*remote sensing*

MDPI

*Article*

# County-Level Poverty Evaluation Using Machine Learning, Nighttime Light, and Geospatial Data

**Xiaoqian Zheng [1], Wenjiang Zhang [2,*], Hui Deng [2] and Houxi Zhang [3]**

[1] School of Finance, Fujian Business University, Fuzhou 350016, China; 12304096004@fafu.edu.cn
[2] College of Geography and Planning, Chengdu University of Technology, Chengdu 610059, China; dengh@cdut.edu.cn
[3] Forestry College, Fujian Agriculture and Forestry University, Fuzhou 350028, China; zhanghouxi@fafu.edu.cn
* Correspondence: zhngwenjiang13@cdut.edu.cn

**Abstract:** The accurate and timely acquisition of poverty information within a specific region is crucial for formulating effective development policies. Nighttime light (NL) remote sensing data and geospatial information provide the means for conducting precise and timely evaluations of poverty levels. However, current assessment methods predominantly rely on NL data, and the potential of combining multi-source geospatial data for poverty identification remains underexplored. Therefore, we propose an approach that assesses poverty based on both NL and geospatial data using machine learning models. This study uses the multidimensional poverty index (MPI), derived from county-level statistical data with social, economic, and environmental dimensions, as an indicator to assess poverty levels. We extracted a total of 17 independent variables from NL and geospatial data. Machine learning models (random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM)) and traditional linear regression (LR) were used to model the relationship between the MPI and independent variables. The results indicate that the RF model achieved significantly higher accuracy, with a coefficient of determination ($R^2$) of 0.928, a mean absolute error (MAE) of 0.030, and a root mean square error (RMSE) of 0.037. The top five most important variables comprise two (NL_MAX and NL_MIN) from the NL data and three (POI_Ed, POI_Me, and POI_Ca) from the geographical spatial data, highlighting the significant roles of NL data and geographical data in MPI modeling. The MPI map that was generated by the RF model depicted the detailed spatial distribution of poverty in Fujian province. This study presents an approach to county-level poverty evaluation that integrates NL and geospatial data using a machine learning model, which can contribute to a more reliable and efficient estimate of poverty.

**Keywords:** nighttime light (NL); random forest (RF); multidimensional poverty index (MPI); Fujian province

check for updates

**Citation:** Zheng, X.; Zhang, W.; Deng, H.; Zhang, H. County-Level Poverty Evaluation Using Machine Learning, Nighttime Light, and Geospatial Data. *Remote Sens.* **2024**, *16*, 962. https://doi.org/10.3390/rs16060962

Academic Editor: Lefei Zhang

Received: 13 January 2024
Revised: 5 March 2024
Accepted: 5 March 2024
Published: 9 March 2024

**Correction Statement:** This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

**Copyright:** © 2024 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 represents a major global challenge, posing a substantial barrier to sustainable human development and exhibiting intricate interconnections with economic growth, ecological restoration, education levels, and resource utilization [1]. The accurate and effective identification and assessment of poverty, coupled with the analysis of its geographical distribution, are critical for formulating rational policies to allocate resources and alleviate poverty [2].

In conventional perspectives, "poverty" typically refers to an economic phenomenon, describing situations where individuals or households cannot meet basic living standards [3]. Consequently, key socioeconomic indicators such as gross domestic product (GDP), per capita income, or per capita government expenditure are widely used for poverty identification and evaluation [4,5]. As society evolves, the definition of poverty

has broadened to encompass broader social, cultural, and human development dimensions, moving from a singular economic focus to a multidimensional concept that includes factors such as nature, society, and ecology [6]. In 2011, the economists Alkire and Foster proposed the multidimensional poverty index (MPI), bringing significant methodological innovation to the field of poverty measurement [7]. Unlike indicators that are solely focused on income or economic factors, the MPI serves as a comprehensive poverty measurement tool, considering multiple dimensions such as natural conditions, health, education, and living standards. It provides a more global and in-depth analysis of poverty, offering a comprehensive reflection of individual or household poverty across various aspects [8]. Consequently, the MPI has become one of the most widely used tools to analyze regional poverty characteristics. Numerous studies have sought to identify and assess poverty at different scales, including the national, provincial, and municipal levels, using the MPI and achieving significant results [9–11]. However, previous research on the MPI has frequently depended on statistical data, a process that is inherently time-consuming due to the collection and organization involved, leading to data lag. Consequently, outdated data often lead to a struggle to accurately depict the current poverty situation [3,12]. Therefore, it is imperative to explore the use of new types of data for MPI calculations to improve the accuracy and timeliness of poverty assessments.

In contrast to conventional statistical or survey data, remote sensing data offer distinct advantages, such as extensive spatial coverage, frequent updates, cost-effectiveness, and the ability to capture detailed information from otherwise inaccessible or remote areas [13,14]. This enables a more comprehensive and dynamic understanding of the surface of the Earth and socioeconomic phenomena [3]. Nighttime light (NL) data, a prevalent type of remote sensing data, have significant potential to track human-related socioeconomic activities [14]. The intensity of nighttime lights reflects the economic prosperity of a country or region, with numerous studies demonstrating a strong correlation between NL brightness and economic indicators such as GDP and gross regional product (GRP) [14–16]. The spatial characteristics of the NL data provide new opportunities for the quantitative assessment of regional poverty levels. Many studies have been conducted to evaluate poverty based on NL data [16–21]. For instance, Elvidge et al. [16] calculated a global poverty index using Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS) nightlight time data, demonstrating the utility of NL remote sensing data for poverty estimation. Jean et al. [17], utilizing DMSP-OLS nighttime light images to assist in the training of their machine learning model, obtained poverty indices for five African countries, further affirming the effectiveness of NL data for the national assessment of poverty. The use of NL data for the estimation of poverty has also been extended to provincial and municipal scales [18], district and county scales [19–21], and even campus scales [22]. However, most studies have mainly relied on NL data as the sole data source for their evaluation. Considering the complexity of poverty characteristics, it is challenging to effectively identify poverty in theory and practice using a single type of NL data. To address this limitation, integrating additional diverse geographical spatial data is necessary [23]. Various types of geographical spatial data, such as land use and land cover (LULC), points of interest (POI), terrain characteristics, and road network density (RND), not only provide information beyond NL data, but are also often closely associated with economic development [24,25]. Therefore, incorporating them into the model can contribute to improving the accuracy of a poverty regression model. However, the inclusion of additional data, such as geographic spatial data, significantly increases the volume of the data. Furthermore, the relationship between multi-source data and poverty is highly complex, often exhibiting nonlinear characteristics [3,26]. Therefore, traditional linear regression models face challenges in adequately expressing these relationships [19,26]. In recent years, the rapid development of machine learning technology has provided new methods for addressing these challenges [27]. Machine learning techniques can extract relationships between various factors from vast and complex data, thereby establishing relevant prediction models [28–30]. Thus, the introduction of machine learning models

and the fusion of NL and geographic data are critical in assessing poverty. However, most current studies use either a single machine learning model or multiple models with a single data source for poverty assessment modeling, leading to considerable uncertainty in the evaluation results. Therefore, it is necessary to explore methods for poverty assessment using machine learning models combining NL and multiple-source geographic data.

In this study, we propose a novel approach to assess poverty by integrating NL and geospatial data using machine learning models. The MPI, derived from county-level statistical data on the social, economic, and environmental dimensions, serves as an indicator for evaluating poverty levels. Specifically, our objectives are to (1) develop an optimal regression model for the MPI through the integration of multi-source data using machine learning models, (2) determine the relative importance of the features in MPI modeling, and (3) generate a poverty map at the county level, analyzing the spatial distribution of poverty in Fujian province. We anticipate that these findings will provide new insights into poverty estimation and offer vital support for the development of effective poverty alleviation policies.

## 2. Materials and Methods

### 2.1. Study Area

The study area is Fujian province (115°40′~120°50′E, 23°20′~28°40′N), located along the southeast coast of China (Figure 1). Covering around 124,000 km$^2$ of land and an additional 136,000 km$^2$ of maritime territory, the province boasts a coastline stretching over 3700 km. Characterized by varied topography, with higher elevations in the northwest and lower elevations in the southeast, Fujian experiences a warm and humid climate, showcasing a diverse landscape that includes mountains, hills, and plateaus. Fujian province comprises nine prefecture-level cities: Fuzhou, Xiamen, Putian, Quanzhou, Zhangzhou, Longyan, Sanming, Nanping, and Ningde, along with the Pingtan Comprehensive Experimental Zone (Pingtan county), totaling 84 counties, including county-level cities and districts. These administrative divisions underscore the noticeable economic development discrepancies between economically advanced coastal regions (such as Xiamen, Quanzhou, and Fuzhou) and comparatively less developed inland areas (such as Ningde and Nanping). The economic structure in the region has changed from manufacturing and tourism in coastal areas to agriculture and ecological tourism resources in the inland regions, contributing to the intricate socioeconomic landscape of Fujian province. The diverse natural environments, social structures, and economic development of Fujian's counties provide an ideal case study for evaluating and analyzing county-level poverty conditions.

