

Peningkatan Akurasi Peramalan Trafik Udara Bandara Hang Nadim dengan ARIMA-Neural Network

Enhancing Air Traffic Forecasting Accuracy at Hang Nadim Airport Using ARIMA-Neural Network

¹Masparudin*, ²Abdullah, ³Raymond Erz Saragih, ⁴Yonky Pernando, ⁵Ilwan Syafrinal

^{1,5}Program Studi Teknik Perangkat Lunak, Fakultas Komputer, Universitas Universal

²Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Islam Indragiri

^{3,4}Program Studi Teknik Informatika, Fakultas Komputer, Universitas Universal

^{1,3,4,5}Pasir Putih, Sadai - Batam, Kepulauan Riau, Indonesia

²Jl. Provinsi Parit 1 Tembilahan Hulu, Riau, Indonesia

*e-mail: masparudin.mahmud@gmail.com

(received: 28 March 2026, revised: 23 April 2026, accepted: 24 April 2026)

Abstrak

Fluktuasi arus penumpang di Bandara Internasional Hang Nadim memiliki volatilitas ekstrem yang dipengaruhi oleh karakteristik unik kawasan *Free Trade Zone* (FTZ). Metode statistik tunggal sering gagal menangkap pola *non-linear* pada data dengan variabilitas tinggi ini. Oleh karena itu, penelitian ini mengusulkan model Hibrida *ARIMA-Neural Network* untuk meningkatkan akurasi peramalan. Variabel utama yang digunakan adalah total volume penumpang bulanan (kedatangan dan keberangkatan). Tahapan penelitian diawali dengan prapemrosesan data (rasio latih-uji 80:20), pemodelan komponen linear menggunakan ARIMA, ekstraksi residu, dan pemodelan komponen non-linear menggunakan Multi-Layer Perceptron (MLP) untuk mengoreksi sisa kesalahan secara one-step-ahead. Hasil evaluasi menunjukkan model tunggal ARIMA lambat mengantisipasi lonjakan ekstrem dengan *Mean Absolute Percentage Error* (MAPE) sebesar 23,75%. Integrasi model hibrida terbukti sukses mengompensasi kelemahan tersebut, menurunkan nilai MAPE menjadi 12,51%. Pencapaian ini merepresentasikan reduksi *error* sebesar 47,33% dari *baseline*. Secara kebaruan, pendekatan hibrida ini memberikan solusi komputasional yang sangat andal bagi manajemen bandara dengan karakteristik ganda (wisata dan industri) dalam memitigasi ketidakpastian perencanaan kapasitas.

Kata kunci: ARIMA, peramalan, bandara hang nadim, model hibrida, *neural network*, analisis deret waktu.

Abstract

Passenger traffic fluctuations at Hang Nadim International Airport exhibit extreme volatility influenced by the unique characteristics of the Free Trade Zone (FTZ). Single statistical methods often fail to capture non-linear patterns in this high-variability data. Therefore, this study proposes a Hybrid ARIMA-Neural Network model to enhance forecasting accuracy. The primary variable used is the total monthly passenger volume (arrivals and departures). The research stages began with data preprocessing (80:20 train-test ratio), linear component modeling using ARIMA, residual extraction, and non-linear component modeling using Multi-Layer Perceptron (MLP) to correct residual errors on a one-step-ahead basis. Evaluation results show that the standalone ARIMA model is slow to anticipate extreme surges, resulting in a Mean Absolute Percentage Error (MAPE) of 23.75%. The hybrid model integration proved successful in compensating for these weaknesses, reducing the MAPE value to 12.51%. This achievement represents a 47.33% error reduction from the baseline. In terms of novelty, this hybrid approach provides a highly reliable computational solution for airport management with dual characteristics (tourism and industry) in mitigating uncertainty in capacity planning.

Keywords: ARIMA, forecasting, hang nadim airport, hybrid model, neural network, time series analysis.

1 Pendahuluan

Transportasi udara di kawasan strategis Free Trade Zone (FTZ) Batam memiliki karakteristik unik karena melayani dua profil penumpang utama secara simultan: pelaku bisnis industri dan wisatawan. Interaksi ini menciptakan struktur deret waktu yang sangat dinamis, di mana pola musiman rutin sering kali tumpang tindih dengan fluktuasi mendadak (shocks) akibat kebijakan ekonomi dan iklim investasi kawasan. Kompleksitas tersebut menghasilkan residu volatilitas tinggi yang sering kali gagal ditangkap oleh model linier tradisional seperti ARIMA. Ketidakakuratan dalam prediksi arus ini dapat berdampak pada inefisiensi alokasi sumber daya hingga peningkatan biaya operasional [1], [2], sehingga pengembangan model yang presisi menjadi kebutuhan mendesak untuk mendukung keputusan strategis manajemen bandara [3].

Metode statistik konvensional seperti ARIMA sering dijadikan standar dasar (baseline) karena kemampuannya memodelkan struktur linear [4]. Namun, berbagai literatur menyoroti bahwa model tunggal ARIMA memiliki keterbatasan fundamental dalam menangkap volatilitas non-linear yang dinamis [5], [6]. Untuk mengatasi kelemahan tersebut, pendekatan hibrida yang mengintegrasikan ARIMA dengan Artificial Neural Network (ANN) digunakan guna menangkap residu non-linear yang tidak terdeteksi oleh model statistik [7], [8]. Bukti empiris mengonfirmasi bahwa mekanisme koreksi residu secara konsisten menurunkan tingkat kesalahan dibandingkan model tunggal [9], [10], [11], bahkan pada pola spasial-temporal yang rumit [12], [13]. Meskipun demikian, masih terdapat celah penelitian mengenai penerapan model ini pada bandara di kawasan FTZ yang memiliki karakteristik musiman unik [14], [15].

Berdasarkan kesenjangan tersebut, perumusan masalah dalam penelitian ini berfokus pada efektivitas skema koreksi residu univariat dalam menangani anomali volatilitas ekstrem di kawasan berstatus khusus. Adapun tujuan penelitian ini adalah untuk mengusulkan model Hibrida ARIMA-Neural Network guna meramalkan arus penumpang di Bandara Hang Nadim serta meningkatkan akurasi perencanaan operasional.

2 Tinjauan Literatur

Evolusi metode peramalan telah bergerak dari model statistik murni menuju pendekatan berbasis kecerdasan komputasional. Zhang (2003) dan Khashei & Bijari (2010) meletakkan dasar teoretis bahwa data dunia nyata jarang sekali murni linear, sehingga penggunaan ARIMA tunggal sering kali tidak memadai [16]. Untuk mengatasi hal ini, paradigma hibrida diperkenalkan dengan asumsi bahwa data deret waktu Y_t terdiri dari komponen linear L_t dan komponen non-linear N_t . Dalam kerangka ini, model statistik digunakan sebagai pra-pemrosesan linear, sementara model pembelajaran mesin (*Machine Learning*) berfungsi sebagai korektor residu yang cerdas [14]. Pendekatan probabilistik terkini bahkan mulai mengintegrasikan ketidakpastian (*uncertainty*) ke dalam model jaringan saraf untuk meningkatkan reliabilitas prediksi jangka panjang [15].

