Indonesian Journal of Statistics and Its Applications
DOI: https://doi.org/ 10.29244/ijsa.v10i1p23-36

RESEARCH ARTICLE

Spatial Analysis of Primary-Sector GRDP in Pemalang Regency Using Remote Sensing Vegetation Indices

Khusnudin Tri Subhi[1], Vincentia Anggita Puspitasari[2]

[1]Department of Architecture and Planning, Faculty of Engineering, Universitas Gadjah Mada, Indonesia
[2]Graduate School of Science and Engineering, Ritsumeikan University, Japan

OPEN ACCESS

Check for updates

Article History
Received 05 May 2026
Revised 30 June 2026
Accepted 30 June 2026

Keywords
GRDP downscaling, geographically weighted regression, remote sensing

Corresponding Author:
Khusnudin Tri Subhi

Email:
khusnudin.tri.s@mail.ugm.ac.id

Department of Architecture and Planning, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia.

ABSTRACT

Agriculture has contributed the highest percentage of Pemalang Gross Regional Domestic Product in 2022 and continues to be the regency's primary source of income. However, GRDP figures published at the regency level do not show how agricultural productivity varies spatially, which matters for effective policy decisions. In this paper, this study combine official economic records with satellite-derived vegetation indicators to map where primary-sector GRDP is concentrated using a Geographically Weighted Regression (GWR) approach. This study downscaled GRDP data from BPS (Statistics Indonesia) onto a 500 × 500 m grid as the response variable. Vegetation indices including NDVI, EVI, MSAVI, LAI, and MNDWI, calculated from Landsat and MODIS imagery through Google Earth Engine, served as predictors reflecting crop condition, land productivity, and water availability. The GWR model showed a much better fit than ordinary global regression, with an Adjusted R^2 of 0.77, indicating that spatial variation is significant. The highest GRDP coefficients appeared in northern and northeastern coastal subdistricts such as Ulujami, Comal, and Petarukan, as well as in fertile hinterland areas where irrigated farming and aquaculture thrive. Lower values emerged in more urbanized central and southern areas. A clear dual landscape exists in Pemalang: strong agricultural output in rural regions, alongside a gradual shift toward manufacturing and services in urban centers. Our results highlight how remote sensing, spatial modeling, and economic statistics can be brought together to produce detailed agricultural GRDP maps that are useful for protecting farmland, targeting agricultural support, and planning more balanced regional development.

Introduction

Spatial planning in a regional area can vary depending on the economic data, especially in official statistics. But there is difficulty in the official Gross Regional Domestic Product (GRDP) that is usually created only for one region on a macro scale. Because of this, we lost the benefit of knowing every little part of the regions. So, this indicator cannot show the factual conditions across different areas of the region. This condition makes it impossible to apply the policy process to smaller areas, including between nearby sub-districts and villages (Chen et al., 2022). So, in order to accommodate this situation, we need to make GRDP downscaling into parts, as shown by grids along the region. By this action, we can easily observe the economic activity, such as manufacturing and agricultural activity, in the regency. This downscaling can overcome the limitation of the macro GRDP data that Statistics Indonesia publishes in one big region, assuming that economic value is evenly spread out over the Pemalang regency administrative area. That idea doesn't take into account a lot of local variables. It makes it hard to figure out where economic activity is really happening. This is why it is now necessary to break down GRDP by region in order to get more accurate and precise economic figures (Nashirudien, 2024; Nashrudien & Pramono, 2025).

© 2026 Tri Subhi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY) license, allowing unrestricted use, distribution, and reproduction in any medium, provided proper credit is given to the original authors.
Print only if necessary. Let's protect our trees and the planet together.

Pemalang Regency is a good example of an agrarian area that faces exactly this kind of challenge. Going by official government data, in 2022 the Agriculture, Forestry, and Fisheries sector made up the biggest portion of Pemalang's GRDP, reaching about 26.49% (Badan Pusat Statistik Kabupaten Pemalang, 2023). Studies in academics have confirmed this, stating that while agricultural growth has been somewhat slower than that of other sectors, it still serves as the region's economic foundation (Haryanto & Purnomo, 2023). That is, in part, because of the terrain of Pemalang Regency. With 23,813 ha of arable land and 38,617 ha of rice paddies, farming is the obvious economic driver of the regency. In the end, regencies like Pemalang in Central Java Province will be highly vital for keeping the area's food supply safe (Badan Pusat Statistik Provinsi Jawa Tengah, 2023; Kementerian Pertanian, 2022). The local government wants to strengthen the crucial components of farming as part of its Medium-Term Development goal (RPJMD) for 2021 to 2026 (Bappeda Kabupaten Pemalang, 2021). This is in line with Central Java Province's overall development plan in RPJMD, which promotes economic activity that needs a lot of workers and is based on the strengths of each region, such as farming (Haryanto & Purnomo, 2023).

There is a genuine requirement for innovative analytical methodologies capable of delineating agriculture's economic impact at a more granular regional scale. Several studies have attempted GRDP spatial downscaling using different approaches. Chen et al. (2022) used nighttime light (NTL) combined with NDVI to estimate county-level GDP in Fujian Province, while Li et al. (2023) applied multi-source remote sensing with machine learning for high-resolution agricultural GDP mapping. However, NTL imaging is less effective in rural and agricultural areas due to the minimal artificial light created by farmland at night (Xu et al., 2023). Most existing studies also rely on global regression models that assume a uniform relationship across the entire study area, ignoring spatial heterogeneity in agricultural landscapes (Oshan et al., 2019). In a place like Pemalang Regency, Central Java, where farming is the main industry, vegetation-based indicators such as NDVI, EVI, MSAVI, MNDWI, and LAI are considerably better at capturing how active the primary sector truly is. These vegetation indices have been useful for estimating farm productivity and crop health, especially when applied within spatial analytical frameworks (Xia et al., 2022). To address the limitations of global models, we use a Geographically Weighted Regression (GWR) framework that allows regression coefficients to vary across locations, capturing the spatial non-stationarity that standard OLS models cannot detect (Comber et al., 2022). This approach offers two key advantages over previous studies: it combines multiple vegetation indices rather than relying on NTL alone, and it accounts for local spatial variation in the relationship between vegetation condition and economic output across Pemalang Regency.

