Natural Resources Research, Vol. 11, No. 4, December 2002 (© 2002)

Conditional Independence Test for Weights-of-Evidence Modeling

Frederik P. Agterberg[1,3] and Qiuming Cheng[2]

Received and accepted 5 July 2002

Weights-of-evidence modeling is a GIS-based technique for relating a point pattern for locations of discrete events with several map layers. In general, the map layers are binary or ternary. Weights for presence, absence or missing data are added to a prior logit. Updating with two or more map layers is allowed only if the map layers are approximately conditionally independent of the point pattern. The final product is a map of posterior probabilities of occurrence of the discrete event within a small unit cell. This paper contains formal proof that conditional independence of map layers implies that T, the sum of the posterior probabilities weighted according to unit cell area, is equal to n, being the total number of discrete events. This result is used in the overall or "omnibus test" for conditional independence. In practical applications, T generally exceeds n, indicating a possible lack of conditional independence. Estimation of the standard deviation of T allows performance of a one-tailed test to check whether or not T-n is significantly greater than zero. This new test is exact and simpler to use than other tests including the Kolmogorov-Smirnov test and various chi-squared tests adapted from discrete multivariate statistics.

KEY WORDS: Conditional independence; weights-of-evidence; mineral deposits; map layers; significance test.

INTRODUCTION

Weights-of-evidence modeling was developed originally for medical diagnosis, but was applied subsequently to mineral–potential mapping. A pattern of mineral deposits is related to several map layers representing geoscience data that may be indicative of occurrence of mineral deposits. A comprehensive summary of the theory with applications is given in Bonham-Carter (1994). Arc-WofE is a weights-of-evidence extension for ArcView/Spatial Analyst users that is available freely on the Internet (Kemp, Bonham-Carter, and Raines, 1999). It includes documentation and references.

The number of users of weights-of-evidence is increasing. Users now include scientists in fields such as physical geography, epidemiology, ecology, forestry, and medicine for the study of regional disease occurrence patterns. Examples of recent papers using and evaluating weight-of-evidence are Carranza and Hale (2000, 2002), Cheng and Agterberg (1999), Boleneus and others (2001), Harris and others (2001), Mihalasky and Bonham-Carter (2001), Raines (1999), Scott and Dimitrakopoulos (2001), and Venkataraman and others (2000).

Other pattern recognition techniques with similar objectives include neural networks (Singer and Kouda, 1999) and logistic regression (Agterberg and Bonham-Carter, 1999). Weights-of-evidence also is compared to other techniques in the book by Pan and Harris (2000). One of the advantages of weights-of-evidence is its simplicity, and straightforward interpretation of the weights.

Suppose that n deposits have been discovered in a study region of N equal-area unit cells. The unit cell is

[1] Geological Survey of Canada, Ottawa, Canada.
[2] Department of Earth and Atmospheric Science, York University, Toronto, Canada.
[3] To whom correspondence should be addressed at Geological Survey of Canada, 601 Booth Street, Ottawa, Canada K1A 0E8; e-mail: agterberg@gsc.nrcan.gc.ca

small. It may be a pixel or simply a relatively small unit of area. The prior probability, P_d, that a cell selected at random contains a deposit (or the center of gravity of a deposit) is set equal to n/N. If N_A cells are underlain by binary map layer A representing a theme favorable for occurrence of deposits, $n_A/N_A > n/N$. This relation can be written as $P(d \mid A) > P_d$ with $P(d \mid A) = n_A/N_A$ denoting the probability that a cell on A contains a deposit.

The probability that a randomly selected cell occurs on A is $P_A = N_A/N$. The probability $P(A \mid d) = n_A/n$ that a deposit cell occurs on A satisfied Bayes' rule for the relationship between $P(d \mid A)$ and $P(A \mid d)$ that can be derived as follows. If dA denotes presence of both d and A in a cell, it follows that $P(d \mid A) = P(dA)/P_A$ and $P(A \mid d) = P(dA)/P_d$. Elimination of $P(dA)$ gives $P(d \mid A) = P(A \mid d) \cdot P_d/P_A$ (Bayes' rule). It is convenient to use logits $L = \log_e\{P/(1-P)\}$. In terms of logits, Bayes' rule becomes $L(d \mid A) = W^+(A) + L_d$ where $L(d \mid A)$ and L_d correspond to $P(d \mid A)$ and P_d. The positive weight of A satisfies $W^+(A) = \log_e P(A \mid d)/(P(A \mid {\sim}d)$ with $\sim$ denoting "not." The validity of these statements can be verified readily by using the probabilities $P({\sim}d) = (N-n)/N$ and $P({\sim}d \mid A) = (N_A - n_A)/N_A$ for cells without deposits. A negative weight $W^-(A) = W^+({\sim}A)$ applies to the $N - N_A$ cells not belonging to A.

