



1 **Remote sensing data processing by multivariate regression analysis**
2 **method for iron mineral resource potential mapping: A case study in**
3 **Sarvian area, central Iran.**

4 Edris Mansouri¹, Faranak Feizi^{2*}, Alireza Jafari Rad³, Mehran Arian⁴

5

6 1- Department of Geology, Science and Research Branch, Islamic Azad University, Tehran, Iran

7 2- Department of Mining Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

8 3- Department of Geology, Science and Research branch, Islamic Azad University, Tehran, Iran

9 4- Department of Geology, Science and Research Branch, Islamic Azad University, Tehran, Iran

10 *corresponding author

11 Fax number: +98 021 88830012; Cell: +98 912 3006753; faranakfeizi@gmail.com.

12 **ABSTRACT**

13 This paper used multivariate regression to create a mathematical model (with reasonable accuracy) for iron skarn
14 exploration in the region of the interest and generalizing multivariate regression in Mineral Prospectivity Mapping
15 (MPM) field. The main target of this manuscript is to exert multivariate regression analysis (as a MPM method) to
16 iron outcrops mapping from northeast part of the study area to discover new iron deposits in other parts. Two types of
17 multivariate regression models as two linear equations were employed to discover new mineral deposits. The Aster
18 satellite image bands (14 bands) sets as Unique Independent Variables (UIVs) and iron outcrops map as dependent
19 variables were used for MPM. According to the results of p-value, R^2 and R_{adj}^2 , the second regression model (which
20 was a multiples and exponents of UIVs) was the fitted model versus other models. Also the accuracy of the model
21 was confirmed by iron outcrops map and geological observations. Based on field observation iron mineralization
22 occurs as contact of limestone and intrusive rocks (skarn type). Iron minerals consist dominantly of magnetite,
23 hematite and goethite.

24

25 **Key words:** Multivariate regression, Mineral Prospectivity Mapping, Iron, Sarvian

26



27 1. INTRODUCTION

28 Diagnosing futuristic zones and finding new mineral deposits in the region of interest, is the
29 definitive main of mineral investigation. One way to achieve this aim is Using satellite image
30 processing for identify Mineral Prospectivity Mapping (MPM) (Carranza, 2008; Abedi et al., 2013;
31 Golshadi et al., 2016 and Feizi et al., 2012).

32 The utilization of satellite images for mineral investigation has been extremely effective in
33 indicating out the attendance of minerals. Likewise, remote sensing gives the synoptic view, which
34 is useful in distinguishing proof and delineation of different land frames, linear features, and
35 structural elements (Feizi and Mansouri, 2013b).

36 The main objective of this manuscript is to use multivariate regression analysis (as a MPM
37 method) to pixel values of Aster satellite image from north-east part of the study area to identify
38 new iron deposits in other parts. Two types of multivariate regression models utilized to find new
39 mineral deposits, using pixel values of Aster satellite image bands (14 band) sets as Unique
40 Independent Variables (UIVs) and Iron outcrops surface (digitized by geology map of study area
41 (scale 1:5000) and check field) data as dependent variables.

42 Regression analyses have been utilized as a part of numerous logical fields, such as
43 geosciences.

44 Identification of stream sediment anomalies have been used by multiple regression analyses
45 (e.g., Carranza, 2010a; Carranza, 2010b). Likewise, multivariate regression has been effectively
46 utilized by Granian et al. (2015) to display subsurface mineralization from lithochemical
47 information. Granian et al. (2015) utilized four types of multivariate regression models to depict
48 significant surface geochemical anomalies for acknowledgment subsurface gold mineralization
49 utilizing borehole data as dependent variables and surface lithochemical data as independent
50 variables.

51 This paper utilized multivariate regression to make a mathematical model (with sensible
52 precision) for iron potential zones investigation in the region of the interest and summing up
53 multivariate regression in remote sensing field.

54

55 2. STUDY AREA

56 The Sarvian area is located in the Orumieh-Dokhtar magmatic arc in Central of Iran (Fig. 1a).
57 This magmatic arc is the most imperative metallogenic area inside the district and hosts the
58 majority of the larger metals deposits such as lead, zinc, copper and iron (Feizi et al., 2016 and
59 Feizi et al., 2017).

60 The explored zone determined by Eocene intrusive rocks and carbonates of Qom formation.
61 Several types of metal and non-metal mineral ore deposits have as of now been reported in the



62 study area. According to the 1:100,000 geological map of Kahak, the lithology of this part includes
 63 cream limestone with intercalations of marls (Qom formation), dark green, andesitic-basaltic lava,
 64 volcanic breccia, hyaloclastic limestone, green megaporphyritic andesitic-basaltic lava,
 65 rhyodacitic domes, tonalite-quartzdiorite, microquartzdiorite-microquartzmonzo-diorite, granite-
 66 granodiorite, alteration of light green, grey tuff, tuffaceous sandstone and shale with intercalation
 67 of nummulitic sandy limestone and andesitic lava, grey limestone, orbitolina bearing, thick bedded
 68 to massive (Aptian–Albian) (Feizi et al., 2016) (Fig. 1b).

69 In view of the current confirmations and furthermore contact of intrusive bodies with carbonate
 70 rocks (Qom arrangement) and Iron outcrops in the north-east of study area, Calcic iron skarn ore
 71 (Sarvian mine) is located in the northeast of study area (Feizi et al., 2017) (Fig. 2).

