(This file contains suggested edits, and some comments. Equations were left out for simplicity. Yellow highlights mark places where I cannot understand what is meant.)

Colin Pain

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

Deleted: .

Edris Mansouri[1], Faranak Feizi[2*], Alireza Jafari Rad[3], Mehran Arian[4]

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

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

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

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

*corresponding author

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

ABSTRACT

This paper uses multivariate regression to create a mathematical model for iron skarn exploration in the Sarvian area, central Iran, using multivariate regression for Mineral Prospectivity Mapping (MPM). The main target of this manuscript is to apply multivariate regression analysis (as an MPM method) to mapping iron outcrops in the northeast part of the study area in order to discover new iron deposits in other parts of the study area. Two types of multivariate regression models using two linear equations were employed to discover new mineral deposits. Aster satellite images (14 bands) were used as Unique Independent Variables (UIVs), and iron outcrops were mapped as dependent variables for MPM. According to the results of p-value, R^2 and R^2_{adj} , the second regression model (which was a multiples and exponents of UIVs) fitted better than other models. The accuracy of the model was confirmed by iron outcrops map and geological observations. Based on field observation, iron mineralization occurs at the contact of limestone and intrusive rocks (skarn type). Iron minerals consist dominantly of magnetite, hematite and goethite.

Deleted: used
Deleted: (with reasonable accuracy)
Deleted: region of the interest and generalizing
Deleted: in
Deleted: field
Deleted: exert
Deleted: mapping from
Deleted: as
Deleted: The
Deleted: bands
Deleted: sets
Deleted: were used
Formatted: Highlight
Deleted: was the
Deleted: model
Deleted: versus
Deleted: Also
Deleted: as

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

1. INTRODUCTION

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

The utilization of satellite images for mineral investigation has been extremely effective in indicating the presence of minerals. Likewise, remote sensing gives a synoptic view, which is useful for identifying and delineating different landscapes, linear features, and structural elements (Feizi and Mansouri, 2013b).

The main objective of this research was to use multivariate regression analysis (as a MPM method) to use pixel values from Aster satellite images of the north-east part of the study area to identify new iron deposits in other parts. Two types of multivariate regression models wereused to find new mineral deposits, using pixel values of Aster satellite image bands (14 bands) as Unique Independent Variables (UIVs), and iron outcrop areas (digitized a 1:5000geology map of the study area and field data as dependent variables.

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

Multiple regression analyses have been used to identify stream sediment anomalies (e.g., Carranza, 2010a; Carranza, 2010b). Likewise, multivariate regression has been effectively

Formatted: Highlight
Formatted: Highlight
Deleted: Using
Deleted: identify
Deleted: out
Deleted: attendance
Deleted: the
Deleted: in distinguishing proof
Deleted: delineation
Deleted: of
Deleted: land frames
Deleted: manuscript is
Deleted: of
Deleted: from
Deleted: utilized
Deleted: sets
Deleted: Iron
Deleted: surface
Deleted: by
Deleted: (scale 1:5000)
Deleted: check
Deleted:)
Formatted: Highlight
Formatted: Highlight
Deleted: identification
Deleted: of
Deleted: have been used by multiple regression analyses

used by Granian et al. (2015) to display subsurface mineralization from lithogeochemical information. Granian et al. (2015) used four types of multivariate regression models to depict significant surface geochemical anomalies indicating subsurface gold mineralization utilizing borehole data as dependent variables and surface lithogeochemical data as independent variables.

This paper uses multivariate regression to develop a useful and precise mathematical model of iron potential zones using remote sensing of the region of the interest.

2. STUDY AREA

The Sarvian area is in the Orumieh-Dokhtar magmatic arc in Central of Iran (Fig. 1a). This magmatic arc is the most important metallogenic area inside the district and hosts large metal deposits such as lead, zinc, copper and iron (Feizi et al., 2016 and Feizi et al., 2017).