Penerapan model hibrida pada sektor penerbangan menunjukkan hasil yang menjanjikan, terutama pasca-pandemi COVID-19 di mana pola data menjadi sangat tidak teratur. Rabbouch et al. (2022) menggunakan dekomposisi VMD yang digabungkan dengan LSTM-ARIMA untuk memprediksi lalu lintas udara AS, menunjukkan bahwa teknik dekomposisi membantu model menangani guncangan data ekstrem [17]. Wang et al. (2022) juga menemukan bahwa penggabungan karakteristik musiman dengan model hibrida SARIMA-NAR sangat efektif untuk prediksi aliran lalu lintas yang memiliki pola berulang [13]. Dalam konteks pariwisata yang erat kaitannya dengan penerbangan, Misengo et al. (2023) membuktikan bahwa hibrida ARIMA-MLP memberikan akurasi superior dalam meramalkan kedatangan wisatawan bulanan [18].

Terkait arsitektur komponen non-linear, literatur menunjukkan beragam pendekatan yang efektif. Naheliya et al. (2023) dan Sattarzadeh et al. (2023) menyoroti kemampuan arsitektur *Deep Learning* seperti *Long Short-Term Memory* (LSTM) dalam menangkap dependensi temporal jangka panjang [19], [20]. Namun, penelitian lain juga menunjukkan bahwa arsitektur *Neural Network* standar (seperti MLP) tetap sangat efektif dan efisien, terutama ketika difokuskan pada pemodelan residu [18]. Selain itu, teknik adaptif seperti *fuzzy filters* [21] dan pendekatan berbasis kluster (*clustering*) [22] juga telah diusulkan untuk meningkatkan presisi pemodelan residu pada data yang sangat fluktuatif, seperti konsumsi bahan bakar [23] atau volume inspeksi kargo [24].

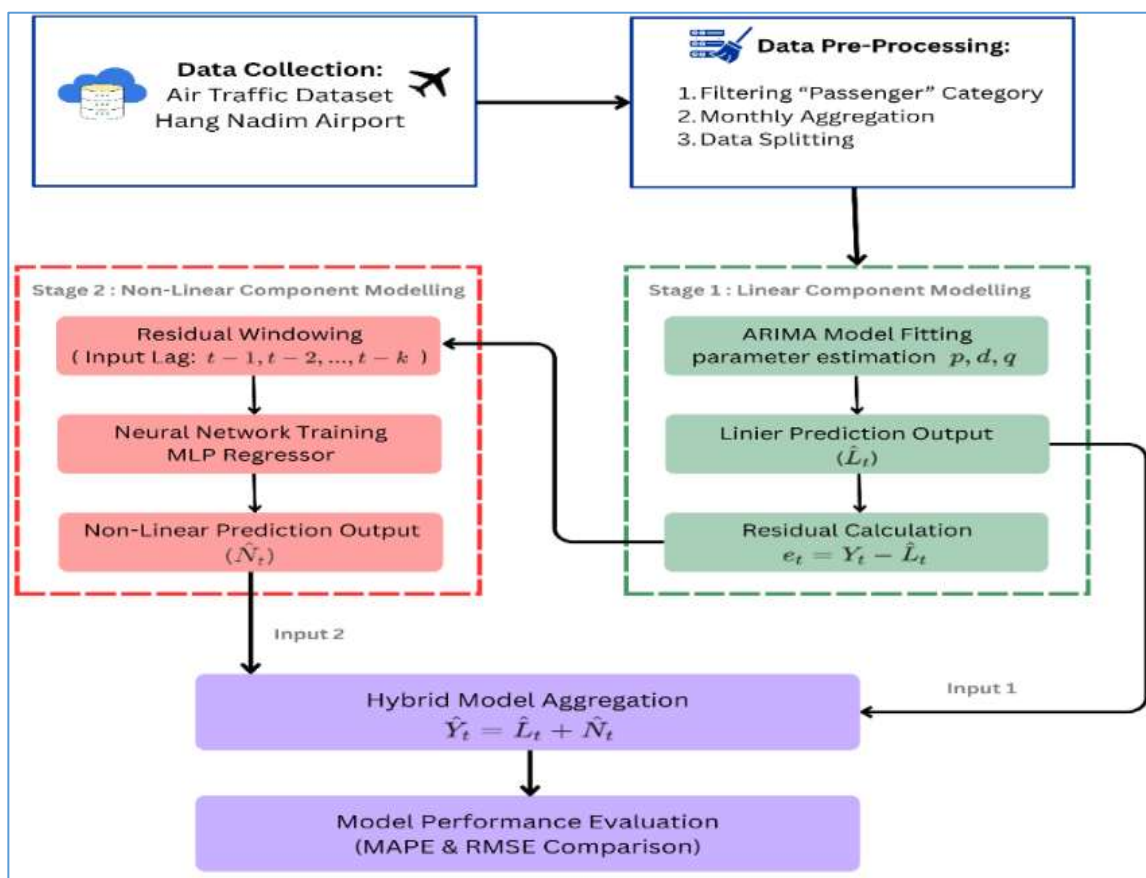
Meskipun berbagai studi telah membuktikan keunggulan arsitektur Hibrida ARIMA-NN, sebagian besar literatur terdahulu masih berfokus pada dataset penerbangan dari bandara komersial standar [25], [26], [27]. Masih terdapat celah penelitian (*research gap*) mengenai efektivitas skema koreksi residu (*residual windowing*) secara univariat pada bandara dengan karakteristik ganda (wisata dan industri). Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan menguji seberapa jauh arsitektur hibrida standar mampu beradaptasi dan mengoreksi anomali volatilitas ekstrem di kawasan berstatus khusus (*Free Trade Zone*).

3 Metode Penelitian

Bagian ini menguraikan kerangka metodologi peramalan arus penumpang menggunakan model Hibrida ARIMA-Neural Network. Penjelasan disusun secara sistematis mengikuti tahapan operasional, yang meliputi rancangan alur penelitian, pengumpulan dan prapemrosesan data, spesifikasi arsitektur pemodelan, serta metrik evaluasi kinerja.