72

73 3. MULTIVARIATE REGRESSION

74 Regression analyses is a good statistical manner for analysing relationships among variables
 75 (Granian et al., 2015). This strategy can show the conduct of an event (a dependent variable) in
 76 light of related variables (some independent variables). In regression analyses, if a dependent
 77 variable called (Y) and independent variables called (x_i), the equation is:

$$78 \quad Y = f(x_i). \quad (1)$$

79 Y could be linear or non-linear function. Linear regression is used for modeling mineral
 80 prospectivity in the Sarvian area. For linear regression Y is defined as follows:

$$81 \quad Y = a_0 + a_1x_1 + a_2x_2 + \dots + a_ix_i + \varepsilon, \quad i = 1, 2, \dots, n. \quad (2)$$

82

83 For this function, the constant factor is a_0 , the random error is ε , and the regression
 84 coefficients are a_i . If there were n samples in a data set, for each sample t variables were
 85 measured. Thus, function (2) can be as follows:

86

$$87 \quad Y_i = \hat{a}_0 + \hat{a}_1X_{i1} + \hat{a}_2X_{i2} + \dots + \hat{a}_tX_{it} + \varepsilon_i \quad i = 1, 2, \dots, n. \quad (3)$$

88

89

90 Equation (3) can be re-written as a matrix. The linear function matrix is:

$$91 \quad [Y] = [X][A] + [\varepsilon]. \quad (4)$$



92

$$[Y] = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; [A] = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_t \end{bmatrix}; [X] = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1t} \\ 1 & X_{21} & X_{22} & \dots & X_{2t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nt} \end{bmatrix}; [\varepsilon] = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}. \quad (5)$$

94

95 The least squares technique is used for estimating $[A]$ as the coefficient matrix, as follows:

96

$$[A] = [\Sigma]^{-1}[C] = ([X]'[X])^{-1}[X]'[Y]. \quad (6)$$

98

99 The inverse of variance-covariance samples matrix is $[\Sigma]^{-1}$ and the covariance matrix among
 100 independent variable and samples is $[C]$. Thus by equation 6, the regression coefficients model is
 101 calculated.

102 In regression analysis, some criteria are required to review. These criteria are as follows:

103 1. The variance and the mean of the random error should be a constant value and zero,
 104 respectively.

105 2. The coefficient of determination value (R^2) should be examined. This value is calculated as
 106 follows (Granian et al., 2015):

107

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (7)$$

109

110 The mean of the variable (\bar{Y}), value of i th sample (Y_i) and estimated value of the i th sample
 111 (\hat{Y}_i) for dependent variables were used in equation 7. The calculated R^2 value determined within
 112 $[0, 1]$ range. The value of R^2 is close to 1 for well fitted models.

113 1. According to the fact that adding independent variables to the model will increasing R^2
 114 value, the adjusted determination coefficient (R_{adj}^2) is defined as follows (Granian et al.,
 115 2015):

$$R_{adjusted}^2 = 1 - \frac{n-1}{n-t}(1 - R^2). \quad (8)$$

116



117 As it was mentioned, n is number of samples (or data) and t is the number of variables (or
118 regression coefficients). If a set of explanatory variables are introduced into a regression one at a
119 time, with the R_{adj}^2 computed each time, the level at which R_{adj}^2 reaches a maximum, and decreases
120 afterward, would be a well fitted model.

121 2. In regression analyses, the model should be fitted to the data. Accordingly, the p-value of
122 the regression model in Analysis of variance (ANOVA) test should be acceptable (less
123 than or equal to 0.05). Also calculating p-value of final coefficients for each model, could
124 help on optimizing and improving the model. This criterion could be considered after
125 choosing the best model.

126 4. GEO-DATA SETS PREPARATION

127 The iron ore skarn type located in the northeastern of Sarvian area. There are several iron vein
128 and outcrop in this area. According to the regional geological conditions of the area, the data set
129 of this iron mine is a good model for exploring the surrounding area. In this paper, satellite imagery
130 and map the geology of the mine is used as a training area. In the training area, this method can
131 model the iron outcrops (a dependent variable) based on Aster satellite image bands (some
132 independent variables) (Fig. 3).

133 **Figure 3 is about here.**

134 4.1. REMOTE SENSING DATA (INDEPENDENT VARIABLES)

135 The ASTER sensor was propelled in December 1999 on board the Earth Observation System
136 (EOS) US Terra satellite to record sun powered radiation in 14 spectral bands (Table 1). ASTER
137 provides high-resolution images of the land surface, water, ice, and clouds using three separate
138 sensor subsystems covering 14 multi-spectral bands from visible to thermal infrared. The
139 significant resolution scales are 15m, 30m, and 90m in the visible, short-wave IR, and thermal IR,
140 respectively. ASTER consists of three different subsystems; the Visible and Near Infrared (VNIR),
141 the Shortwave Infrared (SWIR), and the Thermal Infrared (TIR). To find out more about each
142 module click on the item of interest (Feizi and Mansouri, 2013b and Mansouri and Feizi, 2016).

143 Several factors influence the signal measured at the sensor, for example, float of the sensor
144 radiometric calibration, atmospheric and topographical effects. In this way, Aster data collection
145 was utilized and radiance correlation, such as wavelength, dark subtract and log residual by
146 ENVI5.1 software which is basic for multispectral images, were utilized (Mansouri et al., 2015).

147 In this study after the corrections, pixels size of SWIR and TIR bands based on VNIR3 band
148 (Panchromatic band) convert to 15 meter, than use layer stacking function to build a new multiband
149 file from georeferenced images of various pixel sizes, extents, and projections.

