

Relazione di statistica economica

Michi Leonardo

Matricola:7047635

Corso di Laurea Triennale in Statistica, Anno di Corso: Terzo

07/02/2023

Relazione time series economica: analisi della produzione di energia rinnovabile negli Stati Uniti

OBBIETTIVO DELL'ANALISI

L'analisi di una time series economica può fornire informazioni preziose sulle tendenze di un determinato settore o industria. Nel presente studio utilizziamo il modello Sarima per analizzare la produzione di energia rinnovabile negli Stati Uniti dal 1973 ad oggi. L'obiettivo è utilizzare i dati storici per prevedere l'andamento futuro della produzione di energia rinnovabile negli Stati Uniti.

DESCRIZIONE DEI DATI

Il dataset di riferimento, disponibile sul sito dell'EIA (U.S. Energy Information Administration) all'indirizzo: <https://www.eia.gov/totalenergy/data/monthly/>, è costituito da 595 osservazioni con cadenza mensile e l'unità di misura utilizzata è "trillion btu" (british thermal unit).

In Fase di diagnostica, il test ARCH ha riscontrato un'elevata eteroschedasticità, per questo, è stata analizzata la trasformazione in radice quadrata dei dati per stabilizzare la varianza.

ANALISI PRELIMINARE

L'utilizzo di grafici quali quello dell'andamento della serie, l'ACF e la PACF (*figura 1*), rappresenta un metodo efficace per acquisire informazioni sulla time series in esame. Nel corso degli anni la produzione di energie rinnovabili ha avuto un andamento crescente, con un'accelerazione a partire dagli anni duemila. La componente stagionale risulta essere presente e anche il grafico dell'ACF è indicatore di una possibile componente stagionale. Anche il grafico della PACF fornisce indicazioni sul modello. L'autocorrelazione parziale a lag 1, infatti, è molto vicina ad 1, indicatore di non stazionarietà dovuta, questa volta, a $d \geq 1$.

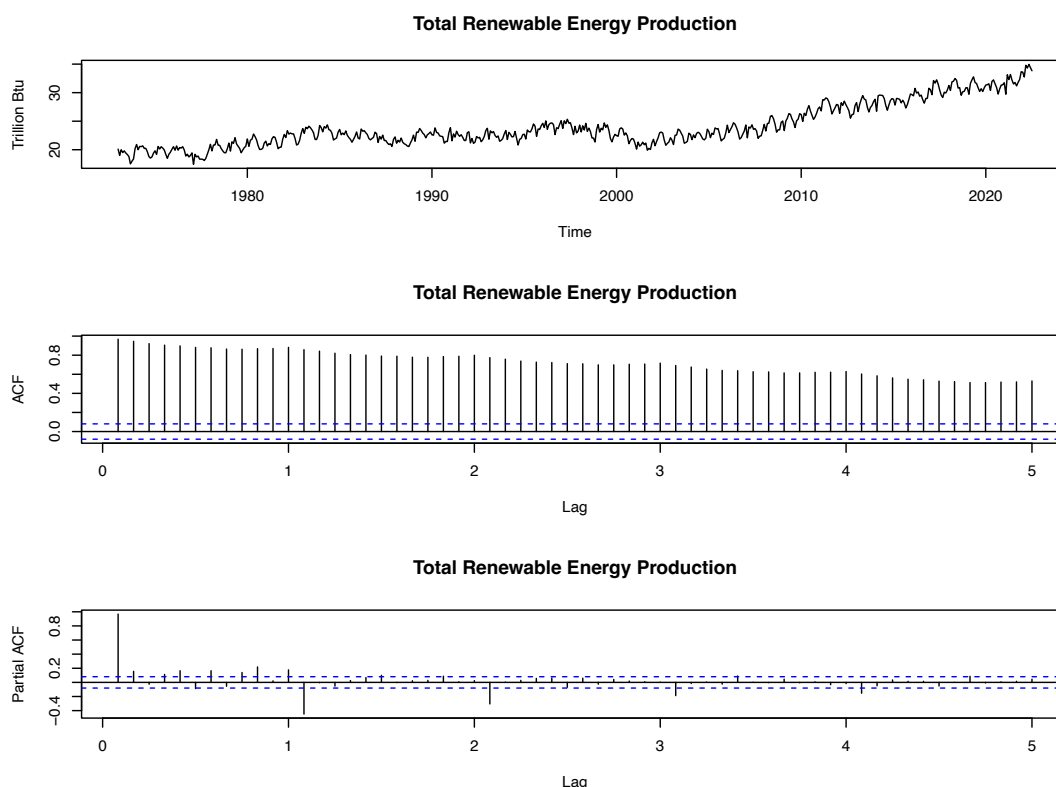


Figura 1: Andamento, ACF, PACF dei dati originali

Per valutare la non stazionarietà è indispensabile controllare la presenza di Radici Unitarie mediante il test ADF (Augmented Dickey-Fuller). Il primo test, su un modello random-walk con trend, porta ad accettare l'Ipotesi nulla e, quindi, a rifiutare la presenza di un possibile trend. Il secondo test, su un modello random-walk con drift, porta ad accettare l'ipotesi nulla e, conseguentemente, a rifiutare la presenza di un possibile drift. L'ultimo test accetta nuovamente l'ipotesi nulla e, pertanto, la presenza di radici unitarie. Un ulteriore test per la presenza di radici unitarie è il KPSS che conduce alla medesima conclusione, anche se le ipotesi sono invertire rispetto all'ADF (*figura 2*).

```

Unit root analysis

-----
Test1: ADF with trend
      tau3      phi2      phi3
statistic -0.572962 2.53156 1.231148
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
phi2  6.09  4.68  4.03
phi3  8.27  6.25  5.34

-----
Test2: ADF with drift
      tau2      phi1
statistic 0.9776132 3.041672
      1pct  5pct 10pct
tau2 -3.43 -2.86 -2.57
phi1  6.43  4.59  3.78

-----
Test2: ADF with none
      tau1
statistic 2.385422
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62

-----
KPSS with tau
[1] 0.5452385
      10pct  5pct 2.5pct  1pct
critical values 0.119 0.146 0.176 0.216

-----
KPSS with mu
[1] 2.506887
      10pct  5pct 2.5pct  1pct
critical values 0.347 0.463 0.574 0.739

```

figura 2: ADF test e KPSS test

il test HEGY conferma l'ipotesi, derivante dalla PACF che d sia uguale ad 1 (figura 3).

HEGY test for unit roots

```

data:  y

      statistic p-value
t_1      0.8304  0.9929
t_2     -2.6907  0.0063 **
F_3:4     6.4186  0.0018 **
F_5:6     7.3287  8e-04 ***
F_7:8    14.3803      0 ***
F_9:10    17.056      0 ***
F_11:12   3.4097  0.0367 *
F_2:12     9.7924      0 ***
F_1:12     9.0373      0 ***

```

Figura 3: test HEGY

Per controllare D , effettuo le analisi sulle differenze prime. Il grafico dell'ACF (figura 3) presenta un decadimento lineare a lag multipli di dodici, indicatore di una possibile presenza di un'altra radice unitaria e che quindi, verosimilmente, D sia uguale ad 1. Il test ADF sulle differenze prime conferma quanto detto.

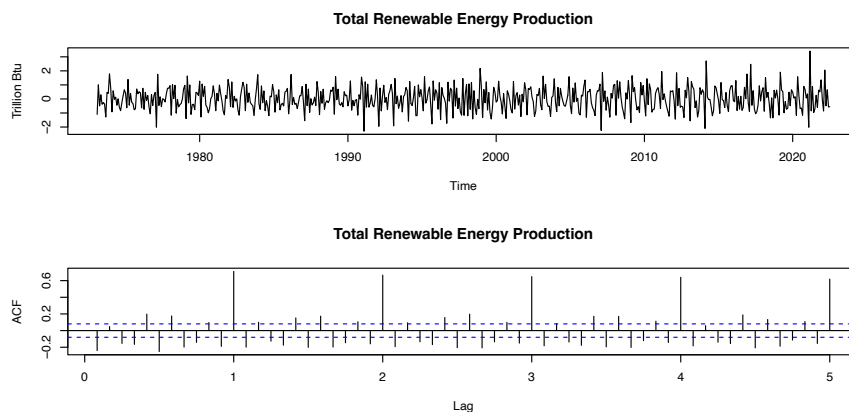


Figura 3: ACF differenze prime

APPLICAZIONE

In base alle analisi effettuate, il modello che più efficacemente rappresenta i dati risulta essere un $ARIMA(0,1,2)X(0,1,1)_{12}$ (figura 4).

