

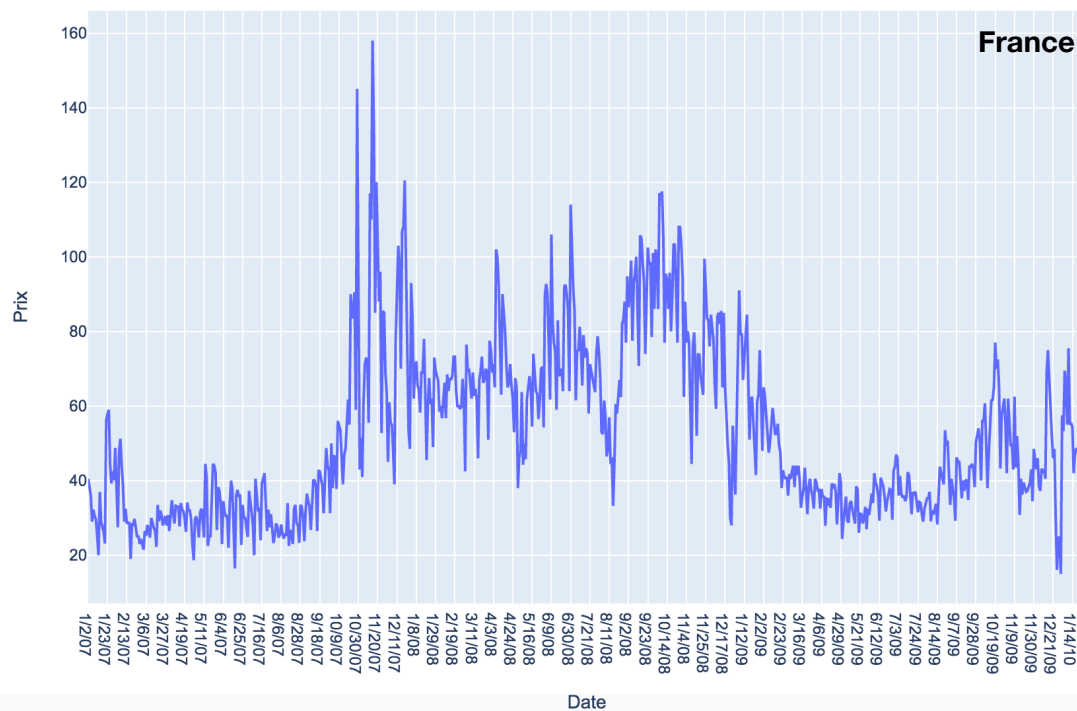
Projet prévision prix de l'électricité

L'objectif est de fournir une prévision des prix de l'électricité dans plusieurs pays sur une base journalière, mensuelle et annuelle. Nous choisirions 2 pays(France et Germany) sur les 5 proposés et nous ferons des analyses sur DAY, MONTH, YEAR.

一. Day

1. France

Séries temporelle de le prix d'électricité en Allemagne et en France



1/2/2007 - 1/28/2010 (Day)

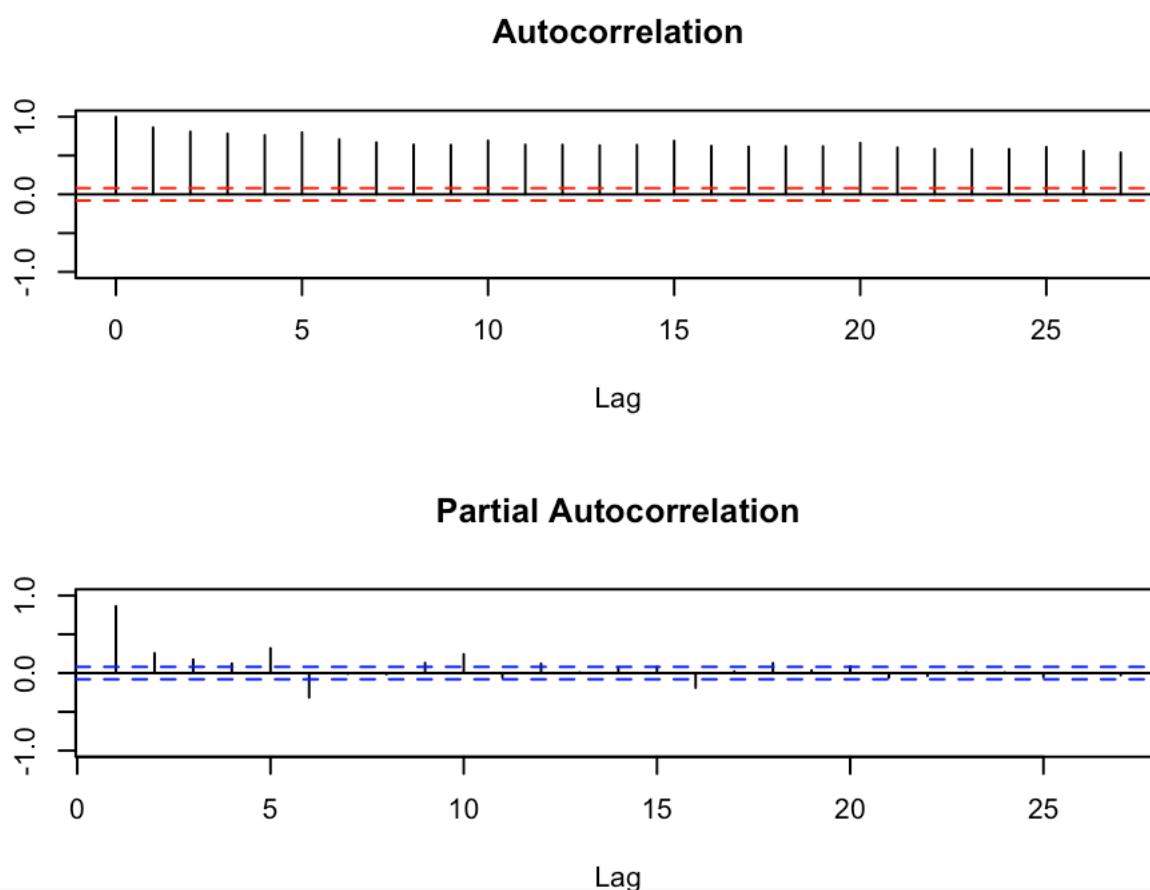
	France
count	776.000000
mean	52.782796
std	22.926363
min	15.000000
25%	34.500000
50%	46.175000
75%	67.500000
max	158.000000

1) Dans le base de donnée, on ne trouve pas les valeurs manquantes et aberrantes.

(Des prix (ou des contrats) de type AHEAD (day-ahead, etc.) sont les prix prévisionnels dans les journées prochaines)

2) Deux façon pour verifier cette série est stationnaire ou pas.

i) ACF et PACF.



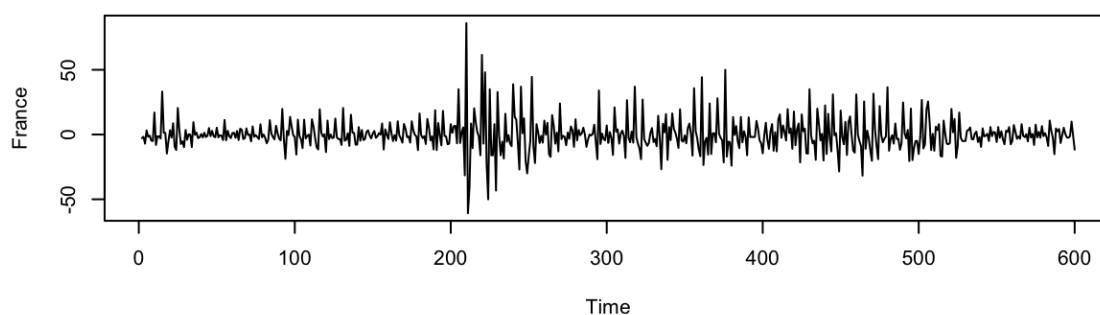
L'ACF montre clairement une très grande persistance. Notez que même si théoriquement la RU devrait donner des autocorrélations qui ne décroissent pas, en pratique elles le font, mais à un rythme très lent. Par contre, notez que pour une tendance déterministe, l'ACF décroît un plus rapidement et se stabilise en général. En tout cas, cette structure suggère une non stationnarité et le symptôme s'apparente à une RU.

ii) RU(racine unitaire)

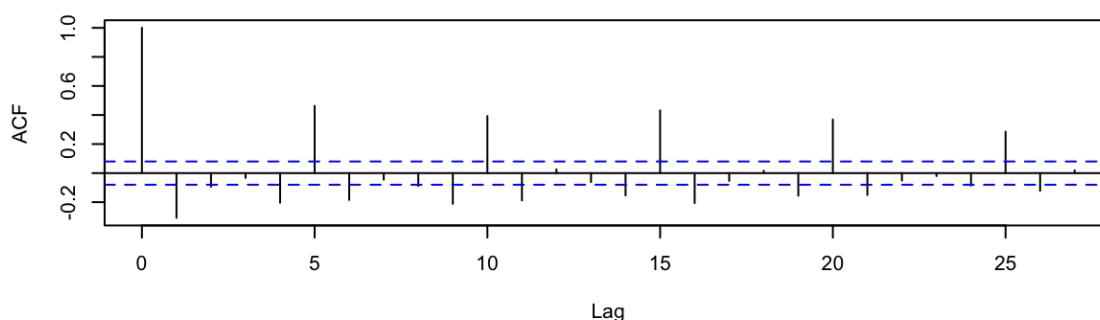
Augmented Dickey-Fuller Test: p-value = 0.2009, du coup, on peut bien dire que cette série est non stationnaire.

3) Différence pour la stabilité de cette série.

