

***FORECASTING DATA PANEL LUAS PANEN DAN PRODUKSI PADI
MENGUNAKAN *HYBRID PANEL VECTOR AUTOREGRESSIVE*
(PVAR)-LONG SHORT TERM MEMORY (LSTM)***

Skripsi

Oleh

**NADIA GHASSANI
NPM. 2217031078**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

ABSTRACT

PANEL DATA FORECASTING OF HARVESTED AREA AND RICE PRODUCTION USING A HYBRID PANEL VECTOR AUTOREGRESSIVE (PVAR)–LONG SHORT-TERM MEMORY (LSTM) MODEL

By

Nadia Ghassani

Rice production is one of the most important indicators of national food security, therefore an accurate forecasting method is required to provide reliable estimates of production development across various regions in Indonesia. This study aims to analyze the relationship between harvested area and rice production as well as to perform prediction and forecasting for 34 provinces in Indonesia. The data used consist of monthly panel time series data for the period 2018–2023, obtained from the *Area Sampling Frame Survey (Kerangka Sampel Area, KSA)* conducted by Statistics Indonesia (BPS). The method applied in this research is the *Panel Vector Autoregression (PVAR)* model with the *Generalized Method of Moments (GMM)* estimation approach. The optimal model is determined based on the *Andrews–Lu Model and Moment Selection Criterion (MMSC)*. In addition, *Impulse Response Function (IRF)* and *Forecast Error Variance Decomposition (FEVD)* analyses are conducted to evaluate shock responses and the contribution of each variable in explaining variations in rice production. Furthermore, the PVAR prediction results are combined with a *Long Short-Term Memory (LSTM)* model as a *hybrid* approach to improve forecasting capability in capturing seasonal patterns and nonlinear fluctuations. The results show that the PVAR model is able to describe the dynamic relationship between harvested area and rice production; however, its predictions tend to be less responsive to sharp production spikes. The *hybrid* PVAR–LSTM approach improves forecasting accuracy compared to the standalone PVAR model, as indicated by lower MAE, RMSE, and MAPE values across most provinces. Therefore, this *hybrid* model can serve as an effective alternative to support rice production planning and food security policy in Indonesia.

Keywords: rice production, harvested area, panel time series, PVAR-GMM, LSTM, *hybrid forecasting model*.

ABSTRAK

FORECASTING DATA PANEL LUAS PANEN DAN PRODUKSI PADI MENGUNAKAN HYBRID PANEL VECTOR AUTOREGRESSIVE (PVAR)-LONG SHORT TERM MEMORY (LSTM)

Oleh

Nadia Ghassani

Produksi padi merupakan salah satu indikator penting dalam ketahanan pangan nasional, sehingga diperlukan metode peramalan yang mampu memberikan estimasi akurat terhadap perkembangan produksi di berbagai wilayah Indonesia. Penelitian ini bertujuan untuk menganalisis hubungan antara luas panen dan produksi padi serta melakukan prediksi dan peramalan pada 34 provinsi di Indonesia. Data yang digunakan merupakan data panel deret waktu bulanan periode 2018–2023 yang bersumber dari Survei Kerangka Sampel Area (KSA) Badan Pusat Statistik. Metode yang digunakan dalam penelitian ini adalah *Panel Vector Autoregression* (PVAR) dengan pendekatan estimasi *Generalized Method of Moments* (GMM). Model optimal ditentukan berdasarkan kriteria Andrews–Lu *Model and Moment Selection Criterion* (MMSC). Selain itu, dilakukan analisis *Impulse Response Function* (IRF) dan *Forecast Error Variance Decomposition* (FEVD) untuk mengevaluasi respons guncangan serta kontribusi variabel dalam menjelaskan variasi produksi padi. Selanjutnya, hasil prediksi PVAR dikombinasikan dengan model *Long Short-Term Memory* (LSTM) sebagai pendekatan *hybrid* guna meningkatkan kemampuan prediksi pada pola musiman dan fluktuasi nonlinier. Hasil penelitian menunjukkan bahwa model PVAR mampu menggambarkan hubungan dinamis antara luas panen dan produksi padi, namun prediksi cenderung kurang responsif terhadap lonjakan produksi yang tajam. Pendekatan *hybrid* PVAR–LSTM memberikan peningkatan akurasi peramalan dibandingkan PVAR murni, ditunjukkan oleh penurunan nilai MAE, RMSE, dan MAPE pada sebagian besar provinsi. Dengan demikian, model *hybrid* ini dapat menjadi alternatif yang efektif untuk mendukung perencanaan produksi padi dan kebijakan ketahanan pangan di Indonesia.

Kata kunci: produksi padi, luas panen, data panel, PVAR-GMM, LSTM, *hybrid forecasting*.

***FORECASTING DATA PANEL LUAS PANEN DAN PRODUKSI PADI
MENGUNAKAN *HYBRID PANEL VECTOR AUTOREGRESSIVE
(PVAR)-LONG SHORT TERM MEMORY (LSTM)****

NADIA GHASSANI

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

Judul Skripsi : **FORECASTING DATA PANEL LUAS
PANEN DAN PRODUKSI PADI
MENGUNAKAN HYBRID PANEL
VECTOR AUTOREGRESSIVE (PVAR)-LONG
SHORT TERM MEMORY (LSTM)**

Nama Mahasiswa : **Nadia Ghassani**

Nomor Pokok Mahasiswa : **2217031078**

Program Studi : **Matematika**

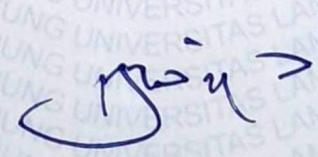
Fakultas : **Matematika dan Ilmu Pengetahuan Alam**




Dr. Dian Kurniasari, S.Si., M.Sc.
NIP 196903051996032001


Riza Sawitri, S.Pd., M.Sc.
NIP 198905042024062001

2. Wakil Dekan Bidang Akademik dan Kerjasama,
FMIPA Universitas Lampung


Mulyono, S.Si., M.Si., Ph.D.
NIP. 197406112000031002

MENGESAHKAN

1. Tim Penguji

Ketua : **Dr. Dian Kurniasari, S.Si., M.Sc.**

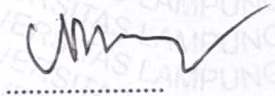


Sekretaris : **Riza Sawitri, S.Pd., M.Sc.**



Penguji

Bukan Pembimbing : **Ir. Warsono, M.S., Ph.D.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.

NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: **19 Mei 2026**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : **Nadia Ghassani**
Nomor Pokok Mahasiswa : **2217031078**
Jurusan : **Matematika**
Judul Skripsi : ***Forecasting Data Panel Luas Panen dan
Produksi Padi Menggunakan Hybrid Panel
Vector Autoregressive (PVAR)-Long Short Term
Memory (LSTM)***

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 19 Mei 2026

Penulis,


Nadia Ghassani

RIWAYAT HIDUP

Penulis memiliki nama lengkap Nadia Ghassani yang lahir di Pagaram Provinsi Sumatera Selatan pada tanggal 7 Oktober 2004. Penulis merupakan anak pertama dari empat bersaudara dari pasangan Bapak Hadi Ismanto dan Ibu Uswatun Hasanah.

Penulis mulai menempuh pendidikan di Taman Kanak (TK) Al-Azhar Kota Pagaram pada tahun 2009-2010, dan dilanjutkan di Sekolah Dasar di SD Muhammadiyah 3 Kota Pagaram dari tahun 2010-2016. Penulis melanjutkan pendidikan sekolah menengah pertama di SMP Negeri 2 Kota Pagaram pada tahun 2016-2019, kemudian melanjutkan ke jenjang menengah atas di SMA Negeri 1 Kota Pagaram pada tahun 2019-2022. Pada tahun 2022, penulis menjadi salah satu mahasiswa jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Selama menjalani studi di perguruan tinggi, penulis telah terlibat aktif dalam kegiatan organisasi kemahasiswaan dengan menjadi bagian dari HIMATIKA UNILA pada tahun 2023 sebagai anggota Bidang Keilmuan. Pada Desember 2024 hingga Februari 2025, penulis melaksanakan Kerja Praktik (KP) di Dinas Komunikasi Informatika dan Statistika (DISKOMINFOTIK) Provinsi Lampung. Selain itu, penulis turut berpartisipasi dalam kompetisi nasional yaitu, Pagelaran Mahasiswa Nasional Bidang Teknologi Informasi dan komunikasi (GEMASTIK) pada tahun 2024 dan kompetisi Statistika Ria dan Festival Sains Data (Satria Data) pada tahun 2025. Kemudian, pada Juni sampai Agustus 2025, penulis melaksanakan Kuliah Kerja Nyata (KKN) sebagai bentuk pengabdian penulis kepada masyarakat di Kelurahan Sumur Putri, Kecamatan Teluk Betung Selatan, Kota Bandar Lampung.

KATA INSPIRASI

”Allah SWT. tidak akan mengatakan hidup itu mudah. Tetapi Allah SWT. berjanji bahwa sesungguhnya bersama kesulitan ada kemudahan”.

(QS. Al-Insyirah: 5-6)

”Jangan pedulikan apa yang dikatakan oleh orang lain mengenaimu. Engkau tahu siapa dirimu dan Allah lebih tahu keadaan dirimu dan niat yang ada didalam hatimu.”

(QS. Al-Qiyamah: 14)

”Tidak ada mimpi yang terlalu tinggi. Tak ada mimpi yang patut untuk diremehkan. Lambungkan setinggi yang kau inginkan dan gapailah dengan selayaknya yang kau harapkan”

(Maudy Ayunda)

”Gagal hanya terjadi jika kita menyerah”

(Bacharuddin Jusuf Habibie)

PERSEMBAHAN

Dengan mengucapkan Alhamdulillah dan syukur kepada Allah SWT atas nikmat serta hidayah-Nya sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya. Dengan rasa syukur dan bahagia, saya persembahkan rasa terimakasih saya kepada:

Ayah dan Ibuku Tercinta

Terimakasih kepada orang tuaku atas segala pengorbanan, motivasi, doa dan ridho serta dukungannya selama ini. Terimakasih telah memberikan pelajaran berharga kepada anakmu ini tentang makna perjalanan hidup yang sebenarnya sehingga kelak bisa menjadi orang yang bermanfaat bagi banyak orang.

Dosen Pembimbing dan Pembahas

Terimakasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan arahan serta ilmu yang berharga.

Sahabat-sahabatku

Terimakasih kepada semua orang-orang baik yang telah memberikan pengalaman, semangat, motivasinya, serta doa-doanya dan senantiasa memberikan dukungan dalam hal apapun.

Almamater Tercinta

Universitas Lampung

SANWACANA

Alhamdulillah, puji dan syukur penulis panjatkan kepada Allah SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "*Forecasting Data Panel Luas Panen dan Produksi Padi Menggunakan Hybrid Panel Vector Autoregressive (PVAR)-Long Short Term Memory (LSTM)*" dengan baik dan lancar serta tepat pada waktu yang telah ditentukan. Shalawat serta salam semoga senantiasa tercurahkan kepada Nabi Muhammad SAW.

Dalam proses penyusunan skripsi ini, banyak pihak yang telah membantu memberikan bimbingan, dukungan, arahan, motivasi serta saran sehingga skripsi ini dapat terselesaikan. Oleh karena itu, dalam kesempatan ini penulis mengucapkan terimakasih kepada:

1. Ibu Dr. Dian Kurniasari, S.Si., M.Sc. selaku Pembimbing I sekaligus dosen pembimbing akademik yang telah banyak meluangkan waktunya untuk memberikan arahan, bimbingan, motivasi, saran serta dukungan kepada penulis sehingga skripsi ini dapat terselesaikan.
2. Ibu Riza Sawitri, S.Pd., M.Sc. selaku Pembimbing II yang telah memberikan arahan, bimbingan dan dukungan kepada penulis sehingga skripsi ini dapat terselesaikan.
3. Bapak Ir. Warsono, M.S., Ph.D. selaku Penguji yang telah bersedia memberikan kritik dan saran serta evaluasi kepada penulis sehingga dapat menjadi lebih baik lagi.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Seluruh dosen, staff dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

6. Bapak yang selalu kuat dan berjuang untuk anaknya, Ibu yang selalu mendengarkan dan memberi nasehat, adik-adik yang selalu memberikan sumber semangat, serta seluruh keluarga yang senantiasa memberikan doa, dukungan, dorongan, motivasi, serta semangat kepada penulis selama ini.
7. Teman penulis di bangku perkuliahan di antaranya Devinta, Pretty, Viki, Zakiah. Terimakasih sudah menemani dan memberi dukungan selama perkuliahan ini.
8. Rekan seperjuangan skripsi, di antaranya Meilina Risdianti, Khusni Sinta Rodiah, Fadillah Pinasti, Anita Caroline, Erin Elfitriani, Ahmad Rizki Munandar, Oja Widiyatama, Fatur Rozak, dan Benaya. Terimakasih sudah bersama-sama menyelesaikan ini.
9. Teman-teman seperjuangan Jurusan Matematika angkatan 2022.

Semoga skripsi ini dapat bermanfaat bagi kita semua. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, sehingga penulis mengharapkan kritik dan saran yang membangun untuk menjadikan skripsi ini lebih baik lagi.

Bandar Lampung, 19 Mei 2026

Nadia Ghassani

DAFTAR ISI

	Halaman
DAFTAR ISI	iii
DAFTAR TABEL	iv
DAFTAR GAMBAR	v
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
II TINJAUAN PUSTAKA	5
2.1 Penelitian Terkait	5
2.1.1 Penelitian Pertama	8
2.1.2 Penelitian Kedua	8
2.1.3 Penelitian Ketiga	9
2.1.4 Penelitian Keempat	9
2.1.5 Penelitian Kelima	10
2.1.6 Penelitian Keenam	11
2.2 Peramalan	11
2.3 Data Panel	12
2.4 Analisis Deret Waktu	12
2.5 Deret Waktu Multivariat	13
2.5.1 <i>Vector Autoregressive</i> (VAR)	13
2.5.2 <i>Panel Vector Autoregressive</i> (PVAR)	14
2.6 Kestasioneran Data	15
2.7 Estimasi Model dengan <i>Generalized Method of Moments</i> (GMM) .	17
2.8 Identifikasi dan Validasi Model PVAR	18

2.8.1	Pemilihan lag optimal	19
2.8.2	Uji Stabilitas Model	20
2.8.3	Uji Validitas Instrumen (Hansen test)	20
2.9	Analisis Dinamis PVAR	21
2.9.1	<i>Impulse Response Function</i> (IRF)	21
2.9.2	<i>Forecast Error Variance Decomposition</i> (FEVD)	22
2.10	<i>Machine Learning</i> dan <i>Deep Learning</i>	22
2.11	<i>Preprocessing</i> Data untuk LSTM	23
2.12	Fungsi Aktivasi	25
2.13	<i>Long Short-Term Memory</i> (LSTM)	27
2.14	<i>Hyperparameter Tuning</i>	29
2.15	Model <i>Hybrid</i> PVAR–LSTM	30
2.16	Evaluasi Model	31
2.17	Produksi dan Luas Panen Padi	32
III	METODE PENELITIAN	34
3.1	Waktu dan Tempat Penelitian	34
3.1.1	Tempat Penelitian	34
3.1.2	Waktu Penelitian	34
3.2	Data Penelitian	34
3.3	Alat Penelitian	36
3.4	Metode Penelitian	38
IV	HASIL DAN PEMBAHASAN	43
4.1	<i>Input</i> Data	43
4.2	Model <i>Panel Vector Autoregressive</i> (PVAR)	49
4.2.1	<i>Preprocessing</i> Data	49
4.2.2	Uji Stasioneritas	50
4.2.3	Pembagian Data	51
4.2.4	Penentuan Lag Optimal	52
4.2.5	Pembentukan Model PVAR	53
4.2.6	Uji Validitas Instrumen Model PVAR	54
4.2.7	Uji Stabilitas Model PVAR	55
4.2.8	<i>Impulse Response Function</i> (IRF)	56
4.2.9	<i>Forecast Error Variance Decomposition</i> (FEVD)	57
4.2.10	Hasil Prediksi Model PVAR	58

4.2.11	Menghitung Residual	64
4.3	Prediksi Model <i>Hybrid</i> PVAR-LSTM	70
4.3.1	<i>Scaling</i> Data	71
4.3.2	Membangun Model P_LSTM	71
4.3.3	Membangun Model E_LSTM	73
4.3.4	Prediksi model P_LSTM	75
4.3.5	Prediksi model E_LSTM	81
4.3.6	Prediksi model <i>Hybrid</i> PVAR-LSTM Skema 1 (EP)	87
4.3.7	Prediksi model <i>Hybrid</i> PVAR-LSTM Skema 2 (E)	93
4.4	Hasil Evaluasi	99
4.5	Peramalan	120
4.6	Perbandingan dengan Penelitian Sebelumnya	126
V	KESIMPULAN DAN SARAN	128
5.1	Kesimpulan	128
5.2	Saran	129
	DAFTAR PUSTAKA	130
	LAMPIRAN	135
	Lampiran 1. Hasil Peramalan Luas Panen Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Januari–Maret)	135
	Lampiran 2. Hasil Peramalan Produksi Padi Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Januari–Maret)	138

DAFTAR TABEL

Tabel	Halaman
1. Ringkasan Penelitian Terkait	5
2. Data Penelitian	36
3. <i>Missing Value</i>	49
4. Data Nol	50
5. Uji Levin–Lin–Chu (LLC)	50
6. Uji IPS	51
7. Pembagian Data Penelitian	52
8. Pemilihan Optimal Lag	52
9. Uji Validitas Instrumen (Hansen Test)	55
10. Uji Stabilitas Model PVAR	55
11. Forecast Error Variance Decomposition (FEVD)	57
12. <i>Hyperparameter Tuning</i> Model P_LSTM	72
13. <i>Hyperparameter Tuning</i> Model E_LSTM	74
14. Nilai MAE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Luas Panen) . . .	99
15. Nilai RMSE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Luas Panen) . .	101
16. Nilai MAPE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Luas Panen) . .	103
17. <i>Accuracy</i> Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Luas Panen)	105
18. Evaluasi MAPE Berdasarkan Kelompok Skala Luas Panen	107
19. Nilai MAE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Produksi Padi) . .	109
20. Nilai RMSE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Produksi Padi) .	111
21. Nilai MAPE Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Produksi Padi) .	113
22. <i>Accuracy</i> Model PVAR dan <i>Hybrid</i> PVAR-LSTM (Produksi Padi) . . .	115
23. Evaluasi MAPE Berdasarkan Kelompok Skala Produksi	118

DAFTAR GAMBAR

Gambar	Halaman
1 Fungsi Aktivasi Sigmoid (Akbar dkk., 2022).	26
2 Fungsi Aktivasi Tanh (Akbar dkk., 2022).	26
3 Arsitektur LSTM (Wang dkk., 2024).	27
4 Alur Penelitian Skema 1	41
5 Alur Penelitian Skema 2	42
6 Data	43
7 Visualisasi Aktual Luas Panen Kelompok Tinggi	44
8 Visualisasi Aktual Luas Panen Kelompok Sedang	44
9 Visualisasi Aktual Luas Panen Kelompok Rendah	45
10 Visualisasi Aktual Produksi Padi Kelompok Tinggi	46
11 Visualisasi Aktual Produksi Padi Kelompok Sedang	47
12 Visualisasi Aktual Produksi Padi Kelompok Rendah	48
13 Estimasi Parameter Model PVAR	53
14 <i>Impulse Response Function</i> (IRF)	56
15 Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Tinggi . .	58
16 Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Sedang . .	59
17 Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Rendah . .	60
18 Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Tinggi .	61
19 Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Sedang .	62
20 Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Rendah	63
21 Visualisasi Residual Kelompok Luas Panen Tinggi	64
22 Visualisasi Residual Kelompok Luas Panen Sedang	65
23 Visualisasi Residual Kelompok Luas Panen Rendah	66
24 Visualisasi Residual Kelompok Produksi Padi Tinggi	67
25 Visualisasi Residual Kelompok Produksi Padi Sedang	68
26 Visualisasi Residual Kelompok Produksi Padi Rendah	69

27	<i>Scaling Data</i>	71
28	Grafik <i>Loss Function</i> Model P_LSTM	73
29	Grafik <i>Loss Function</i> Model E_LSTM	74
30	Visualisasi Prediksi model P_LSTM Kelompok Luas Panen Tinggi . . .	75
31	Visualisasi Prediksi model P_LSTM Kelompok Luas Panen Sedang . .	76
32	Visualisasi Prediksi model P_LSTM Kelompok Luas Panen Rendah . .	77
33	Visualisasi Prediksi model P_LSTM Kelompok Produksi Padi Tinggi .	78
34	Visualisasi Prediksi model P_LSTM Kelompok Produksi Padi Sedang .	79
35	Visualisasi Prediksi model P_LSTM Kelompok Produksi Padi Rendah .	80
36	Visualisasi Prediksi model E_LSTM Kelompok Luas Panen Tinggi . .	81
37	Visualisasi Prediksi model E_LSTM Kelompok Luas Panen Sedang . .	82
38	Visualisasi Prediksi model E_LSTM Kelompok Luas Panen Rendah . .	83
39	Visualisasi Prediksi model E_LSTM Kelompok Produksi Padi Sedang .	84
40	Visualisasi Prediksi model E_LSTM Kelompok Produksi Padi Sedang .	85
41	Visualisasi Prediksi model E_LSTM Produksi Padi Rendah	86
42	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Luas Panen Tinggi	87
43	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Luas Panen Sedang	88
44	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Luas Panen Rendah	89
45	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Produksi Tinggi	90
46	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Produksi Sedang	91
47	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 1 Kelompok Produksi Rendah	92
48	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Luas Panen Tinggi	93
49	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Luas Panen Sedang	94
50	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Luas Panen Rendah	95
51	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Produksi Padi Tinggi	96
52	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Produksi Padi Sedang	97

53	Visualisasi Prediksi <i>Hybrid</i> PVAR–LSTM Skema 2 Kelompok Produksi Padi Sedang	98
54	Visualisasi Peramalan Kelompok Luas Panen Tinggi	121
55	Visualisasi Peramalan Kelompok Luas Panen Sedang	121
56	Visualisasi Peramalan Kelompok Luas Panen Rendah	122
57	Visualisasi Peramalan Kelompok Produksi Padi Tinggi	123
58	Visualisasi Peramalan Kelompok Produksi Padi Sedang	124
59	Visualisasi Peramalan Kelompok Produksi Padi Rendah	125

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Subsektor tanaman pangan merupakan bagian penting dalam pembangunan sektor pertanian di Indonesia, dengan padi sebagai komoditas utamanya. Pada tahun 2023, jumlah penduduk Indonesia diperkirakan mencapai 277,43 juta jiwa dengan laju pertumbuhan sebesar 1%. Dengan jumlah tersebut, tingkat konsumsi beras diperkirakan mencapai 111,32 kg/kapita/tahun, sehingga menjadikan Indonesia sebagai salah satu konsumen beras terbesar di dunia (Kementerian Pertanian, 2024). Kondisi ini menunjukkan bahwa ketersediaan produksi padi perlu dijaga secara berkelanjutan untuk mendukung ketahanan pangan nasional. Oleh karena itu, analisis peramalan luas panen dan produksi padi menjadi hal penting dalam perencanaan dan pengambilan kebijakan pangan.

Data luas panen dan produksi padi di Indonesia memiliki karakteristik yang kompleks, seperti pola musiman (*seasonality*), tren jangka panjang, serta fluktuasi yang dipengaruhi oleh faktor cuaca, kebijakan, dan kondisi sosial ekonomi. Selain itu, terdapat hubungan timbal balik antara luas panen dan produksi, di mana peningkatan luas panen mendorong kenaikan produksi, sementara hasil produksi sebelumnya juga dapat memengaruhi keputusan luas panen pada periode berikutnya (Harahap dkk., 2024). Dalam konteks data panel antar wilayah, heterogenitas yang tinggi juga muncul akibat perbedaan kondisi iklim, ketersediaan air, dan pola musim tanam, sehingga menyebabkan variasi produksi padi di tiap daerah (Buswar dkk., 2025). Metode peramalan tradisional seperti ARIMA dan *Exponential Smoothing* telah banyak digunakan dalam sektor pertanian, tetapi keduanya memiliki keterbatasan karena umumnya bekerja optimal pada data univariat dan pola linier (Mujiyanto dkk., 2024). Pendekatan ekonometrik multivariat seperti *Vector Autoregressive* (VAR) dapat menangkap hubungan dinamis antar variabel,

tetapi masih terbatas pada data runtun waktu tunggal dan belum memperhitungkan variasi antar wilayah. Untuk mengatasi hal tersebut, dikembangkan model *Panel Vector Autoregressive* (PVAR) yang mampu menganalisis hubungan dinamis dalam data panel dengan memperlakukan seluruh variabel sebagai endogen (Yang dkk., 2023).

Estimasi parameter dalam model PVAR umumnya dilakukan menggunakan pendekatan *Generalized Method of Moments* (GMM) yang diperkenalkan oleh Hansen (1982). Metode ini banyak digunakan pada data panel dinamis karena mampu mengatasi permasalahan endogenitas, autokorelasi, serta heteroskedastisitas yang sering muncul dalam model ekonometrik. Pendekatan serupa digunakan oleh Usman dkk. (2022) dalam membangun model GMM–PVAR untuk menganalisis keterkaitan energi terbarukan dan pembangunan keuangan di negara-negara Uni Eropa.

Meskipun PVAR mampu menggambarkan hubungan struktural antar variabel serta heterogenitas antar wilayah, model ini masih berasumsi pada hubungan linier. Padahal, dinamika produksi padi di Indonesia juga dipengaruhi oleh faktor non-linear seperti perubahan iklim, perkembangan teknologi pertanian, serta perilaku petani yang sulit dimodelkan secara sederhana. Di sisi lain, pendekatan *deep learning*, khususnya *Long Short-Term Memory* (LSTM), terbukti efektif dalam menangkap pola non-linear dan dependensi jangka panjang pada data deret waktu. Namun, LSTM cenderung berfokus pada akurasi prediksi tanpa memberikan interpretasi hubungan sebab-akibat antar variabel (Kapetanios dkk., 2021). Oleh karena itu, integrasi kedua pendekatan menjadi alternatif yang potensial, di mana PVAR digunakan untuk menangkap hubungan linier dan struktural, sedangkan LSTM digunakan untuk memodelkan pola non-linear pada residual PVAR.

Model *hybrid PVAR–LSTM* dikembangkan sebagai pendekatan yang menggabungkan kekuatan analisis ekonometrika dan *deep learning*. Setelah hubungan antar variabel seperti luas panen, produktivitas, dan produksi dianalisis menggunakan PVAR, residual yang tersisa kemudian dipelajari dengan LSTM untuk menangkap fluktuasi non-linear yang tidak dijelaskan oleh model linier. Pendekatan ini diharapkan dapat meningkatkan akurasi peramalan sekaligus mendukung analisis kebijakan berbasis data, seperti deteksi dini potensi defisit pangan dan perencanaan distribusi beras antar wilayah.

Penelitian mengenai penerapan PVAR maupun LSTM dalam analisis data panel dan deret waktu telah berkembang pesat. Sun (2021) menunjukkan bahwa PVAR efektif dalam memodelkan hubungan dinamis antar variabel pembangunan ekonomi dan pendidikan di berbagai wilayah Tiongkok. Di sisi lain, Solgi dkk. (2021) membuktikan bahwa LSTM mampu mengenali pola musiman serta fluktuasi non-linear dengan tingkat akurasi tinggi dalam peramalan data lingkungan. Pada sektor pertanian, Firdausi dkk. (2025) menemukan bahwa multivariat LSTM memberikan performa lebih baik dibandingkan metode tradisional dalam memprediksi produksi padi bulanan di Jawa Timur. Selain itu, Savada dkk. (2025) menunjukkan bahwa model *hybrid* VAR–LSTM mampu meningkatkan akurasi prediksi pada data ekonomi dibandingkan model tunggal.

Berdasarkan berbagai penelitian tersebut, dapat disimpulkan bahwa PVAR unggul dalam memodelkan keterkaitan dinamis antar variabel pada data panel dan memodelkan pola linear, sedangkan LSTM efektif dalam menangkap pola non-linear dan musiman pada data runtun waktu. Oleh karena itu, penerapan model *hybrid* PVAR–LSTM dinilai potensial untuk diterapkan pada sistem pertanian Indonesia yang kompleks dan heterogen, khususnya dalam memprediksi luas panen dan produksi padi secara lebih akurat.

1.2 Rumusan Masalah

Adapun rumusan masalah dari penelitian ini yaitu:

1. Bagaimana membangun model panel *Vector Autoregressive* untuk menangkap hubungan dinamis antar provinsi dalam luas panen dan produksi padi?
2. Bagaimana mengkaji model *hybrid* PVAR-LSTM untuk meningkatkan akurasi peramalan luas panen dan produksi padi?
3. Bagaimana performa model *hybrid* PVAR-LSTM dalam prediksi luas panen dan produksi padi di berbagai provinsi di Indonesia?
4. Bagaimana hasil peramalan luas panen dan produksi padi dengan menggunakan *hybrid* PVAR-LSTM?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu:

1. Membangun dan mengestimasi model *Panel Vector Autoregressive* (PVAR) dengan metode GMM untuk menangkap hubungan dinamis luas panen dan produksi padi.
2. Mengkaji penerapan model *hybrid* PVAR–LSTM dalam meningkatkan kemampuan prediksi luas panen dan produksi padi.
3. Mengetahui performa model *hybrid* PVAR-LSTM pada prediksi luas panen dan produksi padi di berbagai provinsi di Indonesia.
4. Melakukan peramalan luas panen dan produksi padi dengan menggunakan *hybrid* PVAR-LSTM.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini yaitu:

1. Memberikan wawasan mengenai pemodelan dan peramalan luas panen dan produksi padi.
2. Memberikan pengetahuan mengenai peramalan produksi padi dan luas panen.
3. Dapat menjadi referensi bagi penelitian selanjutnya mengenai metode *hybrid* PVAR-LSTM.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Penelitian terkait dalam penelitian ini digunakan sebagai bahan acuan untuk hasil analisis. Topik yang menjadi acuan analisis adalah model PVAR, model LSTM, dan model hybrid VAR-LSTM. Beberapa penelitian terkait yang digunakan sebagai acuan dalam penelitian ini dirangkum dalam Tabel 1 sebagai berikut:

Tabel 1. Ringkasan Penelitian Terkait

No.	Penelitian (Tahun)	Data	Metode/Model Penelitian	Hasil
1.	<i>Effects of Domestic Material Consumption, Renewable Energy, and Financial Development on Environmental Sustainability in the EU-28: Evidence from a GMM PVAR</i> (Usman dkk., 2022)	Data panel EU-28 tahun 1995–2019 mencakup konsumsi material domestik, energi terbarukan, dan pengembangan keuangan.	GMM PVAR (<i>Panel Vector Autoregressive Generalized Method of Moments</i>).	Menunjukkan hubungan dinamis dua arah antara energi terbarukan, pembangunan keuangan, dan keberlanjutan lingkungan. Model GMM–PVAR mampu menangkap keterkaitan jangka pendek dan panjang.

No.	Penelitian	Data	Metode	Hasil
2.	<i>An Empirical Study on the Relationship between Education and Economic Development Based on PVAR Model</i> (Sun dkk., 2021)	Data panel provinsi di Tiongkok dengan variabel tingkat pendidikan (jumlah mahasiswa per 100,000 penduduk) dan tingkat pembangunan ekonomi (PDB per kapita).	Model <i>Panel Vector Autoregression</i> (PVAR) dengan estimasi <i>Generalized Method of Moments</i> (GMM).	Menunjukkan adanya hubungan timbal balik antara pendidikan dan pembangunan ekonomi. Pertumbuhan ekonomi mendorong peningkatan pendidikan, sedangkan pengaruh pendidikan terhadap ekonomi muncul setelah periode jeda (<i>incubation period</i>) dan berbeda antar wilayah.
3.	<i>Long Short-Term Memory Neural Network (LSTM-NN) for Aquifer Level Time Series Forecasting Using In-Situ Piezometric Observations</i> (Solgi dkk., 2021)	Data level air tanah dari pengamatan in-situ piezometer selama beberapa tahun.	LSTM <i>Neural Network</i> dengan Adam optimizer.	LSTM–Adam menghasilkan kinerja terbaik dengan RMSE = 0.12, MAE = 0.09, dan $R^2 = 0.94$. Model mampu menangkap tren musiman dan fluktuasi jangka panjang dengan akurat.
4.	<i>A Multivariate LSTM Approach for Monthly Rice Production Forecasting in East Java</i> (Firdausi dkk., 2025)	Data produksi padi bulanan (ton) dan luas panen (hektar) Provinsi Jawa Timur periode 2018–2024 dari BPS.	<i>Long Short-Term Memory</i> (LSTM)	Model LSTM memberikan performa terbaik dibanding <i>Moving Average</i> dan <i>Linear Regression</i> dengan MAE = 95.030,16, RMSE = 120.229,01, dan MAPE = 16,64%. Model mampu menangkap pola musiman produksi padi serta hubungan antara produksi dan luas panen secara efektif.

No.	Penelitian	Data	Metode	Hasil
5.	<i>Forecasting Dengue across Brazil with LSTM Neural Networks and SHAP-Driven Lagged Climate and Spatial Effects</i> (Chen dan Moraga, 2025)	Data mingguan kasus dengue pada 27 federal states di Brasil periode 2016–2023 yang dikombinasikan dengan variabel iklim.	<i>Long Short-Term Memory (LSTM)</i>	Model LSTM dengan integrasi variabel iklim dan spasial menunjukkan performa terbaik dibandingkan model Bayesian baseline berdasarkan MAE, MAPE, dan CRPS. Pendekatan ini efektif dalam menangkap dinamika temporal, pengaruh faktor lingkungan, serta dependensi spasial antar wilayah dalam peramalan kasus <i>dengue</i> .
6.	<i>Economic Data Forecasting Using Hybrid Vector Autoregressive Long Short Term Memory Model</i> (Savada dkk., 2025)	Data harian Indeks Harga Saham Gabungan (IHSG) dan Kurs Rupiah terhadap Dolar AS periode 2004–2024, diperoleh dari Yahoo Finance dan Bank Indonesia.	<i>Model Hybrid Vector Autoregressive (VAR) dan Long Short-Term Memory (LSTM)</i>	Model VAR–LSTM menunjukkan peningkatan akurasi signifikan dibanding VAR tunggal, dengan penurunan MAE sebesar 42,72 untuk IHSG dan 55,82 untuk Kurs Rupiah. Kombinasi VAR–LSTM efektif menangkap pola linier dan nonlinier pada data deret waktu multivariat. Hal ini mengidentifikasikan model <i>hybrid</i> dapat menghasilkan hasil yang lebih baik dari model tunggal.

2.1.1 Penelitian Pertama

Penelitian yang dilakukan oleh Usman, Alola, dan Akadiri (2022) bertujuan untuk menganalisis hubungan dinamis antara konsumsi material domestik, penggunaan energi terbarukan, serta perkembangan sektor keuangan terhadap keberlanjutan lingkungan di negara-negara Uni Eropa. Data yang digunakan berupa data panel tahunan dari 28 negara Uni Eropa selama periode 1995–2019. Metode analisis yang diterapkan adalah GMM–PVAR, yang dipilih untuk mengatasi masalah endogenitas sekaligus menguji keterkaitan jangka pendek dan jangka panjang antar variabel penelitian. Hasil analisis menunjukkan adanya hubungan kausal dua arah antara energi terbarukan, pembangunan keuangan, dan indikator keberlanjutan lingkungan. Selain itu, peningkatan penggunaan energi terbarukan terbukti memberikan dampak positif dalam menurunkan emisi karbon, sedangkan konsumsi material domestik berdampak sebaliknya. Secara keseluruhan, penelitian ini menunjukkan bahwa model GMM–PVAR efektif untuk mengetahui dinamika interaksi variabel ekonomi dan lingkungan.

2.1.2 Penelitian Kedua

Penelitian yang kedua dilakukan oleh Sun (2021) bertujuan untuk menganalisis hubungan dinamis antara tingkat pendidikan dan pembangunan ekonomi di berbagai wilayah di Tiongkok menggunakan pendekatan *Panel Vector Autoregression (PVAR)* dengan metode estimasi *Generalized Method of Moments (GMM)*. Pada penelitian ini memanfaatkan data panel provinsi yang mencakup indikator tingkat pendidikan dan produk domestik bruto (PDB) per kapita untuk menggambarkan tingkat pembangunan ekonomi. Tujuannya adalah untuk menguji hubungan dua arah (*bidirectional*) antara pendidikan dan pembangunan ekonomi. Analisis dilakukan dengan menggunakan uji *Granger causality*, *Impulse Response Function (IRF)*, dan *Variance Decomposition* untuk menangkap dinamika hubungan antar variabel dalam jangka pendek maupun jangka panjang.

Hasil penelitian menunjukkan bahwa pembangunan ekonomi memiliki pengaruh positif dan langsung terhadap peningkatan pendidikan, sedangkan pengaruh pendidikan terhadap pertumbuhan ekonomi bersifat tertunda (*lagged effect*) dan bervariasi antar wilayah. Di wilayah timur yang lebih maju, efek pendidikan

terhadap ekonomi muncul lebih cepat, sedangkan di wilayah tengah dan barat efeknya baru terlihat setelah beberapa periode.

2.1.3 Penelitian Ketiga

Penelitian selanjutnya yang dilakukan oleh Solgi dkk. (2021) mengembangkan model jaringan saraf *Long Short-Term Memory* (LSTM-NN) untuk memprediksi level air tanah dengan menggunakan data historis level air tanah sebagai satu-satunya sumber input, tanpa melibatkan variabel hidrologis atau klimatologis tambahan sebagaimana lazim digunakan dalam penelitian sebelumnya. Tujuan utama penelitian tersebut adalah mengevaluasi kemampuan LSTM dalam melakukan peramalan pada horizon jangka pendek hingga jangka panjang serta membandingkan performanya dengan model *simple neural network* (simple NN). Data yang digunakan terdiri dari data harian, rata-rata bulanan, dan minimum bulanan pada akuifer Edwards di Texas, yang kemudian diproses pada skenario prediksi 1–30 hari dan 1–3 bulan ke depan. Model simple NN dibangun dengan satu *hidden layer* berisi sepuluh neuron sigmoid, sedangkan model LSTM terdiri dari sepuluh sel memori pada satu *hidden layer*. Pelatihan kedua model dilakukan menggunakan algoritma *Adam Optimizer* dengan fungsi objektif *Mean Squared Error* (MSE). Hasil evaluasi menunjukkan bahwa model LSTM memiliki akurasi yang jauh lebih tinggi dan stabil dibandingkan simple NN, dengan nilai koefisien determinasi R^2 mencapai 0.9989 untuk prediksi satu langkah ke depan dan tetap lebih besar dari 0.90 hingga 26 langkah prediksi.

2.1.4 Penelitian Keempat

Penelitian yang dilakukan oleh Firdausi dkk (2025), Penelitian ini bertujuan membangun model prediksi produksi padi bulanan di provinsi Jawa Timur menggunakan pendekatan multivariat *Long Short-Term Memory* (LSTM). Data yang digunakan berupa produksi padi (ton) dan luas panen (hektar) periode 2018–2024 yang bersumber dari BPS. Tahapan penelitian meliputi *preprocessing* data, normalisasi Min–Max, pembentukan *sequence* dengan metode *sliding window*, pembangunan arsitektur LSTM, pelatihan model, kemudian evaluasi yang digunakan adalah MAE, RMSE, dan MAPE. Hasil penelitian menunjukkan bahwa model LSTM mampu menunjukkan kinerja prediksi yang lebih baik dibandingkan metode

tradisional seperti *Moving Average* dan *Linear Regression*. Model menghasilkan nilai MAE sebesar 95.030,16, RMSE sebesar 120.229,01, dan MAPE sebesar 16,64% yang termasuk kategori akurasi baik. LSTM juga mampu menangkap pola musiman produksi padi serta hubungan antara produksi dan luas panen sebagai variabel multivariat. Secara umum, penelitian ini menegaskan bahwa pendekatan *deep learning* khususnya LSTM efektif digunakan untuk peramalan data pertanian yang bersifat nonlinier dan memiliki pola musiman. Namun, model masih memiliki keterbatasan dalam memprediksi nilai ekstrem, sehingga diperlukan pengembangan model lanjutan seperti integrasi variabel tambahan atau penggunaan model *hybrid* untuk meningkatkan akurasi prediksi di masa depan.

2.1.5 Penelitian Kelima

Penelitian yang dilakukan oleh Xiang Chen dan Paula Moraga (2025) mengkaji penerapan metode *Long Short-Term Memory* (LSTM) dalam peramalan kasus *dengue* pada 27 federal states di Brasil. Penelitian ini mengintegrasikan data kasus *dengue* mingguan dengan variabel iklim seperti temperatur, curah hujan, kelembaban relatif, tekanan udara, *thermal range*, serta jumlah hari hujan. Selain itu, penelitian tersebut juga memasukkan efek spasial dari wilayah tetangga untuk menangkap dependensi antar wilayah. Evaluasi model dilakukan dengan pendekatan *moving window* untuk prediksi 1 dan 3 bulan ke depan.

Hasil penelitian Chen dan Moraga (2025) ini menunjukkan bahwa model LSTM multivariat yang mengintegrasikan variabel iklim dan spasial memberikan performa terbaik dibandingkan model Bayesian baseline berdasarkan metrik *Mean Absolute Error (MAE)*, *Mean Absolute Percentage Error (MAPE)*, dan *Continuous Ranked Probability Score (CRPS)*. Penelitian ini menegaskan bahwa integrasi faktor eksternal serta struktur spasial dalam model deret waktu multivariat mampu meningkatkan akurasi peramalan dan memperkuat kemampuan model dalam menangkap dinamika temporal yang kompleks. Oleh karena itu, penelitian tersebut relevan sebagai landasan teoritis dalam pengembangan model peramalan multivariat pada penelitian ini.

2.1.6 Penelitian Keenam

Penelitian yang terakhir dilakukan oleh Gilang Aleyusta Savada, Fitri Bimantoro, dan Diah Risqiwati (2025) menunjukkan efektivitas integrasi antara metode statistik klasik dan *deep learning*. Penelitian ini menggabungkan *Vector Autoregressive (VAR)* yang efektif dalam menangkap hubungan linier antar variabel dengan *Long Short-Term Memory (LSTM)* yang mampu mengenali pola nonlinier serta hubungan jangka panjang. Data yang digunakan berupa indeks harga saham gabungan (IHSG) dan Kurs Rupiah terhadap Dolar AS. Tujuan utama penelitian tersebut adalah meningkatkan akurasi peramalan deret waktu multivariat.

