Peramalan Harga Bahan Pangan pada 34 Provinsi di Indonesia

Arkavidia: Datavidia

Hilmy Rahmadani
Departemen Matematika
Universitas Indonesia
Depok, Indonesia
rahmadanihilmy@gmail.com

Mohammad Raffy Zeidan
Departemen Matematika
Universitas Indonesia
Depok, Indonesia
raffy.zeidan@gmail.com

Farhan Akhtar Gymnasiar
Departemen Matematika
Universitas Indonesia
Depok, Indonesia
f.akhtar26@outlook.com

***Abstract*—Peramalan harga bahan pangan pokok bermanfaat untuk tetap menjaga fluktuasi harga bahan pangan sehingga tidak menyengsarakan masyarakat. Tujuan dari penelitian ini adalah untuk mengevaluasi model gabungan ARIMA, SARIMAX, dan exponential smoothing dalam tugas peramalan harga bahan pangan pada setiap provinsi yang ada di Indonesia. Hasil yang didapat adalah Model gabungan kompleks (ARIMA + SARIMA + Exponential Smoothing), yang diusulkan, dengan parameter yang lebih banyak menunjukkan performa terbaik dengan MAPE terendah sebesar 0.02888. Hasil peramalan menunjukkan bahwa model yang digunakan berhasil menangkap pola harga 13 komoditas di 34 provinsi dengan cukup baik. Meskipun model menunjukkan performa yang baik dalam memprediksi pola harga, terdapat beberapa variasi antarprovinsi yang mengindikasikan adanya faktor lokal yang berperan dalam pergerakan harga.**

***Index Terms*—Forecasting, Peramalan harga bahan pangan, ARIMA, SARIMAX, Time Series Analysis**

I. Pendahuluan

Harga bahan pokok selalu menjadi permasalahan utama setiap pembuat kebijakan. Fluktuasi harga dapat berdampak pada berbagai sektor, seperti inflasi, kesejahteraan masyarakat dan stabilitas ekonomi. Oleh karena itu, peramalan harga menjadi instrumen penting dalam perencanaan ekonomi dan kebijakan publik dari suatu negara [1].

Fluktuasi harga komoditas dipengaruhi oleh berbagai faktor, termasuk kondisi ekonomi global, kebijakan pemerintah, perubahan permintaan dan penawaran, serta faktor musiman. Secara prinsip fluktuasi harga komoditas pangan pada dasarnya terjadi akibat ketidakseimbangan antara jumlah pasokan dan jumlah permintaan yang dibutuhkan konsumen (Bahtiar, 2023). Salah satu karakteristik utama seperti perubahan dalam pola konsumsi dan produksi, serta meningkatnya peran komoditas dalam pasar keuangan, turut memengaruhi volatilitas harga. Faktor siklis, seperti guncangan permintaan dan penawaran global, juga berkontribusi pada fluktuasi harga komoditas bahan pangan dalam jangka pendek maupun panjang.

Peramalan harga dengan akurasi yang tinggi masih menjadi suatu tantangan. Model peramalan berbasis data historis, seperti ARIMA, SARIMAX, telah banyak digunakan untuk memprediksi harga komoditas. Penelitian yang dilakukan oleh Dutta dan Maiti (2021) menunjukkan bahwa model ARIMA dalam tugas meramalkan harga produk pertanian seperti bawang dan kentang di India mendapatkan nilai MAPE sebesar 7.86% untuk kentang dengan parameter ARIMA(3,1,0) dan 9.85% untuk bawang merah dengan parameter ARIMA(1,1,1). Studi tersebut menunjukkan bahwa pemilihan parameter yang tepat sangat penting untuk meningkatkan akurasi peramalan. Model ARIMA juga menunjukkan keunggulan dalam menangani volatilitas harga komoditas dengan tingkat kesalahan yang rendah, meskipun model ini masih memiliki keterbatasan dalam menangani faktor eksternal seperti harga komoditas global, mata uang, dan tren pencarian oleh google [2].

Oleh karena itu, penelitian ini bertujuan untuk menerapkan gabungan model ARIMA, SARIMAX, *exponential smoothing* dalam peramalan harga bahan pangan pada setiap provinsi di Indonesia serta mengevaluasi kinerjanya dalam memberikan estimasi harga yang akurat.