3.1 Alur Penelitian

Penelitian ini dilaksanakan dengan mengikuti kerangka kerja sistematis yang dirancang untuk memastikan validitas dan reproduktivitas hasil peramalan. Secara garis besar, tahapan penelitian dimulai dari pengumpulan data sekunder, dilanjutkan dengan pra-pemrosesan data untuk menjamin kualitas *input*, kemudian masuk ke tahap pemodelan yang terdiri dari pemodelan komponen linear menggunakan ARIMA dan pemodelan komponen non-linear menggunakan *Neural Network* pada residu. Pendekatan sistematis ini mengacu pada kerangka kerja hibrida yang telah divalidasi dalam berbagai studi peramalan deret waktu [6], [10]. Tahap akhir meliputi agregasi model hibrida dan evaluasi kinerja model menggunakan metrik statistik standar. Rincian alur kerja yang diterapkan dalam penelitian ini divisualisasikan pada **Gambar 1**.



Gambar 1 Diagram alur penelitian

3.2 Pengumpulan dan Pra-pemrosesan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari dataset publik arus lalu lintas udara Bandara Internasional Hang Nadim, Batam. Fokus penelitian dibatasi pada variabel jumlah penumpang total (*Total Arrival + Departure*) guna menganalisis beban operasional bandara secara menyeluruh. Dataset ini mencakup rentang waktu bulanan dari **April 2016** hingga **Juni 2025**, dengan total observasi sebanyak **111** bulan. Atribut awal data meliputi tanggal, kategori (penumpang, bagasi, kargo), sub-jenis penerbangan, dan jumlah volume trafik.

Sebelum dilakukan pemodelan, data mentah melalui serangkaian proses pembersihan dan transformasi. Langkah pertama adalah penyaringan (*filtering*) data khusus untuk kategori penumpang, yang kemudian diagregasi dari data transaksional harian menjadi total bulanan (*monthly sum*) untuk membentuk struktur deret waktu yang teratur. Selanjutnya, dilakukan pemeriksaan terhadap kelengkapan data (*handling missing values*) untuk memastikan tidak terdapat indeks waktu yang kosong, mengingat kontinuitas data sangat krusial dalam analisis deret waktu [26]. Apabila ditemukan kekosongan data, teknik imputasi nilai nol atau interpolasi linear diterapkan. Setelah data bersih, dataset dibagi (*data splitting*) menjadi dua bagian dengan rasio 80:20, di mana **80%** data awal (**sebanyak 88 bulan, periode April 2016 – Juli 2023**) dialokasikan sebagai data latih (*training set*) untuk pembentukan model, dan **20%** sisanya (**sebanyak 23 bulan, periode Agustus 2023 – Juni 2025**) digunakan sebagai data uji (*testing set*) untuk mengukur akurasi model pada data yang belum pernah dilihat sebelumnya (*unseen data*). Skema pembagian ini umum digunakan dalam studi peramalan untuk menyeimbangkan proses pembelajaran dan validasi [11], [19].

3.3 Arsitektur Model Hibrida ARIMA-Neural Network

Metode utama yang diusulkan dalam penelitian ini adalah model hibrida seri (*series hybrid*) yang mengintegrasikan metode ARIMA dan *Neural Network*. Pendekatan ini didasarkan pada asumsi bahwa data deret waktu arus penumpang (Y_t) tersusun atas dua komponen utama, yaitu komponen linear (L_t) dan komponen non-linear (N_t) [17], [25]. Hubungan matematis komponen tersebut dinyatakan dalam **Persamaan (1)**.

$$Y_t = L_t + N_t \quad (1)$$

Proses pemodelan diawali dengan Tahap 1, yaitu pemodelan struktur linear menggunakan ARIMA (p, d, q). Pada tahap ini, uji stasioneritas dilakukan menggunakan *Augmented Dickey-Fuller* (ADF) untuk menentukan proses *differencing* (d). Selanjutnya, Penentuan ordo *Auto Regressive* (p) dan *Moving Average* (q) yang paling optimal dilakukan melalui algoritma pencarian otomatis (*hyperparameter tuning*) menggunakan fungsi *auto_arima* dengan mengevaluasi nilai *Akaike Information Criterion* (AIC) terendah. Analisis visual melalui plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) tetap disertakan sebagai instrumen validasi untuk mengonfirmasi kelayakan model terpilih. Model ARIMA terbaik kemudian digunakan untuk menghasilkan prediksi linear (\hat{L}_t). Selisih antara data aktual dan prediksi linear kemudian dihitung sebagai residu (e_t) yang diasumsikan mengandung informasi pola non-linear [14], sebagaimana dirumuskan dalam **Persamaan (2)**.

$$e_t = Y_t - \hat{L}_t \quad (2)$$

Selanjutnya pada Tahap 2, residu (e_t) yang dihasilkan diolah menggunakan *Multi-Layer Perceptron* (MLP) *Regressor*. Data residu ditransformasi menjadi format *supervised learning* menggunakan teknik *sliding window*. Guna mengoptimalkan arsitektur komputasi, konfigurasi parameter *Neural Network* yang digunakan dalam penelitian ini ditetapkan berdasarkan pengujian empiris dan dirangkum pada Tabel 1.

Tabel 1 Spesifikasi parameter neural network

Parameter	Spesifikasi
Model Jaringan	<i>Multi-Layer Perceptron</i> (MLP) <i>Regressor</i>
Teknik Input	<i>Sliding Window</i> (Lag k)
Penskalaan Data	<i>MinMax Scaler</i> [-1, 1]

Hidden Layer	2 lapis (<i>layer</i>)
Jumlah Neuron	50 neuron per hidden layer
Fungsi Aktivasi	Hyperbolic Tangent (Tanh)
Algoritma Optimasi	Adam (<i>Adaptive Moment Estimation</i>)
Batas Iterasi	500 - 1000 epoch

Pada Tahap 3, hasil akhir peramalan hibrida (\hat{Y}_t) diperoleh dengan menjumlahkan kembali hasil prediksi linear dari ARIMA dan hasil prediksi residu non-linear dari *Neural Network*, seperti ditunjukkan pada **Persamaan (3)**. Metode koreksi kesalahan ini terbukti meningkatkan akurasi dibandingkan metode tunggal [27].

