ISSN 2090-3359 (Print)
ISSN 2090-3367 (Online)

ΑΔΣ

Advances in Decision Sciences

Volume 28
Issue 2
June 2024

亞洲大學
ASIA UNIVERSITY

SCIENTIFIC & BUSINESS WORLD

Published by Asia University, Taiwan and Scientific and Business World

Soil Particle Prediction using Spatial Ordinary Logistic Regression

Henny Pramoedyo
Department of Statistics,
Faculty of Mathematics and Natural Sciences,
Brawijaya University, Malang, Indonesia
***Corresponding author** Email: hennyp@ub.ac.id

Atiek Iriany
Department of Statistics,
Faculty of Mathematics and Natural Sciences,
Brawijaya University, Malang, Indonesia
Email: atiekiriany@ub.ac.id

Wigbertus Ngabu
Statistics Studies program,
Faculty of Mathematics and Natural Sciences,
University of San Pedro, Kupang, Indonesia
Email: bertongabu@gmail.com

Sativandi Riza
Department of Soil,
Faculty of Agriculture,
Brawijaya University, Malang, Indonesia
Email: sativandiriza@ub.ac.id

Received: February 2, 2024; First Revision: February 24, 2024;
Last Revision: June 6, 2024; Accepted: June 10, 2024;
Published: June 16, 2025

Abstract

Purpose: This study embarks on a novel journey to create a predictive model for soil texture. We utilize the Spatial Ordinary Logistic Regression model to estimate soil particles in the topsoil with unprecedented accuracy. This involves employing Geographically Weighted Ordinary Logistic Regression to analyze and map the spatial distribution of these particles based on primary data collected from the field.

Design/methodology/approach: This study distinguishes itself by adopting a meticulous approach to gathering soil particulate and geospatial data from various random locations. This method is crucial in addressing the complexity of modelling soil texture, an essential aspect of soil management. The study analyzes soil texture, a combination of sand, silt, and clay, using Digital Elevation Model (DEM) data. By leveraging topographical variations, the study predicts soil texture, employing Geographically Weighted Ordinary Logistic Regression for areas without direct observations. This approach significantly enhances both understanding and prediction in soil science.

Findings: The proposed model will be cross-validated to ensure precision. Aimed at aiding land and resource management, this study focuses on examining spatial variations in topsoil particle sizes and their influencing factors. The Geographic Weighted Ordinary Logistic Regression (GWOLR) model, designed for estimating soil particle sizes using a fixed bi-square weight, demonstrated superior effectiveness with a 90% accuracy rate compared to the standard model's 88%. Further findings show that all topographical predictors exhibit significant spatial autocorrelation (Moran's I, $p < 0.05$), justifying the spatial approach. The GWOLR model also provides localized parameter estimates, revealing spatial heterogeneity in the influence of terrain features. Specifically, vertical curvature and slope positively associate with sandy textures, while lower northness aspects correlate with higher clay and silt presence. Spatial prediction maps generated from the model align closely with actual field data, affirming its practical value in precision agriculture, land-use planning, and targeted soil conservation strategies.

Practical Implications: In 2023, soil particle size data were gathered from the Kalikonto Watershed Area in Batu City, East Java, Indonesia. This data, divided into three categories, was analyzed using the Geographically Weighted Ordinal Logistic Regression method, incorporating spatial factors.

Originality/value: This study presents innovative methods for enhanced spatial analysis, notably the Geographically Weighted Ordinary Logistic Regression technique. This approach improves spatial and statistical data integration for analyzing geographic information, offering insights into how spatial variables influence soil properties. Focusing on estimating particle-size fractions in the soil's top layer, the research underscores the significance of soil attributes on plant growth and agricultural productivity. Furthermore, it provides new perspectives in the crucial field of soil property investigation.

Keywords: Soil, Spatial, Logistic Regression

JEL classification : C21, C31, R32, R12, Q15, Q24, R14

1. Introduction

The advancement of spatial modeling for soil properties has seen significant progress in recent years, driven by the growing need for detailed soil spatial data to enhance precision farming activities (Gao, 2021). Precision farming refers to the application of modern technologies to boost crop production and profitability while minimizing the use of conventional resources such as soil, water, fertilizers, herbicides, and pesticides. Essentially, precision agriculture enables farmers to achieve greater efficiency with fewer resources. In this context, modeling serves as a critical source of quantitative data essential for making informed land management decisions (Plant, 2018).

Accurate modeling of soil texture is crucial for guiding soil management strategies. Its importance is profound, as soil texture, defined by its particle size distribution, significantly affects the movement of water, heat, and nutrients, as well as the soil's capacity to retain water and nutrients and the development and stability of soil structure (Ding & Huang, 2017). However, modeling soil texture is challenging due to its compositional nature, which includes sand, silt, and clay proportions in the soil's mineral component (Peruzzetto, et al., 2021). A significant complication arises from the inherent constraint in compositional data that the sum of sand, silt, and clay percentages must always equal 100%.

The study of particle-size fractions is crucial for understanding soil characteristics and optimizing conditions for agricultural productivity. Particle-size fractionation involves classifying soil particles based on their sizes, a key parameter for assessing soil's physical properties and its suitability for various types of plant growth. Accurate prediction of particle-size fractions in surface soils is essential for farmers and agricultural scientists to develop effective soil management strategies that ensure optimal conditions for plant growth (Koiter et al., 2017). Traditional methods, such as Ordinary Logistic Regression, have been employed to predict these fractions. However, these conventional approaches fail to account for the spatial variabilities that significantly influence soil properties across different geographic locations. The inability of these methods to provide detailed, location-specific models highlights a significant limitation, underscoring the need for a more advanced analytical approach incorporating spatial factors (Minár, et al., 2020).

2. Literature Review

The complexity of soil data, characterized by nonlinear relationships among variables, presents substantial challenges for conventional linear models. These models often struggle to capture the intricate dynamics within the data, where changes in one variable do not result in linear effects on another. This limitation is particularly evident in the context of geographic data, where spatial interactions further complicate the distribution patterns of variables. Traditional linear models fail to account for the influence of spatial relationships, leading to a lack of precision in analyses that require an understanding of spatial dynamics. These models do not adequately address the subtle interactions between variables influenced by geographic proximity, resulting in studies that overlook the spatial heterogeneity inherent in geographic data. While the works of (Leightner & Inoue, 2012) offer alternative and advanced methodologies, focusing on the development and application of statistical and mathematical techniques that enhance data analysis across various contexts, they primarily provide insights into general approaches, such as nonlinear modeling, more sophisticated estimation techniques, or addressing omitted variables (Ruymgaart, et al., 2011). This suggests that while these approaches can expand the analytical framework for complex data, there remains a clear need for techniques specifically designed to handle the unique aspects of spatial data.

Therefore, the proposed model is not just a solution to a problem but a significant advancement in soil science. To overcome these limitations, the Geographically Weighted Ordinary Logistic Regression method emerges as a superior alternative for spatial data analysis. This advanced technique integrates spatial and statistical information, offering a nuanced approach to understanding geographic data (Tian, et al., 2023). Unlike its predecessors, Geographically Weighted Ordinary Logistic Regression is adept at handling the unique challenges posed by geographic data, including the analysis of spatial interactions. Accommodating local variations in the relationships between variables reveals spatial dynamics that remain obscured by linear models (Zhang & Yang, 2020). This method's ability to directly incorporate geographic influence into the model enhances analyses' accuracy and spatial relevance. In soil science, it provides a mechanism to identify and understand the geographic factors influencing soil properties across different locations, offering insights critical for targeted soil management and agricultural planning. (Pramoedyo, et al., 2023).

This constraint is further supported by the study conducted by (Rodrigues, et al., 2018), which presented an in-depth spatial-temporal examination of the factors influencing human-induced wildfires in Spain through the use of Geographically Weighted Regression. This highlighted the significant capability of Geographically Weighted Regression for analyzing spatial data (Rodrigues, et al., 2018). Nevertheless, the literature has paid scant attention to this issue, as evidenced by the work of (Windle, et al., 2010), who investigated the spatial variability of fisheries survey data utilizing Geographically Weighted Regression and provided a case study from the Northwest Atlantic (Windle, et al., 2010).

The refined comprehension of spatial relationships and the utilization of Geographically Weighted Ordinary Logistic Regression (GWOLR) in forecasting the distribution of particle size within soil highlight the significance of this methodology in the evolution of soil science. GWOLR paves the way for tailored soil management approaches specific to distinct geographical areas by offering an intricate understanding of the spatial determinants influencing soil attributes. This methodology enriches the domain of agricultural science and can enhance agricultural yields significantly by customizing soil management practices to align with the unique spatial characteristics of each location. Engaging in spatial analysis through GWOLR, a technique that investigates the interplay between spatial variables and observed phenomena across various locales, allows for a sophisticated blend of logistic regression and spatial analytical methods. This integration facilitates a more precise and effective amalgamation of statistical data and spatial insights in analyzing geographic information (Huang, et al., 2023) and is further elaborated by (Comber, et al., 2020). This advancement in combining logistic regression with spatial analysis techniques marks a pivotal development in harnessing geographic data to enhance the accuracy of soil science research and its applications.