II. Metode Analisis

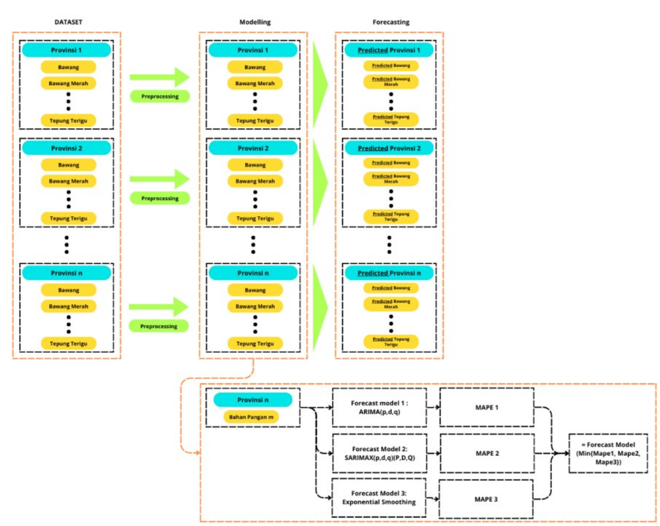

Fig. 1. Research Flow Method

Model ARIMA (*Autoregressive Integrated Moving Average*) dan SARIMAX (*Seasonal ARIMA with Exogenous Variables*) adalah dua model yang sangat populer dalam analisis dan tugas peramalan urutan waktu. Model ini digunakan untuk menangani data urutan waktu yang memiliki tren, dan musiman. ARIMA adalah model yang digunakan untuk data non-musiman, sedangkan SARIMAX adalah perluasan dari ARIMA yang mencakup komponen musiman dan variabel eksogen. Model lain yang lebih cocok pada kasus peramalan suatu data yang tidak memiliki kejelasan tren atau musiman adalah *exponential smoothing* atau disebut *simple exponential smoothing* (SES). Model *PROPHET* adalah model peramalan waktu yang dikembangkan oleh Facebook (S. J. Taylor & Letham, 2018). Awalnya, model ini dirancang untuk meramalkan data harian yang memiliki musiman mingguan dan tahunan, serta efek hari libur. Namun, model ini kemudian diperluas untuk mencakup berbagai jenis data musiman lainnya. Model *PROPHET* bekerja paling baik pada data time series yang memiliki musiman yang kuat dan beberapa musim data historis.

A. ARIMA

Model ARIMA(p,d,q) terdiri dari 3 komponen utama [3], yaitu *AutoRegressive* (AR), *Integrated* (I), *Moving Average* (MA) dengan,

- p adalah orde dari komponen model AR
- d adalah orde *differencing* yang diperlukan untuk membuat data stationer
- q adalah orde dari komponen model MA

*1) **Komponen AR(p)**:* Diasumsikan nilai saat ini dari urutan waktu dapat dijelaskan oleh nilai-nilai sebelumnya. Model AR(p) dapat ditulis sebagai,

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t \tag{1}$$

dengan,

- Y_t adalah nilai variabel pada waktu t,
- ϕ_i adalah koefisien model,
- serta ε_t adalah error term

*2) **Komponen I**:* Komponen I mengacu pada proses *differencing* yang dilakukan untuk membuat data stasioner. *Differencing* orde d dapat ditulis sebagai,

$$\Delta^d Y_t = (1 - L)^d Y_t \tag{2}$$

dengan L adalah operator lag, yaitu $LY_t = Y_{t-1}$

*3) **Komponen MA(q)**:* Diasumsikan nilai saat ini dari urutan waktu dapat dijelaskan oleh error term sebelumnya. Model MA(q) dapat ditulis sebagai,

$$Y_t = \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{3}$$

dengan,

- Y_t adalah nilai variabel pada waktu t,
- θ_j adalah koefisien model,
- serta ε_t adalah error term

*4) **Model ARIMA**:* Model ARIMA adalah gabungan dari ketiga komponen diatas, sehingga dapat ditulis,

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{4}$$

B. SARIMAX

Model SARIMAX adalah turunan ARIMA yang mencakup komponen musiman dan variabel eksogen. Model ini dinotasikan sebagai SARIMAX(p, d, q)(P, D, Q, s), dengan,

- (p, d, q), parameter non-musiman seperti pada ARIMA
- (P, D, Q), parameter musiman
- S, panjang siklus musiman

