

Article

# Developing Relative Spatial Poverty Index Using Integrated Remote Sensing and Geospatial Big Data Approach: A Case Study of East Java, Indonesia

Salwa Rizqina Putri <sup>1</sup> , Arie Wahyu Wijayanto <sup>1</sup>  and Anjar Dimara Sakti <sup>2,\*</sup> 

<sup>1</sup> Department of Statistical Computing, Politeknik Statistika STIS, Jakarta 13330, Indonesia; 221810596@stis.ac.id (S.R.P.); ariewahyu@stis.ac.id (A.W.W.)

<sup>2</sup> Remote Sensing and Geographic Information Science Research Group, Faculty of Earth Science and Technology, Institut Teknologi Bandung, Bandung 40132, Indonesia

\* Correspondence: anjar@itb.ac.id

**Abstract:** Poverty data are usually collected through on-the-ground household-based socioeconomic surveys. Unfortunately, data collection with such conventional methods is expensive, laborious, and time-consuming. Additional information that can describe poverty with better granularity in scope and at lower cost, taking less time to update, is needed to address the limitations of the currently existing official poverty data. Numerous studies have suggested that the poverty proxy indicators are related to economic spatial concentration, infrastructure distribution, land cover, air pollution, and accessibility. However, the existing studies that integrate these potentials by utilizing multi-source remote sensing and geospatial big data are still limited, especially for identifying granular poverty in East Java, Indonesia. Through analysis, we found that the variables that represent the poverty of East Java in 2020 are night-time light intensity (NTL), built-up index (BUI), sulfur dioxide (SO<sub>2</sub>), point-of-interest (POI) density, and POI distance. In this study, we built a relative spatial poverty index (*RSPI*) to indicate the spatial poverty distribution at 1.5 km × 1.5 km grids by overlaying those variables, using a multi-scenario weighted sum model. It was found that the use of multi-source remote sensing and big data overlays has good potential to identify poverty using the geographic approach. The obtained *RSPI* is strongly correlated (Pearson correlation coefficient = 0.71 ( $p$ -value =  $5.97 \times 10^{-7}$ ) and Spearman rank correlation coefficient = 0.77 ( $p$ -value =  $1.58 \times 10^{-8}$ ) to the official poverty data, with the best root mean square error (RMSE) of 3.18%. The evaluation of *RSPI* shows that areas with high *RSPI* scores are geographically deprived and tend to be sparsely populated with more inadequate accessibility, and vice versa. The advantage of *RSPI* is that it is better at identifying poverty from a geographical perspective; hence, it can be used to overcome spatial poverty traps.

**Keywords:** spatial poverty index; remote sensing; geospatial big data; zonal statistics; East Java Indonesia



**Citation:** Putri, S.R.; Wijayanto, A.W.; Sakti, A.D. Developing Relative Spatial Poverty Index Using Integrated Remote Sensing and Geospatial Big Data Approach: A Case Study of East Java, Indonesia. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 275. <https://doi.org/10.3390/ijgi11050275>

Academic Editors: Wolfgang Kainz, Nicholas Hamm, Qian Sun and Keone Kelobonye

Received: 23 January 2022

Accepted: 2 April 2022

Published: 24 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Poverty is a historical problem that almost all countries have not been able to clearly solve [1]. According to the UN, approximately 8.2% of people in the world were living in poverty in 2019 [2]. People are considered poor if they live on less than USD 1.90 a day [3]. To overcome this problem, the UN proposed to “end poverty in all its forms everywhere” as the first goal of their Sustainable Development Goals (SDGs); this goal is expected to be achieved by 2030 [4]. Consequently, the reduction of poverty has become a challenging task for all countries, especially for developing or less developed countries [5]. Indonesia is one of the developing countries that is facing poverty. According to Statistics Indonesia, locally known as Badan Pusat Statistik (BPS), there are approximately 10.14% or 27.54 million Indonesian people living in poverty as of March 2021 [6]. To reduce poverty, through the Indonesia SDGs Roadmap Toward 2030 [7], the government has set targets for the poverty rate of Indonesia to drop to 6.5–7% by 2024 and 4.4–5% by 2030.

To reach Indonesia's poverty reduction target, comprehensive poverty monitoring by policymakers plays a significant role. Poverty monitoring in Indonesia is carried out based on official poverty data obtained through the National Socio-Economic Survey (SUSENAS). SUSENAS data collection is conventionally conducted through an on-the-ground household survey that is conducted every six months. Unfortunately, data collection with this conventional method is limited in scope, expensive, laborious, and time-consuming [8]. This deficiency causes limitations in Indonesia's poverty data representation in terms of coverage and time. Data limitation is widely considered one of the main factors for ineffectiveness in providing social assistance [9] according to Indonesia's poverty reduction strategies, based on the 2020–2024 National Mid-Term Development Plan [10]. Therefore, poverty estimation with better granularity in scope and requiring less cost and time to update is needed in order to ameliorate the limitations of the existing household survey-based poverty data collection.

In comparison to conventional socioeconomic survey data, remote-sensing satellite imagery and geospatial big data are valuable resources [11–13], due to their uniqueness and objectivity, by which to observe socio-economic and physical phenomena from a multiple-scale perspective effectively and accurately [14,15]. The advantages of the use of this data are lower costs, frequent updates, and the granularity of the area coverage representation [15,16]. Recent studies have shown that night-time light satellite imagery (NTL) has great potential to capture socio-economic phenomena [17,18], such as gross domestic product (GDP) [19–22], housing vacancy rate [23], electric power consumption [24], population [25], and others [26]. Indirectly, NTL data show the development of human society, which is correlated with economic activity intensity. This data can identify poverty accurately, effectively, and granularly [27]. Several studies have used NTL to identify poverty. Yu et al. [28] reported that NTL data had great potential in identifying poverty. Despite its great potential as one of the best poverty proxy indicators, NTL alone is insufficient as a single data source because poverty is a complex concept and has regional characteristics that are more related to its topography; hence, integration with other geospatial features is required to improve the quality of estimation [29].

As well as night-time satellite imagery, daytime satellite imagery also has the ability to describe socioeconomic conditions. Daytime satellite imagery captures geographical characteristics that indicate the socio-economic conditions of an area. This is proven by the fact that land use or land cover, built-up growth, and accessibility could be used as proxies for certain socioeconomic indicators, such as poverty [30]. Tian et al. [31] concluded that the causes of regional poverty can be structurally different, based on geographical characteristics. Dawson et al. [32] revealed that the NDVI (normalized difference vegetation index) was significantly correlated with poverty, both positive high and/or negative high, in both shrinking and growing countries. The built-up index (BUI) has the ability to map urban areas [33] that are related to regional poverty [34]. Land surface temperature (LST) also has the ability to indicate regional poverty; it tends to be high in low-income and high-poverty areas [35]. Ahmed [36] stated that the urban thermal field variance index (UTFVI), which is created using NDVI, the normalized difference built-up index (NDBI), and LST, is mostly related to the expansion of urban areas. In addition, the normalized difference water index (NDWI) also has the potential to be used as an approach to detect urban areas. Zheng [37] built a vegetation–water-adjusted NTL urban index (VWANUI) that integrated NTL, NDVI, and NDWI, which provides a more accurate and detailed extraction of urban land.

Remote-sensing satellite imagery can also capture air pollution information, which can indicate poverty. According to Baloch et al. [38] and Sakti et al. [39], the increase in poverty has a detrimental effect on environmental pollution. Carbon monoxide (CO) concentration is related to urbanization and economic growth [40,41]. Nitrogen dioxide (NO<sub>2</sub>) is positively correlated to GDP [42]. Sulfur dioxide (SO<sub>2</sub>) is significantly related to population density and energy consumption [43]. Therefore, CO, NO<sub>2</sub>, and SO<sub>2</sub> can be used to map the environmental pollution related to poverty. Moreover, the use of geospatial

big data, such as point-of-interest (POI) data, is another potential option to identify poverty geographically. Shi et al. [14] stated that POI density and POI cost distance can reflect the convenience of human survival and production supporting the degree of regional socio-economic development, which is closely related to poverty.

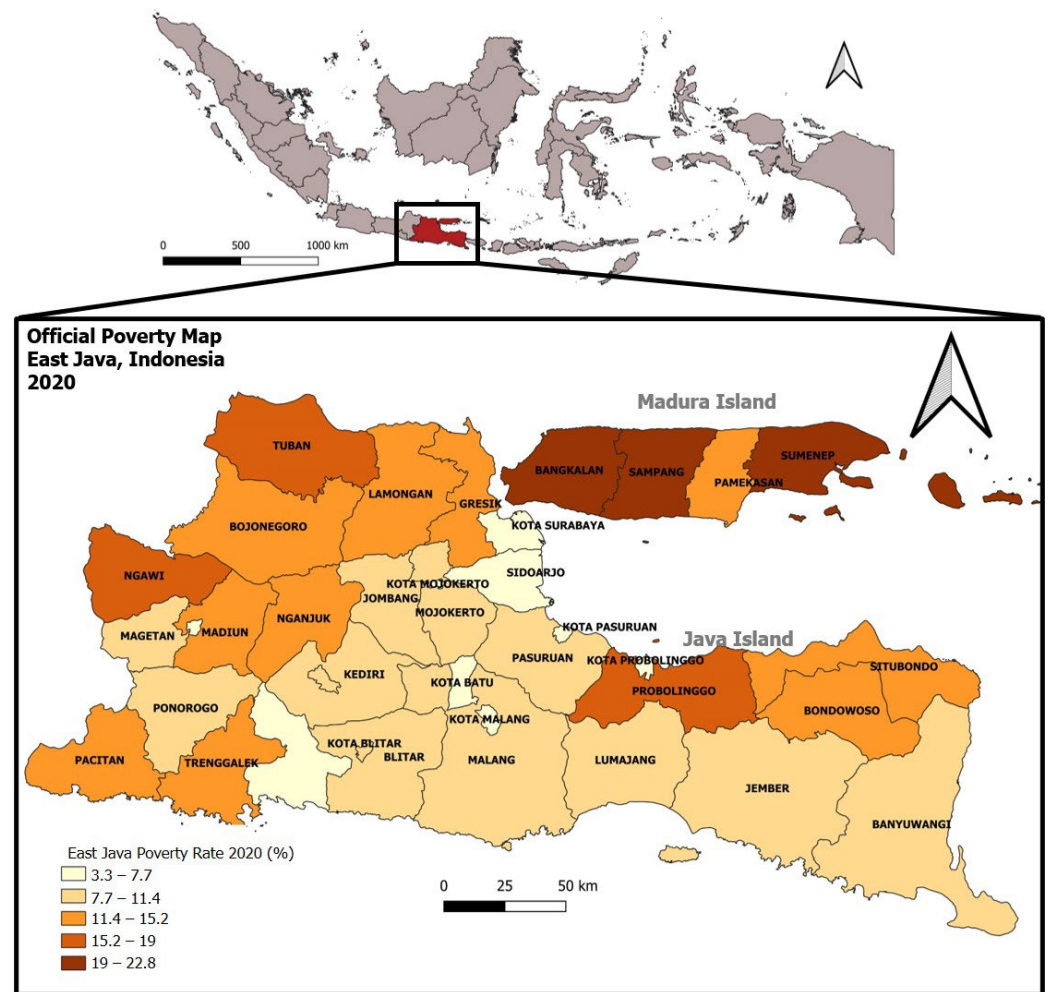
Several studies have identified poverty from various geographic aspects through remote sensing and big data. Duque et al. [44] used high-resolution satellite imagery to produce an intra-urban poverty index using land-cover data. They found that the built-up index can explain up to 59% of the variability in the survey-based slum index. Principal component analysis (PCA) has also been used to summarize the spectral information obtained [45]. Niu et al. [46] developed the multi-source data poverty index (MDPI) based on the characteristics of the built environment obtained from remote sensing satellite imagery, consisting of land cover composition (NDVI, NDBI, and NDWI), NTL, and social conditions obtained from housing rent data. From this study, it was found that there is a high consistency between the MDPI and the official poverty measurement. Shi et al. [14] also developed a poverty index based on remote-sensing satellite imagery and geospatial big data, the comprehensive poverty index (CPI). The CPI was developed by combining NTL data, the digital elevation model (DEM), NDVI, and POI data to map poverty at 500-m spatial resolution. The results suggest that CPI provides a powerful way of identifying poverty distribution.

Multiple sources of remote sensing and geospatial big data have promising potential for capturing the geography of poverty (GOP), which studies poverty from a geographical point of view [47]. To the best of our knowledge, investigation regarding a study identifying the relative spatial poverty of Indonesia that integrates the potential of remote sensing and geospatial big data is still limited, especially in terms of using economic spatial concentration (NTL), infrastructure distribution (BUI), land cover (NDVI and NDWI), air pollution (CO, NO<sub>2</sub>, and SO<sub>2</sub>), and accessibility (POI) as combined poverty proxy indicators. However, the availability of this information can support the limitations of the official poverty data. To answer this problem based on the existing potential, we proposed a new approach to map poverty, with better granularity in scope and at a lower cost. In particular, this study aims to: (1) calculate the relative spatial poverty index (*RSPI*) based on multisource remote sensing and geospatial big data that represent poverty in the case study area, using a median aggregation method at 1.5 km grid level; (2) provide a 1.5-km spatial resolution *RSPI* poverty map; and (3) validate the obtained *RSPI* through numerical and descriptive approaches. We selected East Java, Indonesia as a case study area and focused our research on poverty in 2020 in East Java. In 2020, East Java was the province with the largest number of poor people in Indonesia [48]. The obtained results are expected to provide a lower-cost granular spatial poverty map that needs less time to update, to support the existing official poverty data. Thus, policy decisions are expected to be more effective and efficient as a consequence, so that the poverty reduction target can be achieved.

## 2. Materials and Methods

### 2.1. Study Area

East Java is one of 34 provinces in Indonesia with Kota Surabaya as the capital province. East Java consists of 38 regencies/municipalities. The percentage of East Java in poverty in 2020 has reached 11.09, or approximately 4,419,100 people who are considered to be living in poverty [48]. Figure 1 shows the map of East Java as the case study, along with the distribution of official poverty data at the regency/municipality level in 2020. Official poverty data were calculated, based on a monetary expenditure approach, through the Indonesia National Socio-Economic Survey, SUSENAS [48].



**Figure 1.** East Java, Indonesia as the case study area and its poverty rate (percent).

The poverty of East Java is concentrated in the regencies on Madura Island, the northeast part of the East Java province. Madura Island is lacking in fertile areas and is in the form of a plateau without volcanoes and dry agricultural land [49]. In 2020, the poverty of Sampang Regency was the highest in East Java, reaching 22.78%, followed by Bangkalan (20.56%), and Sumenep (20.18%) [48]. Kota Batu (3.89%) and Kota Malang (4.44%) are the municipalities with the lowest poverty rates in East Java. Both municipalities are well-developed areas with fertile plateaus and tourism areas [50,51]. Kota Surabaya, the largest metropolitan city in East Java and the second-largest in Indonesia, has the fourth-lowest poverty rate, with a value of 5.02% [48,52]. Physiographically, the southern region of East Java is a plateau area with volcanoes in the middle, while in the northwest, there are limestone mountains that are relatively barren.

## 2.2. Data Used in This Study

To identify spatial poverty in East Java, there are two types of geospatial data used in this study. The first type uses remote sensing raster images obtained from multisource satellite images; we used night-time light intensity (NTL) images from NOAA-VIIRS, the normalized difference vegetation index (NDVI), built-up index (BUI), and normalized difference water index (NDWI) from Sentinel-2, land surface temperature (LST) from MODIS, and carbon monoxide (CO), nitrogen dioxide (NO<sub>2</sub>), and sulfur dioxide (SO<sub>2</sub>) from Sentinel-5P. We selected currently publicly available remote-sensing data sources with the highest resolution. Remote-sensing satellite imagery data were collected and preprocessed through the Google Earth Engine, a cloud-based platform designed to store and process

earth data. The Google Earth Engine supported the processing of the satellite image data into classification tasks [53–55]. The area and time scope of the data collection is East Java Province, Indonesia from 1 January 2020 to 31 December 2020.