Regression with $ARIMA(0,1,2)(0,1,1)[12]$ errors

Coefficients:

	ma1	ma2	sma1	sh	A0578
	-0.1962	-0.2158	-0.8759	0.1039	-1.4815
s.e.	0.0408	0.0415	0.0326	0.0481	0.3563

$\sigma^2 = 0.2108$: log likelihood = -379.06

AIC=770.11 AICc=770.26 BIC=796.31

Figura 4: $ARIMA(0,1,2)X(0,1,1)_{12}$

Il modello ARIMA presenta un effetto calendario, le festività singole, che influenza la regolarità della produzione di energia rinnovabile. Oltre a ciò, include un'anomalia, ovvero una variazione dei dati, non totalmente compatibile col processo considerato. Questa anomalia è causata da un evento esogeno, cioè non legato alla time series dei dati in esame. Nel caso specifico, l'anomalia è di tipo AO (additive outlier), ovvero causata da un evento singolo e isolato e dovuta, verosimilmente, alla seconda ondata di Covid-19. A febbraio 2021, i dati mostrano un valore della produzione di energia rinnovabile minore di quello aspettato (figura 5). Ciò, è stato probabilmente causato dalle restrizioni imposte per contenere la diffusione del virus, che hanno avuto un impatto negativo anche sulla produzione di energia rinnovabile.

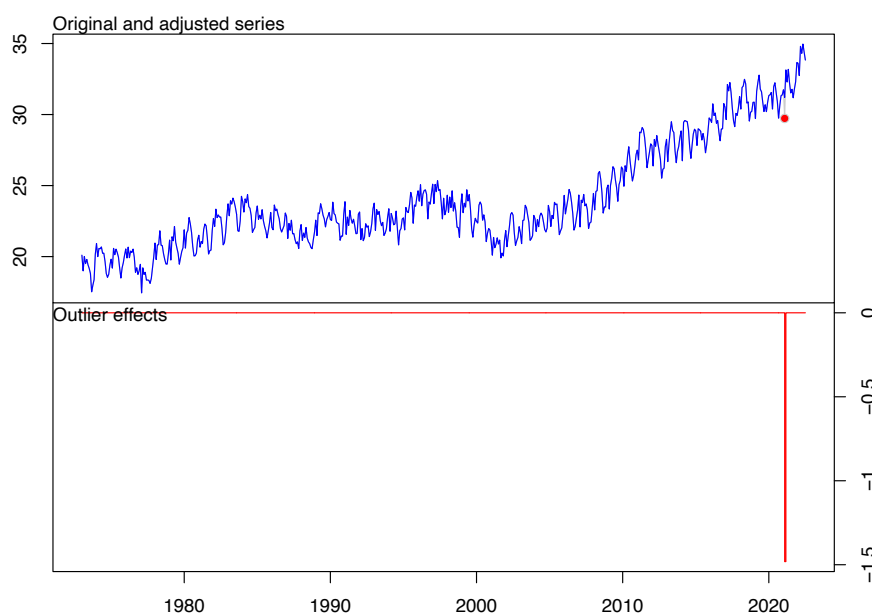


Figura 5: effetto dell'anomalia sull'andamento previsto dei dati,

Tra le ipotesi imposte sul modello in fase di stima troviamo: i residui devono avere media uguale a zero, identica distribuzione normale, indipendenza, varianza costante. Tutte ipotesi che devono essere verificate in fase di diagnostica.

Il grafico dei residui (figura 6) mostra un andamento molto simile a un processo white-noise. Il grafico dell'ACF e della PACF portano alla stessa conclusione, in quanto non presentano lag significativi tranne uno. Ciò conferma l'uguale distribuzione dei residui e, di conseguenza, che il modello stimato è corretto.

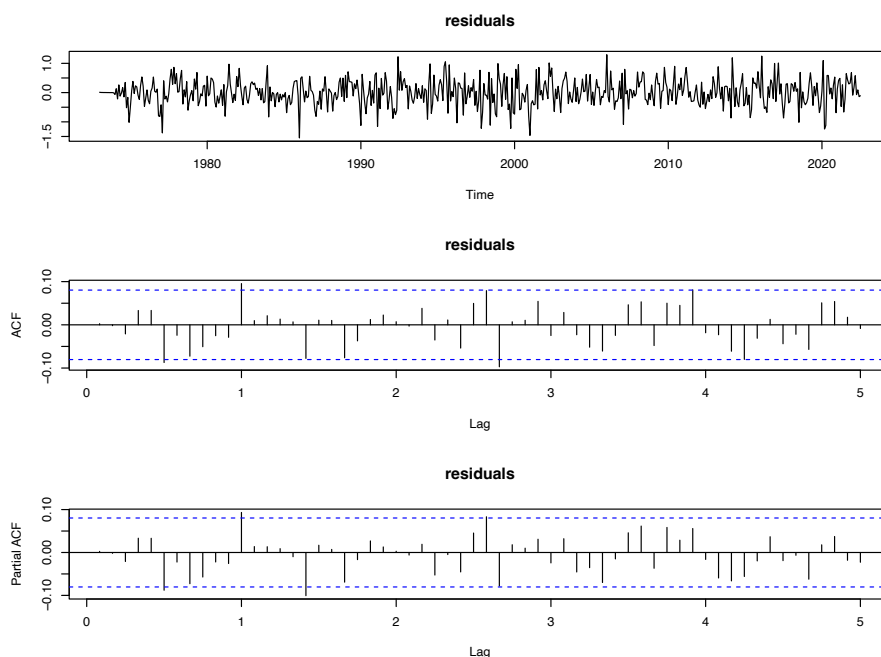


Figura 6: grafico, ACF, PACF dei residui

Dal test Ljung-Box sui residui (figura 7) emerge tuttavia che, essi risultano autocorrelati in blocco. Questo indica che il modello non sta catturando tutte le informazioni presenti nei dati, anche se, all'aumentare dei lag, l'autocorrelazione diminuisce gradualmente.

Ljung-Box of residuals at different lags

lag	6	7	10	15	20	25
statistic	6.191205	6.544906	11.6688	18.16866	25.63278	26.9377
parameter	1	2	5	10	15	20
p.value	0.01283867	0.0379133	0.03961888	0.05218435	0.0420643	0.137026

Figura 7: Ljung-Box test

Per controllare l'omoschedasticità è necessario controllare l'ACF e la PACF dei residui al quadrato e in valore assoluto (figura 8), che presentano qualche lag significativo ma meno rispetto ai grafici dei residui prima della trasformazione. I lag significativi comportano che i residui al quadrato e in valore assoluto possano essere correlati serialmente e che, quindi, “muovendosi”, anche la varianza non sia stabile.

