

## PERFORMA PREDIKSI: Kointegrasi GARCH-SVR VS. GARCH UNTUK VOLATILITAS HARGA KOMODITAS ENERGI GLOBAL

Prajna Pramita Izati<sup>1\*</sup>, Fariz Budi Arafat<sup>1</sup>

<sup>1</sup> Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

\*e-mail: [prajnapramitaizati@lecturer.undip.ac.id](mailto:prajnapramitaizati@lecturer.undip.ac.id)

DOI: [10.14710/j.gauss.14.2.500-511](https://doi.org/10.14710/j.gauss.14.2.500-511)

### Article Info:

Received: 2025-03-23

Accepted: 2025-12-03

Available Online: 2025-12-07

### Keywords:

GARCH; GARCH-SVR; Long memory; SVR; Volatility

**Abstract:** Forecasting provides benefits in decision-making, one of which is forecasting the volatility of global energy commodity prices. However, there are challenges in forecasting volatility due to the presence of heteroskedasticity and long-memory effects in the data. Therefore, a combination of the GARCH and SVR methods is needed as a cointegration-based *machine learning* approach. The aim of this study is to compare the forecasting performance of GARCH and GARCH-SVR for global energy commodity price volatility. The findings indicate that the GARCH-SVR model performs well when volatility data exhibits non-stationary long-memory characteristics, whereas the GARCH model is more suitable when the volatility data shows stationary long-memory characteristics.

## 1. PENDAHULUAN

Komoditas energi merujuk pada sumber daya alam yang menghasilkan energi atau bahan bakar. Komoditas ini sangat penting bagi berbagai industri, sehingga harga mereka menjadi faktor kunci dalam perekonomian global. Harga komoditas energi memiliki peran krusial dalam perekonomian global, dan fluktuasi harga dapat memberikan dampak yang signifikan. Perubahan harga komoditas energi, baik naik maupun turun, mempengaruhi berbagai aspek ekonomi, baik pada tingkat mikro maupun makro. Variasi harga ini dapat memengaruhi kebijakan pemerintah, tren ekonomi, serta keputusan investasi di sektor energi. Oleh karena itu, peramalan harga komoditas energi menjadi hal yang sangat penting bagi investor dan organisasi internasional dalam merancang kebijakan untuk mengelola volatilitas harga di pasar.

Komoditas energi dianggap memiliki hubungan kointegrasi karena fungsinya yang sama sebagai sumber energi. Penelitian yang dilakukan oleh Lin dan Li (2015) menunjukkan adanya hubungan kointegrasi antara harga minyak mentah dan harga gas alam. Pergerakan harga komoditas ini dipengaruhi oleh faktor eksternal yang menyebabkan fluktuasi yang signifikan. Fluktuasi tersebut menunjukkan perilaku non-linear dan non-stasioner, sehingga menghasilkan volatilitas yang sulit diprediksi. Ketidakpastian ini menjadi tantangan besar dalam peramalan volatilitasnya.

Memahami dan meramalkan volatilitas sangat penting untuk merancang kebijakan yang tepat. Perubahan harga produk dari proyek pertambangan, khususnya batu bara, gas alam, dan minyak mentah, telah berdampak besar pada aktivitas pertambangan dan sektor terkait (Ramirez, et al., 2003). Oleh karena itu, pemahaman terhadap volatilitas harga minyak, gas, dan batu bara menjadi sangat krusial, karena fluktuasi yang terus-menerus dapat menciptakan risiko bagi produsen maupun konsumen dalam industri ini. Hal ini juga dapat mempengaruhi keputusan investasi dalam rantai pasokan, produksi, dan transportasi minyak, batu bara, serta gas alam (Pindyck, 2004).

Pemodelan data volatilitas menggunakan *Autoregressive Integrated Moving Average* (ARIMA) menghadapi tantangan karena sifat data yang heterogen. Untuk mengatasi hal ini,

Engle (1982) memperkenalkan model *Autoregressive Conditional Heteroscedasticity* (ARCH) guna merepresentasikan fluktuasi dalam volatilitas. Bollerslev (1986) kemudian mengembangkan model ini dengan menciptakan *Generalized ARCH* (GARCH), yang memberikan fleksibilitas lebih dalam menangkap karakteristik volatilitas. Namun, model GARCH memiliki keterbatasan, terutama dalam ketidakmampuannya menangkap efek asimetris (dikenal sebagai *leverage effect*). Untuk mengatasi keterbatasan ini, diperlukan model yang mampu menangkap efek asimetris dalam volatilitas, khususnya dalam bagaimana berita baik dan buruk mempengaruhi perubahan harga. Salah satu pendekatan yang digunakan adalah model *Asymmetric GARCH*, yang dirancang khusus untuk menangkap efek *leverage*. Model ini dapat menangkap respons volatilitas yang berbeda terhadap berita positif dan negatif, sehingga memberikan gambaran perilaku pasar yang lebih akurat. Penelitian yang dilakukan oleh Charles dan Darne (2018) menunjukkan bahwa model asimetris yang diestimasi berdasarkan *filtered returns* memberikan estimasi *out-of sample* yang lebih baik dibandingkan model GARCH, GAS, dan MSM yang diestimasi berdasarkan *raw return series* untuk Brent dan WTI dengan masing-masing kriteria RMSE sebesar 0,0971 dan 0,2020.

