



UNIVERSITÉ DE LILLE
FACULTÉ DES SCIENCES ET TECHNIQUES

**Prévisions dans les séries
chronologiques :
Projet**

MASTER 2
Ingénierie Statistique et Numérique

Professeur : Laurence Broze

Préparé par : Khalil AL-SAYED

le 21 Decembre 2021

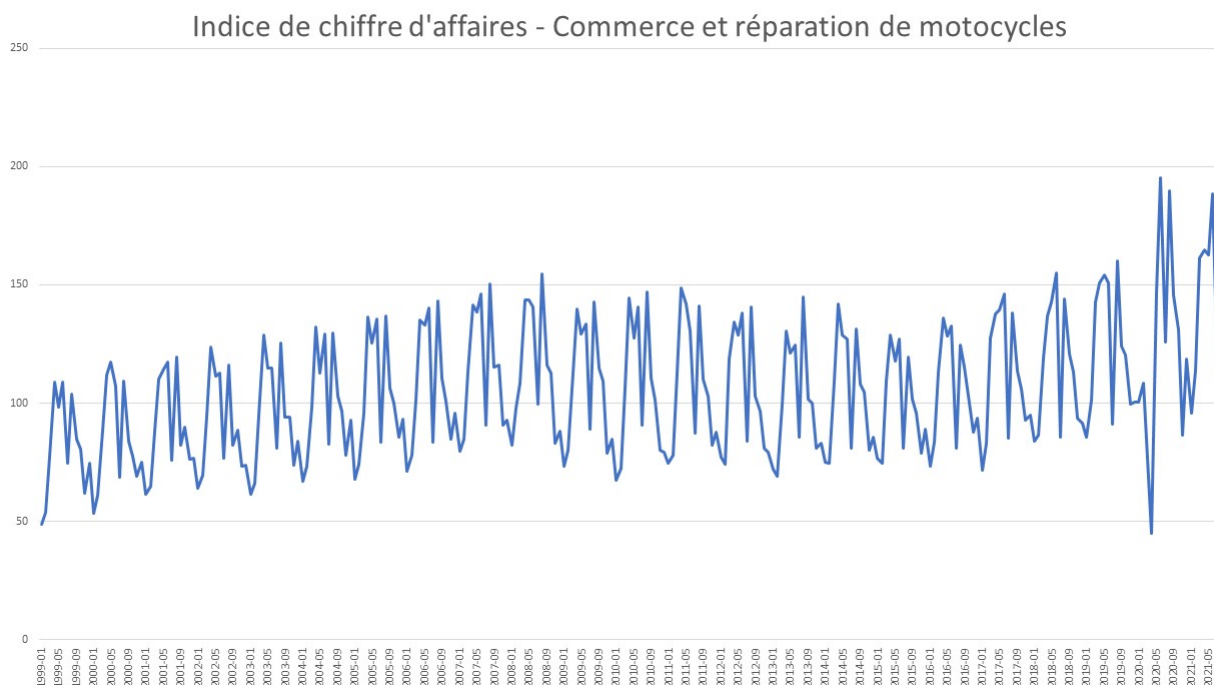
Table des matières

1	Introduction	2
2	Décomposition	2
3	lissage	5
4	Box-Jenkins	7
5	Conclusion	14
6	Annexe	15

.

1 Introduction

La série de ce projet mesure mensuellement (de 01/1999 à 08/2021) l'évolution de l'activité des entreprises (Indice de chiffre d'affaires) dans les secteurs du commerce et réparation de motocycles en france, ces évolutions sont élaborés chaque mois à partir des déclarations mensuelles (CA3) faites par les entreprises relevant du régime fiscal réel normal pour le paiement de la taxe sur la valeur ajoutée (TVA).

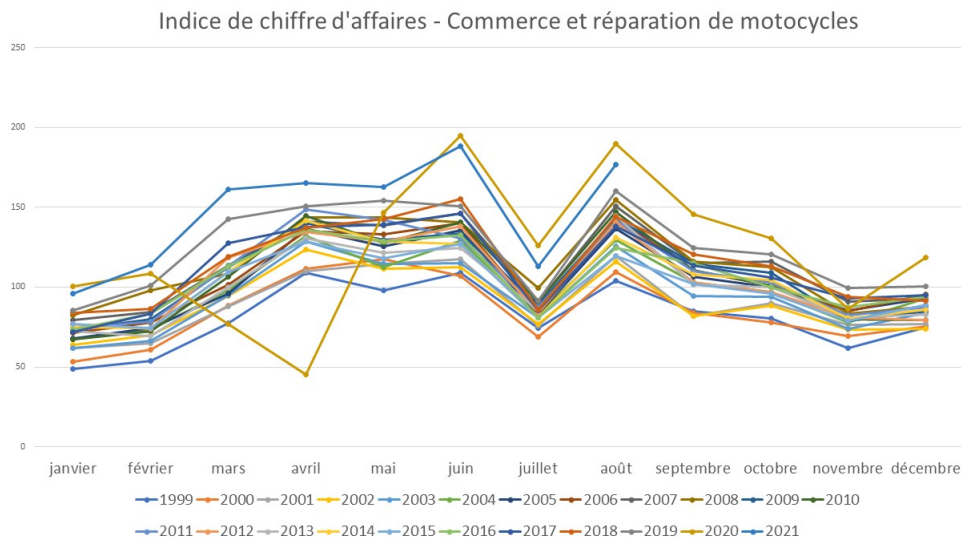


2 Décomposition

D'après le graphique de la série ci-dessus, notre séries présentant à la fois tendance et mouvement saisonnier, alors ont peut tracer le tableaux de Buys-Ballot :

Xt: Mois\Année	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
janvier	48,82	53,45	61,67	63,99	61,68	67,16	67,7	71,37	79,51	82,33	73,32	67,26	74,56	77,08	72,2	74,82	76,81	73,46	71,83	84,05	85,44	100,51	95,72
février	53,66	61,03	64,88	69,68	66,24	73,16	74,14	77,94	84,58	97,71	80,02	72,34	78,12	74,08	69,15	74,46	74,5	83,51	83,24	86,59	101,19	108,34	113,87
mars	77,61	88,43	88,01	94,37	94,74	98,18	95,81	101,3	113,5	108,63	112,5	106,43	112,75	118,54	99,7	109,26	109,35	113,15	127,67	119,19	142,84	76,71	161,22
avril	108,8	111,72	110,1	123,58	128,75	132,17	136,55	135,15	141,3	143,39	139,93	144,41	148,47	134,36	130,65	141,83	128,71	136,07	137,72	136,93	150,71	45,17	164,94
mai	98,14	117,36	114,46	111,53	114,68	112,52	125,35	133,15	138,38	143,63	129,36	127,68	141,96	128,71	121,28	128,61	117,76	128,55	139,17	142,66	154,24	146,6	162,67
juin	108,94	107,18	117,52	112,51	115	129,09	135,48	140,06	146	140,45	133,58	140,82	130,34	138,22	124,75	127,06	127,14	132,52	146,29	154,95	150,63	195,01	188,36
juillet	74,49	68,77	75,7	76,84	81,06	82,8	83,3	83,39	90,56	99,49	88,97	90,74	87,31	84,01	85,43	80,85	80,85	81,09	85,38	85,65	91,25	125,9	113,01
août	103,94	109,49	119,66	116,22	125,59	129,66	136,77	143,02	150,45	154,69	142,85	146,88	141,03	140,51	144,71	131,3	119,56	124,71	137,87	144,16	160,18	189,84	176,93
septembre	84,64	83,76	82,04	82,08	94,25	103,07	106,34	110,65	115,02	116,08	114,61	110,48	110,17	102,79	101,51	108,13	101,77	115,02	113,6	120,74	124,29	145,77	
octobre	80,57	77,75	89,95	88,52	94,02	96,69	99,98	100,98	116,03	112,72	109,21	101,13	103	96,72	99,84	104,66	95,7	100,19	106,03	113,19	120,45	130,79	
novembre	61,99	69,26	76,48	73,44	73,69	78,04	85,42	84,8	90,79	83,18	78,74	79,9	82,2	80,8	80,92	80,14	79,02	87,52	92,95	93,81	99,43	86,62	
décembre	74,68	75,12	76,83	73,84	83,82	92,68	93,07	95,65	92,69	88,1	84,81	79,46	87,62	79,35	82,91	85,72	88,85	93,84	94,71	91,64	100,28	118,43	

et pour savoir le schéma de décomposition (additif, multiplicatif ou mixte) il faut tracer le graphique de notre profil :



