

MODEL ESTIMASI WAKTU TEMPUH MENGGUNAKAN PENDEKATAN PEMODELAN MATEMATIS DAN OPTIMASI

Aisyah Amalia Hamid¹, Anindya Restu Shafara², Siti Naia Hesti Rachmawati³, Anggraini Puspita Sari⁴,
 Sischa Wahyuning Tyas⁵

^{1,2,3,5} Program Studi Sains Data, Universitas Pembangunan Nasional ‘Veteran’ Jawa Timur

⁴ Program Studi Fakultas Ilmu Komputer, Universitas Pembangunan Nasional ‘Veteran’ Jawa Timur

Jl. Raya Rungkut Madya, Gunung Anyar – Surabaya, Indonesia

Email: ¹24083010034@student.upnjatim.ac.id, ²24083010039@student.upnjatim.ac.id,

³24083010047@student.upnjatim.ac.id ⁴anggraini.puspita.if@upnjatim.ac.id

⁵sischa_wahyuning.sada@upnjatim.ac.id

ABSTRAK

Estimasi waktu tempuh yang akurat sangat krusial bagi efisiensi sistem transportasi dan layanan logistik di perkotaan. Penelitian ini mengembangkan model estimasi waktu tempuh untuk wilayah Surabaya dengan mengintegrasikan metode interpolasi, regresi linear, dan teknik optimasi Newton-Raphson. Data yang digunakan bersumber dari rute OpenStreetMap serta variabel cuaca (curah hujan dan suhu) dari BMKG. Hasil analisis menunjukkan bahwa jarak tempuh, intensitas hujan, dan kondisi jam sibuk secara signifikan memengaruhi durasi perjalanan. Model ini memiliki tingkat akurasi yang tinggi dengan koefisien determinasi RSquare (R^2) sebesar 0,92. Adapun tingkat kesalahan model diukur melalui Mean Absolute Error (MAE) sebesar 201,05 detik (sekitar 3,3 menit) dan Root Mean Squared Error (RMSE) sebesar 267,53 detik. Melalui simulasi optimasi rute, model ini mampu memberikan saran perjalanan yang 8–15% lebih cepat dibandingkan strategi pemilihan rute konvensional berbasis jarak terpendek. Dengan hasil tersebut, model ini dapat diimplementasikan pada sistem navigasi adaptif dan responsif terhadap perubahan kondisi lingkungan dan lalu lintas.

Kata kunci: Estimasi waktu tempuh, Cuaca, Pendekatan numerik, Pemodelan matematis

ABSTRACT

Accurate travel time estimates are crucial for the efficiency of transportation systems and logistics services in urban areas. This study developed a travel time estimation model for the Surabaya area by integrating interpolation, linear regression, and Newton-Raphson optimization techniques. The data used was sourced from OpenStreetMap routes and weather variables (rainfall and temperature) from the BMKG. The results of the analysis show that travel distance, rainfall intensity, and rush hour conditions significantly affect travel duration. This model has a high level of accuracy with a coefficient of determination R Square (R^2) of 0.92. The model's error rate is measured by the Mean Absolute Error (MAE) of 201.05 seconds (approximately 3.3 minutes) and the Root Mean Squared Error (RMSE) of 267.53 seconds. Through route optimization simulations, this model is able to provide travel suggestions that are 8–15% faster than conventional route selection strategies based on the shortest distance. With these results, this model can be implemented in adaptive navigation and responsive systems that respond to changes in environmental and traffic conditions.

Keywords: Travel time estimation, Weather effects, Numerical approach, Mathematical modeling

1. PENDAHULUAN

Pertumbuhan layanan transportasi berbasis aplikasi di berbagai kota besar Indonesia, termasuk Surabaya, semakin menegaskan urgensi penyediaan estimasi waktu tempuh yang presisi [1]. Akurasi estimasi tersebut berperan krusial dalam meningkatkan kenyamanan pengguna, efisiensi perjalanan, serta mendukung proses pengambilan keputusan operasional pada sistem transportasi modern [2]. Beragam faktor memengaruhi durasi perjalanan, di antaranya karakteristik jaringan jalan, panjang rute, kondisi cuaca, dan tingkat kepadatan lalu lintas [3]. Hubungan antara arus lalu lintas, kecepatan kendaraan, dan waktu tempuh telah dijelaskan dalam sejumlah studi transportasi di Indonesia, yang menunjukkan bahwa dinamika lalu lintas merupakan komponen dominan dalam membentuk variasi waktu tempuh aktual [4].



Perilaku perjalanan dan durasi waktu tempuh tidak hanya dipengaruhi oleh kondisi lalu lintas, tetapi juga dipengaruhi secara signifikan oleh kondisi cuaca. Studi di Yogyakarta mengonfirmasi bahwa kondisi seperti hujan, panas, dan angin dapat mengubah preferensi perjalanan serta memengaruhi keputusan pemilihan model [2]. Penelitian lain menegaskan bahwa cuaca merupakan variabel penting dalam pemodelan waktu tempuh, baik melalui pengaruh langsung maupun melalui interaksinya dengan tingkat kepadatan lalu lintas [4]. Di samping itu, pengembangan sistem penentuan rute berbasis kondisi cuaca telah dilakukan melalui pendekatan *fuzzy* maupun model probabilistik, mempertimbangkan kendaraan dan kondisi lingkungan dalam menentukan rute optimal [5].

Kajian pemodelan waktu tempuh di Indonesia masih banyak berfokus pada pendekatan algoritmik berbasis graf, seperti pemanfaatan Algoritme Dijkstra untuk menentukan rute tercepat menuju halte TransJakarta [6]. Sementara itu, formulasi matematis berbasis metode numerik masih relatif terbatas, meskipun beberapa studi mulai mengeksplorasi teknik seperti interpolasi dan deret Taylor untuk estimasi waktu tempuh berjalan kaki. Namun, integrasi variabel cuaca secara eksplisit dalam model numerik belum memperoleh perhatian memadai dalam literatur transportasi [7]. Studi internasional dalam rentang tahun 2020–2025 menunjukkan kecenderungan peningkatan penggunaan model *spatio-temporal* dan *machine learning* dalam estimasi waktu tempuh perjalanan. Pendekatan tersebut banyak diterapkan untuk menangkap hubungan nonlinier antara jarak, kondisi lalu lintas, serta variabel lingkungan seperti cuaca, dengan hasil akurasi yang lebih tinggi dibandingkan model konvensional [8]. Namun, sebagian besar penelitian tersebut masih menitikberatkan performa algoritma dan evaluasi data empiris tanpa aspek formulasi matematis atau pendekatan numerik menjelaskan pengaruh antarfaktor dengan terukur.

Penelitian di Indonesia mulai mengeksplorasi pendekatan numerik dan algoritmik, seperti pemanfaatan *Support Vector Regression* dengan variabel cuaca [9] serta model stokastik dan algoritme genetika untuk optimasi arus lalu lintas [10]. Namun, untuk Integrasi eksplisit variabel cuaca ke dalam model numerik transportasi masih terbatas, sehingga menimbulkan kesenjangan penelitian dalam pengembangan model estimasi waktu tempuh yang mampu menjadi hasil representasi pengaruh kondisi cuaca dan lalu lintas secara sistematis dan terukur. Upaya untuk menanggapi kesenjangan tersebut, penelitian ini mengembangkan model estimasi waktu tempuh dengan memanfaatkan empat pendekatan numerik, yakni interpolasi, regresi Kuadrat Terkecil, metode iteratif Newton–Raphson, serta Pemrograman Linier [11]. Pendekatan numerik tersebut digunakan karena estimasi waktu tempuh merupakan permasalahan kuantitatif melibatkan hubungan fungsional antara jarak, kondisi cuaca, dan waktu perjalanan, sehingga dimodelkan dan dianalisis secara matematis. Dengan data jarak dari OpenStreetMap, sementara data cuaca (kondisi cuaca, suhu, dan kelembapan) dikumpulkan melalui observasi BMKG [12].

Kebaruan penelitian terletak pada integrasi model regresi linear dengan teknik numerik (Interpolasi B-Spline dan Newton-Raphson). Berbeda dengan estimasi navigasi konvensional, tidak hanya memprediksi waktu, tetapi juga mengidentifikasi titik transisi efisiensi perjalanan melalui optimasi matematis yang presisi. Penelitian ini menawarkan kontribusi utama berupa kerangka pemodelan numerik terintegrasi yang menggabungkan interpolasi spline, regresi kuadrat terkecil, metode Newton-Raphson, dan pemrograman linear. Keunikan pendekatan terletak pada kemampuan interpolasi spline dalam menangkap non-linearitas hubungan jarak-waktu pada segmen pendek yang sering diabaikan model konvensional, integrasi eksplisit variabel cuaca ke dalam model regresi yang memungkinkan kuantifikasi penalti waktu secara terukur, serta pemanfaatan optimasi Newton-Raphson untuk mengidentifikasi titik efisiensi perjalanan yang tidak dapat diungkap oleh analisis linear. Dengan demikian, penelitian ini tidak hanya menghasilkan model estimasi, tetapi juga menyediakan alat analisis untuk memahami dinamika waktu tempuh secara lebih mendalam.

