes puisque deux pics sont visibles sur toutes les courbes.

Pour les 28 cas étudiés, en comparant les valeurs de la première pseudo-période trouvées sur les figures 3.12 à 3.18, elles se réduisent finalement à trois valeurs :

- S1 égal à 27,30 Lags ;
- S2 égal à 27,84 Lags;
- S3 égal à 28,40 Lags.

Pour la deuxième pseudo-période, avec le même raisonnement, les valeurs trouvées pour la saisonnalité sont au nombre de douze :

- S4 égal à 2,006 Lags;
- S5 égal à 2,020 Lags;
- S6 égal à 2,034 Lags;
- S7 égal à 2,037 Lags;
- S8 égal à 2,049 Lags;
- S9 égal à 2,064 Lags;

- S10 égal à 2,067 Lags;
- S11 égal à 2,073 Lags;
- S12 égal à 2,076 Lags;
- S13 égal à 2,079 Lags;
- S14 égal à 2,082 Lags;
- S15 égal à 2,094 Lags.

En notant H1, H2, ..., H15 les entropies correspondant à la probabilité d'apparition de S1, S2, ..., S15 pour chaque valeur de q fixée entre 19 à 32 pour la première pseudo-période et entre 28 à 32 pour la deuxième pseudo-période, nous avons pu dresser les tableaux 3.5 et 3.6.

Tableau 3 . 5 : Entropie maximale de la première pseudo-période pour q variant de 19 à 32

	Ordre q du modèle autorégressif de BURG														
Période	32	31	30	29	28	27	26	25	24	23	22	21	20	19	Entropie
S1=27,30	0,082	0,084	0,086	0,088	0,090	0,092	0,095	0,097	0,100	0,103	0,106	0,109	0,113	0,117	H1
S2=27,84	0,132	0,135	0,138	0,141	0,144	0,147	0,144	0,139	0,133	0,126	0,117	0,107	0,094	0,078	H2
S3=28,40	0,158	0,159	0,159	0,160	0,160	0,159	0,160	0,159	0,157	0,153	0,145	0,133	0,113	0,078	H3

Tableau 3 . 6 : Entropie maximale de la deuxième pseudo-période pour q variant de 28 à 32

	Valeur de l'entropie pour l'ordre q du modèle autorégressif de BURG égal à					
Période	32	31	30	29	28	Entropie
S4 = 2,006	0,0517	0,0530	0,0544	0,0559	0,0575	H4
S5 = 2,020	0,0517	0,0530	0,0544	0,0559	0,0575	H5
S6 = 2,034	0,0517	0,0530	0,0544	0,0559	0,0575	H6
S7 = 2,037	0,0819	0,0837	0,0857	0,0878	0,0899	H7
S8 = 2,049	0,0819	0,0837	0,0857	0,0878	0,0899	H8
S9 = 2,064	0,0819	0,0837	0,0857	0,0878	0,0899	H9
S10 = 2,067	0,0517	0,0530	0,0544	0,0559	0,0575	H10
S11 = 2,073	0,1554	0,1520	0,1470	0,1398	0,1297	H11
S12 = 2,076	0,0517	0,0530	0,0544	0,0559	0,0575	H12
S13 = 2,079	0,1039	0,1060	0,1082	0,1105	0,1129	H13
S14 = 2,082	0,1336	0,1356	0,1377	0,1398	0,1419	H14
S15 = 2,094	0,0517	0,0530	0,0544	0,0559	0,0575	H15

Les parties en jaune des tableaux 3.5 et 3.6 représentent l'entropie maximale. Elles nous recommandent de choisir pour la première pseudo-période la valeur 28,40 Lags et pour la deuxième la valeur 2,073 Lags.

Finalement la méthode de l'entropie maximale nous suggère donc deux saisonnalités pour les hauteurs des basses mers :

- une première saisonnalité S_p égale à 28,40 Lags
- et une deuxième saisonnalité S_d égale à 2,073 Lags

Les résultats de la ligne S3H3 permettent en outre d'avancer que pour les basses mers l'ordre optimal q du modèle autorégressif de BURG est inférieur ou égal à 21.

b) Cas des hauteurs des marées sur les pleines mers

Pour l'analyse des hauteurs des pleines mers, nous avons adopté le même raisonnement que pour l'analyse des hauteurs des basses pour retrouver la saisonnalité S .

Nous avons donc tracé les courbes de la densité spectrale en fonction de la période pour chaque ordre q du modèle autorégressif de BURG variant de 1 à 32 et nous avons obtenu les résultats des figures 3.19 à 3.26.

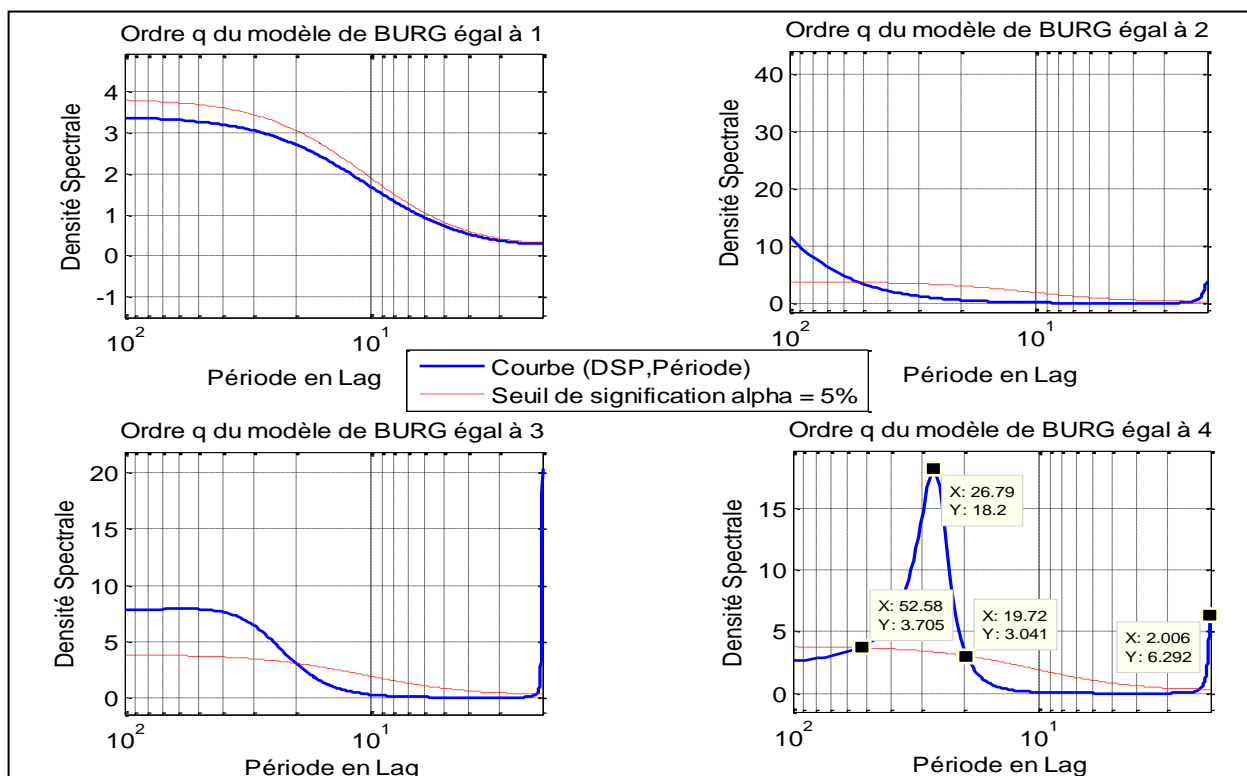


Figure 3.19 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 1 à 4

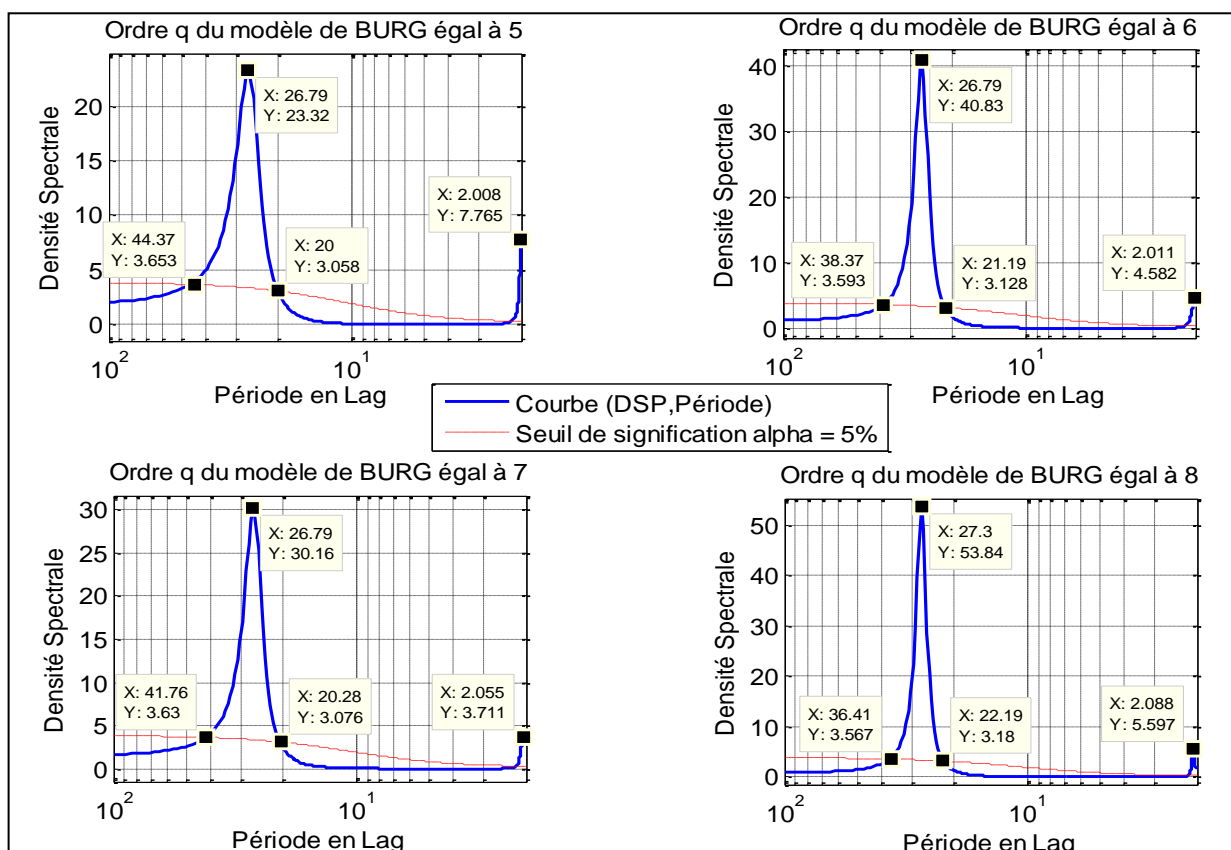


Figure 3.20 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 5 à 8

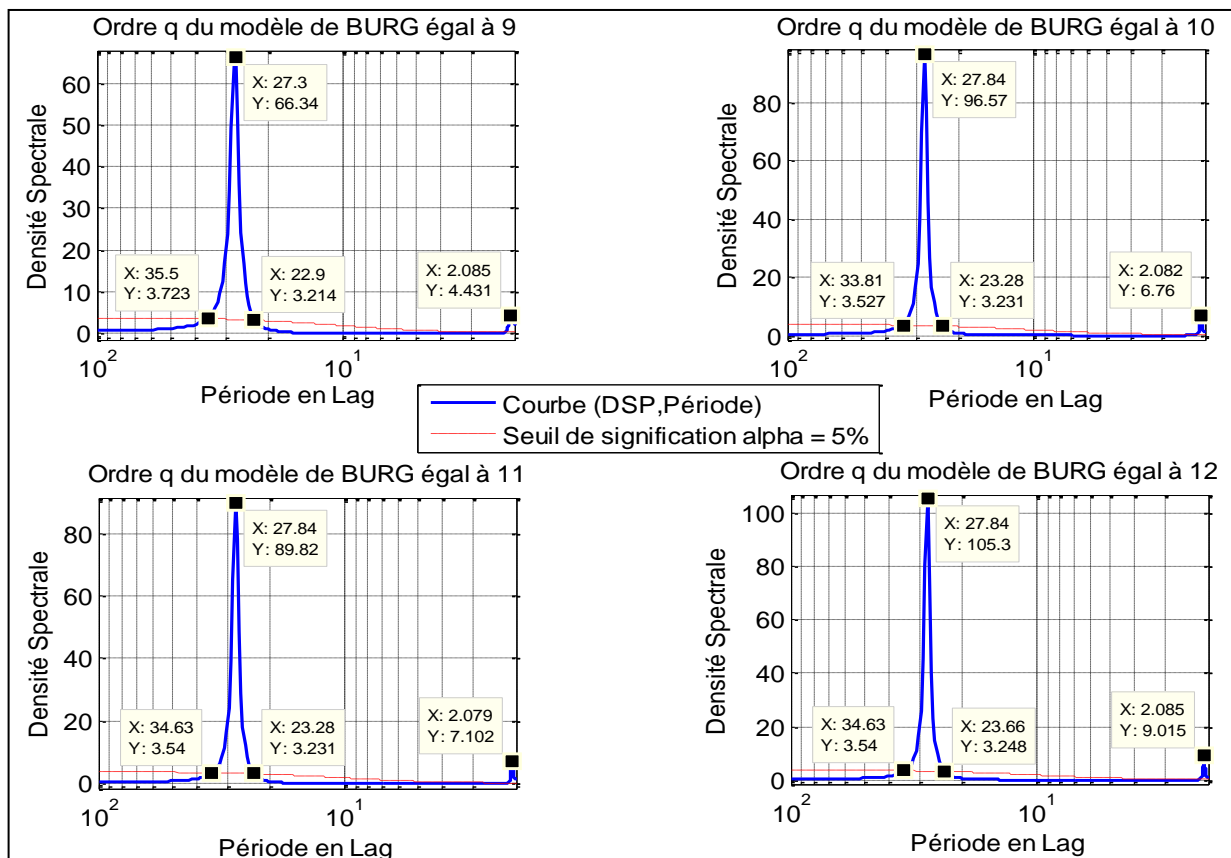


Figure 3.21 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 9 à 12

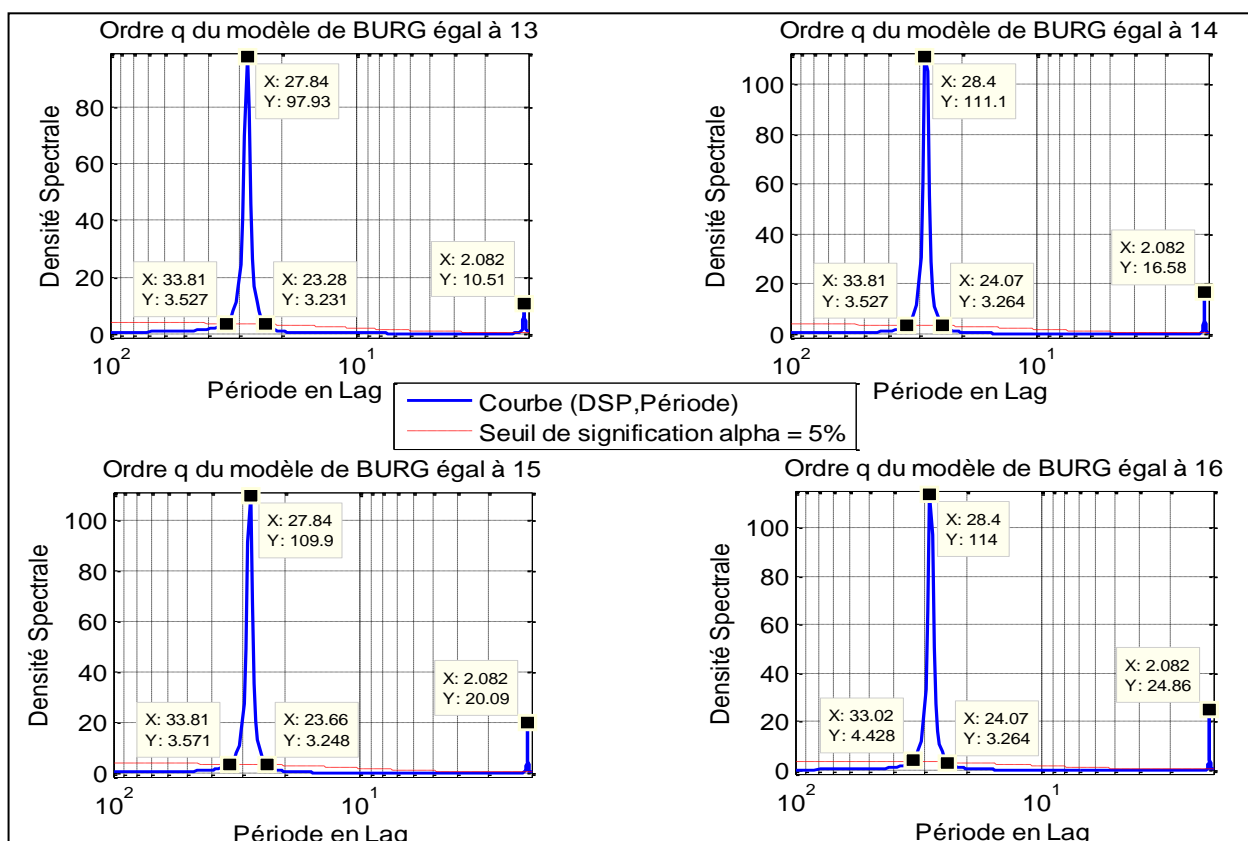


Figure 3.22 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 13 à 16

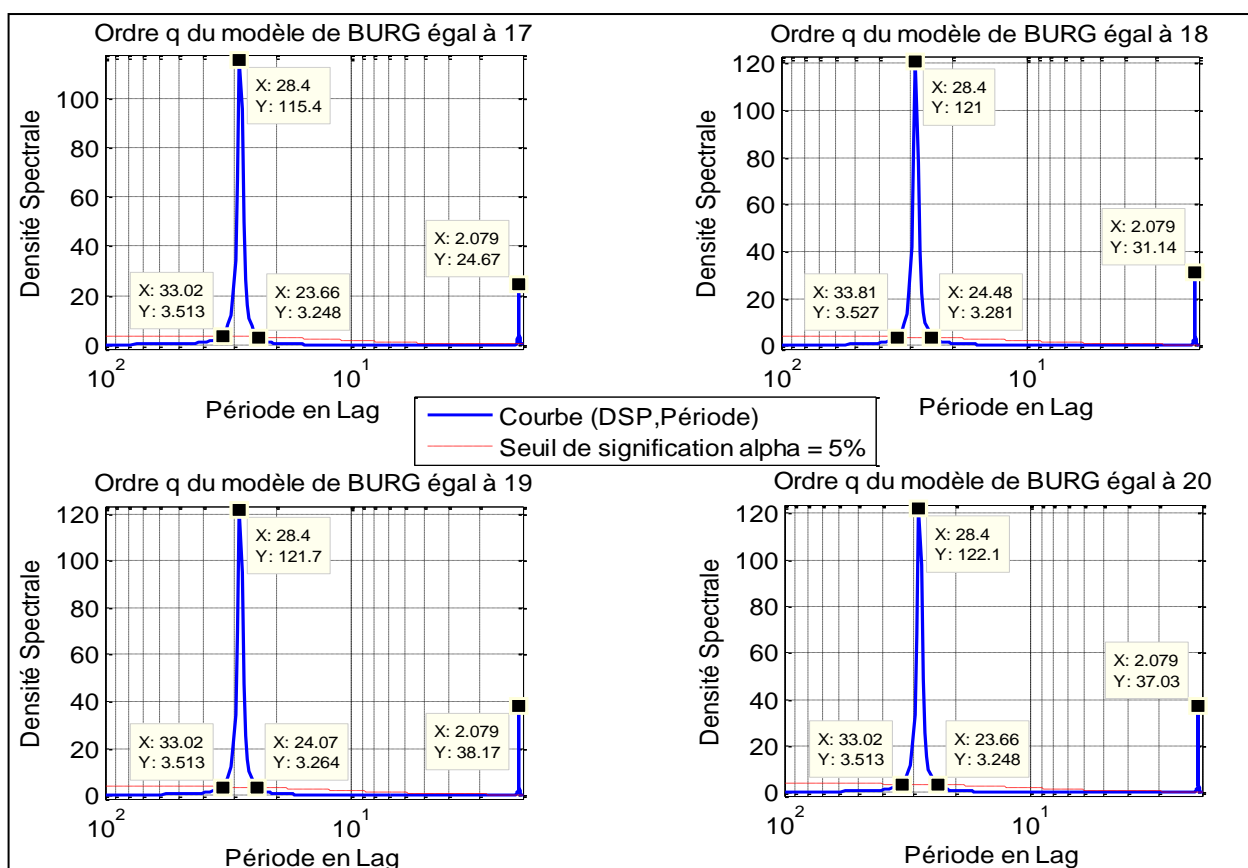


Figure 3.23 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 17 à 20

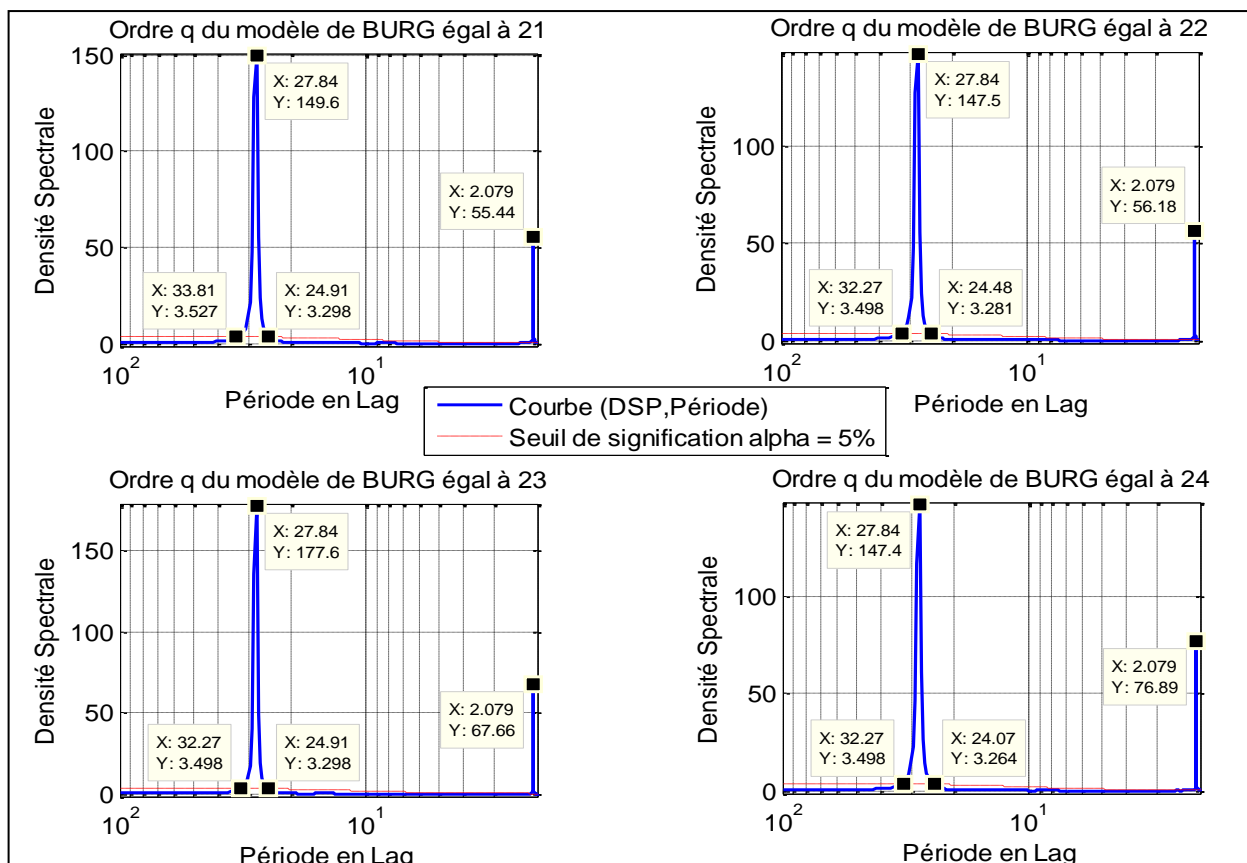


Figure 3.24 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 21 à 24

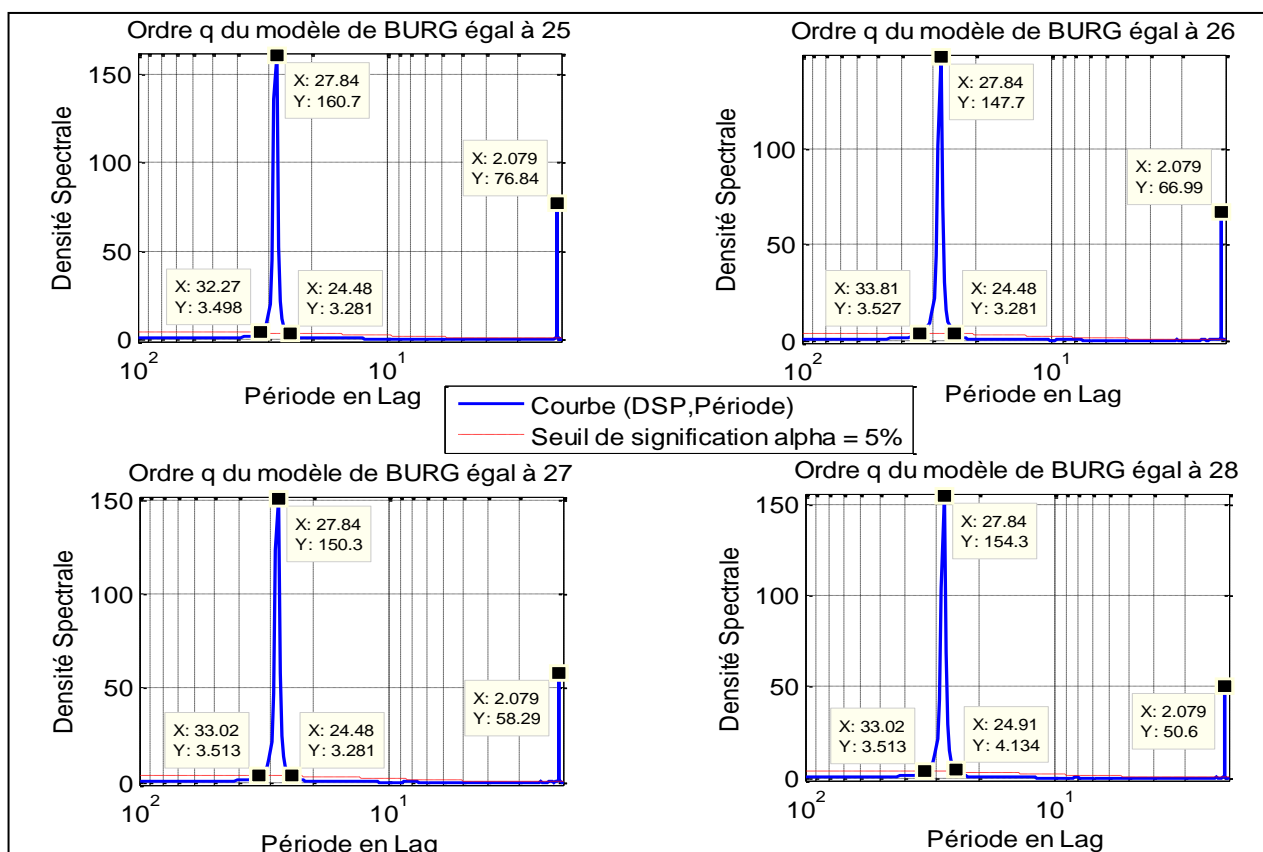


Figure 3.25 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 25 à 28

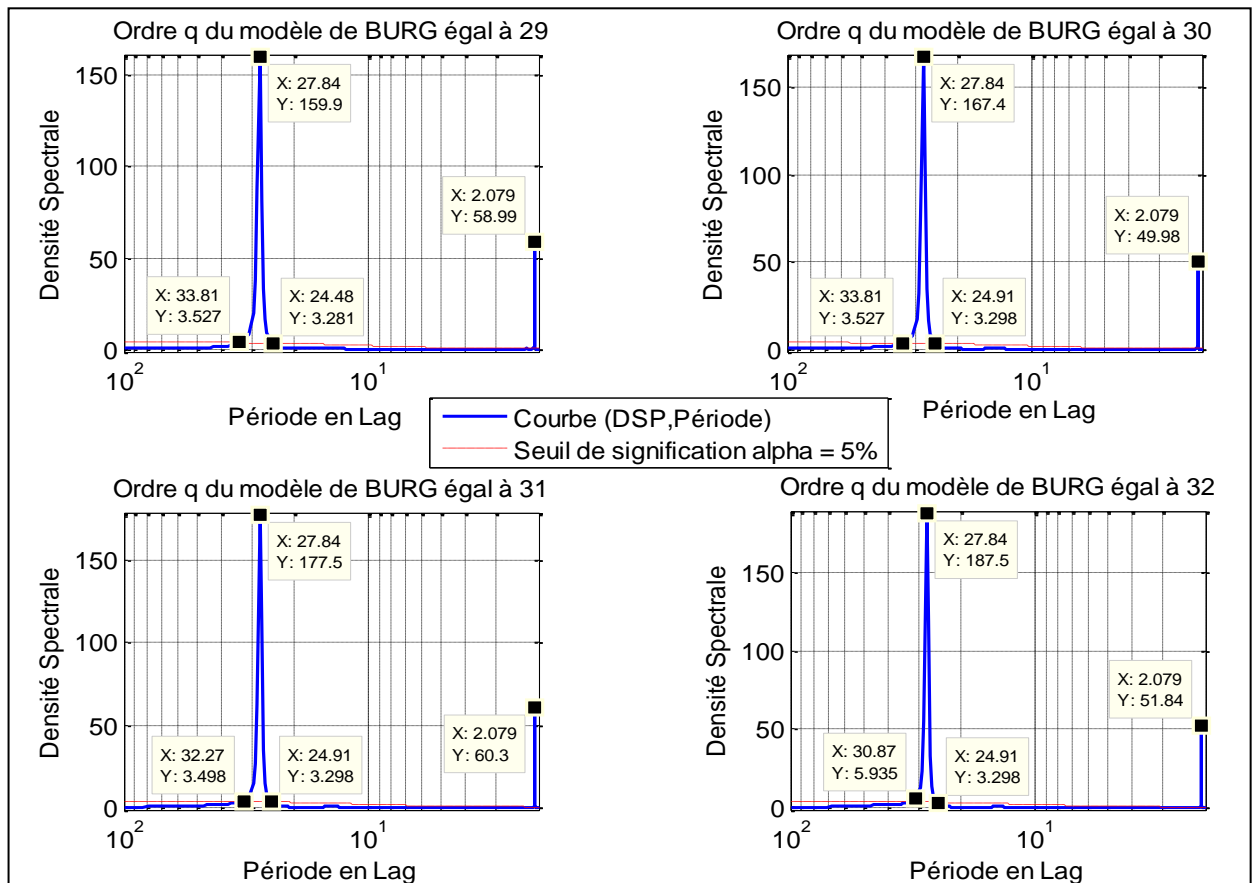


Figure 3.26 : Méthode de l'entropie maximale des hauteurs des pleines mers pour q variant de 29 à 32

D'après les figures 3.19 à 3.26, nous avons constaté l'existence de deux pseudo-périodes puisque deux pics sont toujours visibles sur toutes les courbes.

En comparant les valeurs de la première pseudo-période trouvées pour les 29 cas étudiés sur les figures 3.19 à 3.26, elles se réduisent à quatre valeurs :

- S1' égal à 26,79 Lags;
- S2' égal à 27,30 Lags;
- S3' égal à 27,84 Lags;
- S4' égal à 28,40 Lags.

Pour la deuxième pseudo-période, les valeurs trouvées sont au nombre de douze :

- S5' égal à 2,006 Lags;
- S6' égal à 2,008 Lags;
- S7' égal à 2,011 Lags;
- S8' égal à 2,055 Lags;
- S9' égal à 2,079 Lags;
- S10' égal à 2,082 Lags;

- S_{11}' égal à 2,085 Lags;
- S_{12}' égal à 2,088 Lags.

En notant $H_1', H_2', \dots, H_{12}'$ les entropies correspondant à la probabilité d'apparition de $S_1', S_2', \dots, S_{12}'$ pour chaque valeur de q fixée entre 25 à 32 pour la première pseudo-période, et entre 19 à 32 pour la deuxième pseudo-période, nous avons obtenu finalement les résultats des tableaux 3.7 et 3.8.

Tableau 3.7 : Entropie maximale de la première pseudo-période pour q variant de 25 à 32

Période	Ordre q du modèle autorégressif de BURG								Entropie
	32	31	30	29	28	27	26	25	
$S_1' = 26,79$	0,1187	0,1207	0,1229	0,1251	0,1273	0,1297	0,1321	0,1346	H_1'
$S_2' = 27,30$	0,0801	0,0819	0,0837	0,0857	0,0878	0,0899	0,0922	0,0947	H_2'
$S_3' = 27,84$	0,1360	0,1389	0,1418	0,1448	0,1477	0,1505	0,1532	0,1556	H_3'
$S_4' = 28,40$	0,1416	0,1434	0,1452	0,1470	0,1487	0,1505	0,1522	0,1539	H_4'

Tableau 3.8 : Entropie maximale de la deuxième pseudo-période pour q variant de 19 à 32

Période	Valeur de l'entropie pour l'ordre q du modèle autorégressif de BURG égal à :														Entropie
	32	31	30	29	28	27	26	25	24	23	22	21	20	19	
$S_5' = 2,006$	0,050	0,052	0,053	0,054	0,056	0,058	0,059	0,061	0,063	0,065	0,067	0,070	0,072	0,075	H_5'
$S_6' = 2,008$	0,050	0,052	0,053	0,054	0,056	0,058	0,059	0,061	0,063	0,065	0,067	0,070	0,072	0,075	H_6'
$S_7' = 2,011$	0,050	0,052	0,053	0,054	0,056	0,058	0,059	0,061	0,063	0,065	0,067	0,070	0,072	0,075	H_7'
$S_8' = 2,055$	0,050	0,052	0,053	0,054	0,056	0,058	0,059	0,061	0,063	0,065	0,067	0,070	0,072	0,075	H_8'
$S_9' = 2,079$	0,136	0,139	0,142	0,145	0,148	0,151	0,153	0,156	0,158	0,159	0,160	0,159	0,156	0,151	H_9'
$S_{10}' = 2,082$	0,132	0,134	0,136	0,138	0,140	0,142	0,144	0,146	0,148	0,151	0,153	0,155	0,156	0,158	H_{10}'
$S_{11}' = 2,085$	0,080	0,082	0,084	0,086	0,088	0,090	0,092	0,095	0,097	0,100	0,103	0,106	0,109	0,113	H_{11}'
$S_{12}' = 2,088$	0,050	0,052	0,053	0,054	0,056	0,058	0,059	0,061	0,063	0,065	0,067	0,070	0,072	0,075	H_{12}'

Les parties en jaune des tableaux 3.7 et 3.8 représentent l'entropie maximale. Elles nous recommandent de choisir pour la première pseudo-période la valeur 28,40 Lags et pour la deuxième la valeur 2,079 Lags.

La méthode de l'entropie maximale nous recommande alors deux saisonnalités pour les hauteurs des pleines mers :

- une première saisonnalité S_p' égale à 28,40 Lags,
- et une deuxième saisonnalité S_d' égale à 2,079 Lags.

Les résultats de la ligne $S_4'H_4'$ permettent en outre d'avancer que pour les pleines mers, l'ordre optimal q du modèle autorégressif de BURG est inférieur ou égal à 28.

En résumé, nous avons obtenu à partir **de la méthode du corrélogramme** la même saisonnalité S égale à 28,5219 Lags pour les basses et pleines mers.

A partir **de la méthode de la périodogramme moyennée**, nous avons les mêmes valeurs de la saisonnalité S égales à 28,4 Lags, 28,4645 Lags ou 28,6154 Lags pour les pleines et basses mers.

Et à partir **de la méthode de l'entropie maximale**, nous avons trouvé **deux saisonnalités pour chaque chronique**.

En particulier, pour les basses mers, la première saisonnalité S est égale à 28,40 Lags et la deuxième à 2,073 Lags et pour les pleines mers, la première saisonnalité S est égale à 28,40 Lags et la deuxième à 2,079 Lags

3.2.8. Saisonnalité S en terme journalier

En passant à la moyenne de la différence en jour entre deux temps de prélèvement consécutifs, nous avons trouvé les résultats suivants :

pour les pleines mers :

$$1 \text{ Lag} = 0,51753136 \text{ jour}$$

pour les basses mers :

$$1 \text{ Lag} = 0,51753154 \text{ jour}$$

Ainsi, les résultats du paragraphe précédent deviennent en terme journalier :

- **pour la méthode du corrélogramme :**
 - ⇒ la saisonnalité S des basses mers de 28,5219 Lags est l'équivalent de 14,76098 jours ;
 - ⇒ et la saisonnalité S des pleines mers de 28,5219 Lags est l'équivalent de 14,76097 jours.
- **pour la méthode de la périodogramme moyennée :**
 - ⇒ les saisonnalités S des basses mers de 28,4 ; 28,4645 et 28,6154 Lags correspondent respectivement à 14,69789 ; 14,73127 et 14,80937 jours.

⇒ et les saisonnalités S des pleines mers de 28,4 ; 28,4645 et 28,6154 Lags correspondent respectivement à 14,69789 ; 14,73127 et 14,80936 jours.

• Pour la méthode de l'entropie maximale :

⇒ les deux saisonnalités S des basses mers de 28,40 et 2,073 Lags correspondent respectivement à 14,69789 et 1,07284 jours ;

⇒ et les deux saisonnalités S des pleines mers de 28,40 et 2,079 Lags correspondent respectivement à 14,69789 et 1,07594 jours.

Après la mise en œuvre du test de Student, nous avons constaté d'après les résultats du tableau 3.2.9 une concordance parfaite entre les valeurs de S obtenues par nos méthodes et celles calculées par les méthodes traditionnelles de Newton ou de Laplace (Cf. tableau 3.9).

Tableau 3 . 9 : Test d'égalité de la saisonnalité obtenue par les méthodes traditionnelles et par nos méthodes

	Méthode traditionnelle : Newton, Laplace.	Méthodes proposées : Corrélogramme, périodogramme moyenné, méthode de l'entropie.	Intervalle de confiance et paramètres	Observations
	Saisonnalité : Morte eau, vive eau.			
Pleine mer	S = 14,765 jours	$SN_1 = \begin{cases} S = 14,76097 \text{ jours} \\ S = 14,69789 \text{ jours} \\ S = 14,73127 \text{ jours} \\ S = 14,80936 \text{ jours} \end{cases}$	Moyenne $SN_1 = 14,74988$ Ecart type $SN_1 = 0,05461$ $t_{1-\frac{\alpha}{2}} = 3,182$ S et $SN_1 \in [14,6629, 14,8367]$	Il y a concordance parfaite entre S et les valeurs de SN_1 au seuil $\alpha = 0,05$
Basse mer	S = 14,765 jours	$SN_2 = \begin{cases} S = 14,76098 \text{ jours} \\ S = 14,69789 \text{ jours} \\ S = 14,73127 \text{ jours} \\ S = 14,80937 \text{ jours} \end{cases}$	Moyenne $SN_2 = 14,74988$ Ecart type $SN_2 = 0,05461$ $t_{1-\frac{\alpha}{2}} = 3,182$ S et $SN_2 \in [14,6629, 14,8367]$	Il y a concordance parfaite entre S et les valeurs de SN_2 au seuil $\alpha = 0,05$
	Saisonnalité : Diurne.		Intervalle de confiance	Observations
Pleine mer	S = 1,034 jours	$SN_3 = \begin{cases} 1,06184 \text{ jours} \\ 1,07594 \text{ jours} \end{cases}$	Moyenne $SN_3 = 1,06889$ Ecart type $SN_3 = 0,00997$ $t_{1-\frac{\alpha}{2}} = 12,706$ S et $SN_3 \in [0,94219, 1,19560]$	Il y a concordance parfaite entre S et les valeurs de SN_3 au seuil $\alpha = 0,05$
Basse mer	S = 1,034 jours	$SN_4 = \begin{cases} 1,06443 \text{ jours} \\ 1,07284 \text{ jours} \end{cases}$	Moyenne $SN_4 = 1,06863$ Ecart type $SN_4 = 0,00594$ $t_{1-\frac{\alpha}{2}} = 12,706$ S et $SN_4 \in [0,99307, 1,14419]$	Il y a concordance parfaite entre S et les valeurs de SN_4 au seuil $\alpha = 0,05$

Notons que la valeur de la première ligne du vecteur SN3 et du vecteur SN4 était prise égale à la moyenne des saisonnalités significatives trouvées après l'application de la méthode de l'entropie maximale.

L'inconvénient de notre méthode est qu'il n'arrive pas à détecter la saisonnalité semi-diurne. Ce qui est évident d'après le théorème de Nyquist-Shannon [25] que nous avons défini au paragraphe 2.3.3.3.

Pour détecter la saisonnalité semi-diurne S (12 heures 25 minutes), nous avons donc besoin d'échantillons prélevés tous les $S/2$, soit un prélèvement de données toutes les 6 heures et 12,5 minutes.

Cependant comme la saisonnalité diurne issue des méthodes traditionnelles coïncide avec celle de nos modèles, il nous suffirait d'interpoler ou de diviser par deux la valeur de la saisonnalité diurne pour trouver la valeur de la saisonnalité semi-diurne et aboutir aux mêmes résultats que les méthodes traditionnelles.

3.2.9. Choix de la Saisonnalité S

Les valeurs précédentes trouvées pour S permettent d'encadrer le domaine de son estimation entre $E(S) = 28$ Lags et $E(S+1) = 29$ Lags.

3.2.9.1. Test sur les Résidus

Les tests sur les résidus (Cf. figure 3.27) recommandent, aussi bien pour les pleines mers que pour les basses mers, de choisir la saisonnalité S égale à 29. En effet, nous pouvons immédiatement constater sur la figure 3.27 que la fluctuation résiduelle observée pour $S = 29$ Lags (coloration en rouge) est toujours moindre par rapport à la fluctuation résiduelle observée pour $S = 28$ Lags (coloration en bleu).

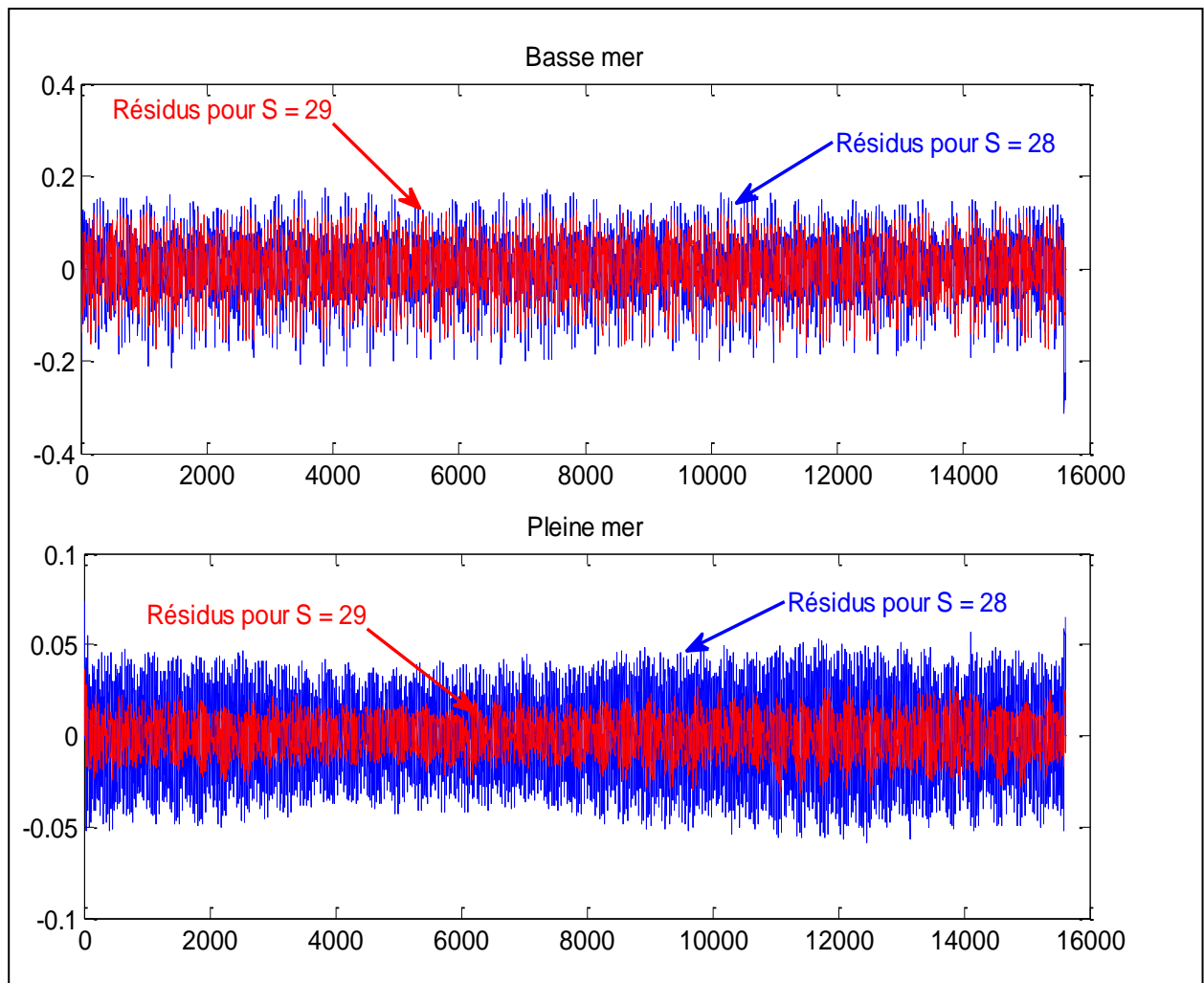


Figure 3 . 27 : Courbes des résidus pour S = 28 et S = 29

3.2.9.2. Analyse de la Variance à un Facteur

Pour pouvoir mettre en œuvre le test de Fisher, nous avons transformé nos données vectorielles en matrice. Nous avons obtenu respectivement pour les pleines mers et pour les basses mers :

- une matrice B = 558 lignes x 28 colonnes ;
- et une matrice A = 539 lignes x 29 colonnes.

Les résultats du tableau 3.10 confirment la possibilité d'utilisation du test de Fisher pour étudier la stabilité de nos chroniques.

Tableau 3 . 10 : Indépendance, homoscedasticité et normalité

	Basse mer	Pleine mer	Décisions
Test	L'indépendance des données résultent de l'indépendance des prélèvements des niveaux de la mer		
Bartlett	$\frac{B}{C} = 5,85 < \chi^2(28) = 41,34$	$\frac{B}{C} = 9,158 < \chi^2(28) = 41,34$	Homoscedasticité acceptée
Levene	$D = 0,37 < F(28, 538) = 1,5$	$D = 3,198 < F(28, 538) = 1,5$	
Normalité totale	$1,96 \leq \min_{1 \leq i \leq 29} (JB_i) = 6$ $1,96 \leq \min_{1 \leq i \leq 29} S_i = 8,8$ $1,96 \leq \min_{1 \leq i \leq 29} K_i = 2,25$	$1,96 \leq \min_{1 \leq i \leq 29} (JB_i) = 6,84$ $ S_i _{1 \leq i \leq 29} \leq 1,96$ pour 13 « i » sur 29 $ K_i _{1 \leq i \leq 29} \leq 1,96$ pour 9 « i » sur 29	Rejet de la normalité totale pour les deux chroniques
Normalité grossière	$0,0018 \leq \mathbf{ad} - 0,7979\sigma \leq 0,01$ $0,002 \leq \mathbf{pd} - 0,6745\sigma \leq 0,04$	$0,00001 \leq \mathbf{ad} - 0,7979\sigma \leq 0,0051$ $0,00008 \leq \mathbf{pd} - 0,6745\sigma \leq 0,019$	Vraisemblance à la loi normale acceptée pour les deux chroniques

Nous avons tracé sur la figure 3.28, pour S = 28 et 29 Lags, les statistiques de Fisher obtenus en appliquant l'analyse de la variance à un facteur aux chroniques sur les pleines et les basses mers.

Les résultats de cette figure 3.28 montrent que les statistiques de Fisher sont significatives :

- pour les basses mers, entre la ligne 1 et la ligne 142 pour la saisonnalité S égale à 28, et entre la ligne 1 et la ligne 161 pour la saisonnalité S égale à 29 ;
- pour les pleines mers, entre la ligne 1 et la ligne 140 pour la saisonnalité S égale à 28, et entre la ligne 1 et la ligne 156 pour la saisonnalité S égale à 29.

Pour chaque chronique, les deux résultats nous recommandent toujours de choisir S égale à 29.

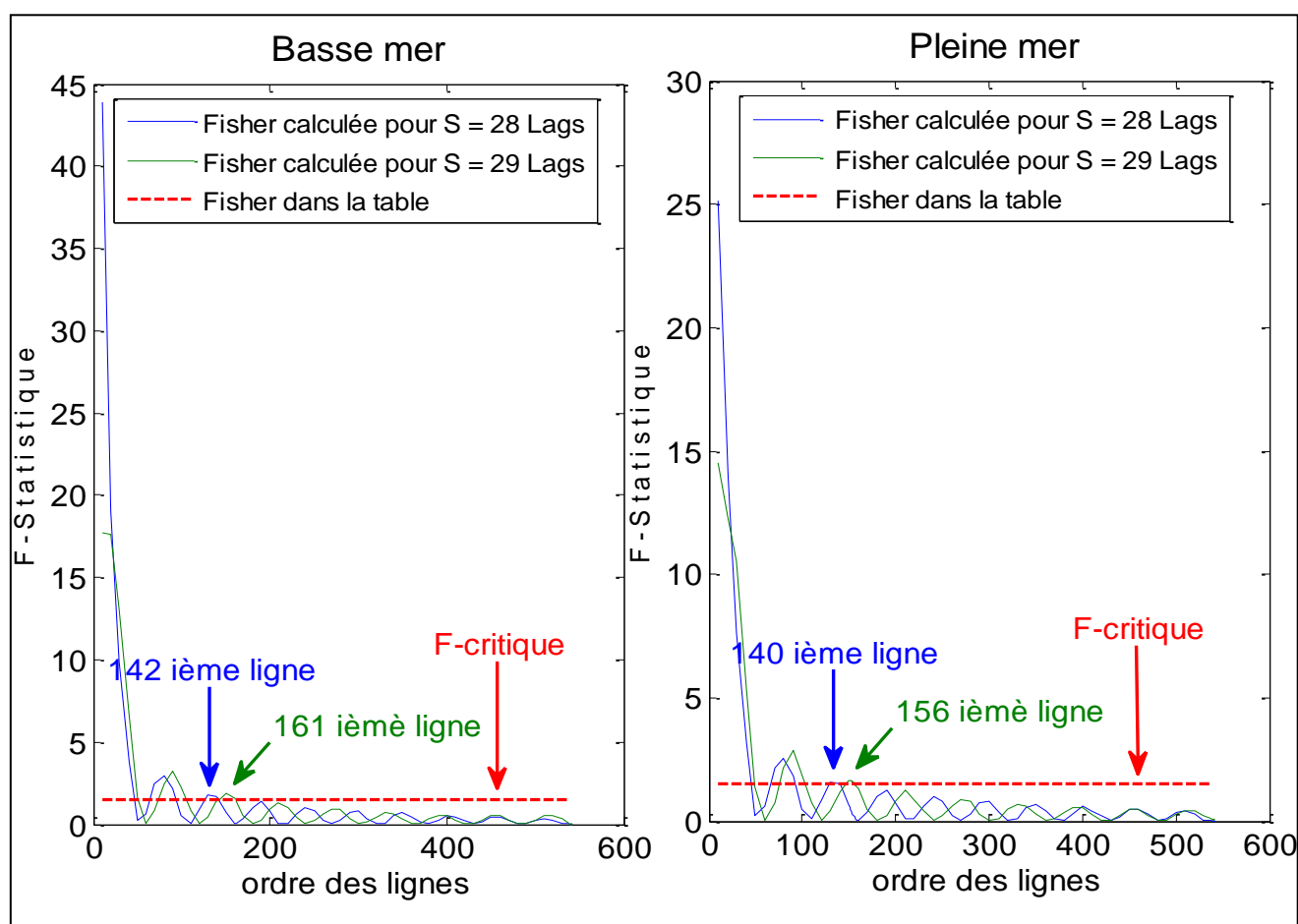


Figure 3 . 28 : Test de Stabilité

En répétant la même opération à partir de la ligne 162 pour les basses mers et à partir de la ligne 156 pour les pleines mers, et ainsi de suite, nous arrivons à établir le tableau 3.11 montrant l'existence d'une stabilité locale au niveau de nos chroniques.

Tableau 3 . 11 : Analyse de la stabilité pour S = 29

Basses mers			Pleines mers		
Ordre des lignes	F-Statistique	F-Critique	F-Statistique	F-Critique	Ordre des lignes
Ligne 1 à 161, 4669 données	1.503	1.478	1,490	1,478	Ligne 1 à 156, 4524 données
Ligne 162 à 323, 6351 données	1.488	1.478	1,533	1,478	Ligne 157 à 314, 4582 données
Ligne 381 à 539, 4611 données	1.527	1.478	1 ,534	1 ,478	Ligne 315 à 468, 4466 données
			1,031	1 ,482	Ligne 469 à 539, 2059 données

Finalement, les deux méthodes nous recommandent d'estimer la saisonnalité S à 29 Lags. Mais, si nous nous référons aux résultats de l'analyse de la variance à un facteur, nous pouvons en plus affirmer que nous devons établir :

- pour les basses mers 03 modèles de calcul,
- et pour les pleines mers 04 modèles.

3.2.10. Analyse de la tendance sur les données stabilisées

3.2.10.1. Test de corrélation de rang de Kendall

La mise en œuvre du test de corrélation de rang de Kendall a permis de dresser le tableau 3.12, donnant les paramètres de corrélation des chroniques sur les pleines et sur les basses mers.

Tableau 3 . 12 : Test de corrélation de rang de Kendall

	τ_b	$E(\tau_b)$	$Var(\tau_b)$	Z_1	$U_{0,1,0,05}$
Pleine mer	-9,98 %	0	$2,8440.10^{-05}$	-18,7082	1,96
Basse mer	11,80 %	0	$2,8440.10^{-05}$	22,1207	1,96

D'après les résultats du tableau 3.12 :

pour le cas des pleines mers, nous pouvons rejeter l'hypothèse H_0 : il n'y a pas de corrélation entre les rangs des niveaux des pleines mers et les rangs des temps de prélèvement car -18,7082 est supérieure à 1,96 en valeur absolue. Comme le Z_1 des pleines mers est négatif, nous pouvons aussi affirmer que les niveaux des pleines mers ont une tendance décroissante en fonction du temps de prélèvement ;

pour le cas des basses mers, nous pouvons également rejeter l'hypothèse H_0 : il n'y a pas de corrélation entre les rangs des niveaux des basses mers et les rangs des temps de prélèvement car 22,1207 est supérieure à 1,96. Comme le Z_1 des basses mers est positif, nous pouvons dire que les niveaux des basses mers possèdent une tendance croissante en fonction du temps de prélèvement.

3.2.10.2. Test de corrélation de rang de Mann-Kendall

De la même manière que dans le paragraphe précédent, nous avons pu définir les paramètres du test de corrélation de rang de Mann-Kendall dans le tableau 3.13.