Hasil penelitian Savada dkk. (2025) ini menunjukkan bahwa model *Hybrid VAR–LSTM* memberikan akurasi yang lebih tinggi dibandingkan model tunggal VAR atau LSTM, dengan penurunan *Mean Absolute Error (MAE)* sebesar 42,72 untuk IHSG dan 55,82 untuk Kurs Rupiah. Oleh karena itu, penelitian tersebut menjadi dasar kuat bagi penggunaan model *Hybrid PVAR–LSTM* dalam penelitian ini, karena *Panel Vector Autoregressive (PVAR)* merupakan bentuk generalisasi dari VAR yang dapat menangani data panel antar wilayah atau komoditas, sedangkan integrasi dengan LSTM membuat model semakin adaptif terhadap pola nonlinier.

2.2 Peramalan

Peramalan (*forecasting*) adalah suatu proses untuk mengetahui apa yang akan terjadi di masa mendatang dengan menggunakan data historis. Data yang digunakan adalah data yang bersifat kuantitatif. Tujuan utama peramalan bukanlah memberikan hasil sepenuhnya pasti, melainkan menghasilkan estimasi yang paling mendekati kondisi sebenarnya yang akan terjadi (Amalia dkk., 2022).

Menurut Montgomery dkk. (2015), peramalan merupakan masalah penting pada berbagai bidang seperti bisnis, pemerintahan, ekonomi, ilmu lingkungan, kesehatan, sosial, politik, dan keuangan. Masalah peramalan sering diklasifikasikan menjadi jangka pendek, jangka menengah, dan jangka panjang. Peramalan jangka pendek dan menengah biasanya digunakan untuk operasional dan penganggaran, sedangkan jangka panjang untuk perencanaan strategis. Pada peramalan seperti jangka pendek

dan menengah biasanya menggunakan pola dari data historis, karena data historis ini biasanya menunjukkan inersia dan tidak berubah secara drastis dengan cepat. Metode statistik akan sangat berguna untuk peramalan jangka pendek dan menengah.

2.3 Data Panel

Data panel adalah data yang terdiri dari observasi terhadap sejumlah unit individu selama suatu periode waktu, yang menggabungkan karakteristik dari data runtun waktu (*time-series*) dan data potong lintang (*cross-section*). Keunggulan utama data panel adalah kemampuannya mengendalikan heterogenitas tak teramati yang bersifat konstan sepanjang waktu tetapi berbeda antar entitas (Arellano, 2003).

Selain itu, menggunakan data panel dapat memungkinkan analisis yang lebih komprehensif terhadap dinamika perubahan antar waktu dan perbedaan antar individu. Dengan memanfaatkan struktur data yang bersifat longitudinal, metode ini tidak hanya mengurangi bias akibat heterogenitas individu yang tidak teramati, tetapi juga meningkatkan efisiensi estimasi melalui jumlah observasi yang lebih besar dan menurunnya kolinieritas antar variabel. Keunggulan ini menjadikan data panel sangat berguna dalam pemodelan perilaku ekonomi yang bersifat dinamis dan kompleks dibandingkan dengan analisis data *time-series* atau *cross-section* murni (Hsiao, 2014).

2.4 Analisis Deret Waktu

Deret waktu adalah urutan pengamatan yang dicatat pada suatu variabel dalam rentang waktu tertentu. Data seperti ini biasanya dikumpulkan pada interval yang teratur baik harian, bulanan, maupun tahunan sehingga memungkinkan analisis terhadap dinamika perubahan secara sistematis. Montgomery menjelaskan bahwa deret waktu dapat berupa nilai sesaat, kumulatif, ataupun statistik yang mencerminkan aktivitas variabel selama periode tertentu. Dalam konteks peramalan, analisis deret waktu berperan penting karena memungkinkan identifikasi dan pemodelan pola historis untuk memperkirakan kondisi di masa mendatang (Montgomery dkk., 2015).

Dari sisi analitis, deret waktu tidak hanya dipahami sebagai urutan data, tetapi juga sebagai struktur yang mengandung berbagai pola dan hubungan internal. Beberapa komponen utama yang umum dikaji meliputi tren jangka panjang, fluktuasi musiman, autokorelasi antar periode, serta kestasioneran yang menunjukkan kestabilan sifat statistik suatu deret sepanjang waktu. Pemahaman terhadap komponen-komponen ini memungkinkan pembangunan model peramalan yang lebih akurat dan relevan terhadap dinamika sistem yang diteliti (Wei, 2019).

2.5 Deret Waktu Multivariat

Deret waktu multivariat merupakan serangkaian data pengamatan yang melibatkan beberapa variabel yang saling berhubungan dan berubah seiring waktu secara bersamaan. Berbeda dengan deret waktu tunggal (univariat), pendekatan ini mampu menggambarkan dinamika temporal serta keterkaitan antar variabel yang terjadi secara simultan (Wei, 2019).

Salah satu deret waktu multivariat adalah *Vector Autoregressive* (VAR) yaitu metode yang menganalisis keterkaitan dinamis antar beberapa variabel endogen dalam deret waktu multivariat (Savada dkk., 2025). Kemudian VAR dikembangkan menjadi *Panel Vector Autoregressive* (PVAR) yang digunakan untuk data panel yaitu data yang memiliki dua dimensi yakni waktu dan wilayah.

2.5.1 *Vector Autoregressive* (VAR)

Model *Vector Autoregressive* (VAR) adalah metode yang menganalisis keterkaitan dinamis antar beberapa variabel endogen dalam deret waktu multivariat. Pada model VAR, setiap variabel direpresentasikan sebagai fungsi dari nilai masa lalunya sendiri serta nilai masa lalu variabel lain dalam sistem, sehingga memungkinkan pemodelan hubungan saling ketergantungan antarvariabel. Pendekatan ini banyak diterapkan dalam bidang ekonomi dan keuangan untuk tujuan peramalan dan analisis dinamika multivariat (Savada dkk., 2025). Model VAR mensyaratkan bahwa seluruh data yang digunakan harus berada dalam keadaan stasioner. Bentuk umum model VAR(p) dapat dituliskan dalam Persamaan (1) sebagai berikut:

$$y_t = \alpha + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \varepsilon_t \quad (1)$$

Keterangan:

y_t : vektor berukuran $m \times 1$ yang berisi m variabel endogen pada waktu t dan periode sebelumnya $t - i$, dengan $i = 1, 2, \dots, p$;

α : vektor konstanta berukuran $m \times 1$;

Φ_i : matriks parameter variabel endogen berukuran $m \times m$ untuk setiap $i = 1, 2, \dots, p$;

ε_t : vektor galat berukuran $m \times 1$.

2.5.2 Panel Vector Autoregressive (PVAR)

Model PVAR pada dasarnya bekerja seperti VAR standar ataupun VAR yang melibatkan variabel eksogen (VARX), di mana seluruh variabel tetap dianggap endogen dan saling berkaitan. Perbedaannya terletak pada penambahan dimensi lintas-unit (*cross-sectional*), sehingga model dapat mengakomodasi analisis bersama untuk beberapa unit $i = 1, \dots, N$. Unit tersebut dapat berupa negara, sektor, pasar, atau kombinasi dari semuanya (Kapetanios dkk., 2021).

Bentuk umum dari model *Panel Vector Autoregressive* (PVAR) dengan panjang lag hingga p dapat dituliskan pada Persamaan (2) sebagai berikut (Irfan dkk., 2023):

$$\begin{aligned} y_{it} &= A_1 y_{i,t-1} + \cdots + A_p y_{i,t-p} + v_{it} \\ i &\in \{1, 2, \dots, N\}, \quad t \in \{1, 2, \dots, T\} \end{aligned} \quad (2)$$

Keterangan:

y_{it} : vektor dari variabel endogen untuk unit ke- i pada waktu ke- t

$y_{i,t-l}$: vektor variabel endogen dengan lag ke- l , $l = 1, \dots, p$

A_l : matriks parameter berukuran $k \times k$, $l = 1, \dots, p$

v_{it} : vektor error berukuran $k \times 1$

i : indeks cross-section, $i = 1, \dots, N$

t : indeks waktu, $t = 1, \dots, T_i$

p : panjang lag model PVAR.

2.6 Kestasioneran Data

Stasioneritas merupakan kondisi ketika data deret waktu memiliki nilai observasi yang cenderung mendatar sepanjang periode pengamatan dimana fluktuasi nilai terjadi di sekitar rata-rata yang konstan (Irfan dkk., 2023). Ada beberapa cara untuk menguji apakah data bersifat stasioner atau tidak diantaranya sebagai berikut:

1. *Levin Lin Chu*

Metode uji akar unit panel Levin–Lin–Chu (LLC) diperkenalkan oleh Levin, Lin, dan Chu (2002). Uji ini merupakan pengembangan dari uji *Augmented Dickey–Fuller* (ADF) yang diterapkan pada data panel untuk mengidentifikasi apakah data mengandung akar unit atau bersifat stasioner secara bersama-sama. Hipotesis yang digunakan dalam uji Levin–Lin–Chu (LLC) adalah sebagai berikut:

$H_0 : \delta = 0$ (data panel mengandung akar unit atau tidak stasioner).

$H_1 : \delta < 0$ (data panel tidak mengandung akar unit atau stasioner).

Levin dkk. (2002) menyatakan bahwa statistik uji akhir yang digunakan dalam pengujian unit root adalah statistik t yang telah disesuaikan (*adjusted t-statistic*), dijabarkan pada Persamaan (3) yaitu:

$$t_{\delta}^* = \frac{t_{\delta} - NT\tilde{S}_N\hat{\sigma}^{-2}STD(\hat{\delta})\mu_{mT}^*}{\sigma_{mT}^*} \quad (3)$$

Keterangan:

N : jumlah individu dalam panel.

T : jumlah periode waktu dalam panel.

$\tilde{S}_N$: rata-rata rasio simpangan baku jangka panjang terhadap jangka pendek.

$\hat{\sigma}^2$: estimasi varians error regresi gabungan.

$STD(\hat{\delta})$: standar error dari koefisien $\hat{\delta}$.

μ_{mT} : faktor penyesuaian rata-rata (*mean adjustment*).

σ_{mT} : faktor penyesuaian simpangan baku (*standard deviation adjustment*).

Statistik uji t_{δ}^* tersebut mengikuti distribusi normal standar secara asimtotik dengan formulasi dalam Persamaan (4) yaitu:

$$t_{\delta}^* \sim N(0, 1) \quad (4)$$

Dengan demikian, keputusan pengujian dilakukan dengan membandingkan nilai t_{δ}^* terhadap nilai kritis distribusi normal standar. Apabila nilai t_{δ}^* lebih kecil daripada nilai kritis pada tingkat signifikansi tertentu, maka hipotesis nol ditolak sehingga data panel dinyatakan stasioner (Levin dkk., 2002).

2. Uji Im–Pesaran–Shin (IPS)

Uji akar unit panel Im–Pesaran–Shin (IPS) diperkenalkan oleh Im, Pesaran, dan Shin (2003). Uji ini merupakan pengembangan dari uji akar unit berbasis *Augmented Dickey–Fuller* (ADF) yang diterapkan pada data panel. Berbeda dengan uji Levin–Lin–Chu (LLC) yang mengasumsikan parameter autoregresif homogen, uji IPS mengizinkan adanya heterogenitas antar individu, sehingga setiap unit dalam panel diperbolehkan memiliki koefisien autoregresif yang berbeda (Im dkk., 2003).

Hipotesis yang digunakan dalam uji IPS adalah sebagai berikut:

$H_0: \beta_i = 0$ (untuk seluruh i data panel mengandung akar unit).

$H_1: \beta_i < 0$ (untuk sebagian sebagian individu dalam panel bersifat stasioner).

Uji IPS dilakukan dengan menghitung statistik uji ADF untuk masing-masing individu, kemudian mengambil rata-rata statistik tersebut. Statistik rata-rata tersebut didefinisikan dalam Persamaan (5) sebagai:

$$\bar{t}_{NT} = \frac{1}{N} \sum_{i=1}^N t_{iT} \quad (5)$$

Selanjutnya, statistik uji akhir IPS dinyatakan dalam bentuk statistik normal standar, pada Persamaan (6) yaitu:

$$Z_{tbar} = \frac{\sqrt{N} (\bar{t}_{NT} - E(t_T))}{\sqrt{Var(t_T)}} \quad (6)$$

Keterangan:

$Z_{t\bar{a}r}$	: statistik uji akhir Im–Pesaran–Shin (IPS).
$\bar{t}_{NT}$	: rata-rata statistik ADF individu dalam panel.
t_{iT}	: statistik uji ADF untuk individu ke- i .
N	: jumlah individu dalam panel.
T	: jumlah periode waktu dalam panel.
$E(t_T)$	: nilai harapan dari statistik ADF.
$Var(t_T)$	: varians dari statistik ADF.
β_i	: koefisien autoregresif individu ke- i .

Statistik uji $Z_{t\bar{a}r}$ mengikuti distribusi normal standar secara asimtotik dengan formulasi pada Persamaan (7), yaitu:

$$Z_{t\bar{a}r} \sim N(0, 1) \quad (7)$$

Keputusan pengujian dilakukan dengan membandingkan nilai $Z_{t\bar{a}r}$ terhadap nilai kritis distribusi normal standar. Apabila nilai statistik uji lebih kecil daripada nilai kritis pada tingkat signifikansi tertentu, maka hipotesis nol ditolak sehingga data panel dinyatakan stasioner (Im dkk., 2003).

2.7 Estimasi Model dengan *Generalized Method of Moments* (GMM)

Generalized Method of Moments (GMM) diperkenalkan oleh Hansen (1982) sebagai metode estimasi berbasis kondisi momen yang memungkinkan pendugaan parameter tanpa menspesifikasikan distribusi penuh dari *error term*. Metode ini tetap konsisten meskipun terdapat heteroskedastisitas maupun pelanggaran asumsi klasik tertentu. Dalam metode momen klasik, jumlah kondisi momen harus sama dengan jumlah parameter yang diestimasi. Namun, ketika jumlah kondisi momen melebihi jumlah parameter (*over-identified*), pendekatan GMM tetap dapat digunakan dengan meminimalkan fungsi kriteria berupa jumlah kuadrat momen berbobot. Oleh karena itu, GMM banyak digunakan dalam analisis data panel dinamis yang berpotensi mengandung endogenitas dan efek individual (Utami dkk., 2014).

Secara asimtotik, estimator GMM bersifat konsisten dan berdistribusi normal, serta mencapai efisiensi apabila matriks pembobot optimal digunakan. Prinsip dasar GMM adalah memanfaatkan kondisi ortogonalitas antara variabel instrumen dan error term, yaitu $E[z_{it}e_{it}] = 0$, dengan z_{it} merupakan variabel instrumen yang tidak

berkorelasi dengan error (Hayakawa, 2016).

Model pada Persamaan (2) ditransformasikan menggunakan metode *Forward Orthogonal Deviation* (FOD). Transformasi ini menghasilkan variabel tertransformasi yang dinotasikan dengan simbol (*), dijabarkan pada Persamaan (9) yaitu:

$$y_{i,t-l}^* = c_t \left[y_{i,t-l} - \frac{y_{i,t-l+1} + \dots + y_{i,T-l}}{T-t} \right], \quad (9)$$

dengan $l = 0, 1, \dots, p$ dan $c_t^2 = \frac{T-t}{T-t+1}$.

Setelah dilakukan transformasi, model yang telah ditransformasikan selanjutnya disusun dengan menumpuk observasi sepanjang dimensi waktu. Melalui proses ini, model dapat dinyatakan dalam bentuk matriks Persamaan (10) sebagai berikut:

$$\mathbf{Y}_i^* = \mathbf{X}_i^* \mathbf{A}' + \mathbf{V}_i^* \quad (10)$$

dengan

$$\begin{aligned} \mathbf{Y}_i^* &= (\mathbf{y}_{i1}^*, \dots, \mathbf{y}_{iT_1}^*)', \\ \mathbf{X}_i^* &= (\mathbf{x}_{i1}^*, \dots, \mathbf{x}_{iT_1}^*)', \\ \mathbf{V}_i^* &= (\mathbf{v}_{i1}^*, \dots, \mathbf{v}_{iT_1}^*)', \\ \mathbf{A} &= (\mathbf{A}_1, \dots, \mathbf{A}_p). \end{aligned}$$

Untuk mempermudah proses estimasi, Persamaan (10) kemudian dinyatakan dalam bentuk vektorisasi dengan formulasi Persamaan (11) sebagai berikut:

$$\text{vec}(\mathbf{Y}_i^*) = (\mathbf{I}_K \otimes \mathbf{X}_i^*) \text{vec}(\mathbf{A}') + \text{vec}(\mathbf{V}_i^*) \quad (11)$$

Persamaan tersebut menunjukkan bahwa model telah dinyatakan dalam bentuk yang siap untuk diestimasi menggunakan pendekatan GMM.

2.8 Identifikasi dan Validasi Model PVAR

Identifikasi model PVAR dimulai dengan menentukan orde lag optimal dari model PVAR. Penentuan lag dilakukan dengan mempertimbangkan indikator statistik seperti MBIC, MAIC, MQIC, serta dengan kriteria pemilihan model dan momen yang konsisten (MMSC) dengan mencari nilai minimumnya (Boruku dkk., 2022),

sehingga dipilih lag yang paling mencerminkan dinamika data.

Setelah model terbentuk, proses yang dilakukan yaitu validasi model melalui pengujian stabilitas model. Model PVAR dinyatakan stabil ketika seluruh nilai modulus akar karakteristik berada dalam radius satu, yang berarti seluruh variabel dalam model tidak menghasilkan respons divergen terhadap guncangan (Haloui dkk., 2025). Kemudian dilakukan uji Hansen test untuk memastikan apakah instrumen yang digunakan valid.

2.8.1 Pemilihan lag optimal

Untuk menentukan panjang lag yang paling sesuai, analisis PVAR membutuhkan pemilihan orde lag yang optimal. Andrews dan Lu pada tahun 2001 memperkenalkan kriteria pemilihan model dan momen yang konsisten (MMSC) untuk model GMM. Kriteria ini didasarkan pada statistik *Joint* dari Hansen tahun 1982 yang terkait dengan pembatasan *over-identifying*. Kriteria pemilihan model dan momen yang konsisten (MMSC) tersebut bekerja serupa dengan kriteria seleksi model berbasis *maximum likelihood* yang umum digunakan, seperti AIC, BIC, dan HQIC (Boruku dkk., 2022).

Ketika kriteria MMSC yang dikembangkan oleh Andrews dan Lu diterapkan pada estimator GMM, kriteria tersebut akan memilih pasangan nilai (p, q) yang meminimalkan fungsi evaluasinya dengan Persamaan (12), (13), dan (14) sebagai berikut (Boruku dkk., 2022).

$$\text{MMSC}_{BIC}(n, k, p, q) = \ln |\hat{\Sigma}| + \frac{qkp \ln n}{n} \quad (12)$$

$$\text{MMSC}_{AIC}(n, k, p, q) = \ln |\hat{\Sigma}| + \frac{2qkp}{n} \quad (13)$$

$$\text{MMSC}_{HQIC}(n, k, p, q) = \ln |\hat{\Sigma}| + \frac{2qkp \ln(\ln n)}{n} \quad (\text{untuk } q > p) \quad (14)$$

Statistik $J_n(k, p, q)$ merujuk pada *J-statistic* untuk pengujian pembatasan *over-identifying* dalam PVAR yang terdiri dari k variabel, dengan lag sebesar p , serta kondisi momen yang dibangun dari q lag variabel dependen, dan ukuran sampel sebesar n .

2.8.2 Uji Stabilitas Model

Uji stabilitas pada model dilakukan untuk memastikan bahwa estimasi koefisien dalam PVAR tetap menunjukkan signifikansi statistik meskipun terdapat variasi dalam proses pengambilan sampel, serta untuk mencegah munculnya singularitas pada penyelesaian model. Stabilitas model merupakan prasyarat agar analisis dinamika selanjutnya termasuk *impulse response* dan *dekomposisi varians* dapat menghasilkan temuan yang valid dan dapat dipercaya (Li dkk., 2024).

Secara teoritis, suatu proses VAR/PVAR dinyatakan stabil apabila seluruh akar dari polinomial karakteristik berada di luar lingkaran satuan, atau ekuivalen dengan seluruh *eigenvalue* sistem memiliki modulus kurang dari satu ($|\lambda| < 1$). (Lütkepohl, 2005) menegaskan bahwa pemenuhan kondisi stabilitas ini diperlukan agar model memiliki representasi yang terdefinisi dengan baik dan interpretasi terhadap *impulse response function* serta analisis variansi dapat dilakukan secara bermakna.

2.8.3 Uji Validitas Instrumen (Hansen test)

Setelah melakukan estimasi pada berbagai model, dilakukan sejumlah pengujian untuk memastikan konsistensi estimator GMM serta menegaskan kelayakan penggunaan pendekatan sistem GMM. Uji Hansen *over-identification* (J Test) diterapkan untuk mengevaluasi keabsahan instrumen yang digunakan dalam model GMM. Uji ini memeriksa apakah instrumen bersifat eksogen atau tidak berkorelasi dengan residual model (Cangombe dkk., 2025).

Hipotesis nol dalam uji Hansen menyatakan bahwa instrumen adalah valid, yaitu tidak berkaitan dengan error, sehingga kondisi *over-indentification* model dianggap terpenuhi. Apabila nilai uji Hansen signifikan, hipotesis nol ditolak, yang berarti instrumen tidak valid dan model berpotensi menghadapi masalah endogenitas (Cangombe dkk., 2025).

2.9 Analisis Dinamis PVAR

Setelah identifikasi dan validasi model PVAR seperti pemilihan lag optimal, uji stabilitas model dan uji validitas instrumen, analisis tambahan dilakukan melalui analisis *Impulse Response Function* (IRF) dan *Forecast Error Variance Decomposition* (FEVD).

Impulse Response Function (IRF) digunakan untuk menganalisis bagaimana suatu variabel dependen/endogen bereaksi ketika variabel dependen/endogen lain mengalami *shock* sedangkan *Forecast-error variance decomposition* (FEVD) menggambarkan proporsi perubahan suatu variabel yang dapat dijelaskan oleh guncangan (*shock*) pada variabel lain (Kazemzadeh dkk., 2023).

2.9.1 *Impulse Response Function* (IRF)

Impulse Response Function (IRF) digunakan untuk menganalisis bagaimana suatu variabel dependen/endogen bereaksi ketika variabel dependen/endogen lain mengalami *shock* (Kazemzadeh dkk., 2023). Melalui IRF, perkembangan respons suatu variabel dapat diamati selama beberapa periode setelah terjadinya satu *shock* tertentu, dengan asumsi bahwa tidak terdapat guncangan lain yang terjadi secara bersamaan pada variabel lain. Dengan asumsi tersebut, respons yang dihasilkan dapat diinterpretasikan sebagai dampak murni dari *shock* yang dianalisis, sehingga hubungan sebab-akibat antar variabel dapat dijelaskan dengan lebih jelas (Love dan Zicchino, 2006).

Dalam praktiknya, data sering menunjukkan adanya korelasi antar error pada variabel-variabel yang dianalisis. Oleh karena itu, digunakan *orthogonalized impulse response function* untuk memisahkan pengaruh masing-masing guncangan. Proses ortogonalisasi dilakukan dengan menetapkan urutan variabel tertentu dalam model VAR yang mencerminkan asumsi identifikasi jangka pendek. Variabel yang ditempatkan lebih awal dalam urutan diasumsikan bersifat lebih eksogen, sedangkan variabel yang ditempatkan lebih akhir dianggap lebih endogen. Dengan pendekatan tersebut, respons impuls yang dihasilkan dapat ditafsirkan sebagai pengaruh dari satu *shock* tertentu tanpa tercampur oleh *shock* lainnya. Analisis IRF diharapkan mampu memberikan gambaran yang lebih komprehensif mengenai

pola respons dinamis antar variabel, serta menjelaskan bagaimana suatu guncangan memengaruhi sistem secara keseluruhan dalam jangka pendek maupun jangka menengah (Love dan Zicchino, 2006).

2.9.2 Forecast Error Variance Decomposition (FEVD)

Forecast-error variance decomposition (FEVD) digunakan untuk mengukur seberapa besar kontribusi guncangan suatu variabel, baik dari dirinya sendiri maupun variabel lain, dalam menjelaskan variansi kesalahan peramalan suatu variabel (Kazemzadeh dkk., 2023). FEVD menggambarkan proporsi variansi kesalahan peramalan variabel j yang disebabkan oleh inovasi pada variabel k didefinisikan dalam Persamaan (15) sebagai:

$$\omega_{i,j}(h) = \frac{\sum_{k=0}^{h-1} (\Psi_k)_{i,j}^2}{\sum_{k=0}^{h-1} \sum_{j=1}^N (\Psi_k)_{i,j}^2} \quad (15)$$

Keterangan:

- $\omega_{i,j}(h)$: kontribusi variabel j terhadap variabel i pada horizon h ,
- Ψ_k : matriks respon impuls untuk lag ke- k ,
- N : jumlah variabel dalam model PVAR.

Rumus tersebut membagi total variansi kesalahan peramalan suatu variabel menjadi bagian-bagian yang disebabkan oleh guncangan terhadap dirinya sendiri (*own shocks*) dan oleh variabel lain (*cross shocks*). Nilai $\omega_{i,j}(h)$ yang semakin besar menunjukkan bahwa *shock* dari variabel k memberikan pengaruh yang lebih besar terhadap variansi kesalahan peramalan variabel j pada horizon waktu tertentu.

2.10 Machine Learning dan Deep Learning

Machine Learning (ML) adalah cabang kecerdasan buatan yang mengembangkan algoritma untuk mengenali pola dalam data dan memanfaatkan pola tersebut guna menghasilkan prediksi atau keputusan terhadap data baru. Proses belajar ini pada dasarnya banyak memanfaatkan teknik statistik sehingga model dapat mengekstraksi pola secara efektif. *Machine learning* memanfaatkan berbagai algoritma baik yang bersifat *supervised* maupun *unsupervised* untuk mempelajari

tugas tertentu, seperti klasifikasi maupun pengelompokkan, berdasarkan data yang diberikan (Barbierato & Gatti, 2024).

Deep learning adalah salah satu konsep dari *machine learning* yang didasarkan pada jaringan saraf tiruan. Dalam banyak kasus, model *deep learning* mampu memberikan hasil yang lebih baik dibandingkan model *machine learning* yang lebih sederhana serta metode analisis data tradisional. Jaringan saraf dalam (*deep neural networks*) memiliki kemampuan lebih unggul dalam belajar dan dapat secara otomatis mengekstraksi fitur penting dari data dengan sedikit campur tangan manusia. Hal ini membuat *deep learning* mampu mengolah data yang besar dan tidak terstruktur dengan lebih efektif dibandingkan metode *machine learning* (Müller & Guido, 2017).

2.11 *Preprocessing* Data untuk LSTM

Preprocessing merupakan tahap awal pengolahan data yang bertujuan menyesuaikan skala, distribusi, dan representasi fitur sehingga data lebih sesuai digunakan oleh algoritma pembelajaran mesin. Tahap ini meliputi teknik seperti *cleaning data*, normalisasi, standarisasi, dan transformasi lainnya yang tidak melibatkan informasi label (*unsupervised*) (Müller & Guido, 2017). Berikut beberapa teknik pada *preprocessing*:

1. *Data Cleaning*

Pembersihan data, yang juga dikenal sebagai *data cleaning*, adalah teknik untuk mendeteksi dan memperbaiki kesalahan serta ketidakakuratan dalam data yang dikumpulkan. Data harus sesuai dengan kebutuhan *input* pada model *machine learning*. Kegiatan utama dalam proses pembersihan data meliputi penyempurnaan data yang berisik (*noisy data*) dan penanganan data yang hilang (*missing data*). Proses ini membantu memastikan bahwa kumpulan data yang diperoleh bersifat lengkap dan bebas dari kesalahan maupun bias yang mungkin muncul selama proses pengumpulan data (Ndung'u, 2022).

2. Normalisasi data

Normalisasi data merupakan salah satu teknik prapemrosesan yang bertujuan untuk menyesuaikan skala data agar setiap fitur memberikan kontribusi

yang seimbang dalam evaluasi kinerja model. Proses ini dilakukan dengan menyeragamkan nilai dari fitur-fitur yang memiliki skala berbeda ke dalam satu skala yang sama, sehingga dapat meminimalkan bias yang timbul akibat adanya variabel tertentu yang memiliki pengaruh lebih besar dibandingkan variabel lainnya (Kim dkk., 2025).

a. Normalisasi Min-Max

Normalisasi minimum–maksimum (*min–max normalization*) merupakan transformasi linear terhadap data asli dengan tujuan membatasi nilai hasil ke dalam rentang 0 sampai 1. Teknik ini digunakan agar perbedaan skala antar variabel tidak mempengaruhi proses pelatihan model. Salah satu metode normalisasi yang umum digunakan adalah *Min–Max Normalization*, dengan rumus Persamaan (16) sebagai berikut:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (16)$$

dengan $x_{\max}$ menyatakan nilai maksimum dan $x_{\min}$ menyatakan nilai minimum dari data sampel (Mahesh dkk., 2024).

b. Normalisasi Z-score

Metode ini dikenal juga sebagai standarisasi, merupakan salah satu teknik normalisasi yang banyak diterapkan dalam bidang kecerdasan buatan dan penelitian statistik. Pendekatan ini bertujuan untuk mentransformasikan data ke dalam bentuk standar, di mana setiap nilai data dinyatakan berdasarkan tingkat penyimpangannya terhadap nilai pusat. Normalisasi Z-score dilakukan dengan mengubah setiap nilai sehingga memiliki nilai rata-rata sebesar 0 dan standar deviasi sebesar 1. Persamaan yang digunakan untuk melakukan standarisasi pada suatu fitur x dirumuskan pada Persamaan (17) sebagai berikut:

$$x' = \frac{x - \text{mean}(x)}{\text{std}(x)} \quad (17)$$

dengan x menyatakan nilai data asli, x' merupakan nilai hasil standarisasi, $\text{mean}(x)$ adalah nilai rata-rata dari fitur x , dan $\text{std}(x)$ menunjukkan standar deviasi dari fitur tersebut (Kim dkk., 2025).

c. RobustScaler

RobustScaler merupakan metode normalisasi data yang bertujuan untuk meminimalkan pengaruh *outlier* terhadap proses pemodelan. Berbeda dengan *StandardScaler* yang melakukan transformasi data berdasarkan nilai rata-rata dan standar deviasi, RobustScaler menggunakan nilai median dan *interquartile range* (IQR) sebagai dasar penskalaan. Pendekatan ini membuat RobustScaler lebih stabil ketika diterapkan pada data dengan distribusi tidak simetris maupun yang mengandung nilai ekstrem. Proses transformasi RobustScaler dirumuskan pada Persamaan (18) sebagai berikut:

$$x_{\text{scaled}} = \frac{x - Q_2}{Q_3 - Q_1} \quad (18)$$

dengan Q_2 menyatakan nilai median, sedangkan Q_1 dan Q_3 masing-masing merupakan kuartil pertama dan kuartil ketiga (Lasiyono dkk., 2025).

2.12 Fungsi Aktivasi

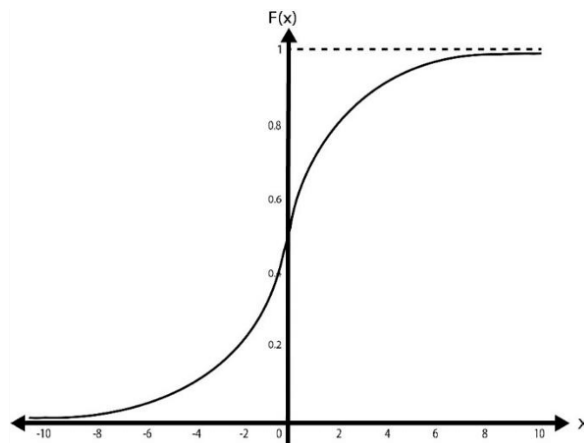
Fungsi aktivasi merupakan mekanisme yang menentukan apakah suatu neuron dalam jaringan saraf akan aktif atau tidak. Fungsi ini bekerja dengan mengolah informasi yang masuk melalui perhitungan gradien tertentu, kemudian menghasilkan *output* yang akan diteruskan ke jaringan saraf berikutnya (Kurniawan dkk., 2024). Beberapa fungsi aktivasi yang umum digunakan yaitu:

1. Fungsi Sigmoid

Fungsi aktivasi sigmoid merupakan fungsi matematis dengan karakteristik kurva berbentuk huruf “S”, yang mengubah nilai masukan menjadi rentang antara 0 hingga 1. Fungsi ini sering disebut sebagai fungsi logistik atau kurva Sigmoid. Sigmoid termasuk salah satu fungsi aktivasi non-linier yang biasa digunakan dalam jaringan saraf. Adapun bentuk matematis dari fungsi Sigmoid dapat dituliskan dalam persamaan (19) sebagai berikut:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (19)$$

Pada persamaan tersebut, $\sigma(x)$ merupakan *output* yang dihasilkan oleh fungsi Sigmoid ketika menerima *input* x , sedangkan e adalah konstanta Euler.



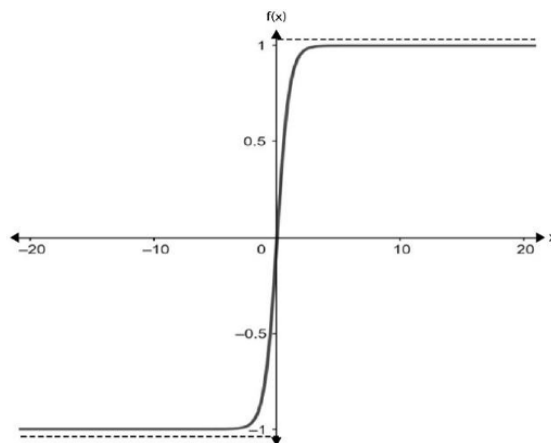
Gambar 1. Fungsi Aktivasi Sigmoid (Akbar dkk., 2022).

2. Fungsi Tanh

Fungsi aktivasi Tanh atau *Tangen Hiperbolik*, merupakan salah satu fungsi aktivasi yang sering digunakan dalam berbagai model *deep learning*. Fungsi ini memiliki karakteristik yang hampir mirip dengan fungsi Sigmoid, tetapi berbeda pada rentang keluarannya. Jika fungsi Sigmoid memetakan nilai pada rentang 0 hingga 1, maka fungsi Tanh memetakan nilai pada rentang -1 hingga 1 dijabarkan pada Persamaan (20) yaitu:

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (20)$$

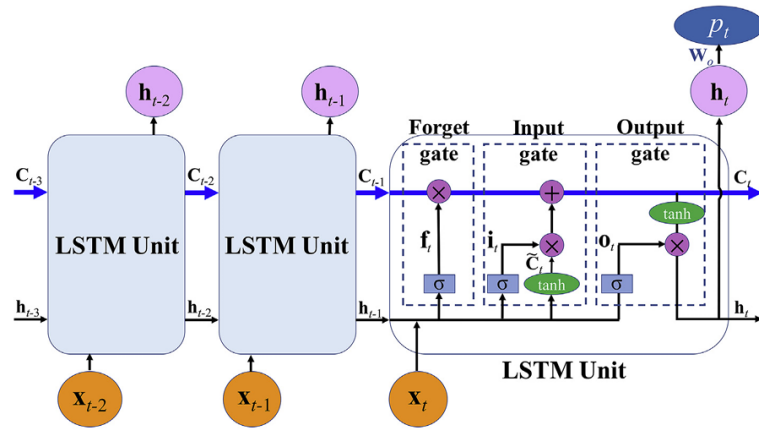
Pada persamaan tersebut, $\text{Tanh}(x)$ merupakan keluaran yang dihasilkan oleh fungsi Tanh ketika menerima masukan x , sedangkan e merepresentasikan konstanta Euler.



Gambar 2. Fungsi Aktivasi Tanh (Akbar dkk., 2022).

2.13 Long Short-Term Memory (LSTM)

LSTM adalah jenis khusus dari jaringan saraf berulang (*recurrent neural network*) yang menggunakan pengamatan berurutan, model ini belajar dari pola sebelumnya untuk memprediksi pola di masa depan, dengan tetap mengingat informasi penting dari urutan data. Dalam LSTM, setiap unit terdiri dari beberapa sel yang menyimpan aliran informasi. Sel-sel ini saling terhubung seperti jalur yang membawa data dari masa lalu ke masa kini. Di dalam setiap sel terdapat beberapa “gerbang (*gate*)” yang bertugas mengatur informasi mana yang perlu dibuang, disimpan, atau diteruskan (Dissanayake dkk., 2021). Gerbang-gerbang tersebut adalah sebagai berikut :



Gambar 3. Arsitektur LSTM (Wang dkk., 2024).

1. *forget gate*

Gerbang pertama adalah *forget gate*. Pada tahap ini, LSTM memutuskan informasi mana yang sudah tidak relevan dan perlu dibuang dengan menggunakan fungsi aktivasi sigmoid. Nilai x_t merupakan *input* pada waktu t , sedangkan h_{t-1} adalah *hidden state* dari langkah sebelumnya. Persamaan (21) digunakan untuk menghitung nilai dari *forget gate* (Rowan dkk., 2022).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (21)$$

Keterangan:

- σ : fungsi aktivasi
- W_f : nilai bobot *forget gate*
- h_{t-1} : vektor *hidden state* pada timestep sebelumnya
- x_t : data *input* pada waktu t
- b_f : nilai bias
- f_t : nilai *forget gate*.

2. *Input gate*

Gerbang kedua adalah *input gate*. Pada tahap ini, LSTM menentukan informasi baru mana yang perlu ditambahkan atau diperbarui ke dalam *cell state* dengan menggunakan fungsi aktivasi sigmoid. Persamaan (22) digunakan untuk menghitung nilai pada *input gate* (Rowan dkk., 2022).

$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i) \quad (22)$$

Keterangan:

- σ : fungsi aktivasi
- W_i : nilai bobot *input gate*
- h_{t-1} : vektor *hidden state* pada timestep sebelumnya
- x_t : data masuk
- b_i : bias
- i_t : nilai *input gate*

3. *Output gate*

Gerbang ketiga yaitu *output gate*, yang menggunakan fungsi aktivasi sigmoid untuk menentukan keluaran pada *hidden state*. Pada tahap ini, nilai *cell state* juga diolah dengan fungsi *tanh* sebelum digabungkan dengan hasil aktivasi sigmoid. Persamaan (23) digunakan untuk menghitung nilai pada *output gate* (Rowan dkk., 2022).

$$o_t = \sigma (W_o \cdot [h_{t-1}, x_t] + b_o) \quad (23)$$

Keterangan:

- σ : fungsi aktivasi
- W_o : nilai bobot *output gate*
- h_{t-1} : vektor *hidden state* pada timestep sebelumnya
- x_t : data masuk
- b_o : bias
- o_t : nilai *output gate*

2.14 Hyperparameter Tuning

Hyperparameter tuning merupakan tahapan yang sangat penting dalam meningkatkan tingkat akurasi serta kemampuan generalisasi model *machine learning*. Oleh sebab itu, proses ini menjadi aspek yang krusial dalam peramalan deret waktu (*time series forecasting*), khususnya ketika menggunakan model yang bersifat kompleks. *Hyperparameter tuning* dilakukan dengan menentukan nilai *hyperparameter* yang paling tepat dalam suatu model *machine learning*, optuna menjadi salah satu alat optimasi *hyperparameter* yang paling sering dimanfaatkan untuk tujuan tersebut. Optuna merupakan *framework* untuk *hyperparameter tuning* yang dirancang dengan mekanisme *sampling* dan *pruning* yang efektif guna memperoleh konfigurasi model yang optimal meskipun berada pada keterbatasan sumber daya komputasi. Melalui penerapan optimasi Bayesian, Optuna mampu menyempitkan ruang pencarian *hyperparameter* secara adaptif dan mengidentifikasi kombinasi parameter terbaik untuk suatu model (Rehman & Iqbal, 2025).

Tahap *hyperparameter tuning* dilakukan untuk menyesuaikan parameter-parameter model dengan tujuan menilai sejauh mana tingkat akurasi yang dapat dicapai oleh model yang diimplementasikan. Proses ini juga berperan dalam melatih data latih dan data validasi sehingga model mampu mencapai performa prediksi yang paling optimal (Isa dkk., 2022). Beberapa *hyperparameter* yang *dituning* untuk memperoleh performa model yang optimal diantaranya yaitu jumlah LSTM *units* menggambarkan banyaknya neuron tersembunyi yang bertugas menyimpan informasi urutan data dari waktu ke waktu. Semakin banyak unit yang digunakan, semakin besar kapasitas model dalam mengenali pola temporal yang kompleks, tetapi berpotensi meningkatkan risiko *overfitting* dan waktu komputasi (Amini

dan Dolatshahi, 2023). Lapisan *Dense* berfungsi menghubungkan seluruh neuron pada lapisan sebelumnya dengan semua neuron di lapisan berikutnya. Lapisan ini bertugas menggabungkan representasi hasil pembelajaran dari jaringan sebelumnya untuk menghasilkan prediksi akhir yang lebih akurat (Liao dkk., 2022). Kemudian Terdapat pula *dropout* Untuk mencegah *overfitting* dengan cara menonaktifkan neuron secara acak selama proses pelatihan, *learning rate* untuk menentukan besar langkah dalam pembaruan bobot model dan *batch size* digunakan untuk mengontrol jumlah sampel yang diproses sebelum setiap pembaruan bobot (Rehman & Iqbal, 2025). *embedding* representasi vektor berdimensi rendah yang dipelajari untuk variabel kategorikal agar bisa digunakan (Lim dkk., 2021).

2.15 Model *Hybrid* PVAR–LSTM

Dalam analisis deret waktu, sering muncul kesulitan dalam menentukan apakah pola data bersifat linear atau non-linear. Pada praktiknya, data deret waktu umumnya mengandung kombinasi pola linear dan non-linear secara bersamaan. Jika hanya menggunakan satu pendekatan peramalan, maka model tersebut sering kali tidak mampu menangkap keseluruhan karakteristik data secara optimal. Oleh karena itu, pendekatan model *hybrid* dikembangkan dengan mengintegrasikan metode statistik klasik dan teknik *deep learning* agar dapat memodelkan pola yang beragam dalam data (Warsono dkk., 2025).

