



# Geographically Weighted Method Integrated with Logistic Regression for Analyzing Spatially Varying Accuracy Measures of Remote Sensing Image Classification

Varun Narayan Mishra<sup>1</sup> · Vivek Kumar<sup>2</sup> · Rajendra Prasad<sup>2</sup> · Milap Punia<sup>3</sup>

Received: 12 April 2019 / Accepted: 7 December 2020 / Published online: 1 February 2021  
© Indian Society of Remote Sensing 2021

## Abstract

The accuracy of thematic information extracted from remote sensing image is assessed recurrently using the confusion matrix method. But the accuracies have been criticized as a consequence of its aspatial nature. The work presented here describes a geographically weighted method combined with logistic regression for producing and visualizing the spatially distributed accuracy measures across the landscape. The outcomes compare the standard confusion matrix-based accuracy measures with those that have been permitted to differ locally. Furthermore, statistical parameters, i.e. Akaike information criterion, adjusted squared correlation coefficient ( $R^2$ ) and residual sum of squares (RSS) were employed to compare the performance of geographically weighted logistic regression (GWLR) with global ordinary least square regression technique. The GWLR technique was found to provide more reliable performance in estimating spatially varying accuracy measures. The results demonstrated that the geographically weighted approach offers additional and valuable insights for examining spatial variation in the context of landscape mapping accuracy.

**Keywords** Geographically weighted method · Logistic regression · Confusion matrix · Accuracy · Remote sensing

## Introduction

In recent years, remote sensing images have been employed broadly to extract thematic information through digital image classification techniques. Assessing the accuracy of regional to global scale thematic maps derived from remote sensing is recognized as an essential requirement to support most of the mapping projects, scientific applications and policy decisions (Foody 2002). In remote sensing, the confusion matrix and its associated measures, for instance, overall accuracy (OA), producer's accuracy (PA) and user's accuracy (UA) have grown as a

conventional standard for describing the accuracy of thematic mapping (Congalton and Green 1999). The descriptions of accuracy can help to appraise the uncertainties coupled with the thematic data or to choose between thematic data sets when there is an immense accessibility of data having distinct properties (See and Fritz 2006). Therefore, accuracy is considered to be a leading feature of every remote sensing data outcome. However, some restrictions are also linked with the paradigm of a confusion matrix. Global estimates of accuracy metrics are inadequate to express the overall quality of thematic maps, as these do not reflect the spatial distribution of errors over the image. It may not be appropriate for local sub-regions because of the possibility of much larger or smaller error rates in comparison with global estimates (McGwire and Fisher 2001; Foody 2005). Several studies have reported different types of spatially distributed errors and methods to conquer this problem (Foody 2005; Steele et al. 1998; Riemann et al. 2010). The existing validation and accuracy assessment methods have been largely disregarded the advances supported by such methods by remote sensing community.

---

✉ Rajendra Prasad  
rprasad.app@itbhu.ac.in

<sup>1</sup> Centre for Climate Change and Water Research, Suresh Gyan Vihar University, Jaipur 302017, India

<sup>2</sup> Department of Physics, Indian Institute of Technology (BHU), Varanasi 221005, India

<sup>3</sup> Centre for the Study of Regional Development, Jawaharlal Nehru University, New Delhi 110067, India

Therefore, the merits of thematic maps derived using remotely sensed imagery needs to be budgeted and require improved methods or tools for approximating and illustrating the spatial distribution of errors in landscape mapping. A phenomenon differs across a landscape because of spatial non-stationarity or heterogeneity. This spatial variability restricts to employ any conventional global regression technique which assumes that the observations are independent of the spatial location. The use of conventional regression techniques, e.g. ordinary least square regression (OLSR) lead to erroneous conclusions in spatial analysis and produces spatially autocorrelated residuals (LeSage and Pace 2001). Alternatively, the challenges caused by spatial non-stationarity are addressed by several local regression methods. One of the best recognized approaches for spatial regression is the geographically weighted regression (GWR) proposed by Brunson et al. (1996). It is a statistical technique that notably covenants with the spatial non-stationarity. GWR is a local regression technique allows for the computation of relationships among variables varying over geographical extent (Fotheringham et al. 2002). It computes the local estimates of regression coefficients for a moving geographic window or kernel at every location. Unlike conventional regression techniques which encapsulate the global relationship among the variables in a single regression equation, GWR creates spatially varying data for the relationships among variables. In several studies, the better performance of GWR has been reported for various applications (Leung et al. 2000; Fotheringham et al. 2002; Zhang and Gove 2005).

The focus of our attention on GWR is motivated by numerous studies which have demonstrated its potential in the investigation of spatially varying relationships, including climatology (Brunson et al. 2001; Khosravi and Balyani 2019; Wehbe et al. 2020), health (Lin and Wen 2011; Ehlkes et al. 2014; Weber 2018; Hong et al. 2018; Hasyim et al. 2018), real estate management (Lu et al. 2014; Liu et al. 2016), urban studies (Faisal and Shaker 2017; Wang et al. 2020) and land change analysis (Maimaitijiang et al. 2015; Dadashpoor et al. 2019). GWR can also be applied in combination with linear, logistic and poisson regression techniques for various applications and reported in many studies (Nakaya et al. 2005; Ehlkes et al. 2014; Mayfield et al. 2018). Lesiv et al. (2016) explored geographically weighted logistic regression technique for comparing the fusion of different remote sensing-based land cover products. Recently Mollalo et al. (2020) have successfully applied GWR and multiscale GWR (MGWR) models for examining the spatial heterogeneity of the rate of COVID-19 occurrence at local level in United States. The geographically weighted method is useful to generate the spatial surfaces of accuracy measures (Comber et al.

2012). It computes a sequence of local diagnostics of accuracies to explore the spatial heterogeneity using weighting functions that represent the distance to the centre of a moving window or kernel (Gollini et al. 2015).