Ketika terdapat hubungan *long-memory* di antara variabel, kombinasi linear dari variabel-variabel tersebut menghasilkan residual yang mengikuti proses *long-memory*. Karakteristik ini menunjukkan adanya hubungan kointegrasi di antara variabel-variabel yang terlibat (Abdullah, et al., 2020). Penelitian sebelumnya mengenai kointegrasi telah dilakukan oleh Bakshi et al. (2021), yang menggunakan metode *machine learning* seperti *Support Vector Regression* (SVR) dan XGBoost. Hasil penelitian mereka menunjukkan bahwa tidak ada hubungan kointegrasi antara variabel pasar dan harga minyak mentah dengan nilai p-value untuk semua pasar  $> \alpha (0,05)$ , sehingga algoritma SVR dan XGBoost tidak dapat membangun model prediksi yang efektif. Dalam penelitian lain, Chuanhe, Zhongwen, dan Ying (2013) melakukan eksperimen kointegrasi non-linear di pasar keuangan menggunakan SVR. Hasil penelitian mereka menunjukkan bahwa SVR unggul dalam menangani hubungan kointegrasi non-linear dengan nilai R/S sebesar 1,2873 pada  $\alpha = 0,05$  antara berbagai variabel keuangan. Selain itu, perbandingan model dengan *Nonlinear Least Squares* (NLS), *Neural Networks* (NN), dan SVR menunjukkan bahwa SVR mencapai koefisien determinasi tertinggi ( $R^2 = 0,8369$ ) di antara ketiga model tersebut. Namun, penelitian yang disebutkan sebelumnya terbatas pada pemodelan menggunakan SVR saja dan belum mencakup peramalan deret waktu. Penelitian ini akan berfokus pada pemodelan dan peramalan dengan menggunakan model GARCH-SVR dan GARCH, yang memperhitungkan heterogenitas serta efek *long memory* dalam data volatilitas.

## 2. TINJAUAN PUSTAKA

Volatility *clustering* terjadi saat harga komoditas mengalami guncangan besar yang diikuti oleh periode tenang (Gujarati, 2004). Pola ini terlihat dari perubahan besar yang cenderung diikuti perubahan besar lainnya, sedangkan perubahan kecil diikuti perubahan kecil (Mandelbrot, 1963). Volatilitas tinggi menyebabkan heteroskedastisitas, yaitu fluktuasi ekstrem naik dan turun. Volatilitas, yang diukur dengan simpangan baku *return*, berperan penting dalam penilaian risiko (Ariefianto, 2012). Volatilitas tinggi menunjukkan risiko besar namun berpotensi memberikan imbal hasil tinggi. Penelitian ini mengusulkan representasi berikut untuk proses rata-rata dalam persamaan (1)

$$r_t = \mu_t + \varepsilon_t \quad (1)$$

dimana  $I^{t-1}$  merepresentasikan informasi  $\varepsilon_t$  dari periode waktu  $t - 1$  and  $\varepsilon_t | I^{t-1} \sim iid(0, \sigma_t^2)$ . Rata-rata diasumsikan bernilai nol, dan residual bergantung pada semua

informasi sebelumnya hingga waktu  $t$ . Sederhananya, jika  $\mu_t = 0$  diketahui, maka nilai  $r_t = \varepsilon_t$ , dan varians  $r_t$  dapat dinyatakan sebagai berikut dalam persamaan (2) dan (3).

$$\begin{aligned} Var(r_t) &= Var(\varepsilon_t) \\ \sigma_{Zt}^2 &= \sigma_{\varepsilon t}^2 \end{aligned} \quad (2)$$

Jika  $E(r_t) = \mu = 0$ , maka  $(E(r_t))^2 = 0$ , sehingga:

$$\begin{aligned} Var(r_t) &= E(r_t^2) - (E(r_t))^2 \\ Var(r_t) &= E(r_t^2) \\ \sigma_{\varepsilon t}^2 &= \sigma_{rt}^2 = E(r_t^2) \end{aligned} \quad (3)$$

Persamaan (3) menunjukkan bahwa volatilitas data dapat diestimasi berdasarkan nilai  $r_t^2$  yang diberikan (Izati, et al, 2024).

Konsep dasar model ARCH adalah bahwa (a) *return* yang dikoreksi terhadap rata-rata  $\varepsilon_t$  tidak berkorelasi secara serial tetapi tetap bergantung, dan (b) ketergantungan  $\varepsilon_t$  ini dapat dinyatakan dalam fungsi kuadratik sederhana dari nilai lag-nya. Secara khusus, model ARCH( $p$ ) menyatakan bahwa:

$$\varepsilon_t = \sigma_t u_t, \quad (4)$$

dimana  $u_t \sim iid(0,1)$ . The ARCH( $p$ ) model secara umum dituliskan sebagai berikut:

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \\ \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2 \end{aligned} \quad (5)$$

dengan  $\omega > 0$  dan  $\alpha_i \geq 0$  (Engle, 1982).

Model GARCH ( $p,q$ ) digunakan untuk menangani heteroskedastisitas dalam data. Varians model dinyatakan sebagai berikut.

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \\ \sigma_t^2 &= \omega + \alpha(L) \varepsilon_t^2 + \beta(L) \sigma_t^2 \end{aligned} \quad (6)$$

di mana  $L$  merupakan lag operator,  $\alpha(L) = \sum_{i=1}^q \alpha_i L^i$  dan  $\beta(L) = \sum_{j=1}^p \beta_j L^j$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  dan  $\sum_{i=1}^q (\alpha_i + \beta_j) < 1$  (Caporin, 2002).

Penanganan asimetri volatilitas terdapat berbagai model *asymmetric* GARCH seperti APARCH, GJR-GARCH, dan EGARCH dikembangkan. Model ini menangkap efek berbeda dari kejutan positif dan negatif pada volatilitas, memberikan representasi lebih akurat untuk data keuangan. Persamaan varians kondisional dalam model GJR-GARCH dinyatakan sebagai berikut:

$$\sigma_t^2 = \omega + \sum_{i=1}^p (\alpha_i \varepsilon_{t-i}^2) + \gamma I_{t-1} \varepsilon_{t-1}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (7)$$

$$I_{t-1} = \begin{cases} 0 & \text{if } Z_{t-1} \geq \mu \\ 1 & \text{if } Z_{t-1} \leq \mu \end{cases} \quad (8)$$

dimana  $\gamma$  merepresentasikan *leverage* parameter;  $I_{t-1}$  didefinisikan variabel *dummy*. APARCH model mengekspresikan  $\sigma_t^2$  sebagai:

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \quad (9)$$

dimana  $\omega > 0$ ,  $\alpha_i > 0$ ,  $\beta_j > 0$ ,  $\delta > 0$ , dan  $-1 < \gamma_i < 1$ . Model EGARCH adalah varian terkenal dari *asymmetric* GARCH. Persamaan varians kondisionalnya dinyatakan sebagai berikut:

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^p (\alpha_i |\frac{\varepsilon_{t-i}}{\sigma_{t-i}}| + \phi_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}) + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2 \quad (10)$$

Faktor asimetris  $\phi_i$  menunjukkan perbedaan dampak berbagai shocks. Jika  $\phi_i > 0$ , *positive shocks* lebih berpengaruh dibandingkan *negative shock* dan sebaliknya. Properti penting dari variabel  $I(1)$  adalah bahwa mungkin ada kombinasi linier dari variabel-variabel yang  $I(0)$ . Jika demikian maka variabel-variabel tersebut dikatakan *cointegrated*. Konsep *cointegration* pertama kali diperkenalkan oleh Granger (1981). Pada variabel  $y_t$  dan  $x_t$  yang  $I(1)$ , maka  $y_t$  dan  $x_{it}$  dikatakan terkointegrasi jika  $u_t$  adalah  $I(0)$ . Proses ini dapat dilambangkan dengan mengatakan bahwa  $y_t$  dan  $x_{it}$  adalah  $CI(1,1)$ . Secara lebih umum, jika  $y_t$  adalah  $I(d)$  dan  $x_{it}$  adalah  $I(d)$ , maka  $y_t$  dan  $x_{it}$  adalah  $CI(d,b)$ . Kombinasi linear dari deret  $y_t$  dan  $x_{it}$  untuk  $i=1,2,\dots,p$  dapat dituliskan sebagai berikut.

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_p x_{pt} + u_t \quad (11)$$

*Support Vector Regression* (SVR) adalah adaptasi *machine learning* dari model klasifikasi SVM (Lu, 2009). SVR mampu menangani estimasi non-linear dan efektif dalam prediksi data deret waktu, terutama di bidang keuangan. Menurut Cortez dan Vapnik (1995), SVR menggunakan model linier yang dapat diterapkan pada kelas non-linear dengan mentransformasikan vektor input  $x$  ke dalam ruang fitur berdimensi tinggi. Tujuan regresi adalah menentukan fungsi  $f(x)$  untuk memprediksi target dengan akurat.

Pengembangan *Support Vector Regression* (SVR) yang awalnya diterapkan pada data independen ( $i = 1,2,\dots,m$ ), telah diperluas untuk data deret waktu ( $t = 1,2,\dots,m$ ), di mana urutan data tidak dapat diubah. Dengan demikian, hasil peramalan mencerminkan data yang mengikuti proses stokastik, sejalan dengan prinsip model regresi deret waktu. Misalkan terdapat  $m$  data pelatihan,  $(\mathbf{x}_t, \mathbf{y}_t)$ , di mana  $t = 1,\dots,m$  dengan data input  $\mathbf{x}_t = (x_{t1}, x_{t2}, \dots, x_{tt}, \dots, x_{tm})$ , di mana setiap  $\mathbf{x}_t$  memiliki dimensi  $n$ . Sementara itu, *output* yang sesuai dengan setiap input adalah  $\mathbf{y}_t = (y_{t1}, y_{t2}, \dots, y_{tt}, \dots, y_{tm})$ , ang merupakan nilai keluaran yang diharapkan. Model SVR kemudian dirumuskan sebagai berikut:

$$f(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + \mathbf{b} \quad (12)$$

Dalam konteks ini,  $\varphi(\mathbf{x})$  merepresentasikan ruang fitur berdimensi tinggi yang diperoleh dari transformasi non-linear terhadap ruang input  $\mathbf{x}$ . Koefisien  $\mathbf{w}$  and  $\mathbf{b}$  dihitung dengan meminimalkan fungsi regularisasi, sebagaimana dijelaskan dalam persamaan (13) dan (14).

$$R(C) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{n} \sum_{t=1}^n L_\varepsilon(\mathbf{y}_t, f(\mathbf{x}_t)) \quad (13)$$

$$L_\varepsilon(\mathbf{y}_t, f(\mathbf{x}_t)) = \begin{cases} |y_t - f(\mathbf{x}_t)| - \varepsilon, & |y_t - f(\mathbf{x}_t)| \geq \varepsilon \\ 0, & \text{lainnya} \end{cases} \quad (14)$$

Estimasi nilai  $\mathbf{w}$  dan  $\mathbf{b}$  diperoleh dengan persamaan (15) direformulasikan ke dalam bentuk primal dengan memasukkan variabel *slack* positif  $\xi$  dan  $\xi^*$  sebagaimana terlihat pada persamaan (15) dan (16). Hal ini memungkinkan masalah optimasi diformulasikan dengan meminimalkan nilai R sebagai berikut:

$$(\mathbf{w}, \xi^*) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{t=1}^n L_\varepsilon(\xi_t, \xi_t^*) \quad (15)$$

$$\text{Subject to} \begin{cases} \mathbf{y}_t - \mathbf{w}\varphi(\mathbf{x}_t) - \mathbf{b}_t \leq \varepsilon + \xi_t \\ \mathbf{w}\varphi(\mathbf{x}_t) - \mathbf{y}_t + \mathbf{b}_t \leq \varepsilon + \xi_t^* \\ \xi_t, \xi_t^* \geq 0 \end{cases} \quad (16)$$

dimana  $\frac{1}{2}\|\mathbf{w}\|^2$  merepresentasikan norma vektor bobot.  $y_t$  adalah nilai yang diharapkan, sedangkan  $C > 0$  merupakan konstanta regularisasi yang mengatur keseimbangan antara kesalahan empiris dan term regularisasi. Selain itu,  $\varepsilon$  mencerminkan tingkat akurasi dalam mendekati data pelatihan. Nilai  $\varepsilon$  yang kecil berhubungan dengan variabel slack  $\xi^*$  yang tinggi serta akurasi aproksimasi yang lebih baik dan sebaliknya. Penerapan metode Lagrange Multiplier dan kondisi dengan kondisi optimalitas, fungsi regresi dapat dinyatakan secara eksplisit didefinisikan pada persamaan 17 berikut:

$$f(\mathbf{x}) = \sum_{t=1}^n (a_t - a_t^*) K(\mathbf{x}, \mathbf{x}_t) + \mathbf{b} \quad (17)$$

dimana  $(a_t - a_t^*)$  merepresentasikan nilai Lagrange multipliers, yang memenuhi persamaan  $(a_t - a_t^*) = 0$ , dengan syarat  $a_t^* \geq 0$  dan  $a_t \geq 0$ , untuk  $t = 1, 2, \dots, n$ . Nilai-nilai ini diperoleh dengan memaksimalkan fungsi dual dalam persamaan (17), yang sesuai dengan format yang dijelaskan dalam persamaan (18).

$$R(a_t, a_t^*) = -\frac{1}{2} \sum_{t=1}^n \sum_{t'=1}^n (a_t - a_t^*) (a_{t'} - a_{t'}^*) K(\mathbf{x}_t, \mathbf{x}_{t'}) + \sum_{t=1}^n \mathbf{y}_t (a_t - a_t^*) - \varepsilon \sum_{t=1}^n (a_t - a_t^*) \quad (18)$$

dimana  $\sum_{t=1}^n (a_t - a_t^*) = 0; 0 \leq a_t \leq C, t = 1, 2, \dots, n; 0 \leq a_t^* \leq C, t = 1, 2, \dots, n$ .

Pada persamaan (18) dengan  $K(\mathbf{x}_t, \mathbf{x}_{t'})$  merupakan *kernel function*. Nilai kernel merepresentasikan hasil perkalian dalam antara dua vektor  $\mathbf{x}_t$  dan  $\mathbf{x}_{t'}$  dalam ruang fitur  $\varphi(\mathbf{x}_t)$  dan  $\varphi(\mathbf{x}_{t'})$ , yang dinyatakan sebagai  $K(\mathbf{x}_t, \mathbf{x}_{t'}) = \varphi(\mathbf{x}_t)\varphi(\mathbf{x}_{t'})^T$ . SVR merupakan algoritma *machine learning* yang dikembangkan oleh Vapnik berdasarkan teori pembelajaran statistik dan banyak digunakan dalam peramalan deret waktu (Datong, et al., 2009). Dalam konteks peramalan *time series*, SVR menunjukkan kemampuan generalisasi yang lebih baik dibandingkan metode pembelajaran lainnya. Namun, akurasi prediksi yang dihasilkan oleh model SVR sangat bergantung pada pemilihan parameter. Perbedaan dalam pemilihan parameter dapat menghasilkan keluaran model yang berbeda secara signifikan.

Kriteria pemilihan model terbaik diperlukan untuk memeriksa ketepatan suatu model *time series* dari model yang telah terbentuk. Berdasarkan residual data *in sample*, kriteria yang biasanya digunakan untuk memilih model terbaik adalah *Akaike's Information Criterion* (AIC) dengan rumus sebagai berikut (Wei, 2006).

$$\text{AIC}(M) = n \ln \sigma_a^2 + 2M \quad (19)$$

Sementara itu, kriteria pemilihan model didasarkan pada residual peramalan dari data *out sample* dengan menggunakan beberapa metode berikut:

$$\text{sMAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(y_t + \hat{y}_t)/2} \times 100 \quad (20)$$

Model terbaik adalah model yang memiliki nilai statistik paling mendekati nol atau yang paling kecil (Wei, 2006).

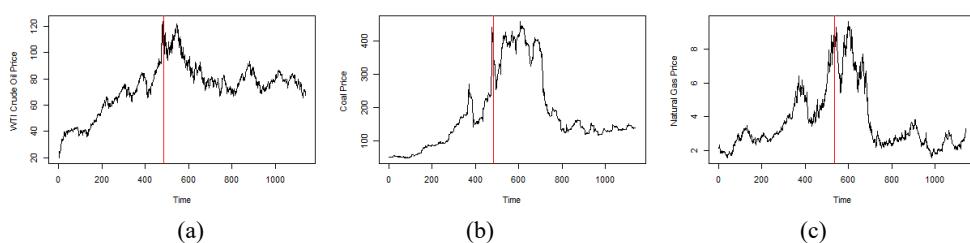
### 3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah dari harga minyak mentah jenis WTI, batu bara jenis Newcastle, dan gas alam mulai dari 30 April 2020 sampai dengan 27 September 2024. Data *in sample* yang digunakan yaitu data harga masing-masing komoditas energi dari 1 Mei 2020 hingga 31 Agustus 2024, dan sisanya merupakan data *out sample*. Data ini bersifat *opensource* yang diperoleh pada website [www.id.investing.com](http://www.id.investing.com). Periode waktu yang digunakan untuk setiap variabel adalah harian dengan jumlah pengamatan sebanyak 1121 pengamatan atau hari. Variabel *return* harga yang digunakan dalam analisis adalah hasil transformasi  $Z_t$  data harga komoditas menjadi bentuk  $R_t$ . Tahapan analisis pada penelitian ini yaitu sebagai berikut:

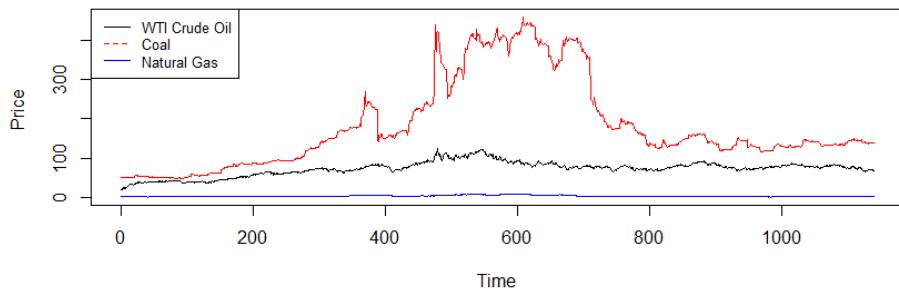
1. Melakukan analisis karakteristik data harga dan volatilitas masing-masing komoditas energi.
2. Melakukan pemodelan volatilitas komoditas energi dengan *asymmetric* GARCH.
3. Melakukan pemodelan volatilitas komoditas energi dengan GARCH-SVR.
4. Membandingkan model terbaik dengan kriteria evaluasi model untuk data *in sample* dan data *out sample* dan melakukan peramalan beberapa periode ke- $h$  dengan model terbaik.