### 2.2. Data

The data used in this study include socioeconomic statistical data, NT remote sensing data, and geospatial data (Table 1). Geospatial data include LULC, digital elevation model (DEM), average building height (ABH), RND, monthly mean temperature (MMT), and POI.

The county-level socioeconomic statistical data for Fujian province in 2022 were sourced from the Statistics Bureau of Fujian Province. The NT data used in this study were derived from National Polar-orbiting Partnership/Visible Infrared Imaging Radiometer (NPP/VIIRS) day/night band (DNB) data, originally provided by the Earth Observation Group of the United States National Oceanic and Atmospheric Administration. These data consist of monthly composite light datasets with a resolution of approximately 500 m, designed for the analysis of human activity [31]. In this study, we utilized the V2.1 NPP/VIIRS NTL composite data for 12 months in 2022, which were processed by the Earth Observation Group of the Colorado School of Mines. LULC data with a spatial resolution of 10 m were acquired from the DAMO Academy Institute's AI Earth team. This dataset includes various types of land cover such as farmland, forests, grasslands, water bodies, and impermeable surfaces. A DEM with a spatial resolution of 30 m was used, obtained through the collaborative efforts of NASA and the National Geospatial-Intelligence Agency (NGA) as part of the Shuttle Radar Topography Mission (SRTM). The average altitude, derived from this

DEM, served as one indicator of natural dimension. The flatland percentage was extracted from this DEM as an independent variable. The average building height data provided a comprehensive nationwide 10 m resolution map of building heights, allowing for a detailed examination of the three-dimensional features of the city. The road network data, including various road types, including highways, urban streets, and rural roads, were downloaded from OpenStreetMap. The monthly mean temperature grid data used in this research were sourced from the National Tibetan Plateau Data Center. This dataset was downscaled for the Chinese region using the Delta spatial downscaling method, based on the global 0.5° climate dataset from the Climate Research Unit (CRU) and the high-resolution climate dataset from WorldClim. The POI data from Amap covered five major categories, including dining services, healthcare facilities, sports and leisure, science and culture, and corporate enterprises.

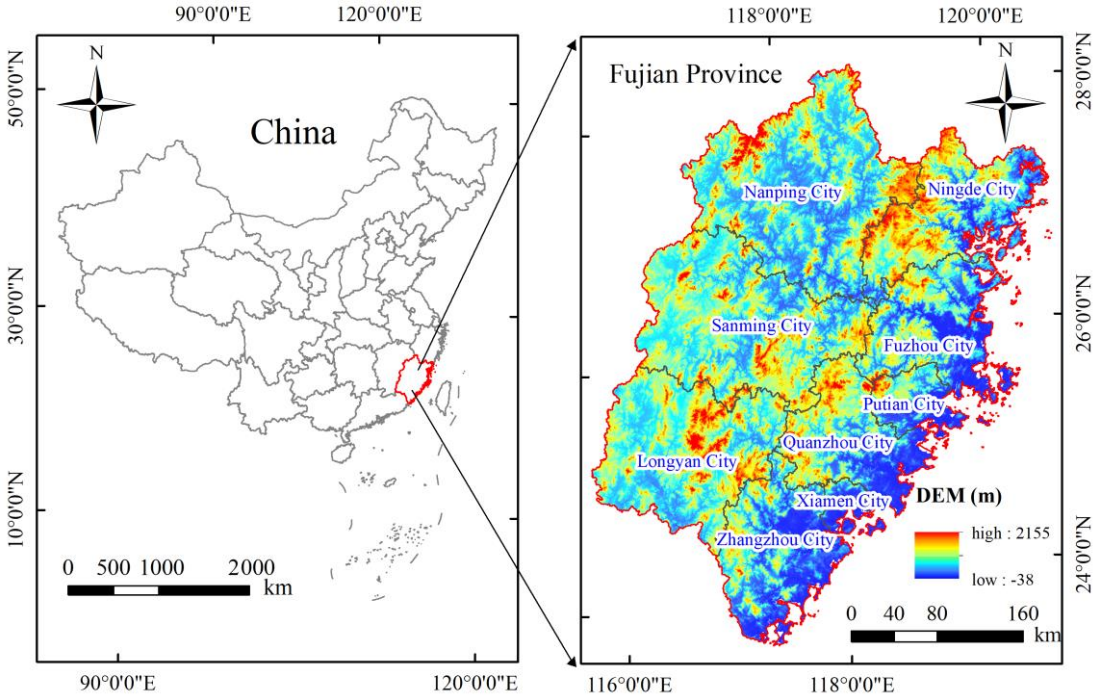



**Figure 1.** The geography of the study area. A digital elevation model (DEM) is a digital representation of ground surface elevation data.

**Table 1.** Descriptions of the datasets.

| Datasets | Description | Year | Source |
|---|---|---|---|
| Socioeconomic statistical data | Indicators of economic and social dimensions | 2022 | Statistics Bureau of Fujian Province (https://tjj.fujian.gov.cn/xxgk/ndsj/ (accessed on 5 July 2023)) |
| DEM | SRTM3, spatial resolution: 30 m, average altitude (natural dimension), and flatland percentage were extracted from DEM | 1999~2000 | Shuttle Radar Topography Mission (http://srtm.csi.cgiar.org/srtmdata/ (accessed on 5 July 2023)) |
| Nighttime light data | National Polar-Orbiting Partnership (NPP)/Visible Infrared Imaging Radiometer Suite (VIIRS) monthly composite light data, V2.1, spatial resolution: ~500 m | 2022 | Colorado School of Mines Earth Observation Group (https://eogdata.mines.edu/nighttime_light/annual/v21/ (accessed on 5 July 2023)) |
| Land use and land cover data | The DAMO Academy Institute's AI Earth team self-researched a 10 m resolution feature classification product for China | 2022 | DAMO Academy (https://engine-aiearth.aliyun.com/#/dataset/DAMO_AIE_CHINA_LC (accessed on 5 July 2023)) |

**Table 1.** *Cont.*

| Datasets | Description | Year | Source |
|---|---|---|---|
| Average building height | Spatial resolution: 10 m | 2022 | Wu et al., 2023 [32] |
| Road network density | Measuring the density of road distribution within a certain area | 2022 | OpenStreetMap (https://www.openstreetmap.org (accessed on 5 July 2023)) |
| Monthly mean temperature | Spatial resolution: 1 km | 2022 | National Tibetan Plateau Data Center (https://data.tpdc.ac.cn/ (accessed on 5 July 2023)) |
| POI | Including catering, leisure, company, education, and medical data | 2022 | Amap (https://www.amap.com/ (accessed on 5 July 2023)) |

### 2.3. Methodology

In this study, we computed the MPI using county-level statistical data from three dimensions (social, economic, and environmental) in Fujian province. Simultaneously, we extracted county-level independent variables from NL and geospatial data. Subsequently, we partitioned the datasets into a training dataset (70%) and a test dataset (30%). Then, we modeled the relationship between MPI and independent variables using both linear regression and machine learning models to identify the most appropriate model. Finally, we applied this optimal model to map the county-level MPI. Figure 2 provides an overview of the workflow.

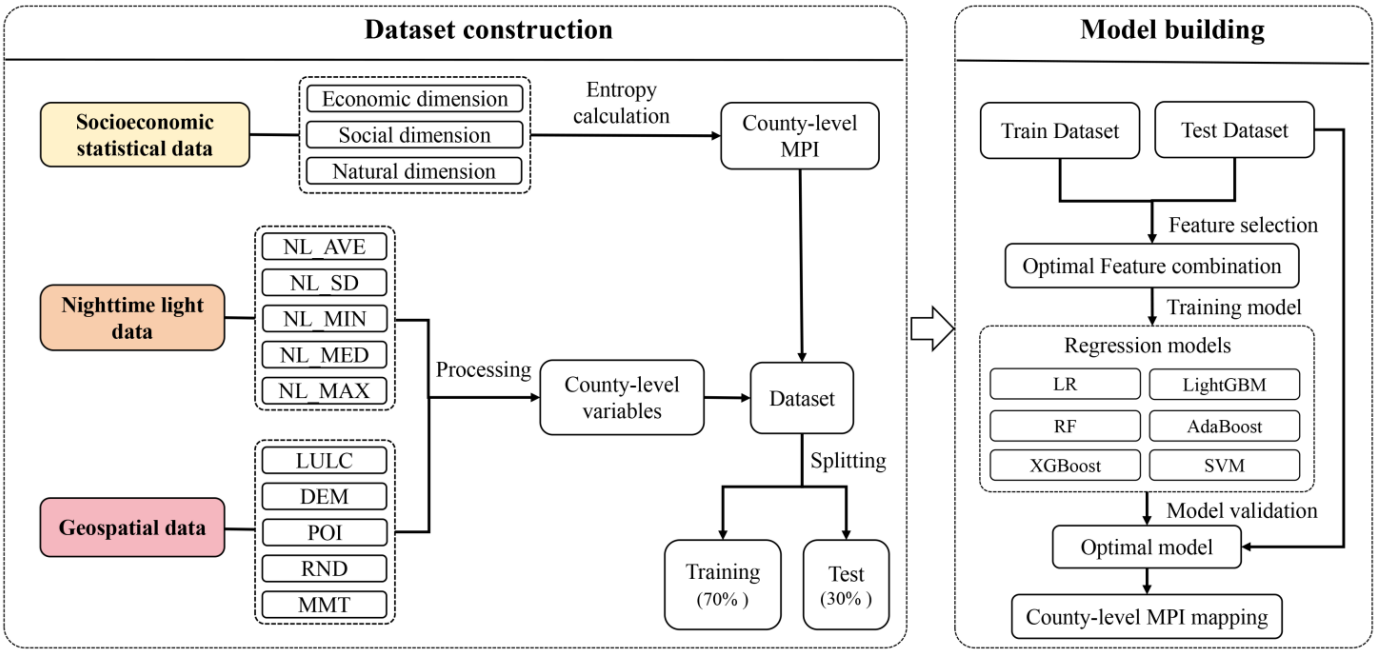



**Figure 2.** Workflow diagram illustrating the processing of the optimal model construction using the multidimensional poverty index (MPI) and independent variables and MPI mapping.

