

Starting Finance Club Cattolica Divisione Data Analytics

La volatilità dell'Euro Stoxx 50: fatti stilizzati, misure e modelli di previsione

Ettore Catenacci, Riccardo Barbieri

Aprile 2026

Contents

1 Introduzione alla volatilità **4**

1.1 Principali applicazioni finanziarie dell'analisi della volatilità 5

1.2 Indice Euro Stoxx 50: introduzione e dataset 6

1.2.1 Dataset 7

1.3 I fatti stilizzati 10

1.3.1 Il volatility clustering 10

1.3.2 Fat tails e deviazione dalla normalità 13

1.3.3 Il leverage effect 14

1.3.4 Long memory 15

1.3.5 Mean reversion 15

2 Misure di volatilità **16**

2.1 Geometric Brownian Motion 16

2.1.1 La volatilità come grandezza latente 17

2.2 Volatilità storica: simple moving average 17

2.3 Realized volatility 19

2.4 Stimatori range-based di volatilità 20

2.4.1 Confronto empirico tra diverse misure di volatilità realizzata 21

2.5 La volatilità implicita dell'Euro Stoxx 50 25

2.5.1 Relazione tra il VSTOXX e l'EURO STOXX 50 27

3 Volatility forecasting **29**

3.1 Introduzione alla previsione di volatilità 29

3.1.1 Dataset e modelli 29

3.1.2 Obiettivi dell'analisi 30

3.2 Impostazione generale dell'analisi dei modelli -ARCH() 30

3.2.1 Metodologia di stima e previsione out-of-sample 30

3.2.2 Dataset . . . 30
3.2.3 Criteri di valutazione della qualità previsiva . . . 30
3.2.4 Studio dei residui . . . 32
3.3 Il modello a media mobile esponenziale (EWMA) . . . 32
3.4 Modelli autoregressivi a eteroschedasticità condizionata: Introduzione . . . 35
3.5 Il modello ARCH(p) . . . 36
3.5.1 ARCH(1) sui rendimenti dell'Eurostoxx 50 . . . 36
3.5.2 Stima dei parametri . . . 37
3.5.3 ARCH(1) vs Garman-Klass . . . 37
3.5.4 Analisi dei residui . . . 39
3.6 Il modello GARCH(p,q)) . . . 40
3.6.1 GARCH(1,1) sui rendimenti dell'Eurotstoxx 50 . . . 41
3.6.2 Stima dei parametri . . . 41
3.6.3 GARCH(1,1) vs Garman-Klass: previsione out of sample . . . 43
3.6.4 Analisi dei residui . . . 45
3.7 Il modello EGARCH(p,o,q) . . . 46
3.7.1 EGARCH(1,1,1) sui rendimenti dell'Eurostoxx 50 . . . 46
3.7.2 Stima dei parametri ed interpretazione . . . 47
3.7.3 Un esempio per capire meglio l'effetto leverage sui dati dell' Eurostoxx 50 . . . 48
3.7.4 EGARCH vs Garma-Klass: previsione out-of-sample . . . 49
3.7.5 Analisi dei residui . . . 50
3.8 Confronto e valutazione conclusiva dei modelli . . . 51
3.8.1 Serie temporali a confronto . . . 51
3.8.2 Metriche di errore a confronto . . . 53

Abstract

La volatilità rappresenta una delle grandezze più rilevanti nell'analisi dei mercati finanziari, poiché riflette il grado di incertezza associato ai prezzi degli asset e svolge un ruolo centrale nel risk management, nel pricing dei derivati e nella costruzione di strategie di investimento.

In questo lavoro, il tema viene affrontato con un obiettivo principalmente applicativo e divulgativo: non si vogliono proporre risultati originali di ricerca, ma approfondire alcuni concetti fondamentali della finanza quantitativa e mostrarne il funzionamento attraverso i dati. L'analisi prende come caso di studio l'indice Euro Stoxx 50 e si sviluppa lungo tre direttrici principali: i) L'esame dei principali fatti stilizzati dei rendimenti finanziari; ii) Il confronto tra diverse misure di volatilità; iii) L'applicazione di alcuni modelli classici di previsione.

Il riferimento teorico di partenza è la letteratura empirica sui rendimenti finanziari e, in particolare la formalizzazione dei fatti stilizzati proposta da Rama Cont (2001, "Empirical properties of asset returns: stylized facts and statistical issues")[3], che mette in evidenza proprietà ricorrenti quali l'assenza di autocorrelazione lineare dei rendimenti, la presenza di heavy tails, il volatility clustering e il leverage effect. In questo quadro, il lavoro non intende verificare in modo esaustivo l'intero insieme dei fatti stilizzati discussi in letteratura, ma concentrarsi su quelle evidenze empiriche che risultano

più direttamente collegate allo studio e alla previsione della volatilità. Dopo aver costruito il dataset e definito i rendimenti logaritmici, il lavoro analizza empiricamente tali regolarità, distinguendo tra caratteristiche dei prezzi, dei rendimenti e della dinamica della volatilità. Successivamente vengono confrontate, stimate e confrontate misure di volatilità storica, realizzata e implicita. Su questa base vengono quindi presentati e valutati, in un contesto di previsione out-of-sample, modelli largamente utilizzati nella letteratura e nella pratica, tra cui EWMA, ARCH, GARCH ed EGARCH.

I risultati ottenuti sono coerenti con quanto atteso sul piano teorico e con la letteratura di riferimento: le specificazioni più flessibili, capaci di cogliere persistenza e asimmetrie della volatilità, mostrano una capacità descrittiva e previsiva superiore rispetto ai modelli più semplici. Pur in un contesto in cui la ricerca più recente ricorre sempre più spesso a tecniche di machine learning e deep learning, il lavoro si concentra volutamente su modelli econometrici classici, scelti perché costituiscono ancora oggi una base essenziale per comprendere la logica della volatility forecasting e il legame tra teoria e applicazione empirica.

1 Introduzione alla volatilità

La volatilità occupa un ruolo centrale nell'analisi dei mercati finanziari, in quanto misura il grado di variabilità dei prezzi o, più precisamente, dei rendimenti degli asset finanziari nel tempo. In termini statistici, essa è strettamente collegata al concetto di dispersione di una variabile casuale attorno al proprio valore atteso: se i rendimenti vengono interpretati come realizzazioni di un processo aleatorio, la volatilità esprime quanto tali realizzazioni tendano a discostarsi dalla loro media.

Se indichiamo con r_t il rendimento al tempo t, con $\bar{r}$ il rendimento medio e con n il numero di osservazioni, possiamo esprimere la volatilità campionaria in questo modo:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2}$$

La volatilità coincide quindi con la deviazione standard dei rendimenti osservati nel periodo considerato. Quanto maggiore è la dispersione dei rendimenti attorno alla loro media, tanto più elevata risulta la volatilità. In termini economici, ciò implica una maggiore incertezza circa il valore futuro dell'attività finanziaria: una volatilità più alta segnala infatti una maggiore probabilità di osservare variazioni di prezzo ampie, sia positive sia negative. Essa non fornisce una descrizione completa della distribuzione dei rendimenti futuri, ma ne sintetizza il grado di dispersione, rappresentando così una delle principali misure del rischio di mercato.

Nel contesto dell'analisi finanziaria, è utile distinguere tra tre principali nozioni di volatilità, che si differenziano per metodo di costruzione e contenuto informativo:

- Volatilità storica: viene calcolata come deviazione standard dei rendimenti passati, generalmente utilizzata come proxy della volatilità vera della popolazione dei rendimenti.

- Volatilità realizzata: rappresenta la variabilità effettivamente manifestata dal prezzo di un'attività finanziaria in un determinato intervallo di tempo, misurata ex post sulla base delle osservazioni di

mercato, usando i rendimenti intraday.

- Volatilità implicita: non viene ricavata dai rendimenti storici dell'asset, ma dai prezzi delle opzioni negoziate su quest'ultimo. Rappresenta quindi l'aspettativa del mercato riguardo alle future fluttuazioni di prezzo di un'attività sottostante.

Queste diverse misure di volatilità verranno approfondite separatamente nella sezione 2. Ora invece, per dare una visione più completa del concetto di volatilità, vediamo alcune applicazioni pratiche.

1.1 Principali applicazioni finanziarie dell'analisi della volatilità

L'importanza della volatilità non si esaurisce nel suo significato teorico o statistico. Al contrario, essa assume un ruolo centrale in molte decisioni operative dei mercati finanziari, in quanto fornisce una misura del grado di rischio associato ad un'attività finanziaria. Vediamo brevemente alcuni degli ambiti principali di applicazione.

- Risk management - In questo contesto, la volatilità entra direttamente nel calcolo delle misure più diffuse, tra cui il Value at Risk (VaR), che esprime la perdita massima attesa entro un determinato orizzonte temporale e a un dato livello di confidenza, e l'Expected Shortfall (ES), che fornisce una stima della perdita media condizionata al verificarsi degli eventi più estremi. La qualità di tali misure dipende in modo significativo dall'accuratezza delle stime di volatilità, che possono essere ottenute mediante approcci storici, modelli parametrici o tecniche basate sulla volatilità implicita.

- Pricing - La volatilità riveste un ruolo centrale nei modelli di valutazione degli strumenti finanziari, in particolare per quanto riguarda i derivati. Un riferimento teorico fondamentale è rappresentato dal modello Black-Scholes (Black and Scholes, 1973)[1], nel quale la volatilità costituisce una variabile determinante per la stima del valore teorico delle opzioni. Il valore di un'opzione dipende fortemente dalla probabilità che quest'ultima venga esercitata da chi la detiene e, in questo senso, maggiore è la volatilità del sottostante, maggiore è la probabilità che l'opzione finisca "in-the-money" prima della scadenza, aumentandone il valore.

- Hedging - La volatilità assume un ruolo cruciale anche nell'ambito delle strategie di copertura, influenzando in modo diretto la sensibilità degli strumenti derivati rispetto alle variazioni dei fattori di rischio sottostanti. In particolare, nelle strategie di copertura dinamica, quali il delta hedging, la frequenza e l'entità degli aggiustamenti richiesti dipendono significativamente dal livello di volatilità del sottostante. Inoltre, le variazioni inattese della volatilità introducono una specifica componente di rischio, comunemente definita volatility risk, che può compromettere l'efficacia delle strategie di copertura e richiede pertanto l'impiego di strumenti e tecniche dedicate.

- Asset allocation - Nel contesto della teoria della selezione di portafoglio, la volatilità è generalmente adottata come misura del rischio associato agli asset finanziari. In particolare, la Modern Portfolio Theory (Markowitz, 1952) utilizza la varianza dei rendimenti come parametro chiave per la determinazione del trade-off tra rischio e rendimento. Tale approccio consente di individuare combinazioni

di portafoglio efficienti, ossia in grado di massimizzare il rendimento atteso dato un certo livello di rischio, o viceversa di minimizzare il rischio per un determinato rendimento atteso. Inoltre, metodologie più recenti, quali gli approcci basati sul risk parity, attribuiscono i pesi agli asset in funzione della loro volatilità, enfatizzando ulteriormente il ruolo di tale grandezza nei processi di asset allocation.

- Derivati sulla volatilità - La crescente rilevanza della volatilità nei mercati finanziari ha condotto allo sviluppo di strumenti derivati specificamente progettati per consentire agli operatori di negoziare e coprire il rischio associato alle sue variazioni. In questa categoria rientrano, da un lato, strumenti over-the-counter come i variance swaps e i volatility swaps; dall'altro, strumenti quotati sui mercati regolamentati (Eurex), come i futures e opzioni su indici di volatilità.

In Europa, il riferimento principale è l'Euro Stoxx 50 volatility index "VSTOXX", il quale misura la volatilità implicita attesa a 30 giorni delle opzioni sull'indice Euro Stoxx 50. Il trading sulla volatilità europea avviene dunque attraverso strumenti derivati che hanno come sottostante l'indice di volatilità VSTOXX. In particolare, i futures sul VSTOXX sono negoziati dal 2009, mentre le opzioni sono state introdotte nel 2010. Essi risultano pertanto particolarmente utili sia per finalità di copertura del rischio di volatilità, sia per l'implementazione di strategie speculative basate sulle aspettative relative all'incertezza dei mercati.

1.2 Indice Euro Stoxx 50: introduzione e dataset

L'Euro Stoxx 50 viene lanciato il 26 febbraio 1998 dalla Stoxx Ltd, ora controllata da Deutsche Börse AG e comprende le 50 società "blue chip" a maggiore capitalizzazione e liquidità dell'Eurozona. L'indice opera come un sottoinsieme dell'EURO STOXX Total Market Index (TMI), da cui i componenti vengono selezionati durante la revisione annuale di settembre basandosi su rigorosi filtri di ranking del sistema di classificazione ICB (Industry Classification Benchmark). Le 50 società vengono classificate in base alla loro capitalizzazione di mercato del flottante, considerando quindi le sole azioni *outstanding*, ovvero le azioni in circolazione disponibili per la negoziazione. Il rischio di concentrazione viene mitigato attraverso un capping del 10% sulla ponderazione di ogni singolo componente, il quale viene ricalcolato trimestralmente. La componente italiana è rappresentata da leader settoriali come Enel, Unicredit, Intesa San Paolo, Ferrari, Eni e Stellantis.

Rappresentando circa il 60% della capitalizzazione totale del mercato azionario dell'Eurozona, Euro Stoxx 50 costituisce un riferimento centrale per l'analisi delle dinamiche macroeconomiche della regione. In questa sede, l'indice sarà l'oggetto d'indagine principale per lo studio della dinamica dei rendimenti e della volatilità, vista anche la presenza del VSTOXX come misura di volatilità implicita.



Figure 1: Prezzi di chiusura EURO STOXX 50.

Ad influenzare l'andamento dell'indice, oltre agli shock globali, pesano i seguenti fattori: i) Politica monetaria della BCE ; ii) Dati macroeconomici tedeschi e francesi: Germania e Francia pesano circa il 60-70% dell'indice; iii) Cambio Euro/Dollaro: molte blue chip sono grandi esportatrici, un euro debole rende i prodotti più competitivi per l'export; iv) Concentrazione settoriale: l'indice è molto esposto a banche, settore del lusso (LVMH) e tecnologia (ASML), quindi shock in questi specifici settori influenzano il livello dell'indice maggiormente.

1.2.1 Dataset

Il dataset è stato creato scaricando i prezzi dell'indice a frequenza giornaliera (fonte: Bloomberg Terminal), per un arco temporale dal 02/03/1998 al 31/03/2026. Per ogni osservazione giornaliera, il dataset contiene quattro variabili di prezzo: apertura, massimo, minimo e chiusura; indicate con l'acronimo OHLC.

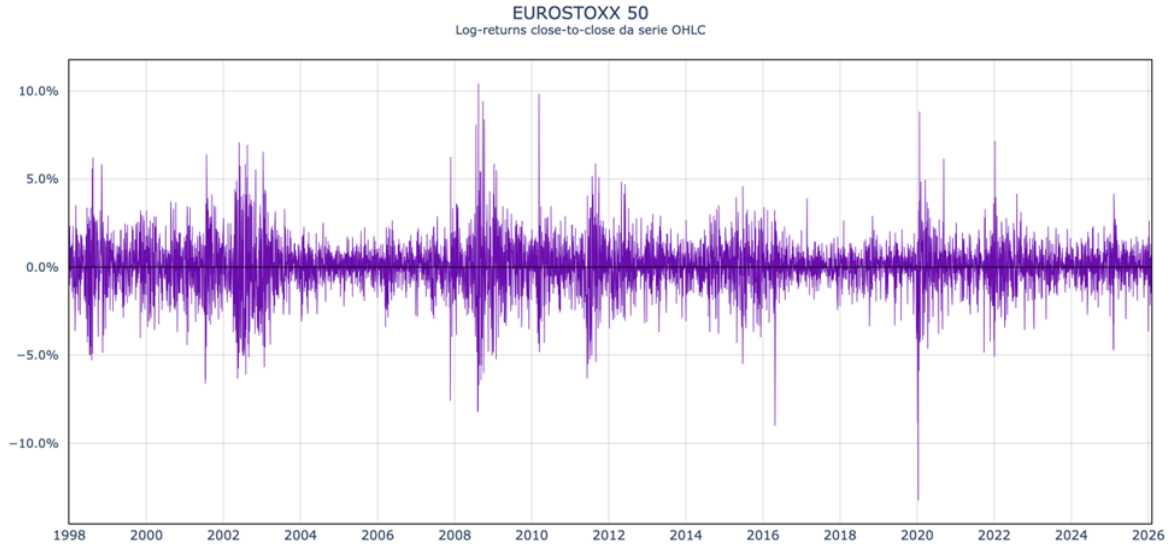


Figure 2: EURO STOXX 50 log-returns giornalieri.

Dai prezzi giornalieri di chiusura, sono stati ottenuti i rendimenti logaritmici (log-returns), che misurano la variazione percentuale continua tra due giorni consecutivi. Se P_t è il prezzo dell'indice al tempo t, mentre $P_t - 1$ è il prezzo dell'indice in t-1, il log-rendimento r_t tra un giorno e l'altro viene calcolato come segue:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

o equivalentemente:

$$r_t = \ln(P_t) - \ln(P_{t-1})$$

Da distinguere dal rendimento aritmetico calcolato come semplice variazione percentuale giornaliera: $R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$

I rendimenti logaritmici sono ampiamente utilizzati in analisi finanziaria e preferiti alle sole osservazioni di prezzo principalmente per i seguenti motivi:

1. Additività nel tempo: la somma dei rendimenti logaritmici giornalieri tra due date equivale al rendimento complessivo dell'intero periodo. Questo semplifica molto l'analisi dei rendimenti cumulati, l'aggregazione temporale e l'interpretazione matematica dei movimenti di prezzo.

   In particolare, questo risolve il problema dell'asimmetria delle variazioni: ad esempio, se consideriamo i rendimenti percentuali semplici, un titolo che perde il 50% deve poi guadagnare il 100% per tornare al prezzo iniziale; al contrario, usando i log-returns, una perdita del –10% è compensata esattamente da +10%.

2. Stazionarietà e analisi statistica ed econometrica: la stazionarietà è la condizione necessaria per cui le proprietà statistiche di una serie (media, varianza e autocorrelazione) non cambiano nel tempo. Se una serie non è stazionaria, i risultati dei modelli sono "spuri", ovvero privi di valore predittivo o descrittivo reale.

Come vedremo in 2.1, il prezzo P_t dipende direttamente da P_{t-1}, più una componente casuale; mostrando una dinamica non stazionaria: la media cambia e la varianza tende all'infinito. Trasformando i prezzi i rendimenti logaritmici, si rimuove questa tendenza, ottenendo una misura approssimativamente stazionaria, che tende ad oscillare attorno ad una media costante e presenta una varianza finita, permettendo lo studio della volatilità.

3. Distribuzione: nonostante l'evidenza empirica (approfondita in 1.3.2) in finanza viene spesso ipotizzato che i rendimenti seguano una distribuzione normale, permettendo l'applicazione del TLC e dei principali modelli di calcolo delle opzioni e del rischio.

   Questo non può essere fatto invece con i prezzi, i quali ovviamente non possono avere valore negativo e per questo si considera la loro distribuzione come log-normale. Le proprietà dei rendimenti logaritmici verranno approfondite insieme ai fatti stilizzati nella sezione 1.3.

   Di seguito vengono riportate le statistiche principali delle osservazioni a disposizione:

	PREZZI DI CHIUSURA (Pt)	LOG RETURNS (rt)
N	7196	7195
mean	3507.988	8.942224e-05
variance	701020.63	0.0001957
std	837.27	0.01398825
skewness	0.7026	-0.2200768
kurtosis	0.0489	5.871
min	1809.98	-0.1324
p05	2343.4975	-0.02215
median	3409.17	0.0003825
p95	5206.59	0.0210813
max	6173.32	0.1043765

Table 1: Statistiche descrittive del dataset.

La serie dei prezzi di chiusura dell'EURO STOXX 50 si colloca nel lungo periodo, su un livello medio intorno a 3508 punti, con una dispersione ampia, coerente con un campione che attraversa fasi di mercato molto diverse tra il 1998 e il 2026. La distanza tra i valori più bassi e più alti osservati segnala infatti una marcata variabilità nel livello dell'indice nel tempo. Per i rendimenti logaritmici, troviamo invece coerenza con l'ipotesi comune di una media nulla giornaliera, infatti, essi oscillano attorno allo 0 non mostrando tendenze medie giornaliere con rilevanza economica.

A dimostrazione di quanto affermato nel punto (2), facendo ricorso al test di Dickey-Fuller aumentato (ADF), si può verificare l'assenza di stazionarietà nella serie di prezzi P_t. Viene testata l'ipotesi nulla H_0 di presenza di unit root, ovvero di non stazionarietà della serie, contro l'ipotesi alternativa H_1 di stazionarietà. I risultati ottenuti evidenziano una netta differenza tra la dinamica dei prezzi e quella dei rendimenti, confermando la motivazione di prima.

SERIE	ADF_STAT	P_VALUE	LAG USATO	DECISIONE	CONCLUSIONE
PREZZI CHIUSURA Pt	-1.4666	0.54994	6	NON RIFIUTA H_0	NON STAZIONARIA
LOG RETURNS rt	-32.7648	0.0	6	RIFIUTA H_0	STAZIONARIA

Table 2: Risultati del test ADF.

Nel caso dei prezzi di chiusura, la statistica del test risulta pari a -1.4666, mentre il p-value è elevato (0.54994). Non esiste quindi sufficiente evidenza per rifiutare H_0, di conseguenza si può affermare che la serie dei prezzi P_t è non stazionaria. Al contrario, per la serie dei log-returns, la statistica ADF assume un valore pari a -32.7648, con un p-value approssimato a zero. In questo caso, l'ipotesi nulla viene ampiamente rifiutata, suggerendo che la serie dei rendimenti è stazionaria. Tale evidenza giustifica l'utilizzo dei log-returns nelle successive analisi empiriche sulla volatilità, consentendo l'applicazione di modelli econometrici adeguati e coerenti con le ipotesi teoriche sottostanti.

