

Università degli Studi di Padova

Dipartimento di Scienze Statistiche

Corso di Laurea Triennale in
Statistica per l'Economia e l'Impresa



RELAZIONE FINALE

Previsione della volatilità di petrolio e gas naturale:
un confronto con modelli di tipo ARCH

Relatore Prof. Luisa Bisaglia
Dipartimento di Scienze Statistiche

Laureando: Gilberto Conte
Matricola 1198846

Anno Accademico 2022/2023

Indice

Introduzione	5
1 Il mercato del petrolio greggio e del gas naturale	7
1.1 Il petrolio greggio e il gas naturale	7
1.2 Cause delle fluttuazioni dei prezzi	8
2 Modelli per la previsione delle volatilità	11
2.1 La volatilità	11
2.2 Il modello ARCH	13
2.3 Il modello GARCH	14
2.4 Il modello EGARCH	16
2.5 Il modello GAS	17
3 Adattamento dei modelli ai dati	19
3.1 Dati di riferimento	19
3.2 Previsioni in-sample	23

4	Previsione della volatilità	29
4.1	Errore quadratico medio (RMSE)	30
4.2	Winning Ratio (WR)	31
4.3	Test di Diebold-Mariano	33
4.4	Somma cumulativa della differenza quadratica degli errori di previsione (CSSFED)	35
5	Conclusioni	39
	Bibliografia	41

Introduzione

Questa relazione finale di laurea si basa sullo studio della previsione della volatilità dei rendimenti per i prezzi di petrolio greggio e gas naturale. Dal punto di vista finanziario, il rendimento si riferisce alla misura dell'aumento o della diminuzione del valore di un investimento nel corso del tempo.

Inizialmente si analizza il mercato del petrolio greggio e del gas naturale descrivendo i fattori che ne influenzano i prezzi. Nel secondo capitolo si descrivono alcuni modelli statistici che possono fornire una soluzione ai problemi relativi alle fluttuazioni di questi rendimenti. In particolare vengono studiate le caratteristiche dei modelli GARCH, EGARCH e GAS. Nel terzo capitolo vengono analizzati i dati che sono stati presi in considerazione per questa indagine. Nel capitolo successivo si effettuano delle previsioni *out-of-sample* per valutare la capacità predittiva dei modelli attraverso quattro metriche distinte. Infine, nell'ultimo capitolo sono state riportate le conclusioni delle analisi svolte.

Capitolo 1

Il mercato del petrolio greggio e del gas naturale

1.1 Il petrolio greggio e il gas naturale

Il petrolio greggio, è una forma non raffinata e non lavorata di petrolio. È un idrocarburo liquido naturale che si trova nel sottosuolo e viene estratto attraverso pozzi petroliferi. La sua composizione chimica può variare a seconda dell'origine geologica e delle caratteristiche del giacimento da cui viene estratto.

Una volta estratto, il petrolio greggio viene sottoposto a processi di raffinazione e lavorazione per separare i diversi componenti e produrre prodotti petroliferi utilizzabili, come benzina, diesel, carburante per aviazione, oli lubrificanti e molti altri. Questo processo di raffinazione consente di ottenere prodotti con diverse caratteristiche e specifiche, che soddisfano le esigenze dell'industria e dei consumatori.

Il gas naturale è una forma di combustibile fossile costituito principalmente da metano.

È un idrocarburo gassoso che si accumula in giacimenti sotterranei porosi e viene estratto attraverso procedure simili a quelle dell'estrazione del petrolio. Dopo l'estrazione, il gas naturale viene trattato e purificato.

Il gas naturale è un combustibile versatile ed efficiente, sfruttato in diversi settori. Viene ampiamente utilizzato per la produzione di energia elettrica e termica, sia nelle centrali elettriche che nei sistemi di riscaldamento

domestico.

Inoltre, viene utilizzato come materia prima nell'industria chimica per la produzione di fertilizzanti, plastica, solventi e altri prodotti chimici.

Il gas naturale è considerato una fonte di energia relativamente pulita rispetto ad altri combustibili fossili, poiché produce meno emissioni di anidride carbonica e altri inquinanti atmosferici durante la combustione.

1.2 Cause delle fluttuazioni dei prezzi

Ci sono molti fattori che influiscono sui prezzi del petrolio e del gas, come la domanda e l'offerta, la stabilità geopolitica dei paesi produttori, il valore delle valute e altri fattori economici e commerciali. Nel corso del XXI secolo, i prezzi relativi al petrolio e al gas naturale hanno subito notevoli cambiamenti, dettati da eventi storici che hanno alternato fasi di recessione e fasi di espansione.

La crisi finanziaria del 2008 ha avuto un impatto significativamente negativo sul settore petrolifero e del gas. Durante la crisi, la domanda di energia è diminuita a causa della recessione globale e del calo dell'attività economica. Ciò ha portato a un surplus di offerta di petrolio e gas nel mercato, il che ha fatto calare i prezzi delle materie prime. Questo causò contrazioni del credito da parte delle banche e da altri intermediari finanziari nei confronti di imprese e famiglie. Queste contrazioni hanno provocato un incremento della disoccupazione, portando quindi ad una diminuzione della domanda di petrolio. Tuttavia, i prezzi del petrolio e del gas hanno iniziato a recuperare gradualmente nel corso del tempo, in parte a causa della ripresa economica globale e della crescente domanda di energia.

Nel marzo del 2020 la pandemia da Coronavirus ha portato una riduzione significativa della domanda dovuta alle misure di lockdown adottate in molti paesi. Con la chiusura di aziende e l'impossibilità per le persone di spostarsi liberamente, la domanda di combustibili è diminuita sensibilmente, portando a un calo dei prezzi del petrolio e del gas.

Sempre nello stesso periodo, si è verificata una guerra dei prezzi tra Russia e Arabia Saudita. Questi due paesi rappresentano i maggiori produttori di petrolio al mondo e la loro decisione di aumentarne la produzione ha contribuito ulteriormente al calo dei prezzi.

Non si può non citare il conflitto tra Russia e Ucraina iniziato nel febbraio 2022. Per quanto riguarda il petrolio, la Russia è al terzo posto per la produzione mondiale, dopo Stati Uniti e Arabia Saudita, e al secondo posto in termini di export. Le esportazioni russe di petrolio sono destinate per il 60% all'Europa e il 20% circa alla Cina. Per quanto riguarda il gas naturale russo, esso rappresenta il 20% del fabbisogno totale di energia dell'Europa. Con la guerra russo-ucraina entrambi i prezzi di questi due beni hanno subito degli shock considerevoli, portando i paesi dell'Unione Europea a coalizzarsi attuando delle azioni nei confronti della Russia per contrastare l'aumento dei prezzi.

Capitolo 2

Modelli per la previsione delle volatilità

2.1 La volatilità

Da un punto di vista finanziario, la volatilità è un indicatore che misura la variabilità del rendimento di un'attività economico-finanziaria all'interno di un determinato arco temporale, viene quindi utilizzata per valutare le fluttuazioni dei futuri movimenti di prezzo di un bene o di un servizio.

La volatilità di per sé, non è direttamente osservabile, tuttavia presenta delle caratteristiche tipiche che si evidenziano osservando le serie dei rendimenti:

- possono verificarsi dei cluster di volatilità (volatility clustering) in cui la volatilità risulta elevata in determinati periodi, bassa in altri (Mandelbrot 1963 e Fama 1965)
- la volatilità rappresenta un fenomeno in costante evoluzione, che si sviluppa continuamente nel tempo, senza però divergere mai verso l'infinito poiché varia all'interno di un intervallo fissato
- la volatilità risulta correlata negativamente con i rendimenti correnti, questo vuol dire che generalmente tende ad essere più elevata negli istanti dove i prezzi calano rispetto a quando i prezzi si alzano. Questa relazione asimmetrica viene definita come effetto di leva finanziaria o *lverage*

La capacità di prevedere gli effetti della volatilità dipende principalmente dalla presenza o meno di una qualche forma di correlazione: quando si analizza una serie di rendimenti, la volatilità risulta molto più complicata da prevedere, poichè essi sono incorrelati.

Data la serie $\{r_t\}$, dove r_t rappresenta il log rendimento di un determinato bene al tempo t , per determinare un modello capace di cogliere la volatilità è necessario introdurre alcuni concetti fondamentali:

$$u_t = E(r_t|T_{t-1}) \quad (2.1)$$

è la media condizionata di r_t dato T_{t-1} e

$$\sigma_t^2 = Var(r_t|T_{t-1}) = E[(r_t - u_t)^2|T_{t-1}] \quad (2.2)$$

è la varianza condizionata di r_t dato T_{t-1} . Si assume che r_t segua un processo autoregressivo a media mobile ARMA(p, q) definito come:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.3)$$

dove ϕ_i , per $i = 1, \dots, p$ e θ_i , per $i = 1, \dots, q$ corrispondono ai parametri di un processo autoregressivo e ai parametri di un processo a media mobile rispettivamente, mentre a_t rappresenta un processo puramente casuale (*white noise*) con media zero e varianza σ_a^2 (quindi $a_t \sim W(0, \sigma_a^2)$). Un processo ARMA lo possiamo anche riscrivere in forma abbreviata come:

$$\phi(B)r_t = \phi_0 + \theta(B)a_t \quad (2.4)$$

dove B (*backward*) è l'operatore ritardo, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ e $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$.