Nevertheless, the application of GWR integrated with logistic regression for generating spatially varying accuracies of a heterogeneous landscape mapping using high resolution remote sensing image is still limited. The error matrix is unable to correspond the varying accuracies of different sub-regions of the area under investigation because of the systematic errors related to a specific sensor, properties of the landscape categories and characteristics (McGwire and Fisher 2001). Even though the studies showed that the GWR results were evaluated with different approaches. The capabilities of GWR are investigated concerning other weighting function, its several bandwidths and selection criteria, etc. The present research work aims to explore geographically weighted logistic regression scheme for estimating spatial variation in accuracy measures and compare it with a conventional global OLSR technique. It may be used in different disciplines, for instance economics, environment, social and earth sciences, where improved understanding about the local behaviour of parameter relations is needed. Studies like this can be viewed as beneficial at different levels such as current mapping projects, future land use planning and management. The information can be used to develop more targeted ground survey interventions, and prioritize the strategies that are most likely to be effective for a specific part of the data, either involving a specific place or definite landscape categories. The consequences of present work may also be helpful for addressing perennial gaps in the investigation and description of spatially explicit reliability accuracy of thematic information of a heterogeneous landscape.

## Study Site and Materials

A portion of Varanasi district in Uttar Pradesh state, India is the study site for this work. It is located along the left semicircular shaped bank of divine River Ganga. This area is highly heterogeneous that provides diverse landscape categories and patterns making it a suitable site to inspect the scheme of spatial heterogeneity. The study site covers a total area of 25,327 ha and geographically extends from 82° 54' 30" to 83° 02' 30" E and from 25° 13' 08" to 25° 20' 43" N. A remote sensing image acquired from high resolution LISS IV sensor of Resourcesat-2 satellite on 6 April 2013, was used in this study. The spatial, spectral, radiometric and temporal resolutions of LISS IV image are 5.8 m, 3 bands, 10 bit, and 24 days, respectively, with a swath width of 70 km. Its spectral resolution covers three

bands including B2 (green, 520–590 nm), B3 (red, 620–680 nm) and B4 (NIR, 770–860 nm). It was classified into six major classes viz. agricultural land, fallow land, vegetation, water bodies, built up and sand according to the landscape of study site using random forest (RF) classifier (Breiman 2001). The mapping and regression analyses were implemented in an open source software R version 3.4.0 (<http://cran.rproject.org>). The precise geolocation of validation data set was collected at 551 locations by random sampling method with the help of Trimble Juno 3B GPS receiver. The study area with the sample point locations is shown in Fig. 1 as viewed on false colour composite (FCC) of LISS IV image.

## Methodology

The regression analysis is the most commonly used statistical method to examine and explore the spatial relationships among the variables. In spatial data analysis, several regression techniques have been described and formulated over the years such as linear mixed models (LMM) (Zhang and Gove 2005), generalized additive models (GAM) (Zhang and Gove 2005), geographically temporal weighted regression (GTWR) (Liu et al. 2016), geographically weighted temporally correlated logistic regression (GWTCLR) (Liu et al. 2018), geographically weighted ordinal regression (GWOR) (Dong et al. 2018) and multiscale GWR (MGWR) (Mollalo et al. 2020). OLSR is a generalized linear regression technique. It may be applied to single or multiple explanatory variables. This regression technique estimates the coefficients by using

ordinary least squares. Since the OLSR technique has been well formulated and documented in literatures, hence only the GWR techniques have been briefly described.

## Confusion Matrix

A confusion matrix was created to assess the accuracies of landscape classes in a conventional manner before starting to explore spatial non-stationarity. A confusion matrix was created using high resolution LISS IV imagery with validation data sets reviews the accuracy of RF classifier. The allocated classes against the reference information for the sample sites are cross-tabulates using the confusion matrix (Foody 2002).

## GWR Method

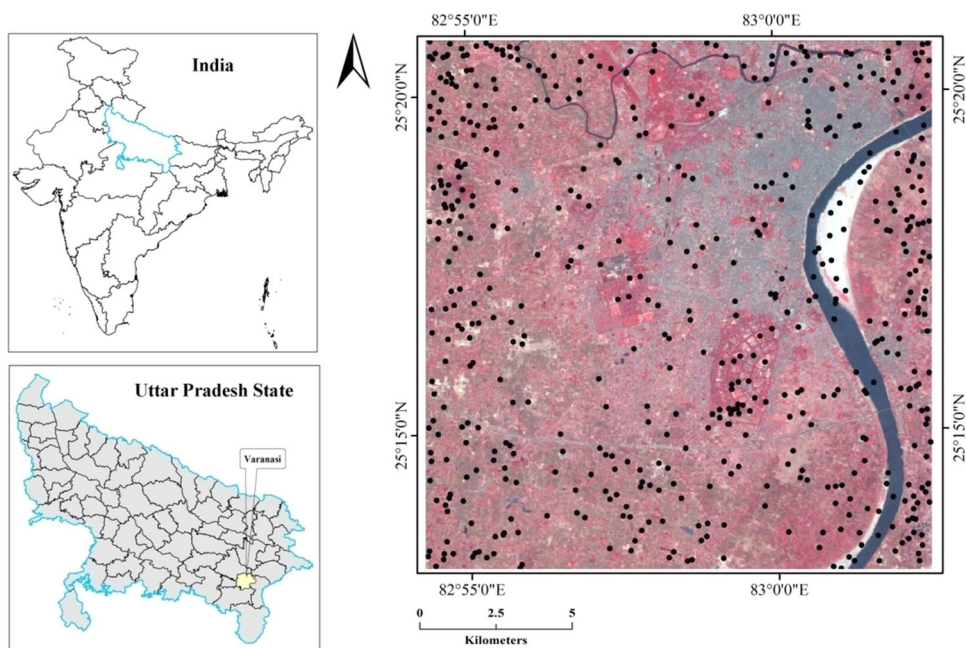
GWR is a spatial statistical technique which extends the framework of conventional regression statistics by revising a globally defined model as a locally estimated model. It enables meaningful analysis in modelling spatially heterogeneous processes (Fotheringham et al. 2002).

