minerals

Article

Recognizing Geochemical Anomalies Associated with Mineral Resources Using Singularity Analysis and Random Forest Models in the Torud-Chahshirin Belt, Northeast Iran

Amirreza Bigdeli, Abbas Maghsoudi * and Reza Ghezelbash

Faculty of Mining Engineering, Amirkabir University of Technology, Tehran 159163-431, Iran; a.bigdeli@aut.ac.ir (A.B.); rezaghezelbash@aut.ac.ir (R.G.)
* Correspondence: a.maghsoudi@aut.ac.ir

Abstract: Identifying the local geochemical anomalies from stream sediment samples is challenging in regional-scale exploration programs. For this purpose, some robust and reliable techniques must be applied to distinguish the geochemical targets from the background values. In this research, a procedure of several tools, including singularity mapping (SM), random forests (RF), success-rate curves, and the *t*-Student method, were employed to analyze the geochemical anomalies within the intrusive-plutonic Torud-Chahshirin belt (TCB), northeast Iran. In this regard, the success-rate curves were initially applied to extract efficient geochemical signatures. Then, singularity analysis was used on the selected geochemical elements (Au, Cu, Pb, and Zn), which were transformed via centered log-ratio (clr) transformation. In the next step, due to the complexity of the ore-forming processes in the TCB, the structural factors (e.g., fault intersection and faults with different orientations) were determined. Based on the success-rate curves, NE-trending faults and fault density were distinguished as critical structural criteria. Afterward, the RF model as a robust machine learning algorithm was executed on the four efficient SM-based geochemical layers and two efficient structural factors. The anomaly map derived by the RF model (Accuracy = 98.85% and Error = 1.15%) illustrates a very high relationship with Cu $\pm$ Au mineral occurrences. Therefore, the RF algorithm assisted by the singularity method is more trustworthy for highlighting the weak geochemical prospectivity areas in the TCB.

Keywords: local geochemical anomalies; singularity mapping; random forest; *t*-Student value; success-rate curve; Torud-Chahshirin belt

check for updates

Citation: Bigdeli, A.; Maghsoudi, A.; Ghezelbash, R. Recognizing Geochemical Anomalies Associated with Mineral Resources Using Singularity Analysis and Random Forest Models in the Torud-Chahshirin Belt, Northeast Iran. *Minerals* **2023**, *13*, 1399. https://doi.org/10.3390/min13111399

Academic Editor: Yongzhang Zhou

Received: 30 August 2023
Revised: 27 October 2023
Accepted: 30 October 2023
Published: 31 October 2023

1. Introduction

The fundamental task of regional and local exploration geochemistry is to make an accurate difference between background and anomaly classes [1–6]. In this regard, a wide range of robust and efficient methods have been successfully expanded into the few past decades for the delineation of geochemically anomalous regions. Geoscientists initially used classical statistical techniques, namely probability plots and univariate and multivariate analysis, to decompose the background and anomalous components within simple geological media [7,8]. However, the approaches mentioned above disregard the spatial distribution of geochemical datasets and are unable to accurately restrict ore-related geochemical patterns within complicated geological media as well. For this reason, spatial statistical techniques are presently applied to distinguish geochemical anomalies [9,10]. Fractal/multifractal methods [11–14] are able to regard both frequency distribution and spatial variability in geochemical data for the identification of ore-related geochemical targets, and so are more applicable within complex geological environments. Furthermore, fractal/multifractal techniques can enhance the weak geochemical patterns of buried sources. Some of the most prominent tools in this specialized field are concentration-area (C-A) [15], spectrum-area (S-A) [11] and singularity mapping (SM) [9]).

Geochemical anomalies are usually linked to ore-forming processes which have sophisticated systems [9,16]. On the other hand, thermodynamic constituents such as tectonics, ore-bearing fluids, and wall rocks are differentiated in various geological environments. Moreover, most orebody deposits are deeply buried below the earth's surface [17]. Hence, for a better understanding and more specific determination of mineralization-related local geochemical anomalies, robust statistical methods based on machine learning algorithms (MLAs) can be constructive in mineral exploration [18–21]. In recent years, MLAs have displayed prominent capabilities as practical tools in addressing multilateral issues related to pattern recognition and classification [19,22–25]. Artificial neural networks (ANNs) [26], deep autoencoder networks [27], support vector machines (SVMs) [28], self-organizing maps (SOM) [5], and random forests (RF) [20,29] are examples of some influential and well-known MLAs, which have frequently been employed in geoscience applications. Specifically, numerous MLAs are practically attractive subjects, which can executed for revealing geochemical anomalies. For example, the SOM method as an unsupervised MLA has been applied to portray geochemical signatures associated with copper mineralization in the Moalleman district [5], and various MLAs such as RF, SVMs, and Ada Boost classifiers have been employed to detect geochemical anomaly classes in Kuh Panji porphyry-Cu district [30].

In this paper, an integrated methodology including the SM and RF models were employed to analyze stream sediment geochemical data collected from the Torud-Chahshirin belt (TCB) to discover Cu $\pm$ Au geochemical anomalies. For this, success-rate curves were initially used to specify efficient geochemical indicators of related to mineralization in the study area. Following the outcomes of the success-rate curves, singularity values of Au, Cu, Pb, and Zn elements were used to delineate the weak geochemical anomalies.

Finally, according to the complex and multiform geological settings present in the study area, the RF model as a robust machine learning algorithm was conducted on the singularity values of efficient geochemical signatures and the mineralization-related structural factors (fault density and distance to NE-trending faults) for precise mapping of mineral prospectivity within the study area.

2. Study Area and Dataset Used

The Torud-Chah Shirin volcano-intrusive belt (TCB), which is located in the south of the Kavir-e-Chah Jam depression (southwestern of Damghan), covers an area of approximately 5000 km^2. It is a part of the Central Iran Magmatic Zone [31] (Figure 1). The existence of diverse mineral deposits/occurrences, particularly base metal veins of epithermal origin, has increased the economic importance of this magmatic belt for economic geology researches [32]. The N-E trending TCB acts as an uplifted block between the Torud and Anjilow regional faults. Moreover, the Tertiary calc-alkaline volcanism and ore-related mineralization took place along the major Anjilow fault in the north and Torud fault in the south. Thus, tectonic processes, especially faults and fractures, are the primary controllers of ore-forming systems in this belt [33] (Figure 2).

The oldest rock units in the TCB include green schist, metamorphosed dolomite, and limestone [34]. Furthermore, the exposed rocks in this complex consist of Eocene volcanic-pyroclastic assemblages, which are underlain by Oligocene intrusive bodies with granodiorite composition. Principal magmatic commodities are formed of andesite and basalt, which over time have tended to be acidic and trachyte. Eventually, they become more alkaline and turn into lava flows, breccias, and andesite tuffs that probably extend to the late Eocene [35]. The volcanic rocks of TCB have been intersected by several effusive rocks, which are of the upper Eocene–Oligocene age. In addition, lithological composition includes granite, micro-granite, granodiorite, micro-granodiorite, micro-quartz-monzonite, micro-monzonite, micro-monzodiorite and micro-quartz-monzodiorite. These rocks have magmatic series, subalkaline to alkaline, and are of type I [36,37]. In addition, these intrusive bodies are considered as one of the crucial factors in the mineralization process in the TCB. There are several significant mineral deposits/occurrences as well as abandoned mines in the TCB district, namely Cheshmeh Hafez (Pb + Zn + Cu), Chah Messi (Cu ± Au), Gandy (Au-Ag ± Pb-Zn), Abolhassani (Pb-Zn ± Ag ± Au), Chalo (Cu ± Au), Darestan (Cu ± Au), Baghu (Au ± Cu), and Reshm (Pb-Zn ± Ag) [38].

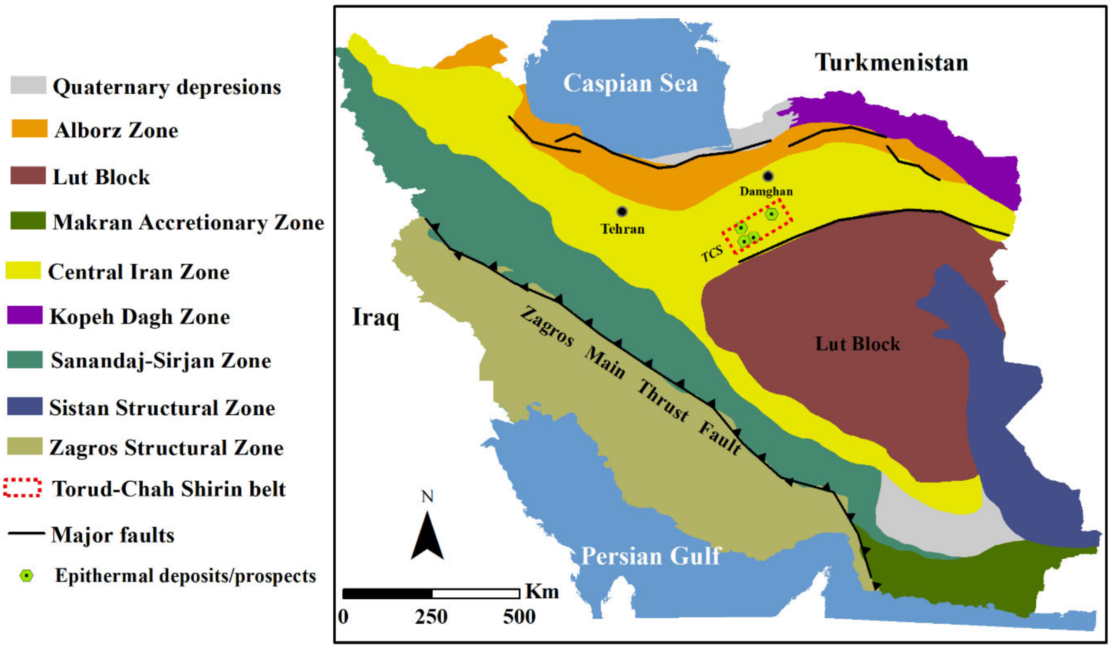

Figure 1. Location of the Torud-Chahshirin volcano-intrusive complex in NE Iran.