Previous researchers have conducted research by using GWR's spatial downscaling to prove the estimation for economic value well with the high resolution by using a grid of 500 x 500 meters (Nashirudien, 2024; Nashrudien & Pramono, 2025). GWR accommodates spatial heterogeneity that offers a comprehensive overview of local variation and helps address the non-stationarity that is faced by the global regression models (Gollini et al., 2015; Wang et al., 2021). Land cover data is usually used in many methodologies, but this data, if standing alone, can not be sufficiently sensitive to show the spatial patterns with the detail for land productivity. These things become the reason why the temporal vegetation index in the GWR downscaling model can be a valuable tool to gain a better understanding of the GRDP growth, especially for the primary sector in a specific area in Pemalang Regency.

With this introduction, this study aims to analyze the heterogeneity of economic conditions in Pemalang Regency in 2022 by using a combination of GIS and remote sensing tools. This work can give two contributions. First, it shows a spatial approach that can generate economic maps in high resolution, especially for agricultural areas. Second, it can provide information spatially, which the local government can use to decide on a strategy for achieving sustainable development goals.

Material and Methods

Materials and Data

As summarized in Table 1, we examined 2022 data from Statistics Indonesia, satellite imagery, and also publicly available spatial sources, which are as follows:

- The Statistics-Indonesia has 2022 Pemalang Regency GRDP data for the Agriculture, Forestry, and Fisheries Sector. The key dependent variable is in this data. We use a 500 × 500 m grid to do geographic distribution. We shall call this information the primary sector. The GRDP data refers to the 2022 fiscal year as published by BPS in 2023
- Google Earth Engine (GEE) gives us access to Landsat 8/9 and MODIS pictures that show land cover and vegetation. We use these pictures to figure out vegetation indices (NDVI, EVI, MSAVI, MNDWI, and LAI) that reveal how much agriculture is going on, if there are wetlands, and how much biomass there is.
- Maps of Pemalang Regency and sub-district boundaries from the Geospatial Information Agency (BIG) for clipping and

This journal article is © Tri Subhi et al. 2026

spatial analysis.

The point of picking these data sources was to combine both direct and indirect signs of economic activity. Because agriculture is the main industry in Pemalang, the vegetation indicators are quite important here.

Table 1. Summary of Data and Sources

No	Variables	Source	Resolution / Scale	Year
1	Primary sector GRDP	Statistics-Indonesia (BPS) Regency Pemalang	Regency	2022
2	Index vegetation (NDVI, EVI, MSAVI, MNDWI, LAI)	Landsat 8/9, MODIS via Google Earth Engine (GEE)	30 m (Landsat), 500 m (MODIS)	2022
3	Administrative boundaries	Geospatial Information Agency (BIG)	Administrative sub-district & village	2022

Research methods

We did the research in four main steps, which are as follows (Nashrudien & Pramono, 2025):

1. Data preprocessing: We put all of the spatial variables into the UTM coordinate system for zone 49S. Then, we put them all on a 500 × 500 m grid. We did normalization to stop bias between variables that was caused by the scale.
2. We downscale the sectoral GRDP data by using the area weighting method to grids. After that, we divide the grid into primary sectors by using the vegetation index.
3. To estimate the model, we use Geographically Weighted Regression (GWR) as a spatial modeling technique. GWR gives us permission to get regression coefficients that change from one place to another. After that, we tested Gaussian kernel functions and both fixed and adaptive bandwidths to find the best setup.
4. To validate whether the model is appropriate or not, we used R^2, MAE, and RMSE to see how well the model worked. We also use satellite images and street view to conclude which areas have high, medium, and low GRDP.

Analysis Method

Figure 1 shows us the workflow for mapping GRDP based on a 500 x 500-meter grid using GRDP downscaling. After that, the GRDP value is estimated for each grid using the GWmodelS software, which was developed by Lu et al. (2023). At this stage, the GWR model is applied to capture spatial variation in the relationship between GRDP values and predictor variables using remote sensing data. A Gaussian kernel function is used to evaluate the best model, which produces the most accurate estimates. The results are as follows: a GRDP estimation map, a local coefficient map, and a local R^2 map that illustrates the model's accuracy at each area in the Pemalang regency.

To downscale GRDP by spatial scale, we need an analytical framework that can handle non-stationary spatial data. This can happen when the correlation between the economic factor and the environmental factor changes from one place to another. Model regression global standard, like Ordinary Least Squares (OLS), assumes that the relationship remains the same over time, but this is not always the case. Every time we use OLS in an area as diverse as the Pemalang agricultural area, big mistakes often happen in estimation results because the model cannot see how processing differs in every area (Oshan et al., 2019; Yu et al., 2024). We made use of Geographically Weighted Regression (GWR) to overcome these challenges that had happened before. This spatial statistical technique permits regression coefficients to differ by location. It works by giving nearby observations distance-decay weights. Because of this, places that are closer to the target point have a bigger effect on the local estimate. This makes GWR a good choice for looking at economic data in the space and for modeling farming activity (Comber et al., 2022; Lu et al., 2023).

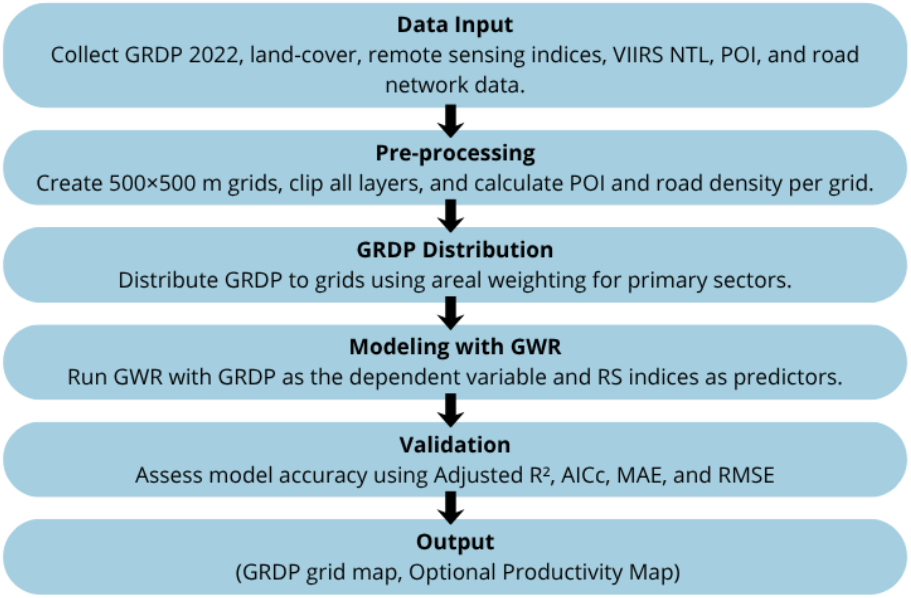

Figure 1. Stages of the process in building a grid-based GRDP map using the Geographically Weighted Regression (GWR) downscaling approach

Figure 1 shows the workflow for mapping GRDP based on a 500 × 500 m grid using the GRDP downscaling approach. The process consists of three main stages: GRDP distribution, GWR modeling, and validation.

