

Implementasi ARIMAX dengan Seleksi Fitur Cross-Correlation dan Granger Causality untuk Prediksi Inflasi

Dian Maharani¹, Anggraini Puspita Sari^{2*}, Muhammad Muharrom Al Haromainy³

^{1,2,3} Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur

¹22081010184@student.upnjatim.ac.id

³ muhammad.muharrom.if@upnjatim.ac.id

*Corresponding author email: anggraini.puspita.if@upnjatim.ac.id

Abstrak— Inflasi merupakan indikator makroekonomi penting yang mencerminkan stabilitas harga dalam suatu perekonomian. Peramalan inflasi menjadi krusial dalam mendukung pengambilan kebijakan ekonomi yang tepat. Penelitian ini bertujuan untuk membangun model prediksi inflasi menggunakan pendekatan ARIMAX dengan seleksi fitur berbasis *Cross-Correlation Function* (CCF) dan uji kausalitas *Granger*. Data yang digunakan meliputi inflasi serta beberapa variabel makroekonomi seperti suku bunga, nilai tukar, jumlah uang beredar, ekspektasi inflasi, *administered price*, dan *volatile food* pada periode 2015–2024, yaitu sebanyak 120 baris data. Hasil penelitian menunjukkan bahwa model ARIMAX(0, 1, 1) merupakan model terbaik dengan nilai AIC dan BIC terendah. Evaluasi model menghasilkan nilai RMSE sebesar 0,056957 dan MAE sebesar 0,046109, yang menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang rendah. Secara keseluruhan, model mampu menangkap pola pergerakan inflasi dengan baik.

Kata Kunci— inflasi, ARIMAX, seleksi fitur, *cross-correlation function*, *granger causality*.

I. PENDAHULUAN

Inflasi adalah indikator makroekonomi fundamental yang mencerminkan tingkat harga umum dalam suatu perekonomian [1]. Pemahaman dan peramalan inflasi penting untuk pemerintah dan bank sentral dalam merancang kebijakan moneter dan menjaga stabilitas ekonomi [2]. Seiring berjalannya waktu, inflasi menunjukkan pola yang kompleks dan dinamis sehingga menjadi tantangan tersendiri untuk model yang menggunakan pendekatan *time series*.

Inflasi dipengaruhi oleh berbagai faktor makroekonomi seperti suku bunga (*BI Rate*), nilai tukar rupiah, jumlah uang beredar, ekspektasi inflasi, dan harga yang ditentukan pemerintah [3]. Variabel-variabel ini bisa memengaruhi inflasi dengan jeda waktu (*lag*) yang berbeda-beda dan dengan kekuatan pengaruh yang tidak sama. Karena itu, memasukkan variabel eksogen ke dalam deret waktu yang penting agar model bisa lebih mencerminkan kondisi ekonomi yang sebenarnya. Namun tidak semua variabel benar-benar berhubungan dengan inflasi. Jika memasukkan variabel yang tidak relevan, model justru akan lebih sulit dipahami dan hasil perhitungannya bisa menjadi bias [4].

Penelitian terdahulu umumnya menggunakan model ARIMA untuk prediksi inflasi karena kemampuannya untuk

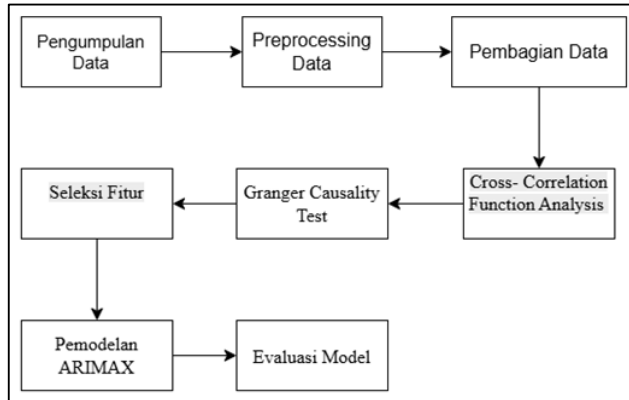
memodelkan data deret waktu univariat. Untuk memasukkan pengaruh variabel eksternal, model ARIMAX banyak digunakan dan diketahui mampu memberikan hasil prediksi yang lebih akurat dibandingkan dengan model ARIMA [5]. Meskipun demikian, banyak penelitian yang menggunakan variabel eksogen tanpa proses seleksi fitur. Untuk mengatasi masalah tersebut, penelitian ini menggunakan pendekatan untuk memilih variabel eksogen yang relevan sebelum pemodelan ARIMAX. Pertama, *Cross-Correlation Function* (CCF) digunakan untuk mengidentifikasi potensi hubungan dan struktur *lag* antara inflasi dan variabel kandidat. Kemudian, *Granger Causality Test* dilakukan untuk menilai variabel yang memiliki hubungan prediktif dengan inflasi [6]. Hanya variabel yang memenuhi kedua kriteria tersebut yang akan dimasukkan ke model ARIMAX.