1.3 I fatti stilizzati

I fatti stilizzati sono un insieme di proprietà statistiche ricorrenti, emerse dallo studio empirico dei rendimenti finanziari in mercati, strumenti e periodi storici differenti. Un riferimento centrale è il contributo di Rama Cont(2001): *"Empirical properties of asset returns: stylized facts and statistical issues"*, che ha raccolto e ordinato in modo sistematico le principali evidenze empiriche osservate nelle serie finanziarie. Queste proprietà non vanno intese come leggi universali, ma come caratteristiche ricorrenti dei dati che offrono un quadro di riferimento essenziale per descrivere il comportamento dei mercati e per valutare la capacità dei modelli econometrici di rappresentarne in modo realistico la dinamica. Il presente lavoro si concentra su alcuni di questi fatti stilizzati, i quali verranno anzitutto presentati sul piano concettuale e successivamente richiamati nell'analisi empirica, dove troveranno riscontro nei risultati ottenuti. In particolare, consideriamo le seguenti proprietà:

- Volatility clustering
- Fat tails e Skewness
- Leverage effect
- Long memory
- Mean reversion

Tali caratteristiche verranno analizzate empiricamente utilizzando i rendimenti giornalieri dell'indice EURO STOXX 50, con particolare attenzione alla distribuzione dei rendimenti, dipendenza temporale (ACF) , clustering della volatilità e infine test statistici (ADF e Ljung-Box).

1.3.1 Il volatility clustering

Una delle evidenze empiriche più robuste riguarda il fenomeno del volatility clustering, secondo cui elevati livelli di volatilità tendono a essere seguiti da ulteriori periodi di elevata volatilità, mentre fasi di bassa volatilità risultano anch'esse persistenti nel tempo. Tale proprietà implica l'esistenza di dipendenza seriale nelle misure di volatilità, nonostante i rendimenti stessi presentino generalmente autocorrelazione nulla. Formalmente, questo fenomeno si manifesta attraverso una significativa autocorrelazione dei rendimenti al quadrato o in valore assoluto, suggerendo che la volatilità sia un processo dinamico e prevedibile almeno nel breve-medio periodo. Questo comportamento è coerente con l'idea che le informazioni arrivino sui mercati in maniera discontinua, generando fasi di turbolenza concentrate nel tempo come possiamo riscontrare in maniera descrittiva nel grafico 1.2 EURO STOXX 50 log-returns giornalieri. L'analisi dell'autocorrelazione consente di verificare se i valori passati di una serie temporale influenzino i valori

futuri. In ambito finanziario, ciò si traduce nella verifica della prevedibilità dei rendimenti: in particolare, ci si chiede se i rendimenti osservati oggi contengano informazione utile per prevedere quelli futuri. A tal fine, si analizza la funzione di autocorrelazione (ACF) sia per i log-returns sia per i quadrati dei rendimenti (squared returns), questi ultimi utilizzati come proxy della volatilità.

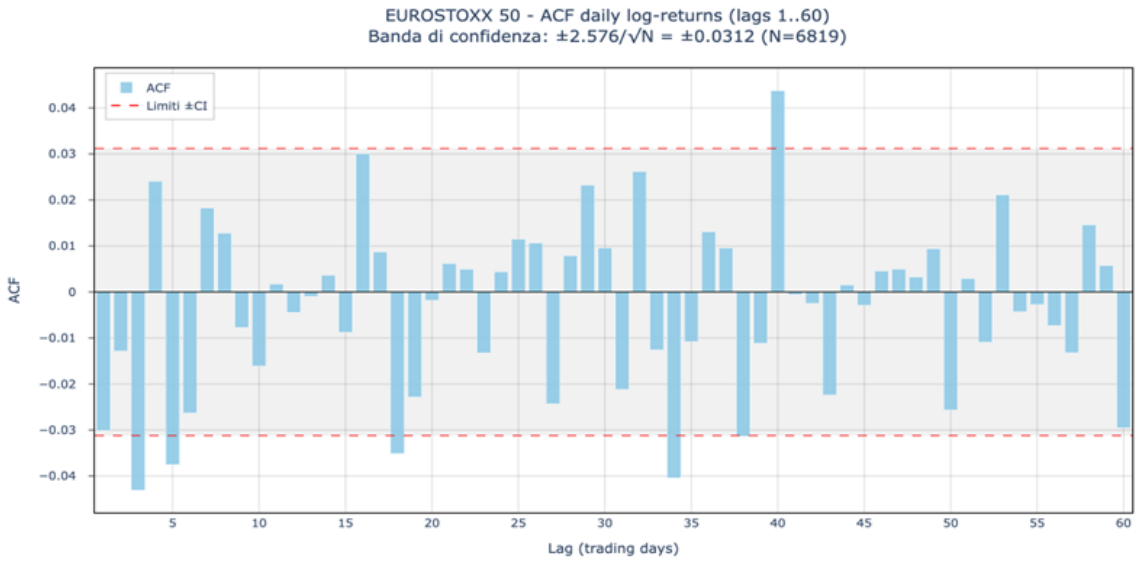


Figure 3: ACF log-returns giornalieri.

Nel caso dei log-returns, i coefficienti di autocorrelazione risultano generalmente prossimi allo zero e, nella maggior parte dei casi, non significativamente diversi da zero, come evidenziato anche dal confronto con le bande di confidenza. Questo risultato è coerente con l'ipotesi di mercato efficiente in forma debole, secondo cui i rendimenti non sono linearmente prevedibili sulla base delle informazioni passate. Diversamente, l'analisi dell'ACF applicata agli squared returns mostra un comportamento profondamente differente. I coefficienti risultano positivi e significativamente diversi da zero per un ampio numero di lag, con una lenta decrescita nel tempo.

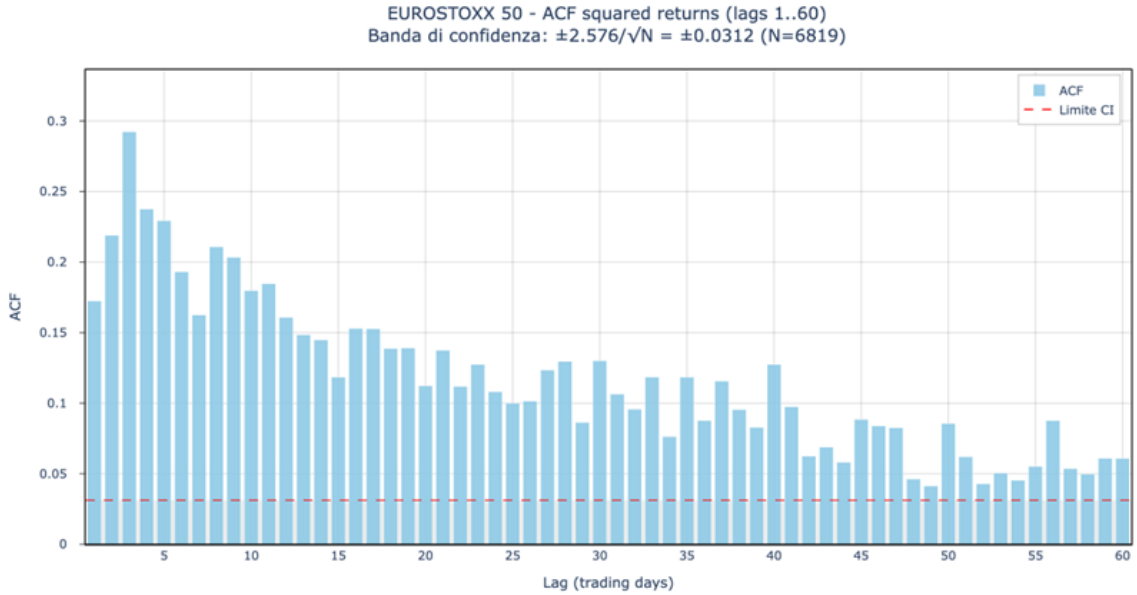


Figure 4: ACF rendimenti al quadrato.

Un aspetto rilevante riguarda la dimensione campionaria elevata, per cui sebbene un ampio numero di osservazioni garantisca maggiore robustezza statistica, esso implica anche che le bande di confidenza dell'ACF siano molto strette in prima approssimazione, dell'ordine di $1/\sqrt{N}$.

Di conseguenza, anche autocorrelazioni di entità ridotta possono risultare statisticamente significative, pur avendo un impatto economico limitato.

LAG	P_VALUE LOG RETURN	P_VALUE SQUARED RETURNS
1	0.013904556	2.9417389862e-46
5	2.77553098e-06	0.0
10	3.61447021e-06	0.0
15	9.04508584e-05	0.0
20	3.18532142e-06	0.0

Table 3: Risultati ACF log-returns giornalieri e ACF rendimenti al quadrato a confronto.

Per testare formalmente la presenza di autocorrelazione si utilizza il test di Ljung-Box, il quale verifica congiuntamente l'ipotesi nulla di assenza di autocorrelazione fino a un certo numero di lag.

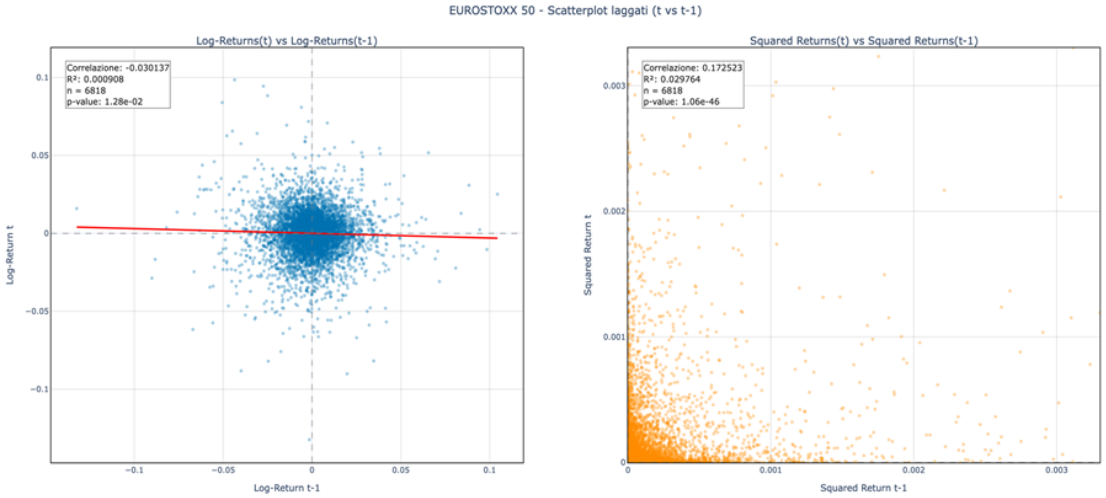


Figure 5: Lag plot dei log-returns e dei rendimenti al quadrato (lag 1).

I risultati mostrano che, mentre per i log-returns i p-value sono relativamente più elevati (pur risultando talvolta significativi), nel caso degli squared returns i p-value sono estremamente bassi per tutti i lag considerati, portando a un netto rifiuto dell'ipotesi nulla di assenza di autocorrelazione. È importante sottolineare che il test di Ljung-Box è di natura cumulativa, all'aumentare del numero di lag considerati, la probabilità di rifiutare l'ipotesi nulla tende ad aumentare, anche in presenza di autocorrelazioni individualmente modeste. Questo rappresenta una limitazione del test, in quanto può amplificare l'evidenza statistica senza necessariamente riflettere una rilevanza economica significativa. Nel complesso, i risultati evidenziano un quadro coerente con la letteratura: assenza di autocorrelazione lineare nei rendimenti, ma forte dipendenza nei momenti di ordine superiore, in particolare nella volatilità. Tale evidenza giustifica l'utilizzo di modelli eteroschedastici condizionati, come i modelli GARCH(sezione 3), per la modellazione della dinamica della volatilità.

1.3.2 Fat tails e deviazione dalla normalità

Un'altra caratteristica fondamentale delle distribuzioni dei rendimenti finanziari è la presenza di fat tails, ovvero code più spesse rispetto a quelle previste da una distribuzione normale. Ciò implica una maggiore probabilità di osservare eventi estremi, sia positivi che negativi, rispetto a quanto suggerito dai modelli gaussiani. Dal punto di vista statistico, questo fenomeno si traduce in valori elevati della curtosi e nella deviazione significativa dalla normalità delle distribuzioni empiriche dei rendimenti. Nel caso dei log-returns dell'Euro Stoxx 50, le statistiche descrittive confermano chiaramente questa evidenza. In particolare, la distribuzione presenta una skewness pari a -0.2201 e una kurtosis pari a 5.871, valori che segnalano rispettivamente una lieve asimmetria negativa e una marcata leptocurtosi. Il segno negativo della skewness suggerisce che la coda sinistra della distribuzione sia leggermente più pronunciata rispetto a quella destra, indicando una maggiore incidenza relativa di movimenti estremi al ribasso. Tuttavia, nel caso considerato, l'evidenza empiricamente più rilevante riguarda soprattutto l'eccesso di curtosi.

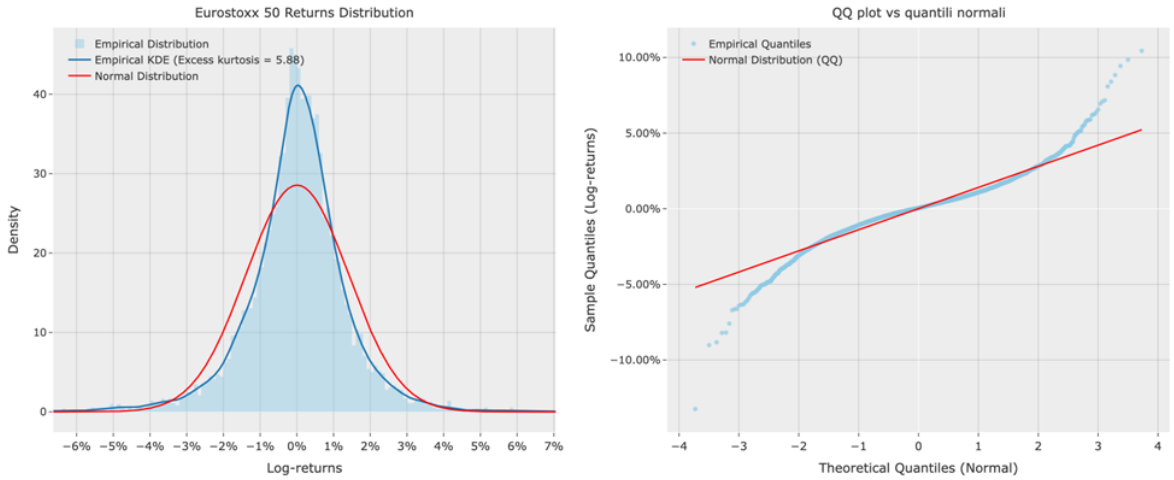


Figure 6: EURO STOXX 50 distribuzione dei rendimenti e QQ plot contro quantili normali.

Questa conclusione è coerente con quanto emerge anche dalla Figura 6. La distribuzione empirica appare infatti più appuntita al centro e più pesante nelle code rispetto alla normale di riferimento. Il QQ-plot fornisce un'ulteriore conferma, infatti, se la distribuzione fosse normale, seguirebbe la linea teorica. Al contrario, mostra scostamenti significativi dalla retta teorica soprattutto nelle regioni estreme.

Per verificare formalmente quanto osservato a livello descrittivo, si può ricorrere al test di Jarque-Bera (JB), che valuta congiuntamente la deviazione di skewness e curtosi dai valori teorici della distribuzione normale.

VOCE	RISULTATO
Ipotesi nulla (H0)	I log-returns seguono una distribuzione Normale
Ipotesi alternativa (H1)	I log-returns NON seguono una distribuzione Normale
Statistica Jarque-Bera	10412.0320
p-value	0.0000e+00
Livello di significatività (alpha)	0.05
Decisione	RIFIUTA H0

Table 4: sultati del test di Jarque-Bera (JB).

L'ipotesi nulla del test stabilisce che i log-returns seguano una distribuzione normale, mentre l'ipotesi alternativa nega tale condizione. Dai risultati ottenuti, la statistica JB è pari a 10412.032 con un p-value praticamente nullo (0.0000). Poiché il p-value è ampiamente inferiore al livello di significatività scelto di 0.05, si rifiuta ampiamente l'ipotesi nulla. Questo implica che i ritorni dell'indice non seguono una distribuzione normale, confermando quanto già osservatoin precedenza.

1.3.3 Il leverage effect

Il leverage effect descrive la relazione asimmetrica tra rendimenti e volatilità, per cui variazioni negative dei prezzi sono generalmente associate a un incremento più marcato della volatilità rispetto a variazioni positive di pari entità. Una possibile interpretazione economica è che una diminuzione del valore di mercato di un'impresa aumenti il suo grado di leva finanziaria, incrementandone il rischio e, di conseguenza,

la volatilità del titolo. Il fenomeno inizialmente è stato studiato da Black (1976), il quale notò come in risposta a "cattive notizie", ovvero quando i rendimenti sono più bassi del loro valore atteso, la volatilità tende ad aumentare e viceversa in risposta a "buone notizie" la volatilità tende a diminuire. Il leverage effect è osservabile direttamente dal grafico del VSTOXX nella sezione 2.5 ottenendo un fenomeno simile a quello di Fabiani(2025)[4] con il VIX. Infine per riuscire a catturare il leverage effect nella sezione 3 utilizzeremo il modello EGARCH.

1.3.4 Long memory

La volatilità finanziaria è caratterizzata da un elevato grado di persistenza poiché shock di volatilità tendono ad avere effetti duraturi nel tempo, con una lenta attenuazione della loro influenza. Dal punto di vista empirico, questa proprietà si manifesta attraverso una lenta decrescita della funzione di autocorrelazione dei rendimenti in valore assoluto o al quadrato. Tale comportamento giustifica l'utilizzo di modelli in grado di catturare dipendenze temporali di lungo periodo, come i modelli HAR o le estensioni dei modelli GARCH. Un processo stocastico stazionario X_t si definisce a memoria lunga se la sua funzione di autocorrelazione ρ(k)decresce lentamente al crescere del lag k. In particolare, esistono una costante $c\rho > 0$ e un parametro $\alpha \in (0,1)$ tali che:

$$\lim_{k\to\infty} \frac{\rho(k)}{c_\rho k^{-\alpha}} = 1 \tag{1}$$

In questo caso, il processo X_t viene definito processo stazionario a memoria lunga, o caratterizzato da dipendenza di lungo periodo (Sewell, 2011)[5].La principale differenza tra processi a memoria corta e a memoria lunga risiede nella dinamica della funzione di autocorrelazione. Nei processi a memoria corta, l'autocorrelazione decresce rapidamente secondo un andamento esponenziale, implicando che gli shock hanno effetti di breve durata. Al contrario, nei processi a memoria lunga l'autocorrelazione presenta un decadimento iperbolico, molto più lento, indicando che gli shock influenzano il sistema anche su orizzonti temporali estesi.

1.3.5 Mean reversion

Infine, la volatilità tende a mostrare un comportamento di mean reversion, ovvero una tendenza a ritornare verso un livello medio di lungo periodo dopo aver subito uno shock. In termini dinamici, questo implica che aumenti improvvisi della volatilità siano seguiti, nel tempo, da una graduale riduzione verso valori più normali. Questo fatto, riscontrabile attraverso la volatilità implicita che nel nostro caso studio coincide con il VSTOXX, è incorporata esplicitamente nella struttura di molti modelli di volatilità come il GARCH che vedremo nella sezione 3, in cui la persistenza è elevata ma inferiore all'unità, garantendo la stazionarietà del processo.

2 Misure di volatilità

2.1 Geometric Brownian Motion

Al fine di introdurre quali sono le principali misure di volatilità, è utile partire da un modello teorico fondamentale nella finanza quantitativa: il moto geometrico browniano. Si tratta di un processo stocastico secondo cui il prezzo P_t di un asset si evolve nel tempo come:

$dS_t = \mu S_t dt + \sigma S_t dW_t$

Dove μ rappresenta il drift del processo (il rendimento atteso), mentre la diffusione $\sigma\ dWt$ costituisce la componente aleatoria, la cui intensità è determinata dalla volatilità σ. In questo modello, il percorso del prezzo di un asset varia a seconda di una componente deterministica, proporzionale a μ, e di una componente stocastica, dove σ regola l'ampiezza delle oscillazioni casuali. Il modello può essere descritto usando anche i log-prezzi, come segue:

$d \ln S_t = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t.$

A scopo illustrativo, è stata condotta una simulazione Montecarlo basata sul modello appena descritto per generare possibili traettorie future dell'indice Euro Stoxx 50.

Il modello è stato calibrato usando il dataset ottenuto da Bloomberg Terminal a nostra disposizione. prezzi di chiusura stimando drift e volatilità annualizzati a partire dai log-rendimenti giornalieri: si sono ottenuti un drift annuo $\mu\hat{} = 0.87\%$ e volatilità annua $\sigma\hat{} = 21.9\%$ con $S_0 = €\ 6131.31$ (prezzo dell'indice al giorno della simulazione).

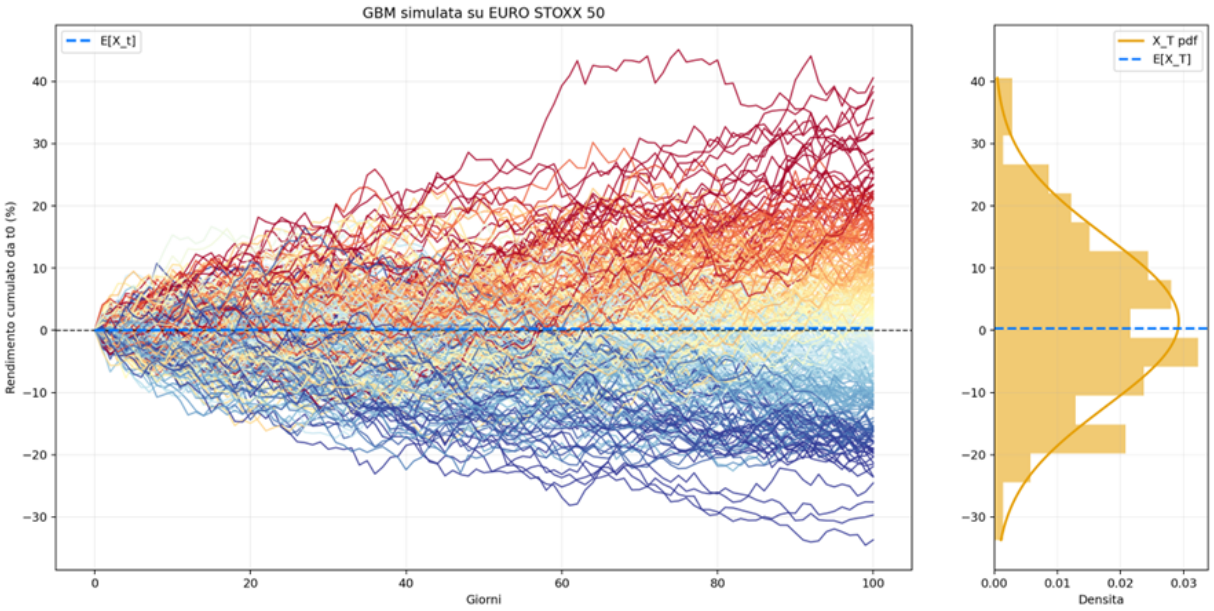


Figure 7: GBM simulazione di Montecarlo su EROSTOXX 50.