#### 2.3.1. Calculation of County-Level MPI from Statistical Data

To assess poverty levels in the study area, we employed the Alkire–Foster method [7] to calculate the county-level MPI based on statistical data from Fujian province in 2022, which cover three key dimensions, social, economic, and environmental, comprising a total of 17 indices (Table 2). Given the diverse measurement units for these indices, data normalization is necessary. The standardization formulas for both positive and negative indicators are as follows:

**Table 2.** Poverty evaluation indices and their corresponding weights.

| Dimension of Poverty | Index | Correlation to Poverty | Weight |
|---|---|---|---|
| Economic dimension | Gross domestic product | + | 0.056 |
| | Secondary industry | + | 0.057 |
| | Total population at year-end | + | 0.072 |
| | Resident population at year-end | + | 0.069 |
| | Urbanization level | + | 0.030 |
| | Budgetary revenue of local government | + | 0.091 |
| | Value-added tax | + | 0.083 |
| | Budgetary expenditure | + | 0.048 |
| | Educational expenditure | + | 0.059 |
| | Expenditure for agriculture, forestry, and water conservancy | + | 0.030 |
| | Net fixed assets of industrial enterprises above a certain scale | + | 0.059 |
| Social dimension | Number of full-time teachers in general junior high school | + | 0.062 |
| | Number of students enrolled in general senior high school | + | 0.063 |
| | Number of beds in health institutions | + | 0.043 |
| | Registered nurses | + | 0.050 |
| | Total retail sales of consumer goods | + | 0.057 |
| Natural dimension | Average altitude | − | 0.045 |

The formula for a positive index:

$$P_{ij} = \frac{P_{ij}^0 - P_{min}}{P_{max} - P_{min}} \quad (1)$$

The formula for a negative index:

$$P_{ij} = \frac{P_{max} - P_{ij}^0}{P_{max} - P_{min}} \quad (2)$$

where $P_{ij}^0$ and $P_{ij}$ represent the values before and after index standardization, $P_{max}$ corresponds to the maximum value of the index, $P_{min}$ denotes the minimum value of the index, and $P_{max} - P_{min}$ serves as the normalization base.

To enhance the objectivity of the county-level MPI, we adopted the entropy weighting method to calculate the weights for each index (Table 2). The specific process of the entropy weighting method is elucidated in the study conducted by Shi et al. [3]. The county-level MPI was computed using the following formula:

$$MPI_i = \sum_{j=1}^{n} W_j \times P_{ij} \quad (3)$$

where $MPI_i$ represents the value for the $i$th county, $W_j$ represents the weight of the $j$th variable, and $P_{ij}$ represents the standard value of the $j$th variable for the $i$th county.

Numerous studies have demonstrated the effectiveness of the MPI calculated from statistical data in reflecting the level of poverty within a specific area [3,8,9]. Consequently, we used the MPI calculated from statistical data as the actual MPI to serve as a reference for the MPI that was predicted by the linear regression and machine learning models. Figure 3 illustrates the county-level MPI in Fujian province, showcasing values ranging from 0.02 to 0.84 across different counties, indicating significant variations in MPI across the region.

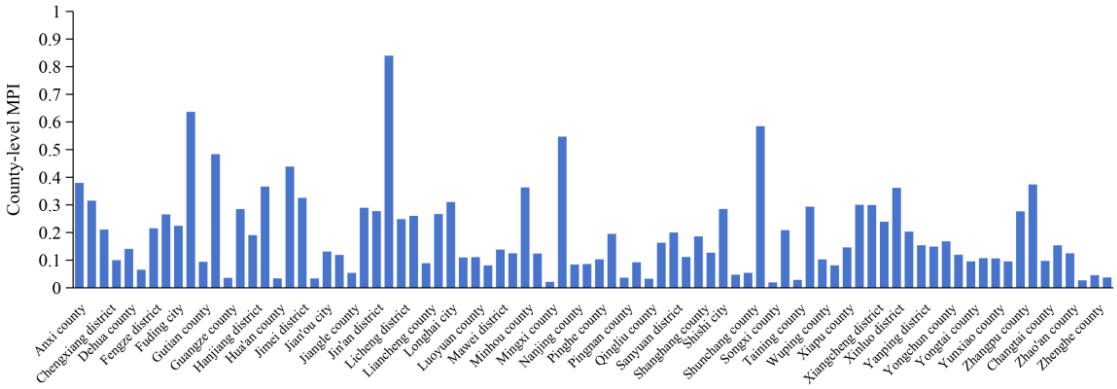



**Figure 3.** County-level MPI map derived from statistical data in Fujian Province.

#### 2.3.2. Independent Variables

In this study, we selected a total of 17 independent variables from nighttime light and geospatial data to model the county-level MPI in Fujian province (Table 3). These 17 features, extensively cited in the literature [3,17–20] as being closely related to poverty levels and demonstrating strong performance in poverty assessment, were selected as explanatory variables for model construction in this study. NL data effectively reflect human activities at night, demonstrating a robust relationship with economic development. Therefore, the average, standard deviation, minimum, median, and maximum of the NL data served as crucial predictors for forecasting the MPI. Other variables were extracted from geospatial data, including LULC, DEM, ABH, RND, MMT, and POI. Specifically, the proportions of forest and arable land in LULC reveal ecological conditions and the development of agriculture and forestry. The percentage of impervious surface from LULC reflects the intensity of human activities in an area, providing information on urbanization levels and socioeconomic conditions. The percentage of flatland is the proportion of land with a slope less than 5 degrees, and generally, land with a smaller slope is more favorable for agricultural production and economic development. The average height of buildings acts as an indicator of the levels of urbanization and regional development. The road network density indicates the development of transportation and regional accessibility. The average temperature illustrates the influence of climate conditions on agriculture and the livelihoods of residents. POI data, including categories such as dining, leisure, companies, education, and healthcare, respectively, represent levels of commercial activity, residents' quality of life, economic development, educational resources, and public health service levels. These multidimensional feature datasets offer a comprehensive and in-depth perspective for analyzing county-level poverty, allowing for more accurate identification and understanding of various factors influencing poverty. All independent variables were resampled to a spatial resolution of 500 m grid data, and the spatial distribution of each independent variable is illustrated in Figure 4.

