

**La prévision du dernier point d'intérim de la série Dares.**

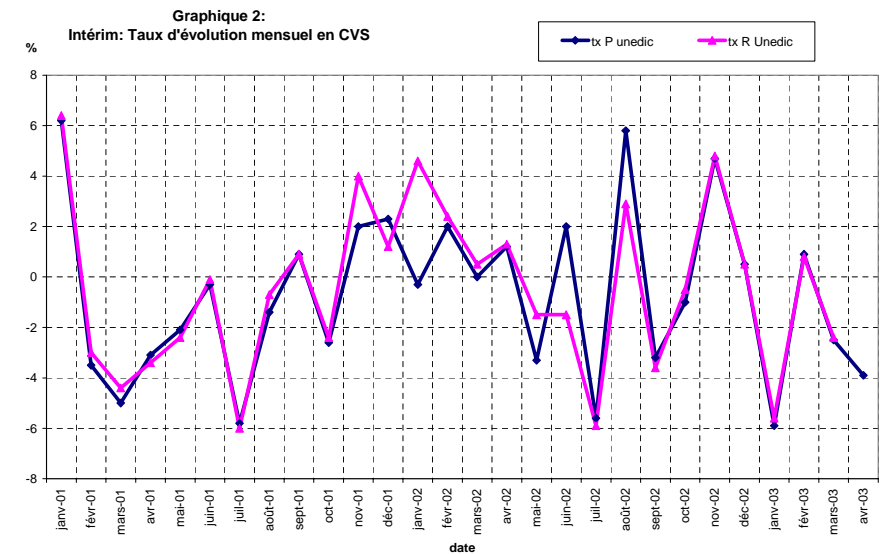
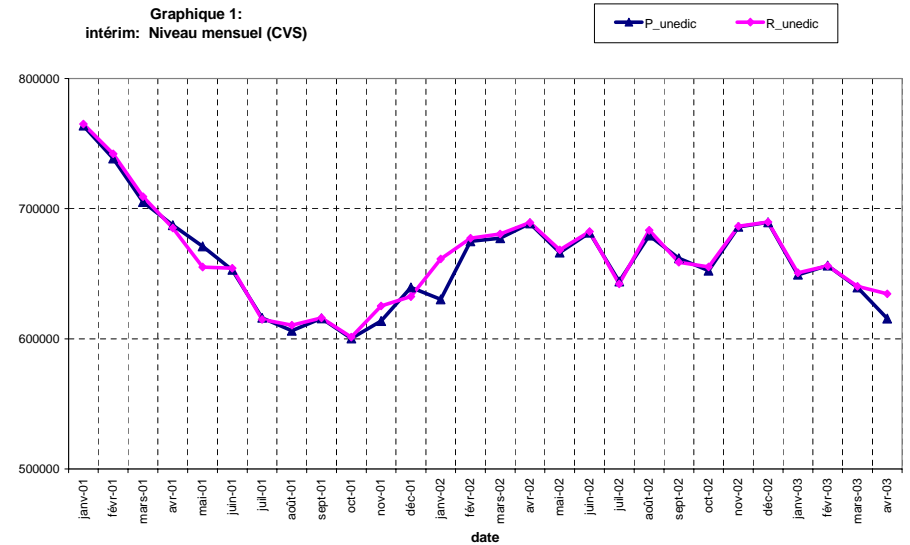
L'objet principal de l'étude est de construire des estimations mensuelles de l'emploi intérimaire. Dans un premier temps, nous avons essayé d'utiliser l'indicateur mensuel Unédic. Suite à l'abandon de l'indicateur Unédic, nous cherchons à mettre en évidence un modèle qui ajuste au mieux la série d'intérim par la DARES.

Ensuite, nous nous intéressons à la prévision du dernier point d'intérim à la publication de DARES.

**1) L'abandonnement les données Unédic**

Nous renoncerons à utiliser les données provisoires de Unédic. Premièrement parce que l'écart entre la variation des séries provisoires et rectifiées est très important. Et deuxièmement, les séries provisoires de l'Unédic et la série de Dares ne sont pas très corrélées. Ceci peut être constater graphiquement ou par la procédure « proc corr » de SAS.