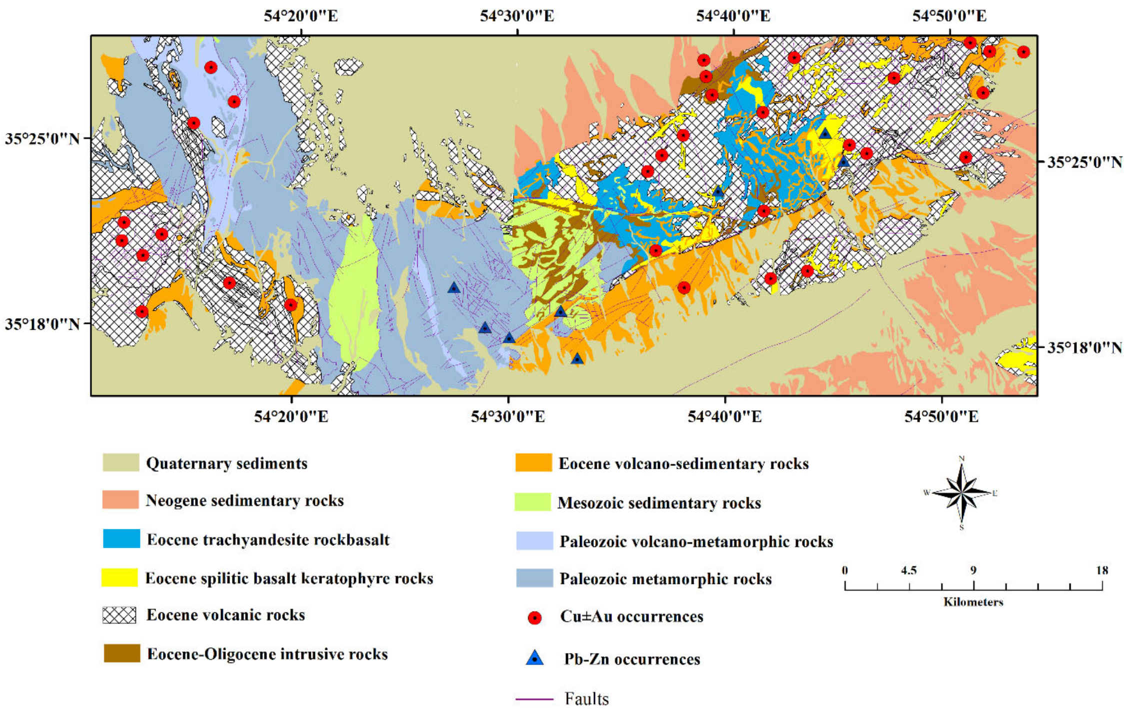

Figure 2. Simplified geological map of TCB [31].

Veins, veinlets, and disseminated mineralization mainly occur in the late Eocene–Oligocene within TCB [32,39]. As a good case in point, mineralization in the Darestan deposit happens in NW-trending veins and is hosted by Eocene volcanic and volcaniclastic rocks. Moreover, a mineral assemblage that includes chalcopyrite, pyrite, minor chalcocite, tetrahedrite, and covellite indicates a high sulfidation epithermal system in the Darestan prospect. More details about other mineral occurrences can be found in TaleFazel et al. (2019) [32].

The stream sediment geochemical survey for mineral exploration in the TCB was conducted by the Jianxi Company in 1993. A systematic net of 1625 composite stream sediment samples, with sampling nets of 1400 m × 1400 m, was designed. Then, subsamples (up to four) were collected from the first- or second-order streams within each cell (Figure 3). Afterward, a sample was allotted to the center of each cell as a representative of collected subsamples. In this case, these composite stream sediment samples can acceptably provide information relevant not only to the upstream sources of the samples but also to the immediate proximity of the sample locations.

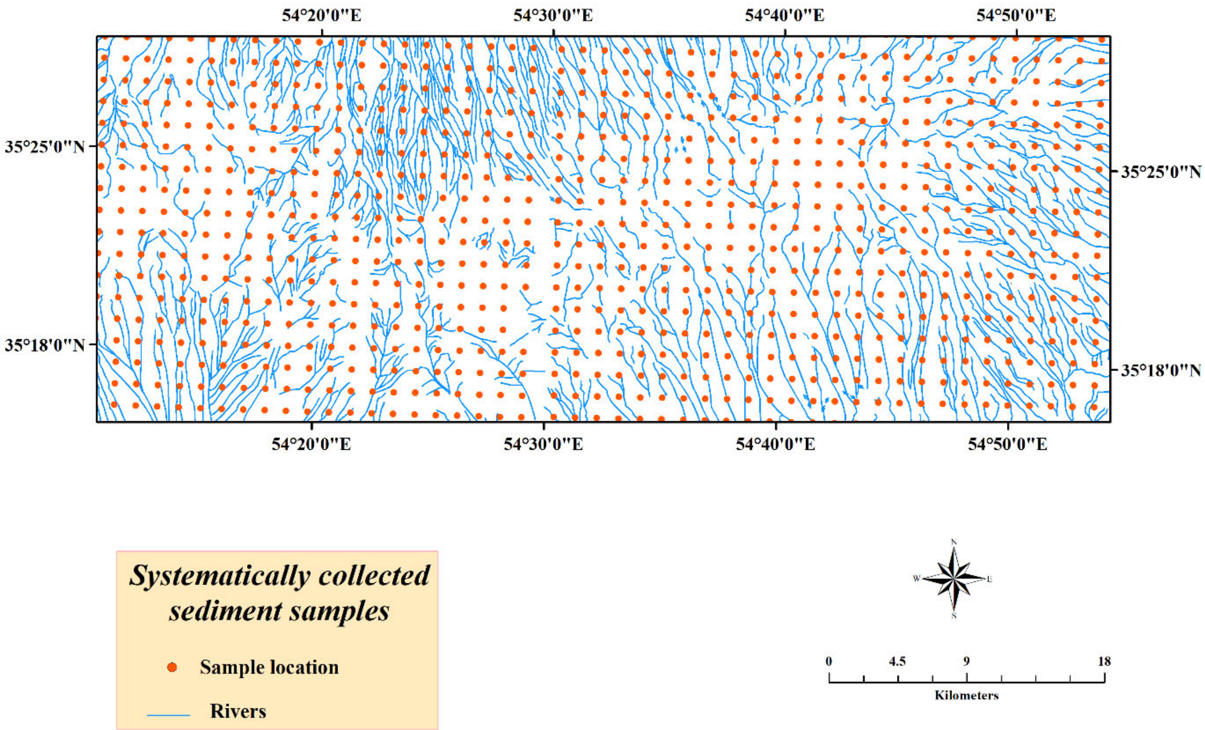

Figure 3. Sampling sites, which are systematically collected from the study area.

In the next step, each representative sample was analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES) for principal and trace geochemical elements, excluding Au, for which fire assay (FA) as a preconcentration technique was applied. Next, the resultant solution was analyzed by atomic absorption spectrometry (AAS). Finally, the concentration contents of six indicator elements (Au, As, Cu, Pb, Sb, and Zn), which are spatially connected to the vein-type epithermal deposits in the TCB [32], were utilized for geochemical investigations.

3. Methodology

3.1. Singularity Mapping

From a geological perspective, hydrothermal mineralization is the result of mineral accumulation in a limitary period of geological time, which usually happens within small amounts of orebodies [1,9]. The idea behind the singularity may be similar to ore-forming processes, which are about releasing enormous energy or gathering material within a spatial–temporal range. Hence, the quantitative measuring of singularity can be a crucial tool for discovering hidden and local geochemical anomalies [9,40,41].

From a multifractal viewpoint, singularity can be expressed through power-law relations. In a 2D model, if the amount of metal within an area (A) is $\mu(A)$, and the average concentration of the metal is $C(A)$, it can therefore be described as $C(A) = \frac{\mu(A)}{A}$. According to relation $C(A)$, $\mu(A)$ and A have a direct relationship (e.g., the amount of metal $\mu(A)$ increases as the area A increases, and vice versa). However, the values of $C(A)$ totally depend on ore-forming processes, and with the variation of area, $C(A)$ can go up, go down, or stay constant. The power-law relationships between these variables are [12,42]:

$$\mu(A) = cA^{\frac{\alpha}{2}} \tag{1}$$

$$C(A) = cA^{\frac{\alpha}{2}-1} \tag{2}$$

where α indicates the singularity index, and c is constant. α can be achieved via the values of average metal concentration, which are measured for several sizes of area by the least-squares method. There are three different conditions based on the α values in 2D modeling: (1) $\alpha < 2$ denotes abnormal enrichment of the element in a specific location, (2) $\alpha > 2$ displays an elemental depletion in that situation, and (3) $\alpha = 2$ illustrates a monofractal behavior in a certain location [9,43].

To calculate the α value for every location of the study area and display the results as a geochemical map, a singularity mapping technique based on moving average window can be applied. Firstly, a group of sliding windows $A(r)$ (usually a square shape) with diverse sizes is defined $(r_i \times r_i)$. In the next step, the algorithm measures the values of average concentration $C[A(r_i)]$ per window, and finally, the α index for each location is estimated through the below equation [9]:

$$\log C[A(r_i)] = C + (\alpha - 2)\log(r_i) \tag{3}$$

The value of $(\alpha - 2)$ is considered as the slope from Equation (3). It can be concluded that the singularity mapping approach is a robustified form of the concentration-area fractal technique [9]. Supplementary descriptions of the phenomenon of the singularity, and also the singularity mapping method, can be found in Cheng (2007) [9].

