Responses and proposed changes to referees for gmd-2017-278

Daojun Zhang, Na Ren, and Xianhui Hou

Dear Editor,

We appreciate both you and the two anonymous reviewers giving our work (ID: gmd-2017-278) positive comments and giving us the chance to make a further modification of our manuscript. We have carefully modified the manuscript according to the suggestions and comments provided by the reviewers and hope our modification could meet with the requirement of GMD. Following are the responds to the reviewers' suggestions and comments one by one (all suggestions and comments are colored in red, and our proposed changes to the manuscript are colored blue). At the end of this file we attached the comparison between the newest edition and the original edition.

Response to Anonymous Referee #1:

1. Line 52-53, these references are so old, please cite more recent references.

Thank you for your comments and suggestions. Here we mainly listed the method research literatures. Unlike application researches, the method researches especially original models (not including modified models) are generally older. Anyway, we have added more recent models here as references, please see lines 57-64 in the comparison edition attached. The new statement is as following.

"(1) Locations are introduced as direct or indirect independent variables. This type of model is still a global model, but space coordinates or distance weights are employed to adjust the regression estimation between the dependent variable and independent variables (Agterberg, 1964; Agterberg and Cabilio, 1969; Agterberg, 1970; Agterberg and Kelly, 1971; Agterberg, 1971; Casetti, 1972; Lesage & Pace, 2009, 2011)."

2. Line 57-61, it is better to show two recent examples.

Thank you for your comments and suggestions. We have added more references here, which are about the new applications of models including locations as direct or indirect independent variables, please see lines 64-71 in the comparison edition attached. The new statement is as following.

"For example, Reddy et al. (1991) performed logistic regression by including trend variables for mappingto map the base-metal potential in the Snow Lake area, Manitoba, Canada. In addition, Casetti (1972) developed a ; Helbich & Griffith (2016) compared the spatial expansion method (SEM) to other methods in modeling the house price variation locally, where the regression parameters are themselves functions of the x and y coordinates as well asand their combinations; Yu & Liu (2016) used the spatial lag model (SLM) and spatial error model to control spatial effects in modeling the relationship between $PM_{2.5}$ concentrations and per capita GDP in China."

3. Line 63-67, there are various applications of GWR in Geosciences, they should be cited here.

Thank you for your suggestion and we have added some new literatures bout the application of GWR in different fields here, please see lines 74-78 in the comparison edition attached. The new statementis as following.

"GWR wasmodels were first developed at the end of the 20th century by Brunsdon et al. (1996) and Fotheringham et al. (1996, 1997, 2002) for modeling spatially heterogeneous processes, and it hashave been used widely in the field of geography.geosciences (e.g., Buyantuyev & Wu, 2010; Barbet-Massin et al., 2012; Ma et al., 2014; Brauer et al., 2015)."

Response to Anonymous Referee #2:

The manuscript presents something that is technically sound. So it can be accepted for publication after addressing the following comments:

1. *The English needs to be improved. It has not been structured well. The statements and propositions have not been organized properly. Reflecting the state of the art is poor as well. The Introduction has not properly been tightened, so the problem and the purpose are not clear.*

Thank you for your suggestions. We have made a major revision to the manuscript. As you can see in the modified manuscript attached, added or subtracted some statements from the original manuscript to clarify the intentions of this work more clearly. We also included the evidential layers in the modified manuscript (please also see Figure R 1). With respect to instruction, we have re-sorted the previous researches in overcoming the non-stationary of spatial variables (especially lines 111-134 in the comparison edition attached), removed the redundant expressions to avoid repetition with later model description parts, and set more natural paragraphs to enhance the level of expression. Some expressions in the summary section have also been modified.

Besides, the English was re-checked thoroughly.

2. *In Fig. 8, two different data sets were bound together and can explicitly be separated by a horizontal line. I think there is something wrong. Perhaps it would be better that the two data sets (A and B) be gridded by the same cell size and the spatial values should not be modeled/mapped individually. You should generate a model similar to the Fig. 5.*

Thank you for your suggestion. We have added that all the raster files in this research are created with the cell size of 1 km x 1 km (lines 481-482 in the comparison edition attached). In fact, it is missing data that caused the sharp differences between the north and south parts (i.e. A and B in Fig. 5) of Fig. 8 (new Fig. 9) rather than data set source, since we have made up a circumstance that there are no geochemical data in region B (lines 485-488 in the comparison edition attached). These expressions are cited following.

"The four independent variables described previously were also used for ILRBSWT modeling in this study (see Figs. 4 (a) to (d)), and they were uniformed in the ArcGIS grid format with a cell size of 1 km × 1 km. To demonstrate the advantages of the new method for missing data processing, we designed an artificial situation in Fig. 5, i.e., grids in region A have values for all four independent variables, while they only have values for two independent variables and no values in the two geochemical variables in region B. "

We acknowledge that the texture looks finer in Fig. 5 (new Fig. 6), and that is because this spatial variable is a continuous variable. However, as a posterior probability layer, Fig. 8 (new Fig. 9) was obtained after the discretizing and integrating the evidence layers, including the buffer layer and the geochemical anomaly layer, which can easily lead to the spatial discontinuity of the grid value. As a result, the texture looks rough, which is not caused by grid size differences.

3. *Weighted evidence layers must be added to the manuscript.*

Thank you for your suggestion and we have accepted it, please see Fig R 4 (Fig. 4 in the attached comparison), which includes all original evidential layers used in this research. Besides, as a sliding window model, ILRBSWT builds predictive model at each local window, and the discretization of original evidential layers and the determination of weights for each class are also based on the local window, thus it is impossible to show the final weights used for modeling.

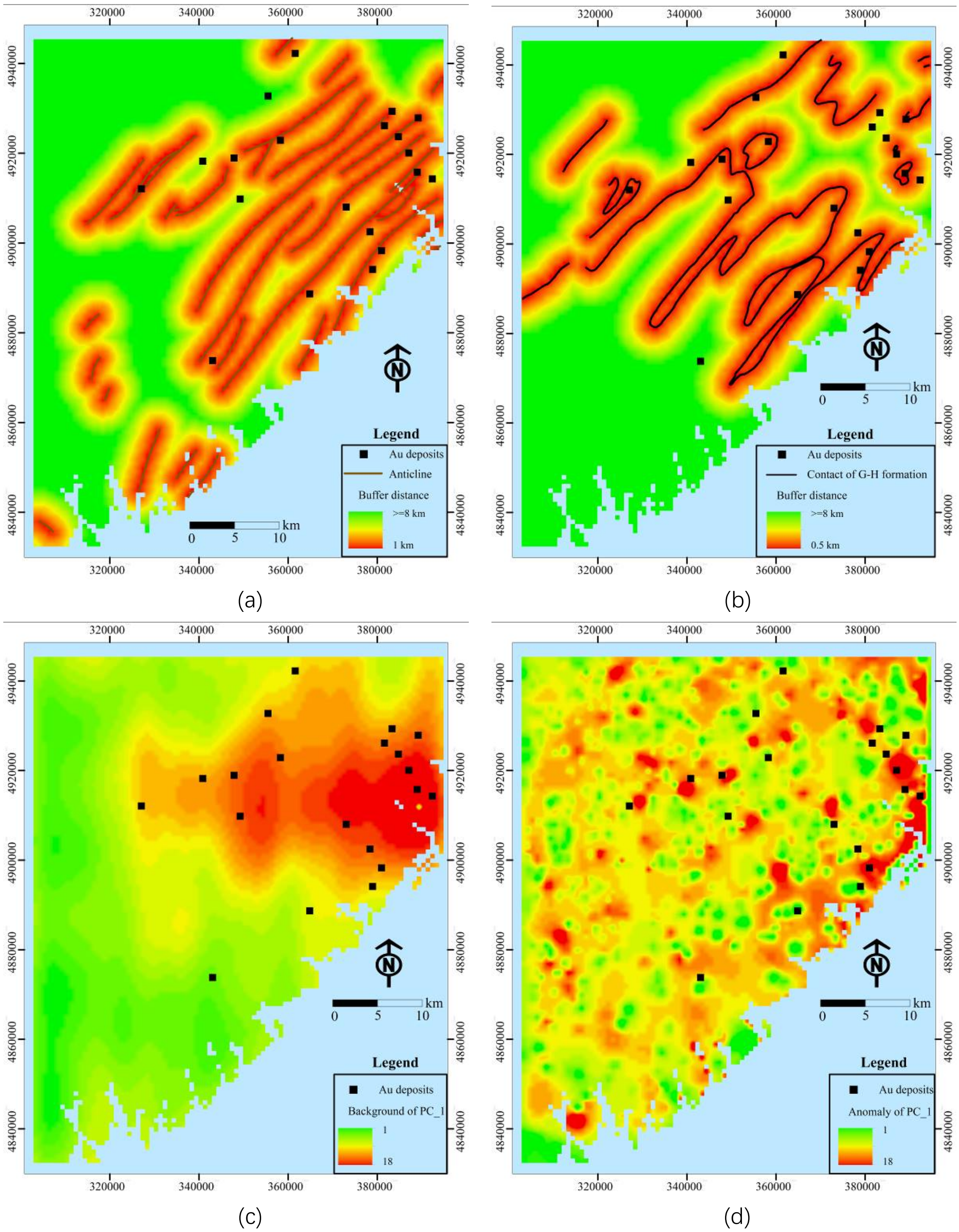


Fig. R 1: Evidential layers used to map Au deposits in this study: buffer of anticline axes (a), buffer for the contact of Goldenville–Halifax Formation (b), and background (c) and anomaly (d) separated with the S-A filtering method based on the ore element loadings of the first component.

4. *The manuscript presents lack of a Discussion section.*

Thank you for your suggestion and we have accepted it. We have added an individual Discussion Section in the new manuscript to discuss the findings and deficiencies of the study (lines 539-602 in the comparison edition attached). Besides, we have added more analyses and discussions in section 5.5 about the comparison of the results of different models (lines 604-638 in the comparison edition attached); please also see details as cited following:

“6 Discussion

Because of potential spatial heterogeneity, the model parameter estimates obtained based on the total equal-weight samples in the classical regression model may be biased, and they may not be applicable for predicting each local region. Therefore, it is necessary to adopt a local window model to overcome this issue. The presented case study shows that ILRBSWT can obtain better prediction results than classical logistic regression because of the former’s sliding local window model, and their corresponding intersection point values are 2.85 and 2.45, respectively. However, ILRBSWT has even advantages. For example, characterizing 26% or 29% of the total study area as promising prospecting targets is too high in terms of economic considerations. If instead 10% of the total area needs is mapped as the target area, the proportions of correctly predicted known deposits obtained by ILRBSWT and logistic regression are 44% and 24%, respectively. The prediction efficiency of the former is 1.8 times larger than the latter.

In this study, we did not separately consider the influences of spatial heterogeneity, missing data, and degree of exploration weight all remain, so we cannot evaluate the impact of each factor. Instead, the main goal of this work was to provide the ILRBSWT tool, demonstrating its practicality and overall effect. Zhang et al. (2017) applied this model to mapping intermediate and felsic igneous rocks and proved the effectiveness of the ILRBSWT tool in overcoming the influence of spatial heterogeneity specifically. In addition, Agterberg and Bonham-Carter (1999) showed that WofE has the advantage of managing missing data, and we have taken a similar strategy in ILRBSWT. We did not fully demonstrate the necessity of using exploration weight in this work, which will be a direction for future research. However, it will have little influence on the description and application of ILRBSWT tool as it is not an obligatory factor, and users can individually decide if the exploration weight should be used.

Similar to WofE and logistic regression, ILRBSWT is a data-driven method, thus it inevitably suffers the same problems as data-driven methods, e.g., the information loss caused by data discretization, and exploration bias caused by the training sample location. However, it should be noted that evidential layers are discretized in each local window instead of the total study area, which may cause less information loss. This can also be regarded as an advantage of the ILRBSWT tool. With respect to logistic regression and WofE, some researchers have proposed solutions to avoid information loss resulting from spatial data discretization by performing continuous weighting (Pu et al., 2008; Yousefi & Carranza, 2015b, 2015c, 2016), and these concepts can be incorporated into further improvements of the ILRBSWT tool in the future."

5. *The methods applied, i.e. "weights of evidence" and "logistic regression" are data-driven MPM methods, which carry exploration bias and uncertainty resulting from using classified spatial data and location of known deposits as training sites. Please add a discussion on the disadvantages of such data-driven MPM methods. There are continuous weighting approaches using logistic functions (e.g., logistic-based weighting methods, geometric average function, continuous fuzzification method, and …) to avoid the aforementioned uncertainty.*

Thank you for your suggestions and we have accepted them. We have included in the Discussion Section a description about the shortcomings of the data-driven MPM method, and reviewed previous efforts in overcoming the issues caused by data discretization; please see details in the third paragraph in the discussion section.

