

Topical Editor Decision: Publish subject to minor revisions (review by editor) (10 May 2018) by Lutz Gross

Comments to the Author:

Daojun

I am happy to inform that the reviewers are happy with your newest version of the manuscript. To be compliant with the GMD code availability I need to ask you to add the source code to the code provided as supplement.

Thanks.

Lutz Gross

GMD Executive Editor

Dear Lutz,

Thanks for your comments, and we have added the source code with necessary notes. The compressed file of *supplement.zip* has been updated, and the new one includes the original data used in this research, the executable programs for ILRBSWT 1.0 which can be run in the window-based operating systems, and the source code in C#. The source code has been recompiled in Visual Studio 2015 in English version so as to assure that all comments and auto-generates by the system are in the English language.

In the new manuscript, we have checked the references and corrected some errors. Besides, some small modifications were made in the sections of *Code Availability*, *Data Availability*, and *Acknowledgments*.

Best regards,

Daojun

May 11, 2018

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

3
4 Daojun Zhang^{1 2*}, Na Ren¹, Xianhui Hou^{1*}

5 ¹College of Economics and Management, Northwest A&F University, Yangling 712100,
6 China

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

9 *Corresponding author: cugzdj@gmail.com (Zhang, D); houXH1019@126.com (Hou, X)

10 **Abstract:** The combination of complex, multiple minerogenic stages and mineral
11 superposition during geological processes has resulted in dynamic spatial distributions and
12 non-stationarity of geological variables. For example, geochemical elements exhibit clear
13 spatial variability and trends with coverage type changes. Thus, bias is likely to occur under
14 these conditions when general regression models are applied to mineral prospectivity
15 mapping (MPM). In this study, we used a spatially weighted technique to improve general
16 logistic regression and developed an improved model, i.e., the improved logistic regression
17 model, based on a spatially weighted technique (ILRBSWT, version 1.0). The capabilities and
18 advantages of ILRBSWT are as follows: (1) it is a geographically weighted regression (GWR)
19 model, and thus it has all advantages of GWR when managing spatial trends and
20 non-stationarity; (2) while the current software employed for GWR mainly applies linear
21 regression, ILRBSWT is based on logistic regression, which is more suitable for MPM
22 because mineralization is a binary event; (3) a missing data processing method borrowed from
23 weights of evidence is included in ILRBSWT to extend its adaptability when managing
24 multisource data; and (4) in addition to geographical distance, the differences in data quality
25 or exploration level can be weighted in the new model.

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

29

30 **1 Introduction**

31 The main distinguishing characteristic of spatial statistics compared to classical statistics is
32 that the former has a location attribute. Before geographical information systems were
33 developed, spatial statistical problems were often transformed into general statistical
34 problems, where the spatial coordinates were similar to a sample ID because they only had an
35 indexing feature. However, even in non-spatial statistics, the reversal or amalgamation
36 paradox (Pearson et al., 1899; Yule, 1903; Simpson, 1951), which is commonly called
37 Simpson's paradox (Blyth, 1972), has attracted significant attention from statisticians and
38 other researchers. In spatial statistics, some spatial variables exhibit certain trends and spatial
39 non-stationarity. Thus, it is possible for Simpson's paradox to occur when a classical
40 regression model is applied, and the existence of unknown important variables may worsen
41 this condition. The influence of Simpson's paradox can be fatal. For example, in geology, due
42 to the presence of cover and other factors that occur post-mineralization, ore-forming
43 elements in Area I are much lower than those in Area II, while the actual probability of a
44 mineral in Area I is higher than that in Area II simply because more deposits were discovered
45 in Area I (Agterberg, 1971). In this case, negative correlations would be obtained between
46 ore-forming elements and mineralization according to the classical regression model, whereas
47 high positive correlations can be obtained in both areas if they are separated. Simpson's
48 paradox is an extreme case of bias generated from classical models, and it is usually not so
49 severe in practice. However, this type of bias needs to be considered and care needs to be
50 taken when applying a classical regression model to a spatial problem. Several solutions to

51 this issue have been proposed, which can be divided into three types.

52 (1) Locations are introduced as direct or indirect independent variables. This type of
53 model is still a global model, but space coordinates or distance weights are employed to adjust
54 the regression estimation between the dependent variable and independent variables
55 (Agterberg, 1964; Agterberg and Cabilio, 1969; Agterberg, 1970; Agterberg and Kelly, 1971;
56 Agterberg, 1971; Casetti, 1972; Lesage & Pace, 2009, 2011). For example, Reddy et al. (1991)
57 performed logistic regression by including trend variables to map the base-metal potential in
58 the Snow Lake area, Manitoba, Canada; Helbich & Griffith (2016) compared the spatial
59 expansion method (SEM) to other methods in modeling the house price variation locally,
60 where the regression parameters are themselves functions of the x and y coordinates and their
61 combinations; [YuHao](#) & Liu (2016) used the spatial lag model (SLM) and spatial error model
62 to control spatial effects in modeling the relationship between PM_{2.5} concentrations and per
63 capita GDP in China.

64 (2) Local models are used to replace global models, i.e., geographically weighted models
65 (Fotheringham et al., 2002). Geographically weighted regression (GWR) is the most popular
66 model among the geographically weighted models. GWR models were first developed at the
67 end of the 20th century by Brunson et al. (1996) and Fotheringham et al. (1996, 1997, 2002)
68 for modeling spatially heterogeneous processes, and have been used widely in geosciences
69 (e.g., Buyantuyev & Wu, 2010; Barbet-Massin et al., 2012; Ma et al., 2014; Brauer et al.,
70 2015).

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

76 et al., 2016; Xiao et al., 2017).

77 GWR models can be readily visualized and are intuitive, which have made them applied
78 in geography and other disciplines that require spatial data analysis. In general, GWR is a
79 moving window-based model where instead of establishing a unique and global model for
80 prediction, it predicts each current location using the surrounding samples, and a higher
81 weight is given when the sample is located closer. The theoretical foundation of GWR is
82 ~~Tobler's~~Tobler's observation that: “everything is related to everything else, but near things are
83 more related than distant things” (Tobler, 1970).

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

93 Another problem with applying GWR 4.09 for MPM is that it cannot handle missing
94 data (Nakaya, 2016). Weights of evidence (WofE) is a widely used model for MPM
95 (Bonham-Carter et al., 1988, 1989; Agterberg, 1989; Agterberg et al., 1990) that mitigates the
96 effects of missing data. However, WofE was developed assuming that conditional
97 independence is satisfied among evidential layers with respect to the target layer; otherwise,
98 the posterior probabilities will be biased, and the number of estimated deposits will be
99 unequal to the known deposits. Agterberg (2011) combined WofE with logistic regression and
100 proposed a new model that can obtain an unbiased estimate of number of deposits while also

101 | avoiding the effect of missing data. In this study, we employed Agterberg (2011) [‘s solution](#)
102 | to account for missing data.

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

105 | Conceptually, this research originated from the thesis of Zhang (2015; in Chinese),
106 | which showed better efficiency for mapping intermediate and felsic igneous rocks (Zhang
107 | et al., 2017). This work elaborates on the principles of ILRBSWT, and provides a detailed
108 | algorithm for its design and implementation process with the code and software module
109 | attached. In addition, processing missing data is not a technique covered in GWR
110 | modeling presented in prior research, and a solution borrowed from WofE is provided in
111 | this study. Finally, ILRBSWT performance in MPM is tested by predicting Au ore
112 | deposits in western Meguma Terrain, Nova Scotia, Canada.

114 | **2 Models**

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

119 | *2.1 Logistic Regression*

120 | In MPM, the dependent variable(Y) is binary because Y can only take the value of 1 and 0,
121 | indicating that mineralization occurs and not respectively. Suppose that π represents the
122 | estimation of Y , $0 \leq \pi \leq 1$, then a logit transformation of π can be made, i.e., $\text{logit}(\pi)$
123 | $= \ln(\pi/(1-\pi))$. The logistic regression function can be obtained as follows:

$$124 | \text{Logit } \pi(X_1, X_2, \dots, X_p) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (1)$$

125 | where X_1, X_2, \dots, X_p , comprises a sample of p explanatory variables x_1, x_2, \dots, x_p , β_0 is the

126 intercept, and $\beta_1, \beta_2, \dots, \beta_p$ are regression coefficients.