In the Geographically Weighted Ordinary Logistic Regression method, logistic regression establishes the link between independent variables (such as soil water content, elevation, and soil classification) and the observed occurrences. Subsequently, spatial analytical methods are employed to integrate the influence of geographical factors into the regression analysis, thereby enhancing the precision and granularity of the outcomes (Hong, et al., 2017; Ngabu, et al., 2023). Several benefits distinguish the Geographically Weighted Ordinary Logistic Regression technique from other spatial analysis methods. One notable advantage is its capacity to precisely account for the impact of geographical elements on the regression analysis at various locations, leading to outcomes that are both more precise and richly detailed (Cao, et al., 2019). Furthermore, the Geographically Weighted Ordinary Logistic Regression approach is versatile and capable of exploring the connections between spatial factors and events or phenomena that traditional regression analyses fail to elucidate.

This study introduces several innovations, specifically employing a spatial analysis method that enhances accuracy and effectiveness: The Geographically Weighted Ordinary Logistic Regression approach enables more precise and efficient spatial and statistical data integration in geographic analysis (Comber, et al., 2023). This advancement facilitates a deeper insight into spatial variables influencing soil attribute variability across different locales. The research prioritizes the forecasting of soil particle sizes, with a particular focus a1 on the prediction of particle-size fractions in surface soil. Such focus is vital as soil properties, including nutrient composition and structure, significantly impact plant development and agricultural yield. This study contributes novel insights into the ongoing research on soil characteristics, underscoring its perennial significance.

Previous research has highlighted limitations in employing Geographically Weighted Regression (GWR) techniques for analyzing soil particles, where each variable (Sand, Silt, Clay) requires individual analysis, thus impeding a unified predictive capability (Pramoedyo, et al., 2021). This study aims to advance soil management practices by leveraging Geographically Weighted Ordinary Logistic Regression analysis. The goal is to furnish precise soil management recommendations to enhance agricultural outputs and farmer livelihoods by predicting soil particle sizes within a designated area (Rodríguez-Lado & Lado, 2017). The intended outcomes include a comprehensive understanding of spatial determinants affecting soil particle-size variation and providing targeted, effective soil management strategies for agricultural professionals.

The primary objective of this research is to develop a predictive model for assessing soil texture through the spatial application of the Ordinary Logistic Regression technique on soil particle sizes. By integrating direct data from field observations, the study aims to map the distribution of soil particle sizes accurately and spatially. It also seeks to assess the model's accuracy and reliability through cross-validation techniques, investigating the elements influencing spatial variation in soil particle sizes. The expected findings include detailed insights into the spatial distribution patterns of soil particle sizes and their influencing factors. Moreover, applying the Spatial Logistic Regression method for spatial analysis is anticipated to yield a spatial prediction model with improved precision, thus providing critical information for land and natural resource management policymakers.

3. Materials and methods

3.1 Data

This study used primary data, encompassing measurements of soil texture conducted directly in the field, and data derived from the analysis of Digital Elevation Models (DEMs). The field data gathered from 50 distinct observation points was pivotal in constructing a model. This included data on soil texture analysis from the Kalikonto watershed area in Batu City, within the East Java Province of Indonesia. The study incorporated eight local morphological variables (LMVs) that illustrate the topographical curvature (Peruzzetto, et al., 2021). These LMVs comprise:

1. Vertical Curvature (K_v)

$$K_v = \frac{p^2 r + 2pqs + q^2 t}{(p^2 + q^2)\sqrt{(1 + p^2 + q^2)^3}};$$

2. Horizontal Curvature (K_h)

$$K_h = \frac{q^2 r - 2pqs + p^2 t}{(p^2 + q^2)\sqrt{1 + p^2 + q^2}};$$

3. Accumulation Curvature (K_a)

$$K_a = \frac{(q^2 r - 2pqs + p^2 t)(p^2 r + 2pqs + q^2 t)}{[(p^2 + q^2)(1 + p^2 + q^2)]^2};$$

4. Ring Curvature (K_r)

$$K_r = \left[\frac{(p^2 - q^2)s - pq(r - t)}{(p^2 + q^2)(1 + p^2 + q^2)}\right]^2;$$

5. Northness Aspects (A_n)

$$A_n = cos\left[-90[1 - sin(q)](1 - |sin(p)|) + 180[1 + sin(p)] - \frac{180}{\pi}\, sin(p) arcos\left(\frac{-q}{\sqrt{p^2 + q^2}}\right)\right];$$

6. Eastness Aspects (A_e)

$$A_e = sin\left[-90[1 - sin(q)](1 - |sin(p)|) + 180[1 + sin(p)] - \frac{180}{\pi}\, sin(p) arcos\left(\frac{-q}{\sqrt{p^2 + q^2}}\right)\right];$$

7. Slope (S)

$$S = arctan\sqrt{p^2 - q^2}.$$

8. Elevation (E_{lev}) (1)

In this study, the variable of interest, or dependent variable, is analyzed to determine the impact of independent variables on it. It is the aspect being measured throughout the research and is influenced by the experimental conditions, encompassing eight dependent variables (X). Specifically, within the

Kalikonto River Basin in Batu City, within the East Java Province of Indonesia, these dependent variables are categorized into Silt, Sand, and Clay. The following sections provide a more comprehensive breakdown of these three dependent variables:

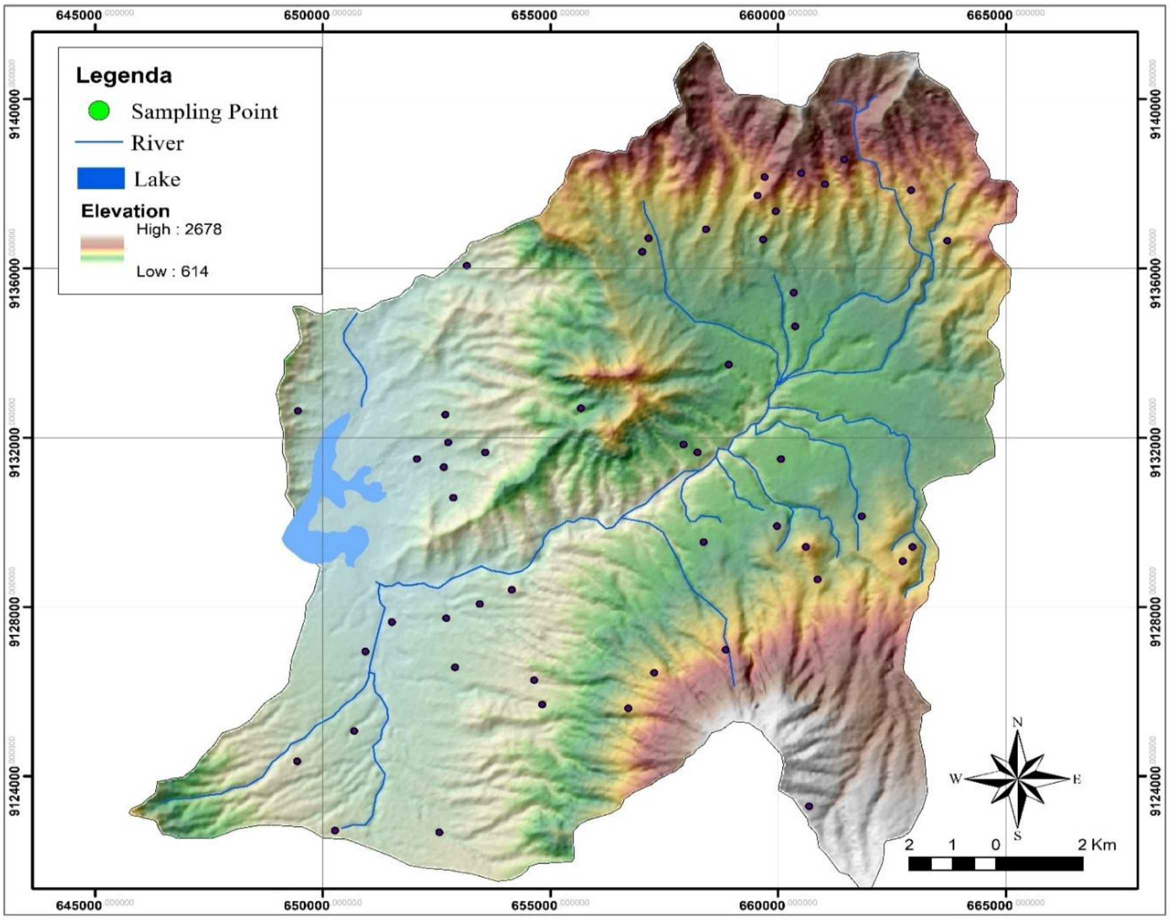
1. **Silt:** Silt is soil particles that are sized between sand and clay (Riza, et al., 2021). Silt particles are relatively fine and can provide a smooth texture to the soil. Silt plays an important role in determining the physical and chemical characteristics of the soil because it has an excellent ability to store water and nutrients (M. Tahat, et al., 2020). In the study, measuring silt can provide information about the soil's capacity to retain water and the availability of nutrients for plants.