The study area is dominated by Eocene intrusive rocks and carbonates of the Qom formation. Several types of metal and non-metal mineral ore deposits have been reported in the study area. According to the 1:100,000 geological map of Kahak, the lithology of this area includes cream limestone with intercalations of marls (Qom formation), dark green, andesitic-basaltic lava, volcanic breccia, hyaloclastic limestone, green megaporphyritic andesitic-basaltic lava, rhyodacitic domes, tonalite-quartzdiorite, microquartzdiorite-microquartzmonzo-diorite, granite-granodiorite, altered light green, grey tuff, tuffaceous sandstone and shale with intercalation

Deleted: utilized
Deleted: utilized
Deleted: for acknowledgment
Deleted: utilized
Deleted: make
Deleted: (with sensible
Deleted: precision) for
Deleted: investigation in
Deleted: and summing up
Deleted: multivariate regression in remote sensing field
Deleted: located
Deleted: imperative
Deleted: the
Deleted: majority of the larger
Deleted: s
Deleted: explored zone determined
Deleted: as of now
Deleted: part
Deleted: alteration of

of nummulitic sandy limestone and andesitic lava, and orbitolina-bearing, thick bedded to massive grey limestone (Aptian-Albian) (Feizi et al., 2016) (Fig. 1b).

These relationships are demonstrated by the calcic iron skarn ore (Sarvian mine) in the northeast of study area (Feizi et al., 2017) (Fig. 2).

3. METHODS AND DATA

3.1 Multivariate Regression

Regression analyses is a good statistical tool for analysing relationships among dependent and independent variables (Granian et al., 2015).

In regression analyses, for dependent variables (Y) and independent variables (x_i), the equation is:

(1)

Y can be a linear or non-linear function.

For linear regression Y is defined as follows:

(2)

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

Deleted: ,

Deleted: orbitolina bearing, thick bedded

Deleted: to massive

Deleted: In view of the current confirmations and furthermore contact of intrusive bodies with carbonate

Deleted: rocks (Qom arrangement) and Iron outcrops in the north-east of study area, Calcic

Deleted: is located

Deleted: manner

Deleted: This strategy can show the conduct of an event (a dependent variable) in

Deleted: light of related variables (some independent variables).

Deleted: if a

Deleted: could

Deleted: Linear regression is used for modeling mineral

Deleted: prospetivity in the Sarvian area.

Deleted:

Deleted: were

Deleted: can be

(3)

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

(4)

(5)

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

(6)

The inverse of variance-covariance samples matrix is $[\sum]^{-1}$and the covariance matrix among independent variable and samples is [*C*]. Thus the regression coefficients model is calculated from equation 6.

The following criteria were used for the regression analysis:

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

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

Deleted: by equation 6,

Deleted: In

Deleted: , some criteria are required to review. these criteria are as follows

(7)

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

1. Given the fact that adding independent variables to the model will increase the R^2 value, the adjusted determination coefficient (R^2_{adj}) is defined as follows (Granian et al., 2015):

(8)

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

2. In regression analyses, the model should be fitted to the data. Accordingly, the p-value of the regression model in the Analysis of variance (ANOVA) test should be acceptable (less than or equal to 0.05). Calculating the p-value of final coefficients for each model, may also help optimize and improve the model. This criterion could be applied after choosing the best model.

3.2. Geo-data Preparation

There are several iron ore bodies and one iron mine in the northeastern Sarvian study area.

The regional geological conditions of the area, suggest that

Deleted: According to
Deleted: increasing
Deleted: Also calculating
Deleted: could
Deleted: on
Deleted: ing
Deleted: improving
Deleted: considered
Deleted: 4
Deleted: SETS
Deleted: The
Deleted: skarn type located
Deleted: of
Deleted: There are several iron vein
Deleted: and outcrop in this area. According to
Deleted: the data set

the Sarvian iron mine is a good model for exploring the surrounding area. In this paper, a geology map of the mine is used as a training area for satellite imagery. In the training area, this method can model the iron outcrops (a dependent variable) based on Aster satellite image bands (independent variables) (Fig. 3).

Figure 3 is about here.

3.2. REMOTE SENSING DATA (INDEPENDENT VARIABLES)

The ASTER sensor was launched in December 1999 on board the Earth Observation System (EOS) US Terra satellite. ASTER provides high-resolution images of the land surface, water, ice, and clouds using three separate sensor subsystems covering 14 multi-spectral bands from visible to thermal infrared (Table 1). Resolutions are 15m, 30m, and 90m in the Visible and Near Infrared (VNIR), Shortwave Infrared (SWIR), and Thermal Infrared (TIR), respectively.

For more information see Feizi and Mansouri, (2013b) and Mansouri and Feizi, (2016).

Several factors influence the signal measured at the sensor, for example, float of the sensor radiometric calibration, and atmospheric and topographical effects. In this way, Aster data were analysed using ENVI5.1 software to provide information such as wavelength, and log residuals which are basic for multispectral analyses (Mansouri et al., 2015).

