E-ISSN: 2962-6110
P-ISSN: 2964-7843
JSRET
jsret.knpub.com
Journal of Scientific Research, Education, and Technology

Spatial Estimation of Land Surface Temperature using Cokriging Approach in Buleleng Regency

Muhammad Daffa Bintang Setyawan, Suliyanto, Dita Amelia, M. Fariz Fadillah Mardianto
Department of Mathematics, Faculty of Science and Technology, Airlangga University, Indonesia

ABSTRACT

Urban Heat Island (UHI) information plays a crucial role in microclimate monitoring and understanding environmental phenomena. However, the availability of observation station data is often spatially limited. This study aims to estimate Land Surface Temperature (LST) distribution in Buleleng Regency using Cokriging, utilizing the Normalized Difference Vegetation Index (NDVI) as a secondary variable due to its physical correlation with temperature. Addressing the significant scale disparity between LST and NDVI, this study applies Z-score transformation prior to modeling to ensure covariance matrix stability. Experimental variograms were constructed using Sturges' rule for lag distance determination and modeled using the Linear Coregionalization Model (LCM) to maintain a positive definite kriging matrix. Model evaluation using Leave-One-Out Cross-Validation (LOOCV) revealed a strong negative correlation between LST and NDVI (Pearson coefficient of -0.890). The Gaussian model was selected as the best fit, indicated by a low Mean Squared Error (MSE) of 2.7680 and a Mean Absolute Percentage Error (MAPE) of 3.63%. These results demonstrate that integrating NDVI through Cokriging significantly improves spatial estimation accuracy. Furthermore, this study supports the Sustainable Development Goals (SDGs), particularly Goal 13 (Climate Action) and Goal 15 (Life on Land), by providing high-precision environmental data essential for effective climate resilience planning.

Keywords: *Cokriging, LST, NDVI, UHI, Buleleng.*

Corresponding author
Name: Muhammad Daffa Bintang Setyawan
Email: muhammad.daffa.bintang-2022@fst.unair.ac.id

INTRODUCTION

Urban Heat Island (UHI) characterizes a thermal phenomenon where urban centers exhibit significantly higher temperatures than their surrounding rural regions, primarily driven by land surface modifications and the reduction of natural vegetation cover. This phenomenon results from the extensive replacement of permeable soil with impervious materials such as asphalt and concrete, which amplifies heat absorption and retention, thereby elevating Land Surface Temperature (LST) (Howard, 1966; Dimitrov et al., 2024). As urbanization accelerates globally, particularly in developing nations, the implications of UHI on environmental quality and human health have become increasingly critical. In the context of Indonesia, rapid infrastructure development and uncontrolled land-use changes have intensified these thermal effects, notably in metropolitan hubs such as Jakarta, Surabaya, and Bandung (Muzaky & Jaelani, 2022; Widyanti et al., 2025). However, emerging evidence suggests that thermal stress is no longer confined to these megacities, as semi-urban and coastal regions undergoing tourism-driven development are now facing similar climatic challenges. Consequently, understanding the propagation of UHI in these developing regions is essential for anticipating future environmental risks in areas previously considered climatically stable.

Buleleng Regency in Bali represents a critical case study for this emerging trend, having recently recorded the highest air temperature in Indonesia at 38.8°C in 2024, signaling a significantly alarming departure from historical norms (BMKG, 2024). To effectively monitor such anomalies, satellite-based remote sensing offers a robust approach for analyzing surface thermal conditions across extensive spatial scales. Specifically, Land Surface Temperature (LST) derived from thermal imagery serves as a primary metric for identifying Surface Urban Heat Island (SUHI) patterns, while the Normalized Difference Vegetation Index (NDVI) acts as a vital proxy for vegetation density. A strong inverse physical relationship exists between these variables; areas with lower NDVI typically exhibit higher LST due to the loss of natural cooling mechanisms such as evapotranspiration and shading (Mitra et al., 2025). Previous local research utilizing Landsat data has confirmed this trajectory, revealing that built-up area expansion in Buleleng has already driven surface temperature increases of 2°C to 7°C (Nugraha, 2020). These findings underscore the urgency of utilizing advanced remote sensing indicators to accurately map and monitor the escalating thermal risks in the region.

While remote sensing provides valuable point-based or raster data, accurate spatial interpolation is essential to estimate environmental variables at unsampled locations to create continuous surfaces. Conventional univariate methods, such as Ordinary Kriging, often lack precision in heterogeneous environments because they rely solely on a single primary variable, thereby ignoring complex environmental interactions. In contrast, Cokriging extends this capability by incorporating secondary variables, such as NDVI, which are spatially correlated with the primary LST data, to refine the estimation structure. Numerous studies have demonstrated that multivariate geostatistical approaches significantly outperform deterministic methods like Inverse Distance Weighting (IDW) by reducing estimation errors through the inclusion of auxiliary vegetation parameters (Bhaduri et al., 2021; Mutiah et al., 2025). Despite this potential, the application of such advanced geostatistical modeling remains scarce in Indonesia's non-metropolitan regions. Based on this background, this study aims to estimate the spatial distribution of Land Surface Temperature (LST) in Buleleng Regency using a multivariate geostatistical approach. The specific objectives of this study are: (1) to examine the spatial relationship between LST and vegetation density represented by the Normalized Difference Vegetation Index (NDVI), (2) to apply the Cokriging method by integrating NDVI as a secondary variable under the Linear Coregionalization Model (LCM), and (3) to evaluate the performance of different variogram models in improving LST estimation accuracy. The novelty of this research lies in the combined application of Z-score transformation and LCM-based Cokriging to address scale disparities and ensure covariance matrix validity in a semi-urban Indonesian context. By focusing on Buleleng Regency, this study extends the use of advanced multivariate geostatistical techniques beyond metropolitan areas and provides new empirical evidence for spatial temperature modeling in developing semi-urban regions.