2. **Sand:** Sand is soil particles larger than silt and clay. Sand has high porosity, meaning it has large pore spaces, allowing water and air to move more freely through the soil. Soil rich in sand usually has good drainage but poorly retains nutrients (Hou, et al., 2020). In the study, analyzing the sand content can indicate how well the soil performs in terms of drainage and how quickly nutrients can be lost from the soil.

3. **Clay:** Clay is tiny soil particles with a high capacity to store water and nutrients due to its extensive surface area. Clay can make the soil more compact and less porous, resulting in poorer drainage than sand-rich soil (Abidin, et al., 2017). However, its ability to retain water and nutrients makes clay an essential component of soil for plant growth. In the study context, measuring clay content can provide insights into the soil's ability to retain water and nutrients and its implications for agriculture or land management (Liu, et al., 2018).

Data collection at 50 points in the Kalikonto River Basin aims to identify variations in the content of silt, sand, and clay across different locations and how these variations can be influenced by eight independent variables not specifically mentioned in your question. Analyzing the relationship between independent variables and these three types of soil content can help in understanding the soil characteristics in the Kalikonto River Basin and their impact on the management of natural resources and agricultural activities in the area. The data collected in 2023 will provide a current snapshot of the soil conditions at the research location. Figure 1 displays the geographic layout of the study area.

Figure 1. Map of Research Locations



Note: sample location point (watershed area of Kalikonto, Batu City)

3.2 Ordinary Logistic Regression

Ordinal logistic regression expands upon binary logistic regression to accommodate response variables on an ordinal scale with three or more categories, using either interval or ratio scales for covariates or nominal or ordinal scales for factors as predictors (Kamberaj, 2021). This method analytically determines the relationship between ordinal dependent variables spanning at least three categories and independent predictor variables, which can be either continuous or categorical, involving at least two variables (Albert & Anderson, 1984). An ordinal logistic regression model, also known as a cumulative logit model, characterizes the ordinal response variable (Y) through cumulative probabilities, effectively expressing the probability $P(Y = 1|x)$ as $\pi(x)$ as a function of (x), as outlined by (Fagerland & Hosmer, 2016):

$$\pi(x) = \frac{\exp(\propto_g + X_i^T \boldsymbol{\beta})}{1 - exp(\propto_g + X_i^T \boldsymbol{\beta})} - \frac{exp(\propto_{g-1} + \boldsymbol{X}_i^T \boldsymbol{\beta})}{1 - exp(\propto_{g-1} + \boldsymbol{X}_i^T \boldsymbol{\beta})}. \quad (2)$$

The logistic regression model falls under the umbrella of generalized linear models, employing the cumulative logit model specifically for ordinal logistic regression purposes.

3.3 Spatial Ordinary Logistic Regression

The Spatial Ordinal Logistic Regression model merges Geographically Weighted Regression and ordinal logistic regression techniques to explore the interaction between ordinal outcomes and explanatory variables across varied geographical settings. This advanced model, known as the Geographically Weighted Ordinal Logistic Regression, offers a nuanced version of logistic regression that considers spatial differences, suggesting that the influence of explanatory variables is not uniform but varies by location. It achieves this by applying a weighting mechanism that allocates specific weights to each observation, thereby incorporating geographical context into the analysis. For example, when categorizing the response variable into G distinct categories, the model formulation for a specific location I, as outlined by (Bertsimas & King, 2017), showcases this approach.

$$\text{logit}[P(Y_i \leq g \leq \mathrm{x}_i] = ln\left[\frac{P\ (Y_i \leq \mathrm{g} \mid \mathrm{x}_i)}{1-P\ (Y_i \leq \mathrm{g} \mid \mathrm{x}_i)}\right] = \alpha_g(u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i), \qquad (3)$$

where g is a category ranging from 1 to G-1. There is an intercept parameter $\alpha_g\ (ui\ , vi\ )$ that varies by location indexed by i with coordinates (u_i, v_i). This parameter is ordered such that $\alpha_1 \leq\ \alpha_2 \leq \cdots\ \leq \alpha_{G-1}$ for any given location. Additionally, there is a vector of regression coefficients $\beta\ (ui\ , vi\ )$ that also varies by location. This vector is denoted as $\beta(u_i, v_i) = \left[\beta_1(u_i, v_i) \quad \beta_2(u_i, v_i)\ \ldots\ \beta_j(u_i, v_i)\right]^T$ where $\boldsymbol{T}$ denotes the transpose of the vector, indicating that the coefficients are arranged in a column vector. The coordinates (u_i, v_i). represent the longitude and latitude of the i^{th} location, respectively.

The aggregate likelihood for the response to fall within the g-th category can be articulated as follows (Mishra, et al., 2021) :

$$P(Y_i \leq g|X_i) = \frac{\exp\left(\propto_g\ (u_i, v_i)\ +\ \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}{1-\ exp\left(\propto_g\ (u_i, v_i)\ +\ \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}, \mathrm{g} =\ 1,2, \ldots \mathrm{G}-1. \qquad (4)$$

If $\pi_g^*(x_i) = P(Y_i \leq g|x_i)$ equals the probability of the response variable at location i falling within or below the j^{th} category given x_i (Nkeki & Asikhia, 2019):

$$\pi_g^*(x_i) = \frac{\exp\left(\propto_g\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}{1-\ exp\left(\propto_g\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)} - \frac{exp\left(\propto_{g-1}\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}{1-\ exp\left(\propto_{g-1}\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}; \qquad (5)$$

$$\text{by: } \frac{\exp\left(\propto_g\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}{1-\ exp\left(\propto_g\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)} = 0;$$

$$\frac{exp\left(\propto_{g-1}\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)}{1-\ exp\left(\propto_{g-1}\ (u_i, v_i) + \boldsymbol{X}_i^T\boldsymbol{\beta}(u_i, v_i)\right)} = 1.$$

In the context of the Geographically Weighted Ordinal Logistic Regression model, the estimation of parameters $\theta(u_i, v_i)$ which is an array comprising $[\alpha_1(u_i, v_i)\ \ \alpha_2(u_i, v_i)$,..., $\alpha_{G-1}(u_i, v_i)\quad \beta\ (u_i, v_i)]^T$, can be determined through the application of the weighted Maximum Likelihood Estimation (MLE) technique (Zuhdi & Saputro, 2017). Consider a scenario where we have a sample of n observations labelled Y_1 , Y_2,... Y_n. Each observation has a probability associated with being in the g-th category, expressed as $\pi_g^*(x_i)$. In this case, Y_i , which can be expanded as (y_{i1}, y_{i2}, ..., y_{iG-1}), follows a multinomial distribution with parameters $\left(1, \pi_1^*(x_i), \pi_2^*(x_i), \ldots, \pi_{G-1}^*(x_i)\right)$. The likelihood function for this distribution is then formulated as follows:

$$l = \prod_{i=1}^{n} \prod_{g-1}^{G} \pi_g^* \, (x_i)^{y_{ig}}$$

$$= \prod_{i=1}^{n} \prod_{g-1}^{G} \left[\frac{\exp\left(\propto_g \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_g \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} - \frac{exp\left(\propto_{g-1} \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_{g-1} \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} \right]^{y_{ig}} . \tag{6}$$

The subsequent phase involves developing the likelihood function's natural logarithm (ln). This is achieved by applying the ln transformation to the likelihood function, resulting in an expression that reads as follows:

$$L\left(\theta(u_i, v_i)\right) = \sum_{j=1}^{n} \sum_{g=1}^{G} y_{ig} \, ln \left[\frac{\exp\left(\propto_g (u_i,v_i) + X_i^T \boldsymbol{\beta}(u_i,v_i)\right)}{1 + exp\left(\propto_g (u_i,v_i) + X_i^T \boldsymbol{\beta}(u_i,v_i)\right)} - \frac{exp\left(\propto_{g-1} (u_i,v_i) + X_i^T \boldsymbol{\beta}(u_i,v_i)\right)}{1 + exp\left(\propto_{g-1} (u_i,v_i) + X_i^T \boldsymbol{\beta}(u_i,v_i)\right)} \right]. \tag{7}$$

hus, the local Spatial Ordinal Logistic Regression model incorporates a specific weight for the log-likelihood function. Let's consider that for each point $(u_i \, , v_i)$, the associated weight is denoted by $\boldsymbol{w}(u_i, v_i)$, for *i* ranging from 1 to *n*. Based on this, we can express the weighted log-likelihood function, as defined by (Pramoedyo, et al., 2024) study:

$$L^* = \sum_{j=1}^{n} \sum_{g=1}^{G} \left\{ y_{ig} \ln \left[\frac{\exp\left(\propto_g \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_g \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} - \frac{exp\left(\propto_{g-1} \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_{g-1} \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} \right] \right\} \boldsymbol{w}_j(u_i, v_i). \tag{8}$$