GRDP Distribution

We divided the regency-level GRDP by the total number of land cover grids relevant to the primary sector to obtain the GRDP density per grid. Land cover categories that do not belong to the primary sector were excluded and treated as having zero contribution to GRDP. The GRDP density formula is defined as follows (Nashirudien, 2024):

$$GRDP\ Density = \frac{GRDP}{Sum\ of\ Grid} \tag{1}$$

GWR Modeling.

After we fit the model that we made before, we need to check to see how accurate the grid-level GRDP estimates really are. We do this by adding up all the grid GRDP values in a region and then comparing them with the total to the actual recorded value. The GWR downscaling method generally figures out local regression models for small areas, like single grid cells (Chu et al., 2019). This method can be articulated in the following (Xu, Y. et al., 2023).

$$Y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)NDVI_i + \beta_2(u_i, v_i)EVI_i + \beta_3(u_i, v_i)MSAVI_i + \beta_4(u_i, v_i)MNDWI_i + \beta_5(u_i, v_i)LAI_i + \varepsilon_i (2)$$

Information:

- $Y_i$is the value of GRDP at location i;
- $(u_i, v_i)$are the geographic coordinates of location i (e.g., longitude and latitude)
- $NDVI_i, EVI_i, MSAVI_i, MNDWI_i, LAI_i$ Are variables free (index vegetation)
- $\beta_k(u_i, v_i)$is the coefficient of the local regression; its value is different in each location i
- ε_i is an error (residual)

When GWR is used to make the best model, it makes an estimate at each location by giving more weight to observations that are close by than to those that are far away. The bandwidth parameter controls how quickly the drop-off happens, and the kernel function reveals the weights that indicate how much influence reduces with distance. In this study, we used a Gaussian kernel, which is defined by the formula (Gollini et al., 2015).

$$w_{ij} = exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right), \tag{3}$$

This journal article is © Tri Subhi et al. 2026

with wij is point regression; dij is the distance between points; b is the bandwidth; i and j are point observations.

Choosing between the right kernel function and bandwidth is one of the most important things to consider in the process of setting up a GWR model. Other researchers have made a lot of observed that an adaptive bi-square kernel performs effectively when spatial observations are unevenly distributed. Especially when, as is the case with grid-level regional economic data, due to its ability to adjust the spatial extent based on data point density (Comber et al., 2022; Lutfiani et al., 2019). We used the corrected Akaike Information Criterion (AICc) to choose the best bandwidth. This method balances model fit with complexity and helps prevent overfitting (Lu et al., 2023). When compared to global models, GWR calibrated with AICc optimization has always been more accurate (higher R-squared) and had less spatial autocorrelation in the residuals. This holds true for mapping agricultural and economic phenomena at a granular level (Oshan et al., 2019).

Validation

We used three statistical measures to check the validity of the model that was used before. They are mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2). These three metrics are used to see how close the model is to the competency to measure the values against the actual values. As follows, the mathematical formulation of these three metrics is (Xu, Y. et al., 2023):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \widehat{Y}_i\right| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2} \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \tag{6}$$

Result and Discussion

Result

Dependent Variables

Our dependent variable is the GRDP density for the primary sector. To generate this, we clipped land cover data against district/city administrative boundaries. The land cover dataset comes from the United States Geological Survey (Type 1 classification) at 500 m × 500 m resolution. It distinguishes 17 land cover categories, including several types of agricultural vegetation, giving us a useful overview of where farming activity is concentrated. Out of these 17 categories, eight represent the agricultural sector, as listed below in Table 2. Out of the 10 land cover types listed, eight are classified under the primary sector, while Urban and Built-up Lands belong to secondary/tertiary sectors, and Barren land is excluded

Table 2. Grouping of Land Cover Types

Code	Land Cover	Count	Sector
1	Water Body	105	primary
11	Permanent Wetlands	89	primary
9	Savannas	905	primary
14	Cropland/N	1404	primary
12	Croplands	801	primary
10	Grasslands	36	primary
13	Urban and Built-up Lands	430	secondary and tertiary
16	Barren	2	Not used
8	Woody Sava	564	primary
2	Evergreen	549	primary

We then calculated primary-sector GRDP density by combining the land cover information with the official GRDP figures for Pemalang Regency, using the following in equation (1).

Applying this formula, we get the per-grid GRDP value for the primary sector in Pemalang Regency, as presented in Table 3.

Table 3. Primary Sector GRDP Allocation per Spatial Grid Unit in Pemalang Regency

GRDP	GRDP value (million rupiah)	Number of Grids	Value per Grid (million rupiah)
Primary Sector	4,735,800	4,453	1063.507635

The table above shows the primary-sector GRDP per grid cell by showing the GRDP density values of each sector. Then, based on how much relevant land cover is in each cell, these values are given to each grid cell in the primary sector.

Independent Variables

The raw raster data that we obtained could not be directly used in the analysis without processing. The independent variables originated from two primary sources: satellite-based indices derived from Google Earth Engine and other spatial data from OpenStreetMap that we extracted. The data processing workflow made use of the spatial data from Google Earth Engine (GEE) and OpenStreetMap (OSM). GEE raster data were converted to a 500 × 500 meter polygon grid and clipped with district/city administrative boundaries to produce variables such as NDVI, EVI, MSAVI, MNDWI, LAI, and NTL. All of these were then included in the analysis as independent variables once processed. If we see the processed data on maps, each independent variable shows a different pattern in space across Pemalang Regency. We will now go through the descriptive analysis for each variable based on the results that were mapped.