METHOD

To improve clarity and reproducibility, the overall methodological framework of this study is summarized in a schematic workflow diagram, illustrating the main steps from data acquisition and preprocessing to geostatistical modeling and validation (Figure 1.)

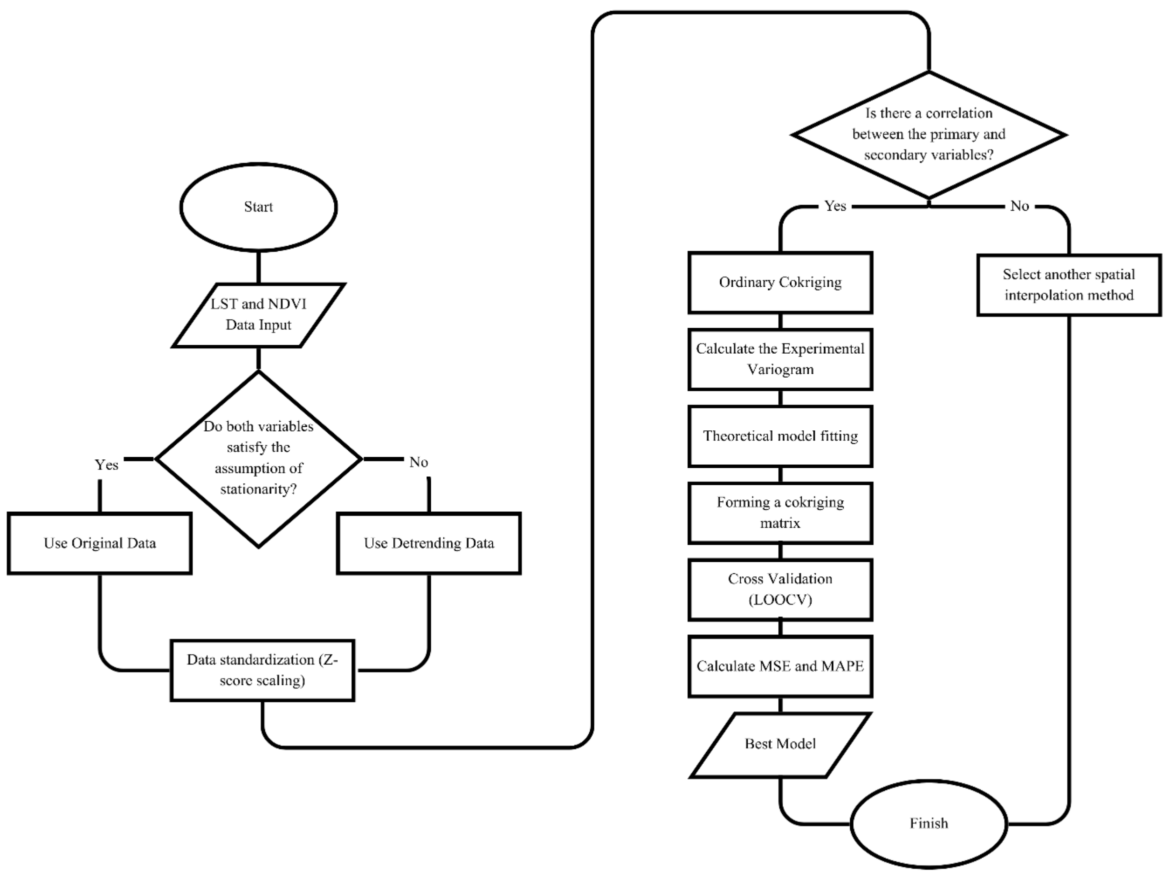

Figure 1: Research Flowchart

A. Study Area

This study was conducted in Buleleng Regency, Bali, Indonesia, which consists of nine districts and represents a coastal semi-urban region experiencing rapid land-use change. The area was selected due to its increasing surface temperature anomalies and recorded extreme air temperatures in recent years. Spatial heterogeneity in land cover, ranging from dense urban areas to agricultural land and forested regions, makes Buleleng suitable for spatial temperature modeling.

B. Data and Data Sources

This study utilized multi-source secondary spatial datasets derived from satellite remote sensing products and administrative boundary layers, all processed within the Google Earth Engine (GEE) platform. The primary variable, Land Surface Temperature (LST), was extracted from the MODIS Aqua MYD11A1 daytime product, which provides daily thermal measurements at a spatial resolution of 1 km. To ensure temporal representativeness, the LST pixel values were aggregated to compute the annual mean temperature for the period of January to December 2023, after converting the digital numbers to degrees Celsius using standard scaling factors. The secondary variable, the Normalized Difference Vegetation Index (NDVI), was derived from high-resolution Sentinel-2 Surface Reflectance imagery by utilizing the near-infrared (Band 8) and red (Band 4) spectral bands. To maintain data quality and minimize atmospheric interference, only imagery with less than 20% cloud cover was selected for the calculation. These datasets provide a comprehensive representation of the thermal and vegetative conditions across the study area for the year 2023.