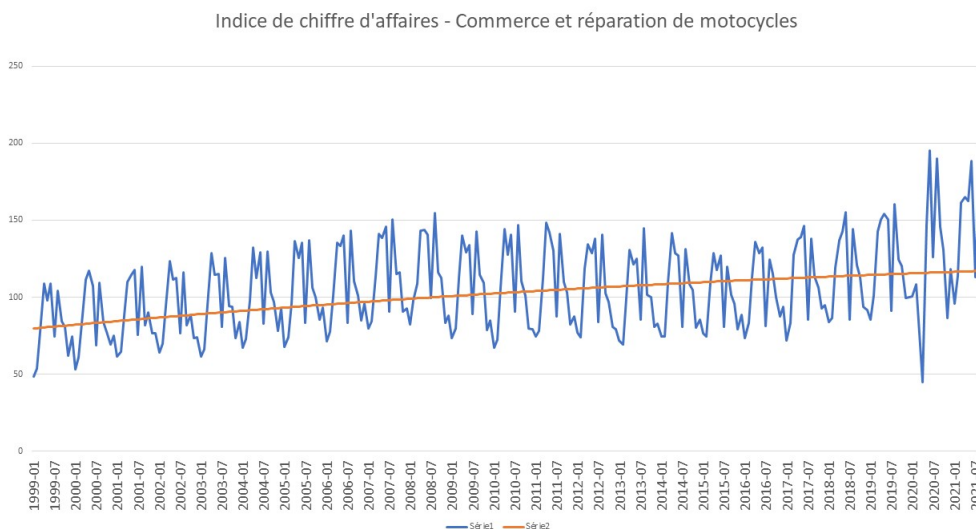
Puisque dans notre cas les saisons représenté par les droits de profil ne sont pas parallèles donc le modèle est multiplicatif :

$$X_t = T_t \times I_t \times \epsilon_t \quad (1)$$

Détermination de la tendance : Dans notre cas on va utiliser la tendance de Gompertz :

$$T_t = e^{ab^t+c} \quad (2)$$

et on trouve (voir mon feuille de calcul excel (Gompertz)) que $a=-0.51$ $b=0.99$ et $c=4.89$, alors le graphique de notre courbe de tendance est :



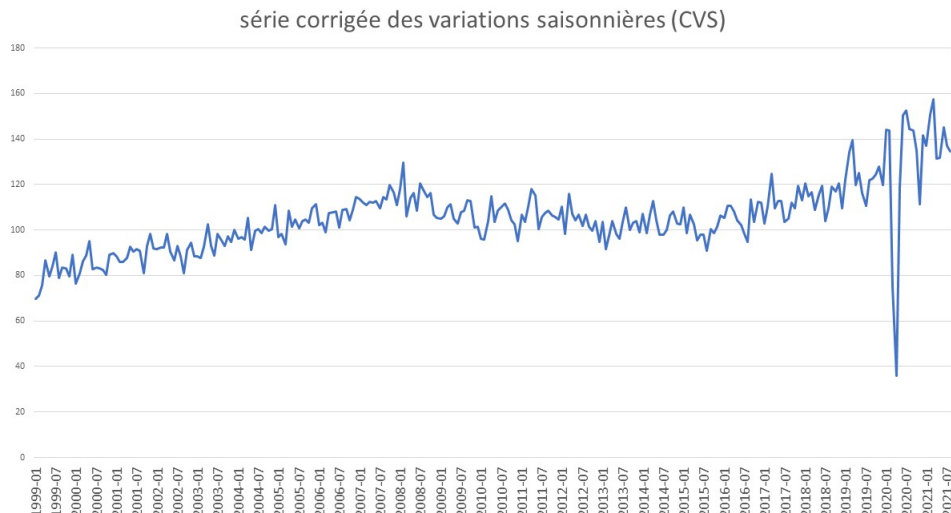
Désaisonnaliser la série : On commence par enlever la tendance : X_t / T_t

Xt/Tt:	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	moyenne It		
janvier	0,612320	0,504220	0,729360	0,736830	0,692500	0,786330	0,725840	0,749260	0,818340	0,831710	0,727800	0,856710	0,716750	0,730210	0,674610	0,690090	0,699840	0,661670	0,640030	0,741130	0,746590	0,870160	0,821680	0,724350	0,69823	
	66	145	925	928	881	774	774	498	451	49	672	232	384	05	969	078	609	617	90	905	32	816	96	78		
février	0,671290	0,740850	0,765580	0,800580	0,742190	0,800580	0,793460	0,816840	0,869130	0,985600	0,793190	0,705370	0,750030	0,700960	0,645400	0,686050	0,678130	0,751500	0,741050	0,763100	0,883300	0,937260	0,976820	0,782530	0,75431	
	60	1398	546	659	439	439	723	727	781	385	05	201	538	237	367	241	820	248	283	517	785	907	1200	703	570	
mars	0,968400	0,708860	0,036150	1,081930	0,059370	1,072330	1,023540	1,059870	1,164460	1,094110	1,113570	1,036460	1,081170	1,203390	0,929510	1,005640	0,994370	1,017290	1,135600	1,049540	1,245920	0,663150	1,382060	1,061110	1,02285	
	6684	404	204	247	953	366	2355	949	079	621	614	163	040	143	064	043	223	825	049	6490	2490	05	193	428	602	011
avril	1,354110	1,349630	1,293270	1,413790	1,436780	1,440840	1,456160	1,411670	1,447390	1,442060	1,383150	1,404390	1,421920	1,268370	1,216720	1,304060	1,169280	1,222230	1,239940	1,204770	1,313560	0,390200	1,412990	1,303530	1,25653	
	273	338	455	07	142	118	486	165	116	92	144	26	656	05	04	946	89	995	555	9	097	994	579	034	001	001
mai	1,218320	1,414360	1,341450	1,273220	1,277200	1,224310	1,334360	1,388450	1,415250	1,442340	1,276880	1,240090	1,357900	1,213620	1,128220	1,181290	1,068770	1,153630	1,235760	1,254170	1,343030	1,682220	1,611450	1,344770	1,29628	
	518294	127	17	562	734	502	853	205	519	964	44	066	207	378	926	472	058	999	469	469	2655	139260	28004	21388	1388	
juin	1,348960	1,288590	1,374210	1,281690	1,278210	1,401970	1,439640	1,458080	1,490840	1,408330	1,316710	1,365890	1,245210	1,301780	1,159240	1,165860	1,152790	1,188180	1,297880	1,361120	1,108700	1,682220	1,611450	1,344770	1,29628	
	827	63	798	168	827	978	918	78	833	458	722	926	368	614	086	663	354	362	213	993	138	548	820	978	15	
juillet	0,920050	0,824830	0,883210	0,873500	0,899190	0,897570	0,883610	0,866690	0,923290	0,996150	0,875780	0,879000	0,833100	0,790310	0,792990	0,741100	0,752370	0,726400	0,756840	0,751770	0,793510	1,085280	0,966170	0,856200	0,82533	
	333	887	680	875	900	147	206	225	216	923	000	875	807	850	040	740	653	058	567	456	303	78	818	921	000	
août	1,280580	1,310130	1,393000	1,318400	1,390410	1,402930	1,448260	1,483990	1,531520	1,546590	1,404220	1,420990	1,344060	1,320310	1,341810	1,202320	1,081990	1,116140	1,221100	1,264320	1,391880	1,635310	1,511650	1,363560	1,31439	
	289	699	582	000	156	882	1428	1462	229	31	154	508	849	587	285	052	040	056	589	638	699	695	558	1200	703	570
septembre	1,040190	0,999890	0,952940	0,929170	1,041390	1,113160	1,124080	1,146240	1,169050	1,158880	1,125080	1,067460	1,048680	0,964770	0,940230	0,989150	0,920120	1,028490	1,005300	1,058080	1,079210	1,254800	1,052560	1,01460	0,952	
	987	927	903	503	793	7058	549	085	472	644	29	574	823	814	708	435	040	093	123	649	415	354	1000	704	985	
octobre	0,987710	0,925990	1,042520	1,036820	1,042340	1,055020	1,044360	1,177510	1,123710	1,070610	0,975870	0,979250	0,906770	0,923780	0,956450	0,864420	0,895090	0,937520	0,991140	1,045090	1,125060	1,004860	0,968630	1,004860	0,96863	
	338	179	06	1	851	377	1028	786	6	585	813	697	227	011	1340	78	939	479	12	513	761	887	14	887	14	
novembre	0,758060	0,822940	0,884460	0,827930	0,811050	0,839750	0,899830	0,875600	0,919990	0,828030	0,770860	0,770030	0,780560	0,756660	0,747920	0,731630	0,713080	0,781200	0,821180	0,820700	0,862070	0,744590	0,807630	0,778520	0,80763	
	031	749	035	112	599	535	559	559	559	559	559	559	559	559	559	559	559	559	559	559	559	559	559	559	559	
décembre	0,910990	0,890500	0,886570	0,830780	0,920760	0,995470	0,978740	0,986040	0,937780	0,875760	0,829170	0,764820	0,831040	0,742250	0,765510	0,781890	0,801040	0,836880	0,836040	0,801180	0,868800	1,017330	0,867690	0,83640	0,86769	