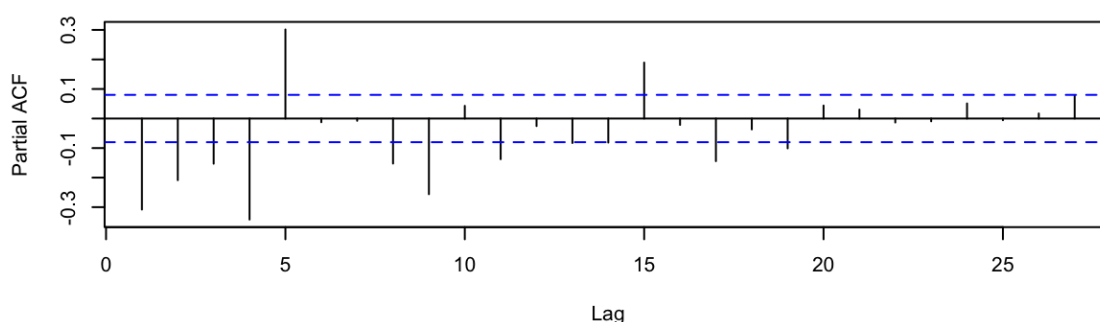
i) Après la différence de premier ordre, on a les graphes de ACF, PACF du différent série temporelle, on peut bien savoir que L'ACF ne montre pas une persistance et il se diminue jusqu'à zéro, par conséquent, on peut bien dire que après la différence de premier ordre, cette série est stationnaire, bien entendu, on va continuer de faire un test de racine unitaire pour confirmer notre jugement.



France



Series disF



ii)

RU(racine unitaire)

Augmented Dickey-Fuller Test: Dickey-Fuller = -12.741, Lag order = 8, p-value = 0.01, du coup, on peut bien rejeter l'hypothèse originelle: non stationnaire, accepter hypothèse alternative, on dit que après la différence de premier ordre, cette série est stationnaire.

4) Vérifiez si la série temporelle après la différence est du bruit blanc ou pas

Box-Ljung test: p-value = 4.063e-14, du coup, on peut bien rejeter l'hypothèse originale et accepter l'hypothèse alternative, la série temporelle après la différence n'est pas le bruit blanc, ensuite on peut faire l'opération prochaine.

5) Choix du modèle

Grâce aux graphes ACF, PAFC du différent série temporelle(procédé 3) et la fonction de R : `auto.arima ()`, on trouve bien un modèle ARMA(5, 1, 5) dont AIC est le plus petit.

```

ARIMA(2,1,2) with drift      : 4639.406
ARIMA(0,1,0) with drift     : 4761.813
ARIMA(1,1,0) with drift     : 4704.961
ARIMA(0,1,1) with drift     : 4661.583
ARIMA(0,1,0)                : 4759.801
ARIMA(1,1,2) with drift     : 4638.029
ARIMA(0,1,2) with drift     : 4649.329
ARIMA(1,1,1) with drift     : 4647.058
ARIMA(1,1,3) with drift     : 4650.777
ARIMA(0,1,3) with drift     : 4648.512
ARIMA(2,1,1) with drift     : 4648.44
ARIMA(2,1,3) with drift     : 4590.829
ARIMA(3,1,3) with drift     : 4634.637
ARIMA(2,1,4) with drift     : 4552.872
ARIMA(1,1,4) with drift     : 4618.077
ARIMA(3,1,4) with drift     : 4546.673
ARIMA(4,1,4) with drift     : 4511.164
ARIMA(4,1,3) with drift     : 4541.53
ARIMA(5,1,4) with drift     : 4530.956
ARIMA(4,1,5) with drift     : 4493.259
ARIMA(3,1,5) with drift     : 4566.616
ARIMA(5,1,5) with drift     : 4482.676
ARIMA(5,1,5)                : 4480.593
ARIMA(4,1,5)                : 4491.208
ARIMA(5,1,4)                : 4528.893
ARIMA(4,1,4)                : 4509.124

```

Now re-fitting the best model(s) without approximations...

ARIMA(5,1,5) : 4483.435

Best model: ARIMA(5,1,5)

Series: x1

ARIMA(5,1,5)

Coefficients:

	ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3	
	-0.2482	-0.2310	-0.2469	-0.2345	0.6608	-0.0694	-0.0190	0.0423	
s.e.	0.0699	0.0707	0.0648	0.0575	0.0529	0.0707	0.0627	0.0509	
	ma4	ma5							
	-0.0193	-0.5926							
s.e.	0.0531	0.0457							

sigma^2 estimated as 101.7: log likelihood=-2230.49

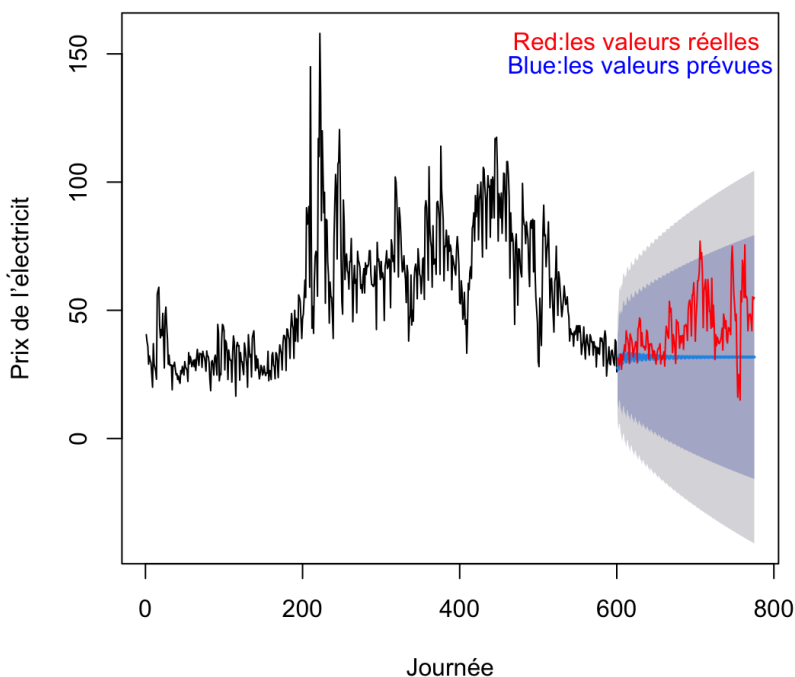
AIC=4482.98 AICc=4483.43 BIC=4531.33

6) Tester si le résidu du modèle est du bruit blanc ou pas

Box-Ljung test: X-squared = 0.61541, df = 1, p-value = 0.4328, on peut bien savoir que le résidu est bruit blanc, du coup, Il n'y a pas d'informations utiles dans les résidus et ils suivent une distribution normale, notre modèle ARMA(5, 1, 5) est bon.

7) Prévision

Forecasts from ARIMA(5,1,5)



Bleu représente les valeurs prévues, rouge représente les valeurs réelles, on peut bien voir que les valeurs prévues se trouvent bien dans 95% d'intervalle de confiance (ombres sombres)

8) Qualité de la prédiction sur l'échantillon restant

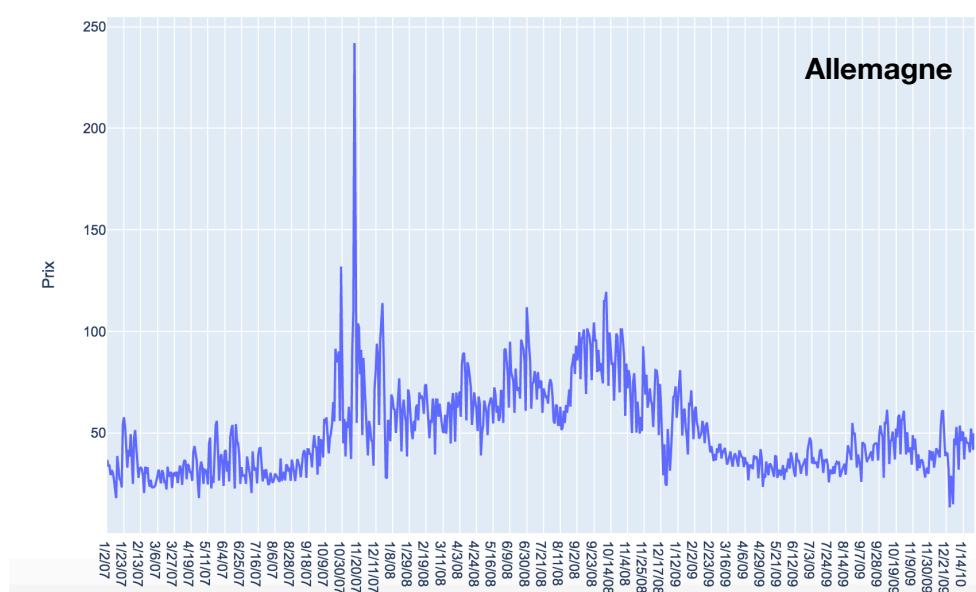
i) L'erreur quadratique moyenne: 12.38

ii) La proportion des erreurs quadratiques moyennes: 25.71%.

(Ici, il semble que la variation des valeurs prévues est trop différente de la variation des valeurs réelles, pourtant, il faut bien noter que les valeurs prévues se trouvent bien dans 95% d'intervalle de confiance (ombres sombres) et le résidu du modèle est bruit blanc, du coup, notre prévision est bonne. L'autre part, on peut bien voir que les première 200 journée dont des prix de l'électricité a l'aire stationnaire, après 200 jusqu'à 550 il y a une grande variation, ensuite après 550 des prix de l'électricité se diminue au niveau de fluctuation de prix d'origine, par conséquent, on peut savoir qu'il y avait une choses significative qui s'est passé durant 200 et 550 journée, qui a provoqué la grande variation des prix de l'électricité et cette période est la crise économique de 2008 (le date de cette base de donnée est de 01/02/2007 à 01/28/2010). En gros, l'explication pour le fait que la variation des valeurs prévues est différente de la variation des valeurs réelles: la deuxième vague de crise économique en 2008 et peu nombre d'observation avec les jours de prévision plus longs(175 journées), mais notre prévision est bonne).