In this study after correction, the pixel size of SWIR and TIR bands based on VNIR3 band

Deleted: of this
Deleted: satellite imagery
Deleted: and map the geology of the mine
Deleted: (some
Deleted: 4.1
Deleted: INDEPENDEBT
Deleted: propelled
Deleted: to record sun powered radiation in 14 spectral bands (Table 1).
Deleted: The
Deleted: significant
Deleted: scales
Deleted: visible
Deleted: short-wave IR
Deleted: thermal fR
Deleted: ASTER consists of three different subsystems; the Visible and Near Infrared (VNIR),
Deleted: the Shortwave Infrared (SWIR), and the Thermal Infrared (TIR). To find out more about each
Deleted: module click on the item of interest (
Formatted: Highlight
Deleted: collection
Deleted: was
Formatted: Highlight
Deleted: utilized
Deleted: and
Deleted: radiance correlation,
Deleted: dark subtract
Deleted: by
Comment [CP1]: Please check to make sure I haven't changed what you want to say.
Deleted: ENVI5.1 software
Deleted: is
Deleted: images, were utilized
Deleted: the
Deleted: s
Deleted: pixles

(Panchromatic band) was converted to 15 m. The layer stacking function was then used to build a new multiband file from georeferenced images of various pixel sizes, extents, and projections.

3.3. MAPPING OF IRON OUPCROPS (DEPENDENT VARIABLE)

There are several iron veins and outcrops around the iron ore skarn mine in the north-eastern part of the Sarvian area. Iron outcrops in the training area were mapped using a geological map at a scale of 1:1000 of the iron ore deposit. The map was then field checked. The shape file layer of iron outcrops was converted to a raster file with a pixel size of 15 m.

4. RESULTS OF REGRESSION ANALYSES

Multiple, factorial, polynomial and response surface regressions have been utilized in many fields including the geosciences (e.g. Granian et al., 2015). In this study; Model 1 (Y_1) was generated as a multiple linear regression model and Model 2 (Y_2) was created from Y_1 plus many UIVs. The formulas for the two models are presented in Table 2. Thus, two linear equations (Y_1 and Y_2) were used to discover new mineral deposits, using pixel values from ASTER satellite data as independent variables and a map of iron outcrops as dependent variables. Of the two models proposed in this paper, model 2 has 106 coefficients (14 for UIVs, 1 as constant, 91 for multiples of UIVs) and model 1 has 15 coefficients (14 for UIVs, 1 as constant, 0 for multiples and exponents of UIVs) (Table 2).

Regression analyses were performed to assess the models in Table 2, and the critical criteria mentioned above, were examined. The values of the R^2, R^2_{adj}

Deleted: eter,
Deleted: than use
Deleted: 4.2
Deleted: The iron ore skarn type located in the northeastern of Sarvian area. There are several iron vein
Deleted: and outcrop in this area. In order to mapping of iron outcrops in the training area used from
Deleted: Geological map (1:1000 scale) of iron ore deposit and check field. For preparing of this layer, the
Deleted: shape file layer of iron outcrops convert to raster file with pixel size of 15 meter.
Deleted: 5
Deleted: IN THE STUDY AREA
Deleted: Regression analyses needs making proper models. Utilizing
Comment [CP2]: Is this what you mean?
Deleted: reaction
Deleted: as a part of numerous logical
Deleted: such as
Deleted: Thus,
Deleted: multiplied
Deleted: of
Deleted: mentioned
Deleted: So in summary
Deleted: utilized
Deleted: of
Deleted: Aster
Deleted: image
Deleted: which were
Deleted: , had become more complexes respectively 1 to 2
Deleted: For assessing
Deleted: which are exhibited
Deleted: regression analyses were performed
Deleted: which are
Deleted: before

and p-value from the ANOVA test of 2 multivariate regression models are provided in Table 3.

Table 4 presents the calculated coefficients of independent variables in regression models. Excluded independent variables are not mentioned in Table 4. Excluded variables were those that had no effect on iron mineralization and the mapped distribution of iron outcrops.

5. DISCUSSION

We used several criteria to review the differences between the two models. Firstly, the variance and the mean of the random error were acceptable for both models. Secondly, based on Table 4, the p-values of ANOVA test of the two models were equal to 0. For regression models the acceptable p-value should be less than or equal to 0.05. Thus, this criterion confirmed the validity of the models without specifying the most appropriate model.