The coefficients in GWR compared to the global regression are functions of varying spatial position. The basic GWR model is given by (Fotheringham et al. 2002) and can be written as:

$$y_i = \beta_{i0} + \sum_{k=1}^m \beta_{ik} x_{ik} + \varepsilon_i \quad (1)$$

where  $y_i$  = dependent variable at location  $i$ ;  $\beta_{i0}$  = intercept variable at location  $i$ ;  $\beta_{ik}$  = local regression coefficient for

**Fig. 1** Location of the study site with the validation points (shaded circle)



the  $k$ th independent variable at location  $i$ ;  $x_{ik}$  = independent variable  $k$  at location  $i$ ;  $m$  = number of independent variables; and  $\varepsilon_i$  = random error at location  $i$ . The coefficients in GWR differ continuously across the study landscape. The coefficients can be computed at any place by the given dependent and one or more independent variables that are measurable at spatial location with known coordinates. For a given coordinates  $(u_i, v_i)$  at the location  $i$ , GWR equation can also be expressed as:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^m \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad (2)$$

GWR determines the implicit relationships in the region of each regression point  $i$ . The weighted least squares approach was used to estimate all the regression coefficients, for which the matrix expression is given as:

$$\hat{\beta}_i = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y \quad (3)$$

where  $X$  represents matrix of independent variables;  $\hat{\beta}_i = (\beta_{i0}, \dots, \beta_{im})^T$  is the vector of  $m + 1$  local regression factors; and  $W(u_i, v_i) = n \times n$  weighting matrix whose diagonal elements indicate geographical weights of each observation for regression point and off-diagonal elements are zero. It can be expressed as:

$$\begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{i3} \end{bmatrix} \quad (4)$$

The weighting system  $W_i$  is computed via a kernel function from the vicinities between regression point  $i$  and its surrounding  $N$  data points.

The main two weighting functions are normal or Gaussian and the Bi-square (Brunsdon et al. 1996; Fotheringham et al. 2002). These two functions differ in the method of choosing the bandwidth ( $b$ ) parameter to be used (whether fixed or adaptive). In this study, a fixed Gaussian kernel is used and can be given as:

$$\text{Gaussian} : w_{ij} = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{b} \right)^2 \right] \quad (5)$$

where  $d_{ij}$  = distance between regression point  $i$  and observation point  $j$  and  $b$  = kernel bandwidth. The fixed bandwidth should be specified by its value in unit of distance when developing a GWR model. The bandwidth at each centre  $i$  is assumed to be constant (fixed kernel) across the study area (Zhang and Gove 2005).

## Bandwidth Selection

It is required to select the bandwidth parameter either by the distance threshold or the number of nearest neighbours while implementing GWR model. An optimal bandwidth can normally be selected using goodness-of-fit measure. There are two ways to automatically select the optimal bandwidth size, namely golden-section search and interval search. In this study, the interval search method is used at a regular interval size bandwidth within a pre-specified range. The interval search applied within a range of 10,000 to 20,000 m at a regular interval of 1000 m using a fixed kernel. The bandwidth selection may be performed by the user or optimized through a process that looks for minimized cross-validation score. Here, the cross-validation (CV) approach (Bowman 1984) was used in which bandwidth is chosen by minimizing CV score. The CV score is calculated by

$$CV = \sum_{i=1}^n (y_i - \hat{y}_{i \neq i})^2 \quad (6)$$

where  $n$  = number of observations. The observation  $i$  is excluded during the calculation process so that the model is not calibrated on  $i$  only in the areas of sparse observations.

## Logistic Regression

Since the existence or absence of a certain landscape category is a binary outcome (1/0, Yes/No, True/False). Therefore, a powerful analytical technique, namely logistic regression or logit model is applied (Peng et al. 2002). The logit function is defined by

$$\log it(Y) = \ln \left( \frac{p}{1-p} \right) = \alpha + \beta X \quad (7)$$

The probability of the occurrence of the result can be calculated by taking the antilog of Eq. (7) and can be written as follows:

$$p = \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}} \quad (8)$$

where  $p$  is the probability of an outcome ( $Y$  = outcome of interest |  $X = x$ , a specific value of  $X$ ),  $\alpha$  =  $Y$  intercept,  $\beta$  = regression coefficient and  $e$  = base of the natural logarithm. Here,  $Y$  is categorical at all times, and  $X$  can be categorical or continuous. The logistic regression for  $Y$  can be extended to multiple predictors by a complex expression given as follows:

$$\log it(Y) = \ln \left( \frac{p}{1-p} \right) = \alpha + \beta_1 X_1 + \beta_2 X_2 \quad (9)$$

Therefore,

$$P = \frac{e^{\alpha + \beta_1 X_1 + \beta_2 X_2}}{1 + e^{\alpha + \beta_1 X_1 + \beta_2 X_2}} \tag{10}$$

where  $p$  is the probability,  $X_s$  = predictors;  $\beta_s$  = regression coefficients; and  $\alpha$  = Y intercept. The logistic model offers a probability varying between 0 and 1, indicates the accurate prediction of the landscape category. A reduced logistic regression model with an intercept term  $\alpha$  only and returns an approximation of its likelihood being equal to 1, can be used to estimate the OA. When the confusion matrix data (Table 1) is examined in the binary outcome, where 482 data points summed up by the diagonal constituents are shown as 1 and 62 data points summed up by the off-diagonal constituents are shown as 0, the outcome is the measure of the OA as 0.875.

### Geographically Weighted Logistic Regression (GWLR)

The logistic regression, coupled with GWR permits to vary local estimation of correctly and incorrectly classified landscape categories spatially over the space. GWLR, geographically weighted extension to the logistic regression, should be applied in case of binary response variables. It is applied to examine how the accuracy of classified outcomes varied across geographical space. GWLR is the geographically weighted extension to the logistic regression model and similar to ordinary regression. But with the help of a moving window or kernel function facility provided by geographically weighted methods, local regression models are computed at locations all over the study region. The GWR Eq. (2) can be extended to GWLR with the help of a logit function in the following way

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0(u_i, v_i) + \sum_{k=1}^m \beta_k(u_i v_i) X_{ik} + \varepsilon_i \tag{11}$$

or it may be written in the following way:

$$p_i = \frac{e^{\beta_0(u_i, v_i) + \sum_{k=1}^m \beta_k(u_i v_i) X_{ik}}}{1 + e^{\beta_0(u_i, v_i) + \sum_{k=1}^m \beta_k(u_i v_i) X_{ik}}} \tag{12}$$

where  $p_i$  is probability of prediction at location  $i$ , and other terms are same as given in Eq. (1).

