

BAB I PENDAHULUAN

1.1 Latar Belakang

Emas merupakan salah satu komoditas strategis dunia yang berperan penting sebagai aset keuangan bagi berbagai negara dan termasuk dalam cadangan moneter global untuk menghadapi krisis keuangan mendadak (Hasan et al., 2023). Selain diperdagangkan secara spot, harga emas juga dipengaruhi oleh kontrak *futures*, yakni kesepakatan harga untuk transaksi yang akan dilakukan di masa mendatang. Meskipun Indonesia belum memiliki bursa *futures* resmi, harga *futures* internasional tetap menjadi acuan utama dalam perdagangan emas domestik karena digunakan oleh investor dan pedagang sebagai dasar penentuan nilai jual beli serta strategi investasi (Chirwa & Odhiambo, 2020). Data harga emas *futures* memiliki karakteristik khusus yang memengaruhi peramalan. Salah satunya adalah nonstasioner, di mana rata-rata dan pola korelasi harga tidak konstan seiring waktu. Kondisi ini membuat tren dan volatilitas harga menjadi dinamis. Selain itu, data juga menunjukkan adanya *long memory*, artinya pergerakan harga di masa lalu (*lagged values*) tetap memengaruhi harga di masa depan (*current values*). Kombinasi antara nonstasioneritas dan *long memory* membuat harga emas *futures* sulit untuk diramalkan secara akurat. Oleh karena itu, diperlukan model peramalan yang adaptif dan mampu menangkap ketergantungan *long memory* serta perubahan struktural (Hasan et al., 2023). Salah satu model yang sering digunakan dalam peramalan deret waktu adalah *Autoregressive Integrated Moving Average*

Model *Autoregressive Integrated Moving Average* (ARIMA) dikembangkan oleh Box dan Jenkins untuk memodelkan hubungan antara *lagged values* dan *current values* suatu variabel (Box et al., 2015). ARIMA terdiri atas tiga komponen utama, yaitu *autoregressive* (AR) yang merepresentasikan pengaruh *lagged values*, *differencing* (I) yang digunakan untuk menjadikan data stasioner, dan *moving average* (MA) yang menangkap pengaruh residual *lagged values* terhadap *current values* (Hamilton, 1994). Kombinasi ketiga komponen tersebut membuat ARIMA efektif dalam memodelkan pola musiman dan tren *short memory* pada data stasioner. Model ini tetap menjadi dasar utama dalam analisis peramalan *modern* karena kemampuannya memberikan hasil yang stabil dan mudah diinterpretasikan pada berbagai jenis data ekonomi (Borrero & Mariscal, 2022). Namun demikian, pada data harga emas yang cenderung bersifat nonstasioner dan memiliki fluktuasi *long memory*, model ARIMA sering kali belum mampu menggambarkan struktur dependensi yang kompleks secara akurat (Hasan et al., 2023). Keterbatasan tersebut kemudian mendorong pengembangan model ARFIMA yang mampu menangkap hubungan *long memory* melalui pendekatan *fractional differencing*.



regressive Fractionally Integrated Moving Average an perbaikan dari model ARIMA yang dikembangkan secara identifikasi ketergantungan *long memory* dalam data deret waktu. kan teknik *fractional differencing* untuk menyesuaikan korelasi s dan *current values* yang berlangsung lebih lama daripada yang n model linear standar. (Hasan et al., 2023). Tidak seperti ARIMA

yang hanya memungkinkan nilai *differencing* bilangan bulat, ARFIMA memperluas konsep tersebut dengan memungkinkan derajat *differencing* (d) bernilai pecahan, sehingga mampu memodelkan proses yang bersifat antara stasioner dan nonstasioner. Nilai d yang berada pada rentang 0 hingga kurang dari 0,5 menunjukkan bahwa proses bersifat stasioner dan memiliki *long memory* dengan ketergantungan positif, di mana pengaruh *lagged values* melemah secara perlahan dalam struktur *long memory*. Sebaliknya, nilai d yang mencapai 0,5 atau lebih mengindikasikan *long memory* yang kuat dan proses yang tidak stasioner, sehingga perubahan mendadak pada data akan memiliki dampak yang bertahan lama (Hosking, 1981). Berkat karakteristiknya yang mampu menangkap ketergantungan *long memory*, model ARFIMA banyak digunakan dalam analisis data harga komoditas, tingkat inflasi, dan indeks keuangan. Namun, meskipun efektif dalam mengidentifikasi struktur *long memory*, ARFIMA memiliki keterbatasan dalam menghadapi perubahan mendadak pada pola pasar serta gangguan *noise* pada data aktual. Kondisi ini menurunkan akurasi model (Hasan et al., 2023). Salah satu pendekatan yang banyak digunakan untuk mengatasi keterbatasan tersebut adalah *Kalman Filter*, yang berfungsi sebagai mekanisme *filtering* informasi untuk memperbarui prediksi secara bertahap. Proses *filtering* ini memungkinkan model menyesuaikan peramalan setiap kali data baru muncul, sehingga hasil prediksi tetap mengikuti perubahan kondisi pasar.

Kalman Filter (KF) adalah algoritma berbasis *state space* yang dikembangkan oleh R. E. Kalman dan digunakan untuk meningkatkan kualitas peramalan deret waktu, termasuk data yang mengandung ketergantungan *long memory*. KF bekerja secara rekursif dengan menggabungkan prediksi *state* sebelumnya dan pengamatan terbaru sehingga mampu mereduksi *noise* serta menyesuaikan prediksi secara *real time*. Mekanisme ini menjadikan KF banyak diterapkan pada berbagai bidang, seperti ekonomi dan keuangan (Al-Gounmeein et al., 2023). KF memiliki kelemahan karena hanya bekerja optimal pada sistem yang linear dan tidak mampu menangani pola data yang bersifat nonlinear. Keterbatasan tersebut membuat KF kurang efektif ketika hubungan dalam data tidak mengikuti pola yang sederhana. Oleh sebab itu, *Extended Kalman Filter* Oleh dikembangkan sebagai pendekatan yang mampu menangani ketidaklinearan sehingga menghasilkan peramalan yang lebih sesuai dengan karakter data (Si et al., 2025)

Extended Kalman Filter (EKF) merupakan pengembangan dari KF yang dirancang untuk mengatasi sistem nonlinear. Model ini melakukan linearisasi di sekitar prediksi *state* saat ini dengan menggunakan deret Taylor untuk mendekati fungsi transisi dan observasi. Proses EKF terdiri atas dua tahap utama, yaitu tahap prediksi dan tahap koreksi berdasarkan data observasi aktual. Pada tahap koreksi, digunakan untuk memperbarui kovariansi *error*. Dengan itu, EKF mempertahankan sifat rekursif dan efisiensi komputasi yang meningkatkan kemampuan adaptasinya terhadap data deret waktu (Ribeiro & Ribeiro, 2014). Namun, EKF an karena pendekatan linearisasinya membuat metode ini kurang menangkap struktur ketergantungan *long memory* yang muncul. Oleh



karena itu, EKF dipadukan dengan ARFIMA agar kedua model saling melengkapi. Kombinasi ini menghasilkan model *hybrid* yang tidak hanya mampu menangkap sifat *long memory* tetapi juga menyesuaikan peramalan terhadap perubahan dinamis pada data deret waktu.