Spatial Aggregation and Unit of Analysis To align the physical environmental data with regional planning units, this study adopted the administrative district as the fundamental spatial unit of analysis. The raster datasets for LST and NDVI were overlaid with the administrative boundaries of the nine districts in Buleleng Regency to perform zonal statistical analysis. Through this process, the pixel values within each district boundary were averaged to obtain a single representative mean value for both LST and NDVI, which effectively captures the macro-scale environmental characteristics of each administrative region. These aggregated values were subsequently converted into point data based on the geometric centroid coordinates of each district. This spatial transformation results in a dataset of nine representative points, which serves as the input for the multivariate Cokriging modeling to estimate the regional spatial distribution of surface temperature based on administrative zoning.

The use of administrative district centroids as spatial representation points was adopted to address the limited number of observation units available at the regional scale. Given that this study aims to support regional spatial planning and policy-oriented analysis, the administrative district was selected as the most relevant spatial unit. Aggregating raster-based LST and NDVI values at the district level allows the integration of environmental information with governance boundaries commonly used in decision-making processes. The centroid-based point representation preserves the relative spatial configuration between districts while enabling the application of geostatistical techniques that require point-based inputs, such as Cokriging.

Nevertheless, the authors acknowledge the potential influence of spatial aggregation bias, commonly referred to as the Modifiable Areal Unit Problem (MAUP), which may arise when continuous spatial data are aggregated into administrative units. Aggregation at the district level may reduce local variability and smooth extreme values of LST and NDVI. However, this limitation is considered acceptable within the scope of this study, as the primary objective is to capture regional-scale spatial patterns rather than fine-scale thermal heterogeneity. Future studies may employ finer spatial units or in situ observations to further reduce aggregation bias.

C. Research Variables

This study uses two main variables consisting of a primary variable and a secondary variable. A summary of the research variables is presented in Table 1.

Table 1: Research Variables Used in This Study

Type	Variable Name	Description	Source
Primary	Land Surface Temperature (LST)	Represents surface thermal conditions and is used as the main indicator for Surface Urban Heat Island (SUHI) analysis	MODIS Aqua MYD11A1
Secondary	Normalized Difference Vegetation Index (NDVI)	Indicates vegetation density and greenness; inversely related to surface temperature	Sentinel-2 Surface Reflectance

The use of NDVI as a secondary variable in the Cokriging model is based on its strong spatial correlation with LST. Previous studies have shown that incorporating vegetation indices as auxiliary variables can significantly improve the accuracy of land surface temperature interpolation compared to univariate geostatistical methods (Mutiah et al., 2025).

D. Data Preprocessing

Prior to geostatistical modeling, several preprocessing steps were applied to ensure data quality and compatibility with the Cokriging framework. First, exploratory data analysis (EDA) was conducted to examine the statistical characteristics and spatial distribution of Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI). This step aimed to identify potential outliers and to understand the initial relationship between the two variables. Second, a spatial trend analysis was performed to evaluate the stationarity assumption required in geostatistical modeling. The presence of large-scale spatial trends was assessed using coordinate-based regression. When a significant trend was detected, a detrending procedure was applied by modeling and removing the deterministic component, allowing the residuals to satisfy second-order stationarity (Wackernagel, 2003). Third, both LST and NDVI values were standardized using Z-score transformation to place the variables on comparable scales. This step is essential in multivariate geostatistical analysis to prevent dominance of one variable over another and to ensure numerical stability during the Cokriging weight estimation (Wackernagel, 2003). Finally, the processed data were converted into point-based spatial datasets using district centroid coordinates. These standardized and stationary datasets were then used as inputs for variogram, cross-variogram, and Cokriging modeling.

E. Variogram Modeling

Experimental Variogram

An experimental variogram is a variogram obtained from observed sample data or measurement data. According to (Kelkar & Perez, 2002), calculating the experimental variogram value is expressed in the following equation:

$$\hat{\gamma}(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z_i - Z_{i+h}]^2$$

with $\hat{\gamma}(h)$ is the semi variance estimate based on sample data at a distance h. Meanwhile, to calculate the variogram that contains a cross-relationship or commonly called cross variance, it can be written as follows:

$$\gamma_c(h) = \frac{1}{2}E\{[Z_i - Z_{i+h}][Z2_i - Z2_{i+h}]\}$$

In practice, to estimate the cross variance between variables z and z2 is as follows:

$$\hat{\gamma}_c(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z_i - Z_{i+h}][Z2_i - Z2_{i+h}]$$

Meanwhile, the cross covariance equation can be written as follows:

$$C_c(h) = E[Z_i Z2_{i+h}] - E[Z_i]E[Z2_{i+h}]$$

In practice, to estimate the cross covariance between variables z and z2 is as follows:

$$c_c(h) = \frac{1}{N(h)}\sum_{i=1}^{N(h)}[Z_i - Z2_{i+h}] - \frac{1}{N(h)}\sum_{i=1}^{N(h)} Z_i \times \frac{1}{N(h)}\sum_{i=1}^{N(h)} Z2_{i+h}$$

In many cases, cross covariance is assumed to be symmetric, namely $C_c(-h) = C_c(h)$.

$$\gamma_c(h) = C_C(0) - C_c(h)$$

where $C_C(0)$ is the cross covariance at distance 0 which can be calculated as follows:

$$c_c(0) = \frac{1}{N}\sum_{i=1}^{N} Z_i - Z2_i - \frac{1}{N}\sum_{i=1}^{N} Z_i \sum_{i=1}^{N} Z2_i$$

where N is the number of locations where the two variables Z and Z2 are measured.

According to Isaacs, E.H., and Srivastava (1989), there are several components in a variogram: sill, nugget, and range. Sill $(C_0 + C)$ is the variogram value at the top (off level) of the variogram. It can also be interpreted as the "amplitude" of a particular component of the variogram. Range (A) is the distance at which the variogram reaches the sill. In theory, the initial value of a variogram is zero. When the lag approaches zero, the variogram value is called a nugget. Nugget (C_0) represents variations at very small distances (lags), including measurement errors.