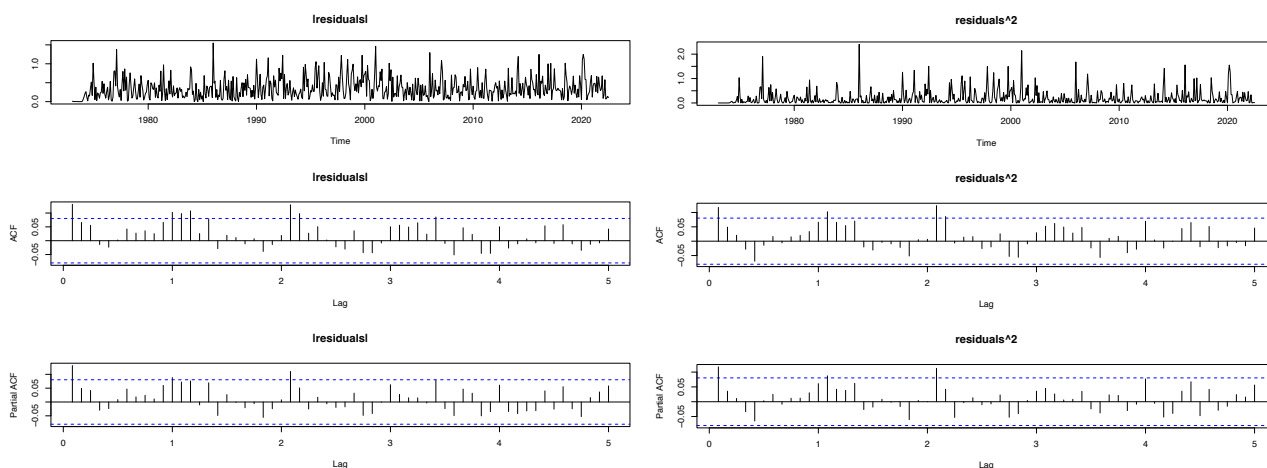


Figura 8: grafico, ACF, PACF dei residui al quadrato e in valore assoluto

Il test ARCH è usato per verificare la presenza o meno di eteroschedasticità condizionata. I risultati (*figura 9*) mostrano assenza di eteroschedasticità, a tutti i lag per cui è testata, a seguito della trasformazione in radice quadrata.

ARCH test on demeaned log-returns

lag	1	2	3	6	12	24
statistic	2.102298	2.113171	2.114102	3.98042	6.312656	15.11108
parameter	1	2	3	6	12	24
p.value	0.147078	0.3476407	0.5490651	0.6793263	0.8995082	0.9174662

Figura 9: test ARCH

L'ultima diagnostica relativa alla distribuzione normale dei residui è effettuata attraverso l'istogramma dei dati, il Q-Q plot e il test di Shapiro-Wilk (*figura 10*). Nell'istogramma, i residui sembrano seguire una distribuzione gaussiana. Inoltre, nel Q-Q plot, i residui sono ben distribuiti lungo la bisettrice del quadrante. A supporto di questa osservazione, il test di Shapiro-Wilk consente di non rigettare l'ipotesi nulla di distribuzione normale.

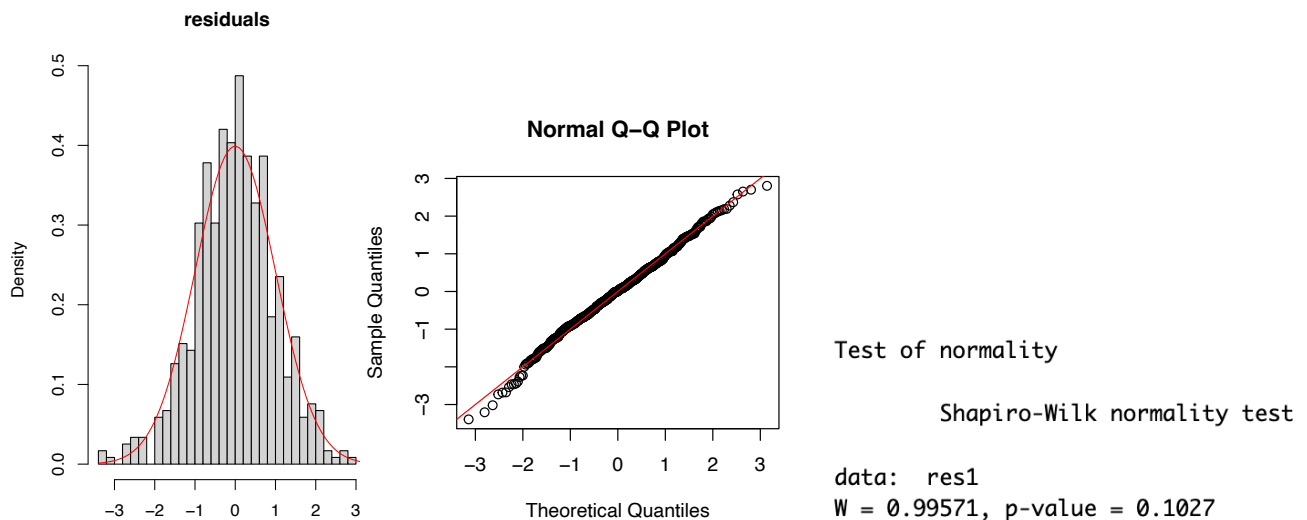


figura 10: istogramma, Q-Q plot, Shapiro-Wilk test

Completata la diagnostica, si procede con le previsioni. In primo luogo, si analizzano le previsioni ex-post, che consentono di confrontare le previsioni, per valutarne la qualità, con l'ultima parte della serie temporale disponibile. Si effettuano 12 previsioni, spostando ogni volta l'origine dei dati. Per effettuare un confronto, si stima anche il modello naive e le bande di confidenza.

Il modello in cui è presente l'effetto calendario e l'anomalia risulta essere il migliore, in quanto mostra le più basse misure di errore (*figura 11*). Il modello è anche notevolmente migliore rispetto al modello naive.

model	h	ME	MAE	RMSE	MPE	MAPE	RMSPE	ScMAE	ScRMSE	LLE	ScLLE
Arima	1	16.21910	20.84930	25.76130	0.01535694	0.01934429	0.02407817	0.1786574	0.2094527	2.000304	0.9965941
Arima + Calendar	1	15.18817	19.89131	24.05766	0.01439892	0.01845572	0.02246607	0.1704484	0.1956013	2.000263	0.9965738
Arima + Outliers	1	15.81051	20.46903	25.65655	0.01483724	0.01884884	0.02367940	0.1753989	0.2086010	2.000294	0.9965891
Arima + Calendar + Outliers	1	15.03057	19.74407	24.35456	0.01411903	0.01818333	0.02244591	0.1691867	0.1980152	2.000263	0.9965738
Naive	1	116.69992	116.69992	122.99339	0.10428329	0.10428329	0.10932803	1.0000000	1.0000000	2.007140	1.0000000

Figura 11: errori stimati

Il Grafico dell'andamento delle previsioni sui dati passati (*figura 12*) conferma nuovamente la bontà del nostro modello, essendo molto vicine ai valori reali, tutti all'interno delle bande di confidenza.

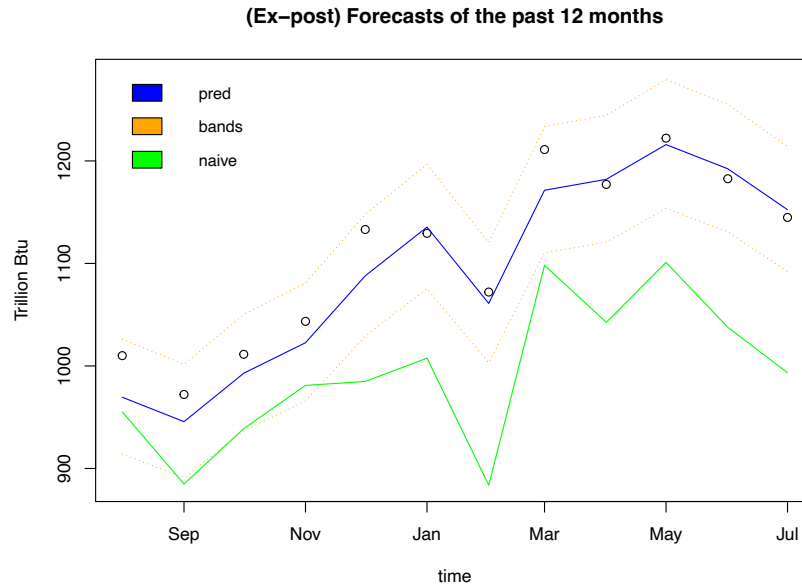


figura 12: previsione ex-post

L'obiettivo principale dell'analisi è la stima della produzione futura di energia rinnovabile negli Stati Uniti. Si prevede che, nei mesi da settembre a dicembre compresi, la produzione di energia rinnovabile crescerà gradualmente, fino ad arrivare ad un decadimento della produzione a febbraio e ad una repentina ripresa a marzo. Dopo un anno, la produzione dovrebbe essere maggiore di quasi il 10% rispetto all'anno precedente (*figura 13*).