Penelitian terdahulu menunjukkan bahwa model ARIMAX mampu meningkatkan akurasi prediksi inflasi melalui pemanfaatan variabel eksogen [5]. Namun, penelitian masih menentukan variabel eksogen berdasarkan pertimbangan teoritis atau ketersediaan data tanpa melakukan proses seleksi fitur yang sistematis. Kondisi tersebut berpotensi menyebabkan masuknya variabel yang kurang relevan ke dalam model, sehingga meningkatkan kompleksitas model dan dapat menurunkan kemampuan prediksi. Selain itu, penelitian yang mengombinasikan analisis *Cross-Correlation Function* (CCF) dan *Granger Causality* sebagai tahapan seleksi variabel eksogen sebelum pembentukan model ARIMAX masih relatif terbatas. Padahal, kedua metode tersebut dapat digunakan untuk mengidentifikasi hubungan *lag* dan hubungan prediktif antarvariabel sehingga pemilihan variabel eksogen menjadi lebih objektif dan berbasis data.

Penelitian ini menggunakan analisis *Cross-Correlation* dan *Granger Causality* sebagai kerangka pemilihan fitur dalam pemodelan ARIMAX. Pendekatan ini membantu mengidentifikasi variabel yang relevan serta hubungan dinamis di antaranya. Dengan demikian, proses pemodelan inflasi memiliki tahapan pemilihan variabel yang jelas dan dapat dijelaskan secara logis.

II. METODOLOGI

Penelitian ini terdiri dari beberapa tahapan yang menjadi bagian utama dalam pembangunan model peramalan, mulai dari pengumpulan data hingga tahap evaluasi model. Tahapan penelitian yang dilakukan dalam studi ini ditunjukkan pada Gbr 1.



Gbr. 1 Tahapan Penelitian

A. Pengumpulan Data

Data pada penelitian ini bersumber pada data yang dipublikasikan oleh Bank Indonesia dan Badan Pusat Statistik. Variabel dependen (target) yang digunakan yaitu inflasi. Sedangkan variabel independen (eksogen) yang digunakan dalam yaitu beberapa variabel makroekonomi seperti suku bunga, nilai tukar rupiah, jumlah uang beredar (M1 dan M2), ekspektasi inflasi 3 bulan dan 6 bulan (EI3 dan EI6), harga yang ditentukan oleh pemerintah, dan *volatile food*. Variabel tersebut dipilih berdasarkan relevansi teoritis terhadap dinamika inflasi dan umum digunakan dalam analisis makroekonomi. *Dataset* dikumpulkan selama periode waktu yang konsisten dengan frekuensi tetap yaitu tahun 2015-2024 dengan jumlah data sebanyak 120 baris.

B. Preprocessing Data

Sebelum pemodelan, *preprocessing* perlu dilakukan untuk memastikan kualitas dan kesesuaian data *time series*. Pertama, mengidentifikasi nilai hilang. Apabila terdapat nilai kosong perlu dilakukan penanganan dengan teknik yang sesuai tergantung pada karakteristik data. Kedua, transformasi data untuk menstabilkan varians dan mencapai stasioneritas. Transformasi dapat dilakukan menggunakan log transformasi dan *differencing*. Proses *differencing* didefinisikan sebagai [7]:

$$\Delta y_t = y_t - y_{t-1}$$

Ketiga, uji stasioneritas dilakukan menggunakan *Augmented Dickey-Fuller (ADF) test* dan *Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test*. Pada uji ADF, data *time series* dianggap stasioner jika hipotesis nol adanya akar unit ditolak (nilai *p-value* < 0,05). Sementara itu, pada uji KPSS, hipotesis nol menyatakan bahwa data bersifat stasioner sehingga data dianggap stasioner apabila nilai *p-value* > 0,05. Penggunaan kedua uji tersebut bertujuan untuk mengevaluasi stasioneritas data dari sudut pandang yang berbeda karena masing-masing memiliki hipotesis nol yang berlawanan. Apabila data belum

stasioner, maka dilakukan transformasi seperti *differencing* hingga diperoleh data yang memenuhi asumsi stasioneritas.

C. Pembagian Data

Data dibagi menjadi data latih (*training set*) dan data uji (*testing set*) dengan rasio 80:20, sehingga diperoleh 96 data untuk pelatihan model dan 24 data untuk pengujian. Evaluasi performa model dilakukan secara *out-of-sample* menggunakan data uji yang tidak dilibatkan selama proses pelatihan, sehingga hasil evaluasi mencerminkan kemampuan model dalam melakukan prediksi pada data yang belum pernah dilihat sebelumnya.