Spatially varying accuracy measures were estimated by applying GWLR which compares the validation and classified data. Using Eq. (11), GWLR was applied to model the relationship between the RF classifier-based classes and the validation class data by logistic transformation (Tsumida and Comber 2015) given as:

$$\ln\left(\frac{P(O_i = 1)}{1 - P(O_i = 1)}\right) = \beta_{0(u_i, v_i)} \tag{13}$$

$$\ln\left(\frac{P(y_i = 1)}{1 - P(y_i = 1)}\right) = \beta_{0(u_i, v_i)} + \beta_1 x_{1(u_i, v_i)} \tag{14}$$

$$\ln\left(\frac{P(x_i = 1)}{1 - P(x_i = 1)}\right) = \beta_{0(u_i, v_i)} + \beta_1 y_{1(u_i, v_i)} \tag{15}$$

where  $P(O_i = 1)$ ,  $P(y_i = 1)$  and  $P(x_i = 1)$  are the probabilities of OA, UA and PA, respectively, at location  $i$ . Here,  $y_i$  and  $x_i$  are the dependent variable (reference data) and independent variable (classified data), respectively. Here,  $\beta_0$  = intercept,  $\beta_1$  = slope and  $(u_i, v_i)$  are the two-dimensional coordinates representing the location of  $i$ . GWLR coefficients are permitted to vary geographically over the two-dimensional space characterized by the coordinates  $(u, v)$  inside the region under investigation. Spatially distributed accuracy measures (OA, UA and PA) were estimated for each landscape categories (viz. agricultural land, fallow land, vegetation, water bodies, built up and sand).

**Table 1** Comparison between the reference and classified landscape categories showing an overall agreement of 87.48%

Classified	Reference						Row total	UA
	Sand	Vegetation	Water bodies	Agricultural land	Fallow land	Built up		
Sand	19	0	0	0	5	10	34	0.559
Vegetation	0	100	1	9	0	0	110	0.909
Water bodies	0	0	91	0	0	0	91	1.000
Agricultural land	0	8	0	101	0	0	109	0.927
Fallow land	0	0	0	0	78	14	92	0.848
Built up	1	1	8	5	7	93	115	0.809
Column total	20	109	100	115	90	117	551	
PA	0.950	0.917	0.910	0.878	0.867	0.795		
OA				0.875				
Kappa coefficient				0.846				

## Comparison of OLSR and GWLR

To compare the model performance between GWLR and OLSR, the statistical parameters, namely  $R^2$ , RSS and AIC were employed in this work. The  $R^2$  measures the goodness-of-fit and ranged between 0 and 1. It is likely to be higher, having more variance explained by dependent variable. The lower AIC value describes the robust capability of the regression technique in reflecting reality. The RSS is used to measure the quantity of variance in the data set that is not explained by the regression technique. It computes the amount of error remaining between the regression function and the data set. A small RSS signifies the robust fit of the model and explains a more considerable amount of the data. The better performance of regression technique is indicated by getting higher  $R^2$  value with lower RSS and AIC values.

## Results

A standard confusion matrix was constructed to calculate UA, PA and OA using validation data set (Table 1). The computation of commission error (inclusion) confers the UA specifying the probability of a classified pixel into a certain class is truly correspond to that class on the field. The computation of omission error (exclusion) delivers the PA that specifies how well the training pixels of an assumed class are classified. The total amount of pixels classified correctly divided by the total amount of the pixels is used to calculate OA of classification results (Lillesand and Kiefer 1999). The UA ranged from 0.559 (sand class) to 1.000 (water bodies class), while the PA ranged from 0.795 (built-up class) to 0.950 (sand class). At the same time, it is perceptible that few landscape categories are classified more presumably than the others. It may be due to imbalance in the selection of training data for RF classifier. The classification results are very much sensitive to the proportion of used training samples. The RF classification-based classified outputs can be inconsistent due to dependency on the input variables and selection scheme of training data sets during the classification process (Millard and Richardson 2015). It is also important to design the training and validation data sets carefully to elude the biasness in classification and inflated evaluation of accuracy measures (Millard and Richardson 2015).

Also, the table does not provide the information on spatially distributed errors associated with different landscape classes. The GWLR was applied to examine the spatially varying relationship between the classified data from the image and the reference data collected through the ground survey. Table 2 illustrates the spatial distribution of

accuracy measures (UA, PA and OA) of landscape categories in terms of minimum, maximum, and median, 1st and 3rd quartiles along with the inter-quartile range (IQR). The IQR is described as a representation of the overall spatial variation in the accuracy measures. The probabilities of the accurate class estimates are shown in terms of the minimum, maximum, median and 1st and 3rd quartiles (Table 2). This information, along with the global and IQR exhibits the probabilities of correctly predicted reference data by the RF-based image classification and their variation.

The larger IQR values signify the higher variation in the spatial distribution of accuracies (Comber et al. 2012). It also exhibits the degree to which the observed ground truth data are inferred by the predicted data based on remote sensing image and the variation in GWLR method. The IQR values of UA, PA and OA for various landscape categories are listed in Table 2.

For agricultural land, OA was varied from 0.004 to 1.000 across the study region in comparison with the conventional value of 0.875. While UA was found to vary from 0.015 to 1.000 as compared to a global number of 0.927, and PA was found to vary between 0.533 and 0.998 as compared to a conventional measure of 0.878 across the study area. For vegetation, OA was found to vary from 0.343 to 1.000 in sub-regions of the study area in comparison with a global figure of 0.875. While UA was found to vary from 0.627 to 1.000 in comparison with a global figure of 0.909, and PA was found to vary from 0.377 to 1.000 in contrast to a conventional measure of 0.917 across the study area. For fallow land, OA was found to vary from 0.017 to 1.000 in diverse regions of study site in comparison with a traditional value of 0.875. While UA was found to vary from 0.041 to 1.000 compared to a traditional measure of 0.848, and PA was found to vary from 0.381 to 1.000 in contrast to a traditional measure of 0.867 across the study area. For built up, OA was found to vary from 0.069 to 0.998 over the study area in contrast to a traditional value of 0.875. While UA was found to vary from 0.271 to 1.000 compared to a conventional measure of 0.809, and PA was found to vary from 0.095 to 1.000 in contrast to a traditional figure of 0.795 across the study area. For water bodies, OA was found to vary from 0.530 to 1.000 in comparison with a global figure of 0.875 across the study area. While UA was found to vary from 1.000 to 1.000 in comparison with a conventional value of 1.000 and PA was found to vary from 0.530 to 1.000 in contrast to a conventional value of 0.910 across the study area. For sand, OA was found to vary from 0.000 to 1.000 in comparison with a conventional value of 0.875 across the study area. While UA was found to vary from 0.000 to 0.932, and PA was found to vary from 0.494 to 0.010 in comparison with global figures of 0.559 and 0.950, respectively, across