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \quad (3)$$

3.4 Evaluasi Kinerja Model

Untuk mengukur tingkat keakuratan dan keandalan model yang diusulkan, penelitian ini menggunakan dua metrik evaluasi standar yang banyak diadopsi dalam literatur transportasi udara [1], [10]. Metrik pertama adalah *Mean Absolute Percentage Error* (MAPE) yang mengukur rata-rata persentase kesalahan absolut, yang dipilih karena kemudahannya untuk diinterpretasikan dalam konteks manajemen kapasitas bandara. Perhitungan MAPE ditunjukkan pada **Persamaan (4)**.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100\% \quad (4)$$

Metrik kedua adalah *Root Mean Square Error* (RMSE) yang digunakan untuk mengukur akar rata-rata kuadrat kesalahan. RMSE memberikan penalti yang lebih besar pada kesalahan prediksi yang ekstrem, sehingga efektif untuk mendeteksi penyimpangan besar pada model [9]. Rumus RMSE disajikan pada **Persamaan (5)**.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (5)$$

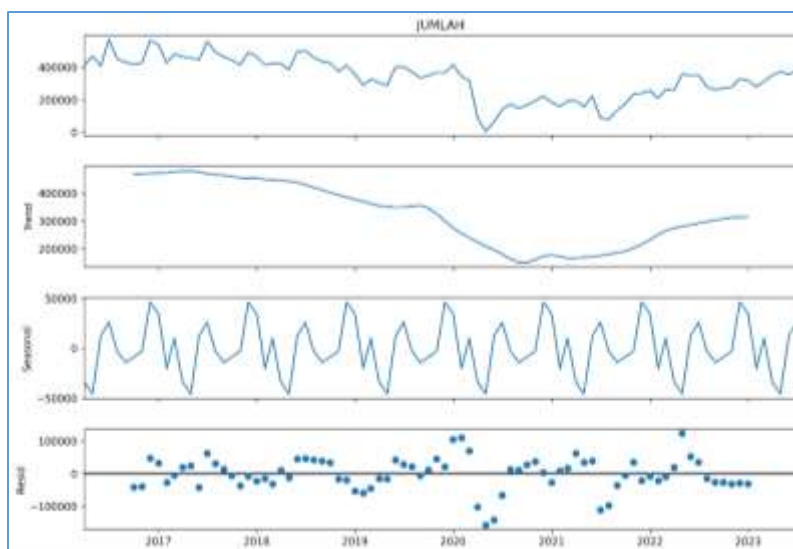
Selain evaluasi kuantitatif menggunakan kedua metrik di atas, analisis visual juga dilakukan dengan membandingkan plot garis prediksi model Hibrida terhadap data aktual untuk memvalidasi kemampuan model dalam mengikuti fluktuasi pola data (*pattern matching*).

4 Hasil dan Pembahasan

Bab ini memaparkan hasil eksperimen penerapan metode Hibrida ARIMA-*Neural Network* pada data arus lalu lintas penumpang Bandara Hang Nadim. Pembahasan menstrukturkan temuan empiris mulai dari analisis pola data, kinerja pemodelan komponen linear dan non-linear, hingga evaluasi komparatif akhir untuk memvalidasi efektivitas model usulan.

4.1 Analisis Data Eksploratif (*Exploratory Data Analysis*)

Langkah awal dalam penelitian ini melibatkan analisis dekomposisi untuk memahami karakteristik intrinsik dari deret waktu arus penumpang. Sebagaimana divisualisasikan pada Gambar 2, data arus penumpang Bandara Hang Nadim menunjukkan pola musiman (*seasonality*) yang kuat disertai dengan tren fluktuatif yang dinamis. Lonjakan penumpang secara konsisten teramati pada periode pertengahan tahun, yang berkaitan dengan libur sekolah dan Hari Raya Idul Fitri, serta pada akhir tahun. Namun, amplitudo lonjakan tersebut tidak selalu konsisten dari tahun ke tahun, yang mengindikasikan adanya komponen non-linear yang kompleks akibat pengaruh faktor eksternal.

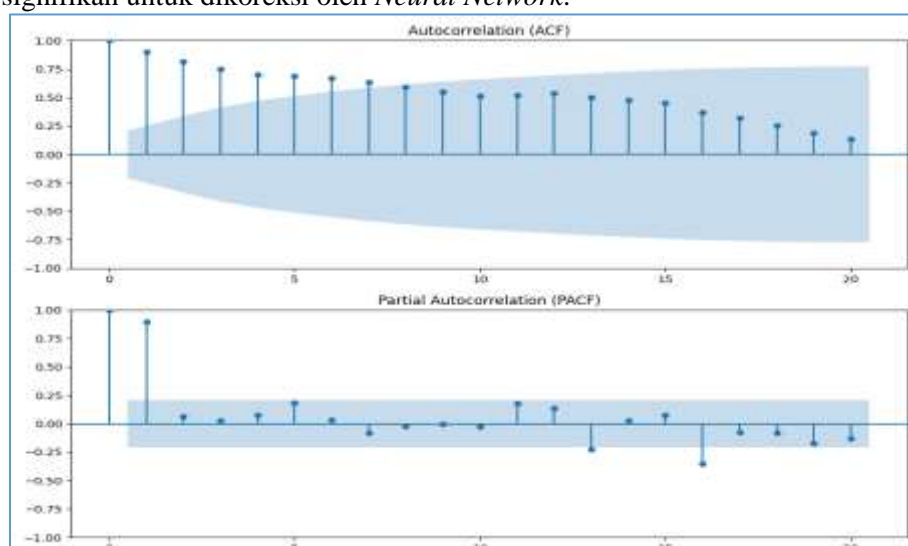


Gambar 2 Dekomposisi data arus penumpang (tren, musiman, dan residu)

Fenomena volatilitas musiman yang ditunjukkan pada Gambar 2 sejalan dengan temuan dalam literatur [10], [13] yang menyatakan bahwa lalu lintas transportasi udara sangat rentan terhadap variasi kalender liburan dan kondisi ekonomi makro. Kompleksitas pola ini menegaskan bahwa penggunaan metode statistik murni mungkin tidak memadai untuk menangkap seluruh dinamika data, sehingga diperlukan pendekatan pemodelan yang mampu mengakomodasi hubungan non-linear dan perubahan pola secara dinamis.

4.2 Pemodelan Komponen Linear dengan ARIMA

Berdasarkan hasil pencarian otomatis, struktur linear data paling optimal dimodelkan menggunakan ordo ARIMA (0, 1, 0). Ordo yang ekuivalen dengan model *Random Walk* ini menunjukkan ketiadaan komponen autokorelasi ($p = 0$) dan *moving average* ($q = 0$). Konfigurasi ini secara matematis menjelaskan mengapa prediksi baseline lambat mengantisipasi titik balik saat terjadi lonjakan ekstrem; model pada dasarnya hanya memproyeksikan nilai observasi terakhir, sehingga menyisakan residu non-linear yang signifikan untuk dikoreksi oleh *Neural Network*.