2. Germany

Séries temporelle de le prix d'électricité en Allemagne et en France



1/2/2007 - 1/28/2010 (Day)	
	Germany
count	776.000000
mean	50.851707
std	21.917824
min	13.500000
25%	34.250000
50%	45.000000
75%	63.750000
max	242.000000

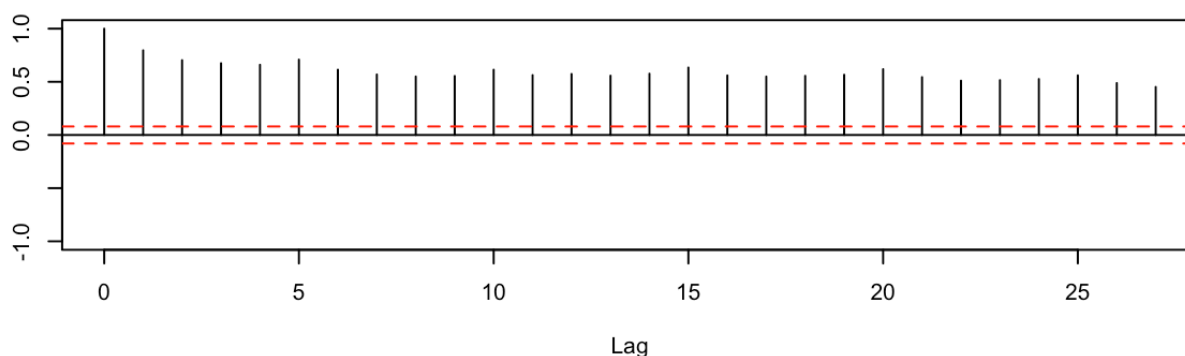
1) Dans le base de donnée, on ne trouve pas les valeurs manquantes et aberrantes.

(Des prix (ou des contrats) de type AHEAD (day-ahead, etc.) sont les prix prévisionnels dans les journées prochaines)

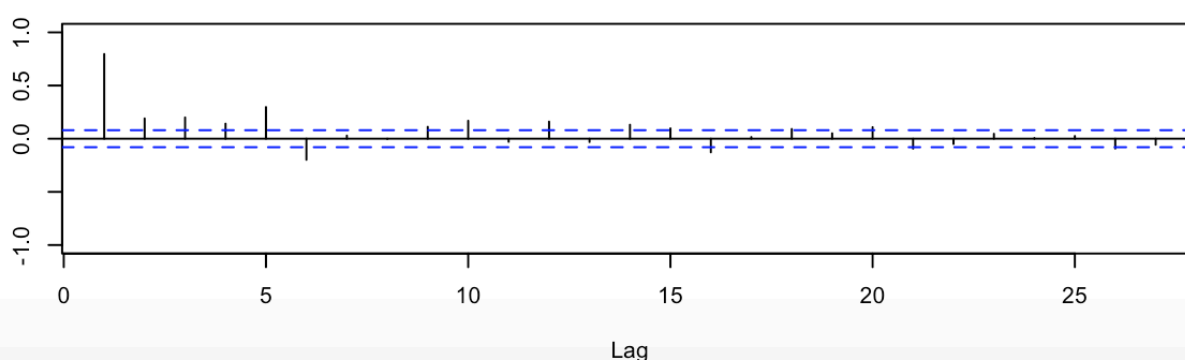
2) Deux façon pour verifier cette série est stationnaire ou pas.

i) ACF et PACF.

Autocorrelation



Partial Autocorrelation



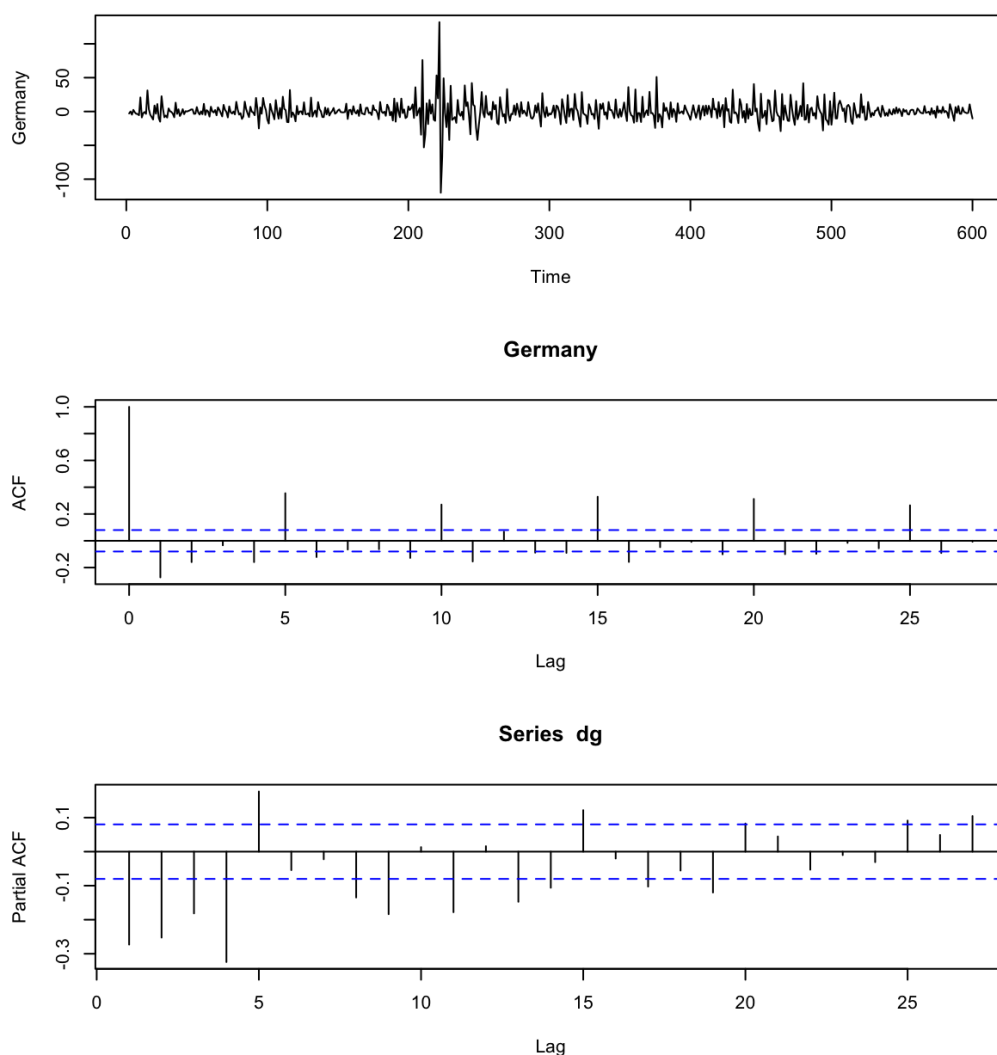
L'ACF montre clairement une très grande persistance. Notez que même si théoriquement la RU devrait donner des autocorrélations qui ne décroissent pas, en pratique elles le font, mais à un rythme très lent. Par contre, notez que pour une tendance déterministe, l'ACF décroît un plus rapidement et se stabilise en général. En tout cas, cette structure suggère une non stationnarité et le symptôme s'apparente à une RU.

ii) RU(racine unitaire)

Augmented Dickey-Fuller Test: $p\text{-value} = 0.09974$, du coup, on ne peut pas rejeter l'hypothèse originale, on peut dire que cette série n'est pas stationnaire.

3) Différence pour la stabilité de cette série.

i) Après la différence de premier ordre, on a les graphes de ACF, PACF du different série temporelle, on peut bien savoir que L'ACF ne montre pas une persistance et PACF se diminue jusqu'à zéro, par conséquent, on peut bien dire que après la différence de premier ordre, cette série est stationnaire, bien entendu, on va continuer de faire un test de racine unitaire pour confirmer notre jugement.



ii) RU(racine unitaire)

Augmented Dickey-Fuller Test: Dickey-Fuller = -12.753, Lag order = 8, p-value = 0.01, du coup, on peut bien rejeter l'hypothèse originelle: non stationnaire, accepter hypothèse alternative, on dit que après la différence de premier ordre, cette série est stationnaire.

4) Vérifiez si la série temporelle après la différence est du bruit blanc ou pas

Box-Ljung test: p-value = 2.063e-11, du coup, on peut bien rejeter l'hypothèse originale et accepter l'hypothèse alternative, la série temporelle après la différence n'est pas le bruit blanc, ensuite on peut faire l'opération prochaine.

5) Choix du modèle

Grâce aux graphes ACF, PAFC du différent série temporelle(procédé 3) et la fonction de R : `auto.arima ()`, on trouve bien les modèle ARMA(5, 1, 3), ARMA(4, 1, 4) dont AIC sont les plus petits.

ARIMA(2,1,2) with drift	: 4793.035
ARIMA(0,1,0) with drift	: 4940.334
ARIMA(1,1,0) with drift	: 4896.836
ARIMA(0,1,1) with drift	: 4840.011
ARIMA(0,1,0)	: 4938.321
ARIMA(1,1,2) with drift	: 4802.947
ARIMA(2,1,1) with drift	: 4801.975
ARIMA(3,1,2) with drift	: 4792.407
ARIMA(3,1,1) with drift	: 4804.554
ARIMA(4,1,2) with drift	: 4762.92
ARIMA(4,1,1) with drift	: 4762.351
ARIMA(4,1,0) with drift	: 4779.484
ARIMA(5,1,1) with drift	: 4763.938
ARIMA(3,1,0) with drift	: 4843.136
ARIMA(5,1,0) with drift	: 4763.478
ARIMA(5,1,2) with drift	: 4750.808
ARIMA(5,1,3) with drift	: 4735.726
ARIMA(4,1,3) with drift	: 4740.64
ARIMA(5,1,4) with drift	: 4740.999
ARIMA(4,1,4) with drift	: 4737.413
ARIMA(5,1,3)	: 4733.658
ARIMA(4,1,3)	: 4738.579
ARIMA(5,1,2)	: 4748.752
ARIMA(5,1,4)	: 4738.926
ARIMA(4,1,2)	: 4760.866
ARIMA(4,1,4)	: 4735.346