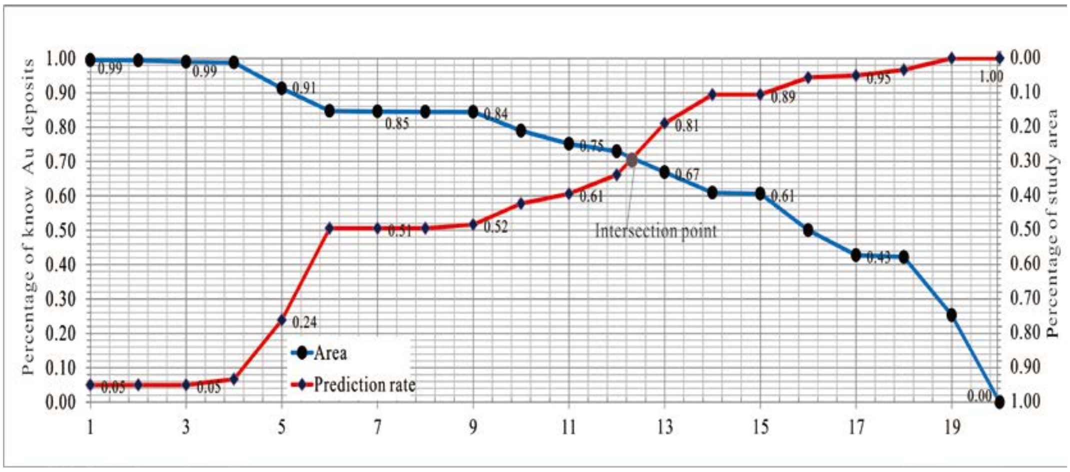
6. *The evaluation method applied could not reflect the efficiency of the two models ade-quately. So you can see that there is no much difference between the models. I think it would be better if you could apply a prediction-area (P-A) plot and calculate normalized density for the two models to compare them.*

Thank you for your suggestion, and we have accepted it. We applied the prediction-area (P-A) plot and normalized density in the new manuscript to replace the previous used *t*-value method for model comparison in *"5.5 Comparison of the mapping results"* (lines 538-558 in the comparison edition attached), as is cited following.

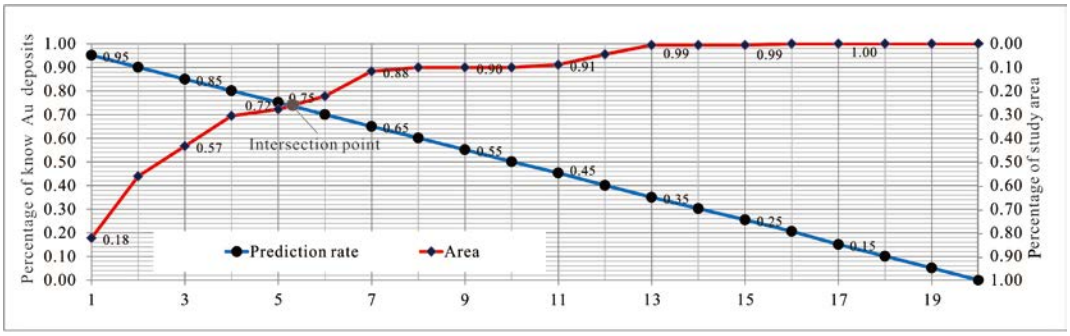
"To evaluate the predictive capacity of the newly developed and traditional methods, the posterior probability maps obtained through logistic regression and ILRBSWT shown in Fig. 9 (a) and 9 (b) were divided into 20 classes using the quantile method. Prediction-area (P-A) plots (Mihalasky & Bonham-Carter, 2001; Yousefi et al., 2012; Yousefi & Carranza, 2015a) were then made according to the spatial overlay relationships between Au deposits and the two classified posterior probability maps in Fig. 10 (a) and (b) respectively. In a P-A plot, the horizontal ordinate indicates the discretized classes of a map representing the occurrence of deposits. The vertical scales on the left and right sides indicate the percentage of correctly predicted deposits from the total known mineral occurrences and the corresponding percentage of the delineated target area from the total study area (Yousefi & Carranza, 2015a). As shown in Figs. 10 (a) and (b), with the decline of the posterior probability threshold for the mineral occurrence from left to right on the horizontal axis, more known deposits are correctly predicted, and meantime more areas are delimited as the target area; however, the growth in the prediction rates for deposits and corresponding occupied area are similar before the intersection point in Fig. 10 (a), while the former shows higher growth rate than the latter in Fig. 10 (b). This difference suggests that ILRBSWT can predict more known Au deposits than logistic regression for delineating targets with the same area, and indicates that the former has a higher prediction efficiency than the latter.

It would be a little inconvenient to consider the ratios of both predicted known deposits and occupied area. Mihalasky and Bonham-Carter (2001) proposed a normalized density, i.e. the ratio of the predicted rate of known deposits to its corresponding occupied area. The intersection point in a P-A plot is the crossing of two curves. The first is fitted from scatter plots of the class number of the posterior probability map and rate of predicted deposit occurrences (the “Prediction rate” curves in Fig. 10). The second is fitted according to the class number of the posterior probability map and corresponding accumulated area rate (the “Area” curves in Fig. 10). At the interaction point, the sum of the prediction rate and corresponding occupied area rate is 1; the normalized density at this point is more commonly used to evaluate the performance of a certain spatial variable in indicating the occurrence of ore deposits (Yousefi & Carranza, 2015a). The intersection point parameters for both models are given in Table 1. As shown in the table, 71% of the known deposits are correctly predicted with 29% of the total study area delineated as target area when the logistic regression is applied; if ILRBSWT if applied, 74% of the known deposits can be correctly predicted with only 26% of the total area delineated as the target area. The normalized densities for the posterior probability maps obtained from the logistic regression and ILRBSWT are 2.45 and 2.85 respectively; the latter performed significantly better than the former.”

The evaluation results supported the conclusions of this research, Please see Fig. R 2 (Fig. 10 in the comparison edition attached).



(a)



(b)

Fig. R 2: Prediction-area (P-A) plots for discretized posterior probability maps obtained by logistic regression and ILRBSWT respectively.

7. *The Conclusion is somewhat repetition of the text body. Please re-think about the Conclusion.*

Thank you for your comment and the conclusion has been reorganized:

"Given the problems in existing MPM models, this research provides an ILRBSWT tool. We have proven its operability and effectiveness through a case study. This research is also expected to provide a software tool support for geological exploration researchers and workers in overcoming the non-stationarity of spatial variables, missing data, and differences in exploration degree, which should improve the efficiency of MPM work.”

~~Improved~~An improved logistic regression model based on a spatially weighted technique (ILRBSWT v1.0) and its application to mineral prospectivity mapping

Daojun Zhang[1 2*], Na Ren[1], Xianhui Hou[1*]

[1]College of Economics and Management, Northwest A&F University, Yangling 712100, China

[2]Center for Resource Economics and Environment Management, Northwest A&F University, Yangling 712100, China

*Corresponding author: cugzdj@gmail.com (Zhang, D); houxh1019@126.com (Hou, X)

Abstract: ~~Due to complexity~~The combination of complex, multiple minerogenic stages~~,~~ and mineral superposition during geological processes~~, the~~ has resulted in dynamic spatial distributions and non-stationarity of geological variables~~ also exhibit specific trends and non-stationarity.~~. For example, geochemical elements exhibit ~~obvious~~clear spatial ~~non-stationarity~~variability and trends ~~because of the deposition of different types of~~with coverage~~.~~ type changes. Thus, bias ~~may clearly~~is likely to occur under these conditions when general regression models are applied to mineral prospectivity mapping (MPM). In this study, we used a spatially weighted technique to improve general logistic regression and developed an improved model~~ called~~, i.e., the improved logistic regression model, based on a spatially weighted technique (ILRBSWT, version 1.0). The capabilities and advantages of ILRBSWT are as follows: (1) ~~ILRBSWT~~it is~~ essentially~~ a geographically weighted regression (GWR) model, and thus it has all ~~its ~~advantages of GWR when ~~dealing with~~managing spatial trends and non-stationarity; (2) while the current software employed for GWR mainly applies linear regression~~ whereas~~, ILRBSWT is based on logistic regression, which is ~~used ~~more ~~commonly~~ ~~in~~suitable for MPM because mineralization is a binary event; (3) a missing data ~~process~~processing method borrowed from weights of evidence is included in ILRBSWT to

extend ~~the~~its adaptability when ~~dealing with~~managing multisource data; and (4) in addition to geographical distance, the differences ~~of~~in data quality or exploration level can ~~also~~ be weighted in the new model ~~as well as the geographical distance~~.

Keywords: anisotropy; geographical information system modeling; geographically weighted logistic regression; mineral resource assessment; missing data; trend variable; weights of evidence.

1 Introduction

The main distinguishing characteristic of spatial statistics compared ~~with~~to classical statistics is that the former has a location attribute. Before ~~the development of~~ geographical information systems were developed, spatial statistical problems were often transformed into general statistical problems, where the spatial coordinates were ~~more like~~similar to a sample ID because they only had an indexing feature. However, even in non-spatial statistics, the reversal ~~paradox~~ or amalgamation paradox (Pearson et al., 1899; Yule, 1903; Simpson, 1951), which is commonly called Simpson's paradox (Blyth, 1972), has attracted ~~much~~significant attention from statisticians and other researchers. In spatial statistics, some spatial variables ~~usually~~ exhibit certain trends and spatial non-stationarity. Thus, it is possible for Simpson's paradox to occur when a ~~global~~classical regression model is applied, and the existence of unknown important variables may ~~make~~worsen this condition ~~even worse~~. The influence of Simpson's paradox can be fatal. For example, in geology, due to the presence of cover and other factors that occur ~~after~~ post-mineralization, ~~the~~ ore-forming elements in Area I are ~~generally~~ much lower than those in Area II, ~~but~~while the actual probability of a mineral in Area I is higher than that in Area II~~, and~~ simply because more deposits ~~may be~~were discovered in Area I (Agterberg, 1971). In this case, ~~a~~ negative ~~correlation will~~correlations would be obtained between ~~the~~ ore-forming elements and ~~the~~ mineralization according to the classical regression model, whereas ~~a~~ high positive ~~correlation~~correlations can be obtained in both areas if they are separated. Simpson's paradox is an extreme case of ~~the~~ bias ~~caused by using a global model~~generated from classical models, and it is usually not so severe in practice. However, this type of ~~biased~~bias needs to be considered and ~~we should take~~ care needs to be taken when applying a classical regression model to a spatial problem. Several solutions to this issue have been proposed ~~previously~~, which can be divided into three types.

(1) Locations are introduced as direct or indirect independent variables. ~~Several studies have employed spatial trend~~This type of model is still a global model, but space coordinates or distance weights are employed to adjust the regression estimation between the dependent variable and independent variables (Agterberg, 1964; Agterberg and Cabilio, 1969; Agterberg, 1970; Agterberg and Kelly, 1971; Agterberg, 1971~~) to express linear or nonlinear trends in space by adding coordinate variables or their functions in predictive models. In these methods, the locations themselves are taken as independent variables as well as the normal independent variables.~~; Casetti, 1972; Lesage & Pace, 2009, 2011). For example, Reddy et al. (1991) performed logistic regression by including trend variables ~~for mapping~~to map the base-metal potential in the Snow Lake area, Manitoba, Canada~~. In addition, Casetti (1972) developed a~~; Helbich & Griffith (2016) compared the spatial expansion method (SEM) to other methods in modeling the house price variation locally, where the regression parameters are themselves functions of the x and y coordinates ~~as well as~~and their combinations; Yu & Liu (2016) used the spatial lag model (SLM) and spatial error model to control spatial effects in modeling the relationship between $PM_{2.5}$ concentrations and per capita GDP in China.

(2) ~~Using local~~Local models are used to replace global models, i.e., geographically weighted models (Fotheringham et al., 2002). Geographically weighted regression (GWR) is the most popular model among the geographically weighted models. GWR ~~was~~models were first developed at the end of the 20th century by Brunsdon et al. (1996) and Fotheringham et al.

(1996, 1997, 2002) for modeling spatially heterogeneous processes, and ~~it has~~have been used widely in ~~the field of geography.~~geosciences (e.g., Buyantuyev & Wu, 2010; Barbet-Massin et al., 2012; Ma et al., 2014; Brauer et al., 2015).