Tableau 3 . 13 : Test de corrélation de rang de Mann-Kendall

	S''	$E(S'')$	$Var(S'')$	Z_2	$U_{0,1,0,05}$
Pleine mer	-12 150 235	0	$4,2437.10^{11}$	-18,6516	1,96
Basse mer	14 361 977	0	$4,2436.10^{11}$	22,0468	1,96

Ces résultats nous permettent :

pour les pleines mers de rejeter l'hypothèse **H₀** : il n'y a pas de corrélation entre les rangs des niveaux des pleines mers et les rangs des temps de prélèvement car -18,6516 est supérieure à 1,96 en valeur absolue. Comme le Z_2 des pleines mers est négatif, nous pouvons aussi affirmer que les niveaux des pleines mers ont une tendance décroissante en fonction du temps de prélèvement ;

pour le cas des basses mers de rejeter l'hypothèse **H₀** : il n'y a pas de corrélation entre les rangs des niveaux des basses mers et les rangs des temps de prélèvement car 22,0468 est supérieure à 1,96. Comme le Z_2 des basses mers est positif, nous pouvons dire que les niveaux des basses mers possèdent une tendance croissante en fonction du temps de prélèvement.

3.2.10.3. Test de corrélation de rang de Spearman

Nous avons pu également déterminer les paramètres de corrélation des rangs de Spearman que nous avons consignés dans le tableau 3.14.

Tableau 3 . 14 : Test de corrélation de rang de Spearman

	r_s	$\left r_s \sqrt{\frac{N-2}{1-r_s^2}} \right $	signe (r_s)	$t_{1-\frac{\alpha}{2}}(N-1)$
Pleine mer	-14,36 %	-18,1415	-	1,96
Basse mer	17,31 %	21,9656	+	1,96

Par rapport à ces résultats, nous pouvons rejeter pour les deux chroniques pleines mers et basses mers, l'hypothèse **H₀** : il n'y a pas de corrélation entre les rangs des niveaux des pleines mers et les rangs des temps de prélèvement car $| -18,1415 |$ et $| 21,9656 |$ sont tous les deux supérieurs à 1,96.

Nous pouvons également affirmer pour la chronique sur :

- **les pleines mers** que la corrélation entre les rangs des niveaux des pleines mers et les rangs des temps de prélèvement est négative et que la tendance est décroissante car $r_s < 0$;
- **les basses mers** que la corrélation entre les rangs des niveaux des basses mers et les rangs des temps de prélèvement est positive et que la tendance est croissante car $r_s > 0$.

En particulier, comme les deux valeurs absolues de r_s pour les deux chroniques sont comprises entre 2,5 % et 50 % d'après les résultats du tableau 3.14 alors la liaison entre les rangs est très faible.

3.2.11. Schéma de prévision

3.2.11.1. Méthode du profil

Dans la figure 3.29, nous avons représenté les courbes représentatives des données correspondant aux lignes 1 à 10 des matrices M1 obtenus en transformant les vecteurs des données initiales homogénéisées des hauteurs des pleines et des basses mers selon la formulation que nous avons définie au paragraphe 2.3.5.

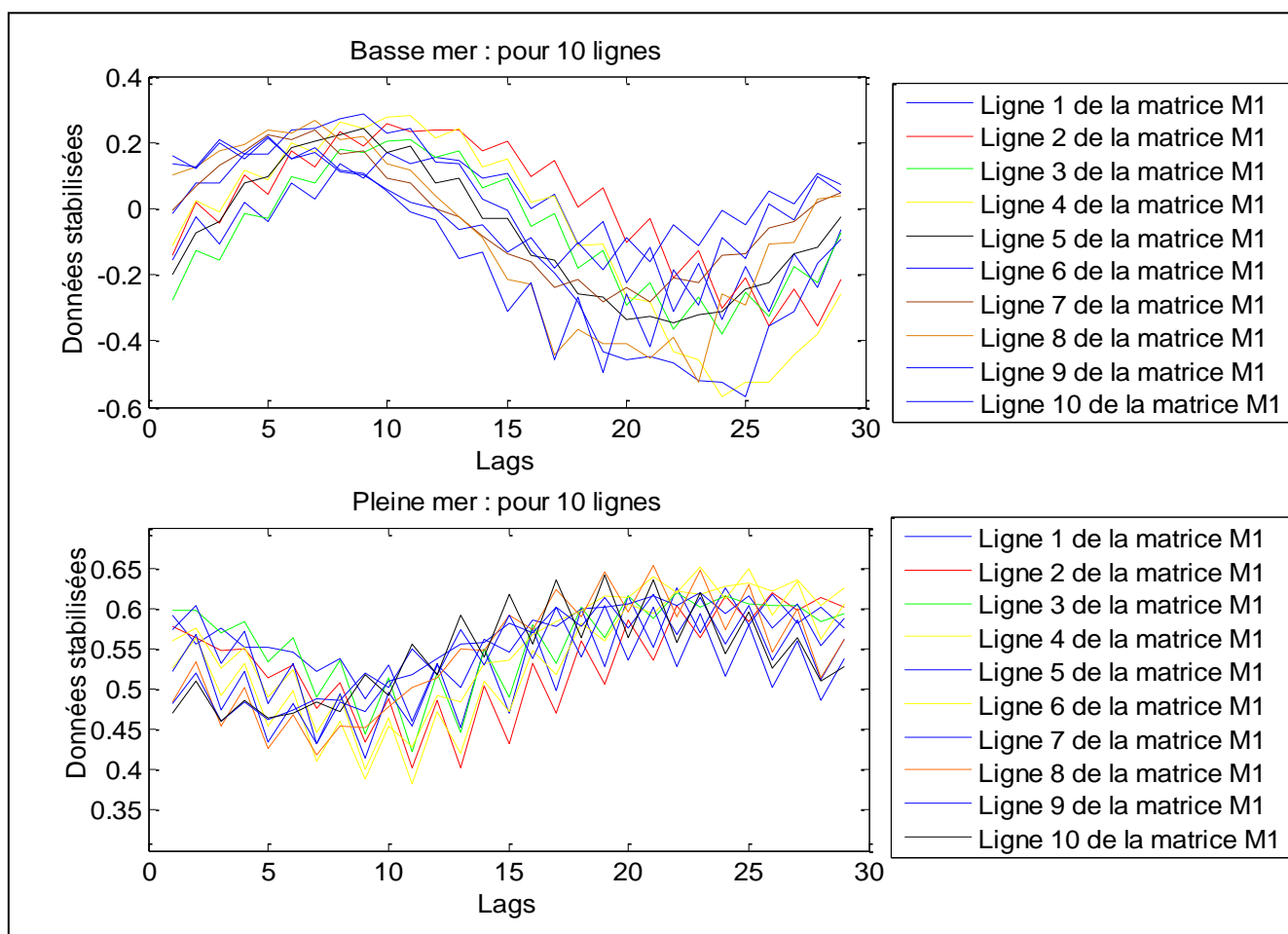


Figure 3 . 29 : Méthode du profil

Les résultats de la figure 3.29 ne nous permettent pas de trancher sur le choix du schéma, puisque plusieurs courbes s'entrecroisent et d'autres sont plus ou moins confondues.

3.2.11.2. Méthode de la bande

Nous avons tracé respectivement les courbes des minima et des maxima des hauteurs des pleines et des basses mers avec les données stabilisées pour obtenir la figure 3.30

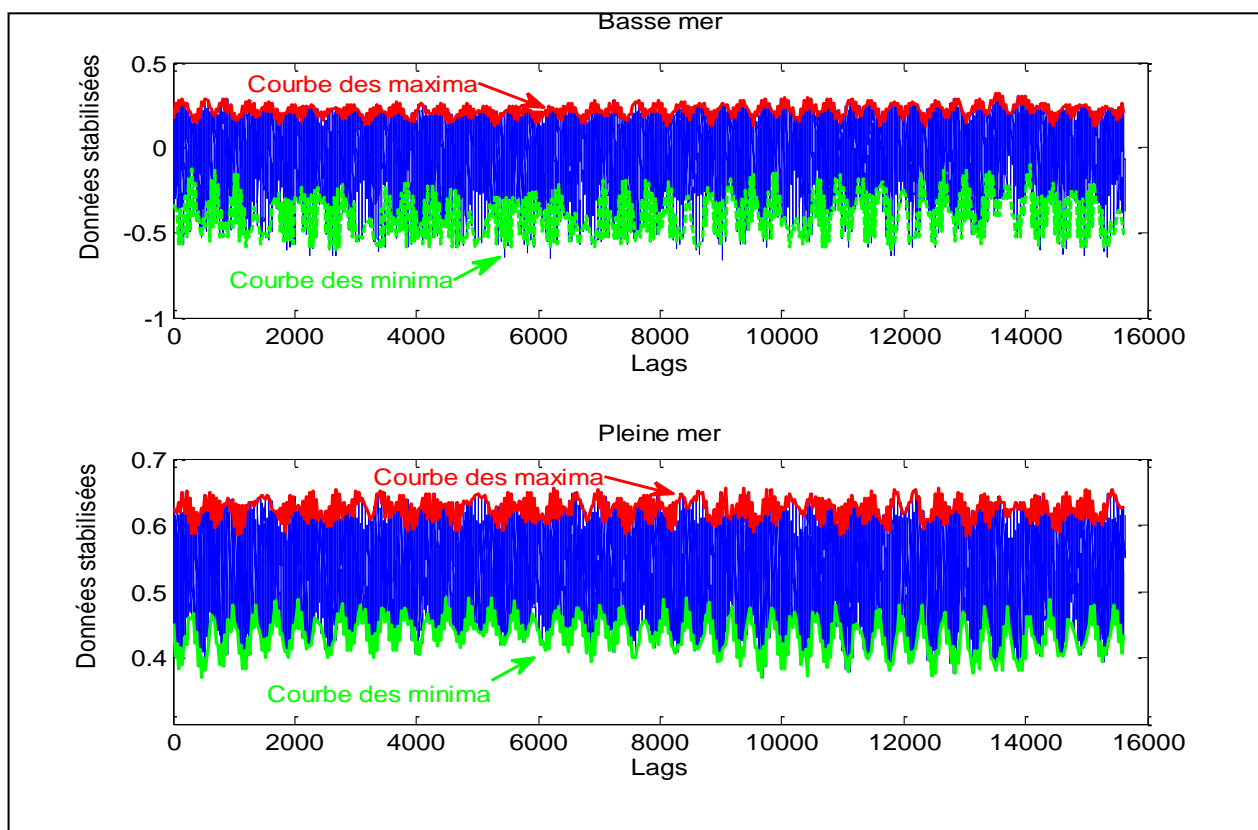


Figure 3 . 30 : Méthode de la bande

Comme les courbes des minima et des maxima ne sont pas parallèles pour les deux chroniques, nous devons choisir un schéma de prévision multiplicatif aussi bien pour les données sur les pleines mers que sur les basses mers.

3.2.11.3. Méthode analytique

Les résultats du tableau 3.15 nous recommandent aussi d'appliquer le schéma multiplicatif puisque -28,11 (t-statistique basse mer) et 2,30 (t-statistique pleine mer) sont supérieurs à 1,96 en valeur absolue.

Tableau 3 . 15 : Régression entre la moyenne et l'écart type

	Basse mer			Pleine mer		
	Coefficient	t -Statistique	Valeur critique	Coefficient	t -Statistique	Valeur critique
g	0,15126582	72,335243	1,96	0,00400041	0,17726521	1,96
μ	-0,84464676	-28,1101052	1,96	0,09609609	2,307198454	1,96

Finalement, sur les trois méthodes testées, deux nous suggèrent de choisir le schéma multiplicatif que nous avons adopté dans cette étude.

3.2.12. Méthode Census X11

La transformation logarithmique a permis :

- pour les basses mers de réduire l'écart-type de 0,4125 à 0,2160
- et pour les pleines mers de réduire l'écart-type de 0,4494 à 0,0566.

Nous avons appliqué la méthode Census X11 aux matrices $A = 539$ lignes x 29 colonnes et nous avons trouvé la tendance en appliquant la moyenne mobile de Henderson à 13 termes puisque nous avons trouvé pour la chronique sur :

- les pleines mers un rapport $\frac{I}{C} = 1,48$, qui est supérieur à 1 ;
- les basses mers le rapport $\frac{I}{C} = 2,24$, aussi supérieur à 1.

Nous remarquons que les courbes des tendances des données initiales et leurs droites de régression linéaire sur la figure 3.31 confirment les résultats trouvés par les différents tests de corrélation des rangs effectués précédemment sur les données stabilisées, à savoir :

- l'existence d'une corrélation négative des rangs entre les niveaux des pleines mers et les temps de prélèvement ;
- l'existence d'une corrélation positive des rangs entre les niveaux des basses mers et les temps de prélèvement.

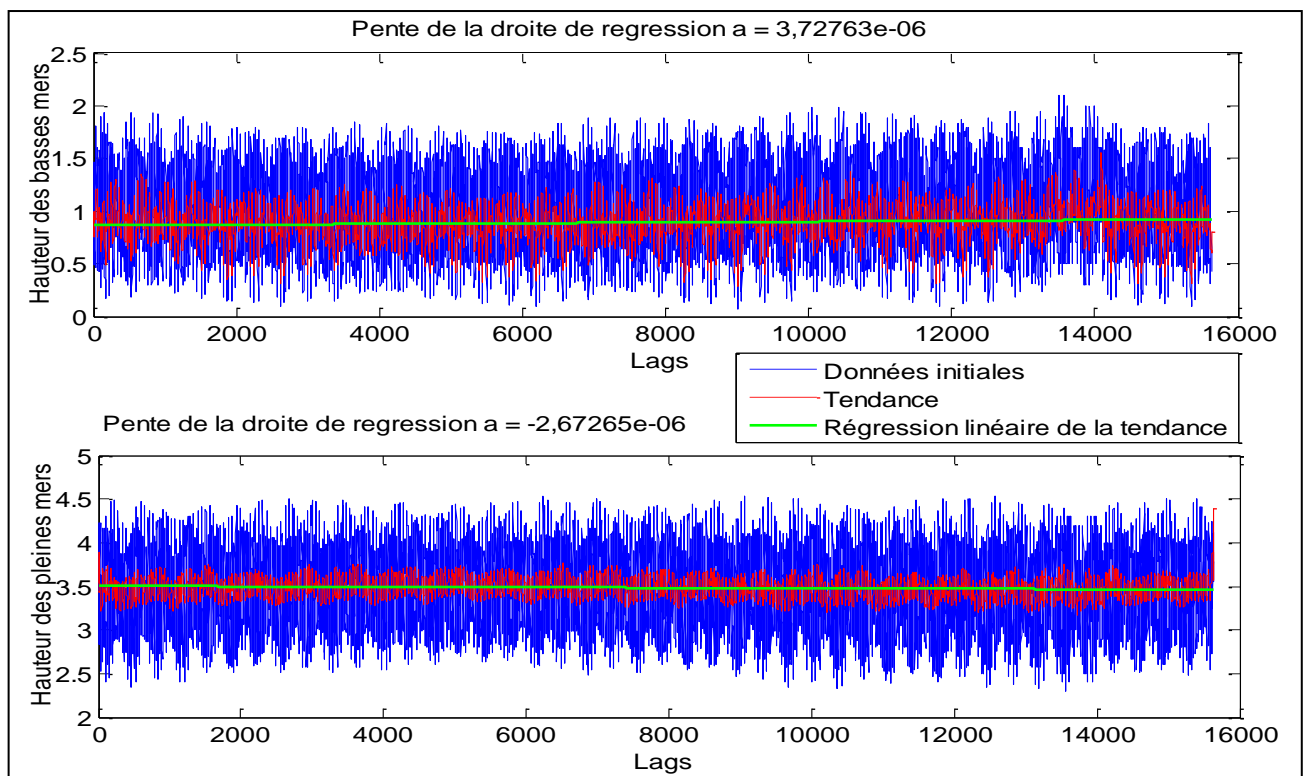


Figure 3 . 31 : Courbes des tendances des données initiales

Les résultats de la figure 3.31 montrent également que nos deux chroniques sur les pleines et basses mers possèdent des tendances non linéaires.

3.2.13. Méthode ARIMA(p+1,d,q)

La mise en œuvre de la méthodologie ARIMA(p+1,d,q) consiste à déterminer l'ordre q de la partie à Moyenne Mobile, l'ordre d d'Intégration l'ordre q de la partie AutoRégressive

3.2.13.1. Identification des ordres d, q et p du modèle ARIMA(p+1,d,q)

a) Minimisation de l'écart type

La mise en œuvre de la technique de minimisation de l'écart type des séries différenciées a donné les résultats du tableau 3.16 suivant :

Tableau 3 . 16 : Minimisation de l'écart type des séries différenciées

	Ecart type : chronique pleine mer				Ecart-type : chronique basse mer		
ordre " d "	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 1	Modèle 2	Modèle 3
d = 0	0,0175	0,0165	0,0194	0,0141	0,0888	0,0924	0,0963
d = 1	0,0106	0,0095	0,0141	0,0133	0,0352	0,037	0,0384
d = 2	0,0198	0,0178	0,0267	0,0249	0,0586	0,0618	0,0637
d = 3	0,0392	0,0352	0,0528	0,049	0,1151	0,1213	0,1245
Minimum écart type	0,0106	0,0095	0,0141	0,0133	0,0352	0,037	0,0384

La partie du tableau 3.16 en jaune nous recommande de prendre pour nos sept modèles l'ordre d'intégration « d égal à 1 ».

b) Stratégie de Dickey Fuller Augmenté

Pour démarrer cette stratégie, nous avons décidé de prendre un nombre de retard initial p_{initial} égal à 14 et de varier d entre 1 et 2.

La mise en œuvre de la stratégie de Dickey Fuller Augmenté aux séries corrigées des variations saisonnières a permis de dresser le tableau 3.17 suivant en considérant des modèles avec tendance et constante :

Tableau 3 . 17 : Stratégie de Dickey Fuller Augmenté avec constante et tendance pour un nombre de retard $p = 14$ [40]

Modèles	Basse mer, nombre de retard initial $p = 14$						Pleine mer, nombre de retard initial $p = 14$						Valeur critique	
	Statistique t pour $d=1$			Statistique t pour $d=2$			Statistique t pour $d=1$			Statistique t pour $d=2$			ϕ	b et a
	ϕ	b	a	ϕ	b	a	ϕ	b	a	ϕ	b	a		
Modèle 1	13,43	0,06	-0,03	-28,24	0,09	-0,06	1,06	-0,02	-0,01	-24,63	-0,05	0,06	-3,41	1,96
Modèle 2	13,43	0,04	-0,05	-32,98	-0,21	0,20	4,77	-0,11	0,11	-23,91	0,05	-0,05	-3,41	1,96
Modèle 3	4,17	-0,05	0,04	-29,00	-0,03	0,03	-3,61	-0,02	0,02	-24,87	0,08	-0,07	-3,41	1,96
Modèle 4							-3,78	0,07	-0,07	-17,76	-0,03	0,02	-3,41	1,96

En nous référant à la partie hachurée en jaune du tableau 3.17, nous pouvons dire que le test de Dickey Fuller augmenté nous recommande :

- ⇒ De prendre pour les trois modèles sur les basses mers « d égal à 2 » et
- ⇒ De prendre pour les deux premiers modèles sur les pleines mers « d égal à 1 » et pour les deux derniers modèles sur les pleines mers « d égal à 2 ».

c) Choix de « d_{final} et de p_{optimal} »

Pour les basses mers, l'ordre d'intégration « d » fourni par les deux méthodes diffère.

Nous avons donc décidé de mettre en œuvre le test de Ljung-Box pour « d compris entre 1 et 2 » et de choisir « un nombre de retard initial p_{initial} égal à 14 » pour débiter la méthode de vérification ex post de la non autocorrélation des résidus.

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.32 et 3.33 pour le modèle 1 basse mer.

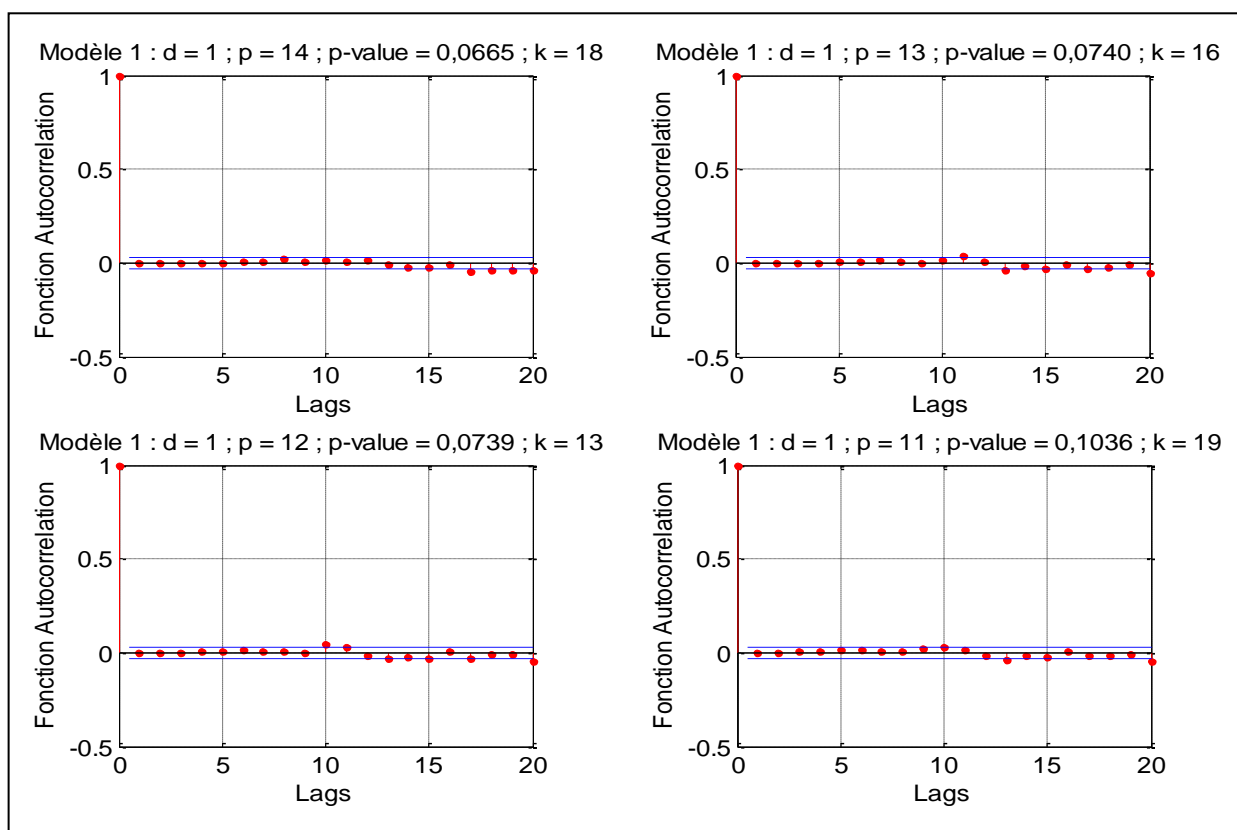


Figure 3 . 32 : Courbes des résidus du modèle 1 basses mers pour $d = 1$ et p variant de 11 à 14

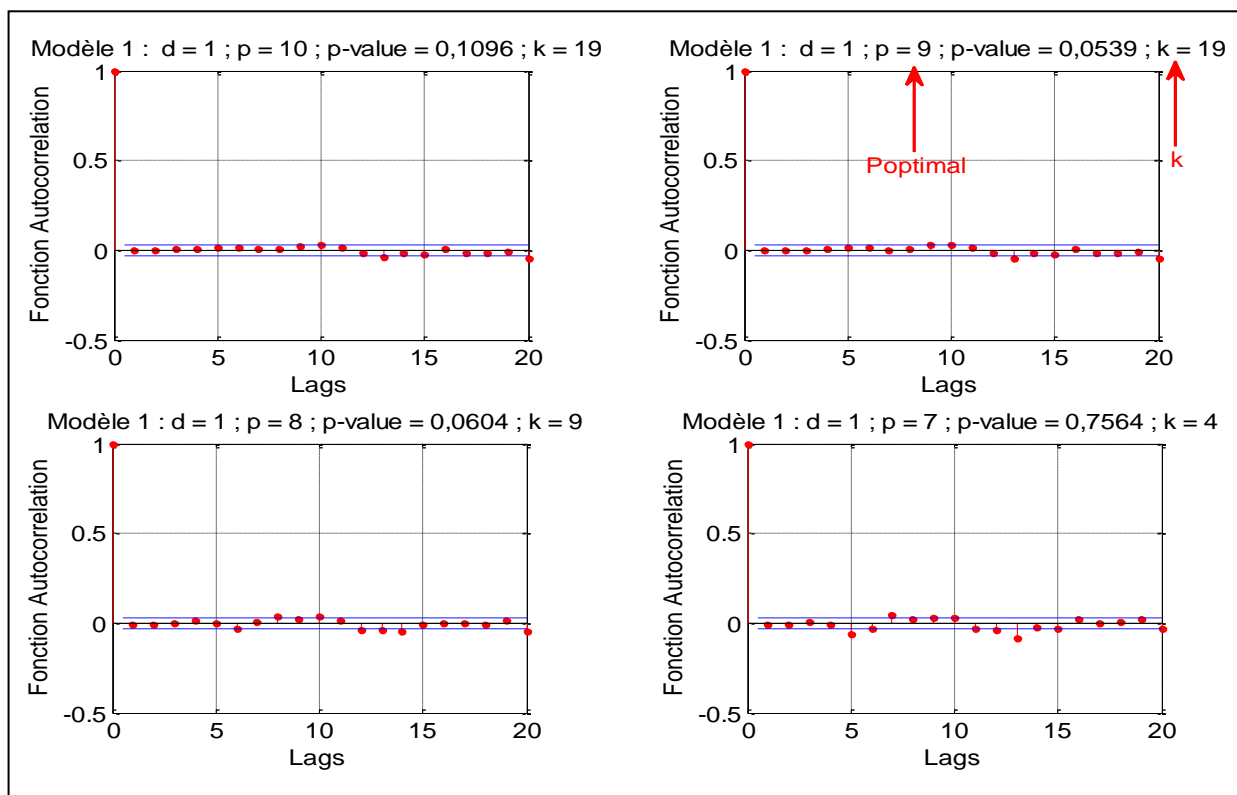


Figure 3 . 33 : Courbes des résidus du modèle 1 basses mers pour $d = 1$ et p variant de 7 à 10

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.34 et 3.35 pour le modèle 2 basse mer.

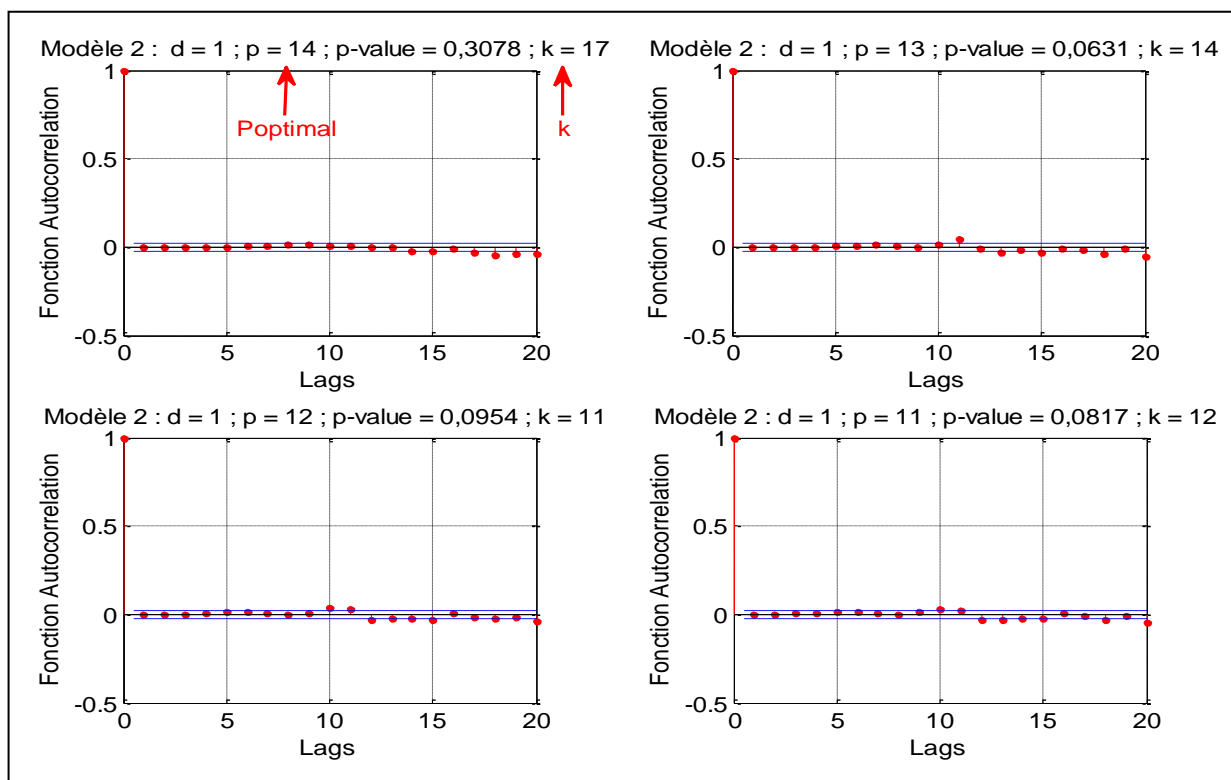


Figure 3.34 : Courbes des résidus du modèle 2 basses mers pour $d = 1$ et p variant de 11 à 14

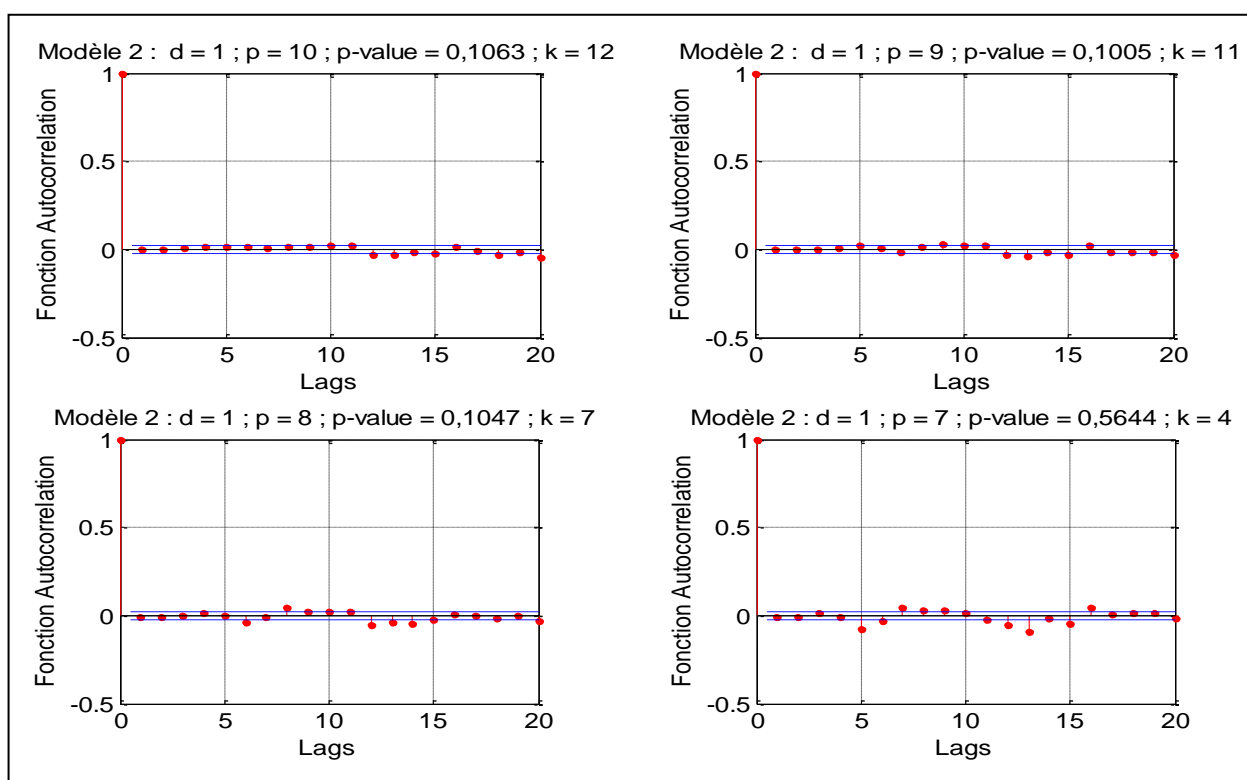


Figure 3.35 : Courbes des résidus du modèle 2 basses mers $d = 1$ et p variant de 7 à 10

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.36 et 3.37 pour le modèle 3 basse mer.

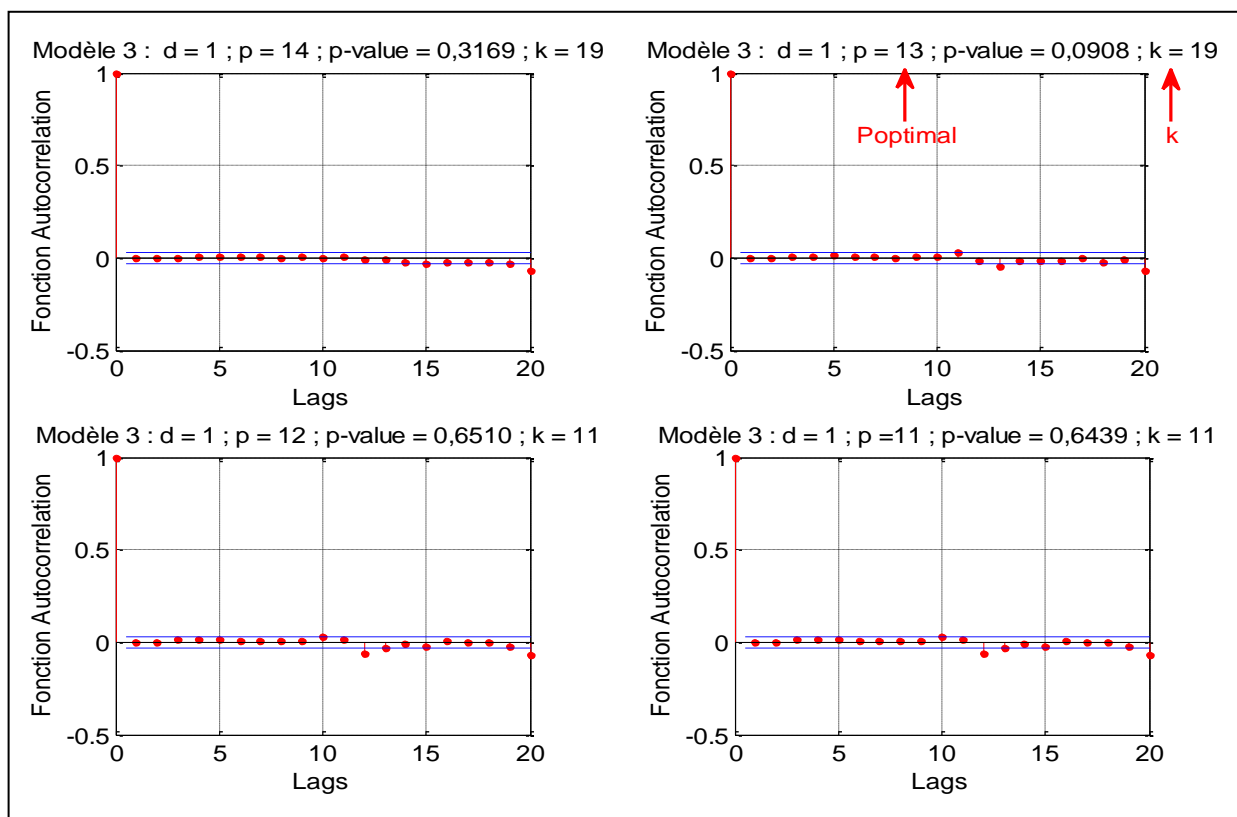


Figure 3.36 : Courbes des résidus du modèle 3 basses mers pour $d = 1$ et p variant de 11 à 14

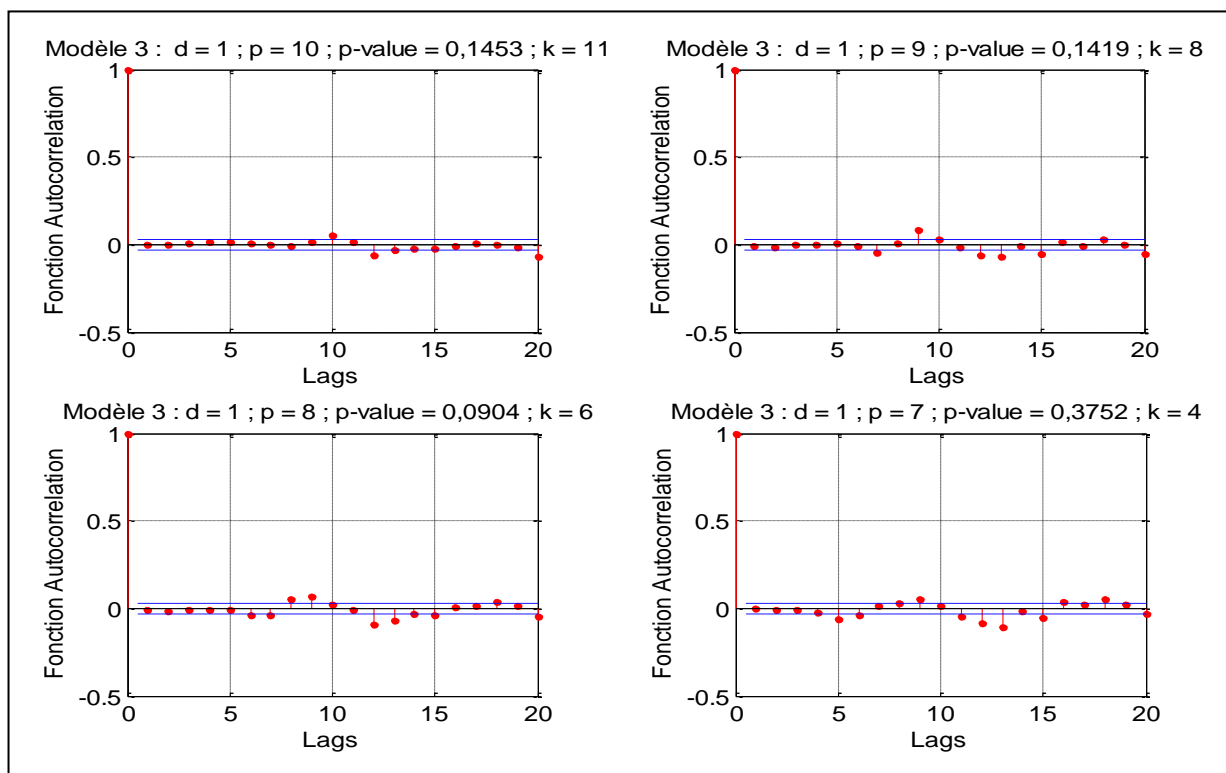


Figure 3.37 : Courbes des résidus du modèle 3 basses mers $d = 1$ et p variant de 7 à 10

En fixant « $d = 2$ » et en diminuant progressivement « p_{initial} » entre 14 et 11, nous sommes arrivés aux résultats de la figure 3.38 pour le modèle 1 basse mer.

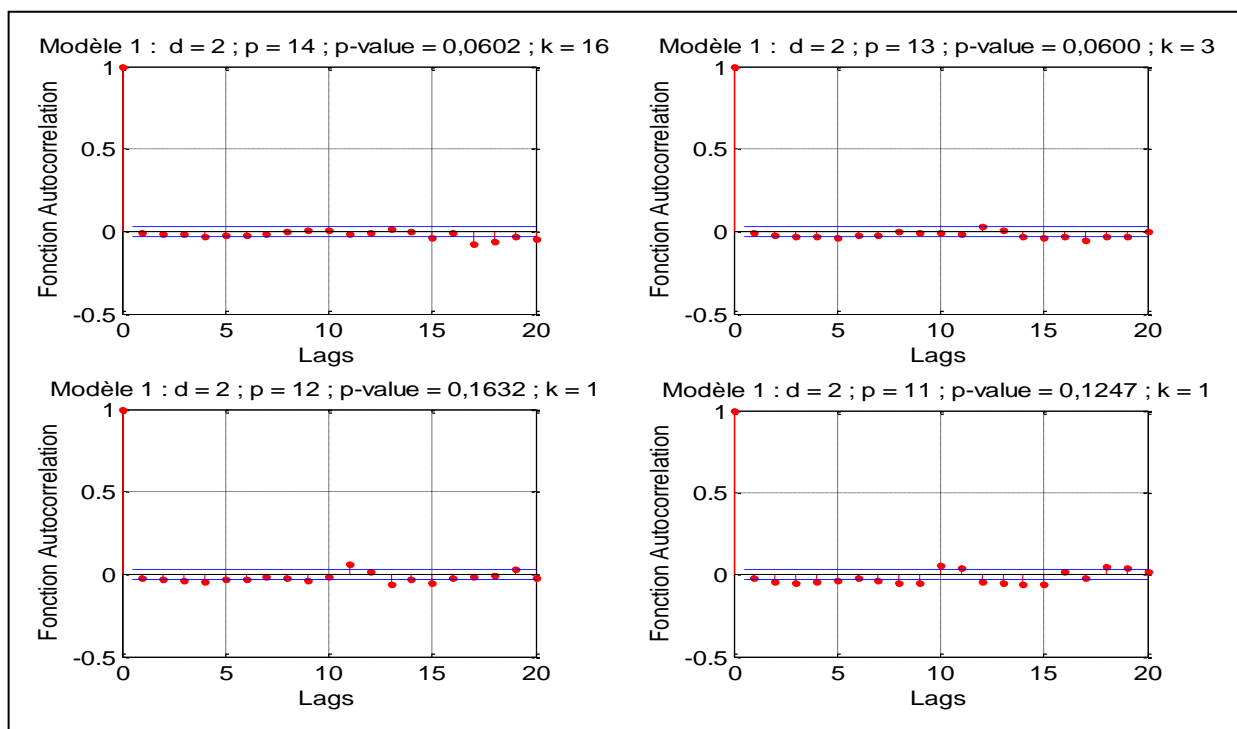


Figure 3.38 : Courbes des résidus du modèle 1 basses mers $d = 2$ et p variant de 11 à 14

En fixant « $d = 2$ » et en diminuant progressivement « p_{initial} » entre 14 et 11, nous sommes arrivés aux résultats de la figure 3.39 pour le modèle 2 basse mer.

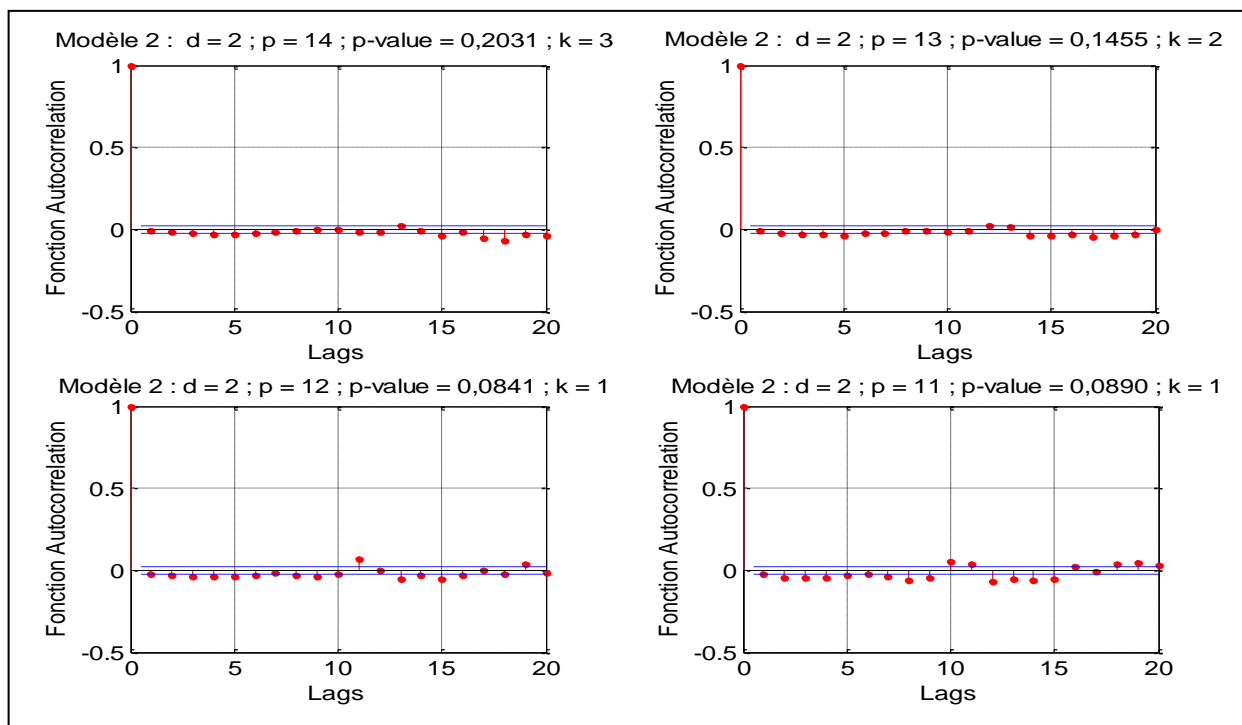


Figure 3.39 : Courbes des résidus du modèle 2 basses mers pour $d = 2$ et p variant de 11 à 14

En fixant « $d = 2$ » et en diminuant progressivement « p_{initial} » entre 14 et 11, nous sommes arrivés aux résultats de la figure 3.40 pour le modèle 3 basse mer.

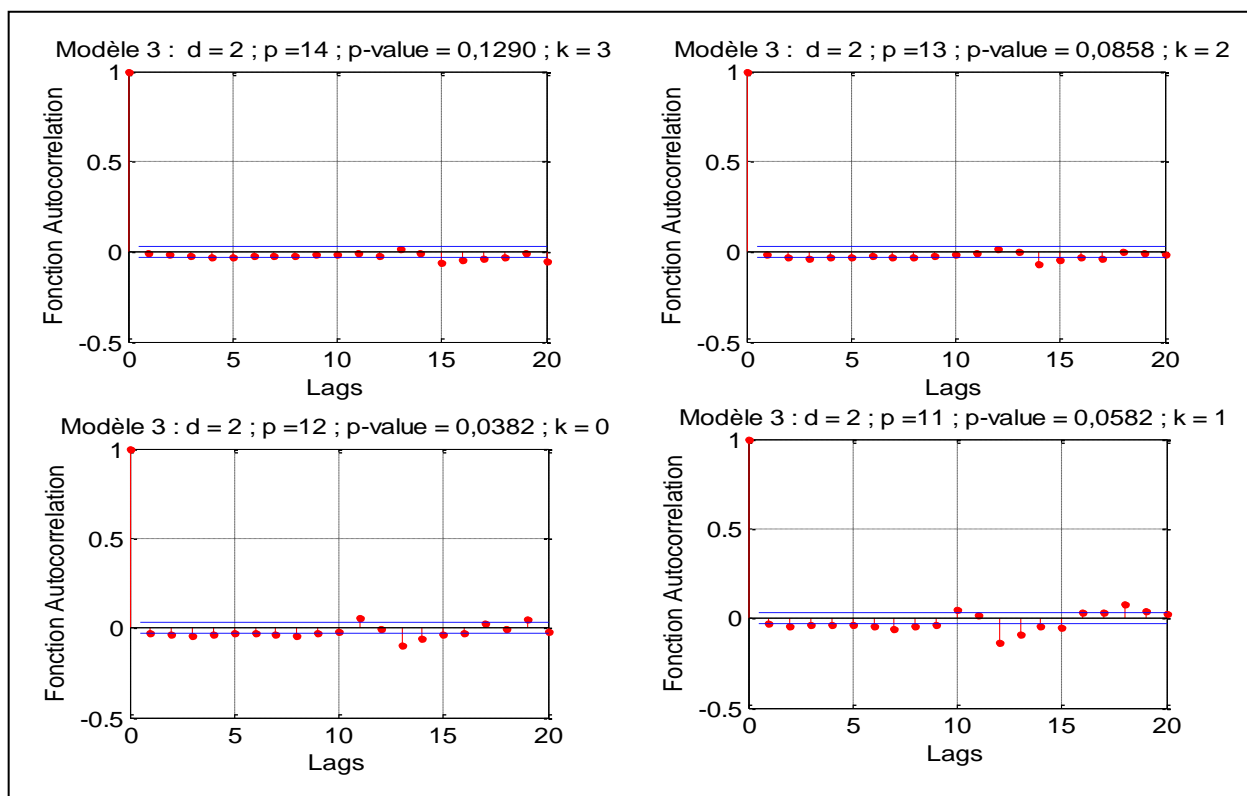


Figure 3 . 40 : Courbes des résidus du modèle 3 basses mers pour $d = 2$ et p variant de 11 à 14

Pour les trois modèles sur les basses mers, les résultats de la méthode de vérification ex post de la non autocorrélation des résidus des figures 3.32 à 3.40 ont été reportés dans le tableau 3.18.

Tableau 3 . 18 : Test de Ljung-Box pour les trois modèles sur les basses mers

	p	d = 1		d = 2				d = 1		d = 2				d = 1		d = 2	
		p-value	k	p-value	k			p-value	k	p-value	k			p-value	k	p-value	k
Modèle 1	14	0,0665	18	0,0602	16	Modèle 2		0,3078	17	0,2030	3	Modèle 3		0,3169	19	0,1290	3
	13	0,0740	16	0,0600	3			0,0631	14	0,1455	2			0,0908	19	0,0858	2
	12	0,0739	13	0,1632	1			0,0954	11	0,0841	1			0,6510	11	0,0382	0
	11	0,1036	19	0,1247	1			0,0817	12	0,0890	1			0,6439	11	0,0582	1
	10	0,1096	19					0,1063	12					0,1453	11		
	9	0,0539	19					0,1005	11					0,1419	8		
	8	0,0604	9					0,1047	7					0,0904	6		
	7	0,7564	4					0,5644	4					0,3752	4		

Les parties en jaune du tableau 3.18 désignent les modèles qui s'approchent de la réalité et qui minimisent les valeurs de p du modèle ARIMA(p,d,q).

Par conséquent, nous pouvons dire que :

- pour le modèle 1 de la chronique sur les basses mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 9$ » ;
- pour le modèle 2 de la chronique sur les basses mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 14$ » ;
- et pour le modèle 3 de la chronique sur les basses mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 13$ ».

Pour les pleines mers, l'ordre d_{final} des modèles 3 et 4 valent 1 puisque la méthode de minimisation de l'écart type et le test de Dickey Fuller Augmenté ont donné la même valeur « d égal à 1 ». Pour déterminer l'ordre d'intégration finale des modèles 1 et 2 et le p_{optimal} de nos quatre modèles sur les pleines mers, nous avons remis en œuvre le test de Ljung-Box et choisi l'ordre « d et p » qui permettent de mieux blanchir les résidus.

Pour débiter le test de Ljung-Box, nous avons pris pour les pleines mers le même nombre de retard initial $p_{\text{initial}} = 14$ que pour les basses mers.

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.41 et 3.42 pour le modèle 1 pleine mer.

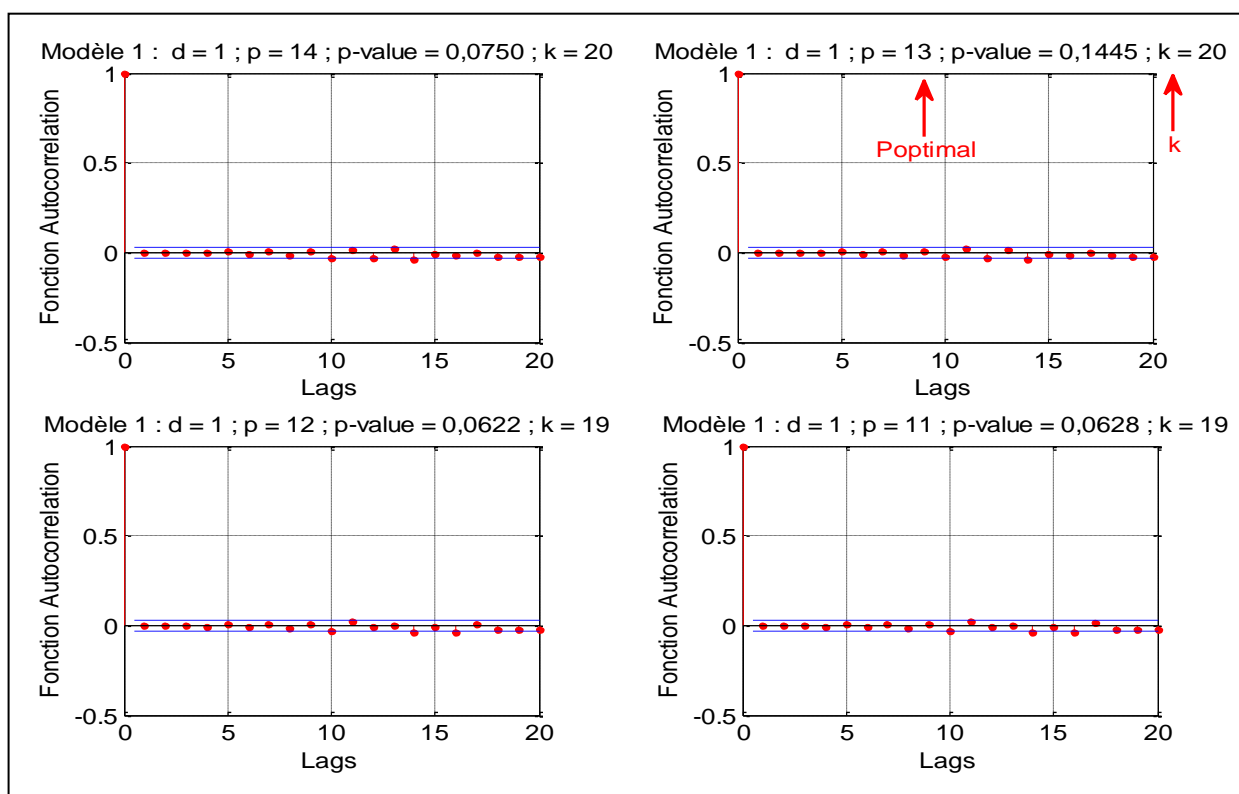


Figure 3.41 : Courbes des résidus du modèle 1 pleine mers pour $d = 1$ et p variant de 11 à 14

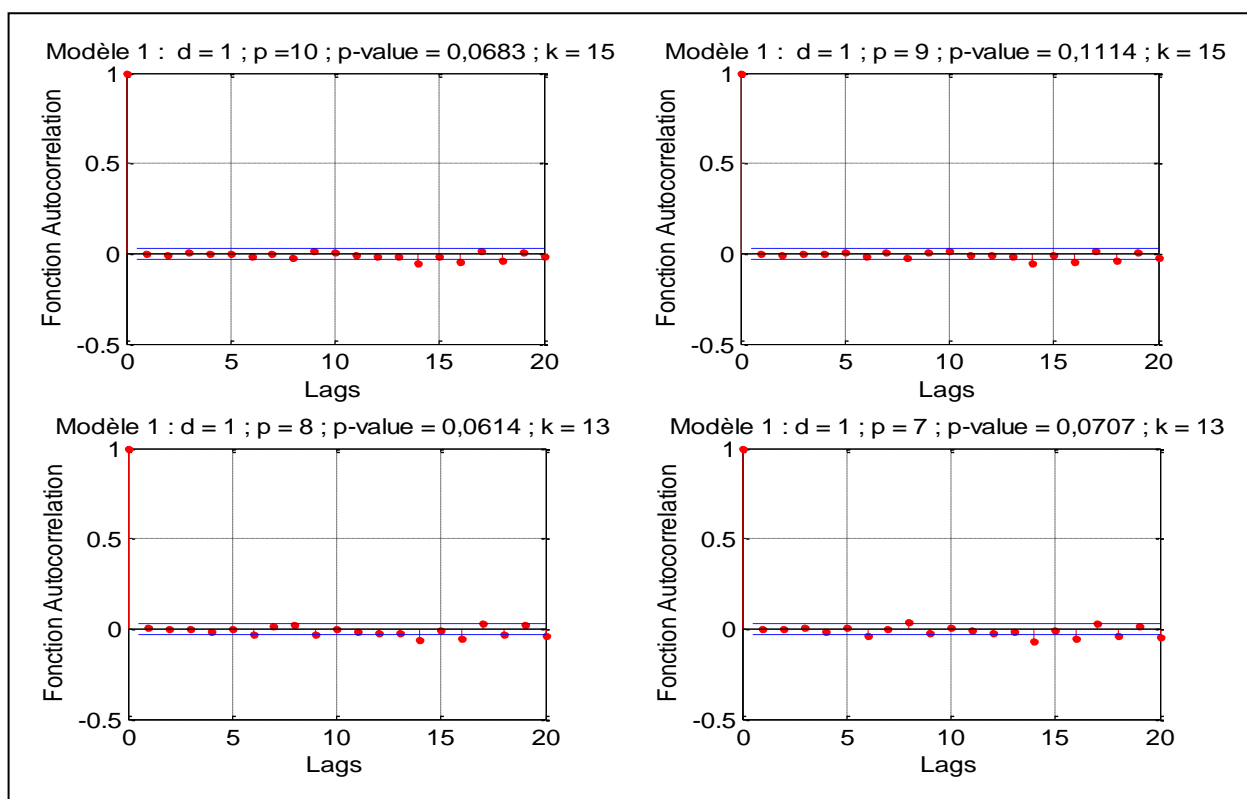


Figure 3 . 42 : Courbes des résidus du modèle 1 pleine mers pour $d = 1$ et p variant de 7 à 10

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.43 et 3.44 pour le modèle 2 pleine mer.