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

$$131 \quad L(\beta) = \prod_{i=1}^n (\pi_i^{Y_i} (1 - \pi_i)^{1-Y_i}), \quad (2)$$

132 where $\pi_i = \pi(X_{i1}, X_{i2}, \dots, X_{ip}) = \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}$. The best estimate can be obtained

133 only if equation (2) takes the maximum. Then the problem is converted into solving
 134 $\beta_1, \beta_2, \dots, \beta_p$. Equation (2) can be further transformed into the following log-likelihood
 135 function:

$$136 \quad \ln L(\beta) = \sum_{i=1}^n (Y_i \pi_i + (1 - Y_i)(1 - \pi_i)) \quad (3)$$

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

$$139 \quad \begin{cases} f(\beta_0) = \sum_{i=0}^n (Y_i - \pi_i) X_{i0} = 0 \\ f(\beta_1) = \sum_{i=0}^n (Y_i - \pi_i) X_{i1} = 0 \\ \vdots \\ f(\beta_p) = \sum_{i=0}^n (Y_i - \pi_i) X_{ip} = 0 \end{cases} \quad (4)$$

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

$$142 \quad \mathbf{X}^T(\mathbf{Y} - \boldsymbol{\pi}) = \mathbf{0} \quad (5)$$

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

$$144 \quad \hat{\boldsymbol{\beta}}(t+1) = \hat{\boldsymbol{\beta}}(t) + \mathbf{H}^{-1} \mathbf{U}, \quad (6)$$

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

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

149 For a more detailed description of the derivations of equations (1) to (6), see Hosmer et al.
150 (2013).

151 2.2 Weighted Logistic Regression

152 In practice, vector data is often used, and sample size (area) has to be considered. In this
153 condition, weighted logistic regression modeling should be used instead of a general logistic
154 regression. It is also preferable to use a weighted logistic regression model when a logical
155 regression should be performed for large sample data because weighted logical regression can
156 significantly reduce matrix size and improve computational efficiency (Agterberg, 1992).
157 Assuming that there are four binary explanatory variable layers and the study area consists of
158 1000×1000 grid points, the matrix size for normal logic regression modeling would be
159 $10^6 \times 10^6$; however, if weighted logistic regression is used, the matrix size would be 32×32 at
160 most. This condition arises because the sample classification process is contained in the
161 weighted logistic regression, and all samples are classified into classes with the same values
162 as the dependent and independent variables. The samples with the same dependent and
163 independent variables form certain continuous and discontinuous patterns in space, which are
164 called “unique condition” units. Each unique condition unit is then treated as a sample, and
165 the area (grid number) for it is taken as weight in the weighed logistic regression. Thus, for
166 the weighted logical regression, equations (2) to (5) in section 2.1 need to be changed to
167 equations (7) to (10) as follows.

168

169 $L_{new}(\beta) = \prod_{i=1}^n (\pi_i^{N_i Y_i} (1 - \pi_i)^{N_i (1 - Y_i)}),$ (7)

170 $\ln L_{new}(\beta) = \sum_{i=1}^n (N_i Y_i \pi_i + N_i (1 - Y_i) (1 - \pi_i))$ (8)

171
$$\begin{cases} f_{new}(\beta_0) = \sum_{i=0}^n (Y_i - \pi_i) X_{i0} = 0 \\ f_{new}(\beta_1) = \sum_{i=0}^n (Y_i - \pi_i) X_{i1} = 0 \\ \vdots \\ f_{new}(\beta_p) = \sum_{i=0}^n (Y_i - \pi_i) X_{ip} = 0 \end{cases}$$
 (9)

172 $\mathbf{X}^T \mathbf{W} (\mathbf{Y} - \boldsymbol{\pi}) = \mathbf{0}$ (10)

173 where N_i is the weight for the i -th unique condition unit, i takes the value from 1 to n , and n
 174 is the number of unique condition units. \mathbf{W} is a diagonal matrix that is expressed as follows:

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

175 In addition, new values of \mathbf{H} and \mathbf{U} should be used in equation (6) to perform Newton
 176 iteration as part of the weighted logistic regression, i.e., $\mathbf{H}_{new} = \mathbf{X}^T \mathbf{W} \mathbf{V}(t) \mathbf{X}$, $\mathbf{U}_{new} =$
 177 $\mathbf{X}^T \mathbf{W} (\mathbf{Y} - \boldsymbol{\pi}(t))$.

178 2.3 Geographically Weighted Logistic Regression

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

185 $\text{Logit } \pi(\mathbf{x}, \mathbf{u}) = \beta_0(\mathbf{u}) + \beta_1(\mathbf{u})x_1 + \beta_2(\mathbf{u})x_2 + \dots + \beta_p(\mathbf{u})x_p,$ (11)

186 where, $\beta_0(\mathbf{u}), \beta_1(\mathbf{u}), \dots, \beta_p(\mathbf{u})$ indicate that these parameters are obtained at the location
 187 of \mathbf{u} . $\text{Logit } \pi(\mathbf{x}, \mathbf{u})$, the predicted probability for the current location \mathbf{u} , can be obtained under
 188 the condition that the values of all independent variables are known at the current location and
 189 all parameters are also calculated based on the samples within the current local window.

190 According to equation (6) in section 2.1, the parameters for GWLR can be estimated with
191 equation (12):

$$192 \hat{\boldsymbol{\beta}}(\mathbf{u})_{t+1} = \hat{\boldsymbol{\beta}}(\mathbf{u})_t + (\mathbf{X}^T \mathbf{W}(\mathbf{u}) \mathbf{V}(t) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(\mathbf{u}) (\mathbf{Y} - \boldsymbol{\pi}(t)), \quad (12)$$

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

199 *2.4 Improved Logistic Regression Model based on the Spatially Weighted Technique*

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

203 *2.4.1 Integration of Different Weights*

204 If a diagonal element in $\mathbf{W}(\mathbf{u})$ is only for one sample, i.e., the grid point in raster data, section
205 2.3 is an improvement on section 2.1, i.e. samples are weighted according to their location. If
206 samples are first reclassified according to the unique condition mentioned in section 2.2, and
207 corresponding weights are then summarized according to each [sample's sample's](#) geographical
208 weight, we can obtain an improved logistic regression model considering both sample size
209 and geographical distance. The new model both reflects the spatial distribution of samples and
210 reduces the matrix size, which is discussed in the following section.

211 In addition to geographic factors, representativeness of a sample, e.g., differences in the
212 level of exploration, is also considered in this study.

213 Suppose that there are n grid points in the current local window, S_i is the i -th grid, $W_i(g)$
214 is the geographical weight of S_i , and $W_i(d)$ represents the individual difference weight or

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

$$220 \quad W_j(t) = \sum_{i=1}^n [W_i(g) * W_i(d) * df_i], \quad (13)$$

221 where $\begin{cases} df_i = 1 & \text{if } S_i \in C_j \\ df_i = 0 & \text{if } S_i \notin C_j \end{cases}$, i takes a value from 1 to n , j takes a value from 1 to N , and $W_j(t)$

222 represents the total weight (by combining both $W_i(g)$ and $W_i(d)$) for each unique condition
 223 unit. We can use the final weight calculated in equation (13) to replace the original weight in
 224 equation (12), which is an advantage of ILRBSWT.