(3) Reducing ~~the~~ trends in spatial variables. For example, Cheng developed a local singularity analysis technique and spectrum-area (S-A) model based on fractal/multi-fractal theory (Cheng, 1997; Cheng, 1999). These methods can remove spatial trends and ~~prevent~~mitigate the strong effects on predictions of the ~~original~~variables starting at high and low values ~~of the variables on predictions~~, and thus they are used widely to weaken the effect of spatial non-stationarity ~~to some degree~~ (e.g., ~~Zuo~~Zhang et al., 2016; ~~Zhang~~Zuo et al., 2016; Xiao et al., 2017).

GWR models can be readily visualized and ~~understood, and it is particularly valid for dealing with spatial non-stationarity, thus it has been used widely~~are intuitive, which have made them applied in geography and other ~~areas~~disciplines that require spatial data analysis. In general, GWR is a moving window-based model where instead of establishing a unique and global model for prediction, it ~~makes a prediction for~~predicts each current location using the surrounding samples, and a higher weight is given when the sample is located closer. The theoretical foundation of GWR is ~~based on~~ Tobler's observation that: "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970).

In mineral prospectivity mapping (MPM), the dependent variables are binary and logistic regression is used instead of linear regression~~, and~~; therefore, it is necessary to apply geographically weighted logistic regression (GWLR) instead. GWLR ~~belongs to~~is a type of geographically weighed generalized linear regression model (Fotheringham et al~~.~~., 2002) ~~and it~~that is included in the software module GWR 4.09 (Nakaya, 2016). However, the function module for GWLR in current software can only ~~deal with the~~manage data in the form of a tabular dataset containing the fields ~~of~~with dependent and independent variables~~,~~ and ~~the~~ x-y coordinates. Therefore, the spatial layers ~~must~~have to be re-processed into two-dimensional tables and the resulting data needs to be transformed back into a spatial form.

Another problem with ~~the application of~~applying GWR 4.09 for MPM is that it cannot ~~deal with~~handle missing data (Nakaya, 2016). Weights of evidence (WofE) is a widely used model for MPM (Bonham-Carter et al., 1988, 1989; Agterberg, 1989; Agterberg et al., 1990~~),~~ ~~which can avoid~~) that mitigates the ~~effect~~effects of missing data. However, WofE was developed ~~based on the premise that an assumption of~~assuming that conditional independence is satisfied among ~~the~~ evidential layers with respect to the target layer; otherwise, the posterior probabilities will be biased, and the number of estimated deposits will ~~not~~ be ~~equal~~unequal to the known deposits. Agterberg (2011) combined WofE with logistic regression and proposed a new model that can obtain an unbiased ~~estimated of the number of deposits as well as avoiding the effect of missing data. In the present study, this concept is employed to deal with missing data and we propose the improved logistic regression model based on spatially weighted technique (ILRBSWT v1.0) for MPM. The main features of ILRBSWT include the following: (1) a spatial *t* statistics method (Agterberg et al., 1993) is introduced to determine the best binary threshold for independent variables, where binarization is performed based on a local window instead of the global level, which can increase the effect of indicating the independent variables to the target variable; and (2) a mask layer is included in the new model to deal with the data quality and exploration level differences among samples.~~ estimate of number of deposits while also avoiding the effect of missing data. In this study, we employed Agterberg (2011) 's to account for missing data.

~~The idea~~One more improvement of the ILRBSWT is that a mask layer is included in the new model to address data quality and exploration level differences between samples.

Conceptually, this research ~~is origin~~originated from the ~~first author's doctoral~~ thesis ~~(~~of Zhang~~,~~ (2015~~)~~; in Chinese~~,~~), which ~~has been shown to have~~showed better efficiency

Formatted: Font: 12 pt

Formatted: Font: 12 pt

for mapping intermediate and felsic igneous rocks (Zhang et al., 2017). ~~The contribution of this research is to elaborate the principle~~This work elaborates on the principles of ILRBSWT, and ~~provide~~provides a detailed algorithm for its design and implementation process with the code and software module attached. In addition, ~~the~~ processing ~~of~~ missing data is not a technique covered in GWR modeling presented in prior research, and a solution borrowed from WofE is provided in this study. Finally, ILRBSWT performance in MPM is tested by ~~former researches. At last, the prediction of~~predicting Au ore deposits in western Meguma Terrain, Nova Scotia, Canada~~, is chosen as case study to show the performance of ILRBSWT in MPM~~.

2 Models

Linear regression is commonly used for exploring the relationship between a response variable and one or more explanatory variables. However, in MPM and other fields, the response variable is binary or dichotomous, so linear regression is not applicable and thus a logistic model ~~can be~~is advantageous.

2.1 Logistic Regression

In MPM, the dependent variable(Y) is binary ~~since~~because Y can only take the value of 1 and 0, ~~which means the~~indicating that mineralization occurs ~~or~~and not respectively. Suppose that π represents the estimation of Y, $0 \leqslant \pi \leqslant 1$, then a logit transformation of π can be made, i.e., logit (π) $=\ln(\pi/(1-\pi))$. ~~Logistic~~The logistic regression function can be obtained as ~~following.~~follows:

$$\text{Logit}\ \pi\left(X_1, X_2, \cdots, X_p\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p \tag{1}$$

where $X_1, X_2, \cdots, X_p$, comprises a sample of p explanatory variables $x_1, x_2, \cdots, x_p$, β_0 is the intercept, and $\beta_1, \beta_2, \cdots, \beta_p$ are regression coefficients.

If there are n samples, we can obtain n linear equations with p+1 unknowns based on equation (1). Furthermore, if we suppose that the observed values for Y are $Y_1, Y_2, \cdots, Y_n$, and these observations are independent of each other, then a likelihood function can be established:

$$\mathrm{L}(\beta) = \prod_{i=1}^{n} (\pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i}), \tag{2}$$

where $\pi_i = \pi(X_{i1}, X_{i2}, \cdots, X_{ip}) = \frac{e^{\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \cdots + \beta_p X_{ip}}}$. The best estimate can be obtained ~~if and~~ only if equation (2) takes the maximum. Then the problem is converted into solving $\beta_1, \beta_2, \cdots, \beta_p$. Equation (2) can be further transformed into the following log-likelihood function~~.~~:

$$\ln \mathrm{L}(\beta) = \sum_{i=1}^{n} (Y_i \pi_i + (1 - Y_i)(1 - \pi_i)) \tag{3}$$

The solution can be obtained by taking the first partial derivative of β_i (i = 0 to p), which should be equal to 0~~.~~:

$$\begin{cases} f(\beta_0) = \sum_{i=0}^{n} (Y_i - \pi_i) X_{i0} = 0 \\ f(\beta_1) = \sum_{i=0}^{n} (Y_i - \pi_i) X_{i1} = 0 \\ \vdots \\ f(\beta_p) = \sum_{i=0}^{n} (Y_i - \pi_i) X_{ip} = 0 \end{cases} \tag{4}$$

where $X_{i0} = 1$, i takes the value from 1 to n, and equation (4) is obtained in the form of matrix operations.

$$\mathbf{X}^{\mathbf{T}} (\mathbf{Y} - \boldsymbol{\pi}) = \mathbf{0} \tag{5}$$

The Newton iterative method can be used to solve the nonlinear equations:

$$\hat{\boldsymbol{\beta}}(t + 1) = \hat{\boldsymbol{\beta}}(t) + \mathbf{H}^{-1} \mathbf{U} \ , \tag{6}$$

where $\mathbf{H} = \mathbf{X}^{\mathbf{T}} \mathbf{V}(t) \mathbf{X}$, $\mathbf{U} = \mathbf{X}^{\mathbf{T}} (\mathbf{Y} - \boldsymbol{\pi}(t))$, t represents the number of iterations, and $\mathbf{V}(t)$, $\mathbf{X}$, $\mathbf{Y}$, $\boldsymbol{\pi}(t)$, and $\hat{\boldsymbol{\beta}}(t)$ are obtained as follows:

$$\mathbf{V}(t) = \begin{pmatrix} \pi_1(t)(1 - \pi_1(t)) & & & \\ & \pi_2(t)(1 - \pi_2(t)) & & \\ & & \ddots & \\ & & & \pi_n(t)(1 - \pi_n(t)) \end{pmatrix},$$

$$\mathbf{X} = \begin{pmatrix} X_{10} & X_{11} & \cdots & X_{1p} \\ X_{20} & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n0} & X_{n1} & \cdots & X_{np} \end{pmatrix},\ \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_1 \\ \vdots \\ Y_n \end{pmatrix},\ \boldsymbol{\pi}(t) = \begin{pmatrix} \pi_1(t) \\ \pi_2(t) \\ \vdots \\ \pi_n(t) \end{pmatrix}, \text{ and } \hat{\boldsymbol{\beta}}(t) = \begin{pmatrix} \hat{\beta}_1(t) \\ \hat{\beta}_2(t) \\ \vdots \\ \hat{\beta}_n(t) \end{pmatrix}.$$

~~Hosmer et al. (2013) provided~~For a more ~~information about~~detailed description of the ~~derivation from~~derivations of equations (1) to (6), see Hosmer et al. (2013).

2.2 Weighted Logistic Regression

In practice, vector data is ~~popularly~~often used, and sample size (area) has to be considered. In this condition, weighted logistic regression modeling should be used instead of a general logistic regression. ~~In addition, it~~It is also preferable to use a weighted logistic regression model when a logical regression should be performed for large sample data~~, since~~ because weighted logical regression can ~~greatly~~significantly reduce ~~the size of the~~ matrix size and improve ~~the~~ computational efficiency (Agterberg, 1992). Assuming that there are four binary explanatory variable layers and the study area consists of 1000×1000 grid points, the matrix size for normal logic regression modeling would be $10^6 \times 10^6$; however, if weighted logistic regression is used, the matrix size would be 32×32 at most. ~~That is~~This condition arises because the sample classification process is contained in the weighted logistic regression, and all samples are classified into ~~the~~ classes ~~which own~~with the same values ~~at~~as the dependent and ~~each~~ independent variables. The samples with the same dependent and independent variables form certain continuous and discontinuous patterns in space, which are called "unique condition" units. Each unique condition unit is then treated as a sample, and the area (grid number) for it is taken as weight in the weighed logistic regression. Thus, ~~in~~for the ~~case of~~ weighted logical regression, equations (2) to (5) in section 2.1 need to be changed ~~as following Equations~~to equations (7) to (10) ~~respectively~~as follows.

$$L_{new}(\beta) = \prod_{i=1}^{n} \left(\pi_i^{N_i Y_i} (1 - \pi_i)^{N_i (1 - Y_i)}\right), \tag{7}$$

$$\ln L_{new}(\beta) = \sum_{i=1}^{n}(N_i Y_i \pi_i + N_i(1-Y_i)(1-\pi_i)) \tag{8}$$

$$\begin{cases} f_{new}(\beta_0) = \sum_{i=0}^{n}(Y_i - \pi_i)X_{i0} = 0 \\ f_{new}(\beta_1) = \sum_{i=0}^{n}(Y_i - \pi_i)X_{i1} = 0 \\ \qquad\vdots \\ f_{new}(\beta_p) = \sum_{i=0}^{n}(Y_i - \pi_i)X_{ip} = 0 \end{cases} \tag{9}$$

$$\mathbf{X}^{\mathbf{T}}\mathbf{W}(\mathbf{Y} - \boldsymbol{\pi}) = \mathbf{0} \tag{10}$$

where N_i is the weight for the *i*-th unique condition unit, *i* takes the value from 1 to *n*, and *n* is the ~~total~~ number of ~~grid points. And~~unique condition units. **W** is a diagonal matrix ~~which can be~~that is expressed as ~~following.~~follows:

$$\mathbf{W} = \begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_n \end{pmatrix}$$

~~Besides~~In addition, new values of H and U should be used in equation (6) to perform Newton ~~iterative under~~iteration as part of the weighted logistic regression, i.e., $\mathbf{H}_{new} = \mathbf{X}^{\mathbf{T}}\mathbf{W}\mathbf{V}(t)\mathbf{X}$, $\mathbf{U}_{new} = \mathbf{X}^{\mathbf{T}}\mathbf{W}(\mathbf{Y} - \boldsymbol{\pi}(t))$.