D. Cross-Correlation Function (CCF) Analysis

CCF digunakan sebagai langkah awal untuk mengidentifikasi hubungan potensial antara inflasi dan masing-masing variabel eksogennya. Tujuan analisis ini adalah untuk mendeteksi hubungan temporal antar variabel dan untuk mengidentifikasi struktur *lag* yang memiliki hubungan terkuat. Prosedur ini dimulai dengan menghitung nilai korelasi silang antara inflasi dan se tiap variabel kandidat selama beberapa periode waktu. Secara matematis, fungsi korelasi silang pada *lag* ke-*k* dinyatakan sebagai [8]:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sqrt{\gamma_x(0)\gamma_y(0)}}$$

$\gamma_{xy}(k)$ adalah kovarians silang pada *lag* ke-*k*, sedangkan $\gamma_x(0)$ dan $\gamma_y(0)$ adalah varians masing-masing deret waktu. Nilai korelasi yang signifikan menunjukkan adanya hubungan antar variabel pada *lag* tertentu. Berdasarkan hal tersebut, variabel dengan korelasi signifikan dipilih sebagai kandidat fitur beserta struktur *lag*-nya. Namun, karena CCF hanya menunjukkan hubungan korelasional dan bukan kausalitas, maka diperlukan uji lanjutan seperti kausalitas Granger untuk memastikan hubungan prediktif antar variabel.

E. Granger Causality Test

Granger Causality Test digunakan untuk memvalidasi relevansi variabel yang dipilih secara statistik, digunakan uji kausalitas Granger. Uji ini mengevaluasi apakah nilai-nilai masa lalu dari variabel eksogen dapat meningkatkan prediksi inflasi. Setiap variabel diuji terhadap inflasi menggunakan beberapa panjang *lag* (misalnya, *lag* 1 hingga *lag* 10). Hipotesis nol (H_0) menyatakan bahwa variabel tersebut tidak menyebabkan inflasi menurut kriteria *Granger*. Jika nilai *p* kurang dari 0,05, hipotesis nol (H_0) ditolak, yang menunjukkan bahwa variabel tersebut memiliki daya prediksi. *Granger Causality* didefinisikan sebagai [9]:

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j x_{t-j} + \varepsilon_t$$

F. Seleksi Fitur

Pemilihan fitur dalam penelitian ini dilakukan dengan menggabungkan hasil dari beberapa metode analisis, yaitu *Cross-Correlation Function (CCF)*, *Granger Causality Test*, dan regresi linear (OLS). Tahap awal dilakukan dengan menggunakan CCF untuk mengidentifikasi potensi hubungan *lag* antara variabel eksogen dan inflasi. Selanjutnya, *Granger Causality Test* digunakan untuk memvalidasi hubungan tersebut memiliki kemampuan prediksi secara statistik. Variabel yang memenuhi kriteria ini dipertimbangkan sebagai kandidat variabel dengan efek *lag*. Secara matematis dapat dituliskan sebagai berikut:

$$X = \{X_i | CCF_i \text{ significant} \cap \text{Granger}_i \text{ significant}\}$$

Selain itu, regresi linear (OLS) digunakan untuk mengidentifikasi pengaruh langsung (*contemporaneous effect*) dari variabel-variabel eksogen terhadap inflasi. Variabel eksogen dalam model ARIMAX dapat dimasukkan baik dalam bentuk *lag* maupun tanpa *lag*. Beberapa penelitian menunjukkan bahwa penggunaan nilai kontemporer (tanpa *lag*) dari variabel eksogen dapat memberikan hasil yang lebih baik dibandingkan penggunaan *lag* [10]. Oleh karena itu, regresi linear (OLS) digunakan untuk mengidentifikasi variabel yang memiliki pengaruh langsung terhadap inflasi dan variabel yang signifikan dalam model OLS dipertimbangkan sebagai kandidat variabel tanpa *lag*.

G. Pemodelan ARIMAX

Model peramalan yang digunakan dalam penelitian ini adalah *Autoregressive Integrated Moving Average with Exogenous Variables* (ARIMAX). Model ini menggabungkan nilai inflasi masa lalu dan variabel eksogen terpilih. Bentuk umum model ARIMAX didefinisikan sebagai [7]:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=1}^m \beta_k X_{k,t} + \varepsilon_t$$

Proses pemodelan melibatkan penentuan parameter optimal (p,d,q) menggunakan kriteria seperti AIC atau BIC. Variabel eksogen yang dipilih kemudian dimasukkan ke dalam model dengan struktur *lag* masing-masing.

H. Evaluasi Model

Evaluasi model pada penelitian ini dianalisis menggunakan metrik kesalahan, sebagai berikut:

- *Root Mean Squared Error (RMSE)*

RMSE merupakan metrik evaluasi yang digunakan untuk menghitung rata-rata kesalahan (*error*) prediksi dengan memberikan bobot lebih besar terhadap kesalahan yang besar. Kesalahan (*error*) dengan nilai besar akan memiliki pengaruh lebih signifikan terhadap nilai RMSE keseluruhan karena menggunakan kuadrat *error*. RMSE dirumuskan sebagai [11]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

- *Mean Absolute Error (MAE)*

MAE mengukur rata-rata selisih absolut antara nilai aktual dan prediksi, sehingga lebih mudah dipahami karena tidak dipengaruhi oleh kuadrat *error*. MAE dirumuskan sebagai [11]:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

- *Koefisien Determinan (R²)*

R² digunakan untuk mengukur kemampuan model dalam menjelaskan variasi yang terdapat pada data aktual. Nilai R² berada pada rentang 0 hingga 1, Semakin tinggi nilai R² menunjukkan model memiliki kemampuan penjelasan yang semakin baik. R² dirumuskan sebagai [12]:

$$R^2 = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y})^2}$$

Metrik tersebut memberikan evaluasi komprehensif dengan mengukur baik besarnya kesalahan prediksi.