Pemilihan pendekatan numerik dalam penelitian ini didasarkan pada kelebihan masing-masing metode dalam memodelkan hubungan kuantitatif antara jarak, kondisi cuaca, dan waktu tempuh. Interpolasi digunakan untuk membentuk hubungan jarak–waktu yang kontinu pada titik pengamatan yang tidak tersedia [13]. Regresi Kuadrat Terkecil dimanfaatkan untuk mengestimasi kontribusi variabel secara kuantitatif melalui meminimalkan galat kuadrat [14]. Metode iteratif Newton–Raphson digunakan karena efisiensi konvergensinya dalam menyelesaikan formulasi nonlinier [15]. Serta, Pemrograman Linier digunakan untuk memperoleh solusi optimal dengan meminimalkan waktu tempuh di bawah batasan tertentu [16]. Penelitian ini membangun data sintetis yang disusun berdasarkan sensitivitas terhadap jarak dan variabel cuaca sebagai pendekatan awal pengembangan model numerik, sebagaimana direkomendasikan untuk tahap awal pengembangan model numerik [17]. Penelitian ini bertujuan untuk memberikan kontribusi. Pertama, menghadirkan kerangka numerik yang mampu merepresentasikan pengaruh cuaca terhadap estimasi waktu tempuh secara sistematis dan terukur [18]. Kedua, menawarkan pendekatan yang stabil, transparan, serta dapat direplikasi pada kondisi dengan keterbatasan data empiris [19]. Ketiga, membuka peluang integrasi metode numerik dengan pendekatan optimisasi dalam mendukung pengembangan sistem transportasi cerdas pada skala perkotaan [20]. Sehingga, penelitian ini membuktikan bahwa pendekatan numerik mampu menangkap dinamika non-linear pada jarak pendek yang sering terabaikan oleh model linear murni, memberikan dimensi baru dalam optimasi rute perkotaan. Dengan gabungan pendekatan interpolasi spline untuk menangkap pola non-linear pada jarak pendek, regresi OLS untuk mengkuantifikasi pengaruh cuaca dan kemacetan, metode Newton-Raphson untuk menemukan titik optimum efisiensi perjalanan, serta pemrograman linear untuk optimasi pemilihan rute.

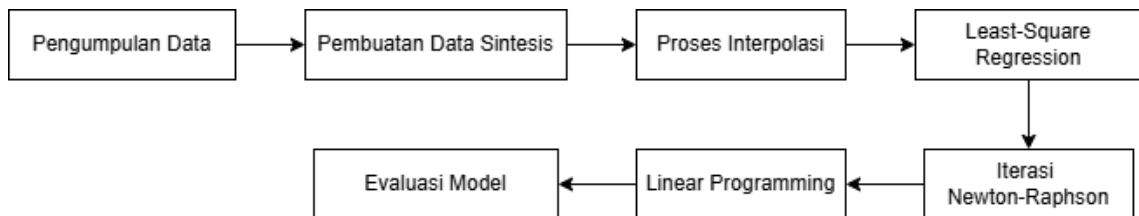
2. METODE PENELITIAN

Sumber Data

Dataset penelitian ini terdiri atas 500 perjalanan sintetis yang merepresentasikan mobilitas kendaraan di Kota Surabaya, dengan setiap observasi mencakup jarak tempuh, waktu perjalanan, kondisi cuaca, serta indikator *rush hour*. Data jarak rute diperoleh dari OpenStreetMap (OSM) melalui pembangkitan 500 pasangan titik asal–tujuan secara acak dalam batas administratif kota. Data cuaca bersumber dari laporan observasi harian Stasiun Meteorologi Juanda yang dipublikasikan oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) periode Januari–Desember 2024, meliputi kondisi cuaca kategorikal, suhu rata-rata harian (°C), dan kelembapan relatif (%). Periode *rush hour* ditetapkan pukul 06.00–09.00 dan 16.00–18.00 WIB pada hari kerja untuk keperluan validasi.

Data sintetis dibentuk dengan menghitung waktu tempuh berdasarkan kecepatan dasar pada kondisi cuaca cerah sebagai kondisi acuan. Kecepatan dasar ini tidak dimaknai sebagai kecepatan aktual tiap ruas jalan, melainkan diperlakukan sebagai nilai rata-rata representatif jaringan jalan perkotaan pada skala makro. Dengan menggunakan faktor penurunan bertahap dengan representasi kondisi cuaca, suhu udara, dan kelembapan yang relatif. Kemudian disesuaikan variasi waktu tempuh untuk jarak dalam kondisi lingkungan yang berbeda.

Sehingga menjaga konsistensi antara jarak, kecepatan, dan waktu tempuh. Keseluruhan tahapan pengolahan dan analisis data tersebut dirangkum dalam diagram alir penelitian yang ditampilkan pada Gambar 1.



Gambar 1. Diagram Alir Analisis

Tahapan-tahapan yang ditunjukkan pada diagram alir selanjutnya menjadi dasar penyusunan dalam variabel penelitian. Setelah proses pengumpulan dan pembentukan data sintetis dilakukan, langkah berikutnya adalah mengidentifikasi karakteristik serta peran tiap variabel dalam pemodelan. Oleh karena itu, ringkasan variabel yang digunakan dalam penelitian ini disajikan secara terstruktur dalam Tabel 1, Tabel 2, Tabel 3, dan Tabel 4.

Tabel 1. Karakteristik Dataset Penelitian

Karakteristik Dataset Penelitian	
Rentang jarak (km)	0,8 – 18,3
Median jarak (km)	6,1
Kondisi cuaca (kategori BMKG)	5
Suhu (°C)	26 – 32
Kelembapan (%)	57 – 92

Tabel 2. Data Jarak Rute Surabaya (OSM) – Nodes

osmid	y	x	street_count	highway	ref	railway	geometry
257690921	-7.258684	112.733157	3	NaN	NaN	NaN	POINT (112.73316 -7.25868)
257690940	-7.259299	112.742743	4	NaN	NaN	NaN	POINT (112.74274 -7.2593)
...
13361044883	-7.259238	112.752803	1	NaN	NaN	NaN	POINT (112.7528 -7.25924)
13368718420	-7.274835	112.741138	1	NaN	NaN	NaN	POINT (112.74114 -7.27484)

Tabel 3. Data Jarak Rute Surabaya (OSM) - Edges

u	v	key	osmid	highway	...	ref	tunnel
257690921	5579670363	0	1026679598	secondary	...	NaN	NaN
257690940	1682372817	0	361505909	secondary	...	NaN	NaN
...



13361044883	3651961063	0	1069455248	living_street	...	NaN	NaN
13368718420	4280987875	0	1457666081	living_street	...	NaN	NaN

Tabel 4. Data Cuaca Surabaya (Laporan Harian BMKG)

flex href	flex	flex 2	...	flex 30	flex 31
https://www.bmkg.go.id/cuaca/prakir-aan-cuaca/3...	Karang Pilang	Hujan Petir	...	23–23 °C	87–87 °C
https://www.bmkg.go.id/cuaca/prakir-aan-cuaca/3...	Wonocolo	Hujan Ringan	...	24–24 °C	85–85 °C
...
https://www.bmkg.go.id/cuaca/prakir-aan-cuaca/3...	Pakal	Hujan Petir	...	23–23 °C	88–88 °C
https://www.bmkg.go.id/cuaca/prakir-aan-cuaca/3...	Sambikerep	Hujan Petir	...	23–23 °C	88–88 °C

Pembuatan Data Sintetis

Pembuatan data sintetis dilakukan dengan merepresentasikan jaringan jalan Kota Surabaya berdasarkan data *OpenStreetMap* (OSM), di mana setiap segmen jalan memiliki atribut jarak dan kecepatan dasar. Waktu tempuh dihitung menggunakan kecepatan acuan pada kondisi cuaca cerah, kemudian disesuaikan dengan faktor penurunan yang merepresentasikan pengaruh kondisi cuaca, suhu udara, dan kelembapan sehingga menghasilkan bobot waktu tempuh yang bersifat dinamis. Struktur data jaringan disimpan dalam bentuk daftar ruas jalan yang saling terhubung berdasarkan koordinat asal-tujuan, sehingga memungkinkan perhitungan total waktu tempuh pada beberapa alternatif lintasan dan mendukung proses pemilihan rute berdasarkan waktu tempuh minimum.

Penggunaan data sintetis memungkinkan kontrol eksperimental yang ketat untuk menguji reliabilitas model matematis, meskipun belum sepenuhnya merepresentasikan anomali lapangan seperti kecelakaan atau penutupan jalan. Oleh karena itu, pada penerapan data riil diperlukan parameter tambahan seperti indeks kepadatan lalu lintas *real-time* untuk menjaga akurasi, sehingga model ini lebih ditujukan untuk analisis skenario dan perencanaan strategis dibandingkan prediksi operasional *real-time*.

Proses Analisis Numerik

Proses analisis numerik dilakukan melalui beberapa tahapan metode yang saling terintegrasi, meliputi interpolasi, regresi kuadrat terkecil, metode Newton–Raphson, dan pemrograman linear. Setiap metode digunakan untuk menjawab tujuan analisis yang berbeda, mulai dari estimasi nilai, pemodelan hubungan antarvariabel, hingga optimasi keputusan. Uraian dari proses analisis numerik dijelaskan sebagai berikut:

1. Interpolasi

Tahap interpolasi dilakukan untuk mengestimasi nilai waktu tempuh pada jarak tertentu yang tidak tersedia dalam data sintetis. Proses ini bertujuan untuk membentuk kurva hubungan jarak-waktu tempuh yang kontinu dan halus serta terhindar dari osilasi yang tidak diinginkan [21]. Data jarak dan waktu tempuh digunakan sebagai titik diskrit, kemudian dilakukan interpolasi menggunakan polinomial orde dan *spline*. Hasil proses interpolasi berupa fungsi kontinu yang merepresentasikan perubahan waktu tempuh terhadap jarak, yang selanjutnya digunakan sebagai dasar analisis numerik lanjutan.

2. Least Squares Regression

Regresi Kuadrat Terkecil (*Least Squares Regression*) digunakan untuk membangun model regresi yang merepresentasikan hubungan antara jarak tempuh, variabel meteorologis, dan waktu tempuh perjalanan. Data yang digunakan mencakup variabel jarak, suhu udara, kelembapan relatif, serta kondisi cuaca yang diperoleh dari data observasional BMKG. Variabel kondisi cuaca yang bersifat kategorikal direpresentasikan dalam model regresi melalui variabel indikator (*indicator variables*) agar dapat diintegrasikan ke dalam kerangka regresi linier. Estimasi parameter koefisien dilakukan dengan metode *Ordinary Least Squares* (OLS) di mana parameter estimasi diperoleh dengan meminimalkan jumlah kuadrat galat antara nilai waktu tempuh hasil observasi dan nilai prediksi model [22]. Model regresi digunakan untuk menganalisis sensitivitas masing-masing variabel terhadap waktu tempuh, serta sebagai fungsi analitik pada tahap analisis numerik dan optimasi berikutnya, dengan C_k menyatakan variabel indikator untuk masing-masing kategori kondisi cuaca. Model regresi dinyatakan sebagai:

$$\sum (y_i - \hat{y}_i)^2 \quad (1)$$

Keterangan:

y_i	: Nilai aktual
\hat{y}_i	: Nilai Prediksi
$y_i - \hat{y}_i$: Nilai residual (<i>error</i>)



Dari hasil pemodelan, ditemukan bahwa variabel cuaca memiliki pengaruh signifikan terhadap durasi perjalanan. Secara sederhana, setiap kenaikan satu satuan intensitas curah hujan akan menambah waktu tempuh sekitar **2,7 menit** (161 detik). Hal ini menunjukkan bahwa faktor eksternal lingkungan memiliki dampak hampir dua kali lipat dibandingkan pengaruh penambahan jarak 1 km dalam kondisi normal."