The second type used is geospatial big data. We used point-of-interest (POI) data obtained from the Open StreetMap. A huge number of point locations of important places are included in the POI data, with 1 July 2020 as the time reference. We filtered several POI categories that are related to poverty according to previous studies [56,57]. Generally, four POI categories were included in this study: education, health, finance, and tourism. There were more than 13,000 points collected, including those representing a hotel, restaurant, hospital, tourist attraction, post office, cafe, school, theatre, mall, university, etc. This data describes the accessibility of an area. The data used in this study are presented systematically in Table 1. From Table 1, we can see that the highest satellite spatial resolution is Sentinel-2 (10 m) and the lowest is Sentinel-5P (1.1 km). To accommodate the lowest resolution, we chose a 1.5 km spatial resolution poverty map as the output of this study. The detailed data pre-processing and calculation of the variables are described in Section 2.3.

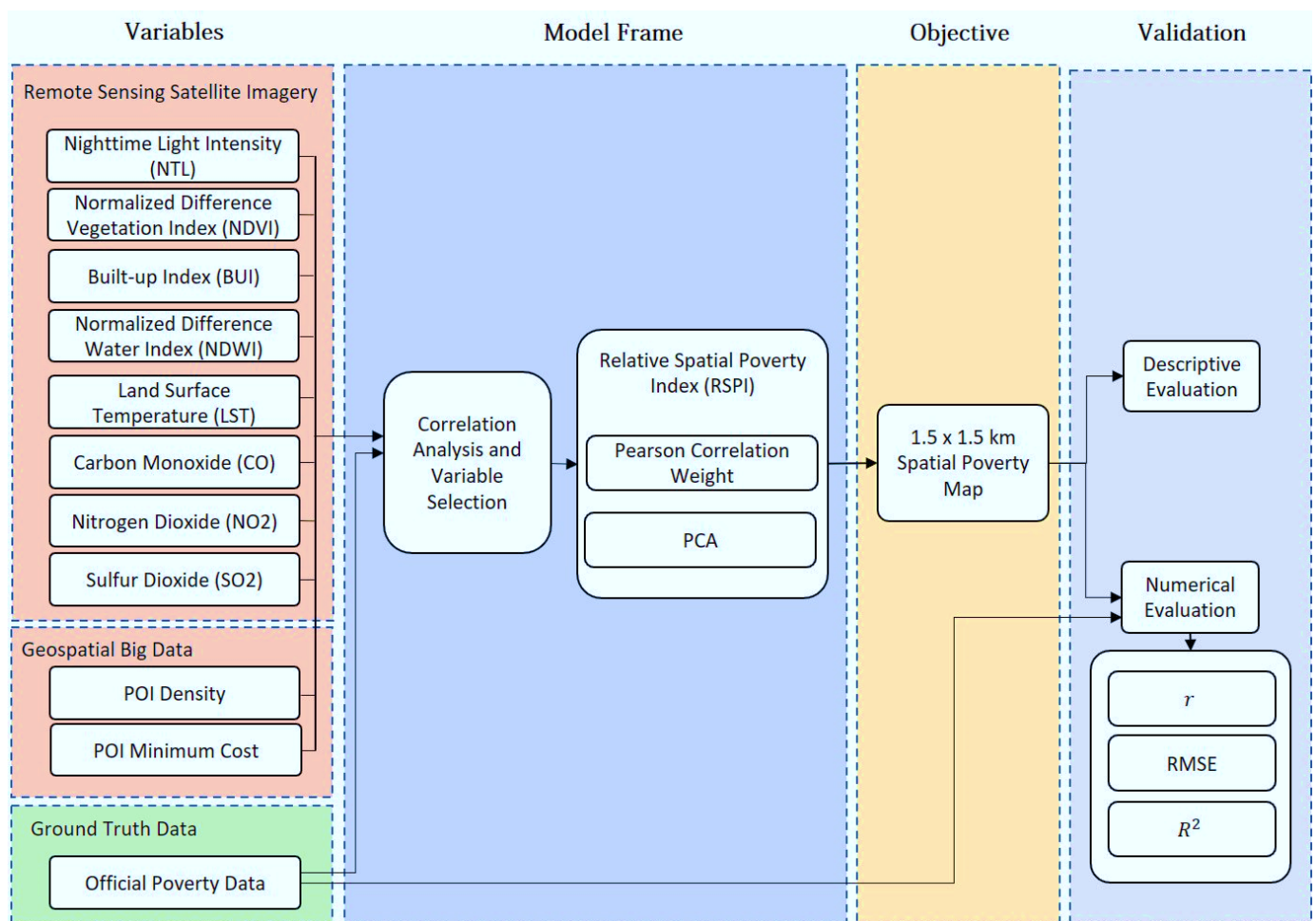
**Table 1.** Summary of data sources and variables.

| Source               | Spatial Resolution | Variable                                           | Band Used                             | Year Data Analysis                                                             | Units                         | Socio-Economic References |
|----------------------|--------------------|----------------------------------------------------|---------------------------------------|--------------------------------------------------------------------------------|-------------------------------|---------------------------|
| NOAA-VIIRS [58]      | 750 m              | Night-time Light Intensity (NTL)                   | avg_rad                               | One-year NTL value                                                             | nanowatts/cm <sup>2</sup> /sr | [17,24–28,59,60]          |
| Sentinel-2 [61]      | 10 m               | Normalized Difference Vegetation Index (NDVI) [62] | B4 (Red) and B8 (NIR)                 | The median value of 2315 cloud masked images.                                  | index                         | [32,36,37]                |
|                      |                    | Built-Up Index (BUI) [63]                          | B4 (Red), B8 (NIR), and B11 (SWIR 1)  |                                                                                | index                         | [33,63]                   |
|                      |                    | Normalized Difference Water Index (NDWI) [64]      | B3 (Green) and B8 (NIR)               |                                                                                | index                         | [36,37]                   |
| MODIS [65]           | 1000 m             | Land Surface Temperature (LST)                     | LST_Day_1 km                          | The median value of 365 cloud-masked images.                                   | Kelvin                        | [35,36]                   |
| Sentinel-5P [66–68]  | 1113.2 m           | Carbon Monoxide (CO)                               | CO column number density              | The median value of the obtained images in 2020.                               | mol/m <sup>2</sup>            | [40]                      |
|                      |                    | Nitrogen Dioxide (NO <sub>2</sub> )                | NO <sub>2</sub> column number density |                                                                                | mol/m <sup>2</sup>            | [42]                      |
|                      |                    | Sulfur Dioxide (SO <sub>2</sub> )                  | SO <sub>2</sub> column number density |                                                                                | mol/m <sup>2</sup>            | [43]                      |
| Open Street Map [69] | -                  | POI Density                                        | -                                     | The number of points in a 1.5 km × 1.5 km grid.                                | points                        | [14]                      |
|                      | -                  | POI Distance                                       | -                                     | The distance from the center of the 1.5 km × 1.5 km grid to the nearest point. | meter                         | [56,57]                   |

For validation purposes, we used the official poverty data at the regency/municipality level, as published by BPS. The data was obtained through the 2020 Indonesia National Socio-Economic Survey, SUSENAS. In this data, poverty is seen as an economic inability to meet basic food and non-food needs, as measured from the expenditure approach. A population is categorized as poor if their average monthly per capita expenditure on basic food and non-food needs is below the poverty line [48].

### 2.3. Methodology

In this study, we built a relative spatial poverty index (*RSPI*) for additional information that describes poverty with better granularity in scope and lower cost, with less time to update. *RSPI* is supposed to enhance the limitations of the existing household survey-based poverty data collection methods. The research framework of this study is schematically illustrated in Figure 2. This study started by collecting and pre-processing data, transforming data, integrating data, performing correlation analysis, and selecting variables, calculating *RSPI*, then validating and interpreting the result. In performing our analysis and visualization, we utilized Python 3.6.9 and QGIS 3.10.4. The expected output was the 1.5 km  $\times$  1.5 km *RSPI* spatial poverty map and its validation. Further detailed explanations are outlined below.



**Figure 2.** The research framework of this study.

### 2.3.1. Data Collection and Pre-Processing

The data collected from the sources that have been described were then pre-processed. Pre-processing data is one of the most important tasks for this index, including preparing the data and converting it to the proper format [70,71]. In this study, data pre-processing mostly aims to clean and improve the quality of the data to be analyzed. Data pre-processing that is performed for remote sensing satellite imagery is different from geospatial big data; the details are explained as follows.

#### Remote Sensing Satellite Imagery Data Pre-Processing

The remote-sensing satellite imagery data used in this study is a collection of the images from 1 January 2020 to 31 December 2020. The obtained images are then pre-processed through four stages: cloud selection, cloud masking, median reducing, and band compositing. Cloud selection and cloud masking of Sentinel-2 and MODIS satellite imagery is performed, based on the quality assessment described in Table 2.

**Table 2.** Sentinel-2 and MODIS cloud selection and cloud masking.

| Satellite Data Source | Band   | Value                                                            |
|-----------------------|--------|------------------------------------------------------------------|
| Sentinel-2            | QA60   | Bit 10 (opaque clouds) equal 0<br>Bit 11 (cirrus clouds) equal 0 |
| MODIS Terra LST Daily | QC_Day | Bits 0–1 (Mandatory QA flags) equal 0                            |

The collected NTL data are the composite image that has been corrected for cloud cover using the VIIRS Cloud Mask (VCM) product. Median reducing is then performed to obtain one value for each observation that represents the satellite images from one year. To get NDVI, BUI, and NDWI values from Sentinel-2 images, band compositing is conducted using the following formulas [62–64,72]:

$$NDVI = \frac{NIR_{band\ 8} - Red_{band\ 4}}{NIR_{band\ 8} + Red_{band\ 4}} \quad (1)$$

$$NDBI = \frac{SWIR_{band\ 1} - NIR_{band\ 8}}{SWIR_{band\ 1} + NIR_{band\ 8}} \quad (2)$$

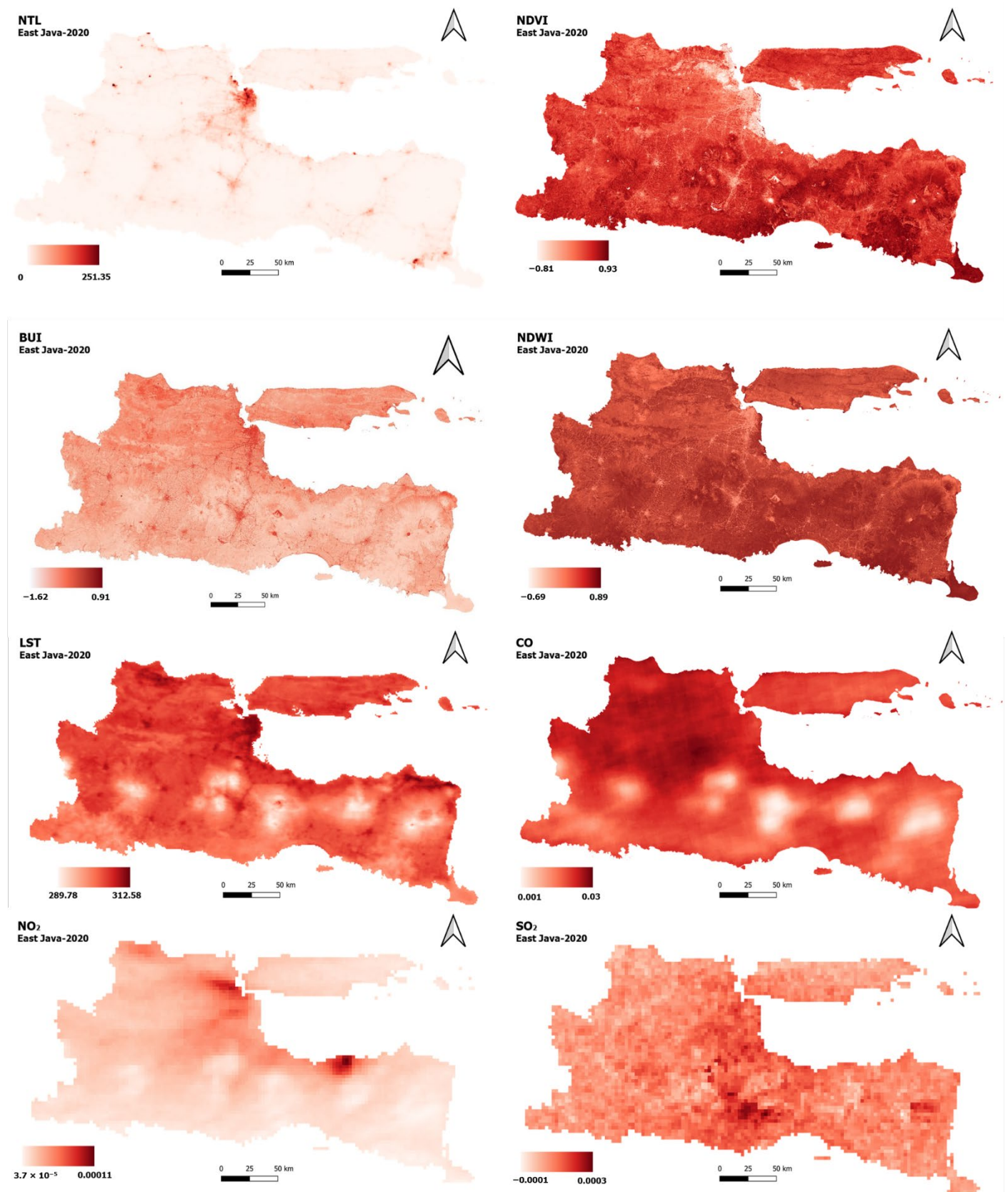
$$BUI = NDBI - NDVI \quad (3)$$

$$NDWI = \frac{Green_{band\ 3} - NIR_{band\ 8}}{Green_{band\ 3} + NIR_{band\ 8}} \quad (4)$$

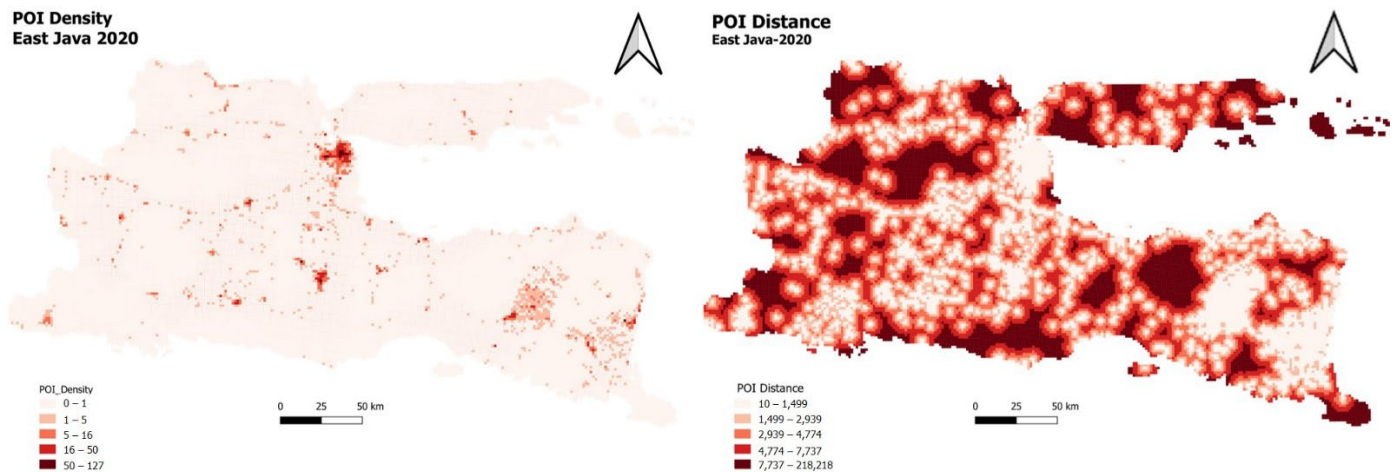
The obtained remote sensing satellite imagery data are shown in Figure 3.