Nous étudions les graphiques des séries provisoires et rectifiées de l'Unédic en utilisant les données mensuelles en CVS. Nous distinguerons deux types graphiques étudiés : Graphe des séries en niveau ou graphe des séries en évolution.



Source : Unédic

Visuellement, nous trouvons que le graphique 1 est mieux corrélé que le graphique 2.

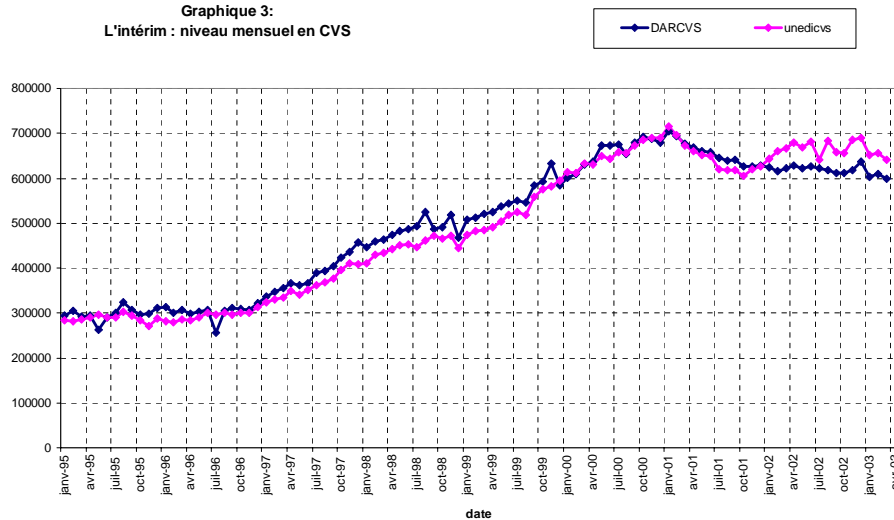
Ensuite, nous étudions les séries d'intérim provenant de la Dares (cec5jbrut). Avec lesquelles nous essayons de faire une liaison avec la série provisoire de l'Unédic. Nous pourrions les comparer graphiquement et aussi en utilisant la procédure « proc corr » de SAS pour obtenir des résultats. Ainsi nous pouvons distinguer 4 cas principaux :

- Le niveau de Séries Brute
- Le niveau de CVS (corrigé les variables saisonnières)
- L'évolution de Séries Brutes
- L'évolution de CVS

La publication de l'Unédic ne dispose que de données en CVS. Pour une question de simplicité, nous devons étudier aussi la série intérim de la Dares en CVS. Dans ce cas, nous devons utiliser la « procédure X11 » de SAS pour obtenir une série d'intérim en CVS à partir de la série brut (cec5jbrut). Il a deux cas principaux et nous utiliserons la « procédure corr » de SAS pour déterminer quel cas qui est le meilleur

Tout d'abord, nous étudions le niveau d'intérim mensuel en cvs.

Graphique 3:  
L'intérim : niveau mensuel en CVS



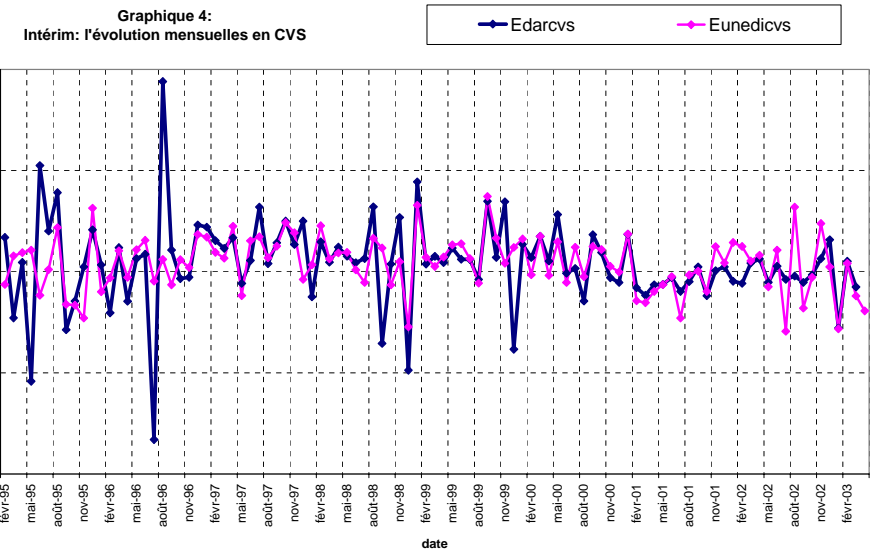
Pearson Correlation Coefficients  
Prob > |r| under H0: Rho=0  
Number of Observations

