

Perbandingan Model Value-at-Risk (VaR) *Hybrid* GARCH-EVT dan Model Standar dalam Pengukuran Risiko Ekstrem pada Portofolio Saham Sektoral di Indonesia

Annisa Syalsabila*, Nur Ikhwana, Agung Tri Utomo, Lalu Ramzy Rahmanda, Zulkifli Rais

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan, Universitas Negeri Makassar

Keywords: Value-at-Risk (VaR), GARCH-EVT, Extreme Value Theory (EVT), Portofolio Optimization, Backtesting, Risk Management, Saham BEI

Abstract:

This study aims to construct an optimal portfolio and compare the accuracy of various Value-at-Risk (VaR) models in measuring the risk of stock portfolios in the Indonesia Stock Exchange (IDX). The optimal portfolio is formed using the Minimum Variance Portfolio (MVP) method based on 11 sector-representative stocks for the period 2019–2025. The risk performance of this portfolio is then evaluated using six VaR models: Variance–Covariance (VC), Historical Simulation (HS), Monte Carlo (MC), GARCH (1,1), Extreme Value Theory (EVT-GPD), and the hybrid GARCH–EVT model. Model accuracy is assessed through backtesting using the Kupiec Proportion of Failures (POF) test and the Christoffersen Conditional Coverage (CC) test at 95% and 99% confidence levels. The optimization results indicate that the MVP portfolio is dominated by defensive sectors such as consumer non-cyclicals (ICBP.JK) and large-cap banking (BBCA.JK). Back testing results show that although all models perform adequately at the 95% level, standard models (VC, MC, GARCH) fail to capture extreme risk at the 99% level. In contrast, the GARCH–EVT model satisfies the backtesting criteria and emerges as the most accurate and superior model for predicting extreme losses.

1. Pendahuluan

Manajemen risiko merupakan pilar utama dalam pengambilan keputusan investasi, khususnya pada pasar modal yang memiliki volatilitas tinggi seperti Bursa Efek Indonesia (BEI). Sebagai salah satu emerging market, BEI menawarkan peluang imbal hasil yang besar, namun di saat yang sama menyimpan risiko fluktuasi harga yang tajam akibat sentimen domestik maupun global. Dalam kondisi pasar yang dinamis seperti ini, pengukuran risiko yang akurat menjadi kebutuhan mendasar bagi investor dan manajer portofolio.

Model-model risiko konvensional, seperti pendekatan Variance–Covariance, umumnya dibangun di atas asumsi bahwa *return* aset mengikuti distribusi normal. Namun, berbagai penelitian telah menunjukkan bahwa data keuangan cenderung memperlihatkan karakteristik leptokurtosis dan *skewness*, yang menghasilkan fenomena *fat-tails* atau ekor gemuk pada distribusi *return*. Kondisi ini membuat kejadian ekstrem (misalnya gejolak pasar akibat pandemi COVID-19) terjadi jauh lebih sering dibanding prediksi distribusi normal. Ketidaktepatan asumsi ini menyebabkan metode seperti *Value-at-Risk* (VaR) tidak memperhitungkan potensi kerugian ekstrem (Peiris et al., 2024)

* Corresponding author.

E-mail address: annisa.syalsabila@unm.ac.id



Untuk menangkap sifat volatilitas pasar yang kompleks, model *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) telah menjadi standar dalam memodelkan volatilitas yang berubah seiring waktu. Namun, meskipun efektif untuk menangkap *volatility clustering*, model ini masih terbatas dalam memodelkan kejadian ekstrem pada ekor distribusi. Hal ini mendorong penggunaan *Extreme Value Theory* (EVT), yang secara khusus dirancang untuk memodelkan perilaku ekor distribusi. Penelitian terbaru mengonfirmasi bahwa integrasi keduanya atau model *hybrid GARCH-EVT* dapat meningkatkan akurasi VaR dan memberikan estimasi risiko ekstrem yang lebih realistis, terutama pada level kepercayaan tinggi (Kaibuchi et al., 2021)

Dalam konteks pasar Asia Tenggara, dinamika volatilitas semakin dipengaruhi oleh perbedaan institusional dan guncangan krisis. Analisis volatilitas pada pasar ASEAN menunjukkan bahwa periode krisis meningkatkan persistensi volatilitas dan ketebalan ekor distribusi, sehingga memperkuat urgensi penggunaan model risiko yang adaptif terhadap *tail events* (Yang, 2025). Temuan ini selaras dengan kondisi pasar Indonesia yang kerap dilanda ketidakpastian akibat faktor makroekonomi dan peristiwa global.

Penelitian dalam lingkup Indonesia juga memberikan bukti bahwa pendekatan GARCH-EVT relevan dan efektif untuk pasar Indonesia. Studi oleh Maulana, Dharmawan, dan Srinadi (2022) menunjukkan bahwa integrasi GJR-GARCH, EVT, dan copula mampu menghasilkan estimasi *Conditional Value-at-Risk* (CVaR) portofolio yang lebih konservatif dan akurat. Riset lain menggunakan TGARCH-GPD untuk mengukur VaR saham domestik juga menegaskan pentingnya menangkap *heavy tails* dan asimetri volatilitas dalam pengukuran risiko (Ariwibowo, 2020). Hal ini memperlihatkan bahwa pendekatan tradisional berbasis distribusi normal semakin tidak memadai untuk menggambarkan perilaku *return* di pasar saham Indonesia.

Berdasarkan kontribusi penelitian terdahulu tersebut, studi ini berupaya menguji secara komprehensif performa model GARCH-EVT dibandingkan model risiko standar dalam mengukur risiko ekstrem pada portofolio saham sektoral di BEI. Dengan membangun portofolio terdiversifikasi yang terdiri dari 11 saham perwakilan sektor utama, penelitian ini bertujuan menilai apakah kombinasi GARCH-EVT mampu memberikan pengukuran VaR yang lebih akurat serta lebih sensitif terhadap pergerakan ekstrem, sehingga dapat menjadi pendekatan unggulan dalam manajemen risiko portofolio di pasar berkembang seperti Indonesia.

2. Tinjauan Pustaka

2.1 Teori Portofolio Modern

Teori Portofolio Modern (*Modern Portfolio Theory*/MPT) pertama kali diperkenalkan oleh Harry Markowitz pada tahun 1952 dan menjadi fondasi utama dalam manajemen portofolio modern. Markowitz menunjukkan bahwa keputusan investasi tidak dapat dibuat dengan menilai suatu aset secara terpisah, melainkan harus mempertimbangkan kontribusi aset tersebut terhadap keseluruhan portofolio. Dengan kata lain, risiko yang relevan bagi investor adalah risiko portofolio secara keseluruhan, bukan risiko individual setiap aset. Hal ini disebabkan karena volatilitas portofolio turut dipengaruhi oleh kovarians antar aset di dalamnya, sehingga kombinasi aset dengan korelasi rendah dapat menurunkan risiko total portofolio secara signifikan (Markowitz, 1952).

