



UJMAS

UMBARA JOURNAL OF MATHEMATICS, ACTUARIAL SCIENCE AND STATISTICS
<https://journal.umbogorraya.ac.id/index.php/ujmas>

Kerangka Preprocessing Data Terstruktur Untuk Pemodelan Harga Emas Berbasis Distribusi

Siska Nurmalasari¹, Dyah Prita Anggraini², Aldila Nur Indah Berliana Ratam³

^{1,2,3} Program Studi Sains Aktuaria, Fakultas Kesehatan dan Sains

Universitas Muhammadiyah Bogor Raya

Jln. Raya Leuwiliang No. 106 Kabupaten Bogor Barat Jawa Barat

Volume 1 Nomor 1
Desember 2024 : 9 - 15

Article History

Submission: 09-12-2024

Revised: 09-12-2024

Accepted: 15-12-2024

Published: 20-12-2024

Kata Kunci:

Harga emas; Preprocessing data;
Deret waktu; Pemodelan distribusi;
XAUUSD.

Keywords:

Gold price; Data preprocessing; Time
series; Distribution modeling;
XAUUSD.

Korespondensi:

(Siska Nurmalasari)

(Telp.-)

(siskanurmala99@gmail.com)

Abstrak: Harga emas dunia (XAUUSD) merupakan salah satu deret waktu keuangan yang banyak digunakan dalam kajian statistik keuangan dan analisis risiko. Namun, data mentah harga emas memiliki karakteristik khusus berupa ketidakstasioneran, volatilitas tinggi, distribusi berekor tebal, serta adanya hari tidak bertransaksi akibat akhir pekan dan hari libur pasar internasional. Karakteristik tersebut berpotensi menyebabkan pelanggaran terhadap asumsi dasar pemodelan distribusi apabila tidak ditangani melalui pengolahan data yang tepat. Penelitian ini bertujuan menyusun dan mengevaluasi kerangka preprocessing data terstruktur sebagai fondasi metodologis bagi pemodelan harga emas berbasis distribusi. Penelitian menggunakan data harian XAUUSD jangka panjang yang diperoleh dari sumber terbuka. Metode yang diterapkan meliputi validasi struktur deret waktu, penyesuaian kalender perdagangan, transformasi harga menjadi log-return, pembersihan anomali struktural, serta evaluasi sifat statistik data sebelum dan sesudah preprocessing. Evaluasi dilakukan melalui analisis deskriptif, uji stasioneritas, pengukuran skewness dan kurtosis, serta analisis karakteristik volatilitas. Hasil penelitian menunjukkan bahwa data mentah XAUUSD bersifat tidak stasioner, memiliki distribusi miring dengan ekor tebal, serta varians yang tidak stabil. Setelah preprocessing, data berubah menjadi stasioner, lebih simetris, dan memiliki varians yang lebih homogen sehingga memenuhi asumsi dasar pemodelan distribusi probabilitas dan risiko. Penelitian ini menghasilkan kerangka preprocessing data harga emas yang terdokumentasi secara sistematis dan dapat direplikasi, serta menyediakan landasan metodologis awal bagi pengembangan pemodelan statistik lanjutan pada data komoditas berekor tebal.

Abstract: World gold prices (XAUUSD) are widely used in financial statistics, quantitative finance, and risk analysis. However, raw gold price data exhibit specific characteristics such as non-stationarity, high volatility, heavy-tailed distributions, and systematic non-trading days due to weekends and international