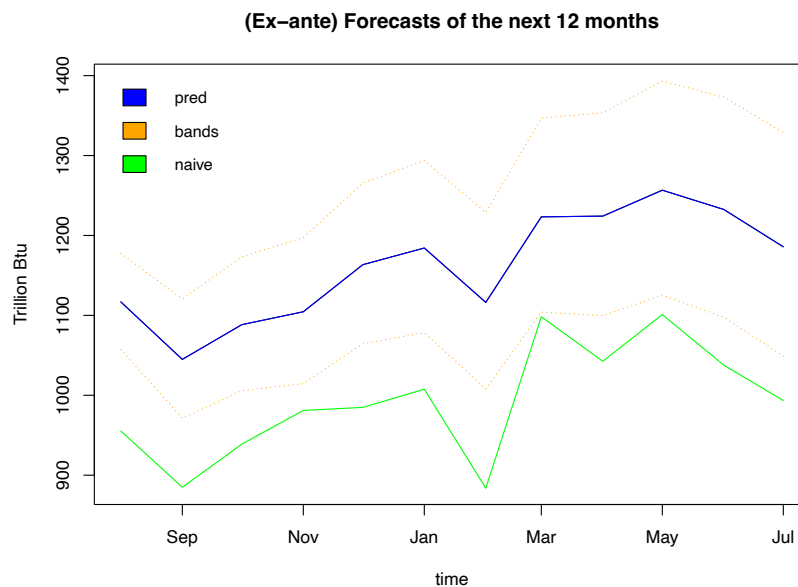


Figura 13: previsione ex-ante

CONCLUSIONI

L'analisi preliminare ha portato ad individuare una componente stagionale e la presenza di radici unitarie, confermate dal test ADF. Il modello migliore è risultato l'ARIMA (0,1,2)X(0,1,1)₁₂ con effetto di calendario e un'anomalia legata al covid-19.

In fase di diagnostica, è stata verificata l'ipotesi che i residui si distribuiscono come un normal white-noise. Qualche problema è arrivato col test di Ljung-Box per testare l'autocorrelazione dei residui, mentre il test ARCH ha accettato l'ipotesi nulla di omoschedasticità.

In fase di previsione ex-post, il modello è risultato soddisfacente, dato che nessuno dei valori reali esce fuori dalle bande di previsione. Questo ha permesso di effettuare la previsione ex-ante per i prossimi mesi.

Relazione Times Series Finanziaria: Azioni di Generali Assicurazioni

OBIETTIVO ANALISI

La times series riguarda prezzi delle azioni di Assicurazioni Generali S.P.A. L'obiettivo è studiare il comportamento della varianza condizionata e della volatilità e, conseguentemente ottenerne previsioni.

DESCRIZIONE DEI DATI

I dati della time series provengono dal sito Yahoo Finance, al link: <https://it.finance.yahoo.com/quote/G.MI/history?p=G.MI>. Essi riguardano i prezzi, in dollari, delle azioni di Assicurazioni Generali S.P.A., tra il 02/01/2015 e il 18/01/2023, con cadenza giornaliera.

Il dataset comprende il prezzo di apertura, di chiusura, di chiusura aggiustata, massimo giornaliero, minimo giornaliero e il volume giornaliero di azioni scambiate.

Inizialmente, i dati presi in considerazione partivano dal 02/01/2013, ma sono stati tagliati in seguito al Nyblom test per la verifica della stabilità dei parametri.

ANALISI PRELIMINARI

I grafici delle chiusure e delle chiusure aggiustate (*figura 14*) non mostrano grandi differenze nell'andamento. le chiusure aggiustate tengono conto dei dividendi. Il *Close* in scala logaritmica è utilizzato poiché le variazioni sono interpretabili con più facilità e i rendimenti hanno come riferimento i prezzi su scala logaritmica. I prezzi decrescono velocemente nel 2016 e nel 2020, con molta probabilità in conseguenza del covid-19, ma crescono fino a raggiungere il massimo nel 2022. Hanno un andamento, in qualche misura, simile a un RW.

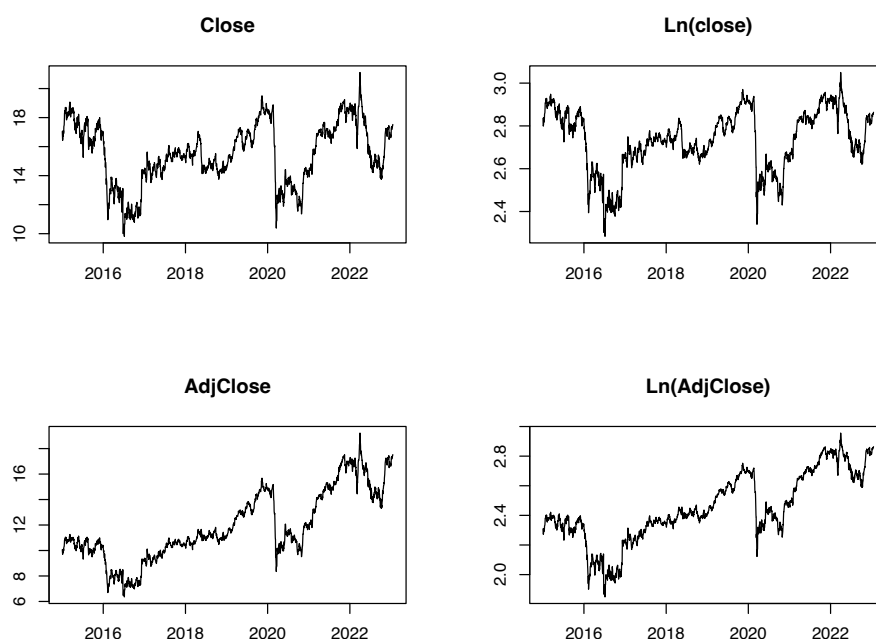


Figura 14: andamento dei prezzi e dei prezzi aggiustati su scala tradizionale e logaritmica

L'ACF e la PACF (*figura 15*) sono tipiche di quasi tutte le serie finanziarie. L'ACF presenta, infatti, decadimento lineare e la PACF ha il primo lag pari quasi a 1, chiari indizi di non stazionarietà.

Il test ADF conferma tali osservazioni.

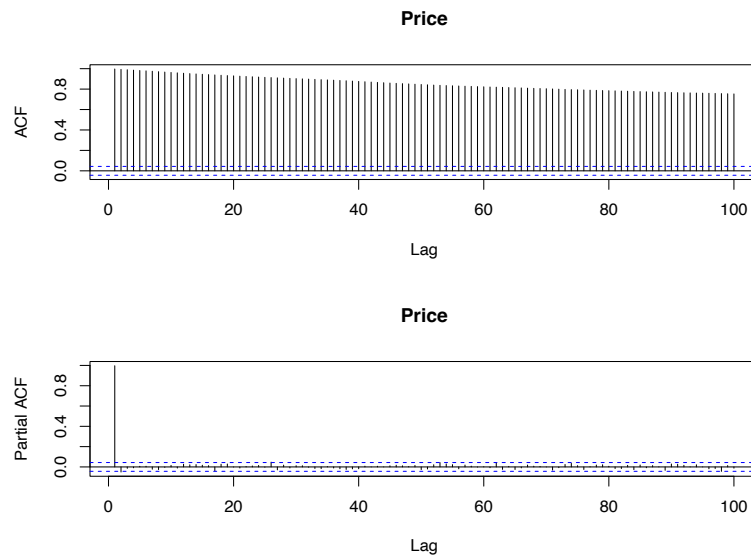


Figura 15: ACF e PACF

Poiché i prezzi non sono stazionari, si analizza una loro trasformazione che sia stazionaria: il logaritmo dei rendimenti (*figura 16*). tale variabile si muove intorno a una media zero. La differenza sostanziale rispetto ad un white noise è che il logaritmo dei rendimenti presenta una variabilità che cambia (si vedano i picchi nel 2016 e nel 2020): è il fenomeno del volatility clustering.