Beberapa penelitian sebelumnya telah mengkaji penerapan ARFIMA dan EKF dalam peramalan harga komoditas dan aset keuangan. Penelitian oleh Mahmad Azan et al. (2021) membandingkan model ARIMA dan ARFIMA untuk meramalkan harga Kijang Emas di Malaysia. Hasilnya menunjukkan bahwa ARFIMA mampu menangkap ketergantungan *long memory* sehingga menghasilkan peramalan yang lebih akurat dibandingkan ARIMA. Sementara itu, studi pada indeks saham di Amerika Serikat dan China oleh Alp et al. (2023) menerapkan EKF untuk menangani data nonlinear dan volatil, serta memperbarui prediksi *state* secara adaptif berdasarkan data observasi aktual, sehingga prediksi menjadi lebih responsif dan akurat. Temuan dari kedua penelitian tersebut menjadi dasar pengembangan model *hybrid* ARFIMA–EKF, yang memadukan keunggulan ARFIMA dalam menangkap sifat *long memory* dengan kemampuan EKF dalam melakukan koreksi dinamis terhadap nonlinearitas dan volatilitas pasar. Model ini diterapkan pada data harga emas periode Januari 2014 hingga Januari 2024, yang mencakup berbagai rezim pasar seperti krisis global, pandemi COVID-19, dan fase pemulihan ekonomi, sehingga memungkinkan pengujian ketahanan model pada kondisi yang beragam. Melalui model tersebut, penelitian ini diharapkan dapat meningkatkan akurasi peramalan harga emas sekaligus mendukung pengambilan keputusan investasi yang lebih responsif dan adaptif terhadap perubahan pasar.

1.2 Batasan Masalah

1. Data yang digunakan dalam penelitian ini terbatas pada harga emas mingguan dari Januari 2014 hingga Januari 2024 yang bersumber dari *website Investing.com*, sehingga hasil peramalan hanya mencerminkan perilaku harga emas pada periode dan frekuensi tersebut.
2. Model yang digunakan terbatas pada *hybrid* ARFIMA-EKF, di mana nilai parameter d ditetapkan berdasarkan hasil estimasi *fractional differencing* terbaik melalui analisis *Rescaled Range* (R/S), tanpa eksplorasi terhadap variasi nilai d lainnya maupun model peramalan lain.

1.3 Tujuan Penelitian

1. Mendapatkan estimasi parameter model ARFIMA terbaik untuk data harga emas periode Januari 2014 hingga Januari 2024.
2. Mendapatkan hasil peramalan harga emas periode Januari 2022 hingga 2024 berdasarkan model *hybird* ARFIMA-EKF



Penelitian

Penelitian ini bertujuan untuk meningkatkan pengetahuan bagi penulis dan pembaca mengenai penerapan model *hybird* ARFIMA-EKF untuk memperoleh hasil peramalan harga emas yang lebih akurat.

2. Memberikan contoh penerapan nyata model *hybrid* ARFIMA-EKF sebagai bahan referensi untuk penelitian selanjutnya pada bidang peramalan data ekonomi dan komoditas.

1.5 Landasan Teori

1.5.1 Peramalan

Peramalan adalah proses sistematis untuk memperkirakan kejadian atau nilai di masa depan dengan memanfaatkan *lagged values* dan *current values* guna mendukung pengambilan keputusan yang lebih efisien (Armstrong, 2001). Metode peramalan dibagi menjadi dua jenis. Pertama, pendekatan kuantitatif, yang mencakup model deret waktu seperti ARIMA dan eksponensial *smoothing*, serta model kausal seperti regresi dan VAR. Kedua, pendekatan kualitatif, yang bergantung pada penilaian ahli, mencakup model Delphi (Petropoulos et al., 2022). Pada analisis harga emas, pendekatan kuantitatif lebih sesuai karena data historis dapat dianalisis untuk mengenali tren, volatilitas, dan ketergantungan *long memory*, sehingga memudahkan penerapan model seperti ARFIMA dan EKF untuk meningkatkan akurasi peramalan.

1.5.2 Harga Emas *Futures*

Harga emas diperdagangkan dalam dua bentuk utama, yaitu harga *spot* dan harga *futures*. Harga *spot* adalah harga emas untuk transaksi saat ini, sedangkan harga *futures* adalah harga yang disepakati untuk transaksi di masa depan dan sering dimanfaatkan untuk keperluan *hedging* maupun aktivitas *trading* berisiko oleh investor dan pedagang. Pergerakan harga emas *futures* dipengaruhi oleh berbagai faktor ekonomi makro, termasuk nilai tukar mata uang, tingkat suku bunga, inflasi, serta ketidakpastian ekonomi dan geopolitik yang meningkatkan permintaan emas sebagai *safe haven*. Selain itu, model *cost-of-carry* menunjukkan bahwa harga emas *futures* dipengaruhi oleh ekspektasi pasar terhadap harga *spot* di masa depan, biaya penyimpanan, dan tingkat bunga, sehingga selisih antara harga *futures* dan harga *spot* mencerminkan faktor-faktor tersebut. Pemahaman terhadap mekanisme ekonomi ini menjadi landasan penting dalam peramalan harga emas *futures* karena dapat membantu investor, pedagang, dan pembuat kebijakan mengambil keputusan yang lebih tepat dan meminimalkan risiko investasi (Radev et al., 2022). Dalam penelitian ini, data harga emas *futures* diambil dari *Investing.com* yang menyediakan data historis dan *real-time* yang dapat diandalkan untuk analisis pasar keuangan.

1.5.3 Data Deret Waktu



Deret waktu adalah sekumpulan data yang dicatat secara berurutan pada waktu tertentu, yang digunakan untuk menganalisis pola, tren, dan variasi data. Model deret waktu mencakup model klasik seperti ARIMA yang banyak digunakan untuk data keuangan dengan karakteristik seperti harga emas (X. Zhang & Cao, 2025). Namun, untuk meningkatkan akurasi, dikembangkan model *hybrid* yang menggabungkan

keunggulan beberapa model dalam menangani kompleksitas dan ketidakpastian data (Petroopoulos et al., 2022). Dalam penerapannya, model deret waktu perlu bersifat stasioner karena ketidakmampuan mengubah data menjadi bentuk stasioner dapat menghasilkan output yang tidak sah (Baidya & Lee, 2024). Oleh karena itu, pemeriksaan dan penyesuaian stasionaritas menjadi tahap penting dalam menganalisis model deret waktu.

1.5.4 Uji Stasioneritas

Uji stasioneritas dilakukan untuk memastikan bahwa data memiliki karakteristik statistik yang konstan, baik dalam rata-rata maupun variansi. Data yang tidak stasioner dapat menyebabkan hasil estimasi model menjadi bias. Uji stasioneritas yang digunakan adalah uji *Augmented Dickey–Fuller*.