Suppose that the prior logit L_d is updated using map layer A. The unit cells are assigned new values representing posterior logits for presence and absence of A. In practical applications, information on presence or absence of A may be missing for a cell. The pattern of A then is ternary instead of binary. The prior logit is left unchanged when information on A is missing.

If $\boldsymbol{X_A}$ and $\boldsymbol{X_{\sim A}}$ represent binary random variables for presence and absence of a deposit in a unit cell on A, respectively, then the two posterior probabilities satisfy $E\boldsymbol{X_A} = P(A \mid d)$ and $E\boldsymbol{X_{\sim A}} = P({\sim}A \mid d)$ where E represents mathematical expectation. This pair of random variables can be written as $\boldsymbol{X_I}(I = A, {\sim}A)$.

The posterior logits of the unit cells can be updated further using the weights of a second binary or ternary map layer B if A and B are conditionally independent of the mineral deposits. This condition implies $P(AB \mid d) = P(A \mid d) \cdot P(B \mid d)$ where AB represents presence of both A and B. If $\boldsymbol{X_A}$ and $\boldsymbol{X_B}$ are binary random variables for deposits on the two map layers, conditional independence exists if $E(\boldsymbol{X_I X_J}) = E\boldsymbol{X_I} \cdot E\boldsymbol{X_J}$. $(I = A, \sim A;\ J = B, {\sim}B)$.

A characteristic feature of this type of conditional independence is that it does not necessarily imply relations such as $P(AB \mid {\sim}d) = P(A \mid {\sim}d) \cdot P(B \mid {\sim}d)$. This is because mineral deposits occur at points. In this respect, weights-of-evidence modeling differs from other applications of discrete multivariate analysis.

The updating process can be continued using other binary map layers. For example, if a third pair of binary random variables $\boldsymbol{X_K}$ $(K = C, {\sim}C)$ is added, conditional independence continues to exist if $E(\boldsymbol{X_I}\ \boldsymbol{X_J}\ \boldsymbol{X_K}) = E\boldsymbol{X_I} \cdot E\boldsymbol{X_J} \cdot E\boldsymbol{X_I}$. The order according to which updating takes place is immaterial, and weights for all map layers can be added simultaneously to the prior logit L_d of a cell. The final posterior logits L_f can be transformed into posterior probabilities P_f. In this paper, the sum of posterior probabilities over all unit cells in the study area is written as T.

As discussed in Agterberg, Bonham-Carter, and Wright (1990), asymptotic theory of discrete multivariate analysis can be used to estimate variances of the weights; these can be augmented by variances for missing data and added to the variance of the prior logit. The result is an estimate of the standard deviation of L_f.

The first derivative of L with respect to P is $\{P(1-P)\}^{-1}$. Multiplication of the standard deviation of L_f by $P_f(1-P_f)$ provides an approximate estimate of the standard deviation of P_f. Estimated standard deviations of all posterior probabilities in the study area can be combined with one another to estimate the standard deviation of T.

CONDITIONAL INDEPENDENCE TESTING

The following simple "redundancy" example illustrates that weights-of-evidence modeling can be sensitive to violations of conditional independence. Suppose that binary map layer A has positive weight $W^+(A) = 2$ and that its pattern coincides with that of map layer B. It follows that $W^+(B) = 2$ as well. A situation of this type could, for example, arise when contour maps of trace elements are used to predict occurrences of mineral deposits genetically related to several trace elements.

Application of weights-of-evidence then would yield posterior logits that are too large in places where A and B are present. When the unit cell area is small, this implies that the corresponding posterior probabilities are $e^2 = 7.4$ times as large as they should be. Clearly, situations of this type should be avoided in practical applications.

