VOCATECH: Vocational Education and Technology Journal 8, 2 (2026): 162-171
Website: http://ojs.aknacehbarat.ac.id/index.php/vocatech/index
ISSN 2716-5183 (online) ISSN 2686-4770 (print)

PREDIKSI RSRP 4G LTE KOTA PADANG MENGGUNAKAN RANDOM FOREST REGRESSION BERBASIS DATA *DRIVE TEST*

Afrizal Yuhanef*
Siska Aulia
Herry Setiawan
Azhranadya

Program Studi Teknik Telekomunikasi, Jurusan Teknik Elektro, Politeknik Negeri Padang
Kampus Politeknik Negeri Padang, Limau Manis, Padang, Sumatera Barat 25163, INDONESIA

Abstract

The development of 4G LTE networks in Padang City necessitates accurate evaluation of signal coverage quality to support network optimization and quality of service management. Reference Signal Received Power (RSRP) is the primary parameter for assessing signal strength received by User Equipment (UE) from eNodeB. RSRP values obtained through drive testing are highly fluctuating due to radio propagation conditions, urban obstacles, and inter-cell interference, making conventional approaches insufficient for modeling the nonlinear signal distribution. This study applies a Random Forest Regression model to predict RSRP based on drive test data collected across Padang City, a region with diverse topography ranging from coastal areas to highland terrain. A dataset of 32,454 data points was obtained after preprocessing from 33,682 raw records. Input variables consisted of latitude, longitude, RSRQ, and SINR, with RSRP as the target variable. The dataset was divided into 70% training and 30% testing subsets (random_state = 42). Model performance was evaluated using MAE, RMSE, MSE, and R^2. Results yielded MAE of 3.46 dBm, RMSE of 4.71 dBm, and R^2 of 77.43%, indicating the model reliably captures the majority of actual RSRP variation. Spatial coverage mapping revealed Good-to-Very Good signal quality dominating the city center, while Fair-to-Bad categories concentrated in peripheral areas such as Lubuk Kilangan and Lubuk Minturun. These findings confirm Random Forest Regression as a viable data-driven approach for RSRP prediction and spatial network coverage analysis in geographically diverse urban regions.

Keywords:
G LTE; *drive test*; *prediction model*; *random forest regression*; *RSRP.*

Abstrak

Perkembangan jaringan 4G LTE di Kota Padang mendorong kebutuhan evaluasi kualitas cakupan sinyal secara akurat guna mendukung optimasi jaringan dan manajemen kualitas layanan. RSRP merupakan parameter utama untuk mengukur kekuatan sinyal yang diterima UE dari eNodeB. Nilai RSRP hasil *drive test* bersifat fluktuatif akibat kondisi propagasi radio, hambatan lingkungan perkotaan, dan interferensi antar sel, sehingga pendekatan konvensional tidak memadai untuk memodelkan distribusi sinyal nonlinier. Penelitian ini menerapkan model Random Forest Regression untuk memprediksi nilai RSRP berbasis data *drive test* di Kota Padang, wilayah dengan topografi beragam mulai dari pesisir pantai hingga kawasan perbukitan. Dataset 32.454 titik data diperoleh setelah pra-pengolahan dari 33.682 data mentah. Variabel input: latitude, longitude, RSRQ, dan SINR; variabel target: RSRP. Dataset dibagi 70% latih dan 30% uji (random_state = 42). Kinerja model dievaluasi menggunakan MAE, RMSE, MSE, dan R^2. Hasil: MAE 3,46 dBm, RMSE 4,71 dBm, R^2 77,43%, mengindikasikan model merepresentasikan sebagian besar variasi RSRP aktual. Pemetaan spasial menunjukkan dominasi *Good* hingga *Very Good* di pusat kota, sementara *Fair* hingga *Bad* di wilayah pinggiran seperti Lubuk Kilangan dan Lubuk Minturun. Temuan ini mengonfirmasi Random Forest Regression sebagai pendekatan berbasis data yang andal untuk prediksi RSRP dan analisis cakupan spasial pada wilayah dengan kondisi geografis beragam.

Kata Kunci:
4G LTE; *drive test*; model prediksi; *random forest regression*; RSRP.

DOI: 10.38038/vocatech.v8i2.340
Received: 03 June 2026; Accepted: 25 October 2026; Published: 30 October 2026

***Corresponding author:**
Afrizal Yuhanef, Politeknik Negeri Padang, Kampus Politeknik Negeri Padang, Limau Manis, Kota Padang, Sumatera Barat 25163, INDONESIA

Citation in APA Style: Yuhanef, A., Aulia, S., Setiawan, H., & Azhranadya. (2022). Prediksi RSRP 4G LTE Kota Padang menggunakan random forest regression berbasis data drive test. *VOCATECH: Vocational Education and Technology Journal, 8*(2), 162-171.

Email: afrizal@pnp.ac.id

1. PENDAHULUAN

Jaringan *Long Term Evolution* (LTE) generasi keempat (4G) merupakan standar komunikasi seluler yang dikembangkan oleh *3rd Generation Partnership Project* (3GPP) dan saat ini menjadi tulang punggung infrastruktur komunikasi seluler di Indonesia, termasuk di Kota Padang, Sumatera Barat. Berdasarkan 3GPP TS 36.306, teknologi LTE dirancang untuk mendukung kecepatan unduh puncak hingga 100 Mbps dan unggah hingga 50 Mbps dengan latensi rendah, sehingga memenuhi tuntutan layanan data *broadband mobile* yang terus meningkat (Yonoriadi et al., 2026). Namun demikian, distribusi kualitas sinyal di lapangan tidak selalu merata, terutama di wilayah dengan karakteristik geografis beragam seperti Kota Padang, yang memiliki luas 694,96 km² mencakup 11 kecamatan dengan ketinggian 0 hingga 1.853 mdpl dan lebih dari 51% kawasan berupa hutan (Badan Stastistik Padang, 2025). Keragaman topografi ini berpotensi signifikan memengaruhi karakteristik propagasi sinyal radio, sehingga distribusi kekuatan sinyal LTE tidak seragam di seluruh wilayah layanan.