The 2022 NDVI map, as shown in Figure 2, shows that the plants in the regency are very different from each other. The central and southern parts are shown in purple, which means that the NDVI values are high. Based on this picture, this means that the area has a lot of plants and good farmland. On the other hand, the northern and eastern coastal areas are green and yellow, which means that the NDVI values are lower. There is more built-up land and open space, followed by low vegetation in this region. This means that the central and southern parts of Pemalang have the greatest potential for vegetation and the most productive agricultural operations. The northern part of Pemalang Regency, which is a coastal area, is mostly used for non-agricultural purposes.

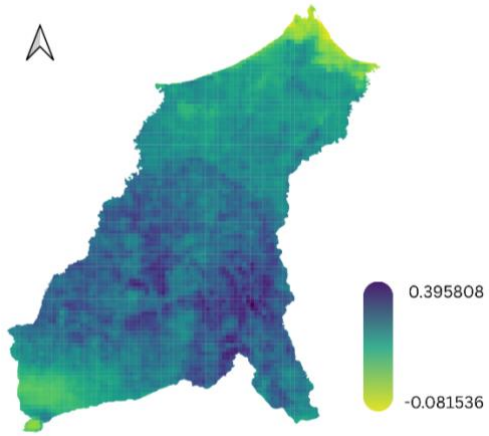

Figure 2. Normalized Difference Vegetation Index (NDVI) map for 2022

Figure 3 presents the MNDWI map, which shows where wetlands and water resources are located in Pemalang. Higher MNDWI values are seen in the northern part of the regency; of course, that's because this area is near the coast and around irrigated rice fields. They look like bluish-green tones. This suggests that there is a lot of farming that needs water, like technical rice paddies, fishponds, and so on. The central and southern areas have lower MNDWI values because they are mostly dryland farming and occupied by residential areas and hilly terrain, and this condition makes sense. That pattern fits very well with the physical characteristics of Pemalang. The northern part depends more on irrigation, and even many pond fish like vannamei and shrimp milkfish exist in the coastal area nearby, while the southern part depends on agriculture, catching rain, namely garden pineapple.

This journal article is © Tri Subhi et al. 2026

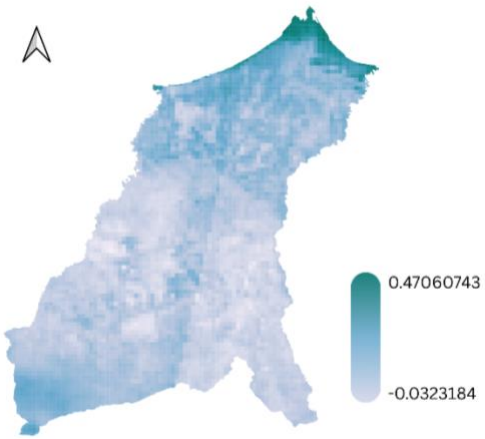

Figure 3. Modified Normalized Difference Water Index (MNDWI) Map of the Regency of Pemalang 2022

As illustrated in Figure 4, EVI Map 2022 is very similar to the NDVI map, but EVI is more useful, showing a bigger difference in area with medium to high vegetation density. The northern part of Pemalang Regency has a big part with a higher EVI value, especially in ricefield irrigation. This shows the productivity of agriculture, which is higher there. The central and southern areas, on the other hand, tend to have lower EVI values. This is because they are more likely to have non-agricultural land use, dryland areas, and hilly areas. These results support the idea that agricultural productivity in Pemalang is not evenly spread out; it is mostly found in the northern lowlands.

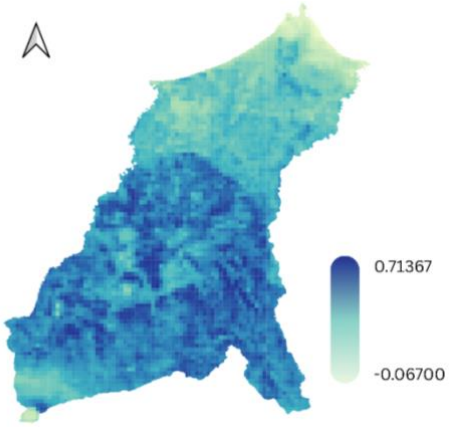

Figure 4. Enhanced Vegetation Index (EVI) Map of the Regency of Pemalang 2022

Figure 5 displays the LAI distribution. The LAI land cover map gives more information about how dense the vegetation has become, especially in Pemalang Regency. In the middle and southern parts of Pemalang, light green and bright yellow tones show us the condition of high LAI values. These areas show us the condition of having a lot of plants and being heavily farmed. The northern and eastern coasts, on the other hand, are dark blue to purple, which means that they have low LAI values because they are mostly residential, open spaces, or sparse vegetation. This means that the central-southern region has the most productive vegetation zones, while the northern coast is under development pressure that is making the vegetation cover less dense. Usually, there are a lot of fish ponds there, and some are affected by flooding and abrasion due to rising sea levels in certain seasons.

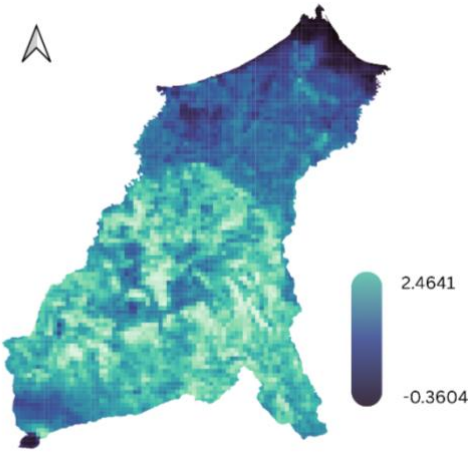

Figure 5. Leaf Area Index (LAI) Map of Regency Pemalang 2022

The MSAVI map in Figure 6 further suggests that inland areas retain considerably more agricultural potential compared to coastal regions. The Modified Soil Adjusted Vegetation Index (MSAVI) map makes it easier to see differences between vegetation in areas with little cover or where vegetation is mixed with bare soil. The northern part of Pemalang has low MSAVI values (shown by light yellow color), which means there are open areas, rice paddies that have just been harvested, or patches of sparse vegetation. The MSAVI values (green) are higher in the central and southern zones, which means that the vegetation cover is more stable in this central area. This shows that inland areas still have more potential for agricultural productivity than coastal regions, which are usually full of fish ponds.