Gagasan utama dalam MPT diwujudkan dalam konsep *Efficient Frontier*, yaitu himpunan portofolio yang memberikan tingkat *return* maksimum untuk suatu tingkat risiko tertentu, atau tingkat risiko minimum untuk suatu tingkat *return* tertentu. Portofolio yang berada di luar kurva ini dianggap tidak optimal karena ada kombinasi aset lain yang mampu memberikan *return* lebih tinggi pada tingkat risiko yang sama. Salah satu titik penting pada kurva *Efficient Frontier* adalah *Minimum Variance Portfolio* (MVP), yaitu portofolio dengan tingkat risiko terendah yang mungkin dicapai dari seluruh kombinasi aset yang tersedia. Pada penelitian ini, MVP digunakan sebagai acuan utama untuk melihat bagaimana diversifikasi dapat menurunkan risiko portofolio, terutama melalui struktur kovarians antar aset.

2.2. Pengukuran Risiko: *Value-at-Risk* (VaR) dan *Expected Shortfall* (ES)

Value-at-Risk (VaR) telah menjadi ukuran standar dalam manajemen risiko pasar sejak pertengahan 1990-an, terutama setelah diperkenalkan secara luas melalui kerangka Risk Metrics oleh J.P. Morgan (1996). VaR memberikan estimasi mengenai kerugian maksimum yang mungkin dialami suatu portofolio dalam rentang waktu tertentu dan pada tingkat kepercayaan $1 - \alpha$. Sebagai ilustrasi, VaR 99% sebesar 1 miliar rupiah menyatakan bahwa dengan keyakinan 99%,

kerugian portofolio pada hari berikutnya tidak akan melebihi nilai tersebut. Secara matematis, VaR pada tingkat kepercayaan $1 - \alpha$ didefinisikan sebagai kuantil dari distribusi kerugian portofolio L :

$$VaR_{1-\alpha} = \inf\{l \in R: P(L > l) \leq \alpha\} \quad (1)$$

Dimana $P(L > l)$ adalah probabilitas kerugian L melebihi nilai l (J.P. Morgan, 1996).

Popularitas VaR sebagian besar disebabkan oleh interpretasinya yang sederhana, namun ukuran ini memiliki kekurangan penting yaitu VaR tidak memberikan informasi mengenai besarnya kerugian ketika kerugian aktual telah melampaui batas VaR tersebut. Penelitian oleh Artzner et al. (1999) menunjukkan bahwa VaR tidak memenuhi sifat *subadditivity*, sehingga tidak selalu mencerminkan risiko secara koheren.

Untuk mengatasi keterbatasan ini, dikembangkan ukuran risiko alternatif yang dikenal sebagai *Expected Shortfall* (ES) atau *Conditional VaR*. ES mengukur rata-rata kerugian ketika kerugian sudah lebih buruk dari nilai VaR. Dengan kata lain, ES menilai perilaku *tail losses*, yaitu kerugian ekstrem yang ketika terjadi seringkali jauh lebih besar dibandingkan batas VaR itu sendiri. Secara matematis, ES pada tingkat kepercayaan $1 - \alpha$ didefinisikan sebagai:

$$ES_{1-\alpha} = E[L|L > VaR_{1-\alpha}] \quad (2)$$

(Acerbi & Tasche, 2002).

Dalam literatur manajemen risiko, ES dianggap sebagai ukuran yang dapat lebih diandalkan karena bersifat koheren (Artzner et al., 1999) dan sensitif terhadap *tail-risk*, sehingga lebih representatif untuk menilai risiko pasar yang bersifat ekstrem dibandingkan VaR.

2.3. Model-Model Value-at-Risk

2.3.1 Model Parametrik: Variance-Covariance (VC) dan Monte Carlo (MC)

Model Variance-Covariance (VC) merupakan pendekatan paling mendasar dalam estimasi VaR. Model ini berasumsi bahwa *return* aset mengikuti distribusi normal multivariat dan korelasi antar aset bersifat linear. Dalam kerangka ini, VaR dihitung secara analitis menggunakan parameter rata-rata (μ) dan standar deviasi (σ) dari portofolio.

Secara matematis, VaR untuk tingkat kepercayaan $(1 - \alpha)$ dirumuskan sebagai:

$$VaR_{VC}^\alpha = W_0(\hat{\mu}_p - z_\alpha \hat{\sigma}_p) \quad (3)$$

Dimana W_0 adalah nilai awal portofolio, $\hat{\sigma}_p$ adalah volatilitas portofolio, dan z_α adalah nilai invers dari fungsi distribusi kumulatif normal standar pada tingkat signifikansi α (Jorion, 2007).

Meskipun model ini menawarkan kemudahan komputasi, model ini memiliki kelemahan signifikan karena ketergantungannya pada asumsi normalitas. Pasar keuangan sering menunjukkan sifat *leptokurtic* (*fat-tails*) dan *skewness* yang tidak dapat ditangkap oleh distribusi normal, sehingga VC cenderung meremehkan risiko saat kondisi pasar ekstrem (Linsmeier & Pearson, 1996).

Meskipun Model Variance-Covariance (VC) menawarkan efisiensi komputasi yang tinggi, kelemahan mendasarnya terletak pada sifat analitisnya yang kaku. Ketika portofolio mengandung instrumen dengan *payoff* non-linear, atau memiliki struktur yang kompleks dengan banyak faktor risiko, asumsi linearitas VC menjadi tidak valid dan memerlukan aproksimasi (*Delta-Approximation*) yang seringkali mengorbankan akurasi. Untuk mengatasi keterbatasan ini, terutama dalam mempertahankan asumsi distribusi (parametrik) namun sekaligus mengakomodasi non-linearitas maka dikembangkanlah Model Monte Carlo Simulation (MCS) yang menggunakan pendekatan numerik stokastik.

Model Monte Carlo Simulation (MCS) dalam implementasi VaR sering menggunakan asumsi distribusi normal untuk *return* portofolio, menjadikannya metode parametrik berbasis simulasi. Perlu dicatat bahwa MCS adalah metode yang sangat fleksibel dan tidak terbatas pada normalitas, dengan kata lain MCS dapat menggunakan distribusi lain (seperti t-Student atau distribusi empiris) untuk memodelkan *fat-tails*. Namun, dalam konteks perbandingan dengan metode estimasi dasar, pendekatan ini (dengan asumsi normalitas) secara konseptual mirip dengan Variance-Covariance (VC)

karena sama-sama didasarkan pada parameter (μ dan σ), tetapi MCS menggunakan ribuan skenario simulasi untuk membangun distribusi hasil, bukan solusi analitis.