For instance, when the outcome variable is categorized into four groups (G=3), the constructed likelihood function is subsequently converted into its logarithmic representation:

$$L^* = \sum_{j=1}^{n} \left\{ y_{j1} ln \left[\frac{\exp\left(\propto_1 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_1 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} \right] \right. \tag{9}$$

$$+ y_{j2} ln \left[\frac{\exp\left(\propto_2 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_2 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} - \frac{\exp\left(\propto_1 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_1 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} \right]$$

$$+ y_{j3} ln \left[1 - \frac{\exp\left(\propto_2 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)}{1 + \, exp\left(\propto_2 \, (u_i, v_i) + \boldsymbol{X}_i^T \boldsymbol{\beta}(u_i, v_i)\right)} \right] \boldsymbol{W}_j(u_i, v_i).$$

Formula (9) can be condensed into

$$L^{*}=\sum_{j=1}^{n}\Big\{y_{j1}\left(\alpha_{1}(u_i,v_i)+\boldsymbol{X}_j^{T}\boldsymbol{\beta}(u_i,v_i)\right)-\left(y_{j1}+y_{j2}\right)\ln\left[1+\exp\left(\propto_{1}(u_i,v_i)+\boldsymbol{X}_i^{T}\boldsymbol{\beta}(u_i,v_i)\right)\right]$$

$$+y_{j2}\ln\left[\exp\left(\propto_{2}(u_i,v_i)+\boldsymbol{X}_i^{T}\boldsymbol{\beta}(u_i,v_i)\right)-\exp\left(\propto_{1}(u_i,v_i)+\boldsymbol{X}_i^{T}\boldsymbol{\beta}(u_i,v_i)\right)\right]$$

$$+\left(y_{j1}-1\right)\ln\left(\alpha_{1}(u_i,v_i)+\boldsymbol{X}_j^{T}\boldsymbol{\beta}(u_i,v_i)\right)\Big\}\boldsymbol{w}_j(u_i,v_i). \tag{10}$$

In the Geographically Weighted Ordinal Logistic Regression model, the factor of geographical location, represented as $\boldsymbol{w_i}(u_i,v_i)$., serves as a weighting element. This factor varies across locations, highlighting the localized characteristic of the Geographically Weighted Ordinal Logistic Regression model.

3.4 Weights

Choosing the correct weighting function is critical for the spatial analysis results, as it factors in the distance between observed locations by using continuous values to form the weighting matrix. This process assigns a weight to each location based on its proximity to the observed location. In our research, we utilized Fixed Bisquare Kernel weights based on the idea that each location has an ideal Bandwidth. Identifying the best bandwidth is crucial because it determines the radius of influence around each observation. A proven approach to finding the best bandwidth is choosing one that minimizes the AIC value (Fotheringham, et al., 2017). The bi-square weighting function was selected due to its incorporation of distance elements between observed locations in a continuous manner when creating a weighting matrix, ensuring that each location is weighted according to its distance from the observed location (Du, et al., 2020).

$$d_{ij}=\sqrt{(u_i-u_j)^2+(v_i-v_j)^2};$$

$$\boldsymbol{W}_{ij}(u_i,v_i)=\begin{cases}[1-(\frac{d_{ij}}{h})^2]^2 & dij\ \le\ h;\\ 0 & dij\ >\ h.\end{cases} \tag{11}$$

3.5 Spatial Effect Testing

A spatial effect examination was conducted to determine a locational effect within the analyzed model (Chen, 2021). The assessment utilized Moran's I, a measure for evaluating spatial autocorrelation values. It aims to pinpoint spatial clusters or global spatial autocorrelation (Ngabu, et al., 2023). This technique is capable of identifying global spatial randomness, which may reveal clustering patterns or spatial trends (Anselin, 2019). The computation of spatial autocorrelation employs Moran's I equation, incorporating a standardized matrix as the weighting matrix (Anselin, 2020).

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}w^{*}{}_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\Sigma_{i=1}\Sigma_{j=1}w^{*}{}_{ij}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{12}$$

3.6 Best Model Selection

A comparison is made between the outcomes of ordinal logistic regression models and Geographically Weighted Ordinal Logistic Regression to identify the optimal regression model. A key criterion for selecting the superior model is identifying which achieves the highest classification accuracy value (Zhang & Yang, 2020). An effective classification accuracy measure is defined as a percentage value exceeding 50% (Murtagh & Heck, 2012).

4. Results and discussion

4.1 Ordinary Logistic Regression Analysis

Ordinal logistic regression analysis was conducted to examine the factors influencing particle-size fraction. The data on particle-size fractions, categorized on an interval scale and encompassing multiple particle-size categories, facilitated the application of ordinal logistic regression analysis. Parameter estimates from the ordinal logistic regression model are presented in Table 1.

Table 1. Ordinal logistic regression model parameters

Parameter	Estimate	Std. error	wald	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
[Y=1]	0.712	0.621	1.146	0.253	-0.514	2.621
[Y=2]	2.143	0.362	5.919	0.002	0.389	4.521
K_v	18.363	1421.28	0.012	0.847	-3278.84	3981.32
K_h	-196.362	1348.36	0.145	0.782	-3148.41	3214.12
K_a	362.029	284.26	1.273	0.362	-161.322	718.21
K_r	128.324	3425.48	0.037	0.487	-5522.36	7116.36
A_n	-112.679	181.42	-0.621	0.421	-5124.36	321.36
A_e	-0.028	0.524	-0.053	0.562	-0.041	0.0342
S	0.268	0.412	0.650	0.824	-0.836	1.925
E_{lev}	-0.482	0.486	0.991	0.362	-1.458	0.621

Note: Vertical Curvature (K_v), Horizontal Curvature (K_h), Accumulation Curvature (K_a), Ring Curvature (K_r), Northness Aspects (A_n), Eastness Aspects (A_e), Slope (S), Elevation (E_{lev})

Based on the above Table 1, an ordinal logistic regression logit model will be created based on the parameter estimates as follows:

$$\text{Logit}\left[\hat{P}(Y \leq 1|x)\right] = 0{,}712 + 18.363\ K_v - 196.362\ K_h + 362.029\ K_a + 128.324\ K_r - 112.679\ A_n - 0.028\ A_e + 0.268\ S - -0.482\ E_{lev}\ ;$$

$$\text{Logit}\left[\hat{P}(Y \leq 2|x)\right] = 2.143 + 18.363\ K_v - 196.362\ K_h + 362.029\ K_a + 128.324\ K_r - 112.679\ A_n - 0.028\ A_e + 0.268\ S - -0.482\ E_{lev}\ .$$

The negative coefficient of the variable K_h (Horizontal Curvature) suggests that a decrease in the value of K_h (Horizontal Curvature) at a specific location point is associated with a reduction in soil particle size, specifically falling into the clay category. With each 0.01 decrease in the K_h (Horizontal Curvature) variable, there is a likelihood of an increase in the proportion of clay in the soil particle size. Conversely, the positive coefficient of the variable S (Slope) implies that an increase in the value of S (Slope) at a location point is likely to be linked to larger soil particles, predominantly sand.

4.2 Accuracy of Ordinal Logistic Regression Classification

Determining the precision of particle-size fraction categorization outcomes within the Kalikonto watershed region of Batu City in 2023 through an ordinal logistic regression model. The computation of odds for each classification is illustrated in the subsequent table attachment:

Table 2. Precision of the Classification by the Logistic Regression Model

Category	Count of Forecasts for Each Classification			Accuracy
	1	2	3	
1	34	3	0	91.89%
2	2	6	0	75%
3	0	1	4	80%

Note: Category 1 (sand), Category 2 (silt), Category 3 (clay)

The precision of the comprehensive classification forecasts is derived from:

$$= \left(\frac{\text{number of correct predictions}}{\text{N}}\right) \text{ x } 100\% \;;$$

$$= \left(\frac{34+6+4}{50}\right) \text{ x } 100\,\% = 88\%.$$

The accuracy of the Particle-size Fraction classification shown in Table 1 is 88%, indicating that the logistic regression model has a high efficacy in categorization, as it achieves an accuracy significantly above the 50% threshold. This outcome underscores the effectiveness of the ordinal logistic regression method applied to classify particle-size fractions within the Kalikonto watershed in Kota Batu for the year 2023. Achieving an accuracy rate of 88%, the model exceeds the basic accuracy criterion of 50% and exhibits exceptional capability in differentiating among various classifications. Such performance attests to the model's strong and dependable nature for guiding decisions in managing and surveilling the Kalikonto watershed.