In the past, two types of conditional independence tests have been applied: (1) contingency table

Table 1. Contingency Table for 2×2 Conditional Independence Test

	Observed frequencies			Expected frequencies		
	A	$\sim A$	Sum	A	$\sim A$	Sum
B	n_{AB}	$n_{\sim AB}$	n_B	$n_A n_B$	$n_{\sim A} n_B$	n_B
$\sim B$	$n_{A\sim B}$	$n_{\sim A\sim B}$	$n_{\sim B}$	$n_A n_{\sim B}$	$n_{\sim A} n_{\sim B}$	$n_{\sim B}$
Sum	n_A	$n_{\sim A}$	n	n_A	$n_{\sim A}$	n

tests, and (2), the overall or "omnibus" test possibly supplemented by a Kolmogorov-Smirnov test for goodness-of-fit. If one or more tests fail in a practical application, new types of map layers can be defined so that there is approximate conditional independence verified by new conditional independence tests. In the preceding example of map layers for intercorrelated trace elements, trace elements would be combined into a single index, for example by factor analysis, prior to combination with other, conditionally independent, map layers.

Contingency Table Tests

Conditional independence of two map layers A and B can be evaluated by the chi-squared test or the G^2-test for goodness-of-fit. Table 1 shows observed and expected frequencies for the 2×2 contingency table for a pair of binary variables $\boldsymbol{X_I}$ ($I = A, \sim A$) and $\boldsymbol{X_J}$ ($J = B, \sim B$) (see previous section). All frequencies in Table 1 are independent of unit cell size provided that unit cell area is sufficiently small.

In the chi-squared test, the observed frequencies are changed slightly when the continuity correction is applied. This well-known refinement consists of adding 0.5 to observed frequencies that are less than expected and subtracting 0.5 from those exceeding expected frequencies. Adding primes as superscripts for this correction, it is possible to compute

$$X^2 = \frac{\boldsymbol{n}\{\boldsymbol{n}'_{A\sim B}\boldsymbol{n}'_{\sim AB} - \boldsymbol{n}'_{AB}\boldsymbol{n}'_{\sim A\sim B}\}}{\boldsymbol{n}_A\boldsymbol{n}_{\sim A}\boldsymbol{n}_B\boldsymbol{n}_{\sim B}}$$

This test statistic is distributed as χ^2 with a single degree of freedom if the two binary variables are conditionally independent. A different test statistic is used for the G^2-test but the two tests generally produce similar results. A possible disadvantage of both tests is that all four expected frequencies of Table 1 should be at least 5 before these tests can be applied.

The theory of the preceding test and its extensions to discrete multivariate statistics are given in several textbooks including Bishop, Feinberg, and Holland (1975). Possible interrelationships between several map layers can be investigated by model comparison. It should be kept in mind, however, that in the resulting test statistics, the map layers are considered only at the points of occurrence of deposits. Consequently, the number of degrees of freedom of the chi-squared value is 1 instead of 2 for a ($2 \times 2 \times 2$) contingency table, arising when two binary variables are tested for conditional independence with respect to a third binary variable in a nonspatial application of discrete multivariate analysis (cf. Agterberg, 1992).

Overall or "Omnibus" Test

The end product of weights-of-evidence modeling is a posterior probability map. If p binary patterns are considered, and there are no missing data, unit cells with the same posterior probability form classes that belong to one of 2^p possible "unique" conditions. Suppose that T represents the sum of posterior probabilities for all unit cells in the study area. Ideally, T should be equal to n representing total number of deposits. In practical applications, T may exceed n. It can be assumed that $T > n$ may be because of lack of conditional independence of map layers. This is the rationale of the overall or so-called "omnibus test" (Kemp, Bonham-Carter, and Wright 1999) for conditional independence. For example, in Bonham-Carter (1994, p. 316), it is argued that T should not exceed n by more than 15%.

The cumulative frequency distribution of the posterior probabilities can be subjected to a Kolmogorov-Smirnov test for goodness-of-fit. This test normally applies to situations in which the maximum observed and expected cumulative frequencies are both equal to 1. If the statistical model for the frequency distribution is correct, the absolute value of the largest difference between calculated and observed cumulative frequencies should not exceed the Kolmogorov-Smirmov test statistic for which statistical tables are available.