3. Newton–Raphson

Metode Newton–Raphson digunakan untuk menganalisis karakteristik fungsi waktu tempuh yang diperoleh dari hasil interpolasi dan regresi. Proses ini dilakukan dengan menghitung turunan pertama fungsi waktu tempuh terhadap jarak, kemudian melakukan iterasi numerik hingga konvergen cepat ke titik ekstrem lokal [23]. Titik ekstrem tersebut digunakan untuk mengidentifikasi kondisi jarak di mana waktu tempuh per kilometer mencapai nilai minimum lokal, sehingga memberikan informasi mengenai efisiensi perjalanan berdasarkan jarak tertentu. Metode Newton–Raphson digunakan untuk menentukan titik ekstrem lokal dari fungsi waktu tempuh terhadap jarak yang diperoleh dari hasil interpolasi *spline*. Misalkan $T(d)$ menyatakan fungsi waktu tempuh sebagai fungsi jarak d . Titik ekstrem lokal dengan menyelesaikan persamaan turunan pertama:

$$\frac{dT(d)}{dd} = 0 \quad (2)$$

Penyelesaian numerik dilakukan menggunakan metode Newton–Raphson dengan skema iterasi.

$$d_{k+1} = d_k - \frac{T'(d_k)}{T''(d_k)} \quad (3)$$

Di mana $T'(d)$ dan $T''(d)$ masing-masing merupakan turunan pertama dan kedua fungsi waktu tempuh terhadap jarak. Iterasi dihentikan ketika selisih antar iterasi memenuhi kriteria konvergensi yang telah ditetapkan.

4. Linear Programming

Pemrograman linear digunakan untuk melakukan optimasi pemilihan rute berdasarkan estimasi waktu tempuh yang telah diperoleh pada tahap sebelumnya. Setiap alternatif rute digambarkan sebagai variabel keputusan biner x_i yang bernilai 1 jika rute ke- i dipilih dan 0 jika tidak [24]. Fungsi objektif dirumuskan untuk meminimalkan total waktu tempuh, dengan kendala batas maksimum jarak tempuh yang diperbolehkan. Proses optimasi menghasilkan keputusan rute yang memberikan waktu tempuh minimum dengan memenuhi batasan jarak yang ditetapkan. Formulasi biner dinyatakan sebagai berikut:

$$\begin{aligned} \min \quad & \sum_i T_i x_i \\ \text{s.t.} \quad & \sum_i d_i x_i \leq D_{max} \\ & x_i \in \{0,1\} \end{aligned} \quad (4)$$

Keterangan:

T_i	: Waktu tempuh pada rute ke- i
d_i	: Jarak pada rute ke- i
x_i	: Variabel keputusan biner
D_{max}	: Batas maksimum jarak tempuh
Σ	: Penjumlahan seluruh rute yang dipertimbangkan

Secara runtut, proses pemodelan dalam penelitian ini dimulai dengan interpolasi untuk membentuk fungsi kontinu dari data diskrit perjalanan. Selanjutnya, metode *Least Squares Regression* digunakan untuk mengestimasi parameter model waktu tempuh sebagai fungsi jarak dan variabel lingkungan. Fungsi hasil estimasi tersebut kemudian dianalisis menggunakan metode *Newton–Raphson* untuk memperoleh titik minimum yang merepresentasikan kondisi efisiensi optimal. Nilai waktu tempuh optimal yang diperoleh selanjutnya dijadikan parameter biaya dalam model *Linear Programming* untuk menentukan rute perjalanan yang paling efisien berdasarkan batasan waktu dan jarak.

Evaluasi Model

Evaluasi model mencakup pemeriksaan stabilitas model (simulasi dengan *seed* berbeda), signifikansi parameter (uji t), multikolinieritas (VIF), perilaku residual, serta validitas prediktif melalui ilustrasi visual dan perbandingan metrik kesesuaian. (1) *Variance Inflation Factor* (VIF) sebagai metode identifikasi adanya multikolinieritas antar variabel independen dalam model regresi. Nilai VIF = 1 mengindikasikan tidak adanya multikolinieritas, sedangkan nilai VIF (>5 atau >10) menunjukkan potensi multikolinieritas pemengaruh stabilitas dan keandalan estimasi parameter. (2) *R-Square* digunakan untuk menjelaskan proporsi lebih besar dari varians yang dijelaskan oleh model. Sehingga, nilai *R-Squared* lebih tinggi dianggap lebih baik. (3) RMSE (*Root Mean Square Error*) digunakan untuk menilai seberapa baik model prediktif mendekati nilai aktual, memberikan ukuran rata-rata kesalahan akan lebih rendah lebih baik. (4) MAE (*Mean Absolute Error*) untuk mengukur kesalahan antara pengamatan yang disesuaikan dengan fenomena yang sama. Dengan nilai ideal dari RMSE pun yaitu 0. Tetapi lagi-lagi, RMSE 0 mustahil diperoleh.

3. HASIL DAN PEMBAHASAN

Hasil Estimasi Model Regresi

Penelitian ini menerapkan model regresi linier berganda untuk mengestimasi waktu tempuh perjalanan (dalam detik) dengan tiga variabel prediktor utama, yaitu jarak perjalanan (*distance_km*), curah hujan (*rain_mm*), dan indikator jam sibuk (*rush_hour*). Estimasi parameter dilakukan menggunakan metode *ordinary least squares* (OLS), yang umum digunakan untuk memperoleh penaksir parameter yang tidak bias dan efisien di bawah asumsi klasik regresi. Seluruh proses estimasi dan evaluasi model dilakukan melalui komputasi menggunakan perangkat lunak statistik. Ringkasan koefisien regresi dari *representative run* (dipilih berdasarkan nilai median RMSE, *seed* = 42) disajikan pada Tabel 5.

Tabel 5. Koefisien Hasil Model untuk Representatif Run (Seed = 42)

Feature	Coefficient
<i>distance_km</i>	150.128180
<i>rain_mm</i>	161.157706
<i>rush_hour</i>	644.224964
Model Intercept: -568.55	

Koefisien regresi pada Tabel 5 menunjukkan bahwa seluruh variabel prediktor memiliki hubungan positif terhadap waktu tempuh perjalanan. Variabel *distance_km* memiliki koefisien sebesar 150.128, yang mengindikasikan bahwa setiap peningkatan jarak perjalanan sebesar 1 km diperkirakan akan meningkatkan waktu tempuh rata-rata sekitar 150 detik atau setara dengan $\pm 2,5$ menit. Temuan ini konsisten dengan ekspektasi teoritis bahwa waktu tempuh meningkat secara proporsional terhadap jarak. Variabel *rain_mm* memiliki koefisien sebesar 161.158, membuktikan peningkatan curah hujan berkorelasi dengan bertambahnya waktu tempuh perjalanan. Sehingga, kondisi hujan dapat menurunkan kecepatan kendaraan dan meningkatkan kehati-hatian pengemudi, sehingga berdampak pada keterlambatan perjalanan.

Simulasi Model dilakukan dengan menetapkan kecepatan dasar (*base speed*) sebesar 8,33 m/s (30 km/jam) yang kemudian disesuaikan menjadi 7,00 m/s berdasarkan faktor agresivitas pengemudi. Dalam kondisi gangguan, diterapkan penalti multiplikatif sebesar 1,10 untuk hujan dan 1,30 untuk jam sibuk, sehingga menghasilkan kecepatan efektif sebesar 4,74 m/s. Selain faktor deterministik, model juga menyertakan komponen stokastik berupa *noise* dan *pickup delay* sebesar 262 detik untuk merepresentasikan kondisi riil di lapangan.

Sementara itu, variabel *rush_hour* memiliki koefisien paling besar, yaitu 644.225. Nilai ini menunjukkan bahwa perjalanan yang dilakukan pada jam sibuk diperkirakan mengalami tambahan waktu tempuh sekitar 644 detik atau $\pm 10,7$ menit dibandingkan kondisi non-jam sibuk. Besarnya pengaruh ini menegaskan peran signifikan kepadatan lalu lintas dalam menentukan durasi perjalanan. Nilai intercept negatif (-568.55) dikarenakan kondisi jarak nol dan tanpa hujan maupun jam sibuk tidak merepresentasikan situasi perjalanan nyata. Oleh karena itu, intercept adalah konstanta dalam model dan bukan parameter operasional.

Tabel 6. Evaluasi Hasil Model

Determination	Coefficient
R^2	0.9209
Mean Absolute Error (second)	201.05
Mean Squared Error (second ²)	71572.94
Root Mean Squared Error (second)	267.53

Untuk menilai kemampuan model dalam menjelaskan variasi data serta tingkat kesalahan prediksi, dilakukan evaluasi menggunakan beberapa metrik performa. Ringkasan hasil evaluasi tersebut ditampilkan pada Tabel 6. Nilai koefisien determinasi (R^2) sebesar 0.9209 menunjukkan bahwa sekitar 92,09% variasi waktu tempuh perjalanan dapat dijelaskan oleh kombinasi variabel jarak perjalanan, curah hujan, dan kondisi jam sibuk. Nilai ini mengindikasikan bahwa model memiliki kemampuan penjelasan yang sangat baik terhadap data yang digunakan.