Sono state condotte 300 simulazioni su 100 giorni ed il grafico mostra i diversi percorsi di rendimento cumulato e la distribuzione di quest'ultimi.

Teniamo presente che l'esercizio ha un valore puramente descrittivo e non predittivo, infatti il modello descritto presenta un limite importante: la volatilità viene considerata come costante. La letteratura sui

fatti stilizzati (Cont 2009) mostra infatti che la volatilità non si comporta come una costante, ma presenta persistenza, clustering, code pesanti e asimmetrie dinamiche. In particolare, come abbiamo visto nelle sezioni precedenti, esiste autocorrelazione positiva e persistente nelle misure di ampiezza delle variazioni (rendimenti al quadrato), suggerendo che la volatilità sia una quantità variabile nel tempo.

Per tener conto di tali caratteristiche, si adotta una formulazione più generale in cui il log-prezzo segue un processo del tipo:

$dP_t = \mu_t \, dt + \sigma_t \, dW_t$

dove la volatilità è ora una funzione del tempo σ_t, rappresentando l'intensità istantanea delle fluttuazioni dei rendimenti. La misura teorica della volatilità su un intervallo temporale è data dalla integrated variance, definita come l'integrale nel tempo della volatilità istantanea al quadrato:

$IV_t = \int_{t-1}^{t} \sigma_s^2 \, ds$

La integrated variance rappresenta quindi la varianza complessiva accumulata nell'intervallo considerato e costituisce il naturale riferimento teorico per l'analisi empirica della volatilità.

2.1.1 La volatilità come grandezza latente

Abbiamo appena visto come la varianza integrata descrive l'accumulazione della variabilità istantanea dei rendimenti lungo un intervallo di tempo finito. Il prezzo di un asset varia istante dopo istante, e così il suo rendimento, la variabilità di questi rendimenti in un determinato intervallo viene definita come volatilità. Il problema però, è che questa grandezza, pur essendo centrale nella modellazione del rischio e nella previsione della volatilità, non è direttamente osservabile, infatti nei dati di mercato si osservano prezzi e rendimenti, ma non la traiettoria della volatilità istantanea che genera tali osservazioni. Si dice quindi, che la volatilità è una grandezza latente poiché esiste come caratteristica del processo che genera i prezzi, ma non è osservabile direttamente nei dati, perciò, deve essere ricostruita indirettamente.

2.2 Volatilità storica: simple moving average

Un metodo comunemente utilizzato per il calcolo della volatilità storica è il Simple Moving Average (SMA), o media mobile semplice, che utilizza i dati storici per calcolare la cosiddetta "rolling standard deviation" su un intervallo temporale fisso, che scivola avanti nel tempo. Prendiamo in considerazione una finestra temporale di n log-rendimenti giornalieri e ne calcoliamo la deviazione standard giorno dopo giorno, seguendo la classica formula:

$$\sigma_{\text{rolling},t} = \sqrt{\frac{1}{n-1} \sum_{i=t-n+1}^{t} (r_i - \bar{r}_t)^2}$$

Dove t e $i = t - n + 1$ sono rispettivamente il giorno più recente e più vecchio dell'intervallo, mentre ri è il log rendimento del giorno i e $\bar{r}_t$ è la media dei log-rendimenti della finestra. Usiamo $n - 1$ in quanto è una stima campionaria.

Ad esempio, per calcolare la rolling volatility a 21 giorni dell'indice Euro Stoxx 50 partendo dal 1 settembre, consideriamo solo i giorni di trading fino al 29 settembre (escludendo i giorni in cui la borsa

è chiusa); il giorno successivo, il calcolo includerà il 30 settembre escludendo il primo del mese, seguendo la stessa logica per i giorni successivi.

In particolare, questa logica di calcolo presenta due implicazioni importanti:

- All'aumentare di n, aumenta il contenuto informativo della stima campionaria, diminuisce la sua reattività ad eventi improvvisi.
- Effetto "ghosting" o eco.

La scelta del numero di osservazioni n: A parità di altre condizioni, l'ampiezza della finestra influenza in modo diretto la rumorosità della stima di volatilità: in una finestra temporale più stretta, il peso relativo di una singola osservazione aumenta. Se calcoliamo la rolling volatility a 7 giorni, il rendimento di una singola giornata incide per il 14,3% sul totale, considerando 21 giorni invece, la stima sarà 3 volte meno sensibile. Nell'immagine che segue, confrontiamo tre diverse stime di volatilità mobile con frequenza diversa: 7, 20 e 30 giorni. Le misure sono calcolate con i rendimenti dell'Euro Stoxx 50.

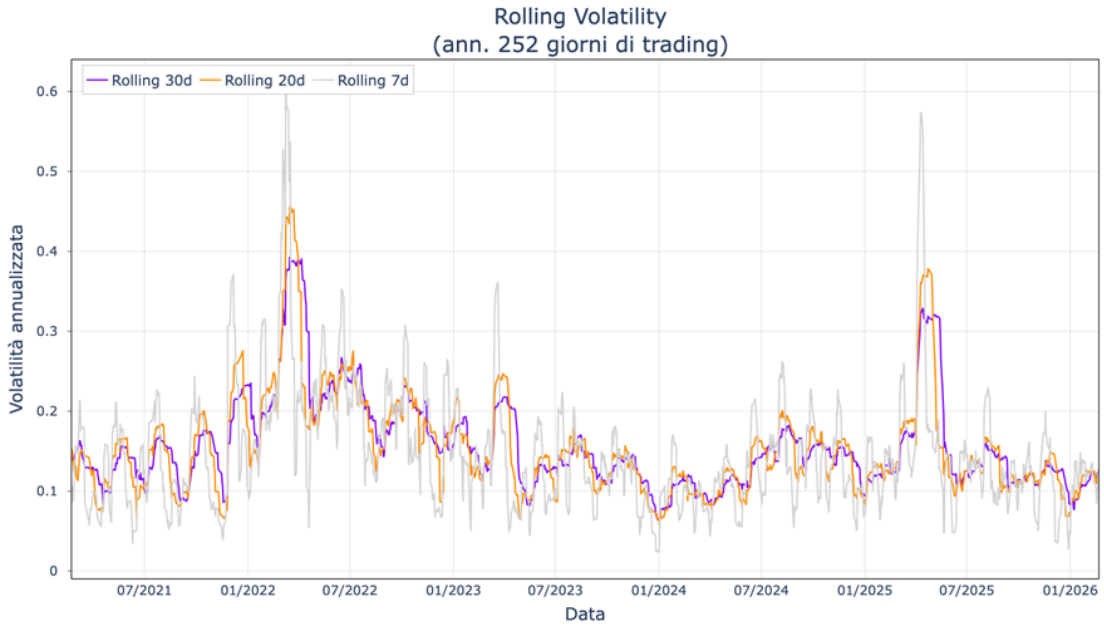


Figure 8: Confronto fra rolling volatility di diversa n.

Esiste quindi un trade-off tra il contenuto informativo della misura e la sua reattività ad improvvise variazioni dei rendimenti. Se il prezzo di un asset subisce uno sbalzo improvviso, una rolling volatility con finestra maggiore attribuisce un peso minore a questo sbalzo. Possibilmente, sbalzi improvvisi possono significare l'inizio ad un nuovo regime di mercato: per questo se la stima della volatilità storica è funzionale alla previsione della volatilità futura, la scelta di un ampio campione di osservazioni può portare a risultati meno validi. La stima giornaliera di volatilità comunemente usata per stime di previsione del rischio utilizza infatti intervalli di tempo contenuti di 20 e 40 giorni.

L'effetto ghosting - Un problema dovuto a questa misura di volatilità si manifesta nel momento in cui il dato relativo ad uno shock esce dalla finestra temporale considerata, provocando una diminuzione

della volatilità che in realtà non si è verificata. Mentre è normale che la rolling volatility aumenti incontrando shock avanti nel tempo, è da considerare imprecisa una stima che viene influenzata quando questi ultimi escono dal campione.

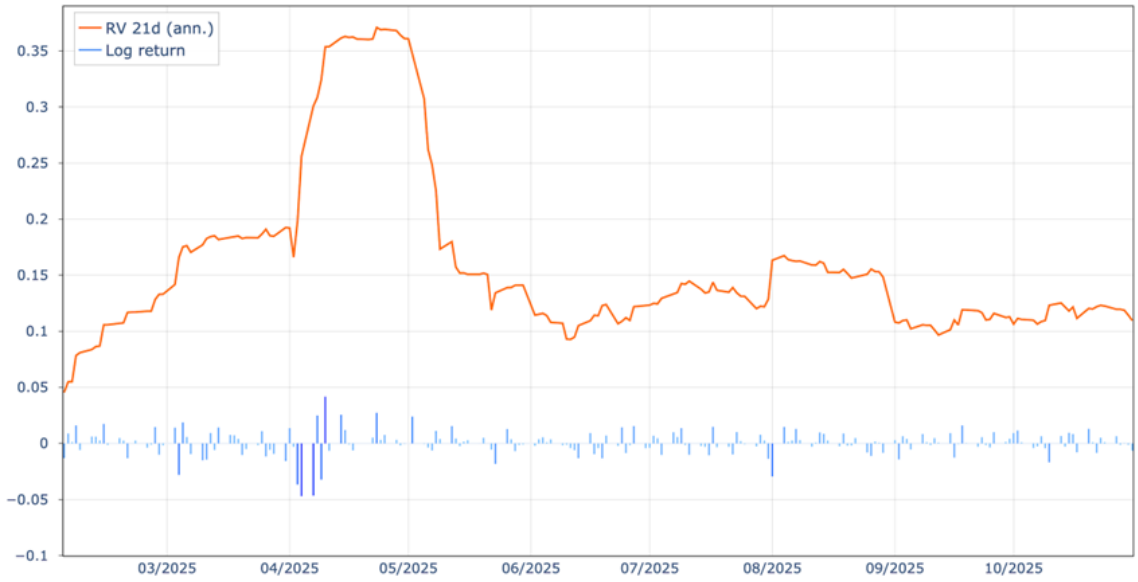


Figure 9: Confronto fra RV 21d e ampiezza log-rendimenti giornalieri.

Vediamo ad esempio, come sarebbe misurare la volatilità dell'indice Euro Stoxx 50 causata dall'annuncio dei dazi di Trump di aprile 2025. In particolare, confrontiamo una rolling volatility a 21 giorni con l'ampiezza dei log-rendimenti giornalieri, sapendo che 21 giorni corrispondono a circa un mese di calendario.

Si nota come ad inizio aprile uno sbalzo nei rendimenti dell'indice comporta un considerevole aumento di volatilità, ma una volta che queste osservazioni escono dalla finestra la volatilità diminuisce drasticamente.

Allo scopo superare i problemi appena descritti, JP Morgan introduce nel 1994 il modello a media mobile esponenziale (EWMA), il quale verrà approfondito nella prossima sezione, insieme ad altri modelli di stima della volatilità.

2.3 Realized volatility

Dopo aver visto come calcolare la volatilità storica su una finestra di n rendimenti giornalieri, vediamo come può essere costruita una misura di volatilità backward-looking che ha lo scopo di rappresentare l'effettiva volatilità realizzata. Volatilità storica e realizzata sono termini che vengono spesso intercambiati tra di loro in quanto entrambe le misure calcolano la deviazione standard dei rendimenti di una serie passata. La chiamiamo realized volatility in quanto si propone di stimare l'effettiva volatilità avvenuta usando i movimenti di prezzo realizzati all'interno di un giorno di trading (intraday). I rendimenti vengono solitamente calcolati ogni 5 minuti, proprio per catturare l'interno percorso del prezzo durante la giornata. La formula per la volatilità realizzata è la seguente, dove $r^2_{t,i}$ indica il rendimento osservato

nel sotto intervallo i-esimo del giorno t, mentre M rappresenta il numero di osservazioni intraday.

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2$$

A primo impatto, la formula potrebbe non sembrare il calcolo della varianza in senso ordinario. Questo perché in presenza di intervalli intraday molto brevi, il rendimento atteso all'interno di ciascun sottointervallo è generalmente considerato trascurabile, ovvero, la media dei rendimenti è approssimativamente nulla.

Questa misura di volatilità intraday costituisce la migliore stima della variabilità dei rendimenti avvenuta in passato: grazie all'alto contenuto informativo, aumentando la frequenza di campionamento, la realized variance converge alla integrated variance rendendo la misura consistente con la volatilità latente.

Nel caso in cui non siano disponibili dati intraday, non è possibile costruire una misura di volatilità realizzata nel suo significato stretto. In tali situazioni, è necessario utilizzare proxy basate su dati giornalieri, che risultano tuttavia meno informative rispetto alle misure ottenute da dati ad alta frequenza.

2.4 Stimatori range-based di volatilità

Per ovviare al problema dell'assenza di dati infragiornalieri è possibile ricorrere all'utilizzo di stimatori che, generalmente, si basano su quattro dati giornalieri: apertura, minimo, massimo, chiusura; comunemente detti OHLC. Rispetto ad una misura di RV che utilizza solo un dato giornaliero (prezzo di chiusura), questi usano il calcolo del range infragiornaliero, portando principalmente due vantaggi:

- Maggiore efficienza: le informazioni aggiuntive permettono di ottenere stime meno rumorose, quindi con minor varianza.
- Robustezza: alcuni stimatori sono in grado di gestire meglio i salti e gli effetti microstrutturali che una misura con un solo dato giornaliero non riuscirebbe a catturare.

  Vediamo nel dettaglio alcuni tre degli stimatori range-based più utlizzati e il loro confronto:

- Stimatore Parkinson: $\hat{\sigma}_{P,t}^2 = \frac{1}{4\ln 2}\left[\ln\left(\frac{H_t}{L_t}\right)\right]^2$

Si tratta dello stimatore range-based più semplice in quanto utilizza solo massimo e minimo giornalieri. L'intuizione è che l'ampiezza del range giornaliero contenga più informazione rispetto al solo movimento tra due prezzi di chiusura. Micheal Parkinson (1980) dimostra infatti, che il suo stimatore è circa 5.2 più efficiente della semplice misura dei rendimenti al quadrato close-to-close.

- Stimatore Garman-Klass: $\hat{\sigma}_{GK,t}^2 = \frac{1}{2}\left[\ln\left(\frac{H_t}{L_t}\right)\right]^2 - (2\ln 2 - 1)\left[\ln\left(\frac{C_t}{O_t}\right)\right]^2$

  Rappresenta un'estensione dello stimatore di Parkinson, che include anche i prezzi di apertura e chiusura, rendendo la misura molto più precisa (fino a 8 volte rispetto ai rendimenti al quadrato).

I due stimatori appena descritti assumono per approssimazione una media dei rendimenti nulla all'interno del giorno di trading. Infatti, entrambi gli stimatori sono derivati da un moto browniano geometrico con drift nullo μ=0 : questo significa ipotizzare che il mercato non stia seguendo una direzione prevalente e che il prezzo si muova casualmente attorno al valore iniziale, senza tendenza di fondo. L'implicazione principale, è che queste stime possano sovrastimare la volatilità nei giorni in cui il prezzo dell'asset sia cambiato in modo consistente tra min/max, mostrando però una tendenza ribassista/rialzista costante. Teniamo conto però, che l'errore nella stima potrebbe essere comunque minimo, stando comunque considerando variazioni infragiornaliere.

- Stimatore Rogers-Satchell: $\hat{\sigma}^2_{RS,t} = \ln\left(\frac{H_t}{O_t}\right)\ln\left(\frac{H_t}{C_t}\right) + \ln\left(\frac{L_t}{O_t}\right)\ln\left(\frac{L_t}{C_t}\right)$

In contrasto con i due stimatori precedenti, quello di Rogers e Satchell (1991) è in grado di fornire una stima della volatilità non influenzata dalla tendenza direzionale dei prezzi. La stima rimane unbiased anche quando il mercato mostra forti trend infragiornalieri.

Il confronto degli stimatori negli studi

Nella letteratura econometrica, la scelta dello stimatore ottimale dipende dal bilanciamento tra efficienza statistica e robustezza. Sebbene il modello di Garman-Klass sia spesso citato come il più efficiente in termini di riduzione della varianza , gran parte della ricerca empirica moderna, tra cui i lavori di Alizadeh et al. (2002), indica Rogers-Satchell come lo strumento più affidabile per gli indici azionari. Poiché asset come l'Euro Stoxx 50 presentano frequentemente trend direzionali e una media dei rendimenti non nulla, la capacità di Rogers-Satchell di non sovrastimare la volatilità in presenza di drift lo rende qualitativamente superiore.

2.4.1 Confronto empirico tra diverse misure di volatilità realizzata

Usiamo ora il dataset dell'indice Euro Stoxx 50, per confrontare gli stimatori range based appena visti con il classico proxy di volatilità r_t^2.

PROXY	MEDIA	SD	SKEWNESS	KURTOSIS
r^2t	0.00019578	0.0005490	11.11	218.61
Parkinson	0.0001360	0.0002655	7.72	96.26
Garman-Klass	0.0001248	0.0002440	8.19	104.66
Rogers-Satchell	0.0001243	0.0002783	10.22	167.53

Table 5: Statistiche descrittive delle diverse stime di volatilità realizzata.

Le statistiche descrittive mostrano una differenza abbastanza chiara tra gli stimatori considerati e la proxy di volatilità calcolata con r_t^2 , la quale risulta molto più rumorosa, con una deviazione standard 0.000549. Gli stimatori producono invece valori medi simili ma con una variabilità molto più contenuta, suggerendo una misura più stabile della volatilità giornaliera. Tra questi, Garman-Klass presenta la SD più bassa, Parkinson mostra invece i valori più bassi di skewness e kurtosis, suggerendo una distribuzione leggermente meno estrema, infine, Rogers-Satchell si colloca in una posizione intermedia, risultando meno disperso dei squared returns ma più irregolare rispetto agli altri due range-based. Le statistiche

confermano che gli stimatori basati su OHLC producono proxy più stabili e meno dominate da spike episodici rispetto alla misura costruita sui soli squared returns.



Figure 10: Serie temporali dei proxy.

Confrontando le serie lungo la linea temporale, si nota come tutte le misure catturano i periodi in cui i mercati finanziari hanno avuto volatilità più alta. Si nota facilmente, dalla misura sull'asse ordinate e dalla forma generale, come queste stime siano significativamente meno rumorose e disperse rispetto agli squared returns (notare l'asse delle ordinate). Dalle statistiche descrittive sono emersi valori di skewness e kurtosi molto alti, questo era in realtà atteso. Infatti, elevata skewness positiva indica che la distribuzione è fortemente asimmetrica a destra: la maggior parte delle osservazioni si concentra su valori bassi, mentre un numero limitato di giornate genera valori molto elevati. L'eccesso di kurtosis segnala invece la presenza di code pesanti e di osservazioni estreme più frequenti di quanto avverrebbe in una distribuzione normale. Nel caso delle misure di volatilità questo comportamento è normale, perché la volatilità è non negativa, tende a restare bassa per molti giorni e aumenta bruscamente solo durante fasi di stress o turbolenza.

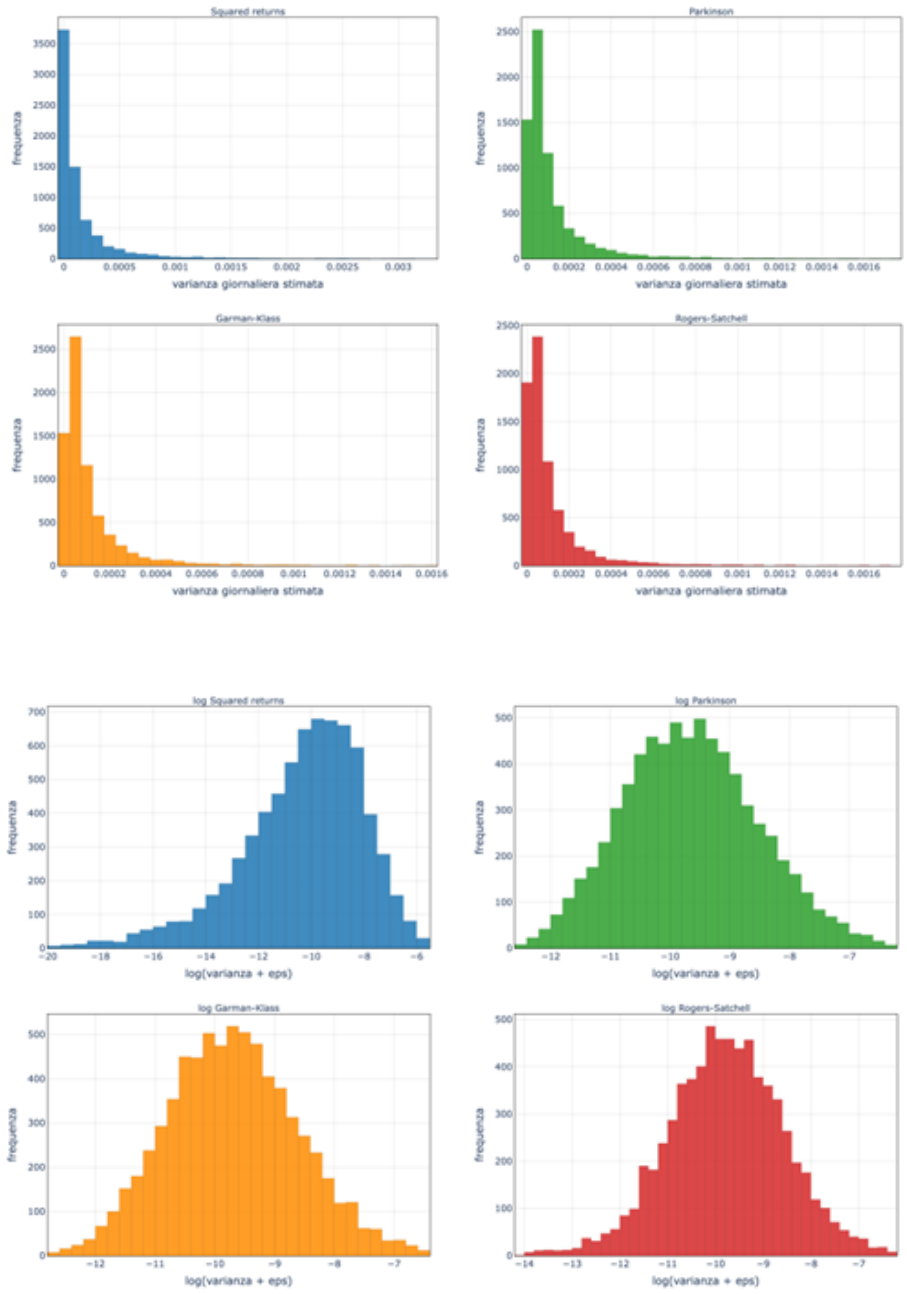


Figure 11: Distribuzioni degli stimatori.