225 2.4.2 Missing data processing

226 Missing data is a problem in all statistics-related research fields. In MPM, missing data are
 227 also prevalent due to ground coverage, and limitations of exploration technique and
 228 measurement accuracy. Agterberg and Bonham-Carter (1999) compared the following
 229 commonly used missing data processing solutions: (1) removing variables containing missing
 230 data, (2) deleting samples with missing data, (3) using 0 to replace missing data, and (4)
 231 replacing missing data with the mean of the corresponding variable. To efficiently use
 232 existing data, both (1) and (2) are clearly not good solutions as more data will be lost.
 233 Solution (3) is superior to (4) in the condition that work has not been done and real data is
 234 unknown; with respect to missing data caused by detection limits, solution (4) is clearly a
 235 better choice. Missing data is primarily caused by the latter in MPM, and Agterberg (2011)
 236 pointed out that missing data was better addressed in a WofE model. In WofE, the evidential
 237 variable takes the value of positive weight (W^+) if it is favorable for the target variable (e.g.,
 238 mineralization), while the evidential variable takes the value of negative weight (W^-) if it is

239 unfavorable for the target variable; and automatically the evidential variable takes the value of
 240 “0” if there is missing data.

$$241 \quad W^+ = \ln \frac{\frac{D_1}{D}}{\frac{A_1 - D_1}{A - D}} \quad (14)$$

$$242 \quad W^- = \ln \frac{\frac{D_2}{D}}{\frac{A_2 - D_2}{A - D}} \quad (15)$$

243 where \mathbf{A} is an evidential layer, A_1 is the area (or grid number, similarly hereinafter) that A
 244 takes the value of 1, and A_2 is the area that A takes the value of 0; A_3 is the area with missing
 245 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
 246 areas where the target variables take the value of 1 in A_1 , A_2 , and A_3 respectively. A_3 and D_3
 247 are not used in equation (15) because no information is provided in area A_3 .

248 If “1” and “0” are still used to represent the binary condition of the independent variable
 249 instead of W^+ and W^- , equation (16) can be used to replace missing data in logistic
 250 regression modeling.

$$251 \quad M = \frac{-W^-}{W^+ - W^-} = \frac{\ln \frac{D}{A - D} - \ln \frac{D_2}{A_2 - D_2}}{\ln \frac{D_1}{A_1 - D_1} - \ln \frac{D_2}{A_2 - D_2}} \quad (16)$$

252

253 **3 Design of the ILRBSWT Algorithm**

254 *3.1 Local Window Design*

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

259 (1) Circular Local Window Design

260 Suppose that W represents a local circular window where the minimum bounding
 261 rectangle is R , then the geographical weights can be calculated only inside R . Clearly, the grid

262 points inside R but outside of W should be weighted as 0, and the weight for the grid with a
 263 center inside W should be calculated according to the distance from its current location.
 264 Because R is a square, we can also assume that there are n columns and rows in it, where n is
 265 an odd number. If we take east and south as the orientations of the x -axis and y -axis,
 266 respectively, and the position of the northwest corner grid is defined as $(x = 1, y = 1)$, then a
 267 local rectangular coordinate system can be established and the position of the current location
 268 grid can be expressed as $O (x = \frac{n+1}{2}, y = \frac{n+1}{2})$. The distance between any grid inside W and
 269 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
 270 j take values ranging from 1 to n . The geographical weight is a function of distance, so it is
 271 convenient to calculate w_{ij} with d_{o-ij} . Figure 1 shows the weight template for a circular
 272 local window with a half-window size of nine grids.

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

273

274

275

276

277

278

279

280

281

282

283

284

Fig. 1 Weight template for a circular local window with a half-window size of nine grids, where w1 to w30 represent different weight classes that decrease with distances and 0 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.

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 $T_j + \frac{n-1}{2}$ from the total research area. Next, we can establish a local rectangular coordinate system according to the previously described steps; we subtract $T_i - \frac{n-1}{2}$ and $T_j - \frac{n-1}{2}$ on the x and y coordinates respectively for all grids in the range. The corresponding relationship can then be established

285 between the weight template and current window. Global weights can also be included via the
 286 matrix product between the global weight layer and local weight template within the local
 287 window. In addition, special care should be taken when the weight template covers some area
 288 outside the study area, i.e., $T_i - \frac{n-1}{2} < 0$, $T_i + \frac{n-1}{2} > T_n$, $T_j - \frac{n-1}{2} < 0$, and $T_j +$
 289 $\frac{n-1}{2} > T_m$.

290 (2) Elliptic Local Window Design

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

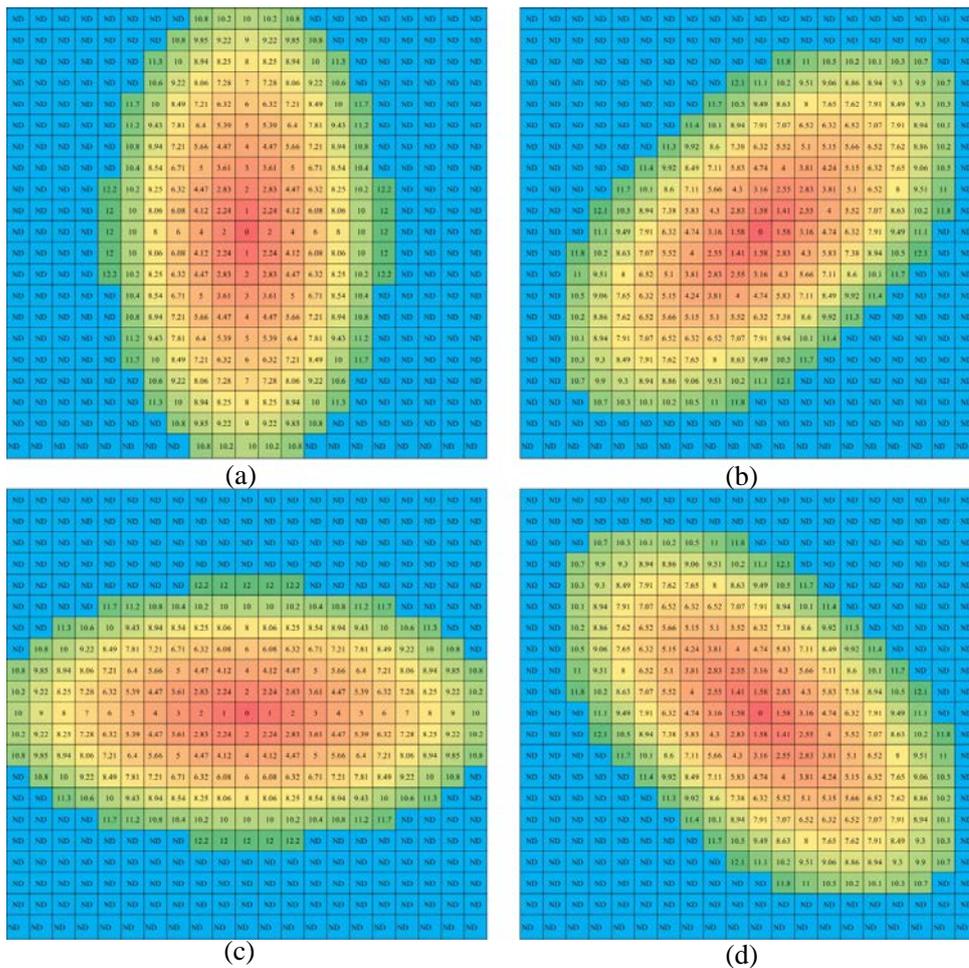
298 We still use W to represent the local elliptic window and a , r , and θ are defined as the
 299 semi-major axis, ratio of the semi-minor axis relative to the semi-major axis, and azimuth of
 300 the semi-major axis, respectively. Then, W can be covered by a square R whose side length is
 301 $2a-1$ and center is the same as W . There are $(2a-1) \times (2a-1)$ grids in R . We establish the
 302 rectangular coordinates as described above and suppose that the center of the top left grid in R
 303 is located at $(x = 1, y = 1)$, and thus the center of W should be $O(x_0 = a, y_0 = a)$. According

304 to the definition of the ellipse, two of the elliptical foci are located at $F_1(x_1 = a +$
 305 $\sin(\theta) \sqrt{a^2 - (a * r)^2}, y_1 = a - \cos(\theta) \sqrt{a^2 - (a * r)^2})$ and