However, it is highly recommended that spatial elements be integrated into the model to make more informed and effective decisions. Adding spatial data, such as geographic information and mapping, will enable users to observe the distribution of particle-size fractions not just numerically or categorically but within a spatial geographic context. This addition will aid in visualizing patterns that may remain hidden when viewed solely through tables and figures. Consequently, spatial mapping will open avenues for further analysis, such as identifying critical zones requiring more intensive management or assessing the impact of interventions that have been implemented. Integrating this spatial data will enhance the model’s capacity to support natural resource management policies and environmental conservation efforts in the Kalikonto watershed, Kota Batu.

4.3 Morans I Test

In our statistical exploration, we initially applied Ordinary Logistic Regression to determine the influence of particle size on our study variables. However, we noted instances where this method did not fully capture the variable's impact. Consequently, we turned to Geographically Weighted Ordinary Logistic Regression for a more detailed analysis that accounts for spatial differences. To determine if these differences significantly influenced our results, we performed Moran's I test, focusing on the p-values to judge if spatial factors had a significant effect, with a significance level set at $\alpha < 0.05$. For Moran's I test, we examined the impact of using different types of weights, including simple distance weights and those based on a fixed kernel bisquare approach, with the latter calculated by the inverse of distance $W_{ij} = \frac{1}{d_{ij}}$, where d_{ij} is the distance between two points in space. The kernel's bandwidth was fine-tuned to 7.2154 to reflect spatial interrelations precisely.

Table 3. The Moran's I test evaluates spatial autocorrelation

Variable	Moran’s I Statistic	probability value
K_v	3.2832	0.00348
K_h	3.8441	0.00412
K_a	4.4127	0.01311
K_r	3.8456	0.00162
A_n	3.9142	0.00312
A_e	4.3621	0.00324
S	2.1124	0.00124
E_{lev}	2.9831	0.00261

Note: Vertical Curvature (Kv), Horizontal Curvature (Kh), Accumulation Curvature (Ka), Ring Curvature (Kr), Northness Aspects (An), Eastness Aspects (Ae), Slope (S), Elevation (Elev)

Table 3 displays the outcomes of Moran's I test for spatial autocorrelation across eight variables labeled Vertical Curvature (K_v), Horizontal Curvature (K_h), Accumulation Curvature (K_a), Ring Curvature (K_r), Northness Aspects (A_n), Eastness Aspects (A_e), Slope (S), Elevation (E_{lev}). Moran's I statistic quantifies spatial autocorrelation, revealing the degree of similarity of a variable within its neighboring locations. Values nearing +1 signify strong positive spatial autocorrelation, indicating similarity across adjacent areas. Conversely, values approaching -1 denote negative spatial autocorrelation, reflecting dissimilarity, while values around 0 suggest a lack of spatial pattern, implying randomness.

The Moran's I values observed range from 2.1124 for variable Slope (S) to 4.4127 for Accumulation Curvature (K_a), indicating distinct levels of positive spatial autocorrelation among the variables. With p-values all falling below the 0.05 threshold, the evidence strongly refutes the hypothesis of spatial randomness for each variable. This outcome substantiates the rejection of the null hypothesis, confirming that the spatial distributions of these variables are not random but exhibit significant spatial autocorrelation.

4.4 Geographically Weighted Ordinary Logistic Regression

The outcomes of parameter estimation for the Geographically Weighted Ordinal Logistic Regression generate a localized model specific to each of the 50 spatial points. This means that a unique model characterizes every location. Consequently, this model yields probability forecasts for various categories. To illustrate, the model corresponding to the initial spatial point is situated at the coordinates (u_i, v_i). Thus, the established Geographically Weighted Ordinal Logistic Regression model is presented as follows:

Probability per category.

$$\pi_1(x_1) = \frac{\exp(3{,}41 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})}{1+\exp(3{,}41 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})};$$

$$\pi_2(x_1) = \frac{\exp(4.321 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})}{1+\exp(4.321 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})};$$

$$\frac{\exp(3{,}41 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})}{1+\exp(3{,}41 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})};$$

$$\pi_3(x_1) = 1 - \frac{\exp(4.321 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})}{1+\exp(4.321 + 18{,}42\ K_v + 21{,}442\ K_h + 29{,}142\ K_a + 9{,}428\ K_r - 63.362\ A_n - 0.421\ A_e + 0.321\ \mathrm{S} - 0.421\ E_{lev})}.$$

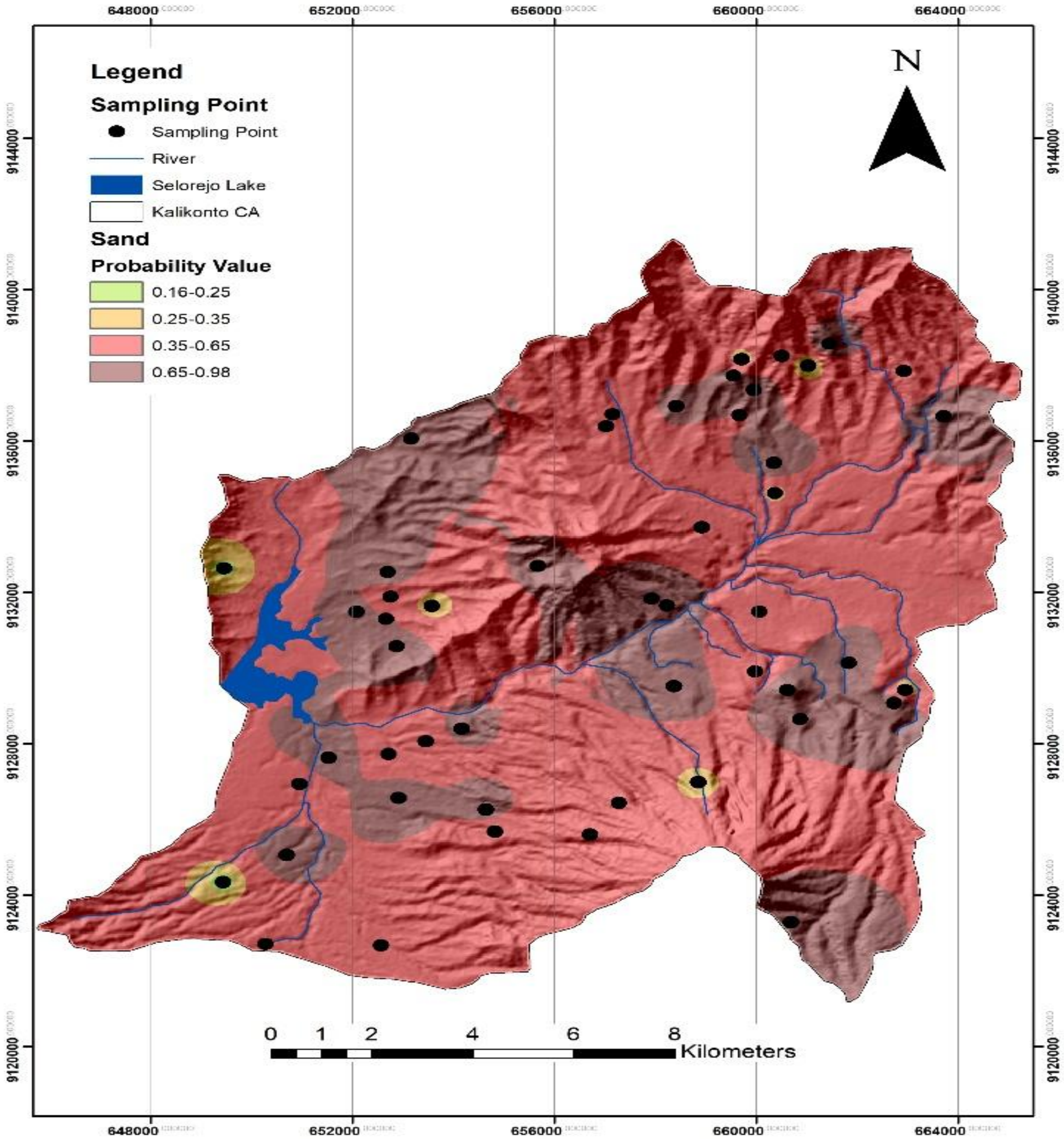
The findings indicate that parameter β_1 are positive suggesting that geographic and topographic features, such as the Vertical Curvature (K_v), significantly influence soil particle size, particularly sand content. This could be due to differential sunlight exposure, affecting soil moisture and, subsequently, weathering processes that determine soil particle size. These findings are supported by the study of (Siqueira, et al., 2023), which explores the use of machine learning for soil mapping in Antarctica, underscoring the importance of spatial analysis in predicting soil texture based on geographic factors.