3.2. Random Forest

Random forest (RF) is one of the ensemble techniques that employs numerous decision trees (DTs) for classification and regression problems [44]. For the training of DTs, RF implements a bootstrapping method [45], which is a resampling technique. It means applying random samples by replacing them with labeled data—values obtained from a group of exploratory criteria at places of presence or absence of mineral occurrences. In other words, in the bootstrapping procedure, about 75% of input data are involved in the training procedure, and the remaining is utilized for calculating the error of generalized trees (out-of-bag error) [20,46]. Each DT is trained by a random sample of the original data, and it also separates the target variable into dichotomous classes. The RF algorithm starts to purify the child nodes from the root nodes via the splitting processes based on multiple input variables [44]. This procedure repeatedly goes on until a predetermined stop criterion is reached. Eventually, the regression model of RF is gained by averaging the outcomes of several DTs [18,29,47].

The outstanding advantage of RF as a robustified version of DTs among the MLAs is random resampling and replacement, which cause several training subsets to be applied to expand DTs, therefore diminishing correlations between trees as well as improving the prediction accuracy by reducing the generalization error [25,44].

More importantly, the RF algorithm is a user-friendly and well-performing machine learning algorithm due to few parameters. There are two influential parameters, namely the ntree and mtry, which have the most considerable effect on the random forest model. (a) ntree is the number of trees to grow, and (b) mtry is the number of variables to be assigned to each node [47]. Selecting the huge number of DTs (ntree) usually raises the model accuracy, but too large values of mtry variable may lead to underfitting. In this regard, the 'tuneRF' function in the 'randomForest' package was utilized to attain the optimum values for learning parameters (e.g., ntree and mtry) and also to decrease the

OOB error. Therefore, the RF model with 215 as the number of trees and with 2 as the number of splits was operated.

To measure the degree of capability of the RF model in the classification procedure, a confusion matrix was implemented [10]. The confusion matrix is comprised of four basic parameters, namely: (1) *TP* (true positive) represents the prospect sites which are correctly predicted, (2) *TN* (true negative) illustrates that the model correctly predicted the non-prospect sites, (3) *FP* (false positive) and (4) *FN* (false negative) illustrate the total number of prospect and non-prospect sites, respectively, which are incorrectly predicted. The total accuracy (*TA*) of the RF model is defined as the ratio of correctly predicted prospect and non-prospect sites to the total number of predicted sites as follows:

$$TA = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

3.3. Success-Rate Curves

Selecting the efficient exploration criteria (e.g., geochemical signatures or controlling geological factors) is very applicable because the high-favorable areas can be recognized with low uncertainty. In other words, by eliminating the inefficient criteria in modeling mineralized zones, trustworthy results can be gained [48,49]. To do this, the success-rate curves were drawn to reveal the efficiency of evidence maps. Furthermore, these curves can be used to assess the performance of different models linked to mineral deposits. To plot a success-rate curve, the x-axis is defined as a section of the study area which is classified as favorable (P_a), and a section of mineral deposits which is estimated correctly (P_o) is displayed on the y-axis [50,51]. In addition, there is a gauge line in the success-rate curve which characterizes the success or failure of each evidence map [28,51]. In this manner, if an evidence layer is located above the gauge line, it demonstrates goodness-of-fit with known mineral deposits in the study area. Conversely, an evidence layer represents a poor relationship with known mineral deposits when it lies under the gauge line in the success-rate curve.

4. Results and Discussion

4.1. Preprocessing of Selected Elements

The first step for applying the different statistical methods in order to reveal geochemical anomalies is to define the type of distribution (e.g., normal or lognormal) of the chemical elements [46,52]. In this regard, the basic statistical properties such as standard deviation, skewness, and kurtosis of raw data belonging to mineralization-related elements were calculated, as shown in Table 1. The outcomes (based on the skewness parameter) exemplify that the values of geochemical elements do not conform to the normal tendency. On the other hand, as Reimann et al. (2002) [53] and Filzmoser et al. (2009) [54] pointed out, the geochemical data of stream sediments are expressed as closed-number systems, which means they have a compositional nature. Hence, to address the closure problem and to modulate the minimum and maximum values of geochemical data, log-ratio transformations have been carried out [55]. In this regard, the original data of six elements, including Au, As, Cu, Pb, Sb, and Zn, have been subjected to the clr-transformation [14,48,55,56]. The clr transformation is appropriate for opening compositional data in statistical analyses because of symmetric results [55] and the feasibility of interpreting resulting values [54].

Aiming to evaluate the normality of clr-transformed data, the Q-Q plots were plotted (Figure 4) and revealed that there were some outliers in the clr-transformed data. It can be deduced that the concentrations of mineralization-related elements display fractal/multifractal distributions with multiple geochemical populations due to complicated geological processes in the TCB [57–59].

Table 1. Statistical parameters of mineralization-related geochemical elements.

	As	Au	Cu	Pb	Sb	Zn
N. of samples, valid	1625	1625	1625	1625	1625	1625
Maximum	85.4	604	1353	6637	24	5171
Minimum	4.3	0.3	10.71	8.86	0.21	36.63
Std. deviation	6.19	19.18	51.6	249.87	1.18	198.76
Skewness	4.6	25.98	15.37	18.83	10.9	18.25
Kurtosis	36.67	737.44	327.22	412.11	162.79	377.74

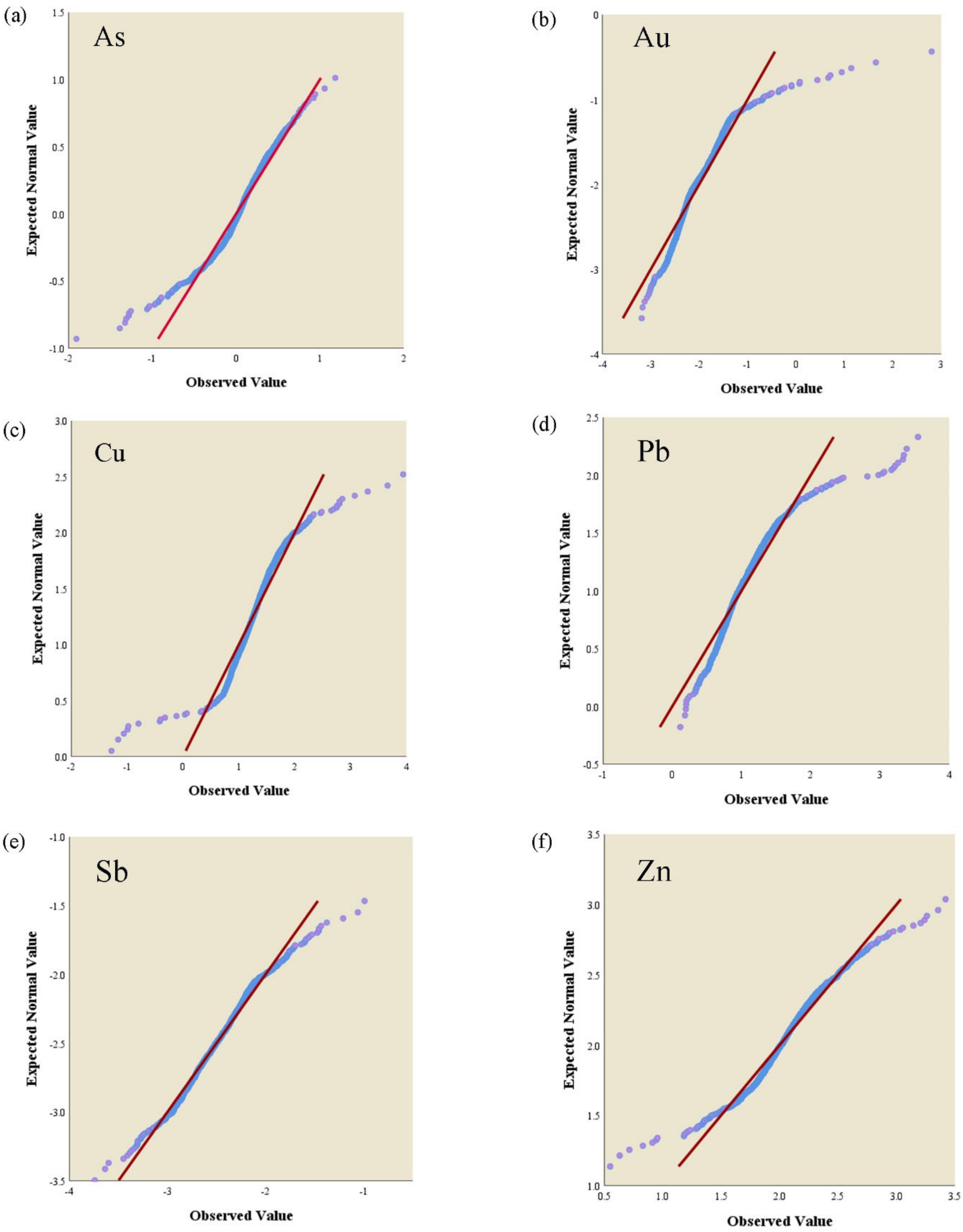

Figure 4. Q-Q plots of clr-transformed elements: (**a**) As, (**b**) Au, (**c**) Cu, (**d**) Pb, (**e**) Sb, and (**f**) Zn.

4.2. Extracting the Most Efficient of Single-Element Geochemical Footprints

From a geochemical point of view, identifying the geochemical anomalies in the stream sediment dataset is vitally relevant to the ability of each single element to discover target areas [1,6,27,60,61].

For this purpose, the success-rate curves were drawn for the determination of efficient clr-transformed geochemical elements (Figure 5a–f). It can be seen that the geochemical values of Cu and Au represent a robust positive relationship with the known mineral occurrences in the TCB, as their success-rate curves lie beyond the gauge line (Figure 5b,c). The Pb and Zn are classified as moderate efficient geochemical elements based on their success-rate curves (Figure 5d,f) as their anomaly classes do not display an appropriate relationship with Cu-Au deposits. Also, Sb (Figure 5e) is the least efficient geochemical element because its success-rate curve lies thoroughly on the diagonal line. Finally, the single-element value of As did not display spatial relations with the locations of mineral deposits in the study area since its success-rate curve is placed entirely under the gauge line (Figure 5a).