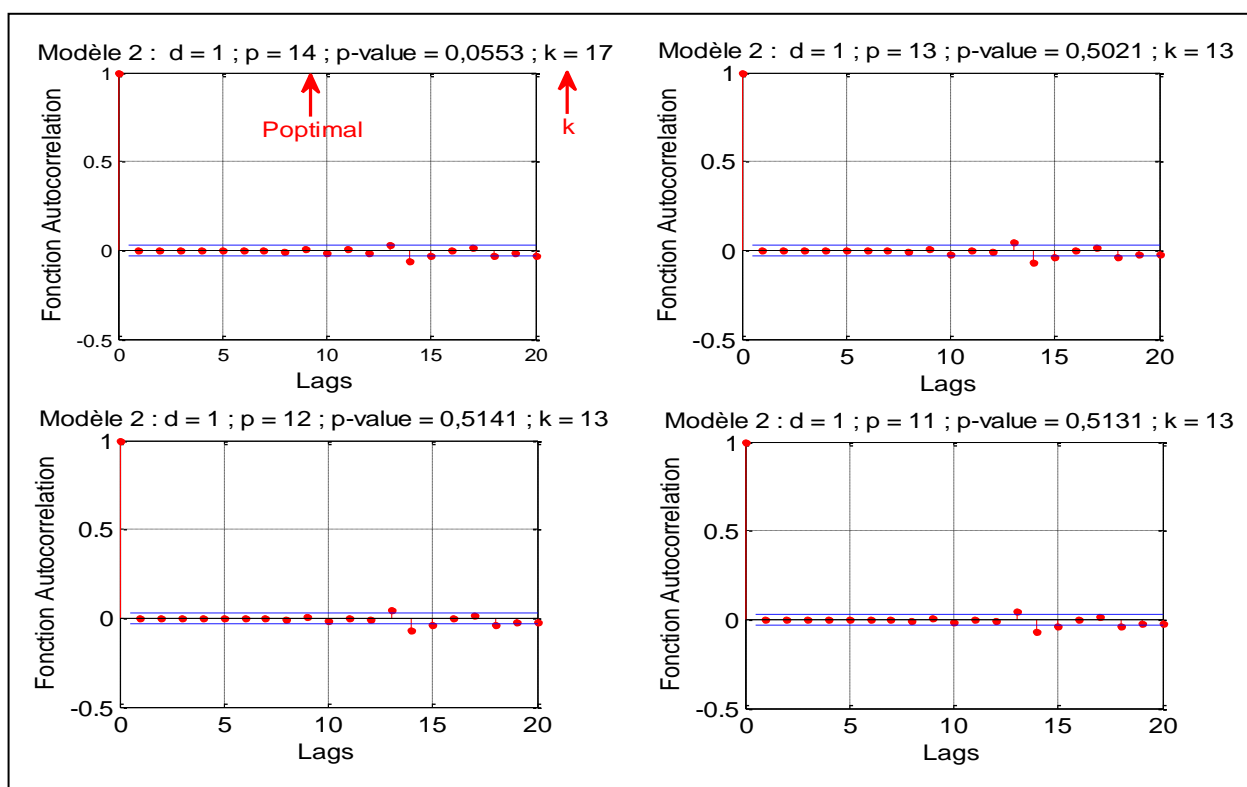


Figure 3 . 43 : Courbes des résidus du modèle 2 pleine mers pour $d = 1$ et p variant de 11 à 14

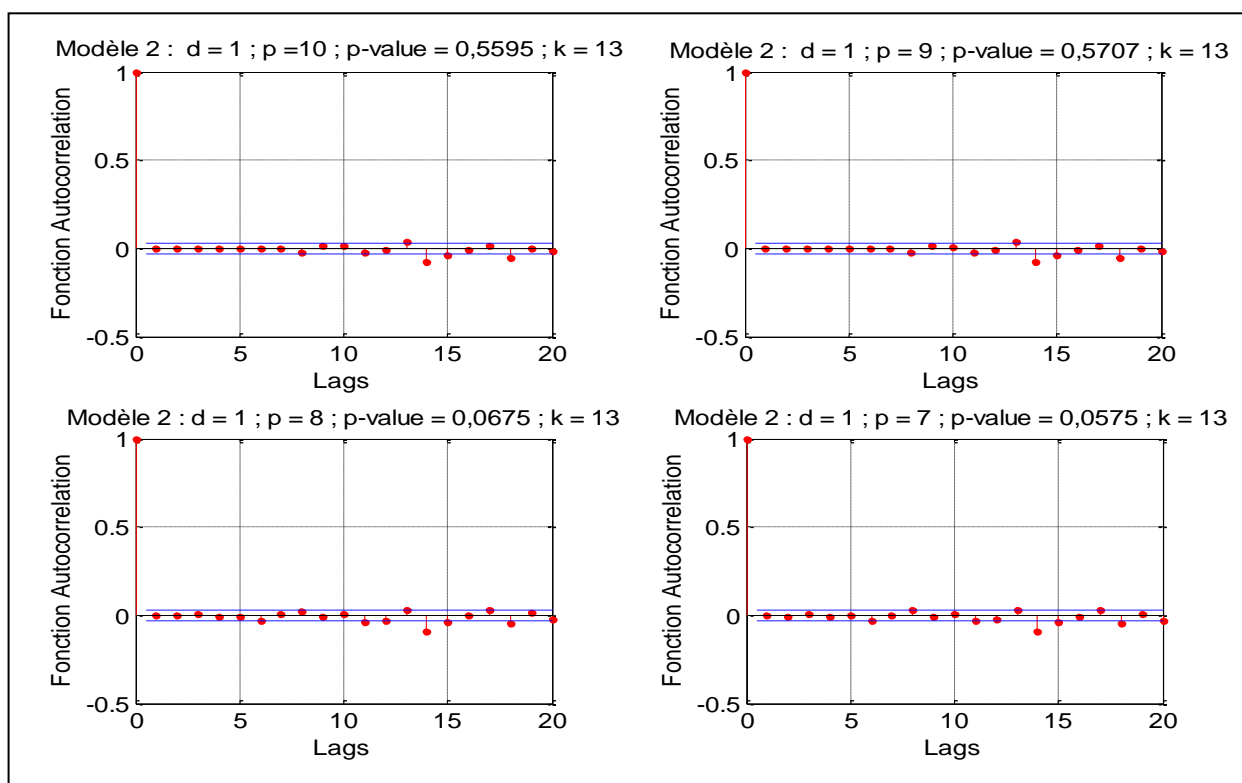


Figure 3.44 : Courbes des résidus du modèle 2 pleine mers pour $d = 1$ et p variant de 7 à 10

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.45 et 3.46 pour le modèle 3 pleine mer.

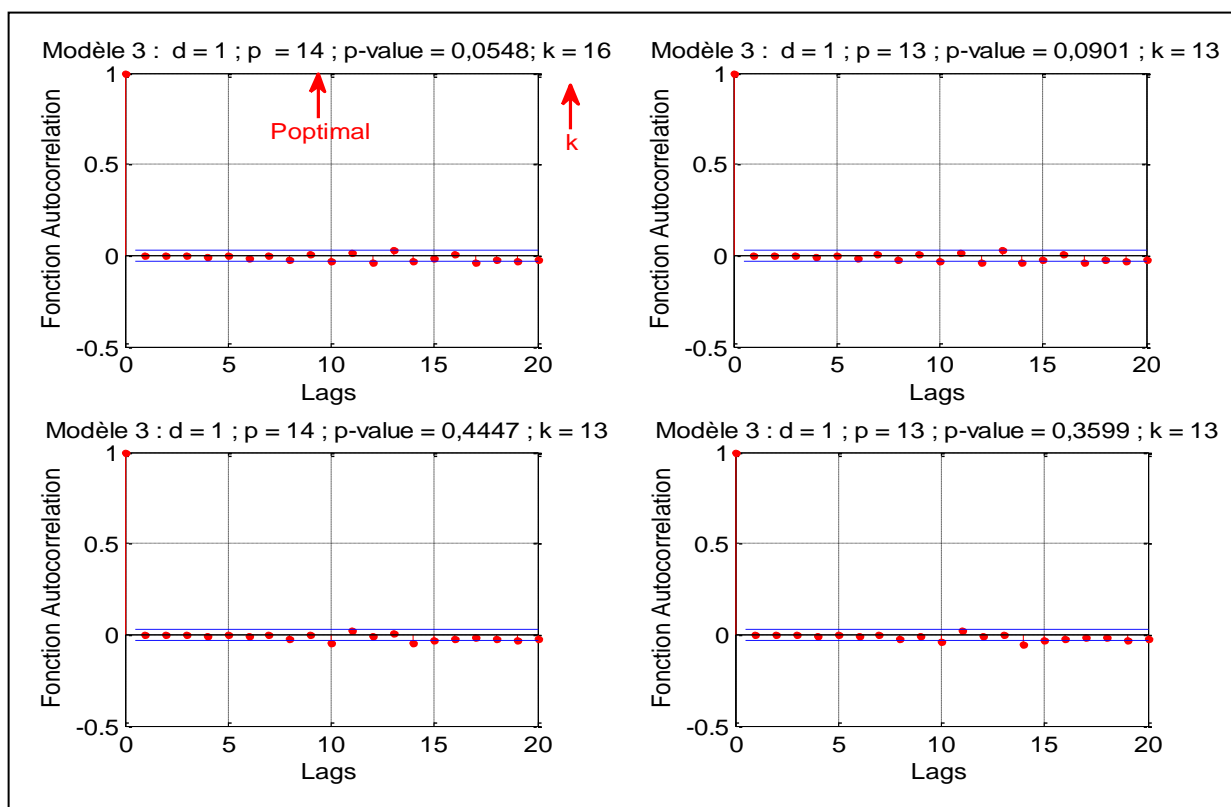


Figure 3.45 : Courbes des résidus du modèle 3 pleine mers pour $d = 1$ et p variant de 11 à 14

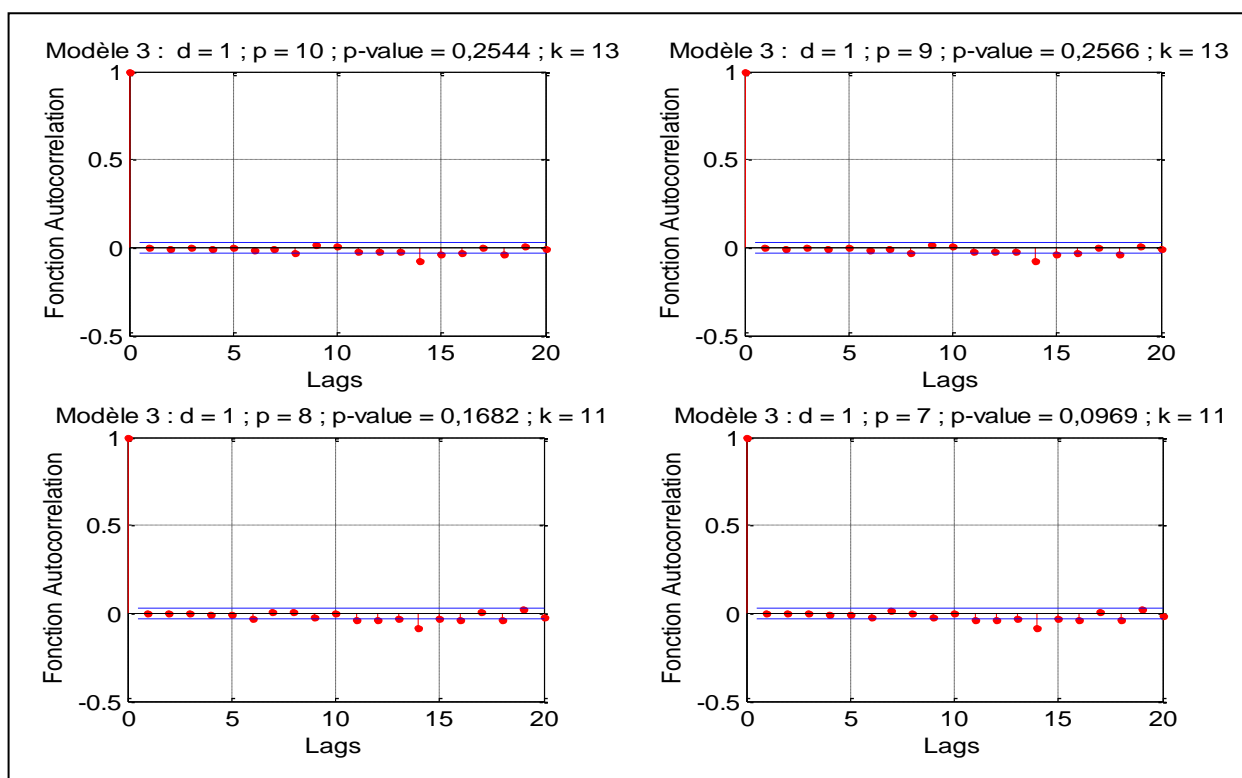


Figure 3 . 46 : Courbes des résidus du modèle 3 pleine mers pour $d = 1$ et p variant de 7 à 10

En fixant « $d = 1$ » et en diminuant progressivement « p_{initial} » entre 14 et 7, nous sommes arrivés aux résultats des figures 3.47 et 3.48 pour le modèle 4 pleine mer.

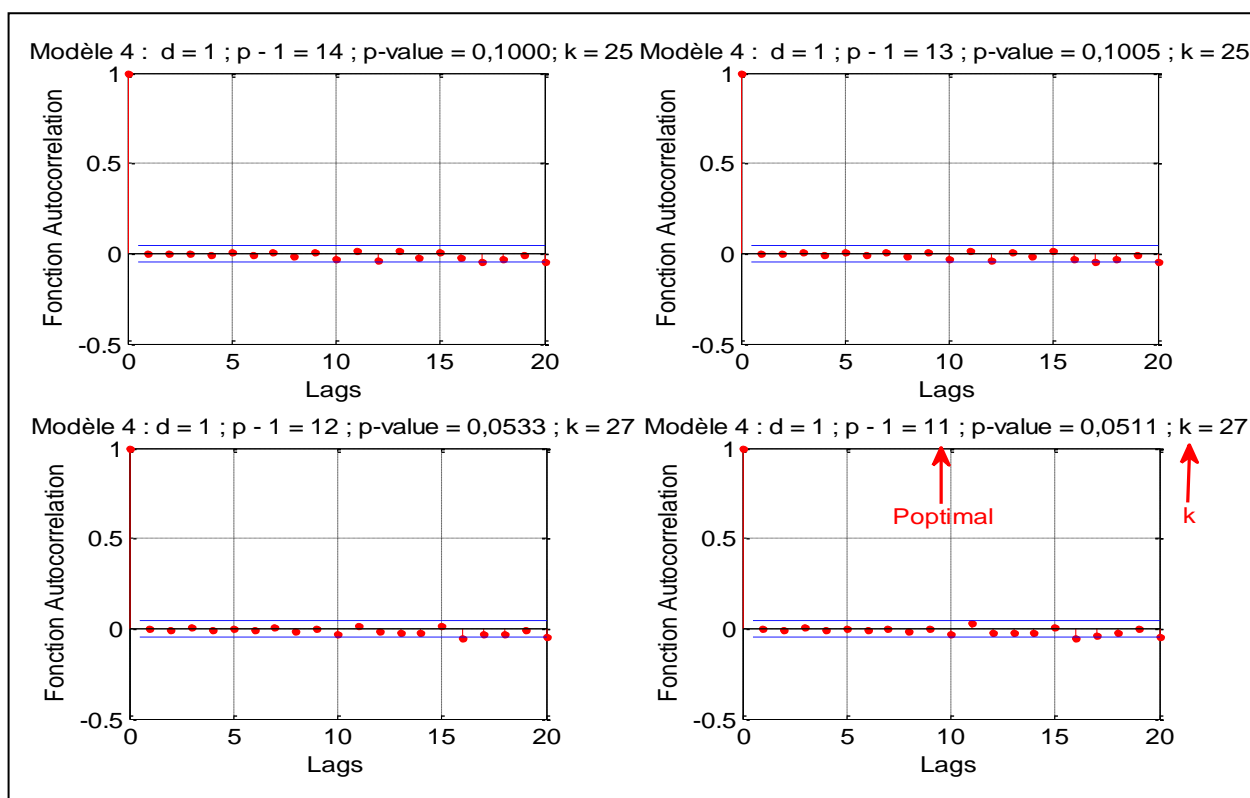


Figure 3 . 47 : Courbes des résidus du modèle 4 pleine mers pour $d = 1$ et p variant de 11 à 14

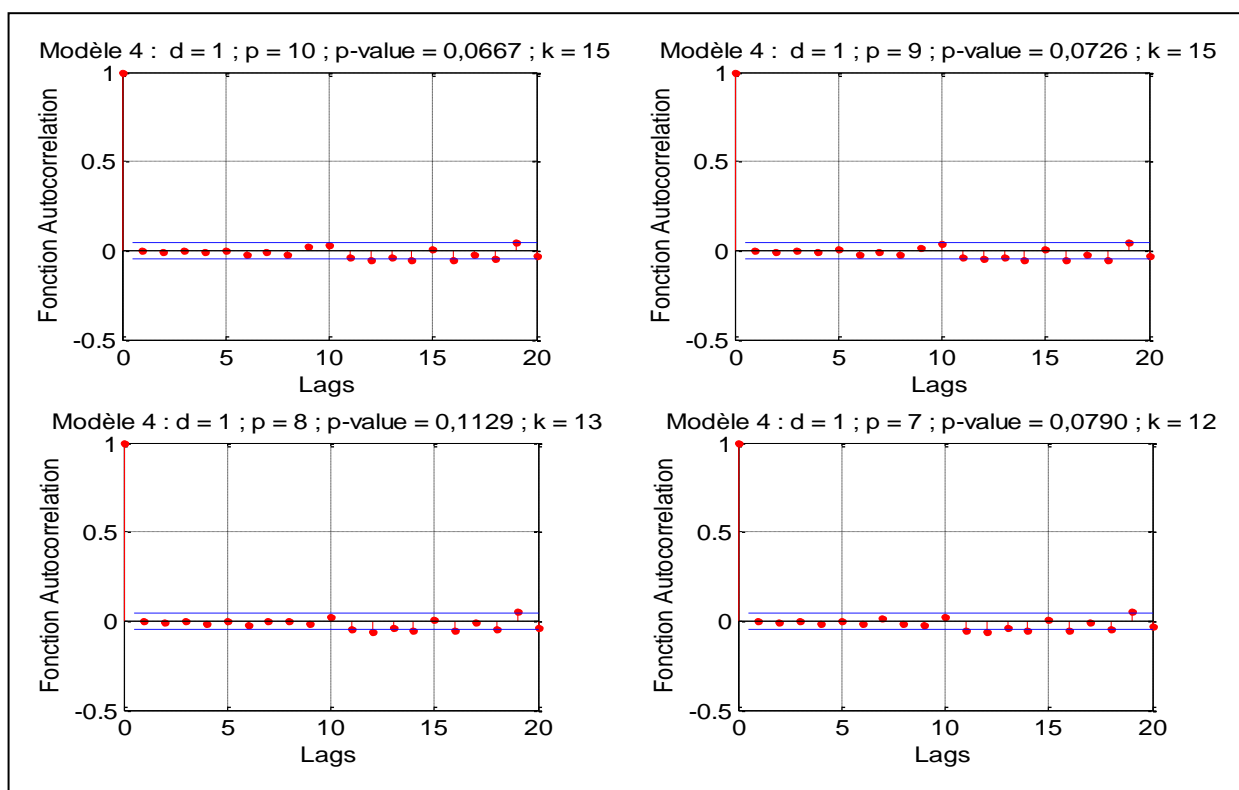


Figure 3.48 : Courbes des résidus du modèle 4 pleine mers pour d = 1 et p variant de 7 à 10

En fixant « d = 2 » et en diminuant progressivement « p_{initial} » entre 14 et 11, nous sommes arrivés aux résultats de la figure 3.49 pour le modèle 1 pleine mer.

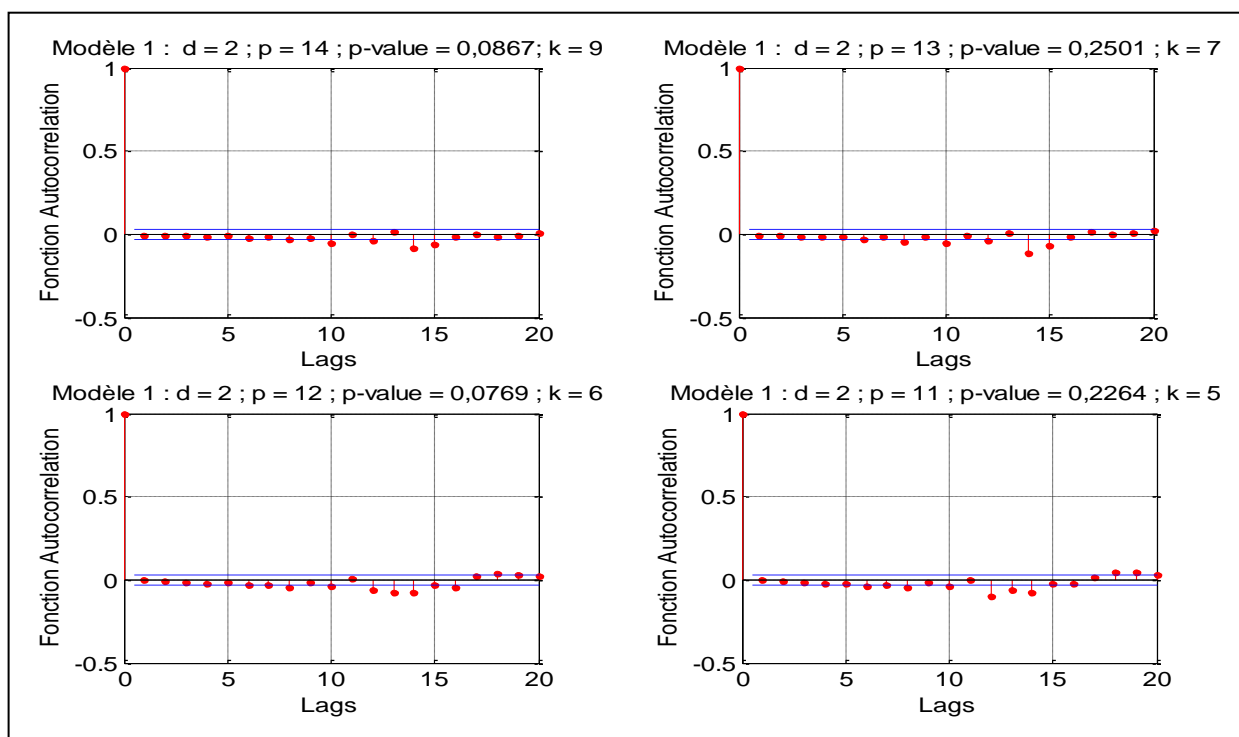


Figure 3.49 : Courbes des résidus du modèle 1 pleine mers pour d = 2 et p variant de 11 à 14

En fixant « $d = 2$ » et en diminuant progressivement « p_{initial} » entre 14 et 11, nous sommes arrivés aux résultats de la figure 3.50 pour le modèle 2 pleine mer.

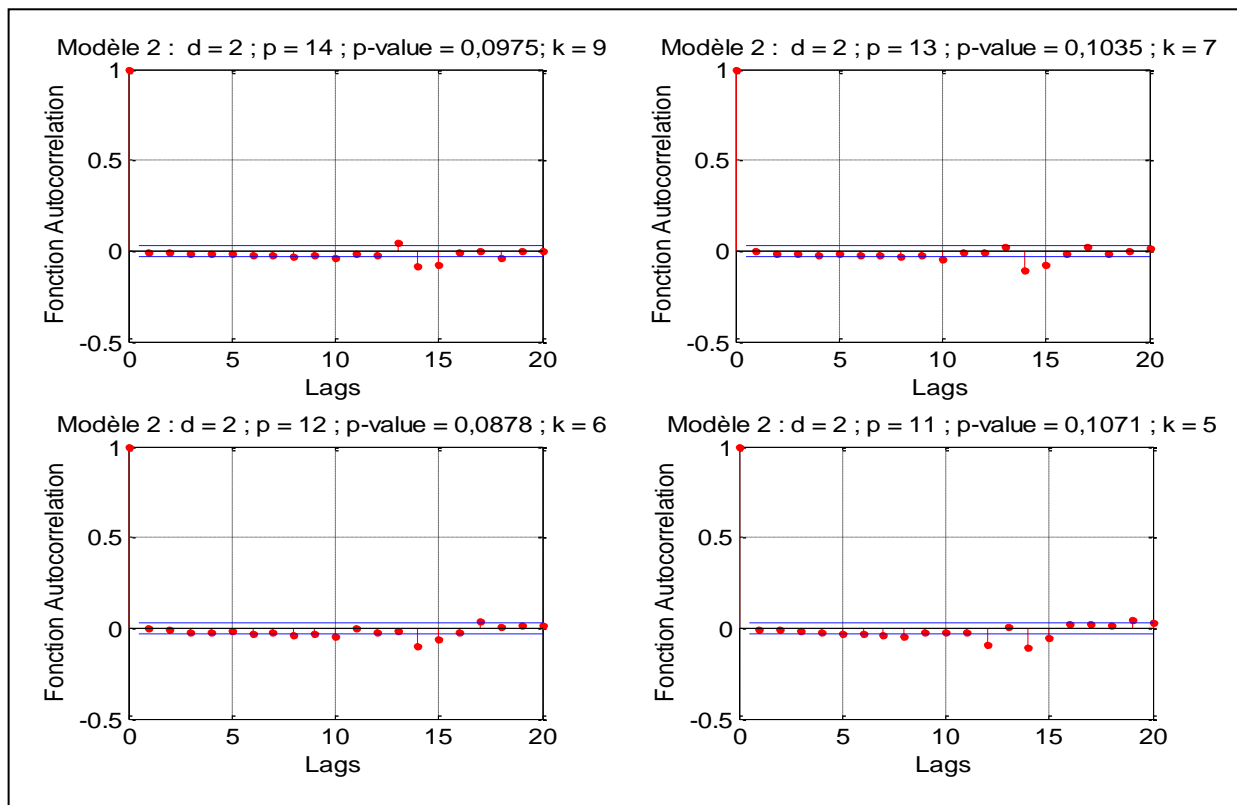


Figure 3 . 50 : Courbes des résidus du modèle 2 pleine mers pour $d = 2$ et p variant de 11 à 14

Nous avons résumé les paramètres définis aux figures 3.41 à 3.50 pour obtenir les résultats du tableau 3.19 suivant :

Tableau 3 . 19 : Test de Ljung-Box pour les quatre modèles sur les pleines mers

	p	d = 1		d = 2			d = 1		d = 2			d = 1			d = 1	
		p-value	k	p-value	k		p-value	k	p-value	k		p-value	k		p-value	k
Modèle 1	14	0,0750	20	0,0867	9	Modèle 2	0,0553	17	0,0975	9	Modèle 3	0,0548	16	Modèle 4	0,3682	20
	13	0,1445	20	0,2501	7		0,5021	13	0,1035	7		0,0901	13		0,4136	20
	12	0,0622	19	0,0769	6		0,5141	13	0,0878	6		0,4447	13		0,4223	20
	11	0,0628	19	0,2264	5		0,5131	13	0,1071	5		0,3599	13		0,3611	20
	10	0,0683	15				0,5595	13				0,2544	13		0,0667	15
	9	0,1114	15				0,5707	13				0,2566	13		0,0726	15
	8	0,0614	13				0,0675	13				0,1682	11		0,1129	13
	7	0,0707	13				0,0575	13				0,0969	11		0,0790	12

Les parties en jaune du tableau 3.19 désigne le modèle qui s'approche de la réalité et qui minimise la valeur de p dans le modèle ARIMA(p,d,q).

D'après ces résultats, nous pouvons donc écrire que :

- pour le modèles 1 de la chronique sur les pleines mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 13$ » ;
- pour le modèles 2 de la chronique sur les pleines mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 14$ » ;
- pour le modèles 3 de la chronique sur les pleines mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 14$ » ;
- et pour le modèles 4 de la chronique sur les pleines mers « $d_{\text{final}} = 1$ et $p_{\text{optimal}} = 11$ ».

En résumé, nous avons trouvé finalement les résultats du tableau 3.20 pour nos 07 modèles avec constantes et tendances :

Tableau 3 . 20 : « Ordre d'intégration final d_{final} et p_{optimal} »

Chronique	Modèles	« d_{final} »	« p_{optimal} »	Ljung-Box test	
				p-value	k
Peine mer	Modèles 1	1	13	0,1445	20
	Modèles 2	1	14	0,0553	17
	Modèles 3	1	14	0,0548	16
	Modèles 4	1	11	0,0511	27
Basse mer	Modèles 1	1	9	0,0539	19
	Modèles 2	1	14	0,3078	17
	Modèles 3	1	13	0,0908	19

Maintenant, nous avons mis en œuvre le test de Ljung-Box pour déterminer si nos modèles possèdent ou non des tendances et des constantes et nous sommes arrivés aux résultats des tableaux 3.21 et 3.22.

Tableau 3 . 21 : Test de significativité des coefficients : Modèles pleines mers

Pleine mer pour " d = 1 "											
Modèle 1			Modèle 2			Modèle 3			Modèle 4		
Paramètres	pvalue	k	Paramètres	pvalue	k	Paramètres	pvalue	k	Paramètres	pvalue	k
$13+ p_b + p_a + p_\Phi$	0,1445	20	$14+ p_b + p_a + p_\Phi$	0,0553	17	$14+ p_b + p_a + p_\Phi$	0,0548	16	$11+ p_b + p_a + p_\Phi$	0,0511	27
$13+ p_a + p_\Phi$	0,1445	20	$14+ p_a + p_\Phi$	0,0553	17	$14+ p_a + p_\Phi$	0,0548	16	$11+ p_a + p_\Phi$	0,0513	27
$13+ p_b + p_\Phi$	0,1445	20	$14+ p_b + p_\Phi$	0,0553	17	$14+ p_b + p_\Phi$	0,0548	16	$11+ p_b + p_\Phi$	0,0513	27
$13+ p_a$	0,0814	20	$14+ p_a$	0,0747	12	$14+ p_a$	0,0547	11	$11+ p_a$	0,1872	27
$13+ p_b$	0,0814	20	$14+ p_b$	0,0747	12	$14+ p_b$	0,0547	11	$11+ p_b$	0,1872	27
$13+ p_\Phi$	0,1445	20	$14+ p_\Phi$	0,0553	17	$14+ p_\Phi$	0,0548	16	$11+ p_\Phi$	0,0514	27
13	0,0814	20	14	0,0747	12	14	0,0547	11	11	0,1872	27

Tableau 3 . 22 : Test de significativité des coefficients : Modèles basses mers

Basse mer pour " d = 1 "								
Modèle 1			Modèle 2			Modèle 3		
Paramètres	p-value	k	Paramètres	p-value	k	Paramètres	p-value	k
$9+p_b+p_a+p_\Phi$	0,0539	19	$14+ p_b + p_a + p_\Phi$	0,3078	17	$13+ p_b + p_a + p_\Phi$	0,0908	19
$9+a+ p_\Phi$	0,0538	19	$14+ p_a + p_\Phi$	0,3078	17	$13+ p_a + p_\Phi$	0,0909	19
$9+ p_b + p_\Phi$	0,0538	19	$14+ p_b + p_\Phi$	0,3078	17	$13+ p_b + p_\Phi$	0,0909	19
$9+ p_a$	0	0	$14+ p_a$	0	0	$13+ p_a$	0,0887	12
$9+ p_b$	0	0	$14 + p_b$	0	0	$13 + p_b$	0,0887	12
$9 + p_\Phi$	0,0538	19	$14 + p_\Phi$	0,3078	17	$13 + p_\Phi$	0,0909	19
9	0	0	14	0	0	13	0,0887	12

où p_b est le nombre de paramètre nécessaire pour estimer la tendance « b » du modèle DFA ;
 p_a le nombre de paramètre nécessaire pour estimer la constante « a » du modèle DFA ;
et p_Φ le nombre de paramètre nécessaire pour estimer le coefficient du premier retard « Φ » du modèle DFA.

Les résultats du tableau 3.21 et 3.22 nous recommandent de prendre pour l'estimation des basses mers et des pleines mers à Hell-Ville des modèles ARIMA(p+1,d,q) sans constante ni tendance.

3.2.13.2. Identification de $q_{initial}$

Après élimination des coefficients non significatifs dans le modèle ARIMA($p_{optimal}+1,d_{final},0$), nous avons pu tracer sur la figure 3.51 les fonctions d'autocorrélation des résidus des 04 modèles sur les pleines mers.

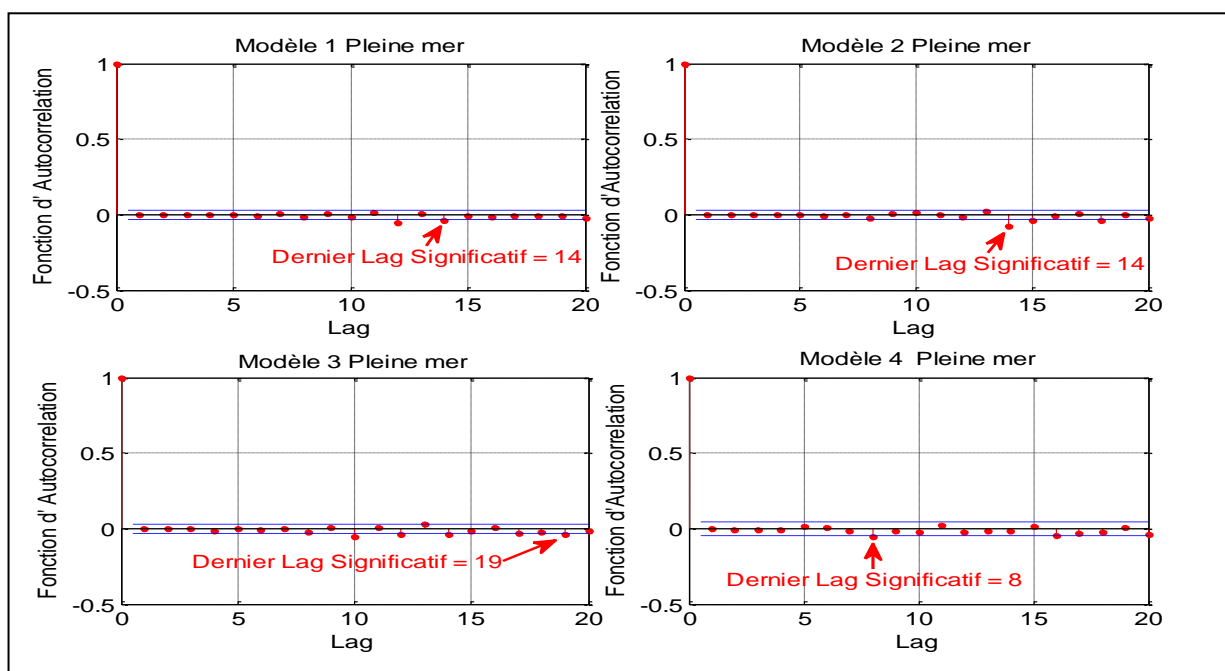


Figure 3 . 51 : Fonctions d'autocorrélation des quatre modèles sur les pleines mers

Les résultats de la figure 3.2.51 nous permettent de déduire que :

- pour le modèle 1 pleine mer q_{initial} est égal à 14 ;
- pour le modèle 2 pleine mer q_{initial} est égal à 14 ;
- pour le modèle 3 pleine mer q_{initial} est égal à 19 ;
- et pour le modèle 4 pleine mer q_{initial} est égal à 8.

Nous avons également pu tracer sur la figure 3.52 les fonctions d'autocorrélation des résidus de nos 03 modèles sur les basses mers.

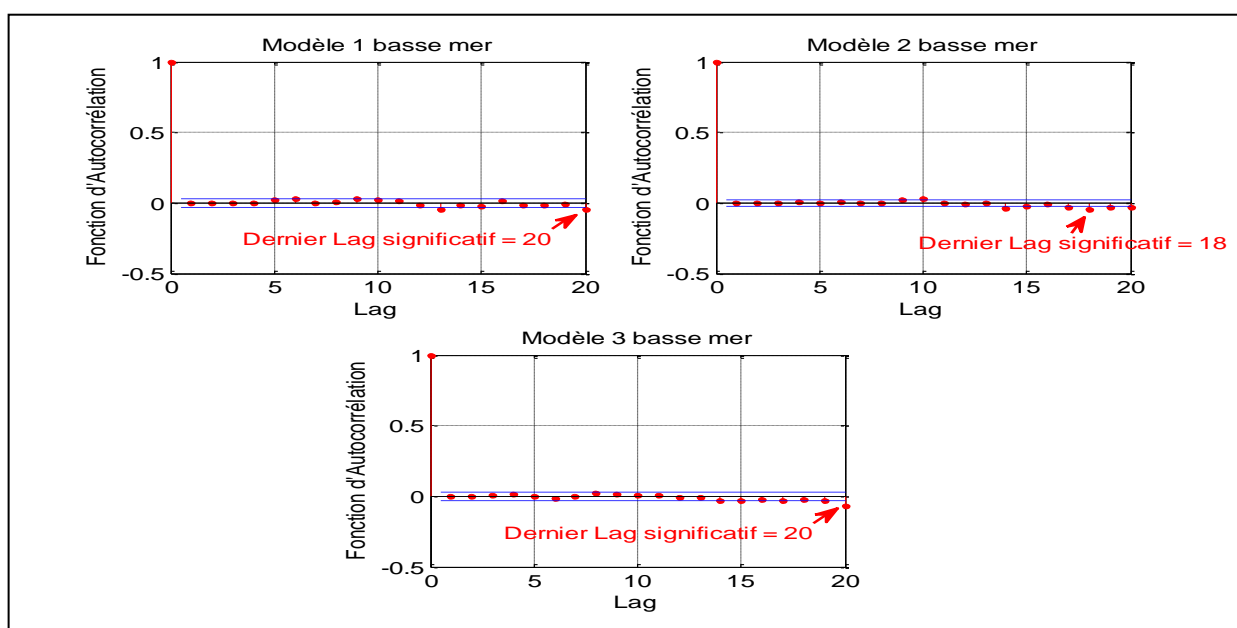


Figure 3 . 52 : Fonctions d'autocorrélation des trois modèles sur les basses mers

Les résultats de la figure 3.52 permet de conclure que :

- pour le modèle 1 basse mer q_{initial} est égal à 20 ;
- pour le modèle 2 basse mer q_{initial} est égal à 18 :
- et pour le modèle 3 basse mer q_{initial} est égal à 20.

3.2.13.3. Choix du modèle ARIMA($p+1,d,q$)

En fixant p_{optimal} et d_{final} et en variant progressivement l'ordre q entre 0 et q_{initial} , nous sommes arrivés aux résultats du tableau 3.23 pour les tests d'identification de q_{optimal} pour les modèles sur les basses mers et aux résultats du tableau 3.24 pour les tests d'identification de q_{optimal} pour les modèles sur les pleines mers.

Tableau 3 . 23 : Identification de q_{optimal} pour les trois modèles sur les basses mers

Basse mer : d = 1											
Modèle 1				Modèle 2					Modèle 3		
q	p_{optimal}	p-value	k	p_{optimal}	p-value	k	AIC	SC	p_{optimal}	p-value	k
20	10	0,2026	18						14	0,2009	17
19	10	0,1315	12						14	0,3024	14
18	10	0,088	8						14	0,958	16
17	10	0,0553	12						14	0,9054	16
16	10	0,1279	24	15	0,3877	12			14	0,0529	19
15	10	0,1292	13	15	0,3038	11			14	0,0967	19
14	10	0,082	15	15	0,093	9			14	0,113	16
13	10	0,2217	7	15	0,6423	14			14	0,4747	12
12	10	0,2397	14	15	0,1056	8			14	0,2354	18
11	10	0,0506	17	15	0,0591	12			14	0,421	18
10	10	0,1269	18	15	0,5094	12			14	0,3691	16
9	10	0,0854	12	15	0,105	5			14	0,2389	24
8	10	0,0683	19	15	0,3325	11			14	0,3233	19
7	10	0,0603	12	15	0,0783	10			14	0,9017	11
6	10	0,1893	17	15	0,6345	12			14	0,2318	19
5	10	0,0725	17	15	0,0913	12			14	0,4068	19
4	10	0,119	19	15	0,079	17	-22450	-22416	14	0,3523	19
3	10	0,0943	19	15	0,311	17	-22440	-22408	14	0,4045	19
2	10	0,1291	19	15	0,2972	17	-22442	-22411	14	0,2931	19
1	10	0,3068	12	15	0,308	17	-22444	-22415	14	0,0511	19
0	10	0,3027	12	15	0,2833	17	-22445	-22418	14	0,0672	19

Tableau 3 . 24 : Identification de q_{optimal} pour les quatre modèles sur les pleines mers

	Pleine mer : d =1																						
	Modèle 1					Modèle 2					Modèle 3			Modèle 4									
q	p _{opt}	pvalue	k	AIC	SC	p _{opt}	pvalue	k	AIC	SC	p _{opt}	pvalue	k	P _{opt}	pvalue	k	AIC	SC					
19											14	0,086	10										
18											14	0,078	9										
17											14	0,332	6										
16											14	0,140	1										
15											14	0,055	5										
14	14	0,076	8			15	0,314	3			14	NON FAISABLE											
13	14	0,151	7			15	0,058	5			14	0,131	7										
12	14	0,442	6			15	0,081	4			14	0,824	3										
11	14	0,646	5			15	0,052	9			14	0,983	1										
10	14	0,108	7			15	0,429	2			14	0,067	20										
9	14	0,398	2			15	0,055	6			14	0,087	3										
8	14	0,379	9			15	0,396	7			14	0,530	5	12	0,074	24							
7	14	0,061	12			15	0,839	13	-21993	-21956	14	0,571	7	12	0,099	17							
6	14	0,000	0			15	0,168	6			14	0,120	5	12	0,151	14							
5	14	0,262	1			15	0,605	2			14	0,066	11	12	0,470	20							
4	14	0,452	10			15	0,412	5			14	0,601	2	12	0,321	20							
3	14	0,967	4			15	0,374	13	-21954	-21924	14	0,727	2	12	0,104	20							
2	14	0,164	11			15	0,172	13	-21871	-21843	14	0,000	0	12	0,119	20							
1	14	0,069	20	-20675	-20650	15	0,964	13	-21851	-21824	14	0,051	8	12	0,075	25	-8773	-8756					
0	14	0,055	20	-20684	-20661	15	0,926	13	-21853	-21828	14	0,133	11	12	0,093	25	-8774	-8758					

En maximisant l'ordre « k » au test de Ljung-Box et les critères d'informations d'Akaike ou les Critères de Schwarz des tableaux 3.23 et 3.24, nous pouvons conclure en nous référant aux parties en jaune que :

- **le modèle 1 pleine mer** est un ARIMA (14,1,1) sans constante ni tendance ;
- **le modèle 2 pleine mer** est un ARIMA (15,1,1) sans constante ni tendance ;
- **le modèle 3 pleine mer** est un ARIMA (14,1,10) sans constante ni tendance ;
- **le modèle 4 pleine mer** est un ARIMA (12,1,1) sans constante ni tendance ;
- **le modèle 1 basse mer** est un ARIMA (10,1,16) sans constante ni tendance ;
- **le modèle 2 basse mer** est un ARIMA (15,1,3) sans constante ni tendance ;
- **et le modèle 3 basse mer** est un ARIMA (14,1,9) sans constante ni tendance.

3.2.13.4. Analyse des résidus

a) Test ARCH

Nous avons pu tracer à la figure 3.53 les fonctions d'autocorrélation des résidus et des résidus au carré.

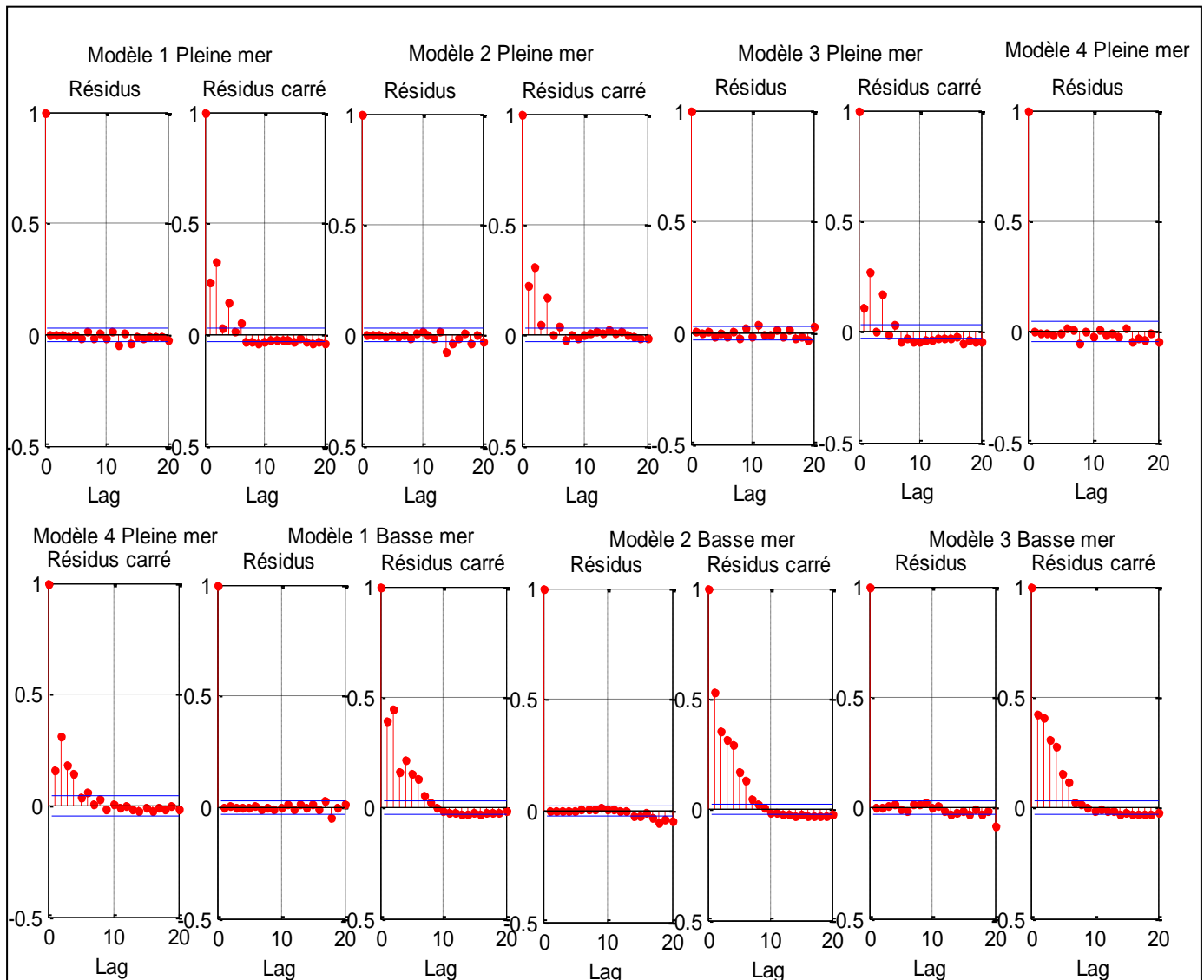


Figure 3 . 53 : Courbe des résidus et des résidus au carré

En mettant en œuvre les tests de l'effet ARCH sous Matlab, nous avons trouvé pour chaque modèle ARIMA($p+1,1,q$) la valeur 1. L'hétéroscedasticité au niveau des résidus est donc confirmée par les deux méthodes.

L'idée d'une modélisation ARIMA($p+1,1,q$) pur des chroniques sur les pleines et les basses mers à Hell-ville est donc rejetée. Il faut inclure à nos modèles une modélisation GARCH(P,Q) des erreurs pour prendre en compte la variance des résidus au cours du temps.

3.2.13.5. Choix du modèle ARIMA(p+1,1,q)-GARCH(P,Q)

En réalité, les p_{optimal} et q_{optimal} trouvés précédemment rendent les résidus les moins corrélés possibles. Ce sont donc des valeurs optimales à partir desquelles nous nous sommes basés pour trouver par maximum de vraisemblance le modèle GARCH (P,Q), EGARCH(P,Q) ou GJR(P,Q) qui permet de confirmer :

- la non autocorrélation des résidus ;
- la non autocorrélation des résidus standardisés qui est défini par le rapport des résidus sur son écart type ;
- la non autocorrélation du carré des résidus standardisés ;
- et autant que possible de confirmer également la normalité centrée et réduite de la distribution des résidus standardisés.

Nous avons assimilé les P_{initial} et Q_{initial} de nos Modèles GARCH(P,Q) aux derniers Lags significatifs de la fonction d'autocorrélation simple et partielle.

Pour le **modèle 1 basse mer**, nous avons trouvé au tableau 3.25 plusieurs modèles ARIMA(10,1,16)-GARCH(P,Q) qui s'approchent de la réalité en faisant varier P et Q entre 0 et P_{initial} et 0 et Q_{initial} et en testant les trois types de modèles GARCH(P,Q) disponibles sous Matlab 2013a.

Tableau 3 . 25 : Choix du modèle ARIMA-GARCH pour le Modèle 1 Basse mer

	p-value	k	p-value	k	p-value	k	p-value	k	p-value	k
résidus	0,2252	2	0,5196	1	0,1396	2	0,4293	2	0,1454	2
résidus/σ	0,8232	1	0,0877	1	0,9393	1	0,0936	1	0,0893	1
(résidus/σ)²	0,0908	4	0,2668	6	0,0656	2	0,1407	15	0,3726	3
LogL	14797		14812		15029		15064		15120	
	EGARCH(0,5)		EGARCH(0,6)		EGARCH(3,5)		EGARCH(5,5)		EGARCH(4,6)	
 S_k 	-9,2314		-7,5850		-3,9104		-6,4833		-5,9586	
 K_u 	43,1637		40,6453		19,6767		21,2897		17,2555	
JB	1948,3		1709,6		402,4631		495,2854		333,2561	
w = Pd - 0,6745σ	-0,1136		-0,1031		-0,0769		-0,0818		-0,0634	
r = ad - 0,7979σ	-0,0601		-0,0532		-0,0385		-0,0385		-0,0332	
Intervalle de Confiance Moyenne	-0,0540	-0,0039	-0,0970	-0,0464	-0,0676	-0,0106	-0,0689	-0,0117	-0,0526	0,0035
Intervalle de Confiance Ecartype	0,8562	0,8916	0,8641	0,8999	0,9741	1,0144	0,9763	1,0167	0,9587	0,9984
Intervalle de Confiance Variance	0,7327	0,7947	0,7463	0,8095	0,9484	1,0286	0,9528	1,0333	0,9188	0,9965
Moyenne	-0,0290		-0,0717		-0,0391		-0,0403		-0,0246	
Ecartype	0,8739		0,8820		0,9942		0,9965		0,9786	
Variance	0,7637		0,7779		0,9885		0,9930		0,9576	

Le tableau 3.25 nous recommande de prendre le modèle ARIMA(10,1,16)-EGARCH(4,6) pour la prévision du cycle 1 Basse mer car c'est le modèle avec le plus grand LogL égal à 15120. Bien que les différents tests inférentiels rejettent l'hypothèse de normalité des résidus standardisés, il semblerait que la méthode graphique de la figure 3.54 basée sur la vraisemblance du QQ-plot de la loi normale et de la distribution des résidus standardisés dans l'intervalle $[-1,96, 1,96]$ nous suggère de l'accepter au seuil de signification $\alpha = 5\%$.

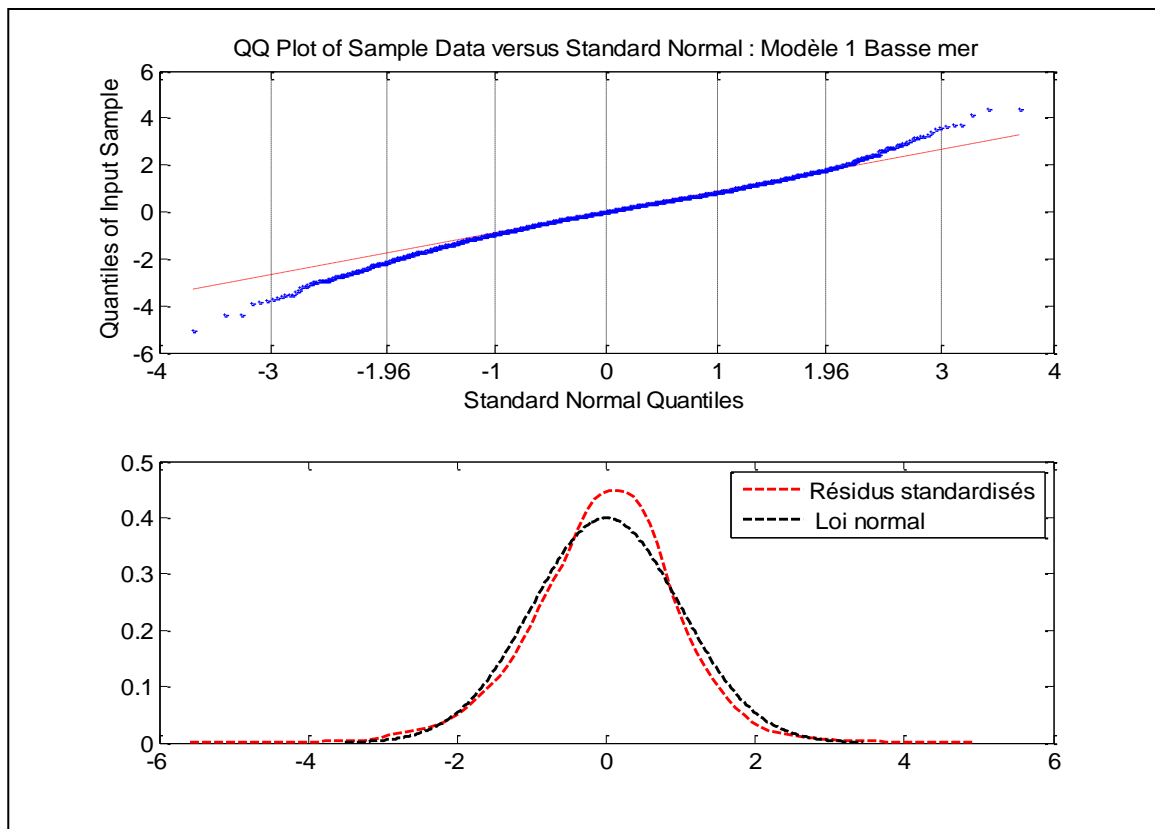


Figure 3.54 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 1 Basse mer

Pour le **modèle 2 basse mer**, nous avons trouvé au tableau 3.26 trois modèles ARIMA(15,1,3)-GARCH(P,Q) qui s'approche de la réalité en faisant varier P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois types de modèles GARCH(P,Q) disponibles sous Matlab 2013a.