DARCVS	
unedicvs	0.98454
unedicvs	<.0001
	100

Spearman Correlation Coefficients  
Prob > |r| under H0: Rho=0  
Number of Observations

DARCVS	
unedicvs	0.94770
unedicvs	<.0001
	100

Puis nous étudions l'évolution de deux séries en CVS, nous obtenons que :



Pearson Correlation Coefficients  
Prob > |r| under H0: Rho=0  
Number of Observations

E_D	
E_U	0.41392
	<.0001
	99

Spearman Correlation Coefficients  
Prob > |r| under H0: Rho=0  
Number of Observations

E_D	
E_U	0.55293
	<.0001
	99

Nous constatons que le niveau d'intérim est plus corrélé que l'évolution d'intérim. Nous nous intéressons donc seulement au cas de niveau en cvs.

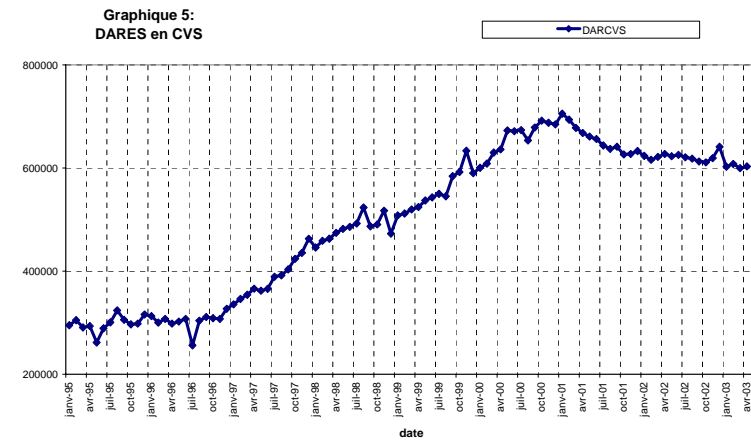
Nous allons étudier maintenant la liaison entre les deux séries en utilisant la « procédure Reg » Deux problèmes apparaissent lorsqu'on teste le résidu par la « procédure ARIMA » de SAS. Le premier est un problème d'autocorrélation et le second, un problème d'intégralité. Pour résoudre le problème d'intégralité, il suffit de différencier une fois (c'est à dire  $\Delta I_{Dares} = I_{Dares_t} - I_{Dares_{t-1}}$ ) les deux séries. Même si le problème d'intégralité est résolu, il reste toujours celui de l'autocorrélation. Pour cela, nous étudions les résultats de la « procédure autoreg » de SAS. Cependant, le test de normalité n'est pas validé car le P- value de test de Shapiro - Wilk est 0,0002. Nous en concluons que ce modèle n'est pas admis car les résidus ne suivent pas une loi normale. Nous pouvons penser que les séries d'intérim de la Dares ont des points atypiques (voir le graphique 4 « d'évolution mensuelle en CVS » ). Ceci montre que le modèle ne suit pas une loi de normale.

Nous allons maintenant essayer le modèle GRACH avec la « procédure autoreg » de SAS pour réduire les points atypiques de séries provenant de la Dares. Et nous essayerons de traiter aussi ces points comme des indicateurs. Puis nous étudierons une liaison entre les deux séries (Provisoire \_ Unédic et actuel \_ Dares). il y a toujours un problème persistant (problème d'autocorrélation ...). De plus, le dernier point de la série provisoire de l'Unédic est différent de la série rectifiée de Unédic.