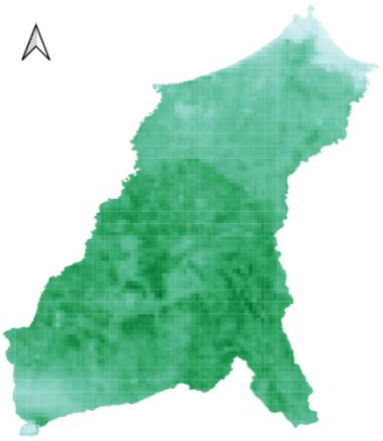

Figure 6. Modified Soil Adjusted Vegetation Index (MSAVI) Map of the Regency Pemalang 2022

Global Regression Analysis (OLS)

Table 4. Analysis Results: Global Regression

Variable	Coefficient (Estimate)	Std. Error	t-value	p-value	Significance
Intercept	1,041.13	1.498	695.01	< 0.001	***
NDVI	5,943.315	177.237	33.53	< 0.001	***
EVI	-2,616,878.858	311,621,751.5	-0.008	0.993	(ns)
MSAVI	-18,366.363	443.391	-41.42	< 0.001	***
MNDWI	-1,858.96	47.245	-39.35	< 0.001	***
LAI	724,583.412	86,130,943.33	0.008	0.993	(ns)
Residual Std. Error					104.7430
Multiple R-squared					0.5392
Adjusted R-squared					0.5386

The global OLS regression result (Table 4) yields an Adjusted R-squared of 0.5387, indicating that approximately 53.9% of

the variation in primary-sector GRDP can be accounted for by the five vegetation indices. Among the predictors, NDVI exhibits the largest positive coefficient (5943.32) and is highly significant ($p < 0.001$), which aligns with the expectation that denser vegetation corresponds to higher agricultural output. MNDWI shows a small but significant ($p < 0.001$) negative coefficient, suggesting that grid cells dominated by standing water tend to have lower primary-sector GRDP, although fishpond areas may behave differently at the local scale. Meanwhile, EVI and LAI display very large standard errors relative to their coefficients and are statistically non-significant, pointing to instability likely driven by multicollinearity among vegetation indices that measure overlapping aspects of canopy condition. The relatively moderate R-squared value, combined with these unstable coefficient estimates, suggests that a single global model is insufficient to represent the spatially varying relationship between vegetation condition and economic output across Pemalang Regency. This finding provides the empirical motivation for adopting a local regression approach. The comparison between OLS and GWR diagnostics (AICc, Adjusted R-squared, and Residual Sum of Squares) presented in the following section serves as the basis for model selection in this study, following the approach recommended by Comber et al. (2022) and Oshan et al. (2019).

By accounting for spatial coordinates, the GWR model significantly improves the overall model fit, increasing the Adjusted R-squared to 0.7722, meaning that 77.22% of the spatial variance in primary-sector GRDP can be explained locally. The core justification for transitioning to GWR lies in its capacity to capture spatial non-stationarity. While the global OLS model assumes a rigid, uniform impact across the entire study area, the local parameter estimations reveal that the relationships between the remote sensing indices (such as NDVI and MNDWI) and GRDP are highly dynamic. The local coefficients fluctuate and even shift directions (from positive to negative relationships) depending on the specific geographic locations within Pemalang Regency. This localized variation provides direct empirical evidence of severe spatial heterogeneity, proving that a single global coefficient oversimplifies the regional economic dynamics and confirming the statistical necessity of the GWR approach.

Geographically Weighted Regression (GWR) Model

Table 5. Results of Applying the Geographically Weighted Regression (GWR) Model

Metric	Value
AICc	128,592.74
Residual Sum of Squares	1,379,122,658,974.07
Adjusted R-square Value	0.77

We used a GWR model to fix these problems. The diagnostics told us a much better story. Based on Table 6, the adjusted R-squared value rose to 0.7722, which means that the spatial model could explain about 77.2% of the differences in primary-sector GRDP at the grid level. And then, the AICc (128,592.74) and the Residual Sum of Squares both went down from OLS, which shows that GWR fits the data better. This gives us solid evidence that the correlation between vegetation and GRDP is only in one place, and not everywhere in the study area. The Adjusted R-square of 0.77, which takes into account the model's local nature and punishes it for having too many parameters, is the best way to judge GWR performance. In earlier GWR applications with strong spatial heterogeneity, similar behavior has been seen (Comber et al., 2022; Oshan et al., 2019).

Although formal spatial diagnostic tests such as Moran's I on OLS residual and Breusch-Pagan heteroscedasticity test were not performed in this study, the substantial improvement in model fit from OLS to GWR provides strong indirect evidence of spatial non-stationarity. The Adjusted R-squared increases from 0.54 to 0.77, the AICc decreases from the OLS baseline, and the reduction in Residual Sum of Squares collectively indicates that the GWR model captures systematic spatial variation that the global model fails to account for. According to Comber et al. (2022), when the improvement in AICc exceeds 3 points and Adjusted R-squared improves substantially, this can be considered sufficient justification for adopting GWR over OLS even in the absence of a formal pre-test, particularly in applied spatial economic studies. Nevertheless, future work should incorporate a formal diagnostic test to further strengthen the methodological rigor.

Furthermore, the variation of local coefficients is interpreted directly through the spatial patterns shown by the coefficient maps. Rather than assuming a uniform relationship across Pemalang Regency, the maps reveal strong spatial heterogeneity and non-stationarity. For instance, the local relationships for primary indicators like NDVI and MNDWI are highly dynamic, with their parameters fluctuating in magnitude and even shifting directions (from positive to negative impacts) across different sub-districts. This spatial variation corresponds to local agricultural conditions, where vegetation density and moisture levels influence economic output differently in coastal versus mountainous zones. Finally, because the ultimate goal of this study is spatial downscaling—disaggregating the regency-level GRDP onto a fine-scale 500 × 500 m grid- the local R^2 analysis demonstrates varying explanatory power across regions. Certain sectors display higher local R^2 values, meaning the remote sensing indices predict agricultural GRDP more accurately in those specific localized contexts. In terms

of research implications, these downscaled maps allow local governments to move away from rigid 'one-size-fits-all' regional planning and instead deploy spatially targeted agricultural interventions that match the exact biophysical and economic characteristics of each grid cell.