I grafici delle distribuzioni in livello confermano chiaramente questa struttura. Tutte le proxy mostrano una forte concentrazione vicino a zero e una coda lunga verso destra, ma nei squared returns tale caratteristica è molto più accentuata: la distribuzione appare più compressa nei valori bassi e più estesa nella coda, segnalando una maggiore sensibilità agli shock estremi. Gli stimatori range-based mantengono la stessa asimmetria generale, ma in forma meno marcata, producendo distribuzioni relativamente più compatte e regolari. I grafici delle distribuzioni in logaritmo mostrano invece che la trasformazione logaritmica attenua sensibilmente l'asimmetria e riduce il peso relativo delle code estreme.

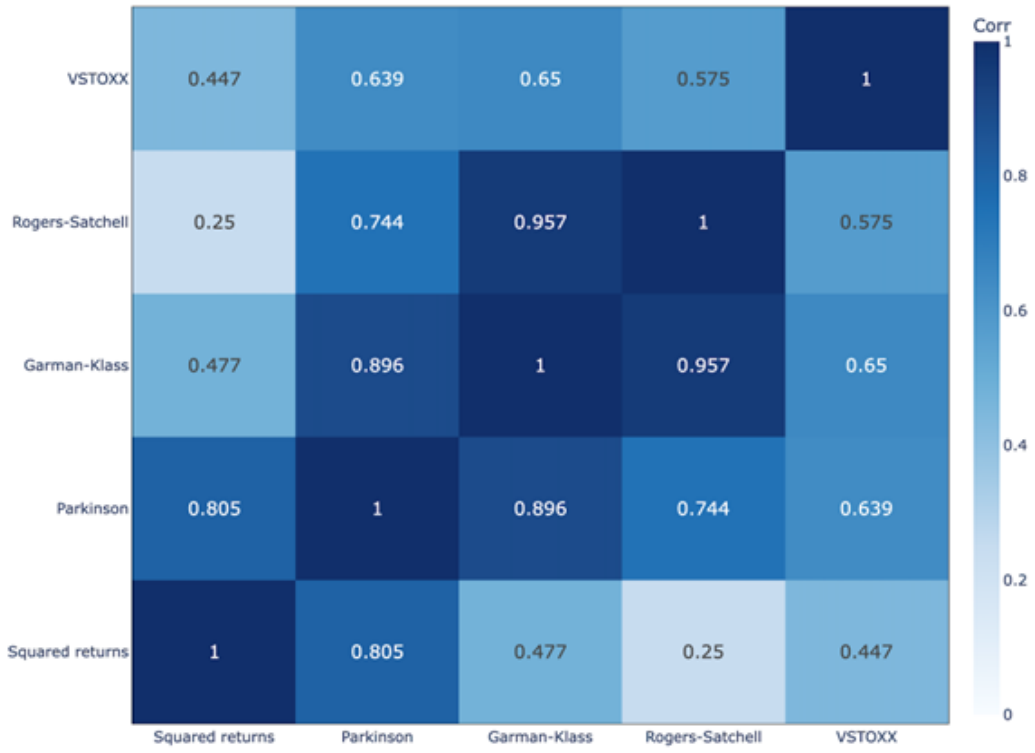


Figure 12: Matrice di correlazione tra proxy.

Studiando la correlazione di Pearson, si può vedere facilmente la relazione tra i vari proxy. Alla matrice è stata aggiunta la misura del VSTOXX, che verrà approfondita subito dopo: per ora, serve per capire quale stimatore si muove in modo più coerente con la percezione del rischio espressa dagli operatori. In questo aspetto, GK e PK sono i migliori.

Guardando alle altre correlazioni, emerge un legame fortissimo tra GK e RS ed una correlazione altrettanto forte tra PK e squared returns.

L'analisi congiunta delle statistiche descrittive, dei grafici delle distribuzioni, della dinamica temporale e della matrice di correlazione suggerisce che i squared returns costituiscano la proxy meno adatta a rappresentare la volatilità giornaliera, in quanto più dispersi, più estremi e meno coerenti con le altre misure considerate. .

Tra gli stimatori considerati, Garman-Klass appare preferibile come proxy principale, poiché combina una minore dispersione complessiva, una forte coerenza con le altre misure range-based e la correlazione più elevata con il VSTOXX.

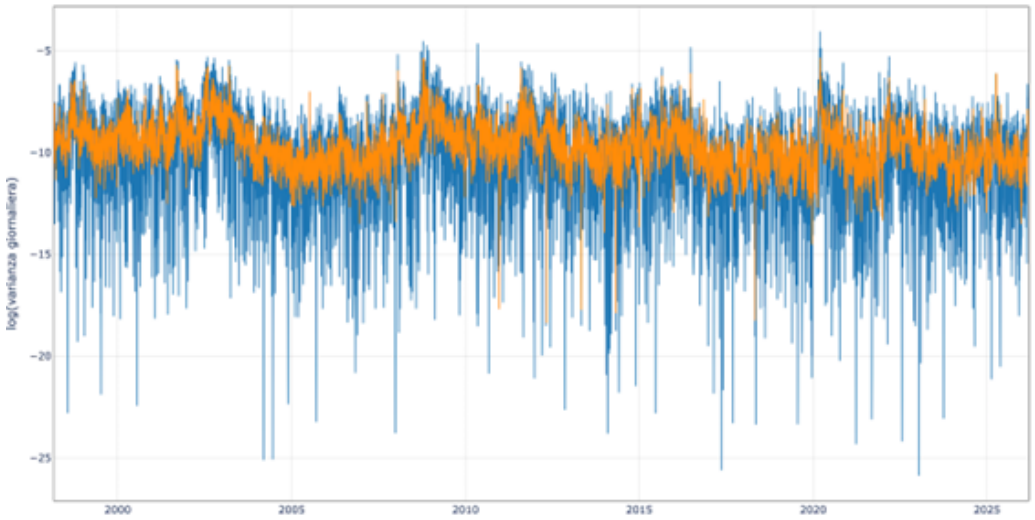


Figure 13: Confronto tra squared returns e GK in log-varianza.

2.5 La volatilità implicita dell'Euro Stoxx 50

Come già citato nella sezione 1 La volatilità implicita è ricavata dai prezzi delle opzioni e riflette le aspettative degli operatori di mercato riguardo alla volatilità futura del sottostante. In questo contesto si inserisce l'EURO STOXX 50 Volatility Index (VSTOXX), che rappresenta uno dei principali indicatori della volatilità attesa nel mercato azionario europeo. A differenza di alcune misure tradizionali di volatilità implicita basate sulle sole opzioni at-the-money, il VSTOXX non misura la volatilità implicita di una singola categoria di opzioni, bensì la varianza implicita aggregata su un insieme di opzioni con una determinata scadenza. Questo approccio consente di incorporare le informazioni provenienti da un ampio spettro di prezzi di esercizio, rendendo la misura più rappresentativa delle aspettative complessive del mercato. Le opzioni sull'indice EURO STOXX 50 costituiscono inoltre uno dei prodotti con il più elevato volume di scambi sulla piattaforma di derivati Eurex, il che contribuisce a garantire una maggiore liquidità e affidabilità delle informazioni contenute nei prezzi delle opzioni utilizzate nel calcolo dell'indice. Il modello di calcolo del VSTOXX è stato sviluppato congiuntamente da Goldman Sachs e Deutsche Börse. Tale metodologia offre importanti vantaggi in termini di attività di trading, copertura del rischio (hedging) e sviluppo di prodotti derivati basati sulla volatilità, rendendo il VSTOXX uno strumento particolarmente rilevante sia per gli operatori di mercato sia per l'analisi della dinamica del rischio nei mercati finanziari europei. L'indice VSTOXX è costruito a partire dalla stima della varianza implicita ottenuta dai prezzi delle opzioni sull'indice Euro Stoxx 50. In particolare, ciascun sotto-indice VSTOXX, associato a una specifica scadenza, è definito direttamente dalla STOXX Strategy Index Guide[6] come:

$$\text{SubIndex}_i = 100 \cdot \sqrt{\sigma_i^2}$$

Dove σ_i^2rappresenta la varianza implicita calcolata come:

$$\sigma_i^2 = \frac{2}{T_i} \sum_j \frac{\Delta K_{i,j}}{K_{i,j}^2} e^{r_i T_i} M(K_{i,j}) - \frac{1}{T_i} \left(\frac{F_i}{K_{i,0}} - 1 \right)^2$$

dove:

- T_i è il tempo a scadenza dell'opzione;
- $K_{i,j}$ rappresenta gli strike delle opzioni considerate;
- $\Delta K_{i,j}$ è la distanza tra strike adiacenti;
- $M(K_{i,j})$ è il prezzo dell'opzione out-of-the-money;
- F_i è il prezzo forward dell'indice;
- $K_{i,0}$ è lo strike più vicino al forward;
- r_i è il tasso di interesse privo di rischio.

L'indice VSTOXX a scadenza fissa (tipicamente 30 giorni) viene quindi ottenuto come media ponderata nel tempo di due sotto-indici con scadenze più vicine, secondo un meccanismo di interpolazione lineare:

$$\text{VSTOXX} = 100 \cdot \sqrt{\frac{T_{st}}{T_{365}}\,\sigma_{st}^2 \cdot \frac{T_{lt} - T_{tm}}{T_{lt} - T_{st}} + \frac{T_{lt}}{T_{365}}\,\sigma_{lt}^2 \cdot \frac{T_{tm} - T_{st}}{T_{lt} - T_{st}}} \cdot \frac{T_{365}}{T_{tm}}$$

dove: T_{st} e T_{lt} sono i tempi a scadenza (in secondi) dei sotto-indici a breve e a lungo termine;

T_{tm} è l'orizzonte temporale target (ad esempio 30 giorni);

T_{365} rappresenta il numero di secondi in un anno.

Come vediamo nella figura 14 il VSTOXX ha attraversato fasi di forte volatilità in corrispondenza delle crisi economiche globali (evidenziate in grigio), con un evidente ritorno verso la media nei periodi di stabilità. Il meccanismo di mean reversion ,già citato nella sezione dei fatti stilizzati, osservato nel VSTOXX è attribuibile a diverse dinamiche di mercato. Durante fasi di estrema incertezza, come nel 2008 o nel crollo indotto dalla pandemia COVID-19, il VSTOXX raggiunge valori molto alti. Tuttavia, man mano che la crisi si attenua e il mercato recupera fiducia, l'indice tende a scendere verso il suo valore medio storico. Questo comportamento riflette l'idea che, sebbene gli shock possano spingere temporaneamente la volatilità a livelli elevati, esiste una forza di attrazione che riporta il sistema verso un equilibrio di lungo periodo.

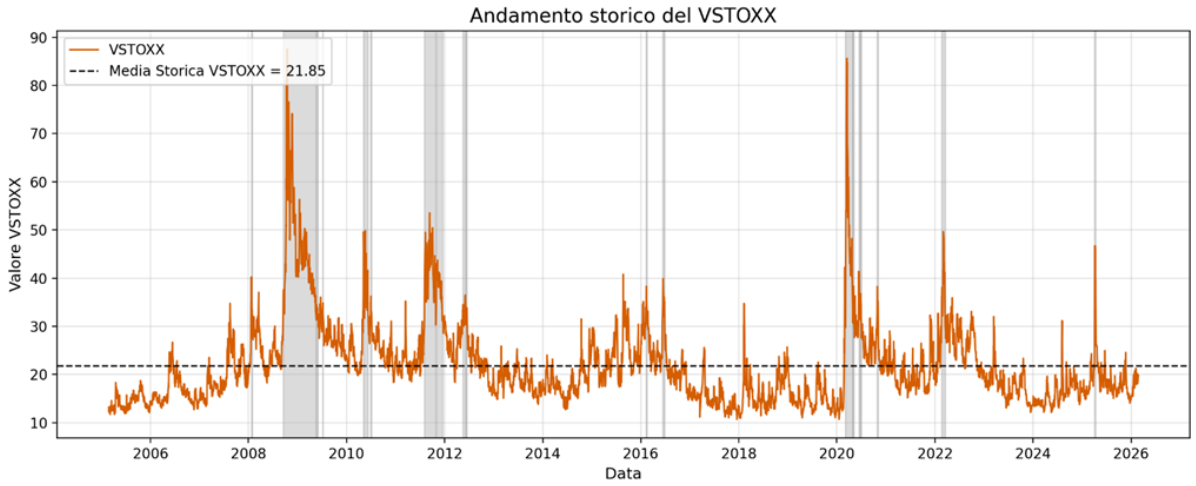


Figure 14: Andamento storico del VSTOXX dal 2005 ad oggi.

Il fenomeno di mean reversion non è soltanto un fenomeno statistico, ma anche un riflesso della psicologia di mercato. Durante i periodi di panico, la paura e l'incertezza spingono gli investitori a reagire in modo eccessivo, facendo lievitare il VSTOXX. Con il tempo, quando la pressione di mercato si attenua e le notizie positive iniziano a riprendere il sopravvento, il panico si dissolve e l'indice ritorna verso i livelli medi.

2.5.1 Relazione tra il VSTOXX e l'EURO STOXX 50

La dinamica tra l'indice azionario e la volatilità implicita è fortemente legata al comportamento degli investitori nei mercati delle opzioni. In presenza di aspettative di ribasso del mercato azionario, gli operatori tendono ad aumentare la domanda di strumenti di copertura, in particolare opzioni put, al fine di proteggersi da possibili perdite. Questo incremento della domanda di protezione determina un aumento del prezzo delle opzioni e, conseguentemente, della volatilità implicita, causando un rialzo dell'indice VSTOXX. In un contesto di mercato ribassista, l'aumento dell'incertezza e della percezione del rischio induce gli investitori a richiedere premi per il rischio più elevati, contribuendo ulteriormente alla crescita della volatilità implicita. Al contrario, nelle fasi di mercato rialzista la domanda di strumenti di copertura tende a ridursi, determinando generalmente una diminuzione del livello del VSTOXX. Questo comportamento genera una relazione inversa e asimmetrica tra i rendimenti dell'indice azionario e le variazioni della volatilità implicita: le variazioni negative dell'indice tendono ad avere un impatto più marcato sulla volatilità rispetto alle variazioni positive.

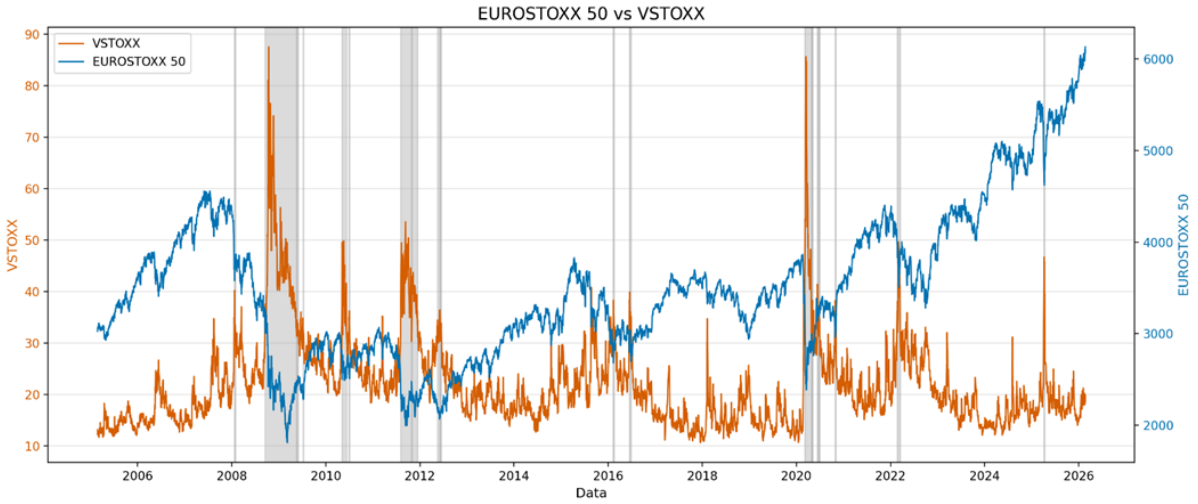


Figure 15: Andamento comparato tra il VSTOXX e l'indice EUROSTOXX 50 dal 2005 ad oggi.

Per analizzare formalmente questa relazione è stata stimata una regressione del tipo:

$$R_{\text{VSTOXX}} = \beta_0 + \beta_1 R_{\text{STOXX}} + \beta_2 R^-_{\text{STOXX}} + \varepsilon$$

dove:

- R_{VSTOXX} rappresenta il rendimento percentuale dell'indice VSTOXX;
- β_0 è il termine costante della regressione;
- R_{STOXX} indica il rendimento percentuale dell'indice Eurostoxx50;
- R^-_{STOXX} rappresenta la componente negativa del rendimento dell'indice, definita come:

$$R^-_{\text{STOXX}} = R_{\text{STOXX}} \cdot \mathbf{I}_{\{R_{\text{STOXX}}<0\}}$$

con la funzione indicatrice:

$$I_{\{R_{\text{STOXX}}<0\}} = \begin{cases} 1 & \text{se } R_{\text{STOXX}} < 0 \\ 0 & \text{altrimenti} \end{cases}$$

- β_1 misura la sensibilità delle variazioni del VSTOXX rispetto ai movimenti dell'indice Eurostoxx50;
- β_2 cattura l'effetto addizionale associato ai movimenti negativi del mercato;
- ε rappresenta il termine di errore.

Parametro	Valore	P-Value	T-Statistic
Intercept	-0.0069	0.074	-1.786
β_1	-16.459	0.000	-4.132
β_2	-16.315	0.037	-2.083

Table 6: Risultati della regressione:$R_{\text{VSTOXX}} = \beta_0 + \beta_1 R_{\text{STOXX}} + \beta_2 R^-_{\text{STOXX}} + \varepsilon$.

I risultati della regressione evidenziano una relazione negativa e statisticamente significativa tra i rendimenti dell'Eurostoxx50 e le variazioni del VSTOXX. Il coefficiente associato al rendimento dell'indice ($\beta_1 \approx$-1.65) indica che, in media, un aumento dell'1% dell'Eurostoxx50 è associato a una riduzione di circa 1.6% del VSTOXX. Inoltre, il coefficiente associato alla componente negativa dei rendimenti ($\beta_2 \approx$-1.63) risulta statisticamente significativo, evidenziando la presenza di una risposta asimmetrica della volatilità implicita ai movimenti del mercato. In particolare, nei giorni in cui il rendimento dell'Eurostoxx50 è negativo, la variazione del VSTOXX risulta più marcata rispetto ai giorni di mercato positivo. Considerando l'effetto complessivo, una diminuzione dell'1% dell'Eurostoxx50 è associata a un incremento del VSTOXX di circa 3.3%, confermando che la volatilità implicita reagisce in maniera più intensa agli shock ribassisti del mercato azionario. Questo risultato è coerente con il fatto stilizzato noto come leverage effect introdotto nella sezione dei fatti stilizzati e che verrà riscontrato nella sezione 3 in particolare nel modello GARCH. Nel complesso, l'analisi empirica conferma che il VSTOXX svolge un ruolo fondamentale come indicatore della percezione del rischio nel mercato azionario europeo, mostrando una forte sensibilità alle variazioni dell'indice Eurostoxx50 e una risposta particolarmente accentuata durante le fasi di mercato negativo.

3 Volatility forecasting

3.1 Introduzione alla previsione di volatilità

utilizzate per approssimarla, il passaggio successivo consiste nel chiedersi se tale grandezza possa essere, almeno in parte, prevista. Parlare di volatility forecasting significa proprio questo: non cercare di anticipare se il mercato salirà o scenderà, ma provare a stimare quanto intensa sarà la variabilità dei rendimenti nel periodo successivo. Infatti, mentre con l'analisi direzionale ci si chiede quale sia il prezzo atteso (la sua media, per intenderci), con la previsione di volatilità ci chiediamo quale sia la magnitudine dei rendimenti futuri, prendendo come dati quelli passati. Questo tipo di analisi è fondamentale per il risk management e altre applicazioni già citate nella sezione 1.1.

L'idea alla base è che la volatilità segue dei comportamenti statistici ricorrenti, i fatti stilizzati (di cui abbiamo parlato nella sezione 1.3), che almeno in parte, possono essere catturati e previsti da modelli econometrici.

3.1.1 Dataset e modelli

In questa sezione, usiamo come dati i log-rendimenti giornalieri dell'indice EURO STOXX 50. Il campione analizzato copre il periodo compreso tra il 4 gennaio 2000 e il 31 marzo 2026, per un totale di 6.727 osservazioni. Il proxy di volatilità utilizzato è il Garman-Klass, poiché come verificato nella sezione 2.4.1, risulta un proxy più stabile e meno dominato da spike episodici rispetto alla misura costruita sui soli redimenti al quadrato.

Verrà introdotto per primo il modello a media mobile esponenziale EWMA, successivamente, i modelli appartenenti alla famiglia ARCH, che verranno analizzati secondo una diversa impostazione di lavoro, descritta nel paragrafo successivo.

3.1.2 Obiettivi dell'analisi

- Presentare EWMA come soluzione alle principali problematiche date dalla stima SMA.

  Per i modelli econometrici della famiglia ARCH, si vuole valutare:

- La capacità di prevedere la volatilità futura.
- L'attitudine a cogliere i fatti stilizzati della volatilità.
- La loro reattività nei periodi di maggiore turbolenza di mercato

3.2 Impostazione generale dell'analisi dei modelli -ARCH()

3.2.1 Metodologia di stima e previsione out-of-sample

Per evitare fenomeni di overfitting e garantire che ogni previsione sia costruita solo con l'informazione disponibile al momento della stima, la valutazione viene condotta in ambito out-of-sample. Il dataset dei log-rendimenti dell'indice Eurostoxx 50 è stato suddiviso in un set iniziale di stima (in-sample) pari all'80% delle osservazioni e un set di valutazione (out-of-sample) pari al restante 20%. La strategia di previsione adottata è quella della expanding window: per ogni istante nel periodo di test, il modello viene ri-stimato utilizzando tutte le informazioni disponibili da inizio campione fino a t, al fine di produrre la previsione per t+1. Tale procedura viene ripetuta iterativamente, aggiornando il set informativo a ogni passo, garantendo così che ogni previsione sia basata sulla massima quantità di dati storici disponibili.

3.2.2 Dataset

Log-rendimenti giornalieri (close-to-close) dell'indice Eurostoxx 50 (fonte: Bloomberg terminal). 6 727 osservazioni giornaliere (dal 4 gen 2000 al 31 mar 2026), suddivise con split 80/20.

