Séries Temporais - Análise da volatilidade da série de preços WTI(16)

Guilherme Souza

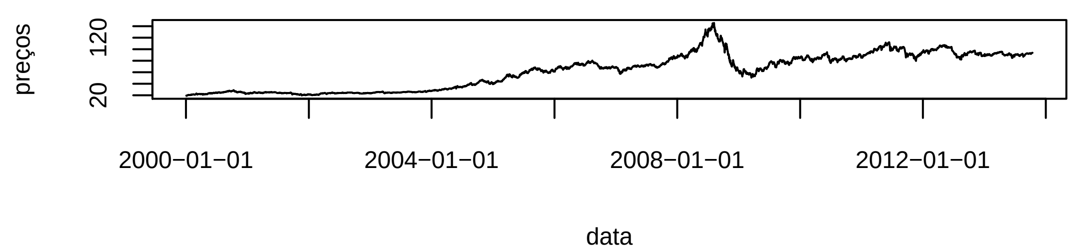
21 de dezembro de 2018

Requisitos formais

- No arquivo "series1":
  - Escolher uma série de retornos diferente das que foram vistas em aulas.
- Realize a estimação da volatilidade condicional com o modelo APARCH(p,q) para os dados de preços da ação escolhida.
  - Utilize a distribuição t Student.
  - Escolha a ordem (p,q) baseando-se no critério BIC.
- Ajuste um dos modelo de volatilidade condicional obedecendo o critério BIC. Escolha o modelo, a ordem, a distribuição dos erros. Mostre todos os modelos analisados e faça os testes diagnósticos e gráficos para o modelo escolhido.
- Escrever a equação do modelo.

Leitura dos dados

Neste trabalho serão analisados os preços do 16º contrato de petróleo WTI, correspondentes à 11ª coluna do arquivo disponibilizado. Após a leitura, os dados foram convertidos em um objeto do tipo "time series" por meio da função `timeSeries()` do pacote timeSeries. Em seguida, foram computados os respectivos retornos pela função `returns()` do mesmo pacote citado. A figura 1 apresenta o comparativo entre os preços originais e os retornos computados.



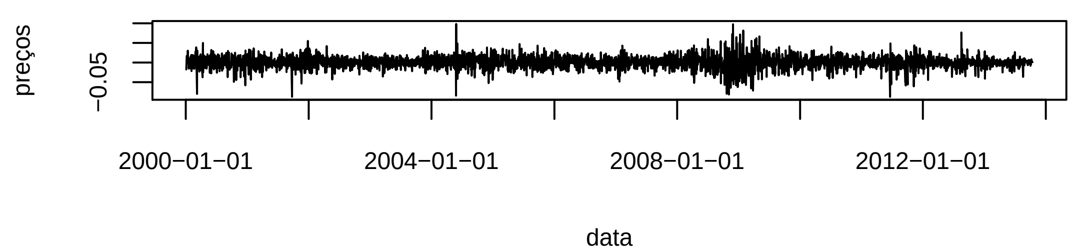


Figura 1: Série dos preços "em natura" do 16º contrato de WTI no primeiro gráfico. Abaixo, os retornos.

Observando o segundo gráfico da figura 1 vemos que a os retornos não são constantes. De acordo com Shumway e Stoffer (2000, p. 256), em séries financeiras o retorno no tempo não tem variância constante, além disso, períodos de alta volatilidade tendem a se agrupar. Na figura 1 podemos observar que a volatilidade encontra-se agrupada em um grande pacote que vai do início de 2008 a 2010.

Análise dos dados

Características da distribuição

A tabela 1 apresenta as principais estatísticas sobre a distribuição da série de retornos. Observamos que o valor da curtose é de 3.35, que é um pouco maior que a da distribuição normal cujo valor é 3. Isso indica que a distribuição possui caldas ligeiramente mais robustas.

Tabela 1: Sumário estatístico da série de retornos.