Model *hybrid* yang menggabungkan metode statistik dengan *deep learning* merupakan salah satu pendekatan modern dalam peramalan deret waktu. Metode statistik umumnya efektif dalam mengidentifikasi hubungan linear antar variabel, sedangkan *deep learning* memiliki keunggulan dalam mengenali pola non-linear dan struktur data yang lebih kompleks. Dengan mengombinasikan kedua pendekatan tersebut, model *hybrid* berupaya mengurangi keterbatasan masing-masing metode sehingga menghasilkan prediksi yang lebih akurat (Warsono dkk., 2025).

Model *hybrid* PVAR–LSTM merupakan gabungan antara model klasik PVAR dan model *deep learning* LSTM. Pendekatan ini mengombinasikan keunggulan PVAR yang mampu menangkap hubungan dinamis dan pola linear dalam data panel dengan kemampuan LSTM dalam mengenali pola data yang kompleks serta ketergantungan jangka panjang. Dengan cara ini, model *hybrid*

dapat menangani data deret waktu yang memiliki hubungan linear dan non-linear secara bersamaan sehingga meningkatkan akurasi prediksi. Proses penggabungan dilakukan dengan menggunakan output dari model PVAR sebagai *input* untuk model LSTM, sehingga menghasilkan prediksi akhir yang lebih akurat dan adaptif terhadap pola data yang dinamis.

2.16 Evaluasi Model

Dalam evaluasi model peramalan deret waktu, beberapa metrik umum digunakan untuk menilai akurasi model, yaitu RMSE, MAE, dan MAPE. *Root Mean Squared Error* (RMSE) merupakan salah satu metrik yang paling umum digunakan untuk menilai tingkat akurasi prediksi suatu model. Metrik ini menghitung selisih antara nilai aktual dan nilai prediksi, kemudian membandingkan besar kesalahan antar model pada suatu data tertentu. Jika y_i menyatakan nilai aktual dan $\hat{y}_i$ menyatakan nilai prediksi untuk sejumlah n observasi, maka nilai RMSE dapat dihitung dengan Persamaan (24) (Dissanayake dkk., 2021).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (24)$$

Mean Absolute Error (MAE) merupakan salah satu metrik yang umum digunakan dalam model regresi multivariat dan berfungsi untuk menghitung rata-rata dari nilai kesalahan absolut antara hasil prediksi dan nilai aktual.

Misalkan x_i adalah nilai aktual dan $\hat{x}_i$ merupakan nilai prediksi untuk sejumlah n observasi, maka nilai MAE dapat dihitung menggunakan Persamaan (25).

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{x}_i - x_i| \quad (25)$$

Mean Absolute Percentage Error (MAPE) merupakan metrik yang digunakan untuk mengukur rata-rata persentase kesalahan absolut antara nilai aktual dan nilai prediksi. Keunggulan utama MAPE adalah sifatnya yang bebas skala (*scale-independent*) sehingga mudah diinterpretasikan dan dapat digunakan untuk membandingkan performa model pada data dengan skala yang berbeda

(Dissanayake dkk., 2021).

Misalkan A_t menyatakan nilai aktual dan F_t adalah nilai hasil prediksi pada waktu ke- t , serta n merupakan jumlah observasi, maka nilai MAPE dapat dihitung menggunakan Persamaan (26).

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (26)$$

2.17 Produksi dan Luas Panen Padi

Produksi pada dasarnya merupakan kegiatan yang bertujuan untuk menciptakan atau meningkatkan nilai suatu barang guna memenuhi kebutuhan masyarakat. Proses produksi meliputi serangkaian tahapan yang diperlukan untuk mengubah berbagai jenis *input* seperti tenaga kerja, bahan baku, dan sarana produksi lainnya menjadi keluaran berupa barang atau jasa. Transformasi tersebut menghasilkan *output* dengan nilai ekonomi yang lebih tinggi daripada *input* yang digunakan. Dalam konteks pertanian, peningkatan produksi tanaman pangan menjadi salah satu fokus penting pemerintah dalam memperkuat sektor pertanian yang berperan sebagai sumber utama penghidupan serta pendapatan bagi masyarakat petani. Upaya peningkatan produksi diharapkan mampu mendukung pembangunan pertanian yang lebih tangguh dan berkelanjutan. Beras sebagai hasil olahan padi merupakan komoditas strategis karena menjadi bahan pokok utama penduduk Indonesia. Oleh sebab itu, ketersediaan produksi padi harus dijaga agar pemenuhan kebutuhan pangan dapat berlangsung secara berkelanjutan. Analisis dan peramalan mengenai potensi produksi padi sangat diperlukan untuk melihat kondisi pemenuhan pangan di masa mendatang dan membantu pemerintah dalam mengambil langkah yang tepat dalam menjaga ketahanan pangan nasional (Rusman dkk., 2023).

Sedangkan luas panen adalah total area lahan pertanian yang telah ditanami dengan tanaman pangan dan kemudian dipanen dalam rentang waktu tertentu, biasanya satu musim atau tahun pertanian. Luas ini menunjukkan seberapa banyak lahan yang berhasil menghasilkan tanaman siap panen dalam suatu wilayah atau negara. Ukuran ini penting untuk mengukur hasil produksi serta memastikan ketersediaan pangan bagi masyarakat.

Dalam konteks padi, produksi dihitung sebagai hasil perkalian antara luas panen dan produktivitas sebagaimana dinyatakan oleh Sofhya (2023). Penelitian tersebut juga menunjukkan adanya perbedaan skala produksi antar provinsi di Indonesia. Berdasarkan pengelompokan terhadap 34 provinsi, diperoleh tiga kelompok utama. Kelompok produksi tinggi terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur. Kelompok produksi sedang yaitu Sulawesi Selatan. Kelompok produksi rendah terdiri dari Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua. Temuan tersebut menunjukkan bahwa produksi padi antar wilayah memiliki skala yang berbeda-beda dan dapat membantu dalam memberikan gambaran umum mengenai karakteristik produksi antar provinsi.

BAB III

METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

3.1.1 Tempat Penelitian

Penelitian ini dilakukan di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang beralamatkan di Jalan Prof. Dr. Ir. Soemantri Brojonegoro, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung, Lampung.

3.1.2 Waktu Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2025/2026 yaitu pada bulan September 2025 - Mei 2026 yang terdiri dari tiga tahap. Tahap pertama melakukan studi literatur, melakukan pengumpulan data, dan menentukan tema penelitian. Tahap kedua adalah tahap pelaksanaan yaitu pengelolaan data yang meliputi input data, *preprocessing*, pembagian data, pemodelan PVAR, membangun model *Hybrid* PVAR-LSTM dan evaluasi kinerja model. Tahap ketiga adalah penyusunan hasil penelitian dan perumusan kesimpulan.

3.2 Data Penelitian

Data yang digunakan merupakan data sekunder yang bersumber dari survei Kerangka Sampel Area (KSA) yang diselenggarakan oleh Badan Pusat Statistik (BPS). Survei KSA bertujuan untuk menghasilkan estimasi luas panen dan produksi

tanaman pangan, khususnya padi, dengan pendekatan berbasis area. Pelaksanaan KSA dilakukan melalui pengamatan langsung terhadap segmen-segmen sampel di lapangan. Hasil pengamatan tersebut kemudian diekstrapolasikan ke populasi untuk memperoleh estimasi luas panen dalam waktu relatif singkat (*rapid estimate*), sehingga informasi luas panen dan produksi padi dapat disajikan secara lebih cepat dan akurat (Arifin dkk., 2023).

Dalam implementasinya, KSA memanfaatkan teknologi digital, di mana petugas lapangan menggunakan perangkat android yang dilengkapi aplikasi KSA untuk mengambil foto dan mengidentifikasi fase pertumbuhan tanaman padi. Setiap titik pengamatan dilengkapi dengan koordinat geografis yang terkunci, sehingga pengambilan data hanya dapat dilakukan apabila petugas berada tepat di lokasi segmen sampel. Selain pengamatan luas panen, KSA juga melakukan pengukuran produktivitas padi menggunakan metode ubinan di sekitar titik segmen sampel. Estimasi produksi padi diperoleh dari kombinasi data luas panen berbasis spasial-digital dan data produktivitas hasil pengukuran lapangan. Proses ini didukung oleh kerja sama lintas lembaga, antara lain Badan Informasi Geospasial (BIG), Badan Riset dan Inovasi Nasional (BRIN), serta Kementerian Agraria dan Tata Ruang/Badan Pertanahan Nasional (ATR/BPN), khususnya dalam penyediaan dan validasi peta lahan baku sawah berbasis citra satelit (Arifin dkk., 2023).

Data penelitian yang digunakan berupa data luas panen dan produksi padi bulanan pada 34 provinsi di Indonesia dengan periode pengamatan Januari 2018 hingga Desember 2023. Data tersebut disimpan dalam *cloud storage* dan dapat diakses melalui tautan <https://tinyurl.com/Datapvar>. Ringkasan data penelitian disajikan pada Tabel 2 sebagai berikut:

Tabel 2. Data Penelitian

Waktu	Provinsi	Luas Panen	Produksi Padi
2018-01	ACEH	14541.53	76118.47
2018-02	ACEH	27075.77	142876.76
2018-03	ACEH	66403.96	375841.91
2018-04	ACEH	58507.03	336742.80
2018-05	ACEH	18687.65	39672.36
2018-06	ACEH	6574.29	37912.07
2018-07	ACEH	16175.29	39281.52
2018-08	ACEH	24947.79	142895.96
2018-09	ACEH	34889.21	192853.32
2018-10	ACEH	25948.67	153129.92
...			
2023-11	PAPUA	2555.38	10348.94
2023-12	PAPUA	480.22	1941.79

3.3 Alat Penelitian

Berikut adalah beberapa alat yang digunakan dalam penelitian ini:

1. Windows 10 64-bit

Sistem operasi ini digunakan sebagai platform kerja karena mendukung komputasi tingkat lanjut, stabil saat menjalankan proses pembelajaran mesin, serta kompatibel dengan berbagai framework berbasis Python.

2. RStudio versi 2025 dan Python versi 3.12.12

RStudio digunakan sebagai lingkungan analisis statistik dan pengolahan data, terutama pada pemodelan berbasis data panel dan verifikasi perhitungan statistik. Python digunakan karena memiliki ekosistem pustaka yang luas, fleksibel, serta banyak diterapkan dalam pengembangan model machine learning dan deep learning.

3. NumPy versi 2.0.2

NumPy digunakan untuk komputasi numerik, operasi array multidimensi,

serta fungsi matematis penting dalam pemrosesan data (Müller & Guido, 2017).

4. Pandas versi 2.2.2

Pandas dimanfaatkan untuk pemrosesan data tabular seperti pembersihan, transformasi, dan penyusunan dataset dalam struktur DataFrame sebelum tahap pelatihan model (Müller & Guido, 2017).

5. Matplotlib versi 3.10.0

Library ini digunakan untuk visualisasi grafik seperti tren deret waktu dan kurva evaluasi model guna mendukung proses eksplorasi data (Müller & Guido, 2017).

6. Seaborn versi 0.13.2

Seaborn digunakan untuk visualisasi statistik tingkat lanjut yang menghasilkan grafik informatif pada proses analisis eksploratif data.

7. SciPy versi 1.16.3

SciPy digunakan untuk komputasi ilmiah lanjutan seperti analisis statistik, optimasi numerik, dan fungsi matematis lainnya yang diperlukan dalam pemodelan data (Müller & Guido, 2017).

8. Scikit-learn versi 1.6.1

Scikit-learn digunakan pada tahap *preprocessing*, normalisasi, pembagian dataset, dan evaluasi model menggunakan metrik kesalahan sebagai pembandingan sebelum diterapkan metode *deep learning* (Müller & Guido, 2017).

9. TensorFlow versi 2.19.0

TensorFlow digunakan sebagai framework dalam pembangunan dan pelatihan model deep learning termasuk arsitektur LSTM karena mampu memproses data skala besar secara paralel.

10. Optuna versi 4.6.0

Optuna digunakan untuk optimasi hiperparameter secara otomatis sehingga meningkatkan efisiensi tuning model dan akurasi hasil prediksi.

11. Statsmodels versi 0.14.5

Statsmodels digunakan untuk analisis statistik dan ekonometrika, termasuk uji

stasioneritas dan pemodelan deret waktu sebagai bagian dari proses validasi sebelum pelatihan deep learning dilakukan.

12. readxl versi 1.4.5

readxl digunakan untuk membaca dan mengimpor data Excel ke dalam lingkungan R secara efisien untuk keperluan analisis statistik dan pemodelan.

13. dplyr versi 1.1.4

dplyr digunakan untuk manipulasi data, termasuk proses seleksi, filter, agregasi, dan transformasi dataset secara cepat dan terstruktur

14. plm versi 2.6-6

plm digunakan sebagai package R untuk analisis ekonometrika berbasis data panel, termasuk model efek tetap dan efek acak dalam pengujian empiris.

15. panelvar versi 0.5.6

panelvar digunakan untuk pemodelan *Panel Vector Autoregressive* (PVAR), termasuk estimasi GMM dan analisis hubungan dinamis antar variabel dalam panel.

16. zoo versi 1.8-14

zoo digunakan untuk pengelolaan dan manipulasi data deret waktu tidak reguler sebagai bagian dari preprocessing sebelum pemodelan dan visualisasi.

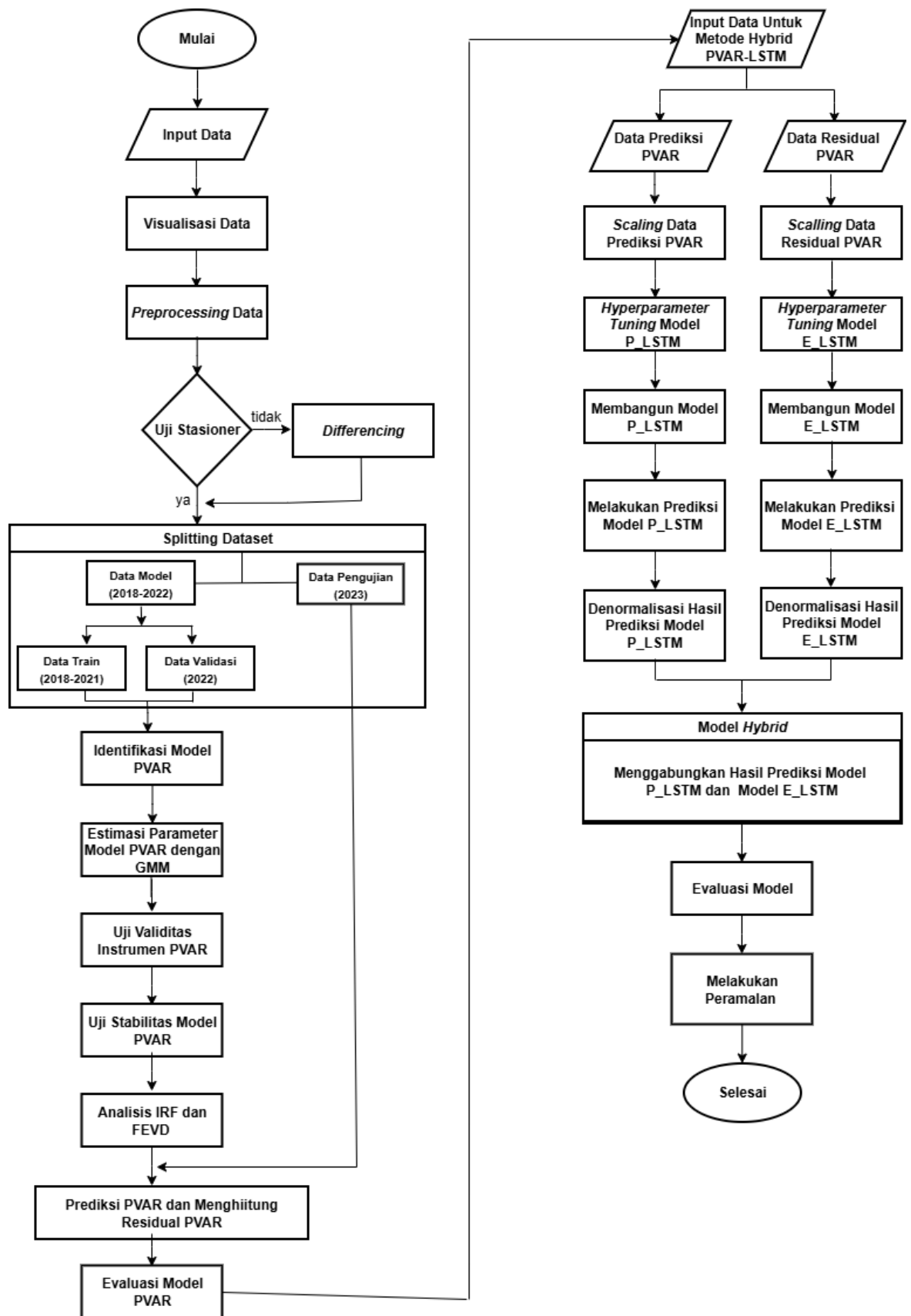
3.4 Metode Penelitian

Langkah-langkah yang dilakukan untuk melakukan peramalan dengan metode *hybrid* PVAR-LSTM sebagai berikut:

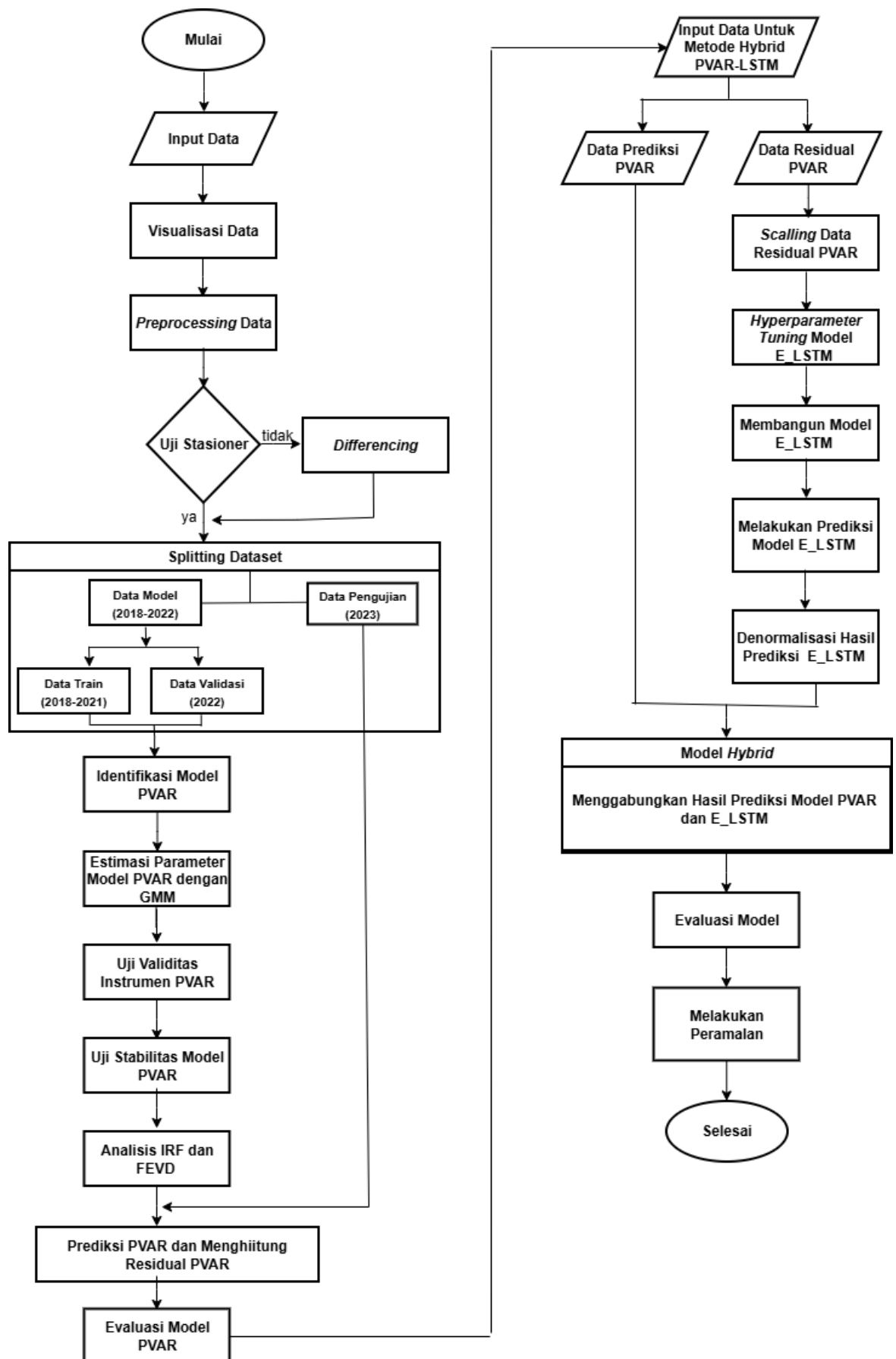
1. Melakukan kajian pustaka mengenai konsep dan penerapan model *hybrid* PVAR–LSTM dengan merujuk pada buku, artikel ilmiah, serta sumber terpercaya lainnya yang relevan dengan topik penelitian.
2. Melakukan visualisasi pada data asli untuk mengetahui tren dan pola pada data.
3. Mengumpulkan data yang diperlukan untuk proses peramalan.
4. Melakukan proses prapengolahan data dengan memeriksa keberadaan *missing values* untuk mengidentifikasi data yang tidak tersedia, serta melakukan interpolasi apabila ditemukan nilai data yang bernilai nol.

5. Melakukan proses pembagian data (*splitting data*) menjadi data model dan data uji menggunakan dua skema proporsi, yaitu tahun 2018-2022 data model dan tahun 2023 sebagai data *testing*.
6. Melakukan pengujian stasioneritas pada data pelatihan menggunakan uji LLC dan IPS. Apabila hasil pengujian menunjukkan bahwa data belum stasioner, maka dilakukan proses *differencing* hingga data memenuhi asumsi stasioner.
7. Menentukan nilai orde p pada model PVAR menggunakan metode MMSC, dengan memilih orde yang memiliki nilai terendah pada kriteria $MMSC_{BIC}$.
8. Menyusun variabel instrumen (Q_i) yang diperlukan dalam proses estimasi model.
9. Melakukan estimasi parameter model PVAR menggunakan metode *Generalized Method of Moments* (GMM).
10. Menguji tingkat signifikansi parameter yang dihasilkan dari estimasi model PVAR.
11. Melakukan validasi terhadap variabel instrumen (Q_i) melalui uji Hansen untuk memastikan keandalan instrumen yang digunakan.
12. Menguji stabilitas model PVAR guna memastikan model memenuhi asumsi kestabilan dinamis.
13. Melakukan analisis *Impulse Response Function* (IRF) untuk menilai respons dinamis antar variabel dalam periode waktu tertentu.
14. Melaksanakan analisis *Forecast Error Variance Decomposition* (FEVD) untuk mengetahui kontribusi relatif masing-masing variabel terhadap variasi kesalahan peramalan.
15. Berdasarkan model PVAR, kemudian dilakukan prediksi dan peramalan data.
16. Menghitung nilai residual dari model PVAR.
17. Melakukan evaluasi performa model PVAR untuk menilai tingkat akurasi dan kelayakan model.
18. Mengintegrasikan data hasil prediksi dan residual dari model PVAR sebagai input untuk tahap pemodelan lanjutan menggunakan metode LSTM.

19. Melakukan proses normalisasi terhadap data prediksi dan residual PVAR sebelum memasuki tahap pelatihan model.
20. Menentukan parameter terbaik untuk kedua model LSTM melalui proses *hyperparameter tuning* menggunakan metode OPTUNA.
21. Mengembangkan dua model utama berbasis LSTM sebagai komponen pembentuk model *hybrid*. Model pertama menggunakan data prediksi hasil PVAR sebagai input (P_LSTM), sedangkan model kedua menggunakan data residual PVAR sebagai input (E_LSTM).
22. Membangun dua model LSTM menggunakan bahasa pemrograman Python sebagai perangkat pemrosesan data.
23. Melakukan proses prediksi menggunakan model P_LSTM dengan memanfaatkan data hasil prediksi dari model PVAR sebagai input.
24. Melaksanakan prediksi menggunakan model E_LSTM yang dibangun berdasarkan data residual model PVAR.
25. Menghasilkan nilai prediksi model *hybrid* PVAR-LSTM melalui dua skema penggabungan, yaitu: untuk skema 1, (1) menjumlahkan hasil prediksi dari model P_LSTM dan E_LSTM, serta (2) menjumlahkan hasil prediksi model PVAR dengan hasil prediksi model E_LSTM yang berbasis residual. Untuk skema 2, (1) menjumlahkan hasil prediksi PVAR, serta (2) menjumlahkan hasil prediksi model PVAR dengan hasil prediksi model E_LSTM yang berbasis residual.
26. Evaluasi Model kedua skema.
27. Melakukan peramalan menggunakan dua skema terbaik.



Gambar 4. Alur Penelitian Skema 1



Gambar 5. Alur Penelitian Skema 2

BAB IV

HASIL DAN PEMBAHASAN

4.1 *Input Data*

Hal pertama yang dilakukan pada penelitian ini yaitu *input* data kedalam *software* R studio. Data yang digunakan merupakan data bulanan luas panen dan produksi padi di 34 provinsi di Indonesia pada tahun Januari 2018 - Desember 2023. Jumlah data yang digunakan sebanyak 2448 dimana masing-masing provinsi sebanyak 72 data. Gambar 6 adalah data yang di input pada *software* R studio.

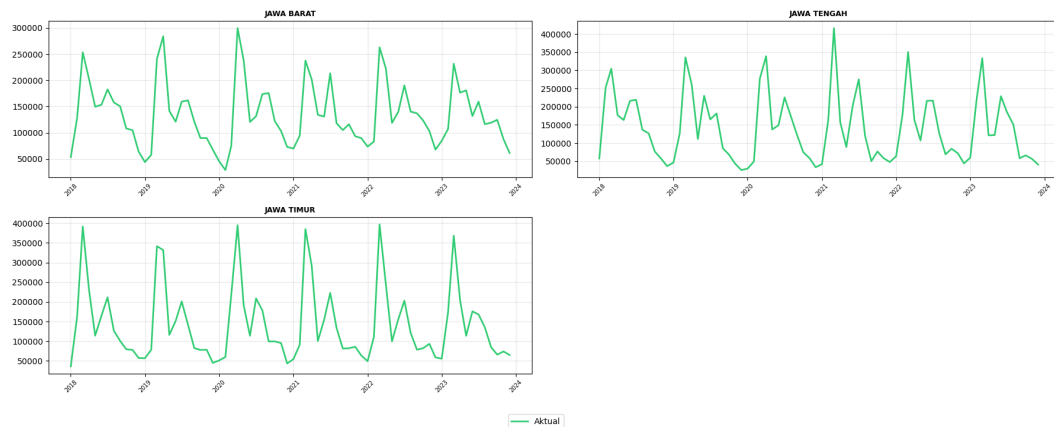
```
# A tibble: 2,448 × 4
  waktu          Provinsi Luas_Panen Produksi_Padi
<dtm>          <chr>      <dbl>      <dbl>
1 2018-01-01 00:00:00 ACEH        14542.        76118.
2 2018-02-01 00:00:00 ACEH        27076.       142877.
3 2018-03-01 00:00:00 ACEH        66404.       375842.
4 2018-04-01 00:00:00 ACEH        58507.       336743.
5 2018-05-01 00:00:00 ACEH        16868.        96372.
6 2018-06-01 00:00:00 ACEH         6574.        37912.
7 2018-07-01 00:00:00 ACEH        16175.        93822.
8 2018-08-01 00:00:00 ACEH        24948.       142896.
9 2018-09-01 00:00:00 ACEH        34889.       192853.
10 2018-10-01 00:00:00 ACEH        25949.       153130.
# i 2,438 more rows
```

Gambar 6. Data

Setelah *input* data langkah selanjutnya adalah visualisasi data dalam bentuk plot di 34 provinsi. Visualisasi data dilakukan untuk melihat pola dan fluktuasi data. Berdasarkan data pada Gambar 6 maka visualisasinya dapat dilihat sebagai berikut:

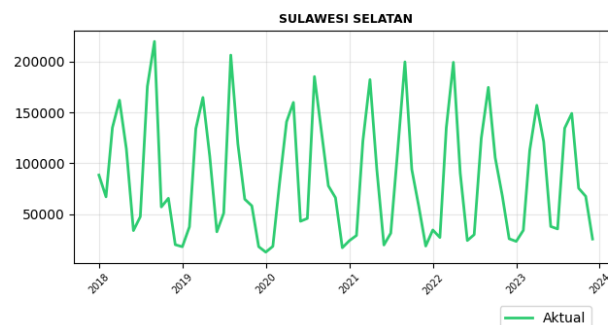
1. Luas Panen

Berdasarkan hasil visualisasi data luas panen pada 34 provinsi, terlihat adanya perbedaan pola dan tingkat fluktuasi antar provinsi. Hal ini menunjukkan bahwa data memiliki karakteristik yang beragam. Adapun pembahasan karakteristik data tersebut disajikan pada Gambar 7, 8 dan 9 sebagai berikut.



Gambar 7. Visualisasi Aktual Luas Panen Kelompok Tinggi

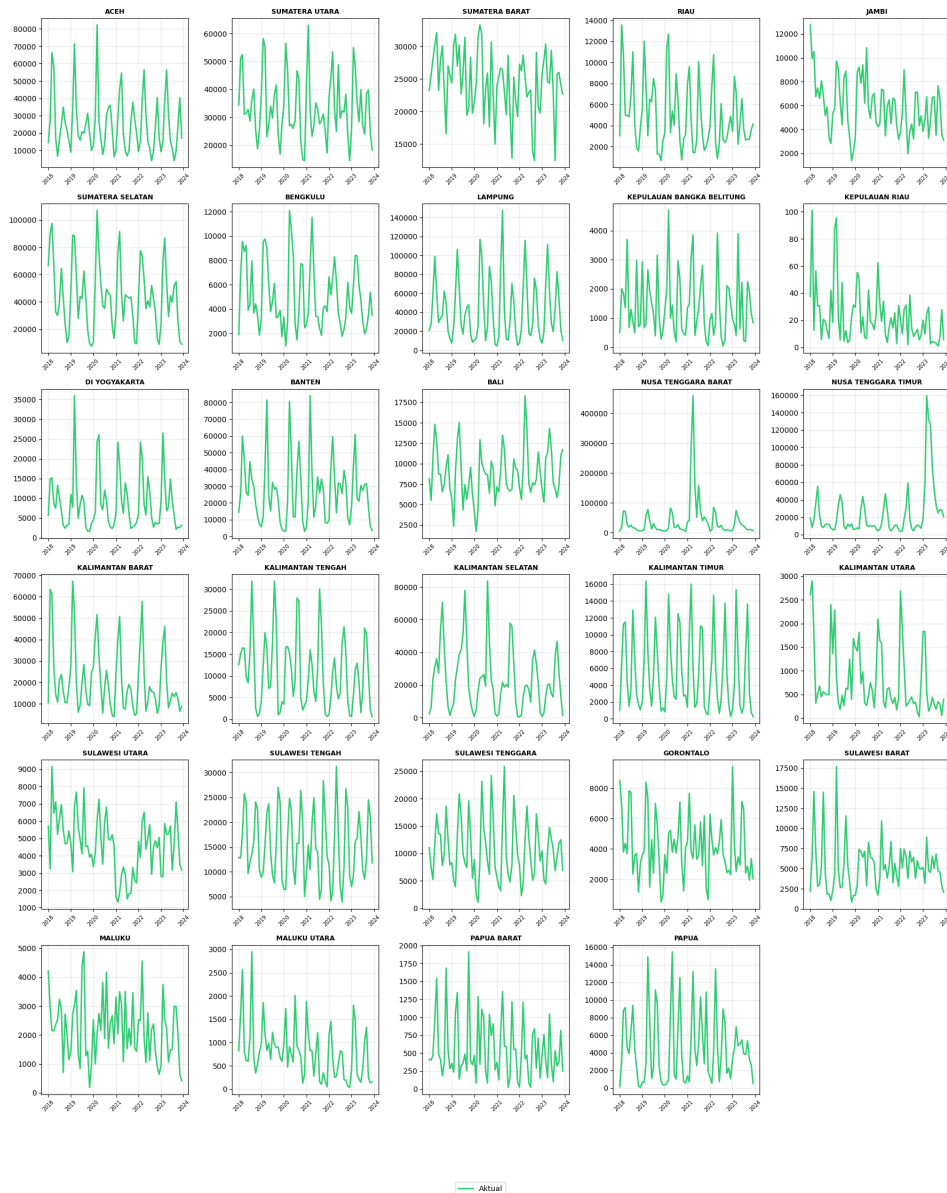
Berdasarkan Gambar 7, pada kelompok luas panen tinggi, seperti Jawa Timur, Jawa Tengah, dan Jawa Barat, terlihat bahwa data memiliki pola dengan bentuk gelombang yang berulang setiap tahun. Kenaikan luas panen terjadi dengan nilai yang tinggi pada awal periode dan diikuti penurunan yang cukup tajam, sehingga membentuk pola naik-turun yang jelas. Selain itu, jarak antara titik terendah dan tertinggi cukup besar, yang menunjukkan bahwa variasi luas panen pada kelompok luas panen ini sangat tinggi. Pola yang terbentuk juga relatif konsisten antar tahun.



Gambar 8. Visualisasi Aktual Luas Panen Kelompok Sedang

Berdasarkan Gambar 8, pada kelompok luas panen sedang yaitu Sulawesi Selatan, terlihat bahwa pola data tetap menunjukkan kenaikan dan penurunan yang berulang secara cukup jelas dari tahun ke tahun. Pola ini membentuk gelombang yang relatif teratur. Jika dibandingkan dengan kelompok tinggi, nilai puncak pada kelompok ini tidak terlalu jauh dari nilai lainnya, sehingga variasinya masih berada pada tingkat menengah. Grafik juga menunjukkan bahwa kenaikan menuju puncak sering terjadi dengan cepat, diikuti penurunan yang cukup tajam, sehingga membentuk

pola naik-turun, pola ini berulang hampir di setiap tahun.



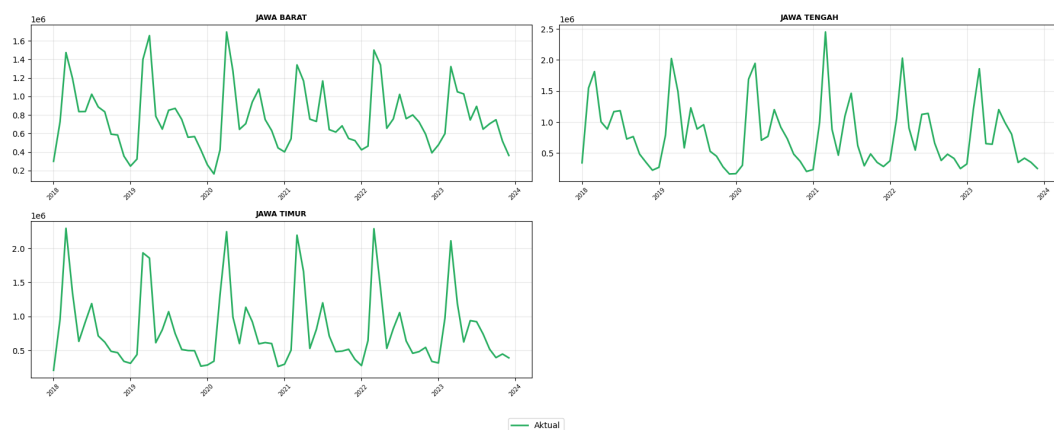
Gambar 9. Visualisasi Aktual Luas Panen Kelompok Rendah

Berdasarkan Gambar 9, pada kelompok luas panen rendah, terlihat bahwa pola data lebih bervariasi antar provinsi dibandingkan kelompok lainnya. Setiap provinsi menunjukkan karakteristik yang berbeda, sehingga tidak terdapat pola umum yang seragam seperti pada kelompok tinggi maupun sedang. Secara umum, nilai luas panen berada pada kisaran yang rendah dengan selisih antara puncak dan lembah yang relatif kecil, sehingga grafik cenderung tampak lebih datar. Beberapa provinsi seperti Jambi, Kalimantan Utara, Kepulauan Riau, Maluku, Maluku Utara,

dan Papua Barat menunjukkan pola fluktuasi kecil yang tidak terlalu tajam dan cenderung berubah-ubah antar periode tanpa pola yang jelas. Di sisi lain, Nusa Tenggara Timur (NTT) terlihat berbeda karena memiliki lonjakan yang sangat tinggi pada satu periode tertentu dibandingkan periode lainnya, meskipun pada sebagian besar waktu nilainya tetap rendah.

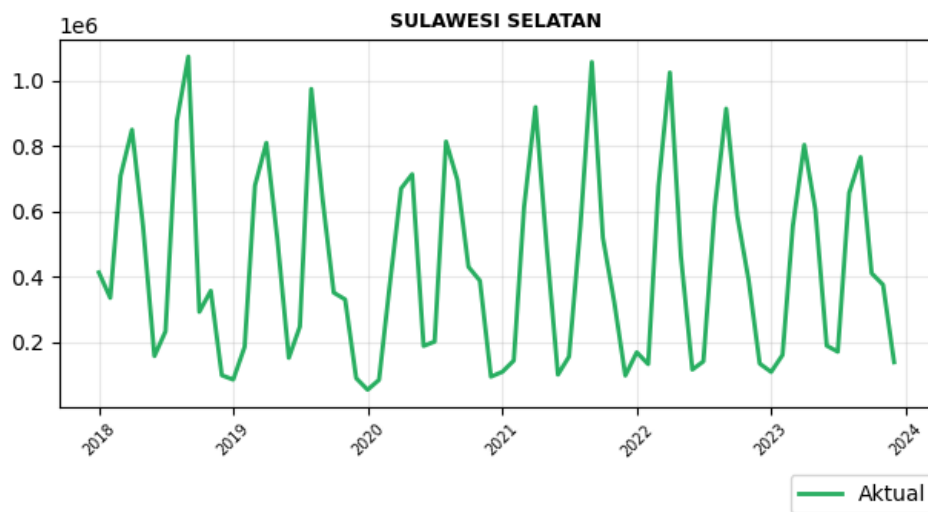
2. Produksi Padi

Berdasarkan visualisasi data produksi padi pada 34 provinsi, terlihat perbedaan pola dan tingkat fluktuasi antar provinsi yang mencerminkan keragaman karakteristik data. Pembahasan tersebut disajikan pada Gambar 10, 11, dan 12 sebagai berikut.



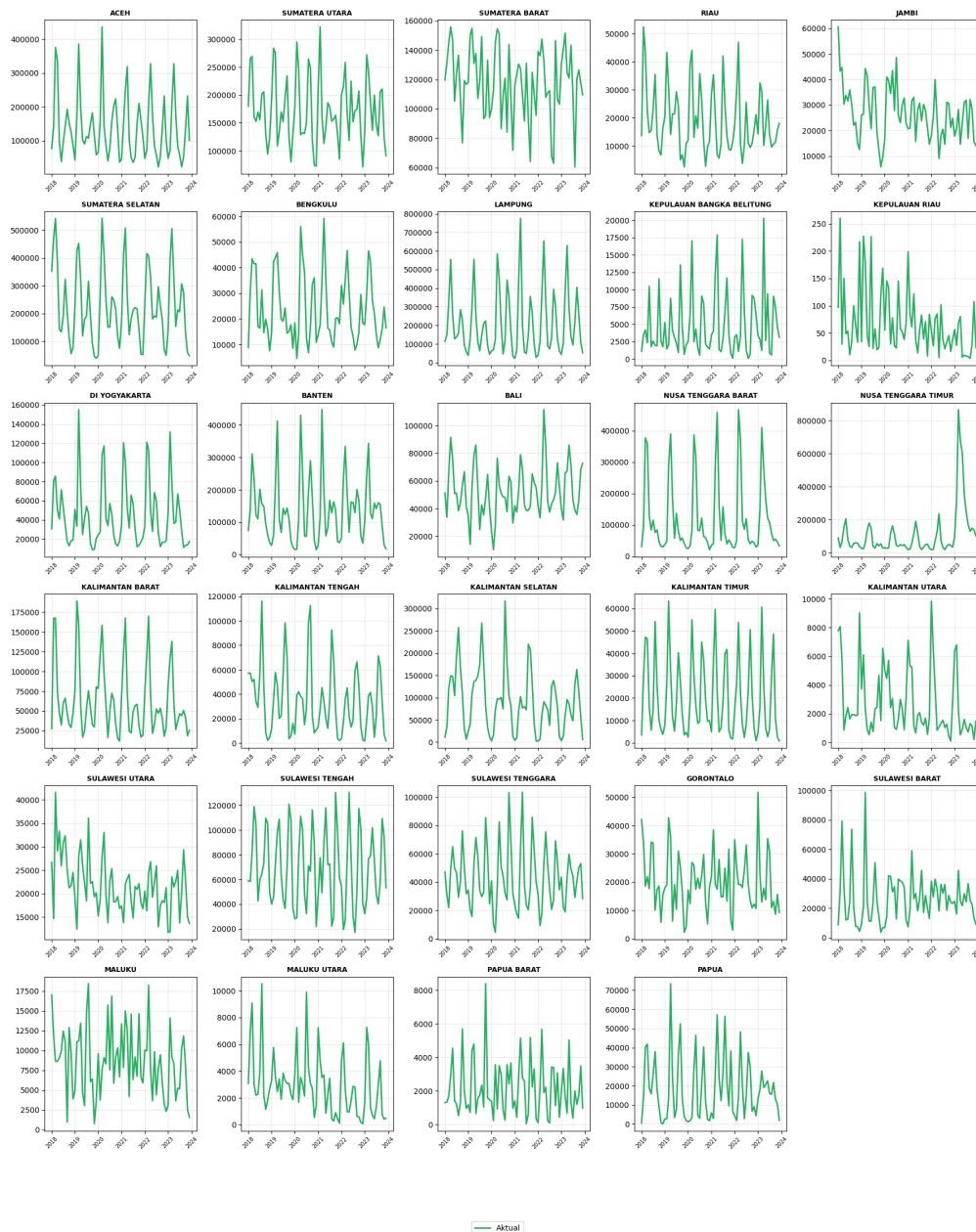
Gambar 10. Visualisasi Aktual Produksi Padi Kelompok Tinggi

Berdasarkan Gambar 10, pada kelompok produksi tinggi, seperti Jawa Timur, Jawa Tengah, dan Jawa Barat, terlihat bahwa data memiliki pola dengan bentuk gelombang yang berulang setiap tahun. Selain itu, selisih antara nilai tertinggi dan terendah cukup besar, sehingga fluktuasi yang terjadi terlihat sangat jelas. Hal ini menunjukkan bahwa variasi produksi padi pada kelompok ini sangat tinggi. Pola yang terbentuk juga relatif konsisten antar tahun.