UJMAS: Umbara Journal Of Mathematics, Actuarial Science And Statistics is licensed under a Creative Commons Attribution-Share Alike 4.0 International License. Copyright © 2025 Prodi Sains Aktuaria Universitas Muhammadiyah Bogor Raya, Indonesia. All Rights Reserved

holidays. If these characteristics are not properly handled, violations of the fundamental assumptions of distribution-based statistical modeling may occur. This study aims to develop and evaluate a structured data preprocessing framework as a methodological foundation for distribution-based gold price modeling. This research employs long-term daily XAUUSD data obtained from open data sources. The proposed methodology consists of time-series structure validation, trading calendar adjustment, transformation of prices into log-returns, structural anomaly cleaning, and statistical evaluation of the data before and after preprocessing. The evaluation includes descriptive analysis, stationarity testing, skewness and kurtosis measurement, and volatility characterization. The results indicate that raw XAUUSD data are non-stationary, highly skewed with heavy tails, and exhibit unstable variance. After preprocessing, the data become stationary, more symmetrically distributed, and display more homogeneous variance, thus satisfying the fundamental assumptions required for probabilistic and risk-based distribution modeling. This study produces a systematically documented and replicable preprocessing framework for gold price data and provides a methodological baseline for further statistical modeling of heavy-tailed commodity time series.

PENDAHULUAN

Harga emas dunia yang direpresentasikan oleh XAUUSD merupakan salah satu deret waktu keuangan yang banyak digunakan dalam penelitian statistik keuangan, manajemen risiko, dan keuangan kuantitatif. Data ini dimanfaatkan sebagai dasar dalam pemodelan distribusi, estimasi volatilitas, serta analisis perilaku pasar komoditas. Namun, data mentah harga emas memiliki karakteristik statistik yang kompleks, antara lain ketidakstasioneran pada level harga, volatilitas yang tidak konstan, distribusi return berekor tebal, serta adanya

observasi ekstrem yang dipengaruhi oleh kondisi pasar global [1][2].

Selain itu, data XAUUSD harian secara struktural mengandung hari tidak bertransaksi akibat akhir pekan dan hari libur pasar internasional. Karakteristik ini menyebabkan pola data hilang yang bersifat sistematis dan tidak dapat diperlakukan sebagai kehilangan data acak [3]. Apabila kondisi tersebut tidak ditangani secara tepat pada tahap pengolahan data, maka asumsi-asumsi dasar dalam pemodelan statistik, seperti normalitas, stasioneritas, dan kestabilan varians, berpotensi tidak terpenuhi [1].

Sebagian besar penelitian terkait harga emas cenderung berfokus pada pengembangan model prediksi atau

analisis risiko, sementara tahap pengolahan data sering diperlakukan sebagai langkah teknis yang dijelaskan secara terbatas. Padahal, literatur deret waktu keuangan menegaskan bahwa kualitas inferensi statistik sangat bergantung pada validitas proses pengolahan data sebelum pemodelan dilakukan [1][4]. Oleh karena itu, diperlukan kajian yang secara khusus membahas pengolahan data harga emas XAUUSD secara metodologis dan terdokumentasi dengan baik.

METODE

Penelitian ini menggunakan pendekatan kuantitatif statistik dengan objek kajian berupa data deret waktu harga emas (XAUUSD). Fokus penelitian diarahkan pada pengolahan data sebagai fondasi metodologis bagi pemodelan statistik, tanpa melakukan segmentasi rezim pasar. Seluruh analisis dilakukan pada data agregat sebagai satu kesatuan deret waktu.

Metode penelitian dirancang secara bertahap, sistematis, dan replikatif untuk memastikan bahwa setiap keputusan pengolahan data memiliki

dasar statistik yang jelas serta dapat dipertanggungjawabkan secara ilmiah.

1. Desain Umum Penelitian

Penelitian dilakukan melalui empat tahapan utama, yaitu:

1. Pengumpulan dan Validasi Data,
2. Pengolahan Data Harga Emas,
3. Evaluasi Sifat Statistik Data, dan
4. Analisis dan Interpretasi Hasil.

HASIL & PEMBAHASAN

Hasil

1. Data yang Digunakan dan Validasi Struktur Deret Waktu

Penelitian menggunakan data harian harga emas dunia (XAUUSD) jangka panjang yang bersumber dari *investing.com*, sesuai rancangan metode penelitian pada proposal:

Pada tahap awal dilakukan validasi struktur data yang meliputi:

Aspek Validasi	Hasil
Kelengkapan observasi	Seluruh data memiliki tanggal unik, tidak terdapat duplikasi
Konsistensi format waktu	Seluruh tanggal telah tersusun kronologis
Hari tidak bertransaksi	Teridentifikasi secara sistematis (akhir pekan & hari libur)
Kelayakan statistik	Layak diproses sebagai deret waktu

Capaian:

→ Terbentuk *dataset XAUUSD tervalidasi* sebagai luaran tahap 1.