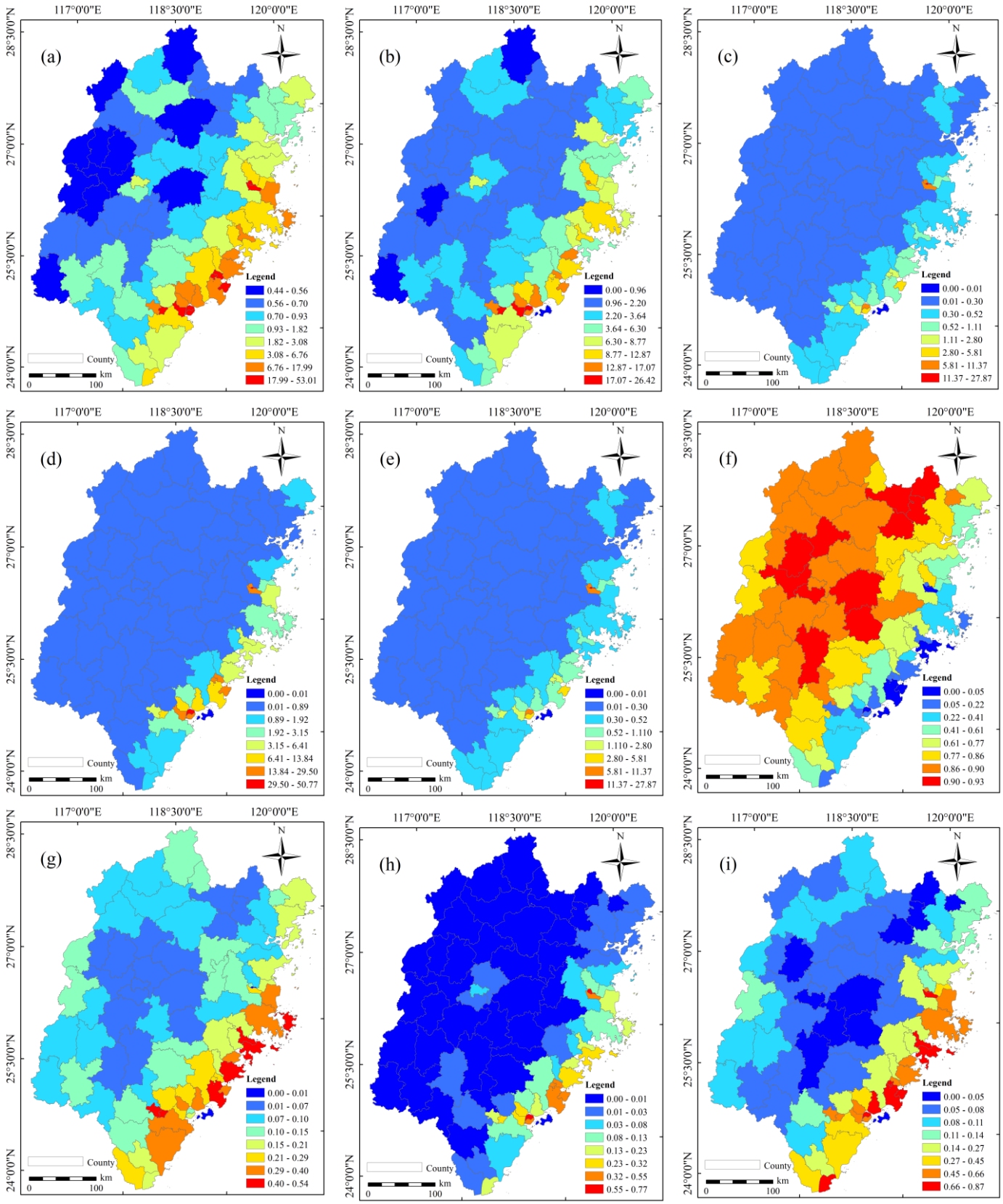



**Figure 4.** *Cont.*

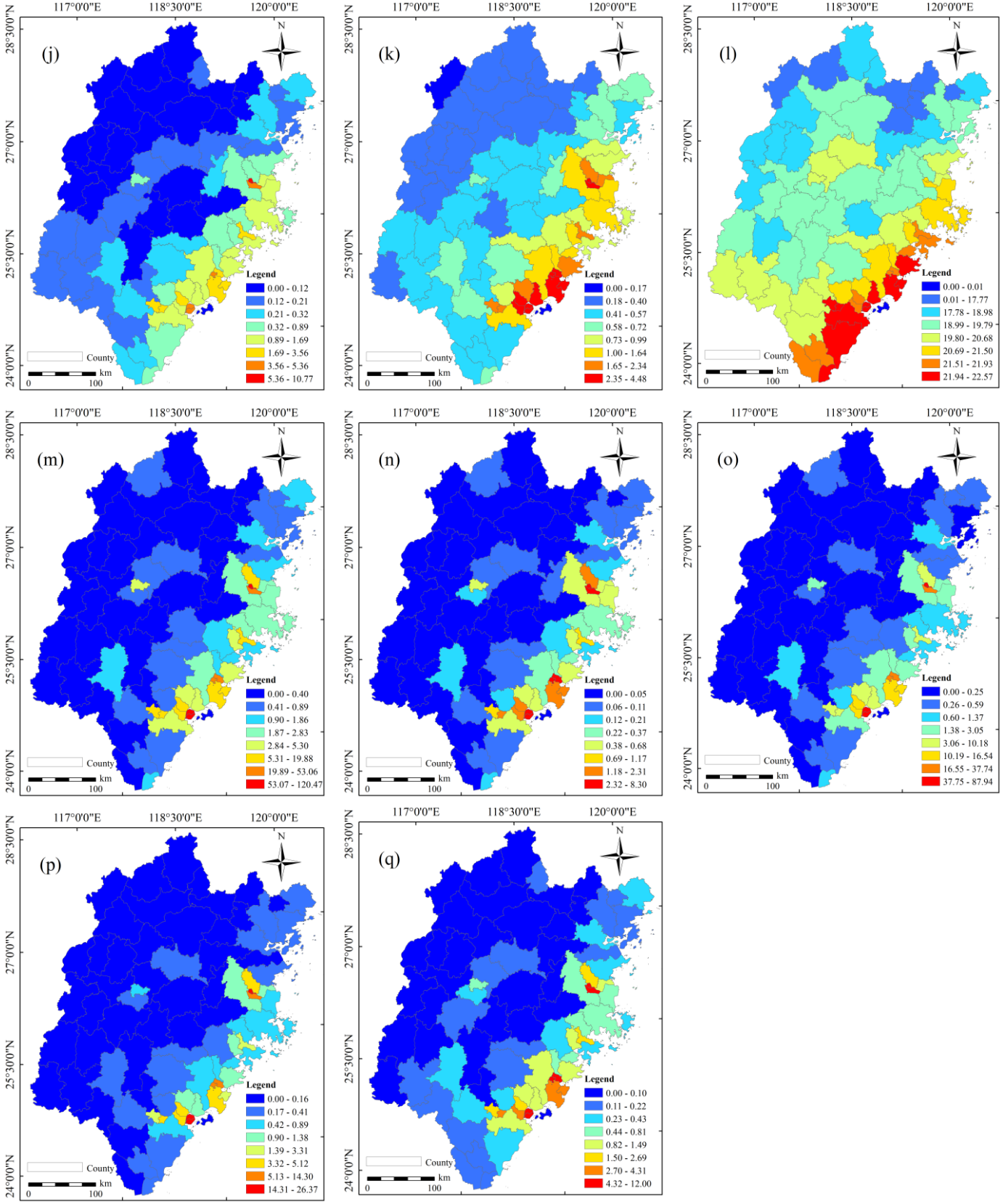



**Figure 4.** Maps of all the independent variables: (**a**) the average of nighttime lights (NL_AVE); (**b**) the standard deviation of nighttime lights (NL_SD); (**c**) the minimum of nighttime lights (NL_MIN); (**d**) the median of nighttime lights (NL_MED); (**e**) the maximum of nighttime lights (NL_MAX); (**f**) the

forest area percentage (LULC_FAP); (**g**) the cropland area percentage (LULC_CAP); (**h**) the impervious surface percentage (LULC_ISP); (**i**) the flatland percentage (FP); (**j**) the average building height (ABH); (**k**) the road network density (RND); (**l**) the monthly mean temperature (MMT); (**m**) catering (POI_Ca); (**n**) leisure (POI_Le); (**o**) company (POI_Co); (**p**) education (POI_Ed); (**q**) medical (POI_Me).

**Table 3.** Summary information of the 17 independent variables.

| Data Type | Variable | Abbreviation | Unit |
|---|---|---|---|
| NL | Average of nighttime lights | NL_AVE | $nW/cm^2/sr$ |
| | Standard deviation of nighttime lights | NL_SD | $nW/cm^2/sr$ |
| | Minimum of nighttime lights | NL_MIN | $nW/cm^2/sr$ |
| | Median of nighttime lights | NL_MED | $nW/cm^2/sr$ |
| | Maximum of nighttime lights | NL_MAX | $nW/cm^2/sr$ |
| LULC | Forest area percentage | LULC_FAP | - |
| | Cropland area percentage | LULC_CAP | - |
| | Impervious surface percentage | LULC_ISP | - |
| - | Flatland percentage | FP | - |
| - | Average building height | ABH | m |
| - | Road network density | RND | % |
| - | Monthly mean temperature | MMT | °C |
| POI | Catering | POI_Ca | PCS |
| | Leisure | POI_Le | PCS |
| | Company | POI_Co | PCS |
| | Education | POI_Ed | PCS |
| | Medical | POI_Me | PCS |

#### 2.3.3. Machine Learning Models and Feature Selection

To establish regression relationships between MPI and independent variables, this study employed traditional linear regression, as well as five commonly used machine learning models, namely, random forest (RF), support vector machine (SVM), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM). The RF model, proposed by Breiman in 2001 [33], is an ensemble learning technique that generates multiple decision trees using randomly selected training samples and features. These decision trees collaborate through a "voting" mechanism to classify the data. RF is known for its ability to handle complex, high-dimensional data and address various linear problems [33]. It offers advantages such as fast computation and resistance to overfitting, making it a popular choice for applications such as remote sensing image classification and regression prediction [34]. SVM, fundamentally a binary classifier, constructs a linear separation hyperplane to categorize data instances. By employing the "kernel trick" to transform the original feature space into a higher-dimensional one, SVM significantly improves its classification capabilities and mitigates overfitting issues in high-dimensional spaces, making it attractive in various application domains [27]. AdaBoost is an ensemble learning algorithm that generates a set of weak learners by assigning weights to the training set. After each weak learning cycle, it increases the weights of misclassified samples, forcing the weak learners to focus on challenging instances. Eventually, it amalgamates all weak learners through weighted aggregation to build a strong learner [35]. It is celebrated for its swift convergence and ease of implementation. XGBoost belongs to the boosting family of ensemble learning. Compared to traditional GBDT algorithms, XGBoost uses a second-order Taylor expansion to approximate the generalization error of the objective function, simplifying the computation of the objective function [36]. It

also introduces regularization terms to reduce model prediction variability and improve resilience against overfitting. XGBoost offers benefits like rapid computation, superior performance, easy parameter tuning, and efficient handling of large datasets [35]. LightGBM is another member of the boosting framework. This algorithm builds on traditional GBDT by introducing gradient-based one-sided sampling and feature merging techniques [36]. It uses a histogram-based algorithm instead of conventional presorted methods, retaining information while reducing the dimension of the features to ensure efficiency and prevent overfitting.