- Training: 5 381 obs (2000-01-04 → 2021-01-04)
- Test: 1 346 obs (2021-01-05 → 2026-03-31)

I log-rendimenti vengono moltiplicati per 100, così da essere espressi in termini percentuali. Tale trasformazione non altera il contenuto informativo della serie, ma migliora la stabilità numerica della procedura di stima, evitando che l'ottimizzazione venga effettuata su valori troppo piccoli in valore assoluto.

3.2.3 Criteri di valutazione della qualità previsiva

Dal punto di vista statistico, l'obiettivo è misurare quanto le previsioni dei modelli si avvicinino ai valori osservati del proxy Garman-Klass. A tale scopo, vengono utilizzate un insieme di metriche complementari, ciascuna delle quali coglie differenti dimensioni dell'errore di previsione. In particolare, vengono considerate tre loss function standard nella letteratura: il Mean Absolute Error (MAE), il Root Mean Squared Error (RMSE) e la QLIKE. Il Mean Absolute Error (MAE) misura l'errore medio in valore

assoluto tra la varianza prevista $\hat{\sigma}_t^2$ e quella realizzata σ_t^2 è definito come:

$$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{\sigma}_t^2 - \sigma_t^2\right|$$

(specificare GK nel secondo termine della formula) Questa metrica penalizza in modo lineare le deviazioni, risultando meno sensibile alla presenza di errori estremi. Il Root Mean Squared Error (RMSE) attribuisce invece un peso maggiore agli errori di ampia entità, essendo basato sul quadrato delle deviazioni:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\sigma}_t^2 - \sigma_t^2\right)^2}$$

Di conseguenza, tale misura è particolarmente utile per valutare la capacità del modello di catturare correttamente i picchi di volatilità. QLIKE (Quasi-Likelihood Loss), una funzione di perdita specificamente progettata per la valutazione di previsioni di varianza (Patton, 2011), definita come:

$$\text{QLIKE} = \frac{1}{T}\sum_{t=1}^{T}\left[\ln\left(\hat{\sigma}_t^2\right) + \frac{\sigma_t^2}{\hat{\sigma}_t^2}\right]$$

Tale misura presenta proprietà teoriche rilevanti, tra cui l'invarianza rispetto a trasformazioni scalari del proxy di volatilità, rendendola particolarmente robusta nel confronto tra modelli. A queste metriche si affiancano ulteriori indicatori volti a cogliere aspetti specifici della qualità previsiva. In particolare, il Mean Error consente di individuare eventuali distorsioni sistematiche nelle previsioni ed è definito come:

$$\text{Mean Error} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\sigma}_t^2 - \sigma_t^2\right)$$

Un valore positivo (negativo) segnala una tendenza alla sovrastima (sottostima) della volatilità. Nota: Le metriche di errore sono calcolate sulla varianza giornaliera, una grandezza nell'ordine di 10^{-4} : i valori assoluti di MAE e RMSE saranno quindi rapportati alla media della varianza realizzata per essere interpretati correttamente.

Metriche high volatility: le misure di valutazione statistica appena descritte, saranno calcolata anche separatamente per i giorni in cui l'indice ha visto volatilità più alta (top 10%). In modo da valutare meglio la capacità previsiva del modello nei periodi con variabilità dei rendimenti maggiore.

Coerenza tra input del modello e metriche di valutazione

Come illustrato in 2.4.1, l'assenza di dati intraday ad alta frequenza ha portato ad adottare il Garman-Klass come proxy della volatilità realizzata a partire dai prezzi OHLC giornalieri. Occore però tenere conto di un importante disallineamento strutturale tra questo benchmark e i modelli di previsione che vedremo in 3.4. Solitamente i modelli ARCH/GARCH vengono stimati sui rendimenti calcolati tra il momento di chiusura del mercato del giorno t e del giorno t-1, seguendo la definizione di rendimento considerata fino ad ora. Lo stimatore Garman-Klass cattura invece esclusivamente la variabilità dei rendimenti intraday, usando i prezzi di apertura, minimo, massimo e chiusura, senza però considerare la chisura del giorno precedente C_{t-1}. Questo produrebbe un bias positivo sistematico nelle previsioni,

documentato nella letteratura (Hansen e Lunde, 2005; Patton, 2011) Per far fronte a questa incoerenza, i modelli sono stati stimati sui rendimenti open-to-close : $r_t^{o2c} = \ln \frac{C_t}{O_t}$ che isolano la sola componente intraday del rendimento. In questo modo la grandezza prevista dal modello diventa omogenea con ciò che il proxy Garman-Klass effettivamente misura.

Le correlazioni analizzate

- Correlazione tra varianza prevista e Garman-Klass: serve a valutare quanto bene il modello riesca a seguire la dinamica temporale della volatilità realizzata. Non misura tanto la precisione puntuale del forecasting, quanto la capacità del modelli di muoversi nella direzione corretta al variare della volatilità.

- Correlazione tra la volatilità prevista e VSTOXX: non serve come misura diretta di accuratezza previsiva, ma come indicatore della coerenza economica delle previsioni del modello. Se la volatilità prevista risulta molto correlata con l'indice di volatilità implicita, ciò suggerisce che il modello si muove in modo coerente con la percezione del rischio del mercato.

Focus sullo stress finanziario della guerra in Iran

Per valutare la qualità previsiva dei modelli, mostreremo come questi ultimi hanno tentato di prevedere la dinamica della volatilità nel periodo di febbraio e marzo 2026. Con lo scoppio della guerra in Iran ad inizio marzo, i mercati hanno visto sempre più volatilità finanziaria, rendendo la previsione più difficile. Per questo, risulta interessante spostare l'analisi anche in questo periodo, per testare la capacità dei modelli di adattarsi a contesti di stress finanziario.

3.2.4 Studio dei residui

Accanto alla valutazione previsiva, viene svolta una verifica diagnostica essenziale sui residui dei modelli, in particolare considerando:

- Grafico dell'errore assoluto out-of-sample (OOS), per osservare come varia nel tempo la qualità della previsione, mostrando in quali momenti il modello sbaglia maggiormente

- ACF dei residui standardizzati al quadrato: permette di capire in che misura il modello ha assorbito la componente di eteroschedasticità condizionata. I quadrati dei residui riflettono l'ampiezza residua degli shock. Se l'ACF dei residui mostra ancora autocorrelazioni significative, questo suggerisce che il modello non ha catturato correttamente il clustering della volatilità

- Test di Ljung Box sui residui standardizzati al quadrato: fornisce una verifica formale di quanto osservato nell'ACF. Lo studio dei residui ha una funzione diversa ma complementare rispetto alla forecast evaluation. Le metriche out-of-sample ti dicono quanto bene il modello prevede; l'analisi dei residui ti dice se il modello è strutturalmente credibile nel modo in cui descrive la volatilità.

3.3 Il modello a media mobile esponenziale (EWMA)

Nella sezione precedente, è stata introdotto il concetto di rolling volatility (SMA) e le principali problematiche relative a questa misura di volatilità storica.

EWMA si propone come soluzione a queste implicazioni, infatti, rispetto al metodo Simple moving average, il modello introdotto da JP Morgan, calcola la media degli scarti quadratici dal rendimento medio attribuendo alle osservazioni una ponderazione non uniforme.

L'idea chiave è che l'informazione contenuta nei rendimenti passati non sia omogenea nel tempo, in quanto gli shock recenti risultano generalmente più informativi rispetto a quelli remoti spiegando che i mercati finanziari hanno una memoria che pesa maggiormente gli eventi vicini e decade nel tempo.

In questo contesto, il modello EWMA mira ad ottenere una stima della volatilità caratterizzata da un alto contenuto informativo e maggiore sensibilità agli shock recenti. La varianza condizionata viene espressa come una media ponderata dei rendimenti passati al quadrato:

$$\sigma_t^2 = \sum_{i=1}^{n} w_i \, r_{t-i}^2$$

dove $w_i > 0$ rappresenta il peso assegnato all'osservazione i-esima e la condizione di normalizzazione impone $\sum_{i=1}^{n} w_i = 1$. Le ponderazioni assegnate sono decrescenti nel tempo, facendo in modo che gli shock di prezzo escano in modo graduale dalla stima di volatilità, evitando quindi l'effetto di ghosting descritto in precedenza. I pesi decrescono esponenzialmente nel tempo secondo una costante di decadimento $\lambda \in (0, 1)$.

In particolare, viene soddisfatta la relazione ricorsiva:

$$w_{i+1} = \lambda w_i$$

Il parametro λ , noto come decay factor, governa il grado di persistenza della volatilità stimata. Valori elevati di λ (prossimi a 1) implicano un lento decadimento dei pesi e quindi una forte memoria del processo, mentre valori più bassi attribuiscono maggiore importanza alle osservazioni recenti, rendendo la stima più reattiva agli shock. Sotto l'ipotesi di rendimenti logaritmici a media nulla, il modello EWMA può essere riscritto in forma ricorsiva come:

$$\sigma_t^2 = (1 - \lambda) \, r_{t-1}^2 + \lambda \sigma_{t-1}^2$$

Dal punto di vista applicativo, il modello EWMA rappresenta un benchmark ampiamente utilizzato nella pratica ad esempio nel framework RiskMetrics, grazie alla sua semplicità e alla capacità di catturare le principali caratteristiche empiriche della volatilità finanziaria.

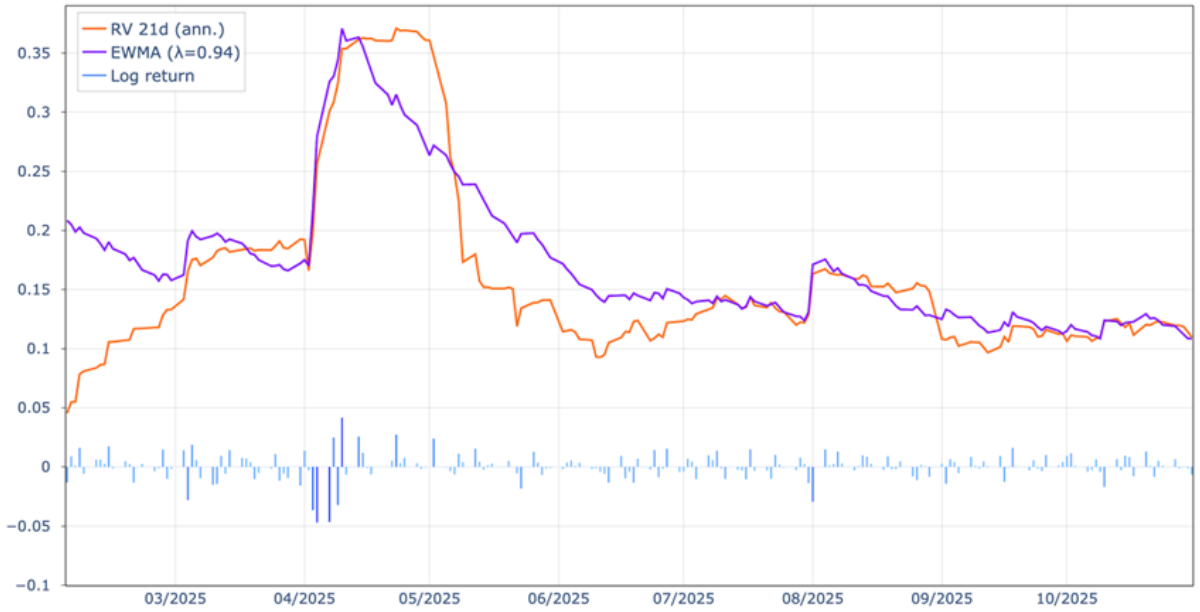


Figure 16: Confronto tra media mobile e modello EWMA $\lambda = 0.94$.

Riprendiamo l'esempio usato in precedenza per mostrare l'effetto ghosting che si verifica utilizzando la misura di rolling volatility a 21 giorni. La volatilità media mobile ora viene confrontata con il modello EWMA λ=0.94.

Si nota molto facilmente come, a differenza della RV 21d, il modello non subisce un calo improvviso, piuttosto, decade progressivamente.

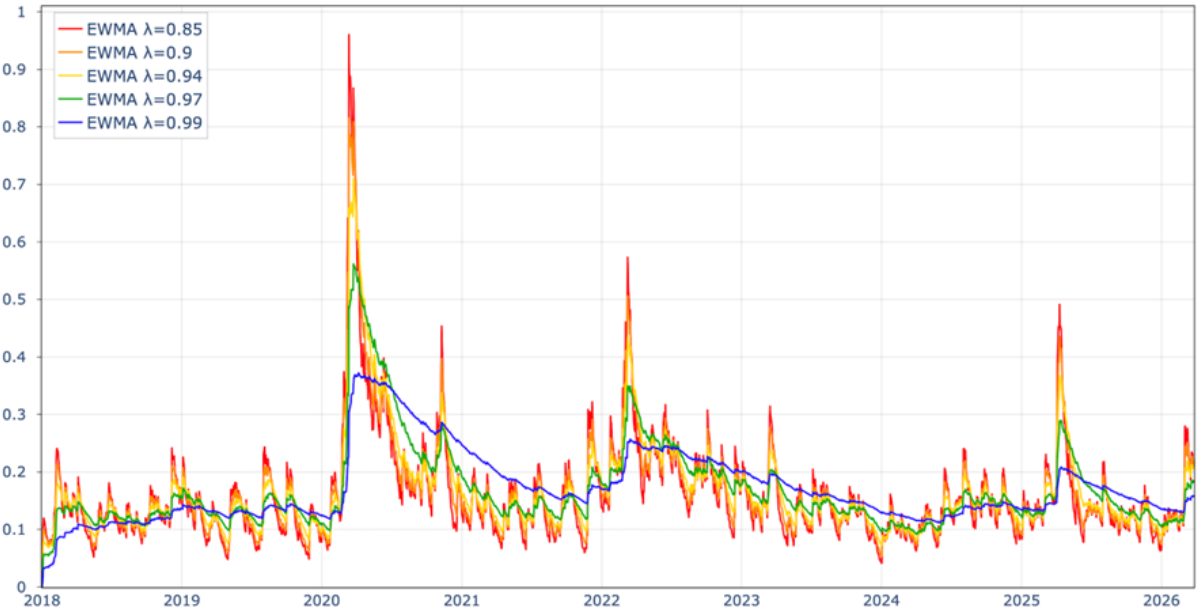


Figure 17: La scelta del decay factor λ.

La scelta del parametro di decadimento λ è l'aspetto più delicato del modello EWMA. Il valore assegnato a λ determina infatti, il compromesso tra reattività e persistenza. Se λ è basso, il modello dà

molto peso agli ultimi dati: la volatilità reagisce rapidamente, ma rischia di diventare più rumorosa. Se λ è alto, il modello smussa maggiormente la serie: la stima è più stabile, ma più lenta nell'adattarsi ai cambiamenti di regime. All'interno del framework RiskMetrics, il valore λ=0.94 viene adottato per le previsioni giornaliere perché rappresenta un compromesso empiricamente efficace tra reattività e stabilità: abbastanza basso da recepire gli shock recenti, ma abbastanza alto da evitare che la volatilità stimata diventi eccessivamente rumorosa. Per orizzonti più lunghi, come quello mensile, RiskMetrics utilizza invece λ=0.97, coerentemente con l'idea che la stima debba incorporare una memoria più lunga.

3.4 Modelli autoregressivi a eteroschedasticità condizionata: Introduzione

Nella sezione 1.3 è stato spiegato come la variabilità dei rendimenti non sia costante nel tempo, ma tenda a concentrarsi in fasi di calma e di turbolenza secondo la dinamica del volatility clustering. I modelli appartenenti alla famiglia ARCH (Autoregressive Conditional Heteroskedasticity), permettono di descrivere la varianza condizionata come una grandezza dipendente in modo autoregressivo dall'informazione passata. Per comprendere l'analisi successiva, è importante avere chiari questi due concetti fondamentali:

- Autoregressivo: indica che la varianza corrente viene modellata come funzione dei suoi determinanti passati, in particolare degli shock osservati in precedenza. Gli stadi futuri sono stimati usando una combinazione lineare dei valori passati.

- Eteroschedasticità condizionata: la varianza degli errori del processo cambia nel tempo in modo condizionato dall'informazione passata.

  Se nel caso di ipotesi di omoschedasticità intendiamo$\text{Var}(\varepsilon_t) = \sigma^2$ costante, nei modelli di eteroschedasticità condizionata, invece, si assume che la varianza sia tempo-dipendente e definita condizionatamente al passato, cioè $\text{Var}(\varepsilon_t \mid \mathcal{F}_{t-1}) = \sigma_t^2$, dove $\mathcal{F}_{t-1}$ rappresenta l'insieme delle informazioni disponibili fino al tempo t-1.

Nel modello ARCH originario, la varianza condizionata dipende dai quadrati degli errori passati. Ciò significa che non è il segno del rendimento a contare direttamente, ma la dimensione dello shock: un rendimento molto elevato in valore assoluto, positivo o negativo che sia, aumenta la varianza attesa del periodo successivo. Questo approccio permette di modellare in modo semplice la persistenza della volatilità, ma presenta anche un limite pratico: per descrivere una dinamica sufficientemente persistente può richiedere un numero elevato di ritardi, rendendo il modello poco parsimonioso.

Per superare tale limite, Bollerslev propose nel 1986 il modello GARCH (Generalized ARCH), che estende l'ARCH facendo dipendere la varianza condizionata non solo dagli shock passati, ma anche dalla varianza condizionata passata. In questo modo il modello riesce a catturare la persistenza della volatilità in maniera più compatta e realistica. Il caso più comune, il GARCH(1,1), è diventato uno standard applicativo proprio perché descrive bene il clustering della volatilità con una struttura molto parsimoniosa.

Un'ulteriore estensione è rappresentata dal modello EGARCH (Exponential GARCH), introdotto per tenere conto del fatto che, nei mercati finanziari, shock positivi e shock negativi non hanno necessariamente lo stesso effetto sulla volatilità futura (leverage effect).

3.5 Il modello ARCH(p)

Formalmente, il processo può essere descritto come:

$$r_t = \mu + \epsilon_t \qquad \epsilon_t = \sigma_t z_t \qquad z_t \sim \text{i.i.d. } (0,1)$$

dove:

- r_t rappresenta il rendimemento osservato al tempo t
- μ è la media costante del processo (ipotizzata come nulla)
- ϵ_t è il termine di errore o innovazione al tempo t
- z_t è una variabile aleatoria, indipendente e identicamente distribuita, con media 0 e varianza 1 (si assume N∼(0,1), o anche altre distribuzioni con code più pesanti).

Il termine di errore viene quindi scomposto in due parti: uno shock standardizzato z_t e la varianza condizionale evolve nel tempo secondo:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2$$

con $\omega > 0$ e $\alpha_i \geq 0$. Questa specificazione implica che la volatilità corrente dipende dai quadrati delle innovazioni passate: grandi shock (positivi o negativi) aumentano la volatilità futura, replicando il fenomeno del clustering. Un aspetto interessante è che, definendo: $\nu_t = \epsilon_t^2 - \sigma_t^2$ il modello può essere riscritto come:

$$\epsilon_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \nu_t$$

ovvero come un modello autoregressivo sui quadrati degli errori. Questo evidenzia come la dinamica della volatilità sia governata da una struttura AR applicata alla variabilità del processo. Dal punto di vista empirico, i modelli ARCH risultano efficaci nel catturare l'eteroschedasticità condizionale; tuttavia, presentano un limite importante: per descrivere adeguatamente la persistenza della volatilità osservata nei dati finanziari, è spesso necessario utilizzare un ordine p elevato, con conseguente aumento del numero di parametri da stimare.

3.5.1 ARCH(1) sui rendimenti dell'Eurostoxx 50

Iniziamo la nostra analisi stimando il modello ARCH di ordine 1, ARCH(1), in cui la varianza condizionata è espressa in questo modo:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2$$

Questo significa che il termine di errore viene stimato con 1 lag, in pratica, la varianza di oggi dipende dallo shock al quadrato di ieri, ovvero dal quadrato del rendimento inatteso osservato. La valutazione e analisi di questo processo, seguirà l'impostazione descritta in 3.2.

3.5.2 Stima dei parametri

I risultati dei coefficienti del modello sono riportati nella seguente tabella:

PARAMETRO	STIMA	STD ERROR	T_STAT	P_VALUE	CI 95%
μ	7.13025e-03	0.018051	0.395011	0.692834	[-0.028249, 0.042510]
ω	1.309794	0.068952	18.995644	1.85302e-80	[1.174648, 1.444941]
α_1	0.303143	0.045679	6.636409	3.21417e-11	[0.213613, 0.392674]

Table 7: Parametri Stimati (In-Sample).

- μ : la stima del rendimento medio è pari a 0.00713, con un p value di 0.69 Il parametro non è statisticamente significativo, inoltre, l'errore standard è maggiore della stima stessa e il CI comprende zero. Non abbiamo quindi evidenza statistica sufficiente per rifiutare l'ipotesi che il rendimento medio giornaliero sia nullo. È un risultato atteso e coerente con la letteratura: su base giornaliera la componente di drift dei prezzi azionari è trascurabile rispetto alla loro volatilità, e il modello conferma che non vi è una tendenza al rialzo o al ribasso statisticamente rilevabile.

- ω: la compente fissa della varianza è stimata a 1.3098 ed è altamente significativa. Questo parametro rappresenta il livello minimo di varianza condizionata: anche nei giorni più tranquilli, in assenza di shock di mercato, la volatilità dell'EURO STOXX 50 non scende mai sotto questa soglia. Economicamente, ciò riflette il fatto che il mercato azionario europeo mantiene sempre un grado irriducibile di incertezza, legato a fattori strutturali come il flusso continuo di informazioni, l'eterogeneità degli operatori e il rischio sistemico.

- α_1: anche la reattività agli shock è altamente significativa e stimata a 0.3031. Si tratta di un risultato importante della nostra stima, in quanto indica che il 30.3% dello shock al quadrato del giorno precedente si trasferisce nella varianza condizionata del giorno successivo. Vediamo un esempio: se ieri il mercato ha registrato un rendimento inatteso del 2%, il contributo alla varianza odierna sarà $0.294 \times 2^2 = 1.212$; mentre se lo shock fosse stato del 4%, il contributo salirebbe a $0.294 \times 4^2 = 4.848$. Quindi, un rendimento inatteso doppio genera un effetto sulla varianza quattro volte maggiore, per via dell'elevazione al quadrato. Il meccanismo spiegato riproduce il fenomeno del volatility clustering: le giornate di forte turbolenza alimentano la volatilità attesa nei giorni successivi, e analogamente le fasi di calma tendono a perpetuarsi.

3.5.3 ARCH(1) vs Garman-Klass

Analizziamo ora la serie temporale, di seguito è mostrato il confronto visivo tra ARCH(1) e il proxy di volatilità realizzata Garman-Klass.