To determine the ideal number of class intervals (k) to properly represent the distance distribution, not too few (too general) and not too many (too random), this study applies the principle of Sturges' Rule. Mathematically, determining the number of classes according to Sturges (1926) is formulated as follows:

$$k = 1 + 3{,}322\ log_{10}\, n$$

With:
k is the number of interval classes (the number of lags that will be formed)
n is the number of pairs of points

Once the k value (number of classes) is obtained, the next step is to calculate the class interval width or lag size that will be used on the horizontal axis of the variogram. The interval width (c) is calculated by dividing the data range (maximum distance minus minimum distance) by the number of classes obtained:

$$c = \frac{X_{max} - X_{min}}{k}$$

With:

c is the number of interval classes (the number of lags that will be formed)

X is the distance between a pair of points

By using this method, the resulting experimental variogram is expected to have proportional spatial resolution, so that the Sill, Range, and Nugget patterns can be identified objectively.

Theoretical Variogram

Since the experimental variogram only yields empirical semi variance values at discrete lag distances, a continuous mathematical function must be fitted to these points to ensure the existence of a solution for the kriging system. The primary objective of this theoretical modeling is to satisfy the positive definite condition of the variance-covariance matrix, which is a prerequisite for minimizing the estimation error variance. In this study, three standard isotropic models were evaluated to describe the spatial structure of LST and NDVI: the Spherical, Exponential, and Gaussian models.

The **Spherical Model** represents spatial phenomena with a clear transition point where autocorrelation ceases at a specific range. The equation is defined as:

$$\gamma(h) = \begin{cases} (C_0 + C)\left[\frac{3}{2}\left(\frac{h}{A_0}\right) - \frac{1}{2}\left(\frac{h}{A_0}\right)^3\right];\ h \leq A_0 \\ \gamma(h) = (C_0 + C); h > A_0 \end{cases}$$

The **Exponential Model** describes spatial continuity that decreases asymptotically, meaning the correlation never completely reaches zero but approaches a limit. It is formulated as:

$$\gamma(h) = (C_0 + C)\left[1 - exp\left(\frac{-3h}{A_0}\right)\right]$$

The **Gaussian Model** characterizes extremely continuous spatial phenomena with smooth variation at short distances. This model is often suitable for variables like temperature that change gradually over space:

$$\gamma(h) = (C_0 + C)\left[1 - exp\left(-3\frac{h^2}{A_0^2}\right)\right]$$

where $\gamma(h)$ is the semivariance at lag distance h, C_0 is the nugget effect, c is the partial sill, and A_0 is the range (or effective range) of spatial dependence. The selection of the optimal theoretical model was based on the lowest Mean Squared Error (MSE) and Root Sum Squared (RSS) values obtained from the cross-validation process. Furthermore, to accommodate the multivariate nature of the study, the parameters were fitted under the Linear Coregionalization Model (LCM) constraint, ensuring that all direct and cross-variograms share the same basic structure and range parameters to maintain mathematical consistency.

F. Cokriging Modeling

Cokriging is a multivariate extension of the Kriging method that incorporates one or more secondary variables that are spatially correlated with the primary variable to improve estimation accuracy (Wackernagel, 2003). In this study, Land Surface Temperature (LST) is treated as the primary variable, while Normalized Difference Vegetation Index (NDVI) is used as the secondary variable.The ordinary Cokriging estimator for LST at an unsampled location $s_0$is defined as:

$$Z_0^* = \sum_{i=1}^{N} \lambda_{Z_i} Z_i + \sum_{k=1}^{M} \lambda_{Z2_k} Z2_k$$

Where

Z_0^* is the estimated value *of cokriging* at the suspected new location
λ_{Z_i} is cokriging weighted for the primer variable LST at the location i

λ_{Z2_k} is cokriging weighted for the secondary variable NDVI at the location k
Z_i is the value of the LST variable at the location i
$Z2_k$ is the value of the NDVI variable at the location k
To ensure unbiased estimation, the weights are subject to the constraint:

$$\sum_{i=1}^{N} \lambda_{Z_i} = 1 \ ; \ \sum_{k=1}^{M} \lambda_{Z2_k} = 0$$

The Cokriging weights are obtained by minimizing the estimation variance:

$$\hat{\sigma}_E^2 = V\left[Z_0^* - \sum_{i=1}^{N} \lambda_{Z_i} Z_i - \sum_{k=1}^{M} \lambda_{Z2_k} Z2_k\right]$$

under the unbiasedness constraint. This leads to a system of linear equations involving auto-variograms and cross-variograms.

G. Linear Coregionalization Model (LCM)

In multivariate geostatistical modeling, ensuring the validity of the covariance matrix is a fundamental prerequisite. To address this, the Linear Coregionalization Model (LCM) was employed to fit the experimental direct and cross-variograms. The LCM assumes that all study variables, LST and NDVI are generated from linear combinations of the same underlying independent spatial processes. This assumption is critical because it imposes a structural constraint where the direct variograms and the cross-variogram must share identical basic structures and range parameters. By enforcing this commonality, the LCM mathematically guarantees that the resulting system of Cokriging equations yields a positive definite matrix. This property is essential to ensure that the estimation variance remains non-negative and the system of linear equations is solvable.