Gambar 11. Visualisasi Aktual Produksi Padi Kelompok Sedang

Berdasarkan Gambar 11, pada kelompok produksi sedang, yaitu Sulawesi Selatan, terlihat pola kenaikan dan penurunan yang berulang secara cukup jelas setiap tahun. Pola ini membentuk gelombang yang teratur, di mana nilai produksi meningkat menuju puncak lalu menurun kembali dalam periode berikutnya. Jika dibandingkan dengan kelompok tinggi, selisih antara nilai tertinggi dan terendah tidak terlalu besar sehingga fluktuasinya berada pada tingkat menengah. Pergerakan data juga terlihat cukup konsisten dari waktu ke waktu dengan pola yang masih stabil.



Gambar 12. Visualisasi Aktual Produksi Padi Kelompok Rendah

Berdasarkan Gambar 12, pada kelompok produksi rendah, terlihat bahwa pola data lebih bervariasi antar provinsi. Setiap provinsi menunjukkan pola yang berbeda sehingga tidak terdapat bentuk gelombang yang seragam. Secara umum, nilai produksi berada pada tingkat yang rendah dengan selisih antara nilai tertinggi dan terendah yang relatif kecil. Beberapa provinsi seperti Jambi, Kalimantan Utara, Kepulauan Riau, Maluku, Maluku Utara, dan Papua Barat menunjukkan fluktuasi kecil yang berubah-ubah antar waktu tanpa pola yang konsisten. Di sisi lain,

Nusa Tenggara Timur (NTT) terlihat cukup berbeda karena memiliki lonjakan produksi yang sangat tinggi pada satu periode tertentu dibandingkan periode lainnya, meskipun sebagian besar nilainya tetap rendah.

4.2 Model *Panel Vector Autoregressive* (PVAR)

Setelah data di *input* hal pertama yang dilakukan menggunakan metode PVAR-LSTM yaitu di analisis menggunakan PVAR terlebih dahulu. Pada model PVAR nantinya akan didapatkan nilai prediksi dan residual yang akan menjadi input pada model LSTM. Berikut langkah yang dilakukan pada bagian PVAR.

4.2.1 *Preprocessing Data*

Pertama dilakukan *preprocessing* data, tahap ini bertujuan untuk memastikan data yang digunakan dalam pemodelan berada dalam kondisi yang baik sehingga model dapat bekerja secara optimal. Pada tahap awal dilakukan pengecekan *missing value* dan data yang memiliki nilai 0.

Pengecekan *missing value* dilakukan untuk mengetahui apakah terdapat data yang hilang pada setiap variabel. Hasil pengecekan *missing value* disajikan pada Tabel 3, yang menunjukkan bahwa variabel luas panen dan produksi padi tidak memiliki *missing value*.

Tabel 3. *Missing Value*

Variabel	<i>Missing Value</i>
Luas Panen	0
Produksi Padi	0

Selanjutnya dilakukan pengecekan terhadap nilai 0 pada data. Apabila ditemukan nilai 0 maka dilakukan interpolasi data. Hasil pengecekan nilai 0 disajikan pada Tabel 4, yang menunjukkan bahwa nilai 0 terdapat pada Provinsi DKI Jakarta dan Kepulauan Riau untuk variabel luas panen dan produksi padi. Oleh karena itu, pada provinsi tersebut akan dilakukan interpolasi data linear, dengan cara mengambil garis lurus di antara dua nilai tetangganya.

Tabel 4. Data Nol

Provinsi	Luas Panen	Produksi Padi
DKI Jakarta	4	4
Kepulauan Riau	1	1

4.2.2 Uji Stasioneritas

Setelah dilakukan pengecekan dan interpolasi terhadap data yang memiliki nilai nol, langkah selanjutnya adalah menguji stasioneritas data panel. Pengujian stasioneritas bertujuan untuk memastikan apakah data telah memenuhi asumsi stasioneritas sebelum dilakukan analisis lebih lanjut. Dalam penelitian ini, uji stasioneritas dilakukan menggunakan dua metode, yaitu uji Levin-Lin-Chu (LLC) dan uji Im-Pesaran-Shin (IPS).

H_0 : data panel tidak stasioner (mengandung akar unit).

H_1 : data panel stasioner (tidak mengandung akar unit).

Apabila asumsi tidak terpenuhi yang artinya data tidak stasioner maka perlu dilakukan *differencing* agar data dapat memenuhi asumsi stasioneritas.

1. *Levin Lin Chu*

Uji Levin Lin Chu (LLC) digunakan untuk mengetahui apakah data panel mengandung akar unit atau tidak. Uji ini mengasumsikan bahwa seluruh unit dalam data panel, seperti setiap provinsi, memiliki pola yang sama. Hasil pengujian menunjukkan bahwa nilai *p-value* untuk variabel yang dianalisis berada di bawah tingkat signifikansi 5%. Hal ini berarti hipotesis nol ditolak. Dengan demikian, berdasarkan uji LLC, variabel luas panen dan produksi padi dinyatakan bersifat stasioner pada level. Hasil uji LLC yang diperoleh dapat dilihat pada Tabel 5 berikut:

Tabel 5. Uji Levin–Lin–Chu (LLC)

Variabel	Statistik Uji	P-Value	Hasil
Luas Panen	-27.377	2.2×10^{-16}	Stasioner
Produksi Padi	-24.84	2.2×10^{-16}	Stasioner

2. Im–Pesaran–Shin (IPS).

Uji Im-Pesaran-Shin (IPS) digunakan untuk menguji apakah data panel bersifat stasioner dengan mempertimbangkan bahwa setiap unit, seperti masing-masing provinsi, bisa memiliki karakteristik yang berbeda. Berbeda dengan uji LLC yang mengasumsikan kesamaan pola antar unit, uji IPS memberikan fleksibilitas karena tidak mengharuskan seluruh unit memiliki pola yang sama. Hasil pengujian menunjukkan bahwa nilai *p-value* untuk variabel luas panen dan produksi padi berada di bawah tingkat signifikansi 5%. Hal ini berarti hipotesis nol ditolak. Dengan demikian, kedua variabel dapat dinyatakan stasioner berdasarkan uji IPS. Hasil uji IPS yang diperoleh disajikan dalam Tabel 6 adalah sebagai berikut:

Tabel 6. Uji IPS

Variabel	Statistik Uji	P-Value	Hasil
Luas Panen	-34.367	2.2×10^{-16}	Stasioner
Produksi Padi	-34.082	2.2×10^{-16}	Stasioner

Secara keseluruhan, berdasarkan hasil uji LLC dan IPS, dapat disimpulkan bahwa data luas panen dan produksi padi telah memenuhi asumsi stasioneritas pada level. Oleh karena itu, data dapat langsung digunakan dalam pemodelan tanpa perlu dilakukan proses *differencing*.

4.2.3 Pembagian Data

Langkah selanjutnya yang dilakukan adalah pembagian data menjadi data model yaitu data *training* dan *validation* serta data *testing*. Data *training* dan *validation* digunakan untuk melatih model serta data *testing* digunakan untuk menguji seberapa baik model bekerja. Skema pembagian datanya yaitu data pada tahun 2018-2022 sebagai data model dan data tahun 2023 sebagai data *testing*. Pembagian data dapat direpresentasikan pada Tabel 7 berikut:

Tabel 7. Pembagian Data Penelitian

Variabel	Pembagian Data		
	Data Model		Data Testing
	Training	Validation	
Luas Panen	1632	408	408
Produksi Padi	1632	408	408

Pembagian data dilakukan berdasarkan urutan waktu. Pendekatan ini sejalan dengan penelitian Chen dan Moraga (2025) yang menerapkan pembagian data berdasarkan urutan waktu. Skema pembagian data dapat dilihat pada Tabel 7. Berdasarkan tabel tersebut, masing-masing variabel memiliki 2.040 pada data model yaitu 1632 sebagai data *training* dan 408 data *validation* dan 408 data pada tahap *testing*.

4.2.4 Penentuan Lag Optimal

Penentuan lag optimal merupakan tahap penting yang dilakukan sebelum mengestimasi model *Panel Vector Autoregressive* (PVAR). Pemilihan panjang lag yang tepat diperlukan agar model mampu mengestimasi dengan optimal. Penentuan lag optimal dilakukan menggunakan *Model and Moment Selection Criteria* (MMSC) yang dikembangkan oleh Andrews dan Lu (2001) pada kerangka estimasi Generalized Method of Moments (GMM). Kriteria yang digunakan dalam penentuan lag optimal meliputi $MMSC_{AIC}$, $MMSC_{BIC}$, dan $MMSC_{HQIC}$.

Lag optimal ditentukan berdasarkan nilai kriteria MMSC yang paling minimum. Hasil pengujian lag optimal untuk model PVAR disajikan pada Tabel 8. Berdasarkan Tabel 8, lag yang dipilih digunakan dalam proses estimasi model PVAR pada tahap berikutnya.

Tabel 8. Pemilihan Optimal Lag

Lag	$MMSC_{AIC}$	$MMSC_{BIC}$	$MMSC_{HQIC}$
1	-5.0196	-38.5404	-18.5524
2	-8.0000	-30.2776	-17.0026

Berdasarkan literatur, Andrews dan Lu (2001) menjelaskan bahwa kriteria

$MMSC_{AIC}$ tidak bersifat konsisten, karena memiliki probabilitas positif dalam memilih spesifikasi model yang tidak tepat, bahkan ketika ukuran sampel semakin besar. Oleh karena itu, penggunaan $MMSC_{AIC}$ kurang disarankan sebagai dasar utama dalam pemilihan lag pada model panel dinamis. Atas dasar tersebut, pemilihan lag optimal dipilih berdasarkan kriteria $MMSC_{BIC}$ dan $MMSC_{HQIC}$ yang terbukti bersifat konsisten. Berdasarkan Tabel 8, kedua kriteria tersebut sama-sama menunjukkan nilai minimum pada lag 1, sehingga lag tersebut dipilih sebagai lag optimal dan digunakan dalam estimasi model PVAR.

4.2.5 Pembentukan Model PVAR

Pada langkah sebelumnya dalam menentukan lag optimal diperoleh bahwa model PVAR(1) merupakan model yang paling optimal dan sesuai untuk digunakan. Berdasarkan orde model yang telah ditentukan pada tahap identifikasi orde sebelumnya maka estimasi parameter serta pengujian signifikansi pada model PVAR (p) dapat dilakukan. Proses estimasi diawali dengan penghilangan efek individual masing-masing wilayah menggunakan metode *forward orthogonal deviation* (FOD). Di mana hasil estimasi parameter untuk model PVAR (1) dapat dilihat pada Gambar 13 sebagai berikut.

	Produksi_Padi	Luas_Panen
lag1_Produksi_Padi	0.7968 *** (0.0605)	0.0226 * (0.0107)
lag1_Luas_Panen	-0.0369 (0.2624)	0.6820 *** (0.0412)

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Gambar 13. Estimasi Parameter Model PVAR

Berdasarkan hasil estimasi parameter model PVAR(1), persamaan model dalam bentuk matriks dapat dituliskan sebagai berikut:

$$\begin{bmatrix} y_{1,it} \\ y_{2,it} \end{bmatrix} = \begin{bmatrix} 0.7968 & -0.0369 \\ 0.0226 & 0.6820 \end{bmatrix} \begin{bmatrix} y_{1,i,t-1} \\ y_{2,i,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it}^{(1)} \\ \varepsilon_{it}^{(2)} \end{bmatrix},$$

Persamaan model PVAR(1) tersebut dapat dituliskan dalam bentuk persamaan 27

dan 28 sebagai berikut:

$$y_{1,it} = 0.7968 y_{1,i,t-1} - 0.0369 y_{2,i,t-1} + \varepsilon_{it}^{(1)}, \quad (27)$$

$$y_{2,it} = 0.0226 y_{1,i,t-1} + 0.6820 y_{2,i,t-1} + \varepsilon_{it}^{(2)}. \quad (28)$$

Berdasarkan hasil estimasi parameter model PVAR(1), diperoleh koefisien hubungan antar variabel luas panen dan produksi padi. Hasil estimasi menunjukkan bahwa masing-masing variabel dipengaruhi oleh nilainya sendiri pada periode sebelumnya serta oleh variabel lainnya pada periode sebelumnya.

Pada persamaan pertama koefisien sebesar 0,7968 menunjukkan bahwa produksi padi periode saat ini dipengaruhi sebesar 0,7968 oleh produksi padi pada periode sebelumnya. Sementara itu, koefisien sebesar $-0,0369$ menunjukkan bahwa luas panen pada periode sebelumnya memiliki pengaruh negatif terhadap produksi saat ini, meskipun besarnya relatif kecil dibandingkan pengaruh produksinya sendiri. Pada persamaan kedua, koefisien sebesar 0,6820 menunjukkan bahwa luas panen periode saat ini juga dipengaruhi sebesar 0,6820 oleh luas panen pada periode sebelumnya. Selain itu, koefisien sebesar 0,0226 menunjukkan bahwa produksi pada periode sebelumnya memberikan pengaruh positif terhadap luas panen saat ini, tetapi dengan pengaruh yang relatif kecil. Secara keseluruhan, hasil ini mengindikasikan bahwa kedua variabel memiliki hubungan dua arah (interdependensi), di mana masing-masing variabel tidak hanya dipengaruhi oleh dirinya sendiri, tetapi juga oleh variabel lainnya. Meskipun demikian, pengaruh terbesar tetap berasal dari nilai variabel itu sendiri pada periode sebelumnya.

4.2.6 Uji Validitas Instrumen Model PVAR

Setelah model PVAR dibentuk, langkah selanjutnya adalah melakukan uji validitas instrumen untuk mengetahui apakah instrumen yang digunakan dalam proses estimasi sudah memenuhi asumsi yang diperlukan. Uji validitas instrumen pada model PVAR dilakukan menggunakan Hansen test untuk mengetahui apakah variabel instrumen yang digunakan bersifat valid, yaitu tidak berkorelasi dengan *error term*. Uji ini memiliki hipotesis nol (H_0) yang menyatakan bahwa instrumen yang digunakan dalam model adalah valid. Hasil uji Hasen test dapat ditampilkan pada Tabel 9.

Tabel 9. Uji Validitas Instrumen (Hansen Test)

Statistik Uji J_n	<i>P-Value</i>
6.98	0.137

Berdasarkan Tabel 9, diperoleh nilai *p-value* sebesar 0,137 dimana nilai *p-value* tersebut lebih besar dari tingkat signifikansi 0,05. Hal ini, dapat disimpulkan bahwa hipotesis nol tidak ditolak, yang berarti tidak terdapat cukup bukti untuk menyatakan bahwa instrumen yang digunakan bersifat tidak valid maka instrumen yang digunakan adalah valid.

4.2.7 Uji Stabilitas Model PVAR

Setelah instrumen dinyatakan valid, kemudian berikutnya adalah melakukan uji stabilitas model PVAR. Uji stabilitas ini bertujuan untuk memastikan bahwa model yang diestimasi telah stabil sehingga nantinya dapat dilanjutkan ke tahap berikutnya. Stabilitas merupakan aspek penting dalam pemodelan PVAR, karena ketidakstabilan dapat menyebabkan hasil estimasi yang tidak realistis dan membuat hasil analisis menjadi kurang dapat dipercaya.

Pengujian stabilitas pada model PVAR dilakukan menggunakan uji stabilitas nilai eigen. Model PVAR dinyatakan stabil apabila seluruh nilai eigen memiliki nilai mutlak kurang dari satu. Sebaliknya, jika terdapat nilai eigen dengan nilai mutlak lebih besar dari satu, maka model tersebut dianggap tidak stabil. Tabel 10 menyajikan hasil uji stabilitas model PVAR berdasarkan uji nilai eigen.

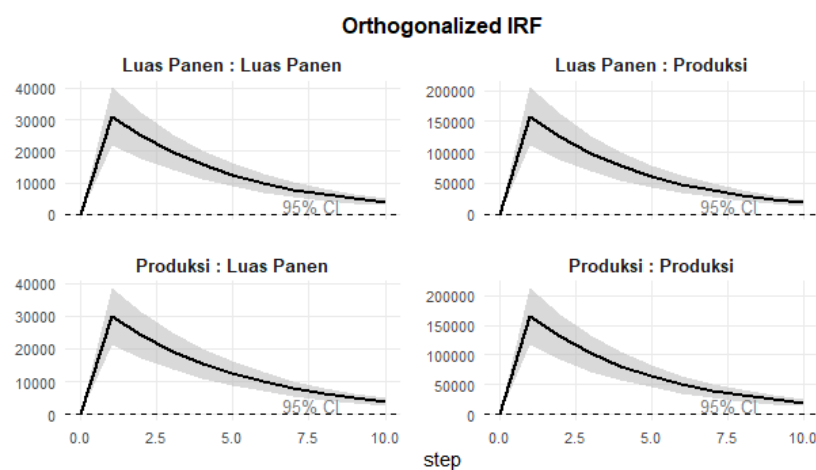
Tabel 10. Uji Stabilitas Model PVAR

Eigenvalue	Modulus
0.7890	0.7890
0.6898	0.6898

Berdasarkan hasil uji stabilitas yang disajikan pada Tabel 10, seluruh nilai modulus eigen memiliki nilai kurang dari satu. Hal ini menunjukkan bahwa model PVAR yang diestimasi telah memenuhi asumsi stabilitas. Hal tersebut berarti model PVAR stabil.

4.2.8 Impulse Response Function (IRF)

Impulse Response Function (IRF) digunakan untuk menganalisis bagaimana suatu variabel endogen merespon adanya guncangan (*shock*) pada variabel endogen lain dalam sistem model PVAR. Analisis ini penting untuk mengetahui pola hubungan antar variabel bagaimana pengaruh suatu *shock* terhadap variabel lainnya. Gambar 14 adalah gambar yang merepresentasikan hubungan IRF antara luas panen dan produksi padi.



Gambar 14. Impulse Response Function (IRF)

Berdasarkan Gambar 14, terlihat bahwa setiap *shock* atau guncangan pada suatu variabel menimbulkan respons positif, baik terhadap variabel itu sendiri maupun terhadap variabel lainnya. Respons tersebut cukup besar pada periode awal, kemudian secara bertahap menurun seiring berjalannya waktu. Penurunan respons ini menunjukkan bahwa dampak dari *shock* bersifat sementara dan tidak bertahan dalam jangka panjang. Artinya, setelah mengalami gangguan, model secara perlahan kembali menuju kondisi keseimbangan. Secara umum, hasil IRF ini menunjukkan adanya hubungan yang saling memengaruhi antara luas panen dan produksi padi. Perubahan pada salah satu variabel dapat memengaruhi variabel lainnya dalam jangka pendek, tetapi pengaruh tersebut akan semakin melemah dalam jangka panjang.

4.2.9 Forecast Error Variance Decomposition (FEVD)

Forecast-error variance decomposition (FEVD) digunakan untuk mengukur seberapa besar kontribusi guncangan suatu variabel, baik dari dirinya sendiri maupun variabel lain, dalam menjelaskan variansi kesalahan peramalan suatu variabel. Hasil perhitungan FEVD untuk produksi padi dan luas panen selama sepuluh periode ke depan disajikan pada Tabel 11 berikut:

Tabel 11. Forecast Error Variance Decomposition (FEVD)

Periode	Produksi Padi		Luas Panen	
	Produksi Padi	Luas Panen	Luas Panen	Produksi Padi
1	51.96096	48.03904	51.96096	48.03904
2	51.97244	48.02576	51.73156	48.26844
3	51.98419	48.01581	51.56043	48.43957
4	51.99148	48.00852	51.43504	48.56496
5	51.99673	48.00327	51.34477	48.65523
6	52.00044	47.99956	51.28089	48.71911
7	52.00301	47.99699	51.23640	48.76360
8	52.00477	47.99523	51.20585	48.79415
9	52.00596	47.99404	51.18516	48.81484
10	52.00675	47.99325	51.17130	48.82870

Berdasarkan Tabel 11, pada periode pertama variasi kesalahan peramalan produksi padi sebagian besar dijelaskan oleh *shock* produksi padi itu sendiri sebesar sekitar 51,96%, sedangkan sebesar 48,04% dijelaskan oleh *shock* luas panen. Pola ini relatif stabil hingga periode ke-10, kontribusi produksi padi terhadap dirinya sendiri meningkat sedikit menjadi sekitar 52,01%, sementara kontribusi luas panen menurun tipis menjadi sekitar 47,99%. Hal ini menunjukkan bahwa dalam menjelaskan kesalahan prediksi produksi padi, pengaruh *shock* dirinya sendiri masih sedikit lebih dominan dibandingkan pengaruh dari Luas Panen. Sebaliknya, untuk variabel Luas Panen, pada periode pertama sekitar 51,96% variasi kesalahan peramalan dijelaskan oleh *shock* luas panen itu sendiri dan 48,04% oleh *shock* produksi padi. Seiring bertambahnya periode, kontribusi *shock* produksi padi terhadap luas panen cenderung meningkat hingga mencapai sekitar 48,83% pada periode ke-10, sementara kontribusi luas panen terhadap dirinya sendiri sedikit menurun menjadi sekitar 51,17%. Hal

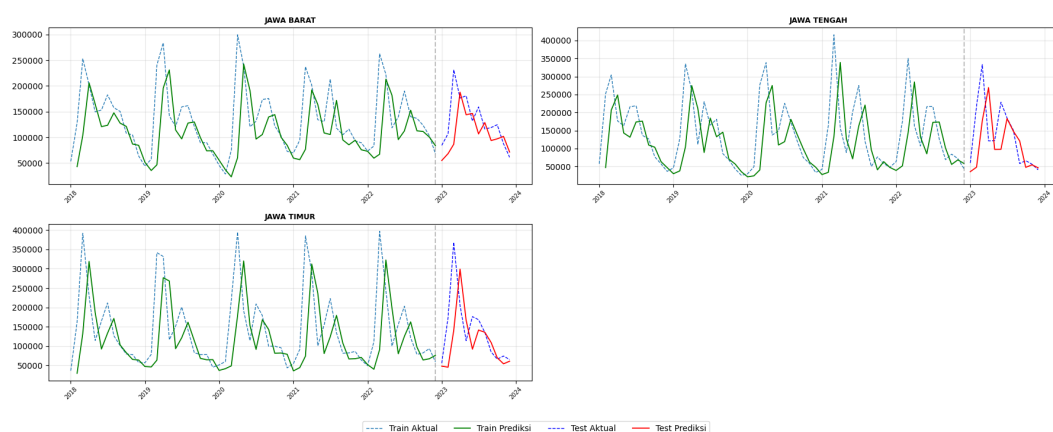
ini menunjukkan bahwa dalam jangka yang lebih panjang, pengaruh produksi padi terhadap luas panen menjadi semakin terlihat. Secara keseluruhan, hasil FEVD menunjukkan bahwa kedua variabel memiliki kontribusi yang hampir seimbang dalam menjelaskan variasi kesalahan peramalan masing-masing. Meskipun setiap variabel masih lebih banyak dipengaruhi oleh *shock*-nya sendiri, tapi variabel lainnya juga mempengaruhi.

4.2.10 Hasil Prediksi Model PVAR

Setelah model *Panel Vector Autoregressive* (PVAR) dibangun dan memenuhi semua asumsi, tahap selanjutnya adalah melakukan prediksi. Prediksi ini bertujuan untuk melihat kemampuan model PVAR dalam menangkap pola data antar waktu. Visualisasi hasil prediksi model PVAR untuk masing-masing variabel disajikan sebagai berikut.

1. Luas Panen

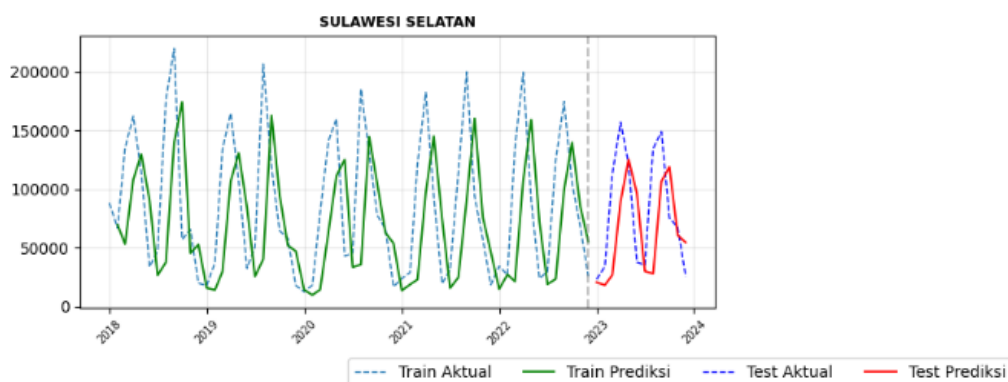
Visualisasi hasil prediksi model PVAR pada variabel luas panen menunjukkan perbandingan antara data aktual dan hasil prediksi pada masing-masing provinsi. Secara umum, hasil prediksi menggambarkan pola pergerakan data yang mengikuti kecenderungan data aktual. Adapun visualisasi tersebut disajikan sebagai pada Gambar 15, 16, dan 17 berikut.



Gambar 15. Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Tinggi

Berdasarkan Gambar 15, pada kelompok luas panen tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur. Hasil prediksi menunjukkan bahwa model

mampu mengikuti pola kenaikan dan penurunan yang berulang dengan cukup baik. Pola gelombang yang terbentuk dapat ditangkap dengan jelas, sehingga arah pergerakan data terlihat searah dengan data aktual. Terlihat bahwa pada beberapa periode terlihat bahwa hasil prediksi belum sepenuhnya mengikuti besarnya nilai puncak, di mana nilai yang dihasilkan cenderung lebih rendah dibandingkan nilai aktual. Secara umum, model sudah mampu merepresentasikan pola pergerakan data pada kelompok ini, tetapi masih terdapat selisih pada antara aktual dan prediksi.



Gambar 16. Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Sedang

Berdasarkan Gambar 16, pada kelompok luas panen sedang yaitu Sulawesi Selatan, hasil prediksi menunjukkan bahwa model mampu mengikuti pola kenaikan dan penurunan yang berulang dengan cukup baik. Pola pergerakan data dapat ditangkap, sehingga arah naik dan turun terlihat searah dengan data aktual. Meskipun demikian, pada beberapa titik masih terdapat selisih, terutama ketika nilai berada pada puncak, di mana hasil prediksi cenderung belum sepenuhnya mengikuti besarnya nilai tersebut. Secara umum, model sudah dapat merepresentasikan pola pergerakan data tetapi masih terdapat perbedaan antara aktual dan prediksi.



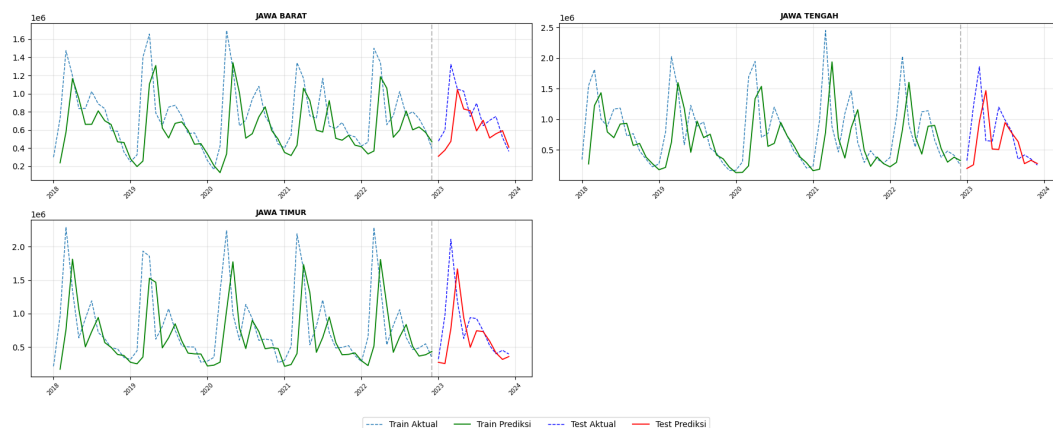
Gambar 17. Visualisasi Data Aktual dan Prediksi Kelompok Luas Panen Rendah

Berdasarkan Gambar 17, pada kelompok luas panen rendah, hasil prediksi menunjukkan bahwa model mampu mengikuti pola pergerakan data. Pola kenaikan dan penurunan masih dapat ditangkap, tetapi pada sebagian besar provinsi terlihat bahwa hasil prediksi tidak sepenuhnya mengikuti perubahan nilai pada data aktual.

Selain itu, pada beberapa periode, terutama ketika terjadi perubahan yang cukup tajam, hasil prediksi belum dapat mengikuti pergerakan tersebut dengan baik. Secara umum, model sudah dapat merepresentasikan pola pergerakan data, tetapi masih terdapat selisih antara aktual dan prediksi.

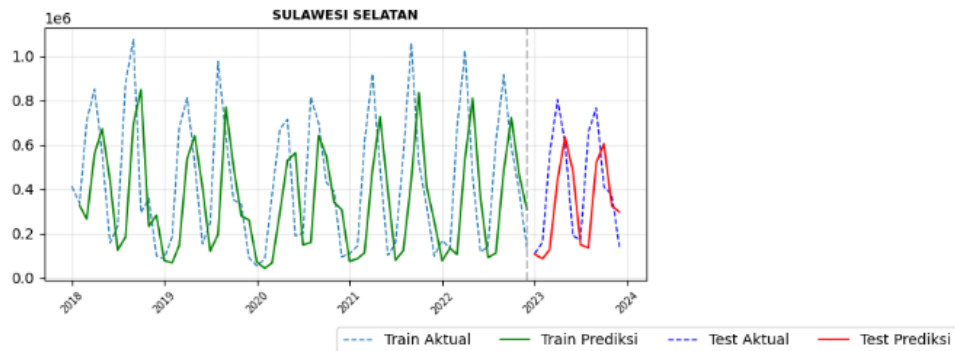
2. Produksi Padi

Visualisasi hasil prediksi model PVAR pada variabel produksi padi menunjukkan perbandingan antara data aktual dan hasil prediksi pada masing-masing provinsi. Secara umum, hasil prediksi menggambarkan pola pergerakan data yang mengikuti kecenderungan data aktual. Adapun visualisasi tersebut disajikan sebagai pada Gambar 18, 19, dan 20 berikut.



Gambar 18. Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Tinggi

Berdasarkan Gambar 18, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur. Hasil prediksi menunjukkan bahwa model mampu mengikuti pola pergerakan data dengan cukup baik. Pola kenaikan dan penurunan yang berulang dapat ditangkap, sehingga arah pergerakan data terlihat searah dengan data aktual. Pada beberapa periode terlihat bahwa hasil prediksi belum sepenuhnya mengikuti besarnya nilai puncak. Selain itu, ketika terjadi perubahan yang cukup tajam, masih terdapat selisih antara hasil prediksi dan data aktual. Secara umum, model sudah dapat merepresentasikan pola pergerakan data pada kelompok ini, tetapi masih terdapat selisih antara aktual dan prediksi.



Gambar 19. Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Sedang

Berdasarkan Gambar 19, pada kelompok produksi sedang, hasil prediksi menunjukkan bahwa model mampu mengikuti pola kenaikan dan penurunan yang berulang dengan cukup baik. Pola pergerakan data masih dapat ditangkap, sehingga arah naik dan turun terlihat searah dengan data aktual. Pada beberapa periode masih terlihat perbedaan antara hasil prediksi dan data aktual, terutama ketika nilai berada pada puncak. Selain itu, pada saat terjadi perubahan yang cukup cepat, hasil prediksi belum sepenuhnya mengikuti pergerakan tersebut. Secara umum, model sudah dapat merepresentasikan pola pergerakan data pada kelompok ini, tetapi masih terdapat selisih antara aktual dan prediksi.



Gambar 20. Visualisasi Data Aktual dan Prediksi Kelompok Produksi Padi Rendah

Berdasarkan Gambar 20, pada kelompok produksi rendah, hasil prediksi menunjukkan bahwa model mampu mengikuti pola pergerakan data. Pola kenaikan dan penurunan masih dapat ditangkap, tetapi pada sebagian besar provinsi terlihat bahwa hasil prediksi belum sepenuhnya mengikuti perubahan nilai pada data aktual. Selain itu, pada beberapa periode terlihat adanya selisih yang cukup jelas, terutama ketika terjadi perubahan yang cukup tajam. Beberapa nilai puncak juga tidak

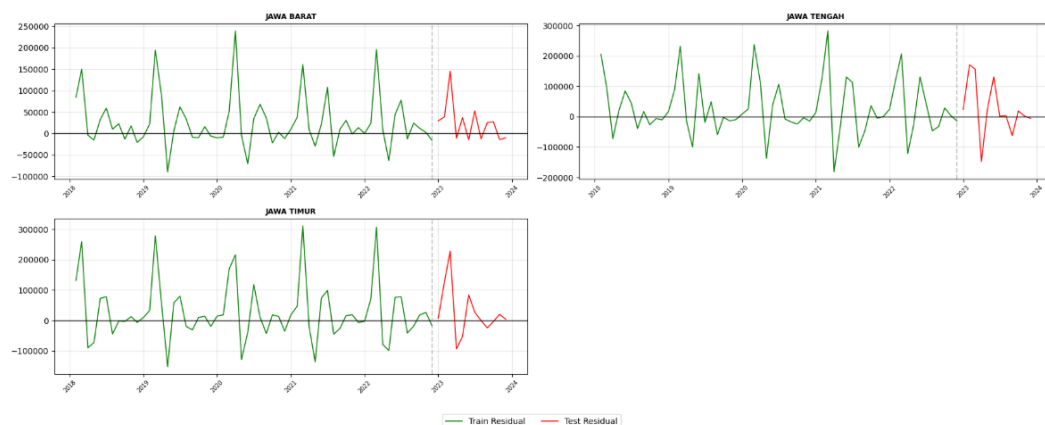
sepenuhnya dapat dicapai oleh hasil prediksi. Secara umum, model sudah dapat merepresentasikan pola pergerakan data, tetapi masih terdapat selisih antara aktual dan prediksi.

4.2.11 Menghitung Residual

Setelah model PVAR diestimasi dan digunakan untuk menghasilkan nilai prediksi, tahap berikutnya adalah menghitung residual model. Residual diperoleh dengan mengurangi nilai prediksi dari nilai aktual pada setiap periode pengamatan. Visualisasi residual untuk masing-masing provinsi disajikan pada Gambar 21, 22, dan 23 sebagai berikut:

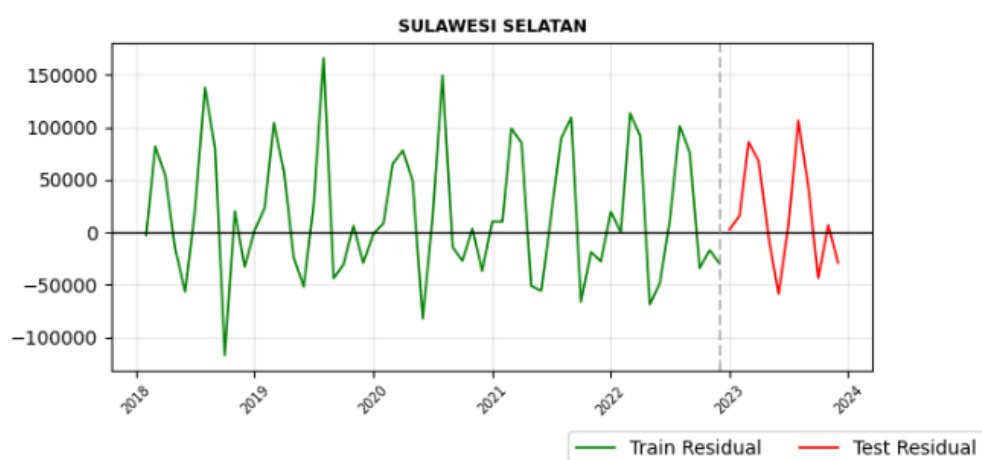
1. Luas Panen

Visualisasi residual pada variabel luas panen menunjukkan selisih antara nilai aktual dan hasil prediksi model PVAR pada masing-masing provinsi. Secara umum, nilai residual berada di sekitar nol yang menunjukkan bahwa hasil prediksi tidak terlalu menyimpang dari data aktual. Adapun visualisasi tersebut disajikan sebagai berikut:



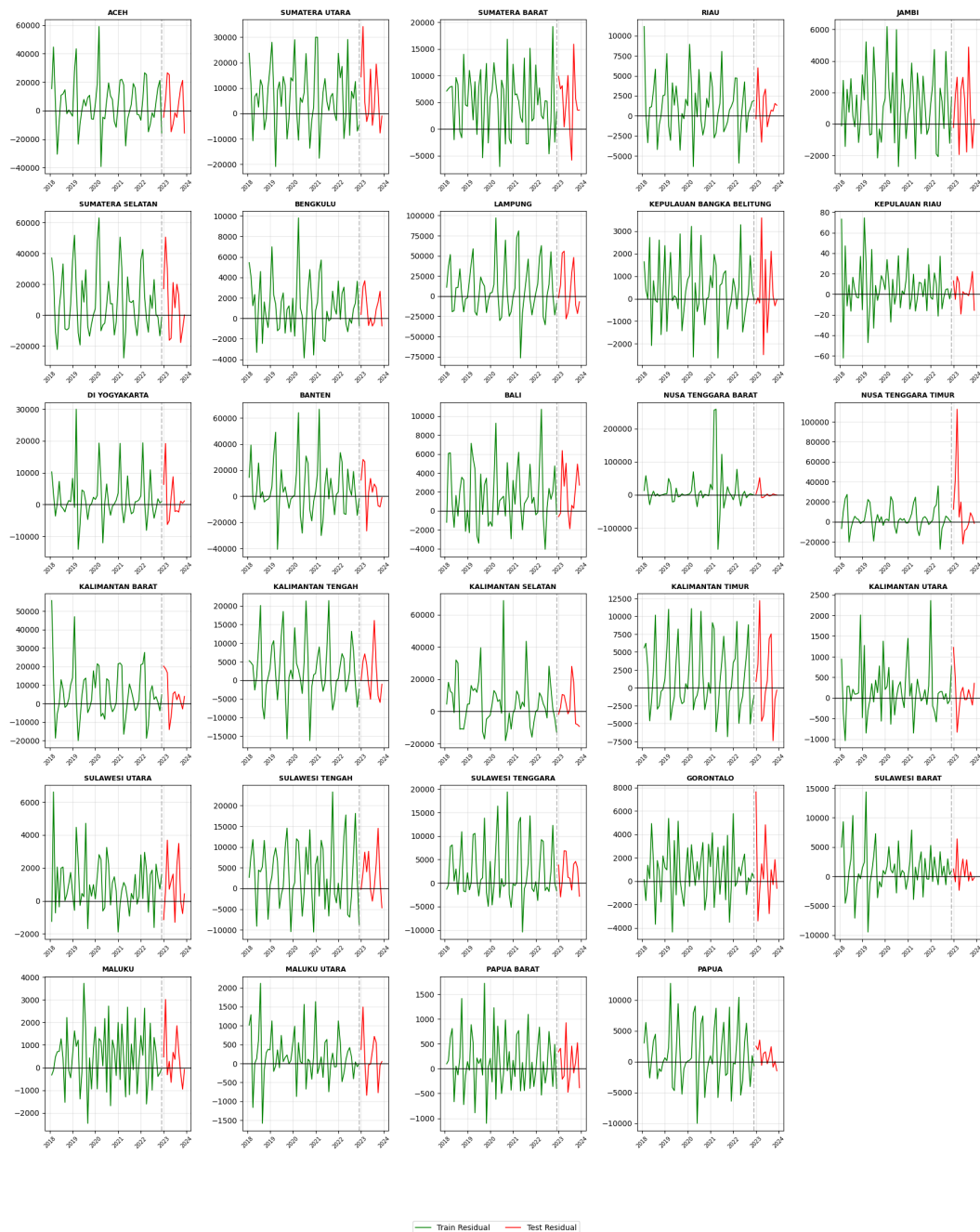
Gambar 21. Visualisasi Residual Kelompok Luas Panen Tinggi

Berdasarkan Gambar 21, pada kelompok luas panen tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, residual memiliki rentang variasi yang besar secara nominal. Hal ini sejalan dengan skala data aktual yang memang jauh lebih tinggi dibandingkan provinsi lain, sehingga selisih antara nilai aktual dan prediksi juga menjadi lebih besar. Residual juga masih menunjukkan pola naik dan turun di sekitar nilai nol. Selain itu, pada beberapa periode panen terlihat adanya lonjakan residual yang menunjukkan bahwa model linier belum sepenuhnya mampu mengakomodasi lonjakan produksi yang ekstrem.



Gambar 22. Visualisasi Residual Kelompok Luas Panen Sedang

Berdasarkan Gambar 22, pada kelompok luas panen sedang, yaitu Sulawesi Selatan, residual cenderung berada di sekitar nilai nol. Variasi yang ditunjukkan tidak sebesar kelompok tinggi. Selain itu, pada beberapa titik terdapat nilai residual yang cukup besar yang mencerminkan adanya selisih antara hasil prediksi dan data aktual. Hal ini menunjukkan bahwa hasil prediksi belum sepenuhnya menggambarkan pergerakan data secara keseluruhan, sehingga masih terdapat bagian pola yang belum terakomodasi.



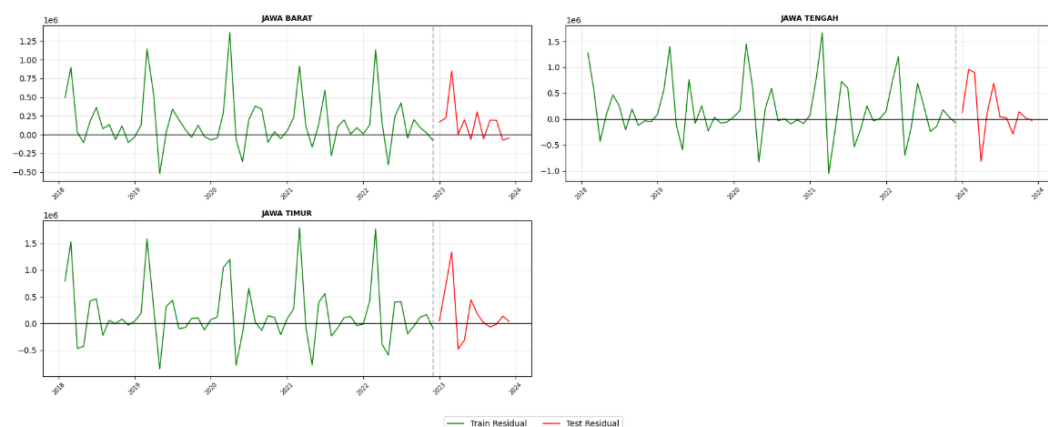
Gambar 23. Visualisasi Residual Kelompok Luas Panen Rendah

Berdasarkan Gambar 23, pada kelompok luas panen rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Riau, Jambi, Kepulauan Riau, hingga Papua, nilai residual menunjukkan variasi yang cukup beragam antar provinsi. Secara umum, residual bergerak di sekitar nilai nol dengan variasi yang relatif kecil. Namun, pada beberapa periode masih terlihat nilai residual yang

menyimpang cukup jauh dari nol, yang menunjukkan bahwa hasil prediksi belum sepenuhnya mengikuti perubahan data aktual. Perbedaan antara nilai prediksi dan aktual masih terlihat pada beberapa provinsi, dengan tingkat variasi yang berbeda-beda, di mana sebagian provinsi menunjukkan pergerakan yang lebih stabil, sementara yang lain cenderung lebih fluktuatif.