#### Geospatial Big Data Pre-Processing

The point-of-interest (POI) data used in this study are in the form of vector data that contain points. The main pre-processing method performed on POI data is calculating POI density and POI distance. To calculate the relative spatial poverty index (RSPi) with 1.5 km × 1.5 km spatial resolution, the calculation of POI density and POI distance are done to fill the value on each grid. POI density is defined as the number of points in a 1.5 km × 1.5 km grid. POI distance is defined as the minimum distance from 1.5 km × 1.5 km grid center to the nearest POI calculated using the Euclidean distance approach. We calculate POI distance in meter units. The obtained remote sensing satellite imagery data are shown in Figure 4.



**Figure 3.** Visualization of the obtained data from NTL (nanowatts/cm<sup>2</sup>/sr), NDVI (index), BUI (index), NDWI (index), LST (Kelvin), CO (mol/m<sup>2</sup>), NO<sub>2</sub> (mol/m<sup>2</sup>), and SO<sub>2</sub> (mol/m<sup>2</sup>).



**Figure 4.** Visualization of the obtained data POI density (points), and POI distance (meters).

### 2.3.2. Data Transformation

There are two main objectives of data transformation application in this study, firstly, getting values with a similar range with the aim of no variable dominating other variables, and secondly, accommodating heteroscedasticity in the data so that better analysis could be obtained. We applied the Yeo-Johnson power transformation to achieve the determined objectives. Yeo-Johnson transformation is a form of Box-Cox transformation that can deal with both positive and negative data; this transformation can be applied to handle the variability of variables that are unequal across the range by making it more Gaussian-like [73]. This data transformation is defined as follows:

$$h_{\lambda}(x) = \begin{cases} \frac{(1+x)^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \text{ and } x \geq 0 \\ \log(1+x) & \text{if } \lambda = 0 \text{ and } x \geq 0 \\ -\frac{(1-x)^{2-\lambda}}{2-\lambda} & \text{if } \lambda \neq 2 \text{ and } x < 0 \\ -\log(1-x) & \text{if } \lambda = 2 \text{ and } x < 0 \end{cases} \quad (5)$$

where  $x$  is the input data and  $\lambda$  is the parameter.

### 2.3.3. Data Integration

To calculate the relative spatial poverty index (*RSPI*) with  $1.5 \text{ km} \times 1.5 \text{ km}$  spatial resolution, we applied aggregation to take the median value of the raster satellite imagery data, based on the  $1.5 \text{ km} \times 1.5 \text{ km}$  grid shapefile, and integrate it with geospatial big data. Besides this, for correlation analysis purposes, we applied zonal statistics to convert raster-based values into administrative-based values by ascertaining the median to make them comparable to the official administrative-based poverty data. The expected final output from this step is a  $1.5 \text{ km} \times 1.5 \text{ km}$  East Java vector grid and administrative-based vector, with the defined attributes shown in Table 1.

### 2.3.4. Correlation Analysis and Variable Selection

In this study, correlation analysis was conducted to determine the relationship between each geospatial variable defined in Table 1 and East Java's official poverty data. We intended to find the relationship at the  $1.5 \text{ km} \times 1.5 \text{ km}$  grid-level but, due to the limitation of official poverty data, we could only measure it at the regency/municipality administrative level. Both Pearson and Spearman correlation analyses were conducted to determine the

relationship. The equation below shows the formula to establish the Pearson correlation coefficient ( $r$ ):

$$r_{xy} = \frac{n \sum_1^n x_i y_i - (\sum_1^n x_i)(\sum_1^n y_i)}{\sqrt{(n \sum_1^n x_i^2 - (\sum_1^n x_i)^2)} \sqrt{(n \sum_1^n y_i^2 - (\sum_1^n y_i)^2)}} \quad (6)$$

where  $r_{xy}$  represents the correlation between  $x_i$  as the first feature,  $y_i$  is the second feature, and  $n$  is the number of observations. The Spearman correlation is calculated in the same way by changing the observation value to its ranking value. The correlation coefficient ( $r$ ) ranges from 0 to 1. The direction of the relationship is indicated by a positive or negative sign. Table 3 shows the guidelines for interpreting the results of the correlation coefficient ( $r$ ) according to Sugiyono [74]. The correlation significance test is then carried out to determine whether the correlation coefficient obtained was statistically significant at the  $\alpha$  (significance level) of 0.05. The defined null hypothesis is that there is no correlation between the two variables, while the alternative hypothesis is defined as assuming that there is a correlation between the two variables.

**Table 3.** Correlation coefficient interpretation adapted from [74].

| Correlation Coefficient    | Interpretation |
|----------------------------|----------------|
| $0.00 \leq  r  \leq 0.199$ | Very Weak      |
| $0.2 \leq  r  \leq 0.399$  | Weak           |
| $0.4 \leq  r  \leq 0.599$  | Moderate       |
| $0.6 \leq  r  \leq 0.799$  | Strong         |
| $0.8 \leq  r  \leq 1.000$  | Very Strong    |

Although several studies have shown that geospatial variables are related to poverty, Wang et al. [75] stated that regional poverty may be varied in terms of spatial differences. From this correlation analysis, we selected variables that are statistically significant and correlated to the East Java official poverty data, according to the hypothesis testing. The selected variables should also be moderately, strongly, or very strongly ( $|r| \geq 0.4$ ) correlated with the East Java official poverty data, according to its correlation coefficient measure. Therefore, the *RSPI* is built based on variables that linearly represent poverty in East Java.

### 2.3.5. Relative Spatial Poverty Index (RSPI) Calculation

The relative spatial poverty index (*RSPI*) is calculated by overlaying selected geospatial variables that represent poverty in East Java. In order to overlay the variables, we implemented a weighted sum model. Several previous studies have implemented this method for constructing a geospatial index [76–78]. The following formula shows the application of the weighted sum model for *RSPI* construction.

$$RSPI = \sum_{i=1}^p w_i x_i \quad (7)$$

where  $p$  is the number of overlaid variables used,  $w$  is the assigned weight, and  $x$  is the observed value.

We used two approaches when performing the weight calculation. First, we established the correlation-based weight; correlation coefficient information was used as the weights. Variables with higher correlations are assumed to represent poverty better. Second, we used the PCA-based weight. These weights are calculated via the first principal component, established using the principal component analysis (PCA) method. Several studies have used PCA as an approach to calculating the socioeconomic index. Uddin et al. [79] used PCA for mapping the socio-economic vulnerability of the coastal region. It was found

that PCA is a very useful method for identifying vulnerable areas in the coastal region of Bangladesh. Cartone and Postiglione [80] used PCA to build a spatial deprivation index.

### 2.3.6. Validation Assessment

Validation assessment is an important way to establish how far *RSPI* could describe poverty in East Java. We used two validation assessment approaches. The first method, numerical evaluation, is an evaluation to measure the similarity between the obtained result and the available ground-truth data numerically. In this evaluation, we calculated the Pearson and Spearman correlation, root mean square error (RMSE), and  $R^2$ . The following formulas show the calculation of RMSE and  $R^2$ :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (9)$$

where  $y_i$  is the true value,  $\hat{y}_i$  is the predicted value,  $\bar{y}_i$  is the mean of true value, and  $n$  is the number of observations. It is not possible to evaluate each pixel due to the limitations of ground-truth data availability. Therefore, we aggregated the pixel values for each regency/municipality by calculating the mean to create administrative-based data for comparison.

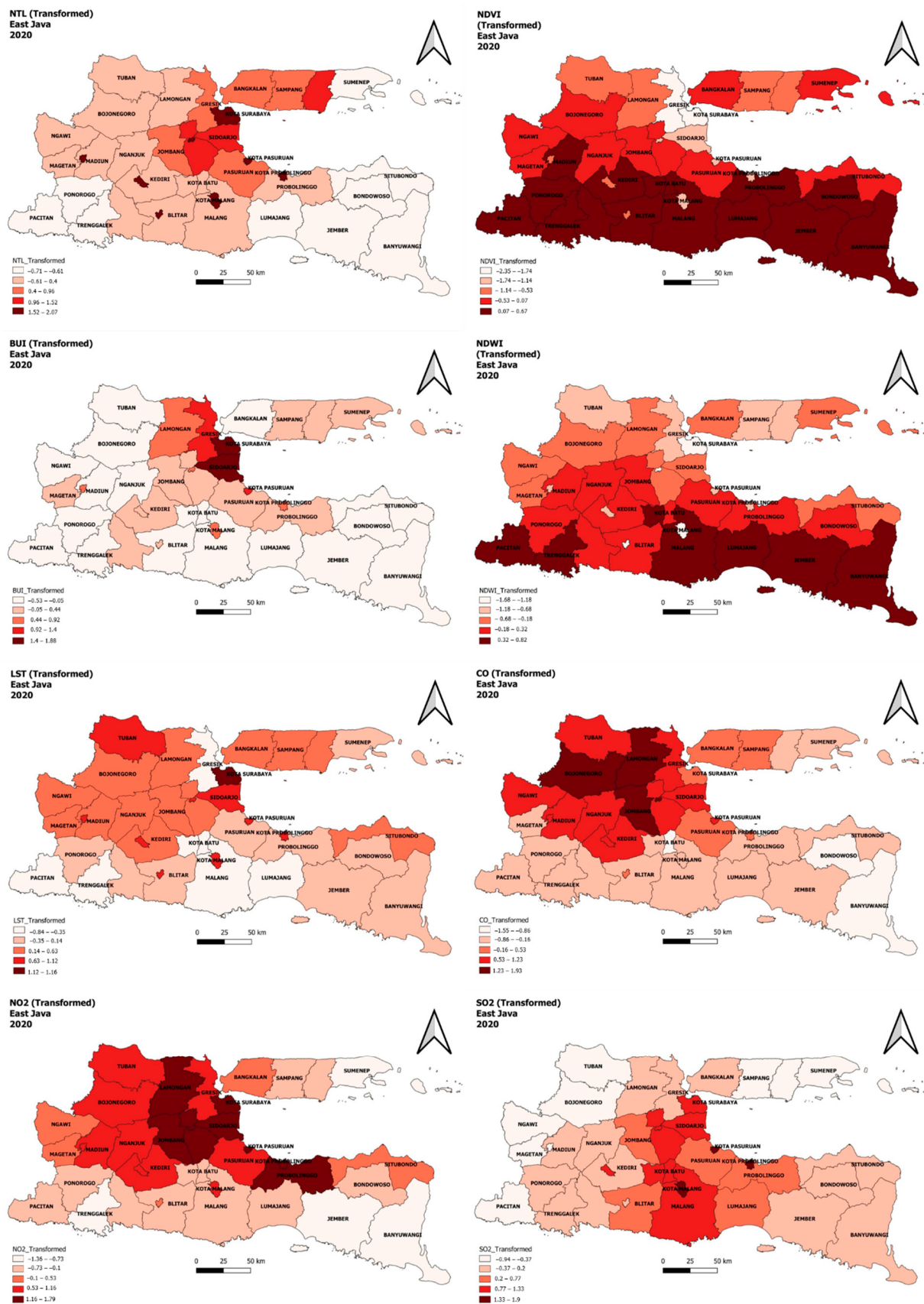
Second, in terms of descriptive evaluation, we visually compared the obtained result with the ground-truth data. Several previous studies have shown that ground-truth identification through high-resolution imagery can offer an evaluation option that cannot be performed numerically on each pixel. For example, Varshney et al. [81] estimated the percentage of roof material through Google Earth images, and Shi et al. [14] randomly chose six points to be identified via high-resolution satellite imagery to recognize poverty through the slope of the land. In this study, we randomly picked six  $1.5 \text{ km} \times 1.5 \text{ km}$  pixels and identified their geographic characteristic areas with the high-resolution Google Earth satellite.

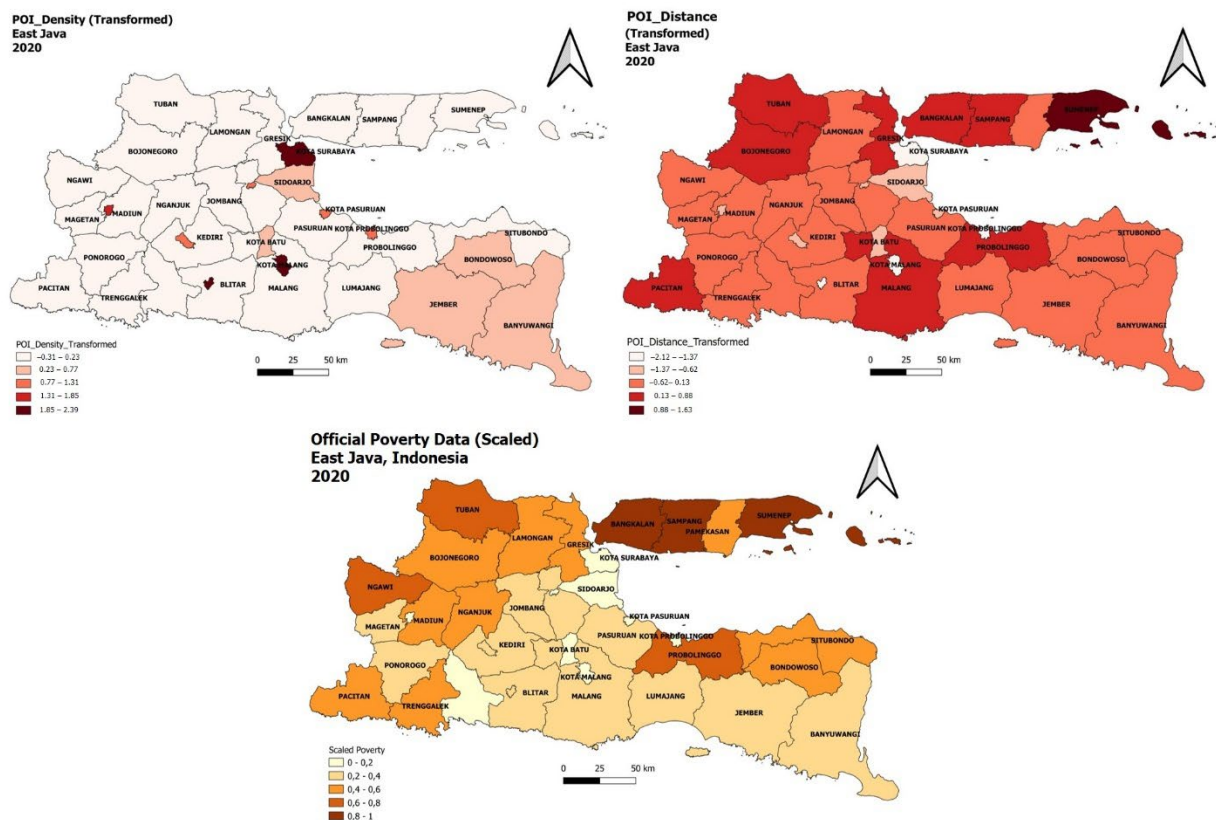
## 3. Results

### 3.1. Correlation Model Development

In this study, correlation analysis was carried out to determine the closeness and direction of the relationship between each geospatial variable and official poverty data. Correlation analysis is carried out after the data has been pre-processed and transformed. Due to the limitations of official poverty data that are only available at the regency/municipality level, we aggregated the pixel-sized geospatial variable data by taking the median value for each regency/municipality. Therefore, 38 observations were obtained for each geospatial variable.

Figures 5 and 6 show the visualization of the geospatial variable maps, along with the official poverty maps, to compare them at the regency/municipality level. It can be seen that each geospatial variable describes a different spatial pattern for each region. The NTL variable, which is a proxy for economic activity, shows that high scores tend to be scattered in municipalities or urban areas with low poverty rates. The variables obtained from the Sentinel-2 satellite imagery (NDVI, BUI, and NDWI), which represent land cover, indicate that there is a homogeneous pattern in the southern regions of East Java, which comprises areas with middle–low poverty rates.