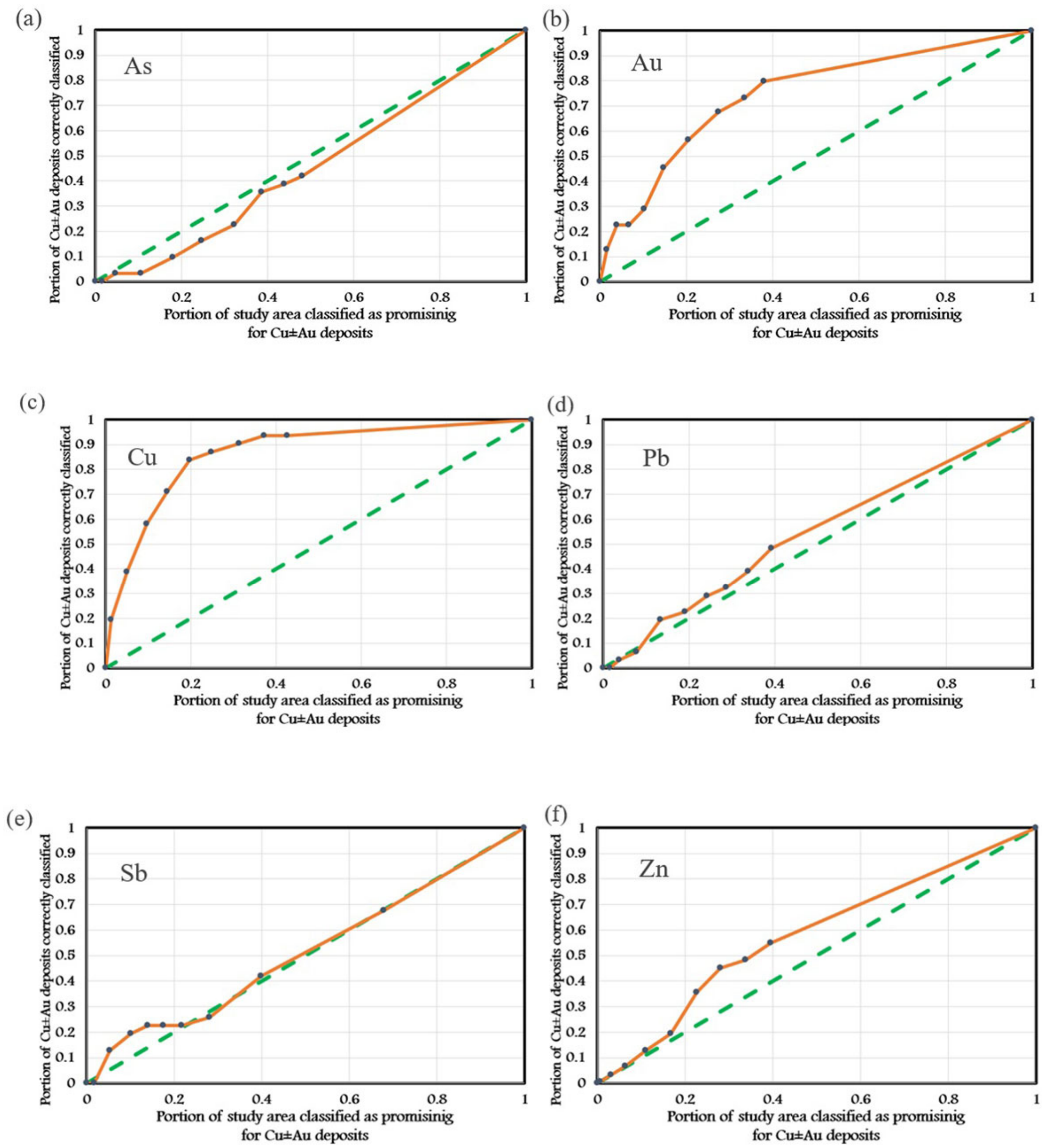

Figure 5. Success-rate curves for clr-transformed values of six indicator elements: (**a**) As, (**b**) Au, (**c**) Cu, (**d**) Pb, (**e**) Sb, and (**f**) Zn.

4.3. Singularity Mapping for Detecting the Local Geochemical Anomalies

The challenging issue of the present research is to identify local and complex geochemical anomalies through stream sediment geochemical data. After recognizing the efficient (most, moderate, and least) and inefficient geochemical elements linked to mineralization sought in the TCB, SM was executed on the clr-transformed geochemical values of Au, Cu, Pb, and Zn in order to distinguish the local anomalies in the study area (Figure 6). However, there was a problem in the application of SM on the clr-transformed geochemical dataset that consisted of negative values, leading to erroneous results. To address this, a positive shift was initially applied to the clr-transformed values of selected elements, aiming at converting them into a positive domain to be used for singularity analysis.

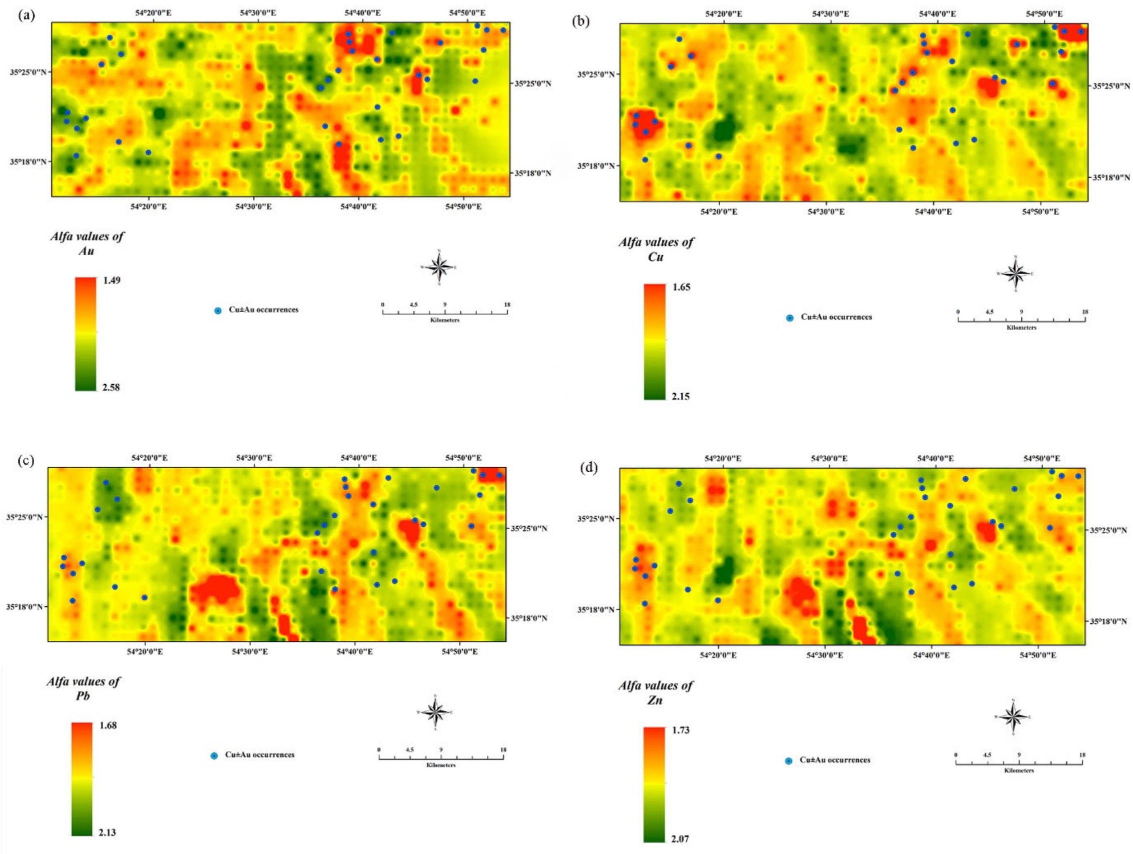

Figure 6. IDW maps of α values related to: (**a**) Au, (**b**) Cu, (**c**) Pb, and (**d**) Zn.

For applying the SM algorithm, firstly, we selected an appropriate cell size (300 m × 300 m) for gridding the clr-transformed geochemical elements scores into a raster map based on the structure of sampling (instead of density of samples) as described by Zuo (2012) [62] and Shuguang et al. (2015) [40]. Next, a group of square sliding windows with different sizes (300 m × 300 m, 600 m × 600 m, . . . , 1500 m × 1500 m) were set. These sizes were calculated by multiples of the first cell size. Afterward, the average concentration of elements with predetermined cell size was computed per window. Then, the singularity indices were attained by applying the least-squares method duo to Equation (3). Eventually, to portray the values of the singularity indices, the IDW was carried out (Figure 6). The SM algorithm was implemented in the MATLAB R2031a program.

4.4. Delineating of Efficient Structural Controlling Factors

The fault systems are the product of stress differences within geological settings. It means that a sudden fall in geostress can form faults and fractures, which are the favorite places for migrating and precipitating hydrothermal fluids [17,63–65]. Moreover, these pathways and their intersection points control important factors such as temperature, pressure, and permeability, which are connected to mineralization processes [17,66,67]. Also, structural controlling factors (e.g., faults and fractures) facilitate the passage of hydrothermal fluids [16,65].

On the other hand, ore-related mineralization processes and the evolution of Tertiary volcanism in the TCS belt are directly influenced by the fault systems. Therefore, the identification of different orientations of faults could be appropriate evidence for specific mineralization. In this regard, four major trends comprising N-trending, E-trending, NE-trending, and NW-trending were extracted as tectonic controls (Figure 7).