III. HASIL DAN PEMBAHASAN

Hasil dan pembahasan pada penelitian ini diuraikan berdasarkan tahapan penelitian yang telah dirancang.

A. Hasil Uji Stasioneritas

Data yang digunakan terlebih dahulu melalui tahap *preprocessing* untuk memastikan kualitas data sebelum dilakukan pemodelan. Tahapan ini meliputi transformasi variabel, seperti *log transformation* dan *differencing* untuk mencapai kondisi yang lebih stabil. Hasil uji stasioneritas menggunakan metode ADF dan KPSS ditunjukkan pada Tabel 1.

TABEL I
HASIL UJI STASIONERITAS

Variabel	ADF Test	KPSS Test
inflasi	0.153	0.100
$\Delta \log(M1)$	0.102	0.100
$\Delta \log(M2)$	0.074	0.100
$\Delta \log(\text{kurs})$	0.000	0.100
$\Delta(\text{suku bunga})$	0.000	0.099
$\Delta(EI3)$	0.000	0.100
$\Delta(EI6)$	0.000	0.100
adm price	0.000	0.100
volatile food	0.000	0.100

Hasil uji stasioneritas menunjukkan bahwa sebagian besar variabel telah memenuhi asumsi stasioneritas. Berdasarkan uji *Augmented Dickey-Fuller (ADF)*, variabel $\Delta \log(\text{kurs})$, $\Delta(\text{suku bunga})$, $\Delta(EI3)$, $\Delta(EI6)$, adm price, dan volatile food dinyatakan stasioner karena memiliki nilai *p-value* di bawah 0,05. Sementara itu, variabel inflasi, $\Delta \log(M1)$, dan $\Delta \log(M2)$ memiliki nilai *p-value* ADF di atas 0,05, yang mengindikasikan bahwa variabel tersebut belum sepenuhnya stasioner. Namun

demikian, berdasarkan uji KPSS, seluruh variabel memiliki nilai p -value di atas 0,05, sehingga dapat dianggap stasioner. Perbedaan hasil antara uji ADF dan KPSS disebabkan oleh perbedaan hipotesis dasar yang digunakan, ADF menguji keberadaan *unit root*, sedangkan KPSS menguji kestasioneran data. Hasil pengujian menunjukkan adanya perbedaan keputusan antara uji ADF dan KPSS pada beberapa variabel setelah proses *differencing*. Perbedaan tersebut dapat terjadi karena kedua uji memiliki hipotesis nol yang berbeda. Dalam penelitian ini, penentuan stasioneritas mengacu pada uji KPSS karena uji tersebut secara langsung menguji kondisi stasioneritas data melalui hipotesis nol bahwa data bersifat stasioner. Hasil uji KPSS menunjukkan bahwa seluruh variabel telah memenuhi asumsi stasioneritas sehingga dapat digunakan pada tahap pemodelan selanjutnya.

B. Hasil CCF

Analisis *Cross-Correlation Function* (CCF) dilakukan untuk mengidentifikasi hubungan awal antara variabel eksogen dan variabel inflasi serta menentukan *lag* yang berpotensi berpengaruh dalam model. Hasil analisis CCF untuk masing-masing variabel ditunjukkan pada Tabel 2.

TABEL II
HASIL CCF

Variabel	Lag Signifikan
$\Delta\log(M1)$	1, 7
$\Delta\log(M2)$	1, 7
$\Delta\log(\text{kurs})$	-
$\Delta(\text{suku bunga})$	5
$\Delta(EI3)$	1, 3, 4, 10
$\Delta(EI6)$	3, 4, 6, 9
adm price	10
volatile food	1, 3, 4, 6, 8, 9

Hasil analisis *Cross-Correlation Function* (CCF) ditunjukkan pada Tabel 2. Berdasarkan hasil tersebut, variabel $\Delta\log(M1)$ dan $\Delta\log(M2)$ menunjukkan korelasi signifikan pada *lag* ke-1 dan ke-7, yang mengindikasikan adanya pengaruh dalam jangka pendek (*lag* 1) serta potensi efek tertunda dalam jangka yang lebih panjang (*lag* 7). Variabel $\Delta(\text{suku bunga})$ menunjukkan korelasi pada *lag* ke-5, yang mencerminkan adanya pengaruh dalam jangka menengah terhadap inflasi. Variabel $\Delta(EI3)$ menunjukkan korelasi pada *lag* ke-1, ke-3, ke-4, dan ke-10, sedangkan variabel $\Delta(EI6)$ menunjukkan korelasi pada *lag* ke-3, ke-4, ke-6, dan ke-9, yang mengindikasikan bahwa ekspektasi inflasi memiliki pengaruh, baik dalam jangka pendek maupun jangka menengah hingga panjang.