150 4.2. MAPPING OF IRON OUPCROPS (DEPENDENT VARIABLE)



151 The iron ore skarn type located in the northeastern of Sarvian area. There are several iron vein
152 and outcrop in this area. In order to mapping of iron outcrops in the training area used from
153 Geological map (1:1000 scale) of iron ore deposit and check field. For preparing of this layer, the
154 shape file layer of iron outcrops convert to raster file with pixel size of 15 meter.

155 5. REGRESSION ANALYSES IN THE STUDY AREA

156 Regression analyses needs making proper models. Utilizing multiple, factorial, polynomial and
157 reaction surface regressions have been utilized as a part of numerous logical fields such as
158 geosciences (e.g. Granian et al., 2015). Thus, in this study; Model 1 (Y_1) was generated as a
159 multiple linear regression model and Model 2 (Y_2) was created from Y_1 plus multiplied UIVs. The
160 formulas of two mentioned models are presented in Table 2. So in summary, two linear equations
161 (Y_1 and Y_2) were utilized to discover new mineral deposits, using pixel values of Aster satellite
162 image as independent variables and map of iron outcrops as dependent variables. The models
163 which were proposed in this paper, had become more complexes respectively 1 to 2, model 2 has
164 106 coefficients (14 for UIVs, 1 as constant, 91 for multiples of UIVs) and model 1 has 15
165 coefficients (14 for UIVs, 1 as constant, 0 for multiples and exponents of UIVs) (Table 2).

166 For assessing the models which are exhibited in Table 2, regression analyses were performed
167 and the critical criteria which are mentioned before, were examined. The values of the R^2 , R^2_{adj}
168 and p-value of ANOVA test of 2 multivariate regression models are provided in Table 3.
169

170 Also, Table 4 is presented the calculated coefficients of independent variables in regression
171 models. Other independent variables which are not mentioned in Table 4 were excluded variables.
172 The excluded variables have no effect on the models. This means that, excluded variables didn't
173 have any effect on iron mineralization and behavior of iron outcrop map.

174

175 6. DISCUSSION

176 For distinguishing the best model among 2 models, a few criteria are required to review.

177 Firstly, the variance and the mean of the random error were acceptable for all of regression
178 models. Secondly, based on Table 4, the p-values of ANOVA test of 2 multivariate regression
179 models were equal to 0. For regression models the acceptable p-value should be less than or equal
180 to 0.05. Thus, this criterion confirmed the validity of models without specifying the most
181 appropriate model.

182 On the other hand, the value of R^2 is close to 1 for well fitted models. The R^2 values of
183 regression models are presented in Table 3. The lowest R^2 belongs to Y_1 and the highest belongs to
184 Y_2 . Thus, Y_2 model is better from Y_1 model.



185 According to the fact that adding independent variables to the model will increasing R^2 value,
186 the R_{adj}^2 value should be checked. The R_{adj}^2 values of regression models are presented in Table 3.
187 As it was mentioned before, if a set of variables are introduced into a regression, with the
188 R_{adj}^2 computed each time, the level at which R_{adj}^2 reaches a maximum, and decreases afterward,
189 would be a well fitted model. So, according to Table 3, Y_2 would be the fitted model versus other
190 models. Thus, Y_4 regression model is the most appropriate model for Mineral Prospectivity
191 Mapping.

192 Thus according to the results of p-value (ANOVA test), R^2 and R_{adj}^2 , the Second regression
193 model (Y_2) would be the fitted model versus other models. For generating the mineral prospectivity
194 map the formula of Y_2 was performed in ArcGIS software by raster calculator tool. The normalized
195 mineral prospectivity map of the study area is presented in Fig. 4.

196 To assess the exactness of the selected model, the created prospectivity map was checked by
197 the iron outcrops map in the northeast part of the study area (Fig. 5). The locations of iron outcrops
198 have appropriate adoption with favorable areas of mineral prospectivity map. In addition the
199 adaption of prospectivity map with the iron outcrops in the northeast part of the study area, three
200 target areas with very high favourability, were checked and the prospectivity map was confirmed
201 by geological observations (Fig. 6). Based on field observation iron mineralization occurs as
202 contact of limestone and intrusive rocks (skarn type). Iron mineralizations consist dominantly of
203 magnetite, hematite and goethite. Therefore, the accuracy of mineral prospectivity map confirmed
204 in the Sarvian area.

205

206 7. CONCLUSION

207

208 The conclusions of this manuscript are presented in summary as follows.

209 1) The application of multivariate regression analysis (as a MPM method) was confirmed in
210 the Sarvian area. This paper used multivariate regression to create a mathematical model (with
211 reasonable accuracy) for iron mineral exploration in the region of the interest and generalizing
212 multivariate regression in MPM field.

213 2) Two types of multivariate regression models as two linear equations were employed to
214 discover new mineral deposits. According to the results of p-value, R^2 and R_{adj}^2 , the second
215 regression model was the fitted model versus other models.

216 3) Also the accuracy of the model was confirmed by iron outcrops map and geological
217 observations. Based on field observation iron mineralization occurs contact of limestone and
218 intrusive rocks (skarn type). Iron mineralizations consist dominantly of magnetite, hematite and
219 goethite.



220 **ACKNOWLEDGEMENTS**

221 The authors would like to thank Amirabbas KarbalaeiRamezanali for his helpful suggestions.

222

223

224 **REFERENCES**

225 Carranza, E.J.M., 2008, Geochemical anomaly and mineral prospectivity mapping in GIS,
226 Handbook of Exploration Environmental Geochemistry. Elsevier, Amsterdam, 368 p.