Parameter utama dalam evaluasi kualitas cakupan sinyal LTE adalah *Reference Signal Received Power* (RSRP), yaitu rata-rata daya sinyal referensi yang diterima *User Equipment* (UE) dari eNodeB pada subcarrier tertentu (Putra et al., 2024). RSRP berperan krusial dalam analisis *Quality of Service* (QoS), pengambilan keputusan *handover*, serta perencanaan dan optimasi jaringan (Al Sulaiman & Saragih, 2024). Nilai RSRP yang diperoleh melalui pengukuran *drive test* di lapangan bersifat sangat fluktuatif, dipengaruhi oleh efek *multipath*, *shadowing*, *fading*, dan interferensi antar sel yang bersifat nonlinier, sehingga pendekatan pemodelan statistik linier konvensional tidak memadai untuk menangkap kompleksitas distribusinya.

Machine Learning (ML) menawarkan solusi yang lebih efektif dalam memodelkan hubungan nonlinier kompleks pada data jaringan seluler (Iliev et al., 2024). Algoritma berbasis *ensemble* terbukti menunjukkan performa unggul dalam memprediksi distribusi kualitas sinyal (Elmezughi et al., 2023). Kemampuan *machine learning* dalam memprediksi parameter kekuatan sinyal nirkabel telah dibuktikan melalui berbagai studi yang menunjukkan akurasi tinggi pada prediksi RSSI dan metrik radio lainnya, sehingga menjadi pendekatan yang menjanjikan untuk pemodelan kualitas sinyal seluler (Azoulay et al., 2023). *Random Forest Regression*, yang membangun sejumlah N *decision tree* secara independen melalui teknik *bootstrap aggregating* (bagging) dan *random feature selection*, dikenal stabil dan tahan terhadap *noise* dalam memodelkan data kualitas sinyal yang kompleks (Sagir & Tugcu, 2024). Beberapa penelitian terdahulu telah menerapkan *machine learning* (ML) untuk memprediksi parameter kualitas jaringan nirkabel, seperti *path loss*, *throughput*, dan cakupan sinyal, dengan hasil yang menunjukkan peningkatan akurasi dibandingkan pendekatan konvensional. Ringkasan penelitian-penelitian tersebut disajikan pada Tabel 1.

Tabel 1. Ringkasan Penelitian Terdahulu tentang Penerapan *Machine Learning* untuk Prediksi Kualitas Jaringan Nirkabel

No	Judul Penelitian	Penulis(Tahun)	Tema/Temuan
1	A Machine Learning Approach for Path Loss Prediction Using Combination of Regression and Classification Models	(Iliev et al., 2024)	Model prediksi *path loss* dengan kombinasi regresi dan klasifikasi ML; lebih unggul dari model empiris konvensional
2	Path loss modeling based on neural networks and ensemble method for future wireless networks	(Elmezughi et al., 2023)	Jaringan saraf dan metode *ensemble* untuk pemodelan path loss pada jaringan nirkabel masa depan
3	Prediksi Downlink Throughput Menggunakan Metode Random Forest Regression pada Jaringan 4G LTE Berdasarkan Data Drive Test	(Adhim et al., 2025a)	Random Forest Regression pada data drive test LTE untuk prediksi *downlink* throughput dengan tingkat kesalahan kecil
4	Hyperparameter Optimization of Random Forest for 5G Coverage Prediction	(Hajar Yuliana, 2024)	Optimasi *hyperparameter Random Forest* untuk prediksi 5G
5	A machine learning framework for predicting downlink throughput in 4G-LTE/5G cellular networks	(Al-Thaedan et al., 2024)	Kerangka kerja ML multi-algoritma untuk prediksi throughput 4G-LTE/5G
6	Comparative Analysis Of 4G LTE Network Quality At 900 Mhz And 2100 Mhz Frequencies.	(Putra et al., 2024)	Random Forest Regression untuk prediksi redaman jaringan fiber optik, R^2 = 98,32%

Meskipun demikian, terdapat kesenjangan (*gap*) penelitian yang signifikan: mayoritas studi terdahulu berfokus pada pemodelan *path loss* atau prediksi *throughput*, sementara penelitian yang secara khusus memodelkan dan memetakan distribusi RSRP secara spasial berbasis data *drive test* aktual—terutama pada wilayah dengan kondisi geografis beragam seperti Kota Padang—masih sangat terbatas. Kebaruan (*novelty*) penelitian ini terletak pada integrasi model prediksi RSRP berbasis ML dengan pemetaan cakupan spasial yang dapat langsung dimanfaatkan sebagai dasar evaluasi dan optimasi jaringan 4G LTE secara praktis. Pendekatan prediksi cakupan berbasis *machine learning* juga telah digunakan untuk memperkirakan distribusi sinyal pada area yang sulit dijangkau pengukuran langsung, namun implementasi pada data *drive test* aktual untuk prediksi RSRP masih relatif terbatas (Tarhuni et al., 2023).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk: (1) membangun model prediksi nilai RSRP menggunakan *Random Forest Regression* berbasis data *drive test* di Kota Padang; dan (2) menghasilkan peta cakupan spasial jaringan 4G LTE berdasarkan nilai RSRP yang diprediksi oleh model, sehingga dapat memberikan kontribusi metodologis dan praktis bagi perencanaan serta optimasi jaringan seluler di wilayah.

2. STUDI PUSTAKA

2.1. Jaringan 4G LTE

Long Term Evolution (LTE) merupakan standar jaringan seluler generasi keempat yang dikembangkan 3GPP sebagai evolusi dari teknologi 3G dan 3.5G (HSPA). Berdasarkan spesifikasi 3GPP TS 36.306, LTE dirancang untuk mendukung kecepatan unduh puncak 100 Mbps dan unggah 50 Mbps dengan latensi *user plane* di bawah 5 ms (Farras et al., 2025). Arsitektur LTE terdiri atas *Evolved Universal Terrestrial Radio Access Network* (E-UTRAN) yang menghubungkan UE dengan eNodeB, *Evolved Packet Core* (EPC) yang terhubung melalui antarmuka S1, serta antarmuka X2 antar eNodeB untuk mendukung proses *handover* (Haghrah et al., 2023).