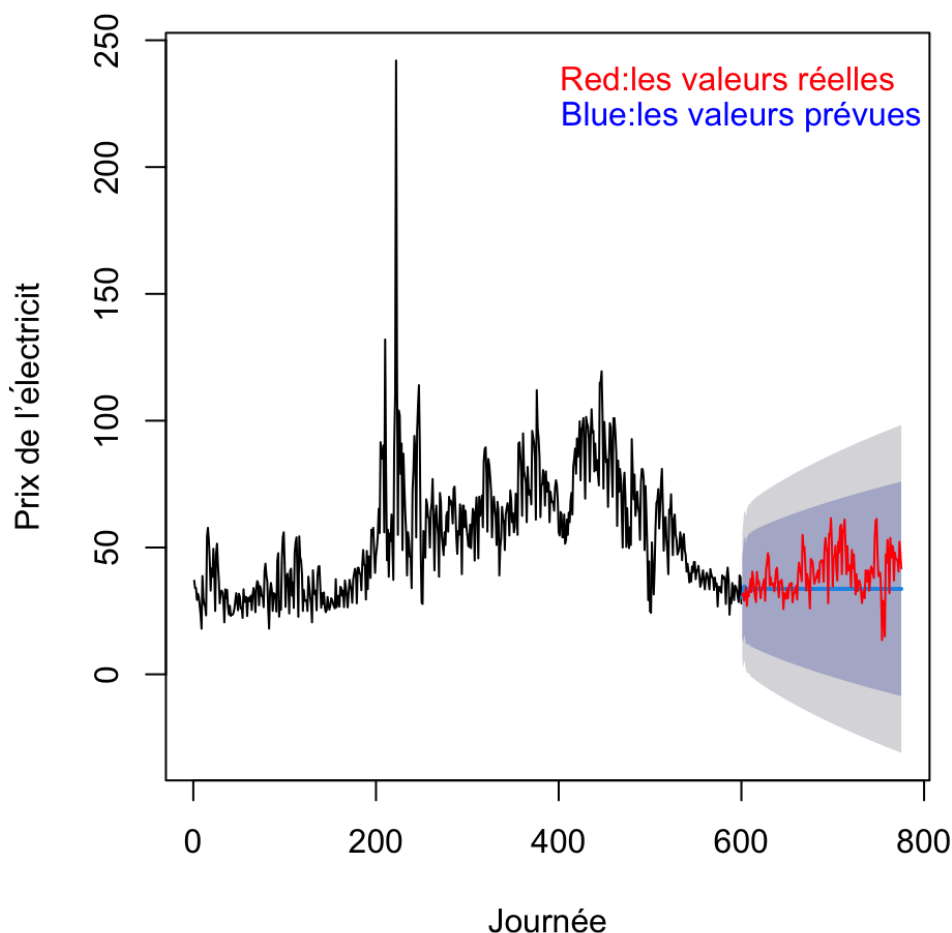
Selon des graphes d'ACF et PACF, on peut savoir que le modèle ARMA(5, 1, 3) est un pas mal modèle pour notre série, du coup, on va l'utiliser, pourtant, on va à la fois faire comparaison du autre modèle ARMA(4, 1, 4).

Comparaison des modèles			
	log likelihood	AIC	Résidus Box-Ljung test(p-value)
ARMA(5, 1, 3)	-2358.87	4733.66	Bruit blanc(0.9861)
ARMA(4, 1, 4)	-2365.31	4735.35	Bruit blanc(0.619)

on peut bien voir que les modèles ARMA(5, 1, 3) et ARMA(5, 0, 4) dont AIC sont les plus bas, leurs résidus sont à la fois bruit blanc, du coup, on peut pour l'instant comparer leurs max-vraisemblance, en raison que le modèle ARMA dont le méthode à calculer repose sur max-vraisemblance, par conséquent, on va choisir le modèle ARMA(5, 1, 3) comme notre modèle à prévoir.

6) Prédiction

Forecasts from ARIMA(5,1,3)



Bleu représente les valeurs prévues, rouge représente les valeurs réelles, on peut bien voir que les valeurs prévues se trouvent bien dans 95% d'intervalle de confiance (ombres sombres)

7) Qualité de la prédiction sur l'échantillon restant

i) L'erreur quadratique moyenne: 7.91

ii) La proportion des erreurs quadratiques moyennes: 18.80%.

(Ici, il semble que la variation des valeurs prévues est trop différente de la variation des valeurs réelles, pourtant, il faut bien noter que les valeurs prévues se trouvent bien dans 95% d'intervalle de confiance (ombres sombres) et le résidu du modèle est bruit blanc, du coup, notre prévision est bonne. L'autre part, on peut bien voir que les première 200 journée dont des prix de l'électricité a l'aire stationnaire, après 200 jusqu'à 550 il y a une grande variation, ensuite après 550 des prix de l'électricité se diminue au niveau de fluctuation de prix d'origine, par conséquent, on peut savoir qu'il y avait une choses significative qui s'est passé durant 200 et 550 journée, qui a provoqué la grande variation des prix de l'électricité et cette période est la crise économique de 2008 (le date de cette base de donnée est de 01/02/2007 à 01/28/2010). En gros, l'explication pour le fait que la variation des valeurs prévues est différente de la variation des valeurs réelles: la deuxième vague de crise économique en 2008 et peu nombre d'observation avec les jours de prévision plus longs(175 journées), mais notre prévision est bonne).

二. Month

1. France

Séries temporelle de le prix d'électricité en Allemagne et en France



1/2/2007 - 1/28/2010 (Month)	
	France
count	776.000000
mean	55.075290
std	20.822288
min	26.000000
25%	36.250000
50%	53.450000
75%	69.500000
max	122.750000

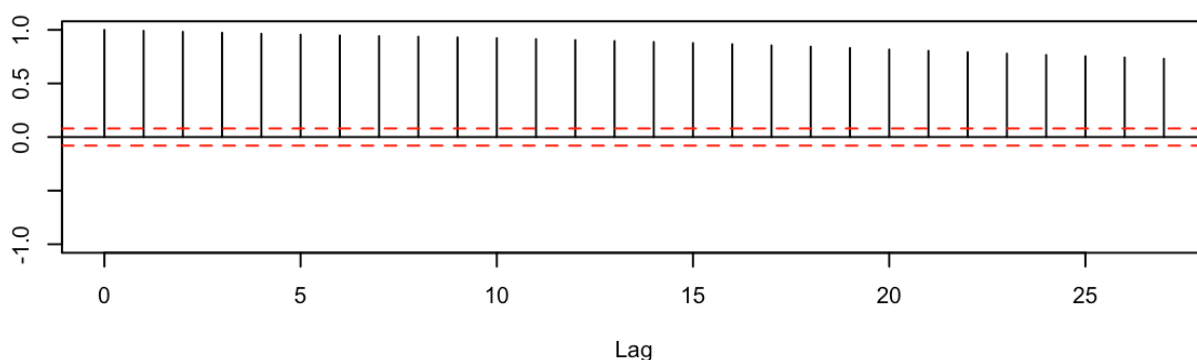
1) Dans le base de donnée, on ne trouve pas les valeurs manquantes et aberrantes.

(Des prix (ou des contrats) de type AHEAD (month-ahead, etc.) sont les prix prévisionnels dans les mois prochains)

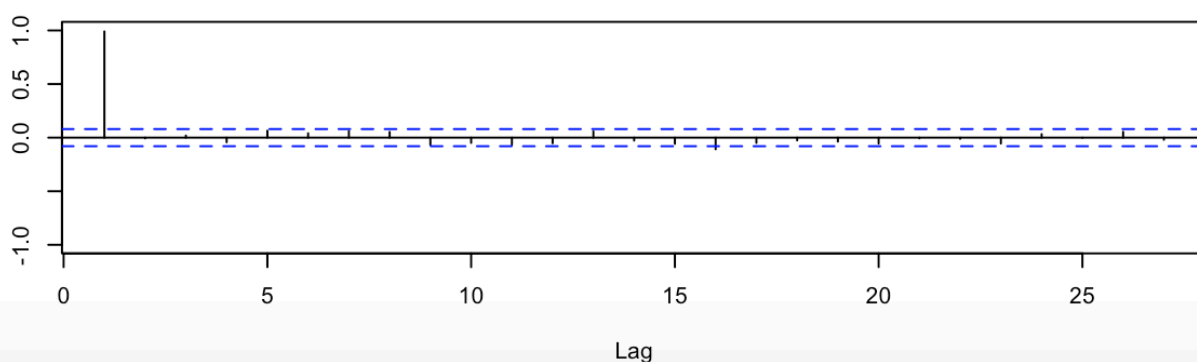
2) Deux façon pour verifier cette série(1- 600) est stationnaire ou pas.

i) ACF et PACF.

Autocorrelation



Partial Autocorrelation



L'ACF montre clairement une très grande persistance. Notez que même si théoriquement la RU devrait donner des autocorrélations qui ne décroissent pas, en pratique elles le font, mais à un rythme très lent. Par contre, notez que pour une tendance déterministe, l'ACF décroît un plus rapidement et se stabilise en général. En tout cas, cette structure suggère une non stationnarité et le symptôme s'apparente à une RU.

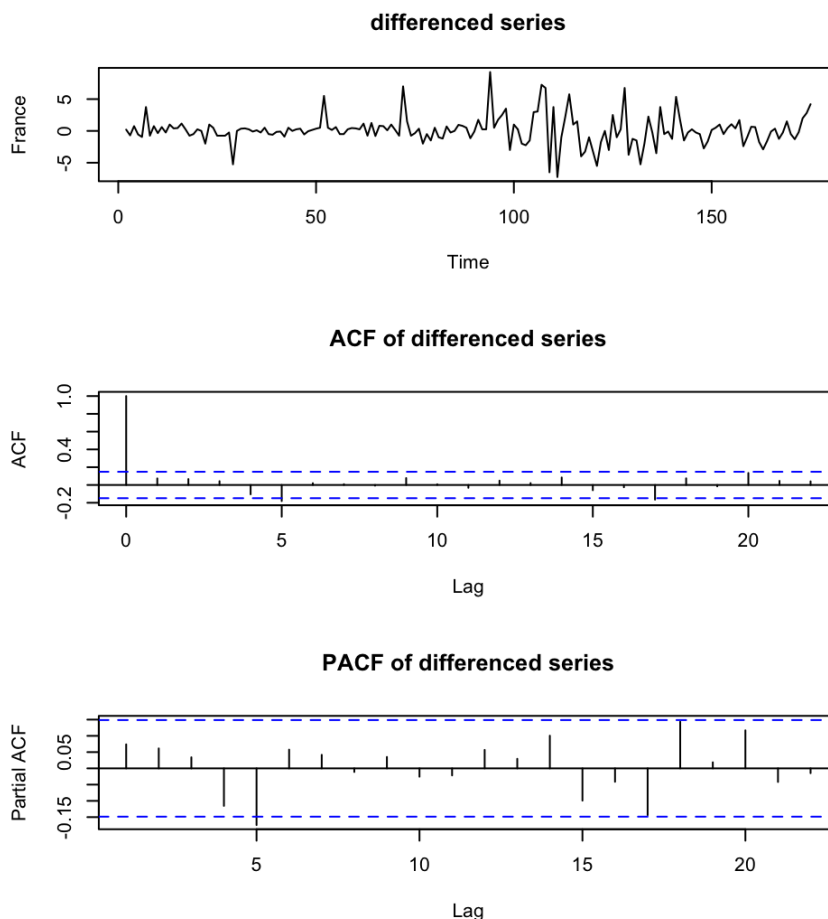
ii) RU(racine unitaire)

Augmented Dickey-Fuller Test: $p\text{-value} = 0.9545$, du coup, on ne peut pas rejeter l'hypothèse originale, on peut dire que cette série n'est pas stationnaire.