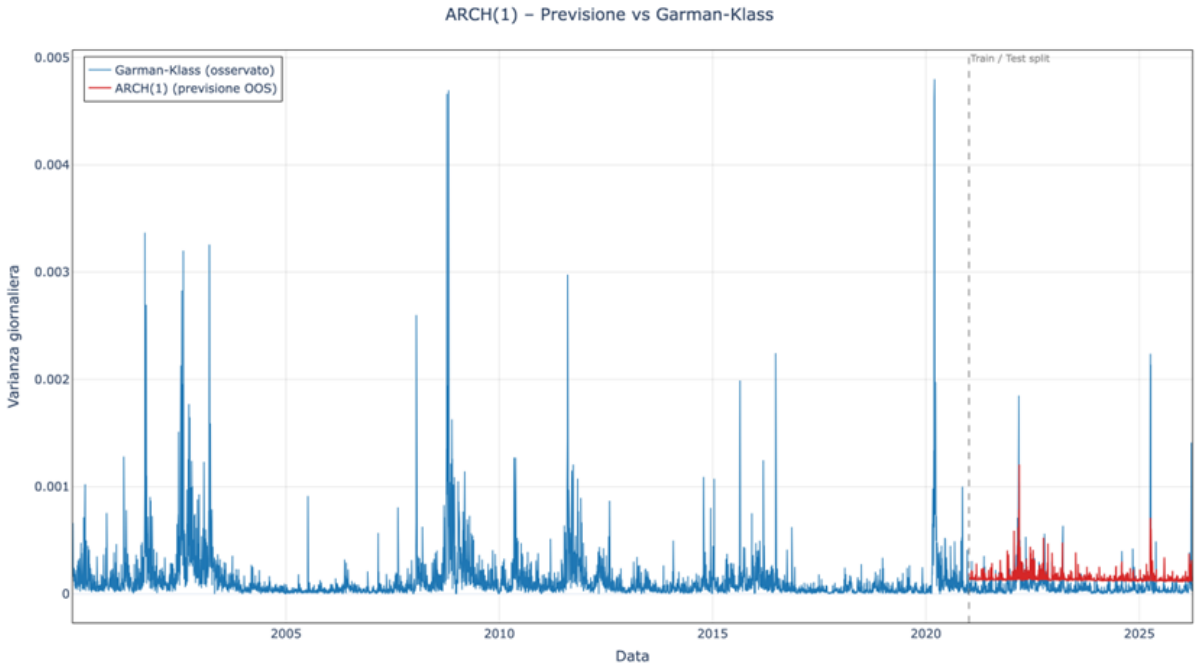


Figure 18: ARCH(1) vs GK, serie temporale completa.

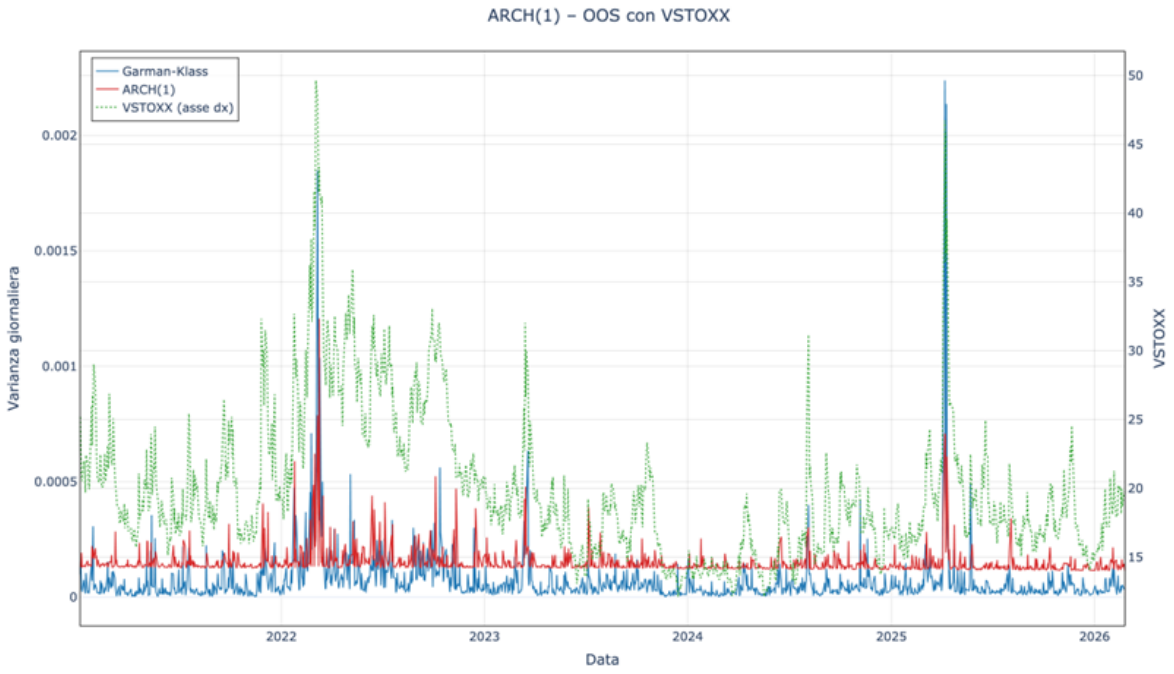


Figure 19: ARCH(1) vs GK, focus out of sample.

Le linea rossa traccia l'andamento della previsione del modello stimato, seguendo le metodologie viste in 3.2.1

In entrambi i grafici, è facile notare che la varianza prevista dal modello, oscilla attorno ad un valore medio più alto rispetto alla volatilità realizzata. In particolare, il modello prevede in media circa 2.5

volte la varianza effettivamente realizzata: la varianza media prevista è 1.81 x10^{-4} , mentre la varianza media GK osservata è 7.22 x 10^{-5} . Allo stesso tempo, quando GK presenta spike acuti (fine 2021, marzo 2022, marzo 2025, marzo 2026), la previsione ARCH non riesce a raggiungerli. Il modello è quindi contemporaneamente troppo alto in media e troppo basso nei picchi.

Possiamo vedere anche come il modello ha cercato di prevedere la volatilità dell'indice durante l'escalation del conflitto tra US e Iran. In questo caso, il modello non riesce a catturare i picchi rilevanti, in particolare quello del 23 marzo. Si nota come gli sbalzi catturati siano inoltre laggati, oltre a sovrastimare la volatilità in periodi più calmi.

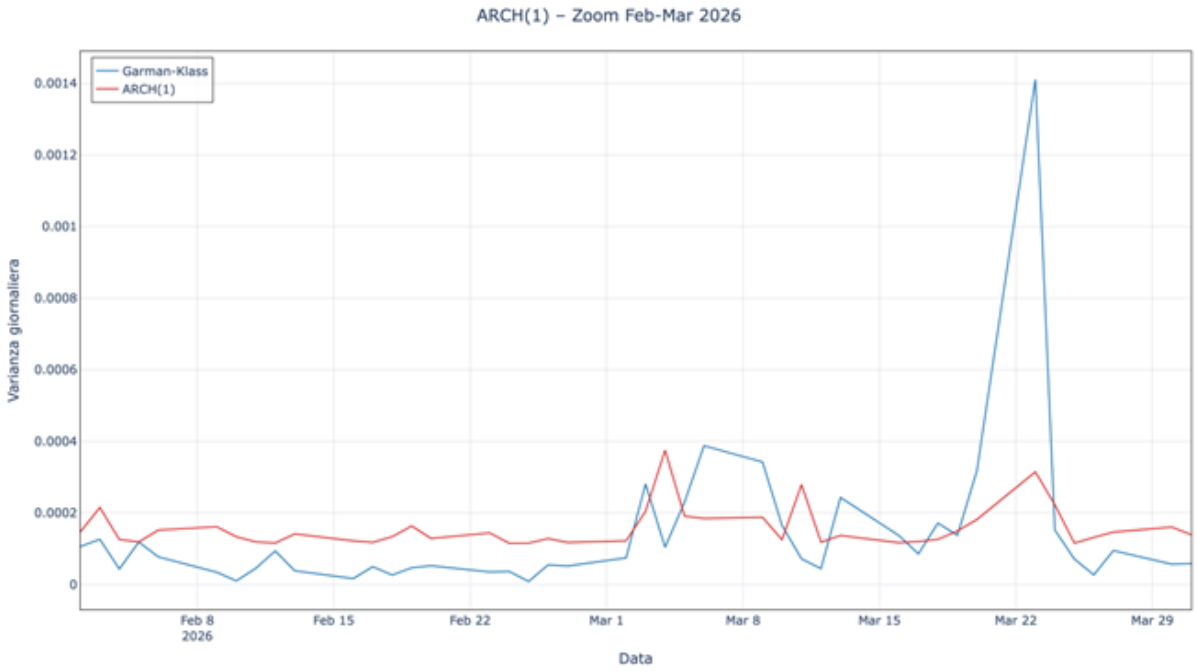


Figure 20: ARCH(1) zoom su febbraio e marzo 2026.

3.5.4 Analisi dei residui

Per verificare se il modello ARCH(1) cattura adeguatamente la dinamica della volatilità, si esaminano i residui standardizzati al quadrato $\hat{z}_t^2 = \frac{\hat{\epsilon}_t^2}{\hat{\sigma}_t^2}$ dovrebbero comportarsi come un processo i.i.d., con autocorrelazioni non significativamente diverse da zero. Il correlogramma ACF di questi residui rivela, tuttavia, un quadro nettamente diverso: tutte le autocorrelazioni fino al lag 30 risultano significativamente positive e ben al di sopra della banda di confidenza al 95%. I valori più elevati si registrano ai primi lag, (circa 0,16 al lag 2 e 0,15 al lag 3, con un picco di circa 0,185 al lag 8). L'autocorrelazione non si esaurisce ai lag brevi: anche oltre il lag 20 i coefficienti restano nell'ordine di 0,05–0,10, ben al di sopra della soglia critica. Questo decadimento lento e persistente è il segno distintivo di una struttura di dipendenza a lungo raggio nella varianza condizionata, long memory della volatilità, che un singolo termine autoregressivo non è in grado di assorbire.

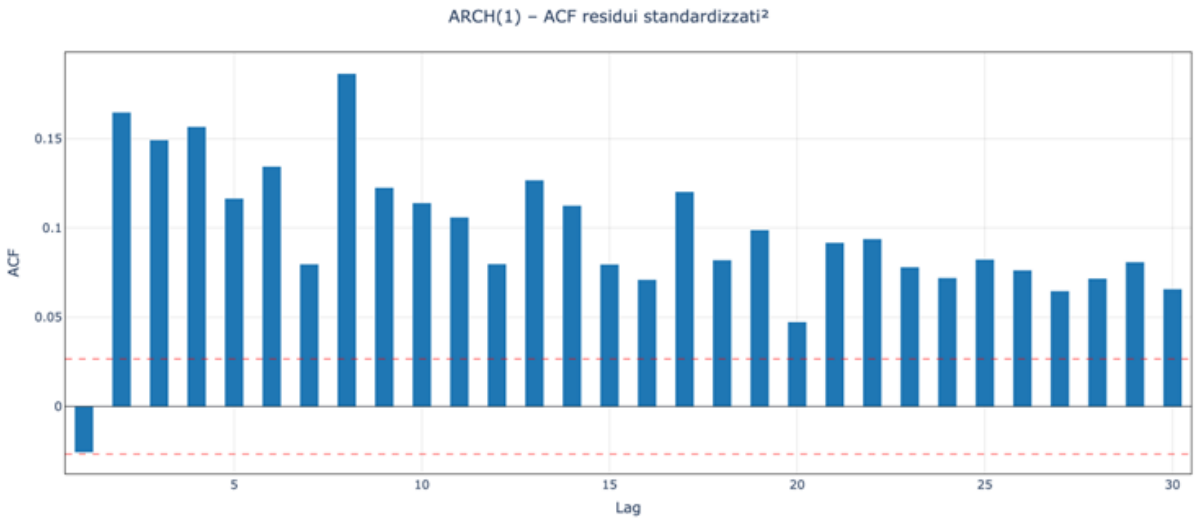


Figure 21: ACF residui standardizzati ² - ARCH(1).

I test statistici confermano questa evidenza visiva. Il test di Ljung-Box, che verifica formalmente se le autocorrelazioni sono congiuntamente nulle, restituisce p-value prossimi a zero a tutti i lag esaminati. Inoltre, il test ARCH-LM verifica se nei residui di un modello rimane ancora eteroschedasticità condizionata (cioè volatilità che varia nel tempo in modo prevedibile), ci si chiede quindi se i residui al quadrato sono ancora spiegabili da un modello ARCH. La risposta è positiva, il p value conferma che nei residui permangono effetti ARCH non modellati.

TEST	LAG	Q_ STATISTIC	P_VALUE
Ljung-Box	5	475.28	1.723e-100
Ljung-Box	10	945.553	9.946e-197
Ljung-Box	15	1229.86	6.270e-253
Ljung-Box	20	1436.54	1.627e-292
ARCH-LM	10	594.76	2.331e-121

Table 8: Valori dei test statistici.

3.6 Il modello GARCH(p,q))

Per superare le limitazioni dei modelli ARCH, Tim Bollerslev propone nel 1986[2] una generalizzazione nota come GARCH (Generalized ARCH). Nel modello GARCH(p,q), la varianza condizionale dipende sia dalle innovazioni passate sia dalle varianze passate:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

In questa formulazione:

- i termini ϵ_{t-i}^2 rappresentano la componente ARCH,
- i termini σ_{t-j}^2 r appresentano la componente GARCH.

Il modello che mostreremo nella pratica è il GARCH(1,1):

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Questa specificazione, pur essendo estremamente parsimoniosa, è in grado di catturare efficacemente la forte persistenza della volatilità osservata nei dati finanziari. Un risultato importante è che il GARCH(1,1) può essere interpretato come un ARCH infinito:

$$\sigma_t^2 = \omega(1 - \beta_1) + \alpha_1 \sum_{i=1}^{\infty} \beta_1^{(i-1)} \epsilon_{t-i}^2$$

Questo implica che la volatilità dipende da tutte le innovazioni passate, con pesi che decrescono esponenzialmente, analogamente a quanto osservato nel modello EWMA. Dal punto di vista dinamico, il modello GARCH(1,1) genera una rappresentazione ARMA per i quadrati degli errori:

$$\epsilon_t^2 = \omega + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 - \beta_1 \nu_{t-1} + \nu_t$$

Un ulteriore elemento di rilievo riguarda la previsione della volatilità. Indicando con:

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

la varianza non condizionale, la previsione a h passi della varianza condizionale è:

$$\hat{\sigma}_{t+h|t}^2 = \sigma^2 + (\alpha_1 + \beta_1)^{h-1} \left(\sigma_{t+1}^2 - \sigma^2\right)$$

Questo risultato mostra che la volatilità prevista converge esponenzialmente verso la varianza di lungo periodo, con una velocità determinata dal parametro $\alpha_1 + \beta_1$. Quando tale somma è prossima a uno, la volatilità risulta altamente persistente, in linea con l'evidenza empirica dei mercati finanziari.

3.6.1 GARCH(1,1) sui rendimenti dell'Eurotstoxx 50

Seguendo l'impostazione precedente, miriamo a superare i limiti del modello ARCH(1). In particolare la sua incapacità di catturare adeguatamente la persistenza della volatilità, migliorando la previsione complessiva. Stimiamo quindi un modello GARCH(1,1) sul medesimo dataset dell'indice Eurostoxx 50, modellando la varianza condizionata come segue:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

La varianza condizionata dipende sia dallo shock che dalla varianza osservata in t-1.

3.6.2 Stima dei parametri

I risultati dei coefficienti del modello sono riportati nella seguente tabella:

PARAMETRO	STIMA	STD ERROR	T_STAT	P_VALUE	CI 95%
μ	0.031846	0.013335	2.388167	0.016933	[5.709e-03, 0.0579]
ω	0.022306	5.50e-03	4.055545	5,001E-02	[0.0115, 0.0331]
α_1	0.105140	0.015685	6.703378	2,036E-08	[0.0744, 0.1359]
β_1	0.883782	0.015575	56.743261	0.000000	[0.8533, 0.9143]

Table 9: Parametri Stimati (In-Sample).

Tutti i parametri del modello di varianza sono altamente significativi, con p-value prossimi a zero.

- μ la stima del rendimento medio è pari a 0,0318, con un p-value di 0,017 e un intervallo di confidenza che non include lo zero. Risulta debolmente significativo al livello del 5%, a differenza di quanto osservato nell'ARCH(1). Tuttavia, questa significatività è probabilmente un effetto indiretto della specificazione più accurata della varianza condizionata: modellando meglio la volatilità, si riduce il "rumore" attorno alla stima della media, rendendo più facile rilevare un drift che resta comunque economicamente trascurabile.

- ω la componente fissa della varianza è stimata a 0,0223 ed è altamente significativa. Il valore è circa 60 volte più piccolo rispetto all'ω dell'ARCH(1) (1,31).

  La differenza non indica un cambiamento nei dati, ma un diverso ruolo del parametro: nell'ARCH puro, ω deve da solo ancorare la varianza al suo livello medio, perché non esiste un meccanismo di propagazione. Nel GARCH, invece, è β_1 la mantenere il livello di varianza nel tempo, e ω funge solo da "pavimento". Il valore così basso indica che, in assenza completa di shock e con varianza passata nulla, il contributo "fisso" alla volatilità giornaliera sarebbe minimo

- α_1 il coefficiente indica che il 10,5% dello shock al quadrato del giorno precedente si trasferisce nella varianza del giorno successivo. l valore è nettamente inferiore ad $\alpha_1 = 0{,}303$ dell'ARCH(1): questo non significa che il GARCH reagisca meno agli shock, ma che il compito di catturare la dinamica della volatilità è ora ripartito tra due canali α_1 per lo shock recente e β_1 per la memoria della varianza.

- β_1 è il parametro che distingue il GARCH dall'ARCH, stimato a 0,8838 con la statistica t più alta tra tutti i parametri (56,7). Indica che l'88,4% della varianza condizionata di ieri si trasmette nella varianza di oggi. In pratica: se ieri il modello stimava una volatilità elevata, oggi ne conserva quasi il 90%, a cui si aggiunge l'eventuale nuovo shock. Questo meccanismo genera esattamente la persistenza che mancava all'ARCH(1).

- $\alpha_1 + \beta_1 = 0{,}989$, molto vicina a 1, segnala una persistenza quasi integrata degli shock: un evento estremo come una crisi produce effetti sulla volatilità che si attenuano molto lentamente nel tempo. La varianza incondizionata $\approx 2,03$, resta finita e coerente con l'ordine di grandezza osservato nei dati.

3.6.3 GARCH(1,1) vs Garman-Klass: previsione out of sample

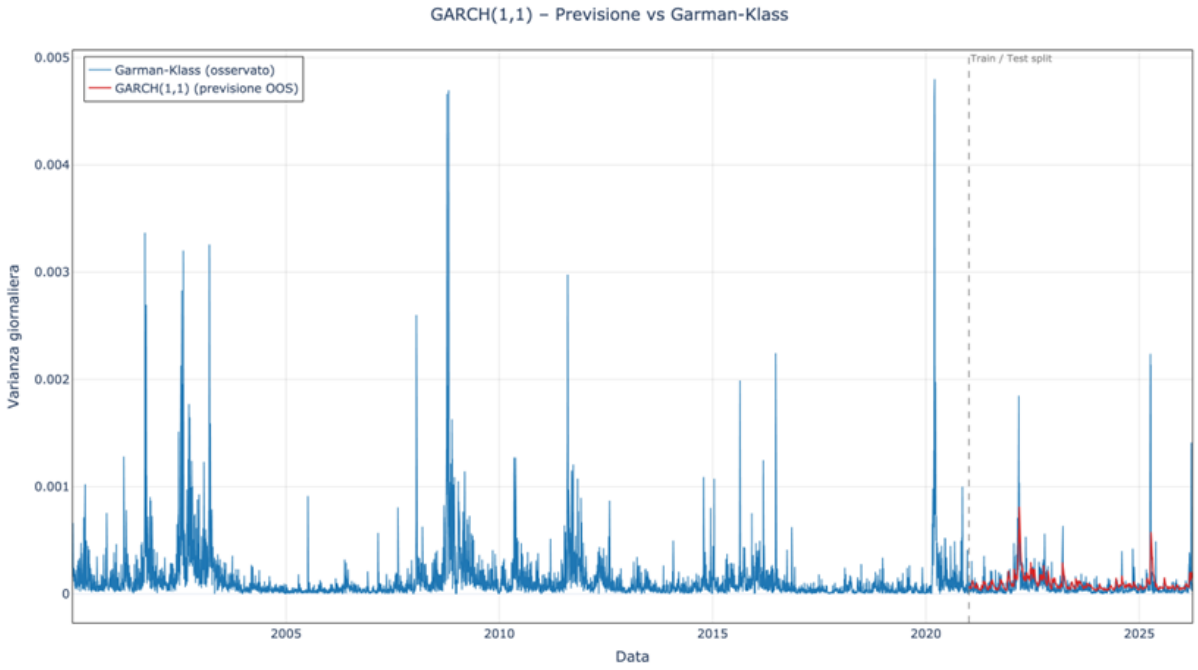


Figure 22: GARCH(1,1) vs GK, serie temporale completa.

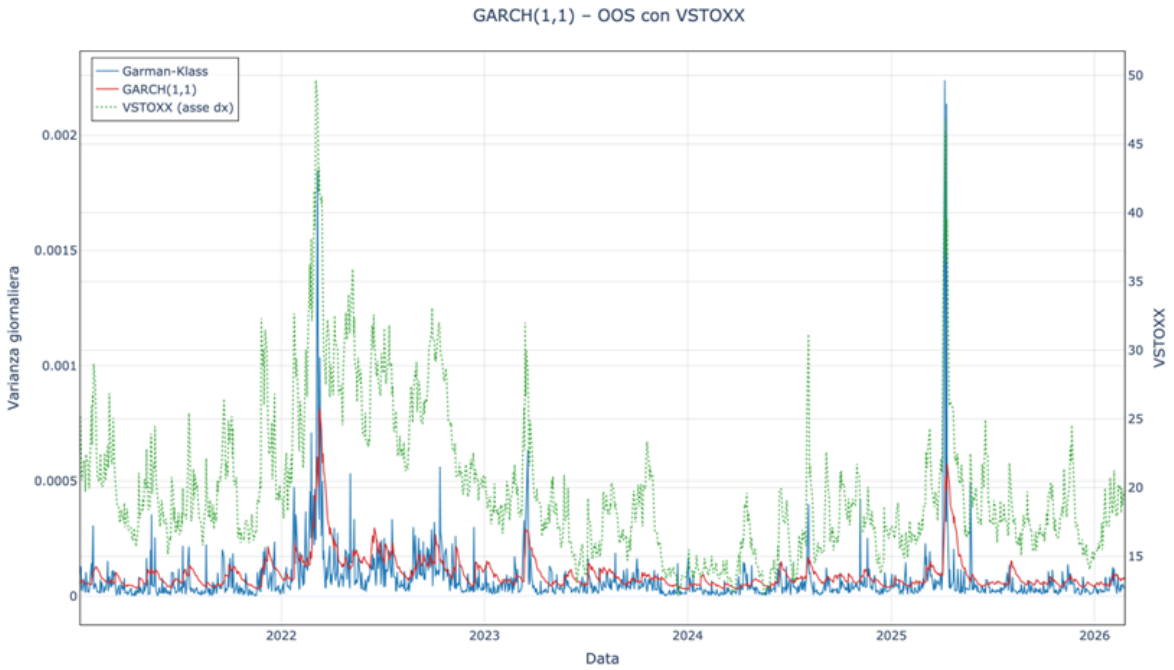


Figure 23: GARCH(1,1) vs GK, focus out of sample.