I diversi modelli di volatilità si contraddistinguono in base al modo con cui la varianza condizionata si evolve nel tempo. Andremo quindi a considerare tre tipi di modelli: il modello GARCH (*Generalized Auto-Regressive Conditional Heteroskedasticity*; Bollerslev, 1986), il modello EGARCH e il modello GAS (*Generalized autoregressive score models with applications*; Creal, Koopman, 2013).

2.2 Il modello ARCH

Prima di descrivere i modelli di nostro interesse, è necessario introdurre il modello autoregressivo a eteroschedasticità condizionata ARCH (AutoRegressive Conditional Heteroskedasticity) di Engle (1982), utile per l'analisi di serie storiche finanziarie.

Data la serie $\{a_t\}$ un modello ARCH(m) è definito come:

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2 \quad (2.5)$$

dove ϵ_t rappresentano variabili casuali indipendenti e identicamente distribuite come una normale standard, σ_t segue un processo autoregressivo AR(p), $\alpha_0 > 0$, $\alpha_i \geq 0$ per ogni $i > 0$. Questo tipo di modello ci permette di capire come grandi shock del passato portino ad una maggiore varianza condizionata σ_t^2 .

Concentriamoci sul modello ARCH(1) definito in questo modo:

$$a_t = \sigma_t \epsilon_t \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 \quad (2.6)$$

dove $\alpha_0 > 0$, $\alpha_1 \geq 0$. In questo caso il valore atteso condizionato risulta essere:

$$E(a_t) = E[E(a_t | I_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0 \quad (2.7)$$

mentre la varianza condizionata si calcola come segue:

$$Var(a_t) = E(a_t^2) = E[E(a_t^2 | I_{t-1})] = E(\alpha_0 + \alpha_1 a_{t-1}^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2) \quad (2.8)$$

Data la stazionarietà di a_t con $E(a_t) = 0$, $Var(a_t) = Var(a_{t-1}) = E(a_{t-1}^2)$, possiamo riscrivere la varianza di a_t come $Var(a_t) = \alpha_0 + \alpha_1 Var(a_t)$, per cui $Var(a_t) = \alpha_0 / (1 - \alpha_1)$.

Tuttavia questo tipo di modello risulta piuttosto restrittivo: infatti il valore di α_1^2 in un modello ARCH(1) deve essere compreso nell'intervallo $[0, \frac{1}{3}]$ se la serie presenta un quarto momento finito. Questo vincolo diventa complicato per modelli ARCH di ordine superiore. Nella pratica, ciò limita la capacità dei modelli ARCH con innovazioni gaussiane di catturare l'eccesso di curtosi. Un altro aspetto da considerare è che il modello ARCH non permette di comprendere le cause delle variazioni all'interno di una serie storica temporale. (*Analysis of Financial Time Series*; Tsay, 2005)

2.3 Il modello GARCH

Il modello ARCH ha bisogno di molti parametri per catturare in maniera efficace l'andamento della volatilità all'interno di una serie di rendimenti. Per questo l'economista e matematico danese Tim Bollerslev (1986) propose un modello più parsimonioso detto modello GARCH (ARCH generalizzato).

Data la serie logaritmica dei rendimenti r_t definiamo $a_t = r_t - u_t$ l'innovazione al tempo t . Il processo a_t segue un modello GARCH(m, s) nel momento in cui:

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (2.9)$$

dove $\{\epsilon_t\}$ è una successione di variabili casuali indipendenti e identicamente distribuite con media nulla e varianza unitaria, $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$ e $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$. Ponendo $\eta_t = a_t - \sigma_t^2$ in modo tale che $\sigma_t^2 = a_t^2 - \eta_t$ si può riscrivere il modello GARCH nella seguente forma:

$$a_t^2 = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^s \beta_j \eta_{t-j} \quad (2.10)$$

dove in questo caso η_t non è indipendente e identicamente distribuita. In questo modo abbiamo sviluppato l'equazione come un processo ARIMA. Il modello GARCH(1,1) è definito come:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2.11)$$

dove i parametri α e β sono compresi tra 0 e 1, mentre $(\alpha_1 + \beta_1) \leq 1$. Si osserva che se a_{t-1}^2 o σ_{t-1}^2 sono elevati, allora anche σ_t^2 lo sarà. Questo ci porta a concludere che un elevato a_{t-1}^2 verrà seguito da un altro grande a_t^2 , generando quindi un cluster di volatilità. Il processo GARCH(1,1), in maniera equivalente con i modelli di tipo ARCH, ha una distribuzione con le code più pesanti rispetto a quelle che ritroviamo in una normale

Su questo tipo di modello possiamo effettuare delle previsioni basandoci sulle metodologie prese in considerazione per i modelli di tipo ARIMA. Prendiamo come esempio un GARCH(1,1) e indichiamo come h l'origine della previsione:

- la previsione 1 passo in avanti si ottiene come:

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2 \quad (2.12)$$

- per previsioni più passi in avanti, usiamo $a_t^2 = \sigma_t^2 \epsilon_t^2$ e risciviamo l'equazione (2.11) come:

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2 + \alpha_1 \sigma_t^2 (\epsilon_t^2 - 1) \quad (2.13)$$

quando $t = h + 1$, l'equazione diventa

$$\sigma_{h+2}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (\epsilon_{h+1}^2 - 1). \quad (2.14)$$

Dato che $E(\epsilon_{h+1}^2 - 1 | F_h) = 0$, la previsione della volatilità 2 passi in avanti è

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(1). \quad (2.15)$$

In generale, abbiamo

$$\sigma_h^2(t) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(t-1), \quad t > 1 \quad (2.16)$$

2.4 Il modello EGARCH

Nel 1991, Nelson propose un'estensione del modello GARCH, detto GARCH-esponenziale (EGARCH). Il modello EGARCH riesce a tenere in considerazione la asimmetria degli effetti di shock positivi e negativi, considerando l'innovazione ponderata:

$$g(\epsilon_t) = \theta\epsilon_t + \gamma [|\epsilon_t| - E(|\epsilon_t|)] \quad (2.17)$$

dove θ e γ sono costanti reali e le $\{\epsilon_t\}$ sono indipendenti e identicamente distribuite. Questa innovazione può essere riscritta in modo da evidenziare un'eventuale asimmetria, come:

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \epsilon_t \geq 0 \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \epsilon_t < 0 \end{cases} \quad (2.18)$$

Il modello EGARCH(m, s) può essere scritto nella seguente forma.

$$a_t = \sigma_t \epsilon_t, \quad \ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B - \dots - \alpha_m B^m} g(\epsilon_{t-1}) \quad (2.19)$$

dove B corrisponde all'operatore ritardo e $1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}$ e $1 - \alpha_1 B - \dots - \alpha_m B^m$ sono polinomi che hanno tutti gli zeri al di fuori della circonferenza unitaria. Anche in questo caso l'equazione (2.17) utilizza una riparametrizzazione di un processo ARMA per descrivere l'evoluzione della varianza condizionale di a_t .

Il modello EGARCH si differenzia dal GARCH poichè utilizza la varianza condizionata logaritmica e $g(\epsilon_t)$ consentendo al modello di rispondere in modo asimmetrico ai valori positivi e negativi di a_t .

Il modello EGARCH(1, 1) risulta quindi:

$$a_t = \sigma_t \epsilon_t, \quad (1 - \alpha B) \ln(\sigma_t^2) = (1 - \alpha) \alpha_0 + g(\epsilon_{t-1}) \quad (2.20)$$

dove gli ϵ_t sono indipendenti e identicamente distribuiti secondo una normale standard.

Una forma alternativa del modello EGARCH(m, s) è la seguente:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^s \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^m \beta_j \ln(\sigma_{t-j}^2) \quad (2.21)$$

Per fare previsioni utilizzando il modello EGARCH(1,1) definiamo h il punto di inizio delle previsioni, quindi:

- la previsione un passo in avanti si ottiene come:

$$\sigma_{h+1}^2 = \sigma_h^{2\alpha_1} \exp[(1 - \alpha_1)\alpha_0] \exp[g(\epsilon_h)] \quad (2.22)$$

- la previsione due passi in avanti sarà:

$$\sigma_{h+2}^2 = \sigma_{h+1}^{2\alpha_1} \exp[(1 - \alpha_1)\alpha_0] \exp[g(\epsilon_{h+1})] \quad (2.23)$$

- la previsione t passi in avanti sarà:

$$\sigma_{h+t}^2 = \sigma_{h+t-1}^{2\alpha_1} \exp[(1 - \alpha_1)\alpha_0] \exp[g(\epsilon_{h+t-1})] \quad (2.24)$$

2.5 Il modello GAS

Creal (2013) propose un nuovo modello utile a generare previsioni migliori rispetto ai modelli della famiglia ARCH: il modello GAS (*Generalized Autoregressive Score*).