*1) **Komponen Musiman**:* Komponen musiman SARIMAX mirip dengan ARIMA, tetapi diterapkan pada data musiman. Model musiman dapat ditulis sebagai,

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{s=1}^{S} \Phi_s Y_{t-s} + X_t \beta + \varepsilon_t \tag{5}$$

dengan,

- Y_t adalah nilai observasi pada waktu t.
- c adalah konstanta.
- ϕ_i adalah parameter autoregressive (AR) untuk komponen non-musiman.
- θ_j adalah parameter moving average (MA) untuk komponen non-musiman.
- Φ_s adalah parameter autoregressive (AR) untuk komponen musiman.
- S adalah periode musiman (misalnya, 12 untuk data bulanan).
- X_t adalah variabel eksogen pada waktu t.
- β adalah koefisien untuk variabel eksogen.
- ϵ_t adalah error term pada waktu t.

C. Exponential Smoothing

Model *Exponential Smoothing* adalah model yang menggunakan perataan eksponensial untuk menangkap pola tren dan musiman yang ada dalam data.

*1) **Simple Exponential Smoothing (SES)**:* Simple Exponential Smoothing (SES) adalah bentuk paling dasar dari exponential smoothing. Model ini digunakan ketika data tidak memiliki tren atau musiman. Persamaan dasar SES adalah,

$$S_t = \alpha Y_t + (1 - \alpha) S_{t-1} \tag{6}$$

dengan,

- S_t adalah nilai yang dihaluskan (smoothed value) pada waktu t.
- Y_t adalah nilai aktual (observed value) pada waktu t.
- α adalah *smoothing factor* (faktor penghalusan), dengan $0 \leq \alpha \leq 1$.
- S_{t-1} adalah nilai yang dihaluskan pada waktu sebelumnya $(t - 1)$.

Nilai α menentukan seberapa besar bobot yang diberikan pada observasi terbaru:

- Jika α mendekati 1, model akan lebih responsif terhadap perubahan terbaru dalam data.
- Jika α mendekati 0, model akan lebih mengandalkan nilai historis dan kurang responsif terhadap perubahan terbaru.

*2) **Triple Exponential Smoothing (Holt-Winters Method)**:* Jika data memiliki tren dan musiman, metode *Triple Exponential Smoothing* (atau *Holt-Winters Method*) dapat digunakan. Metode Holt-Winters terdiri dari empat persamaan utama,

$$\begin{cases} S_t = \alpha \frac{Y_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + T_{t-1}) \\ T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \\ I_t = \gamma \frac{Y_t}{S_t} + (1-\gamma)I_{t-L} \end{cases}$$

dengan,

- I_t adalah komponen musiman pada waktu t.
- L adalah panjang musiman (misalnya, 12 untuk data bulanan).
- γ adalah smoothing factor untuk musiman.

$$\begin{cases} S_t = \alpha \frac{y_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}) & \text{Overall Smoothing} \\ b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} & \text{Trend Smoothing} \\ I_t = \beta \frac{y_t}{S_t} + (1-\beta)I_{t-L} & \text{Seasonal Smoothing} \\ F_{t+m} = (S_t + m b_t) I_{t-L+m} & \text{Forecast} \end{cases}$$

dengan,

- y_t adalah observasi pada waktu t.
- S_t adalah nilai yang dihaluskan (*smoothed observation*) pada waktu t.
- b_t adalah faktor tren (*trend factor*) pada waktu t.
- I_t adalah indeks musiman (*seasonal index*) pada waktu t.
- F_{t+m} adalah peramalan untuk m waktu ke depan.
- L adalah panjang musiman (misalnya, 12 untuk data bulanan).
- α, β, γ adalah konstanta yang harus diestimasi untuk meminimalkan MSE (*Mean Squared Error*) dari kesalahan peramalan.

Penjelasan Komponen

- **Overall Smoothing** (S_t): Menghaluskan observasi dengan mempertimbangkan tren dan musiman.
- **Trend Smoothing** (b_t): Menghaluskan komponen tren.
- **Seasonal Smoothing** (I_t): Menghaluskan komponen musiman.
- **Forecast** (F_{t+m}): Menghasilkan peramalan untuk periode ke depan dengan mempertimbangkan tren dan musiman.

