

ABSTRAK

Bener Meriah merupakan sentra nasional kopi arabika yang produktivitasnya dipengaruhi oleh variabilitas iklim. Penelitian ini bertujuan memprediksi produktivitas kopi bulanan menggunakan data iklim *Google Earth Engine*, menguji pengaruh rekayasa fitur *lag* 6-9 bulan dan fitur historis, membandingkan *Random Forest* dan *XGBoost*, serta menerapkan sistem *constraint* untuk menjaga interpretabilitas. Data produksi 219 desa didistribusikan ke bulanan menggunakan pembobotan hasil panen aktual, sementara data iklim harian diagregasi dengan jendela *lag* 6-9 bulan. Dua himpunan fitur diuji: *8_raw_climate_lag6to9* dan *9_with_prev_yield_lag6to9*. Pemodelan menggunakan *TimeSeriesSplit* dan *RandomizedSearchCV* dengan *constraint* dominasi fitur $\leq 50\%$ dan MAPE validasi $< 40\%$. Hasil menunjukkan model tanpa fitur historis hanya mencapai R^2 validasi 0,43–0,44 dengan MAPE $> 43\%$, mengindikasikan keterbatasan data iklim homogen secara spasial. Penambahan *prev_yield_kg_ha_month* meningkatkan akurasi signifikan; model terbaik *XGBoost* mencapai R^2 0,9870, RMSE 5,29 kg/ha, MAE 2,88 kg/ha, dan MAPE 5,71% serta memenuhi seluruh *constraint* dengan dominasi fitur 47,1%. Radiasi matahari dan selisih suhu harian menjadi variabel iklim paling dominan, sementara curah hujan berpengaruh rendah, sesuai karakteristik wilayah tropis basah. Perbandingan dengan penelitian terdahulu mengonfirmasi keunggulan *XGBoost* dalam menangani data non-linear dan peran krusial fitur historis. Disimpulkan bahwa *XGBoost* dengan fitur historis dan sistem *constraint* merupakan model paling optimal, dengan model *persistence* sebagai fondasi utama, dan direkomendasikan penambahan data iklim spasial serta variabel agronomis spesifik desa untuk prediksi jangka panjang.

Kata kunci: *Prediksi produksi kopi, Google Earth Engine, XGBoost, Random Forest, Kabupaten Bener Meriah.*

