

ABSTRACT

NONPARAMETRIC REGRESSION ANALYSIS USING SMOOTHING SPLINES AND TRUNCATED SPLINES IN MODELING THE AIR QUALITY INDEX IN INDONESIA

By

Nadhia Az Zahra

The Air Quality Index (AQI) is a composite indicator that reflects regional air quality conditions and is influenced by multiple determinants with complex and nonlinear relationships. In such circumstances, parametric regression may be restrictive because it requires a predetermined functional form. This study applies spline based nonparametric regression using smoothing splines and truncated splines to model AQI in Indonesia and to compare the performance of both approaches. AQI is treated as the response variable, while population density, land cover area within and outside forest areas, and the number of motor vehicles are considered as predictor variables. For smoothing splines, the optimal smoothing parameter is selected using Generalized Cross Validation, whereas truncated splines are estimated using Ordinary Least Squares under various knot configurations and selected based on the minimum Generalized Cross Validation value. Model performance is evaluated using Generalized Cross Validation, Mean Squared Error, and Adjusted R squared. The study aims to identify the most appropriate model and to determine key factors influencing AQI variation in Indonesia, thereby providing empirical support for environmental policy making.

Keywords: Air Quality Index (AQI), nonparametric regression, smoothing splines, truncated splines, Generalized Cross Validation (GCV), Mean Squared Error (MSE), Adjusted R².

ABSTRAK

ANALISIS REGRESI NONPARAMETRIK MENGGUNAKAN SMOOTHING SPLINE DAN TRUNCATED SPLINE DALAM MEMODELKAN INDEKS KUALITAS UDARA DI INDONESIA

Oleh

Nadhia Az Zahra

Indeks Kualitas Udara (IKU) merupakan indikator komposit yang mencerminkan kondisi kualitas udara suatu wilayah dan dipengaruhi oleh berbagai determinan dengan hubungan yang kompleks dan nonlinier. Dalam kondisi seperti ini, regresi parametrik dapat bersifat terbatas karena memerlukan bentuk fungsi yang telah ditentukan sebelumnya. Penelitian ini menerapkan regresi nonparametrik berbasis spline menggunakan smoothing splines dan truncated splines untuk memodelkan AQI di Indonesia serta membandingkan kinerja kedua pendekatan tersebut. AQI diperlakukan sebagai variabel respon, sedangkan kepadatan penduduk, luas tutupan lahan di dalam dan di luar kawasan hutan, serta jumlah kendaraan bermotor dipertimbangkan sebagai variabel prediktor. Pada smoothing splines, parameter penghalus optimal dipilih menggunakan Generalized Cross Validation, sedangkan truncated splines diestimasi menggunakan Ordinary Least Squares dengan berbagai konfigurasi titik knot dan dipilih berdasarkan nilai Generalized Cross Validation minimum. Kinerja model dievaluasi menggunakan Generalized Cross Validation, Mean Squared Error, dan Adjusted R squared. Penelitian ini bertujuan untuk mengidentifikasi model yang paling sesuai serta menentukan faktor faktor utama yang memengaruhi variasi AQI di Indonesia, sehingga dapat memberikan dukungan empiris bagi perumusan kebijakan lingkungan.

Kata-kata kunci: Indeks Kualitas Udara (IKU), regresi nonparametrik, smoothing splines, truncated splines, Generalized Cross Validation (GCV), Mean Squared Error (MSE), koefisien Determinasi.