Tableau 3 . 26 : Choix du modèle ARIMA-GARCH pour le Modèle 2 Basse mer

	p-value	k	p-value	k	p-value	k
résidus	0,3276	2	0,354	2	0,4383	1
résidus/ σ	0,0834	1	0,0592	1	0,4147	1
(résidus/ σ) ²	0,3192	1	0,1083	3	0,091	12
LogL	19961		19840		20177	
	EGARCH(5,2)		GARCH(0,3)		EGARCH(6,5)	
S _k	-9,8848		-10,5009		-9,2366	
K _u	31,8644		49,5278		31,3861	
JB	1113		2563,3		1070,4	
w = Pd - 0,6745 σ	-0,1101		-0,1202		-0,1109	
r = ad - 0,7979 σ	-0,053		-0,0613		-0,0544	
Intervalle de Confiance Moyenne	-0,0494	-0,0002	-0,0700	-0,0208	-0,0688	-0,0190
Intervalle de Confiance Ecartype	0,9827	1,0175	0,9829	1,0177	0,9946	1,0298
Intervalle de Confiance Variance	0,9654	1,0349	0,9657	1,0353	0,9890	1,0602
Moyenne	-0,0248		-0,0454		-0,0439	
Ecartype	1,0001		1,0003		1,0122	
Variance	1,0001		1,0005		1,0246	

Le tableau 3.26 recommande de prendre le modèle ARIMA(15,1,3)-EGARCH(6,5) pour la prévision du cycle 2 Basse mer car c'est le modèle avec le plus grand LogL égal à 20177.

La figure 3.55 montre la vraisemblance entre le QQ-plot de la loi normale et le QQ-plot de la distribution des résidus standardisés dans l'intervalle [-1,96 , 1,96]. Nous pouvons donc accepter au seuil de signification $\alpha = 5\%$ la normalité des résidus standardisés.

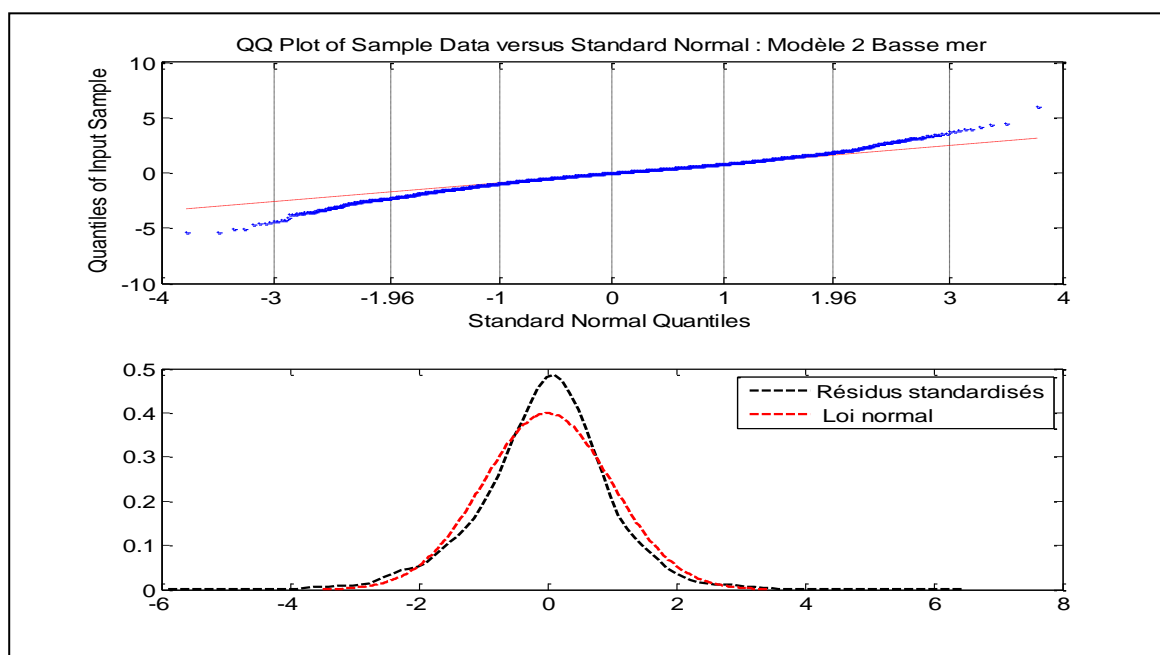


Figure 3 . 55 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 2 Basse mer

Pour le **modèle 3 basse mer**, nous avons trouvé au tableau 3.27 un seul modèle ARIMA(14,1,9)-GARCH(P,Q) qui s'approche de la réalité en faisant varier P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois types de modèles GARCH(P,Q) disponibles sous Matlab 2013a

Tableau 3 . 27 : Choix du modèle ARIMA-GARCH pour le Modèle 3 Basse mer

	p-value	k
résidus	0,0637	1
résidus/σ	0,2993	1
(résidus/σ) ²	0,0882	10
LogL	13376	
	EGARCH(3,7)	
S _k	-6,2265	
K _u	19,9867	
JB	438,2375	
w = Pd – 0,6745σ	-0,085	
r = ad - 0,7979σ	-0,0414	
Intervalle de Confiance Moyenne	-0,0682	-0,0097
Intervalle de Confiance Ecartype	0,9931	1,0345
Intervalle de Confiance Variance	0,9858	1,0697
Moyenne	-0,0390	
Ecartype	1,0006	
Variance	1,0138	

Ainsi le meilleur modèle pour prévoir le cycle 3 Basse mer est donc un ARIMA(14,1,9)-EGARCH(3,7). La vraisemblance de la distribution des résidus standardisés avec la loi normale peut être acceptée au seuil de signification $\alpha = 5\%$ car les QQ-plot des deux distributions de la figure 3.56 sont cohérents dans l'intervalle $[-1,96, 1,96]$.

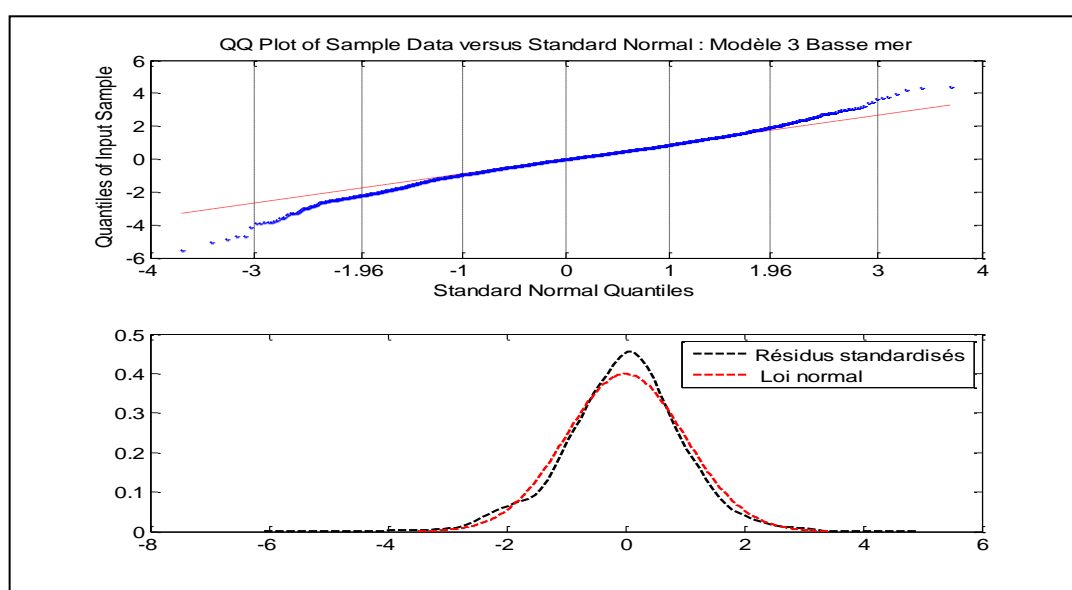


Figure 3 . 56 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 3 Basse mer

Pour le **modèle 1 pleine mer**, le tableau 3.28 montre trois modèles ARIMA(14,1,1)-GARCH(P,Q) qui s'approchent de la réalité en variant P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois modèles GARCH(P,Q) disponibles sous Matlab.

Tableau 3 . 28 : Choix du modèle ARIMA-GARCH pour le Modèle 1 Pleine mer

	p-value	k	p-value	k	p-value	k
résidus	0,8085	1	0,1304	1	0,1055	1
résidus/ σ	0,0597	2	0,2862	1	0,3646	1
(résidus/ σ) ²	0,0532	21	0,407	2	0,3741	2
LogL	18355		18463		18464	
	GJR(7,7)		EGARCH(2,1)		EGARCH(3,1)	
S_k	-11,9795		-11,9192		-12,0305	
K_u	83,3813		94,5871		93,9555	
JB	7096		9088,8		8972,4	
$w = Pd - 0,6745\sigma$	-0,1991		-0,2118		-0,2115	
$r = ad - 0,7979\sigma$	-0,1137		-0,1273		-0,1272	
Intervalle de Confiance Moyenne	-0,0731	-0,0143	-0,1112	-0,0530	-0,1108	-0,0526
Intervalle de Confiance Ecartype	0,9876	1,0291	0,9776	1,0187	0,9776	1,0187
Intervalle de Confiance Variance	0,9748	1,0587	0,9552	1,0373	0,9553	1,0374
Moyenne	-0,0437		-0,0821		-0,0817	
Ecartype	1,0083		0,9981		0,9982	
Variance	1,0167		0,9963		0,9963	

D'après le tableau 3.28, un ARIMA(14,1,1)-EGARCH(3,1) peut être retenu pour prévoir le cycle 1 Pleine mer car c'est le modèle avec le plus grand LogL égal à 18464.

Les résultats de la figure 3.57 nous permettent d'accepter au seuil de signification $\alpha = 5\%$, la normalité de la distribution des résidus standardisés car le QQ-plot des résidus normalisés et le QQ-plot de la loi normale sont cohérents dans l'intervalle $[-1,96, 1,96]$.

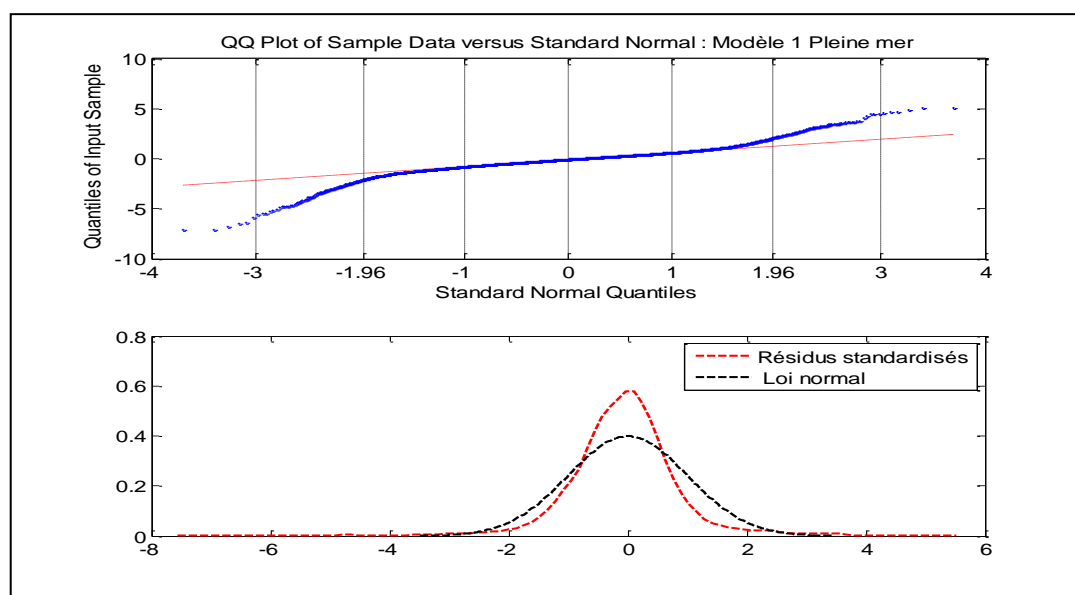


Figure 3 . 57 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 1 Pleine mer

Pour le **modèle 2 pleine mer**, le tableau 3.29 montre deux modèles ARIMA(15,1,1)-GARCH(P,Q) qui s'approchent de la réalité en variant P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois modèles GARCH(P,Q) disponibles sous Matlab 2013a

Tableau 3 . 29 : Choix du modèle ARIMA-GARCH pour le Modèle 2 Pleine mer

	p-value	k	p-value	k
résidus	0,8252	1	0,8546	1
résidus/ σ	0,2512	1	0,1253	3
(résidus/ σ) ²	0,5351	24	0,6208	3
LogL	19428		19595	
	GARCH(1,5)		EGARCH(1,3)	
S_k	-5,4863		-4,6539	
K_u	95,4698		50,3171	
JB	9144,6		2553,5	
$w = Pd - 0,6745\sigma$	-0,1532		-0,1566	
$r = ad - 0,7979\sigma$	-0,0954		-0,0789	
Intervalle de Confiance Moyenne	-0,0898	-0,0320	-0,0543	0,0035
Intervalle de Confiance Ecartype	0,9780	1,0189	0,9777	1,0186
Intervalle de Confiance Variance	0,9561	1,0378	0,9555	1,0371
Moyenne	-0,0609		-0,0254	
Ecartype	0,9985		0,9982	
Variance	0,9970		0,9963	

Le tableau 3.29 montre qu'un ARIMA(15,1,1)-EGARCH(1,3) est valable pour prévoir le cycle 2 Pleine mer car c'est le modèle avec le plus grand LogL égal à 19595.

Les résultats de la figure 3.58 montre encore la vraisemblance de la distribution des résidus standardisés avec la loi normale au seuil de signification $\alpha = 5\%$ car leurs QQ-Plots sont cohérents dans l'intervalle $[-1,96, 1,96]$.

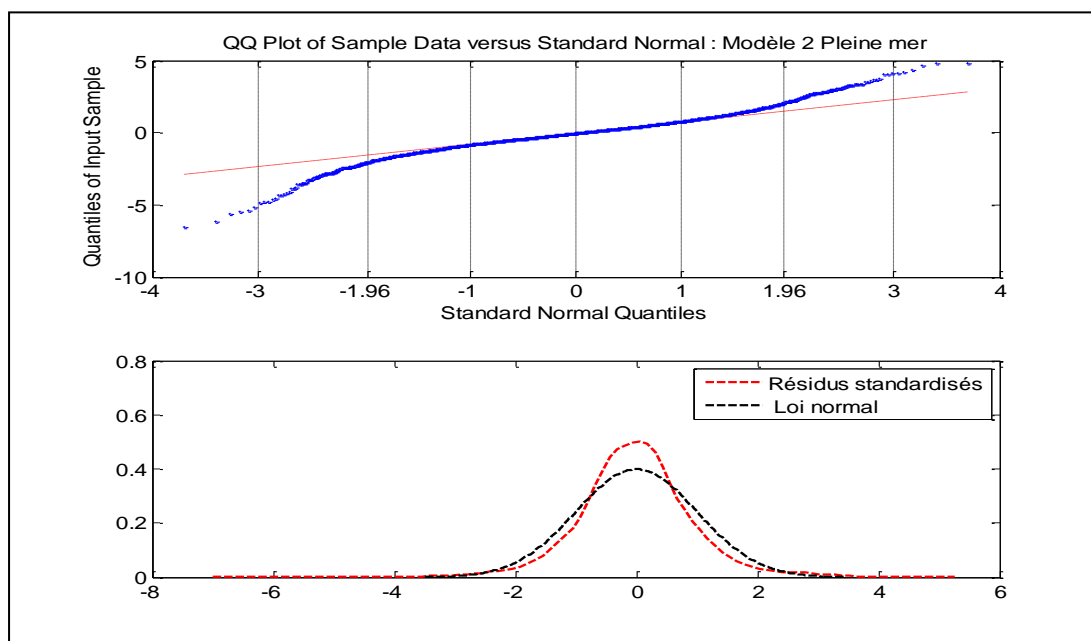


Figure 3 . 58 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 2 Pleine mer

Pour le **modèle 3 pleine mer**, nous avons trouvé au tableau 3.30 cinq modèles ARIMA(14,1,10)-GARCH(P,Q) qui s'approchent de la réalité en variant P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois modèles GARCH(P,Q) disponibles sous Matlab 2013.

Tableau 3 . 30 : Choix du modèle ARIMA-GARCH pour le Modèle 3 Pleine mer

	p-value	k	p-value	k	p-value	k	p-value	k	p-value	k
résidus	0,0725	1	0,2657	1	0,4617	2	0,3595	2	0,3589	2
résidus/ σ	0,9175	1	0,7209	1	0,0812	1	0,3198	1	0,3209	1
(résidus/ σ) ²	0,3427	4	0,1629	3	0,2421	6	0,0843	8	0,084	8
LogL	17059		16960		16984		16998		16998	
	EGARCH(1,5)		GARCH(0,3)		GJR(1,4)		GJR(2,4)		GJR(3,4)	
S_k	-14,489		-26,2015		-23,519		-21,6297		-21,6361	
K_u	42,5869		64,1423		59,2746		55,7401		55,7456	
JB	2023,6		4800,7		4066,6		3574,8		3575,7	
$w = Pd - 0,6745\sigma$	-0,1427		-0,1566		-0,1482		-0,1432		-0,1432	
$r = ad - 0,7979\sigma$	-0,0773		-0,0789		-0,0853		-0,083		-0,083	
Intervalle de Confiance Moyenne	-0,0509	0,0077	-0,0681	-0,0096	-0,0286	0,0302	-0,0373	0,0213	-0,0374	0,0213
Intervalle de Confiance Ecartype	0,9793	1,0208	0,9767	1,0181	0,9809	1,0224	0,9794	1,0209	0,9794	1,0209
Intervalle de Confiance Variance	0,9587	1,0417	0,9535	1,0361	0,9617	1,0449	0,9589	1,0418	0,9588	1,0418
Moyenne	-0,0216		-0,0388		0,0008		0,0080		0,0080	
Ecartype	1,0001		0,9974		1,0016		1,0002		1,0002	
Variance	1,0002		0,9948		1,0033		1,0003		1,0003	

Le tableau 3.30 montre qu'un ARIMA(14,1,10)-EGARCH(1,5) peut être retenu pour prévoir le cycle 3 Pleine mer car il maximise le LogL qui vaut 17059.

Les résultats de la figure 3.59 montre au seuil de signification $\alpha = 5\%$, une concordance entre la loi normale et les résidus standardisés car leurs QQ-plot sont cohérents dans l'intervalle $[-1,96, 1,96]$.

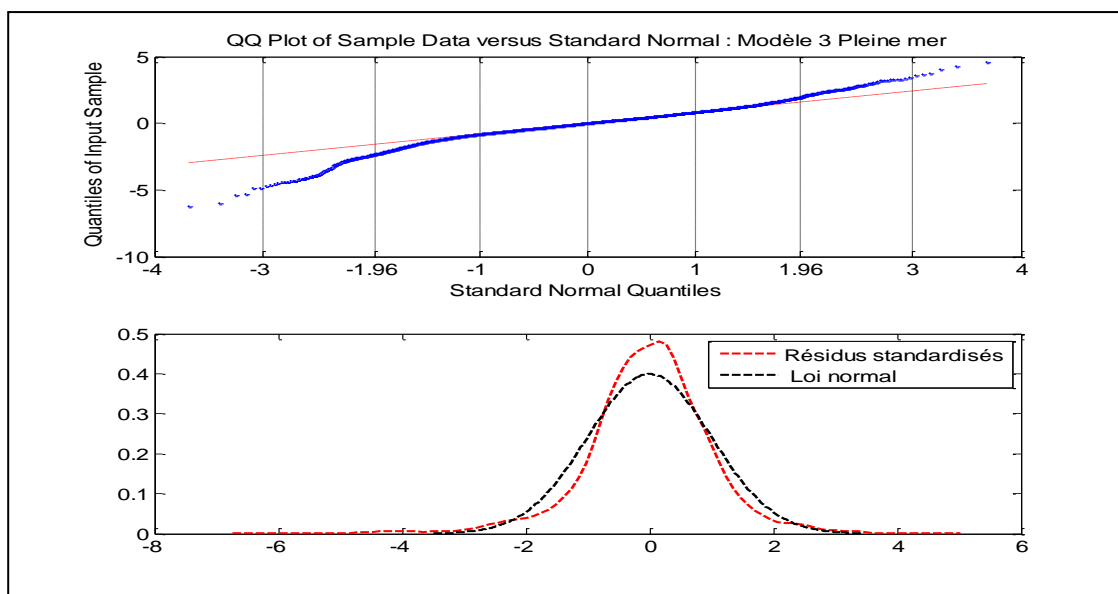


Figure 3 . 59 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale du Modèle 3 Pleine mer

Pour le **modèle 4 pleine mer**, nous avons trouvé au tableau 3.31 quatre modèles ARIMA(12,1,1)-GARCH(P,Q) qui s'approchent de la réalité en variant P et Q entre 0 et P_{\max} et 0 et Q_{\max} et en testant les trois modèles GARCH(P,Q) disponibles sous Matlab 2013a

Tableau 3 . 31 : Choix du modèle ARIMA-GARCH pour le Modèle 4 Pleine mer

	p-value	k	p-value	k	p-value	k	p-value	k
résidus	0,4871	1	0,1683	1	0,2195	1	0,7971	1
résidus/ σ	0,8679	4	0,2474	7	0,1561	7	0,1601	7
(résidus/ σ) ²	0,1061	26	0,4404	2	0,9498	2	0,0858	4
LogL	7705		7744		7747		7832	
	EGARCH(1,7)		EGARCH(3,4)		EGARCH(3,5)		EGARCH(3,6)	
S_k	-10,6047		-3,132		-2,7846		-2,217	
K_u	37,0509		14,8281		15,0094		11,1939	
JB	1485,2		229,6817		233,0348		130,2195	
$w = Pd - 0,6745\sigma$	-0,1401		-0,1249		-0,1256		-0,1118	
$r = ad - 0,7979\sigma$	-0,0803		-0,0566		-0,0567		-0,0451	
Intervalle de Confiance Moyenne	-0,0700	0,0166	-0,0443	0,0420	-0,0436	0,0427	-0,0762	0,0069
Intervalle de Confiance Ecartype	0,9711	1,0323	0,9684	1,0294	0,9680	1,0290	0,9318	0,9905
Intervalle de Confiance Variance	0,9422	1,0648	0,9369	1,0588	0,9361	1,0579	0,8674	0,9803
Moyenne	-0,0267		-0,0011		-0,0005		-0,0347	
Ecartype	1,0017		0,9989		0,9985		0,9612	
Variance	1,0035		0,9979		0,9970		0,9239	

Le tableau 3.31 montre qu'un ARIMA(12,1,1)-EGARCH(3,6) peut être retenu pour prévoir le cycle 4 Pleine mer car c'est le modèle qui maximise le LogL égal à 7832.

Les résultats de la figure 3.60 permet d'assimiler, au seuil de signification $\alpha = 5\%$, la distribution des résidus standardisés à la loi normale car les QQ-plot des deux distributions concordent dans l'intervalle $[-1,96, 1,96]$.

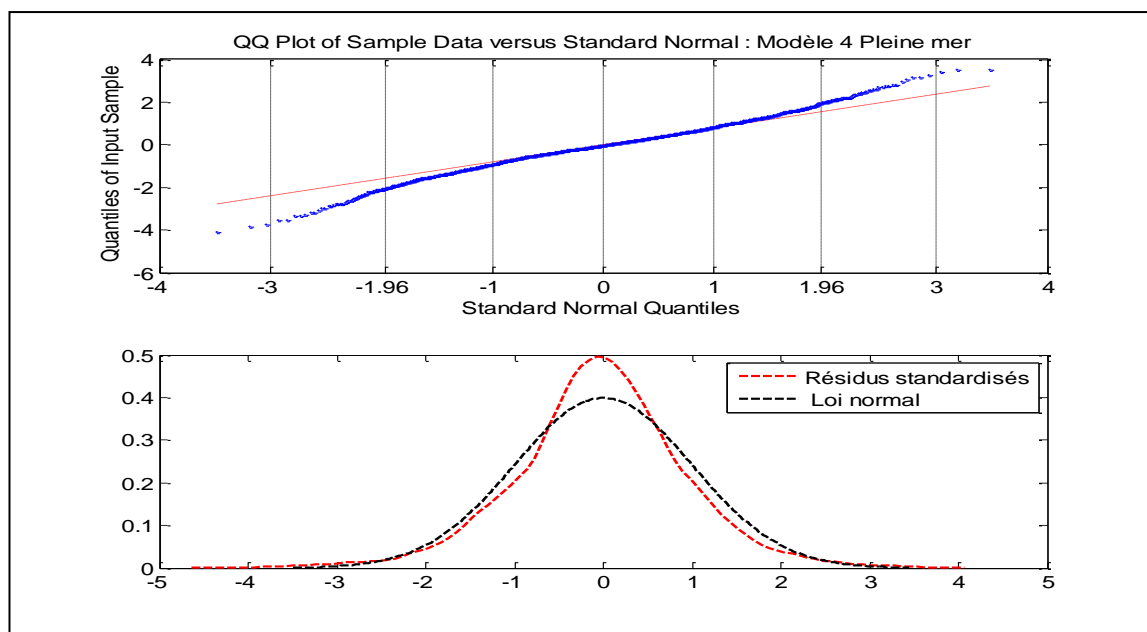


Figure 3 . 60 : QQ-Plot et distribution des résidus standardisés par rapport à la loi normale : Modèle 4 Pleine mer

En résumé, nous avons trouvé que :

- **le modèle 1 pleine mer** est un ARIMA (14,1,1)-EGARCH(3,1) sans constante ni tendance ;
- **le modèle 2 pleine mer** est un ARIMA (15,1,1)-EGARCH(1,3) sans constante ni tendance ;
- **le modèle 3 pleine mer** est un ARIMA (14,1,10)-EGARCH(1,5) sans constante ni tendance ;
- **le modèle 4 pleine mer** est un ARIMA (12,1,1)-EGARCH(3,6) sans constante ni tendance ;
- **le modèle 1 basse mer** est un ARIMA(10,1,16)-EGARCH(4,6) sans constante ni tendance ;
- **le modèle 2 basse mer** est un ARIMA (15,1,3)-EGARCH(6,5) sans constante ni tendance ;
- **le modèle 3 basse mer** est un ARIMA (14,1,9)-EGARCH(3,7) sans constante ni tendance.

Analytiquement, les résidus standardisés de ces sept modèles constituent des bruits blancs centrés et réduits, non gaussiens mais qui s'en rapprochent fortement **et graphiquement**, ces résidus peuvent être assimilés à des bruits blancs à distribution normale, centrée et réduite

3.2.13.6. Analyse de la qualité de la prévision

Pour apprécier la qualité de la prévision, nous avons visualisé graphiquement les résultats de nos prévisions avec les différentes hypothèses énoncées au début du paragraphe 3.2.13.5. Nous avons également comparé les erreurs (RMSE) issues de nos modèles avec les erreurs provenant du logiciel de prédiction du Service Hydrographique Océanographique de la Marine française qui utilise un modèle harmonique [45].

Pour évaluer la performance de nos modèles, nous avons estimé :

- pour chaque cycle sur les basses mers, 268 valeurs futures
- et pour chaque cycle sur les pleines mers, 251 valeurs futures.

En particulier, **pour le modèle 1 Pleines mers**, nous avons obtenu les résultats de la figure 3.61 pour l'estimation de la série corrigée des variations saisonnières différenciée une fois :

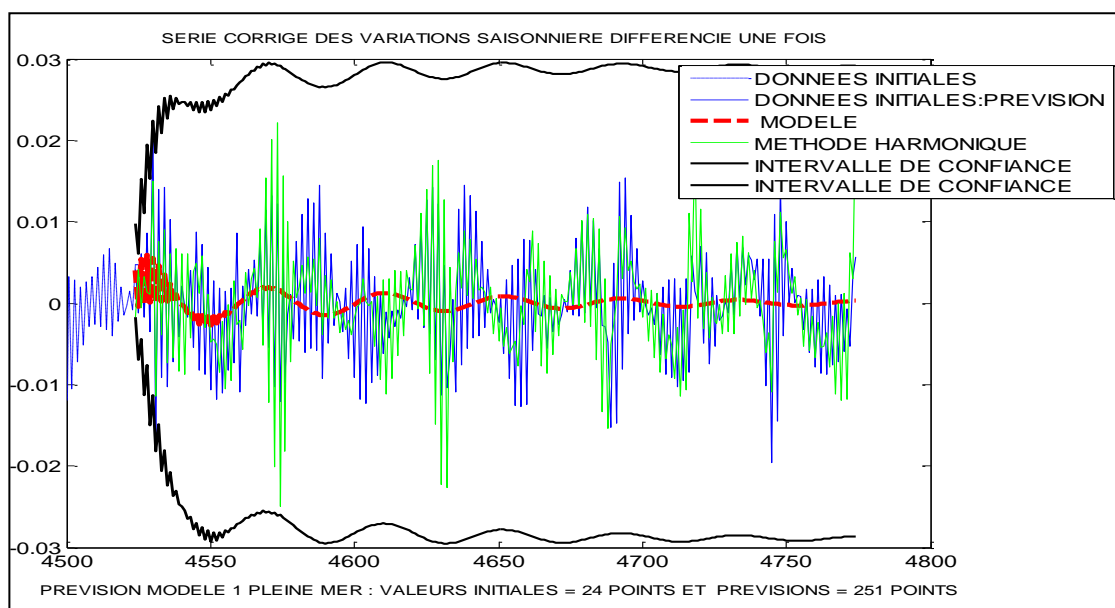


Figure 3 . 61 : Pr vision s rie corrig e des variations saisonni res diff renci e une fois Mod le 1 Pleine mer

En utilisant l  quation 2.96, nous avons pu transformer les r sultats de la figure 3.61 pour arriver   l'estimation de la s rie corrig e des variations saisonni res   la figure 3.62.

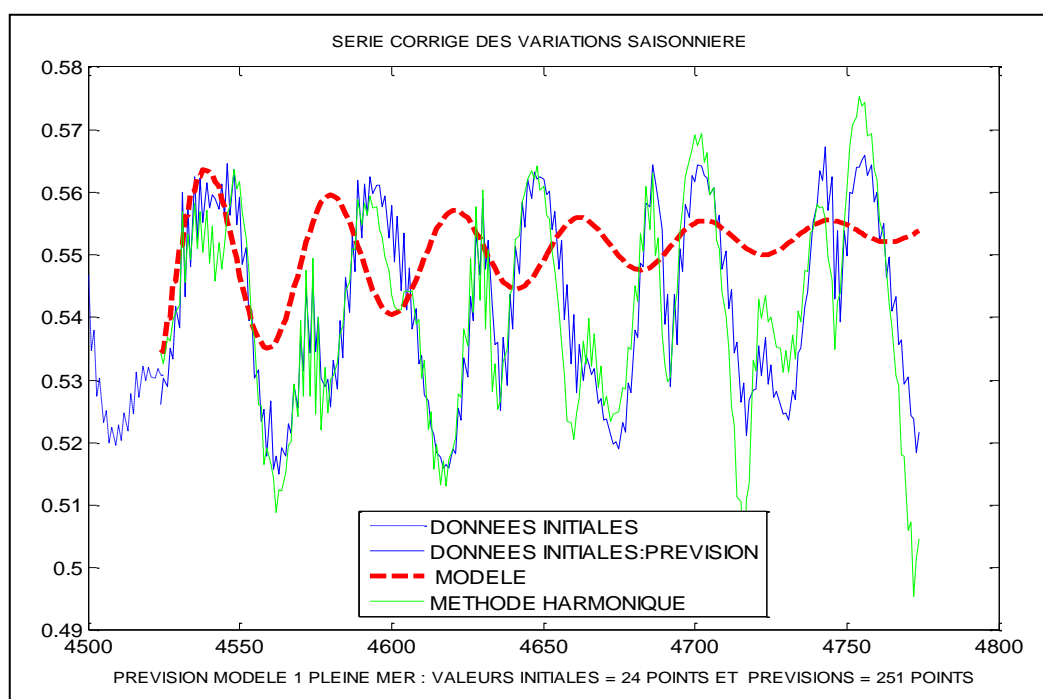


Figure 3 . 62 : Pr vision s rie corrig e des variations saisonni res Mod le 1 Pleine mer

En utilisant l  quation 2.97, nous sommes arriv s   estimer   la figure 3.63 les pr l vements du premier cycle sur les pleines mers.

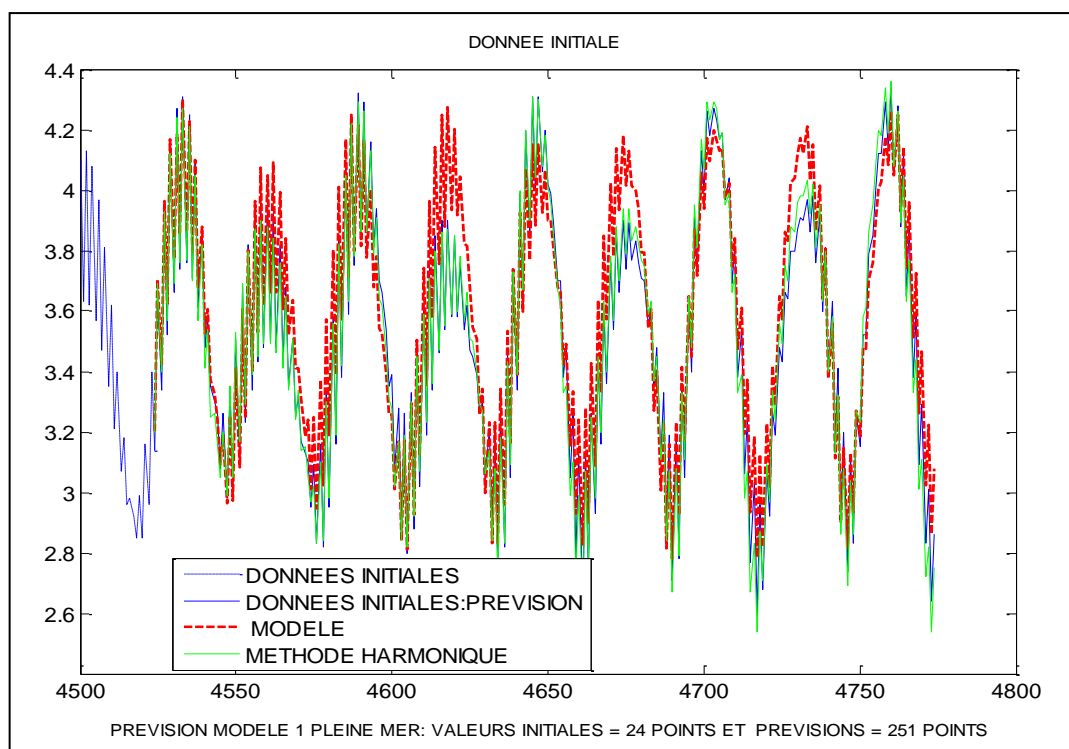


Figure 3 . 63 : Pr vision des donn es initiales Mod le 1

Pour tester la stabilit  du mod le 1 sur les pleines mers et obtenir le graphique de la figure 3.64, nous avons utilis  l' quation 2.101.

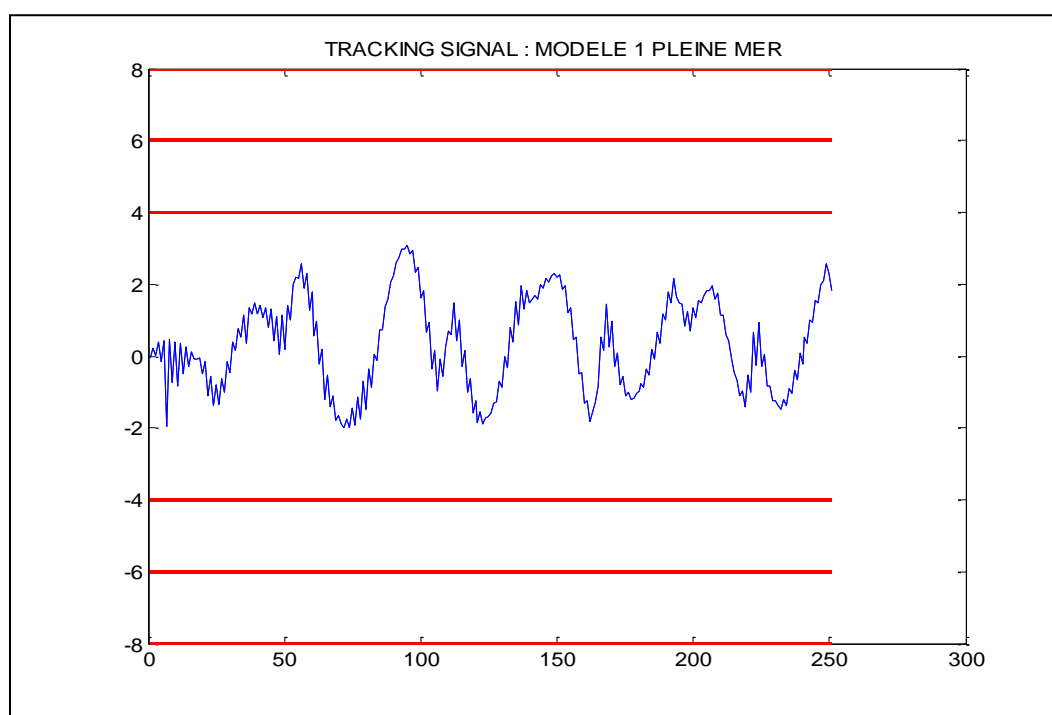


Figure 3 . 64 : Tracking Signal du Mod le 1 Pleine mer

La figure 3.61 sur le modèle 1 Pleine mer montre que l'intervalle de confiance fourni par l'écart type conditionnel arrive à bien cadrer la prévision sur les séries corrigées des variations saisonnières différenciées une fois.

Sur les figures 3.61 et 3.62, nous pouvons apercevoir que notre prévision (couleur en rouge) varie dans le même ordre de grandeur que les données stabilisées (couleur en bleu) et les estimations issues du modèle alternative (couleur en vert).

Sur la figure 3.63, nous pouvons observer une concordance entre les valeurs prévues par notre modèle, les valeurs obtenues par la méthode harmonique et les données initiales.

Comme l'indicateur de Theil qui vaut 2,0251 est supérieur à 1 alors la méthode harmonique offre donc une meilleure précision par rapport à notre modèle.

Mais avec une RMSE moyenne égale à 0,0141, nous pouvons affirmer que notre modèle constitue quand même un excellent modèle puisqu'elle est inférieure à 5%.

De plus, comme la stabilité de notre prévision est confirmé par les résultats du Tracking Signal de la figure 3.64 puisque les valeurs des TS restent à l'intérieur de l'intervalle $[-4, 4]$ alors nous pouvons garder notre modèle pour effectuer la prévision du cycle 1 Pleine mer.

Pour le modèle 2 sur les pleines mers, en procédant de la même façon que pour le modèle 1 sur les pleines mers, nous avons pu tracer sur les figures 3.65 à 3.66, l'estimation de la série corrigée des variations saisonnières différenciée une fois, l'estimation de la série corrigée des variations saisonnières, l'estimation des données initiales et l'estimation du Tracking signal.

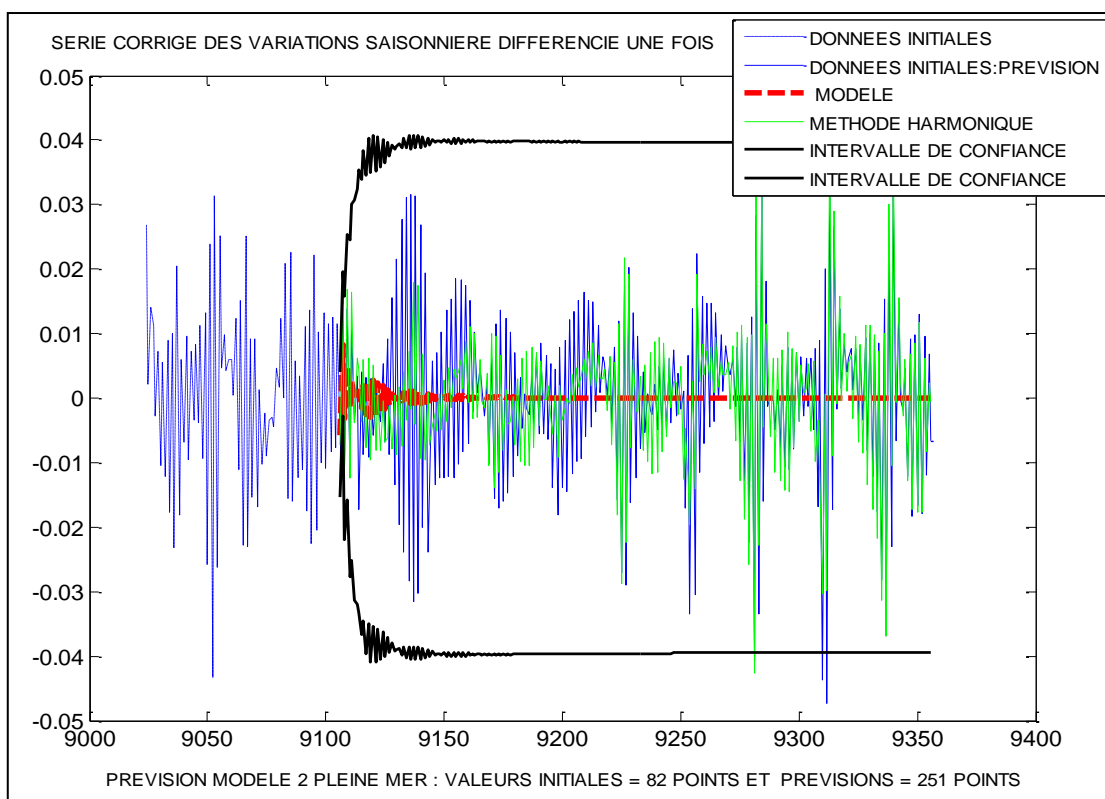


Figure 3 . 65 : Pr vision s rie corrig e des variations saisonni res diff renc e une fois Mod le 2 Pleine mer

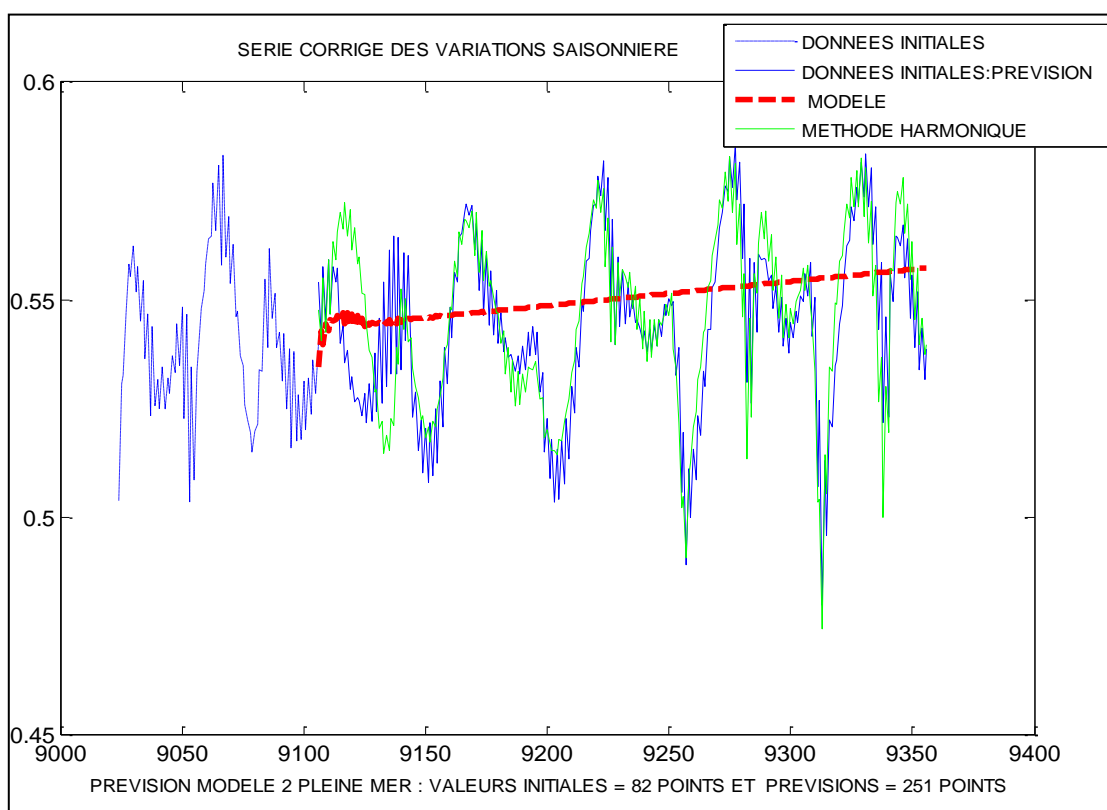


Figure 3 . 66 : Pr vision s rie corrig e des variations saisonni res Mod le 2 Pleine mer

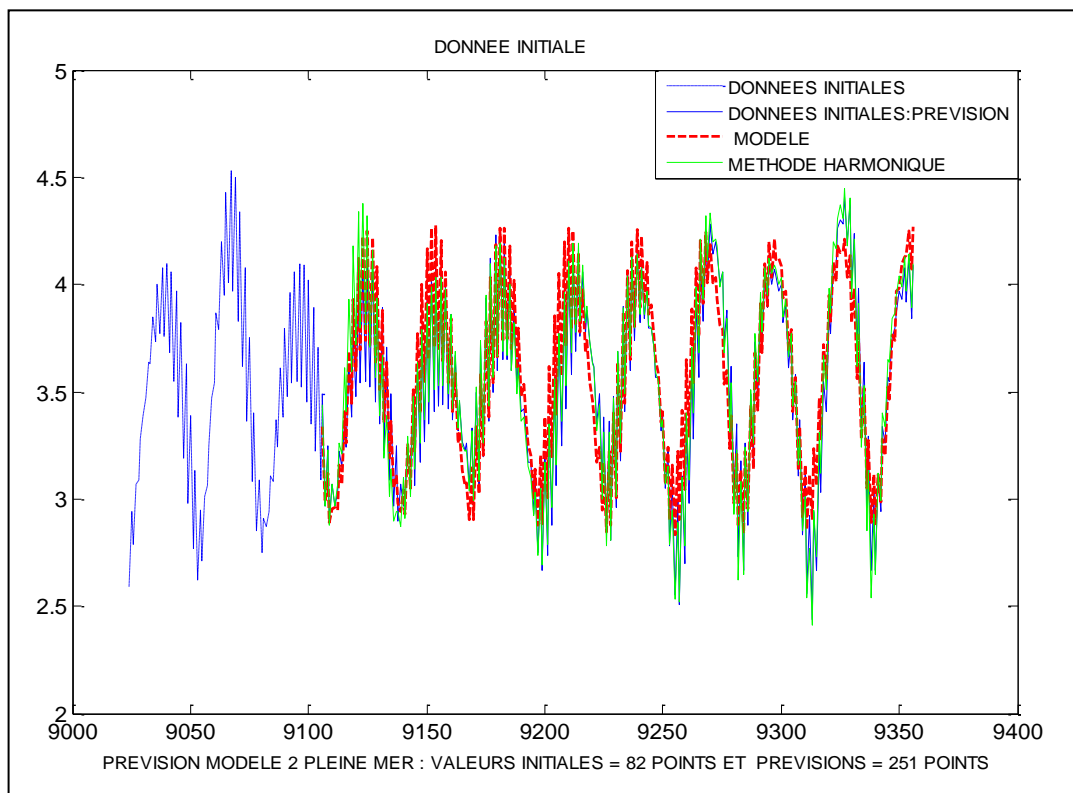


Figure 3 . 67 : Pr vision des donn es initiales Mod le 2 Pleine mer

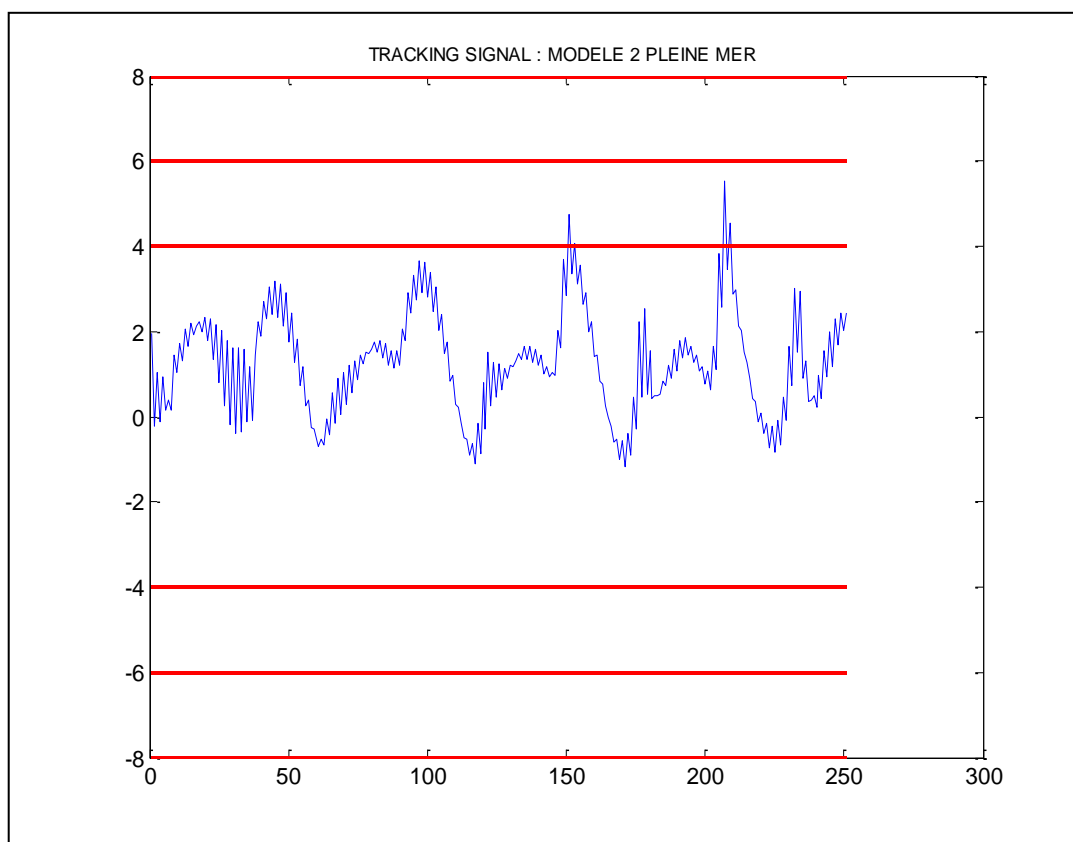


Figure 3 . 68 : Tracking Signal du Mod le 2 Pleine mer

Bien que l'indicateur de Theil ($1,7335 > 1$) montre que notre modèle est moins précis par rapport à la méthode harmonique, nous pouvons toujours affirmer que le modèle 2 pleine mer est un excellent modèle car la RMSE moyenne qui est égale à 0,0199 est inférieure à 5%.

Sur la figure 3.67, nous pouvons apercevoir la forte concordance entre notre prévision (couleur en rouge), les données initiales (couleur en bleu) et les estimations issues du modèle alternative (couleur en vert).

La figure 3.68 confirme également la stabilité du modèle 2 pleine mer puisque les Tracking Signal restent à l'intérieur de l'intervalle $[-6, 6]$

Nous pouvons donc garder notre modèle pour effectuer la prévision du cycle 2 Pleine mer.

Pour le modèle 3 sur les pleines mers, avec la même technique que nous avons utilisée pour les modèles 1 et 2 sur les pleines mers, nous sommes arrivés aux figures 81 à 84, à l'estimation de la série corrigée des variations saisonnières différenciée une fois, à l'estimation de la série corrigée des variations saisonnières, à l'estimation des données initiales et à la construction du Tracking signal.