Discussions

Distribution of GWR Local Coefficients

Table 6. Distribution of Local GWR Coefficients

Variable	Min	1st Qu	Median	3rd Qu	Max
Intercept	-2685209.51	-22654.12	4338.99	37007.40	1561433.69
NDVI	-49475991.97	-7144028.16	-175491.81	3015511.90	32245997.97
EVI	-129611864.47	-7094650.36	69884714.64	126706541.88	277561438.96
MSAVI	-69437480.45	-7559386.48	877512.03	17487166.23	85295005.95
MNDWI	-13436617.62	-528709.71	-8055.45	636967.06	11731553.11
LAI	-76570777.61	-35898737.57	-19558958.19	1917346.82	38271891.80

Based on Table 6, the GWR coefficient summaries show that different variables have a lot of differences. The NDVI values range from around -49.5 million to +32.2 million, and the median is under zero. In practice, more plants not always mean followed by higher GDP. In conservation forests or places with plants that do not produce anything, the opposite can happen. But areas with very high NDVI coefficients usually is agricultural land that much add the GDP. EVI has a big positive median, which means that in most regencies, better vegetation quality related with stronger primary economic activities. Even though it is like that, the very low value spread can be caused by local differences in the weather, the terrain, or the plant types that grow there. MNDWI interesting because the median little negative but the maximum very high. This shows that an exchange exists between agriculture and fishery in coastal waters. Better water conditions actually can help the fishery economy. LAI has a quite big negative median, that means that dense canopy plants like natural forests do not direct to add primary sector GDP in terms of money, although the ecological value is clear.

The MSAVI coefficients have numbers between -69.4 million to +85.3 million, with a small positive median around 877,512. This positive median shows that in most Pemalang Regency areas, MSAVI, that intended to reduce soil background effects in areas with sparse vegetation, catches agricultural productivity signals that in line with NDVI. The situation of the wide range of coefficients across all five vegetation indices shows that a single global coefficient would be wrong. The GWR framework is needed because each index reacts differently depending upon the land cover, topography, and farming practices in the region.

Distribution of Primary Sector GRDP in the Regency of Pemalang 2022

The GWR map in Figure 7 shows us an uneven spatial pattern clearly. The highest local regression coefficients, in the color red, are in the northeastern and northern parts of Pemalang, while the central and southern parts are lower, in green and yellow. This situation can be concluded to indicate that the primary sector is unbalanced. The north-northeast has fertile valleys and coastal areas with a lot of farmland and water bodies like rivers, irrigation canals, and fishponds. They add a lot more to the primary-sector GRDP (Kusnadi, 2020; Haryanto & Purnomo, 2023); the highlands and more urbanized areas in the south, where there isn't as much fertile land, contribute a lot less.

This journal article is © Tri Subhi et al. 2026

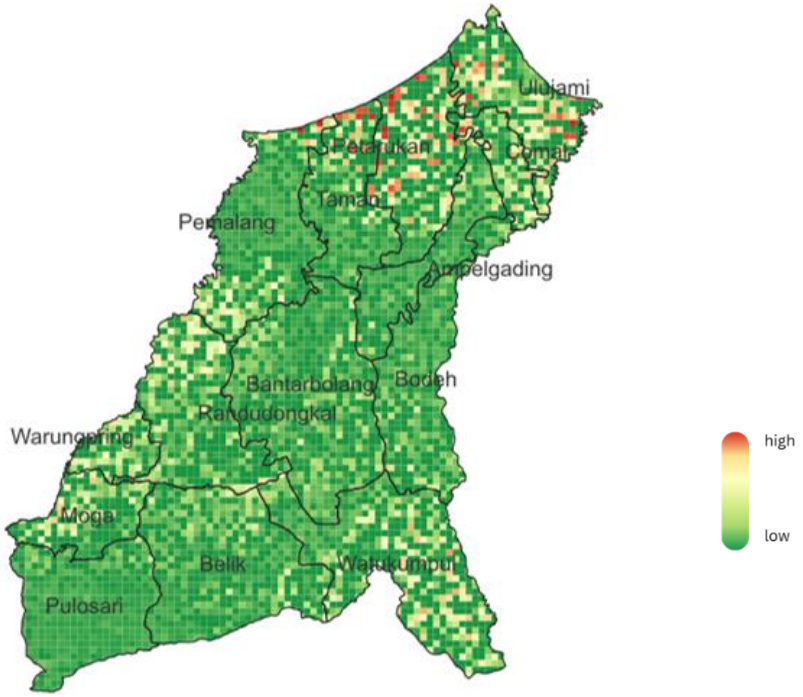

Figure 7. Distribution of Primary Sector GRDP in the Regency of Pemalang 2022

High vegetation indices that are NDVI, EVI, MSAVI, and LAI in the north show that there are a lot of plants and that farming is very productive. The water index represented by MNDWI shows that there are a lot of water resources for irrigation and fish farming. These results are in line with studies from around the world that show a strong link between vegetation and primary economic production. For example, NDVI-crop yield correlations in lowland areas can be as high as 0.7–0.9 (Xia et al., 2022b), and regions with high agricultural economic growth experience an increase in NDVI of around 8.7% (Yue et al., 2020). The red-green pattern on the GWR map shows that areas that consist of lots of plants and water have higher local GRDP, while areas with few plants or difficult terrain (highlands) have lower GRDP from the primary sector. This kind of spatial imbalance is actually pretty common in agricultural development. Access to fertile land and water is often the most important thing for how well agriculture, forestry, and fisheries do economically (Yue et al., 2020; Haryanto & Purnomo, 2023; Kusnadi, 2020)