306 $F_2(x_2 = a - \sin(\theta) \sqrt{a^2 - (a * r)^2}, y_2 = a + \cos(\theta) \sqrt{a^2 - (a * r)^2})$. The summed
 307 distances between a point and two focus points can be expressed as

308 $l_{ij} = \sqrt{(i - x_1)^2 + (j - y_1)^2} + \sqrt{(i - x_2)^2 + (j - y_2)^2}$, where i and j take values from 1 to

309 $2a - 1$. According to the elliptical focus equation, for any grid in R , if the sum of the distances
310 between the two focal points and a grid center is greater than $2a$, the grid is located within W ,
311 and vice versa. For the grids outside of W , the weight is assigned as 0, and we only need to
312 calculate the equivalent distances for the grids within W . As mentioned above, the parameters
313 for the ellipse can be determined using the kriging method. In ellipse W , where the semi-major
314 axis is a , r and θ are maintained as constants, then we obtain countless ellipses centered at
315 the center of W , and the equivalent distance is the same on the same elliptical orbit. Thus, the
316 equivalent distance template can be obtained for the local elliptic window. Figure 2 shows the
317 equivalent distance templates under the conditions that $a = 11$ grids, $r = 0.5$, and the azimuths
318 for the semi-major axis are 0° , 45° , 90° , and 135° respectively.
319



320
321 **Fig. 2 Construction of the distance template based on an elliptic local window: $a = 11$ grid points,**

322 $r = 0.5$, and the azimuths for the semi-major axis are 0° (a), 45° (b), 90° (c), and 135° (d) respectively.

323 3.2 Algorithm for ILRBSWT

324 The ILRBSWT method primarily focuses on two problems, i.e., spatial non-stationarity and
325 missing data. We use the moving window technique to establish local models instead of a
326 global model to overcome spatial non-stationarity. The spatial t -value employed in the WofE
327 method is used to binarize spatial variables based on the local window, which is quite
328 different from traditional binarization based on the global range, where missing data can be
329 handled well because positive and negative weights are used instead of the original values of
330 “1” and “0”, and missing data are represented as “0.” Both the isotropy and anisotropy
331 window types are provided in our new proposed model. The geographical weight function and
332 window size can be determined by the users. If the geographic weights are equal and there are
333 no missing data, ILRBSWT will yield the same posterior probabilities as classical logistic
334 regression; hence, the later can be viewed as a special case of the former. The core ILRBSWT
335 algorithm is as follows.

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

346 Step 2. Establish a loop for all independent variables. In a circular (elliptical) window

347 with a radius (equivalent radius) of r_{\min} , apply the WofE (Agterberg, 1992) model according
348 to the grid weight determined in step 1, thereby obtaining a statistical table containing the
349 parameters W_{ij}^+ , W_{ij}^- , and t_{ij} , where i is the i -th independent variable and j denotes the j -th
350 binarization.

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

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

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

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

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

363 Step 3.3. Apply the WofE model according to the grid weight determined in step 1 in the
364 current local window. If a maximum t_{ij} exists and it is greater than or equal to the standard
365 t -value, record the values of $W_{i-\max_t}^+$, $W_{i-\max_t}^-$, $B_{i-\max_t}$, and r_{current} , which represent the
366 radius (equivalent radius) for the current local window.

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

369 Step 4. Suppose that n_s independent variables still remain.

370 Step 4.1. If $n_s \leq 1$, calculate the mean value for the dependent variable in the current
371 local window with a radius size of r_{\max} and retain it as the posterior probability in the current

372 location. In addition, set the regression coefficients for all independent variables as missing
373 data. Go to step 6.

374 Step 4.2. If $n_s \geq 1$, find the independent variable with the largest local window and
375 apply the WofE model to all other independent variables, and then update the values of
376 $W_{i-\max_t}^+$, $W_{i-\max_t}^-$, and $B_{i-\max_t}$. Go to step 5.

377 Step 5. Apply the logistic regression model based on the previously determined
378 geographic weights, and for each independent variable: (1) use $W_{i-\max_t}^+$ to replace all
379 values that are less than or equal to $B_{i-\max_t}$, (2) use $W_{i-\max_t}^-$ to replace all values that are
380 greater than $B_{i-\max_t}$, and (3) use 0 to replace no data (“-9999”). The posterior probability
381 and regression coefficients can then be obtained for all independent variables at the current
382 location and go to step 6.

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

384

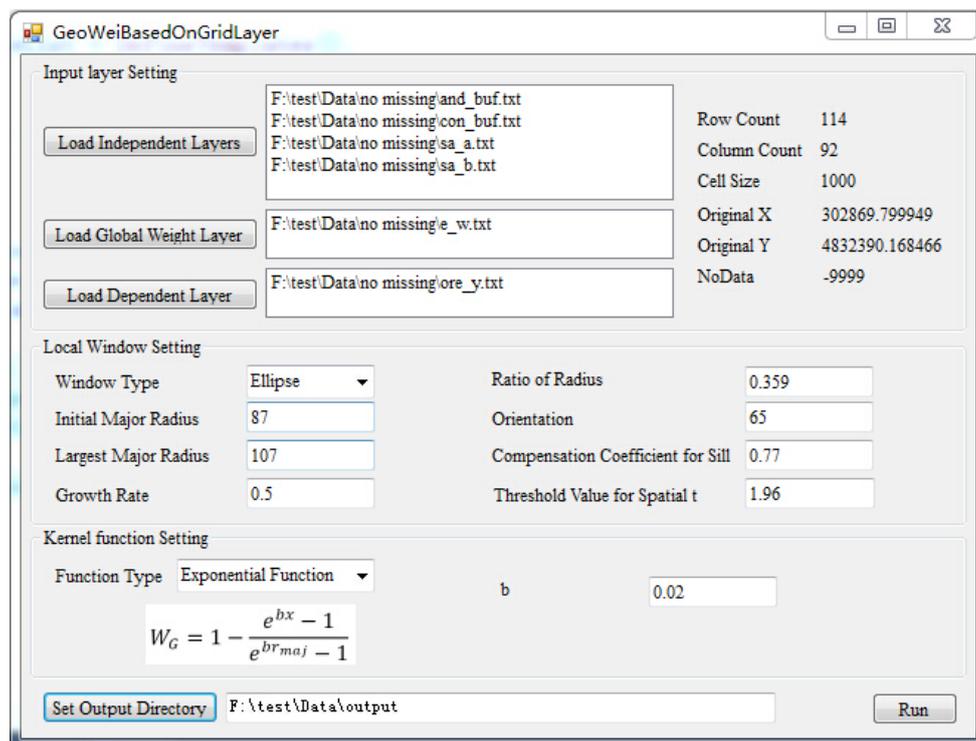
385 **4 Interface Design**

386 Before performing spatially weighted logistical regression with ILRBSWT 1.0, data
387 pre-processing is performed using the ArcGIS 10.2 platform and GeoDAS 4.0 software. All
388 data are originally stored in grid format, which should be transformed into ASCII files with
389 the Arc toolbox in ArcGIS 10.2; after modeling with ILRBSWT 1.0, the result data will be
390 transformed back into grid format

391 As shown in Fig. 3, the main interface for ILRBSWT 1.0 is composed of four parts.

392 The upper left part is for the layer input settings, where independent variable layers,
393 dependent variable layers, and global weight layers should be assigned. Layer information is
394 shown at the upper right corner, including row numbers, column numbers, grid size, ordinate
395 origin, and the expression for missing data. The local window parameters and weight
396 attenuation function can be defined as follows. Using the drop-down list, we prepared a circle

397 or ellipse to represent various isotropic and anisotropic spatial conditions, respectively. The
 398 corresponding window parameters should be set for each window type. For the ellipse, it is
 399 necessary to set parameters composed of the initial length of the equivalent radius (initial
 400 major radius), final length of the equivalent radius (largest major radius), increase in the
 401 length of the equivalent radius (growth rate), threshold of the spatial t -value used to determine
 402 the need to enlarge the window, length ratio of the major and minor axes, orientation of the
 403 [ellipse's ellipse's](#) major axis, and compensation coefficient for the sill. We prepared different
 404 types of weight attenuation functions via the drop-down menu to provide choices to users,
 405 such as exponential model, logarithmic model, Gaussian model, and spherical model, and
 406 corresponding parameters can be set when a certain model is selected. The output file is
 407 defined at the bottom and the execution button is at the lower right corner.