2. Hasil Pengolahan Data (Preprocessing)

Tahapan preprocessing dilakukan sesuai metodologi dalam proposal, yaitu:

1. Penyesuaian kalender perdagangan
2. Transformasi harga menjadi *log-return*
3. Pembersihan anomali struktural non-ekonomis
4. Penyusunan format data siap analisis

Perubahan karakteristik utama data:

Karakteristik	Data Mentah	Data Setelah Preprocessing
Skala	Level harga	Log-return
Stasioneritas	Tidak stasioner	Stasioner
Distribusi	Miring & berekor tebal	Lebih simetris
Varians	Tidak stabil	Lebih homogen

Capaian:

→ Terbentuk *dataset XAUUSD hasil preprocessing* sebagai luaran tahap 2.

3. Hasil Evaluasi Statistik

Evaluasi sifat statistik dilakukan untuk membuktikan dampak preprocessing terhadap pemenuhan asumsi pemodelan distribusi (Tsay; Cont; Hamilton).

Uji Statistik Sebelum	Sesudah	Interpretasi
Uji ADF (stasioneritas)	Tidak stasioner	Stasioner
		Preprocessing efektif

Uji Statistik Sebelum	Sesudah	Interpretasi	
Skewness	Tinggi	Mendekati 0	Distribusi lebih simetris
Kurtosis	Sangat tinggi	Menurun	Tail lebih terkendali
Volatilitas	Heteroskedastik	Lebih stabil	Layak untuk pemodelan distribusi

Hasil ini mengonfirmasi bahwa preprocessing mampu meningkatkan validitas statistik data sebelum pemodelan, sejalan dengan literatur deret waktu keuangan modern.

Capaian:

→ Tersusun *tabel perbandingan sifat statistik sebelum-sesudah preprocessing* (luaran tahap 3).

4. Analisis dan Implikasi Metodologis

Analisis menunjukkan bahwa:

- Data mentah XAUUSD tidak layak langsung digunakan untuk pemodelan distribusi karena tidak memenuhi asumsi dasar statistik.
- Kerangka preprocessing yang dibangun mampu:
 - ✓ Menstabilkan varians
 - ✓ Memperbaiki sifat distribusi
 - ✓ Menghasilkan data yang layak untuk pemodelan distribusi probabilitas dan risiko.

Hal ini memperkuat argumen metodologis bahwa preprocessing bukan sekadar tahapan teknis, melainkan *faktor kunci validitas inferensi statistik*.

Capaian:

→ Terbentuk *analisis metodologis preprocessing harga emas* sebagai luaran tahap 4.

Pembahasan

Hasil penelitian ini memperlihatkan bahwa data mentah harga emas XAUUSD tidak memenuhi asumsi dasar pemodelan statistik karena bersifat tidak stasioner, memiliki varians tidak stabil, serta distribusi berekor tebal. Temuan ini konsisten dengan literatur deret waktu keuangan yang menyatakan bahwa harga aset finansial umumnya tidak stasioner pada level harga dan memiliki karakteristik volatilitas yang tidak konstan [5][6].

Transformasi harga menjadi return yang dilakukan dalam penelitian ini terbukti mampu menghasilkan data yang stasioner dan lebih stabil secara statistik. Hasil ini sejalan dengan temuan Fama dan French yang menegaskan bahwa return aset finansial cenderung lebih mendekati proses

stasioner dibandingkan level harga, sehingga lebih layak digunakan dalam pemodelan distribusi dan analisis risiko [7].