The handling of high-dimensional data presents a challenge to researchers and engineers in the fields of machine learning and data mining. Feature selection offers an effective approach to address this problem by eliminating irrelevant and redundant data [6]. One of the most popular feature selection methods, recursive feature elimination (RFE) [36], is frequently used in conjunction with other machine learning algorithms to construct more efficient classifiers. Recursive feature elimination with cross-validation (RFECV) enhances the process by averaging model performance through cross-validated RFE, helping to automatically determine the optimal number of features that yield the highest classification accuracy [37]. Hence, we used RFECV to perform feature selection in this study. To assess the importance of various environmental variables within our MPI regression model, the relative importance of each variable was evaluated using the mean decrease in Gini (MDG) method [33].

In this study, we used Python 3.8 as the programming environment and implemented relevant algorithms and models using Scikit-learn (sklearn), a well-known machine learning library.

#### 2.3.4. Model Validation

This study adopts the coefficient of determination ($R^2$), the mean absolute error (MAE), and the root mean square error (RMSE) as metrics to quantify model accuracy. The expressions are as follows:

$$R^2 = \frac{\sum_{i=1}^{n} \left(MPI_{p,i} - \overline{MPI}\right)^2}{\sum_{i=1}^{n} \left(MPI_{t,i} - \overline{MPI}\right)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left|MPI_{t,i} - MPI_{p,i}\right| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(MPI_{t,i} - MPI_{p,i}\right)^2} \tag{6}$$

where $n$ is the number of counties, $MPI_{t,j}$ is the actual MPI of the $i$th county (calculated from the county-level statistical data), $MPI_{p,j}$ is the predicted MPI of the $i$th county, and $\overline{MPI}$ is the actual average MPI of the counties.

## 3. Results

### *3.1. Model Performance*

To identify the optimal model, this study conducted variable selection during the modeling process. The RFECV method was used to explore various combinations of the 17 independent variables, with the cross-validation parameter configured to 5 and the scoring metric set to the MAE. The results revealed a notable performance disparity among different models as the number of features varied (Figure 5). It was evident that the optimal variable combinations differed for various models (Figure 5). As the number of features increased, the MAE for the LR model initially decreased, then increased, reaching its lowest value when the number of independent variables was four. The three machine learning models (RF, XGBoost, and AdaBoost) exhibited a similar changing trend in MAE with an increase in the number of variables, reaching the minimum MAE with 10 or 11 independent variables. The MAE fluctuations in the LightGBM model suggested a high sensitivity

to changes in feature number, with the minimum MAE occurring when the number of features was five. Meanwhile, the SVM demonstrated a noticeable trend in MAE variation, characterized by an initial decrease, reaching its minimum when the number of variables was two, followed by an increase, and eventually stabilizing at a relatively higher level.

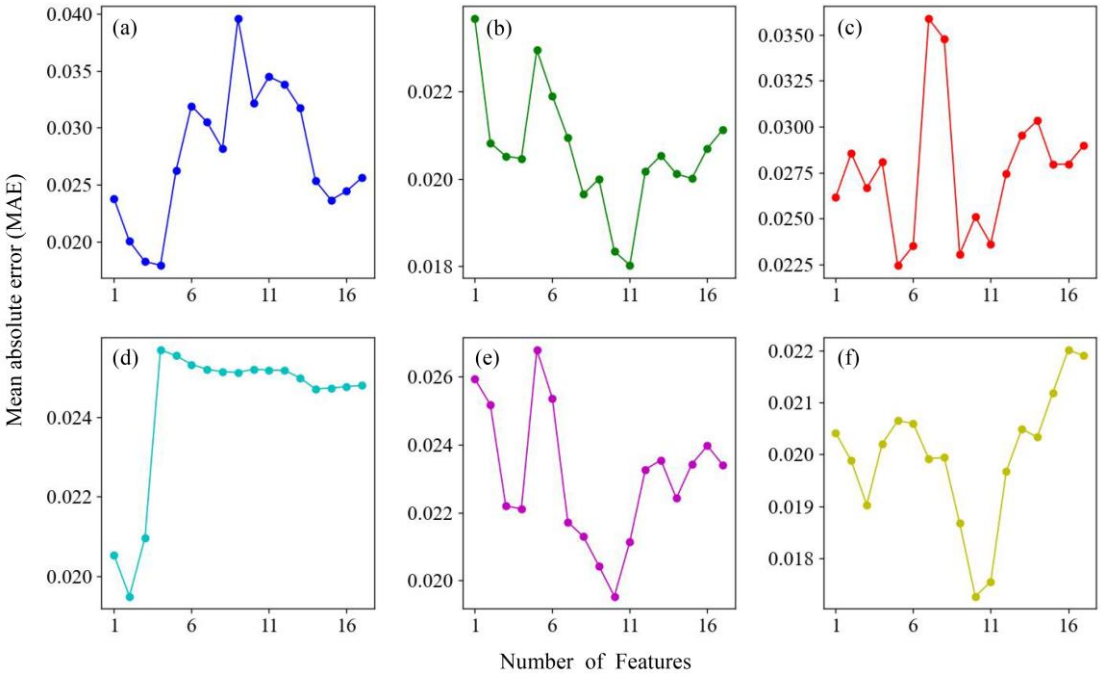



**Figure 5.** Variation in mean absolute error (MAE) for various regression models with the number of features: (**a**) linear regression (LR); (**b**) random forest (RF); (**c**) light gradient boosting machine (LightGBM); (**d**) support vector machine (SVM); (**e**) extreme gradient boosting (XGBoost); (**f**) adaptive boosting (AdaBoost).

Figure 6 shows the scatter plot between the actual MPI and the predicted MPI at the county level, derived from the LR model and five machine learning models. Generally, the machine learning models ($R^2$: 0.555~0.928) exhibited higher accuracy compared to the LR model ($R^2$: 0.517). Among them, the RF model outperformed the others, with an $R^2$ of 0.928, an MAE of 0.030, and an RMSE of 0.037. The scatters of the RF model were considerably closer to the 1:1 line than those of the other five models. Despite a few instances of high-value samples being underestimated by the RF model, the majority of samples aligned closely with the 1:1 line, indicating a strong agreement between the actual and predicted MPI values. In contrast, the LR model exhibited a more scattered plot, revealing relatively high errors in both high-value and low-value samples.

### *3.2. Relative Importance of Variables*

Figure 7a illustrates the contribution and impact trends of different features on the MPI prediction values using the RF model. For the features NL_MAX and POI_Ed, positive SHAP values (in red) are predominantly located to the left of the model's predicted average in the majority of samples. Conversely, negative values (in blue) are situated to the right of the average, indicating high consistency between these feature values and the model's predictions and thus suggesting a positive influence on the prediction values. In contrast, NL_MIN and POI_Le exhibit a color change that is opposite to the aforementioned features, signifying a negative correlation with prediction values across different samples. Additionally, features such as LULC_CAP, MMT, and FP consistently display a blue coloration in most instances, suggesting a relatively minor impact on the model's predictions.