Entrambi i grafici mostrano un miglioramento netto nella previsione out of sample rispetto al modello ARCH(1) visto prima. Mentre la previsione dell'ARCH produceva un profilo "a gradini" che oscillava attorno a un livello quasi costante e reagiva solo allo shock del giorno precedente, la curva del GARCH(1,1)

segue in modo continuo e regolare l'andamento di fondo della volatilità realizzata. ll modello mantiene la previsione su livelli contenuti, rispondendo con sensibilità ai piccoli movimenti della varianza passata grazie al canale di propagazione β_1. Tuttavia, la curva prevista resta sistematicamente più "liscia" della serie osservata: il GARCH(1,1) cattura il livello medio della volatilità ma non riesce a riprodurre i picchi giornalieri estremi del Garman-Klass.

Questo si può notare anche nello zoom fatto nel periodo di febbraio-marzo, dove il GARCH performa relativamente peggio rispetto all'ARCH. Questo risultato, apparentemente paradossale, si spiega con il fatto che la previsione liscia del GARCH "media" gli shock nel tempo: nei giorni di picchi estremi, il modello non può reagire istantaneamente come l'ARCH, che, pur essendo peggiore in media, ha un α_1 più alto (0,303 vs 0,105) e quindi può occasionalmente avvicinarsi a un singolo spike.

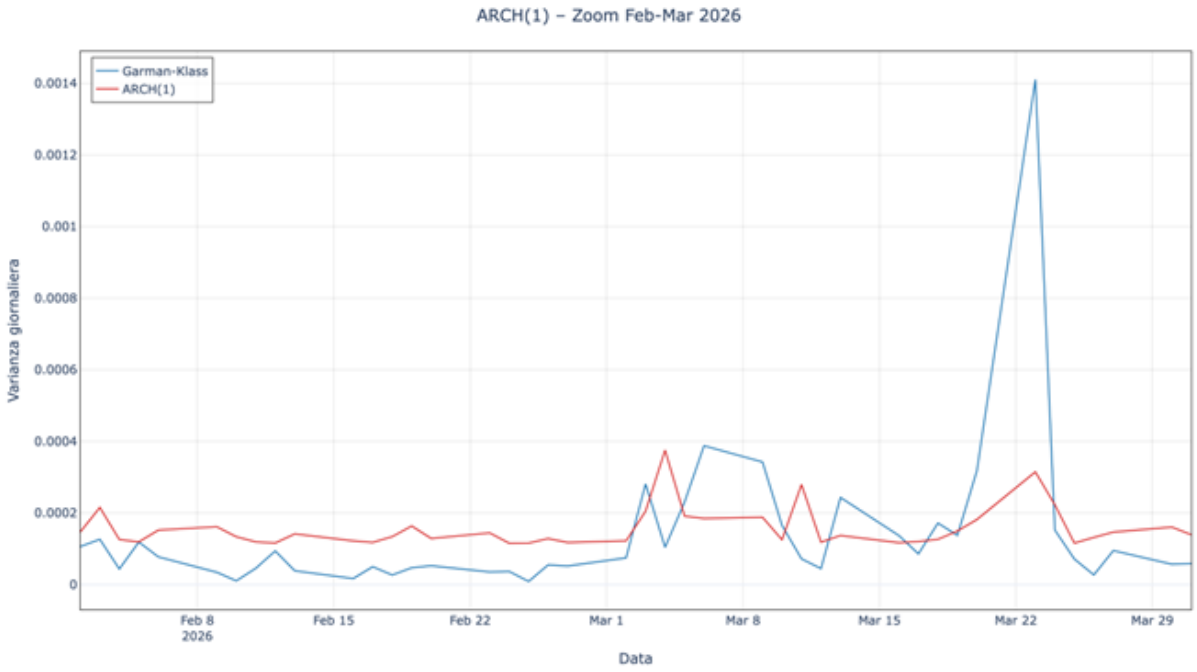


Figure 24: GARCH(1,1) zoom su febbraio marzo 2026.

Perché succede? La ragione è proprio nella struttura del GARCH. Infatti, per calcolare la previsione del 23 marzo, il modello prende l'88,4% della varianza stimata il 22 marzo (che era relativamente bassa) e ci aggiunge il 10,5% dello shock del giorno prima. Anche se lo shock recente è stato grande, il suo peso è solo un decimo del totale: il restante 88% ancora la previsione al passato recente, che era ancora relativamente tranquillo. Si può quindi concludere che la stessa proprietà che rende il modello così efficace nel tracciare le tendenze di fondo, ovvero la "long memomory", diventa un freno quando la realtà cambia da un giorno all'altro.

Bisogna però precisare che l'ARCH non prevede realmente la crisi: semplicemente, sovrastimando la volatilità sempre, capita di essere un po' meno distante quando la volatilità effettivamente esplode. Se si guarda l'intero periodo, il GARCH commette errori complessivamente più piccoli perché non spreca accuratezza nei giorni normali.

3.6.4 Analisi dei residui

Come spiegato in 3.5.4 i residui standardizzati al quadrato dovrebbero comportarsi come numeri casuali indipendenti. In particolare, elevandoli al quadrato, non dovrebbe rimanere correlazione tra il valore di un giorno e quello dei giorni successivi. L'assenza di autocorrelazione nei vari lag dimostrerebbe infatti che il modello ha spiegato tutta la struttura prevedibile della volatilità, lasciando nei residui puro rumore casuale.

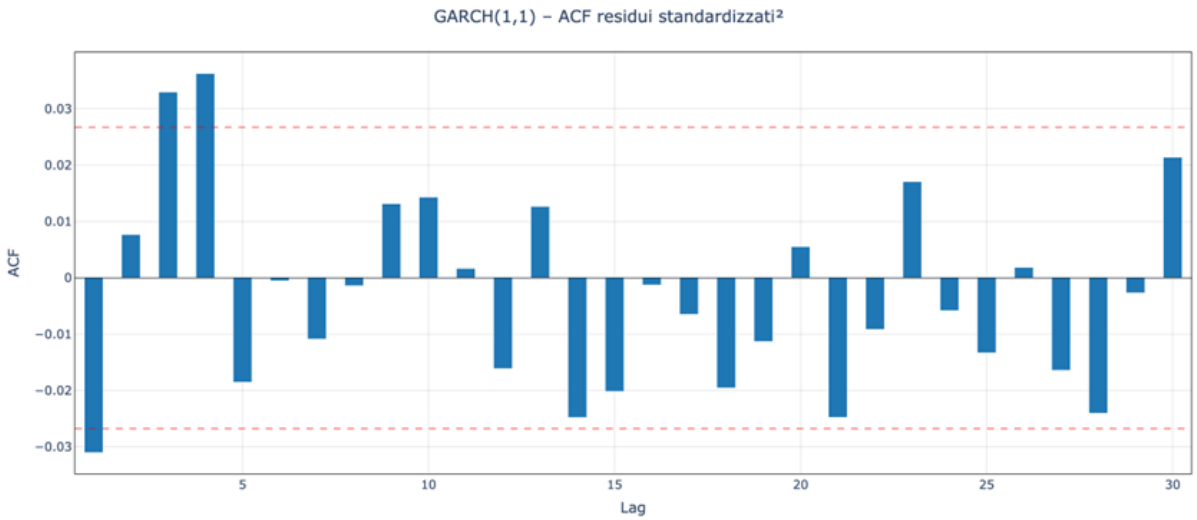


Figure 25: ACF residui standardizzati^2 – GARCH(1,1).

Si nota una differenza netta rispetto all'ACF del modello ARCH(1), le barre oscillano tra –0,03 e +0,036, cioè circa un ordine di grandezza più piccole e solo i lag 1,3 e 4 superano la barra di confidenza al 95%. Inoltre, a differenza dell'ARCH dove tutte le autocorrelazioni erano positive (segno di una dipendenza sistematica non catturata), qui le barre alternano segno in modo irregolare, come ci si aspetterebbe da un segnale in larga parte "ripulito". Il confronto visivo tra i due grafici è forse la prova più immediata del contributo di β_1: il termine di memoria assorbe la quasi totalità della persistenza che l'ARCH lasciava nei residui. Si passa da autocorrelazioni dell'ordine del 10–18% a valori dell'1–3%, una riduzione di oltre il 90%.

TEST	LAG	Q_STATISTIC	P_VALUE
Ljung-Box (squared std resid)	5	20.25	0.00111
Ljung-Box (squared std resid)	10	22.93	0.01097
Ljung-Box (squared std resid)	15	30.70	0.00960
Ljung-Box (squared std resid)	20	33.84	0.02715
ARCH-LM (nlags=10)	10	24.327	0.00677

Table 10: Valori dei test statistici.

Il test di Ljung-Box verifica formalmente la presenza di autocorrelazione tra i vari lag poichè tutti i risultati mostrano p value significativi, rifiutando formalmente l'assenza di autocorrelazione. Consideriamo però, che tale significatività risulta molto più debole rispetto a quella vista prima con il modello ARCH(1). Inoltre, per il GARCH(1,1), la statistica ARCH-LM con 10 lag è 24,3 (p = 0,007), anch'essa

significativa al 5%. Per l'ARCH(1) la stessa statistica era 594,8 ($p \approx 10^{-121}$) la riduzione è del 96%, dimostrando un notevole aumento di eteroschedasticità condizionata spiegata.

3.7 Il modello EGARCH(p,o,q)

Nonostante la capacità dei modelli GARCH di catturare la persistenza della volatilità, essi presentano un limite rilevante: la risposta della volatilità agli shock è simmetrica. In particolare, nel modello GARCH standard la dinamica della varianza dipende dai quadrati delle innovazioni passate, implicando che shock positivi e negativi di uguale intensità abbiano lo stesso impatto sulla volatilità futura. Tale assunzione risulta in contrasto con il fattore stilizzato noto come "leverage effect" introdotto precedentemente in cui si osserva che le variazioni negative dei prezzi tendono a generare un aumento della volatilità maggiore rispetto a variazioni positive di pari ampiezza. Per superare questa limitazione, Daniel B. Nelson (1991) propose il modello EGARCH (Exponential GARCH), che introduce una dinamica asimmetrica nella volatilità condizionale. La principale innovazione del modello EGARCH consiste nel modellare il logaritmo della varianza condizionale:

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \alpha \left(\frac{|\epsilon_{t-1}|}{\sigma_{t-1}} \right) + \gamma \left(\frac{\epsilon_{t-1}}{\sigma_{t-1}} \right)$$

Questa formulazione presenta diversi vantaggi. In primo luogo, il logaritmo garantisce automaticamente la positività della varianza, eliminando la necessità di imporre vincoli sui parametri. In secondo luogo, consente di distinguere tra l'effetto della dimensione dello shock e il suo segno. In particolare, il parametro γ cattura l'asimmetria della risposta della volatilità: un valore negativo di γ implica che gli shock negativi abbiano un impatto maggiore rispetto a quelli positivi, permettendo al modello di replicare il leverage effect osservato empiricamente. Pertanto, il modello EGARCH rappresenta un'estensione naturale del GARCH, in grado di descrivere in modo più realistico la dinamica della volatilità nei mercati finanziari.

3.7.1 EGARCH(1,1,1) sui rendimenti dell'Eurostoxx 50

Procediamo ora a stimare il modello EGARCH(1,1,1) usando il nostro dataset e la stessa impostazione con split iniziale 80/20 e expanding window. L'equazione della media resta la stessa dei modelli precedenti:

$$r_t = \mu + \varepsilon_t, \varepsilon_t = \sigma_t z_t, z_t \sim N(0,1)$$

L'equazione della varianza diventa invece:

$$\ln(\sigma_t^2) = \omega + \alpha_1 |z_{t-1}| + \gamma_1 z_{t-1} + \beta_1 \ln(\sigma_{t-1}^2)$$

Dove $\hat{z}_{t-1} = \hat{\epsilon}_{t-1}/\hat{\sigma}_{t-1}$ è il residuo standardizzato. La notazione (1,1,1) indica rispettivamente: un termine ARCH (α_1), un termine asimmetrico (γ_1) e un termine GARCH (β_1).

Qual è il meccanismo che permette al modello di riprodurre il leverage effect?

La variabile chiave è $\hat{z}_{t-1}$ in quanto permette di mostrare in modo efficiente il leverage effect della volatilità. Infatti, il termine $\alpha_1|\hat{z}_{t-1}|$ cattura l'effetto della dimensione dello shock indipendentemnete

dal segno, mentre il termine $\gamma_1\ \hat{z}_{t-1}$ cattura l'effetto del segno. Quando lo shock è negativo ($\hat{z}_{t-1} < 0$) i due termini si sommano in modulo se $\gamma_1 < 0$, amplificando l'impatto sulla volatilità, mentre quando è positivo si compensano parzialmente.

3.7.2 Stima dei parametri ed interpretazione

PARAMETRO	STIMA	STD ERROR	T_STAT	P_VALUE	CI 95%
μ	-0.016028	0.01234	-1.2994	0.1938	[-0.0402, 8.1497e-03]
ω	6.348e-03	2.9157e-03	2.1772	0.0294	[6.333e-04, 0.0121]
α_1	0.1167	0.01576	7.4018	1.3432e-13	[0.0858, 0.14758]
γ_1	-0.1360	0.01225	-11.1045	1.1922e-28	[-0.16002, -0.1120]
$beta_1$	0.97948	3.4857e-03	281.0003	0.0000	[0.9727, 0.9863]

Table 11: Parametri Stimati EGARCH(1,1,1) (In-Sample).

- μ= –0,0160 (p = 0,194). La stima del rendimento medio è negativa e non significativa: l'intervallo di confidenza al 95% include lo zero [–0,040 ; 0,008]. A differenza del GARCH(1,1), dove μ risultava debolmente significativo (0,0318, p = 0,017), qui il drift scompare. Il cambiamento non è casuale: l'EGARCH, modellando esplicitamente l'asimmetria, ridistribuisce parte della struttura che il GARCH assorbiva implicitamente nella media. Il risultato è coerente con l'attesa teorica che il rendimento medio giornaliero sia indistinguibile da zero su orizzonti brevi.

- ω = 0,00635 (p = 0,029). La componente fissa (volatilità che esiste "a prescindere") del logaritmo della varianza è positiva e significativa al 5%. Il valore è circa 3,5 volte più piccolo dell'ω del GARCH (0,0223), coerentemente con il fatto che qui ω opera sulla scala logaritmica della varianza, non sulla varianza stessa. Come nel GARCH, il suo ruolo pratico è quello di "pavimento": determina il livello base verso cui il logaritmo della varianza tende a convergere in periodi di piena calma dei mercati.

- α_1 = 0,1167 ($p \approx 10^{-13}$). Il coefficiente che misura l'impatto della dimensione dello shock è altamente significativo. Intuitivamente, se ieri si è verificato uno shock grande (indipendentemente dal segno), la volatilità stimata in t sale. Il valore è vicino all'α_1 del GARCH (0,105), il che conferma che la reattività complessiva agli shock è comparabile tra i due modelli. La differenza è che nel GARCH α_1 cattura contemporaneamente dimensione e segno dello shock (confondendoli), mentre qui il compito è diviso tra α_1 (dimensione) e γ_1 (segno).

- γ_1 = –0,1360 ($p \approx 10^{-28}$). Il parametro di asimmetria rappresenta la novità di questo modello ed è altamente significativo, a conferma che l'effetto leva è una caratteristica forte e inequivocabile dei dati dell'Euro Stoxx 50. Il segno negativo ha un significato economico preciso: gli shock negativi (ribassi) aumentano la volatilità più degli shock positivi (rialzi) di pari entità.

- β_1= 0,9795 ($p \approx 0$). Il coefficiente che misura la memoria della volatilità, già introdotto con il GARCH(1,1), risulta molto significativo. Infatti, lapersistenza della log-varianza è la più alta tra i tre modelli (0,884 nel GARCH, 0,980 nell'EGARCH). La statistica t di 281 è anch'essa la più elevata. Il valore alto non significa che la volatilità dimostri maggiore persistenza in senso assoluto rispetto al GARCH: il valore più alto di β_1 compensa l'effetto della formulazione logaritmica e

dell'asimmetria, mantenendo l'half-life degli shock su un ordine di grandezza confrontabile. Ciò che conta è che il meccanismo di memoria resta dominante: circa il 98% della log-varianza di ieri si trasmette nella log-varianza di oggi.

3.7.3 Un esempio per capire meglio l'effetto leverage sui dati dell' Eurostoxx 50

Consideriamo uno shock di una deviazione standard, ovvero il caso in cui $\hat{z}_{t-1} = -1$.

Come visto in precedenza, uno shock è semplicemente la parte inattesa del rendimento giornaliero, cioè lo scostamento del rendimento effettivo dalla sua media:

$$\epsilon_t = r_t - \mu$$

ed il residuo standardizzato

$$z_t = \frac{\epsilon_t}{\sigma_t}$$

esprime quello scostamento in unità di volatilità.

Dato che $\mu \approx 0$, uno shock di $\hat{z}_{t-1} = -1$ corrisponde a un rendimento giornaliero di circa $-0{,}85\%$, cioè un calo dell'indice di poco meno dell'1% in un giorno. Non è un evento raro, infatti è successo circa 1 giorno su 6.

- **Shock negativo** ($\hat{z}_{t-1} = -1$, un ribasso): l'impatto sulla log-varianza è

$$\alpha_1 \cdot 1 + \gamma_1 \cdot (-1) = 0{,}117 + 0{,}136 = 0{,}253.$$

  La varianza viene quindi moltiplicata per

$$e^{0{,}253} \approx 1{,}288.$$

  In altre parole, la varianza aumenta del 29% e la volatilità (che è la radice quadrata della varianza) sale di circa il 13,5%. Se il giorno prima il modello stimava una volatilità giornaliera dell'1%, dopo un ribasso di questa entità la stima passa a circa l'1,14%.

- **Shock positivo** ($\hat{z}_{t-1} = +1$, rialzo della stessa grandezza): l'impatto sulla log-varianza è

$$\alpha_1 \cdot 1 + \gamma_1 \cdot (+1) = 0{,}117 - 0{,}136 = -0{,}019.$$

  La varianza viene quindi moltiplicata per

$$e^{-0{,}019} \approx 0{,}981.$$

  La varianza cala del 2% e la volatilità scende di circa l'1%. Se il giorno prima la volatilità stimata era l'1%, dopo un rialzo di una deviazione standard scende a circa lo 0,99%, una variazione quasi impercettibile.

3.7.4 EGARCH vs Garma-Klass: previsione out-of-sample

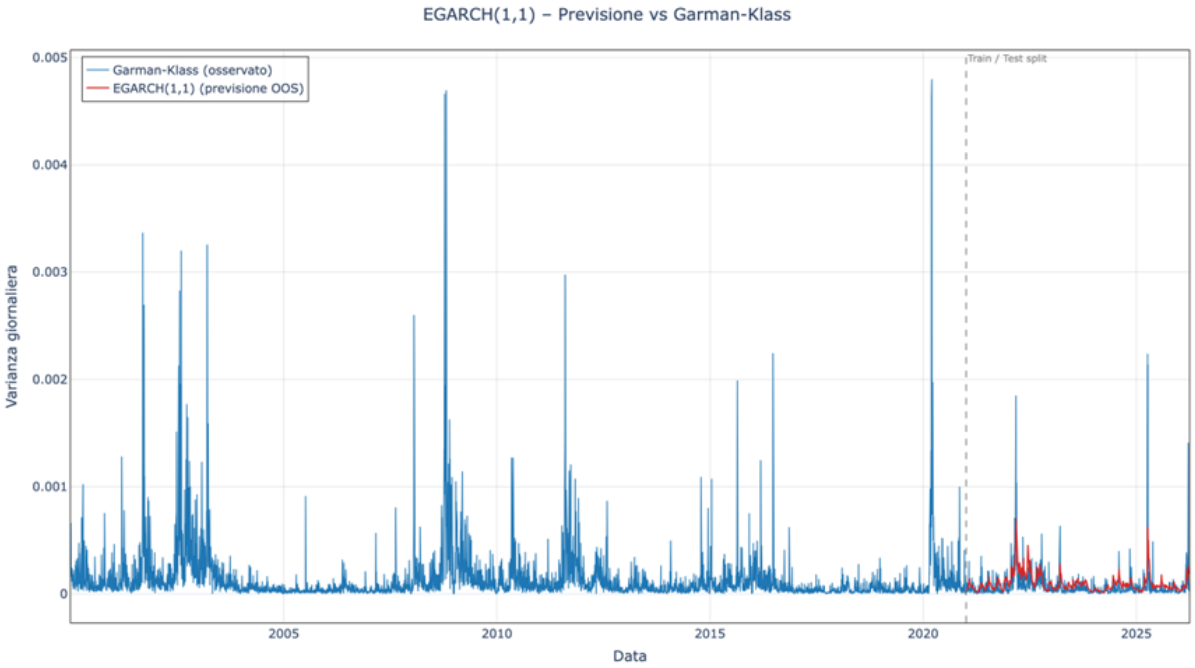


Figure 26: EGARCH(1,1,1) vs GK, serie temporale completa.