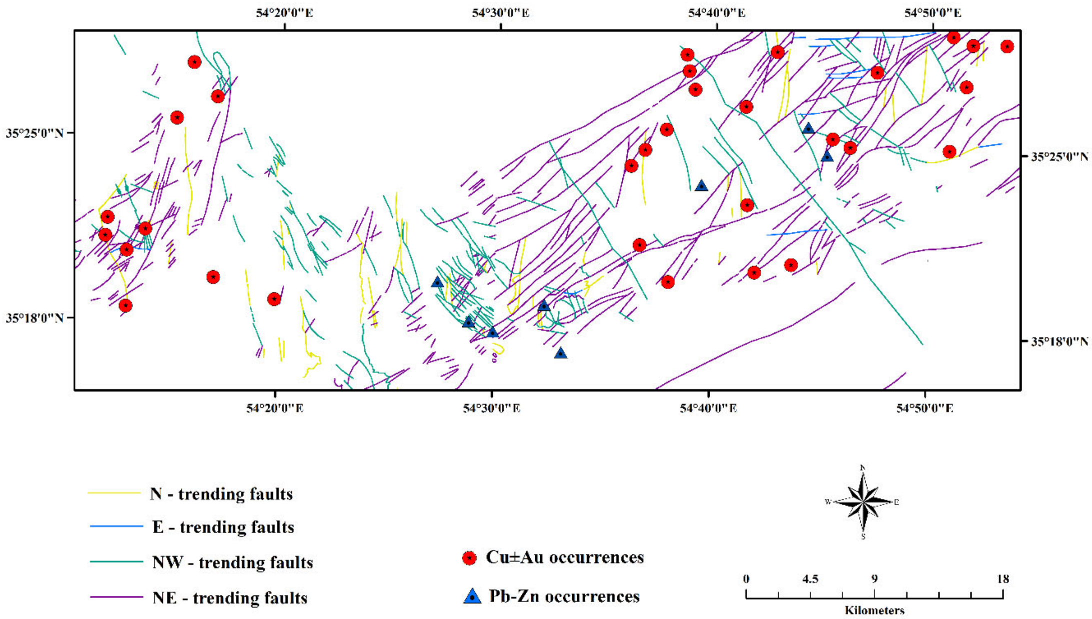

Figure 7. Relations between different geospatial trending faults with known mineral occurrences in the TCB.

In order to specify the relations between structural factors and mineral occurrences, we generated the relevant maps of distance to the faults by Euclidean distance order in ArcMap 10.3 software. For each cell, the Euclidean distance to the closest source was measured. In addition, the maximum distance was set to the edge of the output raster (Figure 8a–d). In addition, the density map of faults as another tectonic control of ore-forming processes in the TCB belt was generated (Figure 8e).

Finally, due to the outcomes of success-rate curves, the maps of distance to NE-trending faults and fault density were recognized as the efficient structural layers (Figure 9b,e). The outcomes of success-rate curves are in compliance with previous studies in the TCB. For instance, Hushmandzadeh et al. (1978) [34] stated that the Eocene–Oligocene volcanism in the TCS district may have formed along an NE-trending fault segment. Furthermore, some prominent deposits (e.g., Gandy, Baghu, and Abolhassani) in the TCS belt are mainly formed parallel to the NE alignment [32]. So, these faults (NE-trending) are directly connected to ore-forming processes in the TCS belt.

The remaining directions (e.g., N-trending, E-trending, and NW-trending) were considered as inefficient structural factors (Figure 9a,c,d).

4.5. Random Forest Model

Aiming at extracting the geospatial relations between geochemical anomalies and Cu $\pm$ Au mineral occurrences, the RF algorithm was applied. For this purpose, four efficient SM-based geochemical layers (the alpha values of Au, Cu, Pb, and Zn), plus two significant structural factors (distance to NE-trending faults and fault density) were utilized as inputs for creating a predictive model of Cu-Au prospectivity. For this purpose, a total of 65,583 pixel values were extracted from six evidence maps in order to obtain test and training data. In this regard, a training dataset was generated based on the prospect and non-prospect locations. These locations should be chosen by the following rules [18,49]:

(1) The number of non-prospect sites must be equal to the number of prospect sites to improve the accuracy of models (here, prospect sites = 31).
(2) In order to differentiate the multiattribute dataset, the non-prospect sites should be selected far from the prospect sites (the selection of the non-prospect locations is conducted based on the point pattern analysis introduced by Carranza et al., 2008 [1], which demonstrated that the non-deposit sites must be selected at least 3 km away from the deposit sites).
(3) Unlike the prospect sites, which have been regularly clustered, the non-prospect sites should have an arbitrary nature.

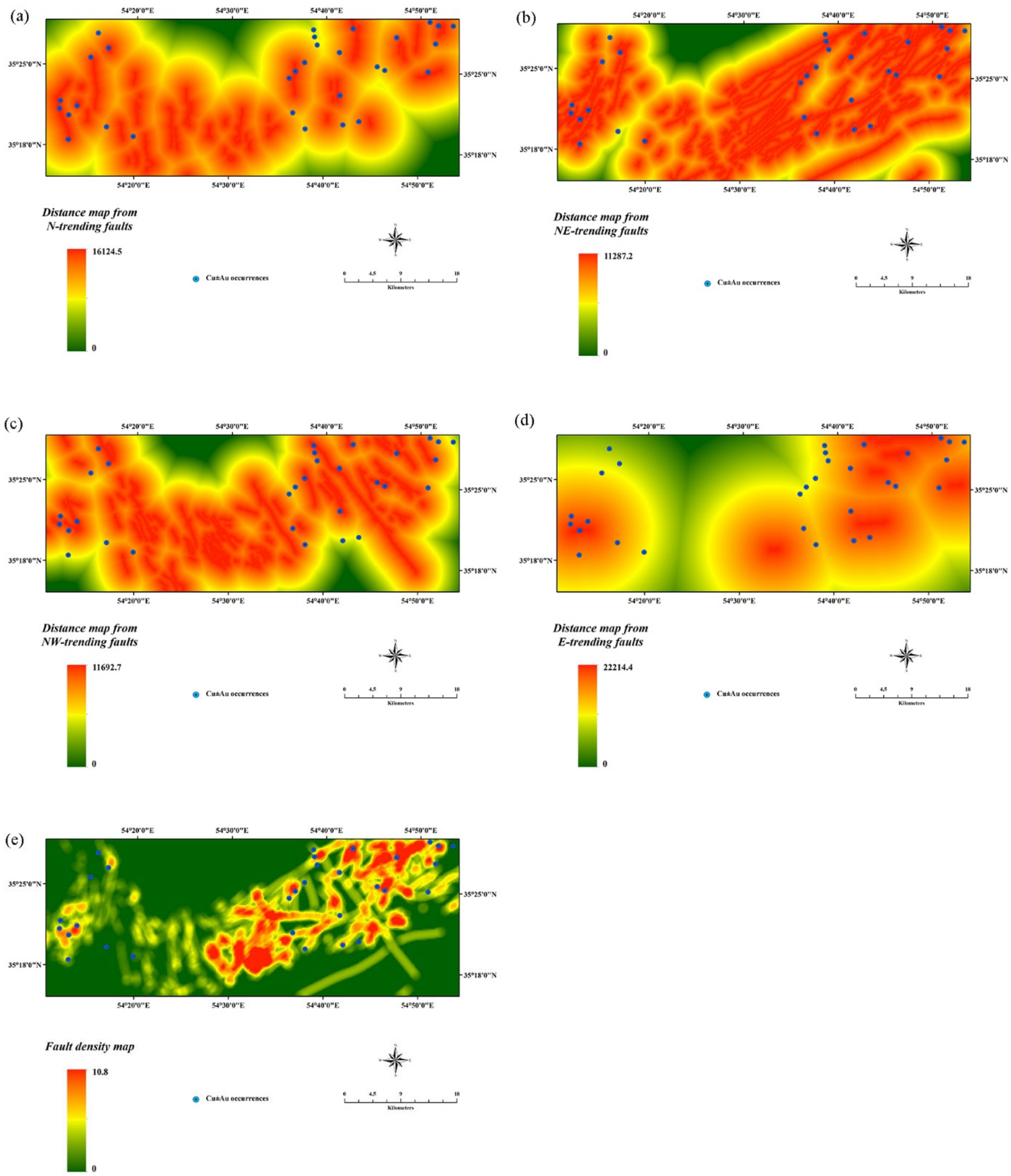

Figure 8. Distance maps from different trending faults: (**a**) N-trending, (**b**) NE-trending, (**c**) NW-trending, (**d**) E-trending, and (**e**) fault density map.