Mekanisme inti dari pendekatan Simulasi Monte Carlo adalah menghasilkan berbagai skenario *return* portofolio fiktif ($\tilde{r}_{p,i}$) yang bersifat acak, namun tetap mencerminkan parameter yang telah diestimasi dari data historis ($\hat{\mu}_p$ dan $\hat{\sigma}_p$). Proses ini dilakukan dalam dua tahap. Pertama, dibangkitkan sejumlah besar (N kali) variabel acak (ϵ_i) yang mengikuti distribusi Normal Standar $N(0,1)$:

$$\epsilon_i \sim N(0,1)$$

Selanjutnya, variabel acak tersebut digunakan untuk menghitung *return* simulasi portofolio $\tilde{r}_{p,i}$ dengan menerapkan parameter estimasi ($\hat{\mu}_p$ dan $\hat{\sigma}_p^2$) yang diperoleh dari data historis:

$$\tilde{r}_{p,i} = \hat{\mu}_p + \hat{\sigma}_p \cdot \epsilon_i \quad (4)$$

Dalam konteks VaR sederhana, *return* simulasi portofolio ($\tilde{r}_{p,i}$) dibangkitkan secara langsung, dimana $\hat{\mu}_p$ adalah rata-rata portofolio estimasi dan $\hat{\sigma}_p$ adalah standar deviasi portofolio estimasi. Kuantil dari distribusi empiris $\tilde{r}_{p,i}$ kemudian digunakan untuk mencari nilai VaR.

Setelah simulasi (misal N skenario) dijalankan, VaR pada tingkat kepercayaan α dihitung sebagai kuantil dari distribusi kerugian yang dihasilkan:

$$VaR_{MC}^{\alpha} = -\text{Quantile}(\tilde{r}_{p,i}, \alpha) \quad (5)$$

Selain VaR, simulasi ini juga memungkinkan perhitungan *Expected Shortfall* (ES) secara akurat, yaitu rata-rata kerugian yang melebihi batas VaR. Kelebihan utama MCS dalam implementasi ini adalah kemampuannya untuk mengestimasi VaR dan ES secara numerik dengan presisi tinggi dan fleksibilitas untuk non-linearitas. Namun, kelemahan mendasarnya dalam studi yang menggunakan asumsi Normalitas tetap pada risiko estimasi yang terlalu rendah (*underestimation*) pada kondisi pasar ekstrem (Glasserman, 2004).

2.3.2. Model Non-Parametrik: Historical Simulation (HS)

Model *Historical Simulation* (HS) menghindari asumsi distribusi teoretis maupun estimasi parameter. Model ini murni menggunakan distribusi empiris dari data masa lalu. VaR pada tingkat kepercayaan α ditentukan secara langsung dari kuantil data historis:

$$VaR_{\alpha} = -\text{Quantile}(r_t, \alpha) \quad (4)$$

Kelebihannya, HS secara otomatis menangkap *fat-tails* dan korelasi yang ada di masa lalu. Kelemahannya, model ini bersifat statis dan berasumsi bahwa pola masa depan akan sama persis dengan masa lalu, serta lambat beradaptasi terhadap perubahan volatilitas pasar (Dowd, 2006).

2.3.3. Model Volatilitas Kondisional: GARCH

Model *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH), yang dikembangkan oleh Bollerslev (1986), dirancang untuk menangani *volatility clustering* dimana periode volatilitas tinggi cenderung berkelompok. Model GARCH(1,1) memodelkan varians kondisional (σ_t^2) sebagai fungsi dinamis:

$$\sigma_t^2 = \omega_0 + \gamma \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (5)$$

Formula ini menghasilkan estimasi volatilitas yang berubah setiap hari, jauh lebih realistis daripada asumsi volatilitas konstan.

2.3.4. Extreme Value Theory (EVT)

Extreme Value Theory (EVT) berfokus secara eksklusif pada pemodelan kejadian langka di bagian "ekor" distribusi, bukan pusat distribusi. Dengan metode *Peaks-Over-Threshold* (POT), kerugian yang melebihi ambang batas (*threshold*) dimodelkan menggunakan *Generalized Pareto Distribution* (GPD):

$$G_{\xi,\beta}(y) = 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-1/\xi} \quad (6)$$

EVT dianggap sebagai metode standar strategis untuk mengukur risiko ekstrem (Embrechts et al., 1997).

2.3.5. Model Hybrid GARCH-EVT

Model *hybrid* GARCH-EVT, yang dipopulerkan oleh McNeil dan Frey (2000), dianggap sebagai salah satu pendekatan paling canggih karena menggabungkan kekuatan pemodelan volatilitas dinamis dari GARCH dan pemodelan ekor ekstrem dari EVT. Proses ini dilakukan dalam tiga tahap:

1. Filtrasi GARCH: Sebuah model GARCH(1,1) difitkan ke data *return* (r_t) untuk menangkap *volatility clustering*. Hasilnya adalah residual standar ($\hat{\epsilon}_t$) yang independen dan identik (*i.i.d*):

$$\hat{\epsilon}_t = \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t} \quad (7)$$

Dimana $\hat{\mu}$ adalah *mean* kondisional dan $\hat{\sigma}_t$ adalah volatilitas kondisional (varians σ_t^2 dihitung menggunakan formula GARCH).

2. Pemodelan Ekor EVT: Model EVT (POT-GPD) kemudian diterapkan pada ekor dari residual standar ($\hat{\epsilon}_t$) untuk mengestimasi kuantil ekornya, $q_{\hat{\epsilon}}(\alpha)$, pada tingkat kepercayaan α .
3. Estimasi VaR Final (Kondisional): VaR satu periode ke depan (VaR_{t+1}) dihitung dengan menggabungkan kuantil ekor dari EVT ($q_{\hat{\epsilon}}$) dengan estimasi volatilitas dinamis dari GARCH (σ_{t+1}):

$$VaR_{t+1}^{\alpha} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} q_{\hat{\epsilon}}(\alpha) \quad (8)$$

Pendekatan *hybrid* ini dianggap paling akurat karena responsif terhadap kondisi pasar terkini (melalui σ_{t+1}) sekaligus mampu mengantisipasi kejadian ekstrem (melalui $q_{\hat{\epsilon}}$) dengan presisi yang lebih tinggi daripada model tunggal (McNeil & Frey, 2000).

2.4. Backtesting VaR

Pada pemodelan risiko, membuat model VaR saja tidak cukup, sehingga perlu dilakukan uji akurasi. Backtesting adalah proses untuk memvalidasi model VaR dengan membandingkan prediksi VaR masa lalu dengan *return* aktual (r_t) yang terjadi (Dowd, 2006). Proses ini mengidentifikasi model yang cenderung meremehkan risiko ekstrem (terlalu banyak pelanggaran) atau melebih-lebihkan (terlalu sedikit pelanggaran) risiko ekstrem.