Selain itu, variabel *administered price* menunjukkan korelasi pada *lag* ke-10, yang mengindikasikan adanya pengaruh dalam jangka panjang, kemungkinan terkait dengan kebijakan harga yang berdampak secara bertahap. Variabel *volatile food* menunjukkan korelasi pada *lag* ke-1, ke-3, ke-4, ke-6, ke-8, dan ke-9, yang mencerminkan adanya pengaruh yang bersifat fluktuatif, baik dalam jangka pendek maupun jangka yang lebih panjang. Sementara itu, variabel $\Delta\log(\text{kurs})$ tidak menunjukkan korelasi yang signifikan pada *lag* yang diuji, sehingga tidak

menunjukkan hubungan awal dengan inflasi. Seluruh *lag* yang diperoleh dari analisis CCF selanjutnya digunakan sebagai kandidat dalam uji kausalitas Granger.

C. Hasil Granger Causality

Uji kausalitas Granger dilakukan untuk mengidentifikasi hubungan prediktif antara variabel eksogen dan variabel inflasi. Uji ini bertujuan untuk mengetahui apakah suatu variabel dapat digunakan untuk memprediksi perubahan inflasi pada periode tertentu. Hasil uji kausalitas Granger ditunjukkan pada Tabel 3.

TABEL III
HASIL GRANGER CAUSALITY

Variabel	Lag Signifikan
$\Delta\log(M1)$	1
$\Delta\log(M2)$	1
$\Delta\log(\text{kurs})$	-
$\Delta(\text{suku bunga})$	-
$\Delta(EI3)$	1, 4
$\Delta(EI6)$	4, 6, 9
adm price	10
volatile food	-

Berdasarkan hasil uji kausalitas Granger, beberapa variabel menunjukkan hubungan prediktif yang signifikan terhadap inflasi pada *lag* tertentu. Variabel $\Delta\log(M1)$ dan $\Delta\log(M2)$ signifikan pada *lag* ke-1, yang mengindikasikan adanya pengaruh dalam jangka pendek terhadap inflasi. Hasil ini menunjukkan bahwa perubahan jumlah uang beredar dapat menjadi indikator awal perubahan inflasi. Secara teori, peningkatan jumlah uang beredar akan meningkatkan likuiditas dalam perekonomian dan mendorong permintaan agregat, yang pada akhirnya dapat menyebabkan kenaikan harga barang dan jasa. Temuan ini sejalan dengan teori kuantitas uang yang menyatakan bahwa pertumbuhan uang beredar memiliki keterkaitan dengan tingkat inflasi [13].

Selain itu, variabel *administered price* menunjukkan hubungan kausalitas pada *lag* ke-10, yang mengindikasikan adanya pengaruh dalam jangka panjang terhadap inflasi. Hal ini menunjukkan bahwa perubahan harga yang diatur pemerintah, seperti tarif listrik, bahan bakar, dan transportasi, juga memiliki pengaruh tidak langsung pada inflasi umum. Dampak kebijakan tersebut cenderung menyebar secara bertahap melalui peningkatan biaya produksi, distribusi, dan harga barang lainnya sehingga efeknya baru terlihat setelah beberapa periode. Kondisi ini sejalan dengan karakteristik komponen *administered prices* yang sering menjadi sumber tekanan inflasi jangka menengah hingga panjang di Indonesia [14].

Sementara itu, variabel $\Delta\log(\text{kurs})$, $\Delta(\text{suku bunga})$, dan *volatile food* tidak menunjukkan hubungan kausalitas yang signifikan terhadap inflasi, sehingga tidak dipertimbangkan dalam tahap selanjutnya. Ketidaksignifikanan variabel kurs dan suku bunga menunjukkan bahwa perubahan kedua variabel tersebut belum memiliki kemampuan prediktif yang konsisten terhadap inflasi selama periode pengamatan. Adapun pada variabel *volatile food*, meskipun sebelumnya menunjukkan korelasi pada analisis CCF, hasil uji Granger menunjukkan bahwa hubungan

tersebut tidak memiliki kemampuan prediktif yang signifikan setelah mempertimbangkan dinamika waktu antarvariabel. Berdasarkan hasil uji kausalitas Granger, hanya variabel yang memiliki hubungan signifikan yang selanjutnya digunakan sebagai variabel eksogen dalam model ARIMA

D. Analisis Pengaruh Langsung Variabel

Selain analisis berbasis *lag*, dilakukan analisis menggunakan regresi linear (OLS) untuk mengidentifikasi pengaruh langsung dari variabel eksogen terhadap inflasi.