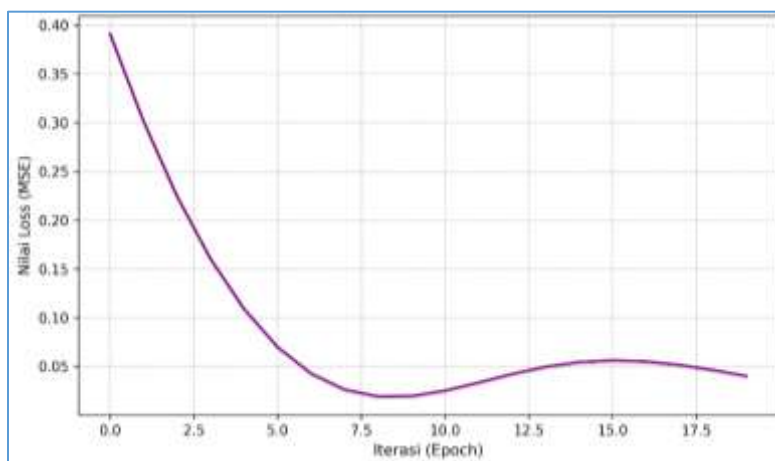
Gambar 3 Plot autocorrelation function (ACF) dan partial ACF

Hasil ini mengonfirmasi keterbatasan fundamental ARIMA yang sering dikutip dalam studi terdahulu [4], [6], di mana model ini cenderung gagal menangkap titik balik (*turning points*) ekstrem pada data dengan volatilitas tinggi. Prediksi ARIMA membentuk garis yang relatif datar (*flat forecast*)

dan tidak mampu mengikuti lonjakan tajam pada periode *peak season*. Secara kuantitatif, keterbatasan ini tercermin dari nilai MAPE sebesar 23,75% dan RMSE 106.374,42. Tingkat kesalahan yang relatif tinggi ini mengindikasikan bahwa masih terdapat informasi pola data yang signifikan yang tertinggal dalam sisaan (residu), yang memerlukan penanganan lebih lanjut menggunakan pendekatan non-linear.

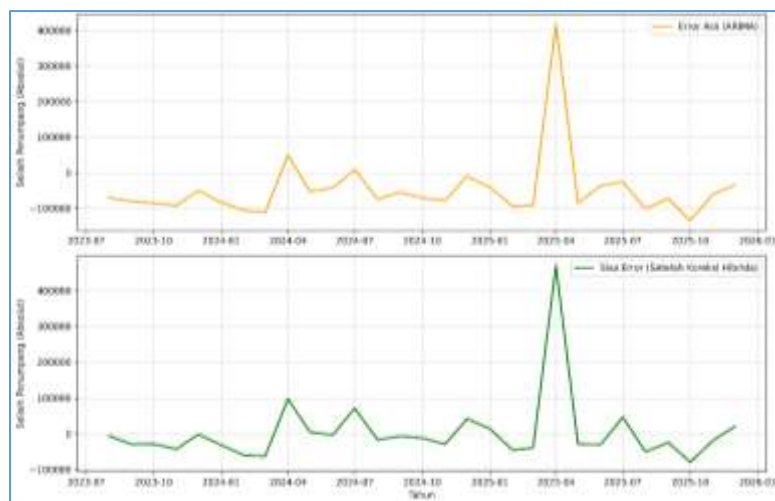
4.3 Pemodelan Komponen Non-Linear dengan Neural Network

Untuk mengatasi kelemahan model linear, residu (e_t) yang diekstrak dari ARIMA kemudian dijadikan target pembelajaran bagi *Neural Network*. Proses pelatihan (*training*) bertujuan meminimalkan fungsi kerugian (*loss function*) agar jaringan mampu mengenali pola kesalahan masa lalu.



Gambar 4 Kurva konvergensi loss pada pelatihan neural network

Pada **Gambar 4** memperlihatkan kurva konvergensi *loss* yang menurun tajam di awal iterasi dan stabil menuju titik optimal, menandakan bahwa model berhasil mempelajari struktur tersembunyi dari data residu.

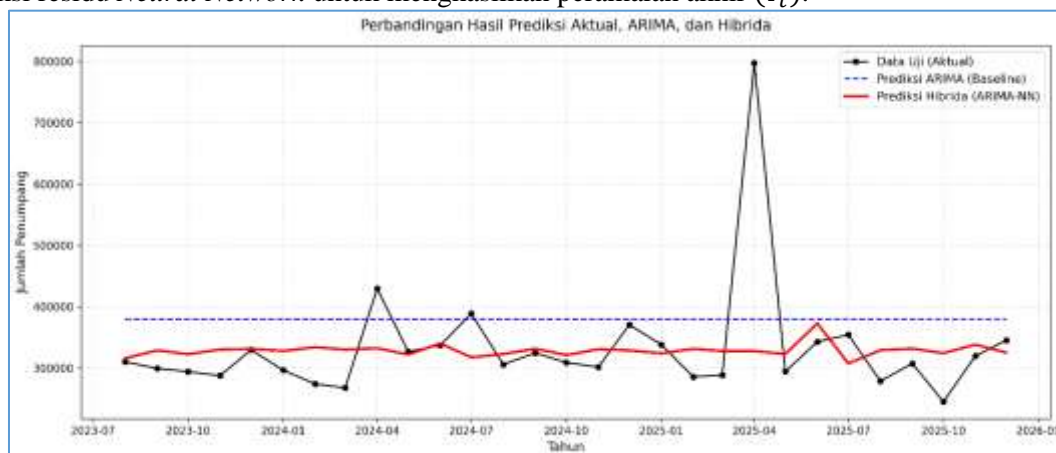


Gambar 5 Analisis kemampuan neural network dalam mengoreksi error

Validasi kemampuan Neural Network dalam memodelkan ketidakpastian ini diperlihatkan pada Gambar 5 (Analisis Kemampuan NN Mengoreksi Error). Gambar tersebut membandingkan fluktuasi error asli bawaan ARIMA (panel atas/oranye) dengan sisa error setelah dikoreksi oleh Neural Network (panel bawah/hijau). Terlihat jelas bahwa amplitudo sisa error pada panel bawah berhasil ditekan mendekati garis nol. Hal ini membuktikan bahwa Neural Network berhasil mengekstraksi informasi non-linear dari residu masa lalu untuk secara aktif mengoreksi kelemahan ARIMA pada setiap langkah peramalan (*one-step-ahead*). Kemampuan “koreksi kesalahan” ini merupakan inti dari keunggulan model hibrida, sebagaimana dijelaskan dalam penelitian seminal [16] dan didukung oleh studi terkini [5].

4.4 Evaluasi dan Komparasi Model Hibrida

Tahap akhir penelitian ini adalah agregasi model, di mana prediksi linear ARIMA dijumlahkan dengan prediksi residu *Neural Network* untuk menghasilkan peramalan akhir (\hat{Y}_t).



