UNIVERSITAS
GADJAH MADA
**Perbandingan Kinerja Model Regresi Linear Berganda dan Algoritma Random Forest untuk Penilaian Tanah
di Kalurahan Sidokarto pada Tahun 2026**
Haliza Zahira, Dr. Ir. Diyono, S.T., M.T., IPU.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# Intisari

Kalurahan Sidokarto, Kapanewon Godean, Kabupaten Sleman, Provinsi Daerah Istimewa Yogyakarta merupakan kawasan suburban yang mengalami perkembangan pesat sehingga nilai tanahnya beragam dan dinamis. Meningkatnya jumlah penduduk dari waktu ke waktu menyebabkan kebutuhan tanah turut bertambah, sedangkan ketersediaan tanah terbatas. Informasi nilai tanah yang akurat menjadi dasar penetapan tarif penerimaan negara bukan pajak layanan pertanahan, penyusunan rencana tata ruang, perhitungan biaya pengadaan tanah, dan transaksi jual beli tanah sehingga diperlukan metode penilaian yang tepat. Penelitian ini bertujuan membandingkan model Regresi Linear Berganda (RLB) log-lin dan algoritma *Random Forest Regression* (RFR) dalam mengestimasi nilai bidang tanah, menguji validasi kedua model, serta menyajikan peta estimasi nilai bidang tanah berdasarkan model dengan kinerja terbaik.

Data yang digunakan berupa 186 data transaksi jual beli tanah tahun 2022 – 2026 yang dikoreksi terhadap waktu penilaian dengan kenaikan nilai tanah 15% per tahun. Data dibagi menjadi 130 data (70%) untuk membentuk model dan 56 data (30%) sebagai data validasi. Sembilan variabel bebas yang diteliti adalah luas tanah (LT), kelas jalan (KJ), tutupan lahan (TL), dan jarak tanah ke jalan utama (JJU), fasilitas pendidikan (JFP), pusat kegiatan ekonomi (JPE), pusat pemerintahan (JPP), sungai (JS), serta pemakaman umum (JPU). Variabel jarak diukur dengan *Euclidean distance* dari batas terluar bidang tanah ke objek terdekat menggunakan *Join Attributes by Nearest* pada QGIS. Dalam pembentukan model, RLB hanya memuat 5 variabel karena mensyaratkan variabel yang signifikan dan bebas multikolinearitas, sedangkan RFR tidak bergantung pada asumsi tersebut sehingga tetap memuat keseluruhan 9 variabel.

Hasil uji validasi menunjukkan bahwa model RFR lebih unggul pada empat indikator, yaitu koefisien determinasi ($R^2$) sebesar 81,72% (RLB 61,6%), *Mean Absolute Error* (MAE) sebesar Rp 114.835/m² (RLB Rp 393.470/m²), *Mean Absolute Percentage Error* (MAPE) sebesar 6,9% (RLB 22,97%), dan *Coefficient of Variation* (COV) sebesar 8,16% (RLB 16,41%). Sementara itu, nilai *Price Related Differential* (PRD) kedua model setara, yakni 1,008 untuk RFR dan 1,005 untuk RLB, yang menunjukkan tidak adanya bias penilaian. Perbedaan kinerja kedua model terbukti signifikan secara statistik melalui *Wilcoxon Signed-Rank Test*. Dengan demikian, *Random Forest Regression* ditetapkan sebagai model dengan kinerja terbaik untuk menyusun peta estimasi nilai 7.242 bidang tanah yang diklasifikasikan ke dalam 8 kelas, dengan rentang nilai Rp 488.297/m² hingga Rp 6.875.862/m².

Kata Kunci: Penilaian Tanah, Nilai Bidang Tanah, Regresi Linear Berganda, *Random Forest*, Wilayah Suburban

UNIVERSITAS
GADJAH MADA
**Perbandingan Kinerja Model Regresi Linear Berganda dan Algoritma Random Forest untuk Penilaian Tanah
di Kalurahan Sidokarto pada Tahun 2026**
Haliza Zahira, Dr. Ir. Diyono, S.T., M.T., IPU.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# *Abstract*

*Sidokarto Village, located in Godean District, Sleman Regency, Special Region of Yogyakarta, is a rapidly developing suburban area characterized by diverse and dynamic land values. A growing population has led to an increased demand for land, even as the available supply remains limited. Accurate land value information is essential for determining non-tax state revenue rates for land services, formulating spatial planning, calculating land acquisition costs, and facilitating land transactions; therefore, an appropriate valuation method is required. This study aims to compare the log-linear Multiple Linear Regression (MLR) model with the Random Forest Regression (RFR) algorithm for estimating the value of land parcels, validate both models, and produce a land value estimation map based on the model that demonstrates the best performance.*

*The dataset consists of 186 land sale and purchase transactions from 2022 to 2026, adjusted to the valuation date by accounting for an annual land value appreciation rate of 15%. The data was split into a training set of 130 data (70%) for model development and a validation set of 56 data (30%). Nine independent variables were examined: land area (LT), road class (KJ), land cover (TL), and the distance from the land parcel to the nearest main road (JJU), educational facility (JFP), economic activity center (JPE), government center (JPP), river (JS), and public cemetery (JPU). Distances were measured using the Euclidean distance from the outermost boundary of each land parcel to the nearest object, calculated via the "Join Attributes by Nearest" function in QGIS. In the model development, the MLR model incorporated only 5 variables as it requires variables to be statistically significant and free from multicollinearity, whereas the RFR model, being independent of such assumptions, utilized all 9 variables.*

*Validation results indicate that the RFR model outperforms the MLR model across four indicators: a coefficient of determination ($R^2$) of 81.72% (compared to 61.6% for MLR), a Mean Absolute Error (MAE) of IDR 114,835/m² (IDR 393,470/m² for MLR), a Mean Absolute Percentage Error (MAPE) of 6.9% (22.97% for MLR), and a Coefficient of Variation (COV) of 8.16% (16.41% for MLR). Meanwhile, the Price-Related Differential (PRD) values for both models are comparable 1.008 for RFR and 1.005 for MLR indicating an absence of appraisal bias. The performance difference between the two models is statistically significant, as confirmed by the Wilcoxon Signed-Rank Test. Consequently, Random Forest Regression is identified as the best-performing model for generating an estimated value map for 7,242 land parcels, which are categorized into 8 classes with values ranging from IDR 488,297/m² to IDR 6,875,862/m².*

*Keywords: Land Valuation, Parcel Land Value, Multiple Linear Regression, Random Forest, Suburban Area*