**Table 2** A summary of the variation in GWLR representing UA, PA and OA for various landscape categories

Categories		Min.	1st Quartile	Median	Mean	3rd Quartile	Max.	IQR
Agricultural land	UA	0.015	0.800	0.999	0.851	1.000	1.000	0.200
	PA	0.533	0.793	0.896	0.875	0.972	0.998	0.179
	OA	0.004	0.644	0.807	0.751	0.976	1.000	0.333
Vegetation	UA	0.627	0.872	0.955	0.921	0.995	1.000	0.124
	PA	0.377	0.882	0.995	0.922	1.000	1.000	0.119
	OA	0.343	0.780	0.911	0.859	0.979	1.000	0.199
Fallow land	UA	0.041	0.802	0.997	0.874	1.000	1.000	0.198
	PA	0.381	0.779	0.879	0.857	0.967	1.000	0.188
	OA	0.017	0.636	0.874	0.782	0.995	1.000	0.359
Built up	UA	0.271	0.741	0.880	0.845	0.975	1.000	0.234
	PA	0.095	0.724	0.861	0.823	0.980	1.000	0.256
	OA	0.069	0.580	0.717	0.724	0.918	0.998	0.338
Water bodies	UA	1.000	1.000	1.000	1.000	1.000	1.000	0.000
	PA	0.530	0.866	0.973	0.923	0.999	1.000	0.133
	OA	0.530	0.866	0.977	0.923	0.999	1.000	0.133
Sand	UA	0.000	0.002	0.374	0.381	0.737	0.932	0.735
	PA	0.494	0.990	1.000	0.944	1.000	1.000	0.010
	OA	0.000	0.000	0.008	0.308	0.721	1.000	0.721

IQR indicates the probability variation among pixels. 1st and 3rd quartiles indicate the probabilities with 25th and 75th percentile

the study area. In the context of UA, the sand class exhibited larger, while the water bodies exhibited smaller spatial variation. In the context of PA, the built-up class exhibited larger, while the sand class exhibited smaller spatial variation. It is remarkable to note that the distribution of UA estimates reveals much larger variation followed by the moderately distributed OA estimates, whereas smaller dissimilarity was observed in the distribution of PA estimates.

The visual representation of spatial variations in UA, PA and OA is illustrated in Figs. 2, 3 and 4, respectively. The legend values in figures signify the probability of the presence of the correct category. Yellow circles represent locations correctly classified. Red circles represent locations incorrectly classified as given category (commission error). Black circles represent locations incorrectly classified to other categories (omission error). For major landscape category agricultural land, the UA and PA are higher in the southern and lower towards the northern parts of the study area (Figs. 2a and 3a). Figure 3a shows the higher variation in the estimates of PA of agricultural land, but the marginally higher trend of variation towards the northern part of the study area. Figure 4a shows that the OA for agricultural land is found to be higher in the southern and lower towards the northern regions of the study area. Figure 2d shows that the UA for other major landscape category built up is higher in the middle of study area while, lower in the south-eastern parts. Figure 3d shows that the

PA for built up is higher in the middle and lower in the western region of the study site. Figure 4d shows that the OA for built up category is found to be higher in the middle and lower in the south-east and western parts of the area under investigation. Other landscape categories are also showing spatial variation in the accuracy measures (Figs. 2, 3, 4). The UA is the probability estimation of the correctly classified pixels, representing the categories on the site. It shows the commission error (inclusion), signifying the probability of locating that category correctly on the ground for the map user. Here, Fig. 2a would imply a smaller amount of confidence in mapping areas as agricultural land truly being that category on the ground. The smaller amount of confidence in mapping land surface as built up actually being that class on the ground would also be implied by Fig. 2a. The PA is the probability estimation of a reference pixel is being appropriately identified in the classified output. It shows the omission error (exclusion), signifying the probability that the classes of attention are omitted from the output product for the map producer. Therefore, it is convinced that the majority of existing agricultural land was mapped with low omission error. However, the commission error for agricultural land category was found to be high in the northern part of the study site as well. In case of built-up category, it is convinced that majority of the area covered with the class of interest was mapped with low omission error. However, the