227 Carranza, E.J.M., 2010a, Catchment basin modelling of stream sediment anomalies revisited:
228 incorporation of EDA and fractal analysis. *Geochemistry: Exploration, Environment, Analysis*,
229 10, 365–381.

230 Carranza, E.J.M., 2010b, Mapping of anomalies in continuous and discrete fields of stream
231 sediment geochemical landscapes. *Geochemistry: Exploration, Environment, Analysis*, 10, 171–
232 187.

233 Feizi, F. and Mansouri, E., 2012, Identification of Alteration Zones with Using ASTER Data
234 in A Part of Qom Province, Central Iran. *Journal of Basic and Applied Scientific Research*, 2, 73–
235 84.

236 Feizi, F. and Mansouri, E., 2013a, Separation of Alteration Zones on ASTER Data and
237 Integration with Drainage Geochemical Maps in Soltanieh, Northern Iran. *Open Journal of*
238 *Geology*, 3, 134–142.

239 Feizi, F. and Mansouri, E., 2013b, Introducing the Iron Potential Zones Using Remote Sensing
240 Studies in South of Qom Province, Iran. *Open Journal of Geology*, 3, 278–286.

241 Feizi, F., Mansouri, E. and KarbalaeiRamezanali, A., 2016, Prospecting of Au by Remote
242 Sensing and Geochemical Data Processing Using Fractal Modelling in Shishe-Botagh, Area (NW
243 Iran). *Journal of the Indian Society of Remote Sensing*, 44, 539–552.

244 Feizi, F., KarbalaeiRamezanali, A. and Mansouri, E., 2017, Calcic iron skarn prospectivity
245 mapping based on fuzzy AHP method, a case Study in Varan area, Markazi province,
246 Iran. *Geosciences Journal*, 21, 123–136.

247 Granian, H., Tabatabaei, S. H., Asadi, H. H. and Carranza, E. J. M., 2015, Multivariate
248 regression analysis of lithochemical data to model subsurface mineralization: a case study from
249 the Sari Gunay epithermal gold deposit, NW Iran. *Journal of Geochemical Exploration*, 148, 249–
250 258.



251 Golshadi, Z., KarbalaeeiRamezanali, A. and Kafaeei, K., 2016, Interpretation of magnetic data
252 in the Chenar-e Olya area of Asadabad, Hamedan, Iran, using analytic signal, Euler deconvolution,
253 horizontal gradient and tilt derivative methods. *Bollettino di Geofisica Teorica ed Applicata*, 57,
254 329–342.

255 Mansouri, E., Feizi, F. and KarbalaeeiRamezanali, A., 2015, Identification of magnetic
256 anomalies based on ground magnetic data analysis using multifractal modelling: a case study in
257 Qoja-Kandi, East Azerbaijan Province, Iran. *Nonlinear Processes in Geophysics*, 22, 579–587.

258 Mansouri, E., Feizi, F., 2016, Introducing Au potential areas, using remote sensing and
259 geochemical data processing using fractal method in Chartagh, western Azarbijan – Iran, E.
260 Mansouri, F. Feizi, *Arch. Min. Sci.*, Vol., No 2, 397–414.

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278



279 **Table 1.** Wavelength ranges and spatial resolutions of ASTER bands (Abrams, 2000).
 280

Module	VNIR	SWIR	TIR
Spectral bandwidth (μm)	Band 1 0.52 - 0.60	Band 4 1.650 - 1.700	Band 10 8.125 - 8.475
	Band 2 0.63 - 0.69	Band 5 2.145 - 2.185	Band 11 8.475 - 8.825
	Band 3 N 0.78 - 0.86	Band 6 2.185 - 2.225	Band 12 8.925 - 9.275
	Band 3 B 0.78 - 0.86 (backward looking)	Band 7 2.235 - 2.285	Band 13 10.25 - 10.95
		Band 8 2.295 - 2.395	Band 14 10.95 - 11.65
		Band 9 2.360 - 2.430	
Spatial resolution (m)	15	30	90

281

282

283 **Table 2.** Formula of regression models used for Aster satellite image bands
 284

Types of Regression	Number of coefficients	Formula
First-Degree	15	$Y_1 = a_0 + a_1x_1 + a_2x_2 + \dots + a_{14}x_{14}$
First-Degree	106	$Y_2 = Y_1 + a_{15}x_1x_2 + a_{16}x_1x_3 + \dots + a_{27}x_1x_{14} + a_{28}x_2x_3 + a_{29}x_2x_4 + \dots$ $+ a_{39}x_2x_{14} + a_{40}x_3x_4 + a_{41}x_3x_5 + \dots + a_{50}x_3x_{14}$ $+ a_{51}x_4x_5 + \dots + a_{60}x_4x_{14} + a_{61}x_5x_6 + \dots + a_{69}x_5x_{14}$ $+ a_{70}x_6x_7 + \dots + a_{77}x_6x_{14} + a_{78}x_7x_8 + \dots + a_{84}x_7x_{14}$ $+ a_{85}x_8x_9 + \dots + a_{90}x_8x_{14} + a_{91}x_9x_{10} + \dots + a_{96}x_9x_{14}$ $+ a_{97}x_{10}x_{11} + \dots + a_{100}x_{10}x_{14}$ $+ a_{101}x_{11}x_{12} + \dots + a_{103}x_{11}x_{14} + a_{104}x_{12}x_{13} + a_{105}x_{12}x_{14}$ $+ a_{106}x_{13}x_{14}$

285

286

287 **Table 3.** The values of the R^2 , R_{adj}^2 and p-value of ANOVA test of 2 multivariate regression
 288 models