1.5.4.1. Uji *Augmented Dickey–Fuller*

Uji *Augmented Dickey–Fuller* (ADF) merupakan pengembangan dari uji *Dickey–Fuller* yang digunakan untuk menguji keberadaan unit *root* dalam data deret waktu. Kehadiran unit *root* menunjukkan bahwa data tidak stasioner karena nilai rata-rata cenderung berubah dari waktu ke waktu. Uji ADF memperhitungkan autokorelasi pada residual dengan menambahkan lag dari differensi pertama, sehingga hasil pengujian lebih akurat. Hipotesis uji yang digunakan adalah (Rosadi D., 2012):

$H_0: \delta = 0$ (terdapat akar unit, data tidak stasioner)

$H_1: \delta < 0$ (tidak terdapat akar unit, data stasioner)

Statistik uji ADF dihitung dengan:

$$t_{hitung} = \frac{\hat{\delta}}{SE(\hat{\delta})} \quad (1)$$

Keterangan :

$\hat{\delta}$: Estimator *least square*

$SE(\hat{\delta})$: *Standard error* dari nilai estimasi δ

Kriteria pengujian uji ADF adalah tolak H_0 Jika $t_{hitung} > t_{tabel}$ dan data dianggap stasioner.

1.5.5 Uji Residual

Uji residual dilakukan untuk memastikan bahwa *error* dari model peramalan memenuhi asumsi dasar statistik, yaitu bersifat *white noise* dan terdistribusi normal. Uji residual penting untuk menilai kelayakan model dan ketepatan peramalan. Uji residual yang digunakan pada penelitian ini adalah uji *Ljung-Box* dan *Kolmogorov-Smirnov*.

1.5.5.1. Uji *Ljung-Box*



Box digunakan untuk memeriksa ada tidaknya autokorelasi pada data deret waktu hingga beberapa lag. Uji ini menyempurnakan uji *Box–Pierce* dengan menambahkan faktor koreksi berbasis ukuran sampel pada rumus. Dengan demikian, hasil pengujiannya menjadi lebih akurat. Dengan demikian, uji *Ljung-Box* memberikan penilaian yang lebih sensitif terhadap keberadaan autokorelasi. Uji ini umum diterapkan pada model seperti ARIMA atau ARFIMA untuk

memastikan kecukupan model. Statistik uji *Ljung-Box* dihitung menggunakan hipotesis berikut (Dare et al., 2022).

$H_0: \rho_1 = \rho_2 = \dots = \rho_t = 0$ (residual bersifat *white noise*)

H_1 : Minimal ada satu nilai $\rho_t \neq 0$, $t = 1, 2, \dots, T$ (residual tidak *white noise*)

Statistik Uji :

$$Q = T(T+2) \sum_{t=1}^h \frac{\rho_t^2}{T-t} \quad (2)$$

Keterangan :

T : jumlah observasi

ρ_t^2 : autokorelasi sampel residual pada lag t

t : jumlah lag yang diuji

h : lag maksimum yang diuji

Keputusan diambil dengan menolak H_0 apabila $Q > X^2$ atau $p - value < 0,05$.

1.5.5.2. Uji Kolmogorov-Smirnov

Uji *Kolmogorov-Smirnov* digunakan untuk menguji kesesuaian distribusi residual dengan distribusi normal. Uji ini bekerja dengan membandingkan bentuk distribusi residual dengan distribusi normal, lalu melihat seberapa besar perbedaannya. Uji *Kolmogorov-Smirnov* efektif untuk mengevaluasi keseragaman distribusi data populasi, karena bersifat nonparametrik dan tidak bergantung pada asumsi bentuk distribusi tertentu. Persyaratan penggunaan KS adalah data bersifat kontinu dan independen. Hipotesis yang digunakan adalah sebagai berikut (E. C. et al., 2025).

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik uji yang digunakan adalah sebagai berikut.

$$D_{cal} = \max |F_0(Y_t) - S_t(Y_t)| \quad (3)$$

Keterangan :

$F_0(Y_t)$: fungsi distribusi kumulatif teoritis

$S_t(Y_t)$: fungsi distribusi kumulatif empiris

Jika $D_{cal} > D_{tabel}$ atau $p - value < 0,05$ maka H_0 ditolak, artinya residual tidak berdistribusi normal.

1.5.6 Uji Linearitas

Pemilihan model yang tepat penting untuk memastikan hasil peramalan maupun interpretasi data yang akurat. Nonlinearitas terjadi ketika hubungan antara variabel tidak mengikuti pola linear sederhana, sehingga perubahan pada variabel



selalu proporsional pada variabel dependen. Penerapan model dengan hubungan nonlinear dapat menimbulkan kesalahan dan menghasilkan peramalan yang tidak akurat. Uji *Ramsey* *in Specification Error Test* (RESET) digunakan untuk mendeteksi akibat nonlinearitas. Tujuan utama uji ini adalah menilai apakah digunakan sudah sesuai atau masih memerlukan transformasi

terhadap variabel peramalan. Hipotesis yang digunakan adalah sebagai berikut (Warsito & Ispiryanti, 2004) .

$H_0: \alpha_2 = \dots = \alpha_k = 0$ (Bentuk hubungan linear)

H_1 : salah satu $\alpha_k \neq 0$ (Bentuk hubungan nonlinear)

Hipotesis tersebut perlu diuji melalui serangkaian tahapan yang disusun dalam prosedur uji RESET. Prosedur ini menjelaskan langkah-langkah yang digunakan untuk menilai kesesuaian model linear dan diawali dengan membangun model linear awal yang dituliskan pada Persamaan (4)

$$Y_t = \hat{Y}_t + \hat{e}_t, \quad \text{di mana } \hat{Y}_t = \tilde{X}_t \hat{\theta} \quad (4)$$

Keterangan :

Y_t : nilai aktual pada waktu ke- t

$\hat{Y}_t$: nilai peramalan model ARFIMA pada waktu ke- t

$\hat{e}_t$: residual ARFIMA

$\tilde{X}_t$: vektor variabel bebas pada waktu ke- t

$\hat{\theta}$: vektor parameter hasil peramalan ARFIMA

Prosedur RESET memperluas model awal dengan memasukkan transformasi nonlinear dari nilai prediksi $\hat{Y}_t$. Pendekatan ini memanfaatkan pangkat-pangkat $\hat{Y}_t$ untuk menilai apakah bentuk nonlinear tersebut mampu menjelaskan variasi yang masih tersisa pada residual model linear. Langkah ini memformulasikan hubungan antara residual awal $\hat{e}_t$ dan transformasi $\hat{Y}_t$ dalam persamaan (5):

$$\hat{e}_t = \alpha_2 (\hat{Y}_t)^2 + \alpha_3 (\hat{Y}_t)^3 + \dots + \alpha_k (\hat{Y}_t)^k + v_t, \quad \text{untuk } k \geq 2 \quad (5)$$

Keterangan :

$(\hat{Y}_t)^k$: transformasi nonlinear dari peramalan ARFIMA

α : koefisien yang mengukur pengaruh transformasi nonlinier terhadap e_t

k : pangkat maksimum yang digunakan dalam uji RESET, dengan $k \geq 2$

$\hat{\theta}$: vektor parameter hasil peramalan ARFIMA

Langkah selanjutnya adalah melakukan estimasi terhadap model alternatif dari Persamaan (5). Estimasi dilakukan dengan memasukkan hasil transformasi nonlinier dari nilai peramalan ARFIMA sebagai variabel tambahan dalam model residual, sehingga proses ini dapat digunakan untuk mendeteksi nonlinearitas melalui model alternatif. Hasil estimasi tersebut kemudian dapat dituliskan dalam Persamaan (6) sebagai:

$$Y_t = \hat{Y}_t + \alpha_2 (\hat{Y}_t)^2 + \alpha_3 (\hat{Y}_t)^3 + \dots + \alpha_T (\hat{Y}_t)^k + \hat{v}_t, \quad \text{untuk } k \geq 2 \quad (6)$$