Pour obtenir la série corrigée des variations saisonnières (CVS), on calcule X_t / T_t

Xt/It:	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
janvier	69,9188	76,5498	88,3223	91,6443	88,3366	96,1849	96,9583	102,214	113,872	117,911	105,007	96,3282	106,783	110,392	103,403	107,155	110,005	105,207	102,873	120,374	122,365	143,948	137,087
février	71,1371	80,9075	86,0115	92,3749	87,8145	96,9883	98,2875	103,325	112,127	129,534	106,082	95,9012	103,563	98,2080	91,6722	98,7117	98,7648	110,708	110,531	114,792	134,147	143,626	150,957
mars	75,8762	86,4544	86,0438	92,2618	92,6235	95,9866	93,6096	99,0369	110,964	106,203	109,986	104,052	110,231	115,891	97,4727	106,819	106,907	110,622	112,4817	116,527	139,648	149,963	157,618
avril	86,5876	88,9114	87,2222	98,3501	102,464	105,186	108,672	107,558	112,452	114,115	111,362	114,927	118,158	106,929	103,974	112,874	102,432	108,290	109,603	108,974	119,941	135,948	131,266
mai	79,5372	95,1140	92,7637	90,3891	92,9420	91,1914	101,589	107,911	112,149	116,404	104,839	103,477	115,051	104,312	98,2910	104,231	95,4382	104,182	112,789	115,618	125,003	118,811	131,835
juin	84,0403	82,6826	90,6593	86,7944	88,7152	95,5848	104,514	108,047	112,629	108,348	103,048	108,633	100,549	106,628	96,2368	98,0188	98,0805	102,230	112,853	119,534	116,201	155,342	145,307
juillet	90,2549	83,3239	91,7205	93,1018	98,2149	100,323	100,928	101,038	109,725	120,545	107,798	109,943	105,787	101,789	103,509	97,9604	97,9604	98,2512	103,449	103,778	110,561	152,544	136,926
août	79,0783	83,3008	91,0382	88,4210	95,5498	98,6463	104,055	108,810	114,463	117,689	108,681	111,747	107,296	106,901	110,096	99,8940	90,9621	94,8803	104,892	109,678	121,866	144,431	134,609
septembre	83,4212	82,5538	80,8586	80,8993	92,8928	101,585	104,808	109,056	113,363	114,408	112,959	108,888	108,583	101,309	100,044	106,578	100,304	113,363	111,964	119,001	122,500	143,671	134,609
octobre	83,1792	90,2678	92,8629	91,3866	97,0647	99,8212	103,217	104,250	119,787	116,370	112,746	104,405	106,335	99,8522	103,073	108,049	98,7991	103,434	109,463	116,851	124,505	135,025	134,609
novembre	79,6249	88,9633	98,2370	94,3322	94,6533	100,240	109,720	108,923	116,618	106,843	101,140	102,630	105,584	103,788	103,942	102,938	101,499	112,417	119,392	120,497	127,713	111,261	134,609
décembre	89,2869	89,8130	91,8574	88,2826	100,214	110,807	111,273	114,358	110,819	105,331	101,398	95,0018	104,757	94,8703	99,1266	102,486	106,228	112,194	113,234	109,564	119,894	141,594	134,609

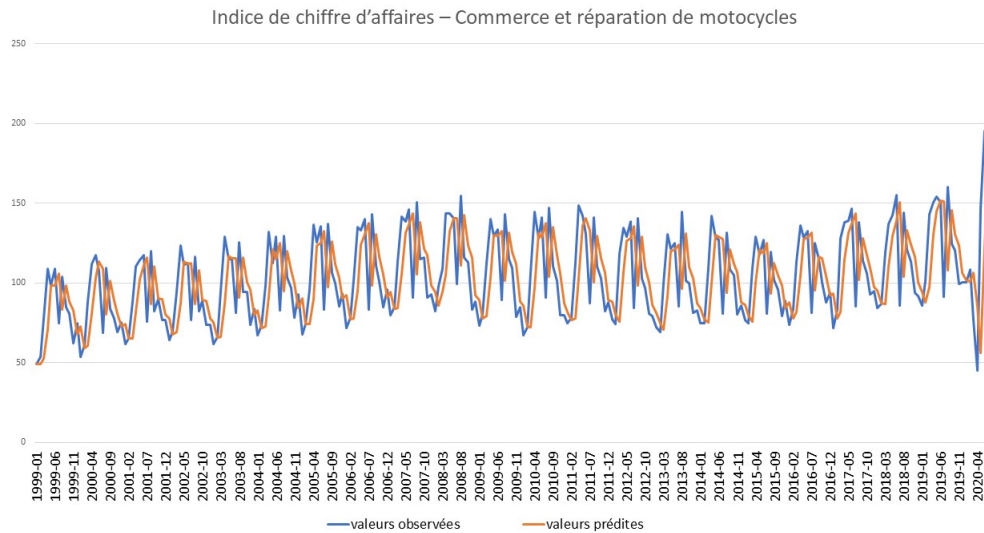
Alors voilà le graphique de notre série corrigée :



3 lissage

Méthode de Holt : D'après mon feuille de calcul Excel (Hold), j'ai trouvé $\alpha=0,72$ et $\gamma=0$ et $MAPE=19.8\%$

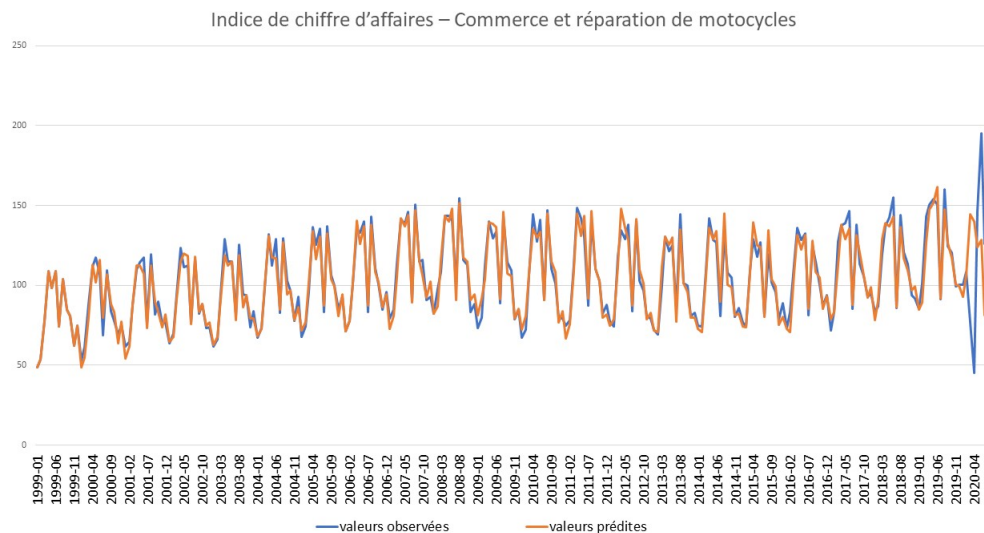
Voici le graphique sur l'échantillon d'estimation représentant les valeurs prédites et les valeurs observées :



Méthode de Holt-Winters additive :