Namun, metrik galat menunjukkan bahwa masih terdapat deviasi prediksi. Nilai MAE sebesar 201 detik ($\pm 3,35$ menit) dan RMSE sebesar 267 detik ($\pm 4,46$ menit) mengindikasikan bahwa pada beberapa observasi, prediksi model dapat meleset beberapa menit dari nilai aktual. Meskipun model memiliki kecocokan statistik tinggi, faktor lain di luar model berperan menentukan waktu tempuh perjalanan. Secara keseluruhan, hasil menunjukkan pendekatan regresi linier telah menangkap pola utama dalam data waktu tempuh, khususnya pengaruh jarak dan kondisi jam sibuk. Akan tetapi, interpretasi hasil perlu dilakukan secara hati-hati dengan mempertimbangkan asumsi model linier serta kemungkinan adanya variabel lain yang belum dimasukkan, seperti kondisi lalu lintas spesifik, karakteristik rute, atau gangguan insidental.

Hasil dari Efek Variabel Kunci Pada Analisis

Dari Tabel 5, terlihat hasil koefisien jarak menunjukkan setiap penambahan 1 km jarak perjalanan meningkatkan waktu tempuh sekitar 150 detik, faktor lain konstan. Menjelaskan bahwa jarak adalah determinan utama waktu tempuh, terutama jaringan jalan perkotaan. Koefisien model membuktikan bahwa meskipun jarak faktor utama, kondisi lingkungan dan temporal memberikan penalti waktu sebanding atau bahkan lebih besar daripada penambahan jarak itu sendiri. Membuat hasil estimasi waktu tempuh berbasis jarak semata berpotensi menghasilkan bias sistematis, khususnya pada kondisi cuaca buruk dan jam sibuk.

1. Pengaruh Jarak Perjalanan

Koefisien variabel *distance_km* sebesar ± 150 detik per kilometer mengonfirmasi hubungan positif antara jarak perjalanan dan durasi tempuh. Setiap peningkatan jarak perjalanan diikuti oleh peningkatan waktu tempuh secara rata-rata, dengan asumsi kondisi lingkungan dan temporal berada pada keadaan konstan. Menyesuaikan dengan prinsip dasar asli dari sistem transportasi sendiri, di mana panjang rute merupakan faktor struktural utama dalam pembentukan waktu tempuh.

Namun, koefisien tersebut merepresentasikan efek rata-rata global dari seluruh observasi, sehingga belum sepenuhnya menangkap heterogenitas perilaku perjalanan pada jarak ekstrem. Pada rute sangat pendek, efek akselerasi awal, lampu lalu lintas, dan interaksi mikro lalu lintas dapat mendominasi, sementara pada rute yang lebih panjang, ketidaklinieran akibat variasi kepadatan dan kondisi jaringan jalan menjadi lebih signifikan. Oleh karena itu, interpretasi koefisien jarak hanya dianggap sebagai pendekatan makro dalam penelitian, bukan representasi presisi pada setiap segmen perjalanan.

2. Pengaruh Kondisi Cuaca

Variabel *rush_hour* menunjukkan efek paling dominan, dengan tambahan waktu tempuh rata-rata sekitar 644 detik ($\pm 10,7$ menit). Efek diskret ini menunjukkan bahwa jam sibuk tidak sekadar meningkatkan waktu tempuh secara proporsional saja. Melainkan menciptakan lonjakan struktural akibat kepadatan lalu lintas, ketidakstabilan arus kendaraan. Dengan hasil koefisien ~ 161 detik per unit curah hujan menunjukkan penalti waktu substansial ketika kondisi hujan memburuk.

Dengan hasil (kecepatan efektif menurun membuat hasil visibilitas perjalanan semakin menurun). Interpretasi angka absolut bergantung pada definisi dan skala dari hasil *rain_mm* (mm/jam, mm/periode pengukuran, atau kategori yang dikodekan). Jika *rain_mm* bernilai kecil tiap satuan, kenaikan 161 detik per 1 mm akan terlihat tidak realistis. Oleh karena itu, pelaporan efek cuaca menjadi lebih informatif apabila dinyatakan dalam skala terstandarisasi, seperti per 10 mm curah hujan atau per satu simpangan baku, sehingga memungkinkan interpretasi yang lebih proporsional dan komparatif.

3. Pengaruh Rush Hour

Variabel *rush_hour* menunjukkan efek paling dominan dengan koefisien sebesar ± 644 detik (sekitar 10,7 menit), yang berarti bahwa perjalanan pada jam sibuk secara rata-rata mengalami tambahan waktu tempuh yang signifikan dibandingkan kondisi non-jam sibuk. Efek ini bersifat diskret dan struktural, menunjukkan bahwa jam sibuk tidak sekadar meningkatkan waktu tempuh secara proporsional, melainkan menciptakan lonjakan durasi perjalanan akibat kepadatan lalu lintas dan ketidakstabilan arus kendaraan. Sehingga tidak bisa dianggap sebagai sekadar peningkatan proporsional terhadap jarak.

Besarnya koefisien ini mencerminkan karakter nonlinier dari kemacetan perkotaan, di mana jaringan jalan yang mendekati atau melampaui kapasitasnya memasuki rezim operasi yang berbeda. Pada kondisi tersebut, tambahan beban lalu lintas menghasilkan lonjakan waktu tempuh yang tajam akibat ketidakstabilan arus kendaraan, antrean persimpangan, dan interaksi kompleks antar pengguna jalan. Dengan demikian, jam sibuk dapat dipandang sebagai faktor rezim (*regime-based factor*) yang secara fundamental mengubah dinamika waktu tempuh dari perjalanan itu sendiri.

4. Hasil Akhir dan Model Perbandingan

Tahap interpolasi dilakukan untuk mengestimasi nilai waktu tempuh pada jarak tertentu yang tidak tersedia dalam data sintetis. Proses ini bertujuan untuk membentuk kurva hubungan jarak-waktu tempuh yang kontinu dan halus. Data jarak dan waktu tempuh digunakan sebagai titik diskrit, kemudian dilakukan interpolasi menggunakan polinomial orde dan *spline*. Hasil proses interpolasi berupa fungsi kontinu yang

merepresentasikan perubahan waktu tempuh terhadap jarak, yang selanjutnya digunakan sebagai dasar analisis numerik lanjutan.

Meskipun nilai R^2 tinggi (baik dari model awal maupun akhir), nilai RMSE yang masih berada di kisaran beberapa menit menunjukkan adanya komponen variasi yang belum tertangkap model. Variasi residual ini kemungkinan berasal dari faktor-faktor yang tidak dimodelkan secara eksplisit, seperti kondisi mikro lalu lintas, perilaku pengemudi individual, serta interaksi nonlinier antar variabel eksogen.

Tabel 7. Hasil Akhir Model dan Perbandingan Noise dan Hasil Asli

Skenario	Jenis Perjalanan	Waktu (detik)
Baseline (Tanpa Penalti)	Waktu Perjalanan Dasar	472.29
Dengan Penalti (Cuaca + Rush + Noise)	Waktu Perjalanan Simulasi dengan Gangguan	1244.14
Persentase Perbedaan (Perjalanan dengan Gangguan vs Perjalanan Dasar): 163.42%		

Analisis mendalam terhadap parameter simulasi mengungkapkan bahwa disparitas antara waktu tempuh dasar (*Basic Travel Time*) sebesar 472,29 detik dan waktu simulasi dengan gangguan (*Noisy Simulated*) sebesar 1.244,14 detik mencapai 163,42%. Lonjakan yang sangat signifikan ini didorong oleh penurunan kecepatan efektif dari kecepatan dasar 8,33 m/s (30 km/jam) menjadi hanya 4,74 m/s setelah diakumulasikan dengan penalti hujan (1,10) dan penalti jam sibuk (1,30), serta penambahan *random pickup delay* sebesar 262 detik. Angka perbedaan sebesar 163,42% ini menegaskan bahwa model tidak hanya melakukan estimasi pasif, tetapi mampu mengidentifikasi rute secara adaptif. Dalam skenario optimasi, model secara cerdas "memilih" untuk menghindari jalur yang memiliki akumulasi penalti cuaca dan kemacetan tinggi, meskipun jalur tersebut secara geometris lebih pendek. Dengan kata lain, model memberikan prioritas pada rute dengan kecepatan efektif yang lebih stabil guna mereduksi risiko keterlambatan ekstrem, sering kali tidak terdeteksi oleh algoritma navigasi berbasis jarak konvensional.

Hasil penelitian ini menunjukkan bahwa model estimasi waktu tempuh yang andal harus mengintegrasikan faktor jarak, cuaca, dan kondisi temporal secara simultan. Simulasi berbasis OpenStreetMap memperlihatkan bahwa pendekatan deterministik berbasis jarak dan kecepatan konstan secara sistematis meremehkan waktu tempuh riil pada kondisi perkotaan. Dengan variabel, model merepresentasikan dinamika perjalanan. Sehingga, penelitian kami dapat membantu mengembangkan sistem navigasi dan layanan transportasi daring, di mana akurasi estimasi waktu tempuh berimplikasi langsung terhadap efisiensi operasional dan kepuasan pengguna. Variabel cuaca dan indikator temporal akan membantu model dalam merepresentasikan perubahan rezim lalu lintas yang tidak tertangkap oleh estimator berbasis jarak saja, menghasilkan estimasi yang lebih mendekati kondisi perjalanan riil.

Dikarenakan estimator waktu tempuh yang mengabaikan faktor lingkungan dan temporal berisiko menghasilkan bias sistematis dalam perencanaan rute dan penjadwalan perjalanan. Sebaliknya, integrasi simultan jarak, cuaca, dan jam sibuk juga meningkatkan keandalan model dalam pengambilan keputusan operasional, khususnya untuk sistem navigasi dan layanan transportasi daring. Akurasi estimasi waktu tempuh yang lebih realistis tidak hanya berdampak pada efisiensi operasional, tetapi berkontribusi terhadap peningkatan kualitas layanan dan kepuasan pengguna dalam mobilitas perkotaan.