TABEL IV
PENGARUH LANGSUNG VARIABEL

Variabel	Koefisien	p-value	Keterangan
$\Delta \log(M1)$	0.1543	0.8667	Tidak signifikan
$\Delta \log(M2)$	0.4408	0.8340	Tidak signifikan
$\Delta \log(\text{kurs})$	-0.3425	0.8080	Tidak signifikan
$\Delta(\text{suku bunga})$	0.2604	0.0584	Tidak signifikan
$\Delta(EI3)$	-0.0047	0.0337	Signifikan
$\Delta(EI6)$	0.0013	0.5981	Tidak signifikan
adm price	0.2305	0.0000	Signifikan
volatile food	0.1925	0.0000	Signifikan

Berdasarkan hasil estimasi OLS pada Tabel 4, hanya beberapa variabel yang menunjukkan pengaruh langsung yang signifikan terhadap inflasi. Variabel *administered price* dan *volatile food* memiliki pengaruh positif dan signifikan terhadap inflasi dengan nilai *p-value* sebesar 0,000. Hasil ini menunjukkan bahwa perubahan harga yang diatur pemerintah serta fluktuasi harga komoditas pangan bergejolak berkontribusi langsung terhadap perubahan inflasi. Kenaikan harga energi, transportasi, maupun komoditas pangan strategis dapat segera meningkatkan tingkat harga secara umum sehingga dampaknya langsung tercermin pada inflasi. Temuan ini sejalan dengan penelitian Farandy (2020) yang menunjukkan bahwa komponen *administered prices* dan *volatile food* merupakan sumber utama tekanan inflasi di Indonesia [14].

Selain itu, variabel $\Delta(EI3)$ menunjukkan pengaruh signifikan terhadap inflasi dengan nilai *p-value* sebesar 0,0337. Hasil ini mengindikasikan bahwa ekspektasi inflasi jangka pendek memiliki hubungan langsung dengan inflasi aktual. Ketika masyarakat dan pelaku usaha memperkirakan kenaikan harga pada periode mendatang, mereka cenderung melakukan penyesuaian konsumsi, harga jual, maupun upah lebih awal sehingga tekanan inflasi dapat muncul pada periode berjalan [15].

Sementara itu, variabel $\Delta \log(M1)$, $\Delta \log(M2)$, $\Delta \log(\text{kurs})$, $\Delta(\text{suku bunga})$, dan $\Delta(EI6)$ tidak menunjukkan pengaruh langsung yang signifikan terhadap inflasi. Hal ini menunjukkan bahwa variabel-variabel tersebut tidak berpengaruh secara langsung, namun kemungkinan memiliki pengaruh melalui mekanisme *lag*.

E. Seleksi Fitur

Seleksi variabel dalam penelitian ini dilakukan dengan menggabungkan hasil analisis berbasis *lag* dan tanpa *lag*. Variabel dengan efek *lag* diperoleh dari hasil *Cross-Correlation Function* (CCF) dan uji kausalitas Granger sebagai metode utama dalam pemilihan variabel. Selain itu, regresi linear (OLS) digunakan sebagai analisis tambahan untuk mengidentifikasi pengaruh langsung dari variabel eksogen terhadap inflasi.

TABEL V
HASIL SELEKSI FITUR

Variabel	CCF Lag	Granger Lag	OLS	Variabel Terpilih
$\Delta \log(M1)$	1, 7	1	Tidak signifikan	Lag 1
$\Delta \log(M2)$	1, 7	1	Tidak signifikan	Lag 1
$\Delta \log(\text{kurs})$	-	-	Tidak signifikan	-
$\Delta(\text{suku bunga})$	5	-	Tidak signifikan	-
$\Delta(EI3)$		1, 4	Signifikan	Lag 1 + tanpa Lag
$\Delta(EI6)$	1, 3, 4, 10	4, 6, 9	Tidak signifikan	Lag 4
adm price	3, 4, 6, 9	10	Signifikan	Tanpa Lag
volatile food	10	-	Signifikan	Tanpa Lag

Hasil seleksi variabel ditunjukkan pada Tabel 5, yang menyajikan hasil analisis CCF, uji kausalitas Granger, dan OLS secara terintegrasi. Berdasarkan tabel tersebut, tidak seluruh variabel dan *lag* yang signifikan digunakan dalam model. Pemilihan variabel akhir dilakukan dengan mempertimbangkan signifikansi, potensi redundansi, serta stabilitas model.

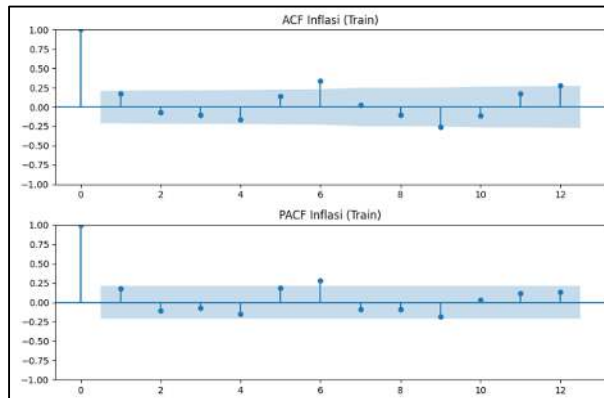
F. Pemodelan ARIMAX

Berdasarkan hasil seleksi variabel, model ARIMAX dibangun dengan memasukkan variabel eksogen yang terdiri dari variabel dengan efek *lag* dan tanpa *lag*. Variabel dengan efek *lag* diperoleh dari hasil analisis CCF dan uji kausalitas Granger, sedangkan variabel tanpa *lag* diperoleh dari hasil regresi linear (OLS) sebagai analisis tambahan.