	Size	Max	Min	Mean	SD	Skewness	Kurtosis	Jarque-Bera
wti.ret	3415	0.0982718	-0.0874053	0.0004615	0.016282	-0.2368714	3.346298	1628.847

Teste de estacionariedade

A tabela 2 demonstra estatísticas dos testes de estacionariedade das séries. No primeiro teste (ADF) o p-valor alto da série permite aceitar a hipótese nula de não estacionariedade. Após computação dos retornos vemos que o teste indica que a série passa a ser estacionária.

O teste KPSS, de forma inversa, revela uma situação atípica pois o p-valor inicial da série original é baixo e permite aceitar a hipótese alternativa de estacionariedade. Contudo, após transformação para retornos o p-valor aumenta, rejeitando a hipótese de estacionariedade e indicando não estacionariedade.

Tabela 2: Testes de estacionariedade aplicados nos preços originais e nos retornos.

Teste	Hipótese alternativa	p-value (série original)	p-value (série de retornos)
ADF	stationary	0.3913955	0.01
KPSS	stationary	0.0100000	0.10

Teste de auto-correlação

Testou-se a auto-correlação *Box-Pierce* da série de retornos por meio da rotina `Box.test()`. A hipótese nula deste teste é de que as observações são independentes umas das outras. O que equivale testar se não há independência linear nos dados da série. Conforme observado na tabela 3, os testes apontam para a hipótese alternativa de que existe correlação linear nos valores da série. Além disso, conforme são aumentados os parâmetros do *lag*, observamos que o p-valor também aumenta.

Tabela 3: Testes de auto-correlação Box-Pierce aplicados na série de retornos variando o parâmetro lag.

X-squared	P-valor	lag	DF
26.39113	0.0000003	1	1
32.89513	0.0000039	5	5
35.53161	0.0001013	10	10

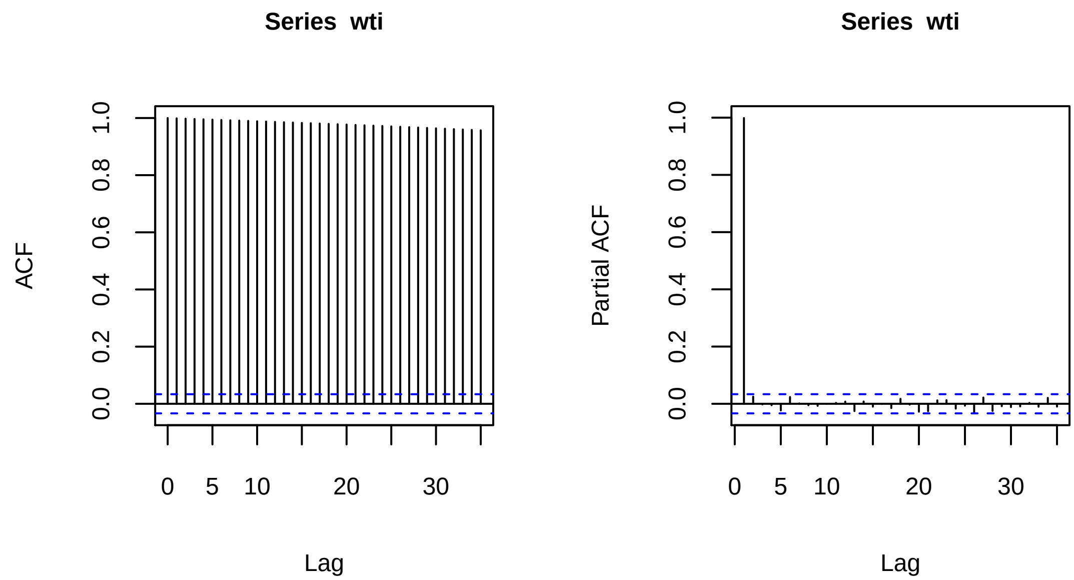


Figura 2: Gráficos de auto-correlação e auto-correlação parcial da série de preços originais.