#### 4. HASIL DAN PEMBAHASAN

Penelitian terdahulu berfokus untuk meramalkan volatilitas tanpa mempertimbangkan adanya efek long memory dan asimetrik yang ada pada data finansial yang berpotensi muncul pada pasar uang, sehingga dilakukan pemodelan dengan kointegrasi GARCH-SVR dan GARCH asli sebagai banding. Sebelum melakukan analisis maka kita perlu mengetahui karakteristik dari data *in sample* yaitu data harga dan *return* masing-masing komoditas energi dunia. Berikut ini adalah karakteristik harga masing-masing komoditas energi dunia yang ditunjukkan pada Gambar 1. Gambar 1 menunjukkan bahwa pada plot harga masing-masing komoditas energi dunia terdapat kecenderungan naik pada beberapa periode kemudian mengalami penurunan setelahnya. Hal ini menunjukkan bahwa harga masing-masing komoditas energi dunia terus meningkat dari waktu ke waktu. Kenaikan dan penurunan harga komoditas yang relatif stabil menunjukkan bahwa data harga komoditas energi stasioner pada variansnya, sementara untuk rata-ratanya tidak. Namun pada periode akhir tahun 2021 hingga tahun 2023 terlihat fluktuasi yang lebih tinggi dan kenaikan nilai yang lebih rendah dibandingkan periode sebelumnya. Fluktuasi dan kenaikan nilai ini kemungkinan disebabkan oleh faktor eksternal, yang akhirnya mempengaruhi harga dari komoditas energi dunia seperti minyak mentah, batu bara, maupun gas alam.

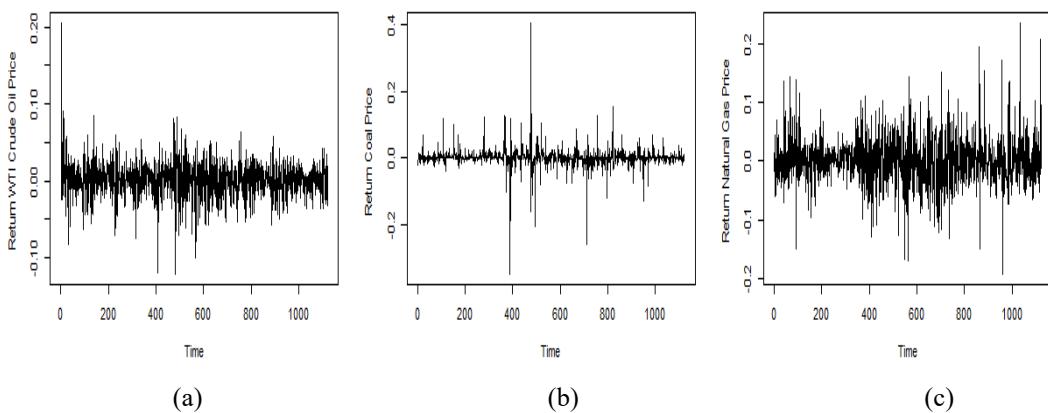


Gambar 1. *Time Series Plot* Harga Komoditas Energi (a) Minyak Bumi (b) Batu Bara (c) Gas Alam



Gambar 2. *Overlay Time Series Plot* Harga Komoditas Energi

Berdasarkan Gambar 2 di atas, diketahui bahwa pergerakan harga komoditas energi minyak mentah WTI, batu bara, dan gas alam memiliki pola yang hampir sama. Pola yang hampir sama pada Gambar 2 mengindikasikan bahwa kemungkinan ada hubungan kointegrasi antar komoditas energi satu dengan lainnya. Hal ini ditunjukkan pada saat invansi Rusia ke Ukraina, di mana harga minyak dunia melambung sebab Amerika Serikat melarang ekspor migas dari Rusia. Kenaikan harga minyak bumi kemudian diikuti dengan kenaikan harga gas alam. Pasalnya, kedua komoditas energi tersebut saling terhubung satu sama lain. Selain itu, harga batu bara juga mengalami kenaikan. Hal ini terjadi lantaran pasokan minyak bumi dan gas alam yang jadi salah satu sumber energi dunia terganggu. Akibatnya meningkatlah pasokan batu bara di tingkat global. Kondisi *return* masing-masing komoditas energi dapat digambarkan seperti pada Gambar 3. Terlihat bahwa data *return* masing-masing komoditas energi telah stasioner dalam mean karena cenderung berada di sekitar titik nol. Pada Gambar 3 menunjukkan bahwa untuk plot *return* harga untuk masing-masing komoditas energi membentuk volatilitas clustering yaitu periode di mana data *return* harga komoditas energi bernilai besar selama periode tertentu dan bernilai kecil untuk selama periode waktu yang lain. Terbentuknya volatility clustering ini menyebabkan munculnya kasus heteroskedastisitas pada data *return* komoditas energi sehingga digunakanlah metode ARCH/GARCH untuk mengatasinya efek dari heteroskedastisitas ini.