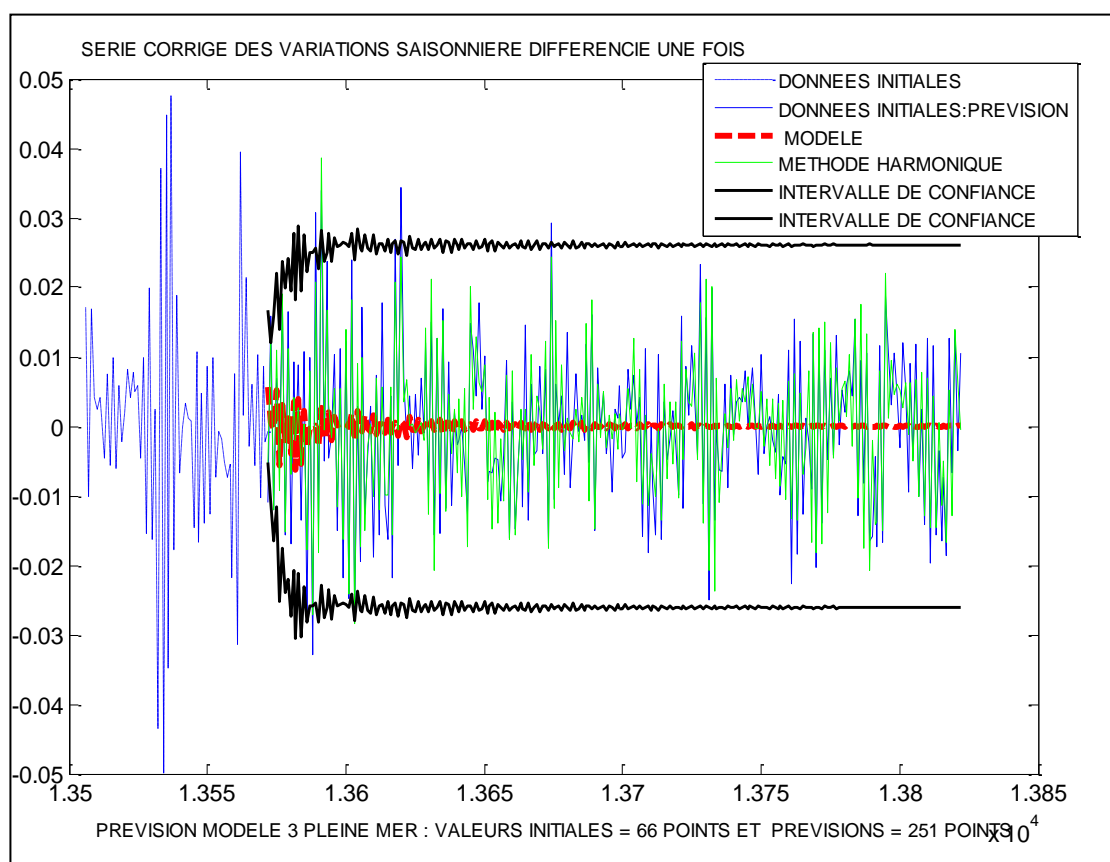


Figure 3 . 69 : Prévision série corrigée des variations saisonnières différenciée une fois Modèle 3 Pleine mer

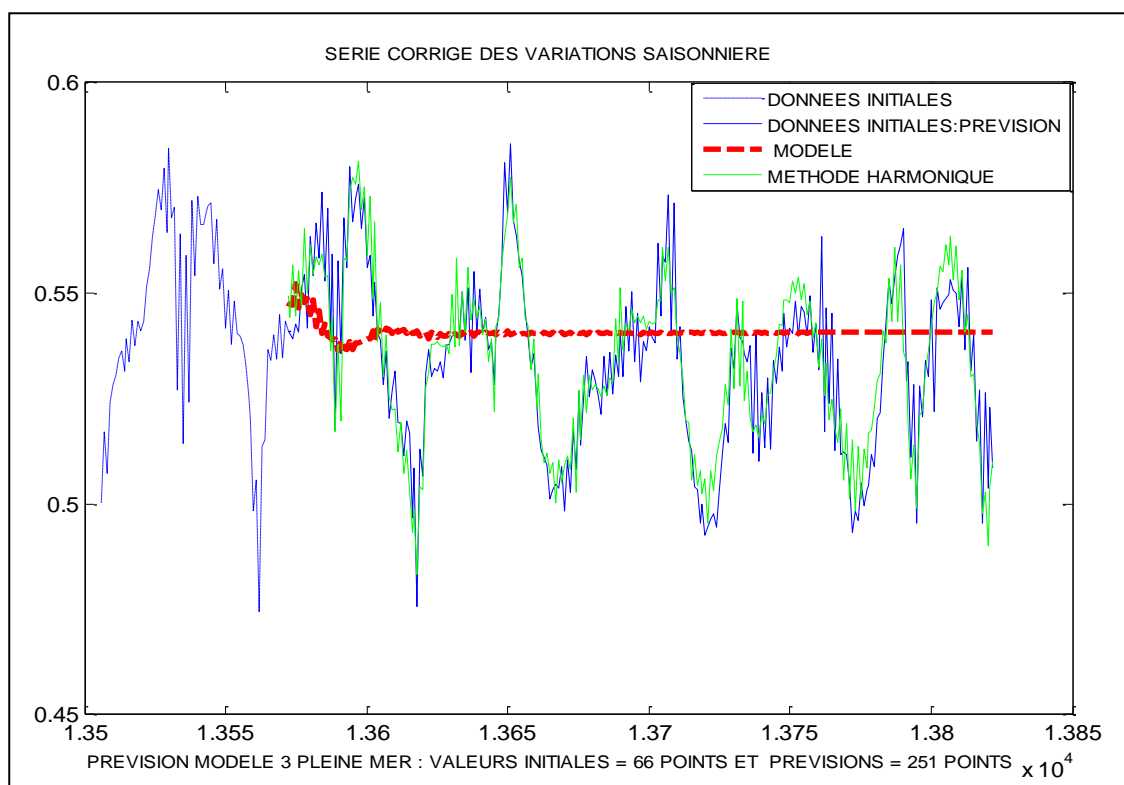


Figure 3 . 70 : Pr vision s rie corrig e des variations saisonni res Mod le 3 Pleine mer

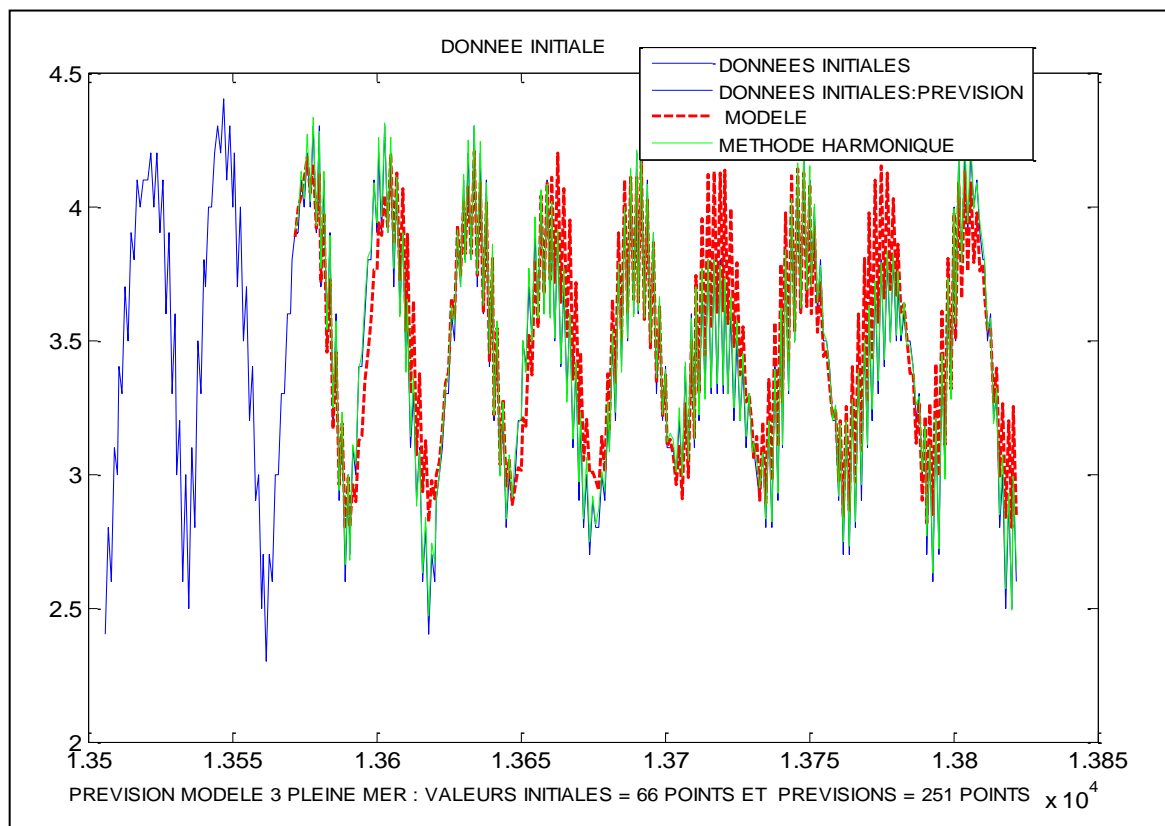


Figure 3 . 71 : Pr vision des donn es initiales Mod le 3 Pleine mer

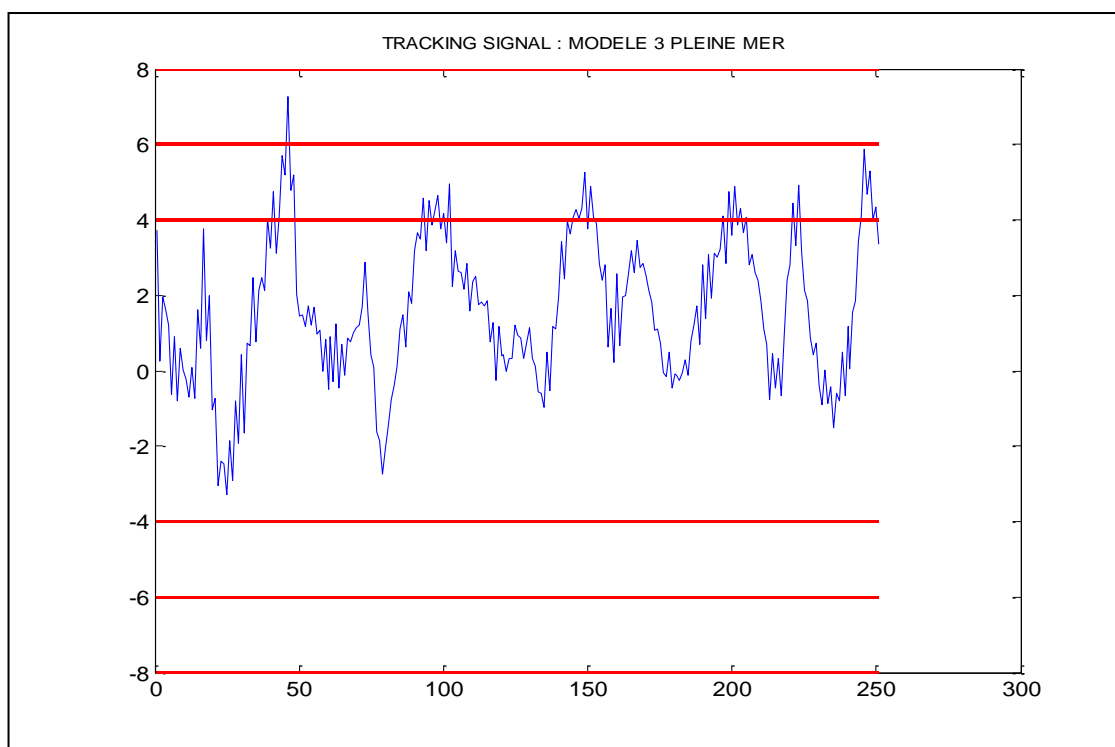


Figure 3 . 72 : Tracking Signal du Modèle 3 Pleine mer

Pour le Modèle 3 pleine mer, puisque l'indicateur de Theil qui est égal à 2,3658 est supérieur à 1 alors la méthode harmonique offre toujours une meilleure estimation par rapport à notre modèle.

La valeur de la RMSE qui vaut 0,0131 permet toujours de classer notre modèle parmi les excellents modèles car cette valeur est inférieure à 5%.

Nous constatons également sur la figure 3.72 du Tracking Signal que le domaine pour nous assurer de la stabilité de notre modèle a un peu augmenté par rapport aux deux précédents modèles mais il reste à l'intérieur de l'intervalle $[-8, 8]$.

A la figure 3.71, la concordance entre notre prévision (couleur en rouge), les données initiales (couleur en bleu) et les estimations issues du modèle alternative (couleur en vert) est toujours confirmée.

Par conséquent nous pouvons retenir notre modèle pour effectuer la prévision du cycle 3 Pleine mer.

Pour le modèle 4 sur les pleines mers, avec les mêmes modes opératoires que pour les modèles 1, 2 et 3 sur les pleines mers, nous avons pu tracer aux figures 3.73 à 3.76, l'estimation de la série corrigée des variations saisonnières différenciée une fois, l'estimation de la série corrigée des variations saisonnières, l'estimation des données initiales et l'estimation du Tracking signal.

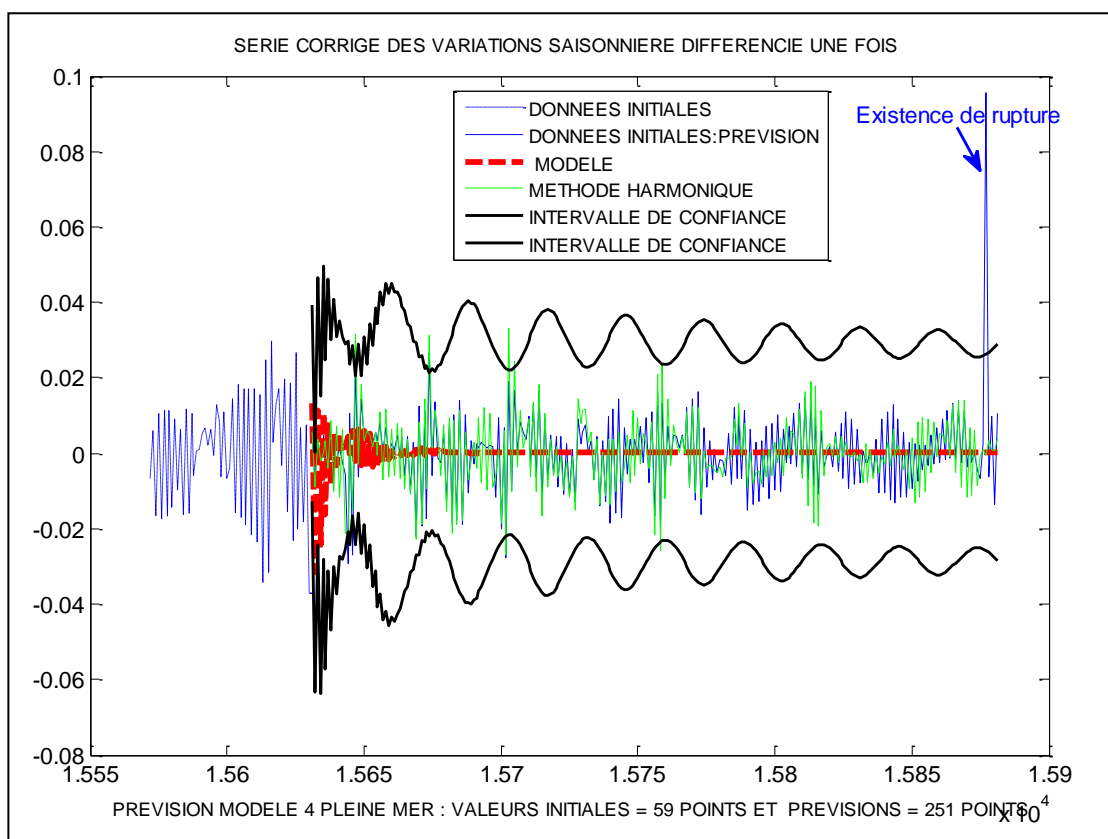


Figure 3.73 : Pr vision s rie corrig e des variations saisonni res diff renci e une fois Mod le 4 Pleine mer

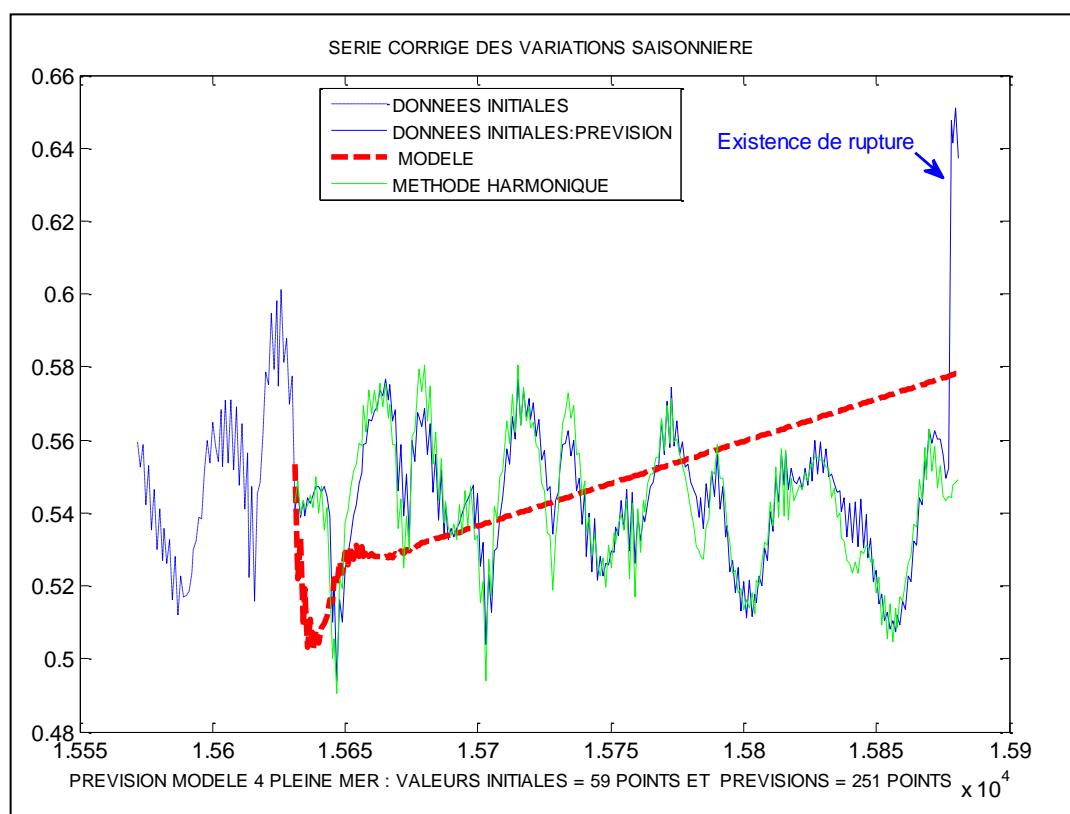


Figure 3.74 : Pr vision s rie corrig e des variations saisonni res Mod le 4 Pleine mer

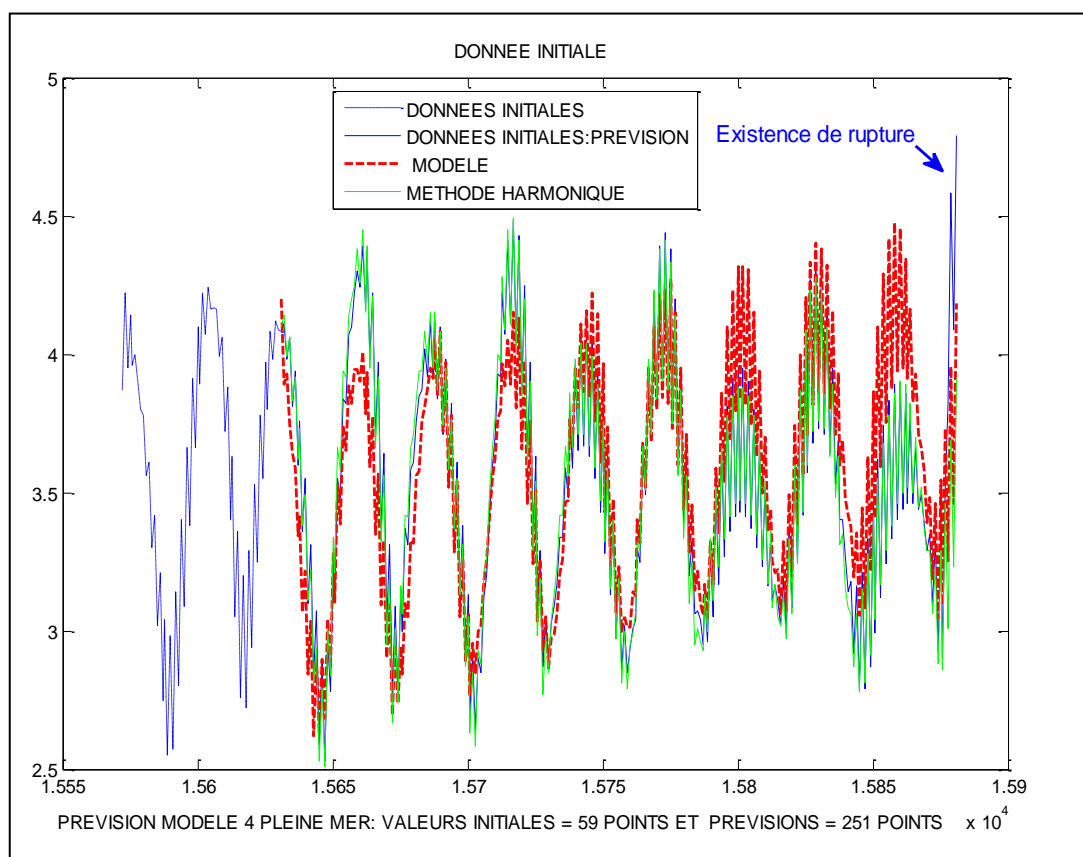


Figure 3.75 : Pr vision des donn es initiales Mod le 4 Pleine mer

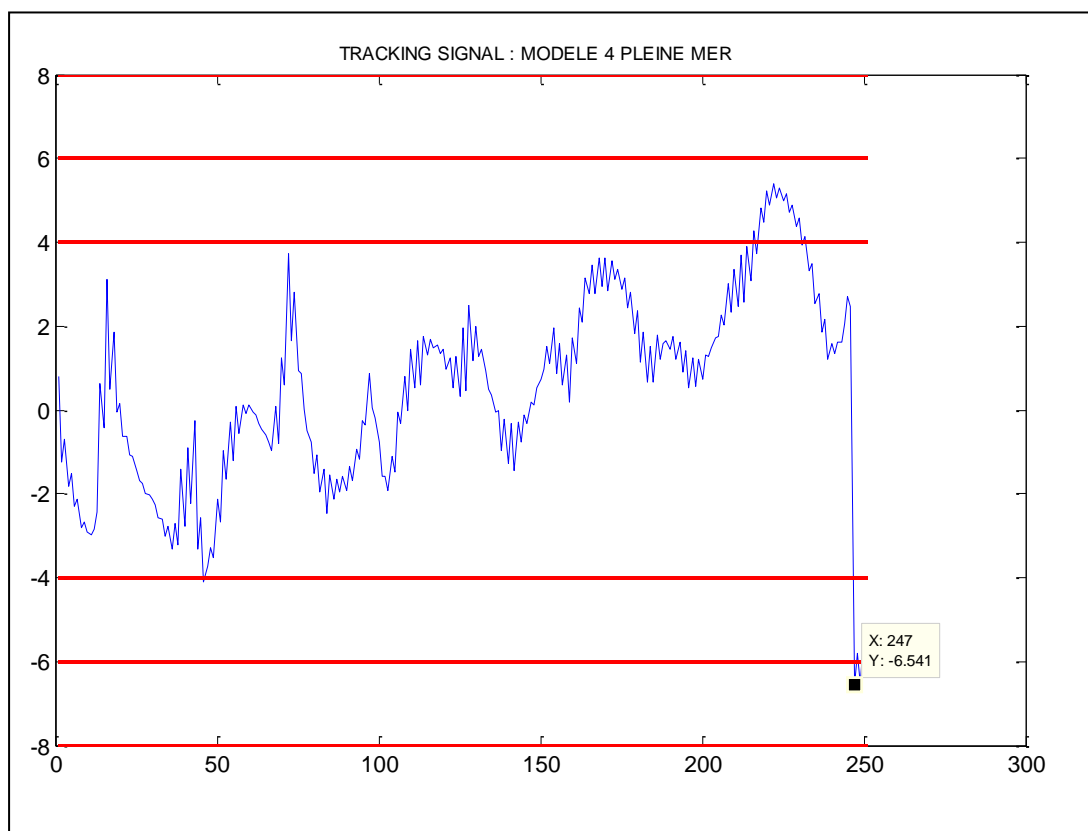


Figure 3.76 : Tracking Signal du Mod le 4 Pleine mer

Nous arrivons pour le Modèle 4 pleine mer à la même conclusion que pour les trois premières modèles sur les pleines mers avec un indicateur de Theil égal à 1,6719.

La méthode harmonique offre donc une meilleure précision par rapport à notre modèle

La valeur de la RMSE moyenne (0,0152) inférieure à 5 % signifie que nous avons un excellent modèle et la courbe Tracking Signal qui reste à l'intérieur de l'intervalle $[-8, 8]$ nous permet d'affirmer que notre modèle est assez stable.

La seule différence avec les trois premières modèles est la détection d'une présence de rupture à partir du 247^{ème} prévision, date qui correspond au prélèvement à partir du 01 Juillet 2012.

Nous avons donc décidé de remettre en œuvre la méthode de la boîte à moustache en utilisant l'équation 2.3.11 pour détecter si les points à partir du changement structurel (01 Juillet 2012) correspondent ou non à des points atypiques.

Pour tester cette hypothèse, nous avons utilisé l'échantillon initial du modèle 4 sur les pleines mers et 355 valeurs situées entre le mois de juillet jusqu'au mois de décembre 2012 et nous sommes arrivés aux résultats du tableau 3.32.

Tableau 3 . 32 : Caractéristique de la boîte à moustache : données initiales modèle 4 pleine mer + 355 points

	Avant correction					
	Q1	Q3	a	b	Min	Max
Basse mer	0,64	1,33	-0,395	2,365	0,07	2,56
Pleine mer	3,17	3,89	2,09	4,97	2,3	5,39

Les résultats du tableau 3.32 et de la figure 3.77 montrent que plusieurs points sur les pleines mers et sur les basses à partir du 01 Juillet 2012 restent en dehors des limites supérieures de confiance de la boîte à moustache que nous avons trouvées, c'est-à-dire 4,97 pour les pleines mers et 2,365 pour les basses mers.

Après calage, les graphiques de la figure 3.78 montrent une concordance parfaite entre les données avant et après la date de rupture détectée au 01 Juillet 2012.

Il semblerait donc que la rupture est due à un problème de calage causant un mouvement de translation uniforme au niveau des prélèvements et qui correspond approximativement à la différence en valeur absolue entre la moyenne des données avant et après rupture. Il nous suffirait alors d'enlever cette différence aux points atypiques pour homogénéiser nos données

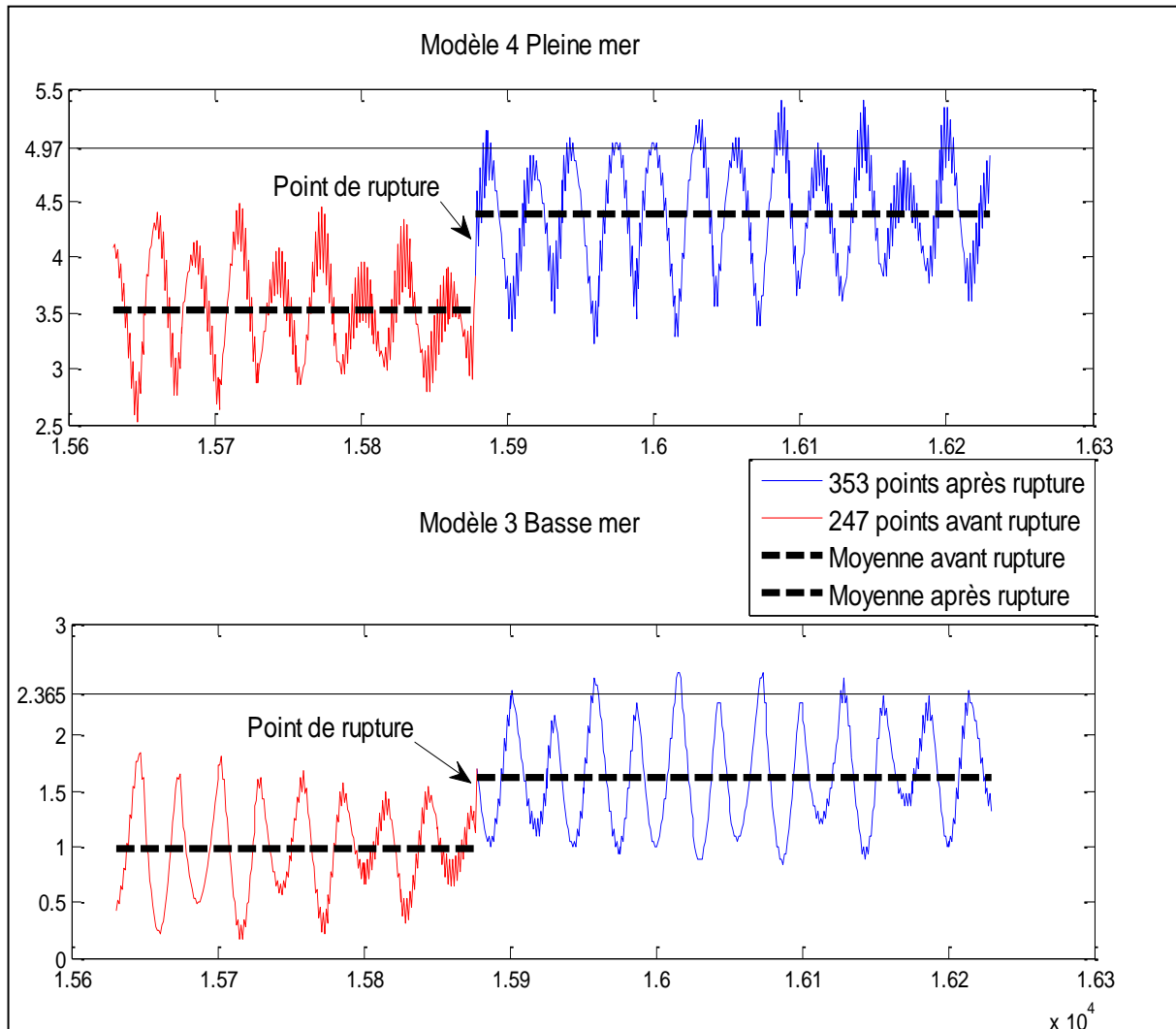


Figure 3.77 : Moyenne des valeurs avant et après rupture

Ainsi, pour arriver à ces résultats, nous avons tout d'abord procédé à l'homogénéisation des données en prenant la différence en valeur absolue entre la moyenne des données avant et après rupture égale à « 0,8595 ».

Ensuite en nous basant sur la concordance entre les valeurs des séries corrigées des variations saisonnières des figures 3.78 et 3.79, nous avons trouvé en comparant les résultats obtenus par les modèles ARIMA(12,1,1)-EGARCH(1,7) et ARIMA(12,1,1)-EGARCH(3,6) que le modèle ARIMA(12,1,1)-EGARCH(1,7) est mieux adapté que le modèle ARIMA(12,1,1)-EGARCH(3,6) que nous avons trouvé auparavant, pour prévoir le cycle 4 sur les pleines mers.

Par conséquent, avant de choisir le modèle qui s'approche de la réalité, il faut vérifier en même temps :

- Par calcul, les hypothèses de validation d'un modèle ARIMA(p+1,d,q)-GARCH(P,Q) énoncés au paragraphe 3.2.13.5 et
- Graphiquement la concordance entre les prélèvements et les prévisions

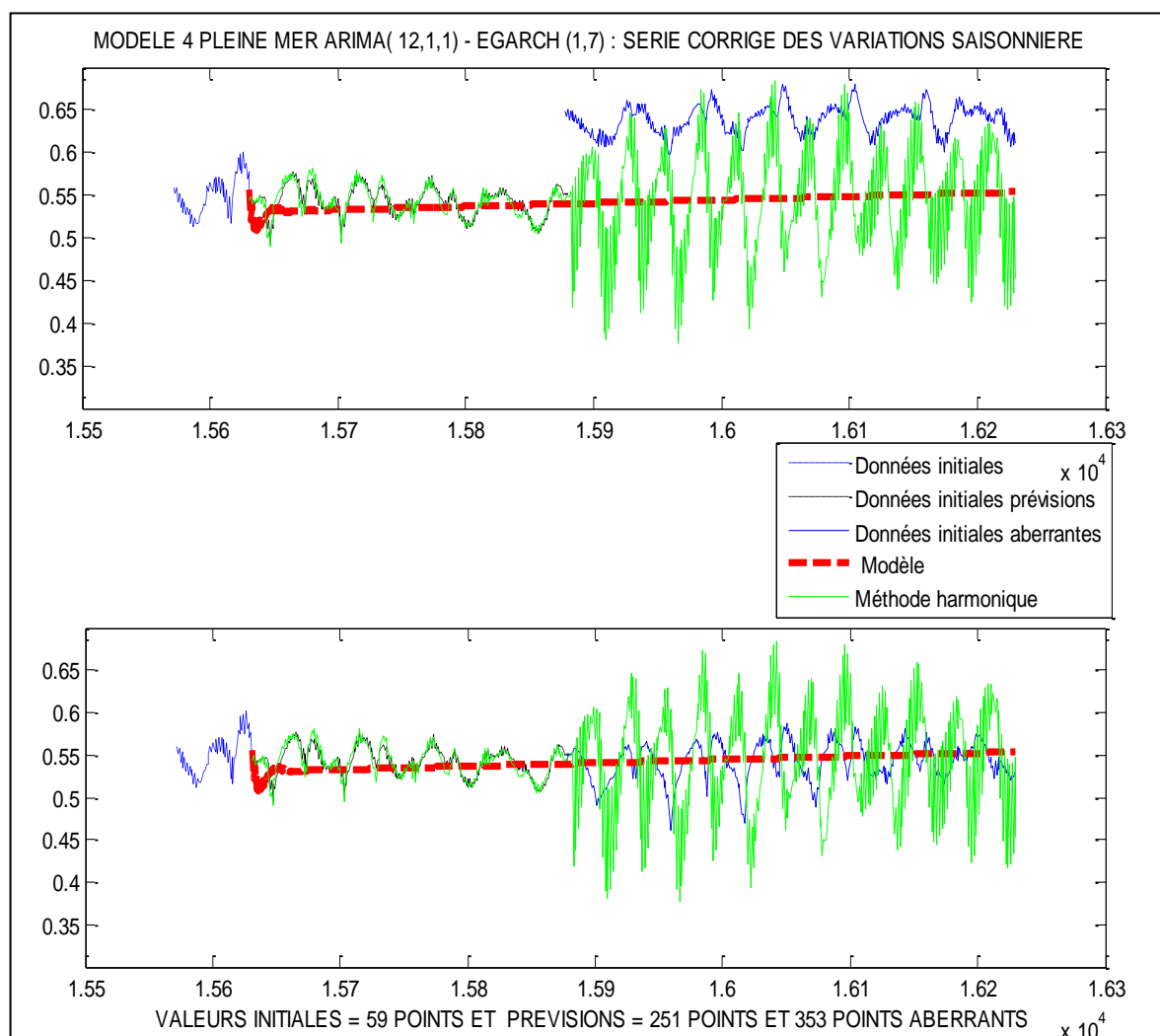


Figure 3.78 : Prédiction série corrigée des variations saisonnières avant et après homogénéisation pour le Modèle 4 Pleine mer ARIMA(12,1,1)-EGARCH(1,7)

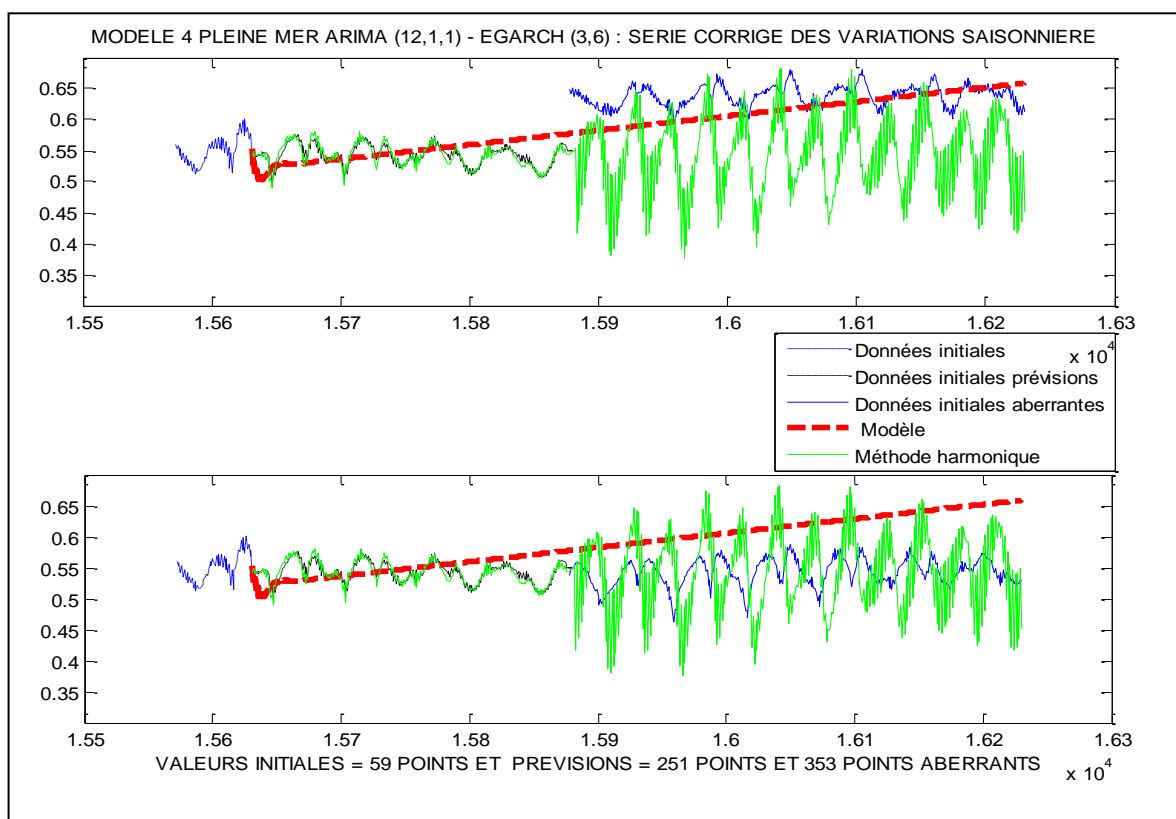


Figure 3.79 : Prédiction série corrigée des variations saisonnières avant et après homogénéisation pour le Modèle 4 Pleine mer ARIMA(12,1,1)-EGARCH(3,6)

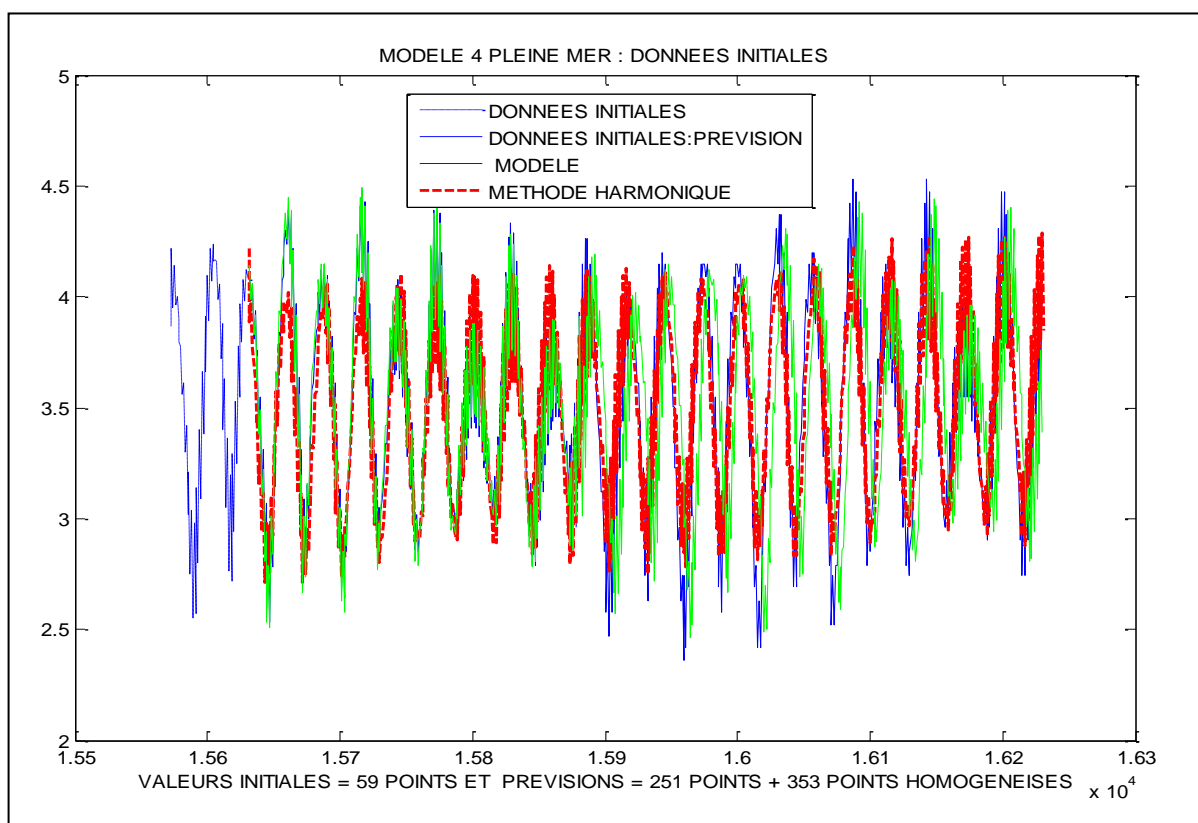


Figure 3.80 : Prédiction données initiales du Modèle 4 Pleine mer après homogénéisation

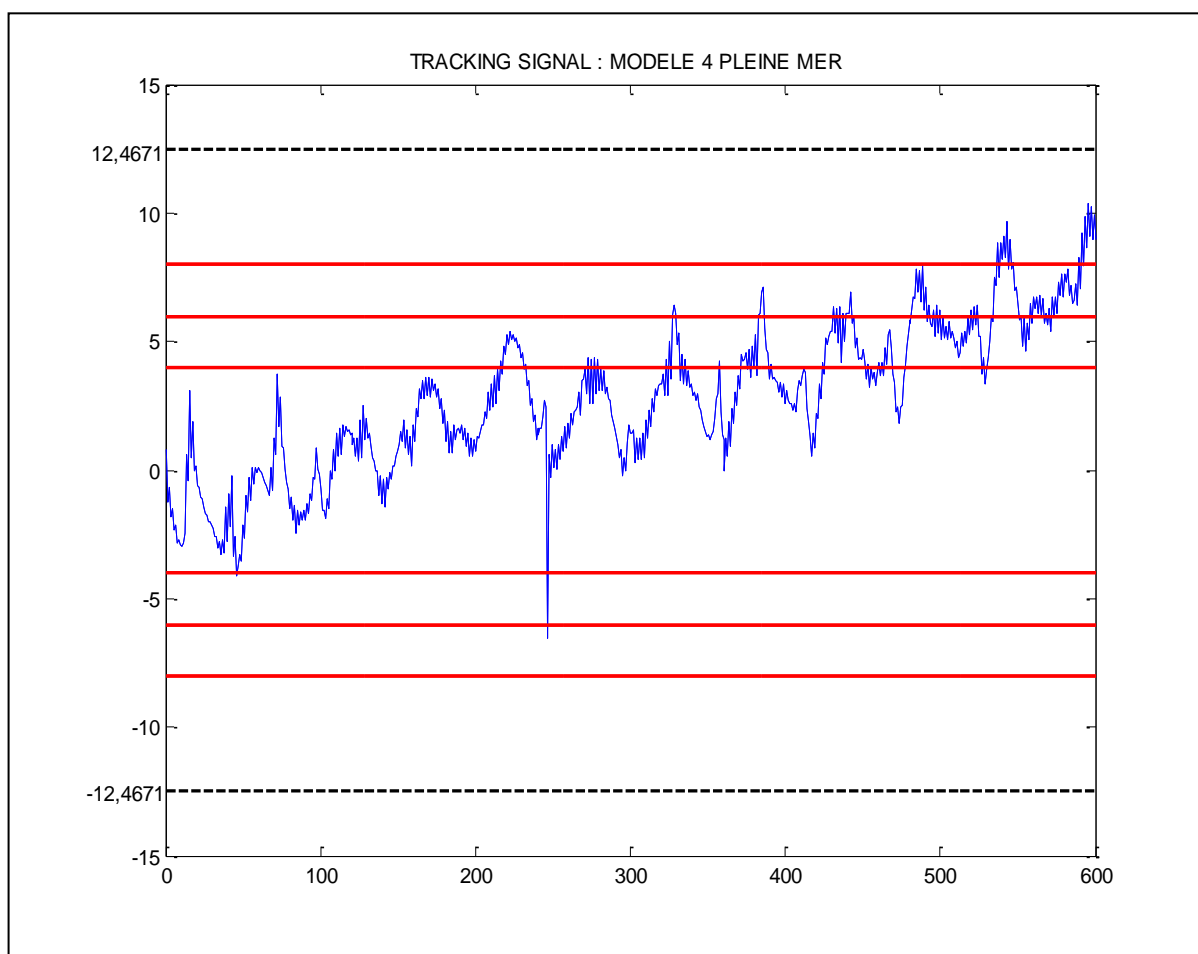


Figure 3 . 81 : Tracking Signal du Modèle 4 Pleine mer après homogénéisation

La prévision obtenue après homogénéisation est assez bonne d'après la figure 3.80. L'indicateur de THEIL et la RMSE sont respectivement égal à 1,6291 et 0,0148.

La prévision n'est pas meilleure que la méthode harmonique car le ratio de Theil est supérieur à 1 mais l'estimation reste excellente puisque la RMSE est inférieure à 5%.

L'intervalle de confiance pour valider la méthode harmonique que nous avons trouvé correspondait à $[-12,4761, 12,4761]$. Cette valeur rend non biaisé le tracking signal de la figure 3.81.

Par conséquent, nous pouvons accepter le modèle $ARIMA(12,1,1)-EGARCH(1,7)$ pour prévoir le cycle 4 sur les pleines mers.

Pour le modèle 1 sur les basses mers, les résultats des figures 3.82 à 3.85 nous permettent d'apprécier la qualité de la prévision du cycle 1 sur les basses mers.

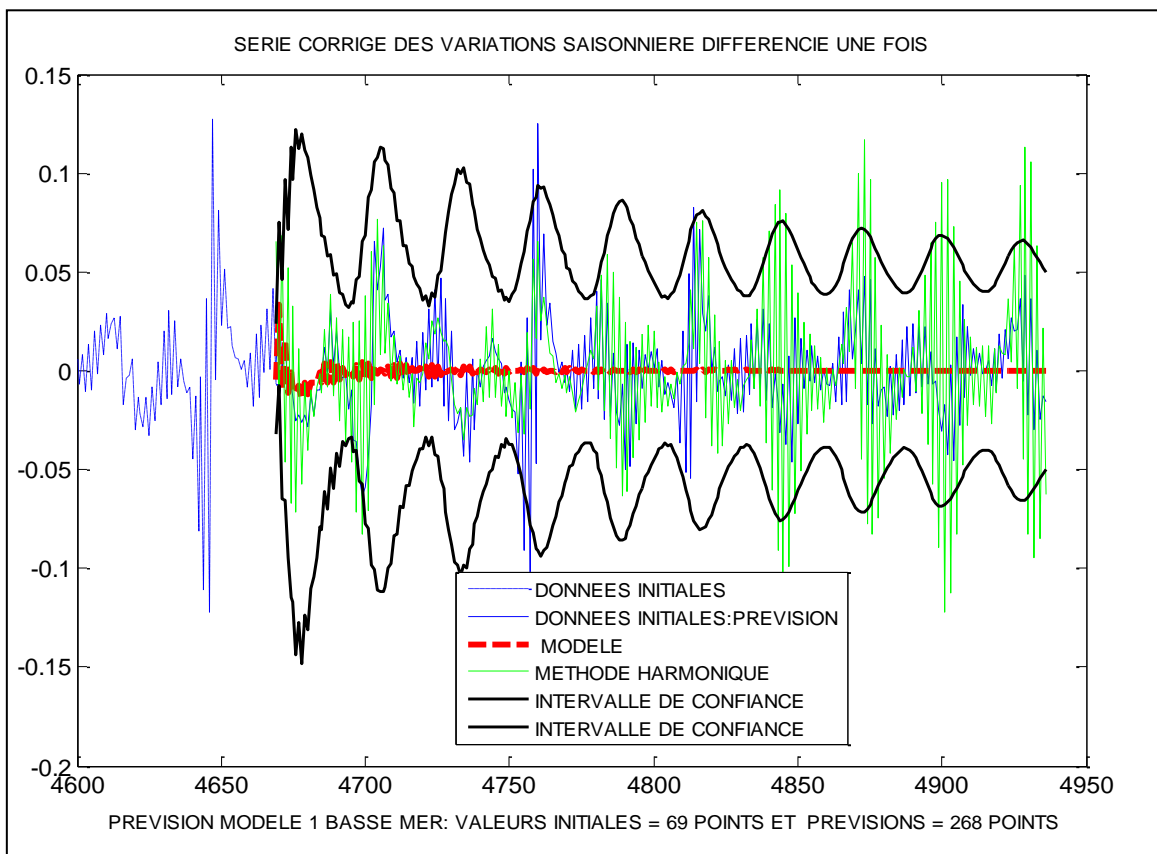


Figure 3 . 82 : Pr vision s rie corrig e des variations saisonni res diff renci e une fois Mod le 1 Basse mer

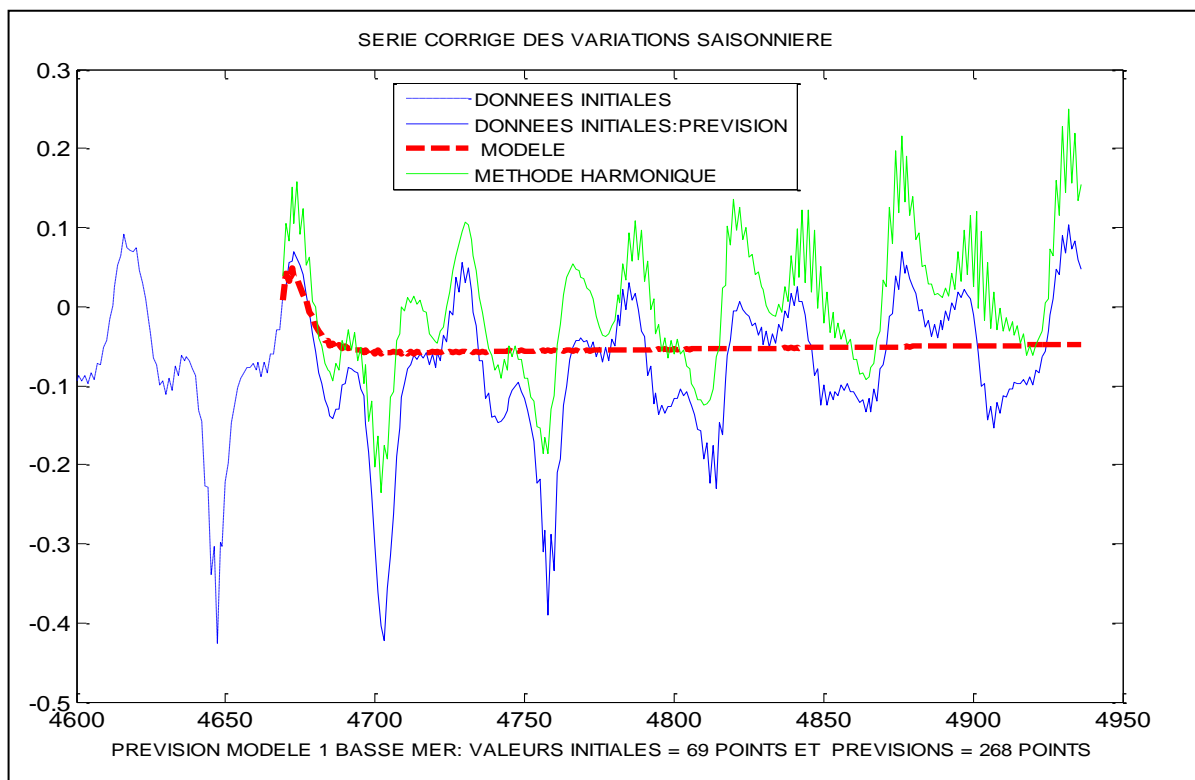


Figure 3 . 83 : Pr vision s rie corrig e des variations saisonni res Mod le 1 Basse mer

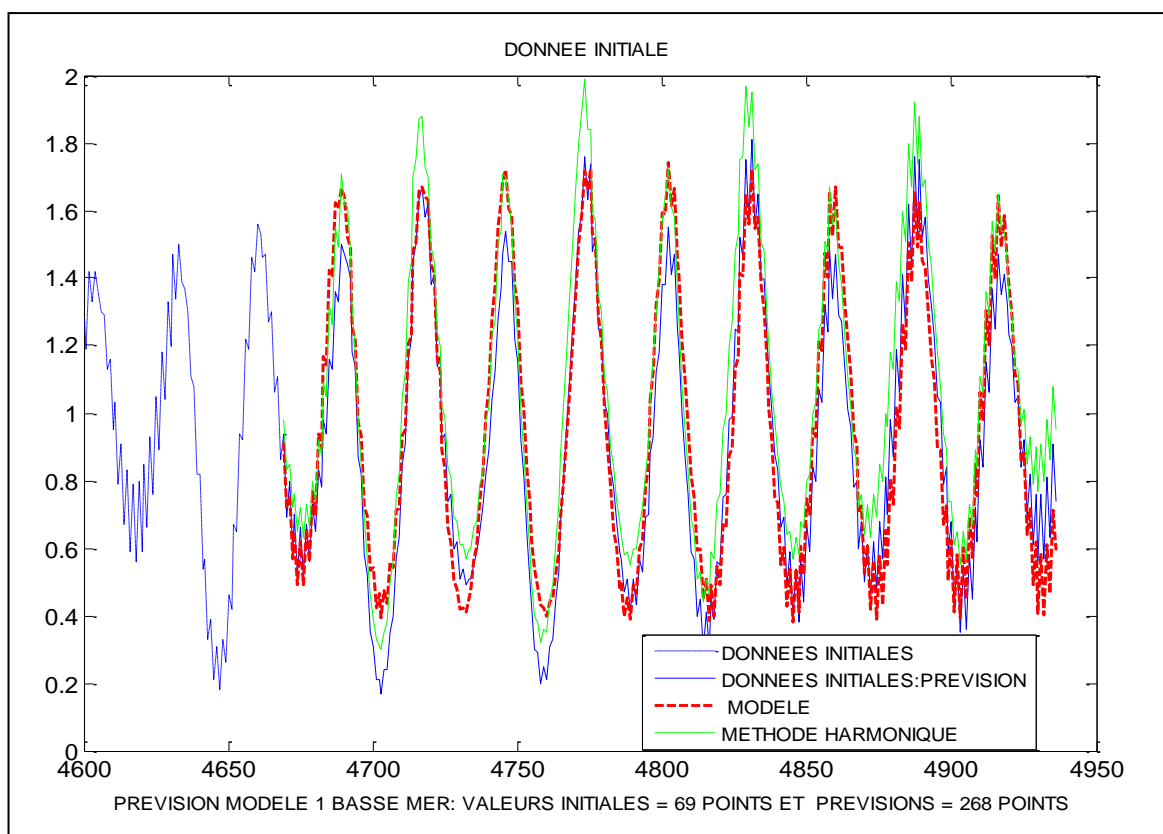


Figure 3 . 84 : Pr vision des donn es initiales Mod le 1 Basse mer

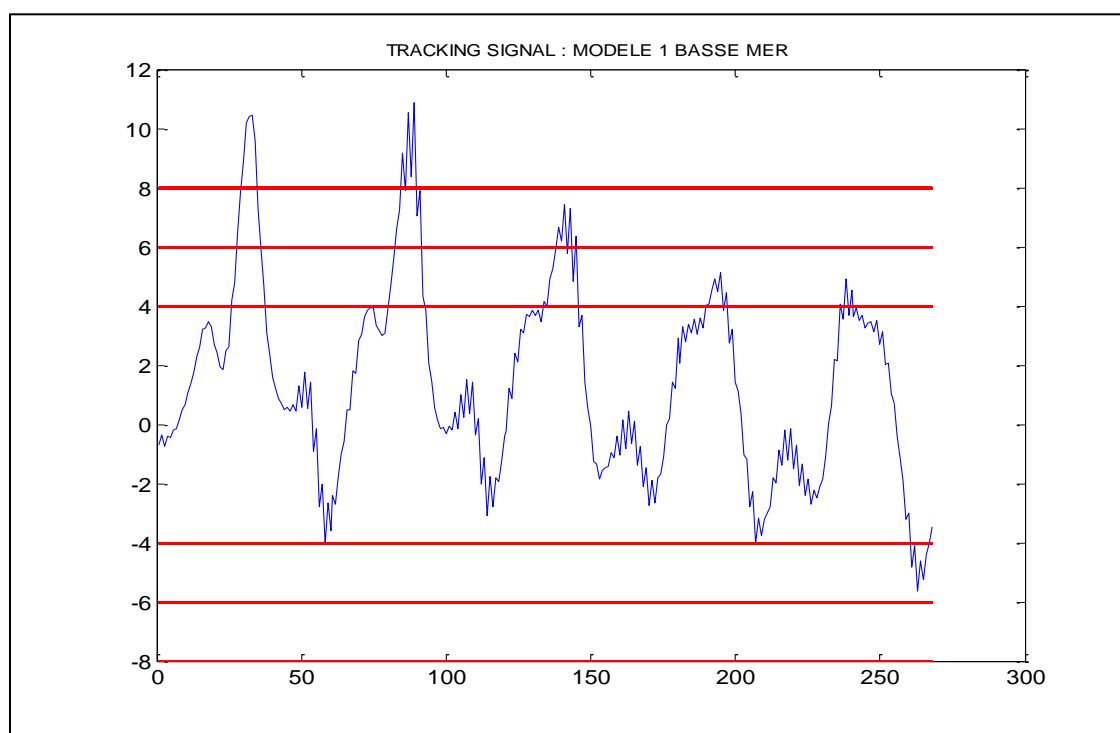


Figure 3 . 85 : Tracking Signal du Mod le 1 Basse mer

En particulier, les figures 3.82 à 3.85 nous permettent d'affirmer la concordance entre notre prévision, la prévision par méthode harmonique et les données initiales.