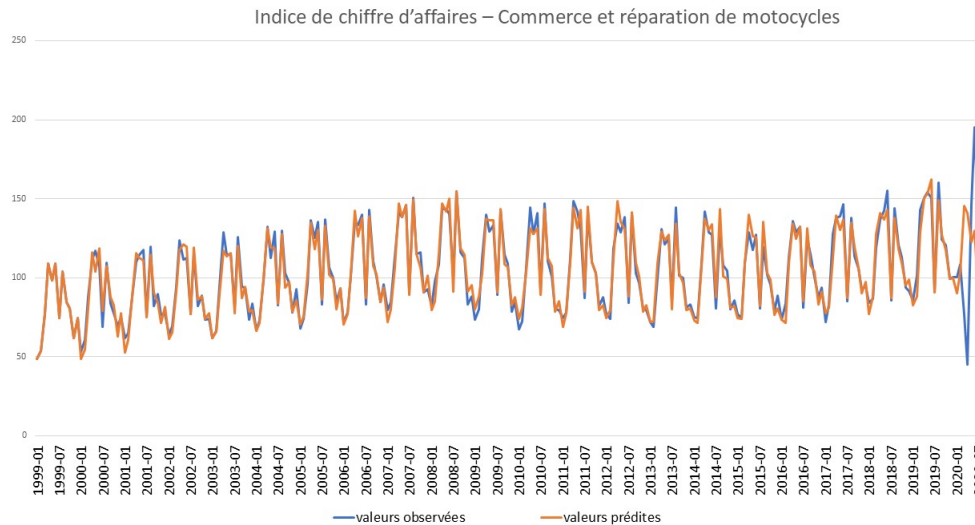
D'après mon feuille de calcul Excel (winters_additif), j'ai trouvé $\alpha=0,20$ et $\gamma=0.00059$ et $\delta=0.84$ et $MAPE=5.66\%$

Voici le graphique sur l'échantillon d'estimation représentant les valeurs prédites et les valeurs observées :



Méthode de Holt-Winters multiplicative : D'après mon feuille de calcul Excel (winters_multiplicative), j'ai trouvé $\alpha=0,21$ et $\gamma=0.00057$ et $\delta=0.68$ et $MAPE=5.63\%$

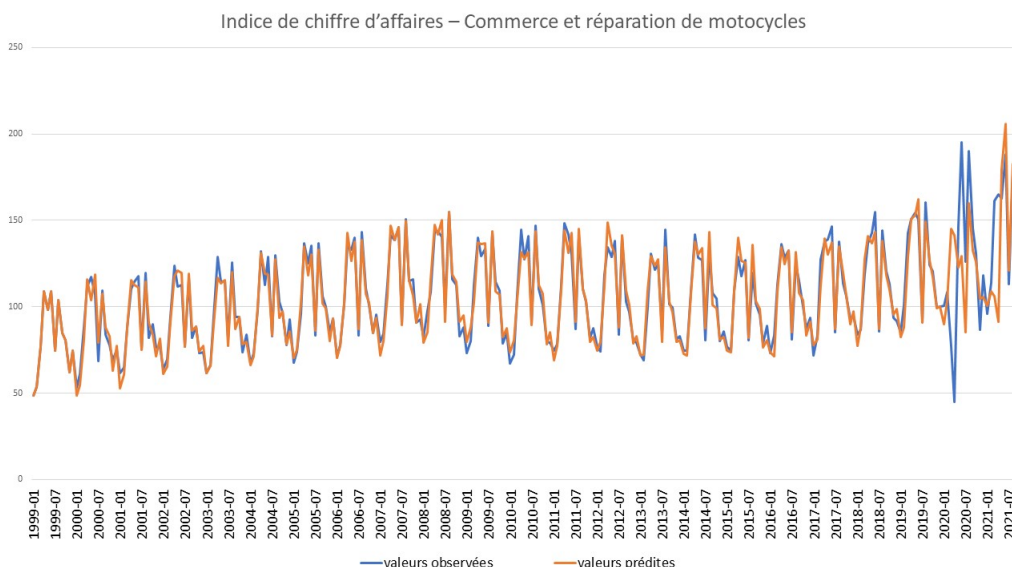
Voici le graphique sur l'échantillon d'estimation représentant les valeurs prédites et les valeurs observées :



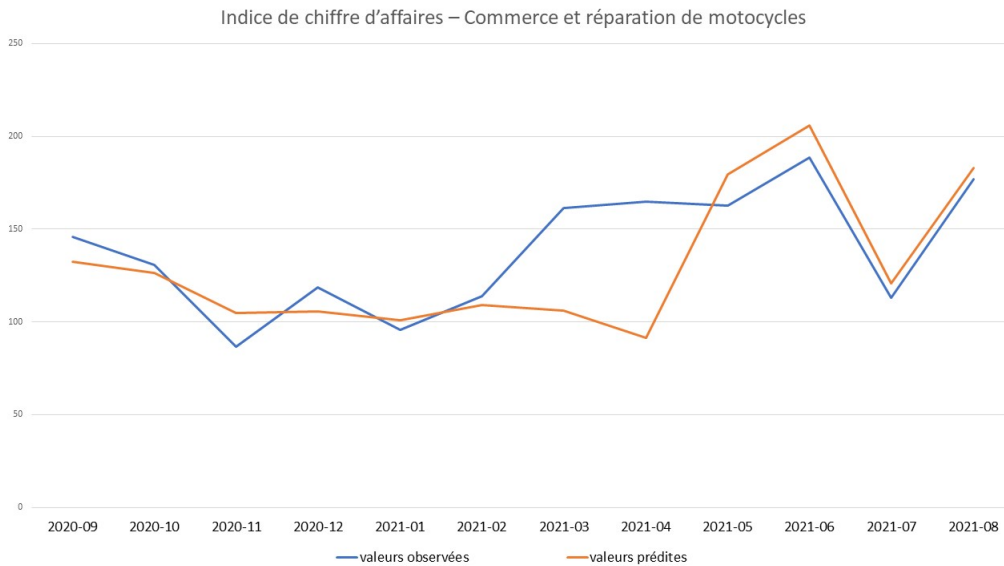
La meilleure méthode : c'est la méthode ayant la plus petit MAPE, donc c'est la méthode de Holt-Winters multiplicative ($MAPE=5,63\%$).

On appliquera cette meilleure méthode de lissage à la prévision des douze dernières données et on calculera le MAPE sur ces 12 données, on obtient alors $MAPE=13,5\%$ (voir MAPE2 dans la feuille de calcul winters_multiplicative).

Voici le graphique sur l'ensemble de l'échantillon représentant les valeurs prédites et les valeurs observées :




Si on zoome sur les 12 dernières mois on trouve :



4 Box-Jenkins

stationnarité de la série : Au debut il faut s'assurer de la stationnarité de la série, on peut utiliser le test de Dickey-fuller augmenté :

 gretl: ADF test

```

Test de Dickey-Fuller augmenté pour vl
test à reculons à partir de 15 retards, suivant le critère AIC
taille de l'échantillon 246
hypothèse nulle de racine unitaire : a = 1

test avec constante
avec 13 retards de (1-L)vl
modèle: (1-L)y = b0 + (a-1)*y(-1) + ... + e
valeur estimée de (a - 1): -0,108952
statistique de test: tau_c(1) = -1,6611
p. critique asymptotique 0,4512
Coeff. d'autocorrélation du 1er ordre pour e: 0,010
différences retardées: F(13, 231) = 103,432 [0,0000]
  
```

notre p-critique est 45,12% alors on ne rejette pas H_0 alors il y a une racine unitaire donc X_t n'est pas stationnaire.

pour la série en différence première $(1 - B)X_t$:

gretl: ADF test



```
Test de Dickey-Fuller augmenté pour d_vl
test à reculons à partir de 15 retards, suivant le critère AIC
taille de l'échantillon 243
hypothèse nulle de racine unitaire : a = 1

test avec constante
avec 15 retards de (1-L)d_vl
modèle: (1-L)y = b0 + (a-1)*y(-1) + ... + e
valeur estimée de (a - 1): -7,73498
statistique de test: tau_c(1) = -5,8306
p. critique asymptotique 3,029e-07
Coeff. d'autocorrélation du 1er ordre pour e: 0,000
différences retardées: F(15, 226) = 109,380 [0,0000]
```

notre p-critique est 0.27% alors on rejette H_0 alors $(1 - B)X_t$ est stationnaire.