2.3 Geographically Weighted Logistic Regression

GWLR is a local window-based model ~~because~~where logistic regression is established at each current location in the GWLR. The current location is changed using the moving window technique with a loop program. ~~If we suppose~~Suppose that $\mathbf{u}$ represents the current location, which can be uniquely determined by a pair of column and row numbers, $\boldsymbol{x}$ denotes ~~that~~ *p* explanatory variables $x_1, x_2, \cdots, x_p$ that take values of $X_1, X_2, \cdots, X_p$ ~~,~~ respectively, and $\pi(\boldsymbol{x}, \mathbf{u})$ is the ~~estimates of~~ ~~*Y*~~*Y* estimate, i.e., the probability that *Y* takes a value of 1, and then the following function can be obtained.

$$\text{Logit}\,\pi(\boldsymbol{x}, \mathbf{u}) \cancel{= \beta_{ui}(\mathbf{u}) + \beta_1} = \beta_0(\mathbf{u}) + \beta_1(\mathbf{u})\cancel{X_1}x_1 + \beta_2(\mathbf{u})\cancel{X_2}x_2 + \cdots + \beta_p(\mathbf{u})\cancel{X_p}x_p \quad , \tag{11}$$

where~~,~~ $\beta_0(\mathbf{u})$, $\beta_1(\mathbf{u})$, $\cdots$, $\beta_p(\mathbf{u})$ ~~denote~~indicate that these parameters are obtained at the location of $\mathbf{u}$. ~~The~~ Logit $\pi(\boldsymbol{x}, \mathbf{u})$, the predicted probability for the current location $\mathbf{u}$, can be

Formatted: English (United States)

obtained under the condition that the values of all ~~the~~ independent variables are known at the current location and all ~~of the~~ parameters are also calculated based on the samples within the current local window. According to equation (6) in section 2.1, the parameters for GWLR can be estimated with equation (12):

$$\widehat{\boldsymbol{\beta}}(\mathbf{u})_{t+1} = \widehat{\boldsymbol{\beta}}(\mathbf{u})_t + (\mathbf{X}^{\mathrm{T}}\mathbf{W}(\mathbf{u})\mathbf{V}(t)\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}(\mathbf{u})(\mathbf{Y} - \boldsymbol{\pi}(t)), \quad (12)$$

Field Code Changed

where t represents the number of iterations; $\mathbf{X}$ is a matrix ~~comprising~~that includes the values of all ~~the~~ independent ~~variable~~variables, and all ~~of the~~ elements in the first column are 1; $\mathbf{W}(\mathbf{u})$ is a diagonal matrix where the diagonal elements are geographical weights, which can be calculated according to distance, whereas the other elements are all 0; $\mathbf{V}(t)$ is also a diagonal matrix and the diagonal element can be expressed as $\pi_i(t)(1-\pi_i(t))$; and $\mathbf{Y}$ is a column vector representing the values taken by the dependent variable.

2.4 Improved Logistic Regression Model based on the Spatially Weighted Technique

As is mentioned in the introduction section, there are primarily two improvements for ILRBSWT compared to GWLR, i.e., the capacity to manage different types of weights, and the special handling of missing data.

2.4.1 Integration of Different Weights

If a diagonal element in $\mathbf{W}(\mathbf{u})$ is only for one sample ~~(~~, i.e., the grid point in raster data~~)~~, section 2.3 ~~can be seen as the~~ is an improvement ~~of~~on section 2.1, i.e. samples are weighted according to ~~its~~their location. If samples are first reclassified ~~firstly~~ according to the unique condition mentioned in section 2.2, and corresponding weights are then summarized according to each sample's geographical weight, we can obtain an improved logistic regression model considering both sample ~~sizes~~size and geographical ~~distances~~distance. The new model ~~can not only~~both reflects the spatial distribution of samples~~, but also reduce~~ and reduces the matrix size, ~~and it~~which is ~~to be~~ discussed in the following section.

In addition to geographic factors, ~~the degree considered in the study can affect the~~

representativeness of a sample, e.g., differences in the level of exploration, is also considered in this study.

Suppose that there are n grid points in the current local window, S_i is the i-th grid, $W_i(g)$ is the geographical weight of S_i, and $W_i(d)$ represents the individual difference weight or non-geographical weight~~ (in~~. In some cases, there may be differences in quality or the exploration level among samples, but $W_i(d)$ takes a value of 1 if there is no difference~~),~~, where i takes a value from 1 to n. Furthermore, if we suppose that there are N unique conditions after overlaying all ~~of the~~ layers ($N \le n$) and C_j denotes the j-th unique condition unit, then we can obtain the final weight for each unique condition unit in the current local window:

$$W_j(t) = \sum_{i=1}^{n}[W_i(\mathrm{g}) * W_i(d) * \mathrm{df}_i], \qquad (13)$$

where $\begin{cases} \mathrm{df}_i = 1 & \text{if } \mathrm{S}_i \in \mathrm{C}_j \\ \mathrm{df}_i = 0 & \text{if } \mathrm{S}_i \notin \mathrm{C}_j \end{cases}$, i takes a value from 1 to n, j takes a value from 1 to N, and $W_j(t)$ represents the total weight (by combining both $W_i(g)$ and $W_i(d)$) for each unique condition unit. We can use the final weight calculated in equation (13) to replace the original weight in equation (12), which is ~~one of the advantages~~an advantage of ILRBSWT.

2.~~5~~4.2 *Missing data processing*

Missing data is a problem~~ existing~~ in all statistics-related research fields. In MPM, missing data are also prevalent due to ground coverage, and limitations of exploration technique and measurement accuracy. Agterberg and Bonham-Carter (1999) ~~once~~ compared the following commonly used missing data processing solutions: (1) removing variables containing missing data, (2) deleting samples with missing data, (3) using 0 to replace ~~the~~ missing data, and (4) replacing ~~the~~ missing data with the mean of the corresponding variable. ~~From the point of utilization efficiency of~~To efficiently use existing data, both (1) and (2) are clearly not good solutions ~~since~~as more data will be lost. Solution (3) is superior to (4) ~~for missing values due~~

Formatted

Formatted

~~to the detection limit of the measuring instrument~~in the condition that work has not been done and real data is unknown; with respect to ~~the~~ missing data caused by ~~the limitation of geographical environment and the prospecting technique~~detection limits, solution (4) is ~~obviously~~clearly a better choice. Missing data is ~~mainly~~primarily caused by the latter in MPM, and Agterberg (2011) pointed out that missing data ~~could be even~~was better ~~dealt with by performing~~addressed in a WofE model. In WofE, the evidential variable takes the value of positive weight (W^+) if it is favorable for the ~~happening of the~~ target variable (e.g., mineralization~~); and~~), while the evidential variable takes the value of negative weight (W^-) if it is unfavorable for the ~~happening of the~~ target variable; and automatically the evidential variable takes the value of "0" if there is missing data ~~happens~~.

$$W^+ = \ln \frac{\frac{D_1}{D}}{\frac{A_1-D_1}{A-D}} \tag{14}$$

$$W^- = \ln \frac{\frac{D_2}{D}}{\frac{A_2-D_2}{A-D}} \tag{15}$$

where **A** is an evidential layer, A_1 ~~means~~is the area (or grid number, similarly hereinafter) that A takes the value of 1, and A_2 ~~means~~is the area that A takes the value of 0; A_3 ~~means~~is the area with missing data, and A_1+A_2 is smaller than the total study area if missing data exists. D_1, D_2, and D_3 are ~~the area that~~areas where the target ~~variable takes~~variables take the value of 1 in A_1, A_2, and A_3 respectively. ~~In fact,~~ A_3 and D_3 are not used in equation (15) ~~since~~because no information is provided in area A_3.

~~However, it is preferred to use~~ If "1" and "0" are still used to represent the ~~positive and negative~~binary condition of the independent variable ~~in logistic regression model. In this case,~~instead of $W^+$and W^-, equation (16) can be used to replace missing data in logistic regression modeling~~, which will cause an equivalent effect just as missing data are processed in WofE~~.

Formatted: Font: Bold
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript
Formatted: Font: Italic
Formatted: Subscript

$$M = \frac{-W^-}{W^+ - W^-} = \frac{\ln\frac{D}{A-D} - \ln\frac{D_2}{A_2-D_2}}{\ln\frac{D_1}{A_1-D_1} - \ln\frac{D_2}{A_2-D_2}} \tag{16}$$

3 Design of the ILRBSWT Algorithm

3.1 Local Window Design

A raster data set is used for ILRBSWT modeling. With ~~a~~ regular ~~grid~~grids, the distance between any two grid points can be calculated easily and ~~we can even obtain~~ distance templates within a certain window scope can be obtained, which is highly efficient for data processing. The circle and ellipse are used for isotropic and anisotropic local window designs, respectively.

(1) Circular Local Window Design

~~If we suppose~~Suppose that *W* represents a local circular window where the minimum bounding rectangle is *R*, then the geographical weights can be calculated only inside *R*. ~~Obviously~~Clearly, the grid points inside ~~of~~ *R* but outside of *W* should be weighted as 0, and the ~~weights~~weight for the grid ~~points~~with a center inside *W* should be calculated according to the ~~distances between themselves and the~~distance from its current location. Because *R* ~~should be~~is a square~~ so~~, we can also assume that there are *n* columns and rows in ~~*R*~~it, where *n* is an odd number. If we take east and south as the orientations of the *x*-axis and *y*-axis, respectively, and the position of the northwest corner grid is defined as (x = 1, y = 1), then a local rectangular coordinate system can be established and the position ~~for~~of the current location grid can be expressed as $O\ (x = \frac{n+1}{2},\ y = \frac{n+1}{2})$. The distance between any grid inside *W* and the current location grid can be expressed as $d_{o-ij} = \sqrt{\left(i - \frac{n+1}{2}\right)^2 + \left(j - \frac{n+1}{2}\right)^2}$, where *i* and *j* take values ranging from 1 to *n*. The geographical weight is a function of distance, so it is convenient to calculate w_{ij} with d_{o-ij}~~ .~~. Figure 1 shows the weight template for a circular local window with a half-window size of nine ~~grid points.~~ grids.

0	0	0	0	0	0	0	0	w30	0	0	0	0	0	0	0	0
0	0	0	0	0	w28	w27	w25	w24	w25	w27	w28	0	0	0	0	0
0	0	0	w29	w26	w23	w21	w20	w19	w20	w21	w23	w26	w29	0	0	0
0	0	w29	w25	w22	w18	w16	w15	w14	w15	w16	w18	w22	w25	w29	0	0
0	0	w26	w22	w17	w14	w13	w11	w10	w11	w13	w14	w17	w22	w26	0	0
0	w28	w23	w18	w14	w12	w9	w8	w7	w8	w9	w12	w14	w18	w23	w28	0
0	w27	w21	w16	w13	w9	w6	w5	w4	w5	w6	w9	w13	w16	w21	w27	0
0	w25	w20	w15	w11	w8	w5	w3	w2	w3	w5	w8	w11	w15	w20	w25	0
w30	w24	w19	w14	w10	w7	w4	w2	w1	w2	w4	w7	w10	w14	w19	w24	w30
0	w25	w20	w15	w11	w8	w5	w3	w2	w3	w5	w8	w11	w15	w20	w25	0
0	w27	w21	w16	w13	w9	w6	w5	w4	w5	w6	w9	w13	w16	w21	w27	0
0	w28	w23	w18	w14	w12	w9	w8	w7	w8	w9	w12	w14	w18	w23	w28	0
0	0	w26	w22	w17	w14	w13	w11	w10	w11	w13	w14	w17	w22	w26	0	0
0	0	w29	w25	w22	w18	w16	w15	w14	w15	w16	w18	w22	w25	w29	0	0
0	0	0	w29	w26	w23	w21	w20	w19	w20	w21	w23	w26	w29	0	0	0
0	0	0	0	0	w28	w27	w25	w24	w25	w27	w28	0	0	0	0	0
0	0	0	0	0	0	0	0	w30	0	0	0	0	0	0	0	0

Fig. 1 Weight template for a circular local window with a half-window size of nine ~~grid points~~grids, where w1 to w30 represent different weight classes that decrease with ~~distance~~distances and 0 ~~denotes~~indicates that the grid is weighted as 0. Gradient colors ranging from red to green are used to distinguish the weight classes for grid points.

~~If we suppose~~Suppose that there are T_n columns and T_m rows in the study area, and *Current* (T_i, T_j) represents the current location, where T_i takes values from 1 to T_n and T_j takes values from 1 to T_m, then the current local window can be established by selecting the range of rows $T\_i - \frac{n-1}{2}$ to $T\_i + \frac{n-1}{2}$ and columns $T\_j - \frac{n-1}{2}$ to $T\_j + \frac{n-1}{2}$ ~~based on~~from the total research area. Next, we can establish a local rectangular coordinate system according to the ~~steps in the last paragraph, where the *x* and *y* coordinates for the northwest corner are defined as the coordinate origin by subtracting~~previously described steps; we subtract $T\_i - \frac{n-1}{2}$ and $T_j - \frac{n-1}{2}$ ~~from~~on the x and y coordinates~~,~~ respectively~~,~~ for all ~~of the grid points~~grids in the range. The corresponding relationship can then be established between the weight template and ~~the~~ current window. Global weights can also be included via the matrix product between the global weight layer and local weight template within the local window. In addition, special care should be taken when the weight template covers some area outside the study area, i.e.~~g.,~~, $T\_i - \frac{n-1}{2} < 0$, $T\_i + \frac{n-1}{2} > T\_n$, $T\_j - \frac{n-1}{2} < 0$, and $T\_j + \frac{n-1}{2} > T\_m$.