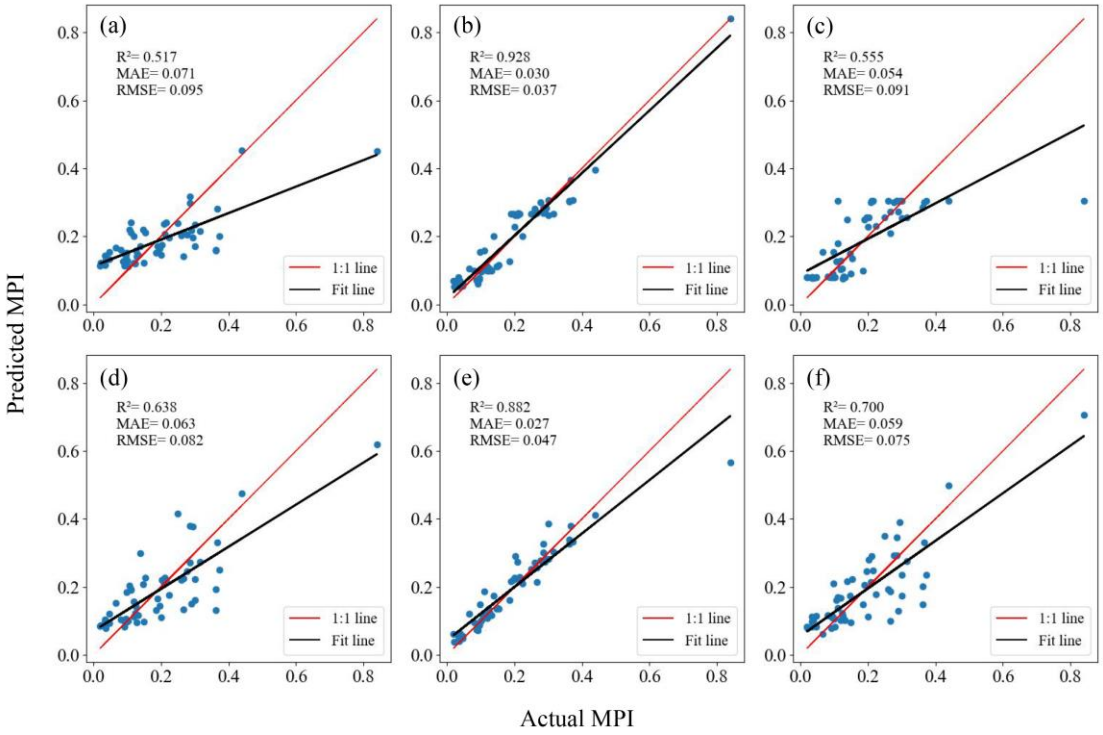



**Figure 6.** Scatter diagrams between the actual multidimensional poverty index (MPI) and the predicted MPI at the county level for various models: (**a**) linear regression (LR); (**b**) random forest (RF); (**c**) light gradient boosting machine (LightGBM); (**d**) support vector machine (SVM); (**e**) extreme gradient boosting (XGBoost); (**f**) adaptive boosting (AdaBoost).

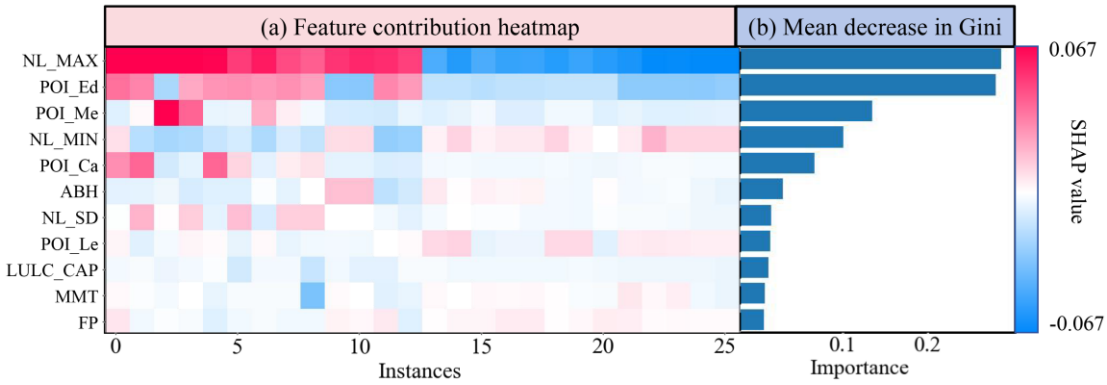



**Figure 7.** The relative importance of the variables in the regression based on the random forest (RF) model.

Figure 7b shows the relative importance of each feature in the MPI modeling using the RF algorithm, based on the mean decrease in Gini. Among the top five most important variables, two (NL_MAX and NL_MIN) were derived from NL data, while three (POI_Ed, POI_Me, and POI_Ca) were from geographical spatial data, emphasizing the significant roles played by nighttime light data and geographical spatial data in MPI modeling. Particularly, NL_MAX, extracted from NL data, has the highest mean decrease value in Gini compared to other variables, making it the most crucial feature for assessing county-level poverty in Fujian province. Following NL_MAX are two POI variables (POI_Ed and POI_Me). It is also noteworthy that among the five nighttime light data variables, three (NL_MAX, NL_MIN, and NL_SD) have relatively high importance rankings. Regarding geographical spatial data features, POI variables have prominent positions, with four of them being in the top eight. The sole variable from LULC, LULC_CAP, ranks relatively

lower, indicating a comparatively smaller impact on the MPI modeling. Among other geographical spatial data features, ABH holds the sixth position, while MMT and FP exhibit relatively lower values for the mean decrease in Gini, signifying their less significant impacts on the MPI estimate compared to other variables.

### *3.3. MPI Mapping and Its Spatial Distribution*

Figure 8 illustrates the actual and predicted MPI at the county level in Fujian province in 2022. Darker shades of blue indicate higher poverty levels, while a deeper red signifies lower poverty, and yellow represents moderate poverty. Despite noticeable color variations in a few counties, particularly in the southern region where underestimation is apparent, the majority of counties exhibit relatively consistent colors between the actual and predicted maps. This suggests that the predicted MPI values for most counties closely align with the actual measurements. Evidently, the RF model demonstrates robust predictive capabilities, producing an accurate county-level MPI distribution map that effectively portrays the spatial distribution of poverty in Fujian province. The MPI prediction map unveils significant spatial variation in the MPI in different counties (Figure 8b). Counties with higher levels of poverty are mainly concentrated in the western, northern, and central inland areas, grappling with various challenges, including inadequate infrastructure and deficiencies in education and healthcare services. In contrast, counties with lower poverty levels are predominantly located in the southeast coastal areas, indicating relative advantages in economic development, education, and health in these regions.

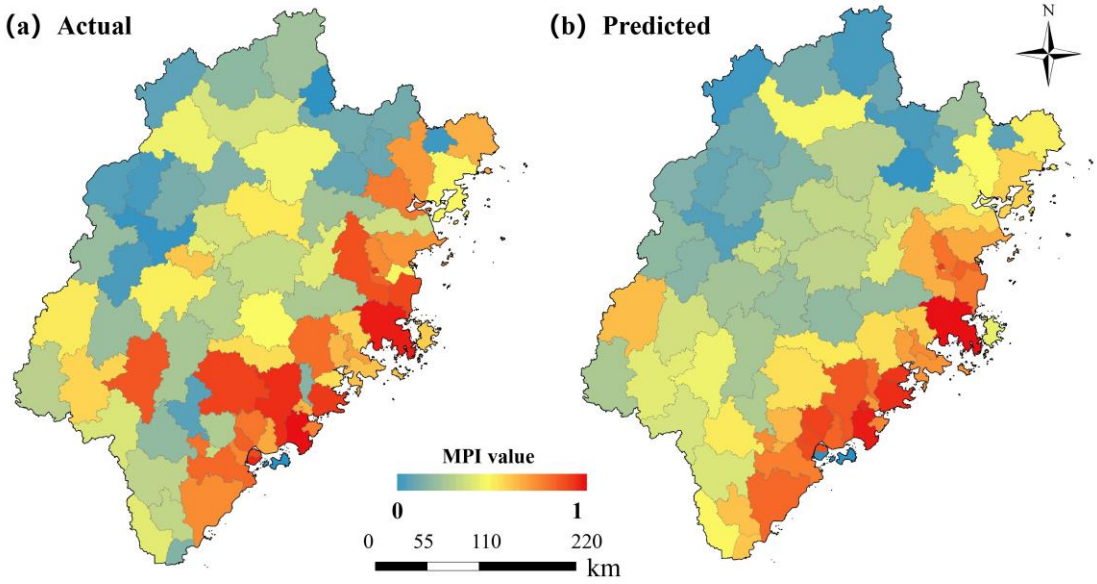



**Figure 8.** Map of the actual (**a**) and predicted (**b**) multidimensional poverty index (MPI) at the county level for Fujian province in 2022. The predicted MPI was generated from the optimal model (random forest).

## 4. Discussion

### *4.1. Effectiveness of RF Model*

The MPI, as a crucial metric for assessing poverty levels, traditionally relies heavily on various statistical data. However, these data sources may suffer from incompleteness or obsolescence, mainly due to factors such as cost constraints [17,20], particularly in recent years. Consequently, this study seeks to model an MPI by integrating NL data and geospatial data using machine learning models. The results indicate that machine learning models exhibit a higher accuracy than the traditional linear regression model, suggesting the effectiveness of the machine learning algorithm in simulating a county-level MPI based on selected independent variables.