pour la série en différence saisonnière $(1 - B^{12})X_t$:

gretl: ADF test

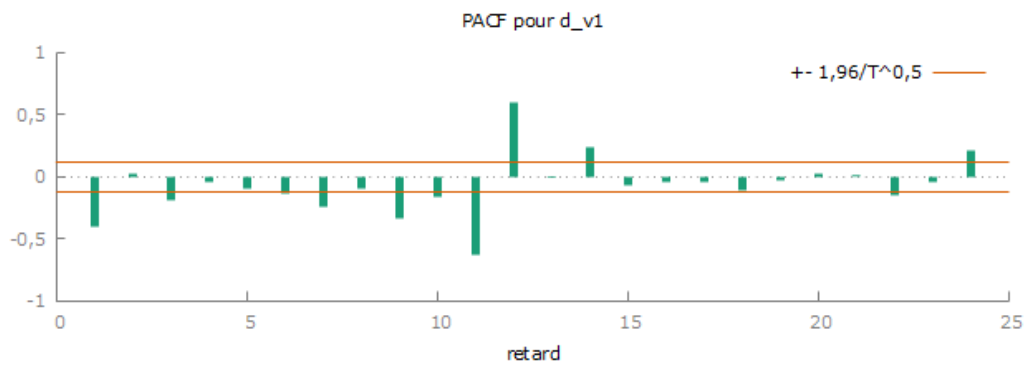
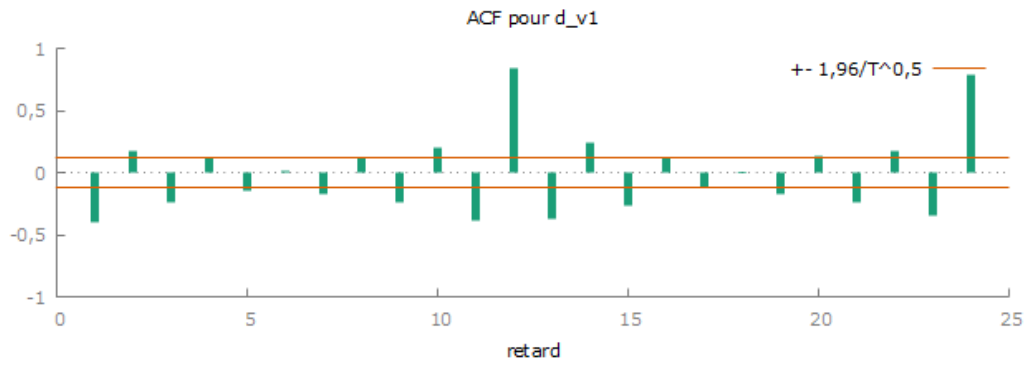


```
Test de Dickey-Fuller augmenté pour sd_vl
test à reculons à partir de 15 retards, suivant le critère AIC
taille de l'échantillon 236
hypothèse nulle de racine unitaire : a = 1

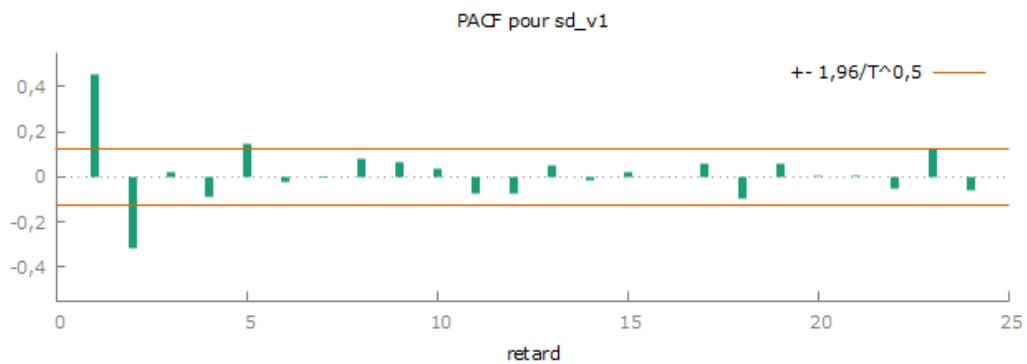
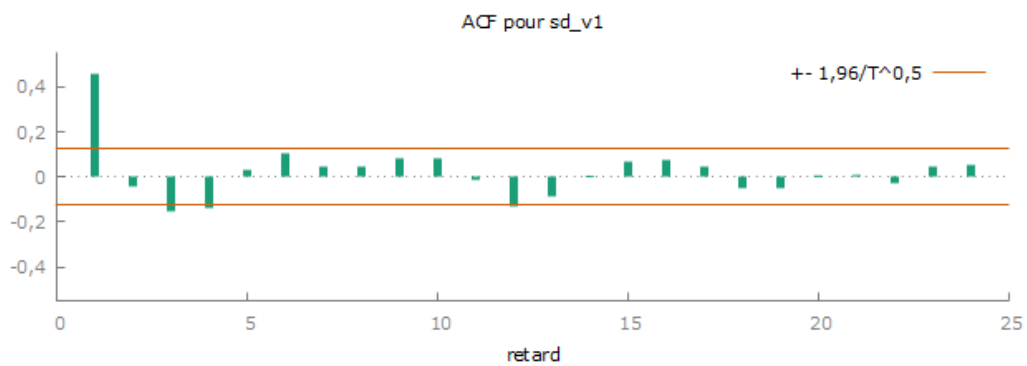
test avec constante
avec 11 retards de (1-L)sd_vl
modèle: (1-L)y = b0 + (a-1)*y(-1) + ... + e
valeur estimée de (a - 1): -0,655404
statistique de test: tau_c(1) = -4,39815
p. critique asymptotique 0,0002939
Coeff. d'autocorrélation du 1er ordre pour e: 0,006
différences retardées: F(11, 223) = 6,584 [0,0000]
```

notre p-critique est 0.029% alors on rejette H_0 alors $(1 - B^{12})X_t$ est stationnaire.

Pour savoir quel Différence je dois choisir $((1 - B)X_t$ ou $(1 - B^{12})X_t$) , il faut tracer le corrélogramme des deux, alors pour $(1 - B)X_t$:



et pour $(1 - B^{12})X_t$

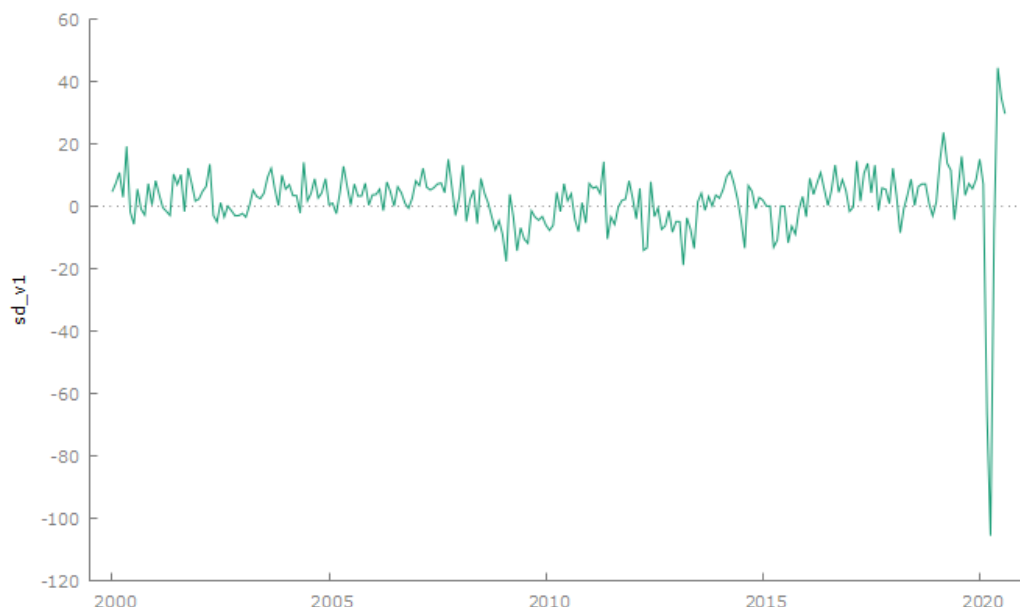


d'après les deux corrélogrammes on trouve que le modèle de $(1 - B^{12})X_t$ est plus claire

(MA(12)) que le modèle de $(1 - B)X_t$ (mixte), Alors je ferais mieux de travailler sur $(1 - B^{12})X_t$.

