

**ANALISIS DATA LONGITUDINAL DENGAN BERBAGAI TINGKAT
AUTOKORELASI DENGAN METODE NONPARAMETRIK *SPLINE
TRUNCATED* DAN *KERNEL REGRESSION***

Oleh

Putu Ratih Yulinda Eka Putri, NIM 2213101015

Jurusan Matematika dan Ilmu Pengetahuan Alam

ABSTRAK

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja metode regresi nonparametrik *Spline Truncated* dan *Kernel Regression* dalam menganalisis data longitudinal dengan berbagai tingkat autokorelasi, serta menilai konsistensi hasil simulasi ketika metode tersebut diterapkan pada data riil. Penelitian ini merupakan penelitian kuantitatif komparatif dengan pendekatan studi simulasi dan penerapan empiris. Data simulasi dibangkitkan pada 30 unit pengamatan dengan 5 titik waktu, menggunakan tiga tingkat autokorelasi, yaitu rendah, sedang, dan tinggi, yang masing-masing direplikasi sebanyak 50 kali. Selanjutnya, kedua metode diterapkan pada data riil persentase penduduk miskin di 34 provinsi di Indonesia periode 2016–2024. Pemilihan parameter optimal dilakukan menggunakan GCV, sedangkan evaluasi kinerja model dilakukan melalui MSE, RMSE, MAE, R^2 , serta diagnostik residual dengan ACF, PACF, dan uji Ljung–Box. Hasil penelitian menunjukkan bahwa peningkatan tingkat autokorelasi menyebabkan kesalahan prediksi pada kedua metode cenderung meningkat, sementara kemampuan model dalam menjelaskan variasi data cenderung menurun. Namun demikian, metode *Spline Truncated* konsisten menghasilkan nilai kesalahan prediksi yang lebih kecil, kurva estimasi yang lebih stabil, serta residual yang lebih terkendali dibandingkan metode *Kernel Regression*, terutama pada kondisi autokorelasi sedang hingga tinggi. Pada penerapan data riil, kedua metode menunjukkan performa in-sample yang relatif berdekatan, tetapi *Spline Truncated* lebih baik dalam stabilitas estimasi dan pengendalian autokorelasi residual. Dengan demikian, dapat disimpulkan bahwa metode *Spline Truncated* lebih sesuai digunakan untuk analisis data longitudinal yang menekankan kestabilan estimasi dan kualitas diagnostik residual. Berdasarkan hal tersebut, rekomendasi perbaikan dalam penelitian ini dapat dikembangkan melalui penggunaan spline orde lebih tinggi, penambahan variabel prediktor, serta pengkajian estimasi interval agar hasil analisis menjadi komprehensif dan mendalam.

Kata-kata kunci: data longitudinal, autokorelasi, *Spline Truncated*, *Kernel Regression*, regresi nonparametrik

**LONGITUDINAL DATA ANALYSIS WITH VARIOUS LEVELS OF
AUTOCORRELATION USING TRUNCATED SPLINE AND KERNEL
REGRESSION NONPARAMETRIC METHODS**

By

Putu Ratih Yulinda Eka Putri, Student ID 2213101015

Department of Mathematics and Natural Sciences

ABSTRACT

This study aims to evaluate and compare the performance of the Truncated Spline and Kernel Regression nonparametric methods in analyzing longitudinal data with various levels of autocorrelation, as well as to assess the consistency of simulation results when these methods are applied to real data. This research is a comparative quantitative study employing both simulation and empirical application approaches. The simulated data were generated for 30 observational units with 5 time points, using three levels of autocorrelation, namely low, moderate, and high, each replicated 50 times. Furthermore, both methods were applied to real data on the percentage of poor population in 34 provinces of Indonesia during the 2016–2024 period. The optimal parameters were selected using Generalized Cross Validation (GCV), while model performance was evaluated using MSE, RMSE, MAE, R^2 , and residual diagnostics through ACF, PACF, and the Ljung–Box test. The results indicate that increasing levels of autocorrelation tend to increase prediction errors in both methods, while the models' ability to explain data variation tends to decrease. However, the Truncated Spline method consistently produces smaller prediction errors, more stable estimation curves, and better-controlled residuals than the Kernel Regression method, particularly under moderate to high autocorrelation conditions. In the application to real data, both methods show relatively similar in-sample performance, but Truncated Spline performs better in terms of estimation stability and residual autocorrelation control. Therefore, it can be concluded that the Truncated Spline method is more appropriate for longitudinal data analysis that emphasizes estimation stability and the quality of residual diagnostics. Based on these findings, future improvements may include the use of higher-order splines, the addition of predictor variables, and the study of interval estimation to produce more comprehensive and in-depth analytical results.

Keywords: *longitudinal data, autocorrelation, Truncated Spline, Kernel Regression, nonparametric regression*