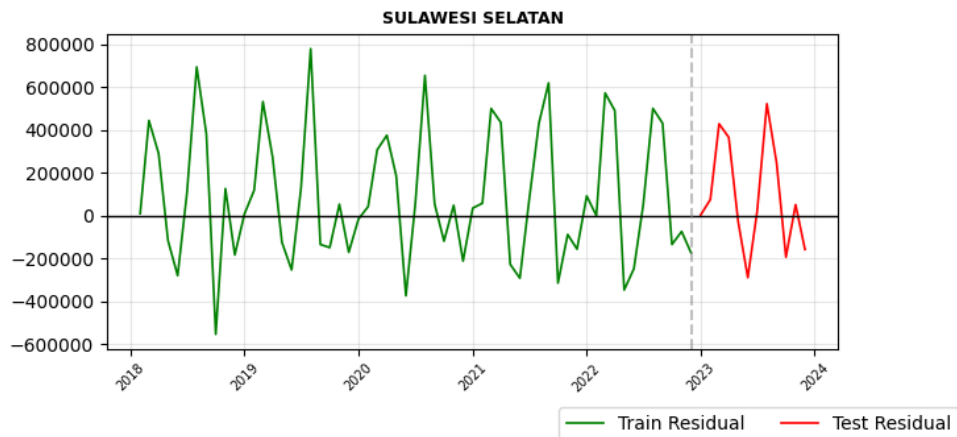
2. Produksi Padi

Visualisasi residual pada variabel produksi padi menunjukkan selisih antara nilai aktual dan hasil prediksi model PVAR pada masing-masing provinsi. Secara umum, nilai residual berada di sekitar nol sehingga hasil prediksi yang dihasilkan relatif mendekati nilai aktual. Adapun visualisasi tersebut disajikan dalam Gambar 24, 25, dan 26 sebagai berikut.



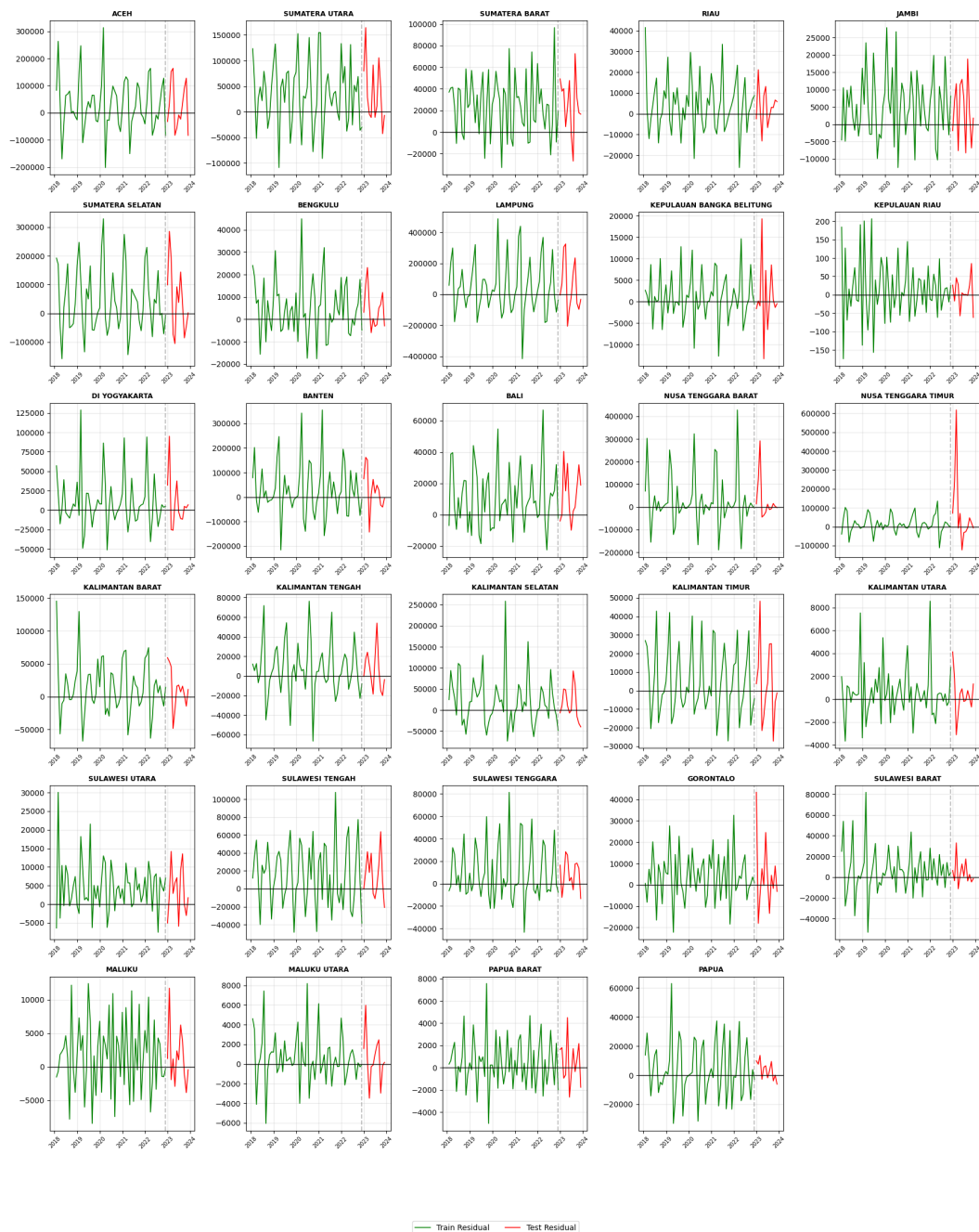
Gambar 24. Visualisasi Residual Kelompok Produksi Padi Tinggi

Berdasarkan Gambar 24, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, nilai residual menunjukkan rentang variasi yang cukup besar secara nominal, sejalan dengan skala data produksi yang tinggi. Nilai residual juga terlihat bergerak naik dan turun di sekitar nol. Pada beberapa periode, terlihat adanya lonjakan nilai residual yang cukup tinggi, yang menunjukkan bahwa hasil prediksi belum sepenuhnya mengikuti besarnya nilai aktual. Perbedaan antara nilai prediksi dan aktual masih terlihat, terutama pada saat terjadi perubahan yang cukup signifikan.



Gambar 25. Visualisasi Residual Kelompok Produksi Padi Sedang

Berdasarkan Gambar 25, pada kelompok produksi sedang, yaitu Sulawesi Selatan, nilai residual menunjukkan pergerakan naik dan turun di sekitar nilai nol dengan variasi pada tingkat menengah. Pada beberapa periode, terlihat adanya nilai residual yang cukup besar yang menunjukkan bahwa hasil prediksi belum sepenuhnya mengikuti perubahan data aktual. Perbedaan antara nilai prediksi dan aktual masih terlihat pada beberapa titik, terutama saat terjadi perubahan nilai. Secara umum, nilai residual masih mengandung bagian yang belum terakomodasi oleh model.



Gambar 26. Visualisasi Residual Kelompok Produksi Padi Rendah

Berdasarkan Gambar 26, pada kelompok produksi rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Riau, Jambi, Kepulauan Riau, hingga Papua, nilai residual menunjukkan variasi yang cukup beragam antar provinsi. Secara umum, residual bergerak di sekitar nilai nol dengan rentang variasi yang relatif lebih kecil dibandingkan kelompok produksi tinggi maupun sedang. Hal ini menunjukkan

bahwa selisih antara nilai prediksi dan nilai aktual tidak terlalu besar pada sebagian besar periode. Meskipun demikian, pada beberapa periode masih terlihat adanya nilai residual yang cukup tinggi, baik dalam bentuk kenaikan maupun penurunan, yang menunjukkan bahwa hasil prediksi belum sepenuhnya mengikuti perubahan data aktual. Selain itu, terdapat perbedaan karakteristik antar provinsi, di mana sebagian provinsi menunjukkan pergerakan residual yang relatif stabil, sementara provinsi lainnya memiliki variasi yang lebih fluktuatif. Hal ini mencerminkan adanya perbedaan dinamika data pada masing-masing wilayah. Secara umum, nilai residual masih mengandung bagian yang belum sepenuhnya terakomodasi oleh model, sehingga dapat dimanfaatkan untuk pemodelan lanjutan, khususnya sebagai input pada model LSTM.

4.3 Prediksi Model *Hybrid* PVAR-LSTM

Setelah diperoleh hasil estimasi dan residual dari model *Panel Vector Autoregressive* (PVAR), tahap selanjutnya adalah melakukan prediksi menggunakan model *hybrid*. Pendekatan *hybrid* ini bertujuan untuk menggabungkan keunggulan PVAR dalam menangkap hubungan linier antarvariabel dengan kemampuan *Long Short-Term Memory* (LSTM) dalam memodelkan pola nonlinier pada data. Pada tahap ini terdapat dua model yaitu model pertama P_LSTM dan model kedua E_LSTM. Pada model pertama P_LSTM, model ini menggunakan data prediksi PVAR sebagai input sedangkan pada model kedua E_LSTM, model ini menggunakan data residual PVAR sebagai input.

Dalam penelitian ini digunakan dua skema *hybrid* PVAR–LSTM. Pada skema pertama, hasil prediksi PVAR dan residual PVAR masing-masing dimodelkan kembali menggunakan LSTM yang kemudian dikombinasikan (P_LSTM + E_LSTM). Pada skema kedua, model PVAR terlebih dahulu digunakan untuk menghasilkan prediksi yang kemudian, residual dari PVAR dimodelkan menggunakan LSTM guna menangkap pola nonlinier yang masih tersisa. Prediksi akhir kemudian diperoleh dengan menggabungkan hasil prediksi PVAR dengan hasil prediksi residual dari LSTM (Hasil prediksi PVAR + E_LSTM). Melalui dua skema *hybrid* PVAR–LSTM tersebut, dilakukan perbandingan kinerja prediksi untuk melihat skema mana yang paling mampu menangkap pola data dengan baik.

4.3.1 *Scaling Data*

Berdasarkan hasil estimasi dan perhitungan residual dari model PVAR, terlihat bahwa nilai data memiliki variasi dan rentang yang cukup lebar. *Scaling* data diperlukan untuk menyeragamkan skala antar data sebelum digunakan sebagai input pada model LSTM. Proses ini bertujuan untuk mengoptimalkan performa model yang digunakan. Metode *scaling* yang digunakan adalah RobustScaler. RobustScaler melakukan *scaling* data berdasarkan nilai median dan *interquartile range* (IQR), sehingga lebih tahan terhadap pengaruh nilai ekstrem. *Scaling* dilakukan pada setiap provinsi dan untuk setiap jenis variabel. Setelah seluruh data *discaling*, diperoleh nilai dengan skala yang lebih seragam sehingga data siap digunakan dalam pembangunan model LSTM pada tahap selanjutnya, yang direspresentasikan pada Gambar 27.

```

===== ACEH =====
PVAR FEATURES (scaled)
Shape : (47, 2)
[[-0.44714623 -0.58670683]
 [ 0.25931754  0.09536176]
 [ 2.51329992  2.4776718 ]
 [ 2.07291142  2.07846228]
 [-0.30354365 -0.37906757]]

RESIDUAL FEATURES (scaled)
Shape : (47, 2)
[[ 0.53733309  0.58826374]
 [ 1.98342208  2.24708793]
 [ 0.00787455  0.19204072]
 [-1.73954971 -1.73621508]
 [-0.57654463 -0.52556108]]

```

Gambar 27. *Scaling Data*

4.3.2 Membangun Model P_LSTM

Setelah data hasil prediksi PVAR melalui proses *scaling* sebagaimana dijelaskan pada Subbab 4.3.1, tahap berikutnya adalah membangun model P_LSTM. Model ini menggunakan data prediksi PVAR sebagai input untuk menangkap pola yang belum sepenuhnya tertangkap oleh model linier. Sebelum dimasukkan ke dalam model, data dibentuk *sequence* dengan panjang jendela waktu (*lookback window*) sebesar 12 periode. Artinya, setiap sampel input terdiri atas 12 observasi

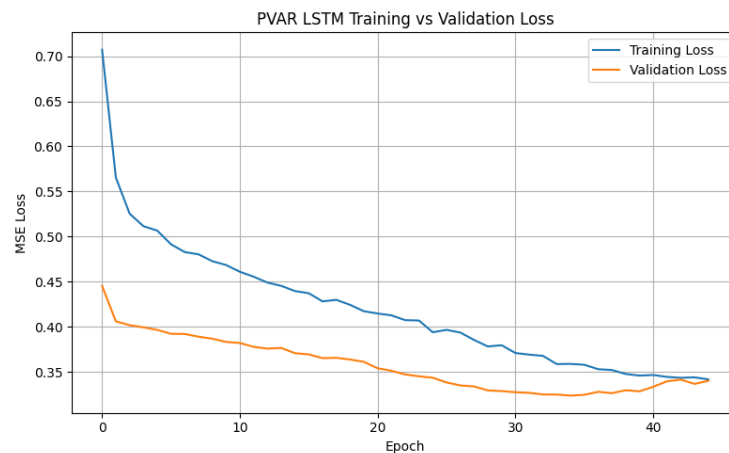
waktu sebelumnya yang digunakan untuk memprediksi nilai pada periode berikutnya.

Dalam pembentukan model ini ditambahkan komponen *province embedding* untuk memahami karakteristik heterogen antar provinsi. Setiap provinsi direpresentasikan dalam bentuk vektor *embedding* dengan dimensi tertentu yang ditentukan melalui proses *hyperparameter tuning*. Representasi *embedding* tersebut dipelajari secara otomatis selama proses pelatihan, sehingga model dapat menangkap perbedaan pola antarwilayah tanpa harus membangun model terpisah untuk setiap provinsi. Arsitektur model disusun dari beberapa parameter, yang kemudian ditentukan melalui proses optimasi menggunakan optuna untuk mendapatkan hasil parameter terbaik. Parameter yang dioptimasi dalam proses *hyperparameter tuning* disajikan dalam Tabel 12 berikut:

Tabel 12. *Hyperparameter Tuning Model P-LSTM*

<i>Hyperparameter</i>	Nilai	<i>Best Parameter</i>
<i>LSTM unit</i>	64, 128, 256	256
<i>Dense</i>	16, 32, 64	32
<i>Drop Out</i>	0.01 - 0.05	0.01032486341259099
<i>Learning Rate</i>	1e-4 - 5e-4	0.00016202403290479947
<i>Batch Size</i>	8, 16	8
<i>Embedding</i>	8, 16	8

Berdasarkan Tabel 12, pada proses pelatihan dilakukan digunakan maksimum 150 epoch dengan menerapkan *early stopping* berdasarkan nilai *validation loss*. Pelatihan akan dihentikan apabila tidak terjadi perbaikan dalam sejumlah epoch tertentu, dan bobot terbaik selama proses pelatihan akan dipulihkan kembali (*restore best weights*). Pendekatan ini bertujuan untuk mencegah terjadinya *overfitting* serta menjaga kemampuan generalisasi model terhadap data validasi, disajikan pada Gambar 28 sebagai berikut:



Gambar 28. Grafik *Loss Function* Model P_LSTM

Perkembangan nilai *training loss* dan *validation loss* selama proses pelatihan ditunjukkan pada Gambar 28. Berdasarkan grafik tersebut, terlihat bahwa nilai *loss* mengalami penurunan secara bertahap dan cenderung stabil pada akhir pelatihan. Hal ini menunjukkan bahwa model mampu mempelajari pola data dengan baik tanpa indikasi *overfitting* maupun *underfitting*. Model P_LSTM yang telah terlatih kemudian digunakan sebagai komponen dalam skema *hybrid* PVAR–LSTM untuk menghasilkan prediksi akhir yang lebih akurat.

4.3.3 Membangun Model E_LSTM

Tahap selanjutnya dalam skema *hybrid* adalah membangun model E_LSTM yang menggunakan residual dari model PVAR sebagai input. Residual merepresentasikan komponen kesalahan prediksi yang tidak mampu dijelaskan oleh model PVAR, sehingga diasumsikan masih mengandung pola nonlinier yang dapat dipelajari lebih lanjut menggunakan LSTM. Sama seperti pada model P_LSTM, data residual terlebih dahulu melalui proses *scaling* menggunakan RobustScaler, kemudian dibentuk *sequence* dengan panjang *lookback window* sebesar 12 periode. Dengan demikian, setiap sampel input terdiri atas 12 observasi residual sebelumnya untuk memprediksi nilai residual pada periode berikutnya.

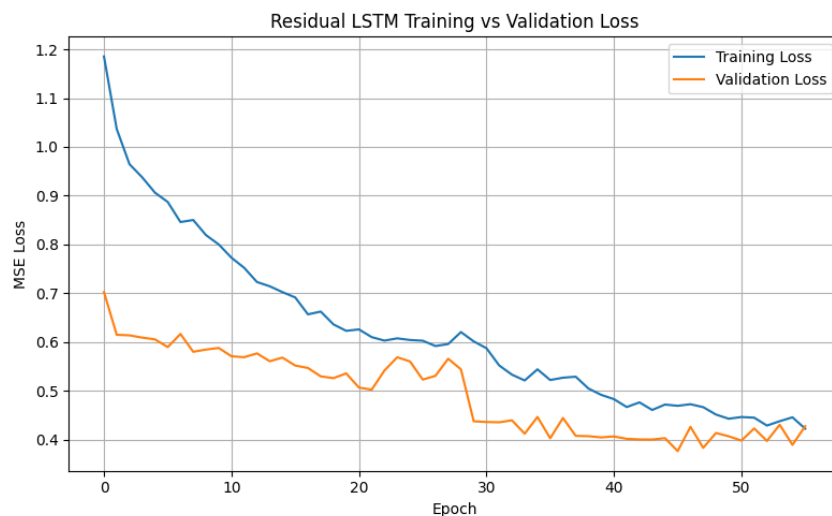
Model ini juga memanfaatkan pendekatan *province embedding* untuk menangkap heterogenitas antar provinsi. Setiap provinsi direpresentasikan dalam bentuk vektor *embedding* yang dimensinya ditentukan melalui proses *hyperparameter*

tuning. Pendekatan ini memungkinkan model memahami karakteristik spesifik masing-masing wilayah tanpa membangun model terpisah. Arsitektur model disusun dari beberapa parameter, yang kemudian ditentukan melalui proses optimasi menggunakan optuna untuk mendapatkan hasil parameter terbaik. Parameter yang dioptimasi dalam proses *hyperparameter tuning* disajikan pada Tabel 13 sebagai berikut:

Tabel 13. Hyperparameter Tuning Model E.LSTM

<i>Hyperparameter</i>	<i>Nilai</i>	<i>Best Parameter</i>
<i>LSTM unit</i>	128, 256	256
<i>Dense</i>	32, 64	32
<i>Drop Out</i>	0.02 - 0.05	0.03984128471637101
<i>Learning Rate</i>	1.5e-4 - 3e-4	0.00023695986825327211
<i>Batch Size</i>	16, 32	16
<i>Embedding</i>	16, 32	32

Berdasarkan Tabel 13, pada proses pelatihan dilakukan digunakan maksimum 150 epoch dengan menerapkan *early stopping* berdasarkan nilai *validation loss*. Pelatihan akan dihentikan apabila tidak terjadi perbaikan dalam sejumlah epoch tertentu, dan bobot terbaik selama proses pelatihan akan dipulihkan kembali (*restore best weights*). Pendekatan ini bertujuan untuk mencegah terjadinya *overfitting* serta menjaga kemampuan generalisasi model terhadap data validasi. Berikut adalah grafik *loss function* model E.LSTM yang disajikan pada Gambar 29.



Gambar 29. Grafik Loss Function Model E.LSTM

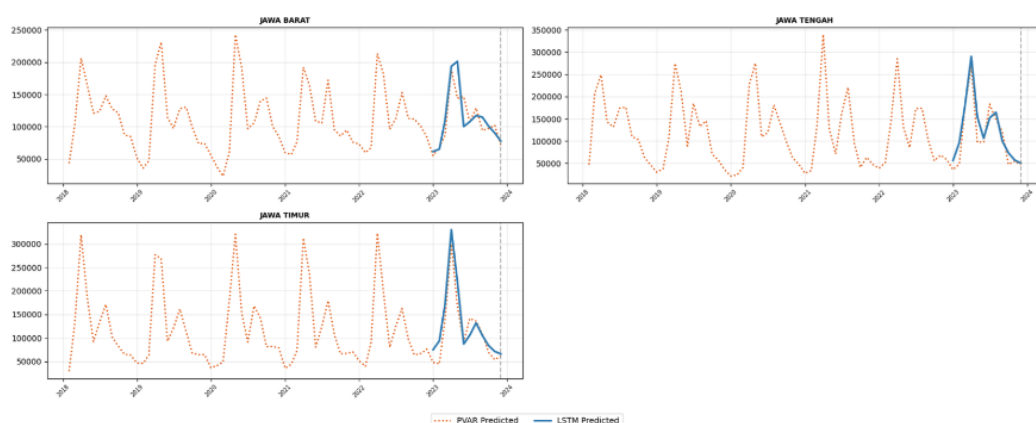
Perkembangan nilai *training loss* dan *validation loss* selama proses pelatihan ditunjukkan pada Gambar 29. Berdasarkan grafik tersebut, terlihat bahwa nilai *training loss* mengalami penurunan yang signifikan pada awal pelatihan dan kemudian menurun secara bertahap. Nilai *validation loss* juga menunjukkan tren penurunan, meskipun terdapat sedikit fluktuasi pada beberapa epoch. Pola ini mengindikasikan bahwa model mampu mempelajari struktur residual dengan baik tanpa mengalami *overfitting* maupun *underfitting*. Model E.LSTM yang telah terlatih kemudian digunakan untuk menghasilkan prediksi residual, yang selanjutnya dikombinasikan dengan hasil prediksi model PVAR dalam skema *hybrid* untuk memperoleh prediksi akhir yang lebih akurat.

4.3.4 Prediksi model P.LSTM

Setelah memperoleh kombinasi parameter terbaik untuk membangun model, tahap selanjutnya adalah melakukan prediksi menggunakan model P.LSTM pada data hasil prediksi model PVAR. Hasil prediksi model P.LSTM kemudian divisualisasikan dalam bentuk plot untuk setiap provinsi.

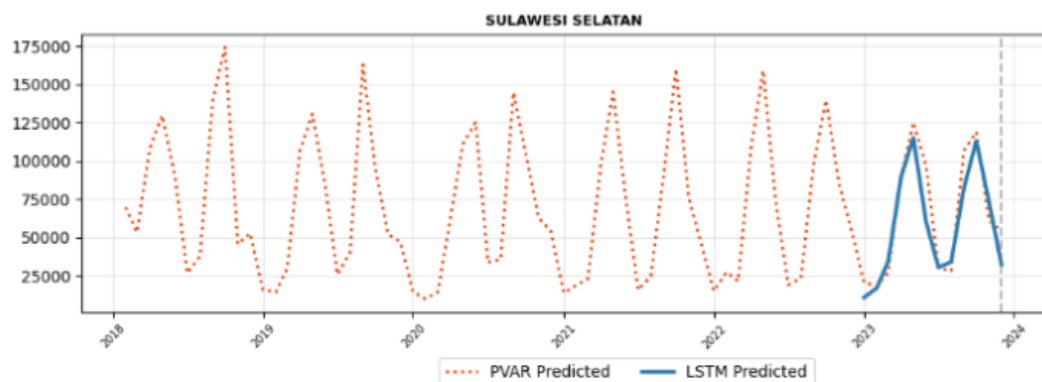
1. Luas Panen

Visualisasi hasil prediksi pada variabel luas panen memperlihatkan hasil prediksi pada masing-masing provinsi. Secara umum, nilai prediksi menunjukkan pola pergerakan tertentu pada setiap periode. Adapun visualisasi tersebut disajikan pada Gambar 30, 31 dan 32 sebagai berikut.



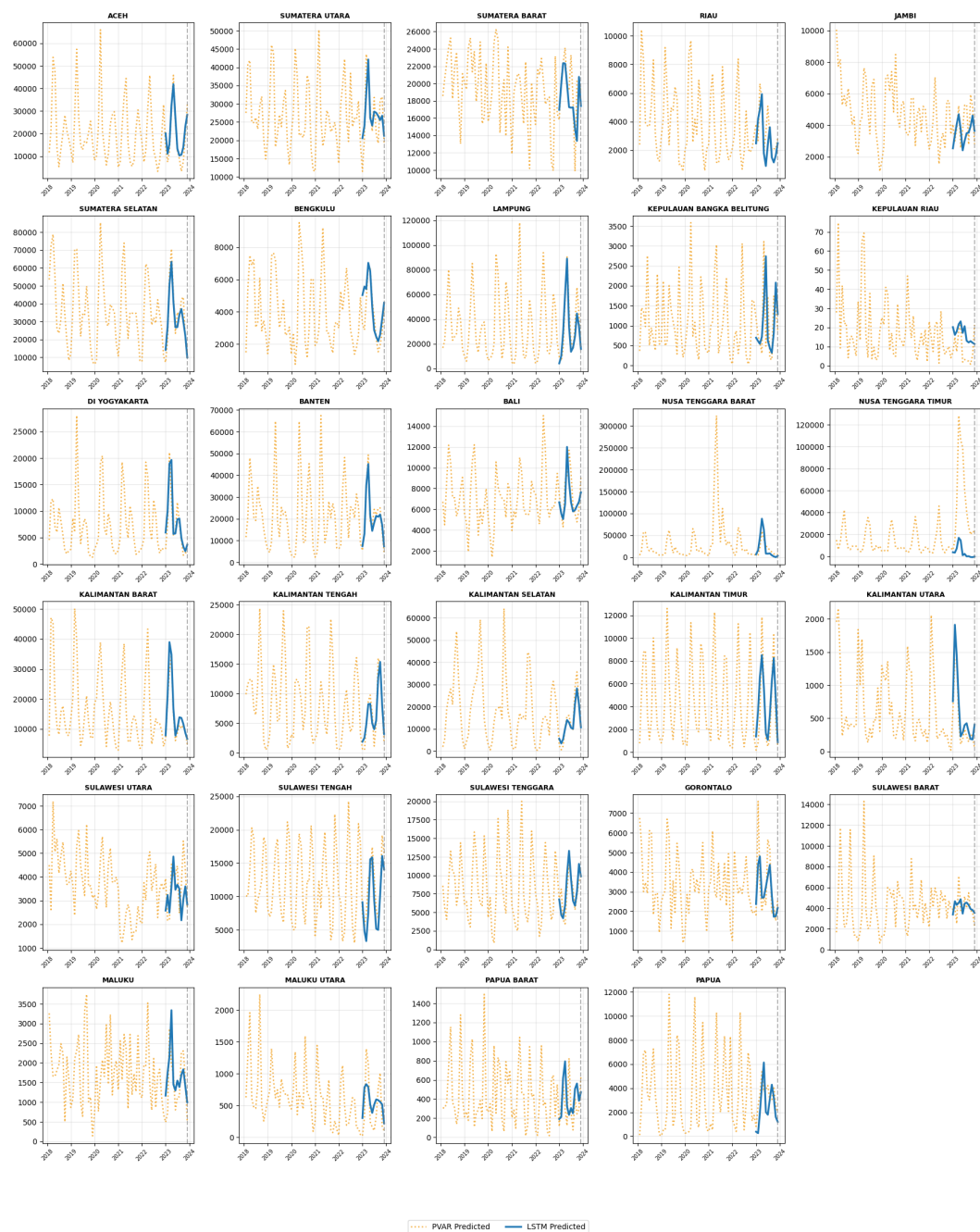
Gambar 30. Visualisasi Prediksi model P.LSTM Kelompok Luas Panen Tinggi

Berdasarkan Gambar 30, pada kelompok luas panen tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, model P_LSTM menunjukkan bahwa pola pergerakan data dapat diikuti dengan cukup baik. Pola kenaikan dan penurunan yang berulang masih dapat ditangkap, sehingga pergerakan yang dihasilkan mengikuti pola hasil prediksi PVAR. Pada beberapa periode, termasuk saat terjadi nilai puncak, hasil prediksi sudah mampu mengikuti pola kenaikan tersebut dengan cukup baik, meskipun masih terdapat sedikit selisih pada beberapa titik.



Gambar 31. Visualisasi Prediksi model P_LSTM Kelompok Luas Panen Sedang

Berdasarkan Gambar 31, pada kelompok luas panen sedang, yaitu Sulawesi Selatan, model P_LSTM menunjukkan bahwa pola pergerakan data dapat diikuti dengan cukup baik. Pola kenaikan dan penurunan yang berulang masih dapat ditangkap, sehingga pergerakan yang dihasilkan mengikuti pola hasil prediksi PVAR. Hasil prediksi juga terlihat mampu mengikuti pola kenaikan pada beberapa periode dengan cukup baik, meskipun masih terdapat selisih pada beberapa titik.



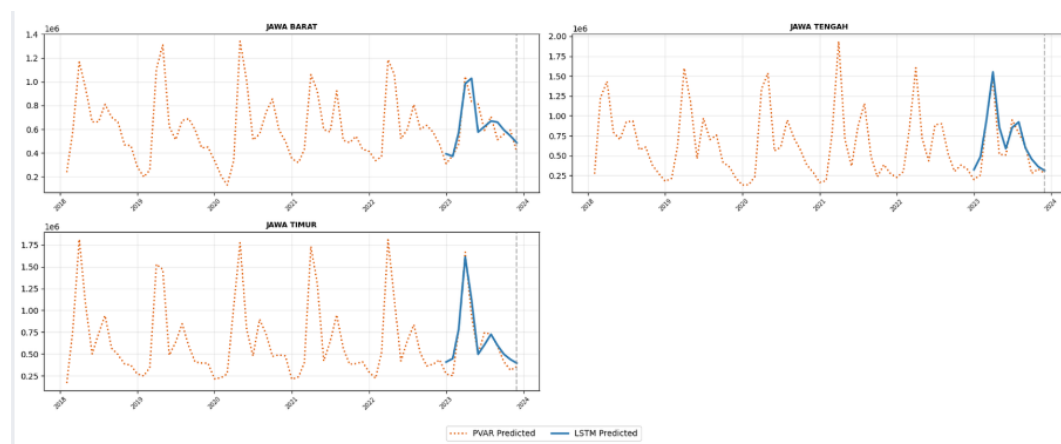
Gambar 32. Visualisasi Prediksi model P_LSTM Kelompok Luas Panen Rendah

Berdasarkan Gambar 32, pada kelompok luas panen rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Riau, Jambi, Kepulauan Riau, Bali, hingga Papua, model P_LSTM menunjukkan pola pergerakan yang bervariasi antar provinsi. Pola kenaikan dan penurunan masih dapat diikuti sesuai dengan

pola hasil prediksi PVAR. Pada beberapa provinsi seperti Aceh, Riau, Sumatra Selatan, Lampung dan Kalimantan tengah, pola pergerakan terlihat cukup berulang dengan fluktuasi yang relatif stabil, sehingga hasil prediksi dapat mengikuti pola tersebut meskipun masih terdapat selisih pada beberapa titik. Pada provinsi seperti Kepulauan Riau, nilai yang relatif kecil menyebabkan perubahan terlihat lebih rapat, sehingga perbedaan antar nilai prediksi masih terlihat pada beberapa periode pengujian. Sementara itu, Nusa Tenggara Timur (NTT) juga menunjukkan pola yang berbeda dengan adanya lonjakan yang cukup tinggi pada salah satu periode.

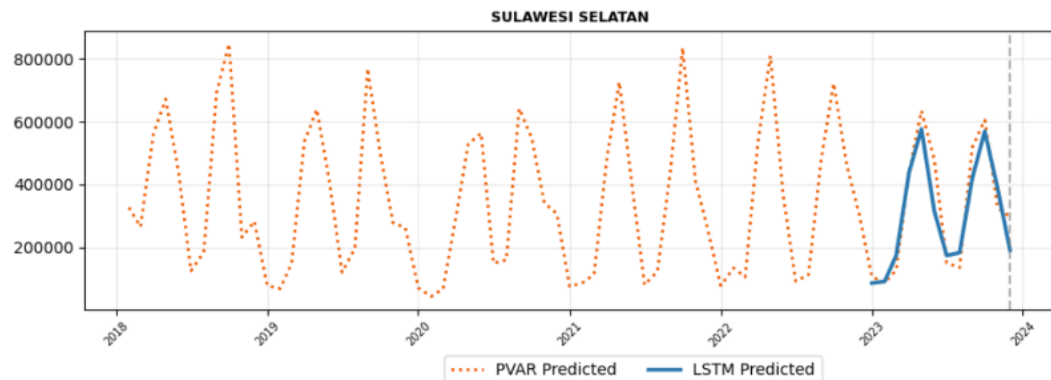
2. Produksi Padi

Visualisasi hasil prediksi pada variabel produksi padi pada masing-masing dapat disajikan pada Gambar 33, 34 dan 35 sebagai berikut.



Gambar 33. Visualisasi Prediksi model P_LSTM Kelompok Produksi Padi Tinggi

Berdasarkan Gambar 33, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, model P_LSTM menunjukkan bahwa pola pergerakan data dapat diikuti dengan baik. Pola kenaikan dan penurunan yang berulang masih dapat ditangkap, sehingga hasil prediksi mengikuti pola hasil prediksi PVAR. Pada beberapa periode pengujian, hasil prediksi terlihat mampu mengikuti pola kenaikan dengan cukup baik, termasuk saat terjadi nilai yang tinggi, meskipun masih terdapat selisih pada beberapa titik.



Gambar 34. Visualisasi Prediksi model P_LSTM Kelompok Produksi Padi Sedang

Berdasarkan Gambar 34, pada kelompok produksi sedang, yaitu Sulawesi Selatan, model P_LSTM menunjukkan bahwa pola pergerakan data dapat diikuti dengan cukup baik. Pola kenaikan dan penurunan yang berulang masih dapat ditangkap, sehingga hasil prediksi mengikuti pola hasil prediksi PVAR. Pada beberapa periode pengujian, hasil prediksi terlihat mampu mengikuti pola kenaikan dengan cukup baik, meskipun masih terdapat selisih pada beberapa titik.



Gambar 35. Visualisasi Prediksi model P LSTM Kelompok Produksi Padi Rendah

Berdasarkan Gambar 35, pada kelompok produksi rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, model P.LSTM menunjukkan pola pergerakan yang cukup bervariasi antar provinsi. Pola kenaikan dan penurunan masih dapat diikuti sesuai dengan pola hasil

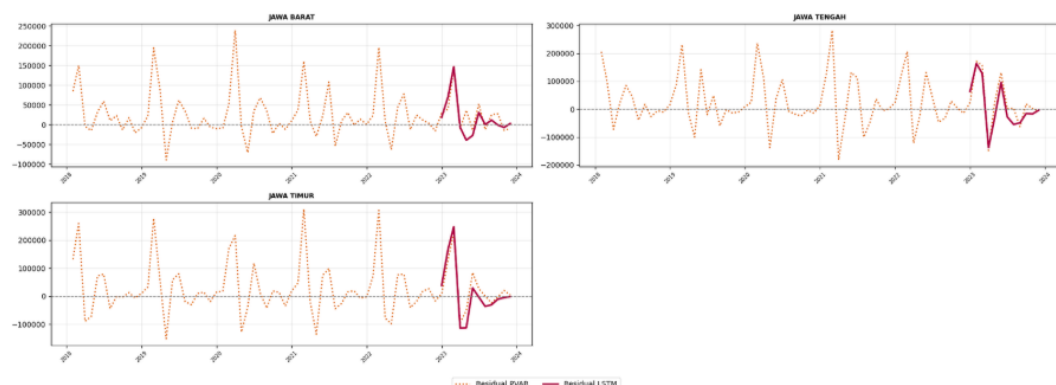
prediksi PVAR. Pada beberapa provinsi seperti Aceh, Sumatera Selatan, Banten, dan Bali, hasil prediksi terlihat mampu mengikuti pola pergerakan dengan cukup baik, meskipun masih terdapat selisih pada beberapa periode pengujian. Pada provinsi seperti Jambi dan Kepulauan Riau, nilai yang relatif kecil menyebabkan perubahan terlihat lebih rapat, sehingga perbedaan antar nilai prediksi masih terlihat pada beberapa periode pengujian. masih menunjukkan selisih pada beberapa periode pengujian. Selain itu, pada provinsi seperti Kalimantan Tengah, Kalimantan Selatan, dan Sulawesi Tenggara, pola pergerakan masih dapat diikuti, meskipun terdapat perbedaan pada beberapa titik. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan pola yang paling berbeda dengan adanya lonjakan yang sangat tinggi pada salah satu periode.

4.3.5 Prediksi model E_LSTM

Langkah selanjutnya adalah melakukan prediksi pada data residual. Hasil prediksi model E_LSTM kemudian divisualisasikan dalam bentuk plot untuk setiap provinsi sebagaimana disajikan sebagai berikut.

1. Luas Panen

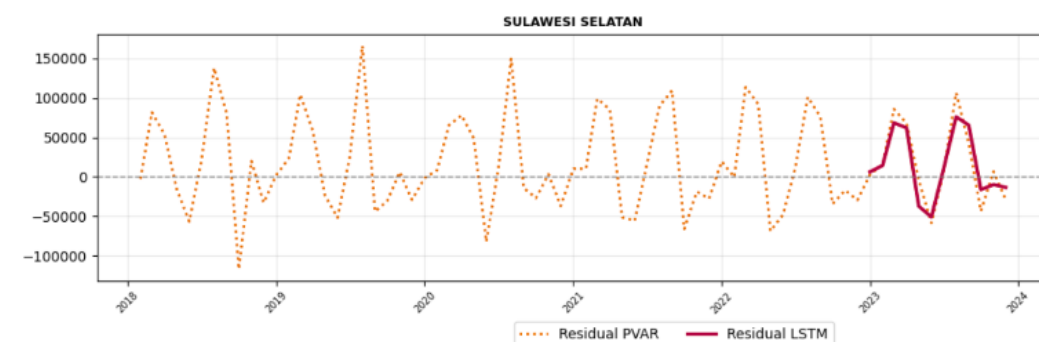
Visualisasi hasil prediksi pada variabel luas panen memperlihatkan hasil prediksi pada masing-masing provinsi. Secara umum, nilai prediksi menunjukkan pola pergerakan tertentu pada setiap periode. Adapun visualisasi tersebut disajikan pada Gambar 36, 37 dan 38 sebagai berikut.



Gambar 36. Visualisasi Prediksi model E_LSTM Kelompok Luas Panen Tinggi

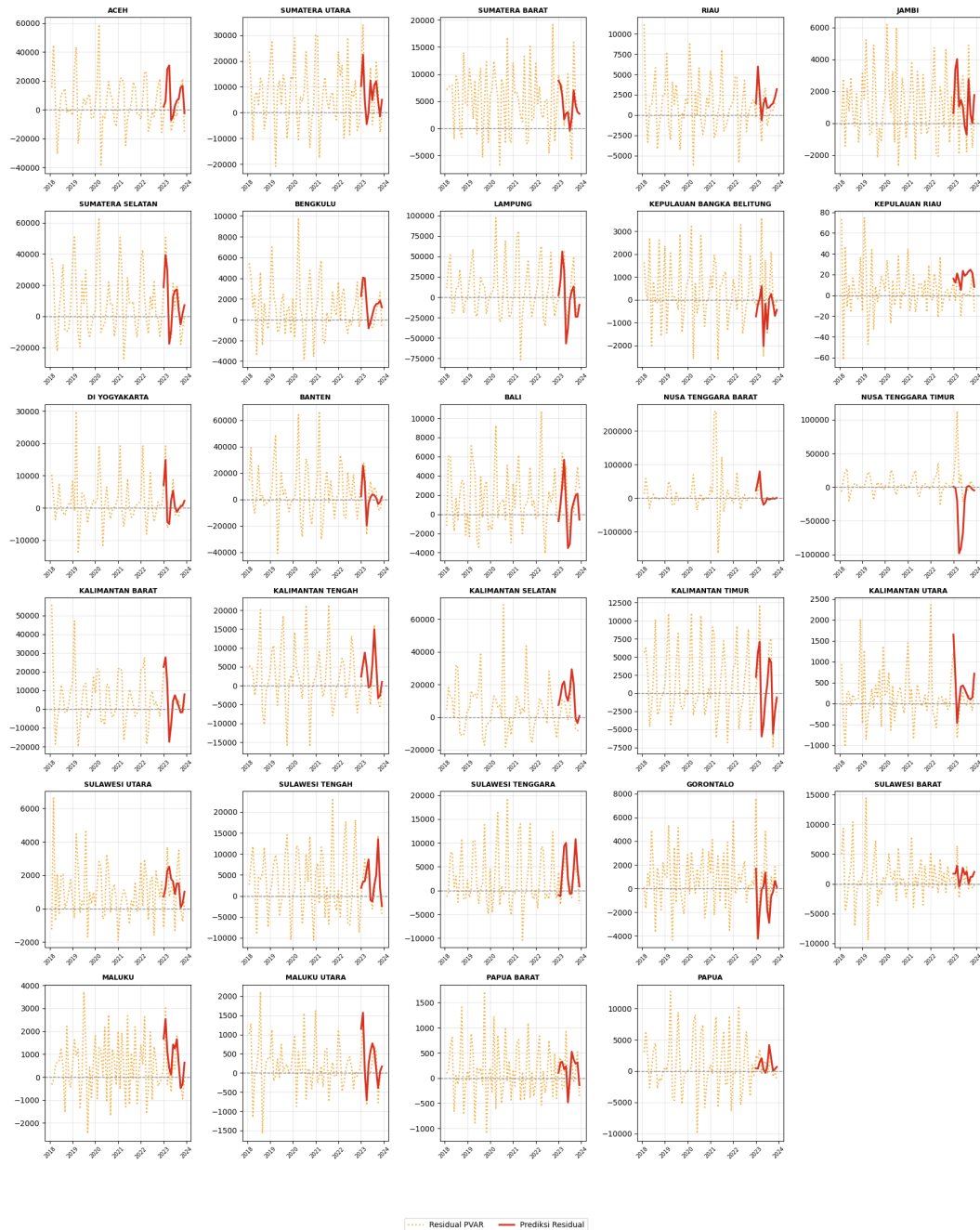
Berdasarkan Gambar 36, pada kelompok tinggi yang terdiri dari Jawa Barat, Jawa

Tengah, dan Jawa Timur, hasil prediksi residual oleh model E_LSTM menunjukkan bahwa pergerakan nilai dapat diikuti dengan cukup baik. Nilai prediksi mampu menangkap arah kenaikan dan penurunan pada beberapa periode pengujian. Pada Jawa Barat dan Jawa Tengah, hasil prediksi terlihat cukup mendekati nilai residual pada beberapa periode pengujian, meskipun masih terdapat selisih pada beberapa titik. Sementara itu, pada Jawa Timur, hasil prediksi juga mengikuti arah pergerakan, namun masih terlihat perbedaan yang lebih jelas pada beberapa periode pengujian.