Models	R^2	R_{adj}^2	p-value (ANOVA)
Y_1	0.738	0.715	0
Y_2	0.847	0.829	0

289



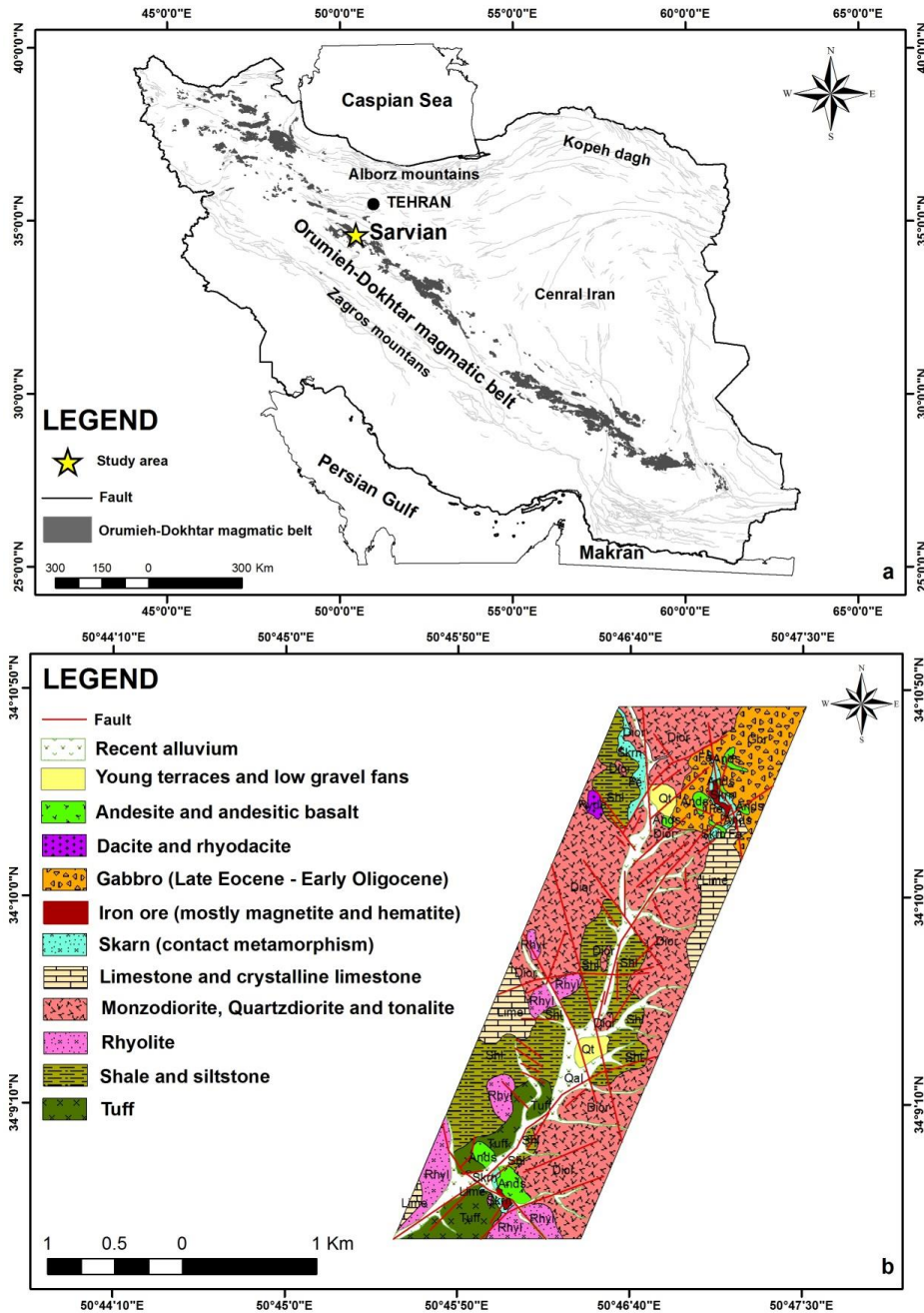
290

291 **Table 4.** The calculated coefficients of regression models 1 and 2.

Model 1		Model 2	
variables	Coefficients (a_i)	variables	Coefficients (a_i)
CST	0.275	CST	0.677
x_1	-0.01	x_1	-0.014
x_2	-0.12	x_2	-0.019
x_3	-0.019	x_3	-0.045
x_4	0.003	x_4	0.022
x_5	-0.006	x_5	-0.017
x_6	-0.005	x_6	-0.001
x_7	-	x_7	-
x_8	-0.004	x_8	-0.02
x_9	-0.005	x_9	-0.006
x_{10}	0.009	x_{10}	-0.014
x_{11}	0.005	x_{11}	0.024
x_{12}	0.016	x_{12}	0.024
x_{13}	0.002	x_{13}	0.018
x_{14}	0.022	x_{14}	0.036
-	-	x_1x_4	-0.0009
-	-	x_1x_6	-0.0002
-	-	x_4x_9	-0.0009
-	-	x_7x_8	0.00082