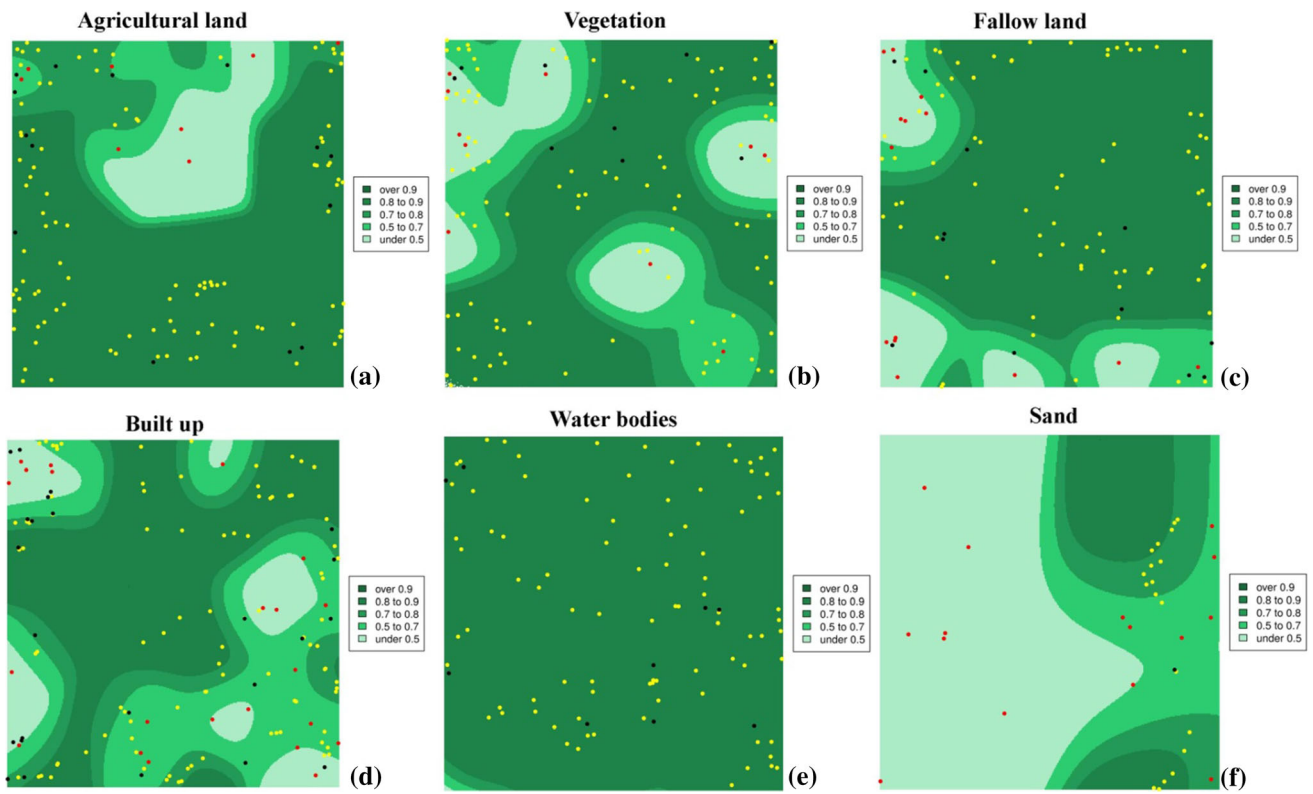


Fig. 2 GWLR-based maps of spatial variation in user's accuracies of landscape categories

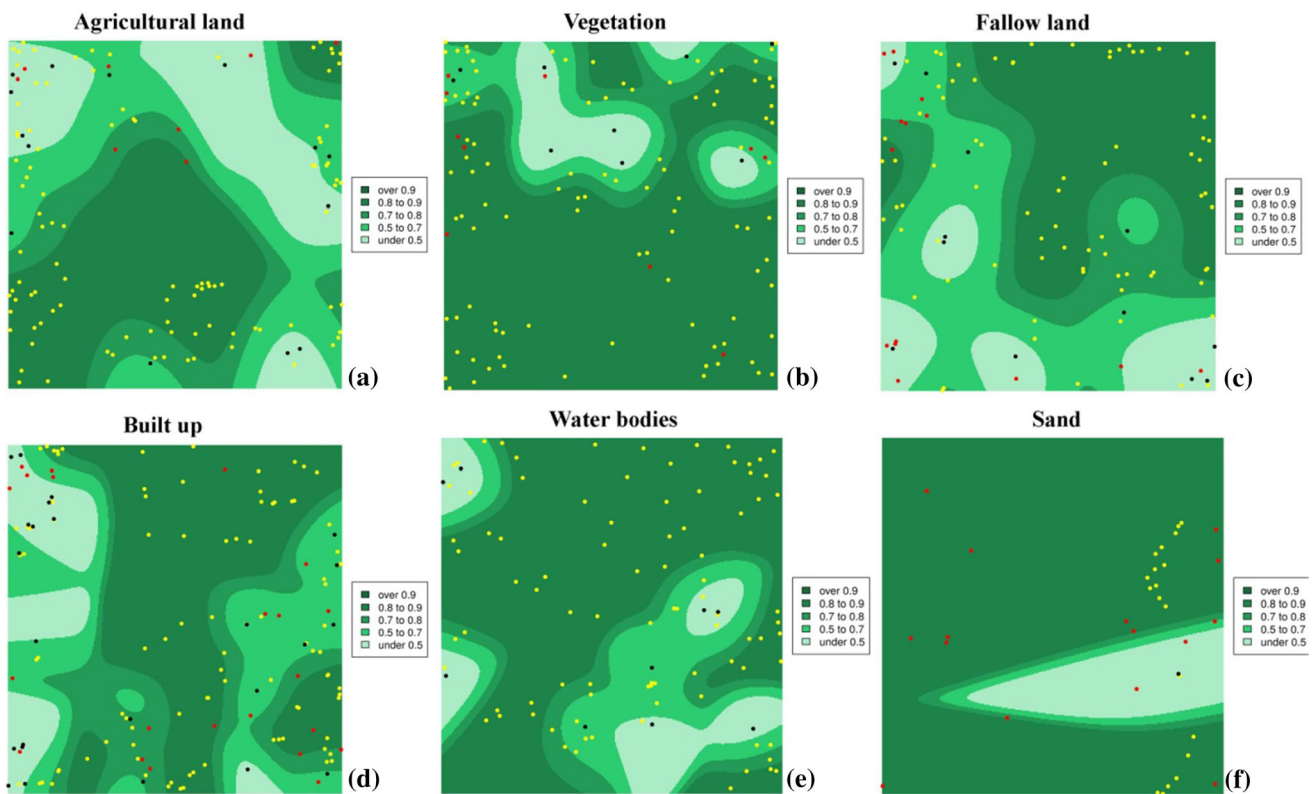


Fig. 3 GWLR-based maps of spatial variation in producer's accuracies of landscape categories