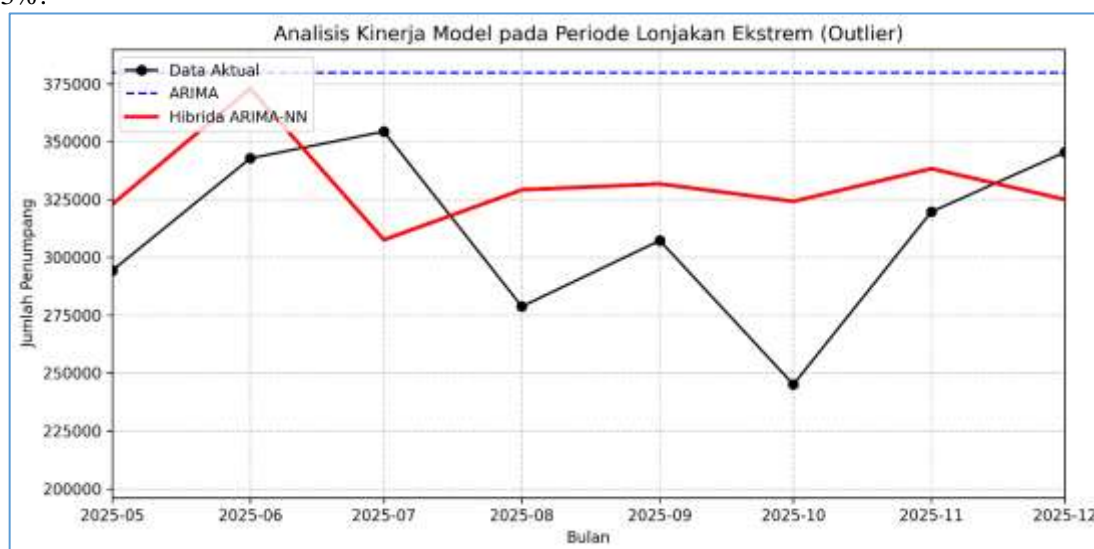
Gambar 6 Perbandingan hasil prediksi aktual, ARIMA, dan hybrid

Perbandingan visual antara data aktual, model ARIMA, dan model Hibrida disajikan pada Gambar 6. Secara visual, model Hibrida (garis merah) menunjukkan peningkatan performa yang signifikan dibandingkan ARIMA (garis biru). Model Hibrida mampu merekonstruksi pola gelombang naik-turun yang sebelumnya hilang pada model *baseline*, sehingga hasil prediksi menjadi jauh lebih mendekati data aktual.

Tabel 2 Perbandingan akurasi model

Metode	RMSE	MAPE
ARIMA (Baseline)	106,374.42	23.75%
Hybrid ARIMA-NN (Usulan)	96,198.25	12.51%
Peningkatan Kinerja	-	47,33%

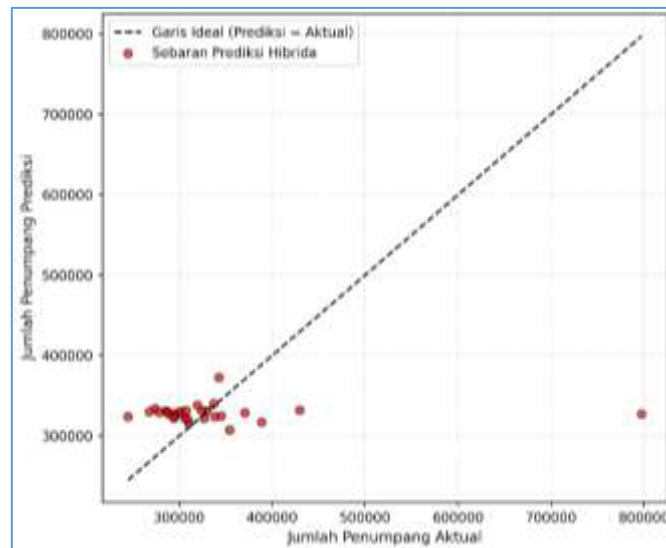
Evaluasi statistik yang dirangkum dalam Tabel 2 memperkuat temuan visual tersebut. Penerapan metode hibrida berhasil menurunkan tingkat kesalahan persentase (MAPE) dari 23,75% menjadi 12,51%. Penurunan nilai MAPE ini setara dengan peningkatan akurasi (*error reduction*) sebesar 47,33%.



Gambar 7 Analisis kinerja pada periode outlier

Lebih lanjut, Gambar 7 menampilkan analisis *zoom-in* pada periode dengan volatilitas ekstrem (*outlier*). Meskipun model Hibrida berhasil mengungguli ARIMA, visualisasi ini mengungkapkan bahwa pada titik lonjakan yang sangat tidak wajar (*external shock*), Hibrida masih menghasilkan selisih error absolut yang cukup terlihat. Fenomena inilah yang secara matematis menjelaskan mengapa metrik RMSE (yang sangat sensitif memberikan penalti kuadratik terhadap *outlier*) hanya mengalami penurunan moderat dari 106.374 menjadi 96.198, meskipun rata-rata persentase kesalahan (MAPE) secara keseluruhan telah turun sangat drastis.

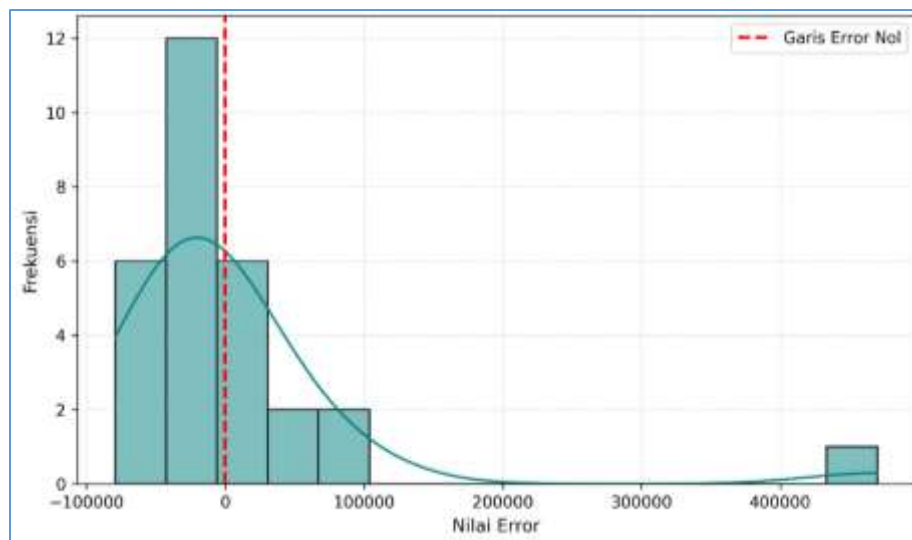
Peningkatan kinerja ini konsisten dengan penelitian [9] dan [11] yang juga melaporkan penurunan error signifikan setelah menerapkan mekanisme koreksi residu pada data transportasi. Korelasi positif yang kuat antara prediksi hibrida dan data aktual juga terkonfirmasi melalui *Scatter Plot* pada Gambar 8, di mana titik-titik persebaran data berkumpul di sekitar garis diagonal, menunjukkan reliabilitas model yang tinggi.