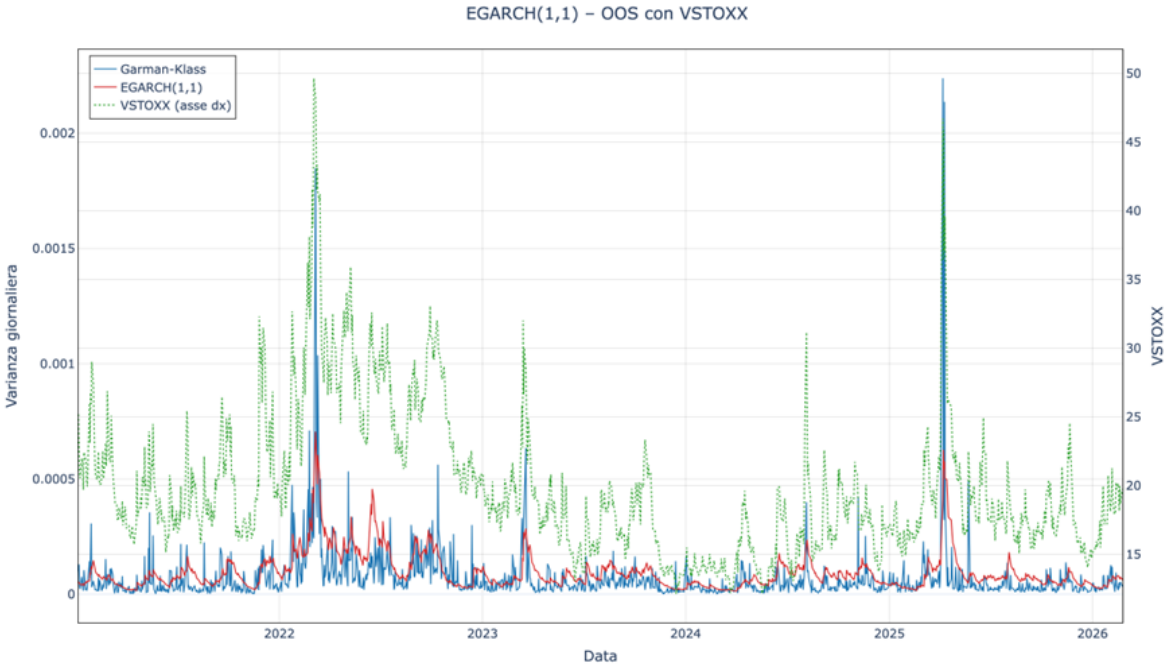


Figure 27: EGARCH(1,1,1) vs GK, focus out of sample.

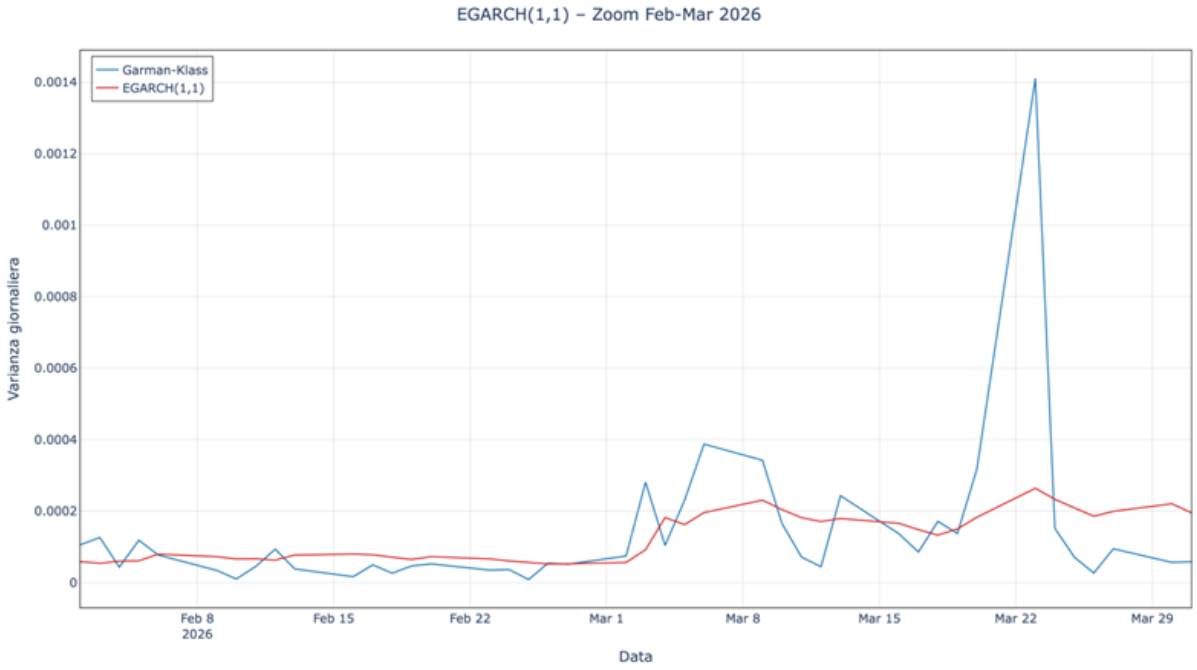


Figure 28: GARCH(1,1,1) zoom su febbraio marzo 2026.

Tutti e tre i grafici non mostrano miglioramenti particolarmente visibili rispetto al GARCH(1,1). Si tratta infatti di un modello ancora molto simile al precedente ed i miglioramenti sono visibili esclusivamente dalle metriche di valutazione mostrate nel confronto finale.

3.7.5 Analisi dei residui

Il grafico ACF dei risidui standardizzati al quadrato mostra risultati simili con quello del modello GARCH(1,1).

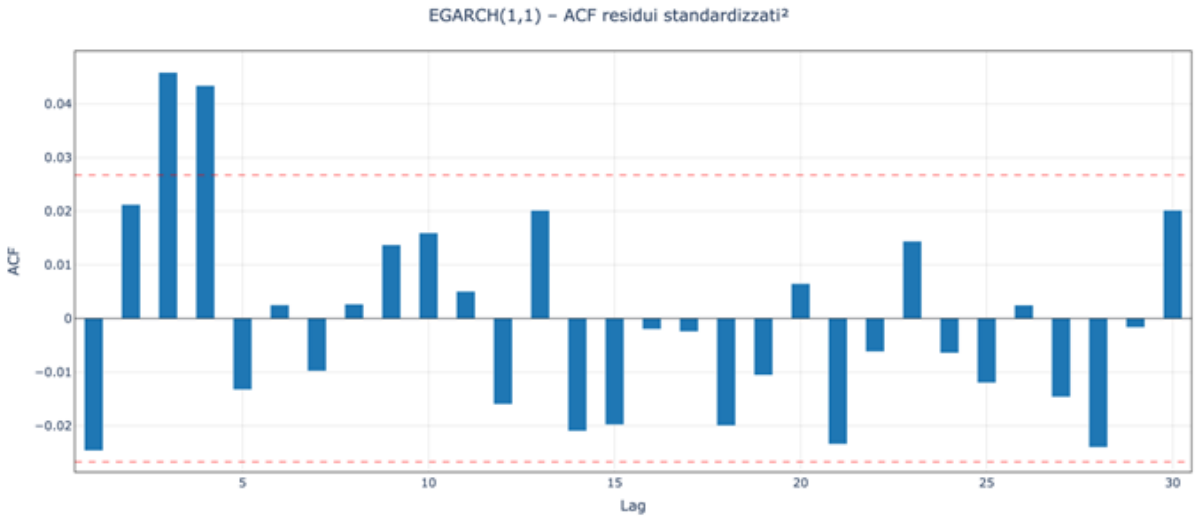


Figure 29: ACF residui standardizzati^2- EGARCH(1,1,1)

La vera conferma è data dal test Ljung-Box:

TEST	LAG	Q_STATISTICS	P_VALUE
Ljung-Box	5	12.389	0.02981
Ljung-Box	10	15.986	0.10002
Ljung-Box	15	22.251	0.10137
Ljung-Box	20	24.367	0.2267
ARCH-LM	10	16.726	0.08063

Table 12: Valori dei test statistici.

I p value calcolati sono aumentati notevolmente rispetto al GARCH(1,1). Infatti, accettiamo l'ipotesi nulla di assenza di autocorrelazione per i lag successivi a 5, mentre quest'ultimo risulta significativo solo al 5%. L'eteroschedasticità condizionata spiegata è quindi aumentata, si può vedere anche dal test ARCH-LM.

3.8 Confronto e valutazione conclusiva dei modelli

3.8.1 Serie temporali a confronto

A questo punto, confrontiamo i tre modelli di previsione: sia dal punto di vista grafico che valutando le metriche di errore e le correlazioni.

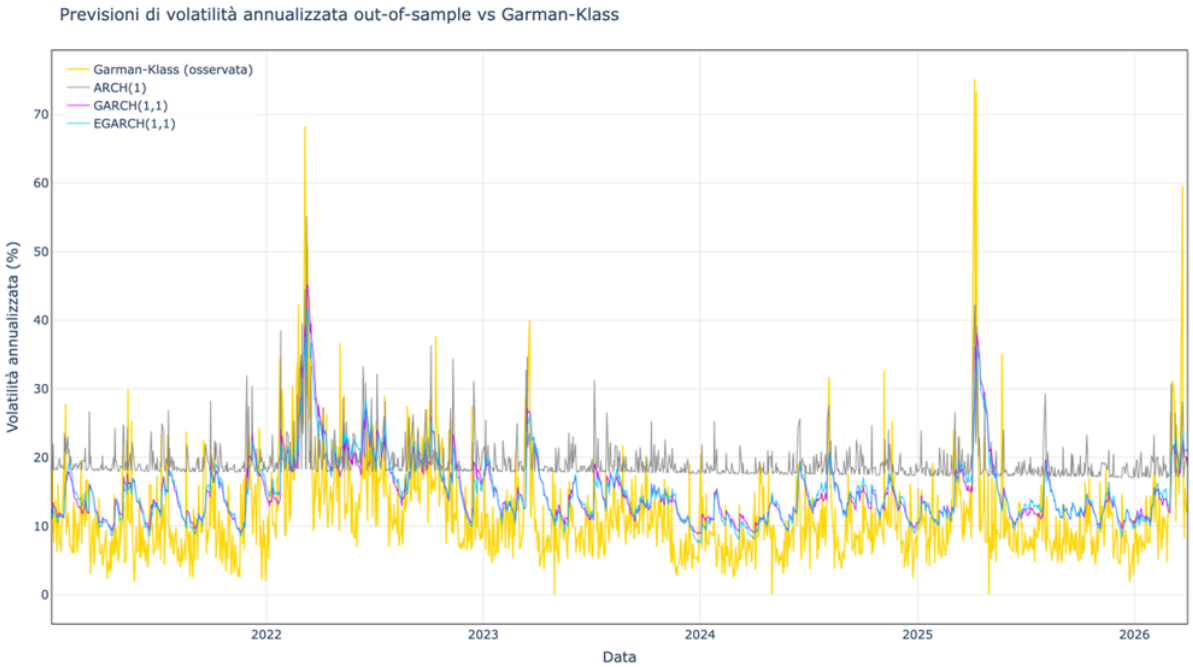


Figure 30: Previsioni di volatilità annualizzata out of sample vs Garman-Klass.

Lo stimatore usato come proxy rappresenta la volatilità effettivamente verificata, i tre modelli cercano di prevederla giorno per giorno attraverso i dati storici e progressivamente integrando informazioni sugli shock ogni t+1. La misura di volatilità sull'asse delle ordinate è stata annualizzata per i giorni di trading $\sqrt{252}$, rendendo il confronto visivo più semplice. Complessivamente, il modello ARCH(1,1) è visivamente

il peggiore, sovrastimando la volatilità nella maggior parte delle previsioni; i modelli GARCH(1,1) e EGARCH(1,1,1) come spiegato in precedenza, sono in realtà molto simili.

Per mostrare le differenze nelle previsioni in modo più preciso, vediamo l'errore assoluto di ciascuna in un grafico a serie temporale:

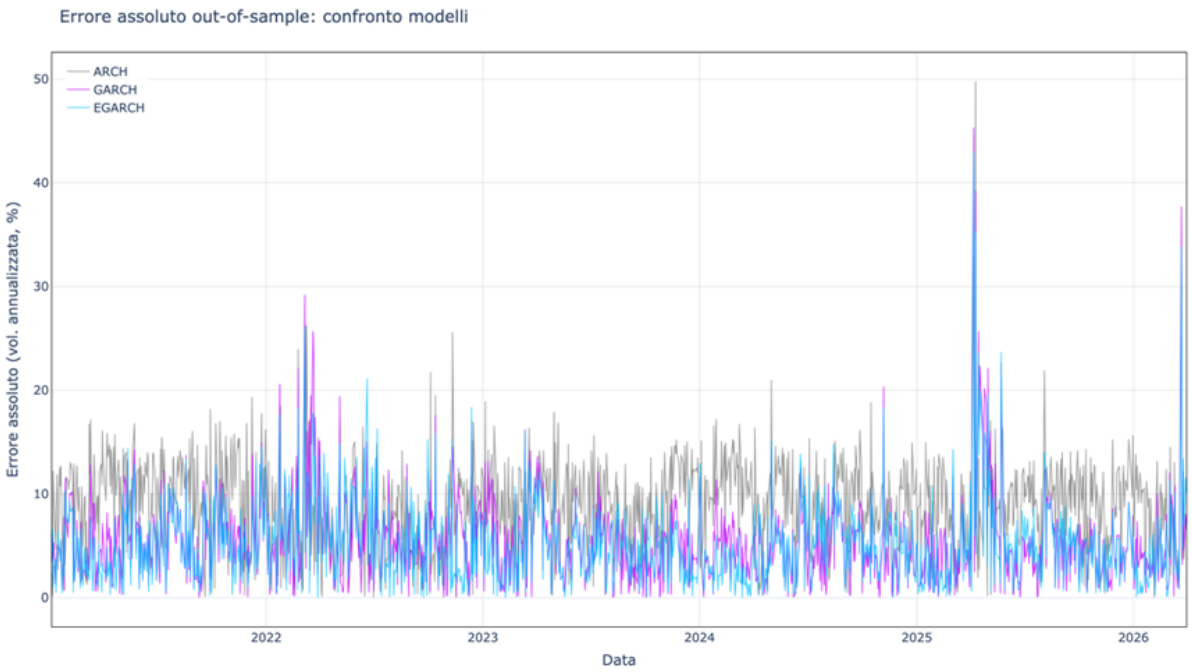


Figure 31: Errore assoluto out of sample: confronto modelli.

Tutti i modelli sbagliano di più nei periodi a volatilità più alta. Questo perché quando la volatilità sale, aumentano insieme tre cose: l'ampiezza dei movimenti, la rapidità dei cambi di regime e l'incertezza stessa sulla varianza futura. Un modello come GARCH, EGARCH, ARCH o simili prevede la volatilità condizionata usando informazione passata, quindi tende a reagire con un certo ritardo agli shock molto forti. Nei picchi di mercato, invece, la volatilità può saltare bruscamente: il modello la insegue, ma raramente la anticipa perfettamente. Possiamo confrontare i modelli negli ultimi due mesi:

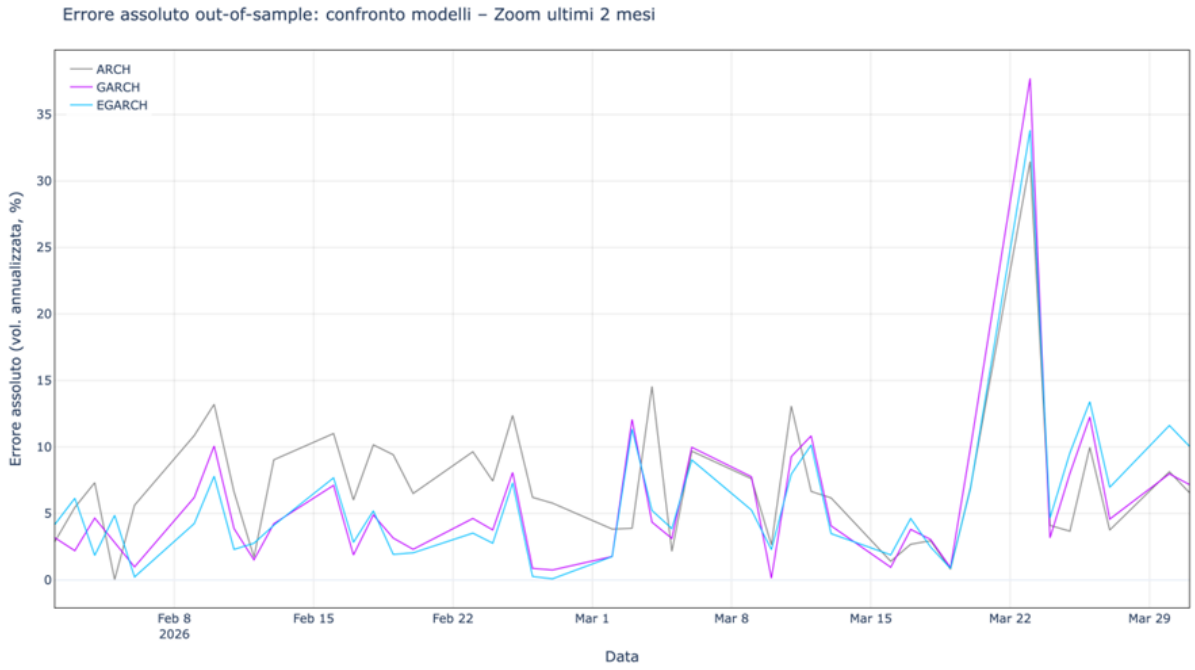


Figure 32: Errore assoluto out of sample: confronto modelli negli ultimi due mesi.

3.8.2 Metriche di errore a confronto

METRICA	ARCH(1)	GARCH(1,1)	EGARCH(1,1,1)
MAE (% GK)	148.2%	86.4%	83.3%
RMSE (% GK)	198.5%	165.1%	156.0%
Mean Error (% GK)	109.5%	36.9%	34.8%
QLIKE	-8.3890	-8.6765	-8.7013
Ratio prevista/GK	2.10x	1.37x	1.35x
Corr(pred, GK)	0.4488	0.5036	0.5716
R^2	0.2015	0.2536	0.3267
MAE high-vol (% GK)	232.8%	239.1%	226.9%
RMSE high-vol (% GK)	432.2%	438.7%	410.4%
Mean Error high-vol (% GK)	-150.3%	-161.5%	-141.0%

Table 13: Metriche di errore.

Il modello EGARCH(1,1,1), come previsto, risulta essere il migliore dalle metriche.

- Accuratezza media: MAE e RMSE indicano che EGARCH è il modello più accurato, seguito da GARCH, mentre ARCH è chiaramente il peggiore. Questo vale sia per l'errore medio sia per gli errori più ampi, quindi EGARCH è quello che si avvicina meglio alla volatilità Garman-Klass nel complesso.

- Bias di previsione: Il Mean Error e il rapporto prevista/GK mostrano che tutti i modelli tendono a sovrastimare la volatilità media, ma con intensità diversa. ARCH sovrastima molto, con una varianza prevista pari a 2.10 volte la GK osservata. GARCH ed EGARCH sovrastimano meno, e risultano molto più equilibrati.

- Capacità di seguire la dinamica: la correlazione con GK e il relativo R^2 mostrano che EGARCH segue meglio l'andamento temporale della volatilità, GARCH è intermedio, ARCH il meno efficace. Quindi EGARCH non solo sbaglia meno nel livello, ma intercetta meglio anche i movimenti della serie.

- Qualità complessiva della loss: il QLIKE penalizza i modelli quando la varianza prevista è lontana da quella osservata, ma soprattutto quando la previsione è mal calibrata rispetto ai picchi. In questo confronto valori più bassi indicano prestazioni migliori, EGARCH è il migliore. Questo significa che, nel complesso, le sue previsioni di varianza sono quelle più ben allineate alla GK osservata secondo una loss specifica per la volatilità.

- Periodi di altà volatilità: nelle giornate più turbolente tutti i modelli peggiorano sensibilmente. È un risultato normale: i picchi di volatilità sono più difficili da prevedere. Anche in questo sottoinsieme EGARCH resta il migliore, ma il vantaggio rispetto a GARCH si riduce.

Conclusione

La volatilità, più di molte altre grandezze finanziarie, permette di guardare il mercato per quello che è: un sistema in cui l'incertezza non è un'eccezione, ma una componente strutturale. Attraverso il percorso svolto in questo lavoro, si è mostrato che dietro l'apparente disordine dei rendimenti dell'Euro Stoxx 50, emergono regolarità precise, che la letteratura ha sintetizzato nei fatti stilizzati e che i modelli cercano di interpretare e organizzare. I risultati ottenuti mettono in luce il collegamento diretto tra i fatti empirici ottenuti dai dati e la logica dei modelli utilizzati per rappresentarli: la persistenza della volatilità, la presenza di asimmetrie e la difficoltà di catturare gli shock estremi rendono infatti evidente perché i modelli più semplici abbiano soprattutto un valore introduttivo, mentre quelli più flessibili offrano una descrizione più aderente alla realtà del mercato. In particolare, il confronto finale suggerisce che EGARCH offra nel complesso la rappresentazione più efficace della dinamica della volatilità nel campione considerato, pur restando evidente un limite comune a tutti i modelli tradizionali: la difficoltà di anticipare con precisione i picchi più estremi, cioè proprio le fasi in cui la volatilità diventa economicamente più rilevante.

Accanto a questi modelli, la letteratura più recente si sta però muovendo in una direzione più ampia, in cui gli strumenti econometrici tradizionali vengono affiancati, e talvolta superati, da tecniche di machine learning e deep learning. In questo ambito rientrano metodi ensemble come XGBoost, reti neurali artificiali, reti ricorrenti e architetture LSTM, che risultano particolarmente adatte quando si vogliono sfruttare grandi quantità di dati, relazioni non lineari complesse o strutture sequenziali difficili da catturare con modelli parametrici standard. La loro applicazione è oggi rilevante soprattutto nella previsione della volatilità realizzata e implicita, nel risk management, nell'asset allocation, nel trading sistematico e, più in generale, in tutti i contesti in cui conta estrarre informazione predittiva da dataset ricchi e ad alta frequenza. Allo stesso tempo, questi approcci introducono nuove criticità, tra cui minore interpretabilità, maggiore sensibilità alla definizione della variabile target e rischio di overfitting, soprattutto quando il segnale è rumoroso. Proprio per questo, i modelli classici analizzati in questo lavoro mantengono un

ruolo essenziale: non rappresentano l'ultima frontiera della ricerca, ma costituiscono ancora oggi la base concettuale necessaria per comprendere come si misura, si modella e si prevede la volatilità.

References

[1] F Black. "' Studies in Stock Price Volatility Changes', Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section". In: (1976).

[2] T Bollerslev. "'Generalized Autoregressive Conditional Heteroskedasticity', Journal of Econometrics, Vol. 31". In: (1986).

[3] Rama Cont. "Empirical properties of asset returns: stylized facts and statistical issues". In: (2001).

[4] Chiara Fabiani. "Strategie di Trading Quantitativo: Previsioni del VIX con tecniche di Machine Learning ed Econometriche". PhD thesis. 2025.

[5] Martin Sewell. "Characterization of Financial Time Series". In: (Jan. 2011).

[6] *STOXX® Strategy Index Guide*. Dec. 2025.