2.4.1. Uji Proporsi Kegagalan

Uji Proporsi Kegagalan, yang dikembangkan oleh Kupiec (1995), adalah uji *backtesting* yang paling mendasar. Uji ini mengukur apakah jumlah pelanggaran (*exceptions*), yaitu hari dimana kerugian aktual melebihi VaR, sesuai dengan tingkat kegagalan yang diharapkan (α). Misalnya, untuk $VaR_{99\%}$ selama $T = 500$ hari, kita mengharapkan tingkat kegagalan $\alpha = 1$, atau $N_{harapan} = 5$ pelanggaran.

Statistik uji Kupiec menggunakan *Likelihood Ratio* (LR_{POF}), yang menguji hipotesis nol (H_0) bahwa proporsi kegagalan aktual ($\hat{\pi}$) sama dengan proporsi kegagalan yang diharapkan (α).

$$LR_{POF} = -2 \ln \left[\frac{(1 - \alpha)^{T-N} \alpha^N}{(1 - \hat{\pi})^{T-N} \hat{\pi}^N} \right] \quad (9)$$

Dimana:

- T = Jumlah total observasi (hari).
- N = Jumlah pelanggaran aktual.
- α = Tingkat kegagalan yang diharapkan (misalnya, 0.01 untuk VaR_{99%}).
- $\hat{\pi} = N/T$ = Proporsi kegagalan aktual.

LR_{POF} berdistribusi Chi-square (χ^2) dengan satu derajat kebebasan. Jika LR_{POF} melebihi nilai kritis, maka H_0 ditolak, dan model dianggap tidak akurat karena proporsi kegagalannya salah (Kupiec, 1995).

2.4.2. Uji Independensi

Sebuah model VaR yang baik tidak hanya harus memiliki jumlah pelanggaran yang benar, tetapi pelanggaran tersebut juga harus terjadi secara acak dan independen. Jika pelanggaran terjadi secara bergerombol (*volatility clustering*), model tersebut gagal beradaptasi dengan kondisi pasar.

Uji Independensi (Independence Test). Uji Christoffersen (1998) menguji hipotesis bahwa terjadinya pelanggaran pada hari tertentu bersifat independen dari pelanggaran hari sebelumnya. Statistik uji *Likelihood Ratio* untuk independensi (LR_{IND}) adalah:

$$LR_{IND} = -2 \ln \left[\frac{(1 - \hat{\pi})^N (1 - \hat{\pi})^{T-N}}{(1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}}} \right] \quad (10)$$

Dimana π_{ij} adalah probabilitas terjadinya status j (0: tidak ada pelanggaran, 1: pelanggaran) pada hari t diberikan status i pada hari $t - 1$, dan T_{ij} adalah jumlah observasi dari transisi status tersebut (Christoffersen, 1998).

Uji Cakupan Bersyarat (Conditional Coverage Test). Uji Christoffersen yang paling ketat adalah Uji Cakupan Bersyarat (LR_{CC}), yang menggabungkan Uji Proporsi Kegagalan (LR_{POF}) dan Uji Independensi (LR_{IND}):

$$LR_{CC} = LR_{POF} + LR_{IND} \quad (13)$$

LR_{CC} berdistribusi χ^2 dengan dua derajat kebebasan. Model yang lolos uji LR_{CC} dianggap akurat (memiliki jumlah pelanggaran yang benar) dan robust (pelanggaran terjadi secara independen), yang merupakan standar validasi model VaR (Christoffersen, 1998).

3. Metode Penelitian

Penelitian ini menggunakan merupakan data sekunder yang diperoleh dari database public Yahoo Finance. Periode data mencakup 1 Januari 2019 hingga 26 November 2025. Periode ini dipilih untuk mencakup kondisi pasar yang beragam termasuk volatilitas tinggi selama pandemi COVID-19. Data yang digunakan merupakan data harga penutupan harian yang disesuaikan (*Adjusted Close Price*) untuk 11 saham di Bursa Efek Indonesia (BEI). Saham terpilih mewakili 11 sektor utama BEI berdasarkan kapitalisasi pasar terbesar dengan data historis memadai. Saham-saham tersebut adalah : BYAN.JK (*Energy*), TPIA.JK (*Basic Materials*), ASII.JK (*Industrials*), ICBP.JK (*Consumer Non-Cyclicals*), MAPI.JK (*Consumer Cyclicals*), KLBK.JK (*Health*), BBKA.JK (*Financials*), BSDE.JK (*Properties & Real Estate*), EMTK.JK (*Technology*), TLKM.JK (*Infrastructure*), dan TMAS.JK (*Transportation & Logistics*).

3.1. Pembagian Sampel

Data dibagi menjadi dua periode:

- Periode *In-Sample*: Data historis yang digunakan untuk mengestimasi parameter model pada setiap iterasi *rolling window*.
- Periode *Out-of-Sample*: Terdiri dari $T = 500$ hari terakhir, yang digunakan murni untuk pengujian (*backtesting*). Semua prediksi VaR akan dihasilkan untuk periode ini.

3.2. Prosedur Analisis Data

Prosedur analisis data dilakukan secara berurutan, dimulai dari pengolahan data mentah hingga pengujian akurasi model.

3.2.1 Perhitungan *Logarithmic Return*

Data harga penutupan harian dikonversi menjadi data *logarithmic return* harian (r_t). Log *return* dipilih karena memiliki sifat aditif waktu (*time-additivity*) dan diasumsikan lebih mendekati distribusi normal, sehingga mendukung asumsi stasioneritas yang diperlukan untuk pemodelan deret waktu (GARCH). Konversi dilakukan menggunakan fungsi R `ROC(prices, type="continuous")`.

3.2.2. Optimisasi Portofolio (*Minimum Variance Portfolio/MVP*)

Tahap ini bertujuan untuk menentukan struktur bobot portofolio yang paling efisien dalam hal risiko, yaitu MVP, yang kemudian digunakan sebagai basis perhitungan risiko:

- Matriks kovarians (Σ) dihitung dari data *log return* 11 saham.
- Optimisasi kuadratik dijalankan menggunakan fungsi `solve.QP` untuk menemukan vektor bobot optimal (w_{opt}) yang meminimalkan varians portofolio. Kemudian untuk kendala yaitu bobot non-negatif ($w_i \geq 0$) dan total bobot sama dengan satu ($\sum w = 1$).
- Simulasi 10.000 portofolio acak dilakukan untuk memvisualisasikan *Efficient Frontier* dan memvalidasi posisi MVP.
- Dibuat *return* portofolio harian (r_p) dengan mengalikan matriks *return* saham dengan vektor bobot optimal (w_{opt}).

3.2.3. Estimasi *Value-at-Risk* dan *Expected Shortfall*

Estimasi Var^α dan ES^α dilakukan untuk tingkat kepercayaan 95% dan 99% menggunakan metode *rolling forecast* satu hari ke depan ($h = 1$) selama periode *out-of-sample* (500 hari). Ukuran jendela bergulir (*rolling window*) ditetapkan sebesar 250 hari (setara dengan kira-kira satu tahun perdagangan).