Agterberg and others (1993) applied the Kolmogorov-Smirnov test after dividing all cumulative frequencies by n. In weights-of-evidence applications, the maximum difference then generally becomes equal to $(T/n - 1)$ for the unique condition with the largest posterior probability. This is because the unique condition with the largest posterior probability in the study area usually has positive weights for all map layers considered. If two or more of these map layers are not conditionally independent, this results in overestimation as illustrated in the "redundancy" example at the beginning of this section. It is not known to what extent the Kolmogorov-Smirnov test statistic can be applied in that situation.

Suppose that $\boldsymbol{T}$ denotes a random variable for the sum of all posterior probabilities. It would be better to first test the hypothesis that $E\boldsymbol{T} = n$ before the Kolmogorov-Smirnov test is applied to cumulative frequencies multiplied by n/T. This correction would force the maximum difference to become equal to zero thus allowing application of the Kolmogorov-Smirnov test. This procedure would be used only after acceptance of the hypothesis of conditional independence.

NEW CONDITIONAL INDEPENDENCE TEST

For a single map layer A, the binary random variable $\boldsymbol{X_I}(I = A, \sim A)$ results in posterior probabilities $E\boldsymbol{X}_A = P(d \mid A) = n_A/N_A$ when A is present, and $E\boldsymbol{X}_{\sim A} = P(d \mid \sim A) = (N_A - n_A)/N_A$ when A is absent. The expected value of $\boldsymbol{T}$ representing the sum of all posterior probabilities in the study area satisfies

$$\begin{aligned} E\boldsymbol{T} &= N_A E\boldsymbol{X}_A + N_{\sim A} E\boldsymbol{X}_{\sim A} \\ &= N_A P(d \mid A) + N_{\sim A} P(d \mid \sim A) \end{aligned}$$

It follows that

$$E\boldsymbol{T} = N_A\{n_A/N_A\} + N_{\sim A}\{n_{\sim A}/N_{\sim A}\} = n$$

The variance is

$$\sigma^2(\boldsymbol{T}) = N_A^2 \sigma^2(\boldsymbol{X}_A) + N_{\sim A}^2 \sigma^2(\boldsymbol{X}_{\sim A})$$

The following results for the sum the four posterior probabilities resulting from two binary patterns A and B are valid only if the two map layers are conditionally independent of the deposits:

$$\begin{aligned} E\boldsymbol{T} &= N_{AB} P(d \mid AB) + N_{A\sim B} P(d \mid A \sim B) \\ &\quad + N_{\sim AB} P(d \mid \sim AB) + N_{\sim A\sim B} P(d \mid \sim A\sim B) \\ &= n_{AB} + n_{A\sim B} + n_{\sim AB} + n_{\sim A\sim B} = n \end{aligned}$$

$$\begin{aligned} \sigma^2(\boldsymbol{T}) &= N_{AB}^2 \sigma^2(\boldsymbol{X}_{AB}) + n_{A\sim B}^2 \sigma^2(\boldsymbol{X}_{A\sim B}) \\ &\quad + N_{\sim AB}^2 \sigma^2(\boldsymbol{N}_{\sim AB}) + N_{\sim A\sim B}^2 \sigma^2(\boldsymbol{X}_{\sim A\sim B}) \end{aligned}$$

In general, $E\boldsymbol{T} = n$ for the sum of 2^p terms of the form $N_{IJK}\ldots P(d \mid IJK\ldots) = n_{IJK\ldots}$ if the p binary map layers with $(I = A, \sim A; J = B, \sim B; K = C, \sim C; \ldots)$ are conditionally independent of the deposits. The corresponding variance is the sum of all possible terms of the form $N_{IJK\ldots}^2 \sigma^2(\boldsymbol{X}_{IJK\ldots})$.

In general, the sum of all posterior probabilities T is probably greater than n. The hypothesis of conditional independence is equivalent to assuming $E\boldsymbol{T} = n$. This hypothesis can be tested because an estimate of $\sigma^2(\boldsymbol{T})$ is available as well.

Theoretically, it is possible that T is less than n for map layers that are not conditionally independent. However, this situaiton is unlikely to arise in practice because map layers are selected in the first place because they are believed to provide positive indicators for the presence of mineral deposits. It is more likely that indicators of this type are correlated positively with respect to mineral occurrences than that they would be negatively correlated.