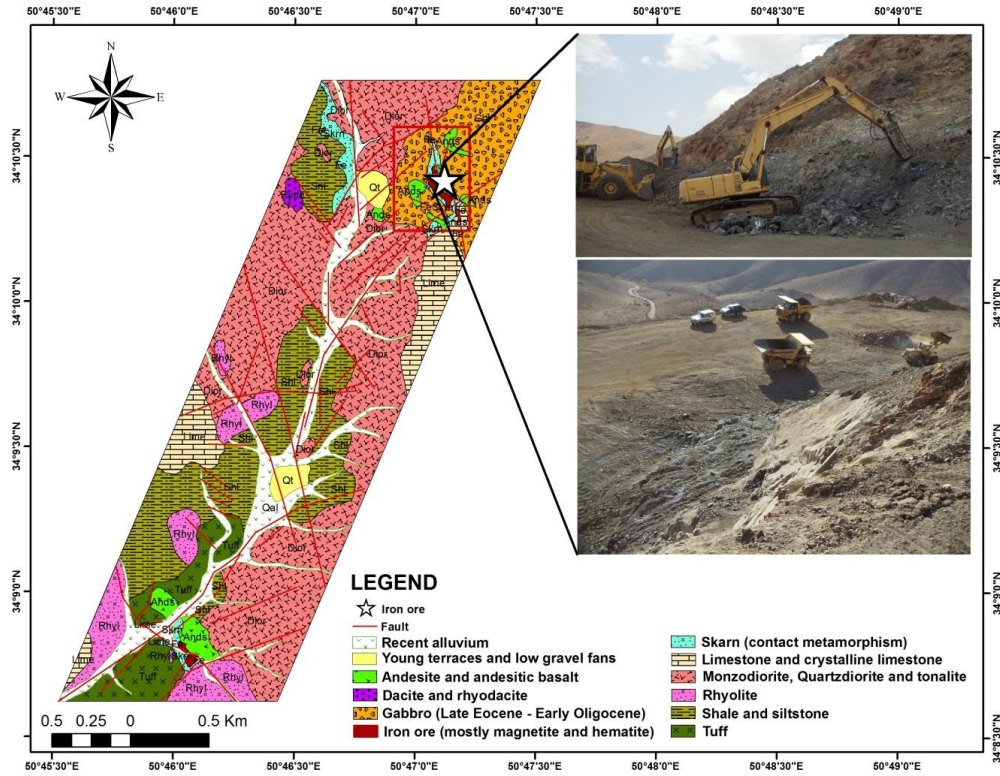
292

293



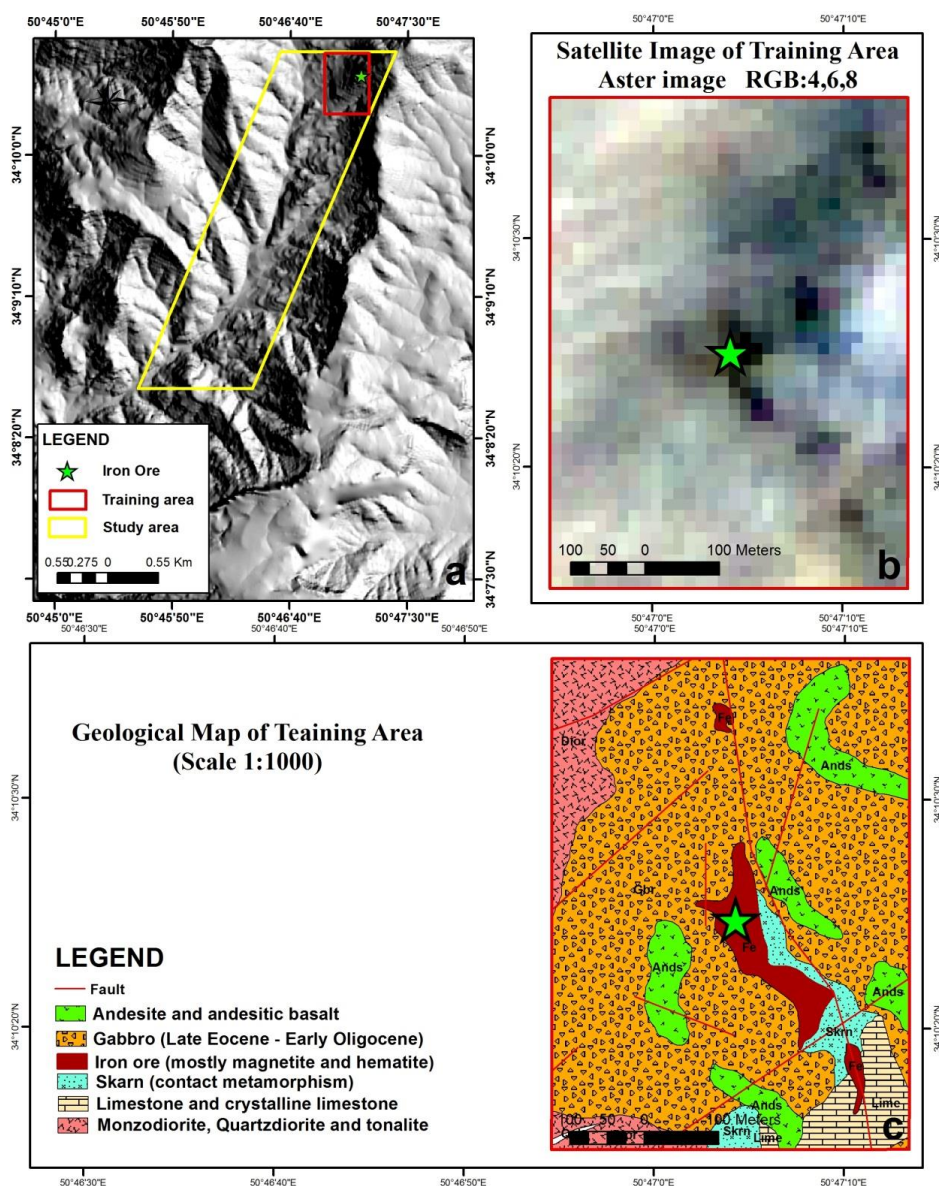
294
 295

296 **Fig. 1. a)** The location of the Sarvian area in the Urumieh–Dokhtar magmatic belt, Iran **b)** Geological map
 297 of Sarvian area (scale 1:25000)