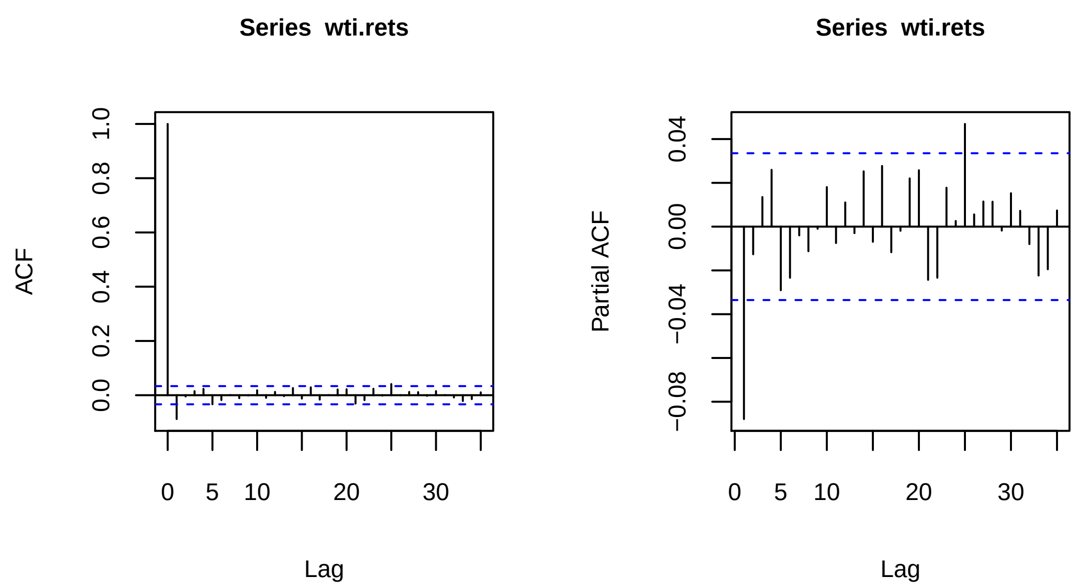


Figura 3: Gráficos de auto-correlação e auto-correlação parcial da série de retornos.

A partir da observação dos gráficos de auto-correlação e auto-correlação parcial da figura 3, é notável a dificuldade em se estimar um modelo ARMA com base apenas nos *lags* significantes. Além disso, as correlações são relativamente próximas de 0. De acordo com Ding (2011, p. 3), modelos dos tipos AR, MA, ARMA e ARIMA são mais aplicáveis a fenômenos de curta duração e não se mostram eficientes em lidar, sozinhos, com a existência do fenômeno denominado *memória longa*.

Teste Jarque-Bera

O teste Jarque-Bera testa a hipótese nula de normalidade nos dados. O resultado do teste é dado a seguir.

```
##
```

```
##  Jarque Bera Test
##
## data:  wti.rets
## X-squared = 1628.8, df = 2, p-value < 2.2e-16
```

Verifica-se um p-valor baixo o que permite descartar a hipótese se nula de normalidade dos dados.

Teste ARCH-LM

O teste Arch-LM verifica a hipótese nula de que não há efeitos de heterocedasticidade condicional auto-regressiva. O teste realizado apresenta os seguintes resultados:

```
##
##  ARCH LM-test; Null hypothesis: no ARCH effects
##
## data:  wti.rets
## Chi-squared = 383.59, df = 12, p-value < 2.2e-16
```

O p-valor baixo sugere que é possível rejeitar a hipótese nula de que não existem efeitos de heterocedasticidade condicional. Em outras palavras, o teste confirma que os dados não se comportam de forma homocedástica e a variância (volatilidade) depende do tempo.

Seleção do modelo ARMA(p, q)

Os parâmetros p e q foram obtidos por intermédio de execuções da função `arima()`. Posteriormente, foram extraídos os critérios de informação AIC (Akaike Information Criterion) e BIC (Bayes information Criterion).

