Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 83 (2016) 1262 – 1267

Procedia Computer Science

The 3rd International Workshop on Machine Learning and Data Mining for Sensor Networks (MLDM-SN)

Combining Weights of Evidence Analysis with Feature Extraction – a Case Study from the Hauraki Goldfield, New Zealand

Leonardo Feltrin[a,b,*], João Gabriel Motta[c], Feras Al-Obeidat[d], Farhi Marir[d], and Martina Bertelli[a]

[a]*Department of Earth Sciences, Western University, London, Ontario, Canada*
[b]*School of Earth and Environmental Sciences, James Cook University, Townsville, Queensland, Australia*
[c]*Graduate School on Geosciences, Universidade Estadual de Campinas, Campinas, São Paulo, Brazil*
[d]*College of Technological Innovation, Zayed University, UAE*

Abstract

In this contribution we combine different image processing and pattern recognition methodologies to map the probability of discovering epithermal mineral deposits in the northern part of the Coromandel peninsula, in New Zealand. The objective of this work is to propose a case-study where the substitution of structural geology GIS themes (commonly developed by humans) with products derived by image processing, computer-based, semi-automatic edge detection analyses, is carried out to reduce subjective input in the prospectivity analysis. Semi-automated lineament extraction results introduced in the mineral favourability statistical modelling can more easily reveal unexpected potentially mineralised target domains, being less subjective. We present initial results of this analysis and explain some of the methodologies adopted. Preliminary results suggest that this approach increases significantly the number of geological discontinuities mapped in the region, with the following implications: (1) prospectivity models are more risk-tolerant and result in an increased number of targets; (2) increments in *posterior* probability affect the statistical validity of the model due to conditional independence violation, requiring careful assessment of probability overestimation; (3) the feature extraction process identifies numerous lineaments that in some instances represent false positives (lineaments determined by a variety of causes, without geological significance); however, we find that *Contrast* calculations in the Bayesian analysis tend to penalize these evidential themes, because of the higher number of pixels (cells) containing a positive pattern (lineament existence = 1, being positive). This aspect reduces the overall impact of these predictors on the analysis, mitigating the effect of false positives (lower positive weights of evidence). Despite the limitations, results obtained are encouraging with a clearly superior and more detailed mapping of potential structural sites and their relative probabilities of hosting epithermal deposits.

Peer-review under responsibility of the Conference Program Chairs

Keywords: Cluster Analysis, Bayesian Learning, Weights of Evidence, Epithermal Gold, Mineral Prospectivity

* Corresponding Author: Leonardo Feltrin. Tel.:+1-519-661-3187
E-mail address: lfeltrin@uwo.ca

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.04.262

1. Introduction

Mineral prospectivity mapping allows mineral endowment estimation for mineral exploration purpose. Its synthetic, geographic representation of endowment distribution represents a solution to the need of identifying prospective regions from a disparate amount of geological GIS 2D information. Mineral endowment results from multiple chemical and physical factors, in a complex 4D multi-scale and multi-stage metallogenic system[1]. Mineral prospectivity mapping exploits the geographic superposition of some these multi-scale processes (the ones amenable to geographic representation). Some of these deterministic components can be characterized numerically in geographic space allowing the mathematical combination of GIS data to estimate the likelihood of finding mineral deposits of the type sought[2,3,4].