H. Model Validation

The performance of the Cokriging model was evaluated using Leave-One-Out Cross-Validation (LOOCV), which is commonly applied in geostatistical studies with limited observation points. In LOOCV, one observation is temporarily removed from the dataset, and its value is estimated using the remaining observations. This process is repeated for all observation points, producing a set of predicted values for validation.
The prediction accuracy was assessed using the following error metrics:

Mean Squared Error (MSE)

$$MSE = \frac{1}{N} \sum_{i=1}^{N} [Z_i - \hat{Z}_i]^2$$

MSE measures the average squared difference between observed and predicted values, with lower values indicating better model performance.

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Z_i - \hat{Z}_i}{Z_i}\right| \times 100\%$$

MAPE expresses the prediction error in percentage form, allowing easier interpretation of model accuracy.

The Cokriging model with the lowest MSE and MAPE values was considered the best-performing interpolation model. These validation results provide quantitative evidence of the effectiveness of incorporating NDVI as a secondary variable in improving LST spatial estimation accuracy (Wackernagel, 2003; Mutiah et al., 2025).

RESEARCH RESULT

This study reveals a clear inverse relationship between Land Surface Temperature (LST) and vegetation density (NDVI) across Buleleng Regency. Descriptive analysis shows substantial spatial variability in both variables, with higher LST values concentrated in coastal and urbanized subdistricts and lower temperatures in vegetated highland areas. Pearson correlation analysis confirms a strong negative association between LST and NDVI ($r = -0.890$, $p < 0.01$), supporting the role of vegetation as a key moderating factor of surface temperature. Variogram and cross-variogram modeling under the Linear Coregionalization Model (LCM) demonstrates significant spatial dependence for both variables, justifying the application of Cokriging. Model validation using Leave-One-Out Cross-Validation (LOOCV) indicates that the Gaussian variogram model provides the highest predictive accuracy, yielding the lowest MSE and MAPE values among the tested models.

A. Description of LST and NDVI

The descriptive analysis shows that Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI) exhibit clear spatial variability across Buleleng Regency. LST values range approximately between 27°C and 33°C, indicating substantial surface temperature differences between districts. Higher LST values are generally observed in urbanized and coastal areas, while lower values are associated with regions dominated by vegetation and higher elevation.

NDVI values display heterogeneous vegetation density across the study area, reflecting variations in land cover types. Areas with higher NDVI values tend to coincide with lower LST, suggesting the cooling effect of vegetation. This descriptive pattern provides initial evidence of a relationship between surface temperature and vegetation cover, which is further examined through correlation and geostatistical modeling.

B. Data Preprocessing

Prior to geostatistical modeling, the assumption of intrinsic stationarity was evaluated to ensure the validity of the Best Linear Unbiased Estimator (BLUE). The first-order stationarity (constancy of the mean) was assessed using Ordinary Least Squares (OLS) regression to detect global spatial trends against latitude and longitude coordinates. The regression analysis for the primary variable, Land Surface Temperature

(LST), yielded an R^2 of 0.4545. However, the p-values for the spatial predictors (Latitude: $p = 0.067$; Longitude: $p = 0.330$) were both greater than the significance level (α=0.05). This indicates that no statistically significant spatial trend exists, confirming that the original LST data satisfies the first-order stationarity assumption. Conversely, the secondary variable, NDVI, exhibited a strong spatial dependence with a higher R^2 of 0.6922. A significant linear trend was detected specifically in the latitudinal direction ($p = 0.010$), reflecting a systematic variation in vegetation density from north to south. Consequently, to satisfy the stationarity requirement, the NDVI data underwent a detrending process using a polynomial regression surface, and the resulting residuals were utilized for subsequent Cokriging modeling.

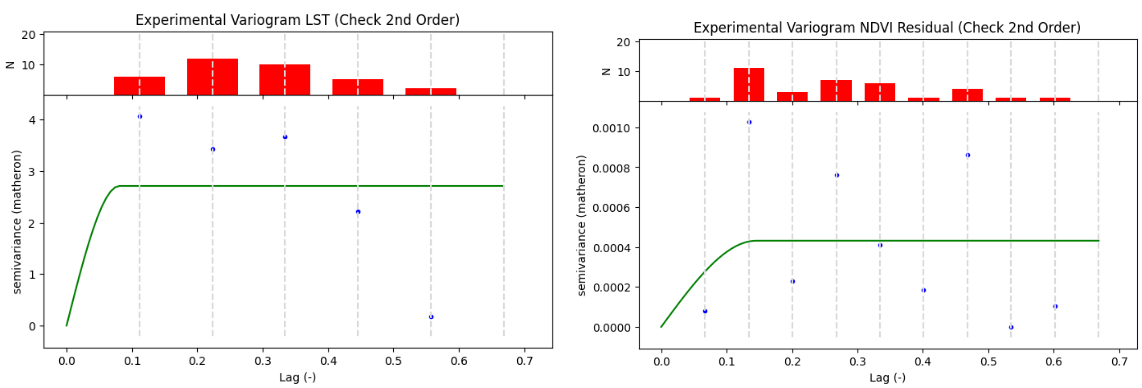

Figure 2: Second-Order Stationarity Test Results

Following the trend analysis, second-order stationarity (constancy of variance) was verified by examining the structure of the experimental semi variograms. As illustrated in Figure 2, the variograms for both the original LST data and the NDVI residuals exhibit a clear "sill", a plateau where the semi variance stabilizes as the lag distance increases. The presence of a defined sill confirms that the variance is bounded and homogeneous across the study area (homoscedasticity), thereby satisfying the second-order stationarity condition required for covariance-based structural analysis.