Perbaikan bentuk distribusi setelah preprocessing, yang ditunjukkan oleh penurunan skewness dan kurtosis, mendukung argumen bahwa pengolahan data merupakan faktor kunci dalam validitas inferensi statistik. Mandelbrot dan kemudian Cont menunjukkan bahwa return aset finansial secara alami bersifat heavy-tailed dan rentan terhadap kejadian ekstrem, sehingga perlakuan data yang tepat sangat diperlukan agar pemodelan risiko tidak menghasilkan estimasi bias [6][8].

Stabilisasi varians yang dicapai melalui preprocessing juga memperkuat kelayakan data untuk pemodelan risiko berbasis distribusi, seperti Value at Risk (VaR) dan Tail Value at Risk (TVaR). McNeil et al. menekankan bahwa validitas pengukuran risiko ekstrem sangat bergantung pada kualitas dan sifat statistik data yang telah diproses dengan benar [9].

Dengan demikian, temuan penelitian ini memperkuat posisi

preprocessing sebagai tahapan metodologis utama dalam pemodelan distribusi harga emas, bukan sekadar langkah teknis, dan memberikan landasan statistik yang kuat bagi penelitian lanjutan di bidang analisis risiko komoditas.

SIMPULAN

Penelitian ini berhasil menyusun dan mengevaluasi kerangka preprocessing data terstruktur sebagai fondasi metodologis bagi pemodelan harga emas berbasis distribusi. Hasil penelitian menunjukkan bahwa data mentah XAUUSD tidak layak langsung digunakan dalam pemodelan distribusi probabilitas karena bersifat tidak stasioner, memiliki distribusi berekor tebal, serta varians yang tidak stabil. Kondisi ini berpotensi menghasilkan inferensi statistik yang bias dan menyesatkan jika tidak ditangani melalui pengolahan data yang tepat. Penerapan kerangka preprocessing yang meliputi validasi struktur deret waktu, penyesuaian kalender perdagangan, transformasi harga menjadi log-return, serta pembersihan anomali struktural terbukti mampu memperbaiki sifat statistik data. Data

hasil preprocessing menjadi stasioner, lebih simetris, dan memiliki varians yang lebih homogen sehingga memenuhi asumsi dasar pemodelan distribusi probabilitas dan pengukuran risiko.

Dengan demikian, preprocessing bukan sekadar tahapan teknis, tetapi merupakan faktor metodologis kunci dalam menjaga validitas inferensi statistik pada pemodelan harga emas. Kerangka yang dihasilkan dalam penelitian ini dapat dijadikan baseline metodologis bagi pengembangan model distribusi lanjutan, analisis risiko ekstrem, serta penelitian deret waktu komoditas berekor tebal di masa mendatang.

UCAPAN TERIMA KASIH

Penulis menyampaikan ucapan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan kontribusi dalam penyusunan penelitian ini. Secara khusus, apresiasi disampaikan kepada institusi dan pihak terkait yang telah menyediakan data dan referensi yang diperlukan sehingga penelitian ini dapat diselesaikan dengan baik. Penulis juga mengucapkan terima kasih kepada para reviewer dan editor

atas masukan, saran, dan kritik konstruktif yang sangat bermanfaat dalam penyempurnaan artikel ini. Semoga hasil penelitian ini dapat memberikan kontribusi positif bagi pengembangan ilmu pengetahuan dan praktik industri asuransi syariah di Indonesia.

DAFTAR PUSTAKA

- [1] R. S. Tsay, *Analysis of Financial Time Series*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2010.
- [2] R. Cont, "Empirical properties of asset returns: Stylized facts and statistical issues," *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.
- [3] H. M. Zangana, "Deep learning-based gold price prediction," *Sistemasi: Jurnal Sistem Informasi*, 2024.
- [4] J. D. Hamilton, *Time Series Analysis*. Princeton, NJ, USA: Princeton University Press, 1994.
- [5] R. F. Engle, "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
- [6] E. F. Fama and K. R. French, "The behavior of stock-market prices," *Journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.
- [7] B. B. Mandelbrot, "The variation of certain speculative prices," *Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.
- [8] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton, NJ, USA: Princeton University Press, 2005.