Prospectivity mapping has two end-members: Data-Driven (DD) and Knowledge-Driven (KD)[2], differing in the way the influence of evidential themes is weighted and how spatial data classification takes place[5]. Any method has subjective or objective counterparts, with the advantage of DD methods being suitable for revealing unexpected target domains and spatial patterns (because they represent the objective end-member of these categories). The weights of evidence in DD methods are calculated according to the statistical correlation between the (binary) spatial distribution of known mineral deposits and the evidential themes considered (predictor variables). Our objective with this contribution, is to apply Weights of Evidence (WofE) analysis[6], in conjunction with Semi-Automated Lineament Extraction (SALE) for the determination of structural features derived from topographic data (NASA SRTM 3 arc second, publicly available data sets). Lineaments derived from this processing are combined with additional geological, geochemical and geophysical GIS themes for the Hauraki Goldfield[7]. SALE substitutes the structural themes commonly defined by a geologist that carries out a structural interpretation and draws linear features on a map layer in a GIS system. Computers with appropriate image processing algorithms[8,9] allow more uniform edge representation and less subjective "structural geology" themes[10]. Despite the power of these methodologies in Earth science applications, it is often difficult to recognize if extracted features represent real structural discontinuities or simple lithological contacts. Contrasts of other nature unrelated to the permeability structure and geological formations also render the solutions of edge detection algorithms heavily affected by false positives, thus limiting their predictive capacity. On the other hand, the use of probabilistic approaches such as WofE can mitigate this problem, by reducing weighting on these features, and by considering other independent data sets to discriminate true from false positives.

2. Geological Setting and Tectonic Controls on Mineralization

The study considers the Hauraki goldfield, in New Zealand's Northern Island (Fig. 1). This mineral province contains more than 50 Miocene, and Pliocene, Au-Ag epithermal deposits, and several Cu-Au-Mo occurrences in a 200-km-long and 40-km-wide north-south trending belt[11]. This belt is the result of the formation of different volcanic zones and associated arcs, with the superposition of distinct magmatic events. This part of New Zealand records a complex history of extensional deformation and associated back-arc magmatism and hydrothermal activity[11]. Epithermal deposits relate to subaerial hydrothermal systems hosted in rocks from early Miocene to late Pliocene (~18–1.9 Ma), of the Coromandel volcanic zone (CVZ)[12]. These are given by thick sequences of volcanic and volcanoclastic units superposed to Late-Jurassic basement[11]. Numerous crustal-scale structures dissect the volcano-sedimentary sequences of the Coromandel Peninsula, some of them linked at depth to older Cretaceous basement structures, predominantly organized in NNW and NE-ENE striking directions, controlling the overlying, volcano-sedimentary architecture. Thickening occurs dominantly along the NE-ENE trends, interpreted as sets of normal structures dipping SE[11]. Structurally important are also a number of calderas predominantly found near more evolved magmatic complexes in the eastern part of the peninsula where the youngest Whitianga group rhyolite sequences are present[11].

In agreement with the different exposure levels in the volcanogenic sequences, diverse styles of mineralization are found in the Coromandel peninsula, both porphyry-style gold and base metals dominate in exposed basement domains. Instead, volcanic zones and associated epithermal systems are more abundant in the preserved shallower sequences, reflecting the erosion that occurred in response to the differential subsidence of segmented sub-basins. Epithermal deposits present strong structural control with classical breccia, stockwork-like, mineralised vein systems[11]. The relationship between volcanic calderas and epithermal systems, as well as their connection to normal structures, is critical with faults and radial ruptures developed during caldera activity being important pathways for hydrothermal

ascent of Au-bearing fluids. The observed architecture is hosting minor porphyry mineralization and (multimillion ounces) epithermal systems predominantly along NE-ENE trends. The mineralization emplacement occurred at different times in the Hauraki field in an overall extensional setting, developed in dip-slip arrays on felsic to intermediate host-rocks. Epithermal mineralization is of variable style with both endmembers (low- and high-sulphidation systems) represented in the area[13,14]. Endmembers differ in the nature of hydrothermal fluid (magmatic vs meteoric input) as well as the relative mineralogy of hydrothermal alteration due to different formation temperatures, which are slightly lower in the low-sulphidation systems (~200 instead of 300 °C).