2.2. *Drive test* dan Parameter Kualitas Sinyal LTE

Drive test adalah metode pengukuran lapangan yang dilakukan dengan membawa perangkat pengukur di dalam kendaraan bergerak untuk mengevaluasi cakupan aktual, parameter radio frequency (RF), tingkat interferensi, dan kualitas layanan jaringan seluler (Akram et al., 2023). Tiga parameter utama yang diukur dalam *drive test* jaringan 4G LTE adalah RSRP, RSRQ, dan SINR, yang disajikan dengan ambang batas klasifikasinya pada Tabel 2.

Tabel 2. Klasifikasi Parameter Kualitas Sinyal Jaringan 4G LTE

No	Parameter	Rentang Nilai	Satuan	Kategori
1	RSRP	≥ −75	dBm	Excellent
		−75 hingga −85	dBm	Very Good
		−85 hingga −100	dBm	Good
		−100 hingga −110	dBm	Fair
		< −110	dBm	Bad
2	RSRQ	≥ −9	dB	Excellent
		−10 hingga −9	dB	Very Good
		−15 hingga −10	dB	Good
		−19 hingga −15	dB	Fair
		< −20	dB	Bad
3	SINR	≥ 20	dB	Excellent
		13 hingga 20	dB	Good
		0 hingga 13	dB	Fair
		< 0	dB	Bad

Berdasarkan Tabel 2, RSRP merupakan indikator kekuatan sinyal dengan kategori *Excellent* pada nilai ≥−75 dBm dan *Bad* pada nilai < −110 dBm. RSRQ menunjukkan kualitas sinyal dengan mempertimbangkan interferensi, sementara SINR menyatakan rasio daya sinyal terhadap total daya interferensi dan *noise*. Ketiga parameter ini digunakan sebagai variabel input model karena masing-masing mencerminkan aspek berbeda dari kondisi kanal radio yang secara bersama-sama memengaruhi nilai RSRP (Imoize et al., 2020).

2.3. Machine *Learning* untuk Prediksi Kualitas Jaringan

Machine Learning (ML) adalah cabang kecerdasan buatan yang memungkinkan komputer mempelajari pola dari data historis dan membuat prediksi tanpa diprogram secara eksplisit (Imoize et al., 2020). Dalam konteks prediksi kualitas jaringan seluler, pendekatan supervised learning digunakan di mana model dilatih menggunakan data berlabel (pasangan input-output) kemudian diaplikasikan pada data baru (Al-Thaedan et al., 2024). Untuk prediksi nilai RSRP yang merupakan variabel kontinu dalam satuan dBm, metode regresi merupakan pendekatan yang tepat (Hajiar Yuliana, 2024).

2.4. Random Forest Regression

Random Forest Regression adalah algoritma ensemble learning yang membangun sejumlah N decision tree secara independen menggunakan dua mekanisme utama: (1) *Bootstrap Aggregating* (*Bagging*)—setiap pohon dilatih pada subset acak dari data latih yang diambil dengan pengembalian (with replacement), sehingga setiap pohon memiliki data latih yang berbeda; dan (2) *Random Feature Selection*—pada setiap pemisahan node, hanya subset acak dari total fitur yang dipertimbangkan, memastikan keberagaman (*diversity*) antar pohon. Nilai prediksi akhir merupakan rata-rata dari seluruh pohon:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} \hat{y}_i \quad \text{.....(1)}$$

Bahwa $\hat{y}$ adalah nilai prediksi akhir, N adalah jumlah pohon (n_estimators), dan $\hat{y}_i$ adalah prediksi pohon ke-i. Mekanisme ganda ini memastikan setiap pohon bersifat beragam dan tidak berkorelasi satu sama lain, sehingga rata-ratanya secara kolektif mengurangi varians dan risiko *overfitting* (Putra et al., 2024).

Keunggulan Random Forest meliputi: kemampuan menangani data berdimensi tinggi, memodelkan hubungan nonlinier yang kompleks, ketahanan terhadap noise dan pencilan, serta tidak memerlukan asumsi distribusi data tertentu (Sagir & Tugcu, 2024). Selain digunakan pada permasalahan regresi umum, Random Forest juga telah diterapkan dalam konstruksi dan pembaruan Radio Environment Map berbasis pengukuran sinyal, menunjukkan kemampuan yang baik dalam menangkap pola spasial pada lingkungan komunikasi nirkabel (Camana et al., 2023).

2.5. Metrik Evaluasi Model

Kinerja model regresi dievaluasi menggunakan empat metrik. Mean Absolute Error (MAE) mengukur rata-rata nilai absolut selisih aktual-prediksi: MAE = $(1/n)\ \Sigma|\hat{y}_i - y_i|$. *Mean Squared Error* (MSE) memberikan bobot lebih besar pada kesalahan besar: MSE = $(1/n)\ \Sigma(\hat{y}_i - y_i)^2$. Root Mean Square Error (RMSE = $\sqrt{}$MSE) memiliki satuan yang sama dengan variabel target sehingga lebih mudah diinterpretasi. Koefisien determinasi (R^2) menunjukkan proporsi variasi data aktual yang dijelaskan model, dengan nilai mendekati 100% yang mengindikasikan kesesuaian sangat baik.v

3. METODE PENELITIAN

3.1. Jenis dan Lokasi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksperimen berbasis data (data-driven). Pengambilan data *drive test* dilaksanakan di Kota Padang, Provinsi Sumatera Barat, pada bulan Januari 2026. Pemilihan Telkomsel sebagai operator didasarkan pada posisinya sebagai operator dengan pangsa pelanggan terbesar di Indonesia dan penetrasi jaringan 4G LTE yang paling luas di Kota Padang, sehingga data yang diperoleh merepresentasikan kondisi jaringan dominan di wilayah tersebut. Total rute *drive test* mencakup jalan-jalan utama dan sekunder di seluruh 11 kecamatan Kota Padang, mulai dari kawasan pusat kota (Padang Utara, Padang Timur, Padang Selatan) hingga wilayah pinggiran (Lubuk Kilangan, Lubuk Minturun, Koto Tangah), dengan durasi pengukuran selama empat hari kerja untuk memperoleh representasi data yang komprehensif.