Mais contrairement aux modèles sur les pleines mers, la prévision par notre modèle sur les basses mers, ici est plus précise par rapport à la prévision par la méthode harmonique étant donné que l'indicateur de Theil qui vaut 0,7177 est inférieur à 1.

La RMSE moyenne égale à 0,0312 (< 5%), nous permet de classer notre modèle parmi les modèles de très bonne qualité.

Le seul problème est que des points de la courbe Tracking Signal de la figure 3.85 dépassent la limite supérieure de l'intervalle de confiance $[-8, 8]$ proposé dans la littérature.

Nous avons donc recherché la limite de confiance utilisée par le service Hydrographique et Océanographique de la marine française et nous avons trouvé :

$$\left[-E \left(\text{Max}(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right), E \left(\text{Max}(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right) \right] = [-E(21,6014), E(21,6014)] = [-21, 21]$$

Avec ces nouvelles limites, la stabilité du modèle 1 Basse mer peut être acceptée.

Pour le modèle 2 sur les basses mers, les résultats des figures 3.86 à 3.89 permettent également de déterminer la qualité de notre modèle à partir des estimations de la série corrigée des variations saisonnières différenciée une fois, de la série corrigée des variations saisonnières, des données initiales et à partir du Tracking signal.

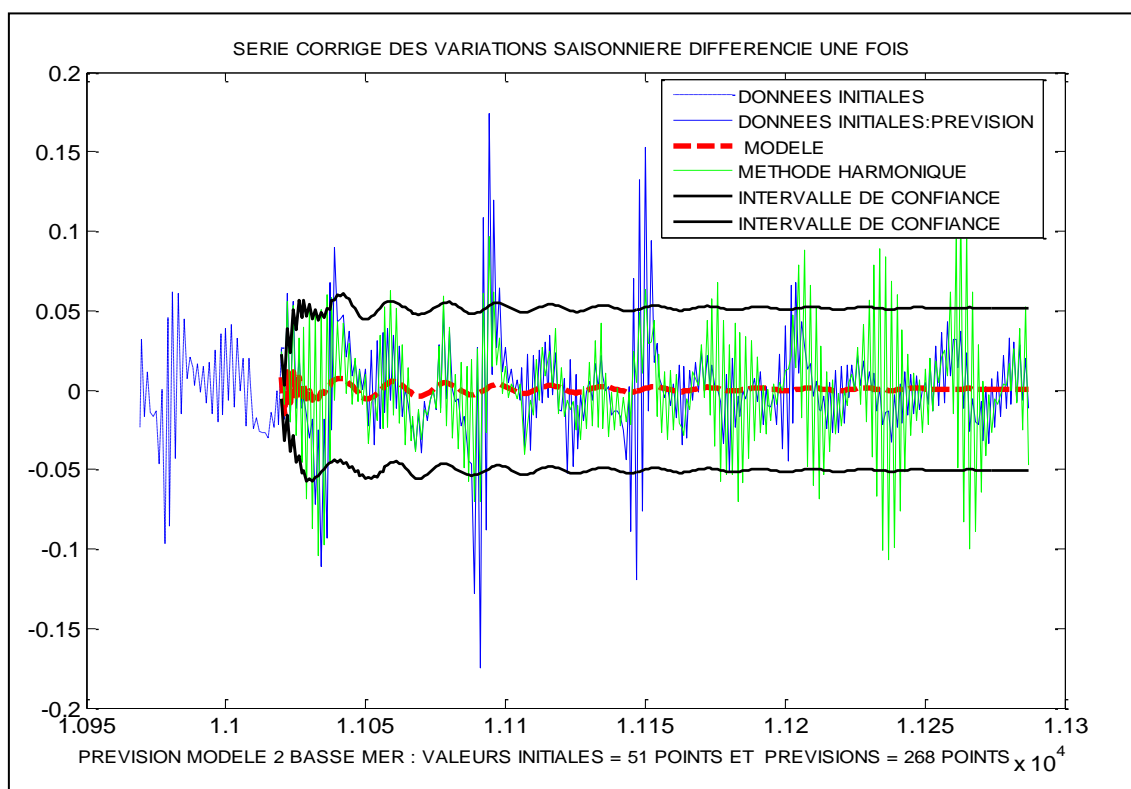


Figure 3.86 : Prédiction série corrigée des variations saisonnières différenciée une fois Modèle 2 Basse mer

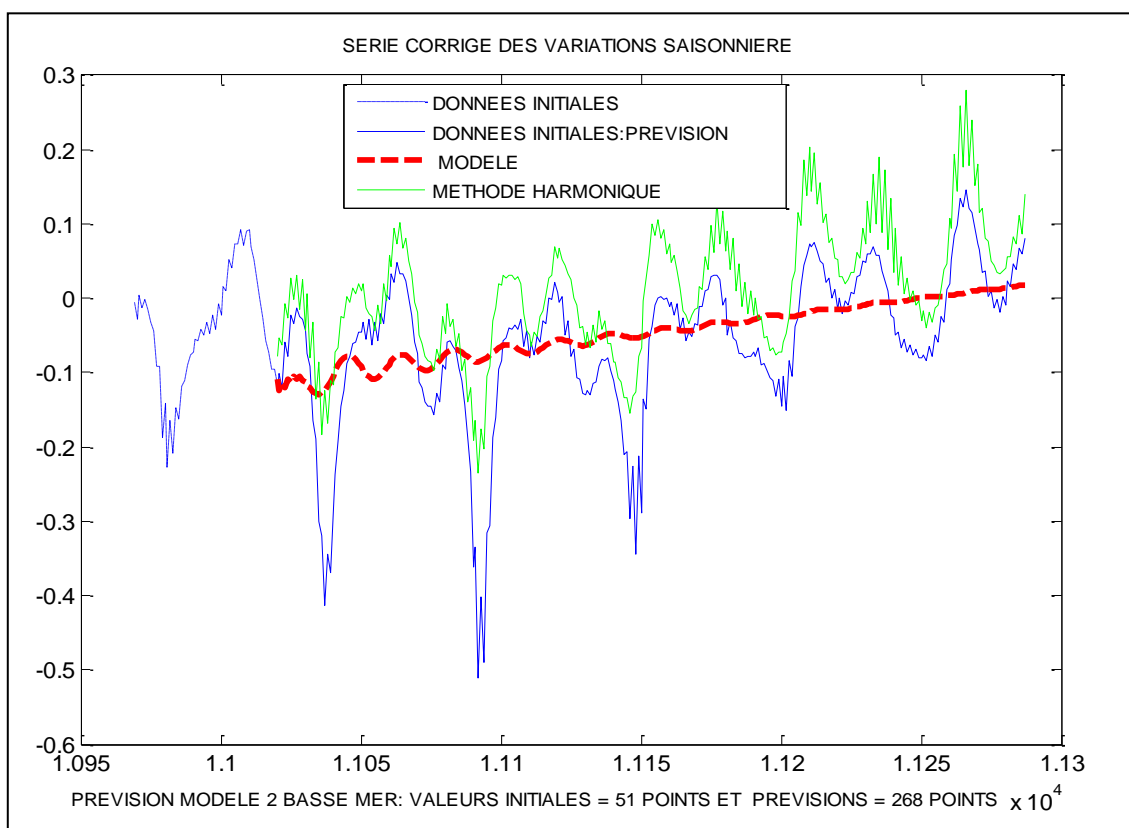


Figure 3 . 87 : Pr vision s rie corrig e des variations saisonni res Mod le 2 Basse mer

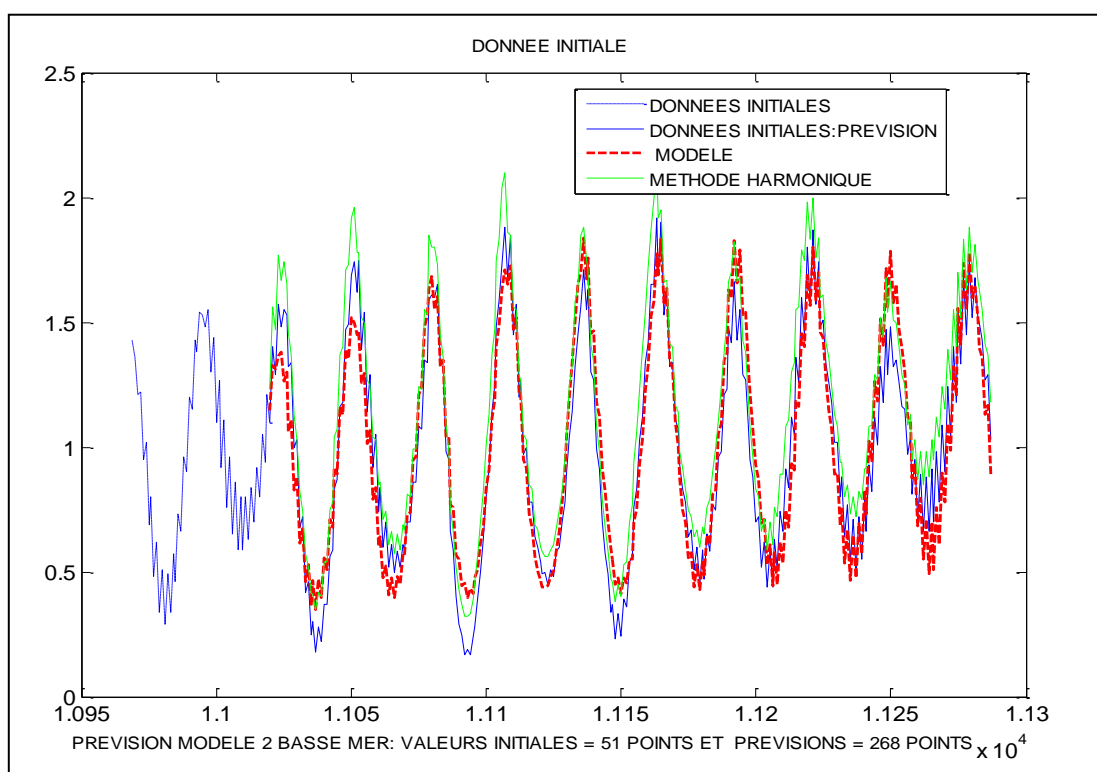


Figure 3 . 88 : Pr vision des donn es initiales Mod le 2 Basse mer

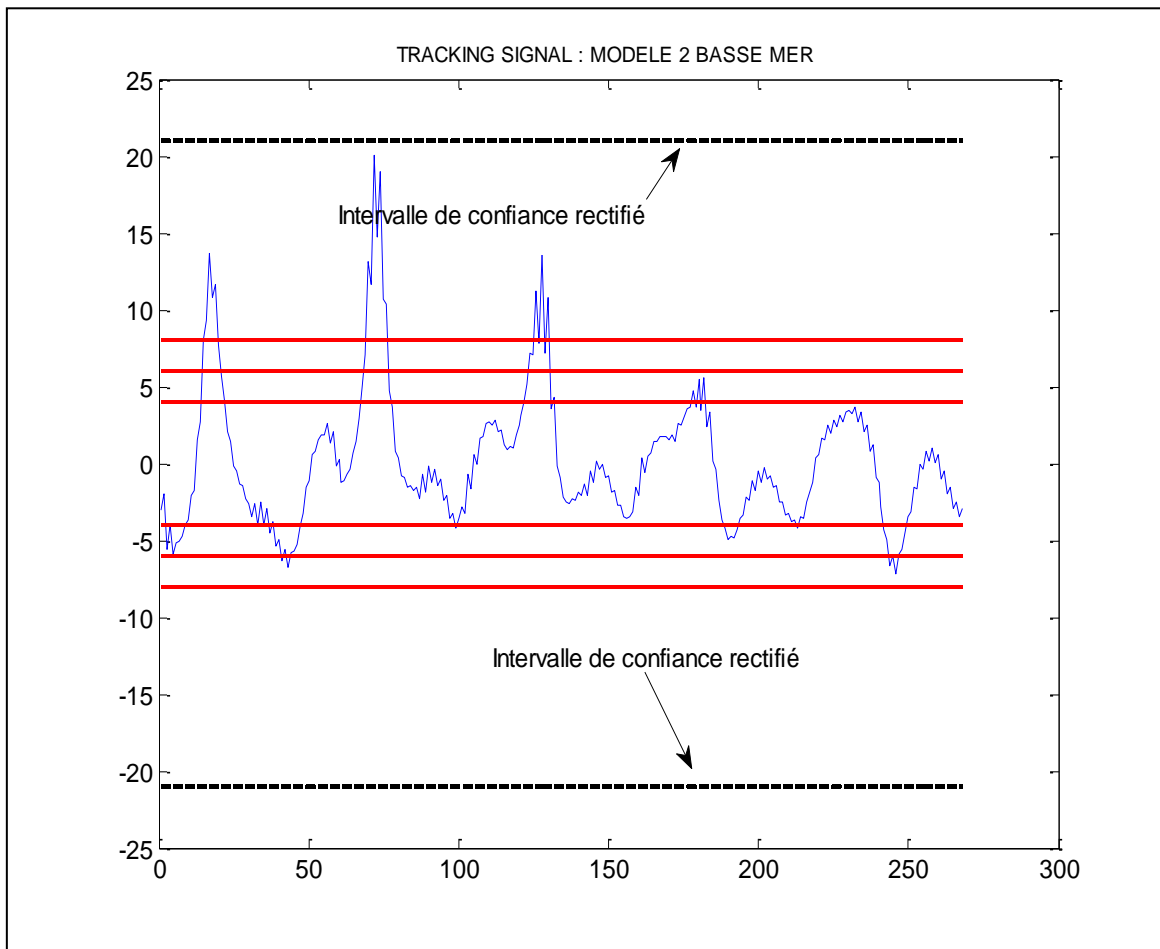


Figure 3 . 89 : Tracking Signal du Modèle 2 Basse mer

Les résultats des figures 3.86 à 3.89 confirment comme dans le cas précédent, la stabilité de notre modèle et la concordance entre les valeurs prédites avec les données initiales.

Nous pouvons aussi dire que notre prévision du cycle 2 basses mers est plus précise par rapport à celle de la méthode harmonique car le ratio de Theil vaut 0,5626 (inférieur à 1).

Comme la RMSE (0,0258) est inférieure à 5% alors notre modèle est de très bonne qualité.

Pour le modèle 3 sur les basses mers, nous avons utilisé les résultats des figures 3.2.89 à 3.2.92 pour déterminer la qualité de notre modèle.

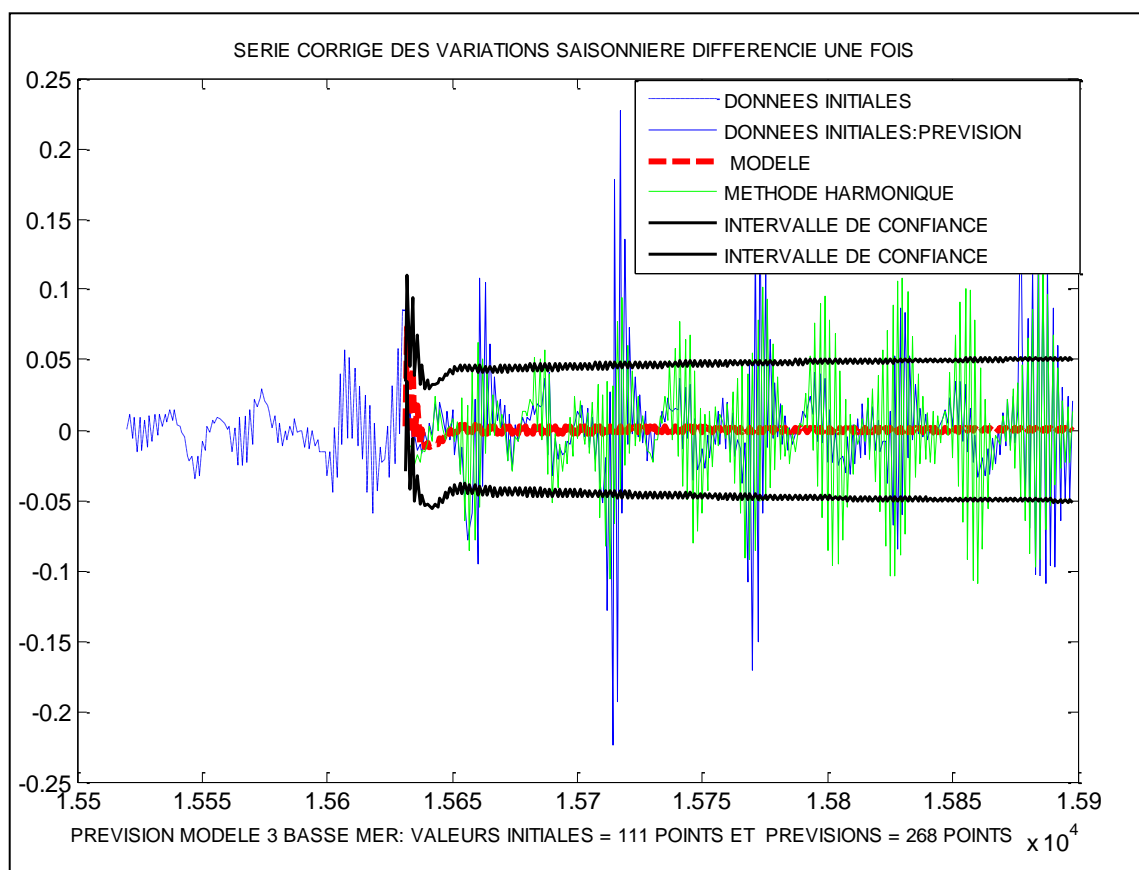


Figure 3 . 90 : Pr vision s rie corrig e des variations saisonni res diff renci e une fois Mod le 3 Basse mer

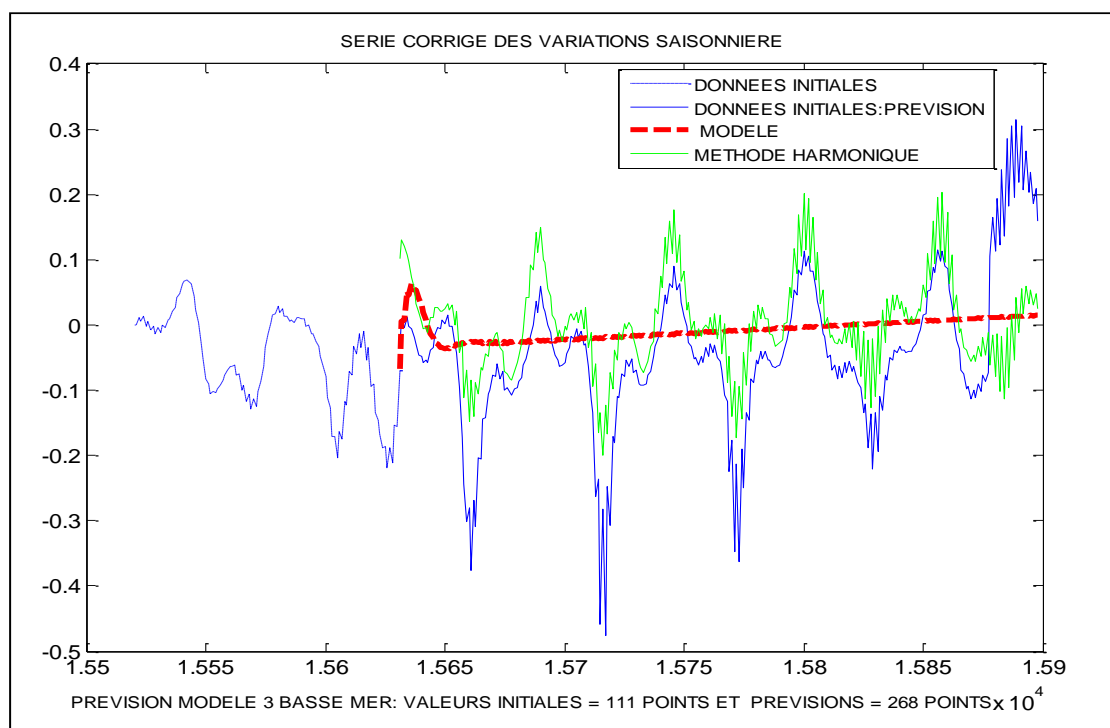


Figure 3 . 91 : Pr vision s rie corrig e des variations saisonni res Mod le 3 Basse mer

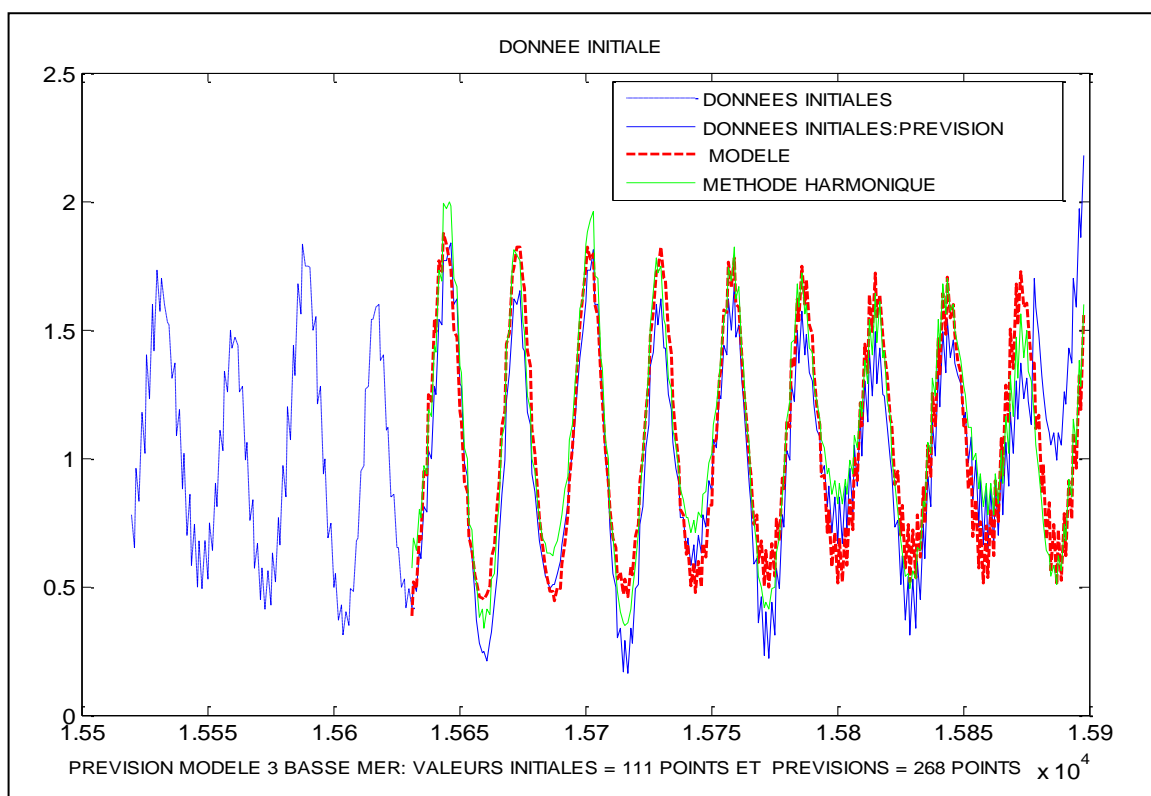


Figure 3 . 92 : Prédiction des données initiales Modèle 3 Basse mer

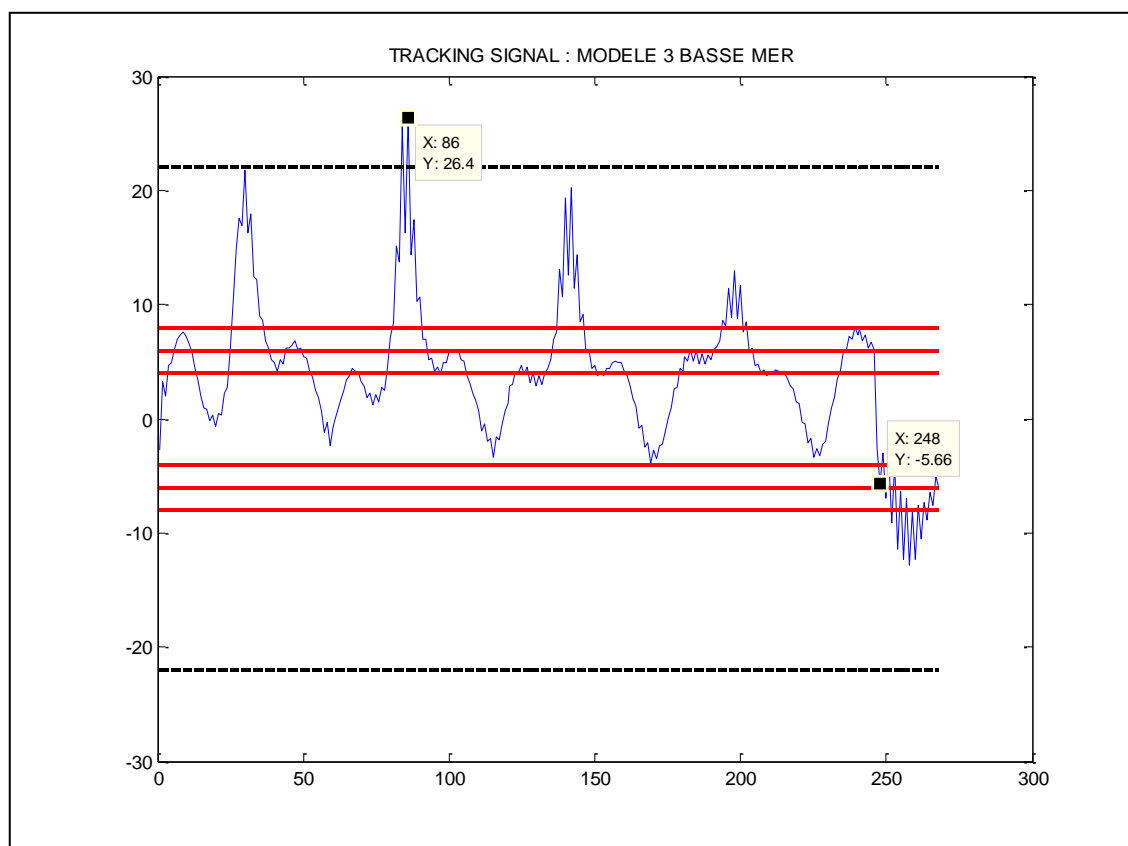


Figure 3 . 93 : Tracking Signal Modèle 3 Basse mer

Pour le modèle 3 Basse mer, avec les mêmes raisonnements qu'auparavant, nous pouvons montrer l'excellente qualité de notre prévision car la RMSE (0,0240) est inférieure à 5%.

Le ratio de THEIL ($0,3402 < 1$) nous permet d'affirmer que la prévision effectuée par notre modèle est beaucoup plus précis par rapport à la méthode harmonique.

Pour la stabilité, un pic sort de l'intervalle de confiance de la figure 3.93 et semble indiquer une instabilité mais après avoir appliqué sous xlstat 2014 le test d'homogénéité de Pettit, nous avons pu confirmer que les résidus normalisés sont homogènes puisque la p-value est égale à 0,3835.

Cela signifie que pour nous assurer de la stabilité de nos modèles sur les basses mers, il faudrait prendre l'intervalle de confiance au moins égal à :

$$\left[-E \left(\text{Max}(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right), E \left(\text{Max}(TS) + \frac{1,96\sigma_{TS}}{\sqrt{N_{TS}}} + 1 \right) \right] = [-E(27,4), E(27,4)] = [-27, 27].$$

Pour les basses mers également, un point de rupture a été observé aux figures 3.90 à 3.93. Il correspond au 248^{ème} prévision ou au prélèvement du 1 Juillet 2012. Nous avons donc procédé à l'homogénéisation des points atypiques comme dans le cas du modèle 4 sur les pleines mers et nous sommes arrivé aux résultats des figures 3.94 à 3.96.

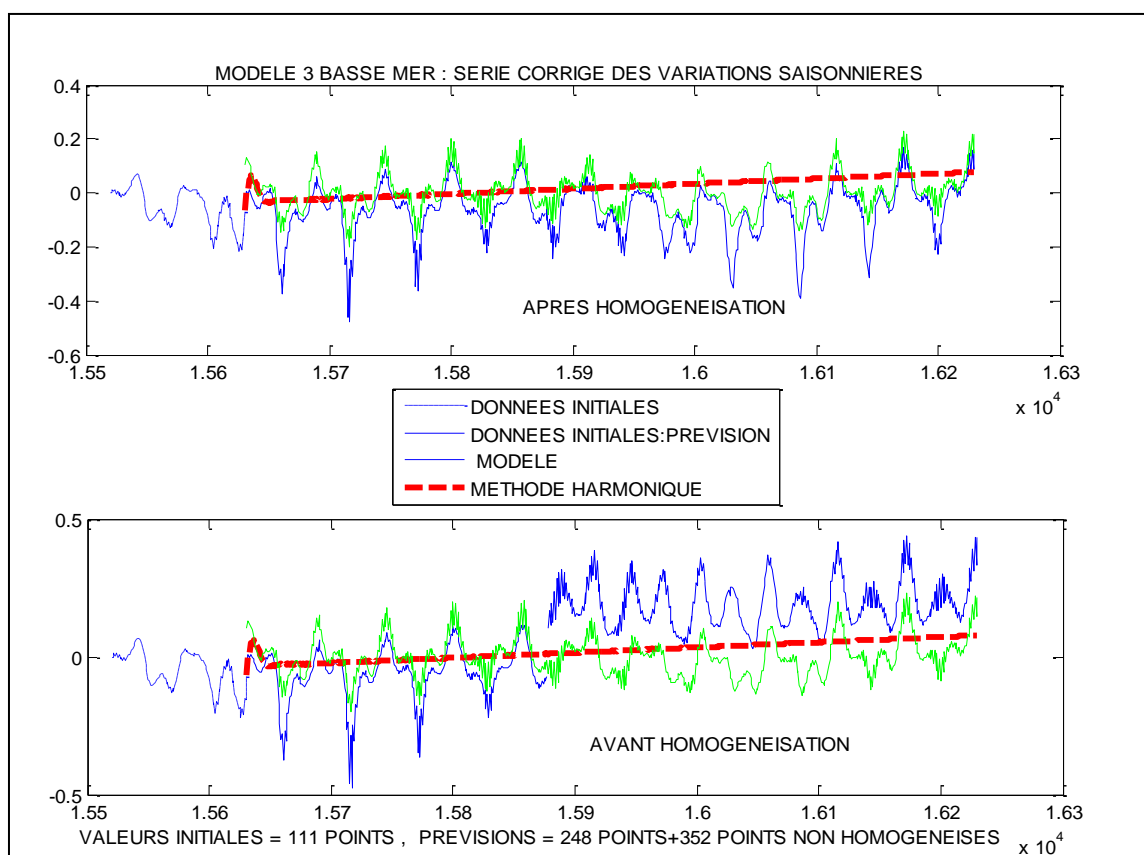


Figure 3 . 94 : Prévision série corrigée des variations saisonnières Modèle 3 Basse mer avant et après homogénéisation

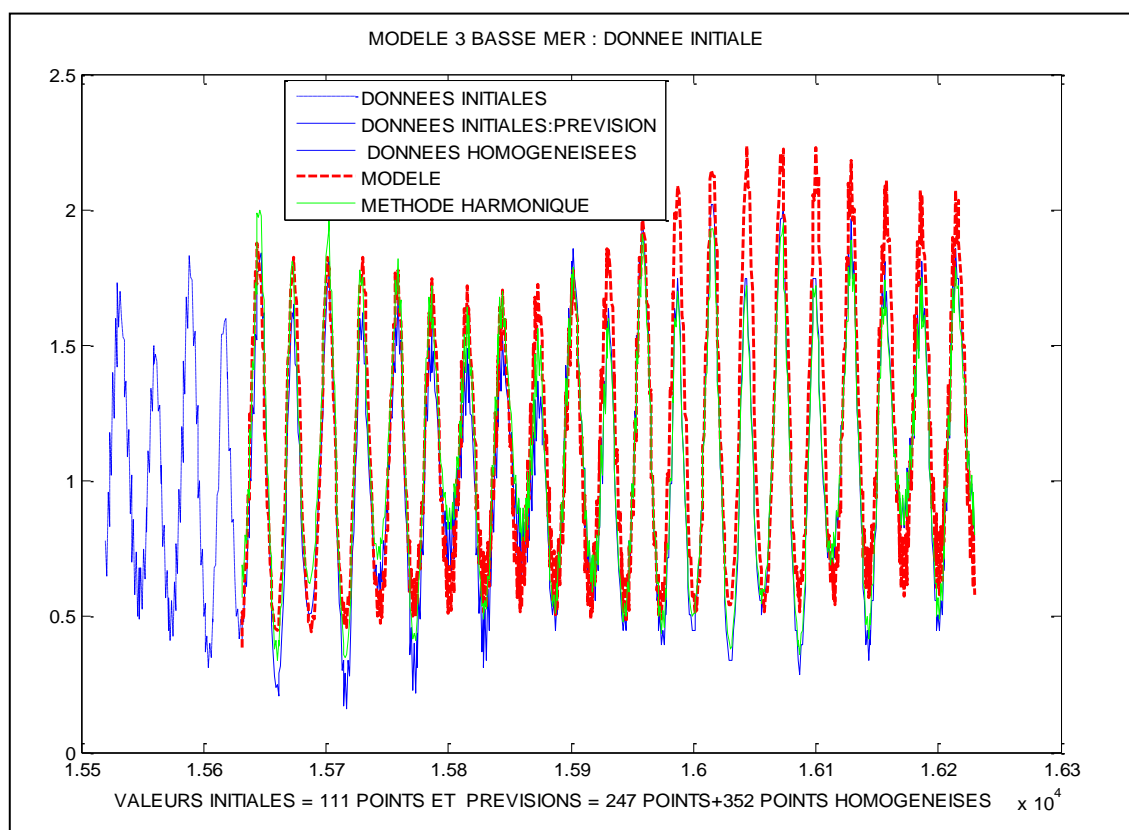


Figure 3 . 95 : Pr vision des donn es initiales Mod le 3 Basse mer apr s homog n isation

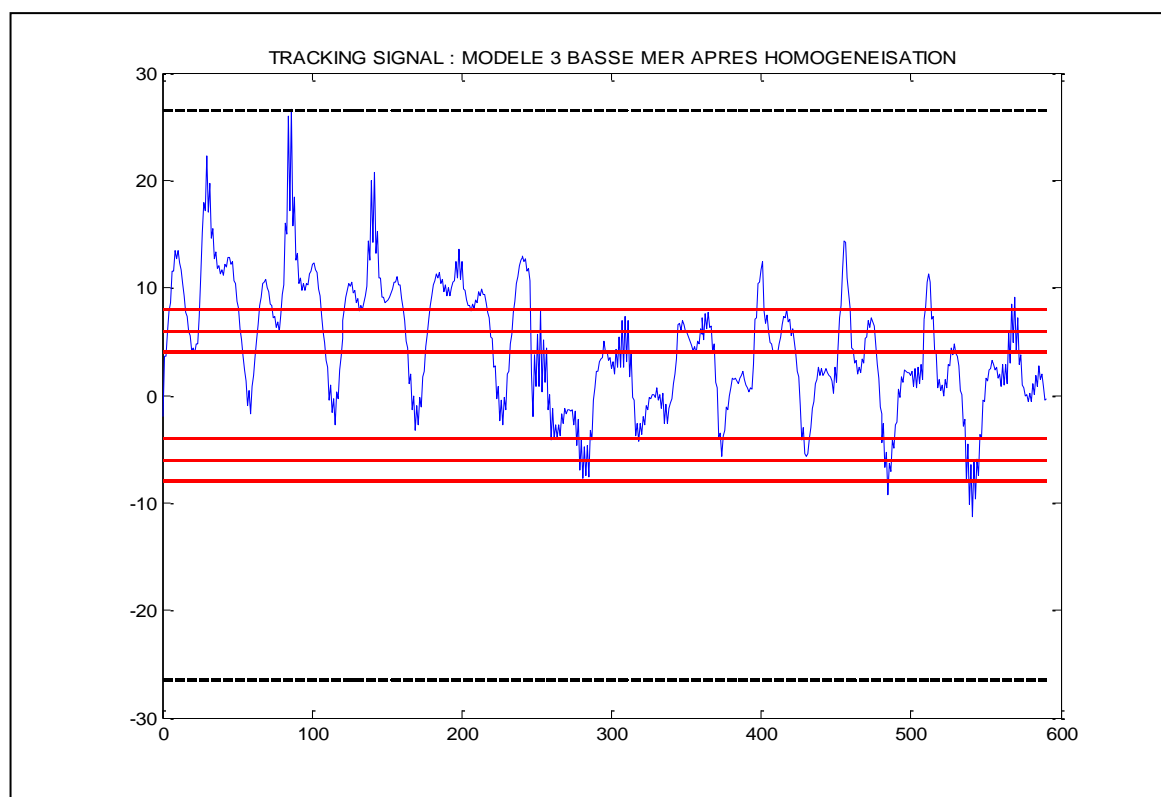


Figure 3 . 96 : Tracking Signal Mod le 3 Basse mer apr s homog n isation

Les graphiques des figures 3.94 à 3.96 montrent que notre modèle arrive à mieux prévoir les basses mers par rapport à la méthode harmonique pour 600 valeurs estimées car le ratio de THEIL qui vaut 0,5103 est inférieur à 1.

Les mêmes figures nous permettent d'affirmer d'une part que notre modèle s'approche de la réalité puisque la RMSE de 0,0253 reste inférieure à 5% et de confirmer d'autre part sa stabilité étant donné qu'aucune valeur du Tracking Signal ne dépasse les limites de l'intervalle $[-27, 27]$. Notre modèle 3 basse mer peut donc être retenu pour prédire le cycle 3 sur les basses mers et les données homogénéisées.

Les sept modèles que nous avons construits arrivent donc tous à prévoir les pleines et les basses mers à Hell-Ville.

Les modèles sur les pleines mers sont plus stables par rapport aux modèles sur les basses mers mais ils sont moins précis par rapport à la méthode harmonique du SHOM

Par contre, les modèles sur les basses mers sont plus précis par rapport à la méthode harmonique.

Nous avons aussi constaté que pour les pleines mers, l'intervalle $[-8, 8]$ suffit pour vérifier la stabilité des modèles à partir de la méthode du Tracking Signal sauf pour le cas homogénéisé ou la limite supérieure devrait être un peu plus grande ou égale à 12,4671.

Pour les basses mers, l'intervalle de confiance que nous avons trouvé pour assurer la stabilité correspond à $[-27, 27]$.

3.2.14. Conclusion

L'analyse des données sur les basses mers et sur les pleines mers à Hell-Ville montre les caractéristiques des séries à mémoire longue (Exposant de Hurst supérieur à 0,5).

Après dessaisonalisation par Census X11 des données stabilisées par transformation logarithmique de base 10, nous avons démontré que le modèle ARIMA-EGARCH arrive à prévoir les données entre 1990 et Juin 2012 et au moins 600 valeurs après le point de rupture (1 Juillet 2012).

Pour confirmer la stabilité des quatre modèles établis sur les pleines mers, l'intervalle de confiance proposé dans la littérature suffit. Par contre pour les basses mers, il est recommandé d'utiliser l'intervalle de confiance du modèle harmonique utilisé par le SHOM pour confirmer la stabilité des trois modèles sur les basses mers.

Dans tous les cas, les modèles établis correspondent à des modèles ARIMA-EGARCH avec des résidus standardisés centrés, réduits et à distribution approximativement normale.

Conclusion générale

Nous avons construit 07 modèles DS (ou processus stationnaire par différenciation), dont 04 sur les pleines mers et 03 sur les basses mers.

Pour les basses mers :

- le modèle 1 est un $ARIMA(10,1,16)$ -EGARCH(4,6), sans constante ni tendance obtenu à partir des données, du cycle 1 basse mer, comprises entre 1 et 4669 ;
- le modèle 2 est à un $ARIMA(15,1,3)$ -EGARCH(6,5), sans constante ni tendance obtenu à partir des données, du cycle 2 basse mer, comprises entre 4670 et 11020 ;
- et le modèle 3 est à un $ARIMA(14,1,9)$ -EGARCH(3,7), sans constante ni tendance obtenu à partir des données, du cycle 3 basse mer, comprises entre 11021 et 15631.

Pour les pleines mers :

- le modèle 1 est un $ARIMA(14,1,1)$ -EGARCH(3,1), sans constante ni tendance obtenu à partir des données, du cycle 1 pleine mer, comprises entre 1 et 4524;
- le modèle 2 est un $ARIMA(15,1,1)$ -EGARCH(1,3), sans constante ni tendance obtenu à partir des données, du cycle 2 pleine mer, comprises entre 4523 et 9106 ;
- le modèle 3 est un $ARIMA(14,1,10)$ -EGARCH(1,5), sans constante ni tendance obtenu à partir des données, du cycle 3 pleine mer, comprises entre 9107 et 13572 ;
- le modèle 4 est un $ARIMA(12,1,1)$ -EGARCH(1,7), sans constante ni tendance obtenu à partir des données, du cycle 4 pleine mer, comprises entre 13573 et 15631.

Ces modèles que nous avons construits sont stables en appliquant la méthode Tracking Signal et permettent d'après notre étude de prévoir au moins 600 données, c'est-à-dire, une prévision de dix mois des phénomènes de marées futurs, après rectification des valeurs aberrantes.

Pour l'établissement des modèles, si nous voudrions utiliser les données à partir du 01 Juillet 2012, alors il faudrait procéder à leurs homogénéisations en retranchant aux points aberrants la différence entre les moyennes des données avant et après le point de rupture.

Nous tenons à signaler que l'homogénéisation doit précéder le choix du modèle qui devrait tenir compte à la fois des résultats graphiques et des résultats par calculs.

Pour les pleines mers, l'intervalle de confiance proposée dans la littérature comprise entre ± 8 suffit pour vérifier la stabilité des modèles sauf pour le modèle 4 contenant des valeurs homogénéisées ou nous avons dû augmenter les limites de confiance entre $\pm 12,467$.

Par contre, les calculs ont montré que pour les basses mers, l'intervalle $[-27, +27]$ convient pour la vérification de la stabilité des modèles par la méthode du Tracking Signal.

Nous avons trouvé que les résidus sont hétéroscédastes et les résidus standardisés ainsi que leurs carrés constituent tous des bruits blancs centrés réduits gaussiens si nous nous référons aux résultats du QQ-plot et non gaussiens mais qui s'en rapproche si nous nous basons aux tests de normalité de Jarque et Bera et de ses variantes.

Ces différents modèles représentent chacun un cycle précis des chroniques sur les basses mers et sur les pleines mers et à l'intérieur duquel les variations saisonnières restent stables.

Pour la pleine mer, ce cycle possède une périodicité entre 4466 (environ 6 ans et 4 mois) et 4582 Lags (environ 6 ans et demi) et pour la basse mer ce cycle est moins précis et se situe entre 4669 (environ 6 ans et 7 mois) et 6351 Lags (environ 9 ans).

Par ailleurs, nous avons également pu démontrer que la tendance est non linéaire décroissante pour les pleines mers et croissantes pour les basses mers si nous utilisons les tests de Mann Kendall mais après visualisation des prévisions, à partir du 01 Juillet 2012, une tendance en hausse a été constatée au niveau des deux chroniques : pleines mers et basses mers.

Nous avons aussi constaté que notre modèle arrive à trouver la même saisonnalité semi-diurne et lunaire trouvée par les méthodes traditionnelles pour les marées de Hell-ville.

Enfin nos modèles sur les basses mers offrent des prévisions plus précises que la méthode harmonique. Ce qui n'est pas le cas de nos modèles sur les pleines mers mais les résultats restent en cohérence avec la réalité.

Ces résultats montrent que la modélisation FARIMA n'est pas utile mais peut être mise en œuvre pour augmenter la précision des modèles.

Avantages et perspectives

Nos modèles présentent l'avantage d'être simple. Le prévisionniste n'a pas besoin de maîtriser les langages de programmation pour exécuter nos modèles, qui sont disponibles directement dans le logiciel Matlab et sous Excel.

En comparaison avec les modèles de Newton et de Laplace, nos méthodes permettent d'effectuer séparément les prévisions, l'analyse des variations saisonnières, l'étude de l'évolution de la tendance. Elles sont également moins restrictives par rapport à la méthode harmonique, qui exige généralement des données d'au moins une année. Or pour notre modèle, si nous voulons effectuer des prévisions mensuelles par exemple, des données mensuelles suffisent.

Bien que le modèle ARIMA-GARCH arrive à prévoir la réalité, les résultats que nous avons obtenus montrent que nous pouvons encore améliorer la précision de nos modèles. L'application du modèle FARIMA-GARCH constitue dans ce cadre une perspective intéressante.

En remarquant que les phénomènes des marées, la température et la pluviométrie à Madagascar possèdent tous une pseudo-période de 6 mois. il est fort probable que notre modèle de calcul peut également être mis en œuvre pour la prévision de la température et de la pluviométrie à Madagascar. Nous pensons que des recherches dans ces domaines peuvent constituer des sujets de thèses futures.

Références bibliographiques

- [1] **NEWTON Isaac, 1687**, « Principes mathématiques de la philosophie naturelle », édition Jacques Gabay, 1990.
- [2] **GUILLAUME Roulet**, Dynamique côtière, *La marée, Master PMMC 2^{ème} année, Spécialité POA, Université de Bretagne Occidentale, ENSTA Bretagne*, 56p,(2011-2012) .
- [3] **SIMON Bernard**, La marée : *la marée océanique cotière*, référence 942-MOG, Collection « Synthèse », Institut océanographique, Fondation Albert I, prince de monaco, 433 p, Janvier 2007.
- [4] **DOODSON A.T**, The Harmonic Development of the Tide-Generating potential, Proceedings of the Royal Society of London, Series A, Containing papers of a Mathematical and Physical Character, volume 100, Issue 704, pp. 305-329, Décembre 1921.
- [5] **KWEMOU Djoukoué Marius**, Réduction de dimension en régression logistique, application aux données Actu-Palu, Thèse de doctorat en mathématique appliquée soutenu le 29 Septembre 2014, effectué en collaboration entre l'Université Evry val-d'essonne et l'Université Gaston Berger de Saint –Louis du Sénégal. p.7.
- [6] **J. LAFFITE**, Mécanique Générale, Institut National des Sciences Appliquées de Toulouse, Edition 1982, 199p.
- [7] Les trois lois de newton et leurs applications
<http://www.tangentex.com/TroisLoisNewton.htm> consulté le 20 septembre 2016.
- [8] **RATIARISON Andriamanga Adolphe**, Cours mécanique Générale II, Filière Mathématiques et Physiques, Faculté des Sciences, Université d'Antananarivo, 1994.
- [9] **QADRI J-Ph**, Référentiel géocentrique et référentiel terrestre, 16p, (2008-2009).
- [10] **BOUYA Ismaël, Chapitre 9**, Changement de référentiels, mécanique, 10p,
<http://melusine.eu.org/syracuse/immae/> consulté le 20 septembre 2016.
- [11] **ELIE Frédéric**, Les marées, version Janvier 2006, 44p,
<http://fred.elie.free.fr> consulté le 10 Janvier 2015.
- [12] https://en.wikipedia.org/wiki/Hurst_exponent, consulté le 20 octobre 2016.
- [13] **PRENAT Michel**, Séries chronologiques, Volume 1 : cours et exercices, Université Paris-Sud, Master Ingénierie Mathématiques, deuxième année, 145p, (2010-2011).
- [14] **CULGIARI Jairo**, Séries chronologiques saisonnières et prévision, L1 IDEA-S, Université Lumière Lyon 2, version compilée le 05 février 2015, 124p, (2014-2015).
- [15] **IMMEDIATO Henri**, Cours de statistique – MASS UE11- Chapitre 2- NTE Lyon 1,
<http://nte-serveur.univ-lyon1.fr/nte/immediato/math2002/mass11/cours/chapitr2b.htm>
consulté le 16 Novembre 2016.

- [16] **THIBAUT Thierry**, 1 introduction et statistique descriptive, Université Nice Sophia Antipolis, 40p,
https://nephi.unice.fr/ftp/users/francour/Cours_Stat_1.pdf consulté le 16 Août 2017.
- [17] **NICOLAU Florence**, Etudes statistiques 2 : Traitement des valeurs manquantes et des valeurs aberrantes, 10p, (2005-2006).
- [18] **BESSON Olivier** ; *Analyse spectrale paramétrique*, 3ème année ENSICA, 102p,
perso.telecom-paristech.fr/~chollet/Biblio/.../obesson_asp.pdf consulté le 16 Août 2017.
- [19] Matlab 2013a, Matlab function.
- [20] **P. GRANJON, J.L LACOUME**, *Introduction à l'analyse spectrale*, 21p
<chamilo2.grenet.fr/inp/courses/.../document/.../IntroAspect.pdf> consulté le 16 Août 2017.
- [21] **J.P.BURG**, *Maximum entropy spectral analysis*. Oklahoma City : in Proc. of the 37th meeting of the society of exploration geophysicists, October 1967.
- [22] **MOHAMMAD-Djafari, Ali**, *Entropie en Traitement du Signal-Entropy in Signal Processing*. Laboratoire de Signaux et Systèmes(cnrs-supélec-ups) Supélec de Moulon, 06 Novembre 2001.
- [23] **J. P. PEREZ**, "*L'Entropie de Boltzmann et l'Entropie de Shannon, même concept*," vol. 92, pp. 145-155, Juin 1998.
- [24] **W.H.PRESS, S.A.TENKOLSKY, W.T. VETTERLING , B. P. Flannery**.*Numerical Recipes in C . s.l. : 2* (1992).
- [25] **THOME Nicolas**, *Base des traitements des images, Transformée de Fourier Avancée- Numérisation*, 87p, 4 Octobre 2016.
- [26] **BERNARD Grais**, Inspecteur général de l'INSEE, Professeur à l'INTEC, Expert auprès de l'Office statistique des Communautés européennes, *Techniques statistiques 2 ; Méthodes statistiques, troisième édition, nouvelle présentation*, DUNOD, 401p, Août 1996.
- [27] **MEYER Fanny, CADRAN Morgane, GAILLARD Margaux**, *Tableau de l'ANOVA*, 86p,
<https://perso.univ-rennes1.fr/valerie.monbet/.../anova.pdf> consulté le 16 Août 2017.
- [28] **RAKOTOMALALA Ricco**, *Comparaison de populations, Tests paramétriques*, Version 1,2, Université Lumière Lyon 2, 101p,
https://eric.univ-lyon2.fr/~ricco/.../Comp_Pop_Tests_Parametrique consulté le 16 Août 2017.
- [29] Dauphine-Eurisko, *Série temporelles linéaires*, Chapitre 4, Estimation, Tests de validation et prévision des processus ARMA, 12 p,
<lacote.ensae.net/.../Dauphine-Eurisko%20-%20Series%20temporelles> consulté le 16 Août 2017.
- [30] **A.VESSEREAU**, Inspecteur Général des Manufactures de l'Etat, *Méthodes statistiques en biologie et en agronomie*, 538p, Septembre 1992.

- [31] **A. JEBRANE**, Test d'égalité des variances, Université de Bourgogne, Master 1 psychologie, 86p, 28 septembre 2016.
- [32] **MANN H.B**, Non parametric tests against trend. *Econometrica*, 13, pp. 245-259, 1945.
- [33] **KENDALL M.G**, *Rank Correlation Methods*. Griffin, London, 202p, 1975.
- [34] **MATIAS Cathérine**, Introduction à la statistique non paramétrique, Centre National de Recherche Scientifique, Laboratoire Statistique et Génome, Evry, 87p, Atelier SFDS du 27 au 28 Septembre 2012,
<http://stat.genopole.cnrs.fr/~cmatias> consulté le 10 Novembre 2016.
- [35] **TSASA VANGU J. Paul, BOSONGA J. Pierre et MUKOKO Daniel**, Econometrie module 1, Université protestante au Congo, centre Congolais –Allemand de Microfinance, 41p, Février 2011.
- [36] **LAGNOUX Agnès**, Séries Chronologiques, Université de Toulouse Le MIRAIL, pp53 ,
www.math.univ-toulouse.fr/~lagnoux/Poly_SC.pdf consulté le 16 Août 2017.
- [37] **GIBERT Thibault**, Etude statistique, Dossier analyse d'une série temporelle, Agriculture – Production laitière bovine en France, Master II Ingénierie de projets, 39p, (2010-2011).
- [38] **LADIRAY Dominique, QUENNEVILLE Benoît**, Comprendre la méthode X11, 192p, Juillet 1999.
- [39] **TENENHAUS Michel**, La méthodologie de Box-Jenkins-Studies2, 109p,
<https://studies2.hec.fr/jahia/.../tenenhaus/.../Box-Jenkins.ppt> consulté le 16 Août 2017.
- [40] Chapitre 2, Tests de non stationnarité et processus aléatoires non stationnaires, 80p,
lacote.ensae.net/.../Dauphine/Dauphine-Eurisko%20-%20Series%20temporelles
consulté le 16 Août 2017.
- [41] **DUCHESNE Pierre**, Méthode de prévision, cours STT-3220, version du 2 Janvier 2017,
- [42] **BOUTHEVILLAIN Karine, MATHIS Alexandre**, Prévision : mesures, erreurs et principaux résultats. In : *Economie et statistique*, N°285-286, pp 88-100, 1995.
http://www.persee.fr/web/revues/home/prescript/article/estat_03361454_1995_num_285_1_5982
consulté le 17 février 2017.
- [43] **RAMUDHIN Amar**, Ingénieur.Ph.D, Prévision de la demande, 38p,
<https://samos.univ-paris1.fr/archives/membres/perraudin/ST/resumechap2.pdf>
consulté le 17 Février 2017.
- [44] **MARIER Philippe**, Ingénieur MBA, coordonnateur de la formation continue et des activités de transfert FORAC, Prévision de la demande, 50p
- [45] Annuaire des marées 2015, Ports d'outre-mer Tome 2, Service Hydrographique et Océanographique de la Marine, 242p, généré le 10 Octobre 2013.

LISTE DE PUBLICATIONS

Afrique science 13(3) (2017) 140 – 168 : Modélisations des pleines et basses mers à Hell-Ville en utilisant la méthode Census X11 et l’approche de Box et Jenkins.