D. PROPHET

Model *PROPHET* adalah sebuah prosedur untuk meramalkan data time series berdasarkan model aditif di mana tren non-linear di-fit dengan musiman tahunan, mingguan, dan harian, serta efek hari libur. Model *PROPHET* memecah time series menjadi tiga komponen utama: tren, musiman, dan hari libur [**?**]. Model ini dapat dinyatakan dalam persamaan berikut,

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \tag{7}$$

dengan,

- $y(t)$, nilai time series pada waktu t.
- $g(t)$, komponen tren yang memodelkan perubahan non-periodik dalam nilai time series.
- $s(t)$, komponen musiman yang memodelkan perubahan periodik (misalnya, musiman mingguan atau tahunan).
- $h(t)$, komponen hari libur yang memodelkan efek dari hari libur atau kejadian khusus.
- ε_t, error term

*1) **Komponen Tren**:* Model Prophet menyediakan dua jenis model tren: pertumbuhan logistik dan tren linear dengan perubahan titik. Model pertumbuhan logistik digunakan untuk data yang menunjukkan pertumbuhan yang mencapai titik jenuh (saturasi). Sehingga persamaan model pertumbuhan logistik *piecewise* adalah,

$$g(t) = \frac{C(t)}{1 - \exp(-(k + \mathbf{a}(t)^\intercal \delta)(t - (m + \mathbf{a}(t)^\intercal \gamma)))} \tag{8}$$

dengan,

- C(t), kapasitas pembawa (carrying capacity) yang dapat berubah seiring waktu.
- k, tingkat pertumbuhan dasar
- m, parameter offset.
- δ, vektor penyesuaian tingkat pertumbuhan pada titik perubahan.
- γ, adalah penyesuaian offset pada titik perubahan.
- vektor indikator $\mathbf{a}(t) \in \{0,1\}^s$, yang menunjukkan apakah waktu t telah melewati titik perubahan s_j

Untuk data yang tidak menunjukkan pertumbuhan jenuh, model tren linear dengan perubahan titik digunakan. Model ini ditulis,

$$g(t) = (k + \mathbf{a}(t)^\intercal \delta)t + (m + \mathbf{a}(t)^\intercal \gamma) \tag{9}$$

*2) **Komponen Musiman**:* Komponen musiman dalam Prophet dimodelkan menggunakan deret Fourier untuk menangani efek musiman yang kompleks.

$$s(t) = \sum_{n=1}^{N} (a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right)) \tag{10}$$

dengan,

- P adalah periode musiman (misalnya, 7 untuk musiman mingguan atau 365.25 untuk musiman tahunan).
- a_n dan b_n adalah parameter yang diestimasi.
- N adalah jumlah komponen Fourier yang digunakan.

*3) **Komponen Hari Libur**:* Komponen hari libur mempunyai efek seperti perubahan signifikan dalam pola data, seperti peningkatan/penurunan tiba-tiba, pada haril libur. Ini dimodelkan sebagai fungsi indikator yang menandai hari-hari tertentu dalam set data.

$$h(t) = Z(t)\kappa \tag{11}$$

Dengan,

- $h(t)$ adalah efek hari libur pada waktu t.
- $Z(t)$ adalah vektor indikator yang menandai apakah waktu t termasuk dalam hari libur tertentu.
- κ adalah parameter yang diestimasi, yang menunjukkan besarnya efek hari libur.

Vektor indikator $Z(t)$ didefinisikan sebagai,

$$Z(t) = [1(t \in D_1), \ldots, 1(t \in D_L)] \tag{12}$$

dengan,

- D_l adalah himpunan hari libur ke-l.
- $1(t \in D_l)$ adalah fungsi indikator yang bernilai 1 jika waktu t termasuk dalam hari libur D_l, dan 0 jika tidak.
- L adalah jumlah hari libur yang dimodelkan.

E. Missing Value Handling

Untuk mengisi nilai-nilai yang hilang (missing values) pada data komoditas di berbagai provinsi, pendekatan yang digunakan adalah metode imputasi berbasis urutan waktu. Analisis awal mengindikasikan bahwa data yang hilang mengindikasikan perubahan yang tidak signifikan, sehingga dapat diasumsikan nilai tersebut identik dengan nilai terdekat dalam rangkaian waktu. Oleh karena itu, proses imputasi dilakukan dengan menerapkan teknik *forward fill* dan *backward fill* secara sekuensial.