408
409 **Fig. 3 User interface design.**
410
411
412

413 **5 Real Data Testing**

414 *5.1 Data source and preprocessing*

415 The test data used in this study were obtained from the case study reported in Cheng (2008).

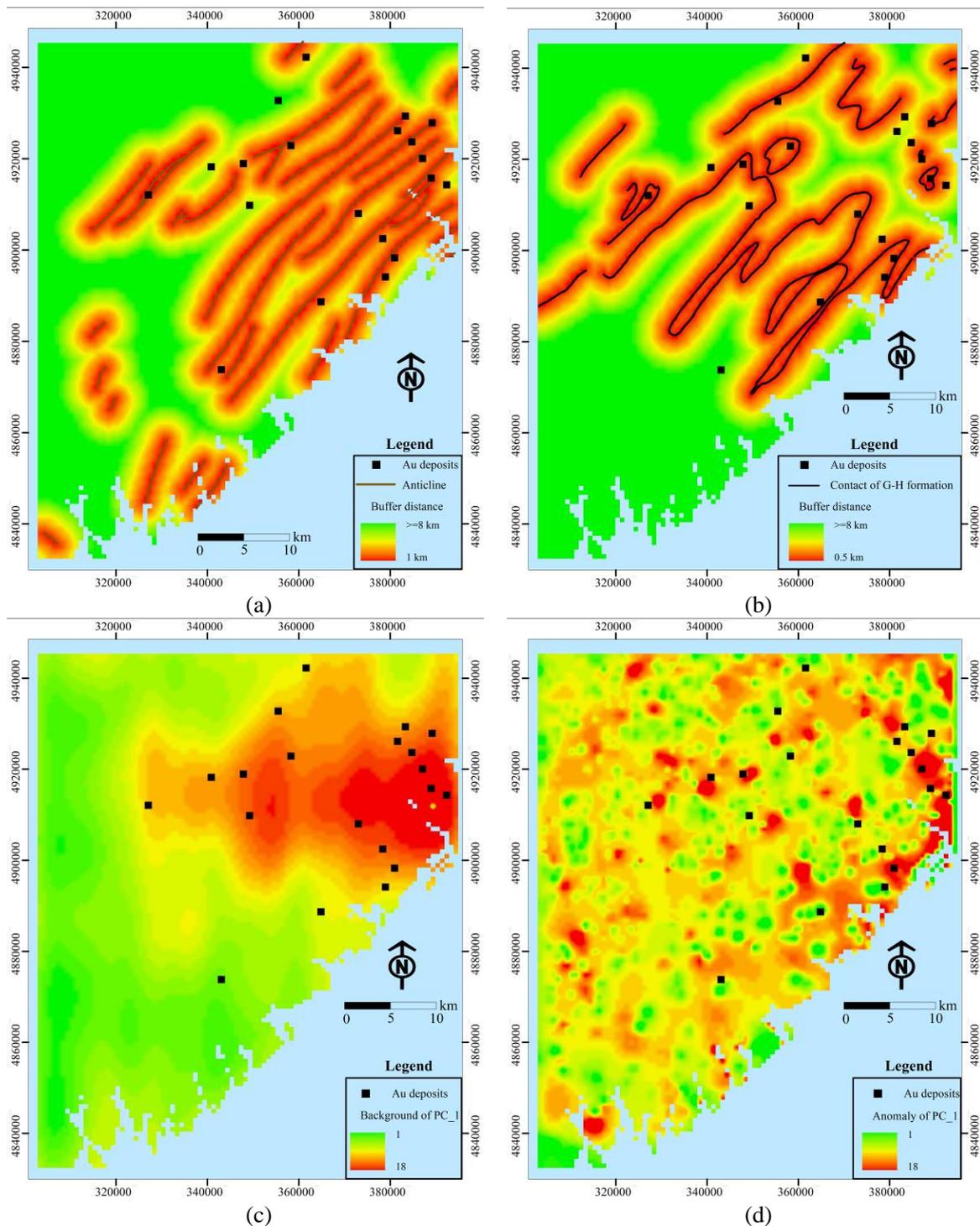
416 The study area ($\approx 7780 \text{ km}^2$) is located in western Meguma Terrain, Nova Scotia, Canada.

417 Four independent variables were used in the WofE model for gold mineral potential mapping

418 by Cheng (2008), i.e., buffer of anticline axes, buffer for the contact of Goldenville–Halifax

419 Formation, and background and anomaly separated with the S-A filtering method based on

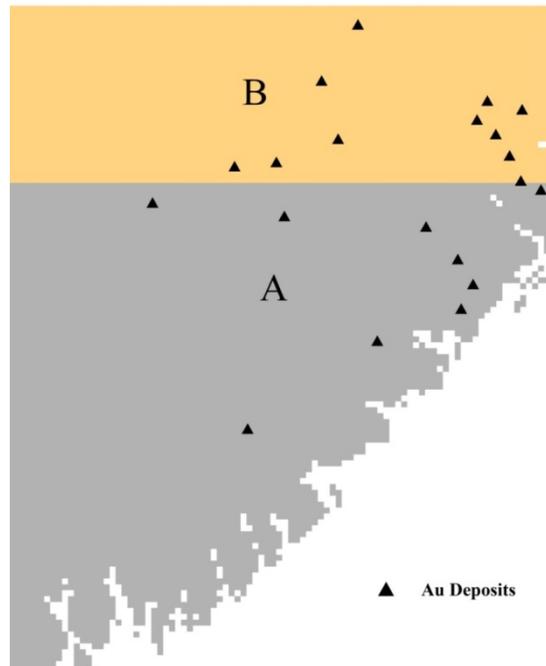
420 ore element loadings of the first component, as shown in Fig. 4.



421
 422 **Fig. 4 Evidential layers used to map Au deposits in this study: buffer of anticline axes (a), buffer for**
 423 **the contact of Goldenville–Halifax Formation (b), and background (c) and anomaly (d) separated**
 424 **with the S-A filtering method based on the ore element loadings of the first component.**

425 The four independent variables described previously were also used for ILRBSWT
 426 modeling in this study (see Figs. 4 (a) to (d)), and they were uniformed in the ArcGIS grid
 427 format with a cell size of 1 km × 1 km. To demonstrate the advantages of the new method for

428 missing data processing, we designed an artificial situation in Fig. 5, i.e., grids in region A
429 have values for all four independent variables, while they only have values for two
430 independent variables and no values in the two geochemical variables in region B.



431
432 **Fig. 5 Study area (A and B) where there is missing geochemical data in area B.**

433 *5.2 Mapping weights for exploration*

434 Exploration level weights can be determined based on prior knowledge about data quality, e.g.,
435 different scales may exist throughout the whole study area; however, these weights can also
436 be calculated quantitatively. The density of known deposits is a good index for the exploration
437 level, i.e., the research is more comprehensive when more deposits are discovered. The
438 exploration level weight layer for the study area was obtained using the kernel density tool
439 provided by the ArcToolbox in ArcGIS 10.2, as shown in Fig. 6.