**Figure 6.** Transformed geospatial variables: POI density (points), and POI distance (meter) and scaled official poverty data at the regency/municipality level.

LST variable values tend to be high in municipalities or urban areas with a low poverty rate. Each variable of air pollution ( $\text{CO}$ ,  $\text{NO}_2$ , and  $\text{SO}_2$ ) obtained from Sentinel-5P satellite imagery shows different patterns.  $\text{CO}$  variable values are high in the northwest of East Java, which is an area with middle–high poverty rates,  $\text{NO}_2$  variable values are high in industrial areas with different poverty rates, while  $\text{SO}_2$  variable values tend to be high in the central area, which is a densely populated mountainous area with relatively a low poverty rate. Accessibility variables (POI density and POI distance) capture the pattern of municipalities or urban areas with high accessibility and low poverty rates.

To achieve a better understanding of the relationship between geospatial variables and the official poverty at the regency/municipality level, we applied correlation analysis. Through this analysis, we acquired  $p$ -values to show the statistically significant correlated variables and correlation coefficients that represent the closeness and direction of the relationship. The correlation coefficient was calculated using two approaches, namely, the Pearson and Spearman rank correlations. Table 4 shows the obtained correlation analysis results. Interpretation of the closeness and direction of the relationship was performed, following the method used by the authors of [74]. As we can see in Table 4, the variables with a positive direction relationship are NDVI, NDWI, and POI distance. This shows that the increment value of these variables will be in line with the increment of the official poverty rate value. Conversely, the variables with a negative direction relationship are NTL, BUI, LST,  $\text{CO}$ ,  $\text{NO}_2$ ,  $\text{SO}_2$ , and POI density. This shows that the increment value of these variables will be in line with the decrement of the official poverty rate value.

From Table 4, it can be seen that there are five variables that are statistically significant when correlated to the official poverty rate: NTL, BUI,  $\text{SO}_2$ , POI density, and POI distance. The  $\text{SO}_2$ , POI density, and POI distance variables are strongly correlated with the official poverty rate, while NTL and BUI are moderately correlated, and the rest are weakly or very weakly correlated. To ensure that the relative spatial poverty index (RSPI) can be

used to represent poverty in East Java, we chose variables that are significantly correlated ( $p$ -value  $< 0.05$ , number of observations = 38) and at least moderately correlated ( $|r| \geq 0.4$ ) to the official poverty rate. Therefore, five variables were selected to calculate the East Java *RSPI*: NTL, BUI, SO<sub>2</sub>, POI density, and POI distance.

**Table 4.** The correlation analysis between geospatial variables and the official poverty rate at the regency/municipality level.

| Variable        | Pearson Correlation     |                       | Spearman Rank Correlation |                       | Closeness | Direction | Statistically Significant |
|-----------------|-------------------------|-----------------------|---------------------------|-----------------------|-----------|-----------|---------------------------|
|                 | Correlation Coefficient | $p$ -Value            | Correlation Coefficient   | $p$ -Value            |           |           |                           |
| NTL             | −0.5                    | 0.001                 | −0.49                     | 0.001                 | Moderate  | Negative  | Yes                       |
| NDVI            | 0.25                    | 0.130                 | 0.21                      | 0.205                 | Weak      | Positive  | No                        |
| BUI             | −0.45                   | 0.004                 | −0.44                     | 0.005                 | Moderate  | Negative  | Yes                       |
| NDWI            | 0.14                    | 0.401                 | 0.14                      | 0.401                 | Weak      | Positive  | No                        |
| LST             | −0.29                   | 0.077                 | −0.31                     | 0.058                 | Weak      | Negative  | No                        |
| CO              | −0.065                  | 0.698                 | −0.054                    | 0.747                 | Very Weak | Negative  | No                        |
| NO <sub>2</sub> | −0.25                   | 0.130                 | −0.26                     | 0.114                 | Weak      | Negative  | No                        |
| SO <sub>2</sub> | −0.6                    | $6.84 \times 10^{-5}$ | −0.63                     | $2.25 \times 10^{-5}$ | Strong    | Negative  | Yes                       |
| POI Density     | −0.64                   | $1.51 \times 10^{-5}$ | −0.72                     | $3.48 \times 10^{-7}$ | Strong    | Negative  | Yes                       |
| POI Distance    | 0.73                    | $1.98 \times 10^{-7}$ | 0.79                      | $3.69 \times 10^{-9}$ | Strong    | Positive  | Yes                       |

### 3.2. Relative Spatial Poverty Index Calculation

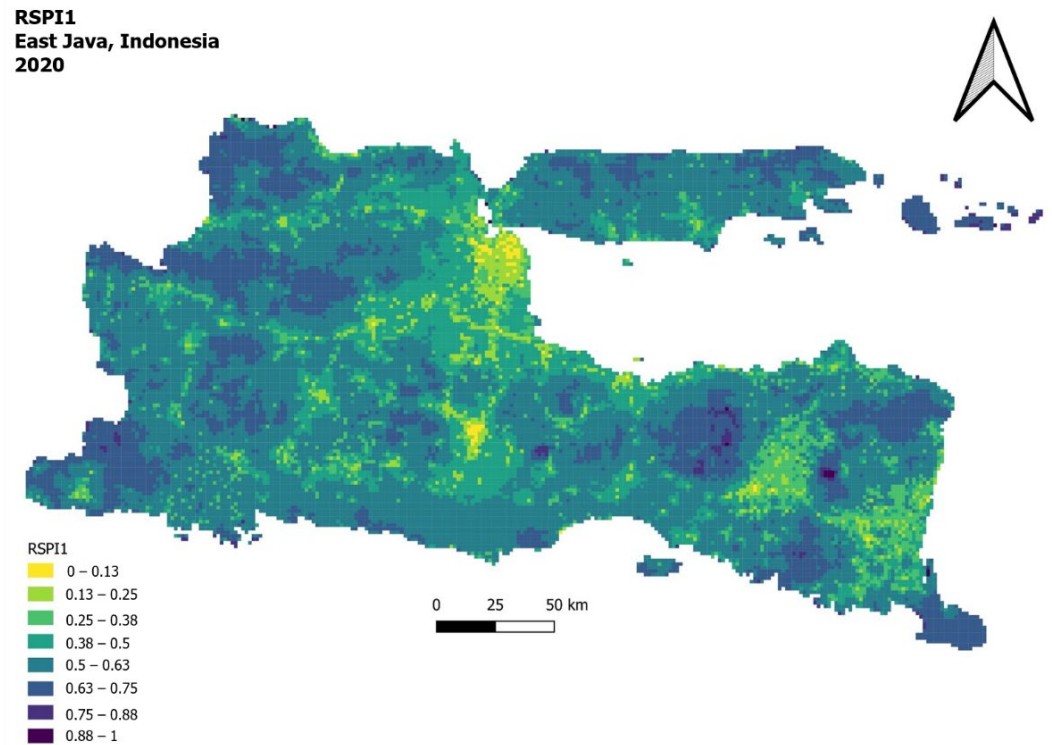
In order to meet the requirement of providing the poverty identification with better granularity in scope and taking the least cost and time to update, and to enhance the limitations of the existing household-based poverty data collection, a relative spatial poverty index (*RSPI*) is proposed by this study. By applying a weighted sum overlay, we calculated the *RSPI* based on variables that were significantly correlated or were at least moderately correlated with East Java official poverty data: NTL, BUI, SO<sub>2</sub>, POI density, and POI distance. Two weight calculation approaches were used in this study: correlation-based weight ( $W_1$ ) and PCA based-weight ( $W_2$ ). The correlation-based weight ( $W_1$ ) was obtained based on the Pearson correlation coefficient. The PCA-based weight ( $W_2$ ) was obtained through the first principal component of PCA. Table 5 shows the derived weight calculations.

**Table 5.** Weight values for the three approaches.

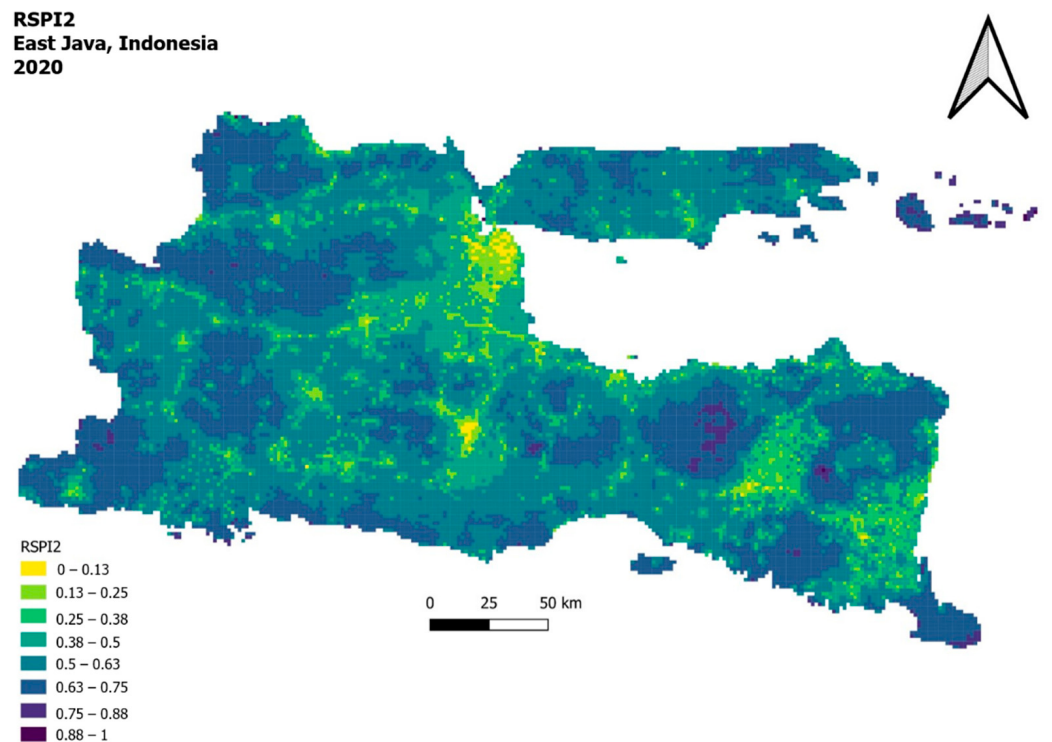
| Variable        | $W_1$ | $W_2$ |
|-----------------|-------|-------|
| NTL             | −0.5  | 0.5   |
| BUI             | −0.45 | 0.26  |
| SO <sub>2</sub> | −0.6  | 0.29  |
| POI Density     | −0.64 | 0.5   |
| POI Distance    | 0.75  | 0.58  |

Of the two types of weighting approaches,  $RSPI_1$  is calculated using  $W_1$  weights and  $RSPI_2$  is calculated using  $W_2$  weights. In this study, both  $RSPI_1$  and  $RSPI_2$  are calculated on a  $1.5 \text{ km} \times 1.5 \text{ km}$  spatial resolution grid to achieve a poverty map, as illustrated in Figures 7 and 8. To simplify the interpretation, we present the min–max-scaled relative spatial poverty map, so that the displayed values are in the range of 0–1. From Figures 7 and 8, it can be seen that the poverty maps generated by  $RSPI_1$  and  $RSPI_2$  have given similar results; low values tend to be concentrated in the central part of East

Java. The southern and northwest parts of East Java, along with Madura Island, tend to have relatively high values. To assess this similarity, the correlation coefficient was calculated; the obtained Pearson correlation coefficient is 0.99 and the obtained Spearman rank correlation is 0.98. This shows that  $RSPI_1$  and  $RSPI_2$  are strongly correlated.



**Figure 7.** The obtained scaled  $RSPI_1$ .



**Figure 8.** The obtained scaled  $RSPI_2$ .

## 4. Discussion

### 4.1. RSPI Numerical Evaluation

The obtained *RSPI* was then validated with two approaches, namely, numerical evaluation and descriptive evaluation. In the numerical evaluation, we focused on calculating how close the *RSPI* values were, numerically, to the official poverty data. The descriptive evaluation will be discussed in the next section. Due to the limitations of the official poverty data, which is only available up to the regency/municipality level, it is not possible to evaluate each pixel. Therefore, we aggregated the obtained *RSPI* pixel values for each regency/municipality by taking the mean value; 38 *RSPI* values were obtained. Numerical evaluation has been performed using two approaches: correlation analysis and *RMSE* calculation. Table 6 shows the obtained correlation analysis.

**Table 6.** The correlation between *RSPI* and the official regency/municipality level poverty rate.

| Index                    | Pearson Correlation     |                       | Spearman Rank Correlation |                       | Closeness | Direction | Statistically Significant |
|--------------------------|-------------------------|-----------------------|---------------------------|-----------------------|-----------|-----------|---------------------------|
|                          | Correlation Coefficient | <i>p</i> -Value       | Correlation Coefficient   | <i>p</i> -Value       |           |           |                           |
| <i>RSPI</i> <sub>1</sub> | 0.71                    | $5.97 \times 10^{-7}$ | 0.77                      | $1.58 \times 10^{-8}$ | Strong    | Positive  | Yes                       |
| <i>RSPI</i> <sub>2</sub> | 0.69                    | $1.64 \times 10^{-6}$ | 0.72                      | $3.48 \times 10^{-7}$ | Strong    | Positive  | Yes                       |

From Table 6, we can see that each *RSPI* is statistically significant (*p*-value < 0.05) correlated to the official poverty data. The Pearson and Spearman rank correlation coefficient shows that *RSPI*<sub>1</sub> and *RSPI*<sub>2</sub> are strongly positively correlated to the official poverty data. The highest correlation coefficient was obtained by *RSPI*<sub>1</sub> (Pearson correlation coefficient = 0.71 (*p*-value =  $5.97 \times 10^{-7}$ ) and the Spearman rank correlation coefficient = 0.77 (*p*-value =  $1.58 \times 10^{-8}$ ) which is calculated using correlation-based weight. The positive direction indicates that the increment value of *RSPI* variables tends to be in line with the increment percentage of official poverty data. We also built a simple linear regression model, with *RSPI* as the independent variable and poverty rate data as the dependent variable. Table 7 shows the obtained model, along with its *RMSE* and *R*<sup>2</sup> value for each *RSPI*.

**Table 7.** Simple linear regression result.

|                       | <i>RSPI</i> <sub>1</sub>      | <i>RSPI</i> <sub>2</sub>      |
|-----------------------|-------------------------------|-------------------------------|
| Model                 | $Pov = 12.603 + 1.7063RSPI_1$ | $Pov = 12.605 + 2.2191RSPI_2$ |
| <i>RMSE</i>           | 3.18%                         | 3.25%                         |
| <i>R</i> <sup>2</sup> | 0.50                          | 0.48                          |

As we can see, the model built by *RSPI*<sub>1</sub> has the lowest *RMSE* value and the highest *R*<sup>2</sup> value. Therefore, we can conclude that *RSPI*<sub>1</sub> is the best index to numerically predict the official poverty data. Figure 9 shows the prediction of the official poverty rate, based on *RSPI*<sub>1</sub> and *RSPI*<sub>2</sub>, and the official poverty rate. It can be seen that the distribution of predictions that are closest to the official poverty is the prediction based on *RSPI*<sub>1</sub>. Hence, we choose *RSPI*<sub>1</sub> as the most representative index of the official poverty data. Figure 9 shows the regression plot of the official poverty rate based on *RSPI*<sub>1</sub> and *RSPI*<sub>2</sub>.