Gambar 3. Plot Time Series *Return* Komoditas Energi (a) Minyak Bumi (b) Batu Bara (c) Gas Alam

Identifikasi adanya heterokedastisitas dalam dilakukan dengan melakukan pengujian efek ARCH pada data *return* kuadrat masing-masing komoditas energi. Dihasilkan bahwa nilai  $LM > \chi^2_{0.05,q}$  untuk semua orde, sehingga volatilitas *return* harga masing-masing komoditas energi mengandung efek heteroskedastisitas. Keberadaan heteroskedastisitas pada data deret waktu dapat diatasi dengan melakukan pemodelan GARCH. Pemilihan model GARCH terbaik pada penelitian ini adalah dengan melihat nilai AIC terkecil. Berikut adalah hasil estimasi parameter model GARCH yang ditunjukkan pada Tabel 1.

Tabel 1 Estimasi Parameter Model GARCH( $p,q$ )

Variabel	Model	Parameter	Estimasi	P-value
Minyak Bumi	GARCH (1,1)	$\mu$	0,0015	0,0186
		$\omega$	0,0000	0,0009
		$\alpha_1$	0,1264	0,0000
		$\beta_1$	0,8373	0,0000
Batu Bara	GARCH (1,1)	$\mu$	0,0019	0,0078
		$\omega$	0,0000	0,0000
		$\alpha_1$	0,0562	0,0000
		$\beta_1$	0,9318	0,0000
Gas Alam	GARCH (1,1)	$\mu$	0,0024	0,0384
		$\omega$	0,0000	0,3342
		$\alpha_1$	0,0336	0,0000
		$\beta_1$	0,9538	0,0000

Model GARCH (1,1) untuk masing-masing *return* komoditas energi dapat digunakan analisis lebih lanjut karena parameter  $\alpha$  dan  $\beta$  signifikan pada taraf  $\alpha=0.05$ . Selanjutnya dilakukan pengujian untuk mengetahui pengaruh asimetris pada volatilitas harga komoditas energi. Berdasarkan hasil Tabel 2, model GARCH (1,1) pada komoditas batu bara memiliki nilai p-value untuk uji *positive size bias* sebesar 0,04282 sehingga keputusannya terdapat efek asimetris dan untuk mengatasinya perlu dilakukan analisis lanjutan dengan GARCH asimetris. Model GARCH asimetris terbaik dapat dilakukan melalui evaluasi kebaikan model dengan AIC sehingga diperoleh bahwa model asimetris terbaik untuk volatilitas harga batu bara yaitu model eGARCH (1,1).

Tabel 2 Estimasi Parameter Model eGARCH( $p,q$ )

Variabel	Model	Statistik	Estimasi	P-value
Batu Bara	e GARCH (1,1)	$\mu$	0,0022	0,0006
		$\omega$	-0,0500	0,0000
		$\alpha_1$	0,0347	0,0000
		$\beta_1$	0,9884	0,0000
		$\gamma_1$	0,1119	0,0000

Ketika terdapat hubungan *long memory* antar variabel, kombinasi linear dari variabel variabel tersebut menghasilkan residual yang mengikuti proses *long memory*. Hal ini menunjukkan adanya hubungan kointegrasi. Selain itu, estimasi parameter d dilakukan pada masing-masing volatilitas komoditas energi. Estimasi ini menggunakan uji Geweke dan Porter-Hudak (GPH) dengan bandwidth sebesar 0,5. Hasil estimasi parameter *long memory* ( $d$ ) *return* kuadrat untuk Minyak bumi, Batu bara, dan Gas Alam masing-masing sebesar 0,2477, 0,1607, and 0,2372. Hasil tersebut menunjukkan bahwa volatilitas *return* kuadrat untuk semua komoditas energi memiliki nilai parameter  $d$  yang terletak pada interval  $0 < d < 0,5$  yang artinya bahwa volatilitas *return* kuadrat harga masing-masing komoditas energi memiliki proses berkorelasi panjang yang stasioner dengan adanya ketergantungan positif antar volatilitas *return* kuadrat tersebut yang terpisah jauh. Untuk lebih meyakinkan, maka dilanjutkan dengan pengujian untuk mengidentifikasi apakah terdapat kointegrasi pada volatilitas masing-masing komoditas energi. Jika dua deret waktu memiliki tren dan pola yang serupa, maka keduanya dianggap terkointegrasi. Uji kointegrasi digunakan untuk menentukan apakah residual dari regresi bersifat stasioner, di mana residual yang stasioner menunjukkan adanya kointegrasi. Dengan demikian, hubungan antar deret waktu bersifat signifikan secara statistik dan bukan sekadar kebetulan. Untuk menguji hal ini, digunakan

Johansen Test, yang dapat mengukur apakah tiga atau lebih deret waktu memiliki hubungan kointegrasi. Hasil pengujian ditunjukkan pada Tabel 3 berikut.

Tabel 3 Hasil Johansen Test

<b>H<sub>0</sub></b>	<b>Test</b>	<b>10 pct</b>	<b>5 pct</b>	<b>1 pct</b>
$r \leq 2$	256,23	6,50	8,18	11,65
$r \leq 1$	517,77	15,66	17,95	23,52
$r = 0$	1021,59	28,71	31,52	37,22

Hipotesis nol menunjukkan tidak terdapat kointegrasi vektor maka hipotesis tersebut ditolak karena  $test(1021,59) > 28,71$ . Pengujian dengan hipotesis nol  $r \leq 2$  dan  $r \leq 1$  menunjukkan hasil bahwa Trace test lebih besar dari critical value sehingga kita kombinasi linear dari tiga deret waktu untuk membentuk deret stasioner. Hasil Johansen cointegration test menunjukkan bahwa terdapat hubungan kointegrasi jangka panjang antar volatilitas masing-masing komoditas energi.