Keterangan :

$(\hat{Y}_t)^k$: transformasi nonlinear dari nilai peramalan ARFIMA

$\hat{v}_t$: residual dari model alternatif RESET pada waktu ke- t

Setelah residual dari model alternatif, langkah berikutnya adalah

melakukan uji RESET untuk menilai keberadaan pola nonlinear yang tidak

terdapat pada model awal. Statistik uji tersebut dirumuskan dalam Persamaan (7):

$$RESET = \frac{(\hat{e}'\hat{e} - \hat{v}'\hat{v}) / (k-1)}{(\hat{v}'\hat{v}) / (T-k)} \quad (7)$$

Keterangan :

$\hat{e}$: residual model ARFIMA

$\hat{v}$: residual dari model alternatif RESET

$\hat{e}'\hat{e}$: jumlah kuadrat residual ARFIMA

$\hat{v}'\hat{v}$: jumlah kuadrat residual model alternatif RESET

H_0 ditolak jika $RESET > F_{t-1, T-t}^\alpha$ atau $p - value < 0,05$.

1.5.7 Model Autoregressive Fractionally Integrated Moving Average

Model ARFIMA merupakan pengembangan dari ARIMA yang memasukkan konsep *fractional differencing*, sehingga mampu menangkap *long memory* pada data deret waktu. Dalam data dengan *long memory*, autokorelasi tetap signifikan pada lag yang jauh dan tidak dapat diakomodasi oleh model ARIMA. Parameter utama ARFIMA adalah orde p (*autoregressive*), d (derajat *fractional differencing*), dan q (*moving average*) (Baillie, 1996). Bentuk umum model $ARFIMA(p, d, q)$ disajikan pada Persamaan (8) sebagai berikut.

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)e_t \quad (8)$$

Keterangan :

B : operator *backshift* ($BY_t = Y_t - 1$)

$\phi_p(B)$: polinomial AR orde p

$\theta_q(B)$: polinomial AR orde q

d : parameter *differencing fractional*, $-0.5 < d < 0.5$

Y_t : nilai aktual data pada waktu $ke-t$

e_t : residual

Jika $d = 0$, model ARFIMA menjadi $ARMA(p, q)$. Jika d bilangan bulat positif, model ARFIMA menjadi $ARIMA(p, d, q)$. Jika d pecahan antara $-0,5$ dan $0,5$, model ARFIMA mampu memodelkan sifat *long memory* dalam data deret waktu. Berdasarkan struktur model tersebut, residual ARFIMA diperoleh dari selisih antara nilai aktual dan nilai yang diperkirakan model, sebagaimana dirumuskan dalam Persamaan (9) sebagai berikut:

$$e_t = Y_t - \hat{Y}_t \quad (9)$$

Di mana e_t adalah residual pada waktu $ke-t$ yang menunjukkan selisih antara data aktual dan hasil prediksi model ARFIMA. Dengan demikian, residual mencerminkan bagian dari pergerakan data yang tidak dapat dijelaskan oleh model, sehingga penting dalam menilai kecocokan model terhadap data deret waktu.



i Long memory

dengan metode ARFIMA memerlukan pemeriksaan awal an sifat *long memory* pada data. Salah satu metode yang umum dengan menghitung *Hurst Exponent* (H). Nilai H menunjukkan

derajat ketergantungan *long memory* dalam data deret waktu (Mandelbrot, 1969). Ukuran nilai H memiliki interpretasi, yaitu:

1. $H = 0,5$ menunjukkan bahwa data memiliki *short memory*
2. $H > 0,5$ menunjukkan bahwa data bersifat *persistent* dan memiliki *long memory*, di mana perubahan nilai deret waktu cenderung melanjutkan arah sebelumnya.
3. $H < 0,5$ menunjukkan bahwa data bersifat *anti-persistent*, yaitu perubahan nilai deret waktu cenderung berbalik arah.

Perhitungan nilai H dimulai dengan menghitung rata-rata dari seluruh data deret waktu. Rata-rata ini menjadi nilai pusat yang mewakili tingkat umum data dan digunakan sebagai titik acuan dalam langkah-langkah selanjutnya. Rata-rata dihitung menggunakan Persamaan (10):

$$(\bar{Y}) = \frac{1}{T} \sum_{t=1}^T Y_t, \quad t = 1, 2, \dots, T \quad (10)$$

Langkah berikutnya adalah menghitung *adjusted mean*, yaitu nilai observasi yang telah disesuaikan terhadap rata-rata. Penyesuaian ini dilakukan dengan mengurangi setiap nilai Y_t dengan $\bar{Y}$, sehingga diperoleh ukuran seberapa jauh suatu pengamatan berada di atas atau di bawah rata-rata. *Adjusted mean* dituliskan pada Persamaan (11):

$$Y_t^{adj} = Y_t - \bar{Y}, \quad t = 1, 2, \dots, T \quad (11)$$

Deret *adjusted mean* tersebut selanjutnya diakumulasi secara bertahap sehingga menghasilkan deviasi kumulatif. Deviasi kumulatif diperoleh dengan menjumlahkan seluruh nilai *adjusted mean* mulai dari awal hingga waktu ke- t , sebagaimana ditunjukkan dalam Persamaan (12).

$$Y_t^* = \sum_{t=1}^T Y_t^{adj}, \quad t = 1, 2, \dots, T \quad (12)$$

Langkah berikutnya adalah menentukan rentang deviasi kumulatif. Rentang ini dihitung sebagai selisih antara nilai maksimum dan minimum dari seluruh nilai deviasi kumulatif pada setiap t , sebagaimana diberikan dalam Persamaan (13).

$$R_t = \max(Y_1^*, Y_2^*, \dots, Y_t^*) - \min(Y_1^*, Y_2^*, \dots, Y_t^*), \quad t = 1, 2, \dots, T \quad (13)$$

Agar rentang pada Persamaan (13) dapat dinormalisasi terhadap tingkat variasi dalam data asli, dihitung standar deviasi dari seluruh observasi menggunakan Persamaan (14). Standar deviasi menggambarkan seberapa besar penyebaran data terhadap rata-rata data deret waktu.

$$s_t = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2}, \quad t = 1, 2, \dots, T \quad (14)$$



kumulatif kemudian dibagi dengan standar deviasi sehingga *Rescaled Range* (R/S) pada setiap t . Rasio ini digunakan sebagai standar dan diformulasikan melalui Persamaan (15):

$$\left(\frac{R}{S}\right)_t = \frac{R_t}{S_t}, \quad t = 1, 2, \dots, T \quad (15)$$

Nilai R/S secara teoritis mengikuti pola hubungan pangkat terhadap ukuran sampel t sesuai Persamaan (16).

$$\left(\frac{R}{S}\right)_t = c \cdot t^H, \quad t = 1, 2, \dots, T \quad (16)$$

Konstanta c pada Persamaan (16) merupakan konstanta proporsionalitas dan H sebagai eksponen *Hurst*. Hubungan tersebut bersifat nonlinear, sehingga penyajiannya dalam bentuk linear dilakukan melalui transformasi logaritmik pada kedua sisi sesuai Persamaan (17) berikut.

$$\log\left(\frac{R}{S}\right)_t = \log c + H \log t \quad (17)$$

Transformasi linear pada Persamaan (17) memungkinkan estimasi nilai H melalui regresi antara $X_t = \log t$ dan $Z_t = \log(R/S)_t$. Nilai H dihitung menggunakan Persamaan (18) berikut.

$$H = \frac{\sum_{t=1}^T (X_t - \bar{X})(Z_t - \bar{Z})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (18)$$

Nilai H yang diperoleh dari analisis R/S digunakan sebagai dasar dalam menentukan derajat *fractional differencing* (d) pada model ARFIMA.