Metode *forward fill* diterapkan dengan mempropagasi nilai dari pengamatan sebelumnya untuk mengisi kekosongan data, sedangkan metode *backward fill* diimplementasikan dengan mengadopsi nilai dari pengamatan berikutnya. Kombinasi kedua pendekatan ini memastikan kelengkapan dataset tanpa menimbulkan distorsi artifisial pada pola perubahan nilai komoditas antar waktu.

III. Hasil dan Pembahasan

A. Prophet

Dalam penelitian ini, kami melaksanakan eksperimen pertama dengan menerapkan model Prophet, yang diasumsikan memiliki tren seragam dalam data. Model Prophet dipilih karena diharapkan dapat memberikan kinerja yang lebih baik dibandingkan dengan model deret waktu lainnya, seperti ARIMA dan SARIMAX.

Tujuan dari analisis ini adalah untuk mengevaluasi efektivitas model tersebut. Selain itu, kami juga mempertimbangkan adanya ketidakstabilan harga komoditas di setiap provinsi selama tahun 2022 dan 2023, yang kami anggap disebabkan oleh masa transisi akibat pandemi COVID-19.

Oleh karena itu, kami akan melakukan perbandingan antara data yang mencakup tahun 2022 dan yang tidak mencakup tahun tersebut.

TABLE I
Perbandingan Kinerja Model Prophet dengan Periode Data Pelatihan Berbeda

Model	Periode Data Pelatihan	MAPE (%)
Prophet	2022-01-01 – 2024-09-30	0.1356
Prophet	2024-01-01 – 2024-09-30	0.07806

B. Exponential Smoothing

Berdasarkan hasil analisis sebelumnya, kami menemukan bahwa asumsi mengenai ketidakstabilan pasar akibat masa transisi pasca pandemi dapat terlihat dari performa model yang dilatih menggunakan data pelatihan tertentu. Temuan ini mendorong kami untuk fokus melatih model hanya pada data dari tahun-tahun yang relevan.

Selain itu, analisis lebih lanjut menunjukkan adanya perbedaan pola pada setiap komoditas di berbagai provinsi, sehingga diperlukan pendekatan yang lebih spesifik untuk masing-masing komoditas.

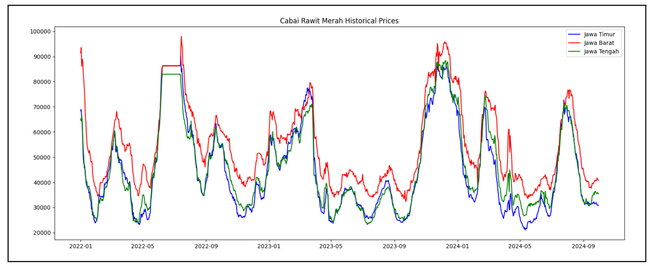

Fig. 2. Visualisasi data harga cabai rawit merah di 3 provinsi (Jawa Timur, Jawa Barat, Jawa Tengah) Januari 2022 – September 2024

Salah satu temuan yang signifikan adalah banyaknya komoditas yang menunjukkan pola tren dan musiman berdasarkan visualisasi data. Oleh karena itu, kami mengusulkan penggunaan model Exponential Smoothing dalam dua mode, yaitu untuk data dengan pola musiman dan tanpa pola musiman, dengan deteksi pola musiman dilakukan menggunakan nilai Autocorrelation Function (ACF).

TABLE II
Kinerja Model Exponential Smoothing

Model	Periode Data Pelatihan	MAPE (%)
Exponential Smoothing	2024-01-01 – 2024-09-30	0.05623

C. Exponential Smoothing + ARIMA (1 Model)

Hasil analisis sebelumnya menunjukkan bahwa ada tingkat signifikansi yang cukup tinggi untuk seluruh provinsi terkait setiap komoditas yang diteliti. Kami kemudian mempertimbangkan kemungkinan penerapan model yang lebih baik daripada yang telah digunakan, dengan menekankan kompleksitas ARIMA, yang diharapkan dapat meningkatkan akurasi prediksi.

Kami berhipotesis bahwa model ARIMA dengan parameter tertentu dapat disesuaikan untuk keseluruhan data di setiap provinsi dan komoditas. Oleh karena itu, kami memutuskan untuk mengganti pendekatan sebelumnya yang menggunakan Exponential Smoothing tanpa mempertimbangkan parameter tren dengan model ARIMA (2,0,15).