The negative value of parameter β_5 suggests that a decrease in the Northness Aspects (A_n), which reflects more level topography, is associated with the presence of finer soil particles, specifically silt, and clay. This observation aligns with the research conducted by (Liu, et al., 2023). Their study delves into the dynamics between climatic conditions, topographical variations, and soil characteristics in relation to crop distribution patterns. By employing remote sensing technology and advanced machine learning techniques, they have demonstrated the significant impact of topographical features on the composition of soil and, consequently, on agricultural practices.

Furthermore, the influence of topography on soil composition is deepened (Gxasheka, et al., 2023). Reviewing the role of topographic and soil factors in woody plant encroachment in mountainous grasslands. This study highlights the impact of topography on vegetation distribution and soil composition, providing additional insights into the dynamics between topography and land-based ecosystems.

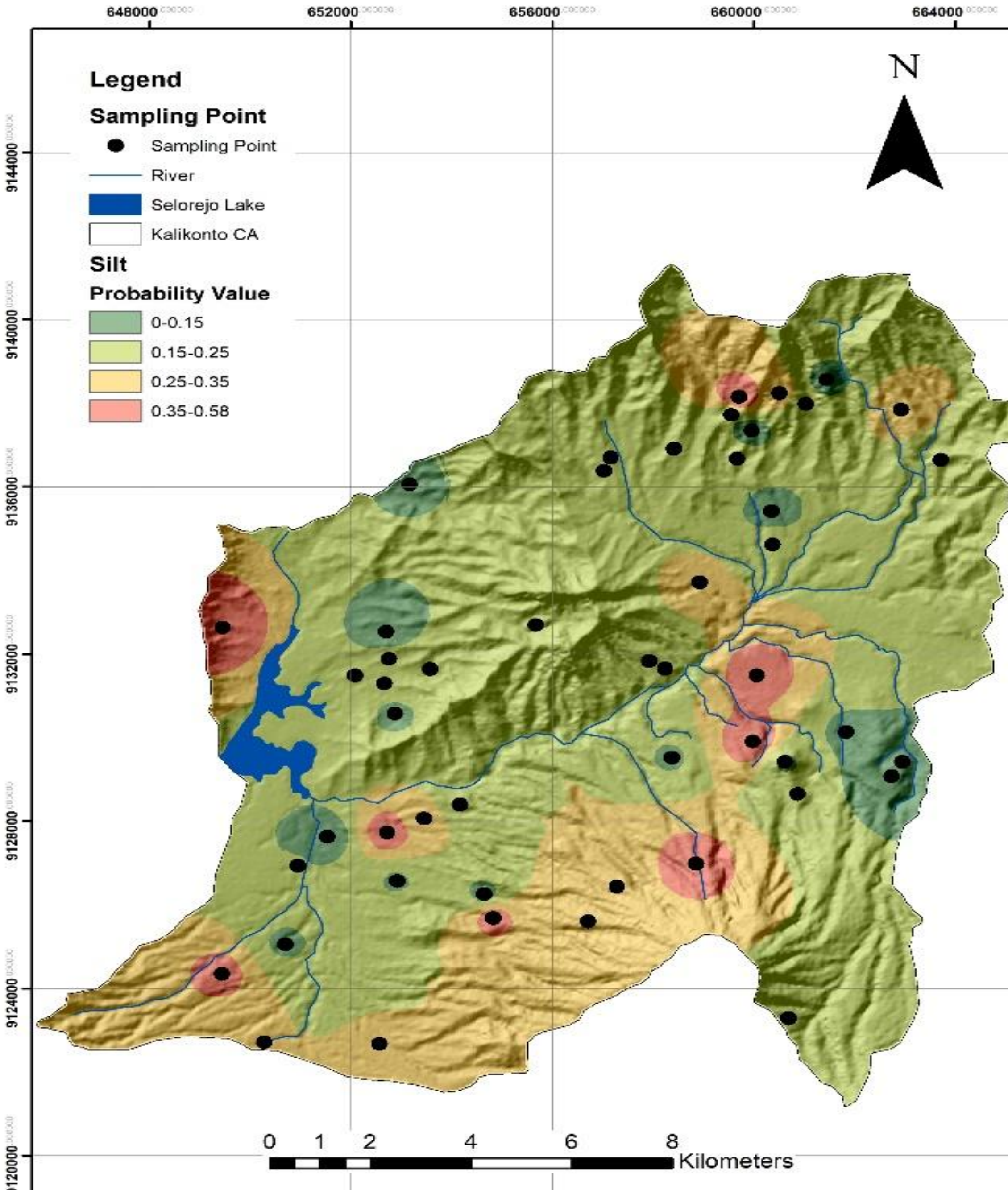
The probability prediction results for each category, namely Sand, Silt, and Clay, are presented in Figure 2. From Figure 2, the forecasted likelihood for each site is derived, and the subsequent chance value is multiplied by 100 to calculate the soil composition value for every site.

Figure 2(a). Actual value mapping and forecasted likelihood



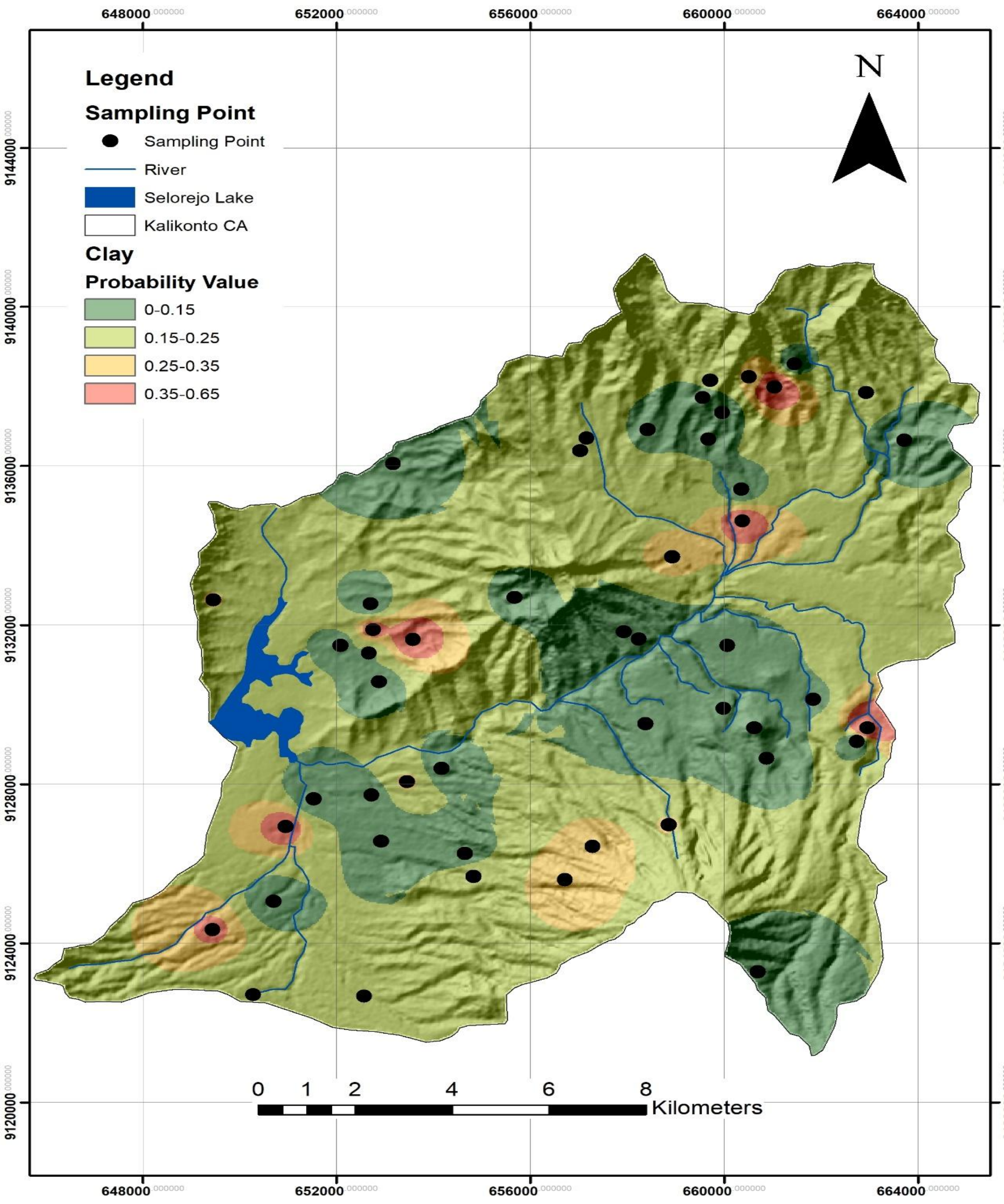
Note: Category 1 Opportunity Prediction, namely the Sand Variable

Figure 2(b). Actual value mapping and forecasted likelihood



Note: Category 2 Opportunity Prediction, namely Silt Variable

Figure 2(c). Actual value mapping and forecasted likelihood



Note: Category 3 Opportunity Prediction, namely the Clay Variable

The spatial visualization of probabilistic predictions for soil particle-size categories, presented in Figure 2(a)–(c), demonstrates a strong spatial agreement between the predictions of the Geographically Weighted

Ordinal Logistic Regression (GWOLR) model and the actual field observations for each category—sand, silt, and clay. These maps depict the spatial distribution of the probability of occurrence for each soil texture class, using color gradients to represent the magnitude of predicted probabilities across the study area. The model's ability to replicate observed spatial patterns through localized probability estimates highlights the robustness of the GWOLR approach, not only in achieving high classification accuracy but also in producing spatially informative and interpretable soil texture maps. This finding emphasizes the model's value as a decision-support tool for site-specific land-use planning, targeted soil conservation, and precision agriculture in topographically complex regions such as the Kalikonto Watershed Area.