To ensure comparability between variables with different measurement scales, Z-score standardization was applied to both LST and NDVI. After transformation, both variables exhibit a mean close to zero and a standard deviation of one (Table 2), indicating that scale-related bias has been effectively removed. This standardization allows the cross-variogram to represent the genuine spatial relationship between LST and NDVI.

Table 2: Parameters of Auto-Variogram and Cross-Variogram Models

Variable	Data Type	Mean	Stdev	Min	Max
LST	Original Data	28.87	1.67	27.09	33.18
LST	Z-Score	-0.00	1.00	-1.05	2.57
NDVI	Residual Data	0.00	0.02	-0.05	0.03

Variable	Data Type	Mean	Stdev	Min	Max
NDVI	Z-Score	-0.00	1.00	-2.11	1.40

These results indicate that the bias caused by differences in measurement units has been successfully eliminated. Thus, spatial fluctuations in LST and NDVI are now of equal magnitude, so the resulting spatial correlation structure (cross-variogram) purely represents the physical relationship between variables, not an artifact of differences in data scale.

C. Correlation Between LST and NDVI

The relationship between LST and NDVI was quantified using Pearson correlation analysis. The results show a strong negative correlation between the two variables, with a correlation coefficient of $r = -0.890$, and $p - value = 0.0013$.

This statistically significant result indicates that areas with higher vegetation density tend to experience substantially lower surface temperatures. The strong inverse relationship confirms the theoretical role of vegetation in regulating surface thermal conditions through evapotranspiration and shading mechanisms (Mitra et al., 2025). This finding provides a solid statistical justification for incorporating NDVI as a secondary variable in the Cokriging framework.

D. Experimental Variogram and Cross-Variogram Modeling

Based on the experimental variogram that has been formed, a theoretical model fitting process is carried out to obtain continuous spatial structure parameters. Referring to the principle of the Linear Coregionalization Model (LCM), the variogram model for LST, NDVI, and cross variogram is constructed using the same basic structure, namely a combination of Nugget and Gaussian models with uniform range parameter limitations. This standardization is carried out to ensure the validity of the matrix. The results of the estimation of model parameters which include nugget, partial sill, and range values for the three components of the variogram are summarized in Table 3.

Table 3: parameters of auto-variogram and cross-variogram models

Component	Nugget Effect	Sill	Range	Interpretation
LST Auto-variogram	0.0022	0.5699	0.0557	Strong spatial dependence with minimal microscale variation
NDVI Auto-variogram	0.1575	0.9179	0.5015	Higher local variability and broader spatial continuity
Cross-Variogram	0.0384	0.1076	0.1672	Moderate inverse spatial correlation between variables

The spatial dependence structures of LST and NDVI were analyzed using experimental variograms and modeled through the Linear Coregionalization Model (LCM). The estimated parameters of the auto-variograms and cross-variogram are summarized in Table 3. The auto-variogram of LST exhibits a very small nugget effect, indicating minimal measurement error and strong spatial continuity at short distances. The relatively short range suggests that surface temperature variations are influenced by local-scale processes, such as land cover composition and urban structure. In contrast, the NDVI auto-variogram shows a larger nugget effect and longer range, reflecting higher local heterogeneity in vegetation distribution and broader spatial continuity. This difference highlights the distinct spatial behavior between thermal and vegetation variables. The cross-variogram parameters confirm a moderate but significant spatial association between LST and NDVI, operating over an intermediate spatial range. This result indicates that vegetation influences surface temperature not only locally but also across neighboring areas. The presence of consistent cross-spatial dependence satisfies the core requirement for applying Cokriging and supports the joint modeling of LST and NDVI within the LCM framework.

E. Model Validation Results Using LOOCV

To assess the predictive performance of the variogram models, a Leave-One-Out Cross-Validation (LOOCV) procedure was conducted. The line plot below presents the absolute error obtained from each iteration across subdistricts for the Spherical, Exponential, and Gaussian models.

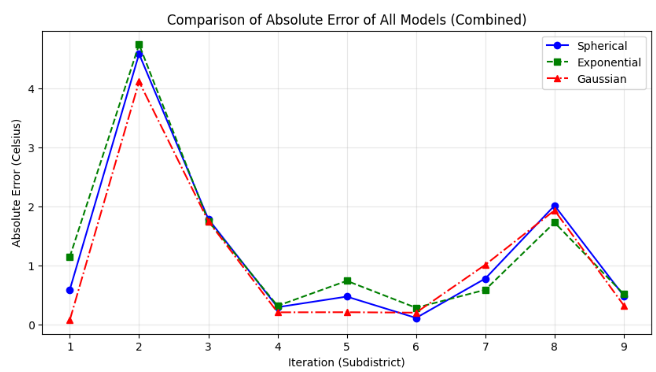

Figure 3: Comparison of Absolute Error of All Models each iteration.

The iterative LOOCV results demonstrate noticeable differences in error magnitude among the three variogram models. The Gaussian model consistently produces smaller prediction errors in most iterations, particularly in locations with moderate LST values. In contrast, the Spherical and Exponential models tend to yield larger errors in locations with extreme LST values, indicating lower adaptability to spatial variability. The largest prediction errors across all models occur in Iteration 2, corresponding to the highest observed LST value. This suggests that extreme temperature conditions are more challenging to predict accurately, especially when the spatial structure is not sufficiently smooth. Nevertheless, the Gaussian model shows relatively smaller error dispersion compared to the other models. When visualized using LOOCV error plots (to be presented in the following figures), the Gaussian model exhibits a more stable error pattern with lower fluctuation across iterations. This behavior indicates better spatial generalization and robustness of the Gaussian variogram in representing the underlying spatial continuity of LST. Model performance was evaluated using Leave-One-Out Cross-Validation (LOOCV) for three variogram models: Spherical, Exponential, and Gaussian. The validation results are summarized in Table 4.