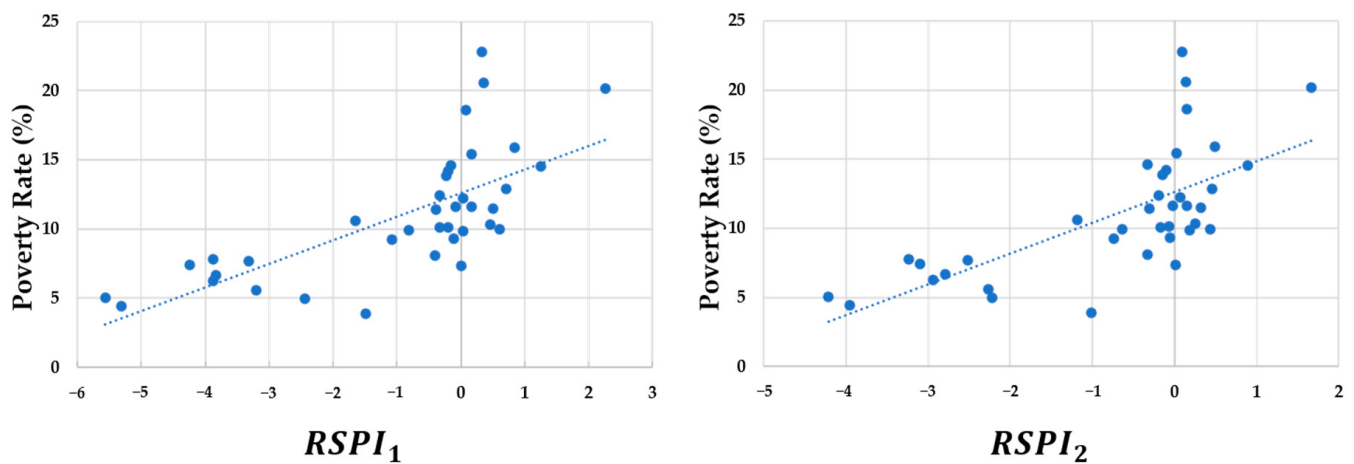


Figure 9. The regression plot referred to in Table 7.

#### 4.2. RSPI Ground Truth Analysis

Many previous studies have shown that ground truth identification through high-resolution imagery can offer an evaluation option that cannot be performed numerically on each pixel. In this study, we randomly picked six  $1.5 \text{ km} \times 1.5 \text{ km}$  pixels and identified their geographic characteristic areas through Google Earth images. The selected *RSPI* plot is *RSPI*<sub>1</sub>, which is the most representative index of the official poverty data. Figure 10 shows the resulting *RSPI* ground-truth check. It can be seen that areas with high *RSPI* scores tend to be sparsely populated areas with inadequate accessibility. This area tends to be a spatially deprived area with limited accessibility. On the other hand, areas with low *RSPI* scores tend to be densely populated areas that have better adequate accessibility. Urban areas tend to have low *RSPI* values.

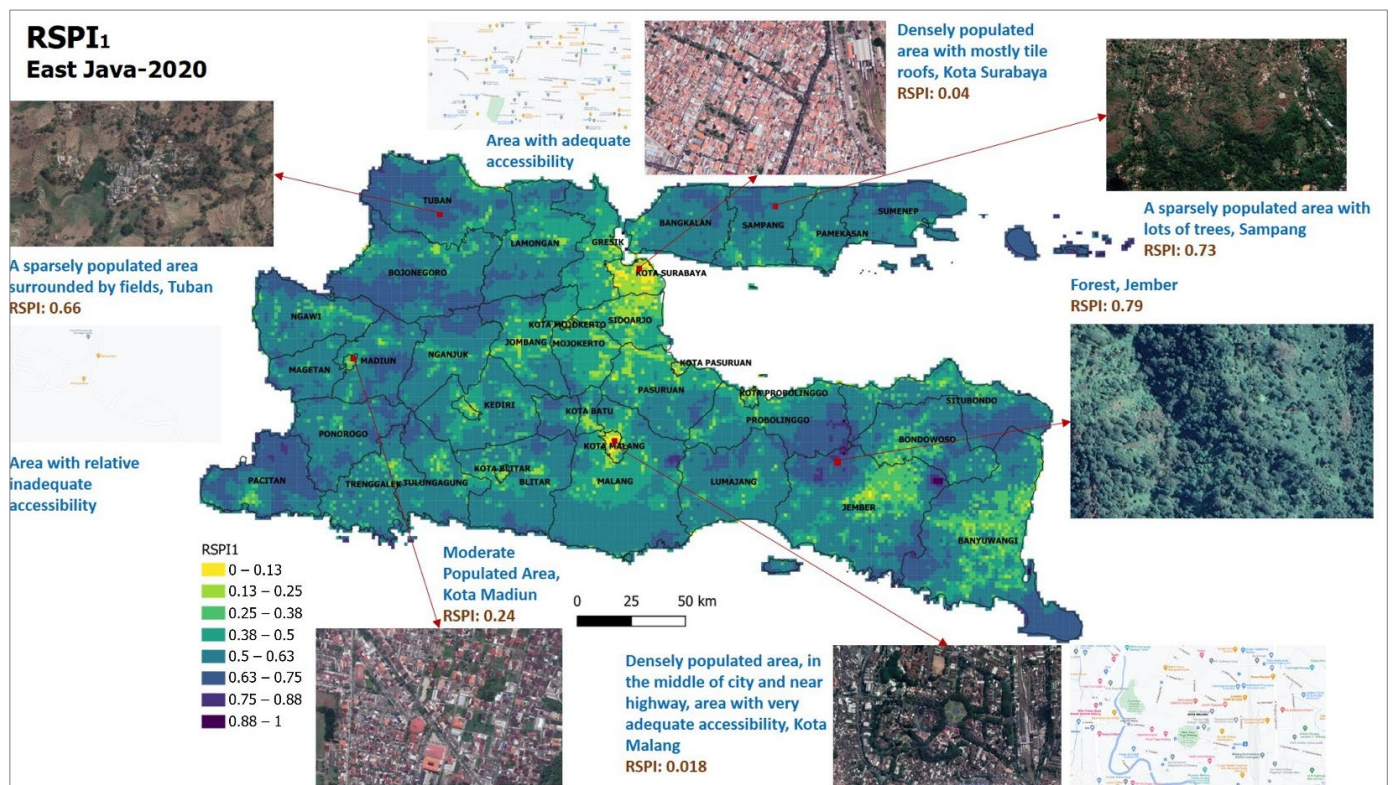
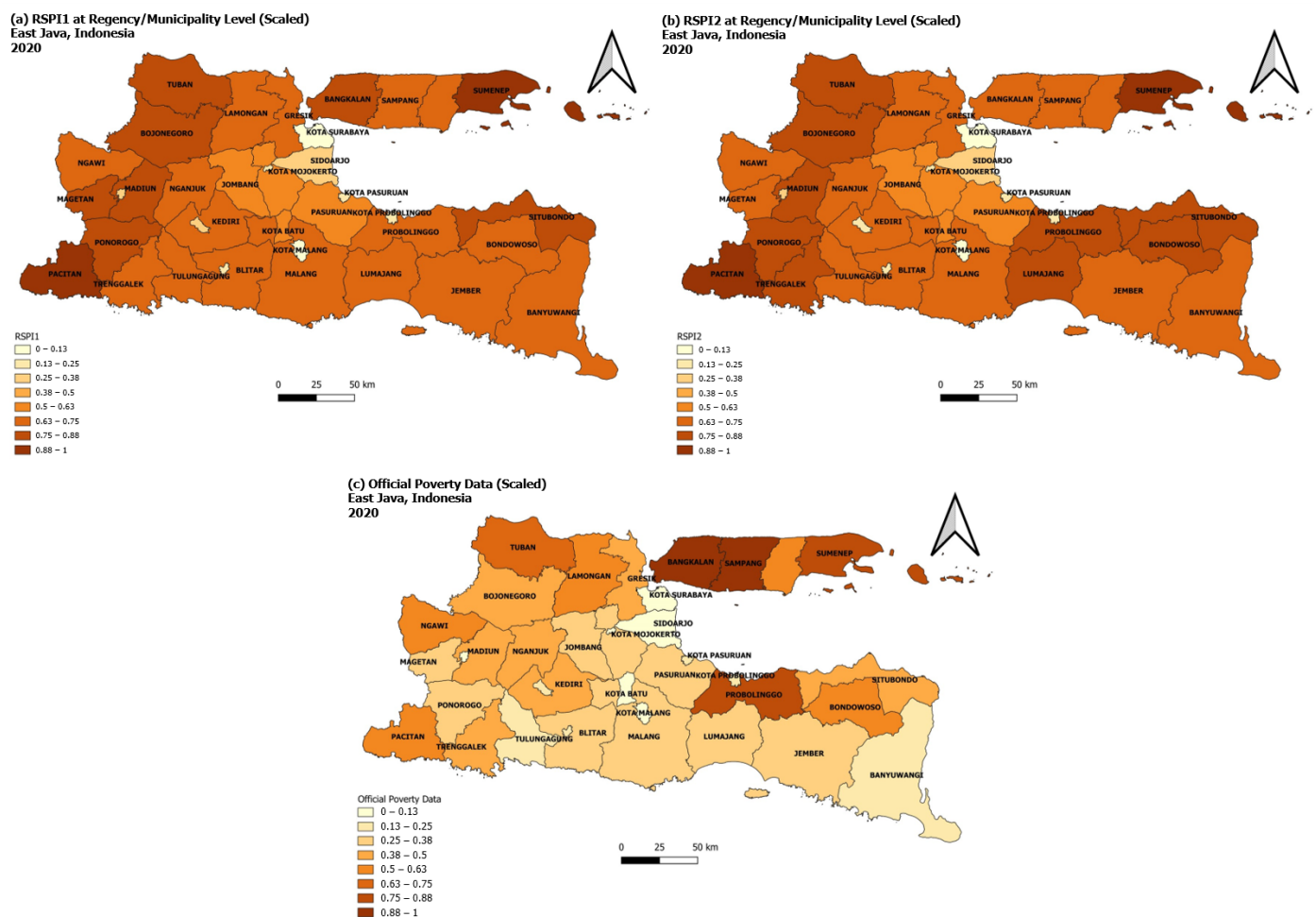


Figure 10. *RSPI* (scaled index) ground-truth check.

#### 4.3. Comparison between the Obtained $RSPI$ and the Official Poverty Data

Although the calculated relative spatial poverty index ( $RSPI$ ) maps poverty from the geographical point of view, we were interested in descriptively comparing the obtained result with the official poverty data, which was calculated using the expenditure approach. Figure 11 shows the comparison between the aggregated regency/municipality-level  $RSPI$  (Figure 11a,b) with the official poverty data (Figure 11c). Aggregation is performed due to the limitations of the official poverty data, which is only available up to the regency/municipality level. It can be seen that  $RSPI_1$  and  $RSPI_2$  present similar results. The areas with the lowest spatial poverty are Kota Surabaya, Kota Malang, and the other urban areas. These areas tend to be non-monetary-deprived areas according to the official poverty data. The southern part of East Java tends to be non-monetary-deprived areas, but it does tend to have high  $RSPI$  scores. Therefore, it can be said that although these areas are less poor (according to the official poverty data), these areas still have the ability to be affected by spatial poverty traps or geographically deprived areas. Moreover, it can be seen that Madura Island's regencies, which are areas with high poverty rates according to the official poverty data, tend to have high  $RSPI$  scores. This indicates that, apart from monetary deprivation, these areas are also spatially deprived.



**Figure 11.** Comparison of (a)  $RSPI_1$  and (b)  $RSPI_2$  (scaled index) and (c) official poverty data (scaled percentage) at regency/municipality level.

#### 4.4. Limitations and Future Possible Directions

Poverty data are usually collected through on-the-ground household-based socioeconomic surveys. Unfortunately, data collection with this method is limited in scope, expensive, laborious, and time-consuming [8]. Indonesia's official poverty data are obtained through the Indonesian National Socio-Economic Survey (SUSENAS), which is conducted every six months at the household level. However, data collection by surveys cannot be separated from the sampling error and non-sampling error [82]. Sampling error is obtained from the use of samples in estimating the population. No matter how many samples are used, there will always be a difference or error between the estimated and the actual value [83]. The non-sampling error is obtained from the inaccuracy of the data collection. Sari et al. [84] revealed that the SUSENAS data collection is still inseparable from human errors, such as inaccuracies during data entry. In addition, the SUSENAS data are only available up to the regency/municipality level, even though decision-making requires more granular data [85]. On the other hand, remote sensing and big data offer datasets that are unlimited, free, easier to obtain and are representative of the population. However, the use of satellite imagery and big data, such as the relative spatial poverty index (*RSPI*), only captures poverty information from the geographical point of view, and there has never been an evaluation of how this data can describe the actual poverty, which is multidimensional. From a comparison of the advantages and disadvantages, it is possible to cover the existing weaknesses with the existing advantages via data integration. In this case, we can obtain monetary poverty information at the regency/municipality level through official poverty data and observe spatial poverty to an accuracy of  $1.5 \text{ km} \times 1.5 \text{ km}$  using *RSPI*.

The integration of the official poverty data and *RSPI* is also useful for describing poverty from another point of view. So far, the calculation of poverty in Indonesia still uses the expenditure approach. A population is categorized as poor if their average monthly per capita expenditure on basic food and non-food needs is below the poverty line [48]. In fact, poverty is a multidimensional problem that cannot only be seen from the point of view of income or expenditure [86,87]. Therefore, studies on measuring poverty from various approaches are still continuously conducted, for example, from the geographic approach. Areas with unique geographic characteristics play an important role in poverty mapping; poor people are found to be more likely to live in certain places [80,88]. The spatial poverty traps are situations where geographical capital, for example, the area's physical nature and socio-political conditions are low and poverty is high, as a result of geographic disadvantage [89]. Spatial poverty traps are areas where people are at higher risk of being trapped in poverty rather than elsewhere, due to spatial disadvantages [90]. In the case of identifying poverty traps, *RSPI* can be a better approach because it is built from various geographic variables.

To sum up, *RSPI* has the ability to capture spatial poverty at a 1.5-km grid level and also has the possibility to be updated every month. The *RSPI* datasets were obtained from free and publicly available data sources, which are faster to access and require fewer human resources. Nevertheless, *RSPI* can only describe poverty from a geographical point of view. Moreover, further and more accurate validation up to a 1.5-km grid level is still needed to evaluate how far the ability of *RSPI* can go in describing poverty, which is a multidimensional problem. Therefore, we suggest the application of data integration between *RSPI* and official poverty data for better poverty identification and monitoring.

In addition, it is possible for *RSPI* calculations to be applied to other regions of Indonesia or even to other countries by paying attention to the spatial characteristics that characterize poverty in a particular area. This is in line with the research of Wang et al. [75], which states that regional poverty can vary according to spatial differences. To ensure that certain geospatial variables can be used to describe poverty in certain areas, further analysis is needed regarding the relationship between the geospatial variables used and poverty in the specific geospatial area.

## 5. Conclusions

To enhance the limitations of the existing household survey-based poverty data collection, this study provides a relative spatial poverty index (*RSPI*) with better granularity in scope, which is less costly and takes less time to update. The *RSPI* calculations utilize the use of multisource remote sensing satellite imagery and geospatial big data. The *RSPI* in the case study area, East Java, Indonesia, is calculated based on geospatial variables that specifically represent poverty in East Java in 2020: night-time light intensity (NTL), the built-up index (BUI), sulfur dioxide (SO<sub>2</sub>), point-of-interest (POI) density data, and POI distance data, which are statistically significant correlated or at least moderately correlated with the official poverty data. These variables are then overlaid by a weighted sum model using two weight calculation approaches: correlation-based weight and PCA-based weight. It was found that the use of multisource remote sensing and geospatial big data has good potential for representing poverty in East Java, Indonesia in 2020. This is evidenced by the strong correlation between the *RSPI* at the regency/municipality level and the official poverty data. The best *RSPI* for representing poverty in East Java 2020, calculated using a correlation-based weight-sum model, is strongly correlated with official poverty data, with a Pearson correlation coefficient of 0.71 ( $p\text{-value} = 5.97 \times 10^{-7}$ ) and a Spearman rank correlation coefficient of 0.77 ( $p\text{-value} = 1.59 \times 10^{-8}$ ). This *RSPI* is also quite promising for use as a predictor variable in the estimation of poverty data. We built a simple linear regression model, estimating the East Java, Indonesia 2020 official poverty rate with *RSPI* as the only predictor variable. The model obtained an RMSE of 3.18%, with an  $R^2$  up to 0.50. *RSPI* is then presented in the form of a non-technical, user-friendly poverty map with a spatial resolution of 1.5 km  $\times$  1.5 km. The results of the descriptive evaluation of this map indicate that areas with high *RSPI* scores tend to be geographically deprived areas that are sparsely populated, with more inadequate accessibility; in contrast, areas with low *RSPI* scores tend to be densely populated areas that have better accessibility. Therefore, the ability of *RSPI* to map spatially deprived areas can be used to support the official poverty data.