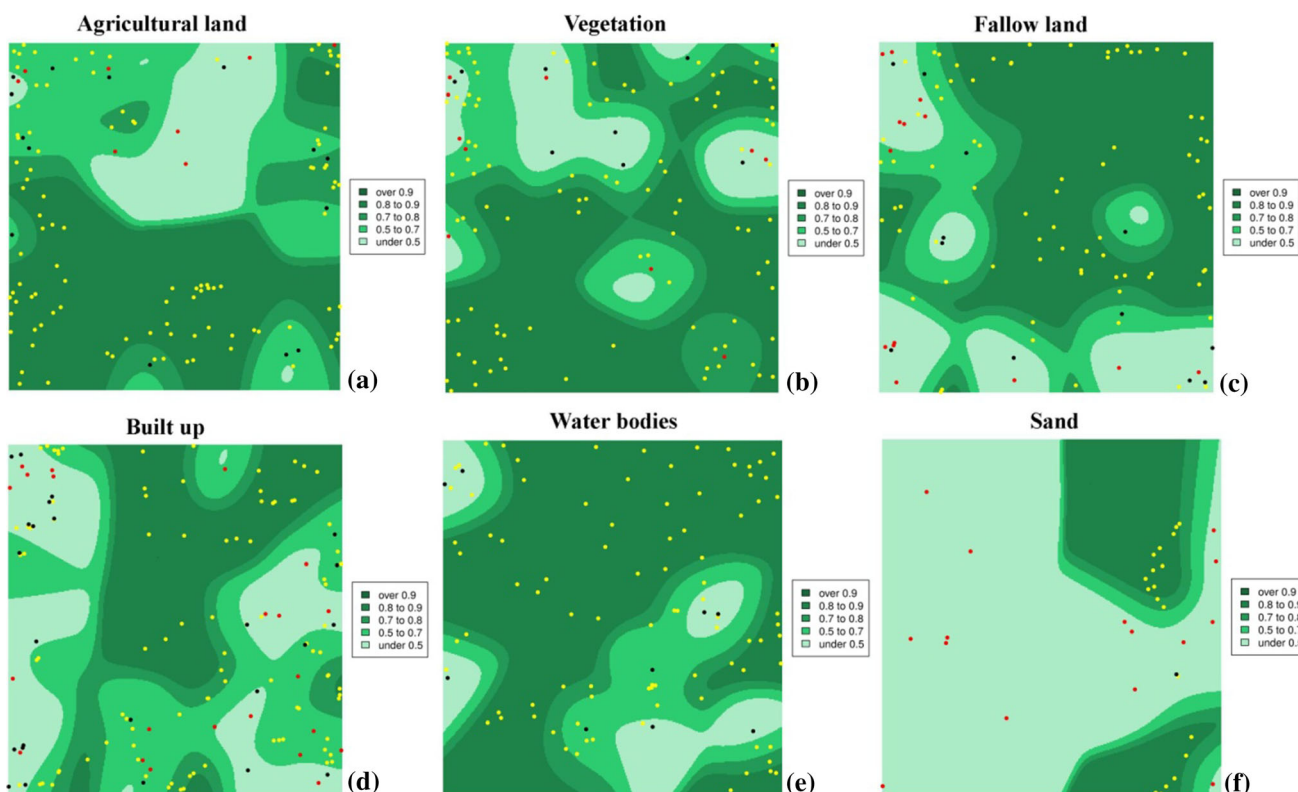


Fig. 4 GWLR-based maps of spatial variation in overall accuracies of landscape categories

commission error for built up category was found to be high in the south-east part of the study reason as well.

The performance of estimating spatial variation in overall accuracies was compared using OLSR and GWLR for different landscape categories. Table 3 reveals that GWLR produced lower AIC and RSS values in comparison with OLSR, indicating a better fit to the observed statistics. The higher  $R^2$  values for GWLR in comparison with OLSR showed that the dependent variables are described by more variance. The detailed information is shown in Table 3. Overall, the performance of GWLR technique was found to be better in comparison with OLSR for estimation of spatially varying accuracy of landscape mapping.

### Discussion

The confusion matrix, a convenient way to summarize errors in landscape mapping, is aspatial in nature. The landscape mapping with comparatively high accuracy may not be necessarily applicable in analyzing spatially varying accuracy at the local level. The spatial representation, however, could be beneficial in landscape mapping with low accuracy or intricate categories having inferior accuracies. Geographical analyses can be used to understand how and why the processes reflect spatially varying nature. The locational characteristics of data are used explicitly to recognize relationships varying at the local scale. It is not similar as analyzing spatial data in itself under the hypothesis that it represents a spatial analysis as the data

Table 3 Comparison of GWLR and OLSR techniques for estimating spatial variation in overall accuracies of various landscape categories

Landscape categories	Model Parameters	GWLR			OLSR		
		Adjusted $R^2$	RSS	AIC	Adjusted $R^2$	RSS	AIC
Agricultural land		0.241	11.81	64.81	0.077	18.07	117.12
Vegetation		0.179	10.48	52.55	0.097	15.97	102.68
Fallow land		0.450	7.64	27.62	0.011	19.50	125.05
Built up		0.197	19.52	125.57	0.032	30.78	188.89
Water bodies		0.155	6.08	7.71	0.049	8.19	37.56
Sand		0.647	1.65	-3.68	0.075	8.69	54.55

are spatial in nature. It is confusing to implement spatially evident approach because the error cluster of land cover features is well identified for a variety of well-reported causes. Therefore, it is needed to evaluate the local error frequently by familiar visual and qualitative techniques (Friedl et al. 2002). This work demonstrates the utility of logistic regression in producing the probabilities of accuracy measures and the ability of its geographically weighted extension to generate spatially distributed variations in probabilities as well. The remote sensing society is perhaps well known with the concepts of OA, PA and UA and the method proposed here may better reflect their needs.

This study does not attempt to surpass the entire limitations associated with the confusion matrix. This study somewhat investigated spatially distinct approach for describing accuracies using geographically weighted approach to identify a relationship vary between classified and reference data over geographical space. The capability of estimating accuracy measures and errors from the data gathered during a standard validation practice, recommends that the maps of distributed accuracies could assist confusion matrices. The geographically weighted methods can be used to produce spatially explicit outputs which point out the potential precedence of incorporating any validation results in conjunction with remote sensing imagery based land cover outcomes. Finally, the method proposed in this study addresses one of the elementary beliefs of geographical analyses that the process under examination varies over space continuously.

## Conclusions

The geographically weighted technique was utilized in describing the spatial variations in the accuracies by random forest classification using high resolution remote sensing data. It addresses major concerns in the analysis and description of accuracy and errors associated with heterogeneous landscape mapping. The information of non-stationarity associated with the frequently varying landscape errors can be observed by analysing its spatial distribution. This study also compared the performance of GWLR with conventional OLSR technique. The investigation showed better performance of GWLR in estimating and analysing spatially varying accuracy measures compared to OLSR. Finally, the denouement of this study is that the spatially explicit accuracy measures are more informative and precise. They offer better support for appraisal of data accuracy compared to the global measures based on the confusion matrix. The results of this work also suggest that there is a need to reconsider the tenets of

accuracy and errors prevalent in the spatial science community.

**Acknowledgement** The authors are grateful to anonymous reviewers for their insightful and constructive comments which assisted in improving the manuscript. The authors are also sincerely thankful to the Editor.