298

299 **Fig. 2.** Location of Sarvian iron mine in the study area

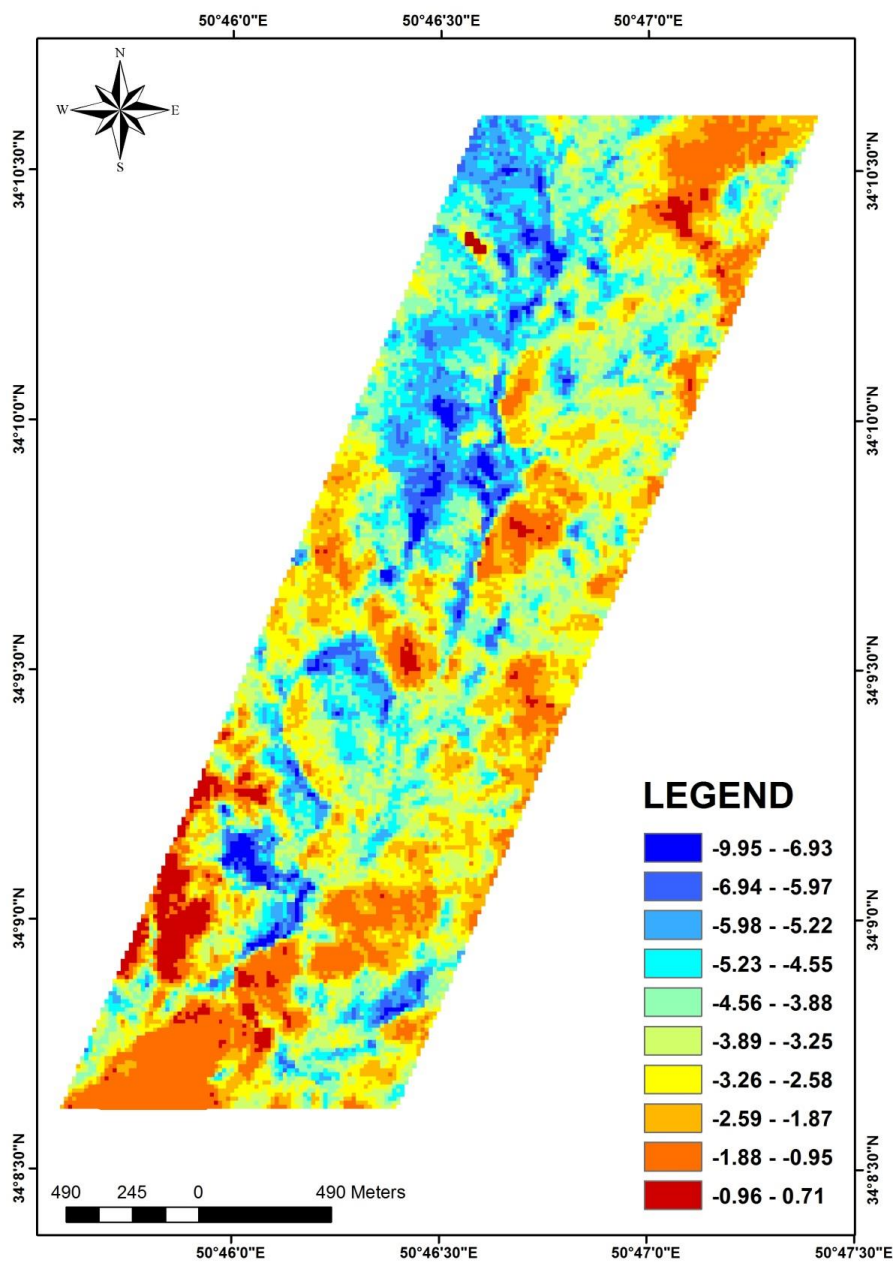


300

301 **Fig. 3.** a) Location of training area in the study area. b) Aster satellite image in the training area