Variabel eksogen yang digunakan dalam model terdiri dari:

- **Variabel dengan efek lag:** $m1_lag1$, $m2_lag1$, $EH3_lag1$, $EH6_lag4$, adm_lag10
- **Variabel tanpa lag:** Adm_Price , Vol_Food , dan $\Delta(EI3)$

Penentuan parameter ARIMA dilakukan dengan menganalisis pola ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*) yang ditunjukkan pada Gbr 2.



Gbr. 2 Grafik ACF dan PACF

Berdasarkan plot ACF dan PACF, tidak terlihat pola *spike* yang signifikan pada beberapa *lag* awal. Nilai autokorelasi pada ACF cenderung menurun secara cepat, sedangkan pada PACF tidak terlihat adanya *cut-off* yang jelas. Hal ini menunjukkan bahwa komponen *autoregressive* (AR) dan *moving average* (MA) dalam model relatif lemah, sehingga parameter yang digunakan cenderung kecil. Oleh karena itu, model ARIMA yang digunakan memiliki orde rendah, seperti ARIMAX(1,1,0), ARIMAX(0,1,1), dan ARIMAX(1,1,1).

Pemilihan model terbaik dilakukan berdasarkan nilai *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC). Model dengan nilai AIC dan BIC yang lebih kecil dipilih sebagai model terbaik.

TABEL VI
PERBANDINGAN AIC DAN BIC SETIAP MODEL

Model	AIC	BIC
ARIMAX(0, 1, 1)	-170.868	-146.325
ARIMAX(1, 1, 0)	-169.319	-144.776
ARIMAX(1, 1, 1)	-167.572	-140.574

Berdasarkan hasil perbandingan model yang ditunjukkan pada Tabel 6, model ARIMAX(0,1,1) memiliki nilai AIC dan BIC paling kecil dibandingkan model lainnya. Hal ini menunjukkan bahwa model tersebut memiliki keseimbangan terbaik antara kompleksitas model dan kemampuan dalam menjelaskan data. Oleh karena itu, model ARIMAX(0,1,1) dipilih sebagai model terbaik dalam penelitian ini.

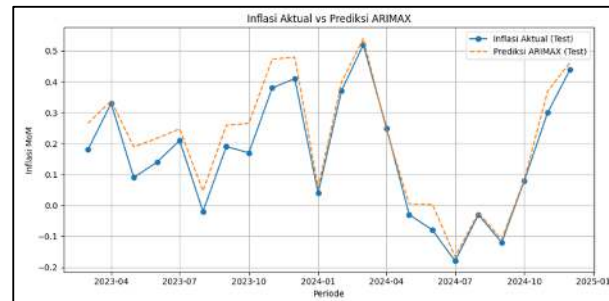
G. Evaluasi Model

Setelah model ARIMAX terbaik diperoleh berdasarkan kriteria AIC dan BIC, selanjutnya dilakukan evaluasi untuk mengukur kinerja model dalam memprediksi inflasi. Hasil evaluasi model ditunjukkan pada Gbr 3.

EVALUASI MODEL ARIMAX	
RMSE :	0.056957
MAE :	0.046109
R ² :	0.911358

Gbr. 3 Hasil Evaluasi Model

Hasil evaluasi model menunjukkan nilai RMSE sebesar 0,053525 dan MAE sebesar 0,040410, yang berarti rata-rata kesalahan prediksi model sekitar 0,04–0,05 poin persentase karena data inflasi dinyatakan dalam satuan persen (%). Selain itu, nilai koefisien determinasi (R^2) sebesar 0,911358 menunjukkan bahwa model mampu menjelaskan sebagian besar variasi data inflasi.



Gbr. 4 Grafik Perbandingan Inflasi Aktual dan Prediksi

Perbandingan antara nilai inflasi aktual dan hasil prediksi model ARIMAX ditunjukkan pada Gbr 4. Secara umum, model mampu mengikuti pola pergerakan inflasi dengan cukup baik, yang terlihat dari kedekatan antara garis prediksi dan nilai aktual pada sebagian besar periode pengamatan. Model berhasil menangkap pola naik dan turun inflasi, terutama pada periode kenaikan signifikan di awal tahun 2024 serta penurunan yang terjadi pada pertengahan tahun 2024. Hal ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mengikuti dinamika perubahan inflasi. Meskipun demikian, masih terdapat beberapa perbedaan antara nilai aktual dan hasil prediksi pada beberapa periode. Pada kondisi tertentu, model terkadang memprediksi terlalu tinggi atau terlalu rendah dibandingkan nilai sebenarnya. Namun, selisih tersebut relatif kecil dan tidak mengubah pola utama dari data inflasi.