3.2. Perangkat Penelitian

Perangkat keras yang digunakan terdiri dari: (a) smartphone Samsung Galaxy S5 sebagai UE pengukuran, dipilih karena kompatibilitas penuh dengan perangkat lunak TEMS Pocket; (b) SIM Card LTE Telkomsel aktif; (c) laptop untuk pemrosesan data lapangan; dan (d) *dongle* untuk konektivitas tambahan.

Perangkat lunak yang digunakan adalah: (a) TEMS Pocket v16.0.3.70 untuk pencatatan parameter jaringan secara *real-time* di lapangan; (b) TEMS Discovery untuk analisis dan ekspor data hasil pengukuran; serta (c) Python 3.x dengan library pandas, numpy, scikit-learn v1.3, dan matplotlib untuk pra-pengolahan data, pemodelan, dan visualisasi.

3.3. Pengumpulan dan Pra-Pengolahan Data

Data mentah hasil *drive test* berjumlah 33.682 baris yang mencatat nilai RSRP, RSRQ, SINR, serta koordinat geografis (latitude dan longitude) setiap titik pengukuran. Pra-pengolahan data dilaksanakan melalui tahapan: (1) Pembersihan duplikat: 1.228 data duplikat berdasarkan koordinat geografis dihapus menggunakan Microsoft Excel, menghasilkan 32.454 baris data unik; (2) Pemeriksaan kelengkapan: pemeriksaan *missing value* menggunakan Python (pandas) menunjukkan data lengkap tanpa nilai kosong; (3) Penetapan variabel: variabel input (fitur) ditetapkan sebagai latitude, longitude, RSRQ, dan SINR; variabel target adalah RSRP; (4) Pembagian dataset: menggunakan *random split* dengan proporsi 70% data latih (22.718 data) dan 30% data uji (9.736 data), parameter random_state = 42 ditetapkan untuk memastikan *reproducibility* hasil.

3.4. Pemodelan Random Forest Regression

Model *Random Forest Regression* dibangun menggunakan library scikit-learn v1.3 pada Python dengan konfigurasi parameter: n_estimators = 100, max_depth = 20, dan random_state = 42. Penetapan n_estimators = 100 didasarkan pada eksperimen awal yang menunjukkan bahwa penambahan pohon di atas 100 tidak memberikan peningkatan R^2 yang signifikan namun meningkatkan waktu komputasi secara proporsional. Nilai max_depth = 20 dipilih sebagai keseimbangan antara kemampuan model menangkap pola kompleks dan pengendalian risiko *overfitting*. Pelatihan model dilaksanakan menggunakan data latih (22.718 data), dan prediksi dihasilkan pada data uji (9.736 data).

3.5. Evaluasi Kinerja dan Pemetaan Spasial

Kinerja model dievaluasi menggunakan MAE, MSE, RMSE, dan R^2 sebagaimana dijelaskan pada Subbab 2.5. Nilai RSRP hasil prediksi kemudian dipetakan secara spasial berdasarkan koordinat latitude dan longitude menggunakan library matplotlib pada Python. Setiap titik data dikategorikan ke dalam kelas kualitas sinyal (*Excellent, Very Good, Good, Fair, Bad*) berdasarkan ambang batas pada Tabel 1.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengolahan dan Statistik Deskriptif Dataset

Proses pra-pengolahan menghasilkan dataset bersih sebanyak 32.454 baris dengan 5 kolom variabel (latitude, longitude, RSRP, RSRQ, dan SINR). Dari total 33.682 data mentah, 1.228 data duplikat berhasil diidentifikasi dan dihapus, serta tidak ditemukan *missing value* sehingga seluruh data siap digunakan dalam pemodelan. Statistik deskriptif ketiga variabel parameter kualitas jaringan disajikan dalam Tabel 2.

Tabel 3. Statistik Deskriptif Variabel Dataset *Drive test* (N = 32.454)

No	Parameter	Min	Max	Mean	Std Dev
1	RSRP (dBm)	−127,6	−48,2	−92,7	9,8
2	RSRQ (dB)	−19,9	−3,1	−12,8	3,1
3	SINR (dB)	−11,2	29,5	7,4	6,5

Berdasarkan Tabel 2, nilai RSRP memiliki rentang lebar dari −127,6 dBm hingga −48,2 dBm, dengan rata-rata −92,7 dBm (kategori Good) dan standar deviasi 9,8 dBm. Variasi yang cukup signifikan ini konsisten dengan karakteristik propagasi sinyal radio di lingkungan perkotaan yang kompleks. Variabel RSRQ memiliki rata-rata −12,8 dB (kategori Good), sedangkan SINR rata-rata 7,4 dB (kategori Fair). Nilai standar deviasi yang moderat pada ketiga variabel mengindikasikan distribusi data cukup tersebar namun tidak skewed ekstrem, sehingga layak digunakan sebagai fitur input model Machine Learning (Imoize et al., 2020).

4.2. Analisis Sebaran RSRP Hasil Drive test

Distribusi kategori RSRP dari hasil *drive test* aktual di Kota Padang disajikan dalam Tabel 3.