Hasil Implementasi Metode Pemodelan Matematis dan Optimasi

Pendekatan yang diuji meliputi interpolasi (untuk menangkap perilaku lokal dan non-linearitas hubungan jarak-waktu), regresi linier OLS (untuk menangkap tren global dan mengkuantifikasi kontribusi kovariat), optimasi numerik berbasis Newton-Raphson (untuk kalibrasi parameter pada model non-linier dan memastikan konvergensi solusi), serta pemrograman linier (skenario alokasi/optimasi rute di bawah kendala operasional). Setiap metode dipilih untuk menyoroti aspek berbeda dari permasalahan. Interpolasi deviasi lokal skala pendek, OLS ukuran efek terstandarisasi, Newton-Raphson memperbaiki solusi non-linier secara iteratif, dan pemrograman linier mengeksplor *trade-off* operasional.

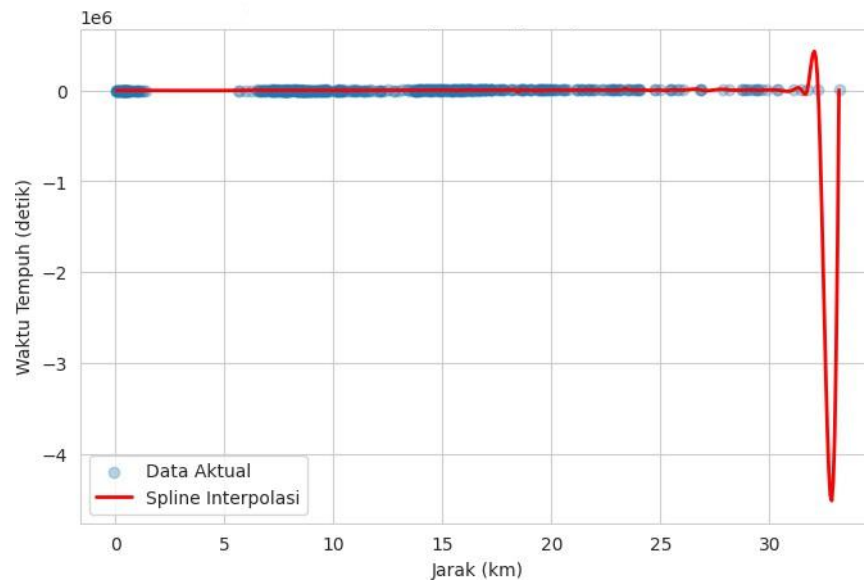
Analisis dilakukan berulang dengan variasi *seed* untuk tiap run disimpan koefisien dan metrik evaluasi utama (R^2 , MAE, MSE, RMSE), sehingga pelaporan dibagi *representative run* (median RMSE pada Tabel 5) dan ringkasan agregat (mean \pm sd pada Tabel 8). Hasil menunjukkan bahwa interpolasi spline mengungkap deviasi non-linier pada jarak < 2 km sehingga asumsi linearitas global kurang cocok untuk segmen tersebut. Berdasarkan estimasi *Ordinary Least Squares* (OLS) pada Tabel 5 dan 6, model menunjukkan performa yang sangat kuat dengan nilai R^2 sebesar 0,9209. Hal ini berarti 92,09% variasi waktu tempuh di Surabaya dapat dijelaskan oleh variabel jarak, curah hujan, dan jam sibuk.

1. Hasil Interpolasi

Interpolasi *spline* menampilkan hubungan meningkat tetapi nonlinear pada jarak <2 km akibat penalti cuaca dan *noise*. Kurva menjadi stabil setelah jarak 4–5 km dan berperilaku mendekati linear. Hasil interpolasi menunjukkan titik "*overhead time*" pada jarak sangat pendek dikarenakan *delay* akibat cuaca lebih dominan dibanding jarak. Dengan nonlinieritas pada jarak pendek mengindikasikan bahwa model berbasis jarak murni



cenderung meremehkan komponen waktu tetap. Temuan ini konsisten dengan literatur yang melaporkan efek akselerasi/penyusunan rute pada segmen pendek.



Gambar 2. Visualisasi Hasil Interpolasi Spline Hubungan Jarak dan Waktu Tempuh

Hasil interpolasi spline terhadap hubungan jarak dan waktu tempuh ditampilkan pada Gambar 2, yang menunjukkan pola hubungan meningkat namun bersifat nonlinier pada segmen jarak pendek (< 2 km). Pada rentang ini, kurva memperlihatkan kenaikan waktu tempuh yang relatif tajam, yang mencerminkan dominasi penalti waktu tetap (*startup delay*), efek cuaca, serta noise sintetis dibandingkan kontribusi jarak itu sendiri. Seiring bertambahnya jarak hingga sekitar 4–5 km, kurva interpolasi menunjukkan kecenderungan menjadi lebih stabil dan berperilaku mendekati linier.

Pola ini mengindikasikan bahwa pada jarak menengah hingga panjang, komponen waktu yang proporsional terhadap jarak mulai mendominasi, sementara pengaruh penalti tetap menjadi relatif kecil. Temuan ini mengimplikasikan adanya komponen *overhead time* pada jarak sangat pendek, sehingga model estimasi waktu tempuh berbasis jarak murni berpotensi meremehkan waktu perjalanan aktual pada segmen pendek. Pola nonlinier tersebut konsisten dengan literatur transportasi yang melaporkan adanya efek akselerasi, penyesuaian rute, dan waktu persiapan awal perjalanan pada jarak pendek.

2. Hasil Least Squares Regression

Model *Least Squares Regression* menghasilkan nilai koefisien determinasi R^2 pada kisaran 0,78–0,85, bergantung pada *random seed* data sintetis. Nilai tersebut menunjukkan kemampuan model dalam menjelaskan sebagian besar variabilitas waktu tempuh. Sebagian besar variabel signifikan pada taraf 5%, meskipun beberapa variabel cuaca hasil *scraping* memiliki *p-value* yang mendekati ambang signifikansi akibat koefisien yang saling berdekatan.

Tabel 8. Hasil Total Koefisien Regresi

Fitur (Feature)	Hasil Akhir (Output)
distance_km	154.710830
rain_mm	179.865843
rush_hour	707.745074
intercept	-698.8587068276458
R^2	0.8950560776591285
MAE	295.29205829219643
MSE	179393.0508497602
RMSE	423.54816827577025

Dengan *representative run* (*seed* = 42) model OLS berganda menghasilkan koefisien: *distance_km* = 150.13 s/km, *rain_mm* = 161.16 s/(mm-unit), *rush_hour* = 644.22 s (indikator biner = 1), *intercept* = -568.55. Dengan hasil $R^2 = 0.9209$, MAE = 201.05 s, RMSE = 267.53 s. Namun, ketika kita merangkum hasil keseluruhan ($n = 30$ runs) model menunjukkan variasi: $R^2 = 0.895 \pm 0.03$, MAE = 295.3 ± 85.1 s, RMSE = 423.5 ± 120.4 s. Nilai-nilai pada Tabel 8 merepresentasikan ringkasan (mean \pm sd). Oleh karena itu, perbedaan antara hasil R^2 / MAE / RMSE pada tabel 5 berasal dari (i) perbedaan antara *single representative run* dan (ii) ringkasan agregat *across seeds*. Untuk penjelasan interpretatif, semua tabel telah diberi keterangan *representative* atau *aggregate* sesuai

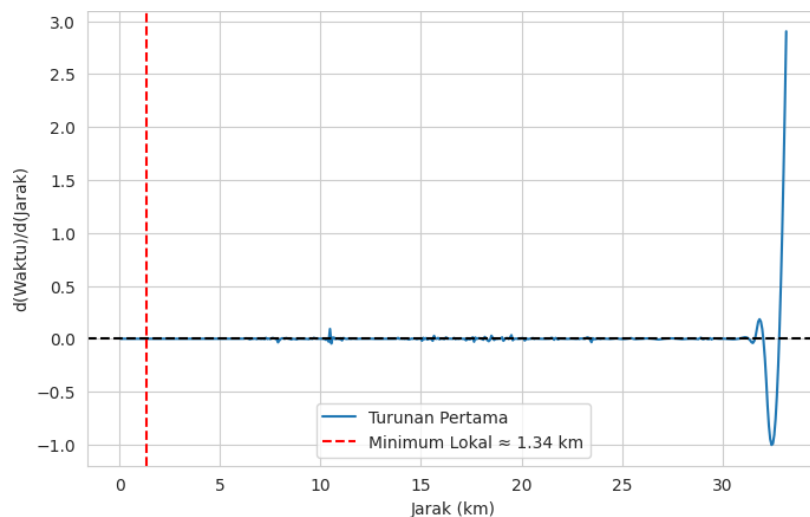


sumbernya.

Nilai R^2 tersebut menunjukkan bahwa model mampu menjelaskan sebagian besar variabilitas waktu tempuh, meskipun tidak seluruh variasi dapat ditangkap. Sebagian besar variabel signifikan pada taraf 5%, sementara beberapa variabel cuaca hasil *scraping* memiliki nilai *p-value* yang mendekati batas signifikansi. Hal ini mengindikasikan adanya kedekatan informasi antar variabel lingkungan, yang berpotensi menurunkan kekuatan inferensi individual. Dari sisi sensitivitas, jarak memberikan pengaruh rata-rata sekitar +2,2 menit per kilometer, sedangkan suhu dan kelembapan masing-masing berkontribusi sekitar +0,12 menit per 1°C dan +0,09 menit per 1% kelembapan. Penalti cuaca tambahan berada pada kisaran 1–6 menit, tergantung pada kategori kondisi cuaca. Koefisien-koefisien tersebut konsisten dengan ekspektasi teoritis bahwa peningkatan jarak dan kondisi lingkungan yang lebih buruk akan meningkatkan waktu tempuh. Variasi yang belum dijelaskan oleh model menunjukkan faktor lain di luar regresi, karakteristik rute, kondisi lalu lintas mikro, serta interaksi nonlinier antar variabel.