Identification d'un modèle :

on a déjà dit que $(1 - B^{12})X_t$ est MA(12), et puisque $(1 - B^{12})X_t$ présente encore de la saisonnalité :



alors pour capter cette saisonnalité on peut multiplier ϵ_t par $(1 - \Theta_1 B^{12})$, donc finalement notre modèle devient :

$$(1 - B^{12})X_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_{12} B^{12})(1 - \Theta_1 B^{12})\epsilon_t \quad (3)$$

Estimation du modèle :

gretl : modèle 1

	coefficient	éc. type	z	p. critique	
const	1,70359	0,816221	2,087	0,0369	**
theta_1	0,593735	0,0611516	9,709	2,75e-022	***
theta_2	0,0686545	0,0806904	0,8508	0,3949	
theta_3	-0,0732422	0,0785835	-0,9320	0,3513	
theta_4	-0,255721	0,0972503	-2,630	0,0086	***
theta_5	-0,0261610	0,0776979	-0,3367	0,7363	
theta_6	0,384269	0,120142	3,198	0,0014	***
theta_7	0,111842	0,114860	0,9737	0,3302	
theta_8	-0,0127860	0,102680	-0,1245	0,9009	
theta_9	0,252194	0,105039	2,401	0,0164	**
theta_10	0,381099	0,121050	3,148	0,0016	***
Theta_1	-0,397317	0,100754	-3,943	8,03e-05	***
Moyenne var. dép.	1,839718	Éc. type var. dép.	11,24189		
Moyenne innovations	-0,001531	Éc. type innovations	8,621229		
R2	0,410929	R2 ajusté	0,386074		
Log de vraisemblance	-891,0369	Critère d'Akaike	1808,074		
Critère de Schwarz	1853,748	Hannan-Quinn	1826,461		

Fichier Édition Tests Enregistrer Graphiques Analyse LaTeX

Évaluations de la fonction : 77
Évaluations du gradient : 26

Modèle 2: ARMA, utilisant les observations 2000:01-2020:08 (T = 248)
Estimation par AS 197 (MV exact)
Variable dépendante: sd_vl
Écart-types basés sur la matrice hessienne

	coefficient	éc. type	z	p. critique	
const	1,74641	0,789136	2,213	0,0269	**
theta_1	0,549377	0,0585435	9,384	6,35e-021	***
theta_4	-0,161249	0,0714645	-2,256	0,0240	**
theta_6	0,294618	0,0839713	3,509	0,0005	***
theta_9	0,309339	0,0938162	3,297	0,0010	***
theta_10	0,363834	0,110111	3,304	0,0010	***
Theta_1	-0,411098	0,0928613	-4,427	9,55e-06	***

Moyenne var. dép.	1,839718	Éc. type var. dép.	11,24189
Moyenne innovations	0,000975	Éc. type innovations	8,766470
R2	0,396205	R2 ajusté	0,383730
Log de vraisemblance	-892,8555	Critère d'Akaike	1801,711
Critère de Schwarz	1829,818	Hannan-Quinn	1813,026

alors notre modèle devient :

$$(1 - B^{12})X_t = \mu + (1 - \theta_1 B - \theta_4 B^4 - \theta_6 B^6 - \theta_9 B^9 - \theta_{10} B^{10})(1 - \Theta_1 B^{12})\epsilon_t \quad (4)$$

avec θ_1 θ_4 θ_6 θ_9 θ_{10} Θ_1 et μ sont tous significatives.

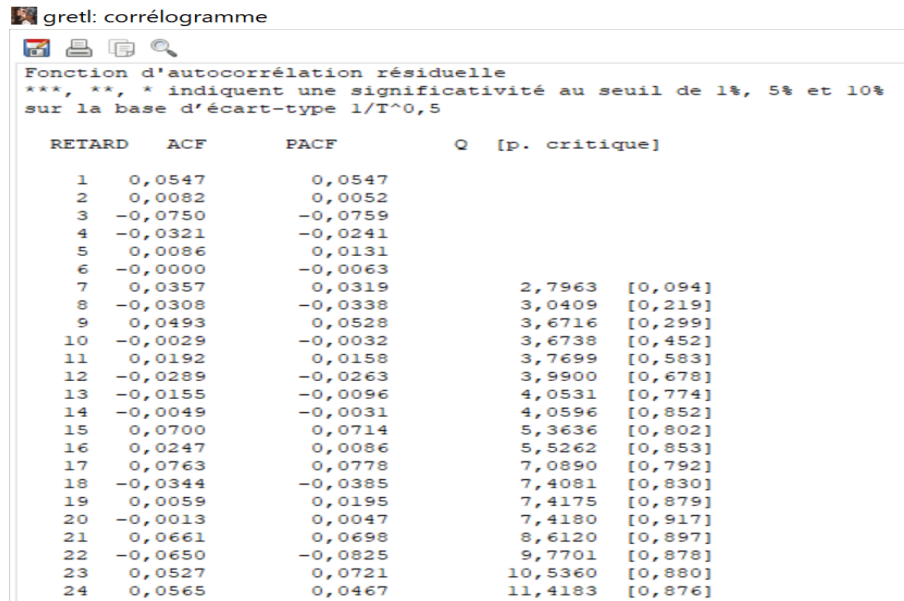
Validation du modèle : Maintenant il faut démontrer que ϵ_t est un bruit blanc.

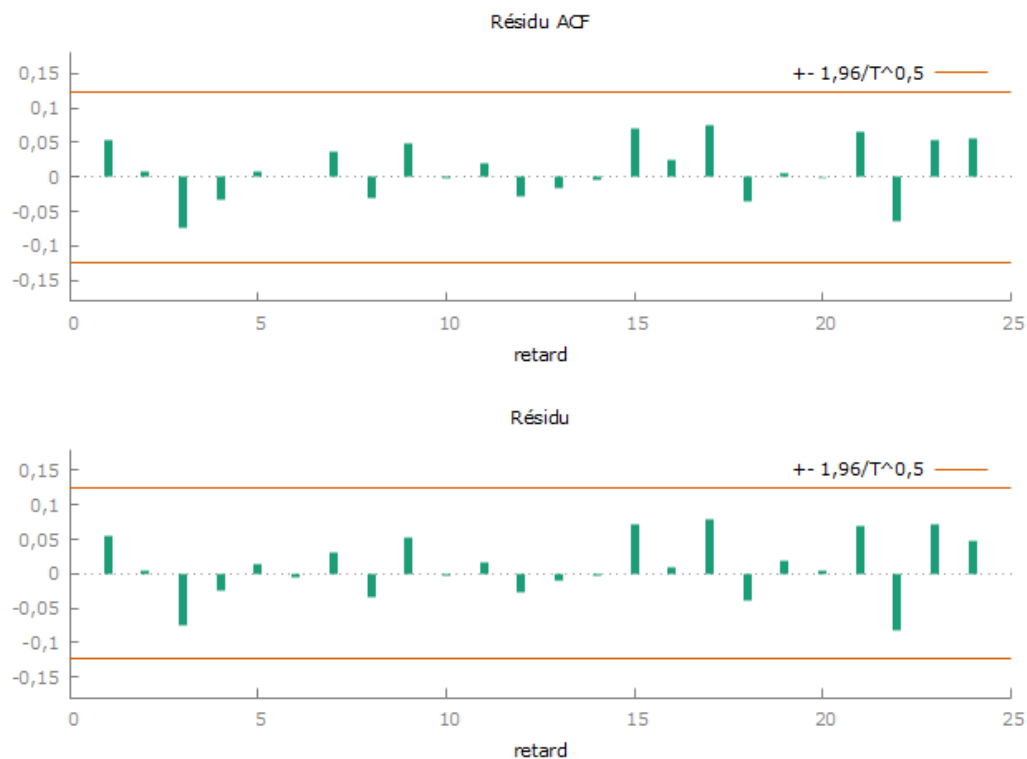
— Pour montrer que $E(\epsilon)=0$ il faut que

$$\sqrt{T} \times \text{Moyenne innovations} / \text{Ec. type innovations} < 1.96$$

et ça marche dans notre cas ($0.00175 < 1.96$).

— pour montrer que ϵ est non-autocorrélé il faut tracer le corrélogramme des résidus :





Donc on peut déduire que ϵ_t est un bruit blanc.