**Author Contributions:** Arie Wahyu Wijayanto conceptualized the framework and methodology; Salwa Rizqina Putri and Arie Wahyu Wijayanto designed the experiments; Salwa Rizqina Putri performed the experiments; Salwa Rizqina Putri, Arie Wahyu Wijayanto and Anjar Dimara Sakti analyzed the data; Salwa Rizqina Putri pre-processed the base datasets; all authors wrote the paper; all authors read the paper and provided revision suggestions. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was funded by the Capacity Building Research Program for ITB Young Scientists by the Institute of Research and Community Service, Institut Teknologi Bandung, and the research grant program by the Politeknik Statistika STIS.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors are grateful to acknowledge the support from Statistics Indonesia (BPS), Politeknik Statistika STIS, and Institut Teknologi Bandung.

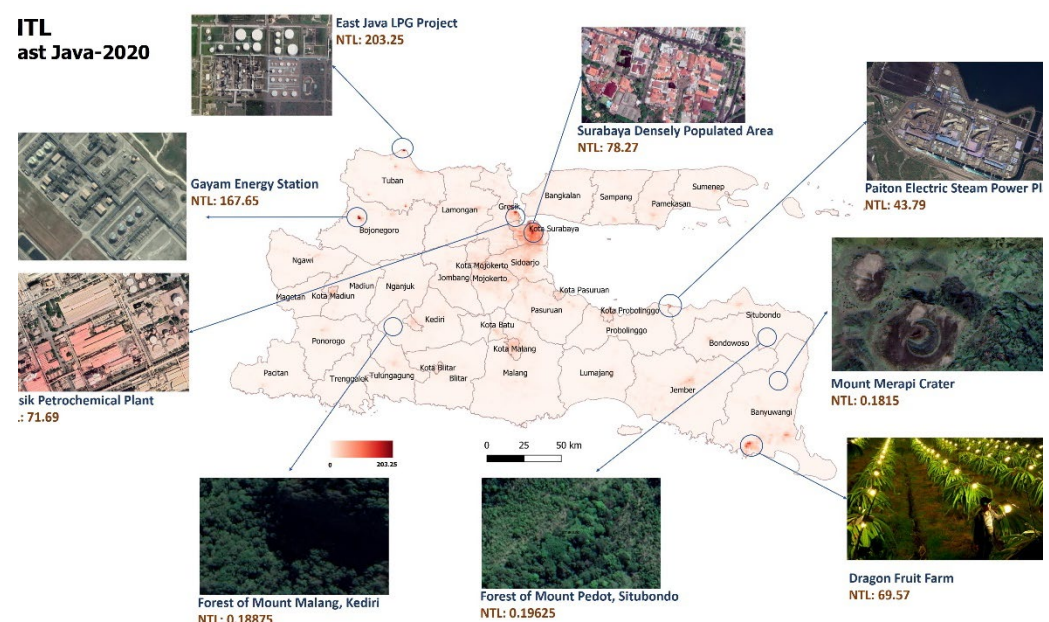
**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### *Remote-Sensing Variables Ground-Truth Check*

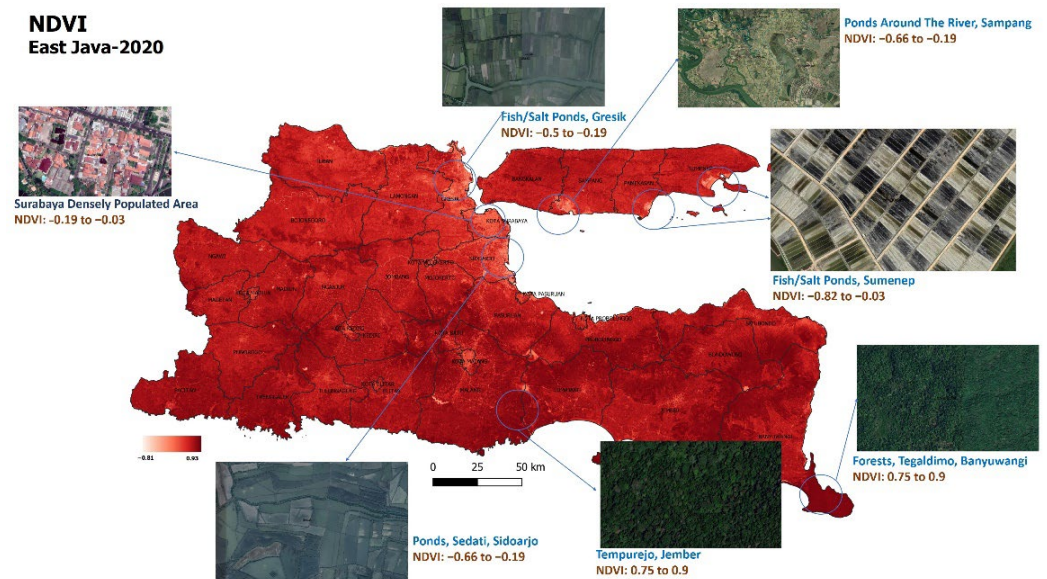
To gain a deeper understanding of the use of multisource remote-sensing satellite imagery, we identified several unique points through high-resolution Google satellites. Night-time light (NTL) is the most commonly used remote sensing imagery for estimating poverty [59]. NTL can record luminosity from human activity at night and has been demonstrated to have a great ability to estimate various socioeconomic parameters, including

poverty. A higher NTL value indicates higher economic activity in an area [60]. Through descriptive analysis, we observed several NTL values, to identify the corresponding geographic features and compare them to the official East Java poverty data. Figure A1 shows some of the NTL values and their ground-truth check. Using Figure A1, we will discuss some of the anomalies that exist in the use of NTL data in identifying poverty in East Java. First, we can see that high NTL scores are clustered in Kota Surabaya. Surabaya is the largest metropolitan city in Indonesia after Jakarta. However, official poverty data shows that Kota Batu, Kota Malang, and Kota Madiun are the three lowest-poverty regencies/municipalities in East Java. Second, we can also see that most of the largest NTL values are not in regencies/municipalities with low poverty. Areas with energy plants and factories that operate late into the night tend to have high NTL values. Third, high values of NTL were also found in plantations with all-night lighting which did not indicate poverty. These anomalies indicate that in some areas, for example, in East Java, the use of NTL values as a poverty approach needs to be supported by the use of other data for better identification. From Figure A1, we can also see that areas with low NTL values are usually the forests around mountains.



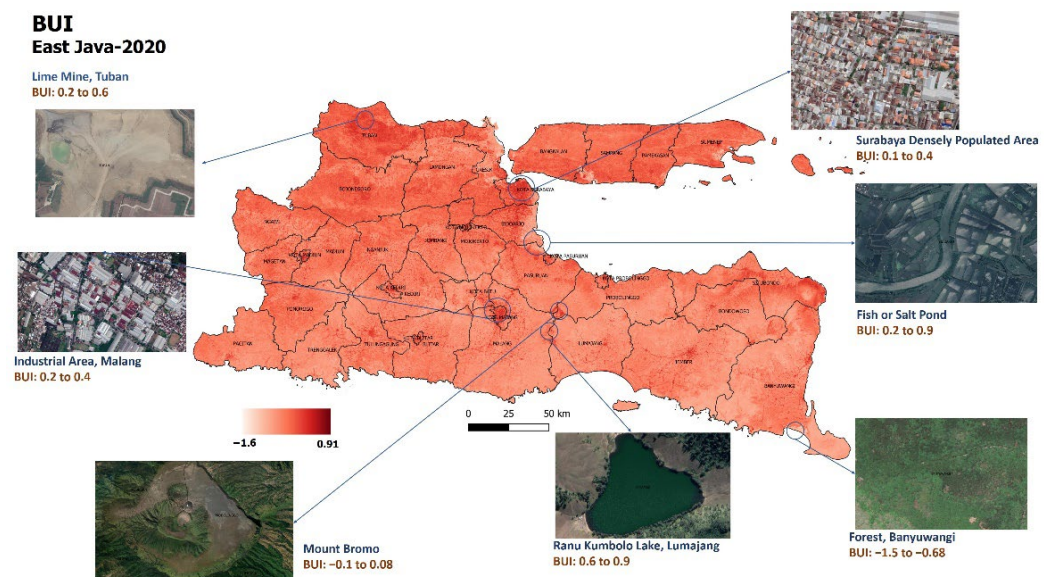
**Figure A1.** East Java NTL values (nanowatts/cm<sup>2</sup>/sr) and their resulted ground-truth checking.

Dawson et al. [32] revealed that the NDVI (normalized difference vegetation index) was significantly correlated with poverty (positive high and/or negative high) in both shrinking and growing countries. From Figure A2, we can see that the very low NDVI value is likely from a water-covered area, such as a pond; this value is usually lower than that of buildings. High NDVI values indicate a dense forest area. This is in line with a study by Maurya R. et al. [91], which stated that negative NDVI values represent non-vegetated areas such as water, while positive NDVI values indicate the vegetated ones. It can be seen that Kota Surabaya and Kota Malang, which are densely built-up areas, tend to have lower NDVI values, and poverty in these areas is also low.



**Figure A2.** East Java NDVI (index) values and their resulted ground truth checking.

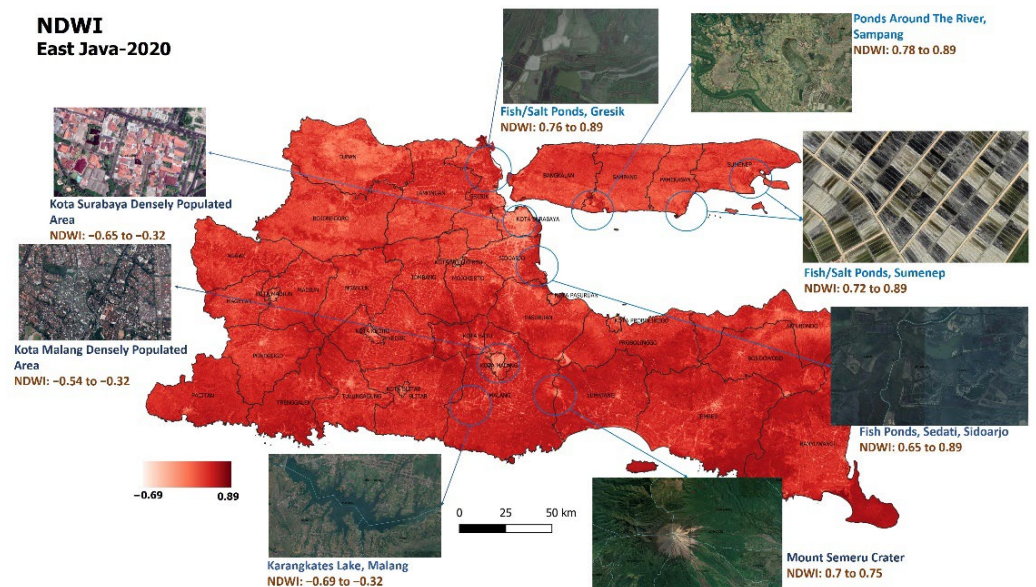
According to Lee et al. [63], the built-up index (BUI) is compatible with classifying urban and non-urban areas. The higher BUI value indicates the higher possibility of a pixel signifying a built-up area. However, in the case of heterogeneous East Java, a very high BUI value was obtained not only in urban built-up areas but also in lime mines and inland water, as we can see in Figure A3. Lime mining is mostly found in the Tuban and Bojonegoro regencies. Therefore, this area also had high BUI values, besides the urban areas. In fact, according to official poverty data, Tuban and Bojonegoro are regencies with high poverty. We can also see that those areas with low BUI values are usually forests and mountain craters. This is why regencies/municipalities with many mountains, such as Kediri, Malang, Lumajang, Jember, and Pasuruan, tend to have low BUI values.



**Figure A3.** East Java BUI values (index) and their resulted ground-truth checking.

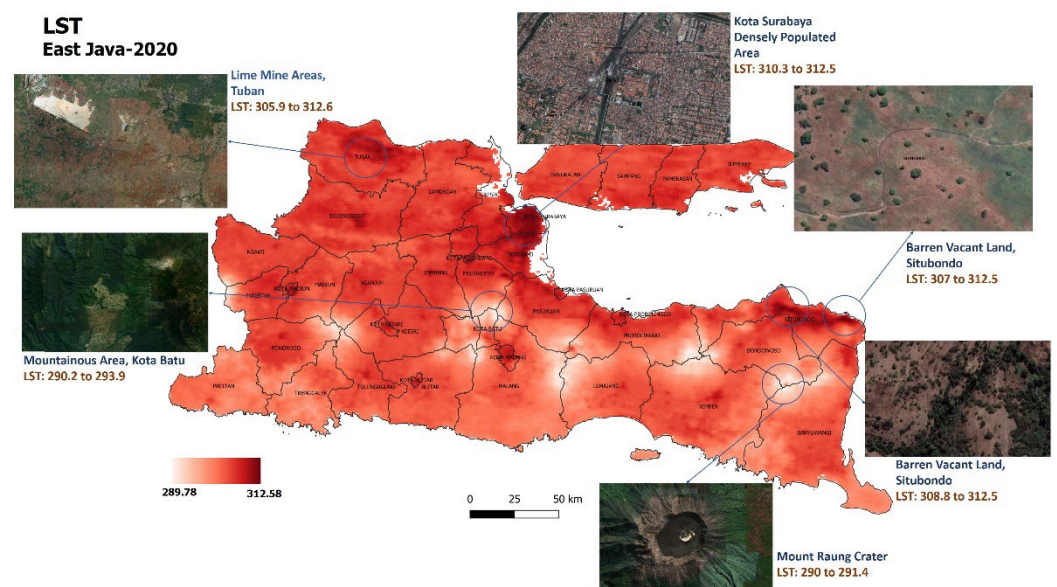
Figure A4 shows the distribution of the normalized difference water index (NDWI). Theoretically, NDWI values above 0 indicate a water body imagery result; otherwise, values below 0 indicate a non-water imagery result [64]. As we can see, areas covered by water are most likely to have an NDWI value of more than zero, except for the Karangates reservoir,

which has an NDWI value below zero. Regencies/municipalities with many mountains, such as Kediri, Malang, Lumajang, Jember, and Pasuruan, tend to have high NDWI values.



**Figure A4.** East Java NDWI values (index) and their resulting ground-truth checking.

Land surface temperature (LST) is often used to identify urban areas. In their research, which examines the correlation between LST and urban heat islands, Mia et al. [92] stated that the higher the LST value, the more extensive the urban heat island. However, as we can see in Figure A5, in the case of East Java, areas with high LST are not only urban areas but also arid areas and limestone mining areas. The northern part of East Java is an area with a great deal of arid and empty land and limestone mines, which tend to have a higher LST value than the southern area, which consists of mountains. We can see that regencies/municipalities with high poverty rates tend to have low LST values, except for Kota Surabaya and the regencies/municipalities around it.



**Figure A5.** East Java LST values (Kelvin) and their resulted ground truth checking.

Air pollution can be one indicator of economic activity that is related to poverty. Wu et al. [93] use CO<sub>2</sub> (carbon dioxide) and SO<sub>2</sub> (sulfur dioxide) to examine the poverty-environmental trap. In this study, we examine CO (carbon monoxide), NO<sub>2</sub> (nitrogen dioxide), and SO<sub>2</sub> to capture air pollution in East Java. Figures A6–A8 show the East Java pollution indicator values and their ground-truth checks. Each pollution indicator's high value captures different information. From Figure A6, we can see that areas with high CO values are industrially dense areas. This value exceeds the CO value of Kota Surabaya, which is the second-largest metropolitan city in Indonesia. From Figure A7, we can see that those areas with high NO<sub>2</sub> values are steam power plants and densely populated and industrial areas. From Figure A8, we can see that the highest SO<sub>2</sub> values we captured are in mountain slope areas; however, the values can be small in other areas. On the other hand, each pollution indicator's low value captures similar information. Areas with low CO, NO<sub>2</sub>, or SO<sub>2</sub> are usually mountain peaks.