TABLE 4. LOOCV Performance of Cokriging Models

Variogram Model	**MSE**	**MAPE (%)**
Spherical	3.2991	4.1225
Exponential	3.4566	4.3870
Gaussian	2.7680	3.6343

Among the tested models, the Gaussian variogram model demonstrates the best predictive performance, yielding the lowest MSE (2.7680) and MAPE (3.6343%). This indicates that the Gaussian model more accurately represents the smooth spatial variation of LST in the study area. The superior performance of the Gaussian model suggests that surface temperature variation in Buleleng Regency exhibits gradual spatial transitions rather than abrupt changes.

F. Spatial Mapping of LST and NDVI

The actual conditions of the research variables at the observation site are presented in Figure 3. This sample point distribution map visually shows the inverse (inversely relative) relationship pattern between surface temperature and vegetation. Areas identified as having high LST (red/yellow color) tend to have low NDVI values, and vice versa. The spatial pattern of this observational data indicates a strong negative correlation between the two variables, which is the basis for considering the use of NDVI as a secondary variable (covariate) in the Cokriging model.

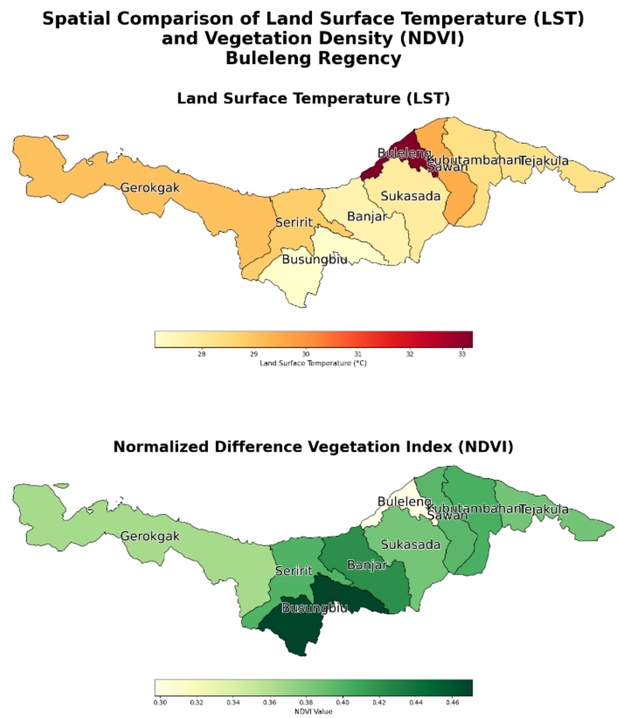

Figure 4: Spatial Comparison of Land Surface Temperature (LST) and Vegetation Density (NDVI) Buleleng Regency

Figure 4 illustrates the spatial distribution of Land Surface Temperature (LST) and the Normalized Difference Vegetation Index (NDVI) across subdistricts in Buleleng Regency. The LST map reveals a clear spatial gradient, with higher surface temperatures concentrated in coastal and urbanized subdistricts such as Buleleng, Sawan, and Kubutambahan, while lower temperatures are observed in vegetated highland areas, including Gerokgak and Busungbiu. In contrast, the NDVI map exhibits an inverse spatial pattern, where subdistricts with higher vegetation density, particularly agricultural and forested areas, display higher NDVI values, while urban and built-up areas show lower NDVI values. A visual comparison of both maps highlights a strong negative spatial correspondence between LST and NDVI, indicating that areas with reduced vegetation cover tend to experience higher surface temperatures. This spatial pattern reflects the Surface Urban Heat Island (SUHI) phenomenon, where vegetation plays a critical role in regulating land surface temperature through evapotranspiration and shading mechanisms.

DISCUSSION

This study demonstrates that incorporating vegetation information through NDVI as a secondary variable in a Cokriging framework substantially improves the spatial estimation of Land Surface Temperature (LST) in a semi-urban coastal region. The strong negative correlation between LST and NDVI confirms that vegetation plays a critical role in moderating surface thermal conditions, consistent with the Surface Urban Heat Island

(SUHI) phenomenon. Areas with dense vegetation cover exhibit lower surface temperatures due to enhanced evapotranspiration and shading effects, whereas urbanized and coastal subdistricts with limited green cover tend to accumulate heat more intensively.

From a methodological perspective, the findings align with international studies that highlight the superiority of multivariate geostatistical approaches over univariate interpolation methods. Previous research in China, Iran, and Southern Europe has shown that Cokriging incorporating vegetation indices or land cover variables consistently outperforms Ordinary Kriging and deterministic methods in estimating surface temperature, particularly in heterogeneous landscapes (Bhaduri et al., 2021; Mutiah et al., 2025). The present study extends this body of knowledge by demonstrating that the combination of Z-score transformation and the Linear Coregionalization Model (LCM) produces a stable and reliable spatial structure, even with a limited number of observation points. This contribution is particularly relevant in data-scarce regions of developing countries, where dense meteorological networks are often unavailable.

The superior performance of the Gaussian variogram model suggests that surface temperature variation in Buleleng Regency exhibits smooth spatial transitions rather than abrupt changes. This spatial behavior is characteristic of climatic variables influenced by gradual land cover gradients and topographical factors. By validating the Gaussian Cokriging model through LOOCV, this study provides empirical evidence that multivariate geostatistical modeling can effectively capture regional thermal patterns in semi-urban environments, which are often overlooked in Urban Heat Island studies that focus primarily on large metropolitan areas.