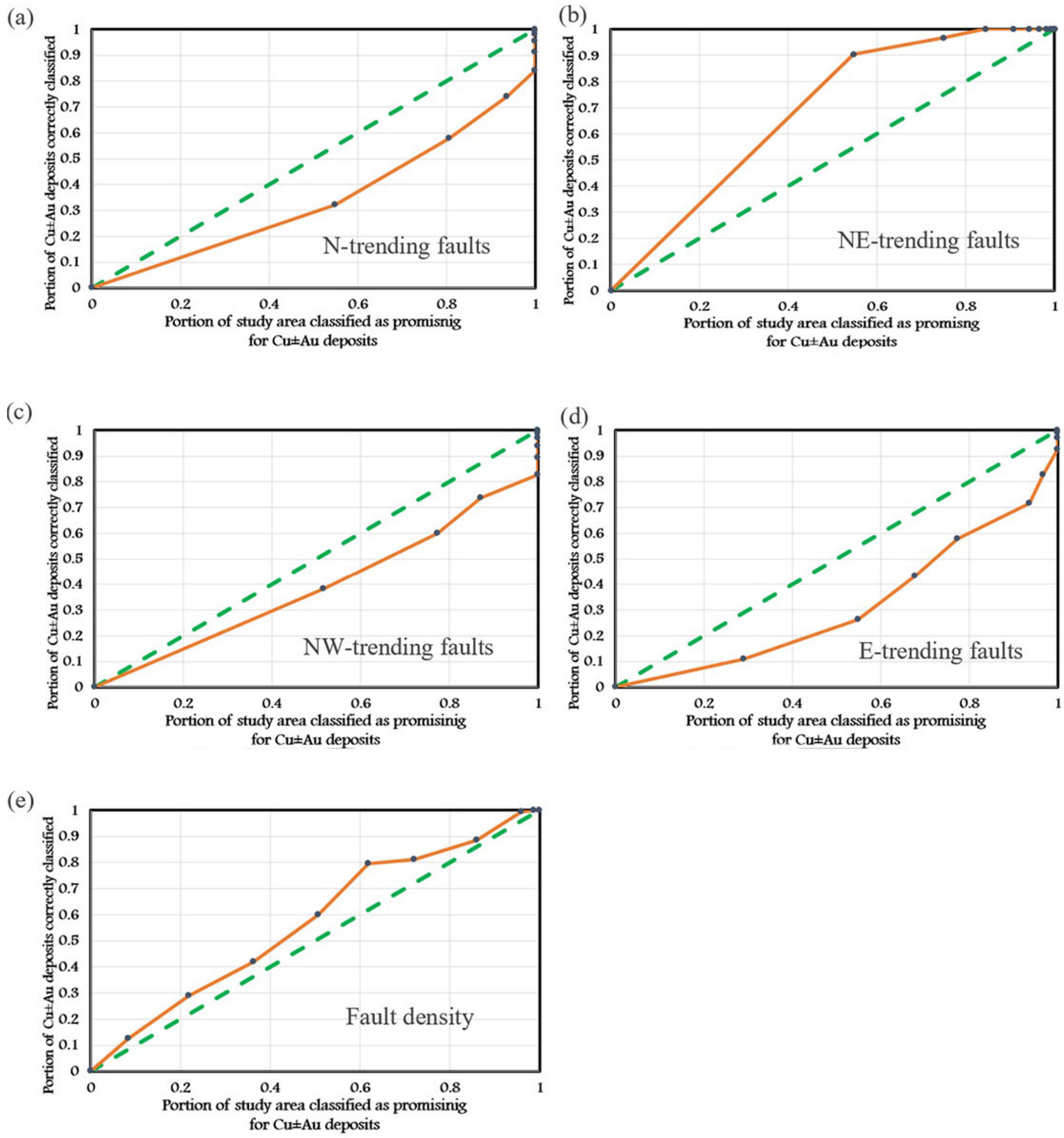

Figure 9. Success-rate curves for different trending faults: (**a**) N-trending, (**b**) NE-trending, (**c**) NW-tending, (**d**) E-trending, and (**e**) fault density map.

Accordingly, the training dataset has been produced (e.g., 2561 pixel values), in which the target variable scores were labeled 1 as prospect and 0 as non-prospect. Then, about three-fourths of the total values (1920 pixels) were chosen to train the RF model and the residue (641 pixels), which were applied as out-of-bag samples (OOB) and were not used for the training procedure. The OOB samples were used to evaluate the accuracy of the trained model. The RF algorithm was carried out using the RStudio-2023.06.0+421.pro1 software.

The OOB error was 1.15%, and as a consequence, the accuracy of the RF model was 98.85% (Table 2), which demonstrate that the algorithm is appropriately operated. Furthermore, based on the confusion matrix, the accuracy of the classification of deposit and non-deposit sites for RF modeling is 98.87% (error = 1.13%) and 96.97% (error = 3.03%), respectively.

Finally, as depicted in Figure 10, the RF model revealed a terrific relationship between high-favorable areas and areas consisting of known Cu $\pm$ Au occurrences.

Table 2. The classification accuracy and error of different RF models for OBB data based on confusion matrix.

	RF via Singularity Values
OBB error	1.15%
Model accuracy	98.85%
Error for the classification of deposit sites	1.13%
Error for the classification of non-deposit sites	3.03%

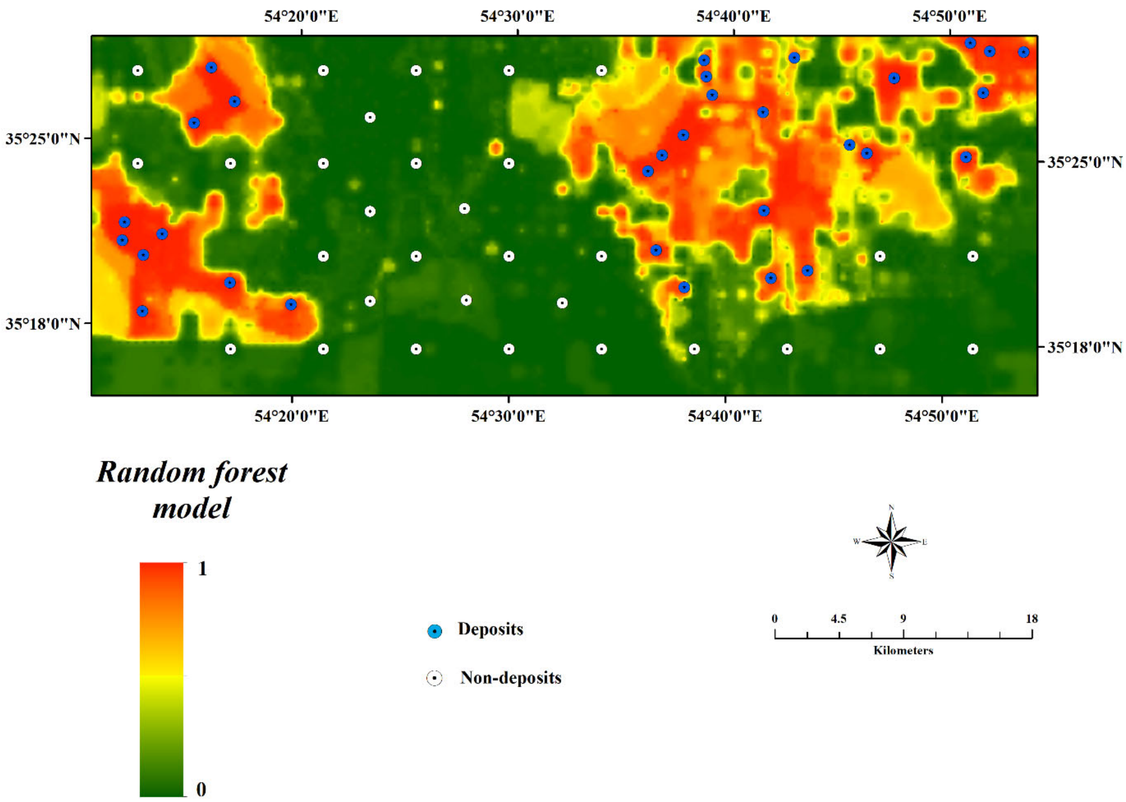

Figure 10. Predictive maps for delineating the potential mineralization derived by RF model using singularity analysis.

4.6. Assessment of Prospectivity Model

In this research, the success-rate curve was applied to measure the robustness of the generated model for delineating derived geochemical anomaly targets in the TCB. In this regard, to upgrade the precision of the analysis, 5-percentile intervals were adopted to create the success-rate curve of the RF model.

As can be seen in Figure 11, the success-rate curve of the RF model has appeared beyond the gauge line, and this means that the predictive model has successfully obtained a positive relationship with the known Cu $\pm$ Au deposits/occurrences.

4.7. Delineation of Exploration Targets by Applying Student's t Method

To specify the appropriate threshold values for separating geochemical anomalies related to the RF model, the *t*-Student spatial statistic was utilized. Student's *t*-value, based on the weights-of-evidence method, measures the spatial association between geochemical values and known mineral occurrences [9,50,68]. Further explanation about the *t*-Student method could be found in Bonham-Carter, 1994 [50]. Basically, the higher the positive values of *t*, the more robust positive relations between known mineral deposits and geochemical anomalies. In this regard, to decompose a continuous-value singularity-based RF map into dichotomous meaningful classes, the highest Student's *t*-value (3.95) (Figure 12a) was selected as an appropriate threshold (0.8) for determining the geochemical anomalies. As shown in Figure 12b, the geochemical targets located in the predictive map of singularity-based RF are spatially correlated with a majority of known mineral deposits in the TCB (Figure 12b).

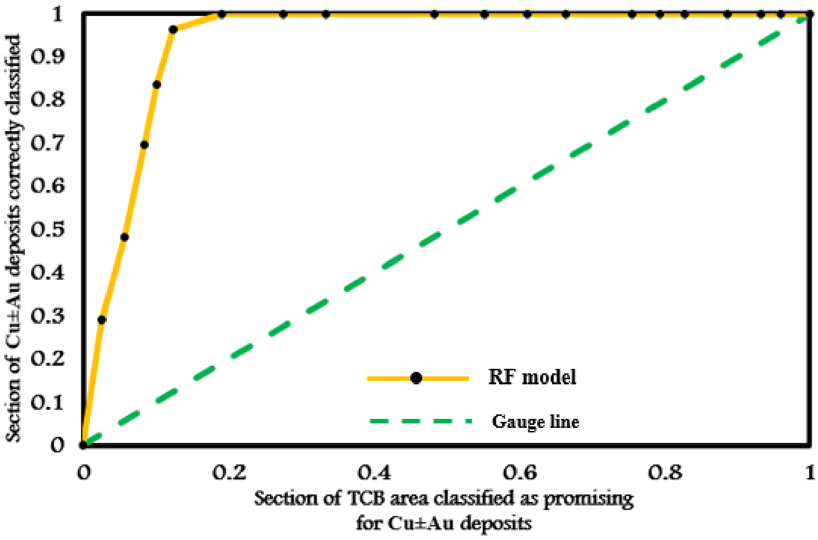

Figure 11. Success-rate curve of generated RF predictive model for identifying Cu ± Au deposits.