The following features are important in the genesis and localization of the low-sulphidation andesite-hosted epithermal Au deposits: (1) Tectonic-volcanic setting: hydrothermal systems that contain andesite-hosted epithermal gold deposits are associated with calc-alkaline magmas, generally in subaerial volcanic arcs. (2) Structural controls: epithermal gold deposits are localised preferentially in extensional fault/fracture systems, with fault jogs providing favourable sites for bigger deposits. Individual quartz-veins occupy steeply dipping normal or oblique-slip faults that show only minor displacement. (3) Lithological controls: andesite lavas host thicker and higher grade veins, whereas less competent flow breccia and pyroclastic rocks carry thinner, less persistent veins. (4) Hydrothermal alteration signature: adularia-sericite or sericite adjacent to quartz veins, with propylitic and argillic alteration in outer zones. (5) Geochemistry: Au and Ag are the main “geochemically anomalous” elements; many deposits also have anomalous zinc, lead and copper, some also have arsenic, antimony, or molybdenum. These factors were translated into GIS themes that were used to perform the probabilistic estimates.

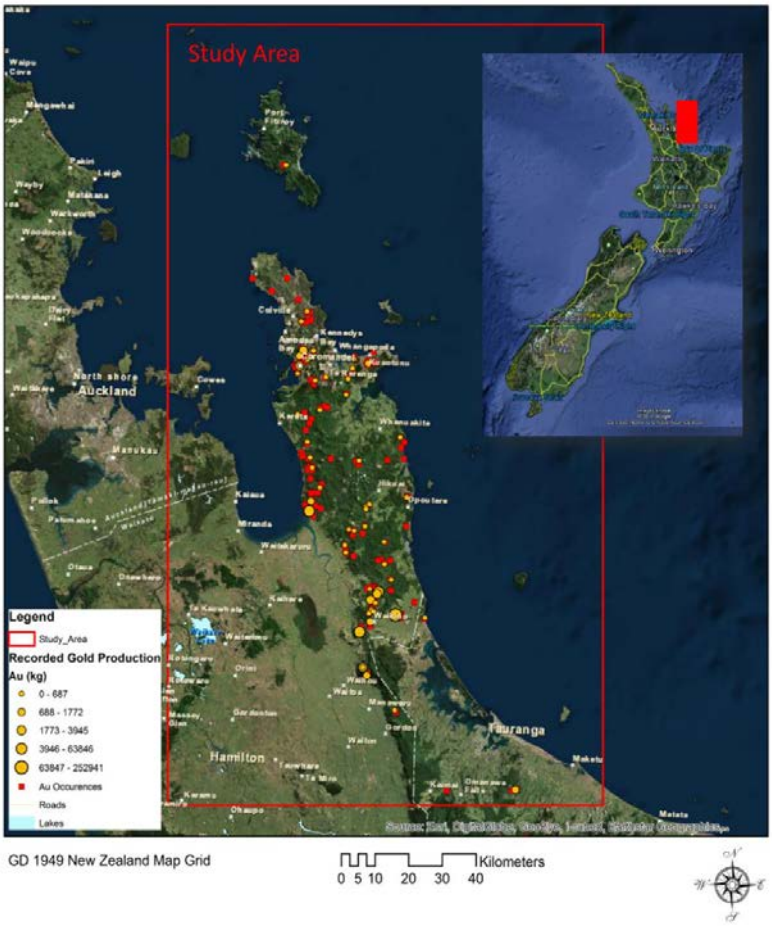

Figure 1. Map illustrating the Coromandel peninsula gold endowment. Data from Rattenbury and Partington[7], GIS compilation MR4343.