Definiamo y_t un vettore casuale n-dimensionale, che indica la variabile dipendente di interesse, f_t il vettore dei parametri variabili nel tempo, x_t un vettore di covariate tutte al tempo t , θ un vettore di parametri statici. Definiamo $Y_t = \{y_1, \dots, y_t\}$, $F_t = \{f_0, f_1, \dots, f_t\}$ e $X_t = \{x_1, \dots, x_t\}$. L'insieme delle informazioni disponibili al tempo t è costituito da $\{f_t, F_t\}$, dove

$$F_t = \{Y_{t-1}, F_{t-1}, X_t\}, \quad t = 1, \dots, n \quad (2.25)$$

Si assume che y_t sia generato dalla seguente densità:

$$y_t \sim p(y_t | f_t, F_t; \theta) \quad (2.26)$$

Inoltre, l'aggiornamento del parametro variabile nel tempo f_t è determinato dall'equazione autoregressiva:

$$f_{t+1} = \omega + \sum_{i=1}^p A_i s_{t-i+1} + \sum_{j=1}^q B_j f_{t-j+1} \quad (2.27)$$

dove ω è un vettore di costanti, le matrici dei coefficienti A_i e B_j hanno dimensioni per $i = 1, \dots, p$ e $j = 1, \dots, q$, mentre s_t è una funzione dei dati passati. I coefficienti incogniti sono funzioni di θ , cioè $\omega = \omega(\theta)$, $A_i = A_i(\theta)$, $B_j = B_j(\theta)$.

Quando viene realizzata un'osservazione y_t , si aggiorna il parametro f_t variabile nel tempo al periodo successivo $t + 1$ utilizzando la (2.27) con

$$s_t = S_t(\theta_t)\nabla_t, \quad \nabla_t = \frac{\delta \log p(y_t|f_t, F_t; \theta)}{\delta f_t}, \quad S_t = S(t, f_t, F_t; \theta) \quad (2.28)$$

dove $S(\cdot)$ è una funzione matriciale.

Considerando $\epsilon_t = \sigma_t \mu_t$ dove $\mu_t \sim N(0, 1)$ e σ_t è una deviazione standard variabile nel tempo, si può mostrare che il modello GAS(1, 1) con $f_t = \sigma_t^2$, si riduce a:

$$f_{t+1} = \omega + A_1(\epsilon_t^2 - f_t) + B_1 f_t \quad (2.29)$$

che è equivalente al modello GARCH(1, 1) dato da

$$f_{t+1} = \alpha_0 + \alpha_1 \epsilon_t^2 + \beta_1 f_t \quad (2.30)$$

dove i coefficienti $\alpha_0 = \omega$, $\alpha_1 = A_1$ e $\beta_1 = B_1 - A_1$ sono sconosciuti e richiedono determinate condizioni di stazionarietà (Bollerslev, 1986).

Per stimare il modello GAS (1,1) è appropriato l'utilizzo del metodo della massima verosimiglianza (Creal, 2013).

Capitolo 3

Adattamento dei modelli ai dati

3.1 Dati di riferimento

In questo capitolo il nostro obiettivo è quello di prevedere prezzi di petrolio e gas naturale attraverso i modelli descritti nei paragrafi precedenti.

Il nostro insieme di dati si basa sui prezzi giornalieri di chiusura di sei diverse serie storiche relative ai due asset energetici:

- Petrolio
 1. Prezzi spot del WTI
 2. Prezzi spot del Brent
 3. *Futures* del WTI
 4. *Futures* del Brent

- Gas naturale
 1. *Futures* dell' IPE Brent
 2. *Futures* di NYMEX

Il periodo di riferimento per la nostra analisi va dal 23 Novembre 2009 al 30 dicembre 2022. I dati sono stati raccolti nel sito [investing.com](https://www.investing.com) e dal database di EIA (*U.S Energy Information Administration*), prendendo in analisi tutti i dati disponibili.

Il WTI (*West Texas Intermediate*) e il Brent sono due tipi di petrolio greggio utilizzati come benchmark per il prezzo del petrolio a livello mondiale.

In particolare il WTI viene estratto principalmente in Texas e in Oklahoma e ha un basso contenuto di zolfo rispetto ad altri tipi di petrolio; il Brent, invece, viene estratto dalla piattaforma petrolifera della Brent, situata nel Mare del Nord, ed è considerato più leggero e meno vischioso del WTI.

L' IPE (International Petroleum Exchange) e il NYMEX (New York Mercantile Exchange), sono due mercati borsistici che si concentrano sulla negoziazione di contratti futures basati sui prezzi delle materie prime, tra cui il petrolio e il gas naturale.

I prezzi spot sono i prezzi correnti di una serie finanziaria, come ad esempio il prezzo di un'azione o di una materia prima, in un determinato momento. Questi rappresentano il prezzo al quale un'azione o una materia prima può essere acquistata o venduta in quel momento specifico. I *futures* sono dei contratti standardizzati che permettono di acquistare o vendere un'azione o una materia prima ad un prezzo predeterminato in una data futura. I contratti future sono utilizzati sia per proteggere contro i rischi di variazione dei prezzi delle materie prime e delle azioni, sia per speculare sui movimenti dei prezzi.

Per questo studio analizzeremo i rendimenti di queste serie storiche. Un rendimento al tempo t viene calcolato come:

$$r_t = 100 * \ln(P_t/P_{t-1}) \quad (3.1)$$

dove P_t è il prezzo di chiusura giornaliero.

Osservando la *Figura 3.1*, *Figura 3.2* e *Figura 3.3* si può vedere come gli andamenti del petrolio risultino molto simili tra di loro, sia per quanto riguarda i prezzi spot, sia per i futures, presentando concentrazioni di volatilità nel 2015, 2020 e nel 2022. Per quanto riguarda il gas naturale vediamo che il grafico dell' IPE risulta praticamente piatto fino alla fine del 2021, subendo nel 2022 uno shock considerevole dovuto all'aumento dei prezzi del gas a causa della guerra in Ucraina.

La *Tabella 3.1* mostra alcune statistiche descrittive per le 6 serie storiche dei rendimenti: media e mediana sono approssimativamente uguali a zero per tutti i rendimenti, la varianza relative al gas naturale risultano molto più grandi rispetto a quelle del petrolio, in particolare l'IPE presenta una dispersione dei dati tre volte superiore rispetto alla maggior parte delle varianze dei rendimenti di petrolio. In generale, il range tra minimo e massimo è più ampio negli spot piuttosto che nelle serie dei *futures*.

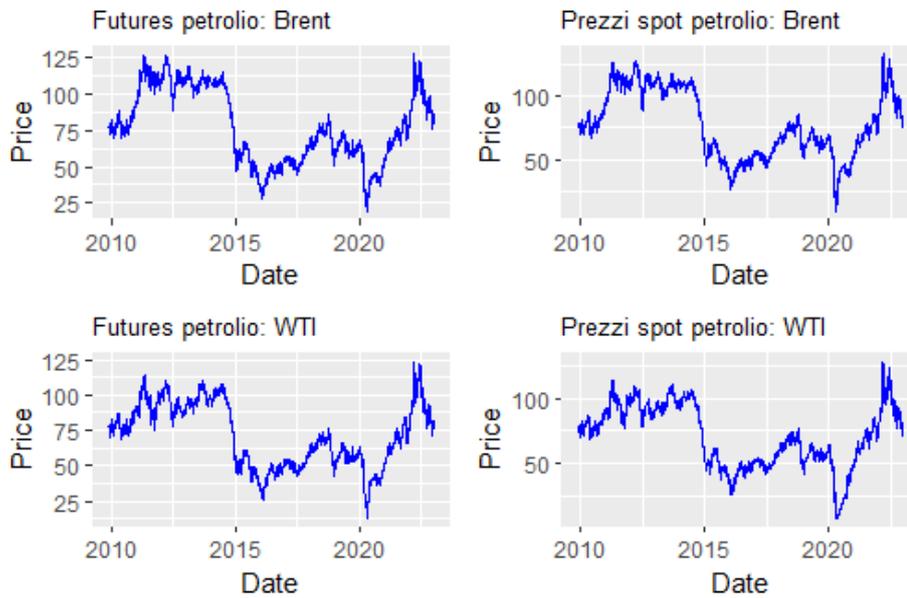


Figura 3.1: *Prezzi spot e futures (giornalieri) del petrolio greggio Brent e WTI*