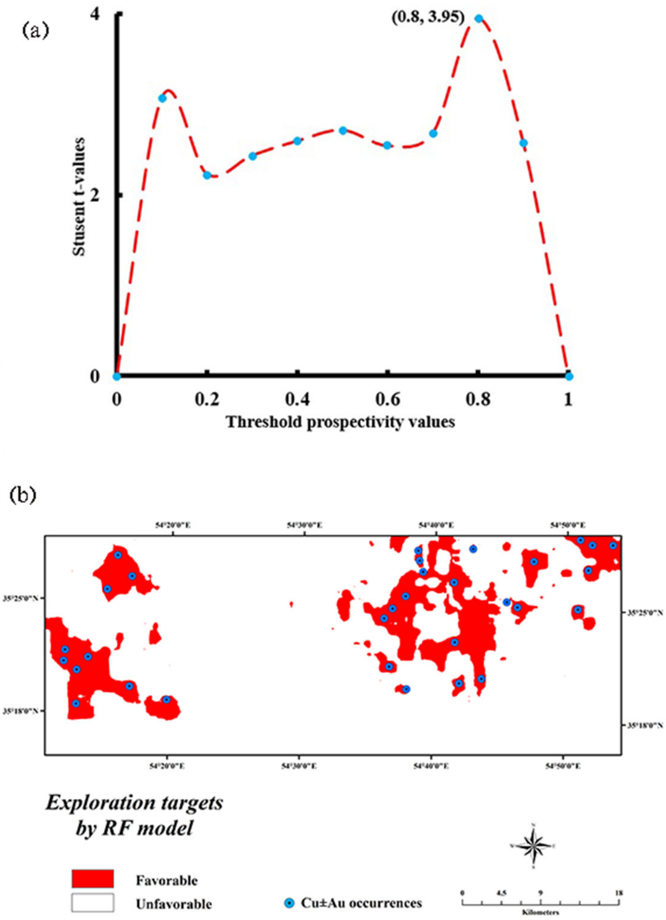

Figure 12. (**a**) Student's *t*-values measured for discretized the predictive model. (**b**) The dichotomous singularity-based RF model illustrates favorable targets derived by the Student's *t* method.

5. Conclusions

In the present study, the RF model with efficient SM-based geochemical layers and critical structural criteria was implemented, aiming to delineate high-favorable geochemical anomalies linked to epithermal Cu ± Au mineralization in the TCB. The substantial outcomes of this research are listed below:

(1) The singularity mapping as a filtering method is significantly successful in outlining the high-favorable geochemical anomalies in the TCB.
(2) The prediction of high-favorable mineralized areas with highly accurate classification (Accuracy = 98.85%) indicates the robustness and effectiveness of the RF model in portraying geochemical anomalies.
(3) The prospectivity model of RF (derived from *t*-Student method) represents the geochemical anomalous areas closely coinciding with Cu ± Au mineral deposits/occurrences.

Author Contributions: Conceptualization, A.B.; methodology, A.B.; software, A.B.; validation, A.B.; resources, A.M.; writing—original draft, A.B.; writing—review and editing, R.G.; supervision, A.M. and R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The regional stream-sediment geochemical data were obtained from the Geological Survey of Iran (GSI).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carranza, E.J.M. *Geochemical Anomaly and Mineral Prospectivity Mapping in GIS*; Elsevier: Amsterdam, The Netherlands, 2008.
2. Zuo, R.; Xia, Q. Application fractal and multifractal methods to mapping prospectivity for metamorphosed sedimentary iron deposits using stream sediment geochemical data in eastern Hebei province, China. *Geochim. Cosmochim. Acta Suppl.* **2009**, *73*, A1540.
3. Grunsky, E.C. The interpretation of geochemical survey data. *Geochem. Explor. Environ. Anal.* **2010**, *10*, 27–74. [CrossRef]
4. Yousefi, M.; Kamkar-Rouhani, A.; Carranza, E.J.M. Geochemical mineralization probability index (GMPI): A new approach to generate enhanced stream sediment geochemical evidential map for increasing probability of success in mineral potential mapping. *J. Geochem. Explor.* **2012**, *115*, 24–35. [CrossRef]
5. Bigdeli, A.; Maghsoudi, A.; Ghezelbash, R. Application of self-organizing map (SOM) and K-means clustering algorithms for portraying geochemical anomaly patterns in Moalleman district, NE Iran. *J. Geochem. Explor.* **2022**, *233*, 106923. [CrossRef]
6. Ghezelbash, R.; Maghsoudi, A.; Carranza, E.J.M. Mapping of single-and multi-element geochemical indicators based on catchment basin analysis: Application of fractal method and unsupervised clustering models. *J. Geochem. Explor.* **2019**, *199*, 90–104. [CrossRef]
7. Stanley, C.R.; Sinclair, A.J. Comparison of probability plots and the gap statistic in the selection of thresholds for exploration geochemistry data. *J. Geochem. Explor.* **1989**, *32*, 355–357. [CrossRef]
8. Luz, F.; Mateus, A.; Matos, J.X.; Goncalves, M.A. Cu-and Zn-soil anomalies in the NE border of the South Portuguese Zone (Iberian Variscides, Portugal) identified by multifractal and geostatistical analyses. *Nat. Resour. Res.* **2014**, *23*, 195–215. [CrossRef]
9. Cheng, Q. Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China. *Ore Geol. Rev.* **2007**, *32*, 314–324. [CrossRef]
10. Carranza, E.J.M. Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. *J. Geochem. Explor.* **2011**, *110*, 167–185. [CrossRef]
11. Cheng, Q.; Xu, Y.; Grunsky, E. Integrated spatial and spectral analysis for geochemical anomaly separation. In Proceedings of the Fifth Annual Conference of the International Association for Mathematical Geology, Trondheim, Norway, 6–11 August 1999.
12. Zuo, R.; Cheng, Q.; Agterberg, F.P.; Xia, Q. Application of singularity mapping technique to identify local anomalies using stream sediment geochemical data, a case study from Gangdese, Tibet, western China. *J. Geochem. Explor.* **2009**, *101*, 225–235. [CrossRef]
13. Ghezelbash, R.; Maghsoudi, A.; Daviran, M. Combination of multifractal geostatistical interpolation and spectrum–area (S–A) fractal model for Cu–Au geochemical prospects in Feizabad district, NE Iran. *Arab. J. Geosci.* **2019**, *12*, 152. [CrossRef]
14. Liu, Y.; Cheng, Q.; Zhou, K. New insights into element distribution patterns in geochemistry: A perspective from fractal density. *Nat. Resour. Res.* **2019**, *28*, 5–29. [CrossRef]
15. Cheng, Q.; Harris, J. GIS-based multifractal anomaly analysis for prediction of mineralization and mineral deposits. In *GIS Applications in Earth Sciences, Geological Association of Canada Special Paper*; Geological Association of Canada: St. John's, NL, Canada, 2006; pp. 289–300.
16. Yu, C. Complexity of earth systems—Fundamental issues of earth sciences (I). *J. China Univ. Geosci.* **2002**, *27*, 509–519.
17. Cox, S.; Etheridge, M.; Wall, V. The role of fluids in syntectonic mass transport, and the localization of metamorphic vein-type ore deposists. *Ore Geol. Rev.* **1987**, *2*, 65–86. [CrossRef]