Ceci nous permet de conclure que nous n'utiliserons pas les données de l'Unédic. Et donc nous n'étudions que la série intérim de la Dares.

## 2) Etude de la série d'intérim de Dares (uni varié)

Nous étudions la série d'intérim mensuelle en cvs.



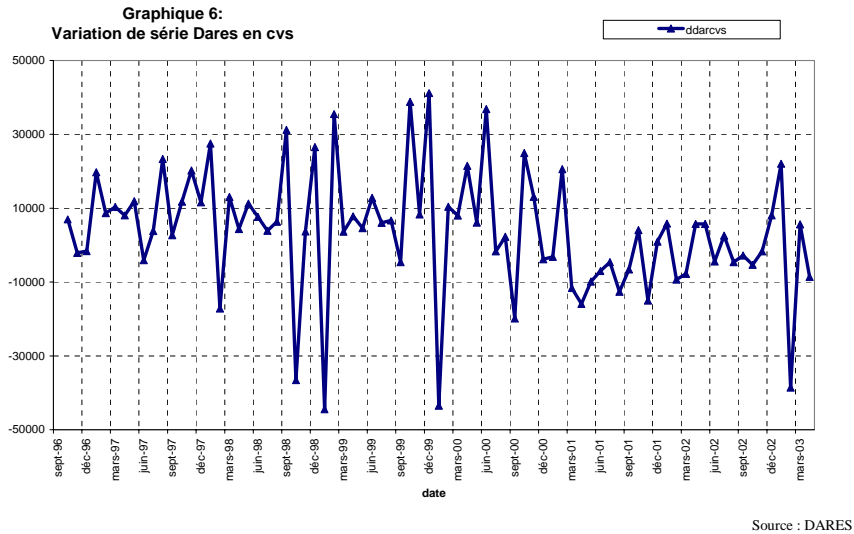
Source : Dares

Sur le graphique 5, la série intérim de la Dares du point de juillet 1996 est très éloigné des autres. Il semble y avoir un point atypique. Dans l'immédiat pour remédier à ce problème, nous tronquons la série en août 1996.

Ensuite, la graphique nous montre que la séries d'intérim en niveau CVS n'est pas stationnaire car l'intérim évolue en fonction du temps. Une façon de régler l'hypothèse de stationnarité d'un processus est de travailler sur le logarithme de la série ou sur la série dite « différenciée » c'est-à-dire :  $\Delta I = I_t - I_{t-1}$ .

Ici, la série d'intérim différenciée une fois est stationnaire d'après le test de DICKEY-FULLER

Nous obtenons le graphique ci-dessous :



Avec ce série, nous utilisons plusieurs modèles par différentes méthodes comme Méthodes Winters, Méthodes de lissage exponentielle et ARIMA. Parmi plusieurs modèles que nous avons étudié. Nous garderons 4 modèles qui sont plus intéressant comme :

**Modèle 1 :**  $(1-\phi_1L)(1-\phi_8L^8) dDares_{cvs} = (1-\theta_6L^6) \epsilon_t$

Avec L= en retard

et dDares = la variation de la série Dares.