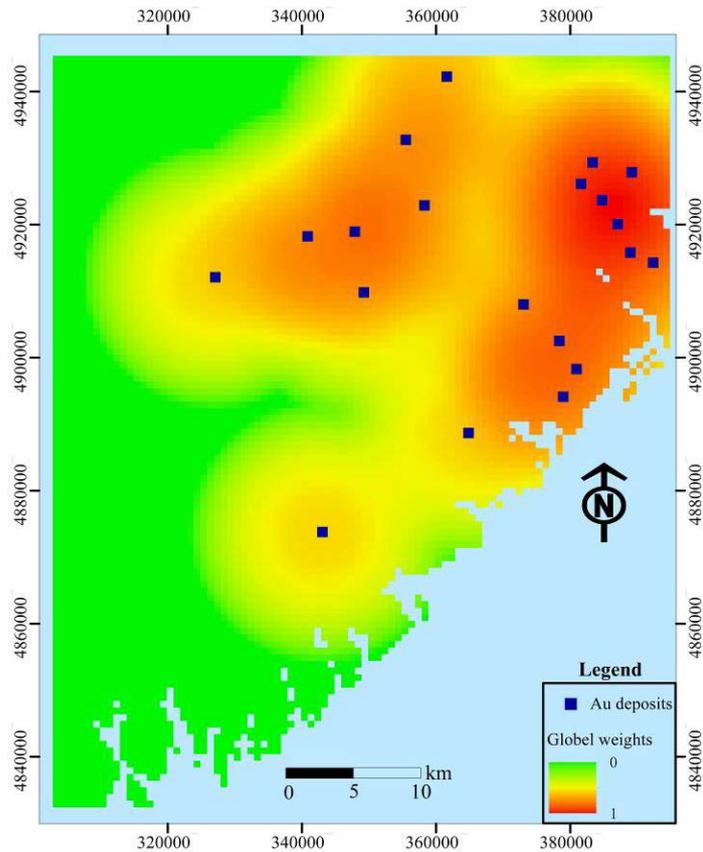
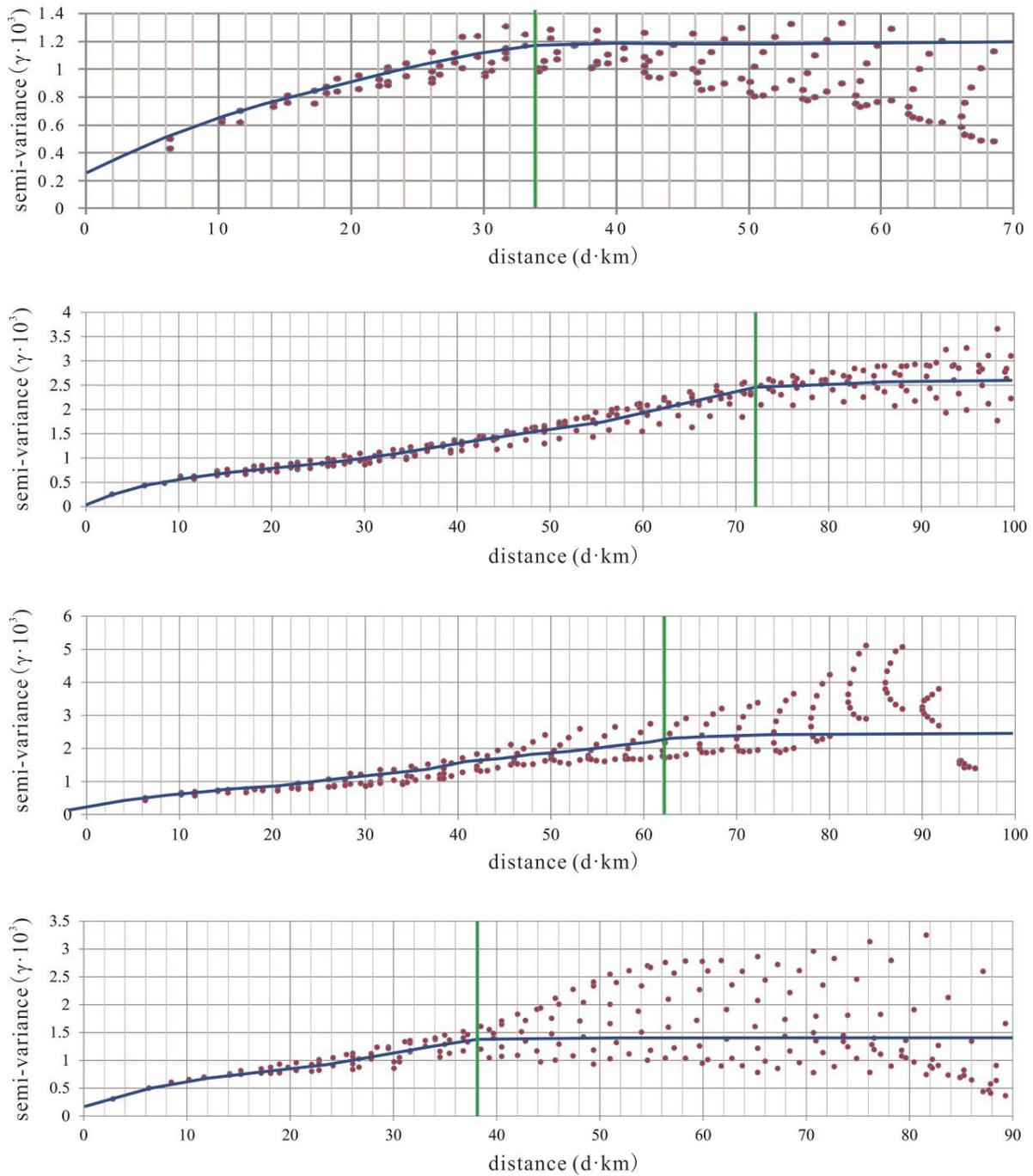


Fig. 6 Exploration level weights.

5.3 Parameter Assignment for local window and weight attenuation function

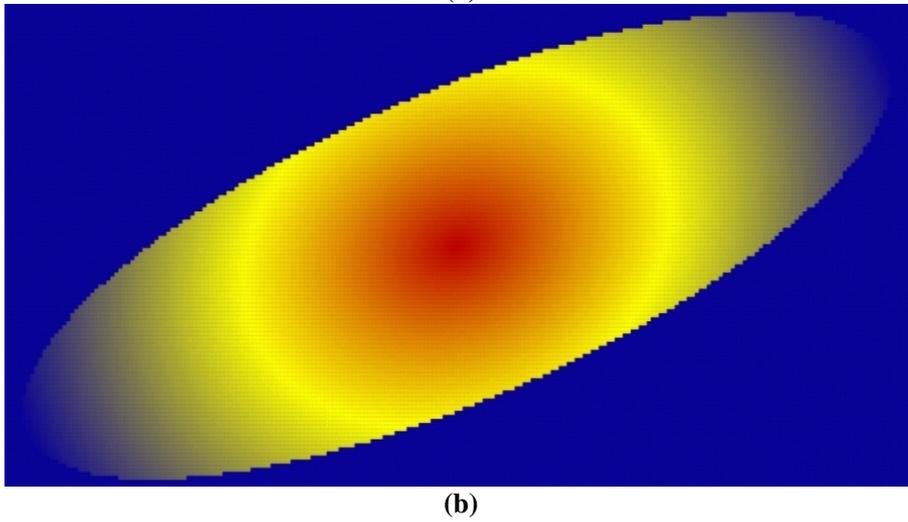
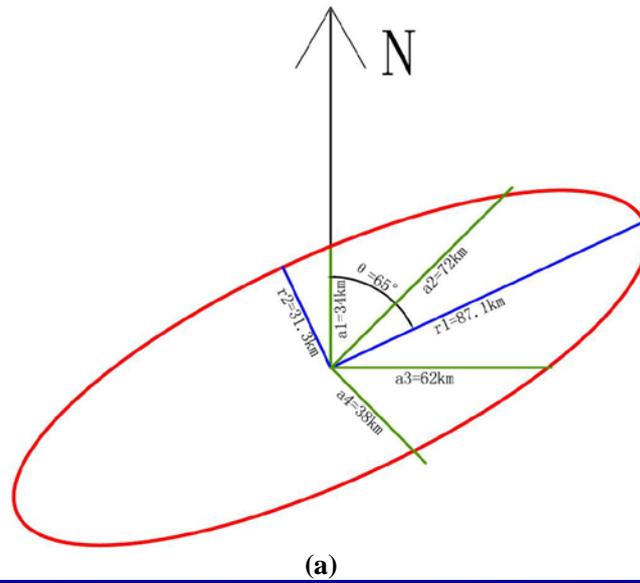
Both empirical and quantitative methods can be used to determine the local window parameters and attenuation function for geographical weights. The variation function in geostatistics, which is an effective method for describing the structures and trends in spatial variables, was applied in this study. To calculate the variation function for the dependent variable, it is necessary to first map the posterior probability using the global logistic regression method before determining the local window type and parameters from the variation function. Variation functions were established in four directions 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. 7, and the final local window and

454 geographical weight attenuation were determined as indicated in Fig. 8 (a) and 8 (b),
455 respectively.