(2) Elliptic Local Window Design

In most cases, the spatial ~~weights change to~~tendency of the spatial variable ~~degrees in~~may vary with different directions and an elliptic local window may ~~be~~ better ~~for describing~~describe the changes in ~~the~~ weights in space. ~~In order to~~To simplify the calculation, we can convert the distances in different directions into equivalent distances, and an anisotropic problem is then ~~becomes~~converted into an isotropic problem. For any grid, the equivalent distance is the semi-major axis length of the ellipse that ~~passes through the grid and that~~ is centered at the current location~~, where~~ and passes through the grid, while the parameters for the ellipse can be determined using the kriging method.

We still use W to represent the local elliptic window and a, r, and θ are defined as the semi-major axis, ~~the~~ ratio of the semi-minor axis relative to the semi-major axis, and ~~the~~ azimuth of the semi-major axis, respectively. Then, W can be covered by a square R~~, where the~~ whose side length is $2a$-1 and ~~the~~ center is the same as W. There are $(2a - 1) \times (2a - 1)$ ~~grid points~~grids in R. We establish the rectangular coordinates as described above and ~~we~~ suppose that the center of the top left grid in R is located at (x = 1, y = 1), and thus the center of W should be $O(x_0 = a, y_0 = a)$. According to the definition of the ellipse, two of the elliptical focuses are located at F_1 $\left(x_1 = a + \sin(\theta)\sqrt{a^2 - (a * r)^2}, y_1 = a -\right.$

Formatted: Font: Italic

$\mathrm{con}(\theta)\sqrt{a^2-(a*r)^2}\big)$ and F_2 $(x_2 = a - \sin(\theta)\sqrt{a^2-(a*r)^2}, y_2 = a + \mathrm{con}(\theta)\sqrt{a^2-(a*r)^2}$~~) .~~). The summed distances between a point and ~~the~~ two focus points can be expressed as $l_{ij} = \sqrt{(i-x_1)^2+(j-y_1)^2} + \sqrt{(i-x_2)^2+(j-y_2)^2}$, where i and j take values from 1 to 2a – 1. According to the elliptical focus ~~formula, we can decide whether a grid in *R* is located in *W*. For~~equation, for any grid in *R*, if the sum of the distances between the two focal points and a grid center is greater than 2*a*, ~~then~~ the grid is located ~~in~~within *W*, and vice versa. For the ~~grid points~~grids outside of *W*, the weight is assigned as 0, and we only need to calculate the equivalent distances ~~should be calculated~~ for the ~~grid points~~grids within *W*. As mentioned above, the parameters for the ellipse can be determined using the kriging method. In ~~the~~ ellipse *W*, where the semi-major axis is *a*, ~~we keep~~ *r* and θ are maintained as constants, ~~so~~then we ~~can~~ obtain countless ellipses centered at the center of *W*, and the equivalent distance is the same on the same elliptical orbit. Thus, the equivalent distance template can be obtained for the local elliptic ~~local~~ window. Figure 2 shows the equivalent distance templates under the conditions that *a* = 11 ~~grid points~~grids, *r* = 0.5, and the azimuths for the semi-major axis are 0°, 45°, 90°, and 135°~~, where the weight template can also be calculated based on Fig. 2.~~ respectively.

Formatted: Font: Italic

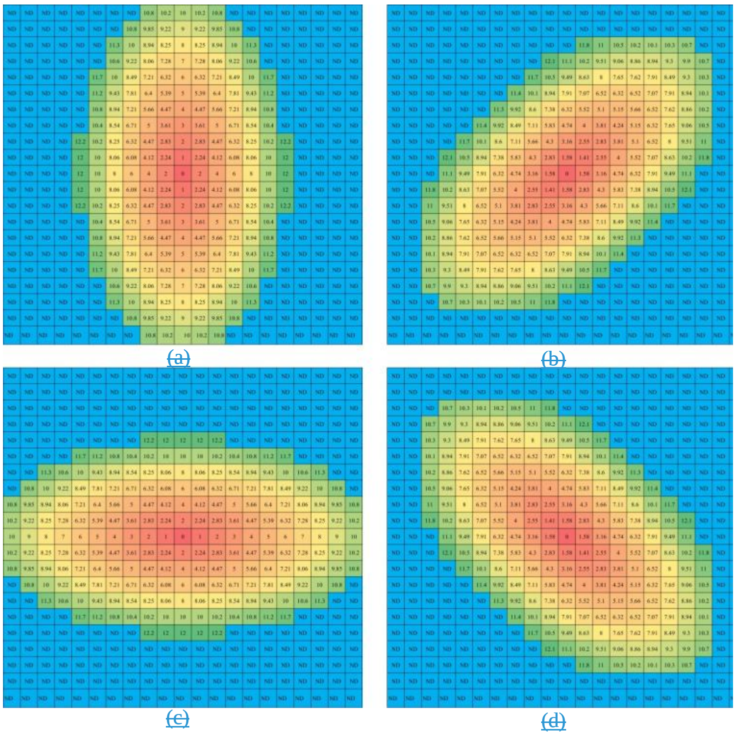


Fig. 2 Construction of the distance template based on ~~a~~an elliptic local window: *a* = 11 grid points, *r* = 0.5, and the azimuths for the semi-major axis are 0° (a), 45°(b), 90° (c), and 135° (d~~).~~) respectively.

3.2 ~~Pseudocode~~Algorithm for ILRBSWT

The ILRBSWT method primarily focuses ~~mainly~~ on two problems, i.e., spatial non-stationarity and missing data. We use the moving window technique to establish ~~a~~ local models instead of a global model~~, which can~~ to overcome ~~the~~ spatial non-stationarity ~~better compared with the global model~~. The spatial *t*-value employed in the WofE method is used to binarize spatial variables based on the local window, which is quite different from traditional binarization based on the global range, where ~~the~~ missing data can be handled well because positive and negative weights are used instead of the original values of "1" and "0" ~~values~~, and ~~the~~ missing data ~~can then be~~are represented ~~well~~ as "0." Both the isotropy and anisotropy window types are ~~possible~~provided in our new proposed model. The geographical ~~weights~~weight function and ~~the~~ window size can be determined by the users ~~themselves~~. If the geographic weights are equal and there are no missing data, ~~then~~ ILRBSWT will yield the same posterior probabilities as classical logistic regression; hence, the later can be ~~treated~~viewed as a special case of the former. The core ILRBSWT algorithm is as follows.

Step 1. Establish a loop for all ~~of the grid points~~grids in the study area according to both the columns and rows. Determine a basic local window with a size of $r_{\min}$ based on a variation function or other method. In addition, the maximum local window ~~with a~~ size ~~of $r_{\max}$~~ is set as $r_{\max}$, with an interval of ΔR. ~~If we suppose~~Suppose that a geographical ~~weight~~weighted model has already been given in the form of a Gaussian curve determined ~~by~~from variations in ~~the~~ geostatistics, i.e., $W(\mathrm{g}) = \mathrm{e}^{-\lambda d^2}$, where d is the distance and λ is the attenuation coefficient, then we can calculate the geographical weight for any grid in the current local window. The equivalent radius should be used in the anisotropic situation. When other types of weights are considered, e.g., the degree of exploration or research, it is also necessary to synthesize the geographical weights ~~and~~with other weights (see equation ~~10~~13).

Step 2. Establish a loop for all ~~of the~~ independent variables. In a circular (elliptical) window with a radius (equivalent radius) of $r_{\min}$, apply the WofE (Agterberg, 1992) model according to the grid weight determined in step 1, thereby obtaining a statistical table containing the parameters ~~of~~ W_{ij}^{+}, W_{ij}^{-}, and t_{ij}, where i is the i-th independent variable and j denotes the j-th binarization.

Step 2.1. If a maximum t_{ij} exists and it is greater than or equal to the standard t-value (e.g., 1.96), record the values of $W_{i-\max\_t}^{+}$, $W_{i-\max\_t}^{-}$, and $B_{i-\max\_t}$, which denote the positive weight, negative weight, and corresponding binarization, respectively, under the condition where t takes the maximum value. Go to step 2 and apply the WofE model to the other independent variables.

Step 2.2. If a maximum t_{ij} does not exist, or it is smaller than the standard *t*-value, go to step 3.

Step 3. In a circular (elliptical) window with a radius (equivalent radius) of $r_{\max}$, increase the current local window ~~based on~~radius from $r_{\min}$ according to the algorithm in step 1.

Step 3.1. If all ~~of the~~ independent variables have already been processed, go to step 4.

Step 3.2. If the size of the current local window exceeds the size of $r_{\max}$, ~~then~~ disregard the current independent variable and go to step 2 to consider the remaining independent variables.

Step 3.3. Apply the WofE model according to the grid weight determined in step 1 in the current local window~~, which has increased.~~. If a maximum t_{ij} exists and it is greater than or equal to the standard *t*-value, record the values of $W^{+}_{i-\max\_t}$, $W^{-}_{i-\max\_t}$, $B_{i-\max\_t}$~~,~~, and ~~r_{curent}~~$r_{current}$, which ~~represents~~represent the radius (equivalent radius) for the current local window.

Step 3.4. If a maximum t_{ij} does not exist or it is smaller than the standard *t*-value, go to step 3.

Step 4. Suppose that n_s independent variables ~~are remainingstill~~still remain.

Step 4.1. If $n_s \leq 1$, ~~then~~, calculate the mean value for the dependent variable in the current local window with a radius size of $r_{\max}$ and retain it as the posterior probability in the current location. In addition, set the regression coefficients for all ~~of the~~ independent variables as missing data. Go to step 6.

Step 4.2. If $n_s \geq 1$, ~~then~~ find the independent variable with the largest local window and apply the WofE model to all ~~the~~ other independent variables, ~~before recording~~and then update the values of $W^{+}_{i-\max\_t}$, $W^{-}_{i-\max\_t}$~~,~~, and $B_{i-\max\_t}$ ~~for this time, and then go~~. Go to step 5.

Step 5. Apply the logistic regression model based on the previously determined geographic weights, and for each independent variable: (1) use $W^{+}_{i-\max\_t}$ to replace all ~~of the~~

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

values that are less than or equal to $B_{i-\max\_t}$~~;~~, (2) use $W^{-}_{i-\max\_t}$ to replace all ~~of the~~ values that are greater than $B_{i-\max\_t}$~~;~~, and (3) use 0 to replace no data (“-9999”). The posterior probability and regression coefficients can then be obtained for all ~~of the~~ independent variables at the current location~~,~~ and go to step 6.

Step 6. Take the next grid as the current location and repeat steps 2–5.

4 Interface Design

~~In addition to the improved GWLR, we developed other modeling processes, where all of the visualization and mapping procedures are performed~~Before performing spatially weighted logistical regression with ILRBSWT 1.0, data pre-processing is performed using the ArcGIS 10.2 platform and GeoDAS 4.0 software. ~~The maps~~All data are originally stored in grid format, which ~~are~~should be transformed into ASCII files ~~based on tools included in~~ with the Arc toolbox ~~before the improved GLWR is performed.~~ in ArcGIS 10.2; after modeling with ILRBSWT 1.0, the result data will be transformed back into grid format

As shown in Fig. 3, the main interface for ~~the improved GLWR comprises~~ ILRBSWT 1.0 is composed of four parts.

The upper left part is for the layer input settings, where independent variable layers, dependent variable layers, and global weight layers should be assigned ~~if they exist.~~. Layer information is shown at the upper right corner, including ~~the~~ row numbers, column numbers, grid size, ordinate origin, and the expression for missing data. The local window parameters and weight attenuation function can be defined ~~in the middle.~~as follows. Using the drop-down list, we ~~can prepare~~prepared a circle or ellipse to represent various isotropic and anisotropic spatial conditions, respectively. The corresponding window parameters should be set for each window type. For the ellipse, it is necessary to set parameters ~~comprising~~composed of the initial length of the equivalent radius (initial major radius), ~~the~~ final length of the equivalent radius (largest major radius), ~~the~~ increase in the length of the equivalent radius (growth rate), ~~the~~ threshold of the spatial *t*-value used to determine the need to enlarge the window, ~~the~~ length ratio of the major and minor axes, ~~the~~ orientation of the ellipse's major axis, and ~~the~~ compensation coefficient for the sill. ~~Next, it is necessary to define the~~We prepared different types of weight attenuation ~~function and a variety of kernel~~ functions via the drop-down menu to provide choices to users, such as exponential model, logarithmic model, Gaussian model, ~~or~~and spherical model, ~~via the drop-down menu. More~~and corresponding parameters can be set when a certain model is selected. The output file ~~settings are~~is defined at the bottom and the execution button is at the lower right corner.