Gambar 37. Visualisasi Prediksi model E LSTM Kelompok Luas Panen Sedang

Berdasarkan Gambar 37, pada kelompok sedang, yaitu Sulawesi Selatan, hasil prediksi residual oleh model E LSTM menunjukkan bahwa pergerakan nilai dapat diikuti dengan cukup baik. Nilai prediksi mampu menangkap arah kenaikan dan penurunan pada beberapa periode pengujian. Pada beberapa periode pengujian, hasil prediksi terlihat cukup mendekati nilai residual, meskipun masih terdapat selisih pada beberapa titik. Pergerakan naik dan turun juga terlihat dapat diikuti, sehingga hasil prediksi tidak menyimpang jauh dari nilai residual.



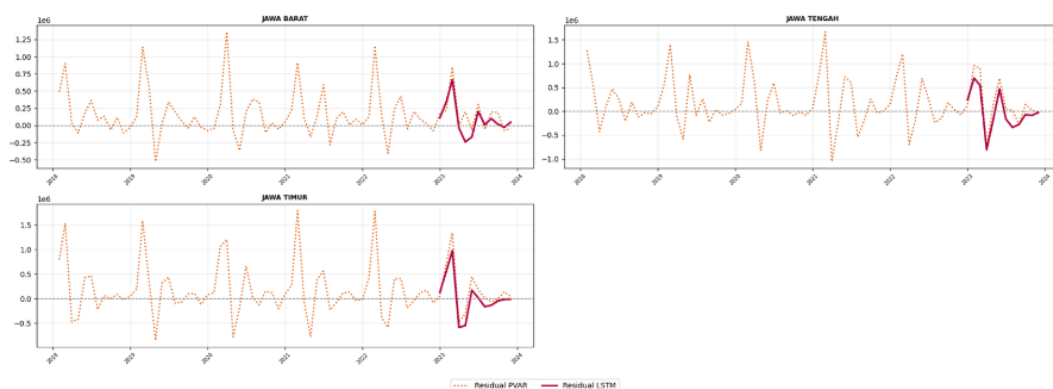
Gambar 38. Visualisasi Prediksi model E LSTM Kelompok Luas Panen Rendah

Berdasarkan Gambar 38, pada kelompok rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, hasil prediksi residual oleh model E LSTM menunjukkan pergerakan yang bervariasi antar provinsi. Secara umum, nilai prediksi mampu mengikuti arah kenaikan dan

penurunan pada beberapa periode pengujian. Pada provinsi seperti Aceh, Sumatera Selatan, Kalimantan Barat dan Kalimantan Tengah, hasil prediksi terlihat cukup mendekati nilai residual pada beberapa periode pengujian, meskipun masih terdapat selisih pada beberapa titik. Pada provinsi seperti Jambi, Kepulauan Riau, dan Bangka Belitung, nilai prediksi juga mengikuti arah pergerakan, namun perbedaan masih terlihat pada beberapa periode pengujian. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan perbedaan yang paling jelas, di mana pada salah satu periode pengujian terdapat nilai yang sangat tinggi, sehingga hasil prediksi belum sepenuhnya mengikuti nilai tersebut dan terlihat selisih yang cukup besar.

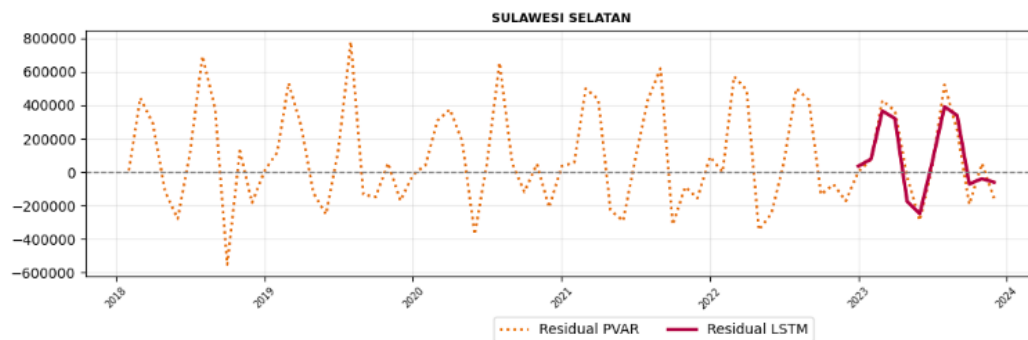
2. Produksi Padi

Visualisasi hasil prediksi pada variabel produksi padi menyajikan nilai prediksi yang diperoleh dari pemodelan residual pada masing-masing provinsi. Secara umum, nilai prediksi menunjukkan pergerakan yang bervariasi pada setiap periode. Adapun visualisasi tersebut disajikan pada Gambar 39, 40 dan 41 sebagai berikut.



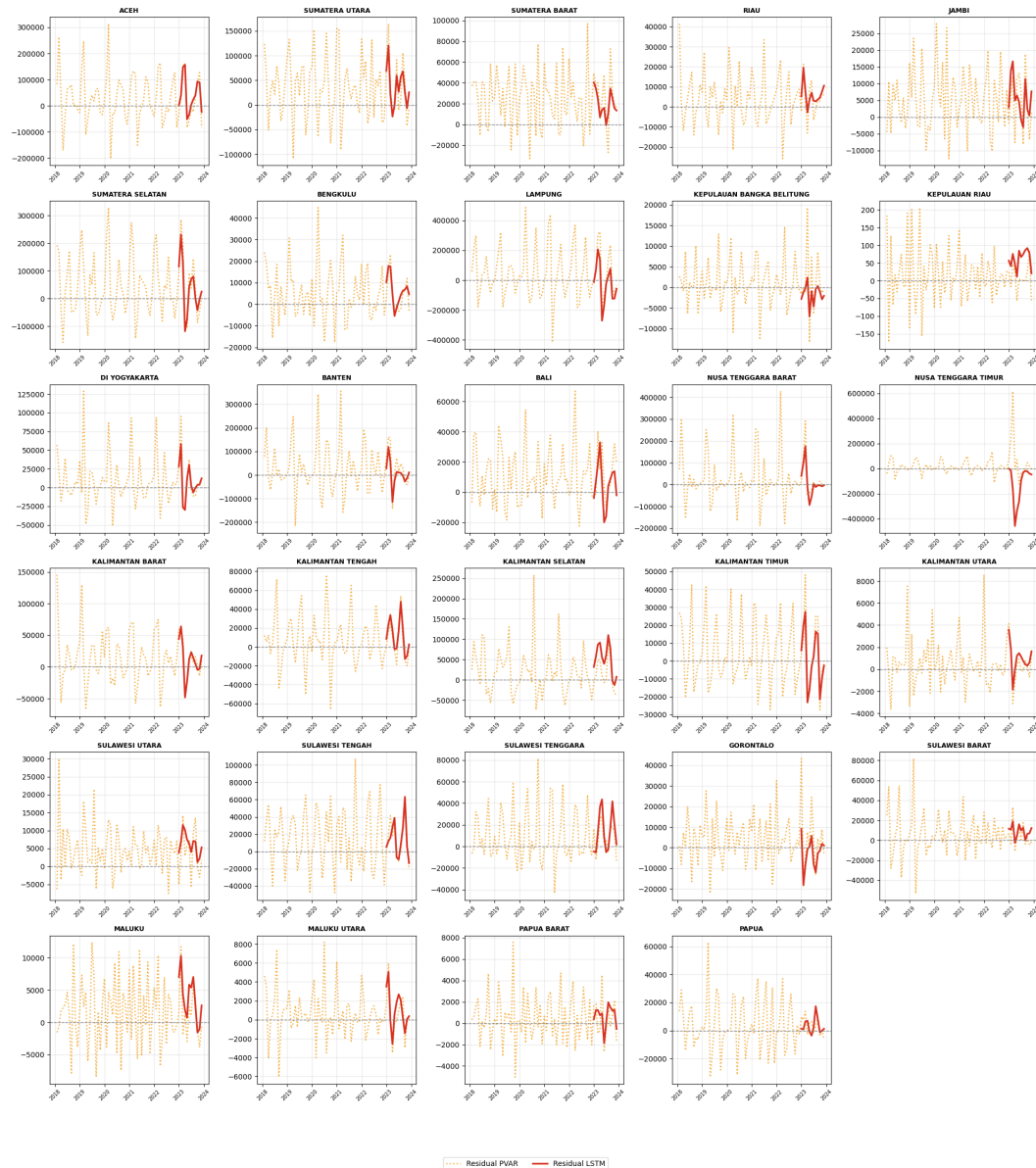
Gambar 39. Visualisasi Prediksi model E.LSTM Kelompok Produksi Padi Sedang

Berdasarkan Gambar 39, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil prediksi residual oleh model E.LSTM menunjukkan bahwa pergerakan nilai dapat diikuti dengan cukup baik. Nilai prediksi mampu menangkap arah kenaikan dan penurunan pada beberapa periode pengujian. Pada Jawa Barat dan Jawa Tengah, hasil prediksi terlihat cukup mendekati nilai residual pada beberapa periode pengujian, meskipun masih terdapat selisih pada beberapa titik.



Gambar 40. Visualisasi Prediksi model E_LSTM Kelompok Produksi Padi Sedang

Berdasarkan Gambar 40, pada kelompok produksi sedang, seperti Sulawesi Selatan, hasil prediksi residual oleh model E_LSTM menunjukkan bahwa pergerakan nilai dapat diikuti dengan cukup baik. Nilai prediksi mampu menangkap arah kenaikan dan penurunan pada beberapa periode pengujian. Pada beberapa periode pengujian, hasil prediksi terlihat cukup mendekati nilai residual, terutama saat terjadi kenaikan, meskipun masih terdapat selisih pada beberapa titik. Selain itu, pada saat terjadi penurunan, hasil prediksi juga mengikuti arah pergerakan, meskipun perbedaan masih terlihat pada beberapa periode pengujian.



Gambar 41. Visualisasi Prediksi model E_LSTM Produksi Padi Rendah

Berdasarkan Gambar 41, pada kelompok produksi rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, hasil prediksi residual oleh model E_LSTM menunjukkan pergerakan yang bervariasi antar provinsi. Secara umum, nilai prediksi mampu mengikuti arah kenaikan dan penurunan pada beberapa periode pengujian. Pada provinsi seperti Aceh, Sulawesi Tenggara, Kalimantan Tengah dan Sumatera Selatan, hasil prediksi terlihat cukup mendekati nilai residual pada beberapa periode pengujian, meskipun masih terdapat selisih pada beberapa titik. Pada provinsi seperti Riau, Jambi, dan

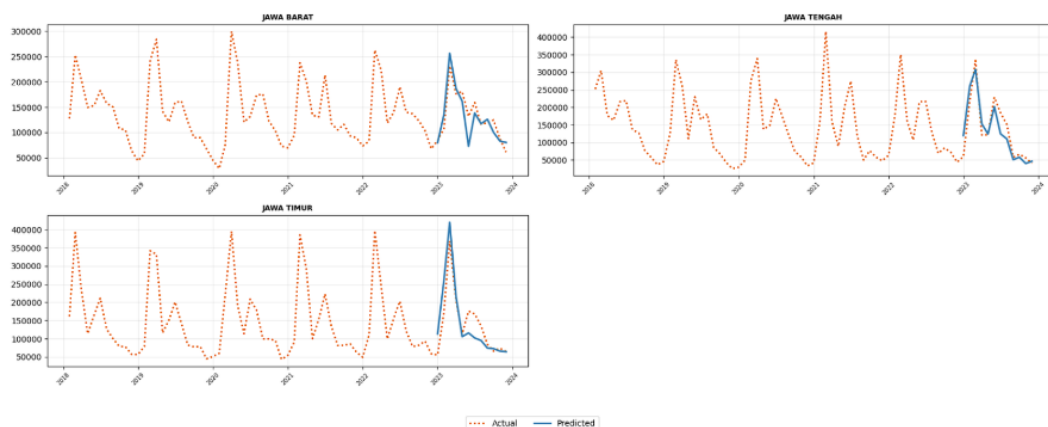
Kepulauan Riau, hasil prediksi juga mengikuti arah pergerakan, namun perbedaan masih terlihat pada beberapa periode pengujian. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan perbedaan yang paling jelas, di mana pada salah satu periode pengujian terdapat nilai yang sangat ekstrem, sehingga hasil prediksi belum sepenuhnya mengikuti nilai tersebut dan terlihat selisih yang cukup besar.

4.3.6 Prediksi model *Hybrid* PVAR-LSTM Skema 1 (EP)

Setelah diperoleh hasil prediksi dari model P_LSTM (berbasis prediksi PVAR) dan model E_LSTM (berbasis residual PVAR), tahap selanjutnya adalah membentuk prediksi akhir pada Skema 1. Pada skema ini, kedua keluaran model LSTM (P_LSTM + E_LSTM) tersebut dikombinasikan untuk menghasilkan nilai prediksi akhir. Berikut merupakan visualisasi skema 1.

1. Luas Panen

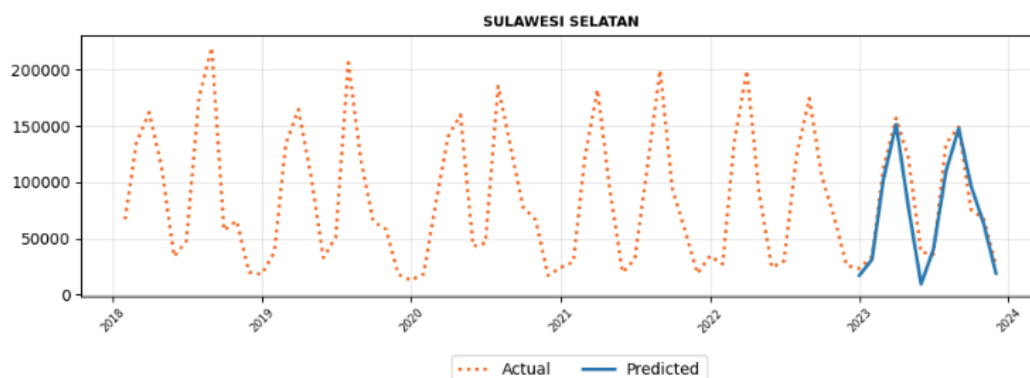
Visualisasi hasil prediksi luas panen pada masing-masing provinsi disajikan untuk menggambarkan pola pergerakan nilai hasil prediksi. Adapun visualisasi tersebut disajikan pada Gambar 42, 43 dan 44 sebagai berikut.



Gambar 42. Visualisasi Prediksi *Hybrid* PVAR-LSTM Skema 1 Kelompok Luas Panen Tinggi

Berdasarkan Gambar 42, pada kelompok luas panen tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pada Jawa Barat, pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, termasuk saat terjadi peningkatan nilai yang cukup tinggi, sehingga nilai prediksi

mendekati data aktual pada beberapa periode. Pada Jawa Tengah, hasil prediksi juga mengikuti pola yang sama dengan data aktual, di mana kenaikan dan penurunan dapat diikuti dengan cukup baik, meskipun masih terdapat selisih pada beberapa titik. Pada Jawa Timur, nilai prediksi menunjukkan kemampuan dalam mengikuti lonjakan yang tinggi serta penurunan setelahnya, sehingga pola pergerakan data dapat diikuti dengan cukup jelas.



Gambar 43. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 1 Kelompok Luas Panen Sedang

Berdasarkan Gambar 43, pada kelompok luas panen sedang yaitu Sulawesi Selatan, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, termasuk saat terjadi peningkatan nilai yang cukup tinggi, di mana hasil prediksi dapat mendekati nilai aktual. Selain itu, saat terjadi penurunan, nilai prediksi juga mengikuti arah perubahan tersebut meskipun masih terdapat selisih pada beberapa titik.



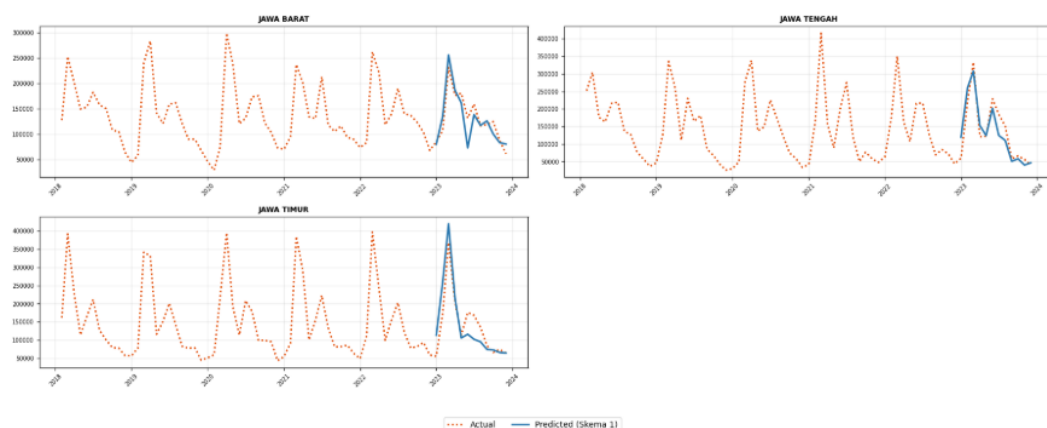
Gambar 44. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 1 Kelompok Luas Panen Rendah

Berdasarkan Gambar 44, pada kelompok luas panen rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual di sebagian besar provinsi. Pada

provinsi seperti Aceh, Kalimantan Barat, Kalimantan Tengah, Sulawesi Tengah dan Sumatera Selatan, pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, sehingga nilai prediksi cukup mendekati data aktual. Pada provinsi seperti Riau, Jambi, dan Kepulauan Riau, hasil prediksi juga mengikuti arah perubahan data, meskipun masih terdapat selisih masih terlihat cukup jelas. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan perbedaan yang paling jelas, di mana terdapat perubahan nilai yang cukup besar pada salah satu periode pengujian, sehingga hasil prediksi belum sepenuhnya mengikuti nilai tersebut.

2. Produksi Padi

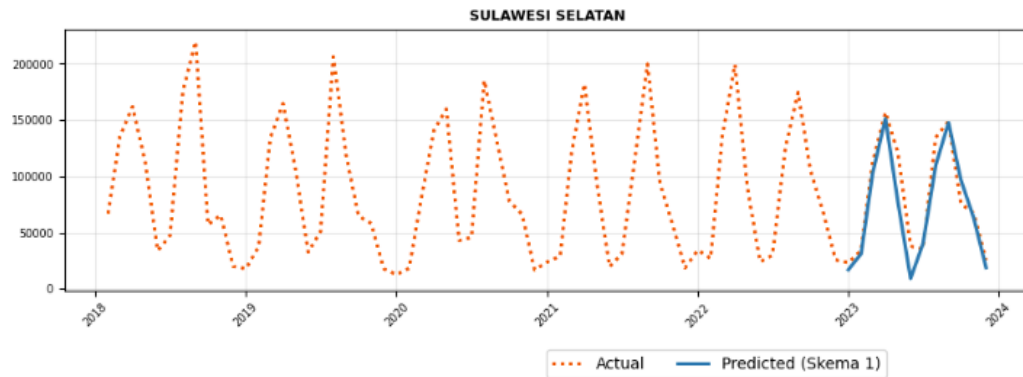
Visualisasi hasil prediksi produksi padi pada masing-masing provinsi disajikan untuk menggambarkan pola pergerakan nilai hasil prediksi. Adapun visualisasi tersebut disajikan pada Gambar 45, 46, dan 47 sebagai berikut.



Gambar 45. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 1 Kelompok Produksi Tinggi

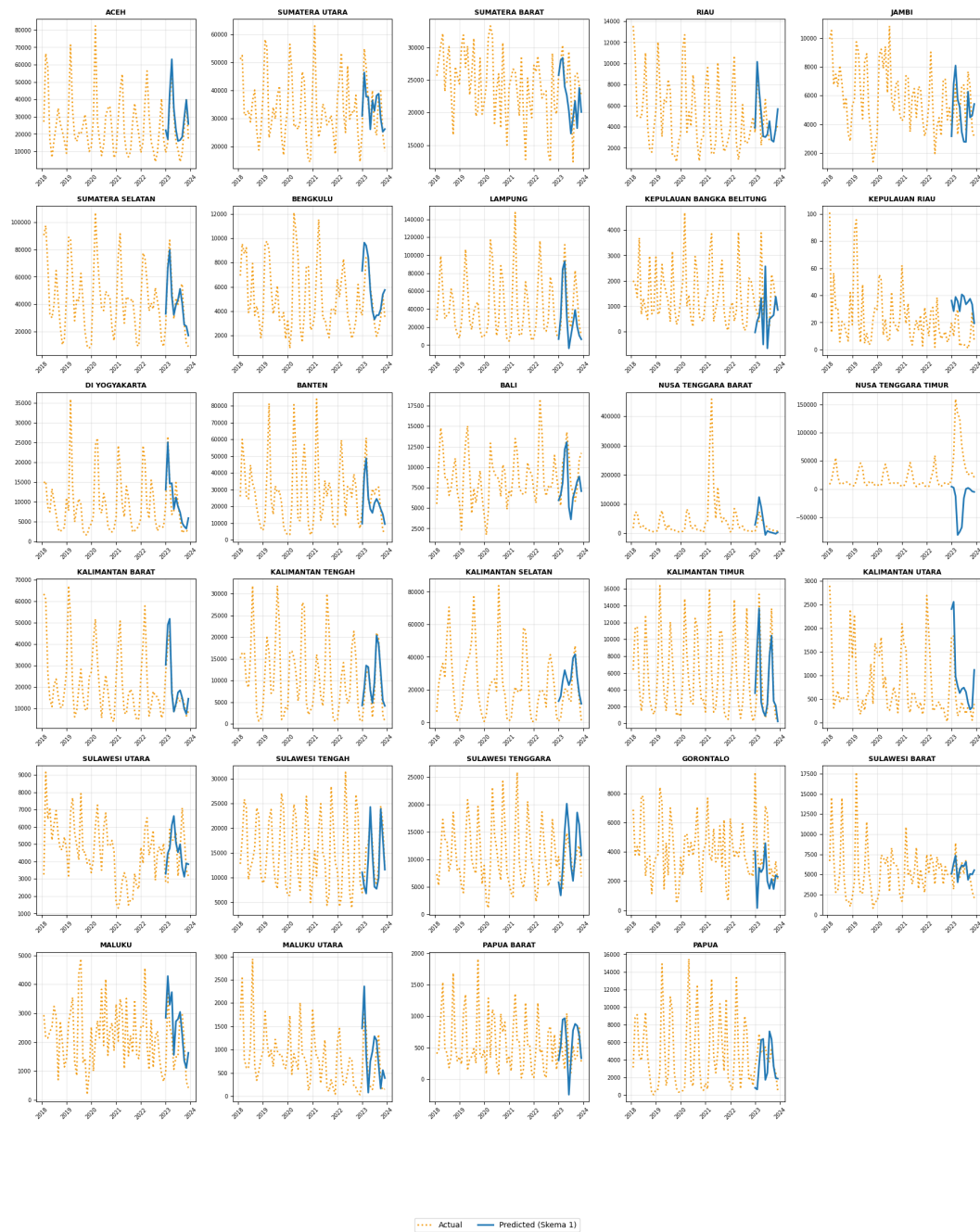
Berdasarkan Gambar 45, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi sudah mengikuti pola data aktual dengan cukup baik. Pada Jawa Barat, hasil prediksi mampu mengikuti pola naik dan turun, termasuk saat terjadi peningkatan yang cukup tinggi, sehingga nilai prediksi terlihat mendekati data aktual. Pada Jawa Tengah, pergerakan prediksi juga mengikuti pola yang sama dengan data aktual, terutama pada saat terjadi kenaikan yang cukup besar, di mana nilai prediksi dapat menangkap perubahan tersebut dengan baik meskipun masih terdapat selisih pada beberapa titik. Pada Jawa Timur, hasil prediksi menunjukkan kemampuan yang baik dalam mengikuti lonjakan nilai yang tinggi, di mana puncak

produksi dapat diikuti dengan cukup jelas, serta penurunan setelahnya juga dapat diikuti oleh model.



Gambar 46. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 1 Kelompok Produksi Sedang

Berdasarkan Gambar 46, pada kelompok produksi sedang yaitu Sulawesi Selatan, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, termasuk saat terjadi peningkatan nilai yang cukup tinggi, di mana hasil prediksi dapat mendekati nilai aktual. Selain itu, saat terjadi penurunan, nilai prediksi juga mengikuti arah perubahan tersebut meskipun masih terdapat selisih pada beberapa titik.



Gambar 47. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 1 Kelompok Produksi Rendah

Berdasarkan Gambar 47, pada kelompok produksi rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, hasil prediksi menggunakan skema 1 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual di sebagian besar provinsi. Pada provinsi seperti Aceh,

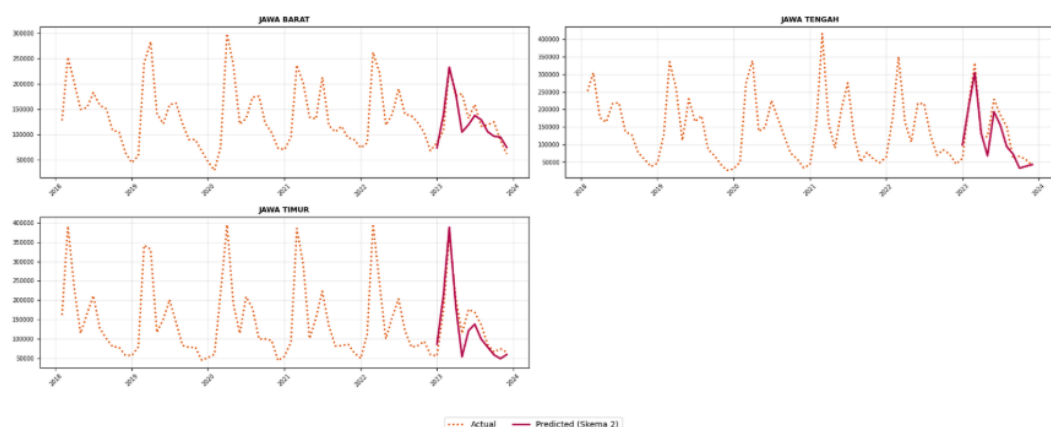
Kalimantan Barat, Kalimantan Tengah, Sulawesi Tengah dan Sumatera Selatan, pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, sehingga nilai prediksi cukup mendekati data aktual. Pada provinsi seperti Riau, Jambi, dan Kepulauan Riau, hasil prediksi juga mengikuti arah perubahan data, meskipun masih terdapat selisih masih terlihat cukup jelas. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan perbedaan yang cukup jelas, di mana terdapat perubahan nilai yang sangat besar pada salah satu periode, sehingga hasil prediksi belum sepenuhnya mengikuti nilai tersebut.

4.3.7 Prediksi model *Hybrid* PVAR-LSTM Skema 2 (E)

Tahap selanjutnya adalah membentuk prediksi akhir pada Skema 2. Pada skema ini, digunakan hasil prediksi PVAR, kemudian dikombinasikan dengan prediksi residual dari LSTM. Hasil prediksi akhir Skema 2 kemudian divisualisasikan dalam bentuk grafik perbandingan antara nilai aktual dan nilai prediksi untuk masing-masing provinsi, baik untuk variabel luas panen maupun produksi padi sebagai berikut.

1. Luas Panen

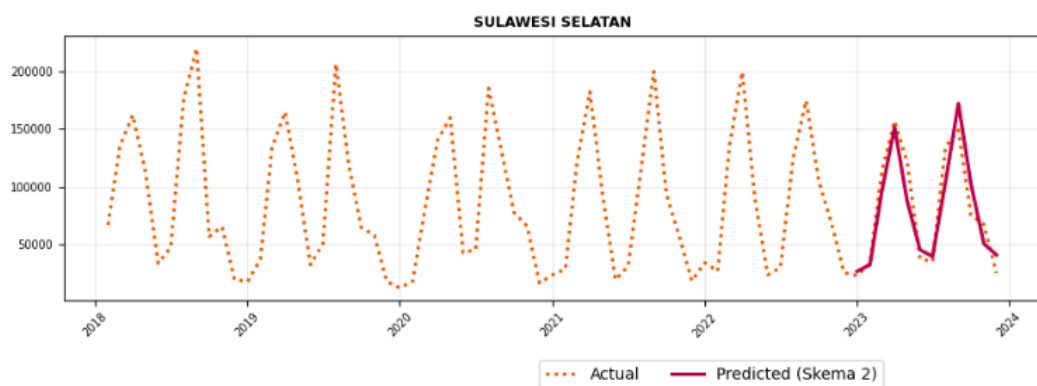
Visualisasi hasil prediksi luas panen pada setiap provinsi ditampilkan untuk melihat bagaimana pola pergerakan nilai hasil prediksi. Adapun visualisasi tersebut disajikan pada Gambar 48, 49, dan 50 sebagai berikut.



Gambar 48. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Luas Panen Tinggi

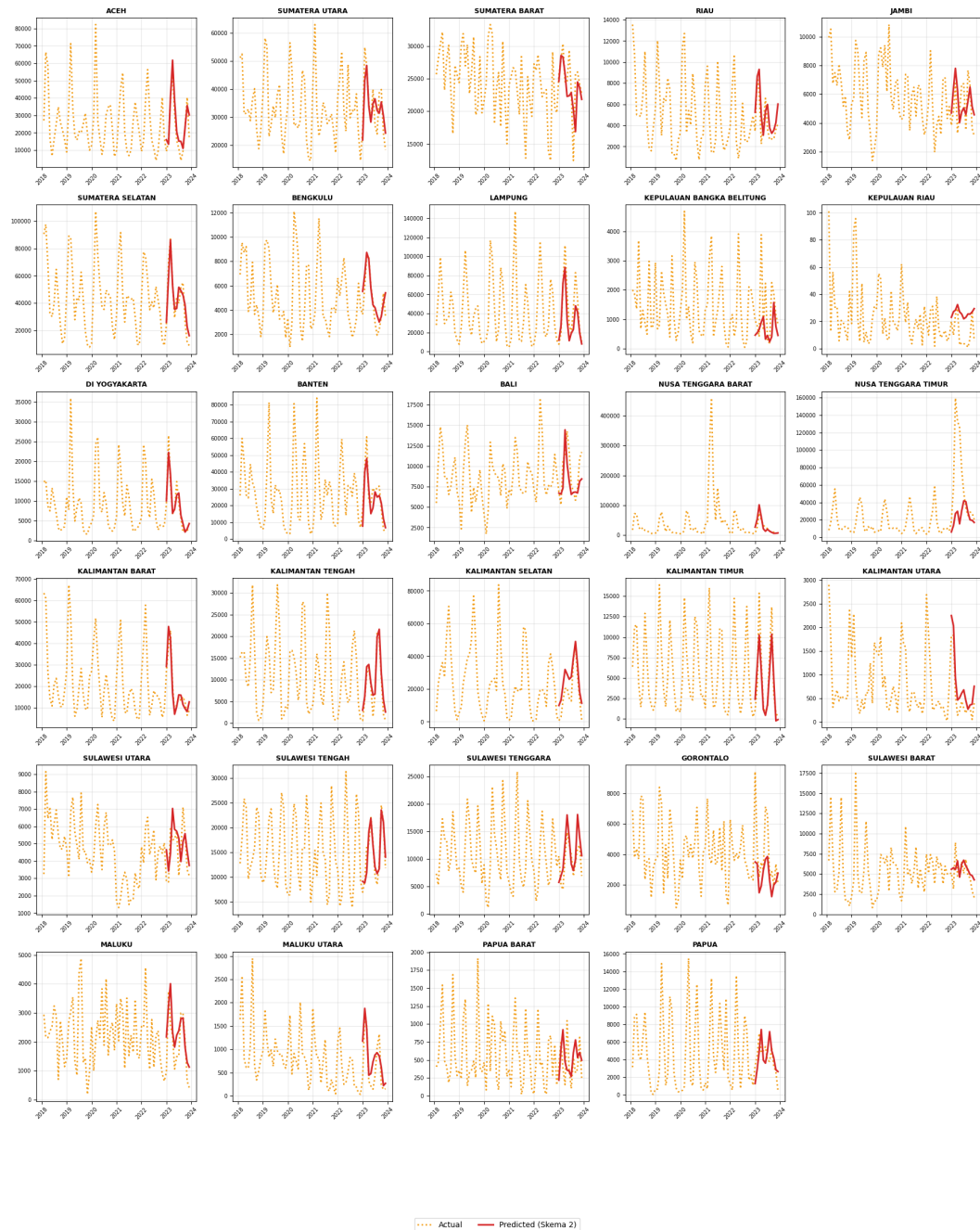
Berdasarkan Gambar 48, pada kelompok luas panen tinggi yang terdiri dari

Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil prediksi menggunakan skema 2 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pada Jawa Barat, pergerakan prediksi terlihat mengikuti pola naik dan turun, termasuk saat terjadi peningkatan nilai yang cukup tinggi, sehingga nilai prediksi mendekati data aktual pada beberapa periode. Pada Jawa Tengah, hasil prediksi juga mengikuti pola yang sama dengan data aktual, di mana kenaikan dan penurunan dapat diikuti dengan cukup baik, meskipun masih terdapat selisih pada beberapa titik. Pada Jawa Timur, nilai prediksi mampu mengikuti lonjakan yang tinggi serta penurunan setelahnya, sehingga pola pergerakan data dapat diikuti dengan cukup jelas.



Gambar 49. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Luas Panen Sedang

Berdasarkan Gambar 49, pada kelompok luas panen sedang yaitu Sulawesi Selatan, hasil prediksi menggunakan skema 2 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pergerakan prediksi terlihat mengikuti pola naik dan turun yang terjadi, termasuk saat terjadi peningkatan nilai yang cukup tinggi, di mana hasil prediksi dapat mendekati nilai aktual. Selain itu, saat terjadi penurunan, nilai prediksi juga mengikuti arah perubahan tersebut meskipun masih terdapat selisih pada beberapa titik.



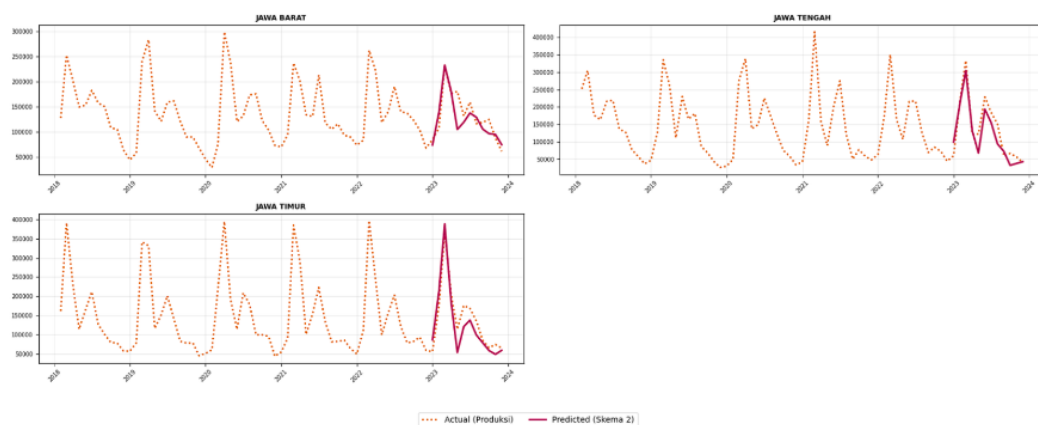
Gambar 50. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Luas Panen Rendah

Berdasarkan Gambar 50, pada kelompok luas panen rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, hingga Papua, hasil prediksi menggunakan skema 2 menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual pada periode pengujian di sebagian

besar provinsi. Pada provinsi seperti Aceh, Kalimantan Barat, Kalimantan Tengah, Sulawesi Tengah dan Sumatera Selatan, pergerakan prediksi terlihat mengikuti pola naik dan turun, sehingga nilai prediksi cukup mendekati data aktual pada beberapa periode pengujian. Pada provinsi seperti Riau, Jambi, dan Kepulauan Riau, hasil prediksi juga mengikuti arah perubahan data, meskipun masih terdapat selisih masih terlihat cukup jelas. Sementara itu, Nusa Tenggara Timur (NTT) menunjukkan perbedaan yang cukup jelas, di mana terdapat perubahan nilai yang besar pada salah satu periode pengujian, sehingga hasil prediksi belum sepenuhnya mengikuti nilai tersebut.

2. Produksi Padi

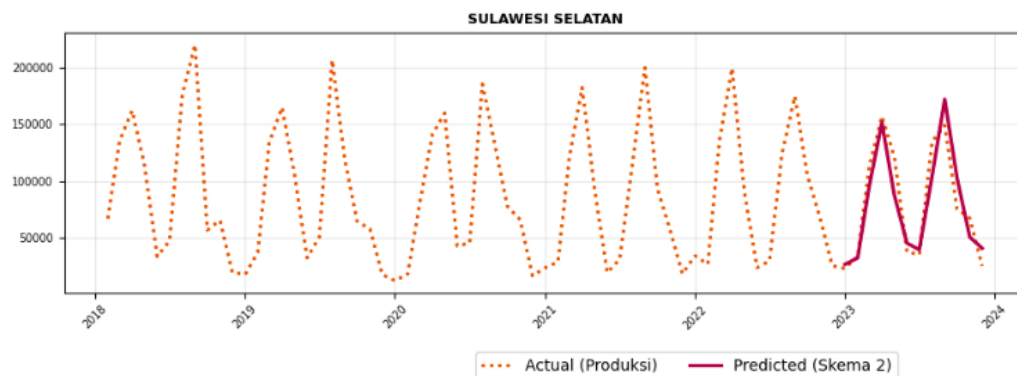
Visualisasi hasil prediksi produksi padi pada setiap provinsi ditampilkan untuk melihat pola pergerakan nilai hasil prediksi. Adapun visualisasi tersebut disajikan pada Gambar 51, 52, dan 53 sebagai berikut.



Gambar 51. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Produksi Padi Tinggi

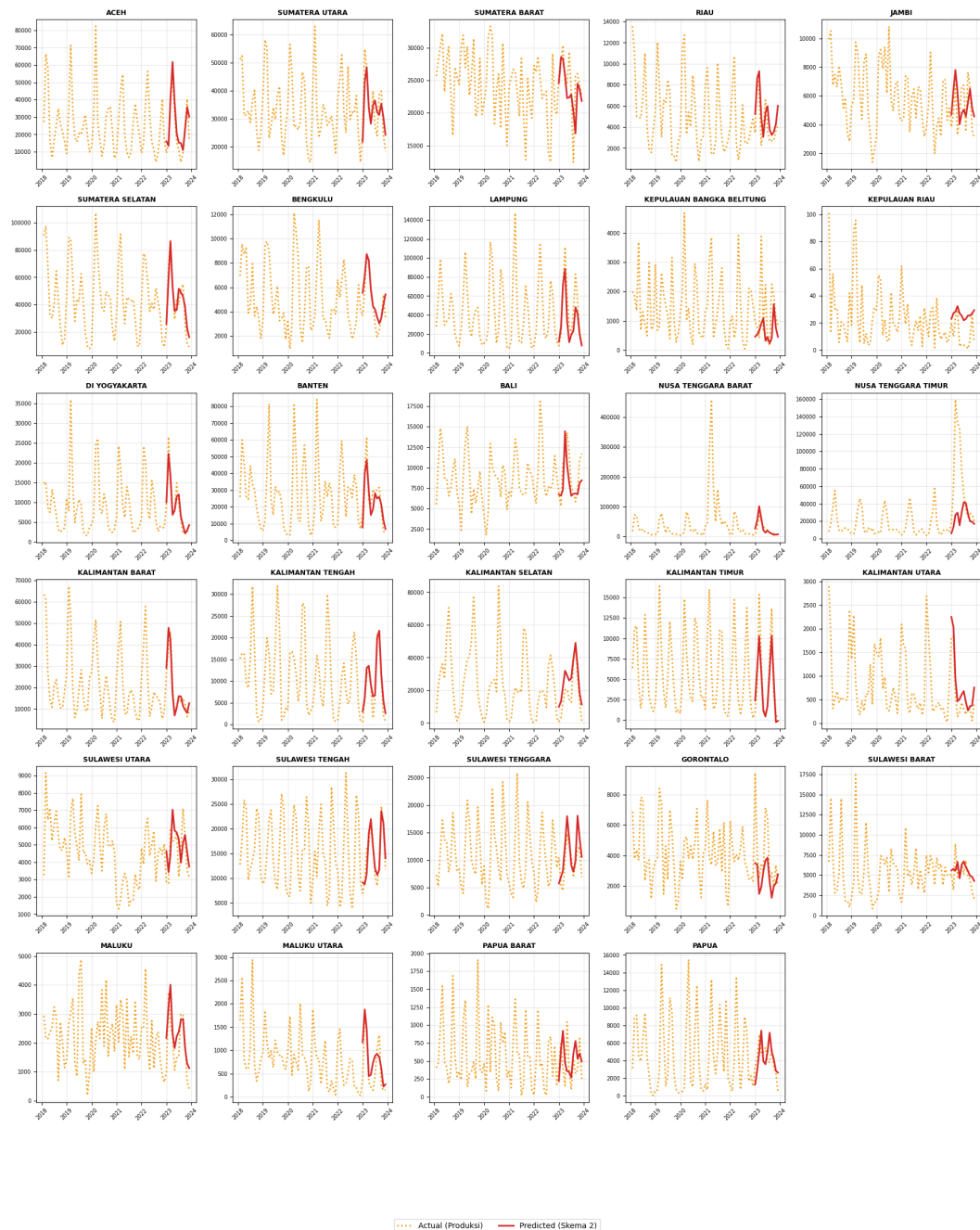
Berdasarkan Gambar 51, pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil prediksi menunjukkan bahwa nilai prediksi mampu mengikuti pola data aktual. Pada Jawa Barat, pergerakan prediksi mengikuti pola naik dan turun data aktual, termasuk saat terjadi peningkatan nilai pada awal periode, kemudian diikuti penurunan secara bertahap hingga akhir periode. Pada Jawa Tengah, hasil prediksi juga menunjukkan pola yang searah dengan data aktual, di mana terjadi kenaikan yang cukup tinggi di awal periode, kemudian menurun dan cenderung stabil pada periode berikutnya. Pada Jawa Timur, pola prediksi terlihat mengikuti pergerakan data aktual, termasuk saat terjadi lonjakan nilai yang tinggi di awal periode pengujian, kemudian diikuti penurunan yang cukup tajam

dan selanjutnya bergerak menurun secara bertahap.