A. Model Parametrik (VC dan MC)

Estimasi *mean* ($\hat{\mu}_p$) dan volatilitas ($\hat{\sigma}_p$) dihitung dari data *return* dalam jendela bergulir (250 hari).

- VC: VaR dihitung secara analitis menggunakan z_α (kuantil normal).
- MC: 100.000 simulasi dijalankan menggunakan `rnorm` untuk membangkitkan $\epsilon_i \sim N(0, 1)$ yang kemudian ditransformasi dengan $\hat{\mu}_p$ dan $\hat{\sigma}_p$ sehingga didapatkan kuantil VaR dan ES.

B. Model Non-Parametrik (HS)

VaR dihitung sebagai kuantil empiris langsung dari data *return* dalam jendela bergulir (250 hari).

C. Model Volatilitas dan Ekor (GARCH, EVT, GARCH-EVT)

- GARCH(1,1): Model GARCH difitkan pada jendela bergulir dengan asumsi residual berdistribusi Student-t untuk memprediksi volatilitas kondisional $\hat{\sigma}_{t+1}$. VaR dihitung dengan mengalikan $\hat{\sigma}_{t+1}$ dengan kuantil Student-t (t_α) pada derajat kebebasan estimasi ($\hat{\nu}$).
- EVT: Model *Generalized Pareto Distribution* (GPD) difitkan pada *tail* data *return*. Ambang batas (*threshold*) u ditetapkan pada persentil ke-95 dari kerugian (*loss*) *in-sample*.
- GARCH-EVT (Model *Hybrid*): Model GARCH(1,1) digunakan untuk mendapatkan residual standar ($\hat{\epsilon}_t$), kemudian EVT diterapkan pada ekor residual standar ($\hat{\epsilon}_t$) untuk mendapatkan kuantil ekor ($q_{\hat{\epsilon}}$), yang kemudian dikalikan dengan $\hat{\sigma}_{t+1}$ GARCH untuk menghasilkan prediksi Var_{t+1}^α .

3.2.4. Uji Validasi Model (Backtesting)

Prediksi VaR harian dari keenam model (untuk tingkat 95% dan 99%) dibandingkan dengan *return* aktual yang terjadi pada periode *out-of-sample* (500 hari). Pengujian dilakukan menggunakan Statistik *Likelihood Ratio* untuk menilai keakuratan model. Uji yang Digunakan: Uji Kupiec (Proporsi Kegagalan/ LR_{POF}) dan Uji Christoffersen (Cakupan Bersyarat/ LR_{CC}) menggunakan fungsi VaRTest dari *package* rugarch.

Hipotesis:

H_0 : Model akurat (jumlah pelanggaran dan independensi pelanggaran sesuai harapan).

H_1 : Model tidak akurat.

Kriteria Kelulusan: Model dinyatakan "Lulus" jika $p - value > 0.05$ pada tingkat signifikansi 5%. Hal ini berarti tidak ada bukti statistik yang cukup untuk menolak keakuratan model.

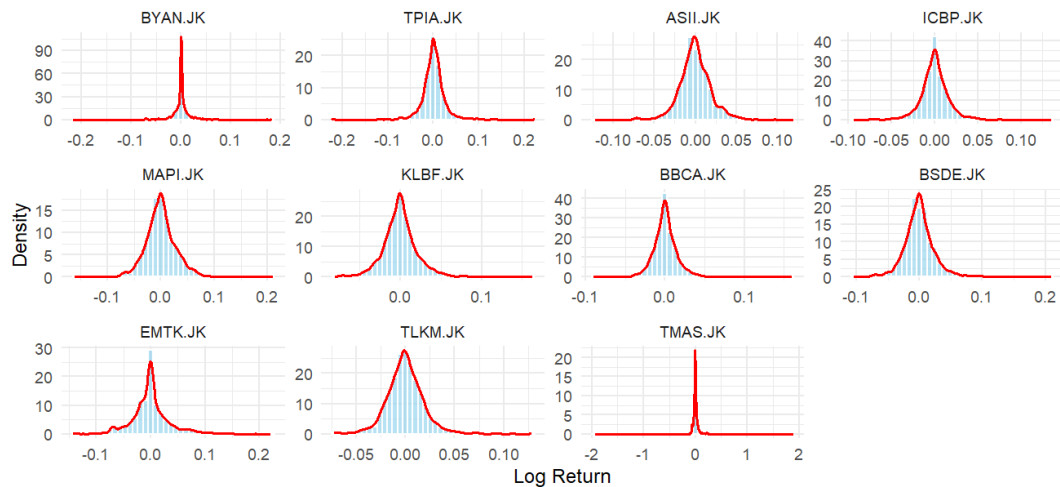
4. Hasil dan Pembahasan

4.1. Analisis Deskriptif.

Statistika deskriptif log *return* dari sebelas saham menunjukkan karakteristik yang konsisten dengan pola umum data keuangan harian. Nilai rata-rata (*mean*) seluruh saham berada sangat dekat dengan nol, yang mencerminkan bahwa perubahan harga harian relatif kecil dibandingkan volatilitasnya. Namun, terdapat perbedaan yang jelas pada magnitudo volatilitas antar sektor. Saham berbasis komoditas dan sektor siklikal seperti BYAN.JK, BSDE.JK, dan TMAS.JK menunjukkan standar deviasi yang secara substansial lebih tinggi dibandingkan saham defensif seperti ICBP.JK atau KLBF.JK. Hal ini mengindikasikan bahwa kontribusi risiko tiap aset dalam portofolio akan sangat heterogen, sehingga struktur korelasi dan proporsi bobot akan memainkan peran penting dalam pembentukan portofolio efisien.

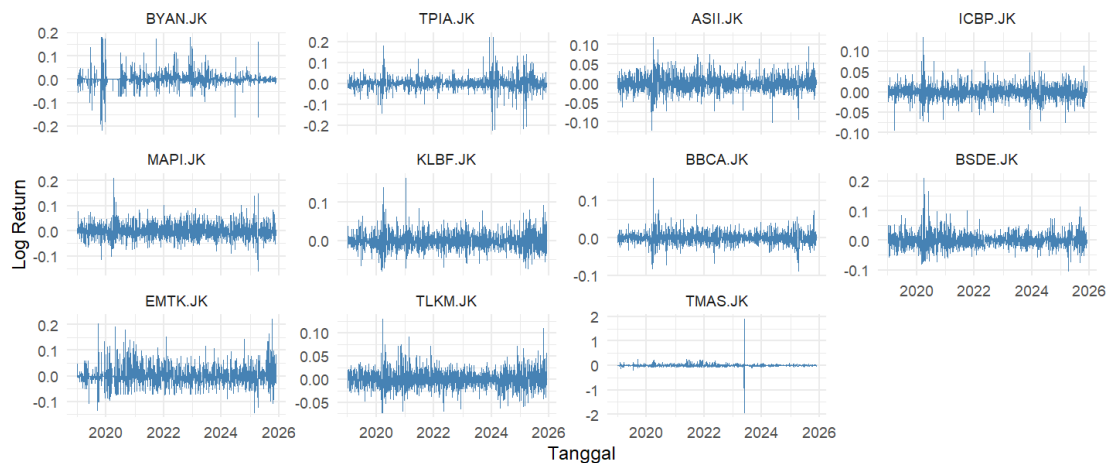
Tabel 1. Statistik Deskriptif Log-Return Harian