4.5 Spatial Ordinal Logistic Regression Classification Accuracy

The precision of particle Soil classification within the Kalikonto Watershed in Batu City was evaluated using the Geographically Weighted Ordinal Logistic Regression model. This model's proficiency in predicting the probability of various particle Soil size categories was recorded in Table 4, indicating the accuracy of the classification as specified below.

Table 4. Precision in Classifying through Spatial Ordinal Logistic Regression

Category	Predictions			Accuracy
	1	2	3	
1	35	1	1	94.59%
2	1	6	1	75 %
3	0	1	4	80%

Note: Category 1 (sand), Category 2 (silt), Category 3 (clay)

The overall predictive accuracy of the classification is derived from the following calculations:

$$=\left(\frac{\text{number of correct predictions}}{\text{N}}\right) \text{ x } 100\%;$$

$$= \left(\frac{35+6+4}{50}\right) \text{ x } 100\ \% = 90\%.$$

Upon examining the data presented in Table 4 regarding the precision of soil particle size categorization, the Geographically Weighted Ordinal Logistic Regression model demonstrates a commendable classification accuracy of 90%. This outcome is notably satisfactory, as the accuracy rate substantially exceeds the 50% threshold, indicating a high level of reliability in the model's predictive capability.

The remarkable precision of the Geographically Weighted Ordinal Logistic Regression model, attaining a 94% accuracy rate in classifying soil particle sizes within the Kalikonto Watershed Area, marks a significant leap forward in the domains of precision agriculture and environmental management. This exceptional accuracy not only underscores the model's robustness and reliability in predicting soil characteristics essential for agronomic decision-making but also heralds a new era in sustainable land management. The model's adeptness at accurately classifying soil particles opens avenues for implementing more customized soil conservation strategies, thereby potentially boosting crop yields and preventing land degradation. Given the pivotal role of soil particle size in influencing water flow and nutrient cycling, it lays a solid quantitative groundwork for subsequent ecological and hydrological research. Moreover, the success of the Geographically Weighted Ordinal Logistic Regression model may spur the creation of tailored soil management practices specific to regions, shaping policy decisions and guiding the distribution of resources. This advancement transcends the academic realm, promising immediate practical applications that stand to benefit agricultural stakeholders, environmental policymakers, and conservationists alike, not only in the Kalikonto Watershed Area but also in similar ecosystems worldwide. By heralding a potential shift towards more productive and environmentally responsible agricultural practices, this research fosters a crucial dialogue among researchers, policymakers, and practitioners, nurturing a mutually beneficial relationship with the land and setting a precedent for future environmental assessments and innovations in agricultural methodologies.

4.6 Best Model Checking

Assessment of Ordinal Logistic Regression and Geographically Weighted Ordinal Logistic Regression Models in Depicting Soil Particle Size Distribution in Kalikonto Watershed Area: An Evaluation of Model Superiority through Classification Accuracy. This study examines the efficiency of both ordinal logistic regression and geographically weighted ordinal logistic regression models in representing the distribution of soil particle size in the Kalikonto watershed for the year 2023. The main criterion for evaluating the performance of these models is their classification accuracy. Thus, the model that achieves the highest classification accuracy is considered the most effective.

Table 5. Best model comparison

Model	Accuracy
Ordinay Logistic Regression	88 %
Spatial Ordinary Logistic Regression	90 %

The analysis presented in Table 5 demonstrates that within the Kalikonto Watershed, Batu City, soil particle size classification accuracy using the ordinal logistic regression and Geographically Weighted Ordinal Logistic Regression models stands at 88% and 90%, respectively. This comparison underscores the Geographically Weighted Ordinal Logistic Regression model's superior accuracy rate of 90%, surpassing the conventional ordinal logistic regression model. The improved performance of the Geographically Weighted Ordinal Logistic Regression model is attributed to its incorporation of spatial variables, highlighting the pivotal influence of geographical factors on the model's predictive accuracy.

Upon reviewing these findings, it becomes evident that applying geographically weighted ordinal logistic regression significantly enhances the precision of soil particle size classification in the Kalikonto watershed. Achieving an impressive 90% classification accuracy compared to the 88% accuracy of the standard ordinal logistic regression model, the Geographically Weighted Ordinal Logistic Regression model asserts its advantage. This advantage is ascribed to its ability to factor in spatial context during analysis, revealing that location-specific variables within the watershed considerably affect the distribution of soil particle sizes.

The implications of this result are twofold. Firstly, it underscores the importance of spatial analysis in environmental modelling, specifically in the context of soil studies. The added accuracy provided by the Spatial Ordinary Logistic Regression model could be vital for more precise soil management and conservation strategies within the watershed. Secondly, this finding invites further investigation into the specific location-based factors that may affect soil composition, opening avenues for targeted research that can leverage the Geographically Weighted Ordinary Logistic Regression model's strengths to yield even more nuanced insights into soil variability across different geographical landscapes.

5. Conclusion

Based on a comprehensive analysis of the soil particle size distribution within the Kalikonto Watershed area, Kota Batu, in 2023, this study employed two distinct regression analysis techniques: ordinal logistic regression and Geographically Weighted Ordinary Logistic Regression (GWOLR). The outcomes demonstrate a notable advantage of GWOLR in terms of classification capability, achieving a classification accuracy of 90%, as opposed to the 88% accuracy attained by ordinal logistic regression. This distinction underlines the enhanced performance of the GWOLR model, accentuating the significance of incorporating geographical considerations into regression analyses for a more profound and precise comprehension of soil particle size distribution patterns.

In addition to its classification performance, several key findings emerge from this study. Firstly, the Moran's I test confirms the existence of significant spatial autocorrelation (p-value < 0.05) among all topographic predictor variables, affirming the necessity of integrating spatial elements into soil modeling. Secondly, the GWOLR model constructs localized models at each observation point, enabling the detection of spatially varying relationships between topographical features and soil texture categories. Notably, positive coefficients for vertical curvature and slope indicate a tendency toward sandy soil in elevated or sloped areas, whereas negative coefficients for the northness aspect suggest greater silt and clay presence in flatter, less exposed terrains.

Furthermore, the spatial prediction maps of soil particle probabilities (sand, silt, and clay) demonstrate high visual and statistical concordance with actual field data, supporting the validity and robustness of the spatial modeling approach. These insights offer valuable implications for designing targeted and location-specific soil conservation and land-use strategies.

This study acknowledges limitations, particularly the potential influence of other environmental variables not included in the current model. Future research is encouraged to explore the inclusion of such variables and expand the study to diverse geographical contexts to validate the broader applicability of GWOLR. These efforts will further strengthen geospatial data analysis in environmental and agricultural decision-making processes.

6. Acknowledgement

The authors express their profound appreciation to the Department of Statistics, Faculty of Mathematics and Sciences, at the University of Brawijaya, for their indispensable financial support, which was essential for realizing this research project. The allocation of funding played a critical role, facilitating the procurement of necessary resources and the successful conduct of our study. We are also immensely grateful for the perceptive feedback and thorough reviews from our peers and colleagues. Their significant input was crucial in enhancing our manuscript, ensuring its scholarly integrity, and augmenting its contribution to the broader scientific discourse. Such cooperative support mirrors our collective dedication to the advancement of knowledge and highlights the significance of joint endeavors in the realm of academic research.

References

Abidin, M. H. Z., Saad, R., Wijeyesekera, D. C., Ahmad, F., Baharuddin, M. F. T., Tajudin, S. A. A., & Madun, A. (2017). The influences of basic physical properties of clayey silt and silty sand on its laboratory electrical resistivity value in loose and dense conditions. *Sains Malaysiana*, *46*(10), 1959–1969.