For this reason, a one-tailed significance test should be used. Approximate normality of T can be assumed if $s(T)$ representing the standard deviation of T is significantly less than T itself. For acceptance of the hypothesis of conditional independence, the difference T-n should be less than $1.645 \cdot s(T)$ with a probability of 95%, or less than $2.33 \cdot s(T)$ with a probability of 99%.

EXAMPLE OF APPLICATION

This example is based on results previously described in Agterberg and others (1993) for 13 occurrences of hydrothermal vents on the ocean floor, along the central axis of the East Pacific Rise near 21°N. Map layers and posterior probabilities for the area (Agterberg and others, 1993, figs. 1–4) were based on material originally published by Ballard and others (1981).

Table 2 shows weights and contrasts for five binary patterns. The contrast C is the difference between positive and negative weight. It measures degree of correlation between point pattern and binary map, and can be tested for statistical significance. Yule's coefficient of association for binary variables is a standardized form of the contrast. The correlation

Table 2. Hydrothermal Vents on East Pacific Rise Example (From Agterberg and others, 1993). Weights and Contrasts With Standard Deviations for Five Binary Patterns

Map layer	W^+	$s(W^+)$	W^-	$s(W^-)$	C	$s(C)$
Age	2.251	1	−1.232	0.29	3.481	1.041
Topography	2.037	1	−0.811	0.29	2.848	1.041
Rock Type	0.632	0.41	−0.338	0.378	0.97	0.558
Contact	2.259	0.41	−0.57	0.378	2.829	0.561
Fissures	0.178	0.448	−0.097	0.354	0.275	0.571

between fissures and hydrothermal vents is not statistically significant.

The first map layer (age) listed in Table 2 was based on relative amount of recent sediments covering the basaltic rocks. These volcanics are of two types: sheet flows and pillow flows forming the binary pattern called "rock type." Because the 13 vents occur near the linear spreading center, they are positively correlated with youngest volcanics. The vents are slightly correlated with occurrence of pillow flows (and negatively correlated with sheet flows).

The map layer for "contact" was based on maximum contrast proximity to the contact between youngest sheet flows and pillow flows. Its pattern strongly resembles the pattern for "age." There is probably significant lack of conditional independence of "contact" and "age." Because of small sample size ($n = 13$) it is not possible to use the contingency table tests.

Although there are significant changes in depth of ocean floor below sea level, 12 of the 13 vents occur within the relatively narrow zone of 2590 m ± 10 m water depth. This is probably for hydrostatic reasons. "Topography" provides a map layer that is probably conditionally independent of the other map layers.

It can be concluded that the most likely locations of hydrothermal vents in this area are in the immediate vicinity of the linear spreading center, at about 2590 m water depth, and in the vicinity of pillow flows.

The study area is 3.7534 km^2. Setting this equal to 37,534 unit areas results in prior probability of (13/37,534 =) 0.00034635 per 10 m $\times$ 10 m cell. Consequently, the prior logit is -7.9677. The sum of all five positive weights in Table 2 is 7.357. Addition of this total weight to the prior logit yields posterior logit of -0.6107 that can be converted into a posterior probability of 0.3519. This would indicate that the most favorable cell in the 5-layer model is approximately 1000$\times$ more likely to contain a hydrothermal vent than when the prior probability would be used only.

However, the sum of positive weights for the first 3 map layers of Table 2 is 4.920. This 3-layer model would result in posterior probability of 0.0453 that is only about 130$\times$ greater than prior. Posterior probabilities resulting from the 5-layer and the 3-layer model obviously are different.

Figure 1 shows observed and estimated number of hydrothermal vents as a function of posterior probability for the 5-layer model. The sum of all posterior probabilities (T) is 37.59 and T-$n = 24.59$. The sum of variances of expected vent frequencies for all unique conditions is $s^2\ (T) = 109.42$. Consequently, $s(T) = 10.46$. Multiplication by 2.33 gives 99% confidence limit of 24.37 for T-$n = 24.59$ indicating that the conditional independence hypothesis should be rejected for the 5-layer model.

On the other hand, the sum of probabilities for the 3-layer model is $T = 14.05$ with $s(T) = 6.45$. The difference T-$n = 1.05$ is not statistically significant and the conditional independence hypothesis can be accepted for the 3-layer model. Details of how T and $s^2(T)$ were estimated for the 3-layer model are provided in the next section.