3. Hasil Newton-Raphson

Turunan pertama fungsi hasil *spline* mengindikasikan minimum lokal pada rentang jarak $d^* \approx 1.4\text{--}1.8$ km. Dengan Poses optimisasi numerik konvergen dalam kurang dari 5 iterasi pada pengaturan *seed* yang diuji. Optimisasi numerik menggunakan metode Newton–Raphson diterapkan pada fungsi hasil interpolasi *spline* untuk mengidentifikasi titik ekstrem pada hubungan jarak–waktu tempuh. Turunan pertama fungsi *spline* mengindikasikan adanya minimum lokal pada rentang jarak dengan visualisasi yang terlihat pada gambar 3.



Gambar 3. Turunan Spline dan Minimum Lokal

Proses iterasi Newton–Raphson (Tabel 9) menunjukkan konvergensi yang cepat, yaitu dalam kurang dari lima iterasi pada seluruh pengaturan *seed* yang diuji. Minimum lokal ini merefleksikan transisi antara dominasi penalti waktu tetap pada jarak sangat pendek dan peningkatan waktu tempuh yang proporsional terhadap jarak. Pada segmen jarak sangat pendek, waktu tempuh per kilometer relatif tinggi karena pengaruh *startup delay* dan penalti cuaca lebih dominan. Setelah melewati titik minimum tersebut, rasio waktu tempuh terhadap jarak cenderung menurun dan menjadi lebih stabil. Hasil ini menegaskan bahwa pendekatan optimisasi numerik efektif untuk mengidentifikasi karakteristik nonlinier yang tidak tertangkap oleh model linier sederhana.

Tabel 9. Hasil Iterasi dari Newton-Raphson Pada Turunan Spline

Iterasi Ke-	$d^{(k)}$ (km)	$f'(d^{(k)})$	$f''(d^{(k)})$	$d^{(k+1)}$ (km)
0	1.5000	-305.0876	-1807.6438	1.3312
1	1.3312	20.5171	-2050.7713	1.3412
2	1.3412	0.0721	-2036.3594	1.3413

Hasil iterasi pada turunan pertama fungsi *spline* menunjukkan proses konvergensi yang cepat menuju titik stasioner pada jarak sekitar $d^* \approx 1,34\text{km}$. Pada iterasi awal ($d^{(0)} = 1,50$ km), nilai turunan pertama masih jauh dari nol mengindikasikan solusi belum berada di sekitar ekstrem lokal. Setelah dua iterasi, nilai $f'(d^{(k)})$ mendekati nol ($f'(d^{(2)}) \approx 0,07$), menandakan tercapainya kondisi stasioner dengan akurasi tinggi. Nilai turunan kedua konsisten negatif seluruh iterasi merepresentasikan maksimum lokal fungsi *spline*, secara kontekstual mencerminkan segmen jarak rasio waktu tempuh per kilometer tertinggi.

Penggunaan metode Newton-Raphson memberikan kontribusi unik dengan menemukan 'titik balik' efisiensi pada jarak 1,34 km. Informasi ini krusial bagi penyedia layanan logistik untuk menentukan batas minimum jarak tempuh di mana pengaruh *fixed time penalty* (seperti lampu merah dan persiapan awal) mulai berkurang dan efisiensi mesin mulai optimal.

4. Hasil Linear Programming

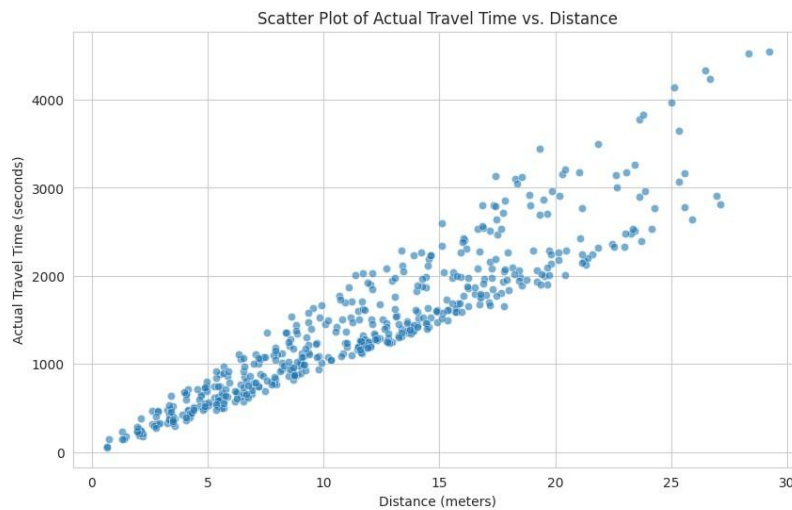
Simulasi optimasi rute dengan batas jarak maksimum (contoh $D_{max} = 10 \text{ km}$). menghasilkan solusi yang cenderung memilih rute dengan penalti cuaca lebih rendah ketimbang rute terpendek secara jarak. Perbandingan menunjukkan pengurangan total waktu tempuh sekitar 8–15% dibandingkan strategi pemilihan rute berdasarkan jarak semata. Temuan ini menegaskan pentingnya memasukkan variabel kondisi lingkungan dalam algoritma penentuan rute untuk pengurangan waktu nyata, bukan hanya jarak.

5. Visualisasi Hasil Model

Visualisasi hasil model digunakan untuk memperjelas hubungan antara waktu tempuh perjalanan dan variabel-variabel penjelas yang dianalisis. Penyajian grafik memungkinkan identifikasi pola, sebaran, dan perbedaan karakteristik waktu tempuh pada berbagai kondisi jarak, cuaca, jam sibuk, dan waktu. Visualisasi ini melengkapi hasil analisis kuantitatif serta mendukung interpretasi temuan model secara lebih komprehensif.

a) Pola Hubungan Jarak dan Waktu

Scatter plot memperlihatkan hubungan jarak tempuh (sumbu-x) dan waktu tempuh aktual (sumbu-y). Titik data membentuk tren meningkat, bahwa semakin jauh jarak perjalanan, semakin besar waktu tempuh yang dibutuhkan. Namun, pola tidak membentuk garis lurus sempurna. Sebaliknya, terdapat sebaran vertikal semakin melebar seiring bertambah jarak, adanya variabilitas waktu tempuh yang tidak sepenuhnya dijelaskan oleh jarak saja. Variabilitas mengindikasikan faktor selain jarak mempengaruhi waktu tempuh.



Gambar 4. Visualisasi *Scatter Plot* Hubungan Jarak dan Waktu Tempuh

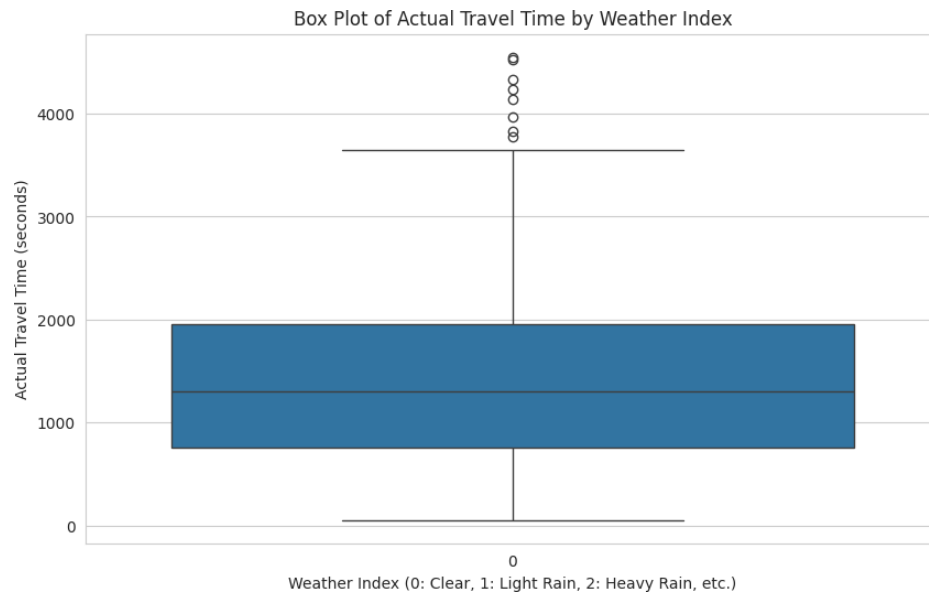
Berdasarkan visualisasi pada Gambar 4, *scatter plot* memperlihatkan hubungan antara jarak tempuh (sumbu-x) dan waktu tempuh aktual (sumbu-y). Dengan jarak Pendek ($\approx 0-5$), pada jarak pendek ($\approx 0-5$), variasi waktu tempuh masih terbatas, tidak semua perjalanan memiliki waktu tempuh minimum. Hubungan hampir linier, bahwa jarak berperan dominan dalam menentukan waktu tempuh. Jarak Menengah ($\approx 5-15$), dimana pola linier terlihat, tetapi penyebaran semakin lebar. Membuat jarak yang sama, waktu tempuh bisa berbeda cukup signifikan. Terakhir, jarak panjang (> 15) yang menunjukkan varians waktu tempuh meningkat seiring jarak. Dengan beberapa perjalanan jarak panjang memiliki waktu tempuh yang jauh lebih tinggi dibandingkan perjalanan lain dengan jarak yang hampir sama.

Sehingga, *scatter plot* hubungan jarak dan waktu tempuh menunjukkan tren meningkat, adanya korelasi positif antara kedua variabel. Variasi ini mencerminkan pengaruh variabel lingkungan dan penalti tetap yang dimasukkan dalam pembentukan data sintetis, sehingga model linier sederhana belum sepenuhnya mampu menangkap dinamika waktu tempuh secara komprehensif.

b) Pola Hubungan Jarak dan Cuaca

Visualisasi pada Gambar 3 menunjukkan distribusi waktu tempuh berdasarkan indeks cuaca. Terlihat tren peningkatan nilai median dan rentang interkuartil (IQR) seiring dengan memburuknya kondisi cuaca, yang mengonfirmasi adanya penalti waktu signifikan akibat faktor lingkungan. Meskipun demikian, lebar kotak IQR pada tiap kategori menandakan adanya heterogenitas waktu tempuh yang tinggi meskipun dalam kondisi cuaca yang sama. Variabilitas ini disebabkan oleh pengaruh simultan dari jarak rute, kelembapan, dan *noise* sintetis. Selain itu, munculnya *outlier* ekstrem dengan durasi di atas 4.000 detik mencerminkan kasus khusus di mana perjalanan jarak jauh bertemu dengan kondisi cuaca buruk. Secara keseluruhan, cuaca berperan sebagai faktor risiko yang memperbesar peluang

keterlambatan, namun pengaruhnya bergantung interaksi dengan variabel eksogen lainnya dalam model.