456
457 **Fig. 7 Experimental variogram fitting in different directions, where the green lines denote the**
458 **variable ranges determined for azimuths of (a) 0°, (b) 45°, (c) 90°, and (d) 135°.**

459

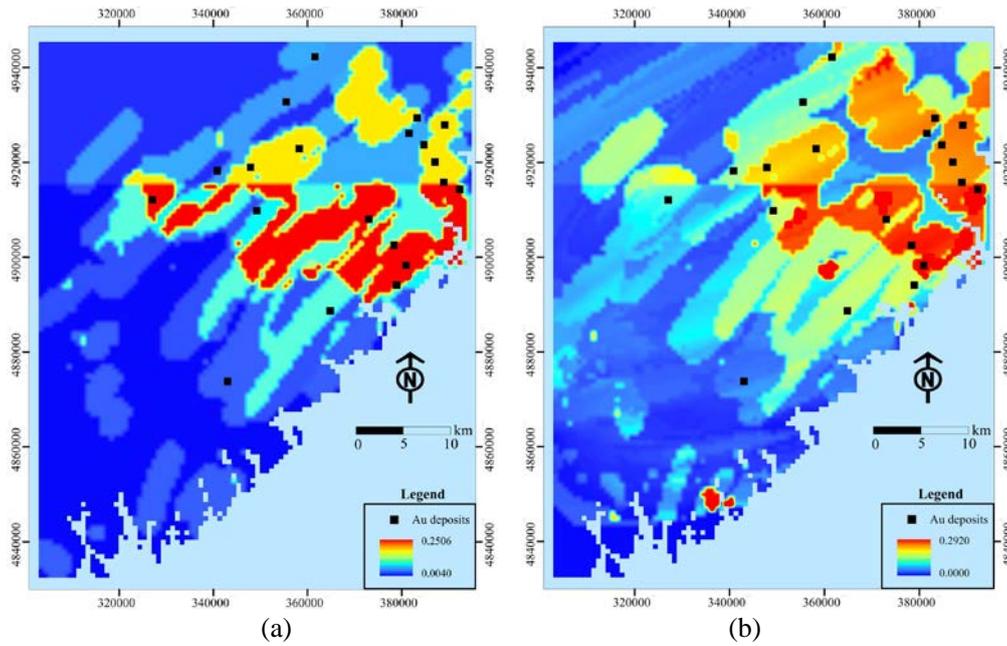


460

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

463 *5.4 Data integration*

464 Using the algorithm described in section 3.2, ILRBSWT was applied to the study area
 465 according to the parameter settings in Fig. 3. The estimated probability map obtained for Au
 466 deposits by ILRBSWT is shown in Fig. 9 (b), while Fig. 9 (a) presents the results obtained by
 467 logistic regression. As shown in Fig. 8, ILRBSWT better manages missing data than logistic
 468 regression, as the Au deposits in the north part of the study area (with missing data) better fit
 469 within the region with higher posterior probability in Fig. 9 (b) than in Fig. 9 (a).



470

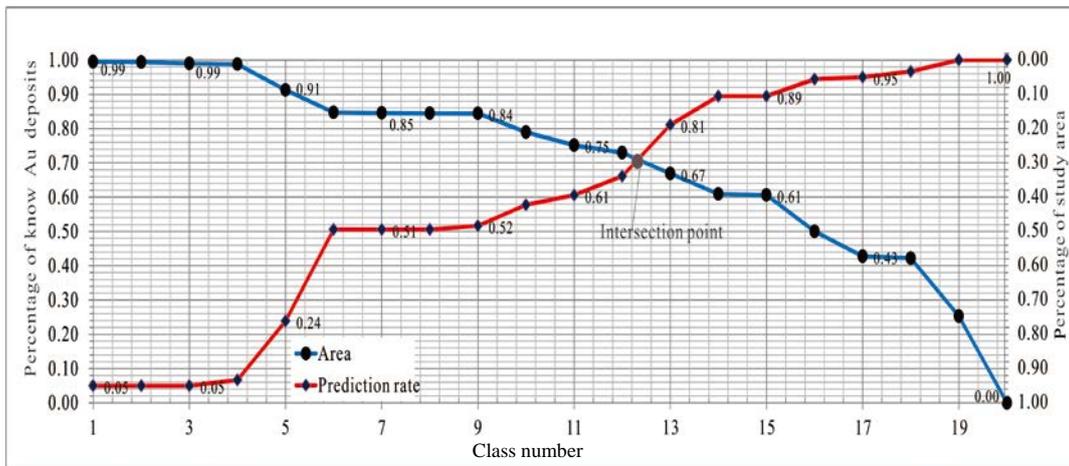
471 **Fig. 9** Posterior probability maps obtained for Au deposits by (a) logistic regression and (b)
 472 **ILRBSWT.**

473 *5.5 Comparison of the mapping results*

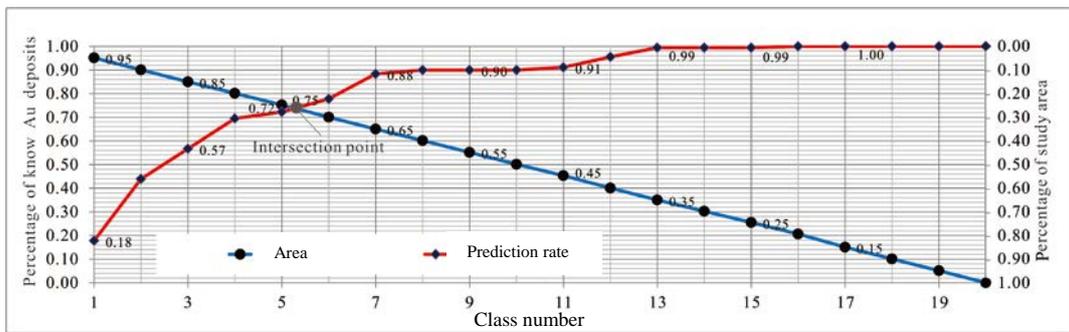
474 To evaluate the predictive capacity of the newly developed and traditional methods, the
 475 posterior probability maps obtained through logistic regression and ILRBSWT shown in Fig.
 476 9 (a) and 9 (b) were divided into 20 classes using the quantile method. Prediction-area (P-A)
 477 plots (Mihalasky & Bonham-Carter, 2001; Yousefi et al., 2012; Yousefi & Carranza, 2015a)
 478 were then made according to the spatial overlay relationships between Au deposits and the
 479 two classified posterior probability maps in Fig. 10 (a) and (b) respectively. In a P-A plot, the
 480 horizontal ordinate indicates the discretized classes of a map representing the occurrence of
 481 deposits. The vertical scales on the left and right sides indicate the percentage of correctly
 482 predicted deposits from the total known mineral occurrences and the corresponding
 483 percentage of the delineated target area from the total study area (Yousefi & Carranza, 2015a).
 484 As shown in Figs. 10 (a) and (b), with the decline of the posterior probability threshold for the
 485 mineral occurrence from left to right on the horizontal axis, more known deposits are
 486 correctly predicted, and meantime more areas are delimited as the target area; however, the

487 growth in the prediction rates for deposits and corresponding occupied area are similar before
488 the intersection point in Fig. 10 (a), while the former shows higher growth rate than the latter
489 in Fig. 10 (b). This difference suggests that ILRBSWT can predict more known Au deposits
490 than logistic regression for delineating targets with the same area, and indicates that the
491 former has a higher prediction efficiency than the latter.