Tabela 4: Modelos e critério de informação. O melhor modelo é ARMA(0, 1)

	AIC	BIC
ARMA(0, 0)	-18429.514518	-18417.2427
ARMA(0, 1)	-18454.40054	-18435.9927
ARMA(0, 2)	-18452.41017	-18427.8664
ARMA(1, 0)	-18454.010078	-18435.6023
ARMA(1, 1)	-18452.407381	-18427.8636
ARMA(1, 2)	-18450.420828	-18419.7412
ARMA(2, 0)	-18452.555065	-18428.0113
ARMA(2, 1)	-18450.499855	-18419.8202
ARMA(2, 2)	-18452.193815	-18415.3782

Baseado nos critérios de infromação, podemos dizer que o modelo que melhor se ajusta é um média móvel de primeira ordem. Confirma-se pela tabela 4 que o modelo com menorer AIC e BIC é o ARMA(0,1). Logo a equação do modelo de média condicional é do seguinte tipo:

$$X_t = \theta Z_{t-1} + Z_t \tag{1}$$

Estimação

As tabelas 5, 6 e 7 apresentam respectivamente as estimações de modelos ARMA(0,1)-SGARCH(p, q), ARMA(0,1)-APARCH(p, q) e ARMA(0,1)-EGARCH(p, q). Para cada estimação foram utilizadas os modelos de distribuição *student-t* (`std`), *ged* e *skew-student* (`sstd`). Ao final de cada tabela são apresentados os critérios de informação Akaike (AIC), Bayes (BIC), Shibata e Hannan-Quinn, tendo sido marcados em negrito os melhores (mínimos) critérios. Os parâmetros ótimos, erro padrão (*standard error*) e p-valor aparecem, nesta ordem, em cada célula.

Nos modelos sGARCH e apARCH, a ordem que apresenta menores critérios de informação é (1,2) com a distribuição `sstd`. Apenas o modelo eGARCH tem mínimo critério de informação na distribuição `sstd` e ordem (2,1). Desta forma, podemos sugerir que, pelo critério de informação, o modelo que melhor se ajusta aos dados é o ARMA(0,1)-APARCH(1,2) com distribuição `sstd`.

Com base na tabela 6, os parâmetros ótimos com p-valores significantes são:

(i) $\alpha_1 = 0.09359$

(ii) $\beta_1 = 0.43981$

(iii) $\beta_2 = 0.47789$

(iv) $\gamma_1 = 0.26282$

(v) $\delta = 1.10465$

A equação do modelo APARCH(1,2) pode então ser escrita:

$$h_t^{1.105} = 0.0936(|y_{t-1}| - 0.263y_{t-1})^{1.105} + 0.434h_{t-1}^{1.105} + 0.478h_{t-2}^{1.105} \tag{2}$$

Tabela 5: Estimação dos parâmetros ARMA(0,1)-sGARCH(p,q)

sGARCH	**(1,1)**			**(1,2)**			**(2,1)**		
	t	**ged**	**sstd**	**t**	**ged**	**sstd**	**t**	**ged**	**sstd**
Optimal Parameters									
ma	-0.098168 $se = 0.017145$ $p = 0.00000$	-0.096960 0.017251 0.000000	-0.112715 0.017024 0.000000	-0.097921 0.017301 0.00000	-0.097203 0.017560 0.000000	-0.112471 0.017168 0.00000	-0.098080 0.017171 0.000000	-0.096953 0.017324 0.000000	-0.112917 0.017062 0.000000
omega	0.000002 0.000001 0.01092	0.000003 0.000003 0.229982	0.000002 0.000001 0.020229	0.000003 0.000002 0.19858	0.000004 0.000003 0.132653	0.000002 0.000001 0.12156	0.000002 0.000001 0.001600	0.000003 0.000002 0.029942	0.000002 0.000001 0.005706
alpha1	0.058298 0.007219 0.00000	0.065790 0.018682 0.000429	0.056825 0.006484 0.000000	0.082679 0.018752 0.00001	0.090345 0.018659 0.000001	0.078878 0.013000 0.00000	0.058237 0.016986 0.000607	0.065822 0.016564 0.000071	0.056762 0.017158 0.000939
alpha2							0.000000 0.015385 0.999988	0.000000 0.010452 0.999992	0.000000 0.015894 0.999998
beta1	0.934281 0.008446 0.00000	0.921930 0.023823 0.000000	0.938857 0.007119 0.000000	0.478188 0.048567 0.00000	0.478191 0.123686 0.000111	0.517817 0.019386 0.00000	0.934409 0.007538 0.00000	0.921895 0.015056 0.000000	0.939050 0.006469 0.000000
beta2				0.428980 0.055880 0.00000	0.415301 0.116135 0.000349	0.397503 0.021228 0.00000			
skew			0.855430 0.020575 0.000000			0.857206 0.020582 0.00000			0.855343 0.020673 0.000000
shape	7.918055 0.911316 0.00000	1.484095 0.045101 0.000000	7.851365 0.931125 0.000000	8.053760 0.921804 0.00000	1.490155 0.046235 0.000000	7.980505 0.948296 0.00000	7.929709 0.917295 0.000000	1.484113 0.045156 0.000000	7.838713 0.929659 0.000000
Information Criteria									
Akaike	-5.6161	-5.6036	-5.6276	-5.6171	-5.6049	**-5.6283**	-5.6155	-5.6030	-5.6270
Bayes	-5.6071	-5.5946	-5.6168	-5.6063	-5.5941	**-5.6158**	-5.6047	-5.5922	-5.6145
Shibata	-5.6161	-5.6036	-5.6276	-5.6171	-5.6049	**-5.6283**	-5.6155	-5.6030	-5.6270
Hannan-Quinn	-5.6129	-5.6003	-5.6238	-5.6133	-5.6011	**-5.6238**	-5.6117	-5.5991	-5.6225