Gambar 52. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Produksi Padi Sedang

Berdasarkan Gambar 52, pada kelompok produksi sedang yaitu Sulawesi Selatan menunjukkan bahwa hasil prediksi mampu mengikuti pola data aktual. Pergerakan prediksi terlihat searah dengan data aktual, dimulai dari nilai yang relatif rendah di awal periode, kemudian mengalami kenaikan hingga mencapai puncak, dan selanjutnya kembali menurun. Pola naik turun yang terjadi pada data aktual juga dapat diikuti oleh hasil prediksi dengan cukup jelas. Selain itu, perubahan nilai antar titik pada periode pengujian dapat direpresentasikan dengan baik oleh model, sehingga bentuk kurva prediksi mendekati kurva data aktual.



Gambar 53. Visualisasi Prediksi *Hybrid* PVAR–LSTM Skema 2 Kelompok Produksi Padi Sedang

Berdasarkan Gambar 53, pada kelompok produksi rendah yang mencakup berbagai provinsi seperti Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, hingga Papua, hasil prediksi menunjukkan variasi performa antar wilayah. Pada sebagian besar provinsi seperti Aceh, Kalimantan Barat, Kalimantan Tengah, Sulawesi Tengah

dan Sumatera Selatan, hasil prediksi mampu mengikuti arah pergerakan data aktual. Nilai prediksi menunjukkan pola naik dan turun yang searah dengan data aktual, meskipun terdapat perbedaan pada besaran nilainya. Sementara itu, pada provinsi seperti Nusa Tenggara Timur (NTT) terlihat adanya perbedaan yang cukup jelas antara nilai prediksi dan data aktual, terutama pada periode akhir yang menunjukkan perubahan nilai yang cukup besar. Pada wilayah lain seperti Kalimantan Timur, Sulawesi Tenggara, Maluku, dan Papua, hasil prediksi tetap mengikuti arah umum pergerakan data aktual, dengan fluktuasi yang relatif searah pada periode pengujian.

4.4 Hasil Evaluasi

Setelah diperoleh hasil prediksi dari model PVAR dan model *hybrid* PVAR–LSTM, Berikut merupakan evaluasi untuk mengetahui tingkat akurasi hasil prediksi yang dihasilkan. Evaluasi dilakukan dengan membandingkan kinerja model PVAR sebagai model dasar dengan model *hybrid* PVAR–LSTM pada Skema 1 dan Skema 2. Ukuran kesalahan prediksi yang digunakan adalah MAE, RMSE dan MAPE, baik untuk variabel produksi padi maupun luas panen. Hasil evaluasi pada setiap provinsi kemudian disajikan dalam bentuk tabel untuk melihat model mana yang memberikan hasil terbaik, sebagaimana ditunjukkan sebagai berikut.

1. Luas Panen

Hasil evaluasi pada variabel luas panen disajikan dalam bentuk tabel yang memperlihatkan nilai kesalahan prediksi antara model PVAR dan model *hybrid* PVAR–LSTM pada masing-masing provinsi, sehingga dapat diketahui model yang memberikan kinerja terbaik. Tabel 14, 15, 16 dan 17 menunjukkan hasil evaluasi setiap model.

Tabel 14. Nilai MAE Model PVAR dan *Hybrid* PVAR-LSTM (Luas Panen)

No	Provinsi	MAE PVAR	MAE <i>Hybrid</i> Skema 1 (S1)	MAE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	12861.0	6521.19	5296.19	Hybrid S2
2	Bali	2320.0	2277.53	2031.12	Hybrid S2
3	Banten	12124.0	6818.59	6352.88	Hybrid S2
4	Bengkulu	1373.0	1300.01	834.74	Hybrid S2
5	DI Yogyakarta	4750.0	2237.28	1504.82	Hybrid S2

No	Provinsi	MAE PVAR	MAE <i>Hybrid</i> Skema 1 (S1)	MAE <i>Hybrid</i> Skema 2 (S2)	Terbaik
6	DKI Jakarta	38.7	23.45	27.77	Hybrid S1
7	Gorontalo	2151.0	1766.11	1603.89	Hybrid S2
8	Jambi	1894.0	1803.66	1426.48	Hybrid S2
9	Jawa Barat	34836.0	18576.85	19475.56	Hybrid S1
10	Jawa Tengah	62330.0	27074.25	27534.23	Hybrid S1
11	Jawa Timur	56195.0	33931.79	28083.28	Hybrid S2
12	Kalimantan Barat	8649.0	3402.03	3015.41	Hybrid S2
13	Kalimantan Selatan	8624.0	8287.86	8340.46	Hybrid S1
14	Kalimantan Tengah	4928.0	2169.05	1691.72	Hybrid S2
15	Kalimantan Timur	4180.0	1686.41	1619.49	Hybrid S2
16	Kalimantan Utara	335.0	416.39	268.20	Hybrid S2
17	Kep. Bangka Belitung	1034.0	828.01	795.19	Hybrid S2
18	Kep. Riau	9.6	22.23	14.79	PVAR
19	Lampung	24332.0	19232.22	13121.82	Hybrid S2
20	Maluku	801.0	904.55	631.47	Hybrid S2
21	Maluku Utara	443.0	518.37	291.69	Hybrid S2
22	Nusa Tenggara Barat	8785.0	20134.47	9995.67	PVAR
23	Nusa Tenggara Timur	20725.0	84619.67	40913.77	PVAR
24	Papua	1449.0	2026.98	1613.59	PVAR
25	Papua Barat	354.0	377.85	265.68	Hybrid S2
26	Riau	1806.0	1172.12	1352.83	Hybrid S1
27	Sulawesi Barat	1668.0	1228.60	1368.14	Hybrid S1
28	Sulawesi Selatan	39004.0	13330.12	15579.93	Hybrid S1
29	Sulawesi Tengah	4717.0	2599.73	1615.22	Hybrid S2
30	Sulawesi Tenggara	3273.0	3467.79	2634.30	Hybrid S2
31	Sulawesi Utara	1429.0	1236.50	1070.23	Hybrid S2
32	Sumatera Barat	6288.0	3493.42	2824.17	Hybrid S2
33	Sumatera Selatan	17911.0	6327.43	7002.50	Hybrid S1
34	Sumatera Utara	9820.0	5915.39	5362.39	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 14 nilai MAE pada variabel luas panen, terlihat bahwa model *Hybrid* PVAR-LSTM secara umum mampu memberikan nilai kesalahan yang lebih kecil dibandingkan model PVAR. Hal ini menunjukkan bahwa pendekatan *hybrid* dapat meningkatkan akurasi prediksi pada sebagian besar provinsi. Secara keseluruhan, *Hybrid* Skema 2 menjadi model terbaik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, Gorontalo, Jambi, Jawa Timur, Kalimantan Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Kep. Bangka Belitung, Lampung, Maluku, Maluku Utara, Sulawesi Tengah, Sulawesi Utara, Sumatera Barat, dan Sumatera Utara. Sementara itu, *Hybrid* Skema 1 menunjukkan kinerja terbaik pada beberapa provinsi seperti Jawa Barat, Jawa Tengah, Kalimantan Selatan, Riau, Sulawesi Barat, Sulawesi Selatan, dan Sumatera Selatan. Di sisi lain, model PVAR masih menjadi yang terbaik pada beberapa provinsi seperti Kepulauan Riau, Nusa Tenggara Barat, Nusa Tenggara Timur, dan Papua. Secara umum, hasil ini menunjukkan bahwa penggunaan model *hybrid*, khususnya Skema 2, memberikan peningkatan akurasi yang signifikan dalam meminimalkan kesalahan absolut pada prediksi luas panen.

Tabel 15. Nilai RMSE Model PVAR dan *Hybrid* PVAR-LSTM (Luas Panen)

No	Provinsi	RMSE PVAR	RMSE <i>Hybrid</i> Skema 1 (S1)	RMSE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	15171.0	7475.62	6467.34	Hybrid S2
2	Bali	3093.0	2915.55	2456.52	Hybrid S2
3	Banten	15293.0	7850.20	7276.56	Hybrid S2
4	Bengkulu	1772.0	1793.63	1065.46	Hybrid S2
5	DI Yogyakarta	6903.0	3054.41	1987.70	Hybrid S2
6	DKI Jakarta	58.1	30.53	36.69	Hybrid S1
7	Gorontalo	3017.0	2520.15	2300.38	Hybrid S2
8	Jambi	2264.0	2076.28	1595.37	Hybrid S2
9	Jawa Barat	49706.0	23884.59	27252.66	Hybrid S1
10	Jawa Tengah	90483.0	33292.26	32374.63	Hybrid S2
11	Jawa Timur	86080.0	44275.97	32850.11	Hybrid S2

No	Provinsi	RMSE PVAR	RMSE <i>Hybrid</i> Skema 1 (S2)	RMSE <i>Hybrid</i> Skema 2 (S2)	Terbaik
12	Kalimantan Barat	10877.0	4395.99	3652.21	Hybrid S2
13	Kalimantan Selatan	11449.0	9420.15	9165.42	Hybrid S2
14	Kalimantan Tengah	6315.0	2474.43	2143.20	Hybrid S2
15	Kalimantan Timur	5471.0	2423.17	2136.57	Hybrid S2
16	Kalimantan Utara	479.0	470.85	290.81	Hybrid S2
17	Kep. Bangka Belitung	1555.0	1095.84	1191.89	Hybrid S1
18	Kep. Riau	12.2	24.69	17.30	PVAR
19	Lampung	30450.0	23029.97	17251.44	Hybrid S2
20	Maluku	1134.0	1032.81	749.56	Hybrid S2
21	Maluku Utara	620.0	594.62	370.88	Hybrid S2
22	Nusa Tenggara Barat	16652.0	24911.02	13916.42	Hybrid S2
23	Nusa Tenggara Timur	36359.0	111621.47	61212.48	PVAR
24	Papua	1740.0	2391.59	1837.33	PVAR
25	Papua Barat	419.0	438.96	329.08	Hybrid S2
26	Riau	2436.0	1622.29	1553.02	Hybrid S2
27	Sulawesi Barat	2381.0	1643.57	1696.05	Hybrid S1
28	Sulawesi Selatan	51336.0	18370.56	18870.12	Hybrid S1
29	Sulawesi Tengah	6253.0	3472.31	2106.01	Hybrid S2
30	Sulawesi Tenggara	3845.0	3881.36	3100.14	Hybrid S2
31	Sulawesi Utara	1808.0	1516.16	1292.12	Hybrid S2
32	Sumatera Barat	7570.0	4738.48	4113.67	Hybrid S2
33	Sumatera Selatan	21766.0	7574.81	8207.43	Hybrid S1
34	Sumatera Utara	13705.0	7218.59	6384.37	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 15 nilai RMSE, terlihat bahwa model *Hybrid* PVAR-LSTM juga menunjukkan kinerja yang lebih baik dibandingkan model PVAR pada sebagian besar provinsi. Nilai RMSE yang lebih kecil menunjukkan bahwa model *hybrid* mampu mengurangi kesalahan yang besar atau ekstrem dalam prediksi. *Hybrid* Skema 2

menjadi model yang paling dominan, dengan hasil terbaik pada sebagian besar provinsi seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, Jambi, Kalimantan Barat, Kalimantan Tengah, Lampung, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, Sumatera Utara dan sebagainya. Sementara itu, *Hybrid* Skema 1 masih unggul pada beberapa provinsi seperti Jawa Barat, Kepulauan Bangka Belitung, Sulawesi Barat, Sulawesi Selatan, dan Sumatera Selatan. Di sisi lain, model PVAR masih menjadi pilihan terbaik pada beberapa provinsi seperti Kepulauan Riau, Nusa Tenggara Timur, dan Papua. Secara keseluruhan, hasil RMSE menunjukkan bahwa model *hybrid*, khususnya Skema 2, lebih efektif dalam menangani kesalahan prediksi yang besar pada data luas panen.

Tabel 16. Nilai MAPE Model PVAR dan *Hybrid* PVAR-LSTM (Luas Panen)

No	Provinsi	MAPE PVAR	MAPE <i>Hybrid</i> Skema 1 (S1)	MAPE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	61.20	63.12	47.83	Hybrid S2
2	Bali	22.70	24.37	19.91	Hybrid S2
3	Banten	52.50	46.00	36.91	Hybrid S2
4	Bengkulu	28.00	35.84	25.25	Hybrid S2
5	DI Yogyakarta	50.70	41.99	22.96	Hybrid S2
6	DKI Jakarta	670.00	548.38	682.03	Hybrid S1
7	Gorontalo	48.20	35.91	32.86	Hybrid S2
8	Jambi	36.30	36.88	31.67	Hybrid S2
9	Jawa Barat	24.30	14.94	15.43	Hybrid S1
10	Jawa Tengah	43.40	24.45	26.08	Hybrid S1
11	Jawa Timur	31.20	26.48	23.69	Hybrid S2
12	Kalimantan Barat	46.10	19.78	18.54	Hybrid S2
13	Kalimantan Selatan	122.00	288.57	244.97	PVAR
14	Kalimantan Tengah	104.00	150.74	111.45	PVAR
15	Kalimantan Timur	121.00	63.68	57.12	Hybrid S2
16	Kalimantan Utara	87.90	158.58	121.17	PVAR
17	Kep. Bangka Belitung	121.00	97.42	56.20	Hybrid S2
18	Kep. Riau	139.00	702.91	504.37	PVAR
19	Lampung	55.20	47.35	31.28	Hybrid S2
20	Maluku	45.90	78.96	56.25	PVAR

No	Provinsi	MAPE PVAR	MAPE <i>Hybrid</i> Skema 1 (S1)	MAPE <i>Hybrid</i> Skema 2 (S2)	Terbaik
21	Maluku Utara	87.90	166.25	102.28	PVAR
22	Nusa Tenggara Barat	27.50	103.32	55.08	PVAR
23	Nusa Tenggara Timur	30.40	121.48	48.77	PVAR
24	Papua	54.30	66.45	70.58	PVAR
25	Papua Barat	86.20	129.21	80.53	Hybrid S2
26	Riau	42.50	31.54	38.54	Hybrid S1
27	Sulawesi Barat	29.00	37.27	34.67	PVAR
28	Sulawesi Selatan	53.30	21.06	21.71	Hybrid S1
29	Sulawesi Tengah	29.50	19.80	13.48	Hybrid S2
30	Sulawesi Tenggara	33.70	37.98	30.89	Hybrid S2
31	Sulawesi Utara	30.30	27.33	26.69	Hybrid S2
32	Sumatera Barat	25.80	15.80	13.37	Hybrid S2
33	Sumatera Selatan	44.40	27.34	28.87	Hybrid S1
34	Sumatera Utara	26.40	19.41	17.74	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 16 nilai MAPE, terlihat bahwa model *Hybrid* PVAR-LSTM juga mampu menurunkan persentase kesalahan pada sebagian besar provinsi. *Hybrid* Skema 2 tetap menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, DI Yogyakarta, Gorontalo, Jambi, Kalimantan Barat, Lampung, Sulawesi Tengah, Sulawesi Utara, Sumatera Barat, Sumatera Utara dan sebagainya. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti DKI Jakarta, Jawa Barat, Jawa Tengah, Riau, Sulawesi Selatan, dan Sumatera Selatan. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku, Maluku Utara, Nusa Tenggara Barat, Nusa Tenggara Timur, Papua, dan Sulawesi Barat. Hal ini menunjukkan bahwa pada beberapa wilayah, kesalahan relatif model *hybrid* masih cukup besar, terutama pada data dengan skala kecil. Secara keseluruhan, hasil MAPE menunjukkan

bahwa meskipun model *hybrid* mampu meningkatkan akurasi, performanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Tabel 17. Accuracy Model PVAR dan Hybrid PVAR-LSTM (Luas Panen)

No	Provinsi	Accuracy PVAR	Accuracy Hybrid Skema 1 (S1)	Accuracy Hybrid Skema 2 (S1)	Terbaik
1	Aceh	38.80	36.88	52.17	Hybrid S2
2	Bali	77.30	75.63	80.09	Hybrid S2
3	Banten	47.50	54.00	63.09	Hybrid S2
4	Bengkulu	72.00	64.16	74.75	Hybrid S2
5	DI Yogyakarta	49.30	58.01	77.04	Hybrid S2
6	DKI Jakarta	-570.00	-448.38	-582.03	Hybrid S1
7	Gorontalo	51.80	64.09	67.14	Hybrid S2
8	Jambi	63.70	63.12	68.33	Hybrid S2
9	Jawa Barat	75.70	85.06	84.57	Hybrid S1
10	Jawa Tengah	56.60	75.55	73.92	Hybrid S1
11	Jawa Timur	68.80	73.52	76.31	Hybrid S2
12	Kalimantan Barat	53.90	80.22	81.46	Hybrid S2
13	Kalimantan Selatan	-22.20	-188.57	-144.97	PVAR
14	Kalimantan Tengah	-3.71	-50.74	-11.45	PVAR
15	Kalimantan Timur	-20.50	36.32	42.88	Hybrid S2
16	Kalimantan Utara	12.10	-58.58	-21.17	PVAR
17	Kep. Bangka Belitung	-20.60	2.58	43.80	Hybrid S2
18	Kep. Riau	-38.50	-602.91	-404.37	PVAR
19	Lampung	44.80	52.65	68.72	Hybrid S2
20	Maluku	54.10	21.04	43.75	PVAR
21	Maluku Utara	12.10	-66.25	-2.28	PVAR
22	Nusa Tenggara Barat	72.50	-3.32	44.92	PVAR
23	Nusa Tenggara Timur	69.60	-21.48	51.23	PVAR
24	Papua	45.70	33.55	29.42	PVAR
25	Papua Barat	13.80	-29.21	19.47	Hybrid S2
26	Riau	57.50	68.46	61.46	Hybrid S1
27	Sulawesi Barat	71.00	62.73	65.33	PVAR

No	Provinsi	<i>Accuracy</i> PVAR	<i>Accuracy</i> <i>Hybrid</i> Skema 1 (S1)	<i>Accuracy</i> <i>Hybrid</i> Skema 2 (S2)	Terbaik
28	Sulawesi Selatan	46.70	78.94	78.29	Hybrid S1
29	Sulawesi Tengah	70.50	80.20	86.52	Hybrid S2
30	Sulawesi Tenggara	66.30	62.02	69.11	Hybrid S2
31	Sulawesi Utara	69.70	72.67	73.31	Hybrid S2
32	Sumatera Barat	74.20	84.20	86.63	Hybrid S2
33	Sumatera Selatan	55.60	72.66	71.13	Hybrid S1
34	Sumatera Utara	73.60	80.59	82.26	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 17 nilai *accuracy*, terlihat bahwa model *Hybrid* PVAR-LSTM mampu meningkatkan tingkat ketepatan prediksi pada sebagian besar provinsi dibandingkan model PVAR. *Hybrid* Skema 2 menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, Gorontalo, Jambi, Kalimantan Barat, Kalimantan Timur, Kepulauan Bangka Belitung, Sulawesi Tengah, Sulawesi Utara, Sumatera Barat, Sumatera Utara dan sebagainya. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti DKI Jakarta, Jawa Barat, dan Jawa Tengah. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kalimantan Selatan, Kalimantan Tengah, dan Kalimantan Utara. Hal ini menunjukkan bahwa pada beberapa wilayah, model dasar masih mampu memberikan hasil lebih baik dibandingkan model *hybrid*. Secara keseluruhan, hasil *accuracy* menunjukkan bahwa model *Hybrid* PVAR-LSTM, khususnya Skema 2, cenderung memberikan performa lebih baik dalam meningkatkan ketepatan prediksi luas panen, meskipun kinerjanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Untuk memperoleh gambaran performa model yang lebih komprehensif, evaluasi selanjutnya difokuskan pada nilai *Mean Absolute Percentage Error* (MAPE) dari model terbaik pada masing-masing provinsi. Penggunaan MAPE bertujuan untuk mengukur tingkat kesalahan relatif terhadap nilai aktual sehingga

hasil evaluasi tetap dapat diinterpretasikan pada provinsi dengan skala produksi yang berbeda. Pengelompokan wilayah dalam analisis ini mengikuti klasifikasi pada Subbab 2.17 berdasarkan Sofhya (2023), yaitu kelompok tinggi, sedang, dan rendah. Nilai MAPE kemudian diklasifikasikan ke dalam empat kategori berdasarkan (Nabillah & Ranggadara, 2020), yaitu MAPE 20%–50% termasuk kategori sangat baik, MAPE 10%–20% termasuk kategori baik, MAPE 20%–50% termasuk kategori cukup, dan MAPE $> 50\%$ termasuk kategori kurang. Klasifikasi tersebut digunakan untuk mempermudah interpretasi hasil prediksi yang disajikan pada Tabel 18.

Tabel 18. Evaluasi MAPE Berdasarkan Kelompok Skala Luas Panen

Kelompok Luas Panen	Provinsi	Rentang MAPE	Interpretasi
Tinggi	Jawa Barat Jawa Tengah Jawa Timur	10%–20% (Baik)	Model menunjukkan akurasi tinggi dan stabil pada wilayah dengan luas panen terbesar.
Sedang	Sulawesi Selatan	20%–50% (Cukup)	Model cukup mampu mengikuti dinamika luas panen pada wilayah berskala menengah.
Rendah	Bali Kalimantan Barat Sulawesi Tengah Sumatera Barat Sumatera Utara	10%–20% (Baik)	Meskipun skala luas panen relatif kecil, pola historis cukup stabil sehingga model tetap akurat.
	Aceh Banten Bengkulu DI Yogyakarta Gorontalo Jambi Jawa Tengah Jawa Timur Kalimantan Timur Lampung Riau Sulawesi Barat Sulawesi Tenggara Sulawesi Utara Sumatera Selatan	20%–50% (Cukup)	Model masih mampu mengikuti pola perubahan luas panen, namun variasi relatif menyebabkan tingkat kesalahan meningkat.

Kelompok Luas Panen	Provinsi	Rentang MAPE	Interpretasi
	DKI Jakarta Kalimantan Selatan Kalimantan Tengah Kalimantan Utara Kep. Bangka Belitung Kep. Riau Maluku Maluku Utara Nusa Tenggara Barat Nusa Tenggara Timur Papua Papua Barat	> 50% (Kurang)	Nilai aktual yang kecil dan fluktuatif menyebabkan MAPE tinggi sehingga akurasi model tergolong rendah.

Berdasarkan Tabel 18, terlihat bahwa pada kelompok luas panen tinggi, sebagian besar provinsi berada pada kategori baik hingga cukup. Hal ini menunjukkan bahwa model mampu menangkap dinamika luas panen secara relatif stabil pada wilayah dengan kontribusi terbesar terhadap total luas panen nasional. Pada kelompok luas panen sedang, nilai MAPE model juga berada pada kategori yang masih dapat diterima, yang menunjukkan bahwa pola perubahan luas panen pada wilayah ini cukup konsisten dan dapat dimodelkan dengan baik. Sementara itu, pada kelompok luas panen rendah, nilai MAPE menunjukkan variasi yang lebih beragam. Sebagian provinsi masih berada pada kategori baik dan cukup, tetapi beberapa provinsi masuk dalam kategori kurang akibat nilai aktual yang relatif kecil dan fluktuatif. Pada kondisi tersebut, kesalahan absolut yang tidak terlalu besar dapat menghasilkan nilai MAPE yang tinggi karena sifatnya yang sensitif terhadap skala data kecil. Secara umum, hasil ini menunjukkan bahwa performa model lebih stabil pada wilayah dengan skala luas panen besar dan sedang.

Secara Keseluruhan, berdasarkan hasil evaluasi pada Tabel 14–17, terlihat bahwa nilai MAE dan RMSE secara umum menunjukkan kecenderungan hubungan yang searah, di mana ketika nilai MAE menurun maka nilai RMSE juga cenderung menurun, begitu pula sebaliknya. Hal ini terjadi karena kedua metrik tersebut sama-sama digunakan untuk mengukur besar kesalahan prediksi model berdasarkan selisih antara nilai aktual dan hasil prediksi. Pada beberapa provinsi seperti Jawa Timur, Aceh, Lampung, Jawa Barat, dan Sulawesi Selatan, penurunan nilai MAE diikuti oleh penurunan nilai RMSE, yang menunjukkan bahwa semakin

kecil rata-rata kesalahan prediksi maka semakin kecil pula error prediksi secara keseluruhan. Selain itu, nilai RMSE cenderung lebih sensitif terhadap error yang besar atau ekstrem karena menggunakan kuadrat error dalam perhitungannya. Oleh sebab itu, apabila terdapat beberapa kesalahan prediksi yang cukup tinggi pada periode tertentu, maka nilai RMSE akan meningkat lebih besar dibandingkan MAE. Dengan demikian, hubungan antara MAE dan RMSE menunjukkan bahwa kedua metrik tersebut saling berkaitan dalam menggambarkan tingkat kesalahan prediksi model.

Sementara itu, nilai *accuracy* dan MAPE menunjukkan hubungan yang berbanding terbalik, di mana ketika nilai MAPE menurun maka nilai *accuracy* cenderung meningkat, begitu pula sebaliknya. Hal ini karena MAPE digunakan untuk mengukur persentase kesalahan prediksi, sedangkan *accuracy* menggambarkan tingkat ketepatan hasil prediksi model. Pada beberapa provinsi seperti Sumatera Barat, Sulawesi Tengah, Jawa Timur, dan Bali, penurunan nilai MAPE diikuti oleh peningkatan nilai *accuracy*, yang menunjukkan bahwa semakin kecil persentase kesalahan prediksi maka tingkat ketepatan prediksi model menjadi semakin baik. Sebaliknya, pada provinsi seperti DKI Jakarta dan Kepulauan Riau, nilai MAPE yang tinggi diikuti oleh nilai *accuracy* yang rendah. Dengan demikian, hubungan antara MAPE dan *accuracy* menunjukkan bahwa kedua metrik tersebut memiliki keterkaitan dalam menggambarkan kualitas hasil prediksi model.

2. Produksi Padi

Pada variabel produksi padi, evaluasi dilakukan untuk mengetahui kinerja model PVAR dengan model *hybrid* PVAR-LSTM dalam menghasilkan prediksi yang lebih akurat. Kinerja disajikan berdasarkan nilai kesalahan prediksi pada masing-masing provinsi untuk melihat model yang memberikan hasil terbaik. Tabel 19, 20, 21 dan 22 menunjukkan hasil evaluasi setiap model.

Tabel 19. Nilai MAE Model PVAR dan *Hybrid* PVAR-LSTM (Produksi Padi)

No	Provinsi	MAE PVAR	MAE <i>Hybrid</i> Skema 1 (S1)	MAE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	74063.0	33685.91	22470.15	Hybrid S2
2	Bali	15097.0	13341.91	12654.50	Hybrid S2
3	Banten	65368.0	44084.06	35959.55	Hybrid S2
4	Bengkulu	7173.0	5779.72	3529.34	Hybrid S2

No	Provinsi	MAE PVAR	MAE <i>Hybrid</i> Skema 1 (S1)	MAE <i>Hybrid</i> Skema 2 (S2)	Terbaik
5	DI Yogyakarta	22028.0	11266.77	7042.63	Hybrid S2
6	DKI Jakarta	188.0	136.51	134.51	Hybrid S2
7	Gorontalo	11110.0	8646.32	8074.24	Hybrid S2
8	Jambi	8213.0	8065.55	6418.89	Hybrid S2
9	Jawa Barat	197876.0	114543.29	126940.75	Hybrid S1
10	Jawa Tengah	346393.0	131900.39	183300.57	Hybrid S1
11	Jawa Timur	316556.0	144882.25	157439.83	Hybrid S1
12	Kalimantan Barat	25987.0	8511.23	8046.98	Hybrid S2
13	Kalimantan Selatan	31075.0	38153.84	35936.32	PVAR
14	Kalimantan Tengah	15989.0	8024.37	7280.54	Hybrid S2
15	Kalimantan Timur	15924.0	5451.88	5472.44	Hybrid S1
16	Kalimantan Utara	1235.0	1172.42	719.86	Hybrid S2
17	Kep. Bangka Belitung	5023.0	3864.36	4154.83	Hybrid S1
18	Kep. Riau	29.6	88.27	55.91	PVAR
19	Lampung	131731.0	99692.32	69318.10	Hybrid S2
20	Maluku	3154.0	4553.89	2818.49	Hybrid S2
21	Maluku Utara	1701.0	1887.13	1097.87	Hybrid S2
22	Nusa Tenggara Barat	48648.0	39347.15	27357.65	Hybrid S2
23	Nusa Tenggara Timur	105466.0	427220.62	196197.46	PVAR
24	Papua	5857.0	7653.87	6549.58	PVAR
25	Papua Barat	1633.0	1611.31	1259.82	Hybrid S2
26	Riau	7442.0	4078.78	4703.42	Hybrid S1
27	Sulawesi Barat	8302.0	7613.32	7841.59	Hybrid S1
28	Sulawesi Selatan	198606.0	56726.42	74545.73	Hybrid S1
29	Sulawesi Tengah	21362.0	11623.98	5709.19	Hybrid S2
30	Sulawesi Tenggara	13358.0	14197.74	10945.19	PVAR
31	Sulawesi Utara	5935.0	5576.54	4511.63	Hybrid S2
32	Sumatera Barat	30884.0	18130.74	13265.53	Hybrid S2
33	Sumatera Selatan	99543.0	39661.69	41409.69	Hybrid S1
34	Sumatera Utara	50331.0	28271.78	27796.73	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 19 nilai MAE, terlihat bahwa model *Hybrid* PVAR-LSTM mampu menurunkan nilai kesalahan prediksi pada sebagian besar provinsi dibandingkan model PVAR. *Hybrid* Skema 2 menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, DKI Jakarta, Gorontalo, Jambi, Kalimantan Barat, Kalimantan Tengah, Kalimantan Utara, Lampung, Maluku, Maluku Utara, Nusa Tenggara Barat, Papua Barat, Sulawesi Tengah, Sulawesi Utara, Sumatera Barat, Sumatera Utara dan sebagainya. Hal ini menunjukkan bahwa kombinasi model pada Skema 2 mampu memberikan hasil prediksi dengan tingkat kesalahan yang lebih rendah pada berbagai wilayah. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Timur, Kepulauan Bangka Belitung, Riau, Sulawesi Barat, Sulawesi Selatan, dan Sumatera Selatan. Hal ini mengindikasikan bahwa pada beberapa provinsi dengan karakteristik data tertentu, pendekatan pada Skema 1 lebih efektif dalam mengurangi kesalahan prediksi dibandingkan Skema 2. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kalimantan Selatan, Kepulauan Riau, Nusa Tenggara Timur, Papua, dan Sulawesi Tenggara. Secara keseluruhan, hasil MAE menunjukkan bahwa model *Hybrid* PVAR-LSTM, khususnya Skema 2, cenderung memberikan performa yang lebih baik dalam menurunkan tingkat kesalahan prediksi produksi padi, meskipun kinerjanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Tabel 20. Nilai RMSE Model PVAR dan Hybrid PVAR-LSTM (Produksi Padi)

No	Provinsi	RMSE PVAR	RMSE <i>Hybrid</i> Skema 1 (S1)	RMSE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	89097.0	38694.59	29300.30	Hybrid S2
2	Bali	19948.0	18106.04	15575.27	Hybrid S2
3	Banten	84914.0	52535.12	42817.10	Hybrid S2
4	Bengkulu	9562.0	7517.73	4380.55	Hybrid S2
5	DI Yogyakarta	33166.0	13930.74	11716.47	Hybrid S2
6	DKI Jakarta	281.0	170.43	166.30	Hybrid S2
7	Gorontalo	16294.0	13032.39	12419.33	Hybrid S2

No	Provinsi	RMSE PVAR	RMSE Hybrid Skema 1 (S1)	RMSE Hybrid Skema 2 (S2)	Terbaik
8	Jambi	9487.0	9494.78	6963.85	Hybrid S2
9	Jawa Barat	291357.0	147214.28	163288.40	Hybrid S1
10	Jawa Tengah	500284.0	176940.05	223539.29	Hybrid S1
11	Jawa Timur	491631.0	185783.62	183298.46	Hybrid S2
12	Kalimantan Barat	32158.0	9762.30	9210.80	Hybrid S2
13	Kalimantan Selatan	41079.0	42727.42	38898.99	Hybrid S2
14	Kalimantan Tengah	20978.0	9269.26	8552.50	Hybrid S2
15	Kalimantan Timur	20848.0	8054.60	7750.06	Hybrid S2
16	Kalimantan Utara	1717.0	1341.05	843.65	Hybrid S2
17	Kep. Bangka Belitung	7787.0	5769.22	6314.96	Hybrid S1
18	Kep. Riau	39.7	96.32	61.57	PVAR
19	Lampung	169051.0	120569.94	90458.40	Hybrid S2
20	Maluku	4360.0	5054.12	3400.78	Hybrid S2
21	Maluku Utara	2418.0	2178.06	1315.40	Hybrid S2
22	Nusa Tenggara Barat	93326.0	55080.84	41292.41	Hybrid S2
23	Nusa Tenggara Timur	197515.0	556566.02	299933.43	PVAR
24	Papua	6927.0	9388.25	7608.17	PVAR
25	Papua Barat	1967.0	1829.68	1543.06	Hybrid S2
26	Riau	9290.0	6042.76	5651.90	Hybrid S2
27	Sulawesi Barat	12220.0	9747.38	8976.88	Hybrid S2
28	Sulawesi Selatan	259307.0	80941.58	86381.83	Hybrid S1
29	Sulawesi Tengah	27974.0	15863.72	8795.34	Hybrid S2
30	Sulawesi Tenggara	15656.0	15990.39	12950.40	Hybrid S2
31	Sulawesi Utara	7309.0	6172.18	5375.04	Hybrid S2
32	Sumatera Barat	36387.0	22193.08	19019.44	Hybrid S2
33	Sumatera Selatan	124500.0	49713.50	43810.24	Hybrid S2
34	Sumatera Utara	69700.0	35503.86	30722.59	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 20, terlihat bahwa model *Hybrid* PVAR-LSTM mampu menurunkan nilai kesalahan prediksi pada sebagian besar provinsi dibandingkan model PVAR. *Hybrid* Skema 2 menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, DKI Jakarta, Gorontalo, Jambi, Jawa Timur, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Lampung, Maluku, Maluku Utara, Nusa Tenggara Barat, Papua Barat, Riau, Sulawesi Barat, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, Sumatera Selatan, dan Sumatera Utara. Hal ini menunjukkan bahwa kombinasi model pada Skema 2 mampu memberikan hasil prediksi dengan tingkat kesalahan yang lebih rendah secara umum. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti Jawa Barat, Jawa Tengah, Kepulauan Bangka Belitung, dan Sulawesi Selatan. Hal ini mengindikasikan bahwa pada beberapa provinsi dengan karakteristik data tertentu, pendekatan pada Skema 1 lebih efektif dalam mengurangi kesalahan prediksi dibandingkan Skema 2. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kepulauan Riau, Nusa Tenggara Timur, dan Papua. Secara keseluruhan, hasil RMSE menunjukkan bahwa model *Hybrid* PVAR-LSTM, khususnya Skema 2, cenderung memberikan performa yang lebih baik dalam menurunkan tingkat kesalahan prediksi produksi padi, meskipun kinerjanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Tabel 21. Nilai MAPE Model PVAR dan *Hybrid* PVAR-LSTM (Produksi Padi)

No	Provinsi	MAPE PVAR	MAPE <i>Hybrid</i> Skema 1 (S1)	MAPE <i>Hybrid</i> Skema 2 (S2)	Terbaik
1	Aceh	62.40	62.99	38.19	Hybrid S2
2	Bali	24.10	22.65	20.04	Hybrid S2
3	Banten	51.30	51.17	34.96	Hybrid S2
4	Bengkulu	29.80	33.84	21.29	Hybrid S2
5	DI Yogyakarta	47.40	37.13	20.48	Hybrid S2
6	DKI Jakarta	627.00	534.47	670.51	Hybrid S1
7	Gorontalo	49.20	32.63	32.86	Hybrid S1
8	Jambi	35.50	38.51	31.75	Hybrid S2
9	Jawa Barat	24.00	16.32	16.67	Hybrid S1
10	Jawa Tengah	42.30	19.25	24.87	Hybrid S1

No	Provinsi	MAPE PVAR	MAPE <i>Hybrid</i> Skema 1 (S1)	MAPE <i>Hybrid</i> Skema 2 (S2)	Terbaik
11	Jawa Timur	30.70	19.19	20.59	Hybrid S1
12	Kalimantan Barat	46.90	21.27	17.77	Hybrid S2
13	Kalimantan Selatan	111.00	308.01	236.11	PVAR
14	Kalimantan Tengah	108.00	169.37	122.49	PVAR
15	Kalimantan Timur	122.00	51.28	52.22	Hybrid S1
16	Kalimantan Utara	91.70	144.52	111.38	PVAR
17	Kep. Bangka Belitung	145.00	101.87	85.15	Hybrid S2
18	Kep. Riau	151.00	954.98	617.42	PVAR
19	Lampung	55.00	48.02	29.11	Hybrid S2
20	Maluku	49.20	109.14	71.21	PVAR
21	Maluku Utara	97.20	182.37	103.36	PVAR
22	Nusa Tenggara Barat	27.40	34.96	21.41	Hybrid S2
23	Nusa Tenggara Timur	29.50	125.24	48.86	PVAR
24	Papua	55.00	66.54	65.42	PVAR
25	Papua Barat	96.30	130.63	88.54	Hybrid S2
26	Riau	43.20	27.43	32.43	Hybrid S1
27	Sulawesi Barat	28.70	49.37	42.88	PVAR
28	Sulawesi Selatan	52.40	17.06	22.33	Hybrid S1
29	Sulawesi Tengah	28.90	19.17	9.73	Hybrid S2
30	Sulawesi Tenggara	33.30	36.05	29.02	Hybrid S2
31	Sulawesi Utara	30.20	33.35	28.21	Hybrid S2
32	Sumatera Barat	25.80	16.78	12.95	Hybrid S2
33	Sumatera Selatan	44.70	20.22	24.61	Hybrid S1
34	Sumatera Utara	26.60	18.60	17.93	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 21, terlihat bahwa model *Hybrid* PVAR-LSTM juga mampu menurunkan persentase kesalahan pada sebagian besar provinsi dibandingkan

model PVAR. *Hybrid* Skema 2 menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, Jambi, Kalimantan Barat, Kepulauan Bangka Belitung, Lampung, Nusa Tenggara Barat, Papua Barat, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, dan Sumatera Utara. Hal ini menunjukkan bahwa kombinasi model pada Skema 2 mampu memberikan hasil prediksi dengan tingkat kesalahan relatif yang lebih rendah pada berbagai wilayah. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti DKI Jakarta, Gorontalo, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Timur, Riau, Sulawesi Selatan, dan Sumatera Selatan. Hal ini mengindikasikan bahwa pada beberapa provinsi dengan karakteristik data tertentu, pendekatan pada Skema 1 lebih efektif dalam mengurangi kesalahan relatif dibandingkan Skema 2. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku, Maluku Utara, Nusa Tenggara Timur, Papua, dan Sulawesi Barat. Secara keseluruhan, hasil MAPE menunjukkan bahwa model *Hybrid* PVAR-LSTM, khususnya Skema 2, cenderung memberikan performa yang lebih baik dalam menurunkan persentase kesalahan prediksi produksi padi, meskipun kinerjanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Tabel 22. Accuracy Model PVAR dan Hybrid PVAR-LSTM (Produksi Padi)

No	Provinsi	Accuracy PVAR	Accuracy Hybrid Skema 1 (S1)	Accuracy Hybrid Skema 2 (S2)	Terbaik
1	Aceh	37.60	37.01	61.81	Hybrid S2
2	Bali	75.90	77.35	79.96	Hybrid S2
3	Banten	48.70	48.83	65.04	Hybrid S2
4	Bengkulu	70.20	66.16	78.71	Hybrid S2
5	DI Yogyakarta	52.60	62.87	79.52	Hybrid S2
6	DKI Jakarta	-527.00	-434.47	-570.51	Hybrid S1
7	Gorontalo	50.80	67.37	67.14	Hybrid S1
8	Jambi	64.50	61.49	68.25	Hybrid S2
9	Jawa Barat	76.00	83.68	83.33	Hybrid S1
10	Jawa Tengah	57.70	80.75	75.13	Hybrid S1
11	Jawa Timur	69.30	80.81	79.41	Hybrid S1
12	Kalimantan Barat	53.10	78.73	82.23	Hybrid S2

No	Provinsi	<i>Accuracy</i> PVAR	<i>Accuracy</i> <i>Hybrid</i> Skema 1 (S1)	<i>Accuracy</i> <i>Hybrid</i> Skema 2 (S2)	Terbaik
13	Kalimantan Selatan	-11.00	-208.01	-136.11	PVAR
14	Kalimantan Tengah	-8.00	-69.37	-22.49	PVAR
15	Kalimantan Timur	-22.00	48.72	47.78	Hybrid S1
16	Kalimantan Utara	8.30	-44.52	-11.38	PVAR
17	Kep. Bangka Belitung	-45.00	-1.87	14.85	Hybrid S2
18	Kep. Riau	-51.00	-854.98	-517.42	PVAR
19	Lampung	45.00	51.98	70.89	Hybrid S2
20	Maluku	50.80	-9.14	28.79	PVAR
21	Maluku Utara	2.80	-82.37	-3.36	PVAR
22	Nusa Tenggara Barat	72.60	65.04	78.59	Hybrid S2
23	Nusa Tenggara Timur	70.50	-25.24	51.14	PVAR
24	Papua	45.00	33.46	34.58	PVAR
25	Papua Barat	3.70	-30.63	11.46	Hybrid S2
26	Riau	56.80	72.57	67.57	Hybrid S1
27	Sulawesi Barat	71.30	50.63	57.12	PVAR
28	Sulawesi Selatan	47.60	82.94	77.67	Hybrid S1
29	Sulawesi Tengah	71.10	80.83	90.27	Hybrid S2
30	Sulawesi Tenggara	66.70	63.95	70.98	Hybrid S2
31	Sulawesi Utara	69.80	66.65	71.79	Hybrid S2
32	Sumatera Barat	74.20	83.22	87.05	Hybrid S2
33	Sumatera Selatan	55.30	79.78	75.39	Hybrid S1
34	Sumatera Utara	73.40	81.40	82.07	Hybrid S2

Keterangan:

Skema 1 : Skema EP

Skema 2 : Skema E

Berdasarkan Tabel 22, terlihat bahwa model *Hybrid* PVAR-LSTM mampu meningkatkan tingkat ketepatan prediksi pada sebagian besar provinsi dibandingkan model PVAR. *Hybrid* Skema 2 menjadi model yang paling baik pada banyak provinsi, seperti Aceh, Bali, Banten, Bengkulu, DI Yogyakarta, Jambi, Kalimantan

Barat, Lampung, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara, Sumatera Barat, dan Sumatera Utara. Hal ini menunjukkan bahwa kombinasi model pada Skema 2 mampu menangkap pola data produksi padi dengan lebih baik pada berbagai karakteristik wilayah. Sementara itu, *Hybrid* Skema 1 menunjukkan hasil terbaik pada beberapa provinsi seperti DKI Jakarta, Gorontalo, Jawa Barat, Jawa Tengah, Jawa Timur, Riau, Sulawesi Selatan, dan Sumatera Selatan. Hal ini mengindikasikan bahwa pada beberapa provinsi dengan pola data tertentu, pendekatan pada Skema 1 lebih mampu menyesuaikan perubahan data dibandingkan Skema 2. Namun demikian, model PVAR masih menjadi yang terbaik pada sejumlah provinsi seperti Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku, Maluku Utara, Nusa Tenggara Timur, Papua, dan Sulawesi Barat. Secara keseluruhan, hasil *accuracy* menunjukkan bahwa model *Hybrid* PVAR-LSTM, khususnya Skema 2, cenderung memberikan performa yang lebih baik dalam meningkatkan ketepatan prediksi produksi padi, meskipun kinerjanya tetap dipengaruhi oleh karakteristik data masing-masing provinsi.