3. Methodology of data pre-processing and analysis

The methodology adopted considers three steps: (1) Multiple artificial, directional illumination of SRTM-derived DEM, processed at 90 m pixel resolution to capture any likely edge arising from sharp variations in image tones due to abrupt changes in elevation. (2) Extraction of lineaments using edge detection analysis. (3) Bayesian integration based on the Weights of Evidence approach (refer to Fig. 2 for details; MODELS 1,2). Edge detection experiments conducted using directional sun-shading to enhance DEM contrast, and Fourier decomposition phase analysis[8,9,15] produced feature-dense outputs that according to the illumination directions highlight either potential linear structures (faults) or circular discontinuities (caldera-rims), methodology adapted from[16]. The sun-shaded DEM images show narrow valleys that extend with northeasterly trends in most of the inland and resemble some of the known fault zones. Four directions of sun-illumination were used in the hill-shading to remove the shadow effects induced by unidirectional illumination. Sun-angle images were coupled together using an index overlay methodology. Vertical sun-illumination, was also used to enhance relatively flat and narrow drainage that was found to correspond to some of the known caldera-rings. These input topographic layers were imported in Matlab to process them with other GIS themes using WofE custom developed routines, to obtain a mineral prospectivity map. Seven predictor layers were considered for WofE integration (magnetite destruction zones - A, fault density - B, faults orientation - C, quartz veins - D, As-anomaly - E, Au-anomaly - F, and favourable lithology - G). Two mineral prospectivity maps were developed for comparison purpose, with (MODEL 2) and without (MODEL 1) the SRTM derived structural layers added as predictors. In particular MODEL 1 uses geologically interpreted (and manually drawn features) buffered using the same methodology of *Contrast* calculation for different distance based predictors. It follows a more extensive discussion of the WofE methodology, which explains the approach adopted to generate the two GIS models.

Geographic mapping of the likelihood of finding a mineral deposit as a function of the distribution of particular unique conditions found to be characteristic of mineralised sites has been exhaustively illustrated in[2] and many other more recent works[17,18]. Despite the variety of existing GIS data integration methods, for spatial data modelling, the Bayesian Weights of Evidence[5,6,19] method has seen the most widespread usage[20]. We choose to apply this methodology because of the simplicity of its implementation. The original FORTRAN programs proposed in Bonham-Carter[2] were translated in Matlab code to facilitate analysis and experimentation. WofE relies on empirical calculations of spatial dependence indices (evidential weights) based on the areal, raster (grid) overlap of known deposits and "independent" evidence layers. Weights of evidence of a feature A in relation to a feature B can be calculated to express numerically positive (positive weights, W^+) and negative correlations (negative weights, W^-). The weights' geospatial distributions are calculated using a statistical *Contrast* measure, which represents the strength of correlation between evidence maps and the mineral deposit themes in bivariate space, these depend on the degree of spatial intersection of evidential themes with the known deposits theme as well their extent with respect to the study area. The procedure of calculating the weights and their geographic distribution on each evidential layer leads to the development of a series of binary or multiclass predictors (Fig. 2). These layers are combined (integrated) to augment *prior* knowledge. This process is represented by a *prior* probability calculation (the initial knowledge we have of the spatial distribution of mineral deposits in the region of interest). The deposit area (number of raster unit cells containing a deposit) is divided by the total number of cells representing the study area obtaining a *prior* probability.

Probabilities can be expressed as odds and if we take the logit of the odds we can simplify the calculation of the *posterior* probabilities derived from each unique condition, resulting from the combination of evidential themes. Odds are defined as the ratio of the probability that an event will occur to the probability that it will not, considered in terms of its logits[2]. Thus, an evidence set can be combined according to its odds into a *posterior* probability mapping, which depicts the likelihood of the phenomena described by the evidence layers to occur in a specific location in space, with each layer of evidence contributing to the final *posterior* probability. *Posterior* probabilities are calculated simply adding the weights of evidence to the logit of the *prior* probability. Once *posterior* probabilities are estimated they can be displayed on a map and they represent a measure of the similarity of each unique condition in a map to the unique conditions most favourable to mineralisation. If the unique conditions determined by the combination of predictors are similar to the unique conditions that are found in mineralised sites, the resulting *posteriors* will be close to or equal to a value of 1. Probability of occurrence is expressed in the interval [0,1] with 1 being the highest. The probability of a deposit to occur in Bayesian probability theory is determined by coexistence of one, or more, favourable factors (evidential themes) expressed as independent probabilities, if we multiply them by the *prior* we obtain a *posterior* estimate[2]. The Bayes' assumption is that probabilities derived from evidence are completely independent processes. For this reason, WofE modelling may lead to significant overestimation of *posteriors*. To

partly mitigate this problem Conditional Independence (CI) tests are conducted[21] since CI violation is a common issue on favourability mapping (e.g., geochemical element affinity, structural and geophysical responses).