### CO East Java-2020

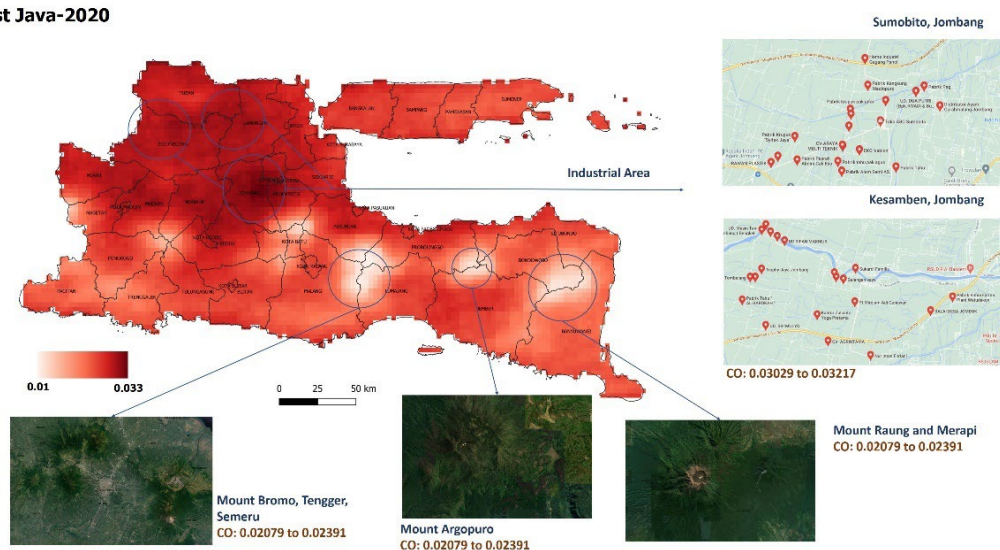


Figure A6. East Java CO values (mol/m<sup>2</sup>) and their resulting ground-truth checking.

### NO<sub>2</sub> East Java-2020

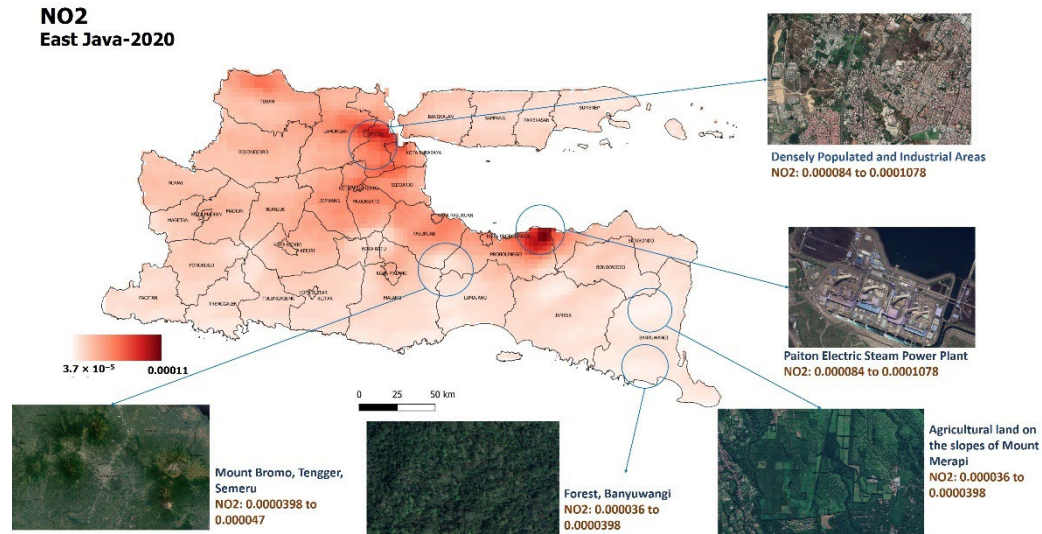
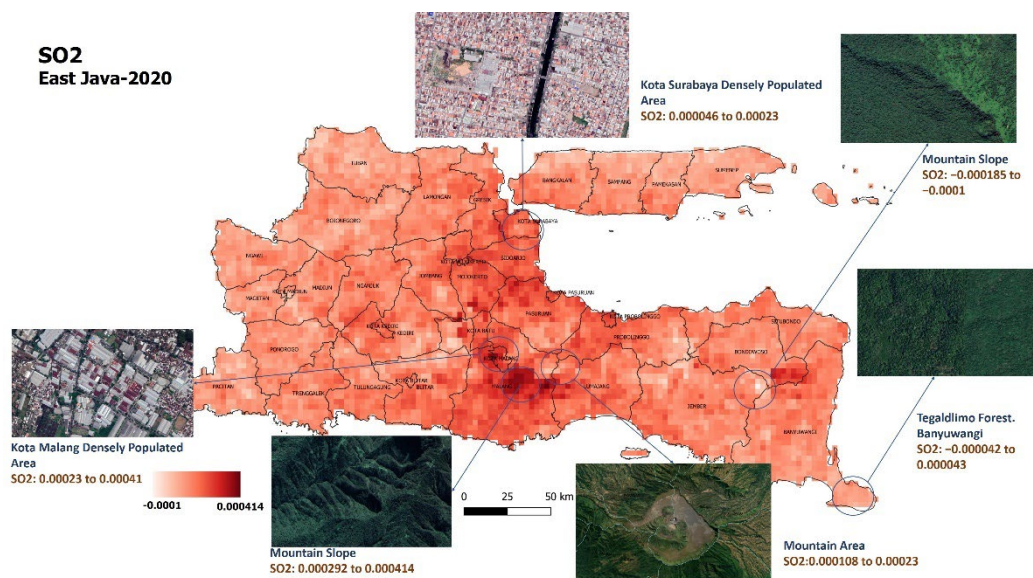


Figure A7. East Java NO<sub>2</sub> values (mol/m<sup>2</sup>) and their resulting ground-truth checking.



**Figure A8.** East Java SO<sub>2</sub> values (mol/m<sup>2</sup>) and their resulting ground-truth checking.

## References

1. Steele, J.E.; Sundsøy, P.R.; Pezzulo, C.; Alegana, V.A.; Bird, T.J.; Blumenstock, J.; Bjelland, J.; Engø-Monsen, K.; de Montjoye, Y.-A.; Iqbal, A.M.; et al. Mapping Poverty Using Mobile Phone and Satellite Data. *R. Soc.* **2017**, *14*, 20160690. [CrossRef] [PubMed]
2. United Nations End Poverty in All Its Forms Everywhere. Available online: <https://unstats.un.org/sdgs/report/2020/goal-01/> (accessed on 23 December 2021).
3. United Nations Ending Poverty. Available online: <https://www.un.org/en/global-issues/ending-poverty> (accessed on 1 March 2022).
4. United Nations about the Sustainable Development Goals. Available online: <https://www.un.org/sustainabledevelopment/sustainable-development-goals/> (accessed on 23 December 2021).
5. Zhao, X.; Yu, B.; Liu, Y.; Chen, Z.; Li, Q.; Wang, C.; Wu, J. Estimation of Poverty Using Random Forest Regression with Multi-Source Data: A Case Study in Bangladesh. *Remote Sens.* **2019**, *11*, 375. [CrossRef]
6. Statistics Indonesia (BPS). *Regency/Municipality Poverty Data and Information in 2021*; Statistics Indonesia: Jakarta, Indonesia, 2021.
7. Indonesia National Development Planning Agency (Bappenas). *Indonesia SDGs Roadmap Towards 2030*; Indonesia National Development Planning Agency: Jakarta, Indonesia, 2017.
8. Jerven, M. Benefits and Costs of the Data for Development Targets for the Post-2015 Development Agenda. *Data Dev. Assess. Pap.* **2014**, *16*, 14.
9. Laurentia, S.; Yusran, R. Evaluation of the Non-Cash Food Assistance Program in Poverty Reduction in Padang District. *J. Civ. Educ.* **2021**, *4*, 7–17. [CrossRef]
10. Indonesia National Development Planning Agency (Bappenas). *National Mid-Term Development Plan (RPJMN) 2020–2024*; Indonesia National Development Planning Agency: Jakarta, Indonesia, 2020.
11. Triscowati, D.W.; Sartono, B.; Kurnia, A.; Domiri, D.D.; Wijayanto, A.W. Multitemporal Remote Sensing Data for Classification of Food Crops Plant Phase Using Supervised Random Forest. In Proceedings of the 6th Geoinformation Science Symposium, Yogyakarta, Indonesia, 26–27 August 2019; Volume 11311, p. 1131102.
12. Triscowati, D.W.; Sartono, B.; Kurnia, A.; Dirgahayu, D.; Wijayanto, A.W. Classification of Rice-Plant Growth Phase Using Supervised Random Forest Method Based on Landsat-8 Multitemporal Data. *Int. J. Remote Sens. Earth Sci.* **2020**, *16*, 187–196. [CrossRef]
13. Wijayanto, A.W.; Triscowati, D.W.; Marsuhandi, A.H. Maize Field Area Detection in East Java, Indonesia: An Integrated Multispectral Remote Sensing and Machine Learning Approach. In Proceedings of the 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 6–8 October 2020; pp. 168–173.
14. Shi, K.; Chang, Z.; Chen, Z.; Wu, J.; Yu, B. Identifying and Evaluating Poverty Using Multisource Remote Sensing and Point of Interest (POI) Data: A Case Study of Chongqing, China. *J. Clean. Prod.* **2020**, *255*, 120245. [CrossRef]
15. Fauzi, A.I.; Sakti, A.D.; Yayusman, L.F.; Harto, A.B.; Prasetyo, L.B.; Irawan, B.; Wikantika, K. Evaluating mangrove forest deforestation causes in Southeast Asia by analyzing recent environment and socio-economic data products. In Proceedings of the 39th Asian Conference on Remote Sensing: Remote Sensing Enabling Prosperity, ACRS 2018, Kuala Lumpur, Malaysia, 15–19 October 2018; Asian Association on Remote Sensing: Kuala Lumpur, Malaysia, 2018; Volume 2, pp. 880–889.
16. Pokhriyal, N.; Zambrano, O.; Linares, J.; Hernández, H. *Estimating and Forecasting Income Poverty and Inequality in Haiti Using Satellite Imagery and Mobile Phone Data*; Inter-American Development Bank: Washington, DC, USA, 2020.

17. Ghosh, T.; Anderson, S.J.; Elvidge, C.D.; Sutton, P.C. Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being. *Sustainability* **2013**, *5*, 4988–5019. [\[CrossRef\]](#)
18. Rajagukguk, Y.S.; Sakti, A.D.; Yayusman, L.F.; Harto, A.B.; Prasetyo, L.B.; Irawan, B.; Wikantika, K. Evaluation of Southeast Asia mangrove forest deforestation using longterm remote sensing index datasets. In Proceedings of the 39th Asian Conference on Remote Sensing: Remote Sensing Enabling Prosperity, ACRS 2018, Kuala Lumpur, Malaysia, 15–19 October 2018; Asian Association on Remote Sensing: Kuala Lumpur, Malaysia, 2018; Volume 2, pp. 931–937.
19. Zhao, N.; Liu, Y.; Cao, G.; Samson, E.L.; Zhang, J. Forecasting China's GDP at the Pixel Level Using Nighttime Lights Time Series and Population Images. *GISci. Remote Sens.* **2017**, *54*, 407–425. [\[CrossRef\]](#)
20. Shi, K.; Yu, B.; Huang, Y.; Hu, Y.; Yin, B.; Chen, Z.; Chen, L.; Wu, J. Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data. *Remote Sens.* **2014**, *6*, 1705–1724. [\[CrossRef\]](#)
21. Putri, S.R.; Suganda, T.G.; Pramana, S. Bayesian Network Implementation for Modelling Indonesia's Green Economy Condition Based on Big Data. In Proceedings of the Seminar Nasional Official Statistics, Jakarta, Indonesia, 25 September 2021; Volume 2021, pp. 1054–1064.
22. Sakti, A.D.; Rinasti, A.N.; Agustina, E.; Diastomo, H.; Muhammad, F.; Anna, Z.; Wikantika, K. Multi-Scenario Model of Plastic Waste Accumulation Potential in Indonesia Using Integrated Remote Sensing, Statistic and Socio-Demographic Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 481. [\[CrossRef\]](#)
23. Chen, Z.; Yu, B.; Hu, Y.; Huang, C.; Shi, K.; Wu, J. Estimating House Vacancy Rate in Metropolitan Areas Using NPP-VIIRS Nighttime Light Composite Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2188–2197. [\[CrossRef\]](#)
24. Shi, K.; Yang, Q.; Fang, G.; Yu, B.; Chen, Z.; Yang, C.; Wu, J. Evaluating Spatiotemporal Patterns of Urban Electricity Consumption within Different Spatial Boundaries: A Case Study of Chongqing, China. *Energy* **2019**, *167*, 641–653. [\[CrossRef\]](#)
25. Sutton, P.; Roberts, D.; Elvidge, C.; Baugh, K. Census from Heaven: An Estimate of the Global Human Population Using Night-Time Satellite Imagery. *Int. J. Remote Sens.* **2001**, *22*, 3061–3076. [\[CrossRef\]](#)
26. Shi, K.; Yu, B.; Hu, Y.; Huang, C.; Chen, Y.; Huang, Y.; Chen, Z.; Wu, J. Modeling and Mapping Total Freight Traffic in China Using NPP-VIIRS Nighttime Light Composite Data. *GISci. Remote Sens.* **2015**, *52*, 274–289. [\[CrossRef\]](#)
27. Elvidge, C.D.; Sutton, P.C.; Ghosh, T.; Tuttle, B.T.; Baugh, K.E.; Bhaduri, B.; Bright, E. A Global Poverty Map Derived from Satellite Data. *Comput. Geosci.* **2009**, *35*, 1652–1660. [\[CrossRef\]](#)
28. Yu, B.; Tang, M.; Wu, Q.; Yang, C.; Deng, S.; Shi, K.; Peng, C.; Wu, J.; Chen, Z. Urban Built-up Area Extraction from Log-Transformed NPP-VIIRS Nighttime Light Composite Data. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1279–1283. [\[CrossRef\]](#)
29. Yin, J.; Qiu, Y.; Zhang, B. Identification of Poverty Areas by Remote Sensing and Machine Learning: A Case Study in Guizhou, Southwest China. *ISPRS Int. J. Geo-Inf.* **2020**, *10*, 11. [\[CrossRef\]](#)
30. Sapena, M.; Ruiz, L.A.; Taubenböck, H. Analyzing Links between Spatio-Temporal Metrics of Built-up Areas and Socio-Economic Indicators on a Semi-Global Scale. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 436. [\[CrossRef\]](#)
31. Tian, Y.; Wang, Z.; Zhao, J.; Jiang, X.; Guo, R. A Geographical Analysis of the Poverty Causes in China's Contiguous Destitute Areas. *Sustainability* **2018**, *10*, 1895. [\[CrossRef\]](#)
32. Dawson, T.; Sandoval, J.S.; Sagan, V.; Crawford, T. A Spatial Analysis of the Relationship between Vegetation and Poverty. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 83. [\[CrossRef\]](#)
33. Kaimaris, D.; Patias, P. Identification and Area Measurement of the Built-up Area with the Built-up Index (BUI). *Int. J. Adv. Remote Sens. GIS* **2016**, *5*, 1844–1858. [\[CrossRef\]](#)
34. Alkire, S.; Chatterje, M.; Conconi, A.; Seth, S.; Vaz, A. Poverty in Rural and Urban Areas: Direct Comparisons Using the Global MPI 2014. *Briefing* **2014**. [\[CrossRef\]](#)
35. Huang, G.; Zhou, W.; Cadenasso, M.L. Is Everyone Hot in the City? Spatial Pattern of Land Surface Temperatures, Land Cover and Neighborhood Socioeconomic Characteristics in Baltimore, MD. *J. Environ. Manag.* **2011**, *92*, 1753–1759. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Ahmed, S. Assessment of Urban Heat Islands and Impact of Climate Change on Socioeconomic over Suez Governorate Using Remote Sensing and GIS Techniques. *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 15–25. [\[CrossRef\]](#)
37. Zheng, Y.; Zhou, Q.; He, Y.; Wang, C.; Wang, X.; Wang, H. An Optimized Approach for Extracting Urban Land Based on Log-Transformed DMSP-OLS Nighttime Light, NDVI, and NDWI. *Remote Sens.* **2021**, *13*, 766. [\[CrossRef\]](#)
38. Baloch, M.A.; Khan, S.U.-D.; Ulucak, Z.S.; Ahmad, A. Analyzing the Relationship between Poverty, Income Inequality, and CO<sub>2</sub> Emission in Sub-Saharan African Countries. *Sci. Total Environ.* **2020**, *740*, 139867. [\[CrossRef\]](#)
39. Sakti, A.D.; Fauzi, A.I.; Takeuchi, W.; Pradhan, B.; Yarime, M.; Vega-Garcia, C.; Agustina, E.; Wibisono, D.; Anggraini, T.S.; Theodora, M.O.; et al. Spatial Prioritization for Wildfire Mitigation by Integrating Heterogeneous Spatial Data: A New Multi-Dimensional Approach for Tropical Rainforests. *Remote Sens.* **2022**, *14*, 543. [\[CrossRef\]](#)
40. Wang, Y.; Li, J.; Wang, L.; Lin, Y.; Zhou, M.; Yin, P.; Yao, S. The Impact of Carbon Monoxide on Years of Life Lost and Modified Effect by Individual-and City-Level Characteristics: Evidence from a Nationwide Time-Series Study in China. *Ecotoxicol. Environ. Saf.* **2021**, *210*, 111884. [\[CrossRef\]](#)
41. Sakti, A.D.; Rahadianto, M.A.E.; Pradhan, B.; Muhammad, H.N.; Andani, I.G.A.; Sarli, P.W.; Abdillah, M.R.; Anggraini, T.S.; Purnomo, A.D.; Ridwana, R.; et al. School Location Analysis by Integrating the Accessibility, Natural and Biological Hazards to Support Equal Access to Education. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 12. [\[CrossRef\]](#)