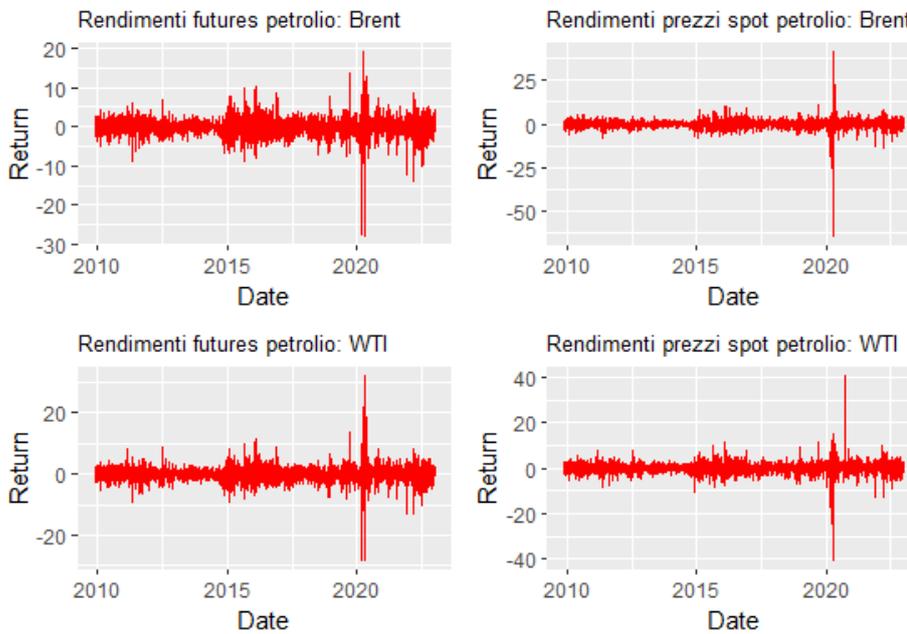


Figura 3.2: *Rendimenti dei prezzi spot e futures (giornalieri) del petrolio greggio Brent e WTI*

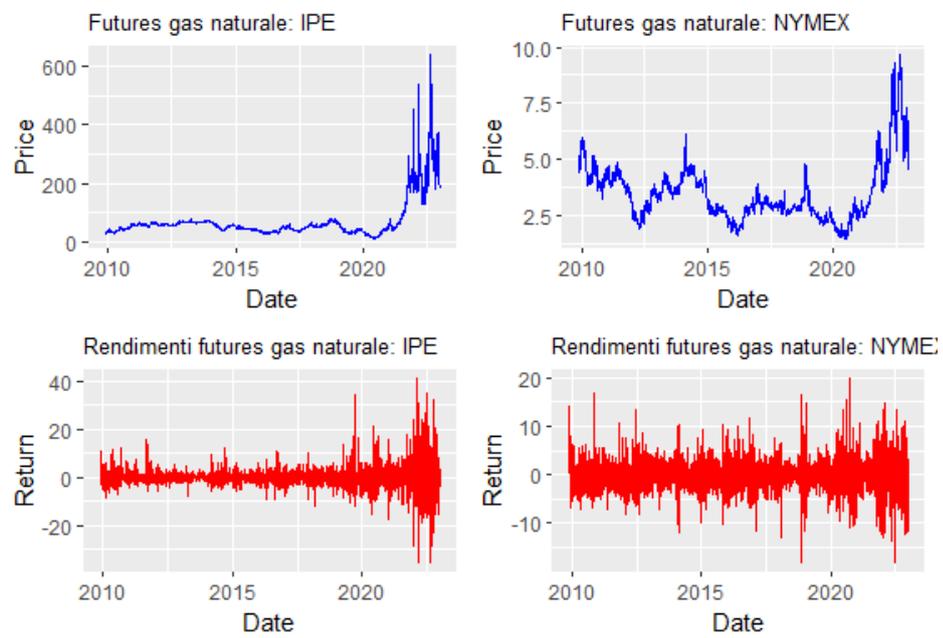


Figura 3.3: *In alto: prezzi spot e futures (giornalieri) del gas naturale IPE e NYMEX. In basso: rendimenti dei prezzi spot e futures (giornalieri) del gas naturale IPE e NYMEX*

	Petrolio				Gas naturale	
	Spot		Futures		Futures	
	Brent	WTI	Brent	WTI	IPE	NYMEX
Media	0.002	0.001	0.003	0.019	0.056	0,000
Varianza	7.806	6.315	5.244	6.785	18.947	10.284
Dev std	2.794	2.513	2.29	2.605	4.353	3.207
Mediana	0.063	0.063	0.086	0.085	-0.021	0
Min	-64.37	-40.463	-27.976	-28.221	-35.592	-18.066
Max	41.202	40.352	19.077	31.963	41.162	19.798
Osservazioni	3317	3609	3386	3410	3329	3409

Tabella 3.1: *Statistiche descrittive delle serie dei rendimenti di petrolio e gas naturale*

3.2 Previsioni in-sample

Dopo aver effettuato un'analisi preliminare dei dati, prendiamo in considerazione dei modelli adatti a catturare le volatilità di queste serie: andiamo quindi a stimare i modelli GARCH(1,1), EGARCH(1,1) e GAS(1,1) con innovazione t di Student. In primo luogo effettuiamo delle previsioni *in-sample* dividendo il campione originale in due sottocampioni: il primo campione che va dal 23 novembre 2009 al 31 dicembre 2019, il secondo dall' 1 gennaio 2020 fino al 30 dicembre 2022. Utilizzeremo il metodo delle *rollings windows* per prevedere le osservazioni del secondo campione. Le tabelle 2, 3 e 4 mostrano i parametri stimati attraverso un GARCH(1,1), EGARCH(1,1) e GAS(1,1) rispettivamente.

Per quanto riguarda i modelli GARCH(1,1) (*Tabella 3.2*) vediamo che per tutte le serie di rendimenti, il parametro μ non risulta significativo e β è sempre maggiore di 0.90: quest'ultimo risultato implica la presenza di un' elevata persistenza di raggruppamenti di volatilità.

I residui standardizzati sono calcolati dividendo i residui ordinari per la deviazione standard condizionata stimata. Questi residui standardizzati vengono utilizzati per valutare il modello, poiché se il modello è valido, né i residui né i loro quadrati devono mostrare autocorrelazione. Il test di Ljung-Box pesato (*Weighted Ljung-Box Test*) è un test statistico utilizzato per valutare la presenza di autocorrelazione nei residui di un modello: più i *p-values* sono elevati, minore è la probabilità che ci sia una correlazione seriale. Nel nostro caso i *p-values* sono estremamente alti, suggerendo che non dovremmo riscontrare una significativa autocorrelazione. Questa conclusione viene confermata dall'analisi dei correlogrammi (*Figura 3.4 e Figura 3.5*).

Per il modello EGARCH(1,1) (*Tabella 3.3*), il parametro μ , come per il

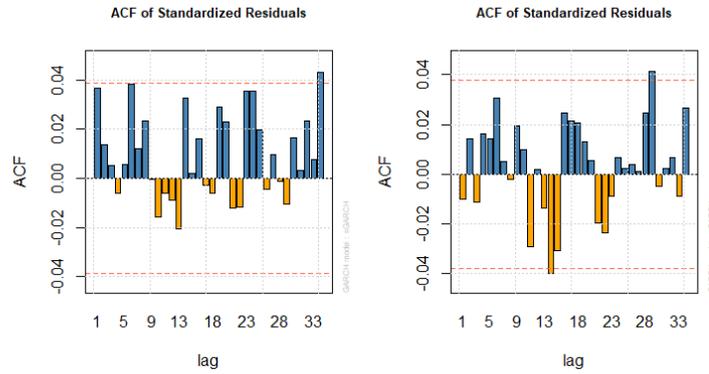


Figura 3.4: Correlogrammi dei residui standardizzati di un $GARCH(1,1)$ delle serie dei rendimenti dei prezzi spot di Brent(a sinistra) e WTI(a destra)