Tabela 6: Estimação dos parâmetros ARMA(0,1)-apARCH(p,q)

apARCH	**(1,1)**			**(1,2)**			**(2,1)**		
	t	**ged**	**sstd**	**t**	**ged**	**sstd**	**t**	**ged**	**sstd**
Optimal Parameters									
	-0.096751	-0.096352	-0.109658	-0.095666	-0.095958	-0.10668	-0.098639	-0.096246	-0.105553
ma1	$se = 0.017472$	0.017662	0.017663	0.017482	0.017713	0.017340	0.016321	0.017475	0.015655
	$p = 0.000000$	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000
	0.000085	0.000104	0.000091	0.000111	0.000138	0.00011	0.000224	0.000091	0.000203
omega	0.000097	0.000106	0.000095	0.000109	0.000127	0.000108	0.000229	0.000095	0.000202
	0.383424	0.328484	0.340641	0.307056	0.279974	0.31073	0.327310	0.336017	0.315915
	0.066172	0.074745	0.065136	0.095944	0.103385	0.09359	0.045030	0.074881	0.044392
alpha1	0.015502	0.019384	0.010645	0.013253	0.014604	0.012577	0.006324	0.012536	0.005804
	0.000020	0.000115	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000
							0.027784	0.000000	0.028529
alpha2							0.005619	0.008765	0.005323
							0.000001	1.000000	0.000000
	0.938722	0.925412	0.942880	0.406501	0.421549	0.43981	0.937590	0.924992	0.941482
beta1	0.018373	0.024468	0.011137	0.027311	0.091872	0.016564	0.003929	0.008173	0.003090
	0.000000	0.000000	0.000000	0.000000	0.000004	0.00000	0.000000	0.000000	0.000000
				0.504698	0.476002	0.47789			
beta2				0.026398	0.088789	0.015951			
				0.000000	0.000000	0.00000			
	0.243082	0.244619	0.238672	0.274820	0.277015	0.26282	0.995402	0.236572	1.000000
gamma1	0.106267	0.094330	0.093027	0.085922	0.081265	0.084717	0.000043	0.084964	0.000066
	0.022169	0.009508	0.010299	0.001381	0.000653	0.00192	0.000000	0.005363	0.000000
							-1.000000	0.999973	-1.000000
gamma2							0.000351	0.000101	0.000351
							0.000000	0.000000	0.000000
	1.143155	1.207057	1.066700	1.160642	1.207540	1.10465	0.913771	1.238081	0.879384
delta	0.331045	0.293609	0.264242	0.217052	0.209411	0.212717	0.229214	0.250526	0.215184
	0.000554	0.000039	0.000054	0.000000	0.000000	0.00000	0.000067	0.000001	0.000044
			0.852595			0.85584			0.852942
skew			0.020884			0.020898			0.020715
			0.000000			0.00000			0.000000
	8.157913	1.499904	8.033829	8.349968	1.509520	8.23559	8.125703	1.500096	8.008912
shape	0.983341	0.047157	0.989505	1.018631	0.047156	1.034643	0.959423	0.047833	0.969357
	0.000000	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000
Information Criteria									
Akaike	-5.6190	-5.6069	-5.6309	-5.6207	-5.6090	**-5.6320**	-5.6193	-5.6056	-5.6312
Bayes	-5.6064	-5.5943	-5.6166	-5.6063	-5.5946	**-5.6159**	-5.6031	-5.5895	-5.6132
Shibata	-5.6190	-5.6069	-5.6309	-5.6207	-5.6090	**-5.6320**	-5.6193	-5.6056	-5.6312
Hannan-Quinn	-5.6145	-5.6024	-5.6258	-5.6156	-5.6039	**-5.6262**	-5.6135	-5.5999	-5.6248