Calcul des prévisions : Notre modèle a le meilleur MAPE sur l'échantillon d'estimation (5.66%)

Note : * indique un résidu supérieur à 2,5 fois l'écart-type
 Statistiques sur la qualité de la prévision utilisant 248 observations

Moyenne Erreur	0,00097529
Moyenne Erreur Carrée (racine)	8,7665
Moyenne Erreur Absolue	5,5617
Moyenne Pourcentage Erreur	-0,73597
Moyenne Pourcentage Erreur Absolue	5,6637
U2 de Theil	0,32294

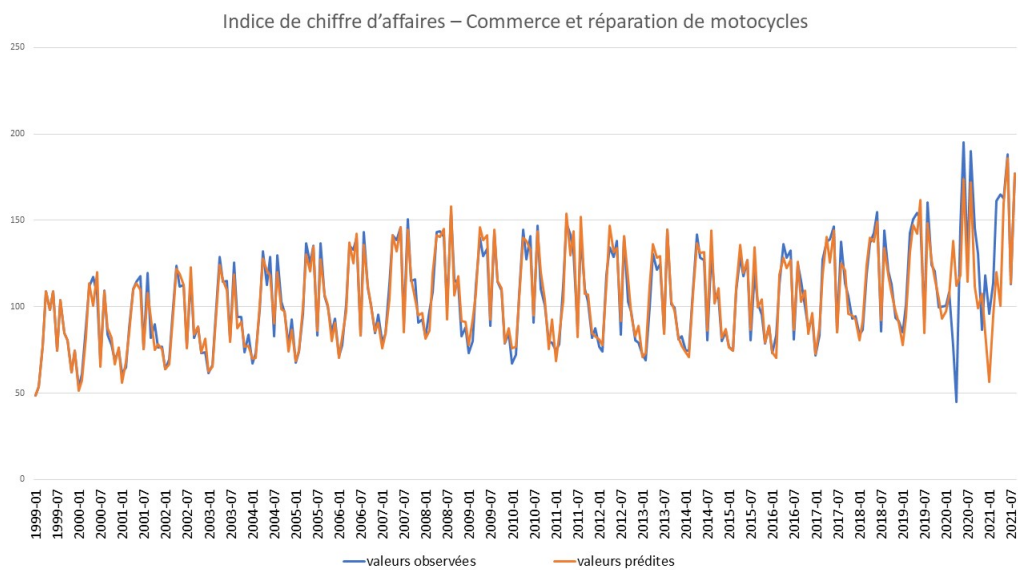
On applique le modèle (modèle 2) pour la prévision des douze dernières données, alors on trouve les valeurs suivantes :

gretl : prévisions

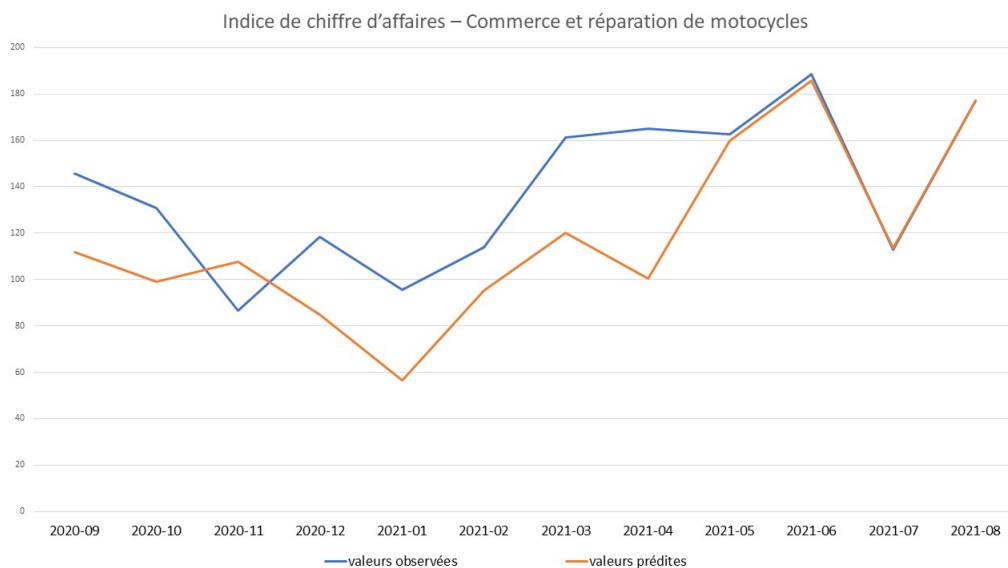
Pour l'intervalle de confiance de 95%, $z(0,025) = 1,96$

	vl	prédiction	éc. type	intervalle de 95%
2020:09		111,658	8,76647	94,4761 - 128,840
2020:10		99,1679	10,0023	79,5638 - 118,772
2020:11		107,502	10,0023	87,8978 - 127,106
2020:12		84,9213	10,0023	65,3171 - 104,525
2021:01		56,5479	10,1017	36,7489 - 76,3468
2021:02		95,2098	10,1017	75,4108 - 115,009
2021:03		119,933	10,4266	99,4975 - 140,369
2021:04		100,447	10,4266	80,0110 - 120,883
2021:05		159,777	10,4266	139,341 - 180,213
2021:06		185,845	10,7735	164,730 - 206,961
2021:07		113,540	11,2357	91,5184 - 135,562
2021:08		177,112	11,2357	155,091 - 199,134

Voilà le graphique comportant les observations et les prévisions :



Si on zoome sur les 12 dernières mois on trouve :



on calculera le MAPE sur ces 12 dernières données, on obtient alors MAPE=18.8%.

Autres modèles : j'ai trouvé plusieurs modèles satisfaisants du point de vue statistique, comme :

$$(1 - B)(1 - \varphi_{24}B^{24})X_t = \mu + (1 - \theta_1B - \theta_2B^2)(1 - \Theta_1B^{12})\epsilon_t \quad (5)$$

avec AIC= 1933,710 et MAPE = 6.3% .

$$(1 - \Phi_1B^{12})(1 - B)(1 - \varphi_{12}B^{12})X_t = \mu + (1 - \theta_1B - \theta_2B^2)\epsilon_t \quad (6)$$

Mais mon modèle 2 c'est le meilleur car AIC et MAPE plus petit.

Remarque : Vous pouvez voir les sorties gretl pour ces deux modèle dans la partie Annexe.

avec AIC = 1924,704 et MAPE = 6.21%

5 Conclusion

La meilleure méthode de l'approche lissage (Holt Winters multiplicative avec MAPE = 5.63%) nous donne une prévision des douze dernières données avec MAPE= 13.5% tandis que la meilleure méthode de l'approche Box-Jenkins (modèle 2 avec MAPE=5.66%) nous donne une prévision des douze dernières données avec MAPE= 18.8%, donc on peut conclure que la meilleure méthode c'est le Holt Winters multiplicative.


Voilà le tableau final qui donne les prévisions par la meilleure méthode de lissage (Holt Winters multiplicative) et par le meilleur modèle de Box et Jenkins (modèle 2) :

	valeurs observées	valeurs prédites (Holt Winters multiplicative)	valeurs prédites (Box et Jenkins modèle 2)
2020-09	145,77	132,5431	111,658
2020-10	130,79	126,27248	99,1679
2020-11	86,62	104,85115	107,502
2020-12	118,43	105,66878	84,9213
2021-01	95,72	100,75436	56,5479
2021-02	113,87	109,12779	95,2098
2021-03	161,22	106,15682	119,933
2021-04	164,94	91,461202	100,447
2021-05	162,67	179,40744	159,777
2021-06	188,36	205,66617	185,845
2021-07	113,01	120,78086	113,54
2021-08	176,93	182,81132	177,112


6 Annexe

les sorties pour

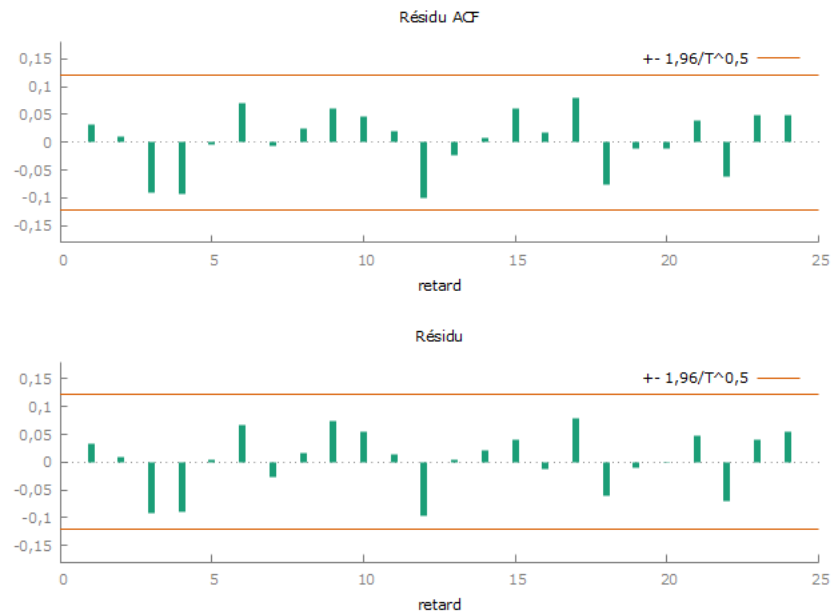
$$(1 - B)(1 - \varphi_{24}B^{24})X_t = \mu + (1 - \theta_1B - \theta_2B^2)(1 - \Theta_1B^{12})\epsilon_t$$