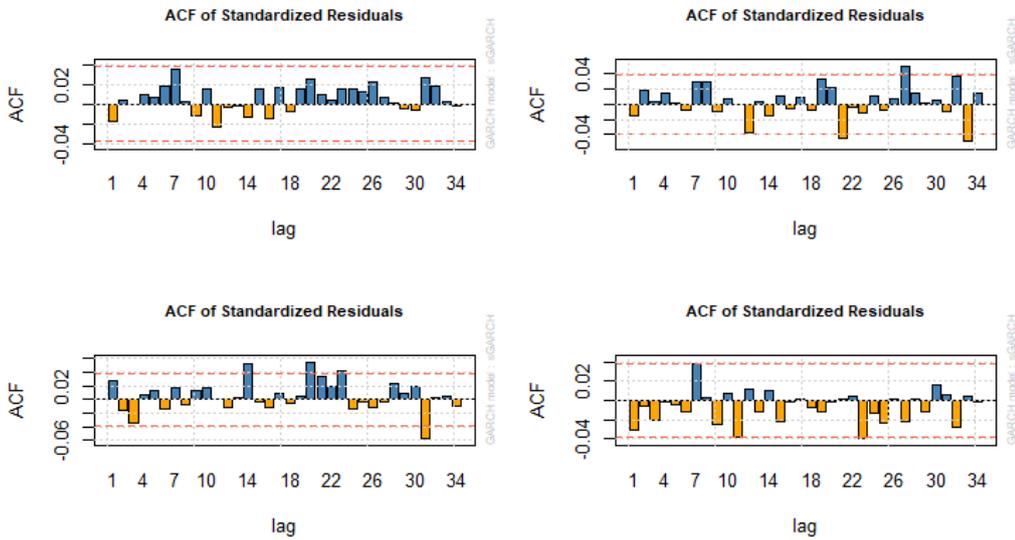


Figura 3.5: Correlogrammi dei residui standardizzati di un $GARCH(1,1)$ delle serie dei rendimenti dei futures di Brent(alto a sinistra), WTI(alto a destra), Ipe Brent(basso a sinistra) e Nymex(basso a destra)

	Petrolio				Gas naturale	
	Spot		Futures		Futures	
	Brent	WTI	Brent	WTI	IPE	NYMEX
μ	0.0175 (0.5489)	0.0437 (0.1355)	0.0350 (0.2056)	0.0386 (0.2074)	-0.0469 (0.1328)	-0.0485 (0.2784)
ω	0.0165 (0.0243)	0.0304 (0.0135)	0.0283 (0.0226)	0.0349 (0.0149)	0.0449 (0.0128)	0.1523 (0.0003)
α	0.0500 (0)	0.0607 (0)	0.0674 (0)	0.0628 (0)	0.0987 (0)	0.0674 (0)
β	0.9470 (0)	0.9330 (0)	0.9281 (0)	0.9301 (0)	0.9003 (0)	0.9125 (0)
ν	6.7036 (0)	6.3391 (0)	5.534 (0)	6.7168 (0)	5.0900 (0)	7.8342 (0)
LL	-4997.3	-5286.9	-5029.7	-5253.6	-5476.1	-6146.7

Tabella 3.2: *Stime di un modello GARCH(1,1) dei rendimenti giornalieri di petrolio e gas naturale. I valori nelle parentesi corrispondono ai p-value*

	Petrolio				Gas naturale	
	Spot		Futures		Futures	
	Brent	WTI	Brent	WTI	IPE	NYMEX
μ	-0.0065 (0.8252)	0.0191 (0.4882)	0.0090 (0.7411)	0.0080 (0.7894)	-0.0545 (0.1004)	-0.0321 (0.4720)
ω	0.0068 (0)	0.0082 (0)	0.0069 (0.0002)	0.0075 (0)	0.0169 (0)	0.0381 (0)
α	-0.0475 (0)	-0.0711 (0)	-0.0627 (0)	-0.0714 (0)	-0.0010 (0.9352)	0.0177 (0.1290)
β	0.9940 (0)	0.9910 (0)	0.9924 (0)	0.9924 (0)	0.9885 (0)	0.9793 (0)
γ	0.0928 (0)	0.0843 (0)	0.0993 (0)	0.0779 (0)	0.1888 (0)	0.1472 (0.0226)
ν	7.2689 (0)	7.0001 (0)	5.9882 (0)	7.5441 (0)	5.0493 (0)	7.5846 (0)
LL	-4985.6	-5258.2	-5012	-5227.8	-5450.5	-6146.9

Tabella 3.3: *Stime di un modello EGARCH(1,1) dei rendimenti giornalieri di petrolio e gas naturale. I valori nelle parentesi corrispondono ai p-value*

modello GARCH(1,1), non risulta significativo per tutte le serie di rendimenti e presenta una bassa correlazione seriale.

Per quanto riguarda il modello GAS(1,1) (Tabella 3.4) le medie condi-

	Petrolio				Gas naturale	
	Spot		Futures		Futures	
	Brent	WTI	Brent	WTI	IPE	NYMEX
μ	0.0757	0.0465	0.1014	0.0382	-0.0510	-0.0415
λ	0.0069 (0.013)	0.0089 (0.009)	0.008 (0.016)	0.0102 (0.008)	0.0134 (0.005)	0.037 (0.000)
ξ	0.1339 (0)	0.1588 (0)	0.1711 (0)	0.1533 (0)	0.2590 (0)	0.1681 (0)
ω	0.9918 (0)	0.9884 (0)	0.9884 (0)	0.9886 (0)	0.9873 (0)	0.9779 (0)
ν	7.248 (0)	6.337 (0)	5.758 (0)	7.105 (0)	4.799 (0)	8.628 (0)
LL	-4998.15	-5285.07	-5028.57	-5253.74	-5439.86	-6146.61

Tabella 3.4: *Stime di un modello GAS(1,1) dei rendimenti giornalieri di petrolio e gas naturale.*

zionate risultato tutte positive, ad eccezione delle serie di rendimenti di gas naturale IPE e NYMEX. Le stime di λ , η e γ sono positive e statisticamente significative per tutte le serie. Un elevato valore di ω indica una forte persistenza della volatilità. Dato che sia i rendimenti del petrolio che quelli del gas naturale presentano valori molto elevati di ω , si può affermare che la volatilità di entrambi i rendimenti è estremamente persistente.

Il modello EGARCH(1,1) presenta valori maggiori della log-verosimiglianza per tutte le serie di rendimenti di petrolio greggio, al contrario per le serie di rendimenti di *futures* di gas naturale, il modello GAS(1,1) presenta valori superiori rispetto agli altri tre modelli. Si può quindi affermare che il modello EGARCH si comporta meglio nel catturare i momenti *in-sample* dei rendimenti di petrolio greggio, mentre il modello GAS per quanto riguarda il gas naturale.

Per queste analisi abbiamo considerato la distribuzione t di Student per i termini di errore. Per verificare la bontà di questa scelta abbiamo effettuato il test dell' *Adjusted Pearson Goodness-of-Fit Test* che confronta la distribuzione teorica della t di Student con la distribuzione empirica dei residui standardizzati. Per tutte le serie di rendimenti i *p-value* di questo test sono abbastanza alti per permetterci di accettare l' ipotesi nulla in base alla quale i residui standardizzati seguono una distribuzione t di Student. Lo si può anche vedere dai grafici della densità empirica e teorica dei residui standardizzati e dal *Q-Q plot*.

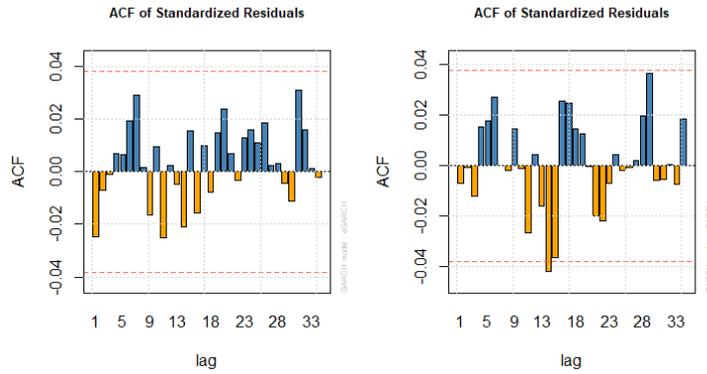


Figura 3.6: Correlogrammi dei residui standardizzati di un EGARCH(1,1) delle serie dei rendimenti dei prezzi spot di Brent(a sinistra) e WTI(a destra)

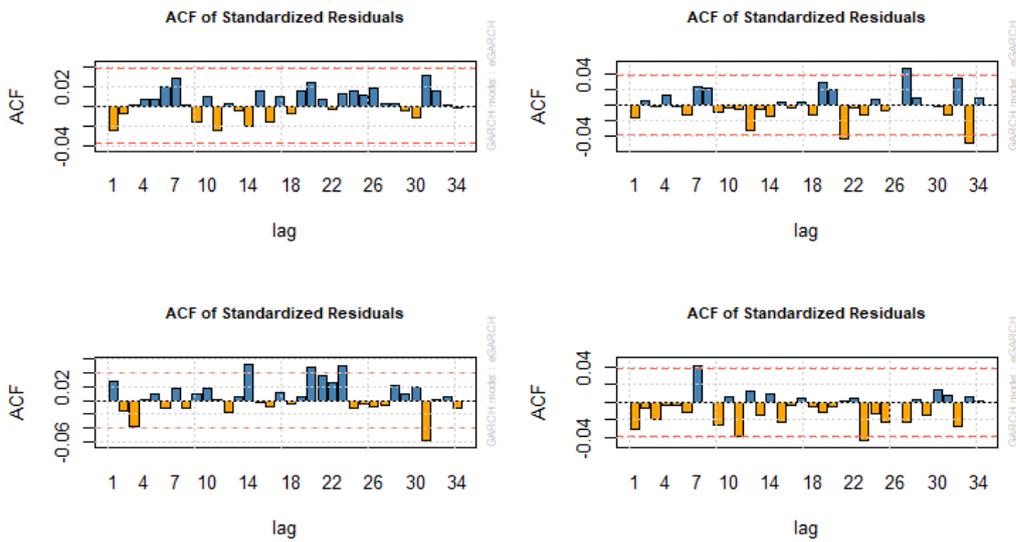


Figura 3.7: Correlogrammi dei residui standardizzati di un EGARCH(1,1) delle serie dei rendimenti dei futures di Brent(altro a sinistra), WTI(alto a destra), Ipe Brent(basso a sinistra) e Nymex(basso a destra)