Tabela 7: Estimação dos parâmetros ARMA(0,1)-eGARCH(p,q)

eGARCH	**(1,1)**			**(1,2)**			**(2,1)**		
	t	**ged**	**sstd**	**t**	**ged**	**sstd**	**t**	**ged**	**sstd**
Optimal Parameters									
ma1	-0.11007 se =0.017247 p =0.000000	-0.096999 0.017662 0.000000	-0.11007 0.017247 0.000000	-0.096847 0.017601 0e+00	-0.097173 0.017581 0.000000	-0.107260 0.017766 0.0e+00	-0.098013 0.017256 0.000000	-0.099577 0.017235 0.0e+00	-0.102154 0.017460 0.000000
omega	-0.06418 0.005158 0.000000	-0.127539 0.020149 0.000000	-0.06418 0.005158 0.000000	-0.115699 0.026050 9e-06	-0.168251 0.085053 0.047906	-0.091407 0.019903 4.0e-06	-0.051187 0.013267 0.000114	-0.070985 0.002877 0.0e+00	-0.042515 0.008474 0.000001
alpha1	-0.03039 0.008373 0.000284	-0.038298 0.009191 0.000031	-0.03039 0.008373 0.000284	-0.056544 0.012341 5e-06	-0.060945 0.010731 0.000000	-0.051444 0.011964 1.7e-05	-0.135160 0.026666 0.000000	-0.136946 0.026100 0.0e+00	-0.129306 0.026675 0.000001
alpha2							0.110082 0.026910 0.000043	0.111150 0.026578 2.9e-05	0.104528 0.026805 0.000096
beta1	0.99229 0.000628 0.000000	0.984818 0.002423 0.000000	0.99229 0.000628 0.000000	0.475996 0.000519 0e+00	0.502197 0.004211 0.000000	0.504630 0.000295 0.0e+00	0.993959 0.001794 0.000000	0.991576 0.000407 0.0e+00	0.994898 0.001129 0.000000
beta2				0.510258 0.000608 0e+00	0.477751 0.004126 0.000000	0.484361 0.000173 0.0e+00			
gamma1	0.12282 0.015849 0.000000	0.138649 0.036595 0.000151	0.12282 0.015849 0.000000	0.184140 0.019932 0e+00	0.195000 0.007210 0.000000	0.180201 0.019310 0.0e+00	0.232752 0.074046 0.001670	0.248412 0.039410 0.0e+00	0.220087 0.063457 0.000524
gamma2							-0.123001 0.104631 0.239770	-0.136680 0.036496 1.8e-04	-0.107751 0.072488 0.137156
skew	0.85103 0.020660 0.000000		0.85103 0.020660 0.000000			0.853967 0.020730 0.0e+00			0.857095 0.020926 0.000000
shape	7.97745 1.272524 0.000000	1.497104 0.053346 0.000000	7.97745 1.272524 0.000000	8.291511 1.001188 0e+00	1.508058 0.046565 0.000000	8.210659 1.025734 0.0e+00	8.567421 3.902756 0.028148	1.527693 0.045230 0.0e+00	8.473932 2.945088 0.004011
Information Criteria									
Akaike	-5.6322	-5.6075	-5.6322	-5.6224	-5.6104	-5.6340	-5.6273	-5.6159	**-5.6383**
Bayes	-5.6196	-5.5967	-5.6196	-5.6098	-5.5978	-5.6197	-5.6129	-5.6016	**-5.6221**
Shibata	-5.6322	-5.6075	-5.6322	-5.6224	-5.6104	-5.6341	-5.6273	-5.6160	**-5.6383**
Hannan-Quinn	-5.6277	-5.6037	-5.6277	-5.6179	-5.6059	-5.6289	-5.6221	-5.6108	**-5.6325**