EXAMPLE OF ESTIMATION OF T AND $S^2(T)$

The 3-layer model was analyzed in a separate experiment for a slightly larger study area of 3.985 km^2. Separate areas for eight unique conditions are listed in Table 3. The first column shows these unique conditions with "1" for presence and "2" for absence of a binary pattern.

The area of the unit cell was set equal to 0.01 km^2 in this experiment. With weights and standard deviations similar to those listed in Table 2 this resulted in the posterior probabilities P_f with standard deviations $s(P_f)$ shown in columns 3 and 4 of Table 3. Multiplication of each P_f by area (number of unit cells) of its unique condition results in the eight predicted vent frequencies $N_{IJK} P_f$. Their sum provides the estimate $T = 14.05$ used in the previous section. The corresponding variance $s^2(T)$ is the sum of the eight values of $N^2_{IJK}\ s^2(P_f)$ listed in the last column of Table 3.

Table 3. Estimation of T and $s^2(T)$ for 3-Layer Model; I – age, J – Topography, K – Rock Type; $N = 100 \times$ Area

IJK	Area (km^2)	P_f	$s(P_f)$	$N_{IJK} P_f$	$N^2_{IJK} s^2(P_f)$
222	1.344	0.0003	0.0005	0.0403	0.0045
212	0.9007	0.0057	0.0065	0.5134	0.3428
221	0.4351	0.0008	0.0013	0.0348	0.0032
122	0.4187	0.0111	0.0126	0.4648	0.2783
112	0.3415	0.1709	0.0907	5.8362	9.5939
211	0.2223	0.0154	0.0176	0.3423	0.1531
111	0.1771	0.3604	0.153	6.3827	7.3421
121	0.1456	0.0297	0.336	0.4324	23.9332
Sum	3.985			14.0470	41.6511