4. Results and Conclusions

Preliminary observations were made by direct visualisation of the prospectivity GIS surfaces. Figure 2, presents two mineral prospectivity maps with different geographic distributions of *posteriors*, indicating the favourability of locating ground that most likely hosts epithermal deposits. These maps illustrate the variation in the large-scale mapping of favourability across the Coromandel area, as well as providing (zoom in) enlargements of some of the localities with economic resources (e.g., Favona, Waihi, Martha, Thames deposits). This is particularly important because the maps need to be analysed at different scales of observation to study their differences. The following preliminary observations and conclusions were drawn from their comparison: (1) Broad scale visualization documents broader *posteriors* with more continue and larger favourability domains in MODEL 1. Instead the *posterior* distributions in MODEL 2 appear more restricted resembling the distribution of SRTM lineaments. (2) If we consider the enlarged sub-sets it is apparent a good overlap of deposits/occurrences with the high favourability domains (*posterior* probability > 0.65).

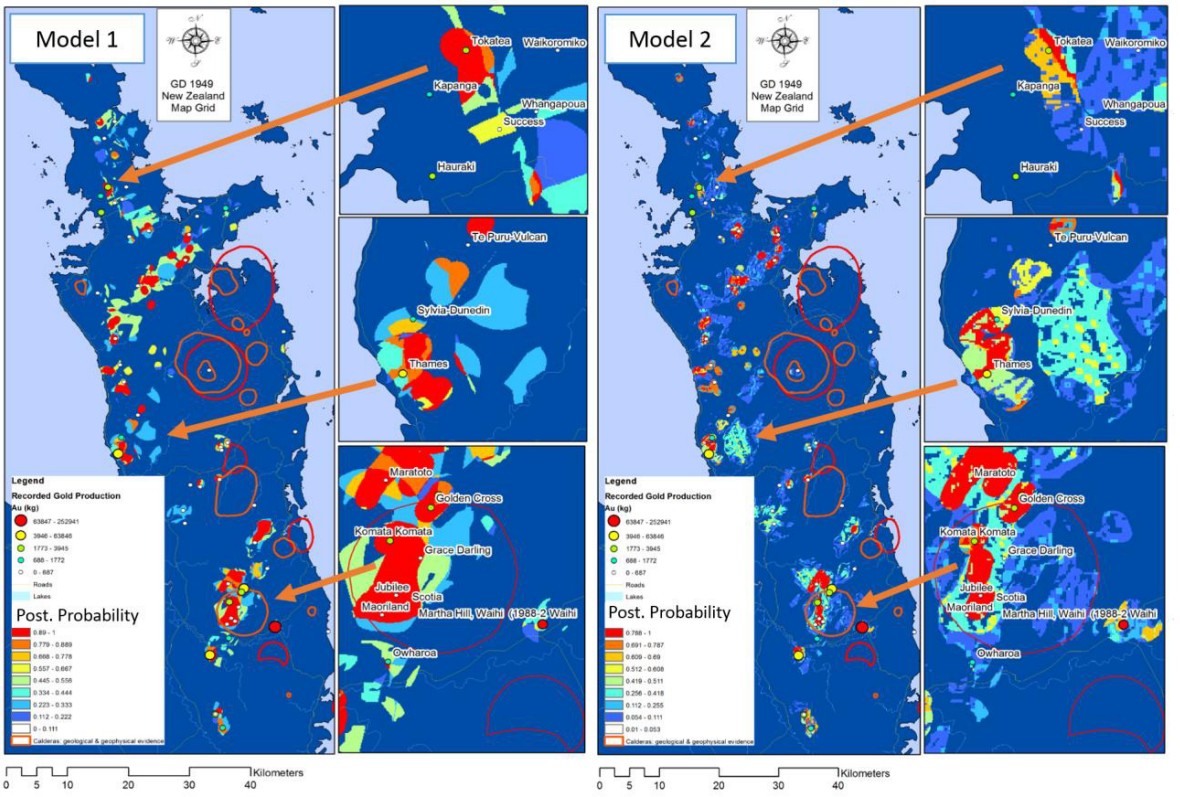

Figure 2. Comparison of WofE analysis results, respectively with conventional structural mapping data (MODEL 1) or inclusion of edge detection products based on the Geosoft Grid Analysis tools (MODEL 2). As it can be seen by comparing sub-sets in MODELS 1,2, at 5 km to 100 m scales MODEL 2 predictions appear to be superior, because potential faults/lineaments or geological contacts correlate well with known mineral deposits, but the areas with the highest favourability (red pixels) are smaller in MODEL 2 and appear to be restricted to the inferred fault zones. Thus, target domains are delineated more accurately and integration with other variables such as alteration and known distribution of quartz veins, can facilitate the interpretation of Au-fertile domains. In particular, the enlargements in the