Spatial and Structural Parameters Weighting

Table 7. Geographically Weighted Regression Bandwidth

Variable	Adaptive	Bandwidth Size	Type	Kernel	Selected	Approach	Distance
Intercept	Fixed	0.34	Fixed	Gaussian	Specified	AIC	CRSDistance
NDVI	Fixed	0.35	Fixed	Gaussian	Specified	AIC	CRSDistance
EVI	Fixed	0.28	Fixed	Gaussian	Specified	AIC	CRSDistance
MSAVI	Fixed	0.31	Fixed	Gaussian	Specified	AIC	CRSDistance
MNDWI	Fixed	0.35	Fixed	Gaussian	Specified	AIC	CRSDistance
LAI	Fixed	0.27	Fixed	Gaussian	Specified	AIC	CRSDistance

The selection of a Gaussian kernel with fixed bandwidth in this study follows the recommendation of Gollini et al. (2015), who note that the Gaussian function provides smooth distance-decay weighting that is well-suited for continuous spatial phenomena such as agricultural productivity. Fixed bandwidth was chosen because the observation grid (500 × 500 m) is regularly spaced, meaning that data density does not vary dramatically across the study area, a condition under which fixed bandwidth tends to perform comparably to adaptive bandwidth (Comber et al., 2022). The optimal bandwidth for each variable was determined by minimizing AICc, which penalizes model complexity more heavily than standard AIC and is therefore more appropriate for GWR calibration with a large sample size (Lu et al., 2023). The resulting bandwidth ranges from 0.27 (LAI) to 0.35 (MNDWI), reflecting the different spatial scale at which each vegetation index relate to economic

output. The narrow bandwidth for LAI and EVI indicates that their relationship with GRDP is highly localized, likely tied to specific crop type and land management practice, whereas the wider bandwidth for MNDWI suggests that water availability influences the economy across broader spatial extents, consistent with the transboundary nature of the hydrological system.

When we look at determining if the downscaling model performs, a GWR has a benefit and clearly does a better job of mapping primary-sector GRDP than traditional global models like OLS. The R-squared values increase rapidly when we let the vegetation index relationships change from one place to another. This matter is in line with recent empirical studies that have been done in various agricultural environments (Han et al., 2024; Oshan et al., 2019; Su et al., 2023), which underline the GWR capability to detect the spatial non-stationarity inherent in agricultural productivity patterns, a feature that is often ignored by global regression models. Our model could reduce estimation biases in the most productive agricultural zones by allowing the relationships between GRDP and variables such as NDVI, EVI, and MNDWI to vary across different areas of Pemalang Regency.

The corrected Akaike Information Criterion (AICc) also shows that GWR is better than OLS in terms of statistics. The AICc value of the GWR model is much lower than that of the OLS model that was constructed before. In spatial econometrics, a decline of over 3 points in AICc they are generally regarded as indicative of a significant enhancement in model performance, achieving a superior equilibrium between data fitting and model simplicity (Li et al., 2021). This big drop in AICc, along with less spatial autocorrelation in the residuals, shows that adding localized distance-decay parameters can help cut down on the over- and under-estimation errors that are common in models used for different types of agricultural landscapes (Rahman & Szabó, 2022; Zhao et al., 2022). Our results back up what other studies have found: employing vegetation indices from different sources in a GWR framework is a truly precise as well as trustworthy way to do high-resolution economic mapping.

Implications of Spatial Primary Sector Productivity

Taking a step back from the specifics, what all of this shows is that the relationship condition between vegetation well-being and primary-sector GRDP isn't the same regardless of location. High NDVI and EVI values may not always indicate high economic productivity. It depends on how the land in one grid cell is actually used for a purpose. These conclusions show how important it is to use spatial methods when performing environmental economic modeling, especially in places with a lot of environmental and economic activity, like coastal areas and farming villages.

Conclusion

This study has demonstrated that the integration of official statistics, official economic data provided by Statistics Indonesia, with satellite-derived vegetation indicators provides a clearer understanding of the variations in primary-sector GRDP across Pemalang Regency. The results show the GWR model is better than the global OLS model, with the Adjusted R^2 increasing from 0.54 to 0.77. This change shows how well the model can catch differences in various places. The GWR coefficients show that the northern and northeastern parts of the Pemalang Regency area, which have much agricultural land and irrigation water, are the places where most primary sector productivity happens. On the other side, the central and southern parts do not give that big contribution because they have more urban land use and less arable land, areas full of mountains. These results show that real spatial differences exist in how well the primary sector works, and that these differences are caused by biophysical conditions, land use patterns, and water availability. The high-resolution spatial data is used not only to make regional GDP estimations more accurate, but also to help society have a better understanding of the environmental economic structure and its connection with agricultural and geographical potentials.

The point is that development policies, especially in regional planning, need to pay more attention to spatial aspects, especially in efforts to close the productivity gap in various Pemalang Regency areas. We suggest that the local government and related agricultural services use these GWR mapping results to focus on productive agricultural land protection in the northern region. They also must take steps to improve dryland agriculture in the southern region. Many actions can be done, for example, build better irrigation systems, plant various types of plants, and offer training programs for farmers. This remote sensing-based spatial method must be more often used in spatial planning and regional economic policies in the development process. Future work can strengthen it by adding socio-economic variables to create a complete and more long-lasting picture of how productivity changes from time to time.

Several limitations should be acknowledged. First, formal spatial diagnostic tests (Moran's I, Breusch-Pagan) on OLS residuals were not conducted, although the substantial improvement in GWR diagnostics over OLS provides indirect justification. Second, local R^2 distribution map and coefficient significance map were not generated due to computational constraints in the original analysis environment. Third, a systematic comparison between Gaussian and bi-square kernel types was not explored. These extensions are strongly recommended for future studies that build upon the framework

This journal article is © Tri Subhi et al. 2026

established here and would further strengthen the validity and interpretability of the spatial downscaling result.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

The author would like to express their deepest gratitude to the supervisors of the Pemalang Regency Regional Planning Studio 1 at the Master of Urban and Regional Planning Program, Universitas Gadjah Mada, for their foundational guidance during this research. Special thanks are also extended to our senior colleagues for their valuable tutorials and technical assistance regarding the spatial data processing and Geographically Weighted Regression (GWR) methodology. Furthermore, sincere gratitude is delivered to the Indonesia Endowment Fund for Education Agency (LPDP) for the scholarship support that enabled the author to pursue postgraduate studies and successfully complete this research.