3) Différence pour la stabilité de cette série.

i) Après la différence de premier ordre, on a trouvé qu'elle était stationnaire ($p\text{-value} = 0.01$: non racine unitaire), mais elle était bruit blanc (Box-Ljung test: $p\text{-value} = 0.9883$), par conséquent, on a même essayé de faire ARCH LM-test sur les résidus: Chi-squared = 7.8275, $df = 20$, $p\text{-value} = 0.9929$, les résidus n'avaient pas du tout l'effet d'ARCH(ici, j'ai essayé $df = 10, 15, 20, 25, 30, 35, 40, 45, 50$, les résultats sont trop similaires = P-valeurs sont 0.9892, ne jamais inférieures à 0.8). Du coup, on analyse l'échantillon restant (601- 775) directement

ii) On a les graphes de ACF, PACF du différent série temporelle, on peut bien savoir que L'ACF ne montre pas une persistance et PACF se saute entre 0.05 et -0.05, par conséquent, on peut bien dire que après la différence de premier ordre, cette série est stationnaire, bien entendu, on va continuer de faire un test de racine unitaire pour confirmer notre jugement.



iii) RU(racine unitaire)

Augmented Dickey-Fuller = -8.3216, Lag order = 8, p-value = 0.01, du coup, on peut bien rejeter l'hypothèse originelle: non stationnaire, accepter hypothèse alternative, on dit que après la différence de premier ordre, cette série est stationnaire.

4) ARCH, GARCH

Les modèles ARCH et GARCH (Les modèles conditionnellement hétéroscédastiques)

ARCH(p):

$$X_t = W_t \varepsilon_t$$

$$W_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2$$

$\varepsilon_t \sim N(0, 1)$ bruit blanc gaussien

GARCH(p, q):

$$X_t = W_t \varepsilon_t$$

$$W_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j W_{t-j}^2$$

$\varepsilon_t \sim N(0, 1)$ bruit blanc gaussien

5) Test d'effet ARCH

i) Établir une équation de valeur moyenne ARMA(2, 0, 2) pour la série temporelle de fluctuation du prix de l'électricité afin d'éliminer toute dépendance linéaire.

ii) Test d'effet ARCH sur le résidu de l'équation moyenne.

ARCH LM-test:

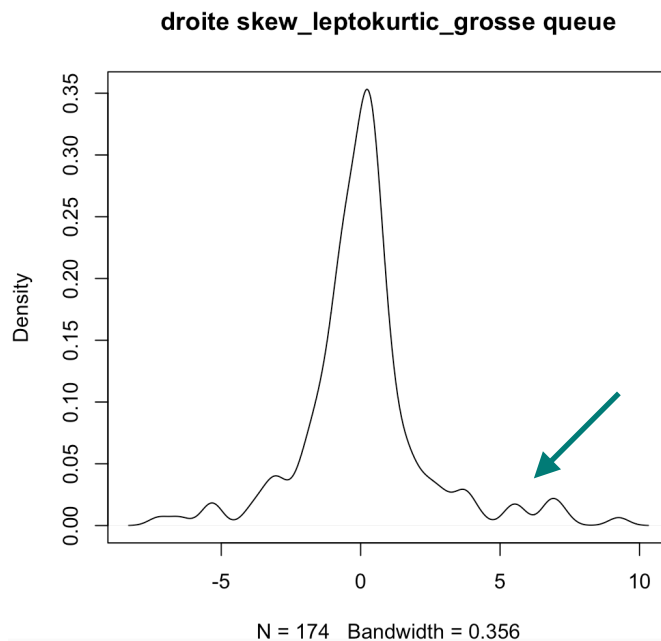
Null hypothesis: no ARCH effects,

Chi-squared = 35.993, df = 20, p-value = 0.02996 < 0.05

Par conséquent, on peut bien rejeter l'hypothèse originale et accepter l'hypothèse alternative, on dit qu'il existe clairement l'effet ARCH au niveau de signification de 5% ($0.02996 < 0.05$)

6) Taux de changement de prix de l'électricité

Description statistique du taux de changement de prix de l'électricité	
Moyen	0.1135057
St	2.260142
Asymétrie	0.6254119
Kurtosis	6.340033
Jarque Bera Test	p-value < 2.2e-16



Selon la description statistique du taux de changement de prix de l'électricité, on peut savoir que le moyen : 0.1135057, l'écart-type: 2.260142, l'asymétrie: 0.6254119, kurtosis: 6.340033 et JB test dont P-valeurs < 2.2e-16, autrement dit, le taux de changement de prix de l'électricité : droite skew, leptokurtic, grosse queue.

7) Estimation des modèles ARCH et GARCH

i) ARCH(1, 0)

ARCH(1, 0)			
	Coefficient	Std. Error	Pr(> t)
mu	0.1053	0.1531	0.4917
omega	3.4071	0.4874	2.75E-12
alpha1	0.4150	0.1715	0.0155
Log Likelihood	-379.5763		
	Standardised Residuals Tests		
	P-Value		
Jarque-Bera Test	0		
Shapiro-Wilk Test	2.963397E-10		
Ljung-Box Test (Q10-R)	0.4987245		
Ljung-Box Test (Q10-R²)	0.9198673		
LM Arch Test	0.9844845		
	AIC 4.397428	BIC 4.451895	SIC 4.396847 HQIC 4.419523

1) Omega et alpha dont les P-valeurs sont tous significatifs (< 0.05), mais mu n'est pas significatif.

2) Les résidus dont JB test montre que ils ne suivent pas la loi normale comme prévu; Ljung-Box tous les deux ne sont pas significatifs, du coup, on peut rejeter l'hypothèse originale et accepter ils sont bruits blancs, LM Arch test montre qu'il n'y a pas d'effet ARCH dans les résidus. À la fin, on peut bien dire que le modèle ARCH peut expliquer efficacement le taux de changement de prix de l'électricité (volatilité de prix de l'électricité).

3) En gros, on a l'équation:

$$X_t = W_t \varepsilon_t$$

$$W_t^2 = 3.407 + 0.415 * X_{t-1}^2$$

$$\varepsilon_t \sim N(0, 1) \text{ bruit blanc gaussien}$$

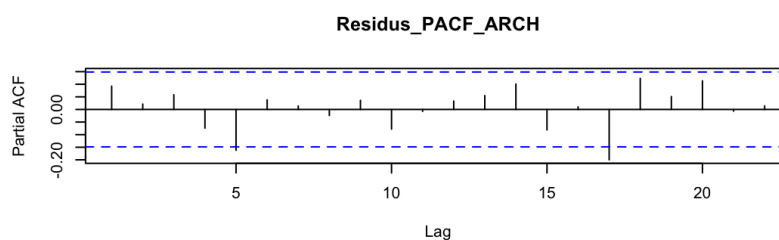
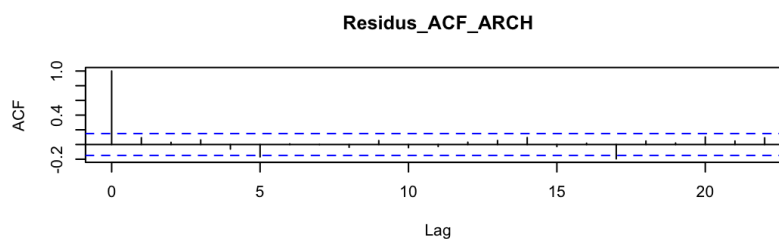
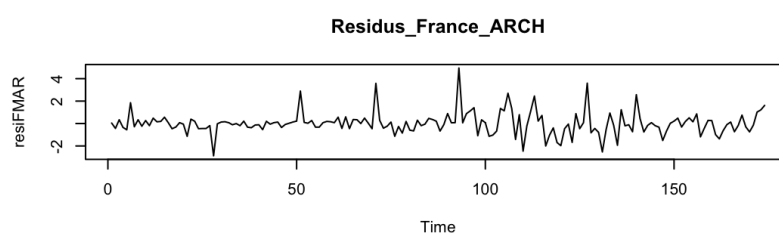
ii) GARCH(1, 1)

GARCH(1, 1)			
	Coefficient	Std. Error	Pr(> t)
mu	0.15637	0.15577	0.3154
omega	0.39932	0.27668	0.1489
alpha1	0.06628	0.1715	0.08853
beta1	0.85176	0.08853	<2e-16
Log Likelihood	-376.7411		
	Standardised Residuals Tests		
	P-Value		
Jarque-Bera Test	0		
Shapiro-Wilk Test	1.070769E-10		
Ljung-Box Test (Q10-R)	0.6574853		
Ljung-Box Test (Q10-R²)	0.9997256		
LM Arch Test	0.9999731		
	AIC 4.376334	BIC 4.448956	SIC 4.375309
			HQIC 4.405794

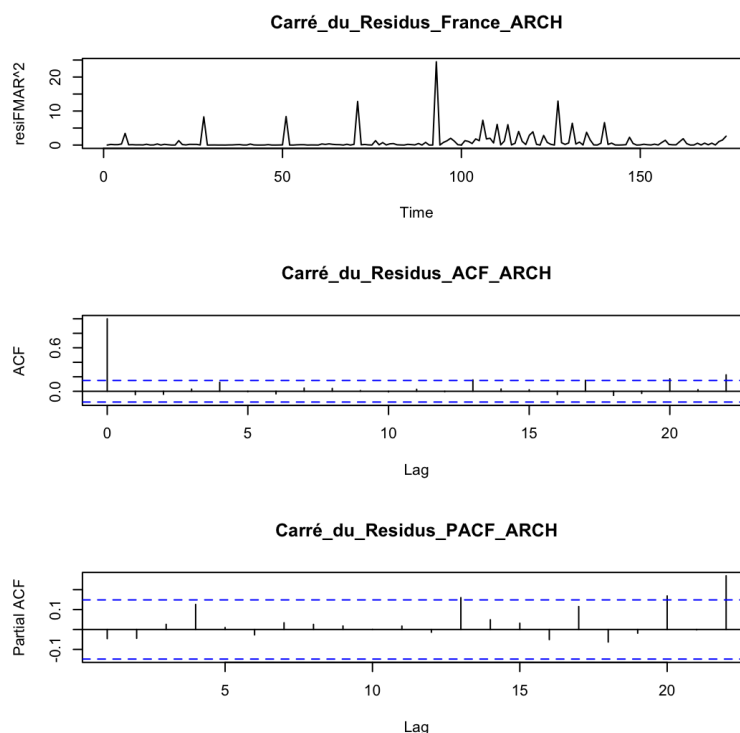
1) on peut bien voir que mu, omega et alpha 1 dont P-valeurs sont tous supérieurs à 0.05(non significatif), seulement, beta1 est inférieur à 0.05 (significatif), par conséquent, le modèle GARCH n'est pas bon, on va utiliser le modèle ARCH(1, 0) comme notre final modèle.