1.5.7.2. Estimasi Parameter *Fractional Differencing*

Identifikasi adanya long memory pada data perlu diikuti dengan penentuan parameter *fractional differencing* (d) karena parameter ini merupakan komponen utama dalam model ARFIMA. Parameter d ditentukan dengan memanfaatkan nilai *Hurst* (H) yang diperoleh melalui analisis *Rescaled Range* (R/S). Hubungan antara nilai H dan parameter d dinyatakan melalui Persamaan (19) sebagai berikut.

$$d = H - 0,5 \quad (19)$$

Pendekatan ini menghubungkan nilai H dengan *fractional differencing* yang diperlukan untuk menjadikan data stasioner sekaligus mempertahankan *long memory*. *Fractional differencing* memungkinkan d bernilai pecahan dan tidak terbatas pada bilangan bulat seperti pada ARIMA. Dengan demikian, model ARFIMA mampu membuat data deret waktu menjadi stasioner sekaligus menangkap sifat *long memory*. Karakteristik deret *fractionally integrated* dijelaskan sebagai berikut (Hosking, 1981).



ong memory dan tidak stasioner

ses *long memory* stasioner dengan ketergantungan positif

oses *long memory* stasioner dengan ketergantungan negatif

es *short memory*

1.5.7.3. Estimasi Parameter Model $ARFIMA(p, d, q)$

Estimasi parameter model $ARFIMA(p, d, q)$ dimulai dengan penentuan struktur awal model melalui identifikasi orde *autoregressive* (p) dan *moving average* (q) menggunakan plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) setelah data berada dalam kondisi stasioner. Plot ACF menunjukkan korelasi antar nilai deret waktu pada berbagai lag, sedangkan PACF menggambarkan korelasi parsial pada setiap lag. Pola *cut off* pada PACF hingga lag ke- p dengan ACF yang melemah mengindikasikan model *autoregressive* $AR(p)$, sementara pola *cut off* pada ACF hingga lag ke- q dengan PACF yang melemah menunjukkan model *moving average* $MA(q)$. Hasil identifikasi ini bersifat awal dan perlu dikonfirmasi melalui estimasi parameter serta pemilihan model terbaik. Oleh karena itu, estimasi parameter dilakukan untuk memperoleh nilai parameter yang paling sesuai dengan pola data deret waktu.

Model $ARFIMA(p, d, q)$ memiliki dua kelompok parameter utama, yaitu ϕ sebagai koefisien *autoregressive* (AR) dan θ sebagai koefisien *moving average* (MA), di mana ϕ menunjukkan pengaruh *lagged values* terhadap *current values*, sedangkan θ menunjukkan pengaruh residual *lagged values* terhadap *current values*. Estimasi parameter dalam penelitian ini menggunakan metode *Exact Maximum Likelihood* (EML) yang bekerja dengan cara memaksimalkan fungsi log-likelihood dari residual model yang diasumsikan berdistribusi normal $N(0, \sigma^2)$ berdasarkan Persamaan (8). Seluruh persamaan pada bagian ini mengacu pada formulasi EML yang dijelaskan dalam Octaviyani Devi (2019). Residual yang digunakan dalam proses pembentukan fungsi likelihood tersebut dihitung sesuai bentuk operasional pada Persamaan (20) berikut.

$$e_t = \theta_q(B)^{-1} \phi_p(B)(1-B)^d Y_t \quad (20)$$

Bentuk lengkap dari persamaan residual dapat dituliskan dengan menguraikan operator AR, MA, dan *fractional differencing* sebagaimana ditunjukkan pada Persamaan (21).

$$e_t = \frac{(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d (Y_t - \mu)}{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q} \quad (21)$$

Residual e_t pada Persamaan (21) diasumsikan mengikuti distribusi normal dengan rata-rata nol dan varians σ_e^2 . Asumsi tersebut digunakan untuk membentuk fungsi kepadatan peluang bersama dari vektor residual $\mathbf{e} = (e_1, e_2, \dots, e_T)$ sebagaimana dinyatakan pada Persamaan (22).

$$P(\mathbf{e} | d, \phi, \mu, \theta, \sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma_e^2} \sum_{t=1}^T e_t^2\right) \quad (22)$$



residual pada Persamaan (22) selanjutnya dinyatakan dalam bentuk log dengan menerapkan transformasi logaritma natural. Untuk menghasilkan fungsi *log-likelihood* sebagaimana dirumuskan pada Persamaan (23).

$$\ln L(\phi, \theta, \sigma_e^2) = -\frac{T}{2} \ln(2\pi\sigma_e^2) - \frac{1}{2\sigma_e^2} \sum_{t=1}^T e_t^2 \quad (23)$$

Fungsi log-likelihood pada Persamaan (23) digunakan sebagai dasar estimasi parameter model $ARFIMA(p, d, q)$ dengan nilai d diasumsikan tetap sehingga estimasi difokuskan pada parameter ϕ , θ , dan σ_e^2 . Sebagai pemisalan, digunakan model $ARFIMA(1, d, 1)$ yang melibatkan parameter ϕ_1 dan θ_1 . Estimasi parameter dilakukan dengan memaksimalkan fungsi log-likelihood. di mana kondisi maksimum dicapai ketika turunan pertama fungsi log-likelihood bernilai nol. Turunan terhadap σ_e^2 dinyatakan pada Persamaan (24).

$$\frac{\partial \ln L}{\partial \sigma_e^2} = -\frac{T}{2\sigma_e^2} + \frac{1}{2(\sigma_e^2)^2} + \sum_{t=1}^T e_t^2 \quad (24)$$

Persamaan (24) menghasilkan penaksir maksimum likelihood bagi variansi error sebagaimana ditunjukkan pada Persamaan (25).

$$\hat{\sigma}_e^2 = \frac{1}{T} \sum_{t=1}^T e_t^2 \quad (25)$$

Selanjutnya, penurunan fungsi log-likelihood terhadap parameter *autoregressive* ϕ_1 menghasilkan kondisi orde pertama yang dinyatakan pada Persamaan (26) berikut.

$$\frac{\partial \ln L}{\partial \phi_1} = -\frac{1}{\sigma_e^2} \sum_{t=1}^T e_t \frac{\partial e_t}{\partial \phi_1} \quad (26)$$

Selanjutnya, turunan pertama fungsi log-likelihood terhadap parameter *moving average* θ_1 dinyatakan pada Persamaan (27).

$$\frac{\partial \ln L}{\partial \theta_1} = -\frac{1}{\sigma_e^2} \sum_{t=1}^T e_t \frac{\partial e_t}{\partial \theta_1} \quad (27)$$

Persamaan (26), Persamaan (27), dan penaksir variansi *error* pada Persamaan (25) membentuk sistem persamaan nonlinear yang diselesaikan secara simultan melalui prosedur optimasi numerik. Penyelesaian sistem tersebut menghasilkan estimasi parameter $\hat{\phi}_1$, $\hat{\theta}_1$, dan $\hat{\sigma}_e^2$ yang memaksimalkan fungsi log-likelihood model $ARFIMA(1, d1)$ dan memberikan kesesuaian terbaik dengan data deret waktu.