Tabel 4. Distribusi Kategori RSRP Hasil *Drive test* di Kota Padang

No	Kategori	Rentang RSRP (dBm)	Jumlah Sampel	Persentase (%)
1	Excellent	≥ −75	2.567	8
2	Very Good	−75 hingga −85	8.161	24
3	Good	−85 hingga −100	17.963	53
4	Fair	−100 hingga −110	4.370	13
5	Bad	< −110	619	2
	Total		**32.454**	**100**

Berdasarkan Tabel 3, kategori *Good* mendominasi distribusi RSRP dengan proporsi 53%, diikuti *Very Good* 24%. Secara kumulatif, kategori *Good* hingga *Excellent* mencakup ≈85% dari total data, mengindikasikan layanan 4G LTE Telkomsel di Kota Padang secara umum dalam kondisi memadai untuk mendukung layanan data. Distribusi spasial menunjukkan kategori *Good* dan *Very Good* terkonsentrasi di kawasan permukiman padat, perkantoran, pusat perbelanjaan, dan lingkungan kampus di pusat kota, yang merupakan area dengan kepadatan eNodeB tinggi dan kondisi *line-of-sight* (LoS) lebih baik. Sebaliknya, kategori *Fair* (13%) dan *Bad* (2%) teridentifikasi di wilayah pinggiran seperti Lubuk Kilangan dan Lubuk Minturun—area *cell edge* dengan kepadatan infrastruktur lebih rendah—yang konsisten dengan karakteristik propagasi radio perkotaan (Barry et al., 2026).

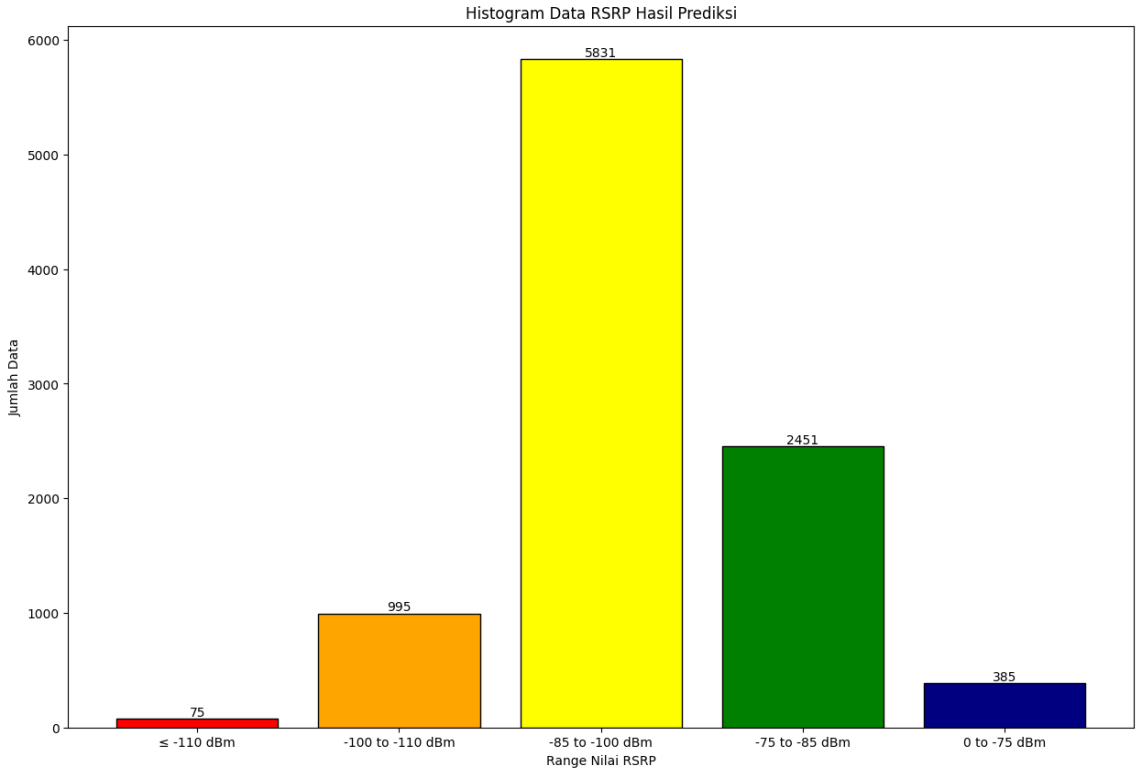
4.3. Hasil Prediksi RSRP Model Random Forest Regression

Model Random Forest Regression menghasilkan prediksi nilai RSRP pada 9.736 data uji. Tabel 4 menampilkan 10 sampel representatif perbandingan nilai RSRP aktual dan prediksi.

Tabel 5. Sampel Hasil Prediksi Random Forest Regression (10 data representatif)

No.	Latitude	Longitude	Aktual (dBm)	Prediksi (dBm)	Selisih (dBm)
1	−0,9574814	100,4032773	−82,4	−82,88	−0,48
2	−0,9231977	100,4480057	−89,0	−86,17	+2,83
3	−0,9634053	100,3586730	−79,1	−78,08	+1,02
4	−0,9429602	100,3569683	−86,5	−83,42	+3,08
5	−0,9826787	100,3766251	−70,9	−71,44	−0,54
6	−0,8796694	100,3938208	−80,8	−83,28	−2,48
7	−0,9543140	100,3643356	−75,8	−78,29	−2,49
8	−0,8928924	100,3948135	−92,3	−91,12	+1,18
9	−0,9226885	100,4103461	−90,5	−98,20	−7,70
10	−0,9154076	100,3501578	−96,9	−88,29	+8,61

Berdasarkan Tabel 4, mayoritas sampel menunjukkan prediksi yang mendekati nilai aktual dengan selisih absolut di bawah 5 dBm (sampel 1–8). Selisih yang lebih besar terjadi pada sampel 9 (selisih −7,70 dBm) dan sampel 10 (+8,61 dBm). Sampel 9 berlokasi di koordinat (−0,9227; 100,4103) yang merupakan area transisi antar sel (cell edge) sehingga nilai RSRP aktual sangat fluktuatif dan sulit diprediksi secara presisi oleh model. Sampel 10 berada di kawasan yang kemungkinan terpengaruh efek temporary shadowing saat pengukuran, sehingga nilai aktualnya menyimpang dari tren umum lokasi tersebut. Secara keseluruhan, selisih rata-rata yang kecil mengindikasikan model mampu menangkap pola umum distribusi RSRP. Perbandingan distribusi secara visual ditampilkan pada Gambar 2.