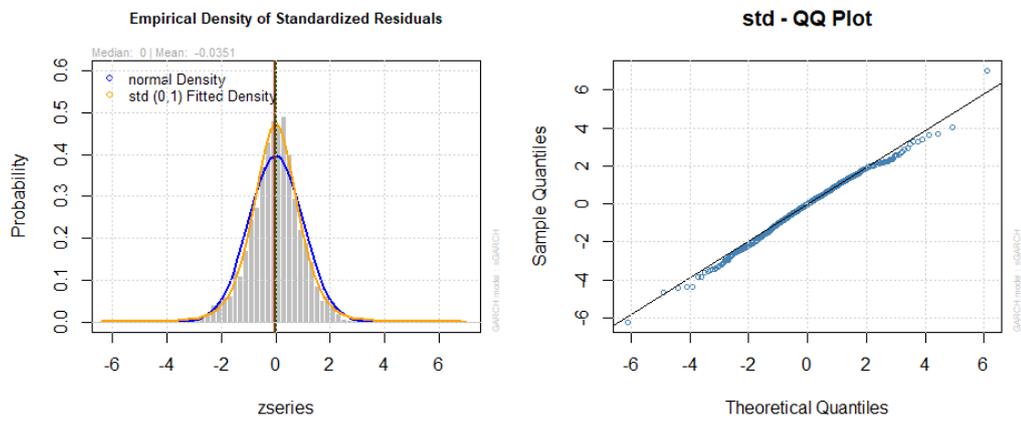


Figura 3.8: *Distribuzione empirica dei residui standardizzati e Q-Q plot della serie dei rendimenti di futures del Brent*

Capitolo 4

Previsione della volatilità

Dopo aver effettuato l'analisi in-sample procediamo con l'analisi out-of-sample per le serie dei rendimenti del petrolio e del gas naturale che abbiamo preso in considerazione fino ad ora.

Effettueremo delle previsioni utilizzando il metodo delle finestre scorrevoli, detto anche rolling-windows. Questa tecnica prevede di suddividere la serie storica in finestre di lunghezza predefinita e utilizzare ciascuna finestra per creare un modello di previsione. Successivamente, la finestra viene spostata di un intervallo di tempo e il processo viene ripetuto per ogni posizione della finestra, fornendo previsioni regolari sulla serie storica. In questo modo, il modello di previsione si adatta ai cambiamenti nella serie storica evitando quindi di essere influenzato da fluttuazioni casuali o temporanee: per la nostra analisi prenderemo in considerazione delle finestre fisse costruendo previsioni ad 1, 5, 20, 60 passi in avanti che corrispondono ad un giorno, una settimana, un mese e un trimestre.

Valuteremo quindi le previsioni del secondo periodo stabilito in precedenza (*out-of-sample*), che va dall' 1 gennaio 2020 fino al 30 dicembre 2022. Per calcolare la previsione della volatilità *out-of-sample* e per valutare le loro prestazioni sono state utilizzate quattro metriche:

- la radice dell'errore quadratico medio (RMSE)
- il Winning Ratio (WR)
- il test di Diebold-Mariano (DM)
- la Somma Cumulata della Differenza degli Errori di Previsione al Quadrato (CSSFED)

4.1 Errore quadratico medio (RMSE)

La radice dell'errore quadratico medio viene utilizzata comunemente per valutare la precisione di un modello di previsione. L'RMSE viene calcolato come la radice quadrata della media delle differenze quadratiche tra le previsioni e i valori effettivi: più basso è il valore dell' RMSE più il modello utilizzato risulta preciso:

$$RMSE = \left[\frac{1}{n} \sum_{t=0}^n (V_t - \hat{V}_t)^2 \right]^{1/2} \quad (4.1)$$

dove t_0 corrisponde all'osservazione iniziale del periodo *out-of-sample* e n al totale delle osservazioni del campione testato, V_t e \hat{V}_t sono la volatilità realizzata e prevista rispettivamente.

L'RMSE (*Root Mean Squared Error*) è una misura dell'errore che permette, elevando le deviazioni al quadrato, di evitare che valori positivi e negativi si annullino. Esso però, è più sensibile ai valori anomali rispetto ad altre misure di accuratezza.

Nella *Tabella 4.2* si possono vedere i valori dell' RMSE relativi ai modelli GARCH, EGARCH e GAS.

Per la previsione 1 passo in avanti, si può vedere come il modello GARCH presenti valori più bassi dell'RMSE per i rendimenti dei prezzi *futures* del petrolio greggio e del gas naturale IPE. I rendimenti dei *futures* di gas naturale NYMEX e i prezzi spot del petrolio greggio WTI sono quelli che hanno valori minori di RMSE per quanto riguarda il modello GAS.

Per quanto riguarda le previsioni 5 passi in avanti, osserviamo che i rendimenti dei prezzi spot WTI e i *futures* di gas NYMEX presentano valori inferiori di RMSE utilizzando il modello EGARCH. Il modello GARCH presenta valori più piccoli per le restati serie: in particolare i *futures* di petrolio greggio Brent presenta il valore più basso in termini di RMSE.

Per le previsioni 20 passi in avanti, il modello GARCH risulta essere migliore per 3 serie di rendimenti su 6, il modello EGARCH per i prezzi spot del petrolio WTI e per i *futures* del gas NYMEX, il modello GAS, invece presenta un valore inferiore di RMSE esclusivamente per la serie di rendimenti dei prezzi spot del petrolio Brent.

Le previsioni 60 passi in avanti sono simili alle previsioni 20 passi in avanti, con l'unica differenza che il modello GAS presenta valori più piccoli

di RMSE anche per i prezzi spot del WTI.

Da questa tabella si evince che, in generale, per le serie di rendimenti dei prezzi *futures* di petrolio greggio e di gas naturale IPE, risulta preferibile il modello GARCH, per i prezzi spot di petrolio WTI è migliore il modello EGARCH, mentre per i prezzi spot del Brent è preferibile il modello GAS fino alle previsioni 20 passi in avanti.

Le previsioni dei rendimenti dei prezzi spot del WTI da parte dei modelli GARCH e GAS risultano quelli meglio previsti di tutte le sei serie di rendimenti.

		Petrolio				Gas naturale	
		Spot		Futures		Futures	
		Brent	WTI	Brent	WTI	IPE	NYMEX
1	GARCH	4,6684	3,6749	3,2813	3,9448	7,8253	4,3593
	E-GARCH	4,6688	3,6746	3,2817	3,9452	7,8266	4,3592
	GAS	4,6685	3,6743	3,2843	3,9449	7,8255	4,3591
5	GARCH	4,6684	3,6743	3,2809	3,9448	7,8254	4,3593
	E-GARCH	4,6688	3,6741	3,2813	3,9454	7,8259	4,3567
	GAS	4,6685	3,6744	3,2842	3,9449	7,8255	4,3591
20	GARCH	4,6686	3,6744	3,2811	3,9448	7,8254	4,3592
	E-GARCH	4,6688	3,6743	3,2814	3,9455	7,8259	4,3591
	GAS	4,6685	3,6745	3,2826	3,9449	7,8256	4,3592
60	GARCH	4,6687	3,6750	3,2813	3,9448	7,8254	4,3591
	E-GARCH	4,6690	3,6748	3,2816	3,9456	7,8258	4,3589
	GAS	4,6686	3,6746	3,2826	3,9449	7,8256	4,3592

Tabella 4.1: *Errori di previsione out-of-sample misurati dall'RMSE.*

L'orizzonte di previsione è 1, 5, 20, 60 giorni. I valori evidenziati indicano che il modello corrispondente è il migliore modello di previsione tra gli altri

4.2 Winning Ratio (WR)

Un altro modo per determinare l'accuratezza delle previsioni di un determinato modello è il WR o Winning Ratio. Questo approccio determina il numero di volte in cui l'errore assoluto di previsione di un modello risulta inferiore a quello di altri modelli, diviso per il numero di previsioni totali.