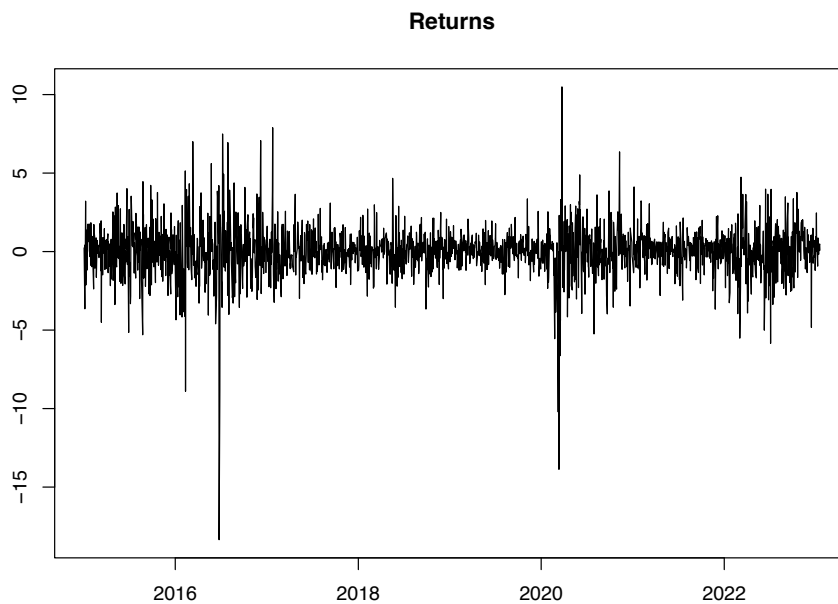


figura 16: logaritmo dei rendimenti

L'affinità a un white noise è illustrata dall'ACF (*figura 17*). Nonostante lag 1 sia di poco significativo, la correlazione seriale è sostanzialmente assente.

Il test di Ljung-Box (*figura 17*) conferma che i test eseguiti sui primi lag risultano non significativi tranne il primo.

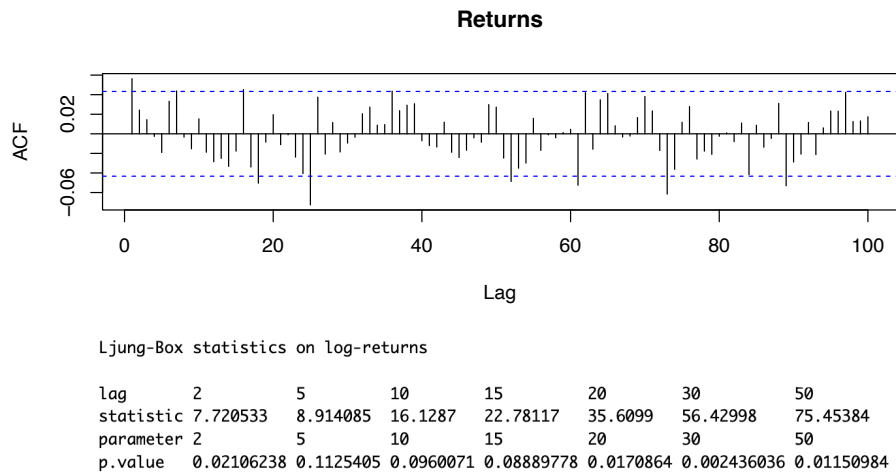


figura 17: ACF e Ljung-Box dei rendimenti logaritmici

i rendimenti logaritmici non risultano, tuttavia, essere generati da variabili casuali indipendenti e identicamente distribuite: il test BDS (figura 18), infatti, riporta p-value altamente minori di 0.05. La causa di tale risultato è, verosimilmente, attribuibile al volatility clustering.

BDS Test

data: x1

Embedding dimension = 2 3 4 5

Epsilon for close points = 0.5247 1.0494 1.5741 2.0987

Standard Normal =

	[0.5247]	[1.0494]	[1.5741]	[2.0987]
[2]	3.9216	3.4569	2.7903	2.1684
[3]	3.5534	3.5194	2.8646	2.2930
[4]	3.4870	3.5018	3.0192	2.6705
[5]	3.9053	3.5575	3.2092	2.9439

p-value =

	[0.5247]	[1.0494]	[1.5741]	[2.0987]
[2]	1e-04	5e-04	0.0053	0.0301
[3]	4e-04	4e-04	0.0042	0.0218
[4]	5e-04	5e-04	0.0025	0.0076
[5]	1e-04	4e-04	0.0013	0.0032

Figura 18: BDS test

Il grafico dell'ACF dei rendimenti al quadrato e, in particolare, di quelli in valore assoluto (figura 19), certifica il volatility clustering in termini di correlazione. Anche il test ARCH mostra, a ogni lag per cui è testato, p-value altamente significativi, portando al rifiuto dell'ipotesi nulla di omoschedasticità.

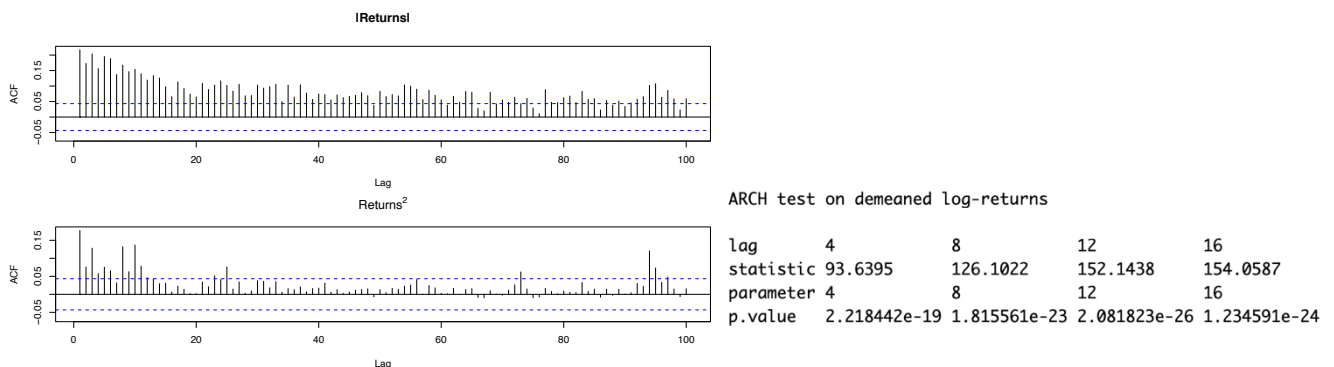


Figura 19: |ACF|, ACF², test ARCH

L'istogramma dei rendimenti (*figura 20*) ha una distribuzione leptocurtica. Questo fenomeno è dovuto al volatility clustering: i periodi di alta volatilità, infatti, tendono a formare code più “massicce”; al contrario, i periodi di bassa volatilità tendono ad addensare la parte centrale della distribuzione. Il Q-Q plot porta alla stessa conclusione. Lo Shapiro-Wilk test supporta quanto detto.

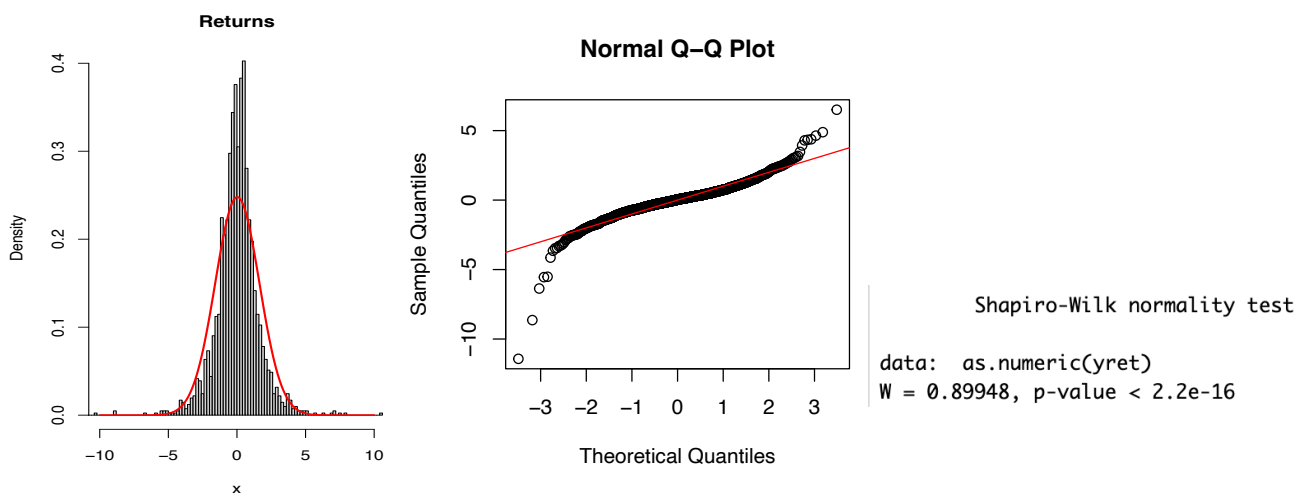


figura 20: istogramma, Q-Q plot, Shapiro-Wilk test

APPLICAZIONE

Il modello ARMA migliore, a causa della scarsa correlazione seriale, risulta essere un ARMA (0,0) distribuzione della componente di errore “std” per cogliere la leptocurtosi (Akaike=3.533, Bayes=3.542).