Gambar 1. Histogram Perbandingan Distribusi RSRP Aktual dan Prediksi (Data Uji: 9.736 titik)

Berdasarkan Gambar 1, pada kategori Good (−85 hingga −100 dBm), jumlah data prediksi (5.831) lebih tinggi dari aktual (5.099). Pada kategori Excellent dan Bad, data prediksi lebih rendah dari aktual (385 vs. 716 dan 75 vs. 178). Kecenderungan underprediksi pada nilai-nilai ekstrem ini merupakan karakteristik inheren Random Forest akibat mekanisme rata-rata prediksi banyak pohon yang menghasilkan regression toward the mean, konsisten dengan temuan (Adhim et al., 2025b; Putra et al., 2024).

4.4. Evaluasi Kinerja Model

Hasil evaluasi kinerja model Random Forest Regression pada data uji disajikan dalam Tabel 5.

Tabel 6. Hasil Evaluasi Kinerja Model Random Forest Regression

No.	Metrik Evaluasi	Nilai	Satuan
1	Mean Absolute Error (MAE)	3,46	dBm
2	Mean Squared Error (MSE)	22,18	dBm^2
3	Root Mean Square Error (RMSE)	4,71	dBm
4	Koefisien Determinasi (R^2)	77,43	%

Berdasarkan Tabel 5, nilai MAE sebesar 3,46 dBm menunjukkan rata-rata penyimpangan absolut prediksi di bawah 4 dBm. Dalam praktik optimasi jaringan LTE, selisih prediksi di bawah 5 dBm masih tergolong dapat diterima karena variasi sinyal radio secara alami dapat berubah akibat kondisi propagasi dan dinamika trafik (Sagir & Tugcu, 2024). Nilai RMSE sebesar 4,71 dBm yang tidak jauh dari MAE mengindikasikan tidak banyak outlier kesalahan prediksi ekstrem, sedangkan MSE sebesar 22,18 dBm^2 melengkapi gambaran distribusi kesalahan.

Nilai R^2 sebesar 77,43% menunjukkan model mampu menjelaskan 77,43% variasi total nilai RSRP aktual hanya menggunakan empat variabel input (latitude, longitude, RSRQ, dan SINR), tanpa memerlukan data infrastruktur jaringan seperti posisi eNodeB atau tinggi antena. Sisa 22,57% yang tidak dapat dijelaskan kemungkinan berasal dari faktor yang tidak dimodelkan: efek shadowing bangunan individual, variasi beban trafik dinamis, perubahan kondisi kanal temporal, dan karakteristik antena eNodeB setempat (Imoize et al., 2020).

Untuk mengontekstualisasikan performa ini, Tabel 6 menyajikan perbandingan dengan penelitian-penelitian terdahulu yang relevan.

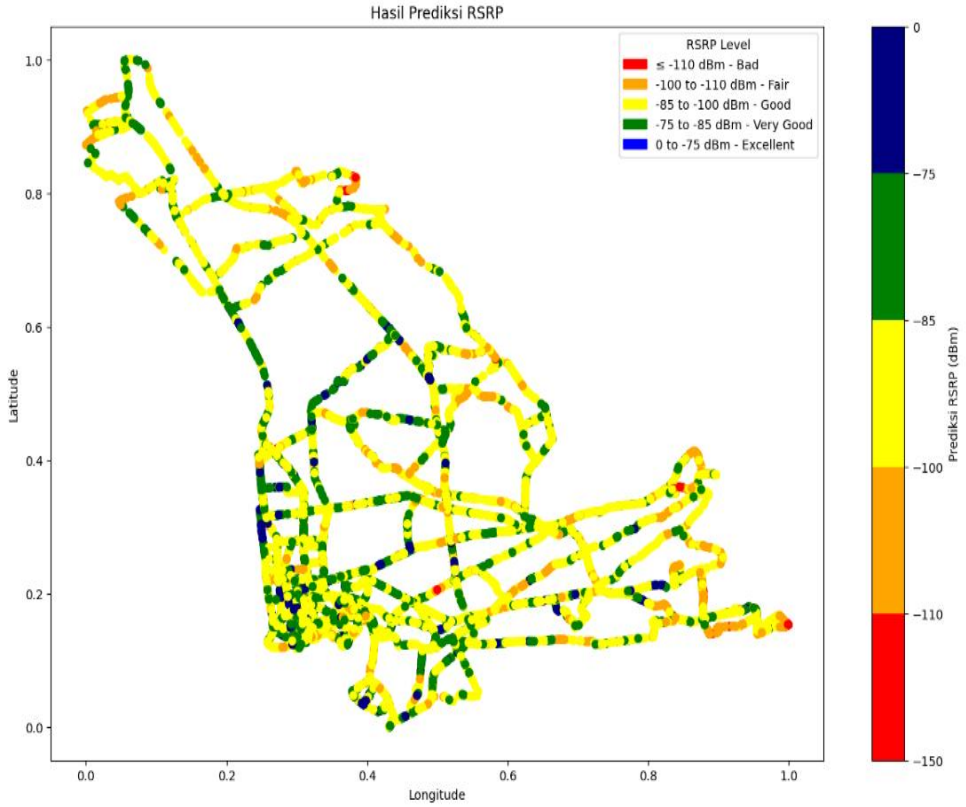
Tabel 1. Perbandingan Hasil Penelitian dengan Studi Terdahulu