AI Writing Statement

During the preparation of this work, the author(s) used Grammarly and Google Gemini to improve English language readability, evaluate the manuscript's flow, and refine the structure of the discussion section. After using these tools/services, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication. The data collection, spatial modeling, interpretation of the GWR results, and formulation of scientific conclusions were conducted entirely by the authors.

References

Badan Pusat Statistik Kabupaten Pemalang. (2023). *Kabupaten Pemalang dalam angka 2023*. Retrieved from Badan Pusat Statistik Kabupaten Pemalang website: https://pemalangkab.bps.go.id

Badan Pusat Statistik Provinsi Jawa Tengah. (2023). *Provinsi Jawa Tengah dalam angka 2023*. Semarang: Badan Pusat Statistik Provinsi Jawa Tengah.

Bappeda Kabupaten Pemalang. (2021). *Rencana pembangunan jangka menengah daerah (RPJMD) Kabupaten Pemalang tahun 2021–2026*. Pemalang: Pemerintah Kabupaten Pemalang.

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., & Liu, Y. (2022). County-level GDP spatialization based on nighttime light and NDVI data: A case study of Fujian Province, China. *Sustainability*, *14*(1): 204. https://doi.org/10.3390/su14010204

Chu, H.-J., Yang, C.-H., & Chou, C. C. (2019). Adaptive non-negative geographically weighted regression for population density estimation based on nighttime light. *ISPRS International Journal of Geo-Information*, *8*(1): 26. https://doi.org/10.3390/ijgi8010026

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., ... Harris, P. (2023). A route map for successful applications of geographically weighted regression. *Geographical Analysis*, *55*(1): 155–178. https://doi.org/10.1111/gean.12316

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, *63*(17): 1–50. https://doi.org/10.18637/jss.v063.i17

Han, H., Yang, C., & Song, J. (2024). Mapping rural economic dynamics and agricultural productivity using geographically weighted models and satellite imagery. *Computers and Electronics in Agriculture*, *216*: 108480.

Haryanto, B., & Purnomo, D. (2023). Spatial distribution of agricultural GDP based on land and water resource availability in Central Java. *Forum Geografi*, *37*(1): 1–14.

This journal article is ©Tri Subhi et al. 2026

Kementerian Pertanian. (2022). *Outlook komoditas pertanian sub sektor tanaman pangan*. Jakarta: Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian Republik Indonesia.

Kusnadi, H. (2020). Analisis spasial kontribusi lahan pertanian terhadap PDRB sektor pertanian di Kabupaten Banyumas. *Jurnal Agribisnis Indonesia*, *8*(2): 122–130.

Li, X., Lu, B., & Dong, G. (2021). Comparison of geographically weighted regression and ordinary least squares in spatial downscaling of socioeconomic indicators. *Spatial Statistics*, *43*: 100504. https://doi.org/10.1016/j.spasta.2021.100504

Lu, B., Hu, Y., Yang, D., Liu, Y., Liao, L., Yin, Z., ... Dong, G. (2023). GWmodelS: A software for geographically weighted models. *SoftwareX*, *21*: 101291. https://doi.org/10.1016/j.softx.2022.101291

Lutfiani, N., Sugiman, S., & Mariani, S. (2019). Pemodelan geographically weighted regression (GWR) dengan fungsi pembobot kernel Gaussian dan bi-square. *Unnes Journal of Mathematics*, *8*(1): 1–9. https://doi.org/10.15294/ujm.v8i1.17103

Nashirudien, A. A. (2024). *Representasi nilai produksi ekonomi wilayah berdasarkan data-data spasial*. Universitas Gadjah Mada, Yogyakarta.

Nashrudien, A. A., & Pramono, R. W. D. (2025). Estimasi distribusi PDRB kawasan Kedungsepur secara spasial menggunakan geographically weighted regression untuk mendukung pengembangan wilayah. *Majalah Geografi Indonesia*, *39*(1).

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python implementation of multiscale geographically weighted regression for investigating spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, *8*(6): 269. https://doi.org/10.3390/ijgi8060269

Rahman, M. M., & Szabó, G. (2022). Spatial analysis of agricultural productivity and its determinants using geographically weighted regression. *Geocarto International*, *37*(25): 7415–7435. https://doi.org/10.1080/10106049.2021.1980618

Su, Y., Wang, X., Li, Z., & Chen, H. (2023). Estimating high-resolution agricultural GDP distribution using multi-source remote sensing data and spatial regression models. *International Journal of Applied Earth Observation and Geoinformation*, *118*: 103227. https://doi.org/10.1016/j.jag.2023.103227

Wang, S., Luo, Y., Li, X., Yang, K., Liu, Q., Luo, X., & Li, X. (2021). Downscaling land surface temperature based on non-linear geographically weighted regressive model over urban areas. *Remote Sensing*, *13*(8): 1580. https://doi.org/10.3390/rs13081580

Xia, Y., Chen, Y., Xu, Y., & Zhao, Y. (2022). Mapping the relationship between NDVI and agricultural productivity in floodplain areas using GWR. *Agricultural Systems*, *197*: 103213. https://doi.org/10.1016/j.agsy.2022.103213

Xu, Z., Wang, Y., Sun, G., Chen, Y., Ma, Q., & Zhang, X. (2023). Generating gridded gross domestic product data for China using geographically weighted ensemble learning. *ISPRS International Journal of Geo-Information*, *12*(3): 123. https://doi.org/10.3390/ijgi12030123

Yu, H., Deng, Y., & Li, Y. (2024). A geographically weighted regression approach to explore the spatial agglomeration and heterogeneity in arable land consumption. *Agriculture*, *14*(10): 1675. https://doi.org/10.3390/agriculture14101675

Yue, W., Zhang, H., Liu, Y., & Huang, Y. (2020). Detecting agricultural growth zones using vegetation indices and economic statistics: A spatial econometric approach. *Ecological Indicators*, *115*.

Zhao, M., Cheng, W., Zhou, C., & Li, M. (2022). Spatial downscaling of gross domestic product at the county level using nighttime light and spatial variables: A case study in agricultural regions. *GIScience & Remote Sensing*, *59*(1): 1056–1075. https://doi.org/10.1080/15481603.2022.2085352

This journal article is © Tri Subhi et al. 2026