 gretl : modèle 10

Fichier	Édition	Tests	Enregistrer	Graphiques	Analyse	LaTeX
Évaluations de la fonction : 137						
Évaluations du gradient : 36						
Modèle 10: ARIMA, utilisant les observations 1999:02-2020:08 (T = 259)						
Estimation par AS 197 (MV exact)						
Variable dépendante: (1-L) vl						
Écart-types basés sur la matrice hessienne						
	coefficient	éc. type	z	p. critique		
const	0,156360	0,0634202	2,465	0,0137	**	
phi_24	0,906743	0,0298963	30,33	4,67e-202	***	
theta_1	-0,409779	0,0489431	-8,373	5,64e-017	***	
theta_2	-0,590221	0,0487492	-12,11	9,66e-034	***	
Theta_1	0,939907	0,0350877	26,79	4,53e-158	***	
Moyenne var. dép.	0,544479	Éc. type var. dép.	28,17018			
Moyenne innovations	0,135870	Éc. type innovations	9,354845			
R2	0,874974	R2 ajusté	0,873503			
Log de vraisemblance	-960,8551	Critère d'Akaike	1933,710			
Critère de Schwarz	1955,051	Hannan-Quinn	1942,291			

 gretl : corrélogramme

RETARD	ACF	PACF	Q	[p. critique]
1	0,0322	0,0322		
2	0,0099	0,0088		
3	-0,0904	-0,0911		
4	-0,0935	-0,0886		
5	-0,0036	0,0036	4,7778	[0,029]
6	0,0711	0,0663	6,1274	[0,047]
7	-0,0056	-0,0258	6,1358	[0,105]
8	0,0247	0,0154	6,3006	[0,178]
9	0,0611	0,0740	7,3085	[0,199]
10	0,0476	0,0549	7,9238	[0,244]
11	0,0189	0,0149	8,0207	[0,331]
12	-0,1011	-0,0963	10,8158	[0,212]
13	-0,0225	0,0048	10,9548	[0,279]
14	0,0069	0,0215	10,9679	[0,360]
15	0,0618	0,0402	12,0244	[0,362]
16	0,0171	-0,0129	12,1054	[0,437]
17	0,0813	0,0796	13,9515	[0,377]
18	-0,0753	-0,0606	15,5436	[0,342]
19	-0,0119	-0,0108	15,5833	[0,410]
20	-0,0119	-0,0007	15,6235	[0,480]
21	0,0383	0,0485	16,0400	[0,521]
22	-0,0618	-0,0711	17,1309	[0,514]
23	0,0497	0,0412	17,8384	[0,533]
24	0,0489	0,0539	18,5274	[0,553]




Note : * indique un résidu supérieur à 2,5 fois l'écart-type
 Statistiques sur la qualité de la prévision utilisant 259 observations

Moyenne Erreur	0,13587
Moyenne Erreur Carrée (racine)	9,3548
Moyenne Erreur Absolue	6,1542
Moyenne Pourcentage Erreur	-0,78239
Moyenne Pourcentage Erreur Absolue	6,301
U2 de Theil	0,34335

Les sorties pour

$$(1 - \Phi_1 B^{12})(1 - B)(1 - \varphi_{12} B^{12})X_t = \mu + (1 - \theta_1 B - \theta_2 B^2)\epsilon_t$$

 gretl : modèle 12

Fichier	Édition	Tests	Enregistrer	Graphiques	Analyse	LaTeX
Évaluations de la fonction : 132						
Évaluations du gradient : 36						
Modèle 12: ARIMA, utilisant les observations 1999:02-2020:08 (T = 259)						
Estimation par AS 197 (MV exact)						
Variable dépendante: (1-L) vl						
Écart-types basés sur la matrice hessienne						
	coefficient	éc. type	z	p. critique		
const	0,158403	0,0535349	2,959	0,0031	***	
phi_12	0,969472	0,0114557	84,63	0,0000	***	
Phi_1	-0,275763	0,0887692	-3,107	0,0019	***	
theta_1	-0,408346	0,0485308	-8,414	3,96e-017	***	
theta_2	-0,591654	0,0483876	-12,23	2,22e-034	***	
Moyenne var. dép.	0,544479	Éc. type var. dép.	28,17018			
Moyenne innovations	0,137867	Éc. type innovations	9,186057			
R2	0,879447	R2 ajusté	0,878029			
Log de vraisemblance	-956,3520	Critère d'Akaike	1924,704			
Critère de Schwarz	1946,045	Hannan-Quinn	1933,284			

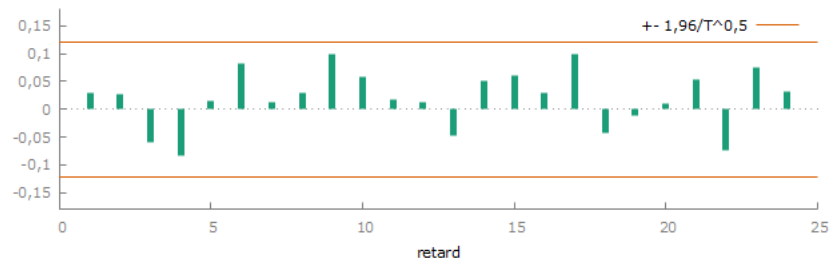


Fonction d'autocorrélation résiduelle

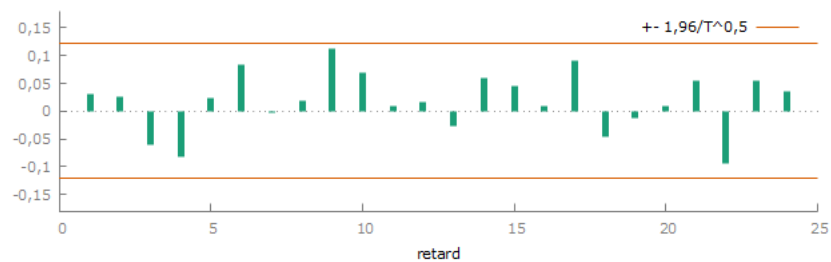
***, **, * indiquent une significativité au seuil de 1%, 5% et 10% sur la base d'écart-type $1/T^{0,5}$

RETARD	ACF	PACF	Q	[p. critique]
1	0,0299	0,0299		
2	0,0260	0,0251		
3	-0,0602	-0,0618		
4	-0,0846	-0,0821		
5	0,0158	0,0240	3,3329	[0,068]
6	0,0828	0,0838	5,1631	[0,076]
7	0,0126	-0,0031	5,2054	[0,157]
8	0,0286	0,0184	5,4264	[0,246]
9	0,0991	0,1125 *	8,0814	[0,152]
10	0,0596	0,0692	9,0457	[0,171]
11	0,0179	0,0084	9,1333	[0,243]
12	0,0115	0,0164	9,1697	[0,328]
13	-0,0481	-0,0259	9,8064	[0,366]
14	0,0524	0,0595	10,5648	[0,392]
15	0,0608	0,0460	11,5892	[0,395]
16	0,0297	0,0090	11,8348	[0,459]
17	0,1000	0,0919	14,6260	[0,331]
18	-0,0418	-0,0455	15,1167	[0,370]
19	-0,0103	-0,0129	15,1465	[0,441]
20	0,0092	0,0095	15,1707	[0,512]
21	0,0537	0,0543	15,9906	[0,525]
22	-0,0731	-0,0935	17,5150	[0,488]
23	0,0760	0,0549	19,1696	[0,446]
24	0,0332	0,0365	19,4860	[0,490]

Résidu ACF



Résidu



Note : * indique un résidu supérieur à 2,5 fois l'écart-type
Statistiques sur la qualité de la prévision utilisant 259 observations

Moyenne Erreur	0,13787
Moyenne Erreur Carrée (racine)	9,1861
Moyenne Erreur Absolue	6,0664
Moyenne Pourcentage Erreur	-0,76433
Moyenne Pourcentage Erreur Absolue	6,2164
U2 de Theil	0,34074