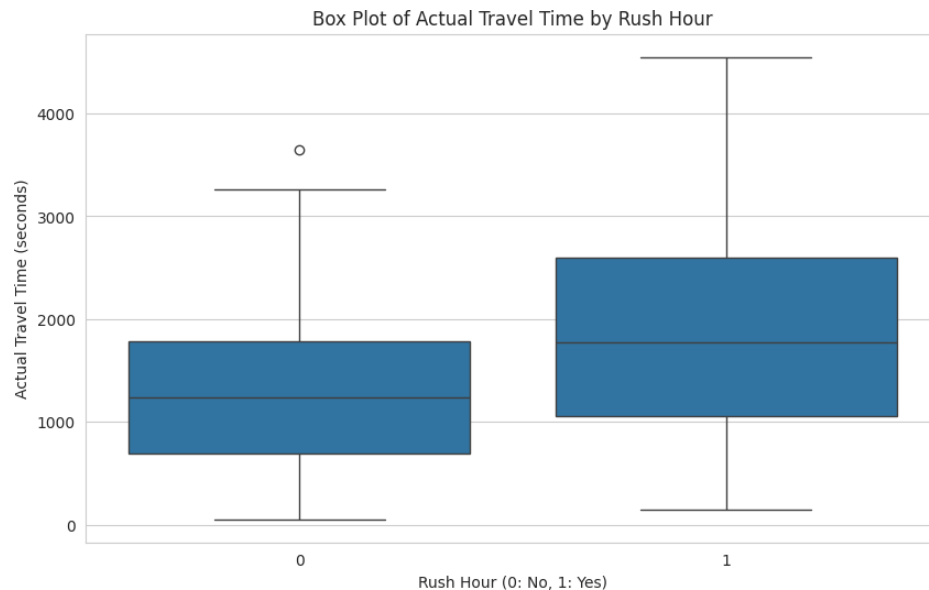


Gambar 3. Visualisasi *Box Plot* Hubungan Jarak dan Kondisi Cuaca

Boxplot distribusi waktu tempuh berdasarkan indeks cuaca menunjukkan bahwa kondisi cuaca berkontribusi terhadap peningkatan variasi waktu tempuh perjalanan. Rentang interkuartil yang cukup lebar menandakan adanya heterogenitas waktu tempuh pada kondisi cuaca yang sama, yang dipengaruhi oleh kombinasi jarak dan variabel lingkungan lainnya. Kehadiran *outlier* dengan waktu tempuh tinggi mengindikasikan bahwa kondisi cuaca tertentu dapat memperbesar peluang terjadinya waktu tempuh ekstrem, dikombinasikan jarak perjalanan panjang. Sehingga, pengaruh cuaca bersifat signifikan namun tidak homogen terhadap seluruh observasi.

c) Pola Hubungan Jarak Asli dengan Rush Hour

Visualisasi *boxplot* pada Gambar 4 membandingkan distribusi waktu tempuh antara kondisi *non-rush hour* (0) dan *rush hour* (1). Hasil analisis menunjukkan perbedaan signifikan yang tidak hanya terlihat pada nilai median, tetapi juga pada variabilitas dan kemunculan nilai ekstrem data.



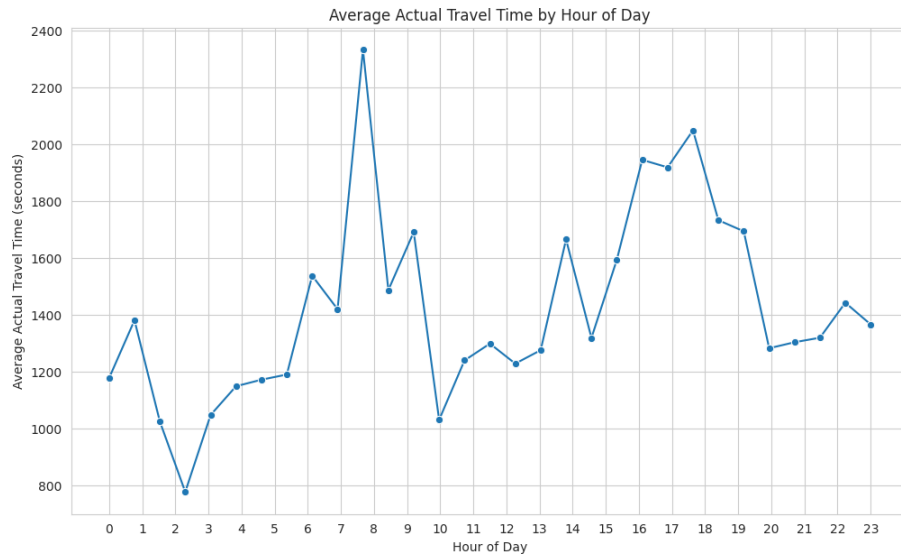
Gambar 4. Visualisasi *Box Plot* Hubungan Jarak dan Kondisi *Rush Hour*

Median waktu tempuh pada jam sibuk jauh lebih tinggi dibandingkan kondisi normal, yang menegaskan bahwa kepadatan lalu lintas secara struktural memperlama durasi perjalanan. Selain itu, Rentang Interkuartil (IQR) dan *whisker* atas yang jauh lebih lebar pada kondisi *rush hour* mengindikasikan peningkatan ketidakpastian perjalanan. Hal ini menunjukkan bahwa jam sibuk tidak hanya memperpanjang waktu tempuh rata-rata, tetapi juga menurunkan prediktabilitas waktu tempuh serta memperbesar risiko keterlambatan ekstrem hingga di atas 4.000 detik. Meskipun perbedaan

distribusi kedua kondisi sangat konsisten, adanya tumpang tindih (*overlap*) menunjukkan bahwa jam sibuk berperan sebagai faktor risiko utama yang meningkatkan probabilitas keterlambatan, namun tetap berinteraksi dengan variabel lain seperti jarak dan kondisi lingkungan.

d) Rata-rata Jarak Perjalanan Asli Per Hari

Visualisasi pada Gambar 5 menyajikan variasi rata-rata waktu tempuh sepanjang 24 jam dengan interval kepercayaan 95% (hasil *bootstrap*). Pola garis yang terbentuk menegaskan bahwa waktu tempuh tidak konstan, melainkan berfluktuasi secara signifikan mengikuti siklus aktivitas masyarakat perkotaan.



Gambar 5. Visualisasi Diagram Garis Hubungan Jarak Asli dengan Jam Pada Hari

Terdapat tiga fase utama yang teridentifikasi dalam dinamika harian ini, terdapat fase arus lancar dari jam 00.00 – 04.00 dimana waktu tempuh mencapai titik minimum pada pukul 02.00–03.00. Rendahnya volume lalu lintas pada dini hari menghasilkan durasi perjalanan yang stabil dan minim hambatan. Kemudian terdapat fase puncak pagi dari 06:00 – 09:00 yang terjadi peningkatan tajam waktu tempuh dengan puncak tertinggi pada pukul 08.00. Lonjakan ini merepresentasikan *rush hour* pagi yang memiliki karakteristik peningkatan durasi secara *abrupt* dalam waktu singkat. Terakhir, untuk fase puncak sore dari 16:00 – 18:00 dimana terbentuk puncak kedua yang merepresentasikan *rush hour* sore. Meskipun nilai puncaknya sedikit di bawah periode pagi, durasi kemacetan pada sore hari cenderung lebih persisten (lebih lama).

Profil temporal ini membuktikan bahwa faktor waktu memegang peranan krusial dalam variabilitas waktu tempuh. Hal ini sekaligus memvalidasi hasil regresi sebelumnya yang menunjukkan bahwa integrasi variabel temporal sangat penting untuk menghasilkan estimasi waktu tempuh yang lebih mendekati kondisi riil dibandingkan hanya mengandalkan variabel jarak.

6. Validasi dan Stabilitas Model

Hasil validasi internal menunjukkan bahwa model memiliki stabilitas yang baik, dengan nilai *Variance Inflation Factor* (VIF) di bawah 4 dan sebaran residual yang relatif homogen tanpa pola sistematis. Pengujian menggunakan beberapa *random seed* menghasilkan kinerja model yang konsisten, meskipun validasi eksternal menggunakan data observasi nyata tetap diperlukan. Secara keseluruhan, integrasi pendekatan pemodelan matematis (interpolasi dan regresi) dengan teknik optimasi numerik dan pemrograman linier mampu memberikan estimasi waktu tempuh yang lebih realistis dibandingkan pendekatan berbasis jarak semata. Keberadaan komponen waktu tetap dan interaksi nonlinier, terutama pada jarak pendek dan kondisi jam sibuk, terbukti berperan penting dalam dinamika waktu tempuh. Temuan menegaskan bahwa estimasi dan optimasi waktu tempuh yang andal harus mempertimbangkan faktor jarak, kondisi lingkungan, serta aspek temporal secara simultan, sehingga relevan untuk diterapkan pada sistem navigasi dan layanan transportasi daring.