Saham	Min	Q1	Median	Mean	Q3	Max
BYAN.JK	-0.2162	-0.0040	0.0000	0.001291	0.004158	0.181217
TPJA.JK	-0.2231	-0.0107	0.0000	0.000939	0.011628	0.223144
ASII.JK	-0.1216	-0.0107	0.0000	-0.000145	0.010444	0.119623
ICBP.JK	-0.0939	-0.0086	0.0000	-0.000118	0.008048	0.135036
MAPI.JK	-0.1603	-0.0161	0.0000	0.000281	0.014467	0.209956
KLBF.JK	-0.0800	-0.0119	0.0000	-0.000126	0.009792	0.163188
BBCA.JK	-0.0892	-0.0080	0.0000	0.000293	0.008060	0.159849
BSDE.JK	-0.1047	-0.0136	0.0000	-0.000172	0.010989	0.209721
EMTK.JK	-0.1425	-0.0175	0.0000	0.000239	0.013187	0.220750
TLKM.JK	-0.0722	-0.0108	0.0000	-0.000016	0.010230	0.128700
TMAS.JK	-1.9289	-0.0155	0.0000	0.001242	0.012121	1.887839



Gambar 1. Distribusi Log-Return Harian untuk 11 Saham Sektoral BEI Beserta Kurva *Density*

Temuan tersebut selaras dengan pola pada histogram log *return*, yang menunjukkan puncak distribusi yang sangat tajam di sekitar nol serta ekor distribusi yang panjang, terutama pada saham-saham dengan volatilitas tinggi. Kurva *density* menegaskan bentuk distribusi yang runcing dan berekor panjang, yang secara visual mengonfirmasi tingginya kurtosis. Grafik *line chart* log *return* juga memperlihatkan dinamika volatilitas yang fluktuatif, termasuk episode volatilitas ekstrem pada saham seperti TMAS.JK dan BYAN.JK. Pola tersebut mencerminkan perilaku *volatility clustering* yang umum terjadi pada *return* pasar modal.

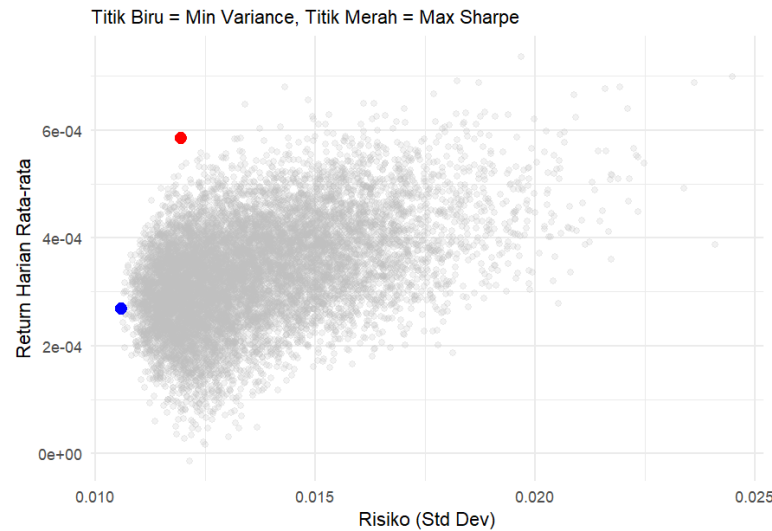


Gambar 2. Line chart Log-Return Harian untuk 11 Saham Sektoral BEI.

Secara keseluruhan, hasil deskriptif menunjukkan bahwa data *return* yang digunakan bersifat non-normal, ber-volatilitas heterogen, dan memiliki risiko ekor yang signifikan. Karakteristik ini mendukung pemilihan kerangka analisis yang sensitif terhadap risiko ekstrem, termasuk penggunaan ukuran risiko seperti *Value-at-Risk* (VaR) dan *Expected Shortfall* (ES), serta justifikasi teoritis bagi pentingnya diversifikasi dalam konteks Teori Portofolio Modern.

4.2. Hasil Optimisasi Portofolio

Analisis portofolio dilakukan dengan pendekatan mean–variance Markowitz untuk memetakan kombinasi risiko dan imbal hasil dari sebelas saham sektoral. Hasil komputasi menghasilkan *efficient frontier*, yaitu himpunan portofolio yang memberikan *return* maksimum untuk tingkat risiko tertentu atau, secara ekuivalen, risiko minimum untuk tingkat *return* tertentu.



Gambar 3. *Effect Frontier* 11 Saham Sektoral BEI

Gambar 3 menunjukkan *efficient frontier* yang diperoleh dari simulasi kombinasi bobot pada 11 saham sektoral. Portofolio minimum-varians (titik biru) menempati posisi risiko terendah di frontier, sedangkan portofolio maksimum-Sharpe (titik merah) menawarkan trade-off terbaik antara imbal hasil dan risiko. Bobot portofolio minimum-varians (Tabel Y) menempatkan porsi terbesar pada ICBP.JK ($\approx 21.7\%$) dan BBKA.JK ($\approx 19.5\%$), yang mengindikasikan peran penting kedua saham tersebut dalam mengurangi varians portofolio, baik melalui volatilitas relatif rendah maupun korelasi yang menguntungkan. Karena bobot-optimal sangat bergantung pada estimasi kovarians dan *return* historis.

Tabel 2. Optimisasi Bobot Portofolio 11 Saham Sektoral BEI

Saham	Bobot (w_{opt})	Bobot Optimal (%)
ICBP.JK	0.217005	21.70%
BBKA.JK	0.194539	19.45%
BYAN.JK	0.144859	14.49%
TLKM.JK	0.098355	9.84%
KLBF.JK	0.074749	7.47%
ASII.JK	0.074418	7.44%
TPIA.JK	0.067018	6.70%
EMTK.JK	0.062123	6.21%
BSDE.JK	0.027533	2.75%
MAPI.JK	0.027308	2.73%
TMAS.JK	0.012094	1.21%
Total	1	100%

Optimisasi lebih lanjut menghasilkan bobot portofolio minimum-variance dengan pembatasan tanpa *short selling* dan jumlah bobot sama dengan satu. Hasil estimasi menunjukkan bahwa aset dengan volatilitas lebih rendah dan korelasi yang menguntungkan terhadap aset lain memperoleh porsi investasi lebih besar. Dalam konteks data ini, alokasi tertinggi diberikan kepada ICBP.JK (21,70%) dan BBKA.JK (19,57%), diikuti BYAN.JK (14,47%) serta TLKM.JK (9,84%). Komposisi tersebut mencerminkan dominasi saham-saham defensif dan berkapitalisasi besar dalam upaya menekan varians portofolio. Sebaliknya, saham dengan volatilitas relatif tinggi atau kontribusi diversifikasi yang lebih kecil memperoleh bobot yang rendah. Secara keseluruhan, struktur bobot ini konsisten dengan karakter portofolio minimum-variance, yakni mengedepankan stabilitas imbal hasil dan pengurangan risiko melalui diversifikasi yang efektif.