42. Han, C.; Gu, Z.; Yang, H. EKC Test of the Relationship between Nitrogen Dioxide Pollution and Economic Growth—A Spatial Econometric Analysis Based on Chinese City Data. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9697. [CrossRef]
43. Bakhsh, K.; Akmal, T.; Ahmad, T.; Abbas, Q. Investigating the Nexus among Sulfur Dioxide Emission, Energy Consumption, and Economic Growth: Empirical Evidence from Pakistan. *Environ. Sci. Pollut. Res.* **2022**, *29*, 7214–7224. [CrossRef] [PubMed]
44. Duque, J.C.; Patino, J.E.; Ruiz, L.A.; Pardo-Pascual, J.E. Measuring Intra-Urban Poverty Using Land Cover and Texture Metrics Derived from Remote Sensing Data. *Landsc. Urban Plan.* **2015**, *135*, 11–21. [CrossRef]
45. Sakti, A.D.; Tsuyuki, S. Spectral Mixture Analysis of Peatland Imagery for Land Cover Study of Highly Degraded Peatland in Indonesia. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*; Copernicus Publications: Göttingen, Germany, 2015; Volume XL-7/W3.
46. Niu, T.; Chen, Y.; Yuan, Y. Measuring Urban Poverty Using Multi-Source Data and a Random Forest Algorithm: A Case Study in Guangzhou. *Sustain. Cities Soc.* **2020**, *54*, 102014. [CrossRef]
47. Zhou, Y.; Liu, Y. The geography of poverty: Review and research prospects. *J. Rural Stud.* 2019, *in press*. Available online: <https://www.sciencedirect.com/science/article/abs/pii/S0743016718303899> (accessed on 30 September 2021). [CrossRef]
48. Statistics Indonesia (BPS). *Regency/Municipality Poverty Data and Information in 2020*; Statistics Indonesia: Jakarta, Indonesia, 2020.
49. Wasonowati, C.; Sulistyarningsih, E.; Indradewa, D.; Kurniasih, B. Physiological Characters of Moringa Oleifera Lamk in Madura. In *AIP Conference Proceedings*; American Institute of Physics: University Park, MD, USA, 2019; Volume 2120, p. 30024.
50. Wisnubroto, E.I.; Rustiadi, E.; Fauzi, A.; Murti Laksono, K. The Dynamic Changes in Peri-Urban Agricultural Area and Typology of Multi-Function Agriculture in Batu City, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *667*, 12093. [CrossRef]
51. Setiawan, F.C.; Rahadian, S. Social, Cultural and Political Conditions in Malang Before Kanjuruhan Kingdom. *Soc. Sci. Stud. Sustain. ISSUES* **2019**, *63*, 63–69.
52. Santoso, E.B.; Aulia, B.U. Ecological Sustainability Level of Surabaya City Based on Ecological Footprint Approach. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *202*, 12044. [CrossRef]
53. Nurmasari, Y.; Wijayanto, A.W. Oil Palm Plantation Detection in Indonesia Using Sentinel-2 and Landsat-8 Optical Satellite Imagery (Case Study: Rokan Hulu Regency, Riau Province). *Int. J. Remote Sens. Earth Sci.* **2021**, *18*, 1–18. [CrossRef]
54. Saadi, T.D.T.; Wijayanto, A.W. Machine Learning Applied to Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production Areas in Nganjuk, East Java, Indonesia. *Int. J. Remote Sens. Earth Sci.* **2021**, *18*, 19–32.
55. Putri, S.R.; Wijayanto, A.W. Learning Bayesian Network for Rainfall Prediction Modeling in Urban Area Using Remote Sensing Satellite Data (Case Study: Jakarta, Indonesia). In *Proceedings of the International Conference on Data Science and Official Statistics*, Online, 13–14 November 2021; Volume 2021, pp. 77–90.
56. Tingzon, I.; Orden, A.; Sy, S.; Sekara, V.; Weber, I.; Fatehkia, M.; Garcia, M.; Dohyung, H. Mapping Poverty in the Philippines Using Machine Learning, Satellite Imagery, and Crowd-Sourced Geospatial Information. *SPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 425–431. [CrossRef]
57. Ledesma, C.; Garonita, O.L.; Flores, L.J.; Tingzon, I.; Dalisay, D. Interpretable Poverty Mapping Using Social Media Data, Satellite Images, and Geospatial Information. *arXiv* **2020**, arXiv:2011.13563.
58. Earth Observation Group Payne Institute for Public Policy Colorado School of Mines VIIRS Nighttime Day/Night Band Composites Version 1. Available online: [https://developers.google.com/earth-engine/datasets/catalog/NOAA\\_VIIRS\\_DNB\\_MONTHLY\\_V1\\_VCMCFG](https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMCFG) (accessed on 10 September 2021).
59. Zhao, N.; Cao, G.; Zhang, W.; Samson, E.L.; Chen, Y. Remote Sensing and Social Sensing for Socioeconomic Systems: A Comparison Study between Nighttime Lights and Location-Based Social Media at the 500 m Spatial Resolution. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *87*, 102058. [CrossRef]
60. Zhou, N.; Hubacek, K.; Roberts, M. Analysis of Spatial Patterns of Urban Growth across South Asia Using DMSP-OLS Nighttime Lights Data. *Appl. Geogr.* **2015**, *63*, 292–303. [CrossRef]
61. Copernicus ESA European Union Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. Available online: [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_SR](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR) (accessed on 10 September 2021).
62. Purevdorj, T.S.; Tateishi, R.; Ishiyama, T.; Honda, Y. Relationships between Percent Vegetation Cover and Vegetation Indices. *Int. J. Remote Sens.* **1998**, *19*, 3519–3535. [CrossRef]
63. Lee, J.; Lee, S.S.; Chi, K.H. Development of an Urban Classification Method Using a Built-up Index. In *Proceedings of the 6th WSEAS International Conference on Remote Sensing*, Takizawa, Japan, 4–6 October 2010; pp. 39–43.
64. McFeeters, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
65. NASA LP DAAC at the USGS EROS Center MOD11A1.006 Terra Land Surface Temperature and Emissivity Daily Global 1 km. Available online: [https://developers.google.com/earth-engine/datasets/catalog/MODIS\\_006\\_MOD11A1](https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD11A1) (accessed on 15 September 2021).
66. Copernicus ESA European Union Sentinel-5P OFFL CO: Offline Carbon Monoxide. Available online: [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S5P\\_OFFL\\_L3\\_CO](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_CO) (accessed on 10 September 2021).
67. Copernicus ESA European Union Sentinel-5P OFFL NO<sub>2</sub>: Offline Nitrogen Dioxide. Available online: [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S5P\\_OFFL\\_L3\\_NO2?hl=en](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_NO2?hl=en) (accessed on 10 September 2021).

68. Copernicus ESA European Union Sentinel-5P OFFL SO<sub>2</sub>: Offline Sulphur Dioxide. Available online: [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S5P\\_OFFL\\_L3\\_SO2](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_SO2) (accessed on 10 September 2021).
69. Open StreetMap Open StreetMap. Available online: <https://www.openstreetmap.org/> (accessed on 15 July 2020).
70. Tamilselvi, R.; Sivasakthi, B.; Kavitha, R. An Efficient Preprocessing and Postprocessing Techniques in Data Mining. *Int. J. Res. Comput. Appl. Robot.* **2015**, *3*, 80–85.
71. Fauzi, A.I.; Sakti, A.D.; Robbani, B.F.; Ristiyani, M.; Agustin, R.T.; Yati, E.; Nuha, M.U.; Anika, N.; Putra, R.; Siregar, D.I.; et al. Assessing Potential Climatic and Human Pressures in Indonesian Coastal Ecosystems Using a Spatial Data-Driven Approach. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 778. [\[CrossRef\]](#)
72. Zha, Y.; Gao, J.; Ni, S. Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [\[CrossRef\]](#)
73. Raymaekers, J.; Rousseeuw, P.J. Transforming Variables to Central Normality. *Mach. Learn.* **2021**, 1–23. [\[CrossRef\]](#)
74. Sugiyono. *Educational Research Methods: Quantitative, Qualitative and R&D Approaches*; Alfabeta: Bandung, Indonesia, 2010.
75. Wang, B.; Tian, J.; Yang, P.; He, B. Multi-Scale Features of Regional Poverty and the Impact of Geographic Capital: A Case Study of Yanbian Korean Autonomous Prefecture in Jilin Province, China. *Land* **2021**, *10*, 1406. [\[CrossRef\]](#)
76. Liu, Y.; Xu, Y. A Geographic Identification of Multidimensional Poverty in Rural China under the Framework of Sustainable Livelihoods Analysis. *Appl. Geogr.* **2016**, *73*, 62–76. [\[CrossRef\]](#)
77. Wang, Y.; Chen, Y. Using VPI to Measure Poverty-Stricken Villages in China. *Soc. Indic. Res.* **2017**, *133*, 833–857. [\[CrossRef\]](#)
78. Goyal, M.K.; Singh, V.; Meena, A.H. Geospatial and Hydrological Modeling to Assess Hydropower Potential Zones and Site Location over Rainfall Dependent Inland Catchment. *Water Resour. Manag.* **2015**, *29*, 2875–2894. [\[CrossRef\]](#)
79. Uddin, M.N.; Islam, A.K.M.S.; Bala, S.K.; Islam, G.M.T.; Adhikary, S.; Saha, D.; Haque, S.; Fahad, M.G.R.; Akter, R. Mapping of Climate Vulnerability of the Coastal Region of Bangladesh Using Principal Component Analysis. *Appl. Geogr.* **2019**, *102*, 47–57. [\[CrossRef\]](#)
80. Cartone, A.; Postiglione, P. Principal Component Analysis for Geographical Data: The Role of Spatial Effects in the Definition of Composite Indicators. *Spat. Econ. Anal.* **2021**, *16*, 126–147. [\[CrossRef\]](#)
81. Sakti, A.D.; Takeuchi, W. A data-intensive approach to address food sustainability: Integrating optic and microwave satellite imagery for developing long-term global cropping intensity and sowing month from 2001 to 2015. *Sustainability* **2020**, *12*, 3227. [\[CrossRef\]](#)
82. Varshney, K.R.; Chen, G.H.; Abelson, B.; Nowocin, K.; Sakhrani, V.; Xu, L.; Spatocco, B.L. Targeting Villages for Rural Development Using Satellite Image Analysis. *Big Data* **2015**, *3*, 41–53. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Assael, H.; Keon, J. Nonsampling vs. Sampling Errors in Survey Research. *J. Mark.* **1982**, *46*, 114–123. [\[CrossRef\]](#)
84. Sarmah, H.K.; Hazarika, B.B.; Choudhury, G. An Investigation on Effect of Bias on Determination of Sample Size on the Basis of Data Related to the Students of Schools of Guwahati. *Int. J. Appl. Math. Stat. Sci.* **2013**, *2*, 33–48.
85. Sari, F.E.K.; Fitria, I.; Hariyanto, D.D.; Fachriansyah, M.A. *Internship Lecture Report Analysis of Internal Control Activities of the National Socio-Economic Survey (Susenas) Central Bureau of Statistics Jombang Regency*; STIE PGRI Dewantara: Surabaya, Indonesia, 2019.
86. Afifah, U.N.; Faradis, R. Sosial Ekonomi Nasional Survey (Susenas) Data Optimization with Small Area Estimation (SAE) Case Study: Village Level Poverty Estimation in Belitung Timur Regency. In Proceedings of the Seminar Nasional Official Statistics, Jakarta, Indonesia, 24 September 2019; Volume 2019, pp. 132–139.
87. Chakravarty, S.R.; Deutsch, J.; Silber, J. On the Watts Multidimensional Poverty Index and Its Decomposition. *World Dev.* **2008**, *36*, 1067–1077. [\[CrossRef\]](#)
88. Jalan, J.; Ravallion, M. Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China. *J. Appl. Econom.* **2002**, *17*, 329–346. [\[CrossRef\]](#)
89. Grant, U. *The Chronic Poverty Report 2004–2005*; Institute for Development Policy & Management, University of Manchester: Manchester, UK, 2004.
90. Addison, T.; Harper, C.; Prowse, M.; Shepherd, A.; Barrientos, A.; Braunholtz-Speight, T.; Evans, A.; Grant, U.; Hickey, S.; Hulme, D.; et al. The Chronic Poverty Report 2008–2009: Escaping Poverty Traps. *Eur. J. Dev. Res.* **2008**, *21*, 159.
91. Maurya, R.; Gupta, P.R.; Shukla, A.S.; Sharma, M.K. Building Extraction from Very High Resolution Multispectral Images Using NDVI Based Segmentation and Morphological Operators. In Proceedings of the IEEE-International Conference on Advances In Engineering, Science and Management (ICAESM-2012), Nagapattinam, India, 30–31 March 2012; pp. 577–581.
92. Mia, B.; Bhattacharya, R.; Woobaidullah, A.S.M. Correlation and Monitoring of Land Surface Temperature, Urban Heat Island with Land Use-Land Cover of Dhaka City Using Satellite Imageries. *Int. J. Res. Geogr.* **2017**, *3*, 10–20.
93. Wu, J.; He, L.-Y.; Zhang, Z. Does China Fall into Poverty-Environment Traps? Evidence from Long-Term Income Dynamics and Urban Air Pollution. *FEEM Work. Pap.* **2019**, *5*, 1–27.