1.5.7.4. Pemilihan Model Terbaik

Tahap pemilihan model terbaik penting dalam analisis deret waktu untuk memastikan model dapat merepresentasikan pola dan karakteristik data secara akurat. Tahap ini melibatkan evaluasi beberapa kandidat model menggunakan ukuran kinerja tertentu, sehingga diperoleh model yang memberikan hasil peramalan paling akurat dan efisien. *Bayesian Information Criterion* (BIC) merupakan salah satu ukuran statistik yang digunakan untuk menilai dan membandingkan beberapa model



bandingan antara tingkat kesesuaian model dan kompleksitasnya. Untuk memilih model yang memberikan likelihood tinggi sekaligus sedikit parameter. Secara matematis, BIC dirumuskan pada bagian berikut (Rossi et al., 2020).

$$BIC = -2 \ln(L) + k \ln(T) \quad (28)$$

dengan L merupakan ukuran seberapa besar kemungkinan model menghasilkan data yang diamati, k jumlah parameter model, dan T jumlah observasi. Model dengan nilai BIC terkecil dianggap sebagai model terbaik. Kriteria BIC sering digunakan dalam pemodelan deret waktu seperti ARIMA, ARFIMA, maupun model *hybrid* untuk menentukan kombinasi parameter model (p, d, q) . Kriteria ini berperan penting sebagai dasar dalam pemilihan model peramalan yang optimal, stabil, dan efisien.

1.5.8 Extended Kalman Filter

Extended Kalman Filter (EKF) merupakan pengembangan dari *Kalman Filter* yang dirancang untuk menangani sistem nonlinear dengan melakukan linearisasi fungsi *state* dan fungsi observasi menggunakan turunan parsial terhadap titik prediksi terakhir (Ribeiro & Ribeiro, 2014). Model ini bekerja secara rekursif melalui dua tahap utama, yaitu tahap prediksi dan tahap koreksi. Tahap koreksi memperbarui prediksi *state* dan kovariansi *error* berdasarkan data observasi aktual, yaitu data pengamatan pada periode berjalan yang digunakan sebagai variabel observasi dalam EKF (Welch & Bishop, 2006). Formulasi EKF yang digunakan dalam penelitian ini disajikan pada Persamaan (29) hingga Persamaan (37) (Agustin, 2019). Model EKF digunakan untuk mengoreksi hasil prediksi ARFIMA melalui pemodelan komponen koreksi sebagai *hidden state*. EKF diawali dengan pendefinisian persamaan *state* yang merepresentasikan dinamika perubahan koreksi terhadap prediksi ARFIMA. Komponen koreksi tersebut dilambangkan sebagai x_t , yang diasumsikan berkembang mengikuti suatu fungsi transisi nonlinear, sehingga model *state* diformulasikan pada Persamaan (29).

$$x_t = f(x_{t-1}) + w_t \quad (29)$$

Keterangan :

- x_t : *hidden state* pada waktu ke- t
- $f(x_{t-1})$: fungsi transisi *state*
- w_{t-1} : proses *noise* pada waktu ke- $t - 1$

Persamaan (26) menghasilkan prediksi *state* yang tidak dapat diamati secara langsung. Informasi mengenai *state* tersebut diperoleh melalui persamaan observasi yang menghubungkan *hidden state* dengan data pengamatan. Dalam penelitian ini, data pengamatan yang digunakan bukan berupa nilai aktual deret waktu, melainkan yang mencerminkan selisih antara nilai aktual dan hasil prediksi dual tersebut diperlakukan sebagai variabel observasi dan *an state* melalui suatu fungsi observasi nonlinear sebagaimana persamaan (30).



$$e_t = h(x_t) + v_t \quad (30)$$

Keterangan :

$h(x_t)$: fungsi observasi

v_t : *noise* observasi

e_t : residual ARFIMA pada waktu ke- t

Mengacu pada model *state* dan observasi pada Persamaan (29) dan Persamaan (30), inialisasi *state* x_0 ditentukan berdasarkan informasi awal, seperti nilai residual atau rata-rata residual ARFIMA. Kovariansi *error* P_0 merepresentasikan tingkat ketidakpastian prediksi awal dan umumnya ditetapkan berdasarkan variansi residual ARFIMA atau sebagai matriks diagonal. Linearisasi terhadap fungsi *state* dan fungsi observasi dilakukan melalui matriks Jacobian, sehingga diperoleh pendekatan linear yang digunakan untuk memprediksi *state* berikutnya sebelum observasi aktual tersedia. Prediksi *state* ini dihitung sesuai dengan Persamaan (31).

$$\hat{x}_t^- = f(\hat{x}_{t-1}) \quad (31)$$

Keterangan :

$\hat{x}_t^-$: prediksi *state* sebelum koreksi pada waktu ke- t

$f(\hat{x}_{t-1})$: prediksi *state* pada waktu ke- $t - 1$

Prediksi *state* menggunakan EKF melibatkan ketidakpastian pada dinamika *state* dan pengukuran, yang direpresentasikan oleh kovariansi *noise* proses Q_t dan kovariansi *noise* pengukuran R . Kovariansi ini menunjukkan seberapa besar kepercayaan terhadap model prediksi dan data observasi, Q_t mencerminkan ketidakpastian evolusi internal *state*, sementara R menggambarkan ketidakpastian pengukuran yang diterima. Estimasi Q_t dan R dilakukan secara bersamaan melalui pendekatan Maximum Likelihood Estimation (MLE), yang menilai konsistensi residual prediksi atau inovasi dengan asumsi distribusi Gaussian. Nilai kovariansi inovasi bergantung secara rekursif pada kovariansi prediksi dan Q_t , sedangkan R muncul langsung dalam perhitungan inovasi, sehingga tidak tersedia solusi analitik tertutup untuk memperoleh keduanya secara terpisah. Selanjutnya, kovariansi *error* prediksi pada waktu ke- t dihitung dengan menggabungkan kontribusi ketidakpastian *state* sebelumnya dan ketidakpastian proses, sebagaimana dirumuskan pada Persamaan (32) sebagai berikut.

$$P_t^- = A_t P_{t-1} A_t^T + W_t \hat{Q} W_t^T \quad (32)$$

Keterangan :

P_t^- : kovariansi *error* prediksi pada waktu ke- t

P_{t-1} : kovariansi *error* prediksi pada waktu ke- $t - 1$

A_t : matriks Jacobian fungsi $f(\cdot)$ terhadap x_{t-1}

$\hat{Q}$: matriks Jacobian dari matriks Jacobian A_t

W_t : matriks Jacobian dari $f(\cdot)$ terhadap *noise* proses w_t ,

$\hat{Q}$: matriks Jacobian W_t

$\hat{Q}$: kovariansi *noise* proses

A_t dan W_t didefinisikan sebagai:



$$A_t = \left[\frac{\partial f_i(x_{t-1})}{\partial x_j} \right]_{x=\hat{x}_{t-1}}, W_t = \left[\frac{\partial f_i}{\partial w_j} \right]$$