Selain itu, kami juga mengubah strategi pelatihan dengan menetapkan tanggal awal pelatihan menjadi 1 April 2024 hingga 30 September 2024. Keputusan ini didasarkan pada analisis kami untuk mengambil sampel yang lebih representatif dalam proses pelatihan.

TABLE III
KINERJA MODEL KOMBINASI EXPONENTIAL SMOOTHING DAN ARIMA

Model	Periode Data Pelatihan	MAPE (%)
Exponential Smoothing + ARIMA (2,0,15)	2024-04-01 – 2024-09-30	0.04878

D. Exponential Smoothing + ARIMA (34 X 13 = 442 Model)

Dalam eksperimen sebelumnya, model ARIMA dengan parameter yang sederhana berhasil meningkatkan akurasi. Ini menunjukkan bahwa model ARIMA sangat efektif, tetapi karena belum dilakukan differencing, model tersebut masih tergolong sebagai model ARMA.

Berdasarkan analisis kami, setiap komoditas di setiap provinsi memiliki pola yang berbeda; oleh karena itu, adalah hal yang wajar jika harga dan pola juga bervariasi. Perbedaan ini dipengaruhi oleh faktor lingkungan, kebiasaan, pendapatan, dan faktor-faktor lainnya.

Oleh sebab itu, kami mulai menganalisis setiap provinsi dan komoditas yang ada di dalamnya. Namun, proses ini memerlukan ketelitian dan pemahaman statistik yang mendalam. Untuk itu, kami hanya mengusulkan parameter **pdq** dari ARIMA dengan pendekatan yang sederhana.

TABLE IV
KINERJA MODEL DENGAN PARAMETER YANG BERVARIASI

Model	p	d	q	MAPE (%)
ARIMA + Exponential Smoothing	0,1,2	0,1	0,1,2,7,15	0.04114

E. ARIMA + SARIMA + Exponential Smoothing

Pada penelitian berikutnya, kami menerapkan model SARIMAX dan mengurangi ketergantungan pada metode exponential smoothing untuk meningkatkan akurasi prediksi (ditinjau melalui MAPE) serta memperkuat kemampuan adaptasi model terhadap komponen musiman harian/mingguan dan tren temporal.

Kami merancang tiga model hibrida dengan tingkat kompleksitas yang berbeda untuk menganalisis data deret waktu. Berikut adalah penjelasan setiap model beserta spesifikasinya:

1) Model 1 (Model Sederhana) + Exponential Smoothing: Model ini merupakan pendekatan awal dengan parameter yang lebih terbatas untuk SARIMA dan ARIMA. Model ini dirancang untuk menangkap pola dasar dalam data dengan kompleksitas yang lebih rendah, sehingga dapat menjadi acuan awal dalam analisis.

Pada model SARIMA, parameter p, d, dan q dibatasi hanya pada nilai 0 atau 1, begitu pula dengan parameter musiman P, D, dan Q. Sementara itu, parameter musiman S divariasikan dengan nilai 2, 3, 4, dan 5 untuk mengakomodasi kemungkinan pola musiman dengan berbagai panjang siklus yang lebih pendek.

Untuk model ARIMA, parameter p dan d dibatasi dalam rentang 0 hingga 2, sedangkan q memiliki nilai yang lebih luas, yaitu 0, 5, 7, 10, dan 15. Model ini memungkinkan eksplorasi awal terhadap pola autoregresif, diferensiasi, dan moving average tanpa terlalu banyak kombinasi parameter yang dapat meningkatkan risiko overfitting.

Sebagai metode pembanding, Exponential Smoothing juga digunakan untuk melihat apakah pola tren dan musiman dalam data dapat ditangkap secara efektif tanpa menggunakan model berbasis ARIMA atau SARIMA. Model ini cocok untuk situasi di mana kita ingin mendapatkan baseline sebelum mengeksplorasi model yang lebih kompleks seperti Model 2 dan Model 3.

TABLE V
PARAMETER MODEL 1

Model	p	d	q	P	D	Q	S
SARIMA	0,1	0,1	0,1	0,1	0,1	0,1	2, 3, 4, 5
ARIMA	0, 1, 2	0, 1, 2	0, 5, 7, 10, 15	-	-	-	-

2) Model 2 (Model Sedang) + Exponential Smoothing: Pada model ini, kami memperluas ruang pencarian parameter dibandingkan dengan Model 1. Model SARIMA memiliki nilai p, d, dan q yang lebih luas, yaitu dari 0 hingga 3, sementara parameter musiman P, D, dan Q bernilai 0 hingga 2. Parameter musiman S juga memiliki lebih banyak kemungkinan nilai, yaitu 5, 7, 10, 12, dan 15.