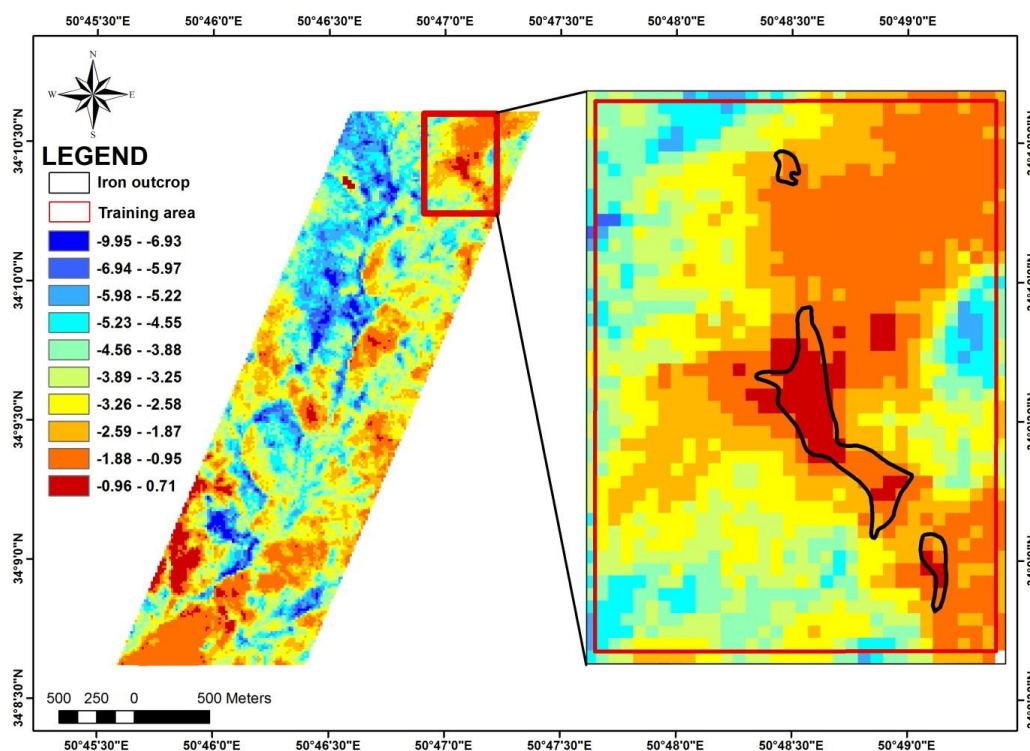
302

(RGB:4,6,8). c) Geological map (scale 1:1000) of training area.



303
304

Fig. 4. Mineral prospectivity map of the Sarvian area.

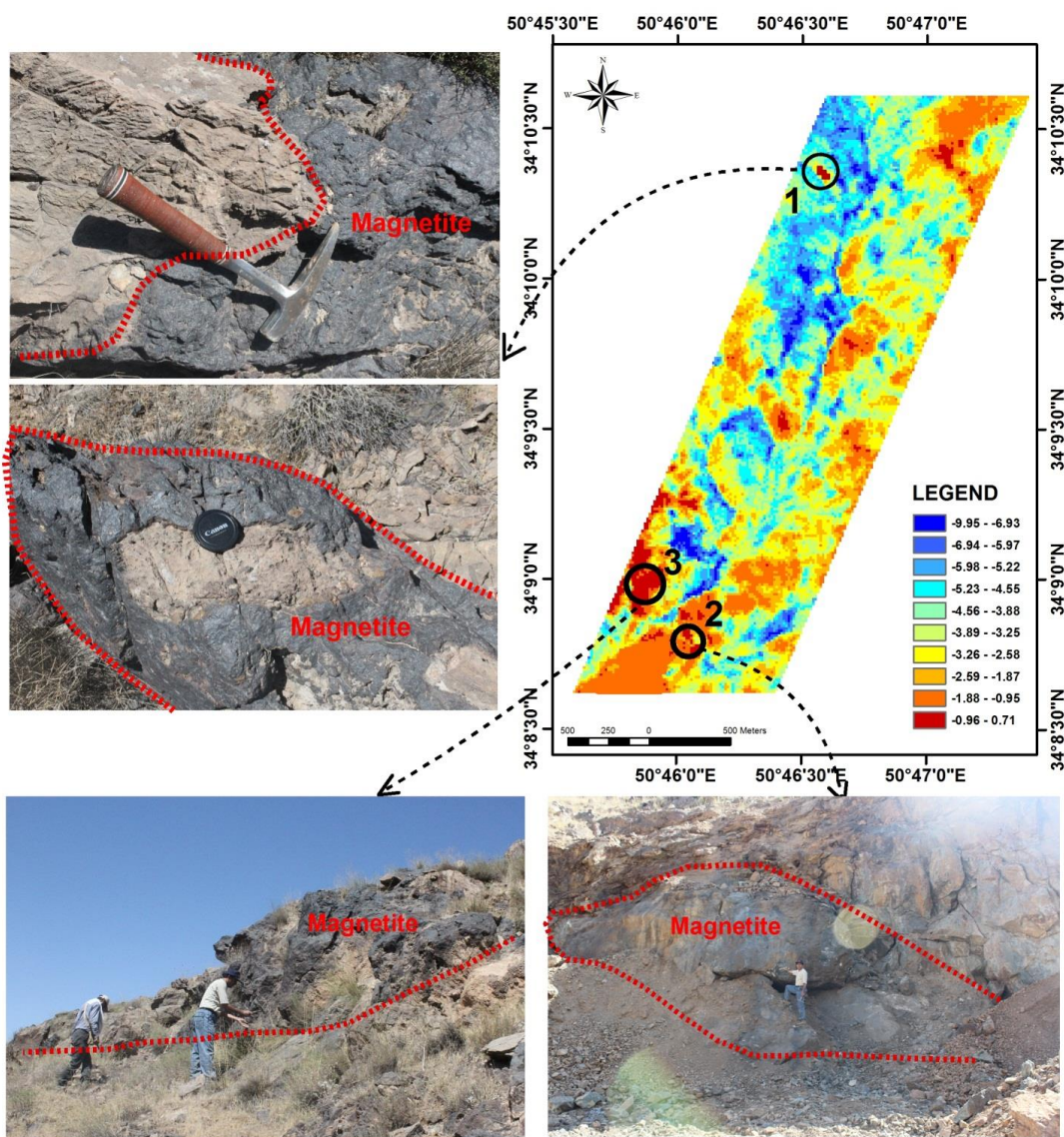


305

306

Fig. 5. Mineral prospectivity map of the Sarvian area which confirmed by iron outcrops.

307



308
309 **Fig. 6.** Mineral prospectivity map of the Sarvian area which confirmed by check field of three target
310 areas.

311
312
313