492 It would be a little inconvenient to consider the ratios of both predicted known deposits
493 and occupied area. Mihalasky and Bonham-Carter (2001) proposed a normalized density, i.e.
494 the ratio of the predicted rate of known deposits to its corresponding occupied area. The
495 intersection point in a P-A plot is the crossing of two curves. The first is fitted from scatter
496 plots of the class number of the posterior probability map and rate of predicted deposit
497 occurrences (the “Prediction rate” curves in Fig. 10). The second is fitted according to the
498 class number of the posterior probability map and corresponding accumulated area rate (the
499 “Area” curves in Fig. 10). At the interaction point, the sum of the prediction rate and
500 corresponding occupied area rate is 1; the normalized density at this point is more commonly
501 used to evaluate the performance of a certain spatial variable in indicating the occurrence of
502 ore deposits (Yousefi & Carranza, 2015a). The intersection point parameters for both models
503 are given in Table 1. As shown in the table, 71% of the known deposits are correctly predicted
504 with 29% of the total study area delineated as target area when the logistic regression is
505 applied; if ILRBSWT is applied, 74% of the known deposits can be correctly predicted with
506 only 26% of the total area delineated as the target area. The normalized densities for the
507 posterior probability maps obtained from the logistic regression and ILRBSWT are 2.45 and
508 2.85 respectively; the latter performed significantly better than the former.



(a)



(b)

509

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

512 **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

513

514 **6 Discussion**

515 Because of potential spatial heterogeneity, the model parameter estimates obtained based
 516 on the total equal-weight samples in the classical regression model may be biased, and they
 517 may not be applicable for predicting each local region. Therefore, it is necessary to adopt a
 518 local window model to overcome this issue. The presented case study shows that ILRBSWT

519 can obtain better prediction results than classical logistic regression because of the
520 ~~former's~~former's sliding local window model, and their corresponding intersection point
521 values are 2.85 and 2.45, respectively. However, ILRBSWT has even advantages. For
522 example, characterizing 26% or 29% of the total study area as promising prospecting targets
523 is too high in terms of economic considerations. If instead 10% of the total area needs is
524 mapped as the target area, the proportions of correctly predicted known deposits obtained by
525 ILRBSWT and logistic regression are 44% and 24%, respectively. The prediction efficiency
526 of the former is 1.8 times larger than the latter.

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

539 Similar to WofE and logistic regression, ILRBSWT is a data-driven method, thus it
540 inevitably suffers the same problems as data-driven methods, e.g., the information loss caused
541 by data discretization, and exploration bias caused by the training sample location. However,
542 it should be noted that evidential layers are discretized in each local window instead of the
543 total study area, which may cause less information loss. This can also be regarded as an

544 advantage of the ILRBSWT tool. With respect to logistic regression and WofE, some
545 researchers have proposed solutions to avoid information loss resulting from spatial data
546 discretization by performing continuous weighting (Pu et al., 2008; Yousefi & Carranza,
547 2015b, 2015c, 2016), and these concepts can be incorporated into further improvements of the
548 ILRBSWT tool in the future.

549

550 **7 Conclusions**

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

556

557 *Code availability*

558 The software tool ILRBSWT v1.0 in this research was developed using C#, and the
559 [main source codes](#) and [key function executable programs \(software tool\)](#) are prepared in the
560 [file “Codes & Key Functions”](#). ~~The executable program files are placed~~ [folders “source code](#)
561 [for ILRBSWT in the folder C#” and “Executable Programs for ILRBSWT”](#) respectively.
562 Please find them in [gmd-2017-278-supplement.zip](#).

563

564 *Data availability*

565 The data used in this research is sourced from the demo data for GeoDAS software
566 (<http://www.yorku.ca/yul/gazette/past/archive/2002/030602/current.htm>), which was also
567 used by Cheng (2008). All spatial layers used in this work are included in the folder “Original
568 Data” in the format of an ASCII file, which is also found in [gmd-2017-278-supplement.zip](#).

569

570 **Acknowledgments**

571 This study benefited from joint financial support from the Programs of National Natural
572 Science Foundation of China (Nos. 41602336 and 71503200), China Postdoctoral Science
573 Foundation (Nos. 2017T100773 and 2016M592840), Shaanxi Provincial Natural Science
574 Foundation (No. 2017JQ7010), and Fundamental Research from Northwest A&F University
575 in 2017 (No. 2017RWYB08). The first author thanks former supervisor Drs. Qiuming Cheng
576 and Frits Agterberg for fruitful discussions of spatial weights and providing constructive
577 suggestions. [Great thanks also to the anonymous referees for their helpful suggestions and](#)
578 [corrections.](#)

579

580 **References**

- 581 Agterberg, F.P., & Cabilio, P., 1969. Two-stage least-squares model for the relationship between mappable geological
582 variables. *Journal of the International Association for Mathematical Geology*, 1(2), 137-153.
- 583 Agterberg, F.P., & Kelly, A.M., 1971. Geomathematical methods for use in prospecting. *Canadian Mining Journal*, 92(5),
584 61-72.
- 585 Agterberg, F.P., 1964. Methods of trend surface analysis. *Colorado School Mines Quart*, 59(4), 111-130.
- 586 Agterberg, F.P., 1970. Multivariate prediction equations in geology. *Journal of the International Association for*
587 *Mathematical Geology*, 1970 (02), 319-324.
- 588 Agterberg, F.P., 1971. A probability index for detecting favourable geological environments. *Canadian Institute of Mining*
589 *and Metallurgy*, 10, 82-91.
- 590 Agterberg, F.P., 1989. Computer Programs for Mineral Exploration. *Science*, 245, 76 – 81.
- 591 Agterberg, F.P., 1992. Combining indicator patterns in weights of evidence modeling for resource evaluation. *Nonrenewal*
592 *Resources*, 1(1), 35–50.
- 593 Agterberg, F.P., 2011. A Modified WofE Method for Regional Mineral Resource Estimation. *Natural Resources Research*,
594 20(2), 95-101.
- 595 Agterberg, F.P., & Bonham-Carter, G.F., 1999. [Logistic regression and weights of evidence modeling in mineral exploration.](#)
596 [In Proceedings of the 28th International Symposium on Applications of Computer in the Mineral Industry. Golden,](#)
597 [Colorado, pp 483-490.](#)
- 598 [Agterberg, F.P., Bonham-Carter, G.F., & Wright, D.F., 1990. Statistical Pattern Integration for Mineral Exploration. in Gaál,](#)
599 [G., Merriam, D. F., eds. Computer Applications in Resource Estimation Prediction and Assessment of Metals and](#)
600 [Petroleum. New York: Pergamon Press: 1-12.](#)
- 601 ~~[Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., & Wright, D.F., 1993. Weights of evidence modeling and weighted](#)~~
602 ~~[logistic regression for mineral potential mapping. *Computers in geology*, 25, 13-32.](#)~~

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

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

607 Brauer, M., Freedman, G., Frostad, J., Van Donkelaar, A., Martin, R.V., Dentener, F., ... & Balakrishnan, K., 2015. Ambient
608 air pollution exposure estimation for the global burden of disease 2013. *Environmental science & technology*, 50(1),
609 79-88.

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

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

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

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

618 Casetti, E., 1972. Generating models by the expansion method: applications to geographic research. *Geographical Analysis*, 4,
619 81-91.

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

622 Cheng, Q., 1999. Multifractality and spatial statistics. *Computers & Geosciences*, 25, 949-961.

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

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

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

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

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

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

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

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

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

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

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

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

645 Pearson, K., Lee, A., & Bramley-Moore, L., 1899. *Mathematical contributions to the theory of evolution*. VI. Genetic
646 (reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred racehorses. *Philosophical*

647 | Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character,-_192,
648 | 257-330.

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

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

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

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

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

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

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

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

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

669 | Yule, G.U., 1903. Notes on the theory of association of attributes in statistics.-_Biometrika,-_2(2), 121-134.

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

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

674 | Zhang, D., Cheng, Q., Agterberg, F.P., & Chen, Z., 2016. An improved solution of local window parameters setting for local
675 | singularity analysis based on excel vba batch processing technology.-_Computers & Geosciences,-_88(C), 54-66.

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