8) Qualité du modèle

i) ACF et PACF (Résidus)

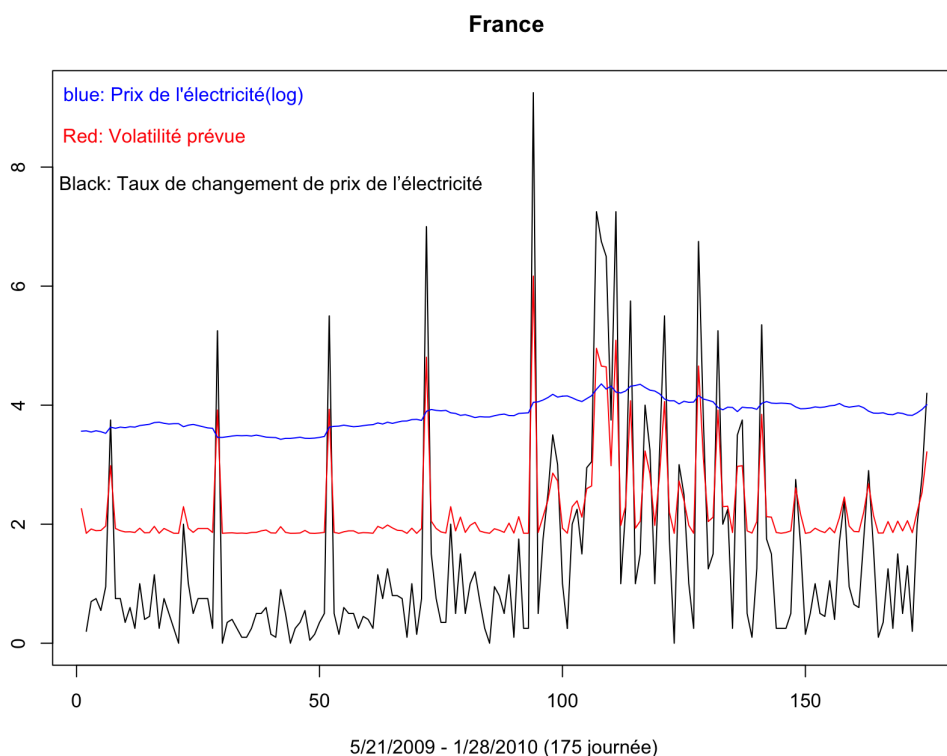


ii) ACF et PACF (Résidus de carré)



Les figures ci-dessus sont le graphe d'auto-corrélation et le graphe d'auto-corrélation partielle des résidus et des carrés résiduels. On peut voir que la plupart des valeurs de fonction des diagrammes ACF et PACF de la séquence résiduelle sautent de haut en bas dans l'intervalle de confiance (la zone en pointillé bleu sur la figure), de sorte que la séquence résiduelle normalisée n'a pas d'auto-corrélation ou a une certaine corrélation faible (Résultat graphique PACF). Cependant, les tracés ACF et PACF de la séquence carrée résiduelle n'ont pas de queue ou de troncature évidente, et toutes les valeurs de fonction sont comprises dans l'intervalle de confiance, de sorte qu'elles n'ont pas de corrélation en série. À la fin, on peut bien dire que le modèle ARCH peut expliquer efficacement le taux de changement de prix de l'électricité (volatilité de prix de l'électricité).

9) Prévision



La ligne en bleu représente les valeurs réelles (par log), la ligne en noir représente la volatilité réelle et la ligne en rouge représente la volatilité prévue (la prévision du modèle ARCH(1, 0)). On peut bien voir que notre volatilité prévue peut bien représenter la volatilité réelle efficacement et notre modèle ARCH(1, 0) dont la qualité est très bonne.

2. Germany

Séries temporelle de le prix d'électricité en Allemagne et en France



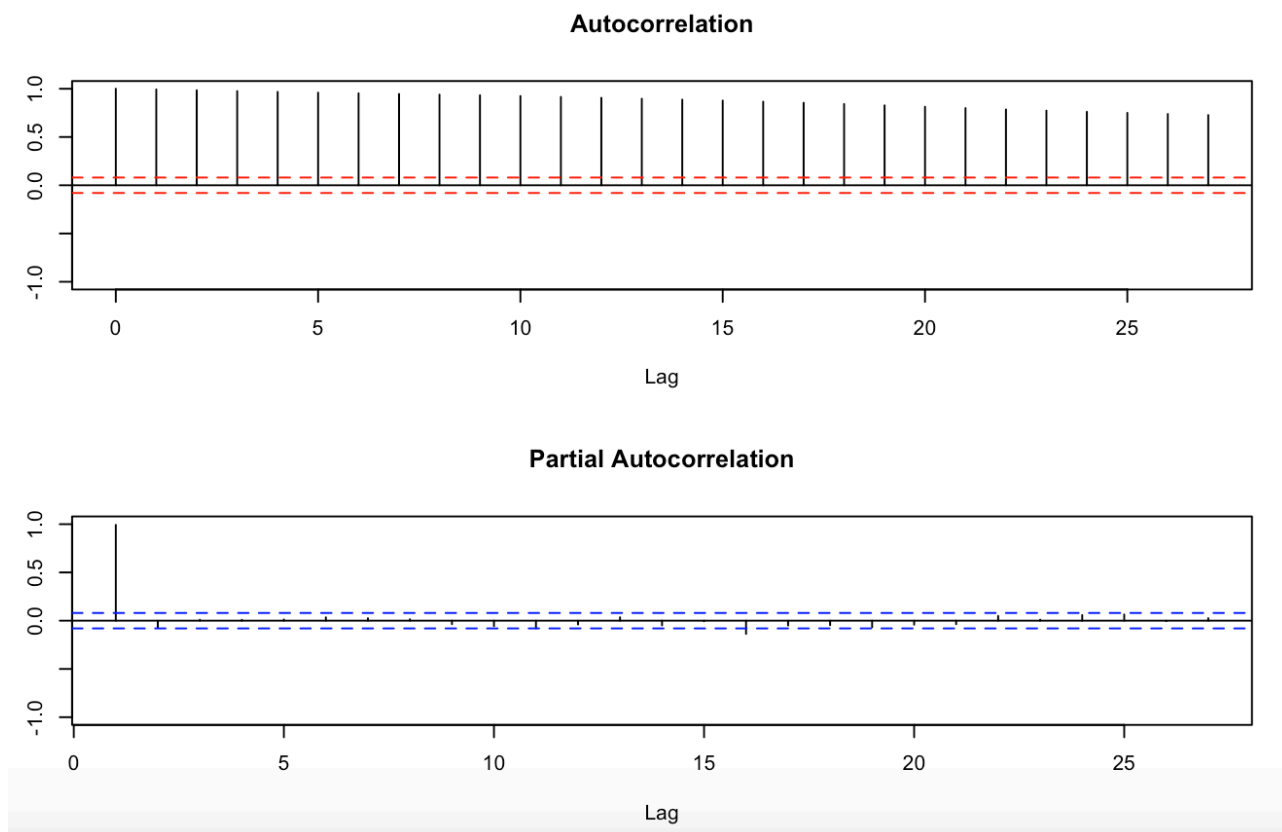
1/2/2007 - 1/28/2010 (Month)	
	Germany
count	776.000000
mean	50.961985
std	17.016688
min	26.700000
25%	36.687500
50%	45.800000
75%	61.812500
max	98.000000

1) Dans le base de donnée, on ne trouve pas les valeurs manquantes et aberrantes.

(Des prix (ou des contrats) de type AHEAD (month-ahead, etc.) sont les prix prévisionnels dans les mois prochains)

2) Deux façon pour verifier cette série(1- 600) est stationnaire ou pas.

i) ACF et PACF.



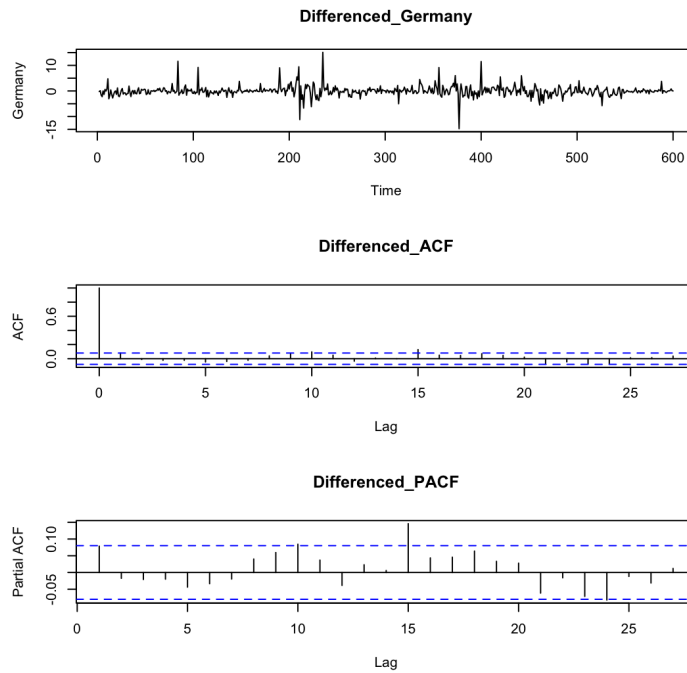
L'ACF montre clairement une très grande persistance. Notez que même si théoriquement la RU devrait donner des autocorrélations qui ne décroissent pas, en pratique elles le font, mais à un rythme très lent. Par contre, notez que pour une tendance déterministe, l'ACF décroît un plus rapidement et se stabilise en général. En tout cas, cette structure suggère une non stationnarité et le symptôme s'apparente à une RU.

ii) RU(racine unitaire)

Augmented Dickey-Fuller Test: p-value = 0.9521, du coup, on ne peut pas rejeter l'hypothèse originale, on peut dire que cette série n'est pas stationnaire.