Gambar 8 Scatter plot korelasi data aktual vs prediksi hibrida

Pencapaian reduksi *error* sebesar 47,33% dalam penelitian ini melampaui tingkat signifikansi pada literatur terdahulu. Sebagai perbandingan, model kombinasi pada penelitian Gu et al. [10] berfokus pada guncangan makro berskala nasional akibat pandemi, sementara penerapan SARIMA-ANN oleh Yollanda & Devianto [11] pada sektor pariwisata hanya menghasilkan penurunan MAPE yang sangat marginal (dari 17,17% menjadi 17,09%). Hal ini membuktikan kebaruan (*novelty*) bahwa mekanisme koreksi *residual windowing* univariat bekerja jauh lebih masif dan responsif ketika dihadapkan pada anomali volatilitas ganda khas kawasan *Free Trade Zone* (perpaduan arus bisnis dan wisata).

Fenomena “garis datar” pada prediksi ARIMA membuktikan model statistik konvensional “buta” terhadap guncangan eksternal kawasan industri. Sebaliknya, mekanisme saling melengkapi (*complementary*) dari *Neural Network* saat terjadi *under-fitting* menjadikan model hibrida ini solusi komputasional yang lebih tangguh (*robust*). Secara praktis, penurunan margin kesalahan hingga ~12% ini dapat langsung diaplikasikan oleh manajemen Bandara Hang Nadim sebagai landasan sistem pendukung keputusan untuk: pertama, **Optimalisasi SDM** yaitu mengatur jadwal staf bandara secara presisi untuk mencegah bottleneck pelayanan saat lonjakan ekstrem. Kedua, **Kapasitas Terminal**, Memitigasi penumpukan penumpang melalui manajemen slot time maskapai secara dinamis. Ketiga, **Efisiensi Biaya**, Meminimalkan pemborosan finansial (*over-capacity*) akibat alokasi sumber daya berlebih pada periode low season yang gagal dideteksi oleh model linear.



Gambar 9 Distribusi residu

Sebagai tahap validasi akhir terhadap kelayakan model secara statistik, Gambar 9 menampilkan histogram distribusi sisa kesalahan (residu akhir) yang dihasilkan oleh prediksi Hibrida ARIMA-NN. Berdasarkan kurva kepadatan probabilitas yang terbentuk, dapat diobservasi bahwa distribusi residu cenderung simetris dan membentuk pola menyerupai lonceng (*bell-shaped*), dengan pemusatan data (*mean*) yang sangat dekat dengan angka nol (ditunjukkan oleh garis vertikal). Karakteristik distribusi ini merupakan indikator kuat bahwa sisaan model telah mendekati asumsi *white noise*, di mana tidak terdapat lagi pola autokorelasi sistematis yang tertinggal di dalam data. Dengan kata lain, agregasi komponen linear dari ARIMA dan kemampuan koreksi non-linear dari Neural Network telah berhasil mengekstraksi secara optimal seluruh informasi historis dari deret waktu penumpang Bandara Hang Nadim. Sisa penyimpangan yang terbentuk murni merepresentasikan gangguan acak (*random noise* atau *unpredictable shocks*) dari lingkungan eksternal FTZ Batam yang memang di luar batas kemampuan pemodelan univariat.

5 Kesimpulan

Penelitian ini mengonfirmasi bahwa model Hibrida ARIMA-*Neural Network* efektif mengatasi keterbatasan ARIMA dalam meramalkan volatilitas ekstrem arus penumpang di kawasan *Free Trade Zone* Bandara Hang Nadim. Kontribusi utama riset ini dibuktikan melalui keberhasilan *Neural Network* mengoreksi pola residu *non-linear* secara *one-step-ahead*, yang menghasilkan penurunan *Mean Absolute Percentage Error* (MAPE) secara drastis dari 23,75% menjadi 12,51%. Hasil kunci ini merepresentasikan peningkatan akurasi (*error reduction*) sebesar 47,33% dibandingkan model dasar. Meski demikian, penurunan moderat pada metrik RMSE (dari 106.374 menjadi 96.198) menunjukkan bahwa arsitektur univariat ini masih memiliki kerentanan terhadap guncangan eksternal yang mendadak. Oleh karena itu, penelitian selanjutnya direkomendasikan untuk mengintegrasikan variabel eksogen (seperti efek *moving holiday* dan indikator makroekonomi) ke dalam pemodelan multivariat, serta mengeksplorasi arsitektur *Deep Learning* bermemori seperti *Long Short-Term Memory* (LSTM) guna mempertajam deteksi anomali jangka panjang.

Referensi

- [1] W. Zeng, J. Li, Z. Quan, and X. Lu, "A Deep Graph-Embedded LSTM Neural Network Approach for Airport Delay Prediction," *J. Adv. Transp.*, vol. 2021, pp. 1–15, Mar. 2021, doi: 10.1155/2021/6638130.
- [2] C. Wu, H. Ding, Z. Fu, and N. Sun, "Air Traffic Flow Prediction in Aviation Networks Using a Multi-Dimensional Spatiotemporal Framework," *Electronics*, vol. 13, no. 19, p. 3803, Sep. 2024, doi: 10.3390/electronics13193803.
- [3] W. Tian, Y. Zhang, Y. Zhang, H. Chen, and W. Liu, "A Short-Term Traffic Flow Prediction Method for Airport Group Route Waypoints Based on the Spatiotemporal Features of Traffic