Setelah kovariansi *error* prediksi diperoleh, EKF memasuki tahap koreksi dengan menghitung *kalman gain* untuk menentukan seberapa besar pengukuran terbaru memperbarui prediksi *state*, sebagaimana dirumuskan pada Persamaan (33):

$$K_t = P_t^- H_t^T (H_t P_t^- H_t^T + V_t R_t V_t^T)^{-1} \quad (33)$$

Keterangan :

K_t : *kalman gain* pada waktu ke- t
 H_t : matriks Jacobian dari fungsi $h(\cdot)$ terhadap state x_t
 V_t : matriks Jacobian dari fungsi $h(\cdot)$ terhadap *noise* observasi v_t
 R_t : kovariansi *noise* observasi pada waktu ke- t
 P_t^- : kovariansi *error* prediksi pada waktu ke- t
 dengan matriks Jacobian H_t dan V_t didefinisikan sebagai:

$$H_t = \left[\frac{\partial h_i(x_t)}{\partial x_j} \right]_{x=\hat{x}_{t-1}}, V_t = \left[\frac{\partial h_i}{\partial v_j} \right]$$

Kalman gain yang diperoleh kemudian digunakan untuk memperbarui prediksi *state* dengan menggabungkan hasil prediksi sebelumnya dan informasi dari observasi aktual. Pembaruan dilakukan berdasarkan selisih antara nilai observasi dan nilai prediksi observasi, sehingga prediksi *state* dapat dikoreksi sesuai dengan informasi baru yang tersedia. Besarnya koreksi ditentukan oleh *Kalman Gain*, yang mengatur seberapa besar pengaruh observasi terhadap hasil prediksi. Proses pembaruan prediksi *state* pada waktu ke- t ditunjukkan pada Persamaan (34).

$$\hat{x}_t = \hat{x}_t^- + K_t(e_t - h(\hat{x}_t^-)) \quad (34)$$

Keterangan :

$\hat{x}_t$: prediksi *state* setelah koreksi pada waktu ke- t
 $\hat{x}_t^-$: prediksi *state* setelah koreksi pada waktu ke- t
 K_t : *kalman gain* pada waktu ke- t
 $h(\hat{x}_t^-)$: prediksi observasi dari *state* sebelum koreksi

Pembaruan *state* tersebut diikuti dengan pembaruan kovariansi *error* P_t yang merepresentasikan tingkat ketidakpastian prediksi *state* setelah proses koreksi, sebagaimana ditunjukkan pada Persamaan (35) berikut.

$$P_t = (I - K_t H_t) P_t^- \quad (35)$$



iansi *error* koreksi
 iansi *error* prediksi pada waktu ke- t
 <s Jacobian dari fungsi $h(\cdot)$ terhadap state x_t
 <s identitas

Secara keseluruhan, EKF bekerja secara rekursif dengan mengulangi tahapan prediksi dan koreksi pada setiap waktu pengamatan, di mana hasil prediksi *state* dan kovariansi *error* digunakan sebagai masukan pada iterasi berikutnya.

1.5.9 Model *Hybrid Autoregressive Integrated Moving Average* dan *Extended Kalman Filter*

Model *hybrid* ARFIMA-EKF merupakan pendekatan gabungan yang memanfaatkan keunggulan ARFIMA dalam menangkap sifat *long memory* pada data deret waktu serta kemampuan EKF dalam memperhalus hasil peramalan secara rekursif. ARFIMA digunakan untuk membangun model linear yang menggambarkan ketergantungan *long memory* antar observasi, sedangkan EKF berfungsi mengoreksi kesalahan peramalan dari ARFIMA berdasarkan data observasi terbaru. Dengan demikian, model ini mampu menghasilkan peramalan yang lebih akurat dibandingkan penggunaan model tunggal. Peramalan awal pada model *hybrid* diperoleh dengan menggunakan model ARFIMA sebagaimana dinyatakan pada Persamaan (8), dengan nilai koefisien autoregressive, moving average, dan differencing fraksional yang diperoleh melalui proses estimasi model berdasarkan fungsi log-likelihood pada Persamaan (24). Berdasarkan model ARFIMA (p, d, q), peramalan satu langkah ke depan dirumuskan pada Persamaan (36).

$$\hat{Y}_{t+1}^{ARFIMA} = \sum_{i=1}^p \phi_i Y_{t+1-i} + \sum_{k=1}^K \pi_k Y_{t+1-k} + \sum_{j=1}^q \theta_j \hat{e}_{t+1-j} \quad (36)$$

Dengan Y_t menyatakan data aktual pada waktu ke- t , ϕ_i merupakan koefisien autoregressive, θ_j merupakan koefisien moving average, π_k merupakan koefisien *fractional differencing*, serta $\hat{e}_t$ menyatakan residual hasil peramalan ARFIMA. Selisih antara data aktual dan hasil peramalan ARFIMA menghasilkan residual sebagaimana didefinisikan pada Persamaan (9). Residual tersebut masih mengandung dinamika nonlinear sehingga selanjutnya dimodelkan dan diperbarui menggunakan EKF sebagai komponen koreksi sesuai Persamaan (37):

$$\hat{e}_{EKF} = e_t - h(\hat{x}_t^-) \quad (37)$$

Di mana e_t merupakan residual ARFIMA, sedangkan $h(\hat{x}_t^-)$ adalah prediksi observasi dari *state* sebelum koreksi. Secara umum, bentuk model *hybrid* dinyatakan pada Persamaan (38) sebagai berikut (G. P. Zhang, 2003).

$$\hat{Y}_{Hybrid} = \hat{Y}_{ARFIMA} + \hat{e}_{EKF} \quad (38)$$



linear $\hat{Y}_{ARFIMA}$ diperoleh dari Persamaan (36), sedangkan nonlinear $\hat{e}_{EKF}$ diperoleh dari Persamaan (37). Dengan kedua komponen tersebut, prediksi akhir $\hat{Y}_{Hybrid}$ mampu memanfaatkan keunggulan ARFIMA dalam menangkap pola *long memory* dengan baik, serta EKF dalam memperhalus dinamika nonlinear, sehingga meningkatkan akurasi peramalan deret waktu.

1.5.10 Akurasi Model Peramalan

Akurasi model peramalan adalah proses menilai ketepatan hasil peramalan dengan membandingkan nilai aktual dan nilai peramalan. Penilaian ini penting untuk mengetahui sejauh mana model mampu merepresentasikan pola data deret waktu dan memberikan peramalan yang dapat diandalkan. Ukuran akurasi yang umum digunakan adalah *Mean Absolute Percentage Error* (MAPE) dan korelasi *Pearson*.

1.5.10.1 Mean Absolute Percentage Error

Mean Absolute Percentage Error (MAPE) merupakan ukuran akurasi peramalan yang dinyatakan dalam bentuk persentase. Indikator ini digunakan untuk mengukur seberapa besar kesalahan rata-rata antara nilai aktual dan nilai hasil peramalan terhadap nilai aktualnya. MAPE memberikan interpretasi yang mudah dipahami karena hasilnya dinyatakan dalam satuan persen, sehingga memudahkan perbandingan antar model. Formulasi nilai MAPE dapat dilihat pada Persamaan (39) berikut (Montgomery et al., 2016).