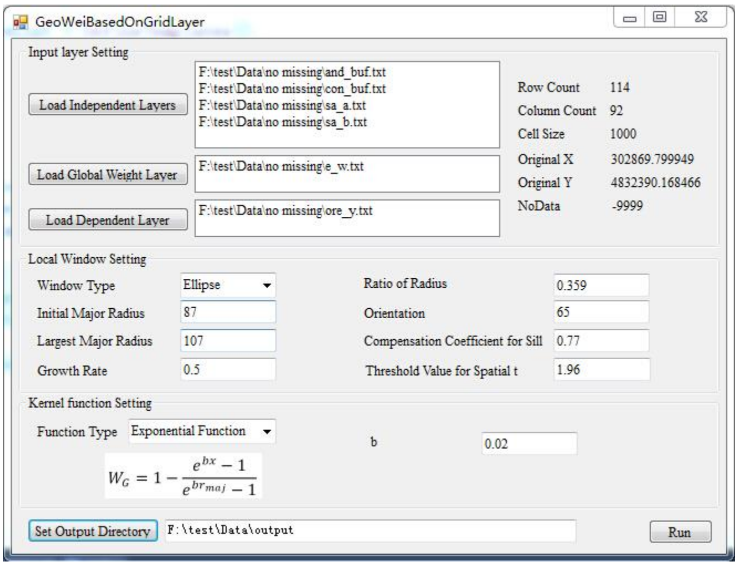


Fig. 3 User interface design.

5 Real Data Testing

5.1 Data source and preprocessing

The test data used in this study were obtained from the case study reported ~~by~~in Cheng (2008). The study area (≈7780 km^2) ~~was~~is located in western Meguma Terrain, Nova Scotia, Canada. Four independent variables were used in the WofE model for gold mineral potential mapping by Cheng (2008), i.e., buffer of anticline axes, buffer for the contact of Goldenville–Halifax Formation, and background and anomaly separated with the S-A filtering method based on ~~the~~ore element loadings of the ~~ore elements of the~~ first component~~. More information about the data set can be found~~, as shown in ~~Cheng (2008).~~Fig. 4.

~~Four~~

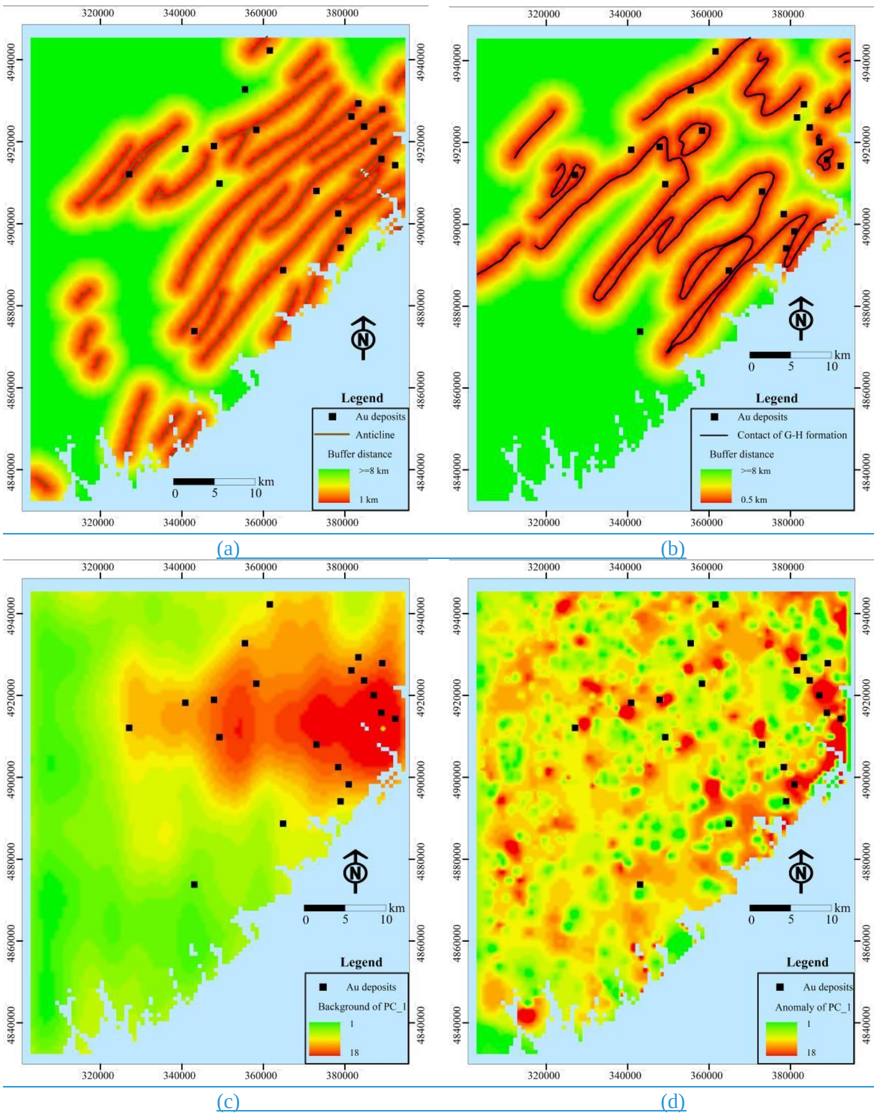


Fig. 4 Evidential layers used to map Au deposits in this study: buffer of anticline axes (a), buffer for the contact of Goldenville–Halifax Formation (b), and background (c) and anomaly (d) separated with the S-A filtering method based on the ore element loadings of the first component.

The four independent variables ~~mentioned above~~described previously were also used for ILRBSWT modeling in this study~~. In order to~~ (see Figs. 4 (a) to (d)), and they were uniformed in the ArcGIS grid format with a cell size of 1 km × 1 km. To demonstrate the advantages of the new method ~~when processing~~ for missing data processing, we designed ~~a~~an artificial situation ~~where the geochemical data were missing for the northern part of the study area, as shown in Fig. 4. In that case,~~in Fig. 5, i.e., grids in region A ~~own~~have values ~~at~~for all ~~of the~~ four independent variables~~; however, grids in region B~~, while they only ~~own~~have values ~~at~~for two independent variables~~,~~ and ~~they have~~ no values in the two geochemical variables~~.~~ in region B.

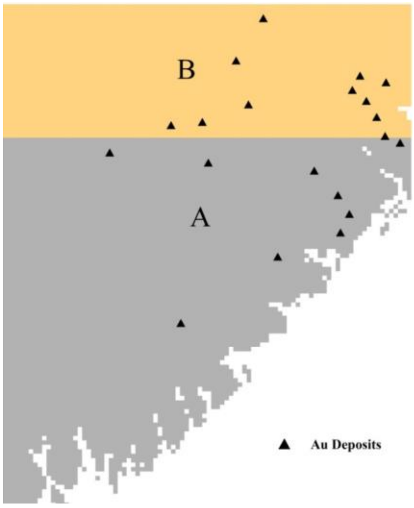


Fig. ~~4~~5 Study area (A and B) ~~and the scope with~~where there is missing geochemical data ~~(~~in area B~~)~~.

5.2 Mapping weights for ~~the~~ exploration ~~level~~

~~These types of~~Exploration level weights can be determined based on prior knowledge ~~according to differences in the exploration data~~about data quality, e.g., different scales may exist throughout the whole study area~~. They~~; however, these weights can also be ~~obtained~~calculated quantitatively. The density of known deposits is a good index for the exploration level, ~~where~~i.e., the ~~degree of~~ research is ~~higher~~more comprehensive when more deposits are discovered. The exploration level ~~weights~~weight layer for the ~~mapped~~ study area was obtained using the kernel density tool provided by the ArcToolbox in ArcGIS 10.2 ~~are~~, as shown in Fig. ~~5~~6.

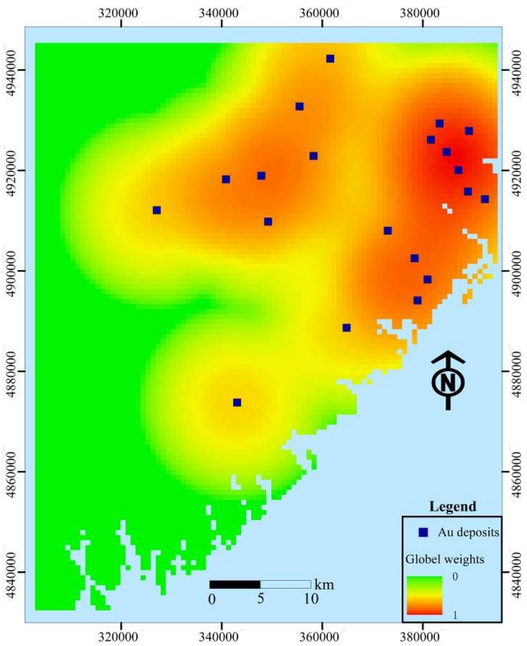


Fig. ~~5~~6 Exploration level weights.

5.3 Parameter Assignment ~~of~~for local window ~~parameters and geographical weights~~and weight attenuation function

~~Empirical~~Both empirical and quantitative methods can be used to determine the local window parameters and ~~the~~ attenuation function for geographical weights. The variation function in geostatistics, which is an effective method for describing the structures and trends ~~of~~in spatial variables, ~~so it~~ was ~~used~~applied in this study. ~~In order to~~To calculate the variation function for ~~a~~the dependent variable, it is necessary to first map the posterior probability using the global logistic regression method~~,~~ before ~~establishing the variation function to determine~~determining the local window type and parameters~~.~~ from the variation function. Variation functions ~~are~~were established in four directions ~~in order~~ to detect anisotropic changes in space. If there are no significant differences among the various directions, a circular local window can be used for ILRBSWT, as shown in Fig. 1; otherwise, an elliptic local window should be used, as shown in Fig. 2. The specific parameters for the local window in the study area were obtained as shown in Fig. ~~6~~7, and the final local window and geographical weight attenuation were determined as indicated in Fig. ~~7~~8 (a) and ~~7~~8 (b), respectively.