Penelitian	Parameter Prediksi	Algoritma	Dataset	MAE/RMSE	R^2 (%)
Iliev et al. (2024)	Path loss	ML Reg.+Klasif.	Lapangan	—	—
Elmezughi et al. (2023)	Path loss	ANN + Ensemble	Lapangan	—	Tinggi
Sagir & Tugcu (2024)	Path loss (V2V)	Random Forest	Highway	< 5 dBm	—
Adhim et al. (2025)	DL Throughput	Random Forest	*Drive test* LTE	Kecil	—
Hajiar Yuliana (2024)	Coverage 5G	RF + Optim.	Simulasi	—	Tinggi
Al-Thaedan et al. (2024)	DL Throughput	Multi-ML	LTE/5G	Kecil	Baik
Putra et al. (2025)	Redaman FO	Random Forest	Data historis	0,10/0,24	98,32
Penelitian ini	**RSRP + Spasial**	**Random Forest Reg.**	***Drive test* Padang**	**3,46/4,71**	**77,43**

Berdasarkan Tabel 6, performa model penelitian ini (MAE = 3,46 dBm, R^2 = 77,43%) berada pada level kompetitif. Nilai R^2 yang lebih rendah dibanding daripada model prediksi redaman fiber optik berbasis data historis (Putra et al., 2024), yang dapat dijelaskan oleh perbedaan karakteristik variabel target: nilai redaman fiber optik memiliki variasi yang lebih terpola dan dapat diprediksi dibanding nilai RSRP sinyal radio yang sangat dipengaruhi kondisi propagasi dinamis. Kebaruan utama penelitian ini adalah fokus pada pemodelan RSRP secara langsung (bukan path loss atau throughput) yang diintegrasikan dengan pemetaan cakupan spasial berbasis data *drive test* aktual dari wilayah dengan kondisi geografis beragam—sebuah pendekatan yang belum banyak dilakukan dalam literatur yang ada.

4.5. Pemetaan Cakupan Spasial Jaringan 4G LTE

Pemetaan spasial nilai RSRP hasil prediksi model ditampilkan pada Gambar 3. Peta ini dibangun menggunakan koordinat latitude dan longitude dari data uji yang dikombinasikan dengan nilai RSRP prediksi, kemudian dikategorikan berdasarkan ambang batas kualitas sinyal pada Tabel 1.



Gambar 2. Peta Cakupan Spasial Prediksi RSRP Jaringan 4G LTE di Kota Padang

Berdasarkan Gambar 2, wilayah pusat Kota Padang dan ruas-ruas jalan utama—termasuk Kecamatan Padang Utara, Padang Timur, dan kawasan Pasar Raya Padang—didominasi kategori Very Good hingga Good yang mencerminkan kepadatan eNodeB lebih tinggi dan kondisi LoS yang relatif baik (Imoize et al., 2020). Kawasan permukiman padat, perkantoran pemerintah, pusat perbelanjaan, serta lingkungan kampus (Universitas Andalas dan Politeknik Negeri Padang) juga menunjukkan kualitas sinyal yang baik.

Sebaliknya, wilayah pinggiran seperti Lubuk Kilangan dan Lubuk Minturun yang berada di kawasan perbukitan dengan tutupan hutan lebat menampilkan dominasi kategori Fair hingga Bad. Kondisi ini berkaitan langsung dengan jarak yang lebih jauh dari eNodeB terdekat (cell edge), hambatan topografi berupa perbukitan, dan kepadatan vegetasi yang menyebabkan pelemahan sinyal signifikan (Elmezughi et al., 2023). Hasil pemetaan ini memiliki nilai aplikatif tinggi: identifikasi visual area Fair dan Bad memungkinkan perencana jaringan memprioritaskan upaya optimasi secara tepat sasaran, seperti penambahan site baru, penyesuaian parameter antena (tilt dan azimuth), atau penerapan teknologi small cell di area cakupan rendah (Farras et al., 2025). Pendekatan pemetaan berbasis prediksi ML ini lebih efisien dibandingkan *drive test* berulang yang memerlukan sumber daya operasional lebih besar.

5. SIMPULAN

Penelitian ini berhasil membangun model prediksi nilai RSRP jaringan 4G LTE di Kota Padang menggunakan algoritma *Random Forest Regression* berbasis data *drive test*, dengan kinerja yang ditunjukkan melalui nilai MAE sebesar 3,46 dBm, RMSE sebesar 4,71 dBm, MSE sebesar 22,18 dBm^2, dan R^2 sebesar 77,43%, mengindikasikan bahwa model mampu menjelaskan lebih dari tiga perempat variasi RSRP aktual hanya menggunakan empat variabel input (latitude, longitude, RSRQ, dan SINR) tanpa memerlukan data infrastruktur jaringan secara eksplisit. Analisis distribusi RSRP menunjukkan kondisi jaringan 4G LTE Telkomsel di Kota Padang didominasi kategori *Good* (53%) dan *Very Good* (24%), dengan kualitas sinyal terbaik di pusat kota dan koridor jalan utama, sementara wilayah pinggiran seperti Lubuk Kilangan dan Lubuk Minturun masih menghadapi kondisi *Fair* hingga *Bad* akibat faktor topografi, jarak dari eNodeB, dan kepadatan vegetasi. Peta cakupan spasial berbasis prediksi yang dihasilkan memberikan visualisasi komprehensif distribusi kualitas sinyal di seluruh wilayah layanan sehingga dapat berfungsi sebagai acuan akurat dan efisien bagi pengambil keputusan dalam perencanaan, evaluasi, dan optimasi jaringan 4G LTE tanpa bergantung sepenuhnya pada kegiatan *drive test* berulang yang memerlukan sumber daya operasional lebih besar. Penelitian selanjutnya disarankan untuk mengeksplorasi penambahan variabel input (jarak UE ke eNodeB, informasi sektor antena), membandingkan dengan algoritma ML lain (*XGBoost*, ANN), serta mengintegrasikan hasil pemetaan dengan sistem informasi geografis (GIS) untuk analisis perencanaan jaringan yang lebih komprehensif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh civitas akademika Program Studi D4 Teknik Telekomunikasi, Jurusan Teknik Elektro, Politeknik Negeri Padang atas dukungan fasilitas, bimbingan, dan sumber daya yang diberikan selama pelaksanaan penelitian ini. Terima kasih juga disampaikan kepada pihak Telkomsel yang jaringannya menjadi objek pengukuran dalam penelitian ini.