Berdasarkan hasil pengujian sebelumnya, maka perlu dilakukan pemodelan cointegration salah satunya menggunakan *machine learning* yaitu SVR. Model terbaik dari GARCH pada Tabel 1 dan Tabel 2 di atas akan kita gunakan untuk melakukan peramalan dengan model SVR, namun sebelum itu kita harus menentukan model terbaik dari SVR untuk volatilitas harga masing-masing komoditas energi dunia. Dalam pemodelan dengan GARCH-Support Vector Regression, langkah pertama yang dilakukan adalah memodelkan data *in sample* dengan Support Vector Regression. Hasil dari model yang telah diperoleh, maka dilakukan peramalan untuk data *out sample* dengan menggunakan hasil peramalan pada GARCH (p,q) yang diperoleh sebelumnya. Berikut merupakan penjabaran dari analisis volatilitas dengan model GARCH-Support Vector Regression untuk volatilitas *return* kuadrat. Pemodelan Support Vector Regression (SVR) pada penelitian ini menggunakan metode grid search untuk menentukan parameter mana yang digunakan. Dalam penggunaannya, Grid Search Optimization biasanya diukur dengan *cross validation* pada data *in sample*. Oleh karena itu, disarankan untuk mencoba beberapa variasi pasangan parameter pada hyperplane. Hal ini bertujuan untuk mendapatkan hasil yang global optimum. Adapun parameter pada model SVR dengan kernel Radial Basis Function ada tiga yaitu parameter *epsilon*, *sigma*, dan *cost*, sehingga untuk mendapatkan parameter yang optimum perlu dilakukan kombinasi nilai dari ketiga parameter agar didapatkan nilai eror yang paling kecil. Rentang nilai parameter yang digunakan dalam pencarian parameter optimal pada tahap *loose grid* dapat dilihat pada Tabel 4.

Tabel 4. Batasan Nilai Range Parameter SVR

<b>Parameter</b>	<b>Nilai Range</b>
$C$	$2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5$
$\gamma$	$2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5$
$\varepsilon_1$	0,1; 0,2; 0,30; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1

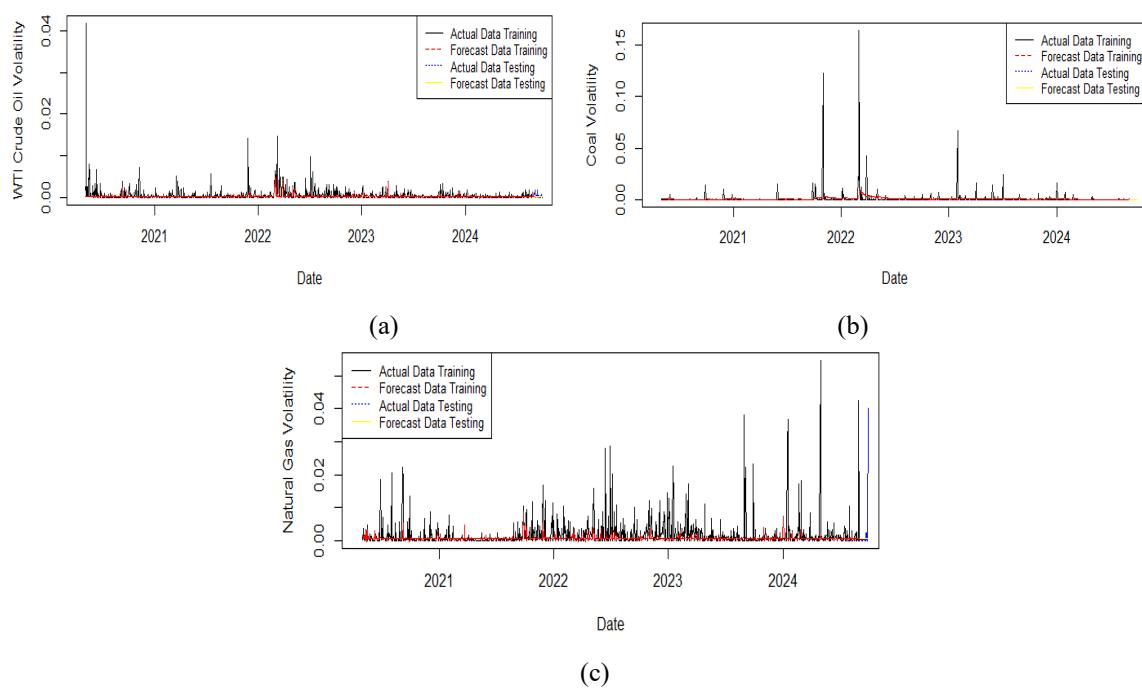
Berdasarkan kombinasi-kombinasi tersebut maka dihasilkan nilai RMSE terkecil untuk pemodelan data vaolatilitas *return* kuadrat. Parameter optimal berdasarkan hasil grid search untuk volatilitas *return* kuadrat minyak mentah dunia adalah  $C = 2$ ,  $\gamma = 8$ , dan  $\varepsilon = 0.2$ . Parameter optimal berdasarkan hasil grid search untuk batu bara adalah  $C = 32$ ,  $\gamma = 0,25$ , dan  $\varepsilon = 0,1$ . Parameter optimal berdasarkan hasil grid search untuk gas alam dunia adalah  $C = 8$ ,  $\gamma = 4$ , dan  $\varepsilon = 0.1$ . Setelah menemukan model terbaik SVR, maka langkah selanjutnya adalah membuat prediksi dengan menggunakan nilai hasil peramalan GARCH

(p,q) sebelumnya. Evaluasi kebaikan model untuk data *in sample* dan data *out sample* pada hasil peramalan GARCH dan GARCH-SVR dapat dilihat pada Tabel 5 berikut.