Untuk model ARIMA, kami memperbolehkan p bervariasi dari 0 hingga 5, d dari 0 hingga 2, dan q tetap pada rentang 0, 5, 7, 10, dan 15. Model ini bertujuan untuk menangkap pola yang lebih kompleks dibandingkan Model 1, dengan tetap menjaga jumlah parameter agar tidak terlalu banyak guna menghindari overfitting.

Exponential Smoothing tetap digunakan sebagai teknik tambahan untuk menangani tren dan musiman dalam data, terutama untuk membandingkan apakah model berbasis ARIMA/SARIMA lebih unggul dibandingkan metode smoothing berbasis eksponensial.

TABLE VI
PARAMETER MODEL 2

Model	p	d	q	P	D	Q	S
SARIMA	0,1,2,3	0,1,2,3	0,1,2,3	0,1,2	0,1,2	0,1	5,7,10,12,15
ARIMA	0-5	0-2	0,5,7,10,15	-	-	-	-

3) Model 3 (Model Kompleks) + Exponential Smoothing: Model 3 merupakan model dengan kompleksitas tertinggi dibandingkan Model 1 dan Model 2. Pada model ini, kami semakin memperluas rentang nilai parameter untuk SARIMA dan ARIMA agar dapat menangkap lebih banyak pola dalam data.

Untuk SARIMA, p, d, dan q dapat bervariasi dari 0 hingga 3, sedangkan parameter musiman P, D, dan Q memiliki nilai hingga 2. Selain itu, parameter musiman S memiliki rentang yang lebih besar, yaitu 5, 7, 10, 12, 15, dan 17. Dengan demikian, model ini dapat menangani pola musiman yang lebih kompleks dan fleksibel.

Untuk ARIMA, parameter p diperluas hingga 6, d hingga 4, dan q hingga 30, memungkinkan model menangkap pola autoregresif dan moving average yang lebih panjang dibandingkan dengan model sebelumnya.

Pendekatan Exponential Smoothing tetap digunakan sebagai metode tambahan untuk membandingkan efektivitasnya terhadap model berbasis ARIMA dan SARIMA. Model 3 sangat cocok untuk dataset dengan pola musiman dan tren yang kuat, tetapi perlu diwaspadai terhadap kemungkinan overfitting akibat kompleksitas yang tinggi.

TABLE VII
PARAMETER MODEL 3

Model	p	d	q	P	D	Q	S
SARIMA	0,1,2,3	0,1,2,3	0,1,2,3	0,1,2	0,1,2	0,1	5,7,10,12,15,17
ARIMA	0,1,2,3,4,5,6	0,1,2,3,4	0,5,7,10,15,20,30	-	-	-	-

TABLE VIII
PERBANDINGAN KINERJA MODEL BERDASARKAN NILAI MAPE

Model	MAPE (%)		
	Eksperimen 1	Eksperimen 2	Eksperimen 3
Model 1	0.03758	0.03699	0.03498
Model 2	0.03484	0.03359	0.03264
Model 3	0.03070	0.02989	**0.02888**

Hasil eksperimen menunjukkan perbedaan signifikan dalam performa ketiga model, dengan Model 3 (kompleks) mencapai akurasi tertinggi (MAPE 0.02888). Hal ini mengindikasikan bahwa:

1) Peningkatan kompleksitas parametrik pada SARIMAX dan ARIMAX secara konsisten meningkatkan kemampuan model dalam menangkap pola multi-seasonal dan nonlinearity data.
2) Kombinasi algoritma hibrida (SARIMAX + ARIMAX + ETS) memberikan sinergi positif, di mana SARIMAX mengoptimalkan komponen musiman, sementara ARIMAX berkontribusi pada pemodelan lag effects jangka panjang

Hasil peramalan menunjukkan bahwa model yang digunakan berhasil menangkap pola harga 13 komoditas di 34 provinsi dengan cukup baik. Secara umum, pola musiman yang khas terlihat pada sebagian besar komoditas, di mana harga mengalami fluktuasi yang berulang sesuai dengan siklus panen, permintaan musiman, serta faktor eksternal seperti cuaca dan kebijakan perdagangan. Tren ini menunjukkan bahwa model mampu mengenali komponen musiman dan tren jangka panjang yang mempengaruhi dinamika harga di berbagai wilayah.