ISSN 1813-518X, <http://www.afriquescience.info>

HEPMAD16 proceedings conference : Modelling tide from Hell-Ville

<http://www.slac.stanford.edu/econf/C161013/>

Modélisations des pleines et basses mers à Hell-Ville en utilisant la méthode Census X11 et l'approche de Box et Jenkins

**Alimanana Nainamahefa RASOLOMANAMPAMONJY *, Adolphe Andriamanga RATIARISON
et Niry Arinavalona RAKOTOVAO**

*Université d'Antananarivo, Laboratoire Dynamique de l'Atmosphère, du Climat et des Océans
(DyACO), Madagascar*

* Correspondance, courriel : nainamahef@gmail.com

Résumé

Ce travail est une initiation à la modélisation Census X11-ARIMA-GARCH des séries chronologiques périodique avec tendance. Avant de mettre en œuvre la méthode Census X11, nous avons montré après l'application d'une transformation logarithmique que nos chroniques sur les pleines mers et sur les basses mers possèdent une saisonnalité commune de 29 Lags (14,765 jours) et des tendances puisque les taux de Kendall de -9,98 % pour les pleines mers et de 11,80 % pour les basses mers sont tous les deux différents de 0. L'analyse de la variance à un facteur de nos deux séries chronologiques nous a permis de découvrir sept cycles compris entre 6 ans et demi et 9 ans à l'intérieur desquels la saisonnalité de 29 lags restait stable. Par conséquent, nous avons décidé de construire sept modèles correspondant à chaque cycle. Pour obtenir les sept modèles, nous avons commencé à appliquer la méthode Census X11 aux données correspondant à chaque cycle. Ensuite, nous avons procédé à la modélisation ARIMA-GARCH des séries corrigée des variations saisonnières obtenues après l'application de Census X11. Finalement, nous avons pu construire sept modèles CensusX11-ARIMA-GARCH avec des résidus non corrélés d'ordre supérieur ou égal à 1 au test de Ljung-Box. Dans cette étude, nous avons utilisé le modèle ARCH(1) ou EGARCH (0,1) pour prendre en compte l'hétéroscédasticité des résidus. Pour valider nos résultats, nous avons comparé nos erreurs avec celles fournies par le logiciel de prédictions des marées du Service Hydrographique et Océanographique de la Marine française (SHOM) de l'édition 2014-2015. Les indicateurs de THEIL de nos modèles sont tous inférieurs à 1 en prenant la méthode harmonique du SHOM comme méthode alternative sauf pour la prévision des pleines mers. Nous pouvons donc affirmer qu'en général, nos modèles offrent une meilleure précision par rapport à la méthode harmonique. Dans le cas de la prévision des pleines mers ou l'indicateur de THEIL est supérieur à 1, nous avons remarqué que la moyenne des erreurs restait inférieure à 5%. Par conséquent, nous pouvons dire que nous avons de très bons modèles pour effectuer la prévision des pleines et basses mers à Hell-ville. Les données des marées que nous avons utilisées dans ce travail proviennent des tables des marées du Centre National de Recherches Océanographiques de Hell-Ville entre 1990 et 2014. La zone d'étude est repérée au point de « latitude 13° 24' Sud et de longitude : 48° 18' Est ». A titre indicatif, le « Lag » sur les figures correspond à la moitié d'une journée lunaire synodique ou à l'équivalent de 12 heures 25 minutes.

Mots clés : *Census X11, ARIMA, GARCH, stratégie de Dickey Fuller Augmenté, Test de Ljung Box, Critères d'Information*

Abstract

Modelling high and low tides from Hell-ville using Census X11 and Box and Jenkins Methodology

This work is an initiation for modeling series with trend and seasonality by Census X11-ARIMA-GARCH. After logarithmic transformation, we showed data have 29 Lags seasonality (14.765 days). Particularly high tides have a decreasing trend and low tides an increasing trend because Kendall's tau have respectively 9.98 % and de11.80 % values. We detect seven local seasonality after applying Analysis of variance one factor located between 6.5 and 9 years. Consequently we decided to establish seven Census X11-ARIMA-GARCH models. Heteroscedasticity of residues was estimated by ARCH(1) or EGARCH(0,1). Finally we established seven models with white noise uncorrelated for order 1 or more after applying Ljung-Box test. To validate our results, we compared our errors with result by the software predictions of tides using by hydrographic and oceanographic service of the french navy on 2014-2015. If alternative method results from software of french navy, our THEIL's ratio were less than 1, except forecasting high tides. Nevertheless the mean absolute percentage error for high tide model was smaller than 5 %. We can conclude that we have excellent model to forecast high and low tides from Hell-ville. Data of tides we used in this work result from tables of tides for the national oceanographic researches center from Hell-ville between 1990 to 2014. The area of study is located in "latitude 13 ° 24 ' South and in longitude: 48 ° 18 ' " East. For information, "Lag" on figures corresponds in the middle of a synodic lunar day or 12 hours and 25 minutes.

Keywords : *Census X11, ARIMA, GARCH, Dickey Fuller Augmented test, Ljung box test, Information Criterion*

1. Introduction

Les prévisions des pleines et basses mers utilisées à Madagascar actuellement proviennent essentiellement des modèles fournis par le service Hydrographique et Océanographique de la Marine Française [1]. Nous avons donc décidé d'effectuer cette étude pour montrer que cette prévision peut être réalisée directement sur place à Madagascar à l'aide d'outils simples et fiables comme Matlab et Excel. Par rapport aux méthodes harmoniques utilisant l'équation du potentiel générateur des marées [2] et les méthodes numériques basés sur l'approximation par éléments finis, différences finis ou volumes finis [3], la méthode que nous proposons qui combine la méthode de dessaisonalisation Census X11 [4] et la méthode ARIMA-GARCH [5, 6] permet d'étudier en détail les différentes composantes du phénomène des marées telles que la saisonnalité, la tendance et cycles et les fluctuations résiduelles. En bref, notre méthode devrait nous permettre d'établir les tables des marées à Hell-ville et d'étudier en même temps l'évolution de différentes composantes des marées au court du temps. Nous aimerons également fournir au lecteur de cette publication une démarche rationnelle pour effectuer la modélisation Census X11 combinée avec la modélisation ARIMA-GARCH.

2. Matériel et Méthodes

2.1 Matériel

Nous avons utilisé les matériels du laboratoire DyaCo, Matlab 2013a et l'utilitaire de Microsoft Office Excel 2007 pour établir et valider nos modèles. L'échantillon est composé de 35 286 individus : 17 643 données sur les pleines mers et 17 643 données sur les basses mers.

2.2 Méthodes

- *Représentation graphique des données initiales*

Elle permet de repérer graphiquement l'existence d'une saisonnalité, d'une tendance et d'une rupture de pente au niveau des chroniques [7].

- *Histogramme*

L'histogramme renseigne sur l'hétérogénéité des données. Pour une donnée hétérogène, nous observerons sur l'histogramme plusieurs modes [8].

- *Exposant de Hurst H [9]*

L'exposant de Hurst informe sur le degré de persistance des données.

Selon Hurst, l'étendue R , l'écart type σ , et H sont reliés par la relation :

$$\frac{R(\tau)}{\sigma(\tau)} = a\tau^H \quad (1)$$

La pente de la régression linéaire de $\log\left(\frac{R}{\sigma}\right)$ sur $\log \tau$ constitue donc une bonne approximation de H

$$\bar{x} = \frac{1}{\tau} \sum_{i=1}^{\tau} x_i \quad (2)$$

$$y_j = x_j - \bar{x} \quad j \text{ variant de } 1 \text{ à } \tau \quad (3)$$

$$z_j = \sum_{i=1}^j y_i \quad j \text{ variant de } 1 \text{ à } \tau \quad (4)$$

$$R(\tau) = \max_{1 \leq j \leq \tau} (z_j) - \min_{1 \leq j \leq \tau} (z_j) \quad (5)$$

$$\sigma(\tau) = \sqrt{\frac{1}{\tau} \sum_{i=1}^{\tau} (x_i - \bar{x})^2} \quad (6)$$

τ : Taille de l'échantillon prise égale à $N, N/2, N/4 \dots$; N : Taille de l'échantillon total

- Si $0,5 < H$, les données possèdent une mémoire courte
- Si $0,5 > H$, les données possèdent une mémoire longue
- Si $0,5 = H$, les données n'ont pas de mémoire

- *Préparation des données*

Nous avons utilisé la méthode de la boîte à moustache pour détecter les points aberrants. Ce sont les points situés en dehors de l'intervalle $[a, b]$ défini par [10] :

$$[a = Q1 - 1,5 * (Q3 - Q1), b = Q3 + 1,5 * (Q3 - Q1)] \quad (7)$$

$Q1$: Premier quartile ; $Q3$: Troisième quartile

Nous avons corrigé nos données en appliquant des moyennes mobiles pondérées définies par les relations suivantes :

$$X(t) = \sum_{i=t-1}^{t+1} c(i) * x(i) \quad \text{pour les données à l'intérieur de la chronique} \quad (8)$$

$$X(t) = \sum_{i=t}^{t+2} c(i) * x(i) \quad \text{pour les données situées au début de la chronique} \quad (9)$$

$$X(t) = \sum_{i=t-2}^t c(i) * x(i) \quad \text{pour les données se trouvant à la fin de la chronique} \quad (10)$$

$x(i)$: Donnée à l'instant « i » ; $c(i)$: Coefficient de pondération égal à $\frac{|b-x(i)|}{b-a}$ pour les points aberrants et égal à 1 pour les autres points ; a et b sont les bornes inférieures et supérieures limitant la boîte à moustache. Nous avons éliminé certains points à la fin de notre chronique puisque nous ne possédons pas les informations nécessaires pour les corriger.

- *Saisonnalité S*

Nous avons recherché la saisonnalité S en utilisant la méthode du corrélogramme et la méthode du périodogramme avec une fonction d'apodisation rectangulaire.

➤ *Méthode du Corrélogramme ou méthode de BLACKMAN-TUCKER [11]*

Cette méthode permet d'estimer la densité spectrale en utilisant les formulations suivantes [12,13]:

$$DSP(f) = \text{fft} [\hat{R}_{xx}(k)] \quad (11)$$

$$\hat{R}_{xx}(k) = \frac{1}{N-k} \sum_{t=0}^{N-k-1} x_{t+k} x_t^* \quad (12)$$

$$\text{fft}(k) = T_e \sum_{t=1}^N x_t e^{-2i\pi(t-1) \frac{(k-1)}{L}} \quad (13)$$

fft : Transformée de Fourier Rapide ; *DSP* : Densité Spectrale ; *N* : Nombre de points temporels ; *t* : Variable temporelle variant de 1 à *N* ; x_t : Données ; *L* : Nombre de points fréquentiels ; *k* : Variable fréquentielle variant de 1 à *L* ; \hat{R}_{xx} : Estimation de la fonction d'autocorrelation R_{xx} par sa forme non biaisée ; T_e : Période de l'échantillon ; F_e : Fréquence de l'échantillon.

➤ *Méthode du Périodogramme moyenné [13]*

Dans cette méthode, la DSP est estimée par la relation :

$$\widehat{DSP}(f) = \frac{1}{M} \sum_{i=1}^M |\text{fft}(k)|^2 \quad (14)$$

M : Nombre d'échantillons égal à L/T_e

➤ *Périodogramme par la Méthode de l'Entropie Maximale [14 -18]*

C'est une variante de la méthode du corrélogramme. Elle consiste à maximiser l'entropie ou la probabilité d'obtenir le maximum d'information sur les données. Selon Burg [19], on peut obtenir une estimation de la DSP en utilisant la fonction d'autocorrélation et la fonction de l'entropie définies par :

$$E(x_{t+k} x_t^*) = \hat{R}_{xx}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} DSP(\omega) e^{jk\omega} d\omega \quad (15)$$

$$H = \int_{-\pi}^{\pi} \ln DSP(\omega) d\omega \quad (16)$$

ω : pulsation ; j : nombre complexe, $j^2 = -1$.

Pour un bruit blanc gaussien centré, en maximisant H, Burg arrivait à formuler la DSP de la manière suivante en utilisant un modèle autorégressif [20]:

$$DSP(\omega) = \frac{\sigma_q^2}{|1 + \sum_{i=1}^q a_i e^{ji\omega}|^2} \quad (17)$$

σ_q^2 : variance du bruit blanc issu du model autoregressive d'ordre 'q'; a_i : Coefficients du model autoregressif d'ordre 'q'

- *Saisonnalité S en terme journalier*

Notons que la saisonnalité S trouvée par les trois méthodes précédentes est fonction des temps de prélèvement. Pour le convertir en unité journalière, il suffit de la multiplier par la moyenne arithmétique des différences des temps de prélèvement successifs en unité journalière sur les pleines et les basses mers. Cette correspondance en terme journalier d'un Lags sur les figures serait définie par la relation suivante.

$$1 \text{ Lag} = \frac{\sum_{i=1}^N D_i}{N} \quad (18)$$

D_i : Différence en jour entre deux temps de prélèvement consécutifs

Nous pourrions ensuite comparer la valeur des saisonnalités S trouvées par les trois méthodes à celles fournies par la théorie statique de Newton en utilisant le « **test t de Student** » qui consiste à vérifier l'égalité de la moyenne d'une série à une valeur donnée et que nous détaillerons à la fin de cette méthodologie quand nous testerons l'égalité de la moyenne des résidus normalisés et la valeur nulle.

- *Choix de la Saisonnalité S*

Comme la méthode du corrélogramme et celle du périodogramme donnent en général des valeurs non entière, il nous faudrait maintenant trouver la meilleure estimation en valeur entière de cette saisonnalité S pour pouvoir mettre en œuvre la méthode Census X-11. Nous avons donc eu recours à :

- Un test sur les résidus et à
- Une analyse de la variance à un facteur pour trouver S
- *Test sur les résidus [21]*

Cette méthode consiste à choisir la saisonnalité S créant moins de fluctuation résiduelle. (Propriété des moyennes mobiles)

- *Analyse de la Variance à un facteur*

Cette technique est utilisée si les données sont indépendantes, homoscedastes et possède une distribution normale ou approximativement normale. [22, 23]. Pour tester la normalité, nous avons utilisé le test de Jarque et Bera et deux autres tests basés sur le coefficient d'aplatissement et le coefficient d'asymétrie [24]. La distribution de la variable h est normale:

$$\text{Si } JB = \frac{N}{6} S_t^2 + \frac{N}{24} K_t^2 \leq 5,99 \quad \text{ou} \quad (19)$$

$$\text{Si } |S| = \sqrt{\frac{N}{6}} \cdot S_t \leq 1,96 \quad \text{ou} \quad (20)$$

$$\text{Si } |K| = \sqrt{\frac{N}{24}} \cdot K_t \leq 1,96 \quad (21)$$

$$S_t = \frac{\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^3}{\left[\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^2 \right]^{3/2}} \quad (22)$$

$$K_t = \frac{\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^3}{\left[\frac{1}{N} \sum_{i=1}^N (h_i - \bar{h}_i)^2 \right]^2} - 3 \quad (23)$$

Pour tester si une distribution s'approche de la loi normale, nous avons utilisé les relations entre l'écart type, l'écart moyen et l'écart probable d'une distribution normale définies par [25]:

$$\mathbf{ad} = \frac{\sum_{i=1}^N |x(i) - \bar{x}|}{N} = 0,7979 \times \sigma \quad (24)$$

$$\mathbf{pd} = \frac{Q3 - Q1}{2} = 0,6745 \times \sigma \quad (25)$$

σ : Ecart type , \mathbf{ad} : Ecart moyen , \mathbf{pd} : Ecart probable

Si l'écart en valeur absolue entre le premier terme et le second terme des **équations 24 et 25** reste inférieur à 5% , nous avons décidé d'accepter la vraisemblance avec la loi normale. L'homoscedasticité a été testée par le test de Bartlett et par le test de Levene [26]:

Test de Bartlett : Si $\frac{B}{c} < \chi^2 (l - 1)$, nous concluons que les données sont homoscedastes

$$B = (N - l) \ln(s_p^2) - \sum_{i=1}^l (N_i - 1) \ln(s_i^2) \quad (26)$$

$$s_p^2 = \frac{\sum_{i=1}^l (N_i - 1) s_i^2}{n - l} \quad (27)$$

$$C = 1 + \frac{1}{3(l-1)} \left(\sum_{i=1}^l \frac{1}{N_i - 1} - \frac{1}{N - l} \right) \quad (28)$$

Test de Levene : Si $D < Fisher (l - 1; n - l)$, nous concluons que les données sont homoscedastes

$$D = \frac{(n-l) \sum_{i=1}^l N_i (\bar{E}_i - \bar{E})^2}{(l-1) \sum_{i=1}^l \sum_{j=1}^{N_i} (E_{i,j} - \bar{E}_i)^2} \quad (29)$$

$$E_{i,j} = |x_{i,j} - \bar{X}_l| \quad (30)$$

$$\bar{E}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} E_{i,j} \quad (31)$$

$$\bar{E} = \frac{1}{N} \sum_{i=1}^l N_i \bar{E}_i \quad (32)$$

N : Taille totale égale à $N_i * l$; l : Nombre de groupe égal à S ; N_i : Taille de chaque groupe ; s_i^2 : Variance de chaque groupe correspondant au colonne « i » ; $x_{i,j}$: Variable correspondant au colonne « i », ligne « j » ; \bar{X}_i : Moyenne du groupe « i » ; Fisher $(l-1; n-l)$: Valeur de F sur la table de Fisher pour « $(l-1; n-l)$ » degré de liberté ; $\chi^2(l-1)$: Valeur de Khi-2 sur table avec « $l-1$ » degré de liberté.

Pour le choix de S entière, nous avons maximisé la fonction $(1/f, DSP)$. Le domaine de définition de f a été estimé en tenant compte de la parité de la fonction DSP et du mode d'itération sous Matlab qui commence toujours par 1. Nous avons alors défini ce domaine par l'intervalle [27] :

$$f = \frac{F_{Nyquist}}{\frac{L}{2}} * \left[1, \frac{L}{2}\right] = \frac{F_e}{\frac{L}{2}} * \left[1, \frac{L}{2}\right] \quad (33)$$

$F_{Nyquist}$: Fréquence de Nyquist

- Analyse de la tendance [28]

L'analyse de l'existence d'une tendance monotonique dans notre travail est basée sur la mise en œuvre du test de corrélation des rangs de Kendall.

Nous avons retenu la définition du coefficient de Kendall τ_b suivante :

$$\tau_b = \frac{S''}{\sqrt{\frac{N(N-1)}{2} E_x} \sqrt{\frac{N(N-1)}{2} E_y}} \quad (34)$$

$$E_x = \sum_{x=1}^{m1} \frac{N_x (N_x - 1)}{2} \text{ et } E_y = \sum_{y=1}^{m2} \frac{N_y (N_y - 1)}{2} \quad (35)$$

$m1$: nombre de groupes d'ex-aequos sur les rangs R_x du grandeur X ; N_x : nombre d'ex-aequos contenu dans le x -ième groupe ; $m2$: nombre de groupes d'ex-aequos sur les rangs R_y du grandeur Y ; N_y : nombre d'ex-aequos contenu dans le y -ième groupe.

Pour le terme S'' correspondant au changement de signe de la différence des rangs, nous avons utilisé la formulation suivante :

$$S'' = \sum_{i < j}^N \text{signe}(X_i - X_j) \text{signe}(Y_i - Y_j) \text{ avec } \text{signe}(G) = \begin{cases} +1 & \text{si } G > 0 \\ 0 & \text{si } G = 0 \\ -1 & \text{si } G < 0 \end{cases} \quad (36)$$

Sous l'hypothèse H_0 : « il n'y a pas de corrélation entre X et Y », la moyenne et la variance de τ_b peuvent être estimées par les relations suivante :

$$E(\tau_b) = 0 \text{ et } \quad (37)$$

$$Var(\tau_b) = \frac{2(2N+5)}{9N(N-1)} \quad (38)$$

En particulier, pour un grand échantillon, la règle de décision est si la valeur absolue de $Z_1 = \frac{\tau_b - E(\tau_b)}{\sqrt{Var(\tau_b)}}$ est supérieure à la valeur critique $U_{0,1,0,05}$ issue de la loi normale centrée- réduite $N(0,1)$, nous rejetons l'hypothèse H_0 en faveur de l'hypothèse H_1 : « il existe une relation monotone entre les deux grandeurs X et Y ».

- *Schéma de prévision [25]*

En général, il existe deux schémas de prévision :

$$\text{Le schéma additif} \quad : x_t = t_t + s_t + r_t \quad (39)$$

$$\text{Le schéma multiplicatif} : x_t = t_t * s_t * r_t \quad (40)$$

t_t : *Tendance* ; s_t : *Composante saisonnière* ; r_t : *Résidus ou innovation*

Nous avons testé les trois méthodes suivantes pour reconnaître le schéma adéquat [30]:

- Méthode du profil
- Méthode de la bande
- Méthode analytique

➤ *Méthode du profil*

Elle consiste à tracer les graphiques des données pour chaque période.

- Si les courbes sont confondues ou parallèles, on choisit le schéma multiplicatif
- Si les courbes s'entrecroisent, on adopte un schéma additif

➤ *Méthode de la bande*

La méthode de la bande consiste à construire pour chaque période, d'une part la courbe des minima et d'autre part la courbe des maxima.

- Si les deux courbes sont parallèles, on opte pour le schéma multiplicatif
- Si les deux courbes s'entrecroisent, on choisit un schéma additif

➤ *Méthode analytique*

La méthode analytique consiste à régresser pour chaque période l'écart type σ_i par rapport à la moyenne \bar{x}_i . Si le terme μ dans l'équation $\sigma_i = \mu \bar{x}_i + g$ est significatif, c'est à dire si la valeur absolue de la statistique de Student correspondant à μ après regression est supérieure à la valeur critique dans la table de Student, on choisit le schéma multiplicatif.

- *Méthode Census X11 [4]*

Les résultats de l'analyse de la saisonnalité et de la tendance nous informent sur les différentes composantes de notre schéma de prévision. Avant la mise en œuvre de la méthode Census X11, nous avons dû stabiliser notre série en appliquant une transformation logarithmique de base 10 qui a également permis de transformer notre schéma multiplicatif en schéma additif.

➤ *Tendance $t_{t,1}$ par moyenne mobile centrée d'ordre S [29]*

La tendance a été estimée en premier lieu en utilisant une moyenne mobile centrée d'ordre S. Si S est paire, c'est-à-dire de la forme $2m$, on utilise la relation :

$$t_{t,1} = \frac{1}{2m} \left(\frac{1}{2}x_{t-m} + x_{t-m+1} + \dots + x_{t+m-1} + \frac{1}{2}x_{t+m} \right) = M_{2m}(x_t) \quad (41)$$

Si S est impaire, c'est-à-dire de la forme $2m + 1$, on utilise la relation :

$$t_{t,1} = \frac{1}{2m+1} (x_{t-m} + x_{t-m+1} + \dots + x_{t+m-1} + x_{t+m}) = M_{2m+1}(x_t) \quad (42)$$

m : un entier positif

➤ *Saisonnalité-Innovation : $s_{t,1} + r_{t,1}$ [4]*

Le terme Saisonnalité-Innovation est obtenu par différence entre les données initiales et la tendance

$$s_{t,1} + r_{t,1} = x_{t,1} - t_{t,1} \quad (43)$$

La saisonnalité $s_{t,1}$ se déduit en appliquant une moyenne mobile 5 termes « 3*3 » à cette valeur de la Saisonnalité-Innovation puis en soustrayant au résultat de $M_{3*3}(x_{t,1} - t_{t,1})$ la valeur correspondant à sa moyenne mobile centrée d'ordre S. $s_{t,1}$ s'obtient alors en utilisant la formule

$$s_{t,1} = [M_{3*3}(x_{t,1} - t_{t,1})] - M_{2m+1}[M_{3*3}(x_{t,1} - t_{t,1})] \quad (44)$$

M_{3*3} : moyenne mobile centrée à 5 termes de coefficient de pondération égal à $\left\{ \frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{1}{9} \right\}$

➤ *Série corrigée des variations saisonnières $scvs_{t,1}$ [4]*

La $scvs_{t,1}$ se déduit en appliquant la relation :

$$scvs_{t,1} = x_{t,1} - s_{t,1} \quad (45)$$

➤ *Tendance $t_{t,2}$ par moyenne mobile de Henderson [4]*

La tendance est ici estimée en utilisant une moyenne mobile de Henderson à 9 ou 13 termes. Donc nous pouvons écrire :

$$t_{t,2} = H_{13 \text{ ou } 9}(scvs_{t,1}) \quad (46)$$

Le principe est simple.

- Si le rapport $\frac{I}{C}$ est supérieur à 1, nous appliquons une moyenne mobile de Henderson à 13 termes
- Si le rapport $\frac{I}{C}$ est inférieur à 1, nous appliquons une moyenne mobile de Henderson à 9 termes.

I et C sont donnés par les formules suivantes :

$$C = \frac{1}{n-1} \sum_{i=2}^n (t_{i,2} - t_{i-1,2}) \quad (47)$$

$$I = \frac{1}{n-1} \sum_{i=2}^n (scvs_{i,1} - t_{i,2}) - (scvs_{i-1,1} - t_{i-1,2}) \quad (48)$$

➤ *Nouvelle expression de la Saisonnalité-Innovation : $s_{t,2} + r_{t,2}$ [4]*

Après connaissance de l'estimation de la tendance $t_{t,2}$, nous avons réestimé la Saisonnalité-Innovation en utilisant la relation :

$$s_{t,2} + r_{t,2} = x_{t,1} - t_{t,2} \quad (49)$$

Et le terme $s_{t,2}$ s'obtient en utilisant la relation :

$$s_{t,2} = [M_{3*5}(x_{t,1} - t_{t,2})] - M_{2m+1}[M_{3*5}(x_{t,1} - t_{t,2})] \quad (50)$$

M_{3*5} : moyenne mobile centrée à 7 termes de coefficient de pondération égal à $\left\{ \frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15} \right\}$

➤ *Nouvelle expression de la série corrigée des variations saisonnière $scvs_{t,2}$ [4]*

Finalement, nous obtenons la $scvs_{t,2}$ en utilisant la relation :

$$scvs_{t,2} = x_{t,1} - s_{t,2} \quad (51)$$

Dans cette étude, l'intervalle de la boîte à moustache a été estimé en utilisant toutes les données. Par contre le logiciel Census X11 utilise en général :

- Les données appartenant à 2 saisonnalités antérieures au point considéré
 - Les données appartenant à la saisonnalité du point considéré et
 - Les données appartenant à 2 saisonnalités futures au point considéré
- *Méthodologie de Box et Jenkins*

La méthodologie expliquée dans cette étude concerne la modélisation ARIMA (p+1,d,q) de Box et Jenkins. Dans ce modèle « q » correspond à l'ordre de la partie à **Moyenne Mobile**, « d » à l'ordre d'Intégration et « p+1 » à l'ordre de la partie **Auto-Regressive**. Le problème se ramène donc à la recherche de p, d et q.

➤ *Identification de d, q et p*

Box et Jenkins utilisaient la formulation suivante pour éliminer la tendance et stationnariser la série [31]

$$r_t' = (I - B)^d scvs_{t,2} \quad (52)$$

B : Opérateur de retard ; r_t' : Processus stationnaire

Nous avons fixé un ordre assez élevé de retard « p » avant de mettre en œuvre la stratégie de Dickey Fuller Augmenté (DFA) en variant « d entre 0 et 2 ». Si le test de Dickey Fuller Augmenté confirme la stationnarité de la série différenciée alors le « d et p » choisis correspondent à l'ordre d'intégration

adéquat $d_{\text{adéquat}}$ et à un nombre de retard maximal acceptable « p_{max} ». Si le test de Dickey Fuller Augmenté rejette l'hypothèse de stationnarité alors nous devrions augmenter le nombre de retard d'une ou de plusieurs unités et répéter l'opération qui cette fois ci devrait confirmer la stationnarité au sens de Dickey Fuller Augmenté. Cette nouvelle valeur de « p » correspondra alors à l'ordre maximal « p_{max} » du modèle DFA. Une fois p_{max} connu, on détermine le nombre de retard optimal « p_{optimal} » des modèles DFA, en fixant « d égal à $d_{\text{adéquat}}$ et p égal à p_{max} », puis en variant progressivement le « p » dans chaque modèle entre 0 et p_{max} et enfin en minimisant le critère de Schwarz défini par la relation [32]:

$$SC(p + q) = N * \log(\sigma_{\varepsilon_t}^2) + (p + q + 3) * \log(N) \quad (53)$$

$\sigma_{\varepsilon_t}^2$: Variance des résidus.

Avec les valeurs de « $d_{\text{adéquat}}$ et de p_{optimal} » précédentes, nous pourrions maintenant rechercher les valeurs de q_{max} et de q_{optimal} en augmentant progressivement q d'une unité. Le q_{max} correspond à la valeur maximale de q rendant non corrélés les résidus au test de Ljung-Box et le q_{optimal} n'est autre que la valeur de q compris entre 0 et ce q_{max} minimisant le critère d'Information d'Akaike défini par la relation [32]:

$$AIC(p + q) = N * \log(\sigma_{\varepsilon_t}^2) + 2(p + q) \quad (54)$$

Le test de Ljung-Box consiste à comparer la p-Value de la statistique de Ljung-Box à la valeur 0,05 [31].

$$\chi_m^2 = N(N + 2) \sum_{k=1}^m \left[\frac{a_k^2(r_t)}{N - k} \right] \quad (55)$$

a_k : Estimation de l'autocorrelation

Si la p-Value ou $Prob[\chi^2(m - a) > \chi_m^2]$ est supérieure ou égale à 0,05 pour un ordre « k » fixé, la série est stationnaire. Finalement, connaissant p_{optimal} , $d_{\text{adéquat}}$ et q_{optimal} , nous pouvons estimer par maximum de vraisemblance les modèles Arima ($p_{\text{optimal}} + 1, d_{\text{adéquat}}, q_{\text{optimal}}$) équivalents à nos modèles DFA et procéder en même temps à l'élimination des coefficients non significatifs. Le test de Dickey Fuller Augmenté consiste à appliquer la stratégie suivante [32] :

$$\text{Modèle DFA 1 : } r_t' = \Phi r_{t-1}' + \sum_{j=1}^p \Phi_j \Delta r_{t-j}' + \varepsilon_t \quad (56)$$

$$\text{Modèle DFA 2 : } r_t' = \Phi r_{t-1}' + a + \sum_{j=1}^p \Phi_j \Delta r_{t-j}' + \varepsilon_t \quad (57)$$

$$\text{Modèle DFA 3 : } r_t' = \Phi r_{t-1}' + bt + a + \sum_{j=1}^p \Phi_j \Delta r_{t-j}' + \varepsilon_t \quad (58)$$

La règle suggère de toujours commencer par tester le modèle DFA 3.

✓ Hypothèse 1

H_0 : $\Phi = 1$, il existe une racine unitaire ;

H_1 : $|\Phi| < 1$, il n'y a pas de racine unitaire, la série est stationnaire

✓ Décision :

- Si la statistique de la loi t de Student pour Φ est inférieur à la statistique de Dickey-Fuller sur table, on rejette H_0 .
- Si la statistique de la loi t de Student pour Φ est supérieur à la statistique de Dickey-Fuller sur table, on accepte H_0

Si la série est stationnaire au sens de Dickey Fuller Augmenté, on teste la significativité de la tendance « b » et de la constante « a » en utilisant la loi t de Student défini par :

✓ Hypothèse 2 :

Ho : « b » = 0 , Pas de tendance

H1 : « b » ≠ 0 , Existence de tendance

✓ Décision :

- Si la statistique de la loi t de Student pour « b » est supérieure à 1,96, on rejette Ho
- Si la statistique de la loi t de Student pour « b » est inférieure à 1,96, on accepte Ho

La significativité de la constante « a » se confirme de la même manière que pour la tendance avec comme hypothèse:

Ho : « a » = 0 , Pas de constante

H1 : « a » ≠ 0 , Existence du terme constant dans le modèle

➤ Analyse des Résidus

L'hétéroscédasticité a été testée en utilisant le test ARCH sous Matlab 2013a défini par [12]

$$r_t^2 = \pi_o + \sum_{k=1}^W \pi_k r_{t-k}^2 + e(t) \quad (59)$$

r_t : Vecteur sur les résidus

Elle est confirmée si la p-value issue du test ARCH est inférieure à 0,05. En cas d'hétéroscédasticité, les résidus seront modélisés par un modèle de type ARCH ou EGARCH.

ARCH model [12]:

$$r_t = \eta_t \sigma_t \quad (60)$$

$$\sigma_t^2 = K + \sum_{j=1}^Q \alpha_j r_{t-j}^2 \quad (61)$$

EGARCH model [12]:

$$r_t = \eta_t \sigma_t$$

$$\log \sigma_t^2 = K + \sum_{i=1}^P \gamma_i \sigma_{t-i}^2 + \sum_{j=1}^Q \alpha_j \left[\frac{|r_{t-j}|}{\sigma_{t-j}} - E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} \right] + \sum_{j=1}^Q \xi_j \left(\frac{r_{t-j}}{\sigma_{t-j}} \right) \quad (62)$$

$$E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} = E \{ |\eta_{t-j}| \} = \sqrt{\frac{2}{\pi}} \text{ pour une distribution gaussienne des résidus}$$

$$E \left\{ \frac{|r_{t-j}|}{\sigma_{t-j}} \right\} = E \{ |\eta_{t-j}| \} = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \text{ pour une distribution de Student avec } \nu \geq 2 \text{ degré de liberté}$$

K : constante du modèle ; γ_i : coefficients GARCH ; α_j : coefficients ARCH ; ξ_j : Leverage

La **normalité** de la distribution des résidus standardisés sera confirmée ou rejetée en appliquant l'une des définitions que nous avons utilisées pour confirmer la normalité dans l'analyse de la variance à un facteur. Pour vérifier que la **distribution** des résidus standardisés est **centrée et réduite**, il suffit de tester l'égalité entre les paramètres (\bar{r}_t, σ^2) et $(0, 1)$. Pour vérifier l'égalité entre la **moyenne des résidus normalisés** \bar{r}_t la **valeur zéro (0)**, nous avons appliqué le test d'égalité de la moyenne d'une série à une valeur donnée ou test de Student. Cette technique permet de déterminer l'intervalle de confiance symétrique à 95% pour zéro [30]. En notant $[c, d]$ cet intervalle la **règle de décision** est la suivante.

- Si $0 \in [c, d] = [\bar{r}_t - t_{1-\alpha/2} \frac{s}{\sqrt{N}}, \bar{r}_t + t_{1-\alpha/2} \frac{s}{\sqrt{N}}]$, alors la moyenne \bar{r}_t est non significativement différente de zéro.
- Si $0 < \bar{r}_t - t_{1-\alpha/2} \frac{s}{\sqrt{N}}$ ou $\bar{r}_t + t_{1-\alpha/2} \frac{s}{\sqrt{N}} > 0$ alors la moyenne \bar{r}_t est significativement différente de zéro.

\bar{r}_t : Moyenne des résidus normalisés ; σ : Écart type biaisé de l'échantillon ; s : Ecart type non biaisé égal à $\sigma * \frac{N}{N-1}$; N : Taille de l'échantillon ; $t_{1-\alpha/2}$: Variable de Student à $N-1$ degré de liberté pour un risque de première espèce α .

De la même manière, on teste l'égalité entre la **variance des résidus normalisés et un (1)** [25]. La règle de décision est la suivante.

- Si $1 \in [s^2 - t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}, s^2 + t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}]$, alors la variance s^2 est non significativement différente de un.
- Si $1 < s^2 - t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}}$ ou $s^2 + t_{1-\alpha/2} s^2 \sqrt{\frac{2}{N}} > 1$ alors la variance s^2 est significativement différente de un.

➤ *Vérification de la qualité de la prédiction et de la prévision [33,34]*

Cette étape consiste à comparer les erreurs de prédiction sur les données initiales et les erreurs de prévision sur les données futures avec celles d'une méthode alternative en utilisant l'indicateur de THEIL et l'Erreur Moyenne Absolue en Pourcentage (EMAP) définis par les formules suivantes.

$$I_{THEIL} = \frac{RMSE_{modèle}}{RMSE_{méthode alternative}} \quad (63)$$

$$RMSE = \sqrt{\frac{1}{N_t} \sum_{t=1}^{N_t} (x_t - x_{modèle})^2} \quad (64)$$

$$EMAP = \frac{\sum_{t=1}^{N_t} \left| \frac{x_t - x_{modèle}}{x_t} \right| * 100}{N_t} \quad (65)$$

N_t : Taille de l'échantillon ; x_t : Donnée initiale à l'instant t ; $x_{modèle}$: Valeur prédite de x_t à l'instant t ; N_t : Taille de l'échantillon sur laquelle on effectue la projection

Si I_{THEIL} est inférieur à 1, le modèle offre une meilleure précision par rapport à la méthode alternative pendant la durée de comparaison. Le modèle est bon si l'EMAP est inférieure à 10% et il devient excellent pour les valeurs en dessous de 5%.

3 Résultats et discussion

3.1 Représentation graphique des données initiales

Sur la **Figure 1**, nous observons une rupture de pente autour du 31 700 Lags. Soit les prélèvements sont erronés, soit un facteur climatique à engendré cette évolution. Mais comme nous n'avons pas les informations sur le protocole de prélèvement des données au Centre National Océanographique de Hell-ville, nous avons préféré enlever ces données aberrantes de nos chroniques pour l'établissement des modèles.

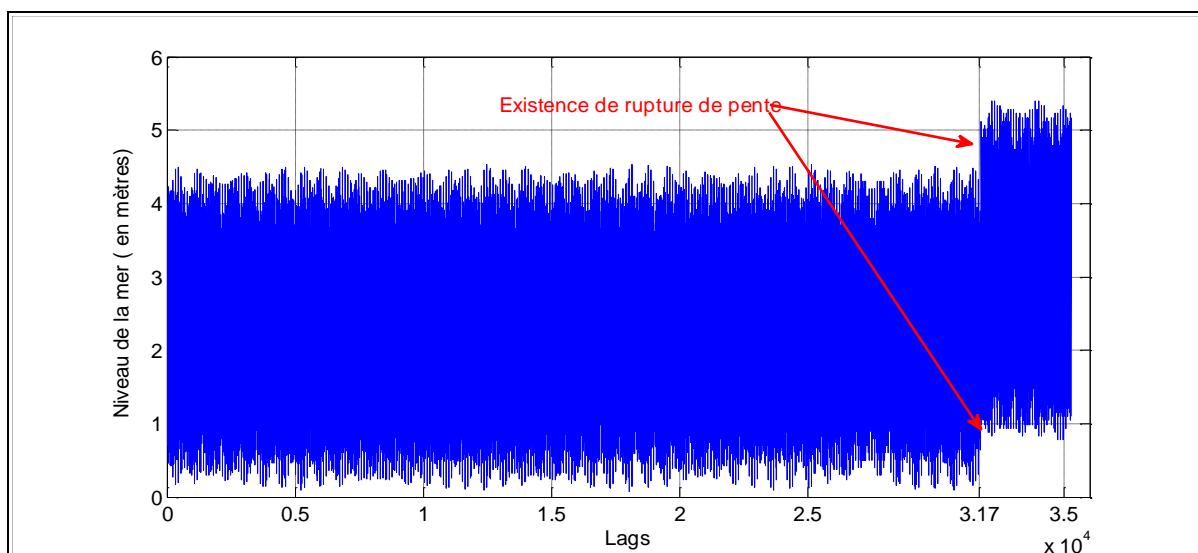


Figure 1 : *Données initiales*

3.2 Histogramme

Sur la **Figure 2**, l'histogramme des données initiales est bimodal et ceux des données séparées (pleines mers d'une part et basses mers d'autre part) sont unimodaux. Il est donc préférable de travailler sur les données séparées pour contourner le problème d'hétérogénéité entre les prélèvements issus des pleines mers et ceux issus des basses mers.

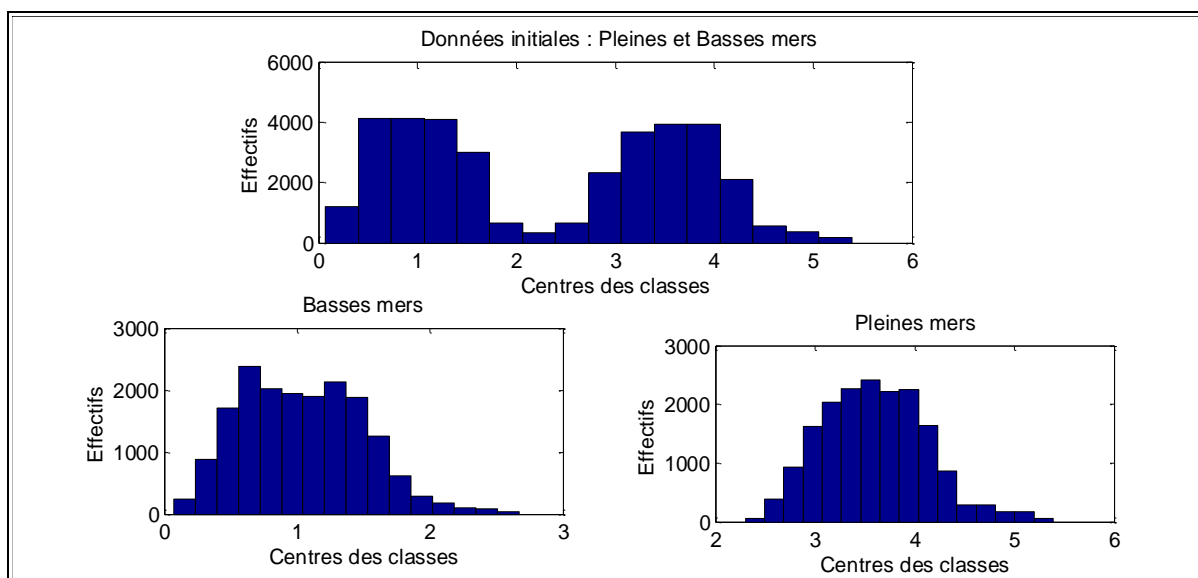


Figure 2 : *Histogrammes*

Par conséquent, nous avons décidé d'effectuer séparément la modélisation des pleines mers et la modélisation des basses mers.

3.3 Exposant de Hurst

Sur la **Figure 3**, nous avons trouvé pour la série corrigée des variations saisonnières un exposant de Hurst H égal à 0,8020 pour les basses mers et H égal à 0,8510 pour les pleines mers.

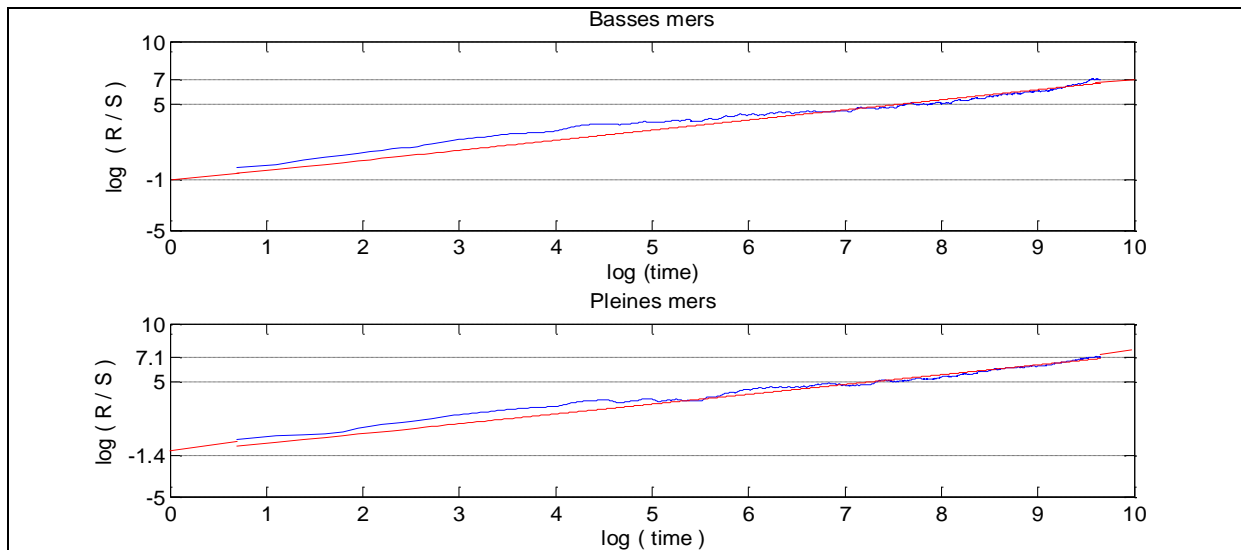


Figure 3 : Exposants de Hurst

Avant d'appliquer directement une modélisation FARIMA propre aux série à mémoire longue et qui est plus compliquée à mettre en œuvre, nous allons tester si une modélisation ARIMA ($p+1,d,q$) marche et suffit à prévoir les pleines et les basses mers dans notre zone d'étude.

3.4 Préparation des données

La **Figure 4** montre que nous avons dû :

- éliminer 1 744 données pour les basses mers pour blanchir notre série des points aberrants ;
- et éliminer 1 761 données pour les pleines mers pour blanchir notre chronique des points aberrants

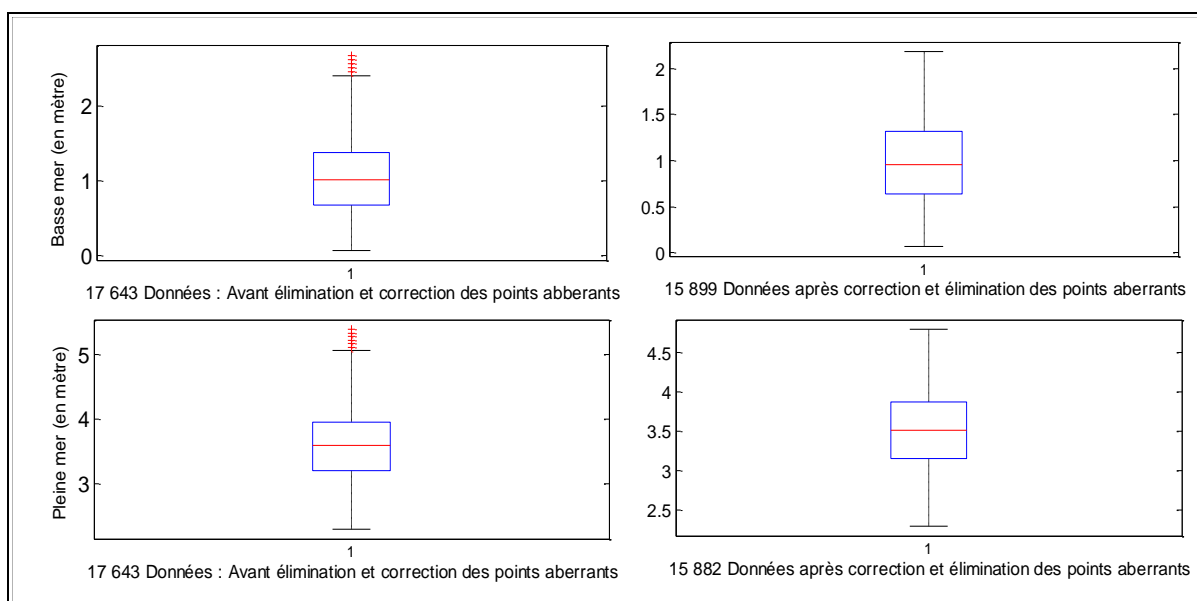


Figure 4 : Boîtes à moustache avant et après correction des points aberrants

3.5 Saisonnalité S

➤ Corrélogramme ou méthode de BLACKMAN-TUCKEY

La méthode du corrélogramme permet de déduire sur la **Figure 5** la même saisonnalité S qui est égale à 28.5219 Lags pour les pleines mers et pour les basses mers.

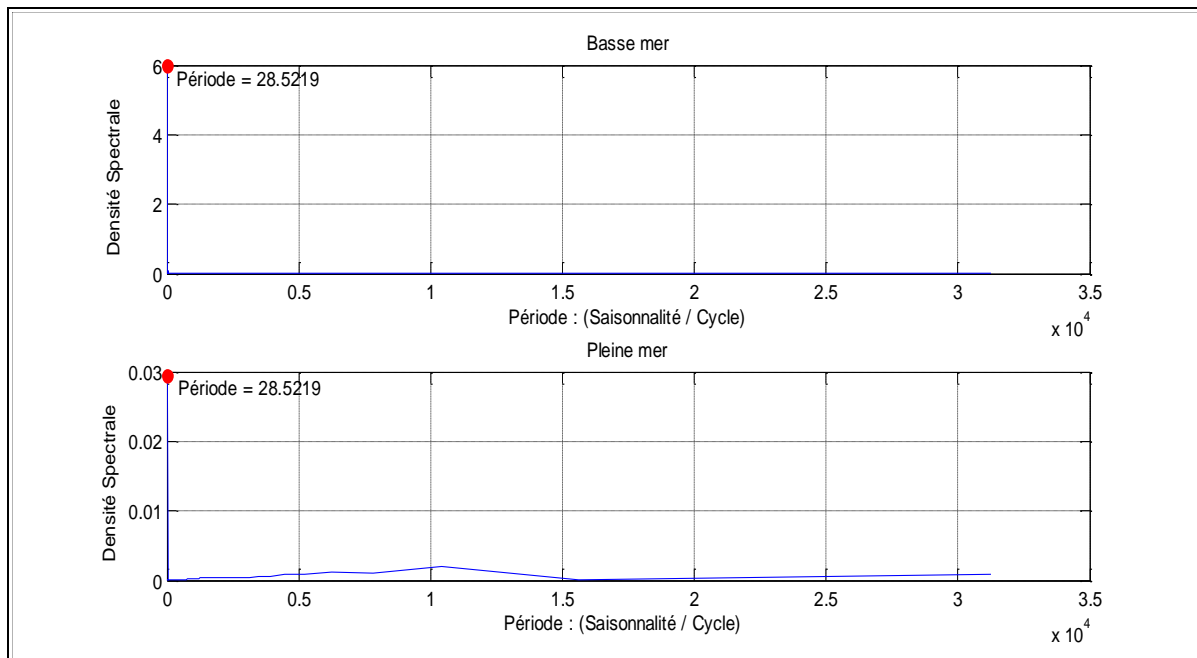


Figure 5 : Corrélogramme des séries stabilisées

➤ Périodogramme moyenné

Sur la **Figure 6 et 7**, les périodogrammes moyennés des pleines et basses mers donnent la même saisonnalité S comprise entre 28.4 et 28.6154 Lags

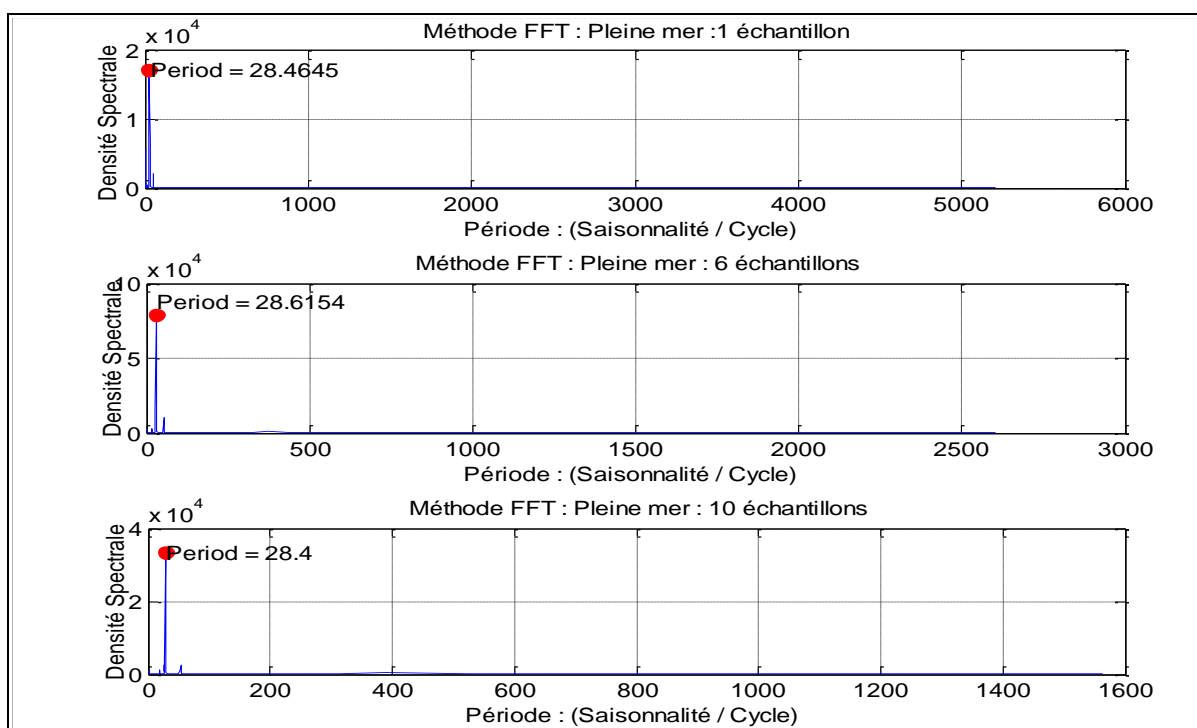


Figure 6 : Périodogramme moyenné : Pleines mers

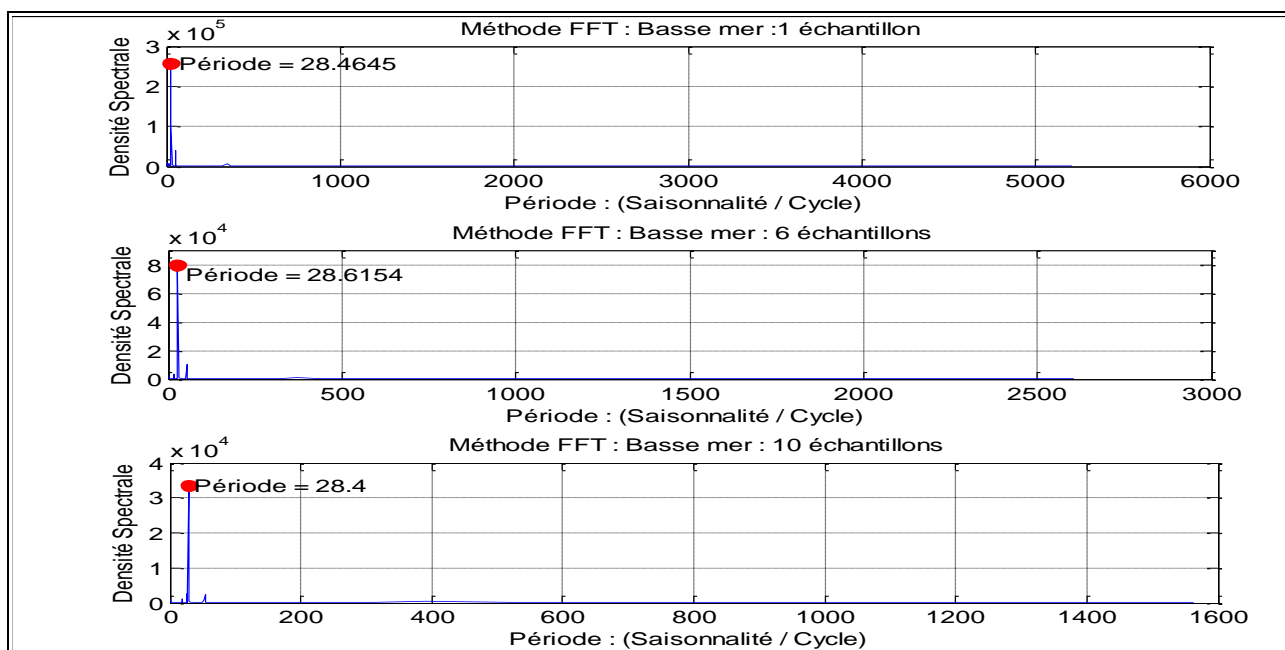


Figure 7 : Périodogramme moyenné : Basses mers

➤ Périodogramme par la Méthode de l'Entropie Maximale

Sur la **Figure 8**, nous trouvons par la méthode de l'entropie maximale la saisonnalité S égale à 28.36 Lags pour les basses mers et égale à 27,8 Lags pour les pleines mers.