Despite its strengths, this study has several limitations that should be acknowledged. First, the use of MODIS LST data with a spatial resolution of 1 km may limit the ability to capture fine-scale urban thermal heterogeneity, particularly in densely built-up areas. Second, the relatively small number of spatial units, resulting from the use of administrative district centroids, may restrict the generalizability of the results and increase sensitivity to spatial aggregation effects. These limitations suggest that future research should explore higher-resolution thermal data, such as Landsat-based LST, and incorporate a greater number of spatial observations to enhance model robustness. Additionally, integrating other auxiliary variables, such as built-up indices or elevation, may further improve the accuracy of multivariate spatial estimation.

In terms of broader implications, the findings of this study are directly relevant to Sustainable Development Goals (SDGs) 13 (Climate Action) and 15 (Life on Land). By providing a spatially explicit assessment of surface temperature patterns and their relationship with vegetation cover, this research supports evidence-based spatial planning strategies aimed at mitigating heat stress through green infrastructure development. The methodological framework presented in this study offers a practical tool for local governments in semi-urban regions to identify priority areas for climate adaptation and sustainable land management.

CONCLUSION

This study demonstrates that incorporating vegetation information through NDVI as a secondary variable in a Cokriging framework provides an effective approach for spatially estimating Land Surface Temperature (LST) in semi-urban coastal regions. The results confirm that vegetation plays a critical role in moderating surface temperature patterns, where areas with lower vegetation density tend to experience higher thermal intensity, consistent with the Surface Urban Heat Island (SUHI) phenomenon. From a methodological standpoint, this research highlights the advantage of multivariate geostatistical modeling over univariate approaches for environmental temperature estimation. The application of Cokriging under the Linear Coregionalization Model (LCM), combined with data standardization, allows for a more reliable representation of spatial dependence between LST and NDVI, even in regions with limited observation points. This finding suggests that the proposed framework is suitable for environmental monitoring in data-scarce regions. In practical terms, the spatial patterns identified in this study provide valuable insights for local governments and spatial planners in Buleleng Regency. Areas consistently exhibiting high surface temperatures and low vegetation cover can be prioritized for climate adaptation strategies, such as the expansion of urban green spaces, coastal vegetation restoration, and the integration of green infrastructure in land-use planning. By supporting targeted interventions, the results of this study contribute to climate-resilient spatial planning and sustainable land management. Overall, this research offers a transferable methodological framework for analyzing surface temperature dynamics in semi-urban regions, particularly in developing countries facing rapid land-use change. Future studies are encouraged to apply this approach using higher-resolution thermal data and additional environmental covariates to further enhance spatial accuracy and policy relevance.

REFERENCES

Badan Meteorologi, Klimatologi, dan Geofisika. (2024). *Catatan iklim dan kualitas udara Indonesia tahun 2024*. BMKG.

Bhaduri, B., Bright, E., Coleman, P., & Dobson, J. (2021). Land use and land cover change impacts on surface temperature. *Remote Sensing of Environment*, *115*(9), 2198–2208. https://doi.org/10.1016/S1002-0160(06)60103-3

Cressie, N. A. C. (1993). *Statistics for spatial data*. John Wiley & Sons.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Howard, L. (1966). The climate of London. *The Geographical Journal*, *132*(1), 84–102.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. Oxford University Press.
Kelkar, M., & Perez, G. (2002). *Applied geostatistics for reservoir characterization*. Society of Petroleum Engineers.
Mitra, S., Mondal, M., Biswas, M., Datta, S., & Roy, A. K. (2025). Managing urban heat islands using greenery and waterscapes. *Sustainable Cities and Society*, *118*, 106033.
Mutiah, S., Aidi, M. N., Saefuddin, A., & Ernawati, F. (2025). Comparison of ordinary kriging and cokriging for spatial estimation based on simulated data. *Cauchy: Journal of Mathematics*, *10*(2), 533–544. https://doi.org/10.18860/cauchy.v10i2.33409
Muzaky, H., & Jaelani, L. M. (2022). Analisis pengaruh tutupan lahan terhadap distribusi suhu permukaan: Kajian urban heat island di Jakarta, Bandung, dan Surabaya. *Jurnal Penginderaan Jauh Indonesia*, *1*(2), 45–51.
Nugraha, A. S. A. (2020). Pemanfaatan citra penginderaan jauh multi-temporal untuk deteksi urban heat island di Kabupaten Buleleng. *Jurnal Penginderaan Jauh dan Sistem Informasi Geografis*, *22*(2), 104–116.
Sobrino, J. A., Jiménez-Muñoz, J. C., & Paolini, L. (2004). Land surface temperature retrieval from Landsat TM 5. *Remote Sensing of Environment*, *90*(4), 434–440. https://doi.org/10.1016/j.rse.2004.02.003
Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0
Wackernagel, H. (2003). *Multivariate geostatistics: An introduction with applications* (3rd ed.). Springer.
Wan, Z. (2008). New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sensing of Environment*, *112*(1), 59–74. https://doi.org/10.1016/j.rse.2006.06.026
Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment*, *89*(4), 467–483. https://doi.org/10.1016/j.rse.2003.11.005
Widyanti, R., Nasrullah, N., & Sulistyantara, B. (2025). Analisis pengaruh ruang terbuka hijau terhadap fenomena urban heat island di Kota Depok, Jawa Barat. *Jurnal Ilmu Lingkungan*, *23*(2), 371–381.