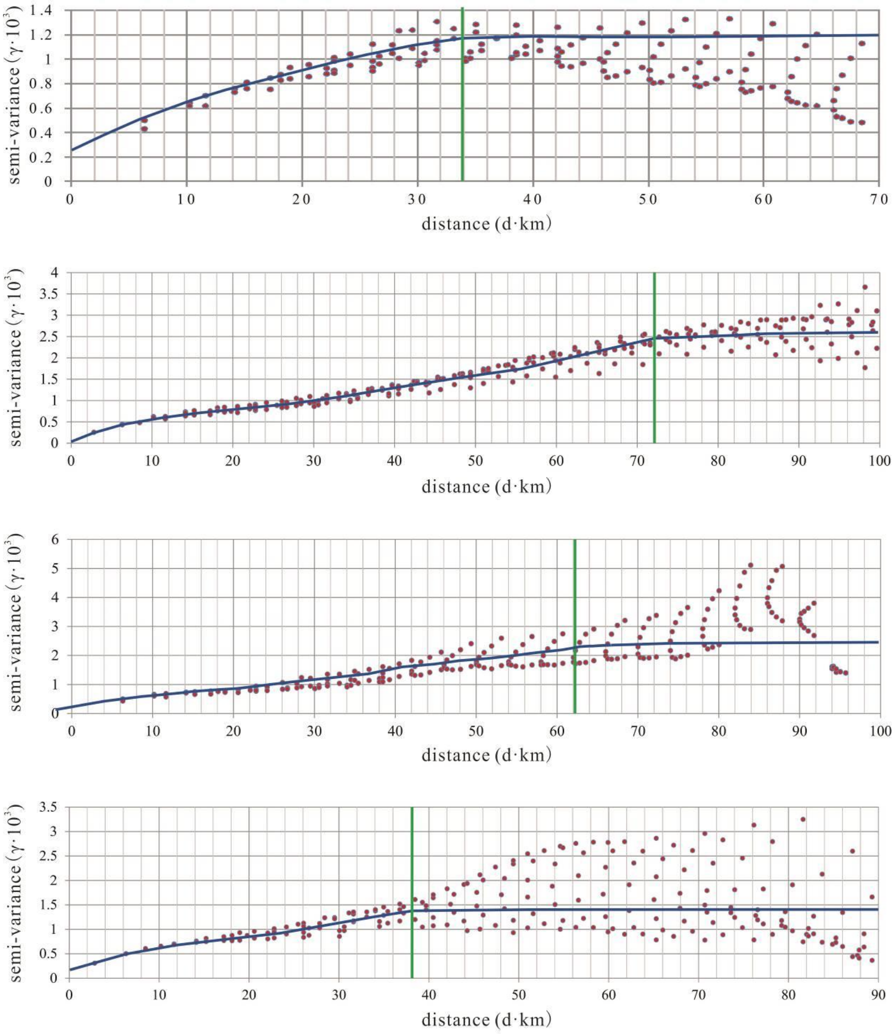
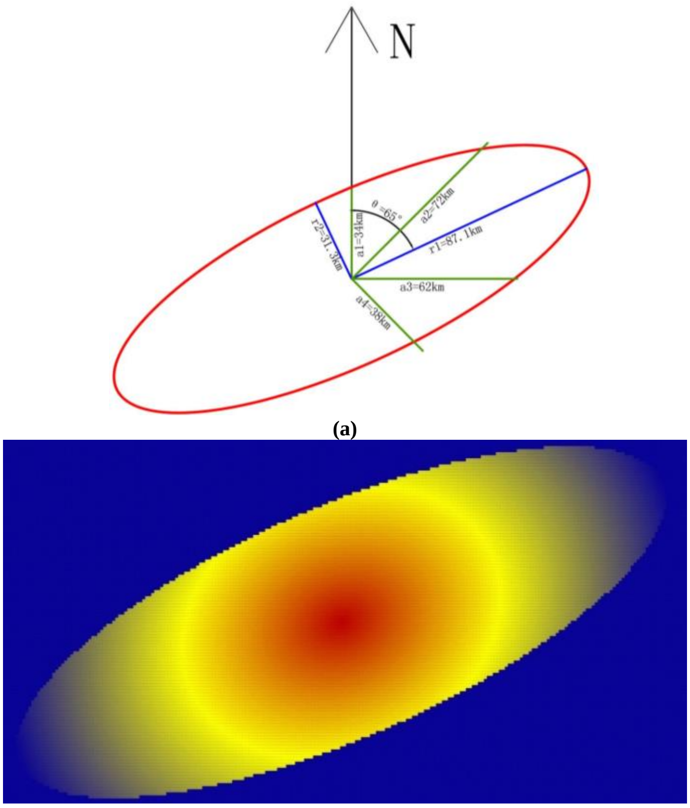


Fig. ~~6~~7 Experimental variogram fitting in different directions, where the green lines denote the

variable ranges determined for azimuths of (a) 0°, (b) 45°, (c) 90°, and (d) 135°.



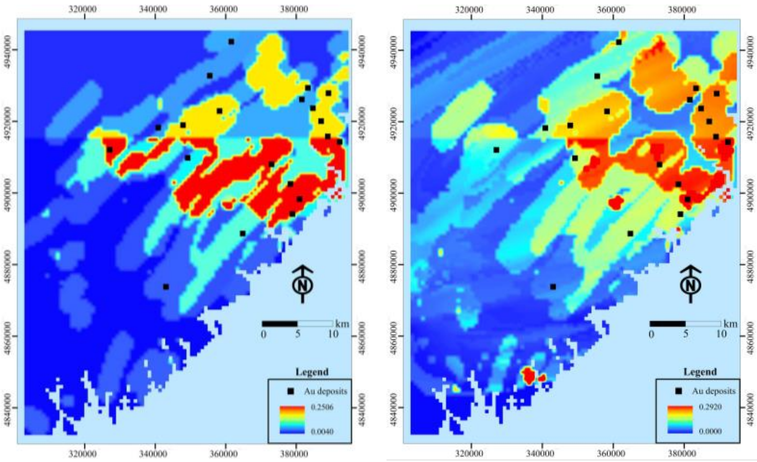
(b)

Fig. ~~7~~8 Nested spherical model for different directions. The green lines in (a) correspond to those in Fig. ~~6~~5, and (b) shows the geographical weight template determined based on (a).

5.4 Data integration

Using the algorithm described in section 3.2, ILRBSWT was ~~performed for~~applied to the study area according to the parameter settings in Fig. 3. The estimated probability map obtained for ~~intermediate and felsic igneous rocks~~ Au deposits by ILRBSWT is shown in Fig. ~~8~~9 (b), while Fig. ~~8~~9 (a) presents the results obtained by logistic regression. ~~It can be seen~~

~~from~~ As shown in Fig. 8 ~~that,~~ ILRBSWT ~~can~~ better ~~weak the effect of~~manages missing data than logistic regression, ~~since~~as the Au deposits in the north part of the study area (~~where~~with missing data ~~exist) are well felled into)~~ better fit within the region with ~~relatively~~ higher posterior probability in Fig. ~~8~~9 (b) than in Fig. ~~8~~9 (a).



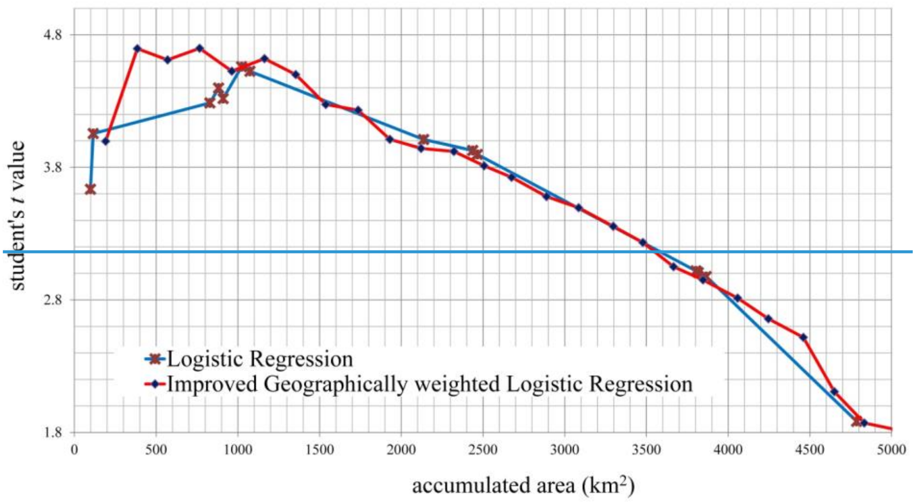
(a) (b)

Fig. ~~8~~9 Posterior probability maps obtained for ~~an~~ Au ~~deposit~~deposits by (a) logistic regression and (b) ILRBSWT.

5.5 Comparison of the mapping results

Formatted: Font: Italic

~~In order to~~To evaluate the predictive capacity of the newly developed ~~method~~ and ~~the~~ traditional ~~method~~methods, the posterior probability maps obtained ~~by~~through logistic regression and ILRBSWT shown in Fig. ~~8(b~~9 (a) and ~~8(a), respectively,~~ 9 (b) were divided into 20 classes ~~by~~using the quantile method ~~and the *t*-values~~ . Prediction-area (P-A) plots (Mihalasky & Bonham-Carter, 2001; Yousefi et al., 2012; Yousefi & Carranza, 2015a) were then ~~calculated using WofE modeling (Fig. 9). Clearly, ILRBSWT performed better because higher *t*-values were obtained, especially when a smaller area was delineated as the target area, which is much more realistic. In the northern part of the study area, the known deposits fitted better to the high~~ made according to the spatial overlay relationships between Au deposits and the two classified posterior probability ~~area shown in Fig. 8(b) than that in Fig. 8(a), which indicates that ILRBSWT can deal with missing data better than logistic regression.~~



~~**Fig. 9 Student's *t* values calculated for the spatial correlation between the known Au deposit layer**~~maps in Fig. 10 (a) and ~~the~~ (b) respectively. In a P-A plot, the horizontal ordinate indicates the discretized classes of a map representing the occurrence of deposits. The vertical scales on the left and right sides indicate the percentage of correctly predicted deposits from the total known mineral occurrences and the corresponding percentage of the delineated target area from the total study area (Yousefi & Carranza, 2015a). As shown in Figs. 10 (a) and (b), with the decline of the posterior probability ~~**layers obtained by logistic regression and ILRBSWT at different**~~ threshold ~~**levels.**~~

Formatted: Font: 12 pt, Not Bold

Formatted: Font: 12 pt, Not Bold

Formatted: Font: 12 pt, Not Bold

Formatted: Font: 12 pt, Not Bold

~~**6 Conclusions**~~

~~In this study, we developed an improved GWLR model ILRBSWT based on logistic regression, WofE, and the current GWR model. Furthermore, a software module was developed for ILRBSWT and a case study demonstrated its capacities~~for the mineral occurrence from left to right on the horizontal axis, more known deposits are correctly predicted, and ~~advantages. Following objectives were achieved:~~

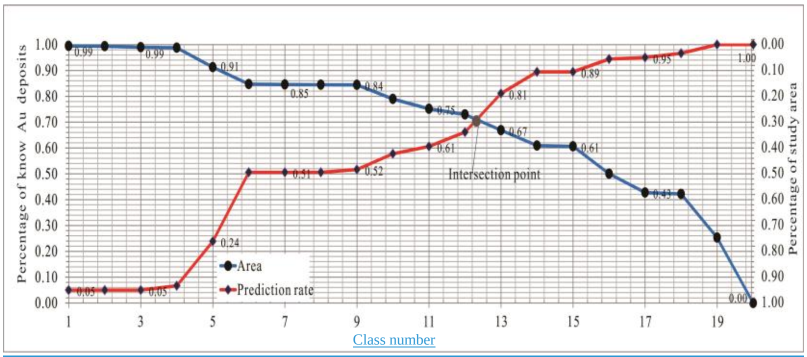
~~(1) A moving window technique is employed for spatial variable parameter logistic regression, which can overcome or weaken the effect of spatial non-stationarity in MPM and improve the accuracy of mineral~~ meantime more areas are delimited as the target area; however, the growth in the prediction~~.~~

~~(2) The variogram model in geostatistics is used to determine the spatial anisotropic parameters~~ rates for deposits and corresponding occupied area are similar before the intersection point in Fig. 10 (a), while the former shows higher growth rate than the latter in Fig. 10 (b). This difference suggests that ILRBSWT can predict more known Au deposits than logistic regression for delineating targets with the same area, and ~~geographical weight attenuation model, which makes the local window parameter design more objective and tenable~~indicates that the former has a higher prediction efficiency than the latter.

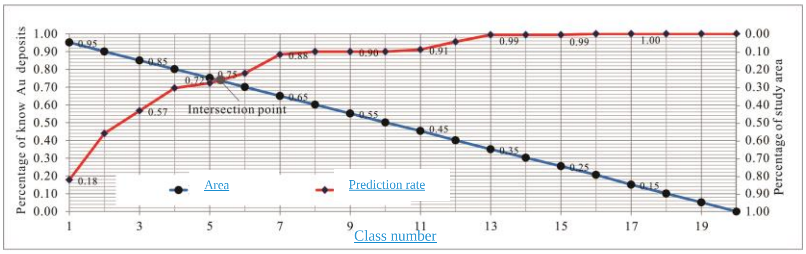
It would be a little inconvenient to consider the ratios of both predicted known deposits and occupied area. Mihalasky and Bonham-Carter (2001) proposed a normalized density, i.e. the ratio of the predicted rate of known deposits to its corresponding occupied area. The intersection point in a P-A plot is the crossing of two curves. The first is fitted from scatter plots of the class number of the posterior probability map and rate of predicted deposit occurrences (the “Prediction rate” curves in Fig. 10). The second is fitted according to the class number of the posterior probability map and corresponding accumulated area rate (the “Area” curves in Fig. 10). At the interaction point, the sum of the prediction rate and corresponding occupied area rate is 1; the normalized density at this point is more commonly used to evaluate the performance of a certain spatial variable in indicating the occurrence of ore deposits (Yousefi & Carranza, 2015a). The intersection point parameters for both models are given in Table 1. As shown in the table, 71% of the known deposits are correctly predicted with 29% of the total study area delineated as target area when the logistic regression is

Formatted: Indent: First line: 0 ch

applied; if ILRBSWT if applied, 74% of the known deposits can be correctly predicted with only 26% of the total area delineated as the target area. The normalized densities for the posterior probability maps obtained from the logistic regression and ILRBSWT are 2.45 and 2.85 respectively; the latter performed significantly better than the former.



(a)



(b)

Fig. 10 Prediction-area (P-A) plots for discretized posterior probability maps obtained by logistic regression and ILRBSWT respectively.

Table 1. Parameters extracted from the intersection points in Figs. 10 (a) and (b).