Análise dos resíduos

A figura 4 exibe o histograma de distribuição e gráfico de auto-correlação dos resíduos. Observa-se dois *lags* onde a correlação é significante: 5 e 25. Contudo, isso não é suficiente para afirmar que os resíduos não apresentam característica de ruído branco.

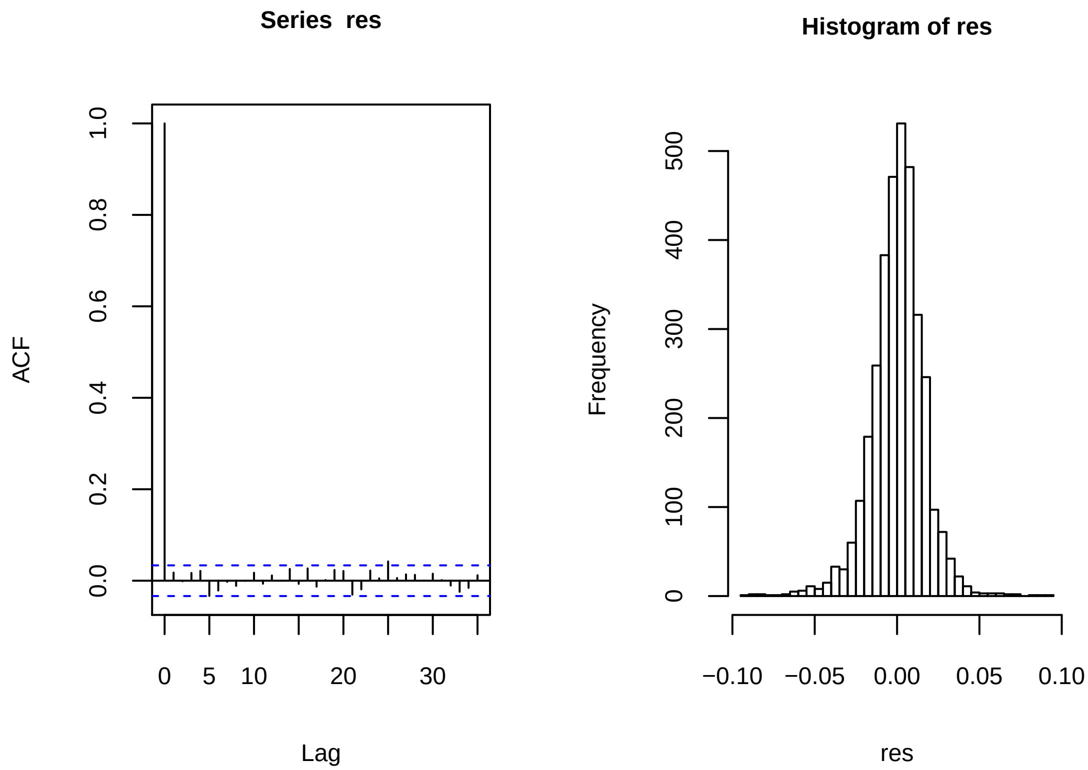
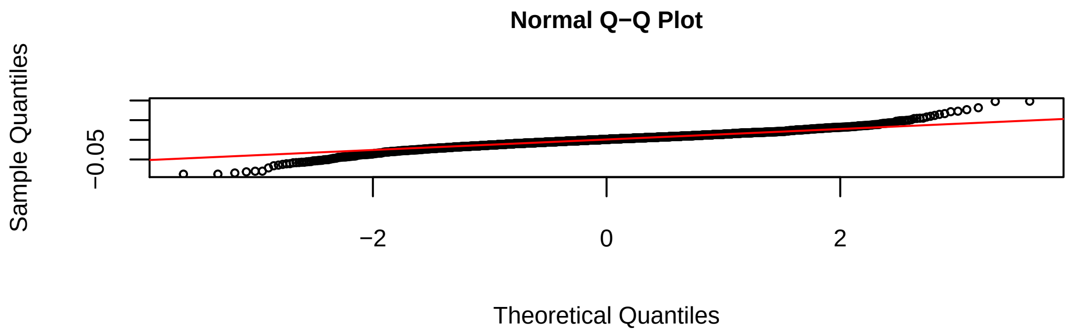


Figura 4: Histograma dos resíduos do ajuste e gráfico de auto-correlação.