3) Différence pour la stabilité de cette série.

i) Après la différence de premier ordre, on a trouvé qu'elle était stationnaire(p-value = 0.01: non racine unitaire) et ACF, PACF ne montrent pas une persistance. mais elle était bruit blanc(Box-Ljung test: p-value = 0.05353), par conséquent, on a même essayé de faire ARCH LM-test sur les résidus: Chi-squared = 7.8275, df = 21, p-value = 0.01423, les résidus avaient l'effet d'ARCH(ici, j'ai essayé df= 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 et seulement, df est entre 21 et 54, l'effet ARCH exist, sinon, pas du tout).



4) ARCH, GARCH

Les modèles ARCH et GARCH (Les modèles conditionnellement hétéroscédastiques)

ARCH(p):

$$X_t = W_t \varepsilon_t$$

$$W_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2$$

$\varepsilon_t \sim N(0, 1)$ bruit blanc gaussien

GARCH(p, q):

$$X_t = W_t \varepsilon_t$$

$$W_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j W_{t-j}^2$$

$\varepsilon_t \sim N(0, 1)$ bruit blanc gaussien

5) Test d'effet ARCH

i) Établir une équation de valeur moyenne ARMA(0, 0, 0) pour la série temporelle de fluctuation du prix de l'électricité afin d'éliminer toute dépendance linéaire.

ii) Test d'effet ARCH sur le résidu de l'équation moyenne.

ARCH LM-test:

Null hypothesis: no ARCH effects,

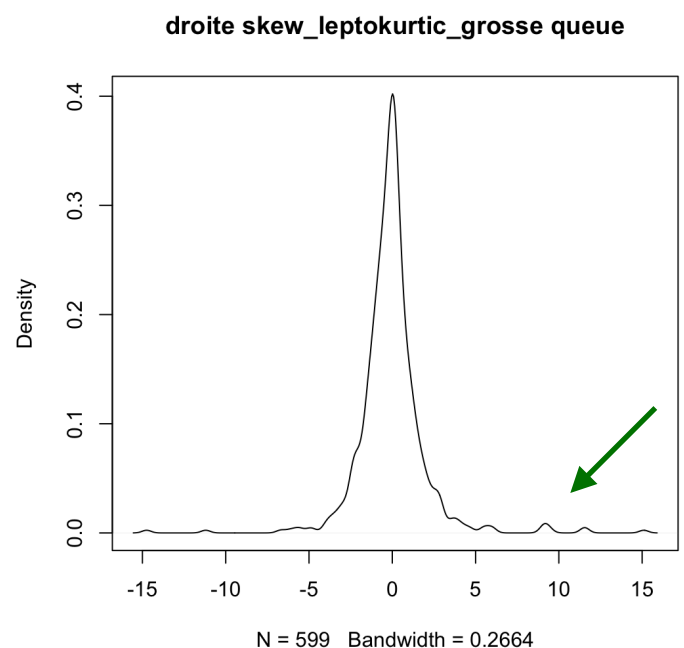
Chi-squared = 35.993, df = 21, p-value = 0.01423 < 0.05

*(Seulement, il' exist l'effet ARCH en condition que df est entre 21 et 54)

Par conséquent, on peut bien rejeter l'hypothèse originale et accepter l'hypothèse alternative, on dit qu'il existe clairement l'effet ARCH au niveau de signification de 5% (0.01423 < 0.05)

6) Taux de changement de prix de l'électricité

Description statistique du taux de changement de prix de l'électricité	
Moyen	-0.02929883
St	2.065721
Asymétrie	0.9348955
Kurtosis	18.24401
Jarque Bera Test	p-value < 2.2e-16



Selon la description statistique du taux de changement de prix de l'électricité, on peut savoir que le moyen : -0.02929883, l'écart-type: 2.065721, l'asymétrie:0.9348955, kurtosis: 18.24401 et JB test dont P-valeurs < 2.2e-16, autrement dit, le taux de changement de prix de l'électricité : droite skew, leptokurtic, grosse queue.

7) Estimation des modèles ARCH et GARCH

i)

	AIC	BIC	Log Likelihood
GARCH(2, 2)	4.124344	4.168370	-1229.241
GARCH(1, 2)	4.185556	4.222244	-1248.574
GARCH(2, 1)	4.185669	4.222358	-1248.608
GARCH(1, 1)	4.182111	4.211462	-1248.542
ARCH(1, 0)	4.273712	4.295725	-1276.977
ARCH(10, 0)	4.116336	4.204388	-1220.843
ARCH(15, 0)	4.130837	4.255578	-1220.186

ii) GARCH(2, 2)

GARCH(2, 2)			
	Coefficient	Std. Error	Pr(> t)
mu	3.354E-02	6.464E-02	0.604
omega	1.084E+00	2.162E-01	5.32E-07
alpha1	1.000E-08	9.361E-03	1.000
alpha2	5.237E-01	2.611E-02	2.94E-05
beta1	1.000E-08	7.779E-02	1.000
beta2	3.909E-01	1.253E-01	5.02E-07
Log Likelihood	-1229.241		
	Standardised Residuals Tests		
	P-Value		
Jarque-Bera Test	0		
Shapiro-Wilk Test	0		
Ljung-Box Test (Q10-R)	0.06915612		
Ljung-Box Test (Q10-R²)	0.995324		
LM Arch Test	0.9984176		
AIC	BIC	SIC	HQIC
4.124344	4.168370	4.124146	4.141484

GARCH(2, 2):

$$X_t = W_t \varepsilon_t$$

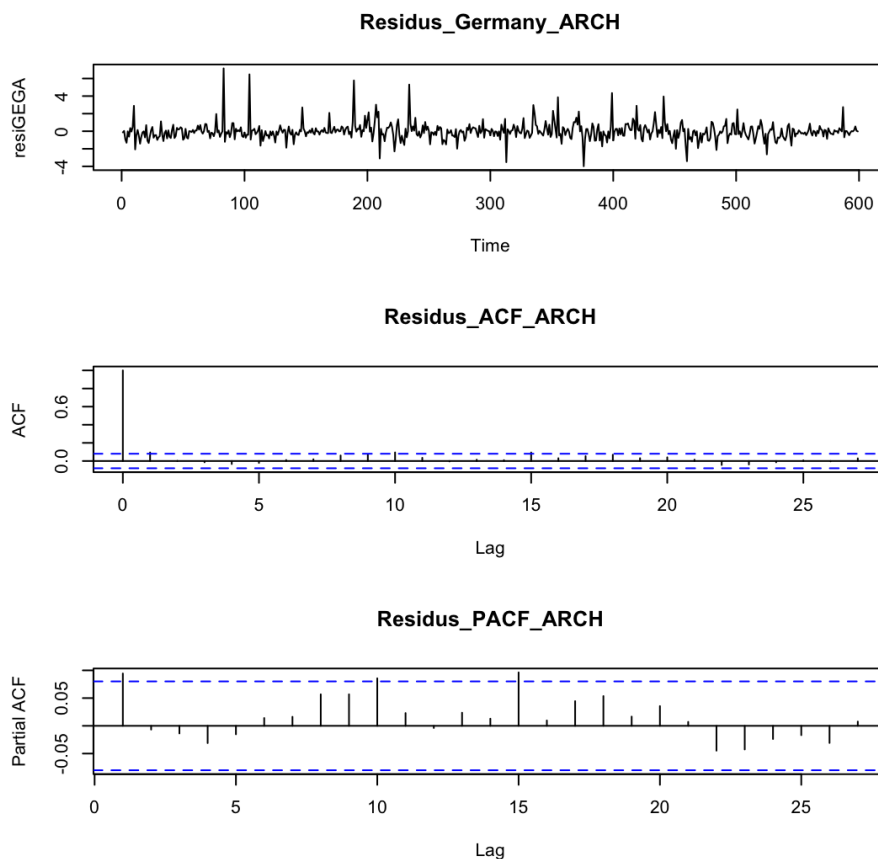
$$W_t^2 = 1.084 + 5.237e^{-01}X_{t-2}^2 + 3.909e^{-01}W_{t-2}^2$$

$$\varepsilon_t \sim N(0, 1) \text{ bruit blanc gaussien}$$

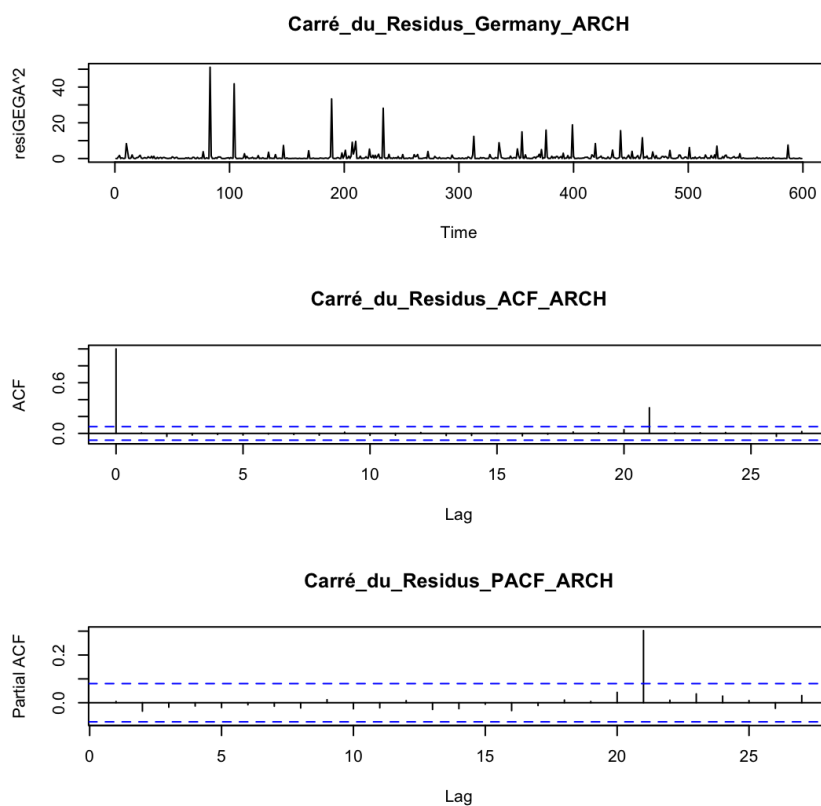
On fait la comparaison sur chaque modèle, on trouve le modèle GARCH(2, 2) dont AIC, BIC et Log vraisemblance sont les plus pertinents, du coup, on va utiliser le modèle GARCH(2, 2) comme notre final modèle.