Albert, A., & Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*, *71*(1), 1–10.

Anselin, L. (2019). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In *Spatial analytical perspectives on GIS* (pp. 111–126). Routledge.

Anselin, L. (2020). Local spatial autocorrelation. *Other Local Spatial Autocorrelation Statistics*.

Bertsimas, D., & King, A. (2017). Logistic regression: From art to science. *Statistical Science*, 367–384.

Cao, X., Liu, Y., Li, T., & Liao, W. (2019). Analysis of spatial pattern evolution and influencing factors of regional land use efficiency in China based on ESDA-GWR. *Scientific Reports*, *9*(1), 520.

Chen, Y. (2021). An analytical process of spatial autocorrelation functions based on Moran's index. *PLoS One*, *16*(4), e0249589.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., & Wang, Y. (2020). The GWR route map: a guide to the informed application of Geographically Weighted Regression. *ArXiv Preprint ArXiv:2004.06070*.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., & Wang, Y. (2023). A route map for successful applications of geographically weighted regression. *Geographical Analysis*, *55*(1), 155–178.

Ding, W., & Huang, C. (2017). Effects of soil surface roughness on interrill erosion processes and sediment particle size distribution. *Geomorphology*, *295*, 801–810.

Du, Z., Wang, Z., Wu, S., Zhang, F., & Liu, R. (2020). Geographically neural network weighted regression for the accurate estimation of spatial non-stationarity. *International Journal of Geographical Information Science*, *34*(7), 1353–1377.

Fagerland, M. W., & Hosmer, D. W. (2016). Tests for goodness of fit in ordinal logistic regression models. *Journal of Statistical Computation and Simulation*, *86*(17), 3398–3418.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265.

Gao, J. (2021). *Fundamentals of spatial analysis and modelling*. CRC Press.

Gxasheka, M., Gajana, C. S., & Dlamini, P. (2023). The role of topographic and soil factors on woody plant encroachment in mountainous rangelands: A mini literature review. *Heliyon*.

Hong, H., Pradhan, B., Sameen, M. I., Chen, W., & Xu, C. (2017). Spatial prediction of rotational landslide using geographically weighted regression, logistic regression, and support vector machine models in Xing Guo area (China). *Geomatics, Natural Hazards and Risk*, *8*(2), 1997–2022.

Hou, D., Bolan, N. S., Tsang, D. C. W., Kirkham, M. B., & O'Connor, D. (2020). Sustainable soil use and management: An interdisciplinary and systematic approach. *Science of the Total Environment*, *729*, 138961.

Huang, Z., Li, S., Peng, Y., & Gao, F. (2023). Spatial Non-Stationarity of Influencing Factors of China's County Economic Development Base on a Multiscale Geographically Weighted Regression Model. *ISPRS International Journal of Geo-Information*, *12*(3), 109.

Kamberaj, V. (2021). Categorical Data Analysis Using Logistic Regression. *Available at SSRN 3921693*.

Koiter, A. J., Owens, P. N., Petticrew, E. L., & Lobb, D. A. (2017). The role of soil surface properties on the particle size and carbon selectivity of interrill erosion in agricultural landscapes. *Catena*, *153*, 194–206.

Leightner, J. E., & Inoue, T. (2012). Solving the omitted variables problem of regression analysis using the relative vertical position of observations. *Advances in Decision Sciences*, *2012*.

Liu, G., Wang, J., Liu, X., Liu, X., Li, X., Ren, Y., Wang, J., & Dong, L. (2018). Partitioning and geochemical fractions of heavy metals from geogenic and anthropogenic sources in various soil particle size fractions. *Geoderma*, *312*, 104–113.

Liu, J., Yang, K., Tariq, A., Lu, L., Soufan, W., & El Sabagh, A. (2023). Interaction of climate, topography and soil properties with cropland and cropping pattern using remote sensing data and machine learning methods. *The Egyptian Journal of Remote Sensing and Space Science*, *26*(3), 415–426.

Minár, J., Evans, I. S., & Jenčo, M. (2020). A comprehensive system of definitions of land surface (topographic) curvatures, with implications for their application in geoscience modelling and prediction. *Earth-Science Reviews*, *211*, 103414.

Mishra, V. N., Kumar, V., Prasad, R., & Punia, M. (2021). Geographically weighted method integrated with logistic regression for analyzing spatially varying accuracy measures of remote sensing image classification. *Journal of the Indian Society of Remote Sensing*, *49*, 1189–1199.

M. Tahat, M., M. Alananbeh, K., A. Othman, Y., & I. Leskovar, D. (2020). Soil health and sustainable agriculture. *Sustainability*, *12*(12), 4859.

Murtagh, F., & Heck, A. (2012). *Multivariate data analysis* (Vol. 131). Springer Science & Business Media.

Ngabu, W., Fitriani, R., Pramoedyo, H., & Astuti, A. B. (2023). CLUSTER FAST DOUBLE BOOTSTRAP APPROACH WITH RANDOM EFFECT SPATIAL MODELING. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *17*(2), 945–954.

Nkeki, F. N., & Asikhia, M. O. (2019). Geographically weighted logistic regression approach to explore the spatial variability in travel behaviour and built environment interactions: Accounting simultaneously for demographic and socioeconomic characteristics. *Applied Geography*, *108*, 47–63.

Peruzzetto, M., Mangeney, A., Bouchut, F., Grandjean, G., Levy, C., Thiery, Y., & Lucas, A. (2021). Topography curvature effects in Thin-Layer models for Gravity-Driven flows without bed erosion. *Journal of Geophysical Research: Earth Surface*, *126*(4), e2020JF005657.

Plant, R. E. (2018). *Spatial data analysis in ecology and agriculture using R*. cRc Press.

Pramoedyo, H., Aini, N. N., Riza, S., & Ariyanto, D. (2021). Prediction of Soil Particle Size Fraction using Geographically Weighted Regression and Random Forest. *WSEAS Transactions on Mathematics*, *20*, 683–693.

Pramoedyo, H., Ngabu, W., Riza, S., & Amaliana, L. (2023). Single Bootstrap Approach With Geographically Weight Regression Modeling Using Particle-Size Fraction. *Journal of Theoretical and Applied Information Technology*, *101*(10), 3749–3756.

Pramoedyo, H., Ngabu, W., Riza, S., & Iriany, A. (2024). Spatial Analysis Using Geographically Weighted Ordinary Logistic Regression (GWOLR) Method for Prediction of Particle-Size Fraction in Soil Surface. *IOP Conference Series: Earth and Environmental Science*, *1299*(1), 012005. https://doi.org/10.1088/1755-1315/1299/1/012005

Riza, S., Sekine, M., Kanno, A., Yamamoto, K., Imai, T., & Higuchi, T. (2021). Modeling soil landscapes and soil textures using hyperscale terrain attributes. *Geoderma*, *402*, 115177.

Rodrigues, M., Jiménez-Ruano, A., Peña-Angulo, D., & De la Riva, J. (2018). A comprehensive spatial-temporal analysis of driving factors of human-caused wildfires in Spain using Geographically Weighted Logistic Regression. *Journal of Environmental Management*, *225*, 177–192.

Rodríguez-Lado, L., & Lado, M. (2017). Relation between soil forming factors and scaling properties of particle size distributions derived from multifractal analysis in topsoils from Galicia (NW Spain). *Geoderma*, *287*, 147–156.

Ruymgaart, F., Wang, J., Wei, S.-H., & Yu, L. (2011). Some asymptotic theory for functional regression and classification. *Advances in Decision Sciences*, *2011*.

Siqueira, R. G., Moquedace, C. M., Francelino, M. R., Schaefer, C. E. G. R., & Fernandes-Filho, E. I. (2023). Machine learning applied for Antarctic soil mapping: Spatial prediction of soil texture for Maritime Antarctica and Northern Antarctic Peninsula. *Geoderma*, *432*, 116405.

Tian, M., Wang, X., Wang, Q., Qiao, Y., Wu, H., & Hu, Q. (2023). Geographically weighted regression (GWR) and Prediction-area (PA) plot to generate enhanced geochemical signatures for mineral exploration targeting. *Applied Geochemistry*, *150*, 105590.

Windle, M. J. S., Rose, G. A., Devillers, R., & Fortin, M.-J. (2010). Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES Journal of Marine Science*, *67*(1), 145–154.

Zhang, C., & Yang, Y. (2020). Modeling the spatial variations in anthropogenic factors of soil heavy metal accumulation by geographically weighted logistic regression. *Science of The Total Environment*, *717*, 137096.

Zuhdi, S., & Saputro, D. R. S. (2017). R programming for parameters estimation of geographically weighted ordinal logistic regression (GWOLR) model based on Newton Raphson. *AIP Conference Proceedings*, *1827*(1).