Meskipun model menunjukkan performa yang baik dalam memprediksi pola harga, terdapat beberapa variasi antarprovinsi yang mengindikasikan adanya faktor lokal yang berperan dalam pergerakan harga. Beberapa provinsi menunjukkan fluktuasi harga yang lebih tajam dibandingkan yang lain, kemungkinan besar akibat perbedaan akses distribusi, infrastruktur pasar, atau faktor eksternal seperti kebijakan subsidi dan kebijakan impor. Hal ini menunjukkan bahwa meskipun model telah menangkap pola umum, masih ada faktor spesifik regional yang dapat mempengaruhi keakuratan prediksi.

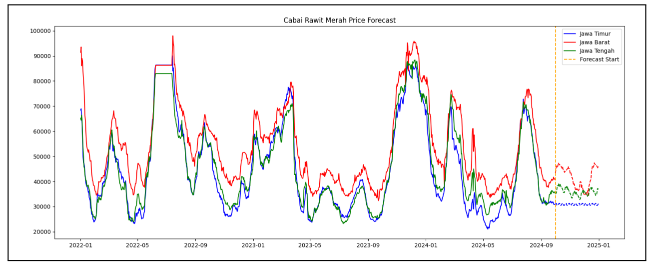

Fig. 3. Visualisasi hasil peramalan harga cabai rawit merah di 3 provinsi (Jawa Timur, Jawa Barat, Jawa Tengah)

IV. KESIMPULAN

Penelitian ini menganalisis efektivitas berbagai model peramalan untuk prediksi harga bahan pangan di 34 provinsi Indonesia. Berdasarkan hasil eksperimen yang dilakukan, dapat disimpulkan bahwa:

1) Model gabungan kompleks (ARIMA + SARIMA + Exponential Smoothing) dengan parameter yang lebih banyak menunjukkan performa terbaik dengan MAPE terendah sebesar 0.02888.
2) Data historis selama masa pandemi COVID-19 (2022-2023) menunjukkan ketidakstabilan yang memengaruhi akurasi model, sehingga penggunaan data yang lebih baru memberikan hasil prediksi yang lebih baik.
3) Setiap komoditas dan provinsi memiliki karakteristik pola data yang berbeda, sehingga pendekatan spesifik untuk masing-masing kasus diperlukan untuk mendapatkan hasil optimal.
4) Teknik penanganan missing value menggunakan forward fill dan backward fill terbukti efektif dalam mempertahankan integritas pola data tanpa distorsi yang signifikan.

Metode gabungan yang diusulkan berhasil menangkap kompleksitas dinamika harga pangan di berbagai provinsi dengan tingkat akurasi yang tinggi, sehingga dapat menjadi alat bantu yang efektif bagi pembuat kebijakan dalam mengantisipasi dan mengelola fluktuasi harga bahan pangan.

REFERENCES

[1] Allayioti, A., & Venditti, F. (2024). ”Working Paper Series The role of comovement and time-varying dynamics in forecasting commodity prices No 2901.” https://doi.org/10.2866/810813
[2] Dutta, S., & Maiti, S. (2021). ”PRICE FORECASTING OF AGRICULTURAL PRODUCTS USING ARIMA MODELS.” Agril. Mktg. (Spl. Issue), Vol. 35, Issue 2. https://www.researchgate.net/publication/363567569
[3] Nau, R. ”Slides on ARIMA models.” https://people.duke.edu/ rnau/Slides_on_ARIMA_models–Robert_Nau.pdf
[4] Adhikari, R., Agrawal, R. K. (n.d.). *An Introductory Study on Time Series Modeling and Forecasting*.
[5] Hyndman, R. J., Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts: Melbourne, Australia. Retrieved from https://otexts.com/fpp3. Accessed on March 15, 2025.
[6] Taylor, S. J., Letham, B. (2017). *Forecasting at Scale*. Retrieved from https://doi.org/10.7287/peerj.preprints.3190v2.
[7] Bahtiar, R., Raswatie, F. D. (2023). Analisis Fluktuasi Harga Pangan di Kota Bogor. Indonesian Journal of Agriculture Resource and Environmental Economics, 1(2), 70–81. https://doi.org/10.29244/ijaree.v1i2.42020