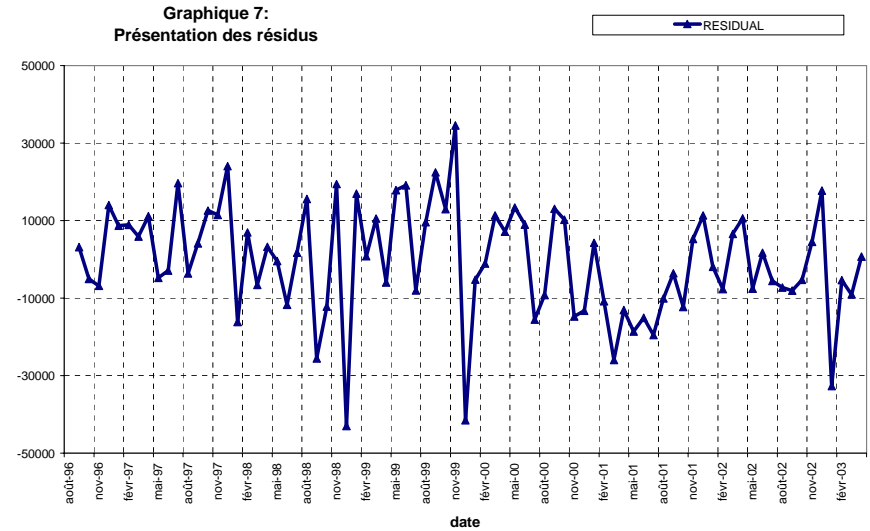
• **Validation du modèle 1**

Il faut vérifier que les résidus sont des bruits blancs, que les paramètres sont significatifs.

**a. Test sur les résidus**

Pour effectuer les tests de validation sur les résidus, on a besoin des résidus estimés que l'on obtient en utilisant la commande forecast.

• **Représentation graphique des résidus**



Pour vérifier les éventuelles autocorrélations des résidus, nous avons procédé à une modélisation ARIMA des résidus :

The ARIMA Procedure  
 Name of Variable = RESIDUAL  
 Mean of Working Series = -352.327  
 Standard Deviation = 14253.45  
 Number of Observations = 80

Lag	Covariance	Correlation	Autocorrelations																			
			-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9
0	203160745	1.00000	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1	-3536939	-.01741	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2	-4597185	-.02263	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	27750902	0.13660	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	21075055	0.10374	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	12361328	0.06085	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	1817085	0.00894	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
7	28200872	0.13881	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
8	8132926	0.04003	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
9	-32231278	-.15865	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
10	-6694087	-.03295	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
11	-6952618	-.03422	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
12	11988593	0.05901	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
13	-26151900	-.12873	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
14	-18971491	-.09338	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
15	30649063	0.15086	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
16	-3602196	-.01773	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
17	-32878220	-.16183	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
18	4426235	0.02179	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
19	22939810	0.11291	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
20	-12814153	-.06307	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

"," marks two standard errors

		Inverse Autocorrelations																				
Lag	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
1	0.07645												**	.	.	.	.	.	.	.	.	.
2	-0.02706												*	.	.	.	.	.	.	.	.	.
3	-0.08824												**	.	.	.	.	.	.	.	.	.
4	-0.07219												*	.	.	.	.	.	.	.	.	.
5	-0.13675												***	.	.	.	.	.	.	.	.	.
6	-0.13822												**	.	.	.	.	.	.	.	.	.
7	-0.13243												*	.	.	.	.	.	.	.	.	.
8	0.01847												.	.	.	.	.	.	.	.	.	.
9	0.17323												***	.	.	.	.	.	.	.	.	.
10	0.04292												*	.	.	.	.	.	.	.	.	.
11	0.08997												**	.	.	.	.	.	.	.	.	.
12	-0.01214												.	.	.	.	.	.	.	.	.	.
13	0.07901												*	.	.	.	.	.	.	.	.	.
14	0.03367												.	.	.	.	.	.	.	.	.	.
15	-0.13974												***	.	.	.	.	.	.	.	.	.
16	-0.04464												*	.	.	.	.	.	.	.	.	.
17	0.08410												**	.	.	.	.	.	.	.	.	.
18	0.00165												.	.	.	.	.	.	.	.	.	.
19	-0.06533												*	.	.	.	.	.	.	.	.	.
20	0.03816												.	.	.	.	.	.	.	.	.	.