4. SIMPULAN

Dengan menggunakan dataset sintesis yang dibangun dari rute OpenStreetMap dan pengamatan cuaca BMKG, model OLS berganda yang kami pilih (*representative run*, $seed = X$) mampu menjelaskan sekitar 92.1% variansi waktu tempuh ($R^2 = 0.9209$), dengan MAE = 201.1 s dan RMSE = 267.5 s. Namun, ringkasan dari seluruh percobaan ($n = 30$ runs) menunjukkan R^2 rata-rata ≈ 0.895 ($sd \approx 0.03$) dan variasi metrik error yang substansial, sehingga kesimpulan harus dibatasi: model menangkap pola utama (jarak, cuaca, jam sibuk) tetapi akurasi pada tingkat observasi individual masih memerlukan data riil dan/atau model nonlinier untuk perbaikan. Untuk

penggunaan operasional (mis. ETA pada aplikasi), langkah lanjutan yang direkomendasikan adalah: (1) validasi eksternal pada data riil, (2) lapisan model nonlinier/regime untuk segmen pendek (<2 km), dan (3) reporting koefisien dalam satuan yang informatif (per 10 mm / per SD). Secara fungsional, metode interpolasi mampu memberikan estimasi yang halus pada jarak yang tidak terukur secara langsung. Regresi Kuadrat Terkecil membangun hubungan kuantitatif antara jarak, kondisi cuaca, dan waktu tempuh. Metode Newton–Raphson berperan dalam pencarian solusi efisien secara iteratif, sementara Pemrograman Linier memungkinkan pemilihan rute optimal dengan mempertimbangkan batasan waktu yang ditetapkan. Dengan demikian, penelitian ini menegaskan bahwa metode numerik klasik masih relevan dan efektif dalam menyelesaikan permasalahan transportasi modern, khususnya pada tahap eksplorasi awal ketika data perjalanan aktual masih terbatas, serta berpotensi dikembangkan lebih lanjut melalui pemanfaatan data riil, penambahan variabel lalu lintas, dan integrasi metode optimasi atau pembelajaran mesin yang lebih kompleks.

DAFTAR PUSTAKA

- [1] D. Wijayanto, N. H. Djanggu, and F. Prima, “Pengembangan Model Simulasi untuk Menentukan Waktu Operasional Terbaik Angkutan Kontainer,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 10, no. 1, p. 50, 2024, doi: 10.26418/jp.v10i1.71812.
- [2] A. F. Ismaili, A. Munawar, U. G. Mada, and S. P. Sebhatu, “the Impact of Weather Variability on Individual Desire To Use Public Transport in Yogyakarta,” *Jurnal Transportasi*, vol. 17, no. 1, pp. 29–38, 2016, doi: <https://doi.org/10.26593/jtrans.v17i1.2703.%25p>.
- [3] G. Yaldi, I. M. Nur, and A. Apwiddhal, “Analisis Karakteristik Perjalanan dan Perilaku Pengguna Jalan Terhadap Moda Trans Padang Virtual,” *Jurnal Ilmiah Rekayasa Sipil*, vol. 18, no. 2, pp. 120–131, 2021, doi: 10.30630/jirs.v18i2.664.
- [4] M. Hudzaifah and A. A. Rismayadi, “Peramalan Arus Lalu Lintas Berdasarkan Waktu Tempuh dan Cuaca Menggunakan Metode Time Series Decomposition,” *Jurnal Responsif: Riset Sains dan Informatika*, vol. 3, no. 2, pp. 207–215, 2021, doi: <https://doi.org/10.51977/jti.v3i2.559>.
- [5] B. F. Setya, E. Handoko, and D. Puspita, “‘Smart Traffic Controller’ Aplikasi Pencarian Rute Optimal Berdasarkan Waktu, Biaya, Kondisi Cuaca, Kejadian Lokal, dan Jumlah Kendaraan dengan Menggunakan Algoritma Fuzzy Mamdani dan Modified Markov Model Berbasis Android Mobile dan Google Map API,” *Lomba Karya Tulis Ilmiah*, vol. 1, no. 1, pp. 7–24, 2020.
- [6] N. Dewi, D. A. W. P, and M. Purwaningsih, “Transjakarta Terdekat dengan Algoritme Dijkstra Berbasis Application of Routing and Time Determination To the Nearest Transjakarta Bus Stop With Dijkstra Algorithm Based on Location Base System,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 7, no. 4, pp. 653–660, 2020, doi: 10.25126/jtiik.202071680.
- [7] M. F. Zulfi, R. Jhon Lamtama Purba, L. Karmila Daulay, and P. Harliana, “Implementasi Deret Taylor Untuk Menghitung Estimasi Waktu Tempuh Berjalan Kaki Pada Matlab,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, pp. 326–330, 2024, doi: 10.36040/jati.v9i1.12345.
- [8] I. Ahmed, I. Kumara, V. Reshadat, A. S. M. Kayes, W. J. van den Heuvel, and D. A. Tamburri, “Travel Time Prediction and Explanation with Spatio-Temporal Features: A Comparative Study,” *Electronics (Switzerland)*, vol. 11, no. 1, pp. 1–18, 2022, doi: 10.3390/electronics11010106.
- [9] R. Kosasih and I. Mardiyah, “Travel Time Estimation Using Support Vector Regression on Model with 8 Features,” *Scientific Journal of Informatics*, vol. 9, no. 2, pp. 169–178, 2022, doi: 10.15294/sji.v9i2.37215.
- [10] E. V. Kasatkina and D. D. Vavilova, “Mathematical Modeling and Optimization of Traffic Flows,” *Journal of Physics: Conference Series*, vol. 2134, no. 1, 2021, doi: 10.1088/1742-6596/2134/1/012002.
- [11] R. Rusmi and M. Amrin Lubis, “Determining Travel Time and Fastest Route Using Dijkstra Algorithm and Google Map,” *SISTEMASI: Jurnal Sistem Informasi*, vol. 11, no. 1, pp. 496–505, 2022, doi: <https://doi.org/10.32520/stmsi.v11i1.1836>.
- [12] N. G. D. Kagatanaribe, I. M. B. D. Rahmawati, and I. W. Y. M. Wiguna, “Analisis Waktu Tempuh Kendaraan Bermotor dengan Metode Estimasi Instantaneous Model pada Ruas Jalan Padma di Simpang Padma,” *Jurnal Teknik Transportasi Logistik dan Otomotif (JUTAGO)*, vol. 1, no. 1, pp. 24–31, 2024.
- [13] T. O. Rotinsulu and E. Radianto, “Metode Interpolasi: Suatu Kajian Memecahkan Data Tahunan Menjadi Data Kuartalan dan Bulanan,” *Jurnal Ilmiah Sains*, vol. 24, no. 2, pp. 120–132, 2024, doi: 10.35799/jis.v24i2.55099.
- [14] D. Rahmawati, T. Kristanto, B. F. Setya Pratama, and D. B. Abiansa, “Prediksi Pelaku Perjalanan Luar Negeri di Masa Pandemi COVID-19 Menggunakan Metode Regresi Linier Sederhana,” *Journal of Information System Research (JOSH)*, vol. 3, no. 3, pp. 338–343, 2022, doi: 10.47065/josh.v3i3.1507.
- [15] N. Rohmawati, I. S. N. Inayah, and A. Wibowo, “Perbandingan Penggunaan Phytion dan Excel dalam Menyelesaikan Persamaan Tak Linier Metode Newton Raphson,” *Numerical: Jurnal Matematika dan Pendidikan Matematika*, vol. 9, no. 1, pp. 87–97, 2025, doi: 10.25217/numerical.v9i1.5798.
- [16] Bayhaqi Yasri, Fauziah Mawaddah Harefa, Maulia Fadila, Nazira Ananda, and Siti Salamah Br Ginting, “Penerapan Program Linier untuk Menyelesaikan Masalah Alokasi Sumber Daya pada Persoalan Penugasan,” *Algoritma : Jurnal Matematika, Ilmu pengetahuan Alam, Kebumihan dan Angkasa*, vol. 3, no. 4, pp. 82–88, 2025, doi: 10.62383/algoritma.v3i4.620.



- [17] N. Z. Wagiu, S. V. Pandey, and S. Y. R. Rompis, “Hubungan Waktu Tempuh dengan Volume Lalu Lintas Pada Jalan Perkotaan (Studi Kasus: Jl. Wolter Mongisidi Segmen Depan Bank BTN Sampai Jl. Sea),” *TEKNO*, vol. 22, no. 88, pp. 749–760, 2024, doi: <https://doi.org/10.35793/jts.v22i88.55301>.
- [18] E. Nursanti, R. M. Suaidy Avief, Sibut, and M. Kertaningtyas, “Peningkatan Efisiensi Waktu dan Biaya Pemeliharaan Overhaul Pesawat Tempur Ellysa,” *Jurnal Teknologi dan Manajemen Industri*, vol. 4, no. 2, pp. 1–6, 2018, doi: 10.36040/jtmi.v4i2.231.
- [19] M. T. Afifudin and D. P. Sahar, “Model Optimisasi Untuk Masalah Minimisasi Waktu Perjalanan Wisata Tur-Tunggal di Daerah Kepulauan,” *Arika*, vol. 15, no. 1, pp. 15–28, 2021, doi: 10.30598/arika.2021.15.1.15.
- [20] B. S. Waloejo and G. Prayitno, “Perhitungan waktu Tempuh Koridor Jl. Sukarno-Hatta,” *Region : Jurnal Pembangunan Wilayah dan Perencanaan Partisipatif*, vol. 15, no. 2, p. 262, 2020, doi: 10.20961/region.v15i2.40565.
- [21] T. Briand and P. Monasse, “Theory and Practice of Image B-Spline Interpolation,” *Image Processing On Line*, vol. 8, pp. 99–141, Jul. 2018, doi: 10.5201/ipol.2018.221.
- [22] O. Suhendra, M. A. Tiro, and Ruliana, “Ordinary Least Square Method in Multiple Regression Analysis to Estimating Coefficients of Factors Affecting Human Development Index,” *ARRUS Journal of Mathematics and Applied Science*, vol. 2, no. 1, pp. 49–59, Feb. 2022, doi: 10.35877/mathscience632.
- [23] L. Xu, “Application of the Newton Iteration Algorithm to the Parameter Estimation for Dynamical Systems,” *Journal of Computational and Applied Mathematics*, vol. 288, pp. 33–43, Nov. 2015, doi: 10.1016/j.cam.2015.03.057.
- [24] R. L. Rosa *et al.*, “Integer Linear Programming for Optimizing Drone-Based Delivery Routes,” *Engineering Journal*, vol. 29, no. 11, pp. 23–38, Nov. 2025, doi: 10.4186/ej.2025.29.11.23.