## References

- Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71(2), 353–360.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Brunsdon, C. F., Fotheringham, A. S., & Charlton, M. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298.
- Brunsdon, C., McClatchey, J., & Unwin, D. (2001). Spatial variations in the average rainfall-altitude relationships in Great Britain: An approach using geographically weighted regression. *The International Journal of Climatology*, 21(4), 455–466.
- Comber, A., Fisher, P., Brunsdon, C., & Khmag, A. (2012). Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, 127, 237–246.
- Congalton, R. G., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: Principles and practices*. Boca Raton: Lewis Publishers.
- Dadashpoor, H., Azizi, P., & Moghadasi, M. (2019). Land use change, urbanization, and change in landscape pattern in a metropolitan area. *Science of the Total Environment*, 655, 707–719.
- Dong, G., Nakaya, T., & Brunsdon, C. (2018). Geographically weighted regression models for ordinal categorical response variables: An application to geo-referenced life satisfaction data. *Computers, Environment and Urban Systems*, 70, 35–42.
- Ehlkes, L., Krefis, A. C., Kreuels, B., Krumkamp, R., Adjei, O., Ayim-Akonor, M., et al. (2014). Geographically weighted regression of land cover determinants of Plasmodium falciparum transmission in the Ashanti Region of Ghana. *International Journal of Health Geographics*, 13(1), 35.
- Faisal, K., & Shaker, A. (2017). Improving the accuracy of urban environmental quality assessment using geographically-weighted regression techniques. *Sensors*, 17, 528.
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, 185–201.
- Foody, G. M. (2005). Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing*, 26, 1217–1228.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester: Wiley.
- Friedl, A. F., Mciver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83, 287–302.
- Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, 63, 85–101.
- Hasyim, H., Nursafingi, A., Haque, U., Montag, D., David, A., Groneberg, D. A., et al. (2018). Spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia. *Malaria Journal*, 17, 87.

- Hong, Z., Hao, H., Li, C., Du, W., Wei, L., & Wang, H. (2018). Exploration of potential risks of hand, foot, and mouth disease in Inner Mongolia Autonomous region, China using geographically weighted regression model. *Scientific Reports*, 8, 17707.
- Khosravi, Y., & Balyani, S. (2019). Spatial modeling of mean annual temperature in Iran: comparing cokriging and geographically weighted regression. *Environmental Modeling and Assessment*, 24, 341–354.
- LeSage, J., Pace, K. (2001). Spatial dependence in data mining. In: *Data mining for scientific and engineering applications*. pp. 439–460. New York: Springer.
- Lesiv, M., Moltchanova, E., Schepaschenko, D., See, L., Shvidenko, A., Comber, A., et al. (2016). Comparison of data fusion methods using crowdsourced data in creating a hybrid forest cover map. *Remote Sensing*, 8, 261.
- Leung, Y., Mei, C. L., & Zhang, W. X. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, 32, 9–32.
- Lillesand, T. M., & Kiefer, R. W. (1999). *Remote sensing and image interpretation*. New York: Wiley.
- Lin, C. H., & Wen, T. H. (2011). Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue. *International Journal of Environmental Research and Public Health*, 8, 2798–2815.
- Liu, J., Yang, Y., Xu, S., Zhao, Y., Wang, Y., & Zhang, F. (2016). A geographically temporal weighted regression approach with travel distance for house price estimation. *Entropy*, 18, 303.
- Liu, Y., Lam, K. F., Wu, J. T., & Lam, T. T.-Y. (2018). Geographically weighted temporally correlated logistic regression model. *Scientific Reports*, 8, 1417.
- Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science*, 28(4), 660–681.
- Maimaitijiang, M., Ghulam, A., Sandoval, J. S. O., & Maimaitiyiming, M. (2015). Drivers of land cover and land use changes in St. Louis metropolitan area over the past 40 years characterized by remote sensing and census population data. *International Journal of Applied Earth Observation and Geoinformation*, 35, 161–174.
- Mayfield, H. J., Lowry, J. H., Watson, C. H., Kama, M., Nilles, E. J., & Lau, C. L. (2018). Use of geographically weighted logistic regression to quantify spatial variation in the environmental and sociodemographic drivers of leptospirosis in Fiji: a modelling study. *Lancet Planet Health*, 2(5), e223–e232.
- McGwire, K. C., Fisher, P. (2001). Spatially variable thematic accuracy: Beyond the confusion matrix. C. T. Hunsaker, M. F. Goodchild, M. A. Friedl, & T. J. Case (Eds.), In: *Spatial uncertainty in ecology: implications for remote sensing and GIS applications* (pp. 308–329). New York: Springer
- Millard, K., & Richardson, M. (2015). On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sensing*, 7, 8489–8515.
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, 728, 138884.
- Nakaya, T., Fotheringham, A. S., Brunson, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, 24, 2695–2717.
- Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *Journal of Educational Research*, 96(1), 3–14.
- Riemann, R., Wilson, B. T., Lister, A., & Parks, S. (2010). An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest inventory and analysis (FIA) data. *Remote Sensing of Environment*, 114, 2337–2352.
- See, L. M., & Fritz, S. (2006). A method to compare and improve land cover datasets: Application to the GLC-2000 and MODIS land cover products. *IEEE Transaction of Geoscience and Remote Sensing*, 44(7), 1740–1746.
- Steele, B. M., Winne, J. C., & Redmond, R. L. (1998). Estimation and mapping of misclassification probabilities for thematic land cover maps. *Remote Sensing of Environment*, 66(2), 192–202.
- Tsutsunida, N., & Comber, A. J. (2015). Measures of spatio-temporal accuracy for time series land cover data. *International Journal of Applied Earth Observation and Geoinformation*, 41, 46–55.
- Wang, Z., Fan, C., Zhao, Q., & Myint, S. W. (2020). A geographically weighted regression approach to understanding urbanization impacts on Urban warming and cooling: a case study of Las Vegas. *Remote Sensing*, 12, 222.
- Weber, A. V. (2018). Exploring local influences on Zika Virus rates in Puerto Rico utilizing geographically weighted regression. *Papers in Applied Geography*, 4(1), 100–112.
- Wehbe, Y., Temimi, M., & Adler, R. F. (2020). Enhancing precipitation estimates through the fusion of weather radar, satellite retrievals, and surface parameters. *Remote Sensing*, 12, 1342.
- Zhang, L., & Gove, J. H. (2005). Spatial assessment of model errors from four regression techniques. *Forest Science*, 51(4), 334–346.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.