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100\% \quad (39)$$

Keterangan :

Y_t : nilai aktual pada waktu ke- t
 $\hat{Y}_t$: nilai peramalan pada waktu ke- t
 T : jumlah observasi

Kriteria nilai MAPE disajikan pada Tabel 1.

Tabel 1. Kriteria Nilai MAPE

Nilai MAPE	Kriteria
< 10%	Kemampuan Peramalan Sangat Baik
10% – 20%	Kemampuan Peramalan Baik
20% – 50%	Kemampuan Peramalan Cukup Baik
> 50%	Kemampuan Peramalan Buruk

Tabel 1 memberikan acuan interpretasi MAPE untuk menilai akurasi peramalan. Semakin kecil nilai MAPE, semakin akurat model dalam merepresentasikan data aktual.

1.5.10.2 Korelasi Pearson

Korelasi merupakan suatu teknik analisis statistik yang digunakan untuk mengetahui tingkat keeratan hubungan antara dua variabel kuantitatif. Analisis ini melihat apakah terdapat hubungan linear antara variabel bebas dan terikat untuk menentukan arah anatar variabel. (Arif Rachman et al., 2024). Formulasi nilai korelasi Pearson disajikan pada Persamaan (40) berikut (Putu et al., 2024).



$$r_{Y_t \hat{Y}_t} = \frac{T \cdot \sum Y_t \hat{Y}_t - (\sum Y_t) \cdot (\sum \hat{Y}_t)}{\sqrt{[T \cdot \sum Y_t^2 - (\sum Y_t)^2] \cdot [T \cdot \sum (\hat{Y}_t)^2 - (\sum \hat{Y}_t)^2]}} \quad (40)$$

Keterangan :

Y_t : nilai aktual pada waktu ke- t

$\hat{Y}_t$: hasil peramalan pada waktu ke- t

T : jumlah observasi

Kriteria penilaian korelasi disajikan Tabel 2 (Jabnabillah & Margina, 2022).

Tabel 2. Kriteria Penilaian Hubungan Korelasi

Interval Koefisien	Kriteria
0,00 – 0,199	Sangat Rendah
0,20 – 0,399	Rendah
0,40 – 0,599	Sedang
0,60 – 0,799	Kuat
0,80 – 0,999	Sangat Kuat

Tabel 2 memberikan pedoman interpretasi nilai koefisien korelasi *Pearson* untuk menilai kekuatan dan arah hubungan antara data aktual dan hasil peramalan. Nilai korelasi yang sangat kuat menunjukkan bahwa model peramalan mampu mengikuti tren data aktual dengan baik, sedangkan nilai korelasi yang sangat rendah menandakan bahwa peramalan kurang mencerminkan pola data aktual. Kriteria ini memudahkan peneliti dalam mengevaluasi kinerja model serta membandingkan akurasi antar model.



BAB II METODOLOGI PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa harga emas *futures* mingguan. Data tersebut diperoleh dari *website* resmi <https://www.investing.com/>. Periode data yang digunakan mencakup Januari 2014 hingga Januari 2024 dengan total 527 observasi. Data tersebut dapat dilihat pada Lampiran 1.

2.2 Identifikasi Variabel

Variabel dependen dalam penelitian ini adalah harga emas *futures* mingguan periode Januari 2014–Januari 2024. Penelitian ini menggunakan pendekatan deret waktu univariat sehingga tidak terdapat variabel independen.

2.3 Analisis Data

Analisis data pada penelitian ini dilakukan untuk membangun dan menguji model peramalan harga emas *futures* menggunakan model *hybrid* ARFIMA-KF. Data yang digunakan dibagi menjadi dua bagian, yaitu data *training* pada periode 5 Januari 2014 hingga 16 Januari 2022 untuk membentuk model, serta data *testing* pada periode 23 Januari 2022 hingga 1 Januari 2024 untuk mengevaluasi kinerja model terhadap data aktual. Seluruh pengolahan data, perhitungan model, dan peramalan dilakukan menggunakan *software RStudio*. Langkah-langkah analisis data yang dilakukan dalam penelitian adalah sebagai berikut.

A. Deskriptif Data

1. Menampilkan plot deret waktu untuk mengidentifikasi pola, tren, dan siklus musiman pada data harga emas.

B. Identifikasi Stasioneritas

1. Menampilkan plot ACF dan PACF untuk melihat adanya korelasi antarperiode.
2. Melakukan uji stasioneritas terhadap rata-rata menggunakan uji ADF sesuai Persamaan (1) dengan menolak H_0 Jika $t_{hitung} > t_{tabel}$ dan data dianggap stasioner.

C. Pemodelan ARFIMA

1. Melakukan identifikasi sifat *long memory* pada data dengan menghitung nilai H menggunakan metode *R/S Analysis* sesuai Persamaan (18).
2. Menentukan nilai dearajat *fractional differencing* (d) berdasarkan hubungan $d = H - 0,5$ untuk mengukur tingkat ketergantungan *long memory* pada data sesuai Persamaan (19).
3. Melakukan identifikasi orde p dan q dengan menganalisis pola pada ACF dan PACF.



estimasi parameter model $ARFIMA(p, d, q)$ menggunakan orde p dan q yang diusulkan oleh metode *Exact Maximum EML* sesuai pada Persamaan (23).

uji residual untuk memastikan bahwa residual model *white noise* sesuai Persamaan (2) dan Persamaan (3). Kriteria uji

yang digunakan adalah menolak H_0 apabila $Q > X^2$ atau $p - value < 0,05$ untuk uji *Ljung-Box* dan menolak H_0 jika $D_{cal} > D_{tabel}$ atau $p - value < 0,05$ untuk uji *Kolmogorov-Smirnov*.

6. Memilih model terbaik berdasarkan nilai BIC sesuai pada Persamaan (25).
- D. Identifikasi Linearitas
 1. Melakukan uji linearitas menggunakan *Ramsey RESET* untuk mendeteksi adanya hubungan nonlinear yang belum tertangkap oleh model ARFIMA sesuai Persamaan (7) dengan menolak H_0 jika $RESET > F_{t-1, T-t}^\alpha$ atau $p - value < 0,05$.
- E. Pemodelan EKF
 1. Menggunakan residual hasil estimasi model ARFIMA sebagai *input* pada proses EKF.
 2. Menentukan bentuk *hidden state* x_t dan observasi z_t sesuai pada Persamaan (29) dan Persamaan (30).
 3. Menetapkan nilai x_0 dan matriks kovariansi P_0 yang digunakan sebagai *input* awal EKF.
 4. Menentukan matriks kovariansi proses Q dan kovariansi observasi R .
 5. Melakukan tahap prediksi sesuai Persamaan (31) dan Persamaan (32).
 6. Melakukan tahap koreksi sesuai Persamaan (33), Persamaan (34), dan Persamaan (35).
- F. Pemodelan *Hybrid* ARFIMA-EKF
 1. Mendapatkan model *hybrid* dengan menggabungkan hasil model ARFIMA dan EKF, di mana residual ARFIMA digunakan sebagai *input* EKF sesuai Persamaan (38).
 2. Melakukan peramalan harga emas dari model *hybrid* ARFIMA–EKF.
 3. Melakukan evaluasi akurasi model *hybrid* ARFIMA–EKF berdasarkan MAPE dan Korelasi sesuai Persamaan (39) dan (40).