Untuk memperoleh gambaran performa model yang lebih komprehensif, evaluasi selanjutnya difokuskan pada nilai *Mean Absolute Percentage Error* (MAPE) dari model terbaik pada masing-masing provinsi. Penggunaan MAPE bertujuan untuk mengukur tingkat kesalahan relatif terhadap nilai aktual sehingga hasil evaluasi tetap dapat diinterpretasikan pada provinsi dengan skala produksi yang berbeda. Pengelompokan wilayah dalam analisis ini mengikuti klasifikasi pada Subbab 2.17 berdasarkan Sofhya (2023), yaitu kelompok produksi tinggi, sedang, dan rendah. Nilai MAPE kemudian diklasifikasikan ke dalam empat kategori berdasarkan (Nabillah & Ranggadara, 2020), yaitu MAPE 20%–50% termasuk kategori sangat baik, MAPE 10%–20% termasuk kategori baik, MAPE 20%–50% termasuk kategori cukup, dan MAPE > 50% termasuk kategori kurang. Klasifikasi tersebut digunakan untuk mempermudah interpretasi hasil prediksi yang disajikan pada Tabel 18.

Tabel 23. Evaluasi MAPE Berdasarkan Kelompok Skala Produksi

Kelompok Produksi Padi	Provinsi	Rentang MAPE	Interpretasi
Tinggi	Jawa Barat Jawa Tengah Jawa Timur	10%–20% (Baik)	Model menunjukkan akurasi yang baik dan konsisten pada sentra produksi padi.
Sedang	Sulawesi Selatan	10%–20% (Baik)	Akurasi baik dan stabil dalam menangkap dinamika produksi.
Rendah	Sulawesi Tengah Sumatera Barat Sumatera Utara Kalimantan Barat	10%–20% (Baik)	Meskipun termasuk wilayah dengan skala produksi kecil, pola data relatif stabil dan mirip dengan provinsi besar dan sedang, sehingga model tetap memberikan akurasi yang baik.
	Aceh Bali Banten Bengkulu DI Yogyakarta Gorontalo Jambi Lampung Nusa Tenggara Barat Riau Sulawesi Tenggara Sulawesi Utara Sumatera Selatan	20% – 50% (Cukup)	Model masih mampu mengikuti pola produksi, namun variasi relatif yang lebih tinggi menyebabkan tingkat kesalahan meningkat.
	DKI Jakarta Kalimantan Selatan Kalimantan Tengah Kalimantan Timur Kalimantan Utara Kep. Bangka Belitung Kep. Riau Maluku Utara Papua Papua Barat	> 50% (Kurang)	Nilai aktual yang kecil dan fluktuatif menyebabkan MAPE menjadi sangat tinggi sehingga akurasi model tergolong rendah.

Berdasarkan hasil pengelompokkan tersebut, terlihat bahwa seluruh provinsi pada kelompok produksi tinggi dan sedang berada pada kategori baik (MAPE 10%–20%). Hal ini menunjukkan bahwa model mampu menangkap dinamika produksi secara lebih stabil pada wilayah dengan kontribusi produksi besar. Volume data yang relatif tinggi dan pola historis yang lebih konsisten memungkinkan model mempelajari tren dan pola musiman dengan lebih optimal. Pada kelompok produksi rendah, nilai MAPE menunjukkan variasi yang lebih beragam. Sebagian provinsi masih berada pada kategori baik, yang mengindikasikan bahwa meskipun skala produksinya kecil, pola historisnya relatif stabil. Namun demikian, sebagian besar provinsi berada pada kategori cukup, yang menunjukkan adanya variasi relatif yang lebih tinggi dalam data produksi. Selain itu, beberapa provinsi masuk dalam kategori kurang (MAPE > 50%), yang umumnya disebabkan oleh nilai aktual yang kecil dan fluktuatif. Pada kondisi tersebut, kesalahan absolut yang tidak terlalu besar dapat menghasilkan persentase kesalahan yang tinggi, sehingga nilai MAPE menjadi lebih sensitif. Secara umum, hasil ini menunjukkan bahwa performa model lebih konsisten pada wilayah dengan skala produksi besar dan sedang, sedangkan pada wilayah berskala kecil, hasil prediksi lebih dipengaruhi oleh karakteristik dan volatilitas data masing-masing provinsi.

Secara Keseluruhan, berdasarkan hasil evaluasi pada Tabel 19–22, terlihat bahwa nilai MAE dan RMSE secara umum menunjukkan kecenderungan hubungan yang searah, di mana ketika nilai MAE menurun maka nilai RMSE juga cenderung menurun, begitu pula sebaliknya. Hal ini terjadi karena kedua metrik tersebut sama-sama digunakan untuk mengukur besar kesalahan prediksi berdasarkan selisih antara nilai aktual dan hasil prediksi model. Pada beberapa provinsi seperti Jawa Timur, Aceh, Lampung, Jawa Barat, dan Sulawesi Selatan, penurunan nilai MAE diikuti oleh penurunan nilai RMSE, yang menunjukkan bahwa semakin kecil rata-rata kesalahan prediksi maka semakin kecil pula error prediksi secara keseluruhan. Selain itu, nilai RMSE cenderung lebih sensitif terhadap error yang besar atau ekstrem karena menggunakan kuadrat error dalam perhitungannya. Oleh sebab itu, apabila terdapat beberapa kesalahan prediksi yang cukup tinggi pada periode tertentu, maka nilai RMSE akan meningkat lebih besar dibandingkan MAE. Dengan demikian, hubungan antara MAE dan RMSE menunjukkan bahwa kedua metrik tersebut memiliki keterkaitan dalam menggambarkan tingkat kesalahan prediksi model.

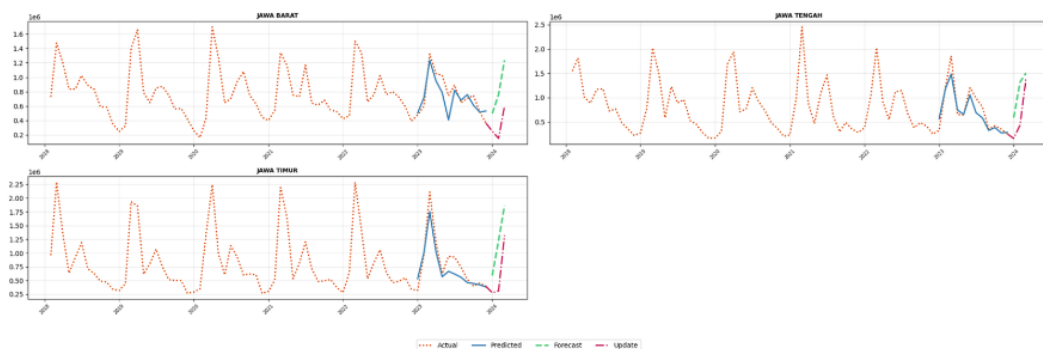
Sementara itu, nilai *accuracy* dan MAPE menunjukkan hubungan yang berbanding terbalik, di mana ketika nilai MAPE menurun maka nilai *accuracy* cenderung meningkat, begitu pula sebaliknya. Hal ini karena MAPE digunakan untuk mengukur persentase kesalahan prediksi, sedangkan *accuracy* menggambarkan tingkat ketepatan hasil prediksi model. Pada beberapa provinsi seperti Jawa Tengah, Jawa Timur, Sumatera Barat, Sulawesi Tengah, dan Bali, penurunan nilai MAPE diikuti oleh peningkatan nilai *accuracy*, yang menunjukkan bahwa semakin kecil persentase kesalahan prediksi maka tingkat ketepatan prediksi model menjadi semakin baik. Sebaliknya, pada provinsi seperti DKI Jakarta, Kepulauan Riau, dan Kalimantan Selatan, nilai MAPE yang tinggi diikuti oleh nilai *accuracy* yang rendah. Dengan demikian, hubungan antara MAPE dan *accuracy* menunjukkan bahwa kedua metrik tersebut memiliki keterkaitan dalam menggambarkan kualitas hasil prediksi model.

4.5 Peramalan

Berdasarkan hasil evaluasi kinerja model menggunakan metrik evaluasi, diperoleh bahwa pada sebagian besar provinsi model *hybrid* PVAR–LSTM Skema 2 menghasilkan nilai kesalahan yang lebih rendah dibandingkan Skema 1 maupun PVAR tunggal. Hal ini menunjukkan bahwa pendekatan pada Skema 2 lebih efektif dalam menangkap pola nonlinier yang belum sepenuhnya terakomodasi oleh model linier PVAR. Oleh karena itu, model *hybrid* PVAR–LSTM Skema 2 dipilih sebagai model terbaik dan digunakan untuk melakukan peramalan tiga periode ke depan.

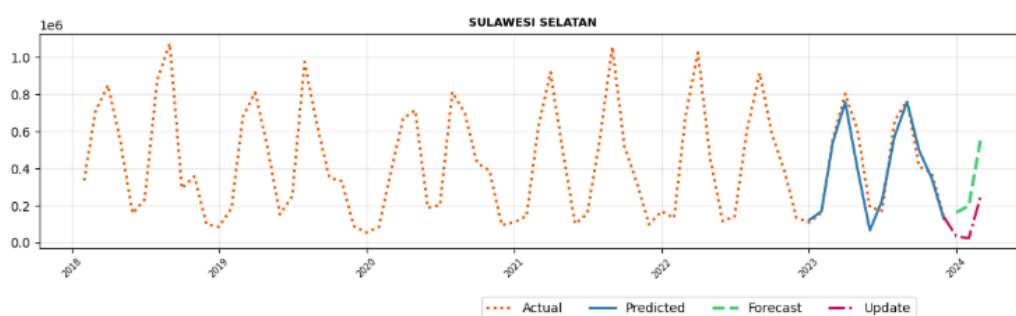
1. Luas Panen

Pada variabel luas panen, dilakukan peramalan menggunakan model terbaik yang diperoleh dari tahap evaluasi sebelumnya. Peramalan ini bertujuan untuk memprediksi nilai luas panen pada beberapa periode ke depan berdasarkan pola data yang telah terbentuk. Hasil peramalan kemudian disajikan dalam Gambar 54, 55 dan 56 sebagai berikut.



Gambar 54. Visualisasi Peramalan Kelompok Luas Panen Tinggi

Berdasarkan Gambar 54, hasil peramalan pada kelompok luas panen tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, terlihat bahwa nilai peramalan secara umum sudah mengikuti arah tren yang sama dengan data *update*, yaitu menunjukkan kecenderungan peningkatan pada periode mendatang. Namun, masih terdapat selisih antara hasil peramalan dan nilai *update*, terutama pada awal periode. Pada Jawa Barat, Jawa Tengah, dan Jawa Timur, nilai *update* cenderung berada sedikit di bawah hasil peramalan di awal periode sebelum akhirnya bergerak naik mengikuti tren yang sama. Hal ini menunjukkan bahwa meskipun model sudah mampu menangkap arah pergerakan dengan baik, masih terdapat perbedaan pada besaran nilai yang dihasilkan, khususnya pada periode awal peramalan. Secara keseluruhan, peramalan sudah selaras dengan data *update* dalam hal tren, namun masih terdapat selisih pada tingkat nilainya.



Gambar 55. Visualisasi Peramalan Kelompok Luas Panen Sedang

Berdasarkan Gambar 55, hasil peramalan pada kelompok luas panen sedang yaitu oleh Sulawesi Selatan, terlihat bahwa nilai peramalan secara umum telah mengikuti arah tren yang sama dengan data *update*, khususnya dalam menunjukkan pola kenaikan setelah periode terakhir *testing*. Namun, masih terdapat selisih antara hasil peramalan dan nilai *update*, terutama pada awal periode. Pada awal periode

peramalan, nilai *update* terlihat lebih rendah dibandingkan hasil peramalan, bahkan sempat mengalami penurunan yang cukup tajam, sebelum kemudian meningkat kembali dan mengikuti arah yang sama. Sementara itu, hasil peramalan menunjukkan pola kenaikan yang lebih stabil. Selisih ini menunjukkan adanya perbedaan pada besaran nilai yang dihasilkan, meskipun arah pergerakan keduanya tetap konsisten. Secara keseluruhan, hasil peramalan sudah sejalan dengan data *update* dalam hal tren, namun masih terdapat perbedaan pada tingkat nilainya, terutama pada periode awal.



Gambar 56. Visualisasi Peramalan Kelompok Luas Panen Rendah

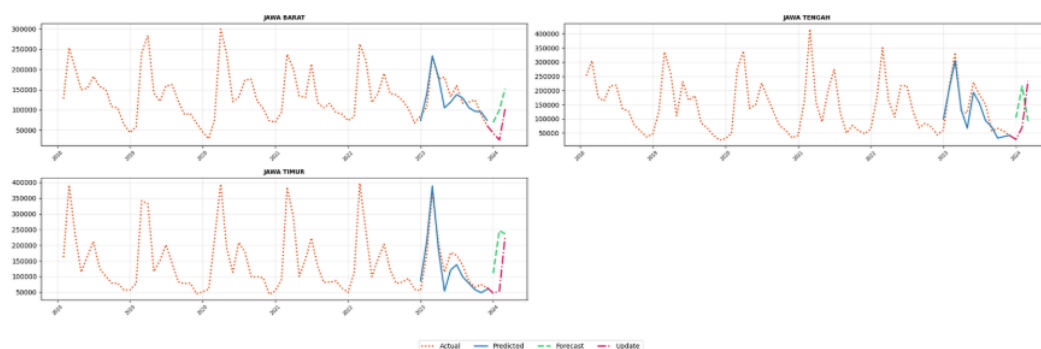
Berdasarkan Gambar 56, hasil peramalan pada kelompok luas panen rendah yang mencakup beberapa provinsi, terlihat bahwa nilai peramalan secara umum telah

mengikuti arah tren yang sama dengan data *update*, meskipun pola pergerakannya cenderung lebih berfluktuasi dibandingkan kelompok lainnya. Pada sebagian besar provinsi, hasil peramalan menunjukkan kecenderungan perubahan yang relatif stabil, sementara nilai *update* pada awal periode seringkali berada di bawah hasil peramalan sebelum kemudian menyesuaikan dan mengikuti arah yang sama. Selisih antara hasil peramalan dan nilai *update* terlihat cukup bervariasi antar provinsi, terutama pada wilayah dengan fluktuasi data yang lebih tinggi. Pada beberapa kasus, nilai *update* menunjukkan penurunan yang lebih tajam di awal periode, sedangkan hasil peramalan cenderung lebih halus dalam mengikuti pola. Meskipun demikian, arah pergerakan keduanya tetap konsisten. Secara keseluruhan, hasil peramalan sudah sejalan dengan data *update* dalam menangkap tren umum, tetapi masih terdapat perbedaan pada tingkat nilai, khususnya pada periode awal dan pada provinsi dengan pola data yang lebih berfluktuasi.

Untuk memberikan gambaran yang lebih rinci terkait hasil peramalan yang telah divisualisasikan, nilai hasil peramalan luas panen untuk masing-masing provinsi pada periode Januari hingga Maret disajikan pada lampiran 1. Lampiran 1 tersebut memuat data *update*, hasil model PVAR, serta hasil pengembangan menggunakan Hybrid Skema 1 dan Hybrid Skema 2. Melalui penyajian ini, dapat diamati secara langsung perbedaan nilai pada setiap metode, sehingga memudahkan dalam mengevaluasi sejauh mana hasil peramalan mampu mengikuti data *update* baik dari sisi tren maupun besaran nilainya.

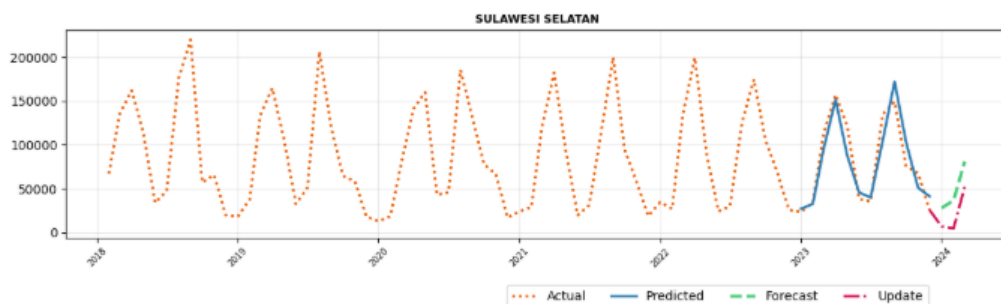
2. Produksi Padi

Peramalan ini bertujuan untuk memprediksi nilai luas panen pada beberapa periode ke depan berdasarkan pola data yang telah terbentuk. Hasil peramalan kemudian disajikan dalam Gambar 57, 58, dan 59 sebagai berikut.



Gambar 57. Visualisasi Peramalan Kelompok Produksi Padi Tinggi

Berdasarkan Gambar 56, hasil peramalan pada kelompok produksi tinggi yang terdiri dari Jawa Barat, Jawa Tengah, dan Jawa Timur, terlihat bahwa nilai peramalan secara umum telah mengikuti arah tren yang sama dengan data *update*, yaitu menunjukkan kecenderungan peningkatan setelah periode terakhir *testing*. Namun, masih terdapat selisih antara hasil peramalan dan nilai *update*, terutama pada awal periode. Pada ketiga provinsi, nilai *update* cenderung berada di bawah hasil peramalan pada awal periode sebelum kemudian meningkat dan mengikuti arah yang sama. Pada Jawa Timur, selisih terlihat lebih jelas karena hasil peramalan menunjukkan kenaikan yang cukup tinggi, sementara nilai *update* pada awal periode masih lebih rendah sebelum akhirnya meningkat tajam. Pola serupa juga terlihat pada Jawa Barat dan Jawa Tengah, meskipun dengan perbedaan yang tidak sebesar Jawa Timur. Secara keseluruhan, hasil peramalan sudah sejalan dengan data *update* dalam menangkap arah tren produksi, tetapi masih terdapat perbedaan pada tingkat nilainya, khususnya pada periode awal peramalan.



Gambar 58. Visualisasi Peramalan Kelompok Produksi Padi Sedang

Berdasarkan Gambar 58, hasil peramalan pada kelompok produksi sedang yang diwakili oleh Sulawesi Selatan, terlihat bahwa nilai peramalan secara umum telah mengikuti arah tren yang sama dengan data *update*, khususnya dalam menunjukkan pola peningkatan setelah periode terakhir *testing*. Namun, masih terdapat selisih antara hasil peramalan dan nilai *update*, terutama pada awal periode. Pada awal periode peramalan, nilai *update* terlihat lebih rendah dibandingkan hasil peramalan dan sempat mengalami penurunan yang cukup tajam, sebelum kemudian meningkat kembali dan mengikuti arah yang sama. Sementara itu, hasil peramalan menunjukkan pola kenaikan yang lebih stabil. Selisih ini menunjukkan adanya perbedaan pada besaran nilai yang dihasilkan, meskipun arah pergerakan keduanya tetap konsisten. Secara keseluruhan, hasil peramalan sudah sejalan dengan data *update* dalam hal tren, tetapi masih terdapat perbedaan pada tingkat nilainya, terutama pada periode awal.



Gambar 59. Visualisasi Peramalan Kelompok Produksi Padi Rendah

Berdasarkan Gambar 59, hasil peramalan pada kelompok produksi rendah yang mencakup beberapa provinsi, terlihat bahwa nilai peramalan secara umum telah mengikuti arah tren yang sama dengan data *update*, meskipun pola pergerakannya cenderung lebih berfluktuasi. Pada sebagian besar provinsi, hasil peramalan menunjukkan perubahan yang relatif lebih halus, sementara nilai *update* pada awal periode sering berada di bawah hasil peramalan sebelum kemudian menyesuaikan dan bergerak mengikuti arah yang sama. Selisih antara hasil peramalan dan nilai *update* terlihat bervariasi antar provinsi, terutama pada wilayah dengan pola data yang lebih tidak stabil. Pada beberapa provinsi, nilai *update* menunjukkan penurunan yang lebih tajam di awal periode dibandingkan hasil peramalan.

Meskipun demikian, arah pergerakan keduanya tetap konsisten. Secara keseluruhan, hasil peramalan sudah sejalan dengan data *update* dalam menangkap tren umum produksi, tetapi masih terdapat perbedaan pada tingkat nilainya, khususnya pada periode awal dan pada provinsi dengan fluktuasi yang lebih tinggi.

Untuk memberikan gambaran yang lebih rinci terkait hasil peramalan yang telah divisualisasikan, nilai hasil peramalan Produksi Padi untuk masing-masing provinsi pada periode Januari hingga Maret disajikan pada lampiran 1. Lampiran 1 tersebut memuat data *update*, hasil model PVAR, serta hasil pengembangan menggunakan Hybrid Skema 1 dan Hybrid Skema 2. Melalui penyajian ini, dapat diamati secara langsung perbedaan nilai pada setiap metode, sehingga memudahkan dalam mengevaluasi sejauh mana hasil peramalan mampu mengikuti data *update* baik dari sisi tren maupun besaran nilainya.

4.6 Perbandingan dengan Penelitian Sebelumnya

Penelitian ini memiliki keterkaitan dengan sejumlah studi terdahulu yang menggunakan pendekatan PVAR, LSTM, maupun model *hybrid* VAR–LSTM dalam analisis deret waktu. Pada penelitian berbasis PVAR, model PVAR umumnya digunakan untuk menangkap hubungan dinamis antar variabel dalam konteks data lintas wilayah. Studi Usman dkk. (2022) dan Sun (2021) menunjukkan bahwa PVAR efektif dalam memodelkan keterkaitan linier antar unit *cross-section*, serta mampu menggambarkan dinamika jangka pendek maupun jangka panjang melalui kerangka panel dinamis. Sejalan dengan temuan tersebut, penelitian ini menggunakan PVAR sebagai model dasar untuk merepresentasikan hubungan struktural antara luas panen dan produksi padi antar provinsi di Indonesia.

Di sisi lain, penelitian berbasis LSTM menunjukkan bahwa pendekatan *deep learning* memiliki keunggulan dalam menangkap pola nonlinier dan musiman yang kompleks pada data deret waktu. Penelitian Solgi dkk. (2021) maupun Firdausi dkk. memperlihatkan bahwa LSTM mampu meningkatkan akurasi peramalan produksi padi dibandingkan model linier konvensional. Namun demikian, sebagian besar penelitian tersebut berfokus pada satu deret waktu atau satu wilayah tertentu, sehingga belum mempertimbangkan antar wilayah dalam kerangka data panel.

Pendekatan *hybrid* VAR–LSTM sebagaimana dikembangkan oleh Savada dkk. (2025) berupaya mengintegrasikan kekuatan model linier dan nonlinier untuk meningkatkan performa prediksi. Meskipun demikian, penerapan model *hybrid* sebelumnya umumnya dilakukan pada konteks multivariat non-panel. Berbeda dengan studi-studi tersebut, penelitian ini mengembangkan pendekatan *hybrid* dalam kerangka panel regional yang melibatkan 34 provinsi di Indonesia. Integrasi PVAR dan LSTM memungkinkan model menangkap keterkaitan linier antar provinsi sekaligus mempelajari pola nonlinier pada masing-masing wilayah.

Hasil evaluasi menunjukkan bahwa model *Hybrid* PVAR–LSTM secara umum menghasilkan nilai MAE, RMSE dan MAPE yang lebih rendah dibandingkan model PVAR tunggal pada sebagian besar provinsi. Temuan ini memperkuat literatur yang menyatakan bahwa kombinasi model linier dan nonlinier dapat meningkatkan akurasi peramalan. Dengan demikian, penelitian ini tidak hanya mengonfirmasi temuan sebelumnya, tetapi juga memperluas penerapan model *hybrid* ke dalam konteks data panel pertanian regional, yang masih relatif terbatas dalam literatur.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan terhadap data luas panen dan produksi padi pada 34 provinsi di Indonesia, baik dalam konteks prediksi maupun peramalan, diperoleh beberapa kesimpulan sebagai berikut.

1. Model *Panel Vector Autoregression* (PVAR) mampu menggambarkan hubungan dinamis yang bersifat linier antara luas panen dan produksi padi antar provinsi. Hal ini menunjukkan bahwa PVAR cukup efektif dalam menangkap keterkaitan struktural pada data panel. Namun demikian, dalam peramalan, model PVAR masih memiliki keterbatasan karena kurang adaptif terhadap perubahan pola terbaru, terutama pada provinsi dengan fluktuasi produksi yang tinggi.
2. Model *Hybrid* PVAR–LSTM dalam penelitian ini memberikan hasil peramalan yang lebih baik dibandingkan model PVAR tunggal pada sebagian besar wilayah. Integrasi kemampuan PVAR dalam menangkap pola linier antar wilayah dengan kemampuan LSTM dalam memodelkan pola nonlinier menjadikan pendekatan *hybrid* lebih fleksibel dalam mengikuti dinamika data pertanian yang kompleks.
3. Evaluasi menggunakan MAE, RMSE dan MAPE menunjukkan bahwa pada sebagian besar provinsi, model *hybrid* menghasilkan nilai kesalahan prediksi yang lebih rendah dibandingkan PVAR. Peningkatan akurasi yang diperoleh berbeda-beda antar provinsi, bergantung pada karakteristik pola musiman serta tingkat fluktuasi data pada masing-masing wilayah.
4. Pada variabel produksi padi maupun luas panen, pendekatan *hybrid* secara umum lebih mampu menyesuaikan pola musiman dan perubahan jangka pendek dibandingkan PVAR. Hal ini menunjukkan bahwa model *Hybrid* PVAR–LSTM

lebih sesuai digunakan untuk peramalan data panel pertanian yang bersifat heterogen antar provinsi.

Secara keseluruhan, penelitian ini menunjukkan bahwa model *Hybrid* PVAR–LSTM merupakan pendekatan yang efektif untuk meningkatkan akurasi peramalan luas panen dan produksi padi di Indonesia, khususnya dalam konteks data panel antar wilayah.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya yaitu menambahkan variabel lain yang relevan dalam sektor pertanian, seperti curah hujan, indeks iklim, penggunaan pupuk, harga gabah, maupun faktor kebijakan pemerintah. Penambahan variabel tersebut diharapkan dapat meningkatkan kemampuan model dalam menangkap perubahan produksi padi yang dipengaruhi oleh faktor eksternal. Kemudian periode data yang digunakan dapat diperpanjang hingga mencakup tahun-tahun terbaru agar model lebih adaptif terhadap perubahan pola produksi yang terjadi setelah tahun 2023. Dengan data yang lebih panjang, peramalan pada periode terbaru diharapkan menjadi lebih akurat. Dengan adanya pengembangan tersebut, diharapkan penelitian mengenai peramalan produksi padi menggunakan pendekatan *hybrid* PVAR-LSTM dapat memberikan kontribusi yang lebih luas dalam mendukung kebijakan pertanian dan ketahanan pangan nasional.

DAFTAR PUSTAKA

- Akbar, R., Santoso, R., dan Warsito, B. 2022. Prediksi Tingkat Temperatur Kota Semarang Menggunakan Metode Long Short-Term Memory (LSTM). *Jurnal Gaussian*. **11**(4): 572–579. DOI: <https://doi.org/10.14710/j.gauss.11.4.572-579>.
- Amalia, E.L., Yunhasnawa, Y., dan Rahmatanti, A.R. 2022. Sistem prediksi penjualan frozen food studi kasus Supermama Frozen Food dengan metode Monte Carlo. *Jurnal Buana Informatika*. **13**(2): 136–145. DOI: 10.24002/jbi.v13i02.6496.
- Amini, A., dan Dolatshahi, M. 2023. Effects of automatic hyperparameter tuning on the performance of multivariate deep learning-based rainfall nowcasting. *Water Resources Research*. **59**(8): 1-27. DOI : <https://doi.org/10.1029/2022WR032789>.
- Arellano, M. 2003. *Panel Data Econometrics*. Oxford University Press, New York.
- Arifin, B., Warsono, dan Kurniasari, D. 2023. Perbaikan Angka Potensi Produksi Padi dengan Metode Hybrid. *Buletin Ringkas Statistical & Policy Brief*. Edisi 22: 53–59. Forum Masyarakat Statistik (FMS).
- Barbierato, E., dan Gatti, A. 2024. The Challenges of Machine Learning: A Critical Review. *Electronics*. **13**(2): 1-30. DOI: <https://doi.org/10.3390/electronics13020416>
- Boruku, P.B., Essi, I.D., dan Dum, D.Z. 2022. Exchange Rate, Interest Rate and Agricultural Export Earnings: An Analysis Using Panel Data Vector Autoregressive Model. *International Journal of Applied Sciences and Management Technology*. **8**(2): 14-37. DOI: 10.56201/ijasmt.v8.no2.2022.pg14.37.
- Buswar, A.S., Ruminta, dan Wicaksono, F.Y. 2025. Rice (*Oryza sativa* L.) production, productivity, and harvested area in Karawang Regency under extreme weather from 1991–2023. *Jurnal Kultivasi*. **24**(2): 121–131. DOI: <https://doi.org/10.24198/kultivasi.v24i2.63735>.
- Cangombe, E.L., Almeida, L.G., dan Tavares, F.O. 2025. Determinants of Banking Profitability in Angola: A Panel Data Analysis with Dynamic GMM Estimation. *Risks*. **13**(7): 1-21. DOI: <https://doi.org/10.3390/risks13070123>.

- Chen, X., dan Moraga, P. 2025. Forecasting dengue across Brazil with LSTM neural networks and SHAP-driven lagged climate and spatial effects. *BMC Public Health*. 25: 973. DOI: <https://doi.org/10.1186/s12889-025-22106-7>.
- Dissanayake, B., Hemachandra, O., Lakshitha, N., Haputhanthri, D., dan Wijayasiri, A. 2021. A Comparison of ARIMAX, VAR, and LSTM on Multivariate Short-Term Traffic Volume Forecasting. *Proceeding of the 28th Conference of FRUCT Association*. 564-570.
- Firdausi, H.M., Utomo, S.B., Rahardi, G.A., dan Prasetyo, D.H.T. 2025. A Multivariate LSTM Approach for Monthly Rice Production Forecasting in East Java. *Jurnal Sistem Cerdas*. 8(3): 364–374. DOI: 10.37396/jsc.v8i3.595.
- Haloui, I., Amzil, H., Yang, G., Fourati, I., dan Li, Y. 2025. The Impact of Renewable Energy Use, Financial Development, and Industrialization on CO₂ Emissions in Middle-Income Economies: A GMM-PVAR Analysis. *Sustainability*. 17(18): 1–25. DOI: <https://doi.org/10.3390/su17188178>.
- Harahap, Q.H., Lubis, Z., dan Supriana, T. 2024. Analysis of rice production trends in ten rice production centers in Indonesia. *TRISS: Jurnal Agribisnis dan Sains Terapan*. 6(2): 233–254. Retrieved from <https://triss.org/index.php/journal/article/view/72>.
- Hayakawa, K. 2016. Improved GMM estimation of panel VAR models. *Computational Statistics & Data Analysis*. 100: 240–264. DOI: <https://doi.org/10.1016/j.csda.2015.05.004>
- Hsiao, C. 2014. *Analysis of Panel Data* (3rd ed.). Cambridge University Press, New York.
- Im, K., Pesaran, M., dan Shin, Y. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics*. 115(2003): 53-74. DOI: [https://doi.org/10.1016/S0304-4076\(03\)00092-7](https://doi.org/10.1016/S0304-4076(03)00092-7)
- Irfan, M., Usman, M., Warsono, dan Wamiliana. 2023. Modeling and forecasting data with panel vector autoregressive GMM estimation on economic growth and inflation data in 10 ASEAN countries. *ICASMI*. AIP Publishing. DOI: 10.1063/5.0208249
- Isa, I.G.T., dan Junedi, B. 2022. Hyperparameter Tuning Epoch dalam Meningkatkan Akurasi Data Latih dan Data Validasi pada Citra Pengendara. *Prosiding Seminar Nasional Sains dan Teknologi*. 12(1): 231-237. DOI: 10.36499/psnst.v12i1.6697.

- Kazemzadeh, E., Fuinhas, J.A., Koengkan, M., dan Shadmehri, M.T.A. 2023. Relationship between renewable electricity consumption, complexity, and financial development: A PVAR approach. *International Economics*. 173(3): 260-275. DOI: 10.1016/j.inteco.2022.12.001.
- Kementerian Pertanian. 2024. *Outlook Padi 2024*. Pusat Data dan Sistem Informasi Pertanian.
- Kapetanios, G., Marcellino, M., Martino, A., Papailias, F., dan Mazzi, G.L. 2021. Panel VAR models for nowcasting. Working Paper No. 2021/5, King's Business School.
- Kim, M. K., Kim, Y.-S., Fu, N., Liu, J., Wang, J., Lee, S., dan Srebrić, J. 2025. Advanced techniques for electricity consumption prediction in buildings using comparative correlation analysis, data normalization, and Long Short-Term Memory (LSTM) networks: A case study of a U.S. commercial building. *Energy Reports*. 11(2025): 56-65. DOI: <https://doi.org/10.1016/j.egyr.2025.05.074>.
- Kurniawan, K., Ceasaro, B., dan Sucipto, S. 2024. Perbandingan Fungsi Aktivasi Untuk Meningkatkan Kinerja Model LSTM dalam Prediksi Ketinggian Air Sungai. *JEPIN*. 10(1): 134–143. DOI: 10.26418/jp.v10i1.72866.
- Lasiyono, M. M., Nurhayati, dan Soares, T. G. 2025. Enhancing Support Vector Machine Performance for Heart Attack Prediction using RobustScaler-Based Outlier Handling. *Bulletin of Informatics and Data Science*. 4(1): 1–9. DOI: 10.61944/bids.v4i1.94.
- Levin, A., Lin, C. F., dan Chu, C. S. J. 2002. Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*. 108(1): 1–24. DOI : [https://doi.org/10.1016/S0304-4076\(01\)00098-7](https://doi.org/10.1016/S0304-4076(01)00098-7).
- Li, H., Yang, X., Meng, F., Hou, Y., Zhang, J., Zhang, L., Yang, G., dan Liu, J. 2024. A Dynamic Impact Evaluation of the High-Quality Development of China's Construction Industry Using the Panel Vector Autoregressive Model. *Buildings*. 14(9): 2871. DOI: <https://doi.org/10.3390/buildings14092871>.
- Liao, L., Li, H., Shang, W., dan Ma, L. 2022. An empirical study of the impact of hyperparameter tuning and model optimization on the performance properties of deep neural networks. *ACM Journals*. 31(4): 1–25. DOI: 10.1145/3506695.
- Lim, B., Arık, S. Ö., Loeff, N., dan Pfister, T. 2021. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. *International Journal of Forecasting*. 37(4): 1748–1764. DOI: 10.1016/j.ijforecast.2021.03.012.

- Love, I., dan Zicchino, L. 2006. Financial Development and Dynamic Investment Behavior: Evidence from Panel VAR. *The Quarterly Review of Economics and Finance*. **46**(2): 190–210. DOI: 10.1016/j.qref.2005.11.007.
- Lütkepohl, H. 2005. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag.
- Mahesh, N., Jagan Babu, J., Nithya, K., dan Arummozhi, S. A. 2024. Water quality prediction using LSTM with combined normalizer for efficient water management. *Desalination and Water Treatment*. **317**(2024): 1-8. DOI: 10.1016/j.dwt.2024.100183.
- Montgomery, D.C., Jennings, C.L., dan Kulahci, M. 2015. *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.
- Mujiyanto, M., Nurindahsari, S., dan Izza, R. N. 2024. Comparison of Linear Regression, ARIMA, Simple Exponential Smoothing, hybrid ARIMA-LSTM, and EWMA in Forecasting Commodity Prices. *Telematika*. **17**(2): 112–127. DOI: DOI:10.35671/telematika.v17i2.2932.
- Müller, A.C., dan Guido, S. 2017. *Introduction to Machine Learning with Python*. O'Reilly Media.
- Nabillah, I., dan Ranggadara, I. 2020. Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut. *Journal of Information System (JOINS)*. **5**(2): 250–255. November 2020. DOI: 10.33633/joins.v5i2.3900.
- Ndung'u, R. N. 2022. Data Preparation for Machine Learning Modelling. *International Journal of Computer Applications Technology & Research*. **11**(6): 231–235. DOI: 10.7753/IJCATR1106.1008.
- Rehman, S., dan Iqbal, M. 2025. Deep learning-based electricity forecasting for smart buildings using Optuna hyperparameter optimization. *Energy Informatics*. **12**(1): 1–19. DOI: <https://doi.org/10.1186/s42162-025-00587-5>
- Rowan, L., Muflikhah, L., dan Cholissodin, I. 2022. Peramalan Kasus Positif COVID-19 di Jawa Timur menggunakan Metode Hybrid ARIMA-LSTM. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (JPTIIK)*. **6**(9): 4146–4153. Tersedia di: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11524/5111>.
- Rusman, M. A. A., Darsono, dan Antriandarti, E. 2023. Analisis Forecasting Produksi Padi di Indonesia. *Prosiding Dies Natalis UNS*. **7**(1): 728–739.

- Savada, A.G.A., Nama, G.F., Yulianti, T., dan Mardiana, M. 2025. Economic data forecasting using hybrid VAR–LSTM. *JuTISI*. **11**(1): 91–104. DOI: <https://doi.org/10.28932/jutisi.v11i1.10066>.
- Sofhya, H. N. 2023. Cluster Analysis of Indonesian Provinces Based on Harvest Area and Rice Productivity Using Single Linkage Method. *Jurnal Sains Dan Teknologi Industri*. **20**(2): 549-558. DOI : <https://doi.org/10.24014/sitekin.v20i2.21737>
- Solgi, R., Alizadeh, M. R., dan Khodabakhshi, M. 2021. Long short-term memory neural network (LSTM-NN) for aquifer level time series forecasting using in-situ piezometric observations. *Journal of Hydrology*. **601**(2021): 1-12. DOI: <https://doi.org/10.1016/j.jhydrol.2021.126800>.
- Sun, Y., Chen, Y., dan Xu, W. 2021. An Empirical Study on the Relationship between Education and Economic Development. *Scientific Programming*. **2021**(9): 1–12. DOI: <https://doi.org/10.1155/2021/5590283>.
- Usman, O., Alola, A. A., dan Akadiri, S. S. 2022. Effects of domestic material consumption, renewable energy, and financial development on environmental sustainability in the EU-28: Evidence from a GMM panel VAR. *Renewable Energy*. 184: 239–251. DOI: <https://doi.org/10.1016/j.renene.2021.11.086>.
- Utami, B. H. S., Warsono, Kurniasari, D., Usman, M., dan Elfaki, F. A. M. 2014. Generalized Method of Moments’ characteristics and its application on panel data. *Science International (Lahore)*. **26**(3): 987–992.
- Wang, B., Tan, Z., Sheng, W., Liu, Z., Wu, X., Ma, L., dan Li, Z. 2024. Identification of Groundwater Contamination Sources Based on a Deep Belief Neural Network. *Water*. **16**(17): 2449. DOI: <https://doi.org/10.3390/w16172449>.
- Warsono., Kurniasari, D., Lestari, D. A. H., Wahyudi, H., Alhani, M., Su’admamaji, A., dan Maulana, M. A. 2025. Forecasting of Multivariate Time Series Data of Export and Import Volumes of Coffee Using the Vector Autoregressive (VAR) - Long Short-Term Memory (LSTM) Hybrid Model. *Proceedings of the 5th International Conference on Applied Sciences, Mathematics, and Informatics (ICASMI 2024)*. Advances in Physics Research, Vol. 13. https://doi.org/10.2991/978-94-6463-730-4_5
- Wei, W.W.S. 2019. *Multivariate Time Series Analysis and Applications*. John Wiley & Sons.
- Yang, R., Li, J., Wang, T., dan Zhou, M. 2023. The knowledge analysis of panel vector autoregression: a systematic review. *SAGE Open*. **13**(4): 1–20. DOI: [10.1177/21582440231215991](https://doi.org/10.1177/21582440231215991).