O histograma dos resíduos é apresentado à direita da figura 4. Uma análise mais atenta mostra que a curtose da distribuição é 3.3337742 e a assimetria é -0.3175134. Isso significa que é uma distribuição que possui as caudas levemente mais pesadas e assimétrica.

A figura 5 apresenta a comparação da distribuição dos quantis dos resíduos com a curva normal. Vemos um alinhamento satisfatório da distribuição de resíduos e desvios em ambas a pontas. Pela inspeção visual, sugere-se que o modelo fez um trabalho razoável.



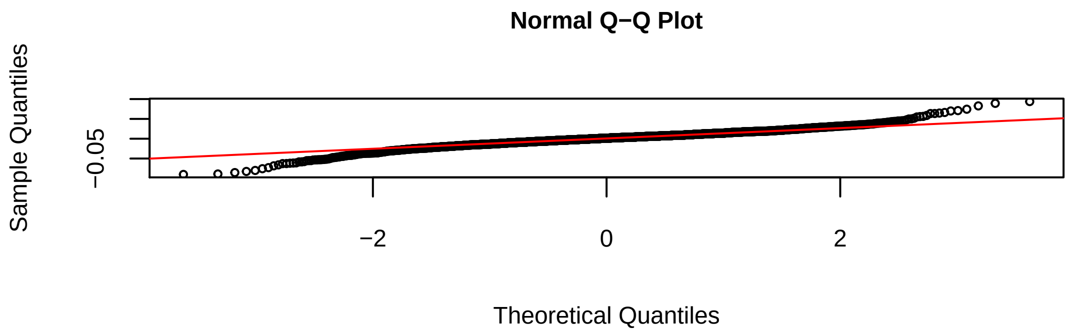


Figura 5: Comparativo da distribuição dos quantis com a normal (qqplot). À esquerda a série de retornos. À direita os resíduos do modelo ARMA(0,1)-APARCH(1,2).

Por fim, a tabela 8 apresenta um sumário dos testes aplicados na série de resíduos. O teste *Box-Pierce* indica correlação linear e que os resíduos não são independentes. O teste Jarque-Bera indica não normalidade. O teste arch aponta que há efeitos de heterocedasticidade condicional. Isso indica que não é possível concluir sobre a fidedignidade do modelo proposto sendo necessário, portanto, uma análise mais adequada de modelos adjacentes.

Tabela 8: Sumário dos testes na série de resíduos.

Teste	Estatística	p-valor
Box.test	1.080138	0.2986667
ArchTest	379.456147	0.0000000
Jarque-Bera	1642.390367	0.0000000

Conclusão

Pelo estudo da série de retornos do primeiro contrato de WTI foram observadas as características das séries financeiras como caudas pesadas, curtose e agrupamento de volatilidade. Foi utilizado o modelo ARMA(0,1) que foi escolhido pelo critério de informação. Para a série de preços foram estimadas três diferentes ordens para os modelos sGARCH, APARCH e EGARCH com três diferentes tipos de distribuição, `std`, `ged` e `sstd`. Foi verificado que o modelo com menor critério de informação é o ARMA(0,1)-APARCh(1,2) com distribuição `sstd`. Apesar de ter se chagada à um modelo, os resíduos não indicam que o mesmo não é incontestável para previsão da volatilidade da série analisada.

References

DING, D. *Modeling of market volatility with APARCH model*. 2011.

SHUMWAY, R. H.; STOFFER, D. S. *Time series analysis and its applications*. [S.l.]: Springer, 2000. v. 9. 375–376 p.