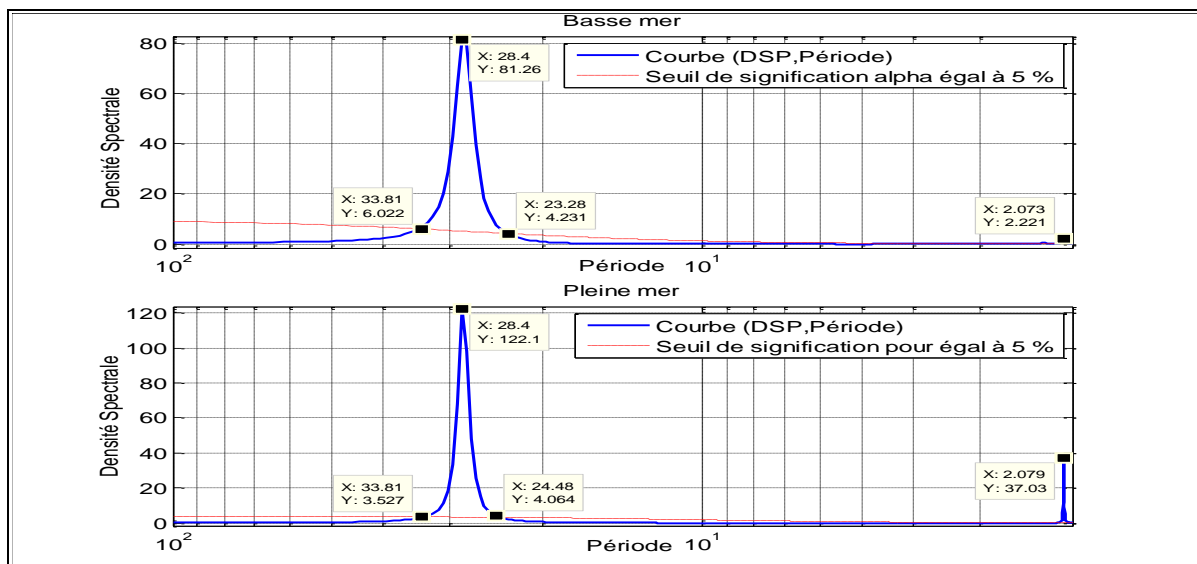


Figure 8 : Périodogramme par la Méthode de l'Entropie Maximale : Pleines mers et Basses mers

Sur les 10 valeurs que nous avons trouvées pour S, 9 résultats nous suggèrent de prendre comme meilleure estimation de la Saisonnalité S, la valeur entière 28 ou 29 Lags. Nous admettrons donc que l'estimation de la saisonnalité est soit 28 Lags ou 29 Lags.

3.6 Saisonnalité S en terme journalier

En passant à la moyenne de la différence en jour entre deux temps de prélèvement consécutifs, nous avons trouvé qu'1 Lag correspond à 0,5175313 jour pour les pleines mers et à 0,51753154 jour pour les basses mers. En portant ces valeurs aux 10 saisonnalités S en Lags que nous avons trouvées

pour les pleines et basses mers et en mettant en œuvre le test t de Student, nous avons pu dresser le **Tableau 1**.

Tableau 1 : Test d'égalité entre la saisonnalité obtenue par les méthodes traditionnelles et par nos méthodes

	Ancienne Méthode	Nos méthodes	Intervalle de confiance et paramètres	Décision
	Saisonnalité : Morte eau, vive eau (longue période)			
Pleine mer	S = 14,765 jours	$SN_1 = \begin{cases} S = 14,76097 \text{ jours} \\ S = 14,69789 \text{ jours} \\ S = 14,73127 \text{ jours} \\ S = 14,80936 \text{ jours} \end{cases}$	Moyenne $SN_1 = 14,74988$ Ecart type $SN_1 = 0,05461$ $t_{1-\frac{\alpha}{2}} = 3,182$ S et $SN_1 \in [14,6629, 14,8367]$	Il y a concordance parfaite entre S et les valeurs de SN_1 au seuil $\alpha = 0,05$
Basse mer	S = 14,765 jours	$SN_2 = \begin{cases} S = 14,76098 \text{ jours} \\ S = 14,69789 \text{ jours} \\ S = 14,73127 \text{ jours} \\ S = 14,80937 \text{ jours} \end{cases}$	Moyenne $SN_2 = 14,74988$ Ecart type $SN_2 = 0,05461$ $t_{1-\frac{\alpha}{2}} = 3,182$ S et $SN_2 \in [14,6629, 14,8367]$	Il y a concordance parfaite entre S et les valeurs de SN_2 au seuil $\alpha = 0,05$
	Saisonnalité : Diurne		Intervalle de confiance	Décision
Pleine mer	S = 1,034 jours	$SN_3 = \begin{cases} 1,06184 \text{ jours} \\ 1,07594 \text{ jours} \end{cases}$	Moyenne $SN_3 = 1,06889$ Ecart type $SN_3 = 0,00997$ $t_{1-\frac{\alpha}{2}} = 12,706$ S et $SN_3 \in [0,94219, 1,19560]$	Il y a concordance parfaite entre S et les valeurs de SN_3 au seuil $\alpha = 0,05$
Basse mer	S = 1,034 jours	$SN_4 = \begin{cases} 1,06443 \text{ jours} \\ 1,07284 \text{ jours} \end{cases}$	Moyenne $SN_4 = 1,06863$ Ecart type $SN_4 = 0,00594$ $t_{1-\frac{\alpha}{2}} = 12,706$ S et $SN_4 \in [0,99307, 1,14419]$	Il y a concordance parfaite entre S et les valeurs de SN_4 au seuil $\alpha = 0,05$

Nous avons pris la valeur de la première ligne du vecteur SN_3 et du vecteur SN_4 égale à la moyenne des saisonnalités significatives trouvées après l'application de la méthode de l'entropie maximale. Les résultats du **Tableau 1** montre qu'il existe une concordance parfaite entre les saisonnalités S longue période et diurne obtenues par nos méthodes et celles calculées par les méthodes traditionnelles, en particulier par la théorie statique de Newton. Nous pouvons affirmer également que seule la Méthode de l'Entropie Maximale (MEM) permet de détecter la saisonnalité diurne. Ce qui semble évident d'après le théorème de Nyquist-Shannon [27] qui stipule que pour des échantillons prélevées tous les S égale à 12 heures 25 minutes, nous ne pourrions détecter que les saisonnalités S de valeur supérieure ou égalé à 24 heures 50 minutes. Nous avons aussi remarqué qu'il suffit d'interpoler la valeur de la saisonnalité diurne pour obtenir la période semi-diurne.

3.7 Choix de la Saisonnalité S

Pour pouvoir mettre en œuvre le test sur les résidus et le test de Fisher, nous avons dû transformer nos données vectorielles sur les pleines mers et sur les basses mers en matrice :

- B à 558 lignes x 28 colonnes et en matrice ;
- A à 539 lignes x 29 colonnes.

➤ Test sur les Résidus

Les tests sur les résidus que nous pouvons observer sur la **Figure 9** recommandent aussi bien pour les pleines mers que pour les basses mers de choisir la saisonnalité S égale à 29.

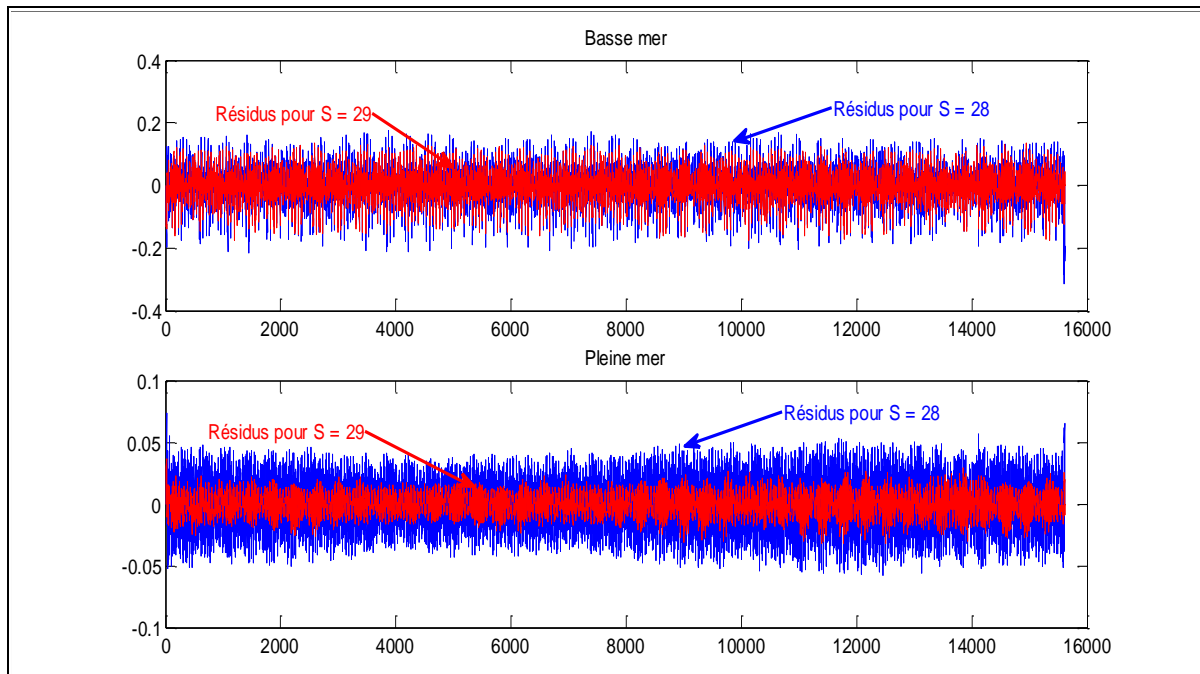


Figure 9 : Courbes des résidus pour S = 28 et S = 29

Cette méthode semble suffisante pour trancher sur le choix de la meilleure saisonnalité S à prendre pour nos prévisions.

➤ Analyse de la Variance à un Facteur

Les résultats du **Tableau 2** confirment la possibilité d'utilisation du test de Fisher pour étudier la stabilité de nos chroniques:

Tableau 2 : Indépendance, homoscedasticité et normalité

	Basse mer		Pleine mer	
	L'indépendance des données résultent de l'indépendance des prélèvements des niveaux de la mer			
Homoscedasticité acceptée	Bartlett	$\frac{B}{C} = 5,85 < \chi^2(28) = 41,34$	Bartlett	$\frac{B}{C} = 9,158 < \chi^2(28) = 41,34$
	Levene	$D = 0,37 < F(28,538) = 1,5$	Levene	D = 3,198 < F(28,538) = 1,5
Normalité	Rejet de la normalité	$1,96 \leq \min_{1 \leq i \leq 29} (JB_i) = 6$ $1,96 \leq \min_{1 \leq i \leq 29} S_i = 8,8$ $1,96 \leq \min_{1 \leq i \leq 29} K_i = 2,25$	Rejet de la normalité	$1,96 \leq \min_{1 \leq i \leq 29} (JB_i) = 6,84$ $ S_i _{1 \leq i \leq 29} \leq 1,96$ pour 13 « i » sur 29 $ K_i _{1 \leq i \leq 29} \leq 1,96$ pour 9 « i » sur 29
	Vraisemblance à la loi normale : oui	$0,0018 \leq \mathbf{ad} - 0,7979\sigma \leq 0,01$ $0,002 \leq \mathbf{pd} - 0,6745\sigma \leq 0,04$	Vraisemblance à la loi normale : oui	$0,00001 \leq \mathbf{ad} - 0,7979\sigma \leq 0,0051$ $0,00008 \leq \mathbf{pd} - 0,6745\sigma \leq 0,019$

Nous apercevons sur la **Figure 10**, que les statistiques de Fisher sont significatives :

- Pour les basses mers :
 - ⇒ Entre la ligne 1 et la ligne 142 pour $S = 28$ et
 - ⇒ Entre la ligne 1 et la ligne 161 pour $S = 29$
- Pour les pleines mers :
 - ⇒ Entre la ligne 1 et la ligne 140 pour $S = 28$ et
 - ⇒ Entre la ligne 1 et la ligne 156 pour $S = 29$

Ce qui nous recommande toujours de choisir S égale à 29.

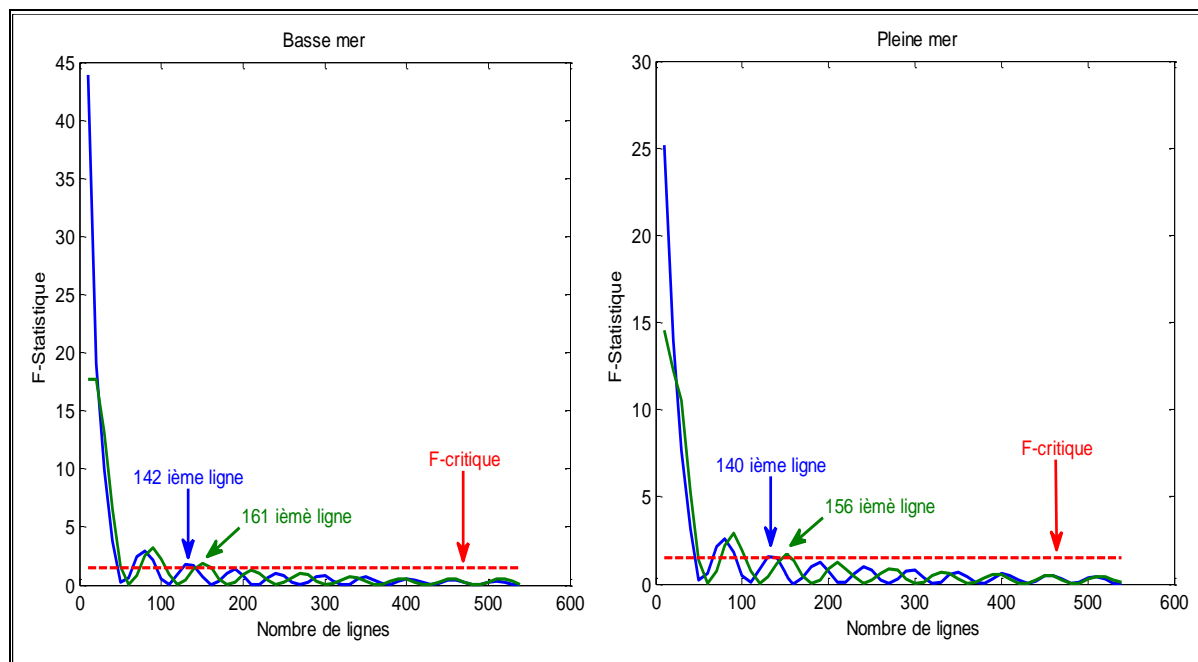


Figure 10 : Test de Stabilité

En répétant la même opération à partir de la ligne 162 pour les basses mers et de la ligne 156 pour les pleines mers et ainsi de suite, nous arrivons à établir le **Tableau 3** qui montre l'existence d'une stabilité locale au niveau de nos chroniques.

Tableau 3 : Analyse de la stabilité pour $S = 29$

Basses mers			Pleines mers		
Nombre de Lignes	F-Statistique	F-Critique	F-Statistique	F-Critique	Nombre de Lignes
Lignes 1 à 161, 4669 données	1.503	1.478	1,490	1,478	Lignes 1 à 156, 4524 données
Lignes 162 à 323, 6351 données	1.488	1.478	1,533	1,478	Lignes 157 à 314, 4582 données
Lignes 381 à 539, 4611 données	1.527	1.478	1,534	1,478	Lignes 315 à 468, 4466 données
			1,031	1,482	Lignes 469 à 539, 2059 données

Finalement, nous constatons que le test sur les résidus et le test de Fisher arrivent à la même conclusion : il faut prendre la saisonnalité $S = 29$ Lags pour effectuer la prévision de nos 03 modèles sur les basses mers et de nos 04 modèles sur les pleines mers.

3.8 Analyse de la tendance

La mise en œuvre du test de corrélation des rangs de Kendall a permis de dresser le **Tableau 4** :

Tableau 4 : Test de corrélation de rang de Kendall

	τ_b	$E(\tau_b)$	$Var(\tau_b)$	Z_1	$U_{0,1,0,05}$
Pleine mer	-9,98 %	0	$2,8440.10^{-05}$	-18,7082	1,96
Basse mer	11,80 %	0	$2,8440.10^{-05}$	22,1207	1,96

Les pleines mers ont donc une faible tendance décroissante en fonction du temps de prélèvement et les basses mers faible tendance croissante en fonction du temps de prélèvement.

3.9 Schéma de prévision

➤ Méthode du profil

D'après la **Figure 11**, cette méthode ne nous permet pas de trancher sur le choix du schéma puisque plusieurs courbes s'entrecroisent et d'autres sont plus ou moins confondues.

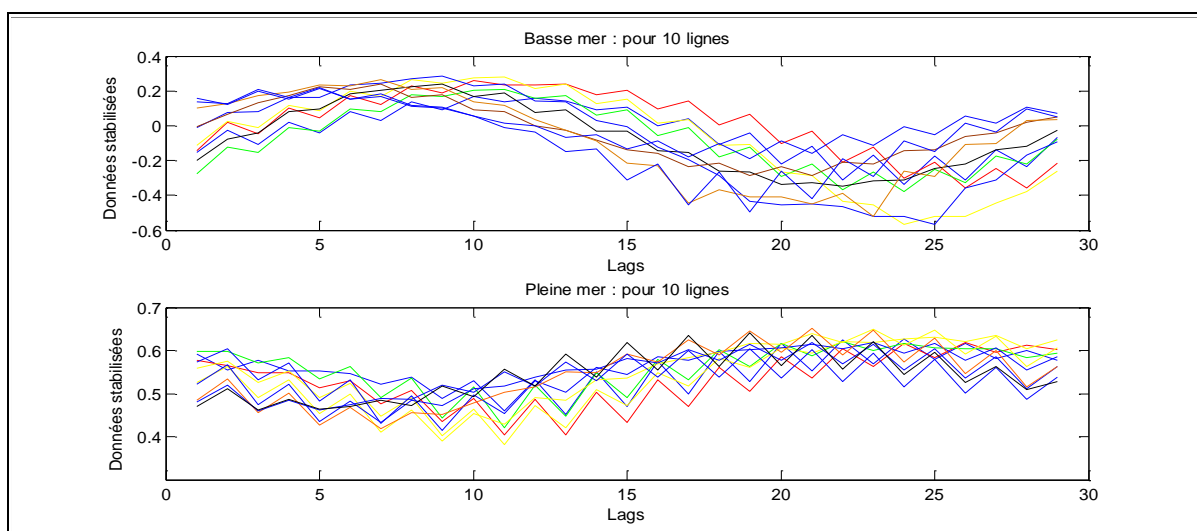


Figure 11 : Méthode du profil

➤ Méthode de la bande

Nous pouvons observer sur la **Figure 12** que la courbe des minima et celle des maxima ne sont pas parallèles pour les deux chroniques. Cela nous conduit à opter pour un schéma multiplicatif.

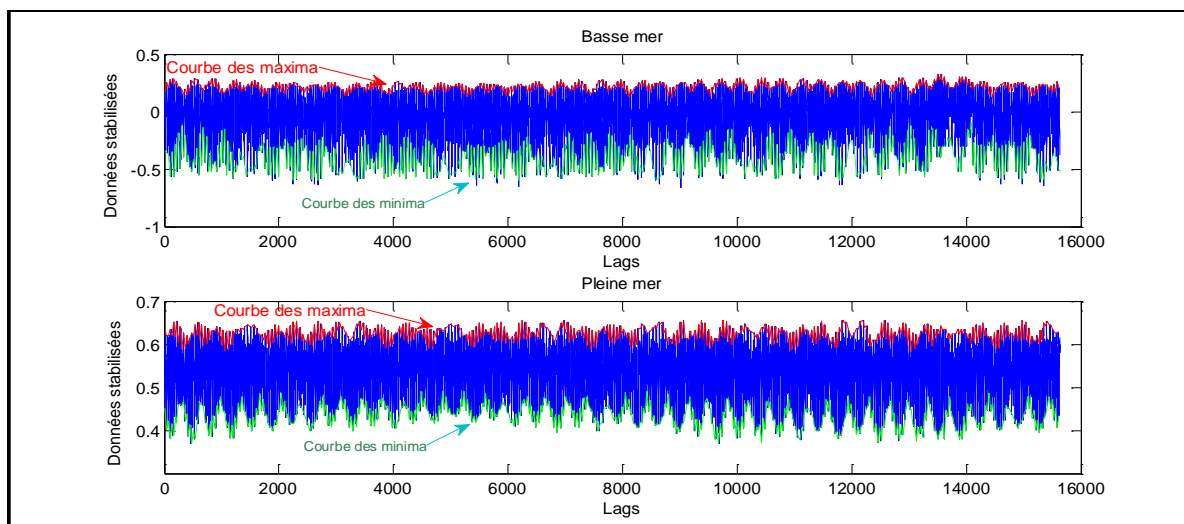


Figure 12 : Méthode de la bande

➤ *Méthode analytique*

Les Résultats du **Tableau 5** nous recommandent aussi d'appliquer le schéma multiplicatif puisque - 28,11 (t-statistique basse mer) et 2,30 (t-statistique pleine mer) sont supérieurs à 1,96 en valeur absolue.

Tableau 5 : Régression entre la moyenne et l'écart type

	Basse mer			Pleine mer		
	Coefficients	t -Statistique	Valeur critique	Coefficients	t -Statistique	Valeur critique
g	0,15126582	72,335243	1,96	0,00400041	0,17726521	1,96
μ	-0,84464676	-28,1101052	1,96	0,09609609	2,307198454	1,96

Finalement, sur les trois méthodes testées, deux nous suggèrent de choisir le schéma multiplicatif que nous avons adopté.

3.10 Méthode Census X11

La transformation logarithmique a permis :

- Pour les basses mers de réduire l'écart type de 0,4125 à 0,2160 et
- Pour les pleines mers de réduire l'écart type de 0,4494 à 0,0566

Nous avons appliqué la méthode Census X11 aux matrices A = 539 lignes *29 colonnes et nous avons trouvé la tendance en appliquant la moyenne mobile de Henderson à 13 termes puisque le rapport $\frac{I}{C}$ reste toujours supérieur à 1 dans les deux cas (1,48 pour les basses mers et 2,24 à pour les pleines mers).

3.11 Méthodologie de Box et Jenkins

- Identification des ordres d , q et $p+1$ du modèle $ARIMA(p+1,d,q)$

Nous avons choisi un nombre de retard $p_{\max}=14$. En appliquant au vecteur $scvs_{t,2}$, la stratégie de Dickey Fuller Augmenté, nous avons trouvé d'après les résultats du **Tableau 6** :

- ⇒ Pour la basse mer « d » égal à 2 et le modèle ne possède ni tendance, ni constante
- ⇒ Pour la pleine mer « d » égal à 2 pour les deux premiers modèles et « d » égal à 1 pour les deux derniers modèles. Les modèles sont toujours dénuer de tendance et de constante

Tableau 6 : Test DFA pour un nombre de retard maximal $p_{max} = 14$ [32]

Basse mer , nombre de retard p = 14									Pleine mer , nombre de retard p = 14									Valeur critique		
Statistique t pour d=0			Statistique t pour d=1			Statistique t pour d=2			Statistique t pour d=0			Statistique t pour d=1			Statistique t pour d=2			φ	b et a	
φ	b	a	φ	b	a	φ	b	a	φ	b	a	φ	b	a	φ	b	a			
1	238	-2 ,10	-6 ,13	13,4	0,06	-0,03	-28,2	0,093	-0,06	133,5	3,36	16,54	1,1	-0,02	-0,006	-24,6	-0,04	0,06	-3,41	1,96
2	266	3,70	-7,23	13,4	0,04	-0,05	-33,0	0,014	-0,02	164,7	-2,74	16,99	4,8	-0,10	0,106	-23,9	0,05	-0,05	-3,41	1,96
3	195	-0,47	-0,40	4,2	-0.04	0,04	-29,8	-0,027	0,03	106,7	-0,47	15,88	-3,6	-0,02	0,022	-24,9	0,07	-0,07	-3,41	1,96
4										64,7	2,56	8,96	-3,8	0,07	-0,069	-17,8	-0,02	0,02	-3,41	1,96

1, 2, 3, 4 représentent le numéro des modèles dans chaque chronique.

3.12 Modélisation

Pour p_{max} égal 14, les résidus des modèles de Dickey Fuller Augmenté sont encore corrélés malgré la confirmation de la stationnarité par le test sauf pour les modèles 3 et 4 pleines mers. Pour blanchir les résidus nous avons dû augmenter p_{max} de trois unités pour les basses mers et de deux unités pour les modèles 1 et 2 sur les pleines mers avant de procéder à la comparaison du Critère de Schwarz du modèle de Dickey Fuller Augmenté en diminuant progressivement p_{max} jusqu'à zéro. Ce qui nous a permis de dresser le **Tableau 7** et de déduire l'ordre optimal du modèle de Dickey Fuller Augmenté que nous avons coloré en bleu.

Tableau 7 : Comparaison des Critères de Schwarz

Nombre de retard p	Pleine mers				Basse mer		
	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 1	Modèle 2	Modèle 3
17					-16 607,75	-22 256,25	-15 141,27
$p_{optimal} = 16$	-20 432,19	-21 711,17			-16 610,95	-22 258,79	-15 142,25
$p_{optimal} = 15$	-20 435,43	-21 714,38			-16 604,43	-22 249,28	-15 136,79
14	-20 423,08	-21 678,83	-19 440,52	-8 753,77	-16 600,87	-22 240,28	-15 126,71
13	-20 420,65	-21 673,17	-19 443,90	-8 756,88	-16 589,40	-22 225,75	-15 114,69
12	-20 405,63	-21 670,71	-19 441,45	-8 758,22	-16 549,82	-22 169,71	-15 066,35
$p_{optimal} = 11$	-20 404,36	-21 658,82	-19 444,97	-8 761,37	-16 513,18	-22 112,30	-14 996,08
10	-20 388,57	-21 642,55	-19 435,46	-8 755,99	-16 464,27	-22 058,05	-14 956,59
9	-20 388,62	-21 640,43	-19 439,11	-8 759,27	-16 428,94	-21 990,95	-14 927,39
8	-20 362,13	-21 621,24	-19 433,37	-8 760,45	-16 390,92	-21 954,77	-14 925,19
7	-20 363,75	-21 621,63	-19 436,62	-8 763,09	-16 378,58	-21 933,61	-14 912,30
6	-20 320,15	-21 590,05	-19 437,71	-8 762,29	-16 374,73	-21 933,94	-14 908,39
$p_{optimal} = 5$	-20 282,91	-21 585,44	-19 441,36	-8 765,59	-16 375,89	-21 927,64	-14 862,33
4	-20 244,37	-21 500,67	-19 422,66	-8 734,53	-16 341,63	-21 882,97	-14 827,82
3	-20 244,84	-21 485,66	-19 406,68	-8 729,20	-16 116,79	-21 569,25	-14 797,98
2	-20 229,12	-21 484,83	-19 245,32	-8 721,02	-16 092,88	-21 548,27	-14 793,98
1	-20 219,49	-21 331,22	-19 132,94	-8 679,18	-15 481,52	-20 800,01	-14 574,88
0	-19 962,01	-21 019,41	-18 380,12	-8 443,81	-15 436,34	-20 732,82	-14 532,27

Nous avons mis en œuvre le test de Ljung Box et la méthode contrôle ex-post de l'absence d'autocorrélation des innovations » [32] pour obtenir les valeurs de $q_{optimal}$ que nous avons consignés dans le **Tableau 8 et 9**.

Tableau 8 : Comparaison des Critères d'Akaike et du Test de Ljung-box pour les pleines mers

Ordre q	Pleine mer											
	Modèle 1			Modèle 2			Modèle 3			Modèle 4		
	Pr	k	AIC	Pr	k	AIC	Pr	k	AIC	Pr	k	AIC
12							0,29	3	-19493,10			
11				0,09	5	-21798,22	0,08	15	-19448,15			
10	0,05	4	-20596,66	0,09	15	-21727,05	0,11	14	-19449,93	0,07	4	-8771
9	0,07	11	-20532,64	0,10	15	-21725,79	0,58	13	-19447,09	0,06	12	-8759
8	0,07	11	-20534,64	0,11	15	-21727,67	0,58	13	-19448,95	0,05	12	-8760
7	0,43	9	-20530,24	0,07	15	-21727,88	0,40	13	-19449,70	0,37	11	-8760
6	0,43	9	-20532,11	0,06	15	-21729,54	0,40	13	-19451,65	0,3	11	-8761
5	0,21	9	-20532,16	0,20	14	-21730,42	0,36	13	-19453,30	0,27	11	-8763
4	0,21	9	-20534,03	0,20	14	-21732,19	0,36	13	-19455,30	0,26	11	-8765
3	0,18	9	-20535,54	0,19	14	-21733,52	0,36	13	-19457,16	0,24	11	-8766
2	0,18	9	-20537,40	0,20	14	-21735,41	0,36	13	-19459,16	0,25	11	-8768
1	0,17	9	-20539,20	0,20	14	-21737,29	0,36	13	-19461,12	0,25	11	-8770
0	0,18	9	-20541,20	0,21	14	-21739,29	0,36	13	-19463,12	0,24	11	-8772

Tableau 9 : Comparaison des Critères d'Akaike et du Test de Ljung-box pour les basses mers

Ordre q	Basse mer								
	Modèle 1			Modèle 2			Modèle 3		
	Pr	k	AIC	Pr	k	AIC	Pr	k	AIC
13	0,12	6	-16483,09	0,06	6	-22361,13			
12	0,07	17	-16407,79	0,46	16	-22266,61			
11	0,07	17	-16409,71	0,64	16	-22267,35			
10	0,07	17	-16411,71	0,67	16	-22269,35	0,11	3	-15458,93
9	0,96	16	-16413,56	0,68	16	-22271,35	0,12	16	-15405,67
8	0,96	16	-16415,56	0,69	16	-22273,26	0,15	16	-15406,51
7	0,96	16	-16417,56	0,69	16	-22275,08	0,14	16	-15407,35
6	0,96	16	-16419,48	0,69	16	-22276,90	0,14	16	-15408,15
5	0,96	16	-16421,41	0,69	16	-22278,54	0,15	16	-15408,95
4	0,96	16	-16423,33	0,69	16	-22280,17	0,15	16	-15409,92
3	0,96	16	-16425,10	0,67	16	-22281,90	0,16	16	-15411,30
2	0,05	17	-16427,10	0,67	16	-22283,81	0,17	16	-15413,13
1	0,05	17	-16429,02	0,66	16	-22285,63	0,19	16	-15415,01
0	0,05	17	-16431,02	0,67	16	-22287,63	0,19	16	-15417,01

Pr représente la p-value à l'ordre k au test de Ljung-box, AIC le Critère d'Information d'Akaike et les parties colorées en marrons et en bleue définissent respectivement q_{\max} et q_{optimal} pour chaque modèle. Les paramètres $d_{\text{adéquat}}$, p_{optimal} et q_{optimal} que nous avons obtenus à partir des **Tableaux 6 à 9** nous permettent de définir maintenant les modèles Arima ($p_{\text{optimal}} + 1, d_{\text{adéquat}}, q_{\text{optimal}}$) pour nos deux chroniques.

➤ En particulier pour la pleine mer :

Le modèle 1 est un ARIMA(16,2,0) sans constante ni tendance:

AR: {-1.32244 -0.460963 -0.22449 -0.266043 -0.339991 -0.520838 -0.496018 -0.37686 -0.33032 -0.311032 -0.251778 -0.237872 -0.219805 -0.212848 -0.185309 -0.0985826} at Lags [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]

Variance: $2,83263 \cdot 10^{-05}$, la **p-value** à l'ordre k = 9 au test de Ljung box est égale à 0,1876

Le modèle 2 est un ARIMA(16,2,0) sans constante :

AR: {-1.30148 -0.207144 0.071424 -0.307404 -0.469021 -0.439674 -0.36106 -0.325447 -0.283225 -0.275951 -0.26728 -0.253469 -0.157726 -0.169205 -0.249312 -0.140137} at Lags [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]
Variance: $1,77285 \cdot 10^{-05}$, la **p-value** à l'ordre $k = 14$ au test de Ljung box est égale à 0.2068

Le modèle 3 est un ARIMA(12,1,0) sans constant :

AR: {-0.445972 0.801033 0.199155 -0.248153 -0.145722 -0.102132 -0.0501472 -0.0709159} at Lags [1 2 3 4 5 6 9 12]
Variance: $4,33812 \cdot 10^{-05}$, la **p-value** à l'ordre $k = 13$ au test de Ljung box est égale à 0.0867

Le modèle 4 est un ARIMA(6,1,0) sans constante :

AR: {-0.509639 0.610855 0.280004 -0.191357 -0.210269} at Lags [1 2 3 5 6]
Variance: $4,99936 \cdot 10^{-05}$, la **p-value** à l'ordre $k = 11$ au test de Ljung box est égale à 0.1593
Résidus : normalité acceptée car $|S| = 0,2959 \leq 1,96$ et $ad -0,7979\sigma = 0,0048$; $pd -0,6745\sigma = -0,0014$

➤ et pour la basse mer,

Le modèle 1 est un ARIMA(17,2,0) sans constante

AR: {-1.14728 0.235627 0.391886 -0.46953 -0.601879 -0.336974 -0.216538 -0.224985 -0.268324 -0.28613 -0.280508 -0.294943 -0.281346 -0.197207 -0.142584 -0.146011 -0.0711499} at Lags [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17]
Variance: $2.60092 \cdot 10^{-05}$, la **p-value** à l'ordre $k = 16$ au test de Ljung box est égale à 0,9520

Le modèle 2 est un ARIMA(17,2,0) sans constante

AR: {-1.15684 0.195224 0.371394 -0.448749 -0.60393 -0.371812 -0.238813 -0.216011 -0.263704 -0.296908 -0.291201 -0.297713 -0.293599 -0.196896 -0.145571 -0.148846 -0.0697773} at Lags [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17]
Variance: 0.000305774, la **p-value** à l'ordre $k = 16$ au test de Ljung box est égale à 0,6461

Le modèle 3 est un ARIMA(17,2,0) sans constante

AR: {-1.15933 -0.102458 0.0580946 -0.30043 -0.410597 -0.473176 -0.424736 -0.336947 -0.278822 -0.261219 -0.321135 -0.383823 -0.381879 -0.262808 -0.191327 -0.162117 -0.0678325} at Lags [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17]
Variance: 0,000490815, la **p-value** à l'ordre $k = 16$ au test de Ljung box est égale à 0,2154

➤ Analyse des résidus standardisés

Le test ARCH indique la présence d'hétéroscédasticité au niveau des résidus car après execution de la fonction « archtest » sous Matlab nous avons obtenu la valeur 1 pour tous les modèles. Nous avons donc modélisé la variance en utilisant soit la fonction garch(0,1) (respectivement (egarch(0,1))), soit la fonction "garchset" sous Matlab. Les résultats du **Tableau 10** montrent que tous les résidus standardisés η_t restent non corrélés à l'ordre 1 et convergent vers la distribution normale centrée réduite. En particulier, pour le modèle 3 des données sur les pleines mers, η_t reste non corrélé jusqu'à l'ordre 9.

Tableau 10 : Paramètres des résidus standardisés

	N°	ARCH (1) OU EGARCH (0,1)	p-value	Test sur l'Intervalle de confiance	
Modèles Pleine mer	1	$K = 1,6869 \cdot 10^{-05}$, $\alpha_1 = 0,4289$	$p = 0,4256$ et $k=1$	$0 \in [-0,0185, 0,0398]$ Moyenne : 0,0106	$1 \in [0,9796, 1,0208]$ Variance : 1,0004
	2	$K = 1,1685 \cdot 10^{-05}$, $\alpha_1 = 0,3345$	$p = 0,2103$ et $k=1$	$0 \in [-0,0161, 0,0418]$ Moyenne : 0,0128	$1 \in [0,9595, 1,0415]$ Variance : 1,0005
	3	$K = -10,0638$ $\alpha_1 = 0,3478$, Leverage = 0,2586	$p = 0,0750$ et $k=9$	$0 \in [-0,0207, 0,0378]$ Moyenne : 0,0086	$1 \in [0,9526, 1,0351]$ Variance : 0,9938
	4	$K = 3,4297 \cdot 10^{-05}$, $\alpha_1 = 0,3603$	$p = 0,2059$ et $k=1$	$0 \in [-0,0626, 0,0238]$ Moyenne : -0,0194	$1 \in [0,9391, 1,0613]$ Variance : 1,0002
Modèles Basse mer	1	$K = 9,4464 \cdot 10^{-05}$, $\alpha_1 = 0,6566$	$p = 0,9948$ et $k=1$	$0 \in [-0,0094, 0,0480]$ Moyenne : 0,0193	$1 \in [0,9593, 1,0404]$ Variance : 0,9999
	2	$K = -8,3557$ $\alpha_1 = 0,5480$, Leverage = 0,1134	$p = 0,5970$ et $k=1$	$0 \in [-0,0385, 0,0106]$ Moyenne : -0,0140	$1 \in [0,9613, 1,0306]$ Variance : 0,9960
	3	$K = 6,8272 \cdot 10^{-04}$ $\alpha_1 = 0,7038$	$p = 0,2518$ et $k=1$	$0 \in [-0,0322, 0,0256]$ Moyenne : -0,0033	$1 \in [0,9594, 1,0411]$ Variance : 1,0002

➤ *Vérification de la qualité de la prédiction et de la prévision*

Nous avons comparé les erreurs issues de nos modèles avec les résidus provenant du logiciel de prédiction du Service Hydrographique Océanographique de la Marine française qui utilise un modèle harmonique [1]. Pour la comparaison, nous avons pris les coefficients saisonniers de l'année 2011 pour effectuer la prévision de 251 points sur les pleines mers et de 268 points sur les basses mers de l'année 2012 ainsi que la prédiction des 58 avant derniers points de chaque chronique. Les calculs de l'indicateur de THEIL et de l'EMAP aboutissaient aux résultats suivants.

Tableau 11 : Comparaison des erreurs et de l'Indicateur de THEIL

Indicateurs	Pleine mer		Basse mer		Prévision avec terme moyenne mobile
	Prédiction	Prévision	Prédiction	Prévision	
THEIL	0,651	2,28	0,22	0,98	1,2
EMAP (modèle)	0,87 %	2,76	2,5 %	16,28 %	18,50 %
EMAP (méthode harmonique)	1,61 %	1,31	17,53 %	18,68 %	18,98 %

En se référant aux indicateurs de THEIL, nous pouvons affirmer que nos modèles offrent de meilleure précision par rapport à la méthode harmonique sauf pour la prévision des pleines mers. Cependant, le modèle reste excellent car l'EMAP est inférieure à 5%. Nous pouvons également dire que tous nos modèles sont excellents puisque leurs EMAP sont inférieurs à 5% sauf pour la prévision des basses mers qui entre dans le classement des modèles à moyenne précision. Malgré cela, ce modèle reste meilleur par rapport à la méthode harmonique car son EMAP est inférieur à celle de la méthode harmonique. Pour confirmer si nos modèles s'ajustent bien aux niveaux des pleines mers et basses mers prélevés à Hell-ville, nous avons passé à la simulation qui montre sur la **Figure 13** qu'il est préférable d'ajouter un terme moyenne mobile au lieu d'utiliser un modèle purement autorégressif pour prévoir les basses mers. Nous allons donc remplacer **le modèle 3** pour la prévision des basses mers précédent par un ARIMA(17,2,8) sans constante défini par :

AR: {-1.20115 -0.155344 0.33759 0.398387 -0.549408 -0.786393 -0.135724 0.384957 -0.273516 -0.253587 -0.27176 -0.143191 0.0691767 0.0410379} at Lags [1 2 3 4 6 7 8 9 11 12 13 14 16 17]
MA: {-0.0712812 -0.33589 -0.446318 0.330744 -0.457627} at Lags [2 3 4 7 8]
Variance: 0,000454712, la **p-value** à l'ordre $k = 19$ au test de Ljung box est égale à 0,5288
K= 1,9294.10⁻⁰⁴; $\alpha_1=0,6798$; **p** = 0,0891 et **k** = 1; **Moyenne** : -0,0109; **Variance** : 1,0001
Résidus : normalité acceptée avec ad -0,7979 σ = -0,0065; pd- 0,6745 σ = -0,0039

Nos modèles et la méthode harmonique ont détecté également un changement structurel anormal (différence d'environ 1 mètre) sur les prélèvements à partir du mois de Juillet 2012. Notre crainte d'utiliser les données à partir de cette date est donc bien fondée.

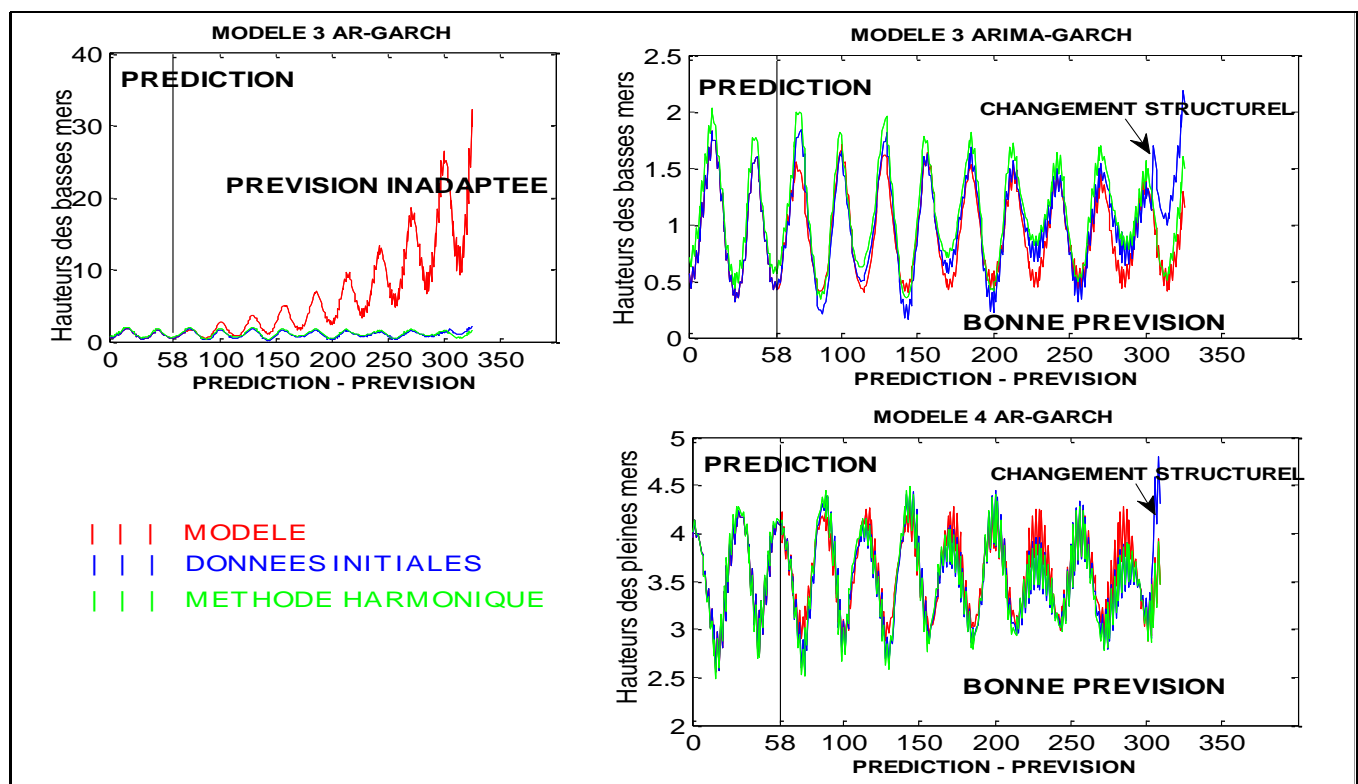


Figure 73 : Prédiction et prédiction

4 Conclusion

Nous avons construit 7 modèles stationnaire par différenciation (modèles DS) dont 4 pour la pleine mer et 3 pour la basse mer. Malgré la stabilité locale, la variation saisonnière n'est pas rigide mais souple. Nous avons donc utilisé les coefficients saisonniers de l'année N pour effectuer la prévision de l'année N+1 qui semblent fournir de bons résultats. Les modèles que nous avons obtenus au départ sont purement autorégressifs, mais après simulation, nous avons constaté qu'il faudrait inclure des termes moyennes mobiles pour obtenir une bonne prévision des basses mers et que le centre de Hellville devrait revoir l'anormalité de leurs données à partir de Juillet 2012. En particulier, nous avons trouvé :

- Pour le modèle 1 basse mer un Arima (17,2,0) correspondant aux données 1 à 4669
- Pour le modèle 2 basse mer un Arima (17,2,0) correspondant aux données 4670 à 11020
- Pour le modèle 3 basse mer un Arima (17,2,8) correspondant aux données 11021 à 15631
- Pour le modèle 1 pleine mer un Arima (16,2,0) correspondant aux données 1 à 4524
- Pour le modèle 2 pleine mer un Arima (16,2,0) correspondant aux données 4525 à 9106

Pour le modèle 3 pleine mer un Arima (12,1,0) correspondant aux données 9107 à 13572
 Pour le modèle 4 pleine mer un Arima (6,1,0) correspondant aux données 13573 à 15631

Nous avons montré que ces modèles possèdent des résidus normalisés non corrélés à l'ordre 1 ou plus et que leurs variances résiduelles peuvent être modélisées par ARCH(1) ou EGARCH(0,1). Le test de Fisher nous informe sur la présence d'un cycle à l'intérieur duquel les saisonnalités restent stables. Pour la pleine mer, ce cycle varie entre 4466 (environ 6 ans et 4 mois) et 4582 Lags (environ 6 ans et demi) et pour la basse mer il est moins précis et varie entre 4669 (environ 6 ans et 7 mois) et 6351 Lags (9 ans). En réalité, nous avons 5 chance sur 6 que le cycle des pleines et basses se situe entre 6,25 et 6,75 ans. Nous avons aussi remarqué que la tendance est non linéaire et que nos méthodes aboutissaient aux mêmes résultats que les saisonnalités de 14,765 jours, de 1,034 jours et de 0,517 jours obtenues par la théorie statique de Newton. Nous avons également constaté que la meilleure approche pour choisir le bon modèle consiste à minimiser le Critère d'Information d'Akaike en variant le q_{\max} dans le modèle ARIMA($p_{\text{optimal}} + 1, d_{\text{adéqua}}, q_{\max}$) et non le q_{optimal} dans le modèle ARIMA($p_{\text{optimal}} + 1, d_{\text{adéqua}}, q_{\text{optimal}}$) obtenu à partir de la stratégie de Dickey Fuller Augmenté. Finalement, nous avons trouvé des modèles AR-GARCH et des modèles ARIMA-GARCH pour prévoir les pleines mers et les basses mers à Hell-ville et qui nous semblent satisfaisants. Il nous paraît donc inutile de procéder à une modélisation FARIMA.

Références bibliographiques

- [1] Annuaire des marées, Ports d'outre-mer Tome 2, Service Hydrographique et Océanographique de la Marine, (2015) 127-150
- [2] **Simon BERNARD**, La marée : la marée océanique côtière, référence 942-MOG, Collection « Synthèse », Institut océanographique, Fondation Albert I, prince de monaco, (2007) 69-236
- [3] **Maher ABDALLAH**, Thèse de doctorat ; sur les méthodes de discrétisation numérique de problèmes hyperboliques non linéaires appliquées aux équations de Barré de Saint Venant pour la modélisation hydraulique en réseau d'assainissement, Ecole Doctorale science pour l'ingénieur, Université Louis Pasteur, Strasbourg I, soutenue le 03 Mai 2005
- [4] **Dominique LADIRAY, Benoît QUENNEVILLE**, Désaisonnaliser avec la méthode X-11, (2000)
- [5] **L. de MONTERA, C.MALLET et L. BARTHES**, Prédiction multi-step de la volatilité : le modèle ARIMA-GARCH appliquée aux séries temporelles d'affaiblissement par la pluie sur les liaisons Terre-Satellite, (2008)
- [6] **Hélène MISULTANE**, Econométrie des séries temporelles, (2002) consulté le 31 mars 2017 à l'adresse <http://halshs.archives-ouvertes.fr/cel-01261174>
- [7] **Jairo CULGIARI**, Séries chronologiques saisonnières et prévision, L1 IDEA-S 2014-2015, Université Lumière Lyon 2, version compilée le 05 février 2015
- [8] **Henri IMMEDIATO**, Cours de statistique – MASS UE11- Chapitre 2- NTE Lyon 1, Consulté le 16 Novembre 2016 à l'adresse <http://nte-serveur.univ-lyon1.fr/nte/immediato/math2002/mass11/cours/chapitr2b.htm>
- [9] https://en.wikipedia.org/wiki/Hurst_exponent
- [10] **Florence NICOLAU**, Etudes statistiques 2 : Traitement des valeurs manquantes et des valeurs aberrantes, (2005-2006)
- [11] **Olivier BESSON** ; Analyse spectrale paramétrique, 3ème année ENSICA
- [12] Matlab 2013a, Matlab fonction
- [13] **P. GRANJON, J.L LACOUME**, Introduction à l'analyse spectrale

- [14] **W.H. PRESS , S.A. TENKOLSKY, W.T. VETTERLING , B. P. Flannery.** Numerical Recipes in C . s.l., 2 (1992)
- [15] **Vincent GOUDARD**, Percussion aérienne : amélioration de la détection des coups dans le jeu d'un percussionniste. Université de la Méditerranée, Aix-Marseille II, (2004).
- [16] <http://www.labri.fr/perso/mraspaud/publications/dafx03.pdf>
- [17] <http://www.dcsc.tudelft.nl/Research/Publication Files/publication-5756.pdf>
- [18] **N. KAYSER , J. ROBST, D.C. YVES TARD** , *Propagation des ondes de sécheresse et d'humidité à travers le monde*. C.R.Acad.Sci.Paris. Série II Hydrologie/hydology. (1990) 757-763
- [19] **J.P.BURG**, *Maximum entropy spectral analysis*. Oklahoma City : in Proc. of the 37th meeting of the society of exploration geophysicists, October (1967)
- [20] <http://www.enee.umd.edu/~afshin/adsp3/report.pdf>
- [21] **Grais BERNARD**, Inspecteur général de l'INSEE, Professeur à l'INTEC, Expert auprès de l'Office statistique des Communautés européennes, Techniques statistiques 2 ; *Méthodes statistiques, troisième édition, nouvelle présentation, DUNOD, (1992)*
- [22] **Fanny MEYER, Morgane CADRAN, Margaux GAILLARD**, Analyse de la variance, M2 statistiques et économétrie.
- [23] **Ricco RAKOTOMALALA**, Comparaison de populations, Tests paramétriques, Version 1,2, Université Lumière Lyon 2, (2013)
- [24] Chapitre 4, Estimation, Tests de validation et prévision des processus ARMA
- [25] **A. VESSEREAU** , Inspecteur Général des Manufactures de l'Etat, Méthodes statistiques en biologie et en agronomie,(1998) 55-158
- [26] **A. JEBRANE**, Test d'égalité des variances, Université de Bourgogne, Master 1 psychologie, 28 septembre 2016
- [27] **Nicolas THOME** ; Bases du traitement des images : Transformée de Fourier Avancée – Numériasion, 4 octobre 2016
- [28] **M.G KENDALL**, *Rank Correlation Methods*. Griffin, London,(1975)
- [29] **Agnès LAGNOUX**, Séries Chronologiques, ISMAG, MASTER1- MI00141X, Université de Toulouse Le MIRAIL
- [30] **Thibault GIBERT**, Etude statistique, Dossier analyse d'une série temporelle, Agriculture – Production laitière bovine en France, Master II Ingénierie de projets, (2010-2011)
- [31] **Michel TENENHAUS**, La méthodologie de Box-Jenkins
- [32] Chapitre 2, Tests de non stationnarité et processus aléatoires non stationnaires
- [33] **Karine BOUTHEVILLAIN, Alexandre MATHIS**, Prévision : mesures, erreurs et principaux résultats. In : Economie et statistique, N°285-286(1995) Doi :10.3406/estat.1995.5982 consulté le 17 février 2017 à l'adresse, http://www.persee.fr/web/revues/home/prescript/article/estat_0336-1454_1995_num_285_1_5982
- [34] **Amar RAMUDHIN**, Ingénieur. Ph.D, Prévision de la demande, 38p, consulté le 17 Février 2017 à l'adresse , <http://samos.univ-paris1.fr/archives/membres/perraudin/ST/resumechap2.pdf>

**Titre : «Analyse et prévision des pleines et des basses mers par la méthode Census X11-ARIMA-GARCH :
Application au cas de Hell-Ville »**

Résumé : Ce travail qui comporte 163 pages est un approfondissement de la modélisation Census X11-ARIMA-GARCH des séries chronologiques périodiques avec tendance non linéaire.

Avant de mettre en œuvre la méthode Census X11, nous avons testé, après stabilisation logarithmique des données, l'existence de saisonnalité et de tendance.

La méthode du corrélogramme et du périodogramme ont permis en premier lieu de trouver la saisonnalité de 29 Lags (14,765 jours) pour les deux chroniques sur les pleines mers et sur les basses mers.

Ensuite, la mise en œuvre des tests de corrélation des rangs de Mann- Kendall, de Kendall et de Spearman nous a permis de mettre en évidence l'existence de faibles tendances monotoniques.

En particulier, les taux τ_b de Kendall dont les valeurs étaient de -9,98% pour les pleines mers et de 11,80% pour les basses mers, confirment l'existence de faible tendance monotonique.

En outre, l'analyse de la variance à un facteur nous a montré l'existence de sept cycles compris entre 6 ans et 9 ans à l'intérieur desquels la saisonnalité de 29 Lags restait stable.

Par conséquent, nous avons décidé de construire sept modèles correspondant à ces sept cycles.

Comme nos séries sont non linéaires et hétéroscédastiques, nous avons utilisé un modèle ARIMA(p,d,q) avec erreur GARCH(P,Q).

Le modèle adéquat est choisi parmi les modèles ARIMA(p,d,q)-GARCH(P,Q), ARIMA(p,d,q) -EGARCH(P,Q) et ARIMA(p,d,q) -GJR(P,Q) et correspond au modèle qui rend non corrélé les résidus, les résidus standardisés, les résidus standardisés au carré et qui maximise le logarithme de vraisemblance.

Finalement nous sommes arrivés à construire sept excellents modèles avec des Racines Carrées des Erreurs Quadratiques Moyennes inférieures à 5%.

Aux vues des indicateurs I de THEIL, nos modèles offrent plus de précision que la méthode harmonique utilisée par le service hydrographique et océanographique de la marine française pour prévoir les marées à Hell-Ville sauf pour les modèles 1 à 4 sur les pleines mers.

Ces modèles établis sur la base des données entre 1990 et juin 2012 arrivent à prévoir au moins 600 données à partir du mois de Juillet 2012 date à partir de laquelle nous avons constaté un changement structurel.

Mots clés : *Census X11, ARIMA-GARCH, Test de Dickey Fuller Augmenté, Test de Ljung-Box, Critère d'information*

**Title : « Analysis and prediction of high and low tides using Census X11-ARIMA-GARCH method:
Application to the case of Hell-Ville »**

Summary: This work have 163 pages. It is an advance study using Census X11-ARIMA-GARCH for modeling series with non linear trend and seasonality.

In first, we detected by corelogram and periodogram methods existence of seasonality.

Particularly, we found seasonality equal to 29 Lags (14,765 days) from high and low tides.

In second, Mann- Kendall, Kendall and Spearman correlation rank methods permitted to detect existence of monotonic trend.

Kendall's τ_b have respectively -9,98% and 11,80% values from high and low tides. These values confirmed existence of monotonic trend.

In third, using analysis of variance one factor, we found seven cycles between 6 and 9 years in witch seasonality of 29 Lags is stable.

Consequently, we decided to create seven models corresponding with each cycle.

As our series is non linear and heteroscedastic, we used ARIMA(p,d,q)-GARCH(P,Q) models.

we obtained several good models after applying ARIMA(p,d,q)-GARCH(P,Q), ARIMA(p,d,q) -EGARCH(P,Q) and ARIMA(p,d,q) -GJR(P,Q) models.

Optimal model correspond for model with uncorrelated residual, uncorrelated standardized residual and uncorrelated square of standardized residual and having also maximum of Logarithmic Likelihood.

Finally, we found seven excellent models because Root Mean Square Errors is low than 5%.

THEIL ratio is less than 1 except model 1 to 4 high tides.

Although our models were based on data between 1990 to June 2012, we can predict at least 600 data beyond the month of July 2012 where we detected a structural change in series.

Key word : *Census X11, ARIMA-GARCH, Augmented Dickey Fuller test, Ljung-Box test, Information criterion.*

Impétrant

Directeur de Thèse

RASOLOMANAMPAMONJY Alimanana Nainamahefa

Pr. ANDRIAMAMPIANINA José Marie Michel

e-mail : nainamahef@gmail.com

Tél : 032 11 083 22