18. Carranza, E.J.M.; Laborte, A.G. Data-driven predictive modeling of mineral prospectivity using random forests: A case study in Catanduanes Island (Philippines). *Nat. Resour. Res.* **2016**, *25*, 35–50. [CrossRef]
19. Yousefi, M.; Kreuzer, O.P.; Nykänen, V.; Hronsky, J.M. Exploration information systems—A proposal for the future use of GIS in mineral exploration targeting. *Ore Geol. Rev.* **2019**, *111*, 103005. [CrossRef]
20. Daviran, M.; Maghsoudi, A.; Ghezelbash, R.; Pradhan, B. A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest approach. *Comput. Geosci.* **2021**, *148*, 104688. [CrossRef]
21. Yousefi, M.; Carranza, E.J.M.; Kreuzer, O.P.; Nykänen, V.; Hronsky, J.M.; Mihalasky, M.J. Data analysis methods for prospectivity modelling as applied to mineral exploration targeting: State-of-the-art and outlook. *J. Geochem. Explor.* **2021**, *229*, 106839. [CrossRef]
22. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised machine learning: A review of classification techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 3–24.
23. Ghezelbash, R.; Maghsoudi, A.; Carranza, E.J.M. Optimization of geochemical anomaly detection using a novel genetic K-means clustering (GKMC) algorithm. *Comput. Geosci.* **2020**, *134*, 104335. [CrossRef]
24. Daviran, M.; Maghsoudi, A.; Cohen, D.R.; Ghezelbash, R.; Yilmaz, H. Assessment of various fuzzy c-mean clustering validation indices for mapping mineral prospectivity: Combination of multifractal geochemical model and mineralization processes. *Nat. Resour. Res.* **2020**, *29*, 229–246. [CrossRef]
25. Ford, A. Practical implementation of random forest-based mineral potential mapping for porphyry Cu–Au mineralization in the Eastern Lachlan Orogen, NSW, Australia. *Nat. Resour. Res.* **2020**, *29*, 267–283. [CrossRef]
26. Zhao, J.; Chen, S.; Zuo, R. Identifying geochemical anomalies associated with Au–Cu mineralization using multifractal and artificial neural network models in the Ningqiang district, Shaanxi, China. *J. Geochem. Explor.* **2016**, *164*, 54–64. [CrossRef]
27. Zuo, R.; Xiong, Y.; Wang, J.; Carranza, E.J.M. Deep learning and its application in geochemical mapping. *Earth-Sci. Rev.* **2019**, *192*, 1–14. [CrossRef]
28. Ghezelbash, R.; Maghsoudi, A.; Bigdeli, A.; Carranza, E.J.M. Regional-scale mineral prospectivity mapping: Support vector machines and an improved data-driven multi-criteria decision-making technique. *Nat. Resour. Res.* **2021**, *30*, 1977–2005. [CrossRef]
29. Wang, Z.; Zuo, R.; Dong, Y. Mapping geochemical anomalies through integrating random forest and metric learning methods. *Nat. Resour. Res.* **2019**, *28*, 1285–1298. [CrossRef]
30. Gonbadi, A.M.; Tabatabaei, S.H.; Carranza, E.J.M. Supervised geochemical anomaly detection by pattern recognition. *J. Geochem. Explor.* **2015**, *157*, 81–91. [CrossRef]
31. Eshraghi, S.A.; Jalali, A. *Geological Map of Moalleman*; Geological Survey of Iran: Tehran, Iran, 2006.
32. TaleFazel, E.; Mehrabi, B.; GhasemiSiani, M. Epithermal systems of the Torud–Chah Shirin district, northern Iran: Ore-fluid evolution and geodynamic setting. *Ore Geol. Rev.* **2019**, *109*, 253–275. [CrossRef]
33. Niroomand, S.; Lentz, D.R.; Sepidbar, F.; Tajeddin, H.A.; Hassanzadeh, J.; Mirnejad, H. Geochemical characteristics of igneous rocks associated with Baghu gold deposit in the Neotethyan Torud-Chah Shirin segment, Northern Iran. *Geol. J.* **2020**, *55*, 299–316. [CrossRef]
34. Hushmandzadeh, A.R.; Naini, M.A.; Haghipour, A.A. Evolution of geological phenomenon in Torud area. In *Geological Survey of Iran*; Geological Survey of Iran: Tehran, Iran, 1978; p. 136.
35. Rashid Nezhad Omran, N. Investigation of Lithological and Magmatic Developments and Its Relationship with Baghu Gold Mineralization. Master's Thesis, Tarbiat Moalem University, Tehran, Iran, 1991.
36. Fard, M.; Rastad, E.; Ghaderi, M. Epithermal gold and base metal mineralization at Gandy deposit, north of Central Iran and the role of rhyolitic intrusions. *J. Sci. Islam. Repub. Iran* **2006**, *7*, 327–335.
37. Mehrabi, B.; Ghasemi, S.; Tale, F. Structural control on epithermal mineralization in the Troud-Chah Shirin belt using point pattern and Fry analyses, north of Iran. *Geotectonics* **2015**, *49*, 320–331. [CrossRef]
38. Shamanian, G.H.; Hedenquist, J.W.; Hattori, K.H.; Hassanzadeh, J. The Gandy and Abolhassani epithermal prospects in the Alborz magmatic arc, Semnan province, Northern Iran. *Econ. Geol.* **2004**, *99*, 691–712. [CrossRef]
39. Akerdi, M.M.; Shafaroudi, A.M.; Karimpour, M.H.; Rahimi, B.; Santos, J.F. Evidence of iron oxide-copper–gold mineralization in the Torud-Chahshirin Magmatic Belt, northern Iran: Insight from the Robaie area. *Ore Geol. Rev.* **2021**, *129*, 103937. [CrossRef]
40. Shuguang, Z.; Kefa, Z.; Yao, C.; Jinlin, W.; Jianli, D. Exploratory data analysis and singularity mapping in geochemical anomaly identification in Karamay, Xinjiang, China. *J. Geochem. Explor.* **2015**, *154*, 171–179. [CrossRef]
41. Liu, Y.; Zhou, K.; Cheng, Q. A new method for geochemical anomaly separation based on the distribution patterns of singularity indices. *Comput. Geosci.* **2017**, *105*, 139–147. [CrossRef]
42. Ersoy, A.; Yunsel, T.Y. Geochemical modelling and mapping of Cu and Fe anomalies in soil using combining sequential Gaussian co-simulation and local singularity analysis: A case study from Dedeyazı (Malatya) region, SE Turkey. *Geochem. Explor. Environ. Anal.* **2019**, *19*, 331–342. [CrossRef]
43. Tan, Q.P.; Wang, X.; Xia, Y.; Liu, Q.; Zhou, J. Identifying ore-related anomalies using singularity mapping of stream sediment geochemical data, a case study of Pb mineralization in the Qinling region, China. *Geochem. Explor. Environ. Anal.* **2018**, *18*, 177–184. [CrossRef]
44. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
45. Efron, B. Bootstrap methods: Another look at the jackknife. In *Breakthroughs in Statistics: Methodology and Distribution*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 569–593.

46. Zhang, S.; Xiao, K.; Carranza, E.J.M.; Yang, F. Maximum entropy and random forest modeling of mineral potential: Analysis of gold prospectivity in the Hezuo–Meiwu district, west Qinling Orogen, China. *Nat. Resour. Res.* **2019**, *28*, 645–664. [CrossRef]
47. Liaw, A.; Wiener, M. Classification and regression by randomforest. *R News* **2002**, *2*, 18–22.
48. Ghezelbash, R.; Maghsoudi, A.; Carranza, E.J.M. Sensitivity analysis of prospectivity modeling to evidence maps: Enhancing success of targeting for epithermal gold, Takab district, NW Iran. *Ore Geol. Rev.* **2020**, *120*, 103394. [CrossRef]
49. Carranza, E.J.M. Controls on mineral deposit occurrence inferred from analysis of their spatial pattern and spatial association with geological features. *Ore Geol. Rev.* **2009**, *35*, 383–400. [CrossRef]
50. Bonham-Carter, G.F. Geographic information systems for geoscientists-modeling with GIS. *Comput. Methods Geosci.* **1994**, *13*, 398.
51. Carranza, E.J.M.; Laborte, A.G. Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. *Ore Geol. Rev.* **2015**, *71*, 777–787. [CrossRef]
52. Zhao, J.; Wang, W.; Cheng, Q.; Agterberg, F. Mapping of Fe mineral potential by spatially weighted principal component analysis in the eastern Tianshan mineral district, China. *J. Geochem. Explor.* **2016**, *164*, 107–121. [CrossRef]
53. Reimann, C.; Filzmoser, P.; Garrett, R.G. Factor analysis applied to regional geochemical data: Problems and possibilities. *Appl. Geochem.* **2002**, *17*, 185–206. [CrossRef]
54. Filzmoser, P.; Hron, K.; Reimann, C. Principal component analysis for compositional data with outliers. *Environmetrics Off. J. Int. Environmetrics Soc.* **2009**, *20*, 621–632. [CrossRef]
55. Aitchison, J.; Barceló-Vidal, C.; Martín-Fernández, J.A.; Pawlowsky-Glahn, V. Logratio analysis and compositional distance. *Math. Geol.* **2000**, *32*, 271–275. [CrossRef]
56. Wang, H.; Zuo, R. A comparative study of trend surface analysis and spectrum–area multifractal model to identify geochemical anomalies. *J. Geochem. Explor.* **2015**, *155*, 84–90. [CrossRef]
57. Zuo, R. Identifying geochemical anomalies associated with Cu and Pb–Zn skarn mineralization using principal component analysis and spectrum–area fractal modeling in the Gangdese Belt, Tibet (China). *J. Geochem. Explor.* **2011**, *111*, 13–22. [CrossRef]
58. Xiao, F.; Chen, J.; Zhang, Z.; Wang, C.; Wu, G.; Agterberg, F.P. Singularity mapping and spatially weighted principal component analysis to identify geochemical anomalies associated with Ag and Pb-Zn polymetallic mineralization in Northwest Zhejiang, China. *J. Geochem. Explor.* **2012**, *122*, 90–100. [CrossRef]
59. Buccianti, A.; Grunsky, E. *Compositional Data Analysis in Geochemistry: Are We Sure to See What Really Occurs during Natural Processes?* Elsevier: Amsterdam, The Netherlands, 2014; pp. 1–5.
60. Carranza, E.J.M.; Hale, M. A catchment basin approach to the analysis of reconnaissance geochemical-geological data from Albay Province, Philippines. *J. Geochem. Explor.* **1997**, *60*, 157–171. [CrossRef]
61. Bai, J.; Porwal, A.; Hart, C.; Ford, A.; Yu, L. Mapping geochemical singularity using multifractal analysis: Application to anomaly definition on stream sediments data from Funin Sheet, Yunnan, China. *J. Geochem. Explor.* **2010**, *104*, 1–11. [CrossRef]
62. Zuo, R. Exploring the effects of cell size in geochemical mapping. *J. Geochem. Explor.* **2012**, *112*, 357–367. [CrossRef]
63. Wyborn, L.; Heinrich, C.; Jaques, A. Australian Proterozoic mineral systems: Essential ingredients and mappable criteria. In Proceedings of the AusIMM Annual Conference, Darwin, Australia, 5–9 August 1994; AusIMM Darwin: Darwin, Australia, 1994.
64. Carranza, E.J.M.; Sadeghi, M. Predictive mapping of prospectivity and quantitative estimation of undiscovered VMS deposits in Skellefte district (Sweden). *Ore Geol. Rev.* **2010**, *38*, 219–241. [CrossRef]
65. Yousefi, M.; Hronsky, J.M. Translation of the function of hydrothermal mineralization-related focused fluid flux into a mappable exploration criterion for mineral exploration targeting. *Appl. Geochem.* **2023**, *149*, 105561. [CrossRef]
66. Kreuzer, O.P.; Blenkinsop, T.G.; Morrison, R.J.; Peters, S.G. Ore controls in the Charters Towers goldfield, NE Australia: Constraints from geological, geophysical and numerical analyses. *Ore Geol. Rev.* **2007**, *32*, 37–80. [CrossRef]
67. Pirajno, F. Intracontinental strike-slip faults, associated magmatism, mineral systems and mantle dynamics: Examples from NW China and Altay-Sayan (Siberia). *J. Geodyn.* **2010**, *50*, 325–346. [CrossRef]
68. Xiao, F.; Chen, J.; Agterberg, F.; Wang, C. Element behavior analysis and its implications for geochemical anomaly identification: A case study for porphyry Cu–Mo deposits in Eastern Tianshan, China. *J. Geochem. Explor.* **2014**, *145*, 1–11. [CrossRef]