This study demonstrates that the predictive accuracy of the RF model ($R^2$ = 0.928) is markedly higher than that of the LR model ($R^2$ = 0.517), suggesting potential differences in their capabilities in handling multidimensional and highly collinear data. The LR algorithm struggles to effectively address the intricate relationships between the MPI and predictors, which are influenced by varying economic development, geographical factors, and natural environments. In contrast, the RF algorithm displays lower sensitivity to data multicollinearity and exhibits robustness in capturing complex nonlinear relationships involving multiple variables [27,33,34]. Moreover, the RF model's resilience to noise and tolerance to outliers make it more reliable than the LR model in the analysis of socioeconomic data [15]. Furthermore, our findings indicate that the RF model's predictive accuracy surpasses that of other machine learning models (SVM, LightGBM, XGBoost, and AdaBoost). This superiority can be attributed to the RF model's generation of a large number of decision trees at each node through the random selection of sample training subsets and variable subsets for splitting [30,34]. This mitigates sensitivity to the quality of training samples and overfitting, thereby enhancing the accuracy and applicability in handling complex feature datasets and limited training samples. While gradient boosting algorithms such as XGBoost and LightGBM typically demonstrate exceptional predictive capabilities when handling high-dimensional and large-scale datasets [35,36], the relatively small sample size in this study may limit the models' ability to fully capture the underlying patterns and complex relationships between features, thus limiting its predictive accuracy.

The diversification of multi-dimensional feature data played a crucial role in achieving high-precision regression results for the model in this study. Previous studies on poverty prediction relied mainly on NL data as the only data source to map the MPI [20,21]. While the nighttime light intensity effectively reflects economic activities, its exclusive use falls short in accurately estimating a county-level MPI due to the intricate interplay of economic, social, and natural factors. In this study, in addition to NL data, geospatial data, including LULC, POI, road network density, monthly mean temperature, and DEM, provided additional vital information for MPI modeling. Therefore, the integration of NL and geospatial data from multiple sources has proven to be effective and efficient in the examination of poverty issues at the county level.

### *4.2. Importance of Variables*

This study reveals that NL_MAX has the greatest influence on the county-level MPI that is predicted by the model. This is mainly because NL_MAX serves as a robust indicator of economic activity. In most cases, the areas with the highest nighttime brightness correspond to regions with concentrated commercial activities, industrial development, or a high population density, which are all indicative of economic prosperity [38]. Therefore, NL_MAX accurately captures crucial information about regional economic vitality, which is particularly essential in predicting poverty levels. NL_MAX also reflects the coverage of robust infrastructure and public services, serving to some extent as a comprehensive socioeconomic indicator. Additionally, NL remote sensing data often exhibit high resolution and extensive spatial coverage, making them more reliable and readily accessible indicators in research and less susceptible to seasonal or short-term fluctuations [10,12,38]. Compared to other indicators, NL_MAX demonstrates higher data quality and availability, contributing significantly more to poverty models than other variables. Furthermore, information related to education (POI_Ed) and healthcare (POI_Me) also influences the model's predictions of poverty. This is attributed to their critical roles in measuring the level of socioeconomic development. The distribution of education resources is directly correlated with the formation of regional human capital and economic potential, while the accessibility of healthcare resources is the cornerstone of residents' health and quality of life [19,39]. Regions lacking education and healthcare services are often associated with high levels of poverty, as these essential services are crucial for enhancing income, promoting employment, and ensuring basic living conditions [1]. In machine learning models, these point-based pieces of information provide a comprehensive view of a re-

gion's infrastructure and social welfare status, becoming significant predictors in poverty predictions. The layout of educational and healthcare resources is also frequently a focal point of policy decisions, reflecting the government's emphasis on poverty alleviation and regional development strategies, further highlighting their importance in predictive models [1,20,21]. Therefore, in analyzing and predicting multi-dimensional indicators of poverty, the feature data of education and healthcare points are indispensable key variables.

In contrast, the impact of forest coverage (LULC_FAP) on poverty prediction in the model is minimal. This may be because the connection between forest coverage and residents' direct economic well-being is not as direct as those of education or healthcare. Although forest ecosystem services have important long-term effects on socioeconomic factors, including the air quality, water source protection, and biodiversity conservation, they may not directly influence the economic output or poverty status in the short term [40,41]. Forest coverage itself more accurately reflects a region's ecological conditions rather than direct economic activities, making it weaker compared to the direct economic driving factors of poverty levels. Furthermore, the economic and social value of forest coverage varies by region; in some areas, it is vital for livelihoods and ecosystem services, while in others, it may not be significantly associated with poverty [40,41]. When assessing poverty levels, forest coverage may not be a strong influencing factor compared to more directly linked socioeconomic indicators such as income, education, health, and infrastructure. Thus, despite the significant impact of forests on environmental protection and ecological balance, their role in specific predictive models of poverty may be overshadowed by other variables with more direct economic connections.

### *4.3. Limitations*

Although the model selected in this study demonstrates several strengths in simulating a county-level MPI, it also exhibits certain limitations. First, although we have included numerous independent variables related to poverty, there are still significant variables that have not been incorporated. The addition of variables such as location-based social media (LBSM) data [42], which are used to analyze user activities and provide context for location-based services, could potentially further enhance the model's precision. Second, it is essential to acknowledge that our study utilized data from only one year (2022) for model construction. This limitation somewhat constrains the model's representativeness and its ability to generalize to different years. Consequently, there is a pressing need to develop temporal models to better capture trends in the MPI and improve the model's applicability. Third, the spatial resolution of the data used in this study is comparatively coarse, especially for the NL data, with a spatial resolution of approximately 500 m. This limitation results in the neglect of many details, consequently restricting the accuracy of our model. Therefore, in the future, adopting NL data with a higher spatial or spectral resolution, such as the Luojia 1-01 products, which boast a spatial resolution of ~130 m and can better represent the spatial differences in light [42], would contribute to further refinement of the accuracy of the model.

## 5. Conclusions

This study proposes an approach that integrates NL remote sensing data and geospatial data using machine learning models to evaluate county-level poverty. The results demonstrate that the RF model achieved a significantly higher accuracy ($R^2$ = 0.928, MAE = 0.030, RMSE = 0.037) compared to other models. The robust accuracy suggests the efficacy of the RF algorithm in stimulating the county-level MPI based on the selected independent variables. Notably, the top five most important variables include two (NL_MAX and NL_MIN) from the NL data and three (POI_Ed, POI_Me, and POI_Ca) from the geographical spatial data, highlighting the significant roles played by NL data and geographical spatial data in MPI modeling. The developed county-level MPI map provides a detailed spatial distribution of the poverty in Fujian province, serving as a vital reference for formulating socioeconomic development strategies and policies. Furthermore,

the methodology presented in this study, employing NL data and geospatial data through the RF model, offers guidance for mapping MPIs at a fine spatial scale in other regions. In conclusion, the combination of machine learning techniques, NL data, and geospatial data enhances the precision of identifying impoverished areas, offering vital support for the development of effective poverty alleviation policies and interventions, especially in regions with limited statistical data.

**Author Contributions:** Conceptualization, X.Z. and W.Z.; Methodology, W.Z., X.Z. and H.Z.; Software, H.D.; Formal analysis, H.D.; Data curation, X.Z.; Writing—original draft, X.Z.; Writing—review and editing, W.Z. and H.Z.; Funding acquisition, H.D. and H.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Fujian Province Education and Scientific Research Project for Young and Middle-aged Teachers (Science and Technology) (Grant No. JAT220263), the Tibet Autonomous Region Science and Technology Plan Project Key Project (XZ202201ZY0003G), Natural Science of the Education Department of Sichuan Province (18ZA0047) and National Natural Science Foundation of China (72073029).

**Data Availability Statement:** See Section 2.2 Data for information about how to access the data used in this study.

**Acknowledgments:** The authors extend their gratitude to Yanzhen Hong from Fujian Agriculture and Forestry University for her invaluable contributions. Hong served as a scientific advisor and provided extensive support in writing and technical editing of the manuscript.