DAFTAR PUSTAKA

Adhim, R. M., Sulistyawan, V. N., & Sukamta, S. (2025a). Prediksi downlink throughput menggunakan metode Random Forest Regression pada jaringan 4G LTE berdasarkan data. *Jurnal Profesi Insinyur Universitas Lampung*, *6*(2). https://doi.org/10.23960/jpi.v6n2.190

Adhim, R. M., Sulistyawan, V. N., & Sukamta, S. (2025b). Prediksi downlink throughput menggunakan metode Random Forest Regression pada jaringan 4G LTE berdasarkan data drive test. *Jurnal Profesi Insinyur Universitas Lampung*, *6*(2). https://doi.org/10.23960/jpi.v6n2.190

Akram, A., Melvandino, F. H., Bragaswara, W. Y., & Ramza, H. (2023). Analisis kinerja jaringan 4G LTE menggunakan metode drive test di Kelurahan Kampung Rambutan, Jakarta Timur. *Jurnal Informatika Dan Teknik Elektro Terapan*, *11*(3). https://doi.org/10.23960/jitet.v11i3.3140

Al-Thaedan, A., Shakir, Z., Mjhool, A. Y., Alsabah, R., Al-Sabbagh, A., Nembhard, F., & Salah, M. (2024). A machine learning framework for predicting downlink throughput in 4G-LTE/5G cellular networks. *International Journal of Information Technology*, *16*(2), 651–657. https://doi.org/10.1007/s41870-023-01678-w

Al Sulaiman, R., & Saragih, Y. (2024). Analisis Quality of Service (QoS) jaringan provider IndiHome melalui drive test di Kabupaten Subang. *Aisyah Journal Of Informatics and Electrical Engineering (A.J.I.E.E)*, *6*(1),

110–117. https://doi.org/10.30604/jti.v6i1.219

Azoulay, R., Edery, E., Haddad, Y., & Rozenblit, O. (2023). Machine learning techniques for received signal strength indicator prediction. *Intelligent Data Analysis*, *27*(4), 1167–1184. https://doi.org/10.3233/IDA-226750

Badan Stastistik Padang, B. P. S. K. P. (2025). *Badan Pusat Statistik daerah Kota Padang*.

Barry, K. A., Camara, M. A., Niang, B., Diallo, M. S., Diakhaté, I., & Soumah, S. (2026). Comparative analysis and modeling of 4G LTE propagation using the Okumura-Hata and COST-231 Hata models, enhanced with spatially correlated log-normal shadowing. *Journal of Electromagnetic Analysis and Applications*, *18*(10), 175–216. https://doi.org/10.4236/jemaa.2026.1810010

Camana, M. R., Garcia, C. E., & Hwang, T. (2023). *applied sciences A REM Update Methodology Based on Clustering and Random Forest*.

Elmezughi, M. K., Salih, O., Afullo, T. J., & Duffy, K. J. (2023). Path loss modeling based on neural networks and ensemble method for future wireless networks. *Heliyon*, *9*(9), e19685. https://doi.org/10.1016/j.heliyon.2023.e19685

Farras, R., Yuhanef, A., & Pamungkas, M. P. (2025). Analisa QoS jaringan 4G LTE operator Telkomsel dan XL di Painan. *SULIWA: Jurnal Multidisiplin Teknik, Sains, Pendidikan Dan Teknologi*, *2*(1), 35–42. https://doi.org/10.62671/suliwa.v2i1.57

Haghrah, A., Abdollahi, M. P., Azarhava, H., & Niya, J. M. (2023). A survey on the handover management in 5G-NR cellular networks: Aspects, approaches and challenges. *EURASIP Journal on Wireless Communications and Networking*, *2023*(1), 52. https://doi.org/10.1186/s13638-023-02261-4

Hajiar Yuliana. (2024). Hyperparameter optimization of Random Forest for 5G coverage prediction. *Buletin Pos Dan Telekomunikasi*, *22*(1). https://doi.org/10.17933/bpostel.v22i1.390

Iliev, I., Velchev, Y., Petkov, P. Z., Bonev, B., Iliev, G., & Nachev, I. (2024). A machine learning approach for path loss prediction using combination of regression and classification models. *Sensors*, *24*(17), 5855. https://doi.org/10.3390/s24175855

Imoize, A. L., Orolu, K., & Atayero, A. A.-A. (2020). Analysis of key performance indicators of a 4G LTE network based on experimental data obtained from a densely populated smart city. *Data in Brief*, *29*, 105304. https://doi.org/10.1016/j.dib.2020.105304

Putra, D., Yuhanef, A., & Chandra, D. (2024). Comparative analysis of 4G LTE network quality at 900 MHz and 2100 MHz frequencies. *International Journal of Wireless And Multimedia Communications*, *1*(1), 39–46. https://doi.org/10.62671/jowim.v1i1.10

Sagir, N., & Tugcu, Z. H. (2024). Machine-learning-based path loss prediction for vehicle-to-vehicle communication in highway environments. *Applied Sciences*, *14*(17), 7545. https://doi.org/10.3390/app14177545

Tarhuni, N., Al Saadi, I., Asif, H. M., Mesbah, M., Eldirdiry, O., & Hossen, A. (2023). Machine-learning-based ground-level mobile network coverage prediction using UAV measurements. *Journal of Sensor and Actuator Networks*, *12*(3), 44. https://doi.org/10.3390/jsan12030044

Yonoriadi, Y., Dharmavid, D. Y. B., & Akbar, R. (2026). Evolusi teknologi komprehensif jaringan komunikasi nirkabel pada sistem 4G menuju 7G. *Jurnal Sistem Komputer (SISKOM)*, *6*(2), 136–149. https://doi.org/10.63447/siskom.v6i2.1888