

ABSTRAK

Penelitian ini memetakan zona agroekologi di Kabupaten Bangkalan berdasarkan karakteristik cuaca dan kesuburan tanah dengan metodologi CRISP-DM. Analisis dilakukan pada 269 desa dari total 273 wilayah administratif yang tercatat oleh Badan Pusat Statistik (BPS) terdiri dari lebih dari 250.000+ titik lahan. Proses praproses data menangani kendala kekosongan informasi pada 4 desa tanpa melakukan penghapusan baris, melainkan menerapkan teknik imputasi spasial. Tujuan utama dari penelitian ini adalah mengelompokkan wilayah menggunakan algoritma K-Means yang didasarkan pada dua kelompok variabel utama: iklim (suhu, curah hujan, kelembapan, kecepatan angin, durasi penyinaran matahari) dan tanah (pH, N, P, K, jenis tanah, elevasi, kemiringan lereng). Dataset diekstraksi secara langsung melalui platform *Google Earth Engine* (GEE) pada periode pengamatan 2025. Untuk menyeimbangkan rentang nilai antarvariabel, diterapkan metode *Z-Score Normalization*, sehingga algoritma tidak didominasi oleh atribut yang berskala besar. Hasil komputasi menunjukkan bahwa pembagian wilayah menjadi empat klaster ($K=4$) merupakan konfigurasi yang paling optimal. Keputusan ini didukung oleh evaluasi matematis yang menunjukkan konvergensi pada *Elbow Method*, capaian *Silhouette Score* sebesar 0,1555, serta *Davies-Bouldin Index* (DBI) pada angka 1,6897. Keempat zona yang terbentuk memiliki profil agronomis yang sangat berbeda. Klaster pertama (lahan lembap) mencakup 6.016 titik di 257 desa yang ditandai dengan tingginya intensitas curah hujan dan kelembapan. Klaster kedua (lahan kering produktif) merupakan wilayah terkecil dengan 1.596 titik di 164 desa; meskipun kelembapannya rendah, zona ini memiliki kualitas hara (N, P, K) tertinggi sehingga potensi panennya tetap unggul. Klaster ketiga (lahan kering marginal) meliputi 5.332 titik pada 254 desa yang menghadapi tantangan defisit unsur hara dan curah hujan minim. Terakhir, klaster keempat (lahan toleran kering) menjadi kelompok terbesar dengan 6.480 titik di 59 desa; wilayah ini merepresentasikan kondisi lingkungan paling ekstrem dengan suhu sangat panas dan curah hujan terendah. Hasil akhir dari penelitian ini berupa web.

Kata kunci: *K-Means*, *Zona Agroekologi*, *CRISP-DM*, Imputasi Spasial Bangkalan.

ABSTRACT

This study maps agroecological zones in Bangkalan Regency based on weather characteristics and soil fertility using the CRISP-DM methodology. The analysis was conducted on 269 villages out of a total of 273 administrative areas recorded by Indonesia's Central Statistics Agency (BPS), covering more than 250,000 land points. The data preprocessing stage addressed missing information in four villages without deleting any records, instead applying a spatial imputation technique. The primary objective of this study was to classify regions using the K-Means algorithm based on two main groups of variables: climate (temperature, rainfall, humidity, wind speed, and sunshine duration) and soil characteristics (pH, N, P, K, soil type, elevation, and slope). The dataset was directly extracted through the Google Earth Engine (GEE) platform during the 2025 observation period. To balance the value ranges among variables, Z-Score Normalization was applied, preventing the algorithm from being dominated by large-scale attributes. The computational results indicate that dividing the study area into four clusters ($K = 4$) provides the most optimal configuration. This decision is supported by mathematical evaluations showing convergence in the Elbow Method, a Silhouette Score of 0.1555, and a Davies-Bouldin Index (DBI) of 1.6897. The four resulting zones exhibit highly distinct agronomic profiles. The first cluster (humid land) consists of 6,016 points across 257 villages and is characterized by high rainfall intensity and humidity levels. The second cluster (productive dry land) represents the smallest zone, with 1,596 points across 164 villages; despite its low humidity, this zone possesses the highest nutrient content (N, P, and K), resulting in superior agricultural productivity potential. The third cluster (marginal dry land) includes 5,332 points across 254 villages and faces challenges related to nutrient deficiencies and limited rainfall. Finally, the fourth cluster (drought-tolerant land) is the largest group, comprising 6,480 points across 59 villages; this zone represents the most extreme environmental conditions, characterized by very high temperatures and the lowest rainfall levels. The final result of this research is a website.

Keywords: K-Means, Agroecological Zones, CRISP-DM, Spatial Imputation Bangkalan.