Tabel 5. Kriteria Kebaikan Model

Variabel	Model	Kriteria Kebaikan Model			
		<i>In sample</i>		<i>Out sample</i>	
		RMSE	sMAPE	RMSE	sMAPE
WTI Crude Oil	GARCH (1,1)	0,00166	106,82987	0,00055	91,82940
	<b>GARCH (1,1)-SVR</b>	<b>0,00168</b>	<b>106,97382</b>	<b>0,00054</b>	<b>87,23044</b>
Coal	eGARCH (1,1)	<b>0,00686</b>	<b>151,13164</b>	<b>0,00028</b>	<b>151,84267</b>
	eGARCH (1,1) - SVR	0,00691	150,78902	0,00062	162,96367
Natural Gas	GARCH (1,1)	0,00407	112,24719	0,00895	137,2553
	<b>GARCH (1,1)-SVR</b>	<b>0,00424</b>	<b>110,08744</b>	<b>0,00907</b>	<b>91,07440</b>

Pada Tabel 5 dilakukan perbandingan sMAPE dan RMSE pada data *in sample* masing-masing komoditas energi antara metode GARCH dan GARCH-Support Vector Regression. Pada komoditas energi minyak mentah WTI dan gas alam, model GARCH (1,1)-SVR memiliki nilai sMAPE dan RMSE lebih kecil pada data *out sample* jika dibandingkan dengan model GARCH (1,1). Sementara itu, pada komoditas batu bara, model yang memiliki nilai sMAPE dan RMSE untuk data *out sample* adalah model eGARCH (1,1).



Gambar 4. Hasil Peramalan dengan Model Terbaik untuk Volatilitas Harga (a) Minyak Bumi (b) Batu Bara (c) Gas Alam

Berdasarkan Gambar 4 diperoleh bahwa pola volatilitas harga hasil peramalan dengan model terbaik untuk masing-masing komoditas energi yang ditandai garis merah juga mengikuti pola aslinya yang ditunjukkan garis hitam. Pada komoditas energi minyak mentah WTI, batu bara, dan gas alam memiliki nilai peramalan volatilitas yang cenderung mengikuti nilai volatilitas aktual dari data *out sample*.

## 5. KESIMPULAN

Penelitian ini menganalisis volatilitas harga komoditas energi, termasuk minyak mentah, batu bara, dan gas alam, dengan mempertimbangkan efek *long memory* dan heteroskedastisitas. Model GARCH digunakan untuk mengatasi heteroskedastisitas, sementara GARCH-SVR digunakan untuk menangani efek *long memory*. Hasil pada *return* batu bara menunjukkan efek asimetris, sehingga diperlukan model *asymmetric* GARCH, dengan EGARCH(1,1) sebagai model terbaik. Dalam analisis data out-of-sample, GARCH lebih unggul dari GARCH-SVR pada volatilitas batu bara karena parameter *long memory* rendah ( $d < 0.2$ ). Namun, GARCH-SVR lebih baik untuk volatilitas minyak mentah dan gas alam, dengan parameter *long memory* yang lebih tinggi ( $d > 0.2$ ). Penelitian ini menyimpulkan bahwa GARCH-SVR lebih disarankan ketika data memiliki efek *long memory* non-stasioner ( $d > 0.2$ ) dan merekomendasikan penggunaannya di masa depan untuk data tersebut.

## DAFTAR PUSTAKA

- Abdullah, A., Kuswanto, H., Prastyo, D. & Suhartono, 2020. Modeling the volatility of gross domestic product based on the volatility of money supply, inflation, interest rate, and exchange rate using the fractional cointegration model. *Journal of Physics: Conference Series*.
- Ariefianto, M. D., 2012. *Ekonometrika : Esensi dan aplikasi dengan menggunakan* (Pertama ed.). Jakarta: Penerbit Erlangga.
- Bakshi, S. S., Jaiswal, R. K. & Jaiswal, R., 2021. Efficiency Check Using Cointegration and *Machine learning* approach: Crude Oil Futures Markets. *Procedia Computer Science*, pp. 304-311.
- Bollerslev, T., 1986. Generalized Autoregressive COnditional Heterocedasticity. *Journal of econometrics*, Volume 31, pp. 307-327.
- Caporin, M., 2002. FIGARCH models: stationarity, estimation methods and the identification problem. *Greta Working Paper*, pp. 1-58.
- Charles, A. & Darne, O., 2017. Forecasting crude-oil market volatility: Further evidence with jumps. *Energy Economics*, Volume 67, pp. 508-519.
- Chuanhe, S., Zhongwen, L. & Ying, L., 2013. Nonlinear Cointegration Analysis Based on Support Vector Machines and Its Application to the Connection between Financial Markets. *Computer Science and Engineering*, pp. 117-121.
- Cortes, C. & Vapnik, V., 1995. Support-vector networks. *Mach. Learn.*, 20(3), p. 273–297.
- Datong, L., Yu, P. & Xiyuan, P., 2009. Fault Prediction based on Time Series with Online Combined Kernel SVR Methods. Singapore, s.n.
- Engle, R., 1982. Autoregressive Conditional Heterocedasticity with Estimates of Variance of UK Infaltion. *Econometrica*, pp. 987-1008.
- Granger, C. J., 1981. "Some Properties of Time Series Data and Their Use in Econometric Model Specification. *Journal of Econometrics*,, pp. 121-130. Gujarati, D., 2004. *Basic Econometrics* (4th ed.). New York: The McGraw-Hill.
- Izati, P. P., Prastyo, D. D., & Akbar, M. S., 2024. Modeling the Volatility of World Energy Commodity Prices Using the GARCH-Fractional Cointegration Model. *Procedia Computer Science*, Volume 234, pp. 412-419.
- Lin, B. & Li, J., 2015. The spillover effect across natural gas and oil markets: Based on the VEC-MGRACH framework. *Applied Energy*, 155(6), pp. 229-241.
- Lu, C. J., 2009. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, Volume 47, pp. 115-125.

- Mandelbrot, B. B., 1963. The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4), pp. 394-419.
- Pindyck, R., 2004. Volatility in natural gas and oil markets, Cambridge: Massachusetts Institute of Technology.
- Ramirez, J. A., Soriano, A., Myriam, C. & Rodolfo, S., 2003. Evaluation of a mining project under the joint effect of commodity price and exchange rate uncertainties using real options valuation. *Physica A: Statistical Mechanics and Its Applications* , 332(C), pp. 583-596.
- Wei, W., 2006. Time Series Analysis Univariate and Multivariate Methods, Second Edition. New York: Temple University.