Per catturare l’eteroschedasticità (fondamentale in ambito finanziario, perché la volatilità è sinonimo di rischio e prevederla ha la sua importanza) che arma non riuscirebbe comunque a spiegare, vengono usati i modelli della famiglia ARCH.

Il modello sGARCH(1,1) (*figura 21*) presenta Akaike e Bayes minori rispetto a quelli del modello ARMA. Il parametro μ è significativo, al contrario di ω che risulta, come spesso accade, vicino a zero. α_1 è minore di 0.1 e β_1 è maggiore di 0.9 (valori classici).

GARCH on log-returns

Information Criteria

Akaike	3.406308
Bayes	3.420034
Shibata	3.406296
Hannan-Quinn	3.411342

Matcoef

	Estimate	Std. Error	t value	Pr(> t)
mu	0.08252740	0.02416277	3.415477	6.367041e-04
omega	0.02645090	0.01185606	2.231002	2.568100e-02
alpha1	0.07090819	0.01789936	3.961494	7.448225e-05
beta1	0.92174381	0.01897705	48.571490	0.000000e+00
shape	4.50916623	0.45219178	9.971801	0.000000e+00

Robust matcoef

	Estimate	Std. Error	t value	Pr(> t)
mu	0.08252740	0.02369132	3.483445	0.0004950051
omega	0.02645090	0.01568118	1.686793	0.0916432185
alpha1	0.07090819	0.02911419	2.435520	0.0148704148
beta1	0.92174381	0.02997585	30.749548	0.0000000000
shape	4.50916623	0.45704081	9.866004	0.0000000000

Figura 21: sGARCH(1,1)

Il Sign Bias Test (*figura 22*) controlla la presenza di asimmetria. Il Sign bias, il Positive Sign bias e il Joint Effect risultano significativi e, quindi, confermano la presenza di asimmetria.

Sign bias test			
	t-value	prob	sig
Sign Bias	2.5062929	0.012277569	**
Negative Sign Bias	0.0746692	0.940485218	
Positive Sign Bias	2.6953798	0.007088449	***
Joint Effect	9.6134250	0.022154825	**

figura 22: Sign Bias Test

per incorporare l'effetto leverage nel modello si modifica l'equazione della varianza condizionata, portando ad un GJR-GARCH (1,1). Akaike=3.390 e Bayes=3.407 denotano un miglioramento rispetto a sGARCH.

Il GJR-GARCH con il Variance Targeting, il processo in base al quale si vincola la stima dei parametri a riprodurre l'unconditional variance stimata, produce leggeri miglioramenti rispetto a quello senza, Akaike=3.389 e Bayes=3.403; inoltre, presenta degli standard errors minori rispetto alla versione senza il Variance Targeting.

Il modello stimato tramite T-GARCH(1,1) (*figura 23*) risulta essere il migliore tra quelli finora proposti. L'Akaike e il Bayes diminuiscono rispetto agli altri modelli testati.

T-GARCH				
Akaike	3.379801			
Bayes	3.396273			
Shibata	3.379784			
Hannan-Quinn	3.385842			
	Estimate	Std. Error	t value	Pr(> t)
mu	5.617388e-02	0.025687516	2.186816e+00	2.875593e-02
omega	1.718133e-02	0.004682116	3.669564e+00	2.429644e-04
alpha1	6.580866e-07	0.011013911	5.975049e-05	9.999523e-01
beta1	9.494027e-01	0.006036775	1.572698e+02	0.000000e+00
gamma1	1.024751e-01	0.015369310	6.667518e+00	2.601652e-11
shape	4.901159e+00	0.537122903	9.124837e+00	0.000000e+00

Figura 23: T-GARCH (1,1)

La diagnostica (*figura 24*) certifica un netto miglioramento rispetto al modello ARMA. L'ACF dei residui standardizzati non presenta più correlazione seriale: il primo lag significativo si trova al lag 25. L'ACF dei residui in valore assoluto è quella in cui si nota il miglioramento maggiore, con pochissimi lag significativamente diversi da zero. Allo stesso modo si comporta l'ACF dei residui al quadrato.

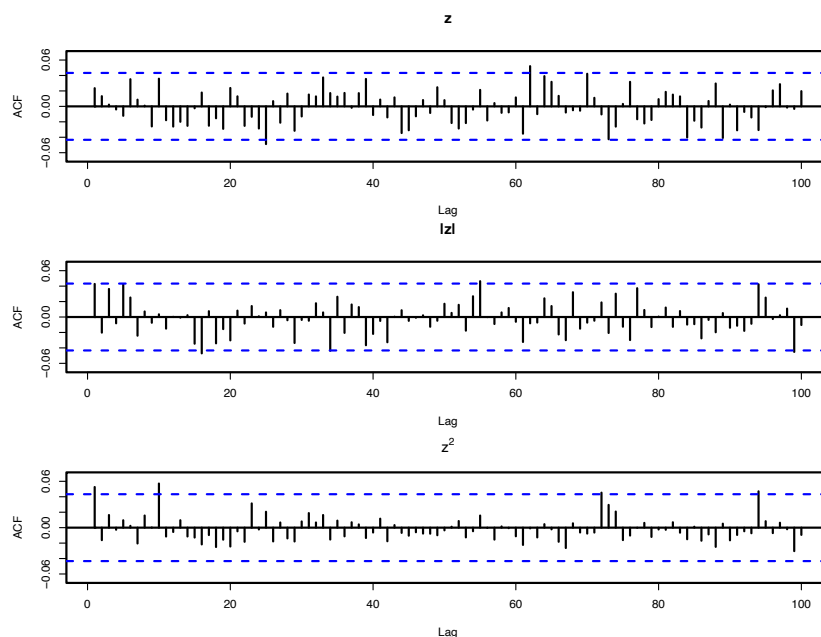


Figura 24: ACF dei residui, in valore assoluto, al quadrato

Il test ARCH (*figura 25*) conferma quanto sopra.

ARCH test on standardized residuals

statistic	7.705597	9.064164	16.91246	18.46407
parameter	4	8	12	16
p.value	0.1029777	0.3369121	0.1529199	0.2974308

figura 25: Test ARCH

Istogramma e Q-Q plot (*figura 26*) confermano che la T-student si adatta molto meglio ai dati di una normale standard. Il BDS test (*figura 27*) supporta l'indipendenza e l'identica distribuzione dei residui standardizzati.