Model	Prediction rate	Occupied area	Normalized density
Logistic regression	0.71	0.29	2.45
ILRBSWT	0.74	0.26	2.85

~~(3) The spatial *t*-statistics method based on WofE is introduced to perform binarization/discretization for the independent variables in each local window, and the new~~

~~model can better handle missing data.~~

~~(4) The global weight layer in ILRBSWT can reflect differences in the data quality or exploration level well.~~

6 Discussion

Because of potential spatial heterogeneity, the model parameter estimates obtained based on the total equal-weight samples in the classical regression model may be biased, and they may not be applicable for predicting each local region. Therefore, it is necessary to adopt a local window model to overcome this issue. The presented case study shows that ILRBSWT can obtain better prediction results than classical logistic regression because of the former's sliding local window model, and their corresponding intersection point values are 2.85 and 2.45, respectively. However, ILRBSWT has even advantages. For example, characterizing 26% or 29% of the total study area as promising prospecting targets is too high in terms of economic considerations. If instead 10% of the total area needs is mapped as the target area, the proportions of correctly predicted known deposits obtained by ILRBSWT and logistic regression are 44% and 24%, respectively. The prediction efficiency of the former is 1.8 times larger than the latter.

In this study, we did not separately consider the influences of spatial heterogeneity, missing data, and degree of exploration weight all remain, so we cannot evaluate the impact of each factor. Instead, the main goal of this work was to provide the ILRBSWT tool, demonstrating its practicality and overall effect. Zhang et al. (2017) applied this model to mapping intermediate and felsic igneous rocks and proved the effectiveness of the ILRBSWT tool in overcoming the influence of spatial heterogeneity specifically. In addition, Agterberg and Bonham-Carter (1999) showed that WofE has the advantage of managing missing data, and we have taken a similar strategy in ILRBSWT. We did not fully demonstrate the necessity of using exploration weight in this work, which will be a direction for future research. However, it will have little influence on the description and application of ILRBSWT tool as it is not an obligatory factor, and users can individually decide if the exploration weight should be used.

Similar to WofE and logistic regression, ILRBSWT is a data-driven method, thus it inevitably suffers the same problems as data-driven methods, e.g., the information loss caused by data discretization, and exploration bias caused by the training sample location. However, it should be noted that evidential layers are discretized in each local window instead of the total study area, which may cause less information loss. This can also be regarded as an advantage of the ILRBSWT tool. With respect to logistic regression and WofE, some researchers have proposed solutions to avoid information loss resulting from spatial data discretization by performing continuous weighting (Pu et al., 2008; Yousefi & Carranza, 2015b, 2015c, 2016), and these concepts can be incorporated into further improvements of the ILRBSWT tool in the future.

7 Conclusions

Given the problems in existing MPM models, this research provides an ILRBSWT tool. We have proven its operability and effectiveness through a case study. This research is also expected to provide a software tool support for geological exploration researchers and workers in overcoming the non-stationarity of spatial variables, missing data, and differences in exploration degree, which should improve the efficiency of MPM work.

Code availability

The software tool ILRBSWT v1.0 in this research ~~is~~was developed ~~by~~ using C#, and the main codes and key functions are prepared in the file “Codes & Key Functions”. The executable program files are placed in the folder "Executable Programs for ILRBSWT". Please find them in gmd-2017-278-supplement.zip.

Data availability

The data used in this research is sourced from the demo data ~~of~~for GeoDAS software (http://www.yorku.ca/yul/gazette/past/archive/2002/030602/current.htm), ~~and this data is~~which was also used by Cheng (2008). All spatial layers used in this work ~~is~~are included in the folder "Original Data" in the format of an ASCII file, which ~~can be~~is also found in gmd-2017-278-supplement.zip.

Acknowledgments

This study benefited from joint financial support ~~by~~from the Programs of National Natural Science Foundation of China (Nos. 41602336 and 71503200), China Postdoctoral Science Foundation (Nos. ~~2016M592840 and~~ 2017T100773 and2016M592840), Shaanxi Provincial Natural Science Foundation (No. 2017JQ7010), and ~~the~~ Fundamental Research ~~Funds for the Central Universities~~from Northwest A&F University in 2017 (No. 2017RWYB08). The first author thanks former supervisor Drs. Qiuming Cheng and Frits Agterberg for fruitful discussions ~~about~~of spatial weights and ~~for~~ providing constructive suggestions.

References

Agterberg, F.P., & Cabilio, P., 1969. Two-stage least-squares model for the relationship between mappable geological variables. Journal of the International Association for Mathematical Geology, 1(2), 137-153.

Agterberg, F.P., & Kelly, A.M., 1971. Geomathematical methods for use in prospecting. Canadian Mining Journal, 92(5), 61-72.

Agterberg, F.P., 1964. Methods of trend surface analysis. Colorado School Mines Quart, 59(4), 111-130.

Agterberg, F.P., 1970. Multivariate prediction equations in geology. Journal of the International Association for Mathematical Geology, 1970 (02), 319-324.

Agterberg, F.P., 1971. A probability index for detecting favourable geological environments. Canadian Institute of Mining and Metallurgy, 10, 82-91.

Agterberg, F.P., 1989. Computer Programs for Mineral Exploration. Science, 245, 76 – 81.

Agterberg, F.P., 1992. Combining indicator patterns in weights of evidence modeling for resource evaluation. Nonrenewal Resources, 1(1), 35–50.

Agterberg, F.P., 2011. A Modified WofE Method for Regional Mineral Resource Estimation. Natural Resources Research, 20(2), 95-101.

Agterberg, F.P., Bonham-Carter, G.F., & Wright, D.F., 1990. Statistical Pattern Integration for Mineral Exploration. in Gaál, G., Merriam, D. F., eds. Computer Applications in Resource Estimation Prediction and Assessment of Metals and Petroleum. New York: Pergamon Press: 1-12.

Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., & Wright, D.F., 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. Computers in geology, 25, 13-32.

Barbet-Massin, M., Jiguet, F., Albert, C.H., & Thuiller, W., 2012. Selecting pseudo - absences for species distribution models: how, where and how many?. Methods in Ecology & Evolution, 3(2), 327-338.

Blyth, C.R., 1972. On Simpson's paradox and the sure-thing principle.Journal of the American Statistical Association, 67(338), 364-366.

Brauer, M., Freedman, G., Frostad, J., Van Donkelaar, A., Martin, R.V., Dentener, F., ... & Balakrishnan, K., 2015. Ambient air pollution exposure estimation for the global burden of disease 2013. Environmental science & technology, 50(1), 79-88.

Brunsdon, C., Fotheringham, A.S., & Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical analysis, 28(4), 281-298.

Bonham-Carter, G.F., Agterberg, F.P., & Wright, D.F., 1988. Integration of Geological Datasets for Gold Exploration in Nova Scotia. Photogrammetric Engineering & Remote Sensing, 54(11), 1585-1592.

Bonham-Carter, G.F., Agterberg, F.P., & Wright, D.F., 1989. Weights of Evidence Modelling: A New Approach to Mapping Mineral Potential. In Agterberg F P and Bonham-Carter G F, eds. Statistical Applications in the Earth Sciences, 171-183.

Buyantuyev, A., & Wu, J., 2010. Urban heat islands and landscape heterogeneity: linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. Landscape Ecology, 25(1), 17-33.

~~Brunsdon, C., Fotheringham, A.S., & Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical analysis, 28(4), 281-298.~~

Casetti, E., 1972. Generating models by the expansion method: applications to geographic research. Geographical Analysis, 4, 81-91.

Cheng, Q., 1997. Fractal/multifractal modeling and spatial analysis, keynote lecture in proceedings of the international mathematical geology association conference, 1, 57-72.

Cheng, Q., 1999. Multifractality and spatial statistics. Computers & Geosciences, 25, 949–961.

Cheng, Q., 2008. Non-Linear Theory and Power-Law Models for Information Integration and Mineral Resources Quantitative Assessments. Mathematical Geosciences, 40(5), 503-532.

Fotheringham, A.S., Brunsdon, C., & Charlton, M.E., 1996. The geography of parameter space: an investigation of spatial non-stationarity. International Journal of Geographical Information Systems, 10, 605-627.

Fotheringham, A.S., Brunsdon, C., & Charlton, M.E., 2002. Geographically Weighted Regression: the analysis of spatially varying relationships, Chichester: Wiley.

Fotheringham, A.S., Charlton, M.E., & Brunsdon, C., 1997. Two techniques for exploring nonstationarity in geographical data. Geographical Systems, 4, 59-82.

Hao, Y., & Liu, Y., 2016. The influential factors of urban pm 2.5, concentrations in china: aspatial econometric analysis. Journal of Cleaner Production, 112, 1443-1453.

Helbich, M., & Griffith, D.A., 2016. Spatially varying coefficient models in real estate: eigenvector spatial filtering and alternative approaches. Computers Environment & Urban Systems, 57, 1-11.

Hosmer, D.W., Lemeshow, S, & Sturdivant, R.X., 2013. Applied logistic regression, 3rd edn. Wiley, New York

LeSage, J.P., & Pace, R.K., 2009. Introduction to spatial econometrics. Chapman and Hall/CRC.

Lesage, J.P., & Pace, R.K., 2011. Pitfalls in higher order model extensions of basic spatial regression methodology. Review of Regional Studies, 41(1), 13-26.

Ma, Z., Hu, X., Huang, L., Bi, J., & Liu, Y., 2014. Estimating ground-level PM2. 5 in China using satellite remote sensing. Environmental science & technology, 48(13), 7436-7444.

Mihalasky, M.J., & Bonham-Carter, G.F., 2001. Lithodiversity and its spatial association with metallic mineral sites, great basin of nevada. Natural Resources Research, 10(3), 209-226.

Nakaya, T., 2016. GWR4.09 user manual. WWW Document. Available online: https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf (accessed on 16 February 2017).

Pearson, K., Lee, A., & Bramley-Moore, L., 1899. Mathematical contributions to the theory of evolution. VI. Genetic (reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred racehorses.Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 192, 257-330.

Pu, L., Zhao, P., Hu, G., Xia, Q., & Zhang, Z. 2008. The extended weights of evidence model using both continuous and discrete data in assessment of mineral resources gis-based. Geological Science & Technology Information, 27(6), 102-106 (in Chinese with English abstract).

Reddy, R.K.T., Agterberg, F.P., & Bonham-Carter, G.F., 1991. Application of GIS-based logistic models to base-metal potential mapping in Snow Lake area, Manitoba. Proceedings of the Canadian Conference on GIS, 18-22.

Simpson, E.H., 1951. The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society. Series B (Methodological), 238-241.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46(2), 234-24.

Xiao, F., Chen, J., Hou, W., Wang, Z., Zhou, Y., & Erten, O., 2017. A spatially weighted singularity mapping method applied to identify epithermal Ag and Pb-Zn polymetallic mineralization associated geochemical anomaly in Northwest Zhejiang, China. Journal of Geochemical Exploration.

Yousefi, M., & Carranza, E.J.M., 2015a. Prediction–area (P–A) plot and C–A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. Computers & Geosciences, 79, 69-81.

Yousefi, M., & Carranza, E.J.M. 2015b. Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. Computers & Geosciences, 74, 97-109.

Yousefi, M., & Carranza, E.J.M. 2015c. Geometric average of spatial evidence data layers: a GIS-based multi-criteria decision-making approach to mineral prospectivity mapping. Computers & Geosciences, 83, 72-79.

Yousefi, M., Kamkar-Rouhani, A., & Carranza, E.J.M., 2012. Geochemical mineralization probability index (GMPI): a new approach to generate enhanced stream sediment geochemical evidential map for increasing probability of success in mineral potential mapping. Journal of Geochemical Exploration, 115(2), 24-35.

Yule, G.U., 1903. Notes on the theory of association of attributes in statistics. Biometrika, 2(2), 121-134.

Zhang, D., 2015. Spatially Weighted technology for Logistic regression and its Application in Mineral Prospectivity Mapping (Dissertation). China University of Geosciences, Wuhan (in Chinese with English abstract).

Zhang, D., Cheng, Q., & Agterberg, F.P., 2017. Application of spatially weighted technology for mapping intermediate and felsic igneous rocks in fujian province, china. Journal of Geochemical Exploration, 178, 55-66.

Zhang, D., Cheng, Q., Agterberg, F.P., & Chen, Z., 2016. An improved solution of local window parameters setting for local singularity analysis based on excel vba batch processing technology. Computers & Geosciences, 88(C), 54-66.

Zuo, R., Carranza, E.J.M., & Wang, J., 2016. Spatial analysis and visualization of exploration geochemical data. Earth-Science Reviews,158, 9-18.