- Flow,” *Aerospace*, vol. 11, no. 4, p. 248, Mar. 2024, doi: 10.3390/aerospace11040248.
- [4] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, “A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks,” *Futur. Internet*, vol. 15, no. 8, p. 255, Jul. 2023, doi: 10.3390/fi15080255.
- [5] A. Atesongun and M. Gulsen, “A Hybrid Forecasting Structure Based on Arima and Artificial Neural Network Models,” *Appl. Sci.*, vol. 14, no. 16, p. 7122, Aug. 2024, doi: 10.3390/app14167122.
- [6] M. Milenković, M. Gligorić, N. Bojović, and Z. Gligorić, “A comparison between ARIMA, LSTM, ARIMA-LSTM and SSA for cross-border rail freight traffic forecasting: the case of Alpine-Western Balkan Rail Freight Corridor,” Dec. 08, 2022. doi: 10.21203/rs.3.rs-2342441/v1.
- [7] K.-L. Li, C.-J. Zhai, and J.-M. Xu, “Short-term traffic flow prediction using a methodology based on ARIMA and RBF-ANN,” in *2017 Chinese Automation Congress (CAC)*, IEEE, Oct. 2017, pp. 2804–2807. doi: 10.1109/CAC.2017.8243253.
- [8] R. V and G. V. S, “Hybrid Time-Series Forecasting Models for Traffic Flow Prediction,” *Promet - Traffic&Transportation*, vol. 34, no. 4, pp. 537–549, Jul. 2022, doi: 10.7307/ptt.v34i4.3998.
- [9] R. Zhou, S. Qiu, M. Li, S. Meng, and Q. Zhang, “Short-Term Air Traffic Flow Prediction Based on CEEMD-LSTM of Bayesian Optimization and Differential Processing,” *Electronics*, vol. 13, no. 10, p. 1896, May 2024, doi: 10.3390/electronics13101896.
- [10] W. Gu, B. Guo, Z. Zhang, and H. Lu, “Civil Aviation Passenger Traffic Forecasting: Application and Comparative Study of the Seasonal Autoregressive Integrated Moving Average Model and Backpropagation Neural Network,” *Sustainability*, vol. 16, no. 10, p. 4110, May 2024, doi: 10.3390/su16104110.
- [11] M. Yollanda and D. Devianto, “Hybrid Model of Seasonal ARIMA-ANN to Forecast Tourist Arrivals through Minangkabau International Airport,” in *Proceedings of the Proceedings of the 1st International Conference on Statistics and Analytics, ICSA 2019, 2-3 August 2019, Bogor, Indonesia*, EAI, 2020. doi: 10.4108/eai.2-8-2019.2290473.
- [12] W. Du, B. Li, J. Chen, Y. Lv, and Y. Li, “A Spatiotemporal Hybrid Model for Airspace Complexity Prediction,” *IEEE Intell. Transp. Syst. Mag.*, vol. 15, no. 2, pp. 217–224, Mar. 2023, doi: 10.1109/MITS.2022.3204099.
- [13] Y. Wang, R. Jia, F. Dai, and Y. Ye, “Traffic Flow Prediction Method Based on Seasonal Characteristics and SARIMA-NAR Model,” *Appl. Sci.*, vol. 12, no. 4, p. 2190, Feb. 2022, doi: 10.3390/app12042190.
- [14] D. S. de O. Santos Júnior, P. S. G. de Mattos Neto, J. F. L. de Oliveira, and G. D. C. Cavalcanti, “A hybrid system based on ensemble learning to model residuals for time series forecasting,” *Inf. Sci. (Ny)*, vol. 649, p. 119614, Nov. 2023, doi: 10.1016/j.ins.2023.119614.
- [15] M. Panja, T. Chakraborty, U. Kumar, and A. Hadid, “Probabilistic AutoRegressive Neural Networks for Accurate Long-Range Forecasting,” 2024, pp. 457–477. doi: 10.1007/978-981-99-8178-6_35.
- [16] M. Khashei and M. Bijari, “An artificial neural network (p,d,q) model for timeseries forecasting,” *Expert Syst. Appl.*, vol. 37, no. 1, pp. 479–489, Jan. 2010, doi: 10.1016/j.eswa.2009.05.044.
- [17] H. Rabbouch, H. Saadaoui, and F. Saadaoui, “VMD-based Multiscaled LSTM-ARIMA to Forecast Post-COVID-19 US Air Traffic,” in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, IEEE, Mar. 2022, pp. 1678–1683. doi: 10.1109/DASA54658.2022.9765132.
- [18] E. E. Misengo, D. D. Prastyo, and H. Kuswanto, “Modeling and forecasting monthly tourist arrivals to the United States and Indonesia using ARIMA hybrids of multilayer perceptron models,” 2023, p. 080001. doi: 10.1063/5.0105680.
- [19] B. Naheliya, P. Redhu, and K. Kumar, “A Hybrid Deep Learning Method for Short-Term Traffic Flow Forecasting: GSA-LSTM,” *Indian J. Sci. Technol.*, vol. 16, no. 46, pp. 4358–4368, Dec. 2023, doi: 10.17485/IJST/v16i46.2520.
- [20] A. R. Sattarzadeh, R. J. Kutadinata, P. N. Pathirana, and V. T. Huynh, “A novel hybrid deep learning model with ARIMA Conv-LSTM networks and shuffle attention layer for short-term traffic flow prediction,” *Transp. A Transp. Sci.*, vol. 21, no. 1, Jan. 2025, doi: <http://sistemasi.ftik.unisi.ac.id>

- 10.1080/23249935.2023.2236724.
- [21] S. Panigrahi and H. S. Behera, "An Adaptive Fuzzy Filter-Based Hybrid ARIMA-HONN Model for Time Series Forecasting," 2019, pp. 841–850. doi: 10.1007/978-981-10-8055-5_74.
- [22] B. S. Shawel, T. T. Debella, G. Tesfaye, Y. Y. Tefera, and D. H. Woldegebreal, "Hybrid Prediction Model for Mobile Data Traffic: A Cluster-level Approach," in *2020 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Jul. 2020, pp. 1–8. doi: 10.1109/IJCNN48605.2020.9207655.
- [23] R. Babazadeh, "A Hybrid ARIMA-ANN approach for optimum estimation and forecasting of gasoline consumption," *RAIRO - Oper. Res.*, vol. 51, no. 3, pp. 719–728, Jul. 2017, doi: 10.1051/ro/2016059.
- [24] J. J. Ruiz-Aguilar, I. J. Turias, and M. J. Jiménez-Come, "A novel three-step procedure to forecast the inspection volume," *Transp. Res. Part C Emerg. Technol.*, vol. 56, pp. 393–414, Jul. 2015, doi: 10.1016/j.trc.2015.04.024.
- [25] Y. Bao, D. Yi, X. Tao, H. Zhongyi, and Z. Shuai, "A Comparative Study on Hybrid Linear and Nonlinear Modeling Framework for Air Passenger Traffic Forecasting," *Int. J. Adv. Inf. Sci. Serv. Sci.*, vol. 3, no. 5, pp. 243–254, Jun. 2011, doi: 10.4156/aiss.vol3.issue5.28.
- [26] Q. Hou, J. Leng, G. Ma, W. Liu, and Y. Cheng, "An adaptive hybrid model for short-term urban traffic flow prediction," *Phys. A Stat. Mech. its Appl.*, vol. 527, p. 121065, Aug. 2019, doi: 10.1016/j.physa.2019.121065.
- [27] Lili Qu, Yan Chen, and Zhenfeng Liu, "Time Series Forecasting Model with Error Correction by Structure Adaptive RBF Neural Network," in *2006 6th World Congress on Intelligent Control and Automation*, IEEE, 2006, pp. 6831–6835. doi: 10.1109/WCICA.2006.1714408.