mineralized domains show different textures defined by more discrete and restricted high-favourability zones in Model 2. It is believed that for this reason although the lineaments may represent false positives, if they are found to be associated with other mineralisation indicators (e.g., distribution of hydrothermal quartz—which indicates hydrothermal activity) in the WofE results, they could be discriminated as real faults and because of the narrower high probability domains, they can be considered superior mineral prospectivity targets, since they provide a more accurate determination of a structural site location.

Comparison of sub-sets illustrates well the textural difference due to inclusion of lineament extraction derivatives within MODEL 2, by exhibiting probability domains that are more heterogeneous. Their appearance in some cases resembles natural patterns. Highest levels of favourability appear to be more limited to narrow linear zones that could be interpreted as fault zones, since they have compatible orientation with known trends as well as being intersected by alteration and other independent structural predictors. Given these preliminary results despite the increase in false positives WofE calculations conducted in MODEL 2 provide a more detailed data-driven mapping of potential structural discontinuities and associated favourability. This could be particularly useful in cases where Au is found in second- and third-order structures[7]. Given the superior resolution of SRTM based themes these features may be detectable by edge extraction methods, as presented in this contribution. In addition to this, in general, use of semi-automated edge detection analysis in conjunction with mineral prospectivity analysis reduced significantly the time needed to produce structural interpretations for the study area, an attractive aspect for fast pace mineral exploration and screening of target domains.

Acknowledgments

We thank Bruce Spencer for providing encouraging comments on earlier drafts of the manuscript.