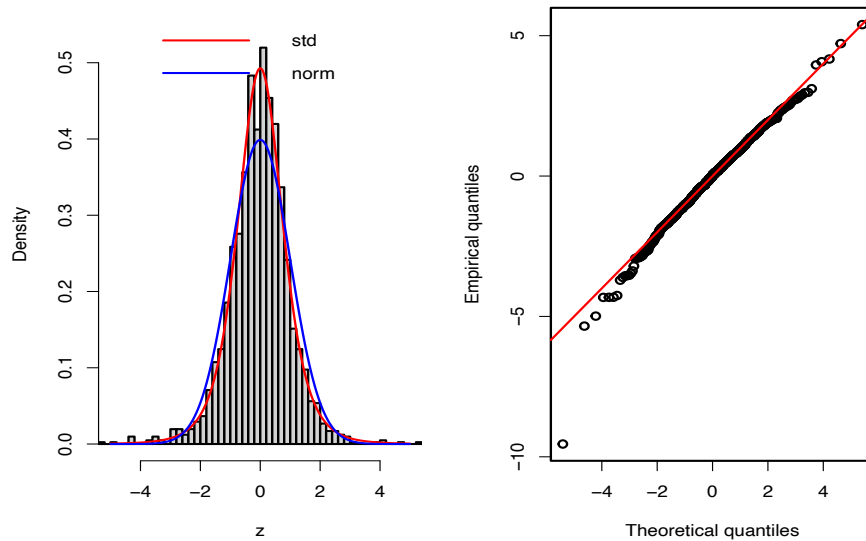


figura 26: istogramma e Q-Q plot

BDS Test

data: x1

Embedding dimension = 2 3 4

Epsilon for close points = 0.5664 1.1328 1.6992 2.2655

Standard Normal =

	[0.5664]	[1.1328]	[1.6992]	[2.2655]
[2]	0.7151	0.4410	0.0672	-0.2378
[3]	-0.0207	-0.2556	-0.6422	-0.7860
[4]	-0.7455	-0.7987	-0.9711	-0.8822

p-value =

	[0.5664]	[1.1328]	[1.6992]	[2.2655]
[2]	0.4746	0.6592	0.9465	0.8121
[3]	0.9835	0.7983	0.5208	0.4319
[4]	0.4560	0.4245	0.3315	0.3777

figura 27: BDS test

Anche in questo caso il Sign bias, il Positive Sign bias e il Joint Effect risultano significativi e, quindi, confermano la presenza di asimmetria (*figura 28*).

Sign bias test

	t-value	prob	sig
Sign Bias	2.3601043	1.836328e-02	**
Negative Sign Bias	0.3254187	7.448975e-01	
Positive Sign Bias	4.4115889	1.079502e-05	***
Joint Effect	19.5692949	2.084526e-04	***

Figura 28: Sign Bias Test

La News Impact Curve evidenzia una netta differenza tra sGARCH, GJR-GARCH e T-GARCH (figura 29). La curva dell'sGARCH ha un andamento simmetrico a forma di parabola, mentre quella del GJR-GARCH e del T-GARCH hanno un andamento asimmetrico. In particolare, la curva del GJR-GARCH appare più alta a sinistra e più bassa a destra rispetto alla curva sGARCH. Questo indica che la volatilità del titolo azionario tende ad essere decisamente più alta quando il titolo è in fase negativa, rispetto a quando è in fase positiva. La curva del T-GARCH ha una forma diversa: è meno incurvata e presenta uno spigolo in corrispondenza del minimo.

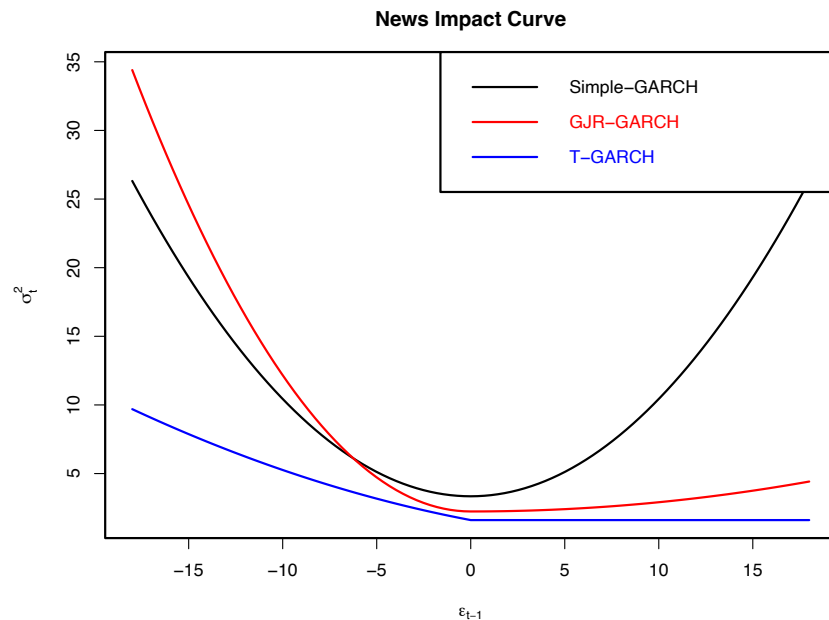


Figura 29. News Impact Curve

L'IGARCH (1,1) (figura 30) presenta Akaike=3.405 e Bayes=3.416, superiori ai precedenti modelli.

Information Criteria
 Akaike 3.405863
 Bayes 3.416844
 Shibata 3.405855
 Hannan-Quinn 3.409890

Robust matcoef

	Estimate	Std. Error	t value	Pr(> t)
mu	0.08392144	0.02411921	3.479444	0.0005024552
omega	0.02151285	0.01197739	1.796122	0.0724751826
alpha1	0.07690517	0.02475598	3.106529	0.0018929757
beta1	0.92309483	NA	NA	NA
shape	4.25388919	0.34847145	12.207282	0.0000000000

figura 30: IGARCH(1,1)

Il benchmark Garman-Klass (figura 31) ha un comportamento simile ai rendimenti in valore assoluto, ma è meno rumoroso. L'accostamento tra il Garman-Klass e la volatilità stimata dai modelli T-GARCH e IGARCH mostra un andamento simile, con le differenze maggiori che si trovano nei periodi di volatilità elevata in cui i modelli stimano meno volatilità del Garman-Klass.

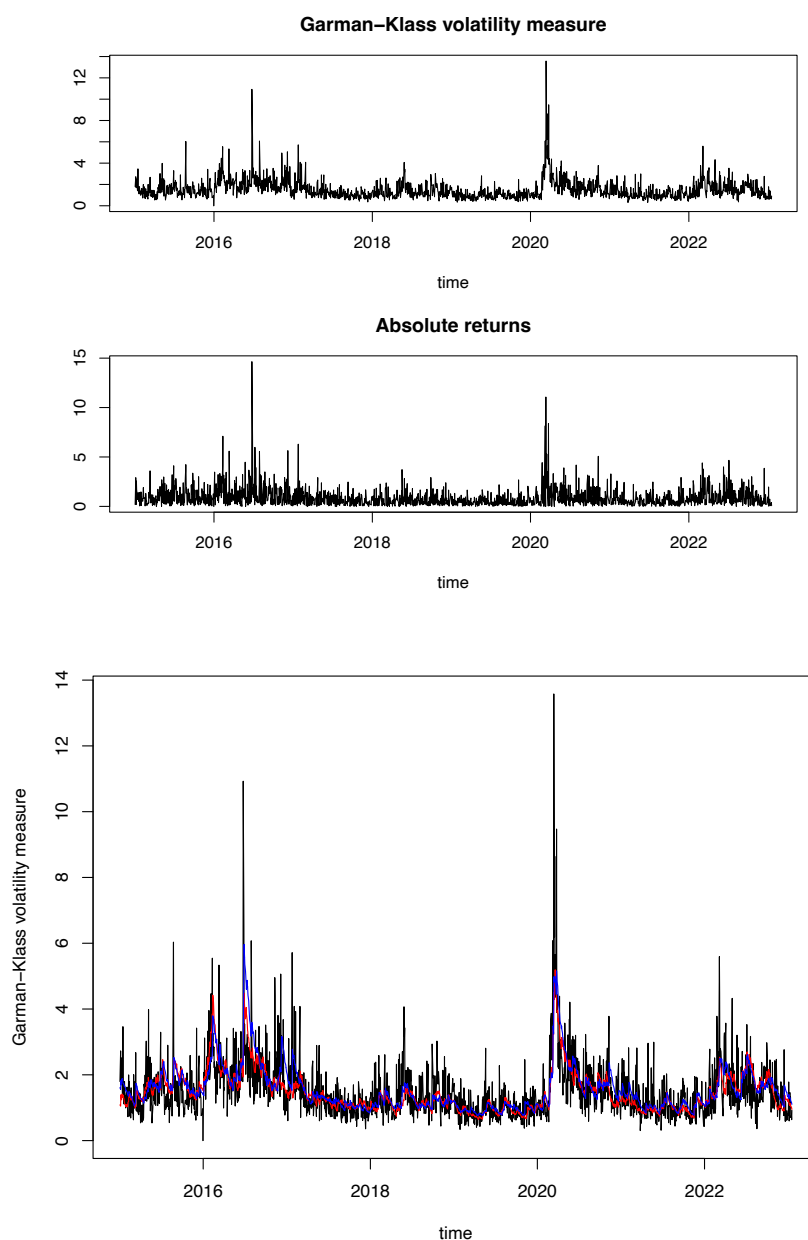


Figura 31: Grafico della volatilità con Garman-Klass

Per quanto riguarda le misure di errore (figura 32), il modello che presenta gli indici più bassi è il T-GARCH, sia per quanto riguarda la volatilità (scMAE=0.7365), sia per la varianza (scMAE=0.7947). Inoltre, l'ME in valore assoluto e il MAE non sono vicini e, quindi non c'è presenza di bias.