8) Qualité du modèle

i) ACF et PACF (Résidus)



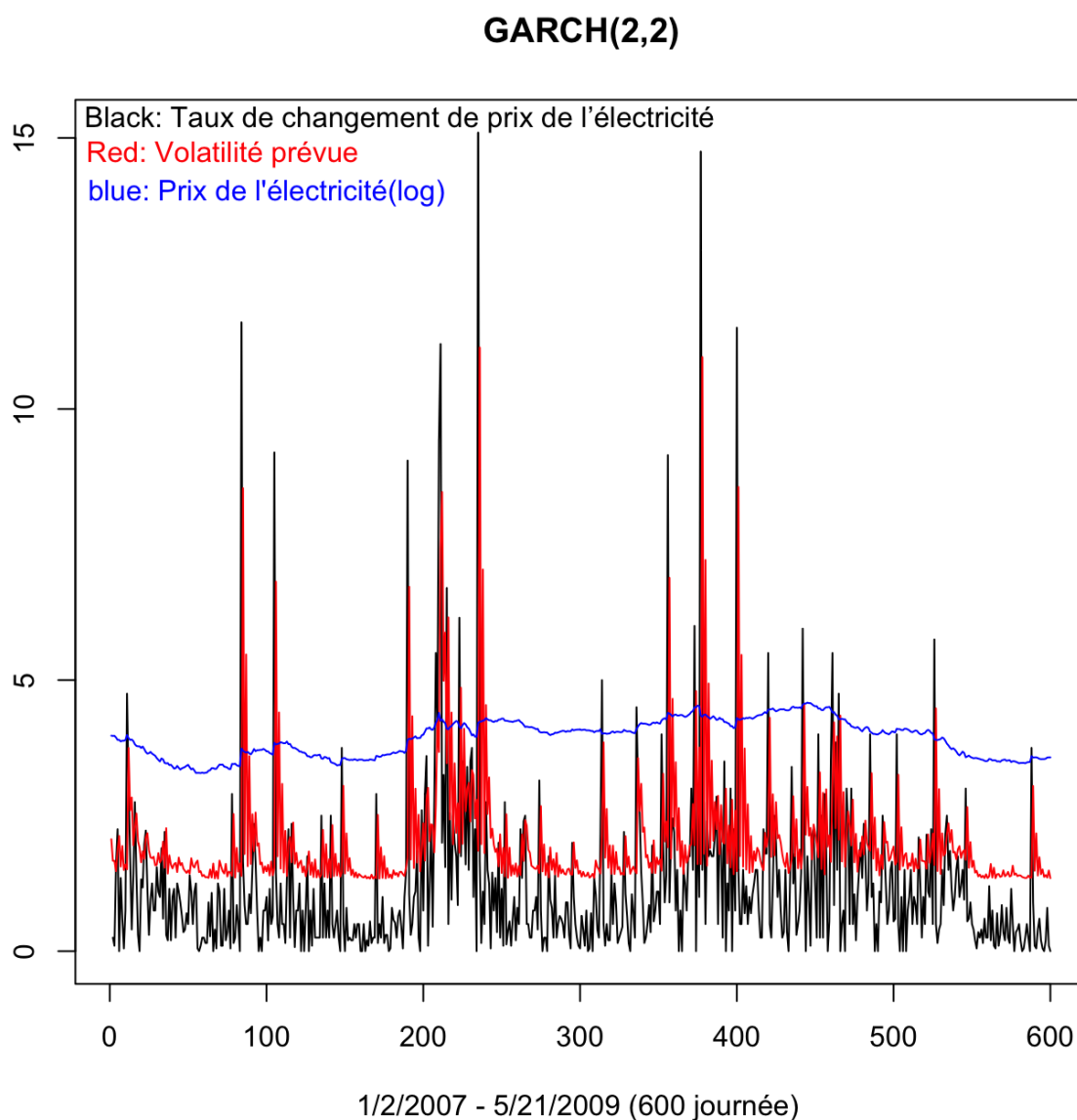
ii) ACF et PACF (Résidus de carré)



Les figures ci-dessus sont le graphe d'auto-corrélation et le graphe d'auto-corrélation partielle des résidus et des carrés résiduels. On peut voir que la plupart des valeurs de fonction des diagrammes ACF et PACF de la séquence résiduelle sautent de haut en bas dans l'intervalle de confiance (la

zone en pointillé bleu sur la figure), de sorte que la séquence résiduelle normalisée n'a pas d'auto-corrélation ou a une certaine corrélation faible (Résultat graphique ACF). Cependant, les tracés ACF et PACF de la séquence carrée résiduelle n'ont pas de queue ou de troncature évidente, et toutes les valeurs de fonction sont comprises dans l'intervalle de confiance, de sorte qu'elles n'ont pas de corrélation en série. À la fin, on peut bien dire que le modèle GARCH peut expliquer efficacement le taux de changement de prix de l'électricité (volatilité de prix de l'électricité).

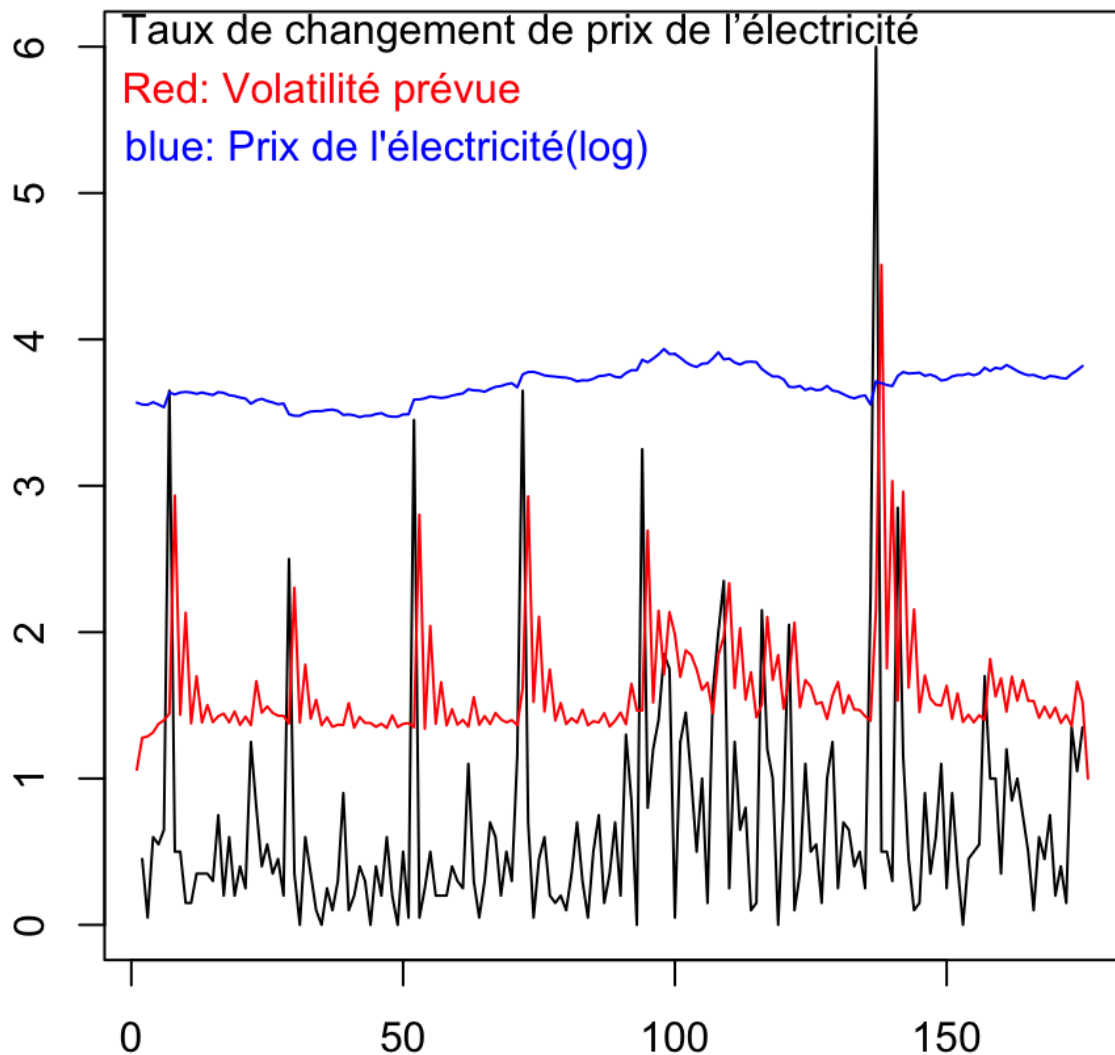
9) Prédiction sur l'échantillon d'apprentissage (600 journées)



La ligne en bleu représente les valeurs réelles (par log), la ligne en noir représente la volatilité réelle et la ligne en rouge représente la volatilité prévue (la prévision du modèle GARCH(2, 2)). On peut bien voir que notre volatilité prévue peut bien représenter la volatilité réelle efficacement et notre modèle GARCH(2, 2) dont la qualité est très bonne.

10) Pr vision sur l chantillon restant (175 journ es).

GARCH(2,2)



5/21/2009 - 1/28/2010 (175 journ e)

La ligne en bleu repr sente les valeurs r elles (par log), la ligne en noir repr sente la volatilit  r elle et la ligne en rouge repr sente la volatilit  pr vue (la pr vision du mod le GARCH(2, 2)). On peut bien voir que notre volatilit  pr vue peut bien repr senter la volatilit  r elle efficacement et notre mod le GARCH(2, 2) dont la qualit  est tr s bonne.