References

1. T. C. McCuaig, S. Beresford, and J. Hronsky, “Translating the mineral systems approach into an effective exploration targeting system,” *Ore Geol. Rev.*, vol. 38, pp. 128–138, 2010.
2. G. F. Bonham-Carter, “Geographic Information Systems for Geoscientists: Modelling with GIS,” *Comput. Methods Geosci.*, vol. 13, p. 416, 1994.
3. F. Agterberg, “Geochemical Anomaly and Mineral Prospectivity Mapping in GIS, vol. 11, Emmanuel John M. Carranza (Ed.), in: Handbook of Exploration and Environmental Geochemistry. Elsevier, Amsterdam, the Netherlands (2008),” *Ore Geol. Rev.*, vol. 35, pp. 455–456, 2009.
4. E. J. M. Carranza, H. Wibowo, S. D. Barritt, and P. Sumintadireja, “Spatial data analysis and integration for regional-scale geothermal potential mapping, West Java, Indonesia,” *Geothermics*, vol. 37, pp. 267–299, 2008.
5. E. J. M. Carranza, “Weights of evidence modeling of mineral potential: A case study using small number of prospects, Abra, Philippines,” *Nat. Resour. Res.*, vol. 13, no. 3, pp. 173–187, 2004.
6. G. F. Bonham-carter and F. P. Agterberg, “Arc-WofE : a GIS tool for statistical integration of mineral exploration datasets,” 1988.
7. M. S. Rattenbury and G. A. Partington, “Prospectivity models and GIS data for the exploration of epithermal gold mineralization in New Zealand.” in Epithermal gold in New Zealand: GIS data package and prospectivity modelling: Crown Minerals, Ministry of Economic Development, and In-stitute of Geological and Nuclear Sciences , p. 68, 2003.
8. P. Kovesi, “Phase congruency: A low-level image invariant,” *Psychol. Res. Forsch.*, vol. 64, pp. 136–148, 2000.
9. P. Kovesi, “Symmetry and Asymmetry from Local Phase,” in *Tenth Australian Joint Converence on Artificial Intelligence*, 1997, pp. 2–4.
10. F. Al-Obeidat, L. Feltrin, and M. Farhi, “Cloud Based Systematic Lineament Extraction of Topographic Lineaments from NASA Shuttle Radar Topography Mission Data,” *Procedia Comput. Sci.*, no. (this volume), 2016.
11. A. B. Christie, M. P. Simpson, R. L. Brathwaite, J. L. Mauk, and S. F. Simmons, “Epithermal Au-Ag and related deposits of the Hauraki Goldfield, Coromandel Volcanic Zone, New Zealand,” *Econ. Geol.*, vol. 102, pp. 785–816, 2007.
12. R. Brathwaite, M. Simpson, K. Faure, and D. Skinner, “Telescoped porphyry Cu-Mo-Au mineralisation, advanced argillic alteration and quartz-sulphide-gold-anhydrite veins in the Thames district, New Zealand,” *Miner. Depos.*, vol. 36, pp. 623–640, 2001.
13. A. B. Christie and R. L. Brathwaite, “Hydrothermal alteration in metasedimentary rock-hosted orogenic gold deposits, Reefton goldfield, South Island, New Zealand,” *Miner. Depos.*, vol. 38, no. 1, pp. 87–107, 2003.
14. N. C. White and J. W. Hedenquist, “Epithermal environments and styles of mineralization: Variations and their causes, and guidelines for exploration,” *J. Geochemical Explor.*, vol. 36, no. 1–3, pp. 445–474, 1990.
15. E.-J. Holden, M. Dentith, and P. Kovesi, “Towards the automated analysis of regional aeromagnetic data to identify regions prospective for gold deposits,” *Comput. Geosci.*, vol. 34, no. 11, pp. 1505–1513, 2008.
16. A. A. Masoud and K. Koike, “Auto-detection and integration of tectonically significant lineaments from SRTM DEM and remotely-sensed geophysical data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 66, pp. 818–832, 2011.
17. A. Joly, A. Porwal, and T. C. McCuaig, “Exploration targeting for orogenic gold deposits in the Granites-Tanami Orogen: Mineral system analysis, targeting model and prospectivity analysis,” *Ore Geol. Rev.*, vol. 48, pp. 349–383, Oct. 2012.
18. A. Joly, A. Porwal, T. C. McCuaig, B. Chudasama, M. C. Dentith, and A. R. A. Aitken, “Mineral systems approach applied to GIS-based 2D-prospectivity modelling of geological regions: Insights from Western Australia,” *Ore Geol. Rev.*, vol. 71, pp. 673–702, Dec. 2015.
19. Q. Cheng, “Application of weights of evidence method for assessment of flowing wells in the greater Toronto area, Canada,” *Nat. Resour. Res.*, vol. 13, no. 2, pp. 77–86, 2004.
20. E. J. M. Carranza, J. C. Mangaoang, and M. Hale, “Application of mineral exploration models and GIS to generate mineral potential maps as input for optimum land-use planning in the Philippines,” *Nat. Resour. Res.*, vol. 8, no. 2, pp. 165–173, 1999.
21. F. Agterberg and Q. Cheng, “Conditional independence test for weights-of-evidence modeling,” *Nat. Resour. Res.*, vol. 11, no. 4, pp. 3–9, 2002.