		Partial Autocorrelations																				
Lag	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
1	-0.01741												.	.	.	.	.	.	.	.	.	.
2	-0.02294												***	.	.	.	.	.	.	.	.	.
3	0.13591												*	.	.	.	.	.	.	.	.	.
4	0.10976												**	.	.	.	.	.	.	.	.	.
5	0.07373												.	.	.	.	.	.	.	.	.	.
6	-0.00100												.	.	.	.	.	.	.	.	.	.
7	0.11707												**	.	.	.	.	.	.	.	.	.
8	0.02061												.	.	.	.	.	.	.	.	.	.
9	-0.17262												***	.	.	.	.	.	.	.	.	.
10	-0.08945												*	.	.	.	.	.	.	.	.	.
11	-0.08950												**	.	.	.	.	.	.	.	.	.
12	0.07514												.	.	.	.	.	.	.	.	.	.
13	-0.09179												*	.	.	.	.	.	.	.	.	.
14	-0.07143												.	.	.	.	.	.	.	.	.	.
15	0.15826												***	.	.	.	.	.	.	.	.	.
16	0.07355												*	.	.	.	.	.	.	.	.	.
17	-0.10548												**	.	.	.	.	.	.	.	.	.
18	-0.01059												.	.	.	.	.	.	.	.	.	.
19	0.08004												*	.	.	.	.	.	.	.	.	.
20	-0.04447												.	.	.	.	.	.	.	.	.	.

To Lag	Chi - Square	DF	Pr > Chi Sq	Autocorrelations																			
6	2.92	6	0.8191	-0.017	-0.023	0.137	0.104	0.061	0.009														
12	7.67	12	0.8104	0.139	0.040	-0.159	-0.033	-0.034	0.059														
18	15.27	18	0.6437	-0.129	-0.093	0.151	-0.018	-0.162	0.022														

Le diagramme des autocorrélations partielles montre que les résidus ne sont pas autocorrélés. De plus, le test du bruit blanc (Ljung Box) confirme l'absence d'autocorrélation des résidus estimés. Les résidus sont des bruits blancs (BB).

• **Test de l'homoscédasticité** (L'égalité des variances des résidus)

On test l'homoscédasticité sur les résidus grâce au test ARCH. On construit alors une régression entre les résidus au carré et les résidus au carrés décalés jusqu'à l'ordre p. Douze tests seront effectués selon les valeurs de p : p=1, p=2, ..., p=12.

Supposons l'hypothèse nulle à tester qui est la nullité conjointe des coefficients (sauf la constante)

$$\epsilon_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + \dots + a_p \epsilon_{t-p}^2 + Y_t \quad \text{pour } p=1,2,\dots, 12$$

Model : m1					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	4.710866E17	4.710866E17	4.12	0.0463
Error	66	7.537585E18	1.142058E17		
Corrected Total	67	8.008671E18			

Model : m2					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	5.992015E17	2.996007E17	2.63	0.0799
Error	65	7.40947E18	1.139918E17		
Corrected Total	67	8.008671E18			

Model : m3					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	6.987502E17	2.329167E17	2.04	0.1172
Error	64	7.309921E18	1.142175E17		
Corrected Total	67	8.008671E18			

Model : m4					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	9.399871E17	2.349968E17	2.09	0.0920
Error	63	7.068684E18	1.122013E17		
Corrected Total	67	8.008671E18			
res42	1	-0.19119	0.13039	-1.47	0.1475

Model : m5					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	9.52425E17	1.90485E17	1.67	0.1543
Error	62	7.056246E18	1.138104E17		
Corrected Total	67	8.008671E18			

Model : m6					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	9.533386E17	1.588898E17	1.37	0.2396
Error	61	7.055332E18	1.156612E17		
Corrected Total	67	8.008671E18			

Model : m7					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	1.008017E18	1.440024E17	1.23	0.2987
Error	60	7.000654E18	1.166776E17		
Corrected Total	67	8.008671E18			

Model : m8					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	1.027102E18	1.283878E17	1.08	0.3863
Error	59	6.981569E18	1.183317E17		
Corrected Total	67	8.008671E18			

Model : m9					
Dependent Variable: res02					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	1.036017E18	1.15113E17	0.96	0.4840
Error	58	6.972654E18	1.202182E17		
Corrected Total	67	8.008671E18			

Model : m10  
Dependent Variable: res02  
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	1.08448E18	1.08448E17	0.89	0.5455
Error	57	6.924191E18	1.21477E17		
Corrected Total	67	8.008671E18			