The value of R^2 is close to 1 for well fitted models. The R^2 values of regression models are presented in Table 3. Model Y_1 has a lower R^2 than Y_2. Thus, the Y_2 model is better than the Y_1 model. Because adding independent variables to the model will increasing the R^2 value, the R^2_{adj} value should be checked. The R^2_{adj} values of regression models are presented in Table 3. As mentioned above, if a set of variables are introduced into a regression, with the R^2_{adj} computed each time, the level at which R^2_{adj} reaches a maximum, and decreases afterward,

Deleted: of
Deleted: Also,
Deleted: is presented
Deleted: Other
Deleted: which are
Deleted: were excluded variables
Deleted: The excluded variables have no effect on the models. This means that,
Deleted: didn't
Deleted: have any
Deleted: behavior
Deleted: map
Deleted: 6
Deleted: For distinguishing the best model among 2 models, a few criteria are required to review.
Deleted: all of regression
Deleted: 2 multivariate regression
Deleted: On the other hand,
Deleted: The
Formatted: Subscript
Deleted: lowest
Deleted: belongs toY_1 and the highest
Deleted: belongs to
Deleted: from
Deleted: According to the fact that
Deleted: it was
Deleted: before

would be a well-fitted model. So, according to Table 3, Y_2 is the fitted model versus other models. Thus, Y_2 regression model is the most appropriate model for Mineral Prospectivity Mapping.

Thus according to the results of p-value (ANOVA test), R^2 and R^2_{adj}, the second regression model (Y_2) would be the fitted model versus other models. For generating a mineral prospectivity map the model Y_2 was implemented in ArcGIS using the raster calculator tool. The normalized mineral prospectivity map of the study area is presented in Fig. 4.

To assess the accuracy of the selected model, the created prospectivity map was checked against the iron outcrops map in the northeast part of the study area (Fig. 5). The locations of iron outcrops is in close agreement with predictions from the mineral prospectivity map. In addition three target areas with very high potential were checked for iron outcrops and the prospectivity map was confirmed by geological observations (Fig. 6). Based on field observation iron mineralization occurs at contacts between limestone and intrusive rocks (skarn type). Iron mineralization consists dominantly of magnetite, hematite and goethite. Therefore, the accuracy of the mineral prospectivity map is confirmed in the Sarvian area.

6. CONCLUSION

Deleted: well
Deleted: would be the
Comment [CP3]: Is this correct?
Deleted: Y_4
Deleted: Second
Deleted: the
Deleted: formula of
Deleted: performed
Deleted: software by
Deleted: exactness
Deleted: by
Deleted: have appropriate adoption with favorable areas of
Deleted: the
Deleted: adaption of prospectivity map with the iron outcrops in the northeast part of the study area,
Deleted: favourability
Deleted: ,
Deleted: as
Deleted: of
Deleted: s
Deleted: 7

The conclusions of this manuscript are as follows.

Deleted: presented in summary

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

Deleted: the interest and generalizing

Deleted: multivariate regression in MPM field.

2) Two types of multivariate regression models, as two linear equations, were employed to discover new mineral deposits. According to the results of p-value, R^2 and R^2_{adj}, , the second regression model best fitted obsservations.

Deleted: was the

Deleted: model versus other models

3) The accuracy of the model was confirmed by iron outcrop mapping and geological observations. Based on field observation iron mineralization occurs in contacts between limestone and intrusive rocks (skarn type). Iron mineralization consists dominantly of magnetite, hematite and goethite.

Deleted: Also

Deleted: s

Deleted: of

Deleted: s

ACKNOWLEDGEMENTS

The authors would like to thank Amirabbas Karbalaei Ramezanali for his helpful suggestions.

REFRENCES

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

Carranza, E.J.M., 2010a, Catchment basin modelling of stream sediment anomalies revisited:

incorporation of EDA and fractal analysis. Geochemistry: Exploration, Environment, Analysis, 10, 365-381.

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

Feizi, F. and Mansouri, E., 2012, Identification of Alteration Zones with Using ASTER Data in a Part of Qom Province, Central Iran. Journal of Basic and Applied Scientific Research, 2, 73-84.

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

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

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

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

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

Deleted: A

258.

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

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

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

Figure captions

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

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

Fig. 3. a) Location of the training area, b) ASTER satellite image of the training area (RGB:4,6,8). c) Geological map (scale 1:1000) of the training area.

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

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

Fig. 6. Mineral prospectivity map of the Sarvian area confirmed by check fields of the three target areas.

Deleted: in the study area

Deleted: Aster

Deleted: in

Comment [CP4]: Please explain th numbers in the legend.

Deleted: which

Deleted: which