4.3. Hasil Backtesting VaR

Hasil *backtesting* terhadap berbagai metode pengukuran risiko pada tingkat signifikansi 1% dan 5% menunjukkan adanya divergensi performa yang substansial, terutama terkait kemampuan model dalam menangkap risiko ekstrem (tail risk). Evaluasi pada Tingkat Signifikansi 1% Pada kuantil ekstrem ($\alpha = 0.01$), kelemahan mendasar dari model parametrik berbasis distribusi normal terlihat jelas. Metode Variance-Covariance, Monte Carlo, dan GARCH (*rolling*) gagal memenuhi kriteria validitas uji Kupiec, dengan mencatatkan jumlah pelampauan (*exceedances*) sebanyak 11 kali. Frekuensi kegagalan yang tinggi ini mengindikasikan bahwa model-model tersebut secara sistematis melakukan *underestimation* terhadap risiko pasar. Kegagalan ini dapat dihubungkan pada ketidakmampuan asumsi normalitas dalam mengakomodasi fenomena *fat tails* (ekor gemuk) yang merupakan karakteristik umum data keuangan. Sebaliknya, pendekatan yang dirancang untuk menangkap perilaku ekor distribusi menunjukkan kinerja yang superior.

Tabel 3. Hasil Uji Backtesting dengan Kupiec Test dan Uji Christoffersen

α	Metode	VaR	ES	Exceedances	P-Value		Keputusan	
					Kupiec	Christ	Kupiec	Christ
1%	Variance-Covariance	-0.0237	-0.0272	11	0.0199178	0.0199178	Gagal	Lulus
	Historical	-0.0278	-0.0400	6	0.6630163	0.84538407	Lulus	Lulus
	Monte Carlo	-0.0238	-0.0272	11	0.0199178	0.05194789	Gagal	Lulus
	EVT (POT-GPD)	-0.0296	-0.0397	4	0.6414349	0.86869882	Lulus	Lulus
	GARCH (<i>rolling</i>)	-0.0221	-0.0254	11	0.0199178	0.05194789	Gagal	Lulus
	GARCH-EVT	-0.0271	-0.0348	7	0.3965697	0.63194787	Lulus	Lulus
5%	Variance-Covariance	-0.0167	-0.0210	19	0.1993852	0.41690375	Lulus	Lulus
	Historical	-0.0148	-0.0235	26	0.8384151	0.92576434	Lulus	Lulus
	Monte Carlo	-0.0167	-0.0210	19	0.1993852	0.41690375	Lulus	Lulus
	EVT (POT-GPD)	-0.0149	-0.0241	25	1.0000000	0.97034299	Lulus	Lulus
	GARCH (<i>rolling</i>)	-0.0155	-0.0195	24	0.8363894	0.30645616	Lulus	Lulus
	GARCH-EVT	-0.0151	-0.0226	25	1.0000000	0.97034299	Lulus	Lulus

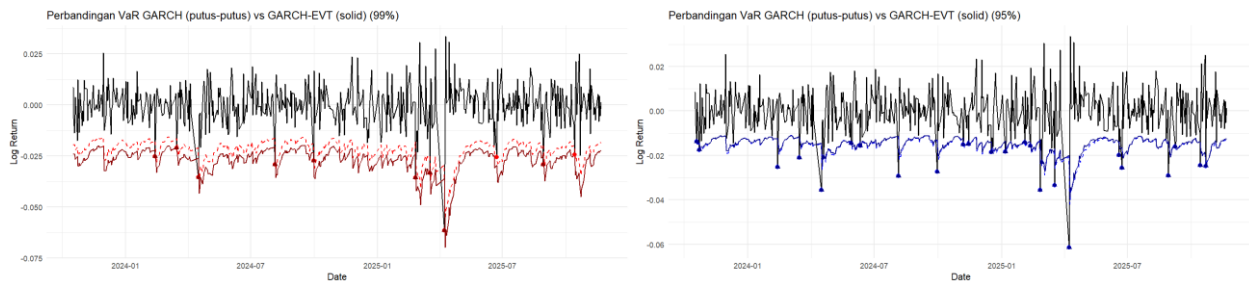
Metode *Extreme Value Theory* (EVT) dengan pendekatan *Peaks Over Threshold* (POT-GPD) serta metode Historis terbukti valid dengan jumlah pelampauan yang berada dalam batas toleransi statistik. Secara khusus, integrasi antara pemodelan volatilitas dan teori nilai ekstrem melalui metode GARCH-EVT berhasil memperbaiki kelemahan GARCH standar, menurunkan jumlah pelampauan dari 11 menjadi 7, dan dinyatakan valid secara statistik. Hal ini menegaskan bahwa pemodelan heteroskedastisitas saja tidak cukup tanpa disertai pemodelan distribusi ekor yang tepat. Evaluasi pada Tingkat Signifikansi 5%.

Berbeda dengan hasil pada kondisi ekstrem, seluruh model yang diuji menunjukkan kinerja yang memuaskan pada tingkat signifikansi 5%. Baik uji *unconditional coverage* (Kupiec) maupun *conditional coverage* (Christoffersen) memberikan keputusan valid (Lulus) untuk semua metode, mulai dari Variance-Covariance hingga EVT. Temuan ini mengimplikasikan bahwa untuk estimasi risiko pada kondisi pasar yang moderat, asumsi distribusi normal masih dapat diterima. Namun, konsistensi ini menyamarkan dampak tersembunyi dimana model yang valid pada kondisi normal ($\alpha = 0.05$) dapat mengalami kegagalan saat diterapkan untuk memprediksi kerugian ekstrem ($\alpha = 0.01$).

Ditinjau dari besaran estimasi kerugian, *Expected Shortfall* (ES) secara konsisten memberikan estimasi yang lebih konservatif dibandingkan *Value at Risk* (VaR), dengan metode EVT menghasilkan proyeksi kerugian terbesar. Mengingat pergeseran standar regulasi perbankan global yang semakin menekankan pada pengukuran *Expected Shortfall* dan ketahanan terhadap *tail events*, hasil ini merekomendasikan penggunaan metode berbasis EVT atau GARCH-EVT. Metode-metode ini terbukti paling robust dalam mengantisipasi kerugian pasar yang jarang terjadi namun berdampak katastrofik, dibandingkan metode konvensional yang cenderung terlalu optimistis di area ekor distribusi.