Model : m11  
Dependent Variable: res02  
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	1.702346E18	1.547587E17	1.37	0.2108
Error	56	6.306325E18	1.12613E17		
Corrected Total	67	8.008671E18			

Model : m12  
Dependent Variable: res02  
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	2.232431E18	1.860359E17	1.77	0.0765
Error	55	5.77624E18	1.050225E17		
Corrected Total	67	8.008671E18			

Il n'y a pas de problème d'hétéroscédasticité au seuil de 5% même si le test de ARCH est validé au seuil de 4% pour le modèle m1. Nous pouvons toujours accepter cette hypothèse.

Le teste de normalité: Le normalité des erreurs: (pour vérifier dans le cas d'indépendance et d'homoscédasticité des erreurs), nous devons procéder à une modélisation UNIVARIATE des résidus :

The UNIVARIATE Procedure  
Variable: RESIDUAL (Residual: Actual - Forecast)

Moments			
N	80	Sum Weights	80
Mean	-352.32683	Sum Observations	-28186.147
Std Deviation	14343.3748	Variance	205732400
Skewness	-0.4880802	Kurtosis	0.58806297
Uncorrected SS	1.62628E10	Corrected SS	1.62529E10
Coeff Variation	-4071.0424	Std Error Mean	1603.63805

Basic Statistical Measures

Location	Mean	Std Deviation	Variability
Mean	-352.327	14343	
Median	860.604	205732400	
Mode	.	74818	
		Range	19974
		Interquartile Range	

Tests for Location: Mu0=0

Test	-Statistic-	Pr >  t	Pr >=  M	Pr >=  S
Student's t	-0.2197	0.8267		
Sign	M	2	0.7376	
Signed Rank	S	31	0.8829	

Tests for Normality

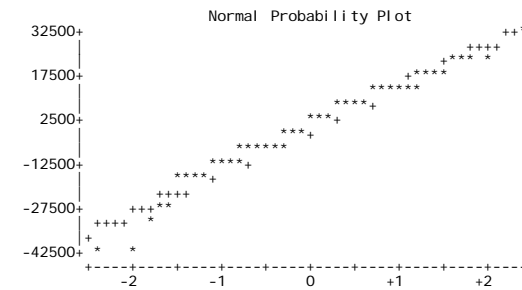
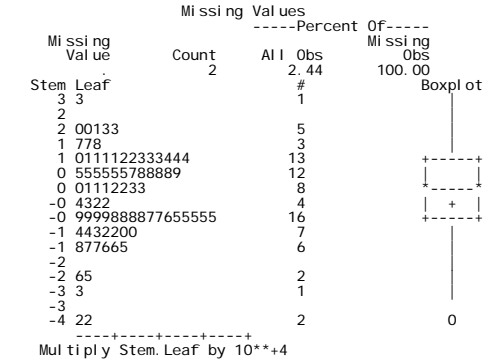
Test	-Statistic-	Pr < W	Pr > D	Pr > W-Sq	Pr > A-Sq
Shapiro-Wilk	W	0.976723	0.1523		
Kolmogorov-Smirnov	D	0.059504	>0.1500		
Cramer-von Mises	A-Sq	0.054101	>0.2500		
Anderson-Darling	A-Sq	0.434496	>0.2500		

Quantiles (Definition 5)

Quantile	Estimate
100% Max	33089.654
99%	32089.654
95%	20691.932
90%	16863.592
75% Q3	10917.213
50% Median	860.604
25% Q1	-9056.692
10%	-16483.457
5%	-25906.194
1%	-41728.439
0% Min	-41728.439

Extreme Observations

Lowest Value	Obs	Highest Value	Obs
-41728.4	29	20114.0	12
-41550.1	41	21269.8	28
-32722.0	78	23196.2	17
-26417.3	56	23476.2	38
-25395.1	26	33089.7	40



Le test de normalité de Shapiro Wilk nous permet d'accepter la normalité des résidus estimés. Et le diagramme « stem and leaf » et le « boxplot » des résidus donnés nous laissent penser que les résidus sont proche d'une loi normale centrée.