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan model ARIMAX dengan pendekatan seleksi fitur berbasis *Cross-Correlation Function* (CCF) dan uji kausalitas Granger untuk memprediksi inflasi. Hasil analisis menunjukkan bahwa tidak semua variabel memiliki pengaruh langsung terhadap inflasi, namun beberapa variabel memiliki hubungan prediktif melalui mekanisme *lag*. Selain itu, analisis tambahan menggunakan regresi linear (OLS) menunjukkan bahwa variabel *administered price*, *volatile food*, dan ekspektasi inflasi jangka pendek memiliki pengaruh langsung yang signifikan terhadap inflasi. Berdasarkan hasil pemilihan model menggunakan kriteria AIC dan BIC, model ARIMAX(0, 1, 1) dipilih sebagai model terbaik. Evaluasi model menunjukkan performa yang baik dengan nilai RMSE sebesar 0,056957, MAE sebesar 0,046109, dan R^2 sebesar 0,911358, serta kemampuan model dalam mengikuti pola pergerakan inflasi secara visual. Dengan demikian, pendekatan seleksi fitur yang digunakan dalam penelitian ini mampu meningkatkan kualitas pemodelan inflasi,

karena dapat mengidentifikasi variabel yang relevan baik dalam bentuk *lag* maupun tanpa *lag*. Model yang dihasilkan diharapkan dapat menjadi alternatif dalam peramalan inflasi yang lebih akurat dan informatif.

REFERENSI

- [1] Mukarramah, M. Zulkarnain, and Arif Ramadhana, "Inflasi Dalam Pertumbuhan Ekonomi: Studi Empiris di Indonesia Periode 1991 – 2024," vol. 25, Jul. 2025, Accessed: Jan. 24, 2026. [Online]. Available: <http://jurnal.umsu.ac.id/index.php/ekawan>
- [2] S. M. Juhro and B. N. Iyke, "Forecasting Indonesian inflation within an inflation-targeting framework: Do large-scale models pay off?," *Buletin Ekonomi Moneter dan Perbankan/Monetary and banking economics bulletin*, vol. 22, no. 4, pp. 423–436, Feb. 2019, doi: 10.21098/bemp.v22i4.1235.
- [3] Afrizal, "Analysis of Macroeconomic Variables on the Existence of Inflation in Indonesia Using the Vector Error Correction Model Approach," *Jurnal Ekonomi Kuantitatif Terapan*, vol. 17, no. 1, pp. 9–35, Oct. 2025, doi: 10.24843/JEKT.2024.v17.i01.p02.
- [4] B. Novianda, "Forecasting Inflation in Indonesia Using ARIMAX : The Role of Money Supply and Exchange Rate," vol. 13, no. 02, 2025, doi: 10.33019/equity.v13i2.370.
- [5] M. Coker, "Short-Term Inflation Forecasting In Sierra Leone: A Comparison of Vector Autoregressive VAR(P), Arimax, And Arima Models," *SSRN Electronic Journal*, 2025, doi: 10.2139/ssrn.5278590.
- [6] J. Olaniyan, D. Olaniyan, I. C. Obagbuwa, B. M. Esiefarienrhe, A. Adebisiyi, and O. P. Bernard, "Intelligent Financial Forecasting with Granger Causality and Correlation Analysis Using Bayesian Optimization and Long Short-Term Memory," *Electronics (Basel)*, vol. 13, no. 22, p. 4408, Nov. 2024, doi: 10.3390/electronics13224408.
- [7] W. W. S. . Wei, *Time series analysis : univariate and multivariate methods*. Pearson Addison Wesley, 2006.
- [8] H. C. Changtzen, D. F. Chang, and Y. H. Lo, "Detecting concurrent relationships of selected time series data for ARIMAX model," *ICIC Express Letters, Part B: Applications*, vol. 10, no. 10, pp. 937–944, Oct. 2019, doi: 10.24507/iceiclb.10.10.937.
- [9] C. F. Akuma, P. Hewage, T. Bren, C. V. Amaechi, and A. Eche, "Forecasting Nigerian Inflation: A SARIMAX and Granger Causality Analysis of Exchange Rate and Crude Oil Price Impacts," *International Journal of Development Mathematics (IJDM)*, vol. 2, no. 4, pp. 200–216, Dec. 2025, doi: 10.62054/ijdm/0204.14.
- [10] S. Ghosh, S. Mukhoti, and P. Sharma, "Impact of rainfall risk on rice production: realized volatility in mean model," Apr. 2025.
- [11] Y. Si, S. Nadarajah, Z. Zhang, and C. Xu, "Modeling opening price spread of Shanghai Composite Index based on ARIMA-GRU/LSTM hybrid model," *PLoS One*, vol. 19, no. 3 March, Mar. 2024, doi: 10.1371/journal.pone.0299164.
- [12] D. C. . Montgomery, C. L. . Jennings, and Murat. Kulahci, *Introduction to time series analysis and forecasting*. Wiley, 2015.
- [13] F. Mishkin, "The Causes of Inflation," Cambridge, MA, Sep. 1984. doi: 10.3386/w1453.
- [14] A. R. Farandy, "Analyzing Factors Affecting Indonesian Food Price Inflation," *Jurnal Ekonomi dan Pembangunan*, vol. 28, no. 1, pp. 65–76, Dec. 2020, doi: 10.14203/JEP.28.1.2020.65-76.
- E. D. Tjahjono, H. Harmanta, and N. M. A. Purwato, "Survey Measures of Inflation Expectation," *Buletin Ekonomi Moneter dan Perbankan*, vol. 14, no. 4, pp. 421–456, Jun. 2012, doi: 10.21098/bemp.v14i4.366.