Error measures							
measure	model	ME	MAE	RMSE	ScMAE	ScRMSE	
1 Volatility	GARCH	-0.005558245	0.4815437	0.7371690	0.7644761	0.8147083	
2 Volatility	GJR-GARCH	-0.003821721	0.4844976	0.7556867	0.7691656	0.8351737	
3 Volatility	T-GARCH	0.032182751	0.4639165	0.7189152	0.7364920	0.7945344	
4 Volatility	IGARCH	-0.048298179	0.4950811	0.7469799	0.7859675	0.8255511	
5 Variance	GARCH	0.369166505	1.9505119	5.9035552	0.8398574	0.8879342	
6 Variance	GJR-GARCH	0.252145032	2.0167559	5.9521936	0.8683809	0.8952498	
7 Variance	T-GARCH	0.520562285	1.8457146	5.8484839	0.7947334	0.8796512	
8 Variance	IGARCH	0.184080297	2.0217645	5.9125290	0.8705375	0.8892840	

Figura 32: Misure di errore

La diagnostica di Mincer-Zarnowitz controlla che il modello stimato non produca previsioni distorte. In particolar modo, nel T-GARCH, effettuando la diagnostica mediante statistica T, risulta che sia β_0 che β_1 accettano l'ipotesi nulla di assenza di bias sistematico, rispettivamente con un HAC.pvalue di 0.1219 e 0.2392. Il test F congiunto su entrambi i parametri presenta un HAC.pvalue=0.027, leggermente significativo ad un livello di confidenza del 95%.

Il test di Diebold-Mariano (*figura 33*) effettua un confronto tra due modelli alla volta e verifica quale produce previsioni significativamente migliori dell'altro, sia in termini di volatilità, che di varianza condizionata. Il primo confronto, tra l'sGARCH e il GJR-GARCH, produce una loss negativa e, quindi, l'sGARCH risulta il modello migliore tra i due. Il vincitore è poi confrontato con il T-GARCH: la statistica test risulta positiva e ciò implica che il T-GARCH produce previsioni più attendibili, rispetto pure all'IGARCH col quale viene effettuato l'ultimo confronto. Stesso risultato avviene sia nel controllo della volatilità, che in quello della varianza condizionata, proclamando il T-GARCH come miglior modello in fase di previsione.

Diebold-Mariano comparison

```
Volatility
GARCH vs GJR-GARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): -1.980295
GARCH vs GJR-GARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): -0.7114207
GARCH vs T-GARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): 2.064493
GARCH vs T-GARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): 3.396593
T-GARCH vs IGARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): -2.709657
T-GARCH vs IGARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): -5.565422
Conditional variance
GARCH vs GJR-GARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): -0.3279954
GARCH vs GJR-GARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): -2.559075
GARCH vs T-GARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): 0.7576304
GARCH vs T-GARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): 3.764686
T-GARCH vs IGARCH -> Horiz: 1 , Loss fct: SE , Stat (L1-L2): -0.6353962
T-GARCH vs IGARCH -> Horiz: 1 , Loss fct: AE , Stat (L1-L2): -5.476027
```

Figura 33: Test di Diebold-Mariano

Il valore del rendimento e la volatilità (*figura 34*), ovvero il fine ultimo della nostra analisi, sono stimati mediante previsione ex-ante, a partire dal giorno dopo l'ultima osservazione a disposizione. La previsione è basata sul modello migliore tra quelli stimati, il T-GARCH. I valori della volatilità crescono, sempre più lentamente, fino a convergere al sigma non condizionato. La previsione ex-post, invece, è effettuata su dati interni a partire dal 21/06/2022, usati come informazione per la valutazione delle previsioni.

-----	*-----*
* GARCH Model Forecast *	* GARCH Model Forecast *
-----	*-----*
Model: fGARCH	Model: fGARCH
fGARCH Sub-Model: TGARCH	fGARCH Sub-Model: TGARCH
Horizon: 10	Horizon: 10
Roll Steps: 0	Roll Steps: 0
Out of Sample: 0	Out of Sample: 10
0-roll forecast [T0=2023-01-18]:	0-roll forecast [T0=]:
Series Sigma	Series Sigma
T+1 0.05617 0.9218	T+1 0.05617 2.062
T+2 0.05617 0.9270	T+2 0.05617 2.052
T+3 0.05617 0.9321	T+3 0.05617 2.043
T+4 0.05617 0.9371	T+4 0.05617 2.033
T+5 0.05617 0.9421	T+5 0.05617 2.024
T+6 0.05617 0.9470	T+6 0.05617 2.015
T+7 0.05617 0.9518	T+7 0.05617 2.006
T+8 0.05617 0.9566	T+8 0.05617 1.997
T+9 0.05617 0.9613	T+9 0.05617 1.988
T+10 0.05617 0.9660	T+10 0.05617 1.979

figura 34: previsioni ex-ante ed ex-post

Trarre conclusioni solo sull'ultima parte delle osservazioni può non risultare esaustivo sul fatto che il modello sia effettivamente il migliore. Per far fronte a ciò, è preferibile effettuare un'analisi ex-post con un metodo di valutazione dinamico: il Rolling window, (figura 35). La previsione della volatilità è effettuata quattro volte con orizzonte fisso e origine che si sposta di giorno in giorno.

	2022-06-21	2022-06-22	2022-06-23	2022-06-24
T+1	2.061890	2.088873	2.055761	1.968929
T+2	2.052198	2.078829	2.046149	1.960449
T+3	2.042632	2.068916	2.036662	1.952079
T+4	2.033190	2.059132	2.027298	1.943818
T+5	2.023872	2.049476	2.018057	1.935664
T+6	2.014676	2.039946	2.008936	1.927617
T+7	2.005599	2.030539	1.999934	1.919675
T+8	1.996640	2.021256	1.991049	1.911836
T+9	1.987798	2.012093	1.982280	1.904100
T+10	1.979072	2.003050	1.973626	1.896464

Figura 35: stima volatilità con Rolling window

CONCLUSIONI

L'ACF e la PACF dei prezzi giornalieri presentano le usuali caratteristiche di non stazionarietà che contraddistinguono le serie finanziarie, confermate dal test ADF per la presenza di radici unitarie. Per questo, sono analizzati i logaritmi dei rendimenti, che sono stazionari.

L'sGARCH(1,1) ha portato miglioramenti rispetto all'ARMA, come evidenziato dalla diagnostica. Il Sign Bias test ha evidenziato asimmetria positiva, ed è stato stimato il GJR-GARCH, con e senza variance targeting, che ha apportato un ulteriore miglioramento. Il modello migliore, però, è risultato il T-GARCH (1,1). L'ultimo modello stimato è stato l'IGARCH.

Il confronto tra la volatilità stimata col GARCH e quella col benchmark Garman-Klass ha evidenziato un comportamento simile tra i due, con qualche differenza nei periodi di alta volatilità. Dopo aver controllato, per mezzo della diagnostica di Mincer-Zarnowitz, che il modello non produca previsioni distorte e, per mezzo del test di Diebold-Mariano, che il T-GARCH sia effettivamente il migliore, si è proceduto a stimare la volatilità.

Sono state effettuate sia previsioni ex-ante, che ex-post, con l'ausilio, anche, della tecnica Rolling window.