Questa misura di accuratezza, come per l'RMSE, penalizza le distanze tra i valori osservati e quelli previsti: in questo caso, a differenza dell'RMSE, dove il modello migliore è determinato dal valore più piccolo, un WR maggiore definisce il modello che genera performance migliori rispetto ai modelli associati.

Nella *Tabella 4.2* sono indicati i WR dei modelli presi in esame. Per le previsioni 1 e 5 passi avanti il modello GARCH assume valori più alti nei rendimenti dei prezzi spot del Brent e futures del WTI, per quanto riguarda il petrolio greggio, e i *futures* di IPE per il gas naturale. Il modello EGARCH, invece presenta valori maggiori per le serie di rendimenti dei prezzi spot di WTI e *futures* di NYMEX. Il modello GAS invece presenta un WR maggiore per la serie dei rendimenti di *futures* del Brent.

Per le previsioni a 20 e 60 passi avanti, il modello EGARCH presenta valori di WR maggiori per quattro serie di rendimenti. Per le restanti, il modello GAS ha valori maggiori per i *futures* di Brent, mentre per i *futures* di IPE il WR è maggiore per il modello GARCH.

Si può quindi osservare che, con passi di previsione pari a 1 e 5 il modello GARCH presenta valori maggiori di WR, mentre, aumentando i passi di previsione, il modello EGARCH risulta preferibile.

		Petrolio				Gas naturale	
		Spot		Futures		Futures	
		Brent	WTI	Brent	WTI	IPE	NYMEX
1	GARCH	0.4882	0.2121	0.04651	0.4987	0.5328	0.4628
	E-GARCH	0.4421	0.4564	0.4483	0.4621	0.4501	0.5142
	GAS	0.0697	0.3315	0.5052	0.0391	0.0171	0.0230
5	GARCH	0.4645	0.1916	0.0556	0.4722	0.5249	0.4596
	E-GARCH	0.4421	0.4693	0.4457	0.4596	0.4449	0.5050
	GAS	0.0934	0.3391	0.4987	0.0657	0.0302	0.0354
20	GARCH	0.3197	0.2196	0.0426	0.4306	0.5184	0.4217
	E-GARCH	0.4447	0.4532	0.4470	0.4596	0.4213	0.5000
	GAS	0.2355	0.3272	0.5103	0.1098	0.0603	0.0783
60	GARCH	0.2776	0.2120	0.0349	0.4154	0.4449	0.3296
	E-GARCH	0.4434	0.4306	0.4470	0.4596	0.4278	0.5227
	GAS	0.2776	0.3574	0.5181	0.1250	0.1273	0.1477

Tabella 4.2: Valori di Winning Ratio (WR) delle previsioni out-of-sample. L'orizzonte di previsione è 1, 5, 20, 60 giorni. I valori evidenziati indicano che il modello corrispondente è il migliore modello di previsione tra gli altri

4.3 Test di Diebold-Mariano

Il test di Diebold-Mariano è un test statistico utilizzato per confrontare le prestazioni di due modelli di previsione, al fine di determinare se uno dei modelli sia significativamente migliore dell'altro nella previsione di una serie temporale. Il test è stato proposto da Francis X. Diebold e Robert S. Mariano nel 1995 e si basa sulla differenza tra le funzioni di perdita per le due previsioni: esse sono accurate nella stessa misura se il valore atteso della differenza è pari a zero.

Sia:

- $L(e_{it})$ la funzione di perdita dell'errore di previsione
- $d_t = L(e_{1t}) - L(e_{2t})$ la differenza tra due funzioni di perdita

Le due ipotesi risultano essere:

$$\begin{cases} H_0 : E(d_t) = 0 \\ H_1 : E(d_t) \neq 0 \end{cases}$$

dove l'ipotesi nulla H_0 indica che le due previsioni hanno la stessa accuratezza previsiva, mentre l'ipotesi alternativa H_1 indica che le due previsioni non hanno la stessa accuratezza.

Il test di Diebold-Mariano è costruito basandosi sul teorema del limite centrale, considerando la seguente distribuzione:

$$\sqrt{h}(\bar{d} - \mu) \xrightarrow{d} N(0, \hat{\omega}_m(d_t)) \quad (4.2)$$

dove:

- $\bar{d} = \frac{1}{h} \sum_{t=1}^h d_t$ è la media campionaria della differenza d_t
- $T - T_0$ è il numero di passi in avanti considerati nell'effettuare le previsioni
- $\hat{\omega}_m(d_t)$ è la varianza di lungo periodo di d_t
- $\mu = E(d_t)$

La statistica test risulta quindi essere: $DM = \frac{\hat{d}}{\sqrt{\frac{\hat{\omega}_m(d_t)}{h}}} \xrightarrow{d} N(0, 1)$

Dato che questo test è un test a due code, si rifiuta l'ipotesi nulla quando: $|DM| > z_{\frac{\alpha}{2}}$ oppure quando $pvalue = 2Pr(Z > DM) < \alpha$ con $Z \sim N(0, 1)$ e α è il livello di significatività del test.

Nel caso in esame, viene utilizzato il test di Diebold-Mariano per valutare la significatività previsiva dei modelli presi in considerazione.

Nella *Tabella 4.3* i modelli GARCH, EGARCH e GAS sono stati classificati basandosi sull' RMSE: il numero 1 è associato al modello che presenta il valore di RMSE più basso, questo viene poi accoppiato con gli altri modelli dove è stato calcolato il test DM a una coda.

Per le previsioni 1 passo avanti il modello GARCH presenta la migliore capacità predittiva per la serie dei rendimenti dei *futures* di WTI, mentre il modello GAS, per le serie dei prezzi spot di WTI e *futures* di NYMEX. Per i rendimenti dei *futures* di gas naturale IPE, dei prezzi spot e *futures* di petrolio greggio Brent , nonostante l'RMSE del modello GARCH risulti minore rispetto agli altri, presenta una capacità predittiva identica ai modelli GAS ed EGARCH.

Le previsioni 5 passi in avanti non presentano sostanziali cambiamenti, tranne per i prezzi spot di WTI e *futures* di NYMEX dove il modello EGARCH presenta una capacità predittiva migliore rispetto agli altri.

Le previsioni 20 passi in avanti il modello GAS è migliore per le serie di rendimenti dei prezzi spot di petrolio Brent. Le serie dei *futures* di Brent e IPE non presentano sostanziali differenze in termini di capacità predittiva dei tre modelli, mentre per il gas NYMEX il modello EGARCH risulta più efficiente.

Infine le previsioni 60 passi in avanti sono relativamente simili alle previsioni 20 passi in avanti tranne che, per la serie di rendimenti di NYMEX, la coppia GARCH-GAS non presenta una sostanziale differenza. Inoltre il modello GAS risulta migliore per tutte le serie dei prezzi spot di petrolio greggio.

		Petrolio				Gas naturale	
		Spot		Futures		Futures	
		Brent	WTI	Brent	WTI	IPE	NYM
1	GARCH	1	3	1	1*	1	3
	E-GARCH	3	2	2	3	3	2*
	GAS	2*	1*	3	2*	2	1*
5	GARCH	1	2*	1	1*	1	3
	E-GARCH	3	1*	2	3	3	1*
	GAS	2*	3	3	2*	2	2*
20	GARCH	2*	2*	1	1*	1	2*
	E-GARCH	3	1*	2	3	3	1
	GAS	1*	3	3	2*	2	3
60	GARCH	2*	3	1	1*	1	2
	E-GARCH	3	2*	2	3	3	1*
	GAS	1*	1*	3	2*	2	3

Tabella 4.3: *Test di capacità predittiva out-of-sample (test Diebold-Mariano)*
L'orizzonte di previsione è 1, 5, 20, 60 giorni. I numeri nella tabella indicano il grado del modello nella previsione delle volatilità in termini di errori al quadrato. 1* significa che il modello è il miglior modello di previsione con previsioni significativamente più accurate rispetto ai modelli classificati al secondo e terzo posto. 2* significa che è il modello di previsione subottimale con previsioni significativamente più accurate rispetto al terzo modello classificato.

4.4 Somma cumulativa della differenza quadratica degli errori di previsione (CSSFED)

Questa metrica è stata utilizzata per monitorare graficamente nel tempo l'evoluzione della performance di previsione di un determinato modello rispetto ad un altro (Siliverstoto, 2017 ; Welch & Goyal, 2008).

Il CSSFED (Cumulated Sum of Squared Forecast Error Differential) è definito come la seguente equazione:

$$CSSFED_{t_0,t_1} = \sum_{t=t_0}^{t_1} [(V_{M1,t} - \hat{V}_t)^2 - (V_{M2,t} - \hat{V}_t)^2] \quad (4.3)$$

dove $CSSFED_{t_0,t_1}$ misura la somma cumulata della differenza degli errori di previsione al quadrato tra il modello M1 e il modello M2 durante il periodo in cui sono state calcolate, ovvero t_0 e t_1 .