Perbandingan batas risiko antara model GARCH konvensional dan model *hybrid* GARCH-EVT menunjukkan perbedaan yang konsisten dalam mengestimasi *tail risk*. Pada tingkat kepercayaan 99%, batas VaR berbasis GARCH-EVT berada lebih dalam (lebih negatif) dibandingkan model GARCH saja, sebagaimana tampak dari garis solid yang

selalu berada di bawah garis putus-putus. Hal ini mencerminkan kemampuan EVT (*Extreme Value Theory*) untuk menangkap distribusi ekor yang lebih tebal, sehingga menghasilkan estimasi kerugian ekstrem yang lebih konservatif. Perbedaan ini menjadi semakin jelas pada periode volatilitas tinggi, terutama saat terjadi lonjakan *return* negatif yang tajam. Dalam beberapa kejadian ekstrem, VaR GARCH tampak melampaui realisasi kerugian aktual (*underestimation*), sementara VaR GARCH-EVT tetap berada pada posisi yang lebih akurat dalam mengantisipasi besarnya kerugian.



Gambar 4. Perbandingan VaR GARCH vs GARCH-EVT pada tingkat kepercayaan 99% dan 95%

Pada tingkat kepercayaan 95%, pola serupa tetap muncul namun dengan intensitas lebih rendah. Garis VaR GARCH-EVT tetap menunjukkan batas risiko yang lebih ketat dibandingkan GARCH, tetapi jaraknya tidak sedalam pada level 99%. Pergerakan batas risiko GARCH tampak lebih responsif terhadap fluktuasi harian, sedangkan GARCH-EVT menunjukkan pola yang lebih halus namun tetap memberikan estimasi ekor yang lebih dalam. Pada beberapa titik kerugian besar, pelampauan VaR (*violations*) lebih sering terjadi pada model GARCH dibandingkan model GARCH-EVT, mengindikasikan bahwa penambahan blok EVT memberikan keunggulan dalam mendeskripsikan dinamika risiko ekstrem.

Secara keseluruhan, hasil visual ini memperkuat literatur bahwa model berbasis EVT lebih unggul dalam memodelkan risiko ekor pada data finansial yang menunjukkan kurtosis tinggi dan volatilitas berkelompok. Kombinasi GARCH untuk dinamika volatilitas dan EVT untuk penanganan ekor distribusi memberikan estimasi VaR yang lebih stabil dan konservatif, terutama pada level kepercayaan yang lebih tinggi dan kondisi pasar yang mengalami tekanan kuat.

5. Kesimpulan

Berdasarkan analisis yang telah dilakukan, ditarik tiga kesimpulan utama:

- Portofolio Varians Minimum (MVP) yang dibentuk dari 11 saham perwakilan sektor di BEI menghasilkan portofolio yang terdiversifikasi dengan baik, namun memiliki fokus bobot yang signifikan pada sektor-sektor yang dianggap defensif dan stabil, yaitu *consumer non-cyclicals* (ICBP.JK) dan perbankan *big-cap* (BBCA.JK).
- Terdapat perbedaan signifikan dalam akurasi model VaR berdasarkan tingkat kepercayaan. Pada tingkat VaR 95% (mewakili risiko pasar normal), semua model yang diuji (VC, HS, MC, GARCH, EVT, GARCH-EVT) terbukti memadai dan lulus uji *backtesting*.
- Untuk pengukuran risiko ekstrem (VaR 99%), model-model standar yang berasumsi normalitas (Variance-Covariance, Monte Carlo) serta model GARCH standar terbukti gagal total dan tidak aman digunakan. Model *hybrid* GARCH-EVT terbukti menjadi model yang superior, yang secara akurat menangkap volatilitas dinamis (dari GARCH) sekaligus memodelkan kejadian ekstrem (dari EVT), dan berhasil lulus semua uji *backtesting*.

References

- Acerbi, C., & Tasche, D. (2002). On the Coherence of Expected Shortfall. *Journal of Banking & Finance*, 26(7), 1487–1503. [https://doi.org/10.1016/S0378-4266\(02\)00283-2](https://doi.org/10.1016/S0378-4266(02)00283-2)

- Ariwibowo, A. (2020). *Analisis Value at Risk menggunakan pendekatan Threshold GARCH (TGARCH) dan Generalized Pareto Distribution (GPD)*.
[https://repository.uinjkt.ac.id/dspace/bitstream/123456789/51563/1/ALDILAH ARIWIBOWO-FST.pdf](https://repository.uinjkt.ac.id/dspace/bitstream/123456789/51563/1/ALDILAH%20ARIWIBOWO-FST.pdf)
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3), 203–228. <https://doi.org/10.1111/1467-9965.00068>
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862. <https://doi.org/10.2307/2527339>
- Dowd, K. (2006). *Measuring market risk* (2nd ed.). John Wiley & Sons.
- Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events: for insurance and finance*. Springer-Verlag.
- Glasserman, P. (2004). *Monte Carlo methods in financial engineering* (Vol. 53). Springer-Verlag. <https://doi.org/10.1007/978-0-387-21617-1>
- Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
- J.P. Morgan. (1996). *RiskMetrics---Technical Document*. <https://www.msci.com/documents/10199/5915b101-1905-4392-85a3-71f3a9e8d3f0>
- Kaibuchi, H., Kawasaki, Y., & Stupfler, G. (2021). GARCH-UGH: A bias-reduced approach for dynamic extreme Value-at-Risk estimation in financial time series. *ArXiv Preprint ArXiv:2104.09879*. <https://arxiv.org/abs/2104.09879>
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2), 73–84.
- Linsmeier, T. J., & Pearson, N. D. (1996). Risk measurement: An introduction to value at risk. *University of Illinois at Urbana-Champaign*.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- Maulana, D., Dharmawan, K., & Srinadi, I. G. A. M. (2022). Estimasi CVaR pada portofolio saham menggunakan metode GJR-EVT dengan pendekatan D-Vine Copula. *E-Jurnal Matematika*, 11(2), 127–139. <https://doi.org/10.24843/MTK.2022.v11.i02.p372>
- McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3–4), 271–300. [https://doi.org/10.1016/S0927-5398\(00\)00012-8](https://doi.org/10.1016/S0927-5398(00)00012-8)
- Peiris, R., Wang, C., Gerlach, R., & Tran, M.-N. (2024). Semi-parametric financial risk forecasting incorporating multiple realized measures. *ArXiv Preprint ArXiv:2402.09985*. <https://arxiv.org/abs/2402.09985>
- Yang, J. (2025). Institutional Differences, Crisis Shocks, and Volatility Structure: A By-Window EGARCH/TGARCH Analysis of ASEAN Stock Markets. *ArXiv Preprint ArXiv:2510.16010*. <https://arxiv.org/abs/2510.16010>