En conclusion, on a montré que les résidus estimés sont centrés, sans autocorrélation, mais homoscédastique et normaux. Donc c'est bien des bruits blancs.

### • Test de la significativité des paramètres

Parameter	Estimate	Standard Error	t Value	Pr >  t	Lag
MU	3823.0	2327.5	1.64	0.1046	0
MA1, 1	-0.23660	0.11710	-2.02	0.0469	6
AR1, 1	-0.29963	0.11220	-2.67	0.0093	1
AR2, 1	0.40022	0.11534	3.47	0.0009	8

Le test de student montre que les coefficients sont statistiquement significatifs au seuil de 5% sauf la constante.

**Modèle 2 :** *le modèle est le même que le précédent mais sans la constante*

Nous trouvons que ce modèle est validé lorsque nous acceptons au seuil de 10% de test de ARCH. Les avantages de ce modèle sont qu'il n'y a pas de constante non significative.

**Modèle 3 :** *la méthode de Winters*

Les avantages de ce modèle montre que nous n'avons pas besoin de passer toutes les hypothèses.

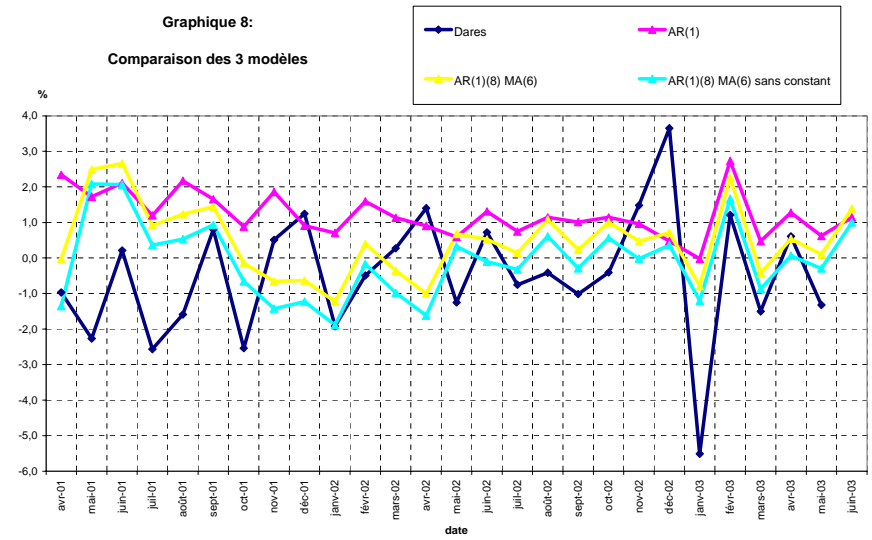
**Modèle 4 :** AR(1) :  $(1-\phi_1L) dDares_{cvs} = \epsilon_t$

Les inconvénients de ce modèle, c'est le diagramme des autocorrélations partielles n'est pas bien vérifié.

**3) Validation ces modèles**

Nous devons faire une simulation en remontant 2 ans en arrière. Puis nous essaierons de faire les Ecarts Quadratiques Moyens (EQM) pour comparer les 3 modèles choisis. Nous trouvons que les deux premières et quatrième modèles montrent des EQM est autour de 2. Par contre le modèle 3 a un EQM autour de 3.9. Dans ce cas, nous étudierons seulement les trois modèles.

Nous faisons un graphique pour comparer les trois modèles :



Nous trouvons que les deux premiers modèles estimés sont toujours supérieures à l'autre. Nous pouvons imaginer c'est parce que ces deux modèles sont très ressemblant. Parmi les trois modèles choisis, je préfère surtout la première modèle car il ressemble bien la série d'intérim de la DARES.

## **Annexe**

**BEULANT Christelle** « Travaux de séries temporelles » de deuxième année de Magistère en 2003, Université de Paris 1 Panthéon - Sorbonne.

**BRESSON Georges** « Econométrie des séries temporelles »

**DEGERINE Serge** « Séries chronologiques »