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

## References

1. Liu, M.; Feng, X.; Zhao, Y.; Qiu, H. Impact of Poverty Alleviation through Relocation: From the Perspectives of Income and Multidimensional Poverty. *J. Rural Stud.* **2023**, *99*, 35–44. [CrossRef]
2. Li, G.; Chang, L.; Liu, X.; Su, S.; Zhou, C.H.; Cai, Z.; Huang, X.; Li, B. Monitoring the Spatiotemporal Dynamics of Poor Counties in China: Implications for Global Sustainable Development Goals. *J. Clean. Prod.* **2019**, *227*, 392–404. [CrossRef]
3. 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]
4. Ferreira, F.H.G.; Leite, P.G.; Ravallion, M. Poverty Reduction without Economic Growth?: Explaining Brazil's Poverty Dynamics, 1985–2004. *J. Dev. Econ.* **2010**, *93*, 20–36. [CrossRef]
5. Labar, K.; Bresson, F. A Multidimensional Analysis of Poverty in China from 1991 to 2006. *China Econ. Rev.* **2011**, *22*, 646–668. [CrossRef]
6. Vollmer, F.; Alkire, S. Consolidating and Improving the Assets Indicator in the Global Multidimensional Poverty Index. *World Dev.* **2022**, *158*, 105997. [CrossRef]
7. Alkire, S.; Foster, J. Counting and Multidimensional Poverty Measurement. *J. Public Econ.* **2011**, *95*, 476–487. [CrossRef]
8. Alkire, S.; Nogales, R.; Quinn, N.N.; Suppa, N. On Track or Not? Projecting the Global Multidimensional Poverty Index. *J. Dev. Econ.* **2023**, *165*, 103150. [CrossRef]
9. Alkire, S.; Roche, J.M.; Vaz, A. Changes over Time in Multidimensional Poverty: Methodology and Results for 34 Countries. *World Dev.* **2017**, *94*, 232–249. [CrossRef]
10. Alkire, S.; Apablaza, M.; Chakravarty, S.; Yalonetzky, G. Measuring Chronic Multidimensional Poverty. *World Dev.* **2017**, *39*, 983–1006. [CrossRef]
11. Ahmed, A.; Asabere, S.B.; Adams, E.A.; Abubakari, Z. Patterns and Determinants of Multidimensional Poverty in Secondary Cities: Implications for Urban Sustainability in African Cities. *Habitat Int.* **2023**, *134*, 102775. [CrossRef]
12. Lin, Y.; Zhang, T.; Liu, X.; Yu, J.; Li, J.; Gao, K. Dynamic Monitoring and Modeling of the Growth-Poverty-Inequality Trilemma in the Nile River Basin with Consistent Night-Time Data (2000–2020). *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102903. [CrossRef]
13. Ma, Y.; Chen, S.; Ermon, S.; Lobell, D.B. Transfer Learning in Environmental Remote Sensing. *Remote Sens. Environ.* **2024**, *301*, 113924. [CrossRef]
14. Li, S.; Cao, X.; Zhao, C.; Jie, N.; Liu, L.; Chen, X.; Cui, X. Developing a Pixel-scale Corrected Nighttime Light Dataset (PCNL, 1992–2021) Combining DMSP-OLS and NPP-VIIRS. *Remote Sens.* **2023**, *15*, 3925. [CrossRef]
15. Liang, H.; Guo, Z.; Wu, J.; Chen, Z. GDP Spatialization in Ningbo City Based on NPP-VIIRS Night-Time Light and Auxiliary Data Using Random Forest Regression. *Adv. Space Res.* **2020**, *65*, 481–493. [CrossRef]
16. 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]
17. Jean, N.; Burke, M.; Xie, M.; Davis, W.M.; Lobell, D.B.; Ermon, S. Combining Satellite Imagery and Machine Learning to Predict Poverty. *Science* **2016**, *353*, 790–794. [CrossRef] [PubMed]

18. Wang, W.; Cheng, H.; Zhang, L. Poverty Assessment Using DMSP/OLS Night-Time Light Satellite Imagery at a Provincial Scale in China. *Adv. Space Res.* **2012**, *49*, 1253–1264. [CrossRef]
19. 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.* **2021**, *10*, 11. [CrossRef]
20. Shen, Y.; Chen, X.; Yao, Q.; Ding, J.; Lai, Y.; Rao, Y. Examining the Impact of China's Poverty Alleviation on Nighttime Lighting in 831 State-Level Impoverished Counties. *Land* **2023**, *12*, 1128. [CrossRef]
21. Xu, J.; Song, J.; Li, B.; Liu, D.; Cao, X. Combining Night Time Lights in Prediction of Poverty Incidence at the County Level. *Appl. Geogr.* **2021**, *135*, 102552. [CrossRef]
22. Li, X.; Levin, N.; Xie, J.; Li, D. Monitoring Hourly Night-Time Light by an Unmanned Aerial Vehicle and Its Implications to Satellite Remote Sensing. *Remote Sens. Environ.* **2020**, *247*, 111942. [CrossRef]
23. 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]
24. Shao, Z.; Li, X. Multi-scale estimation of poverty rate using night-time light imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *121*, 103375. [CrossRef]
25. Li, M.; Lin, J.; Ji, Z.; Chen, K.; Liu, J. Grid-Scale Poverty Assessment by Integrating High-Resolution Nighttime Light and Spatial Big Data—A Case Study in the Pearl River Delta. *Remote Sens.* **2023**, *15*, 4618. [CrossRef]
26. Hu, S.; Ge, Y.; Liu, M.; Ren, Z.; Zhang, X. Village-level Poverty Identification Using Machine Learning, High-Resolution Images, and Geospatial Data. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102694. [CrossRef]
27. Ye, Z.; Yang, K.; Lin, Y.; Guo, S.; Sun, Y.; Chen, X.; Lai, R.; Zhang, H. A Comparison Between Pixel-Based Deep Learning and Object-Based Image Analysis (OBIA) for Individual Detection of Cabbage Plants Based on UAV Visible-Light Images. *Comput. Electron. Agric.* **2023**, *209*, 107822. [CrossRef]
28. Meyer, H.; Pebesma, E. Machine Learning-Based Global Maps of Ecological Variables and the Challenge of Assessing them. *Nat. Commun.* **2022**, *13*, 2208. [CrossRef] [PubMed]
29. Ye, Z.; Wei, J.; Lin, Y.; Guo, Q.; Zhang, J.; Zhang, H.; Deng, H.; Yang, K. Extraction of Olive Crown Based on UAV Visible Images and the U2-Net Deep Learning Model. *Remote Sens.* **2022**, *14*, 1523. [CrossRef]
30. Feng, C.; Zhang, W.; Deng, H.; Dong, L.; Zhang, H.; Tang, L.; Zheng, Y.; Zhao, Z. A Combination of OBIA and Random forest Based on Visible UAV Remote Sensing for Accurately Extracted Information about Weeds in Areas with Different Weed Densities in Farmland. *Remote Sens.* **2023**, *15*, 4696. [CrossRef]
31. Elvidge, C.D.; Zhizhin, M.; Ghosh, T.; Hsu, F.-C.; Taneja, J. Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019. *Remote Sens.* **2021**, *13*, 922. [CrossRef]
32. Wu, W.-B.; Ma, J.; Banzhaf, E.; Meadows, M.E.; Yu, Z.-W.; Guo, F.-X.; Sengupta, D.; Cai, X.-X.; Zhao, B. A First Chinese Building Height Estimate at 10 m Resolution (CNBH-10 m) Using Multi-Source Earth Observations and Machine Learning. *Remote Sens. Environ.* **2023**, *291*, 113578. [CrossRef]
33. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
34. Guo, Q.; Zhang, J.; Guo, S.; Ye, Z.; Deng, H.; Hou, X.; Zhang, H. Urban Tree Classification Based on Object-Oriented Approach and Random Forest Algorithm Using Unmanned Aerial Vehicle (UAV) Multispectral Imagery. *Remote Sens.* **2022**, *14*, 3885. [CrossRef]
35. Li, S.; Jiang, B.; Liang, S.; Peng, J.; Liang, H.; Han, J.; Yin, X.; Yao, Y.; Zhang, X.; Cheng, J.; et al. Evaluation of Nine Machine Learning Methods for Estimating Daily Land Surface Radiation Budget from Modis Satellite Data. *Int. J. Digit. Earth.* **2022**, *15*, 1784–1816. [CrossRef]
36. Fu, B.; Zuo, P.; Liu, M.; Lan, G.; He, H.; Lao, Z.; Zhang, Y.; Fan, D.; Gao, E. Classifying Vegetation Communities Karst Wetland Synergistic Use of Image Fusion and Object-Based Machine Learning Algorithm with Jilin-1 and UAV Multispectral Images. *Ecol. Indic.* **2022**, *140*, 108989. [CrossRef]
37. Xing, H.; Niu, J.; Feng, Y.; Hou, D.; Wang, Y.; Wang, Z. A Coastal Wetlands Mapping Approach of Yellow River Delta with a Hierarchical Classification and Optimal Feature Selection Framework. *Catena* **2023**, *223*, 106897. [CrossRef]
38. Liu, S.; Wang, C.; Chen, Z.; Li, Q.; Wu, Q.; Li, Y.; Wu, J.; Yu, B. Enhancing Nighttime Light Remote Sensing: Introducing the Nighttime Light Background Value (NLBV) for Urban Applications. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *126*, 103626. [CrossRef]
39. Zuo, H.; Li, S.; Ge, Z.; Chen, J. The Impact of Education on Relative Poverty and Its Intergenerational Transmission—Causal Identification Based on the Compulsory Education Law. *China Econ. Rev.* **2023**, *82*, 102071. [CrossRef]
40. Briner, S.; Elkin, C.; Huber, R. Evaluating the Relative Impact of Climate and Economic Changes on Forest and Agricultural Ecosystem Services in Mountain Regions. *J. Environ. Manag.* **2013**, *129*, 414–422. [CrossRef] [PubMed]
41. Qi, H.; Sun, L.; Long, F.; Gao, X.; Hu, L. Does Forest Resource Protection Under the Carbon Neutrality Target Inhibit Economic Growth? Evidence of Poverty-Stricken County from China. *Front. Environ. Sci.* **2022**, *10*, 858632. [CrossRef]
42. Wang, L.; Fan, H.; Wang, Y. Improving Population Mapping Using Luojia 1-01 Nighttime Light Image and Location-Based Social Media Data. *Sci. Total Environ.* **2020**, *730*, 139148. [CrossRef] [PubMed]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.