A differenza dell' RMSE, che rappresenta una variabile scalare, il CSSFED corrisponde ad una serie temporale che mostra l'accuratezza previsionale dei modelli $M1$ e $M2$.

Quando il CSSFED ha una tendenza ad alzarsi significa che il secondo modello della coppia produce errori di previsione più piccoli rispetto all'altro fino a quel punto della serie storica. Al contrario, se si abbassa vuol dire che il modello $M2$ produce errori di previsione più grandi di $M1$.

Nel caso in questione, sono state selezionate, per brevità, le serie dei rendimenti con previsioni a 5 e 60 passi in avanti. Nella *Figura 4.1* e nella *Figura 4.2* sono riassunti i risultati per le coppie GARCH-EGARCH, GAS-GARCH, GAS-EGARCH per gli orizzonti a 1 e 60 passi, rispettivamente.

Studiando la *Figura 4.1*, si può osservare che per i prezzi spot e *futures* del Brent, le coppie GARCH-EGARCH e GAS-EGARCH hanno un crollo considerevole ad aprile 2020 ed iniziano ad assumere valori negativi da febbraio 2021. Questo è dovuto al crollo dei prezzi del petrolio e greggio quando, a Marzo, l'Arabia Saudita ha ridotto in maniera significativa il prezzo del petrolio venduto ai mercati esteri. Decise, infatti, di aumentarne la produzione in un momento in cui la domanda di petrolio era molto bassa, a causa della pandemia da Covid-19. Il CSSFED di queste coppie raggiunge il punto minimo a febbraio 2022 con il conflitto russo-ucraino.

Per quanto riguarda il gas naturale, analizzando l'andamento dei CSSFED dei *futures* di Nymex il modello EGARCH sembra essere superiore rispetto ai modelli GARCH e GAS, mentre per i *futures* di Ipe prevale il modello GARCH.

La coppia GAS-GARCH, per 5 serie di rendimenti su 6, assume per tutto l'out-of-sample valori prossimi allo zero dimostrandoci che i due modelli hanno errori di previsione simili, invece per i *futures* di Ipe il CSSFED della coppia GAS-GARCH tende ad aumentare.

Osservando invece la *Figura 4.2*, relativa alle previsioni a 60 passi in avanti, si può vedere che le serie di rendimenti dei prezzi spot e *futures* di petrolio Brent presentano più o meno le stesse tendenze di CSSFED calcolato con orizzonte a 5 passi in avanti: in questo caso il modello EGARCH prevale sugli altri modelli fino agli inizi del 2021, successivamente predomina il modello GAS. Lo stesso si può dire per le serie di rendimenti dei prezzi spot e *futures* di WTI.

Valutando i grafici relativi al gas naturale, il modello GARCH sembra essere migliore per i *futures* di gas naturale IPE, mentre per i *futures* di NYMEX prevale il modello EGARCH.

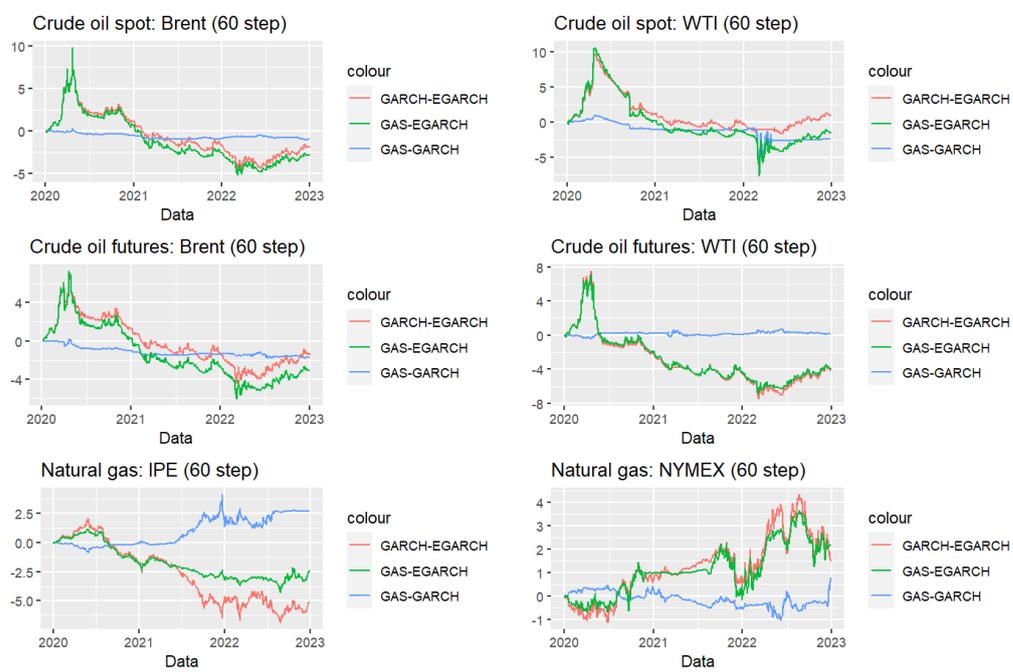


Figura 4.1: *Valutazione dell'accuratezza delle previsioni con l'utilizzo della CS-SFED all'orizzonte di previsione di 60 passi in avanti*

Capitolo 5

Conclusioni

L'indagine condotta ha permesso di analizzare e confrontare tre modelli statistici (GARCH, EGARCH e GAS), al fine di determinare quale sia il migliore in termini di capacità predittiva della volatilità per 6 serie di rendimenti di prezzi del petrolio greggio e gas naturale a orizzonti multipli. Per calcolare la previsione della volatilità *out-of-sample* e per valutare le loro prestazioni sono stati utilizzati quattro approcci differenti: l'RMSE, il WR, il test di Diebold-Mariano e il CSSFED. Dopo un'attenta valutazione e confronto dei risultati ottenuti, è emerso che in generale, il modello GARCH risulta il modello migliore per le previsioni 1 e 5 passi in avanti, mentre per previsioni a più passi in avanti, i modelli GAS ed EGARCH sono preferibili.

In particolare per le serie dei rendimenti dei *futures* di petrolio greggio di WTI e di gas naturale IPE, il modello GARCH ha dimostrato di possedere una maggiore accuratezza e affidabilità nella previsione dei dati, superando gli altri due modelli in termini di performance.

Tuttavia, alcune volatilità possono essere ben descritte anche dagli altri modelli: in particolare il modello EGARCH per i *futures* del gas NYMEX e il modello GAS per i prezzi spot del petrolio greggio e WTI per previsioni 60 passi in avanti.

In conclusione, l'utilizzo del modello GARCH consente di ottenere risultati più accurati e affidabili per previsioni nel breve termine. Tuttavia è importante sottolineare che la scelta del modello dipende dal contesto specifico dell'analisi e dalla natura dei dati in esame.

Bibliografia

Agnolucci, P. (2009). *Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models*. *Energy Economics*, 31(2), 316–321.

Ardia D, Boudt K and Catania L (2016). "Generalized Autoregressive Score Models in R: The GAS Package."

Bollerslev, T. (1986). *Generalized autoregressive conditional heteroskedasticity*. *Journal of Econometrics*, 31(3), 307–327.

Cheong, C. W. (2009). *Modeling and forecasting crude oil markets using ARCH-type models*. *Energy Policy*, 37(6), 2346–2355.

Chun, D., Cho, H., & Kim, J. (2019). *Crude oil price shocks and hedging performance: A comparison of volatility models*. *Energy Economics*, 81, 1132–1147.

Creal D, Koopman SJ, Lucas A (2013). "Generalized Autoregressive Score Models with Applications." *Journal of Applied Econometrics*, 28(5), 777-795.

Harvey, A., & Sucarrat, G. (2014). *EGARCH models with fat tails, skew-*

ness and leverage. *Computational Statistics & Data Analysis*, 76, 320–338.

Herrera, M. A., Hu, L., & Pastor, D. (2018). *Forecasting crude oil price volatility. International Journal of Forecasting*, 34(4), 622–635.

Kang, H. S., Kang, S. M., & Yoon, S. M. (2009). *Forecasting volatility of crude oil markets. Energy Economics*, 31(1), 119–125.

Kang, H. S., & Yoon, S. M. (2013). *Modeling and forecasting the volatility of petroleum futures prices. Energy Economics*, 36(3), 354–362.

Nelson, D. B. (1991). *Conditional heteroskedasticity in asset returns: A new approach. Econometrica*, 59, 347–370.

Tsay (2005). "*Analysis of Financial Time Series. Second edition*", 97-130.

Wei, Y., Wang, Y., & Huang, D. (2010). *Forecasting crude oil market volatility: Further evidence using GARCH-class models. Energy Economics*, 32(6), 1477–1484.

Zhang, J. Y., Yao, T., He, Y. H., & Ripple, R. (2019). *Volatility forecasting of crude oil market: Can the regime switching GARCH model beat the single-regime GARCH models? International Review of Economics & Finance*, 59, 302–317.