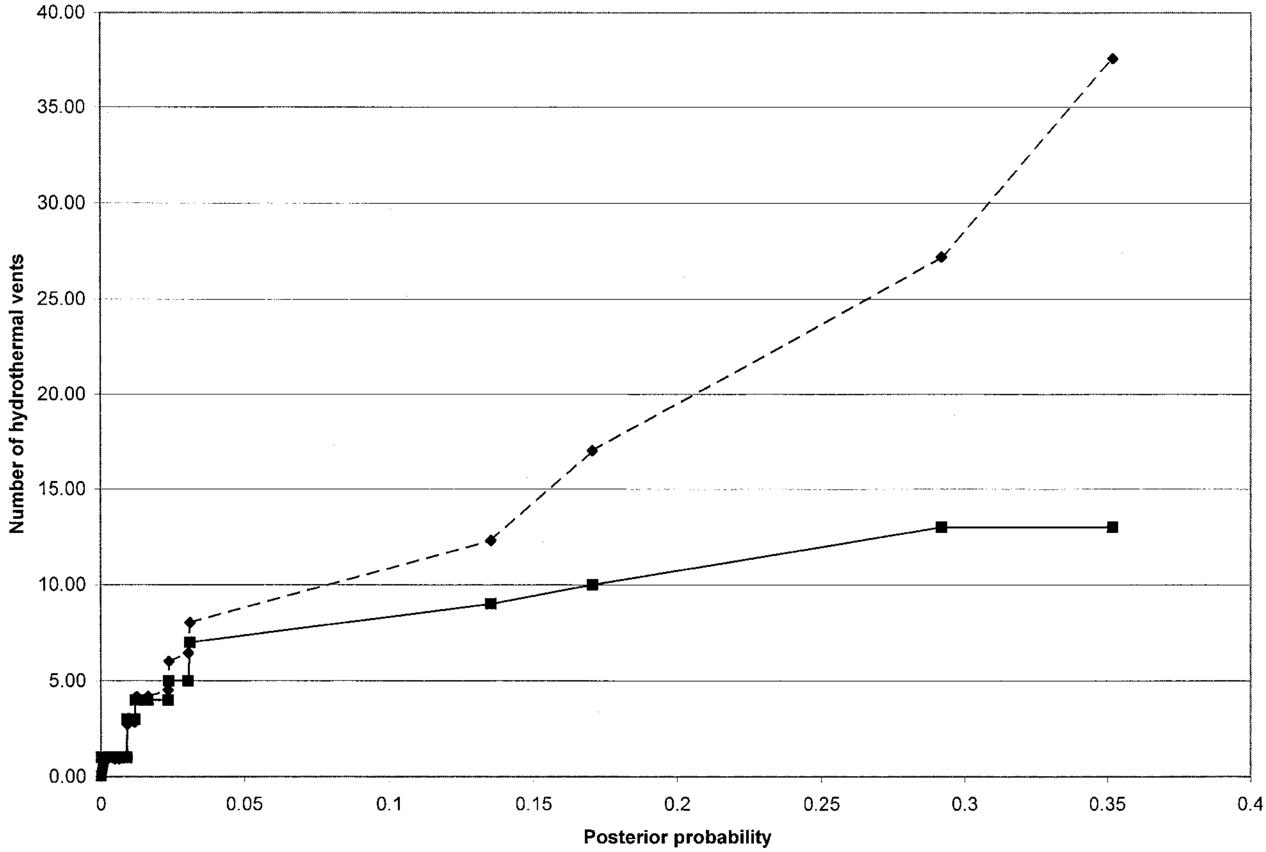

Figure 1. Hydrothermal vents on East Pacific Rise example. Estimated and observed numbers of vents in 5-layer model (adapted from Agterberg and others, 1993, Fig. 2c).

CONCLUDING REMARKS

Conditional independence of all map layers implies that the sum of posterior probabilities (T) is equal to total number of discrete events ($= n$). In practical applications, T generally exceeds n and the difference increases when additional map layers are included, indicating possible lack of conditional independence. In the past, this result was used in the overall or "omnibus test" for conditional independence.

This paper provided formal proof that the expected value of the sum of all posterior probabilities is equal to total number of discrete event if all map layers are conditionally independent. Estimation of $s(T)$ representing the standard deviation of T allows performance of a one-tailed test to check whether T-n is significantly greater than zero. This new test is exact and simpler to use than other tests including the Kolmogorov-Smirnov test and various chi-squared tests adapted from discrete multivariate statistics.

The new test was applied to occurrence of 13 hydrothermal events along the East Pacific Rise indicating that the conditional independence hypothesis can be accepted for a 3-layer model (age, topography, and rock type). This hypothesis must be rejected for a 5-layer model (addition of layers for proximity to contact between youngest pillow and sheet flows and proximity to fissures). It is noted that this conclusion was drawn previously on the basis of the Kolmogorov-Smirnov test approximation (Agterberg and others, 1993). It is also in accordance with the informal 15% rule of the earlier version of the omnibus test.

This indicates that the new test provides good results in situations that contingency table tests do not apply because sample size is too small. For larger samples (e.g. $n = 100$), and many map layers, it remains easy to apply the new omnibus test. During estimation of the standard deviation $s(T)$ it is tacitly assumed that several approximate formulas produce valid results. Possible shortcomings resulting from small sample size (e.g. approximations based on asymptotic theory) are diminished when sample size (n) is increased.

If the map layers are not conditionally independent, it is likely that the standard deviation $s(T)$ will be underestimated because positive association would result in increased variance. The effect of underestimation of $s(T)$ on the significance test would be that the hypothesis of conditional independence is more likely to be rejected than accepted.

Other approaches may not be subject to problems associated with lack of conditional independence. For example, in logistic regression the condition $T = n$ is satisfied automatically and need not be verified. However, logistic regression generally produces weights with greater standard deviations because of approximate linear relationships between map layers, lack of transparency of calculations because an iterative process is used for calculating the coefficients, and it is more difficult to cope with missing data.

The problem that T may exceed n in weights-of-evidence is unrelated to possible existence of undiscovered mineral deposits. The purpose of using this method in mineral-potential mapping is to locate new target areas for further exploration. Coloring the posterior probability map may provide a simple way to achieve this objective. Areas with high density of known deposits and low-density target areas would have the same color assigned to polygons with the largest posterior probabilities.

If there are undiscovered deposits in a region, it is likely that the prior probability is underestimated. On average, the weights are not be affected by this type of bias, provided that associations between deposits and map layers remain the same for known and unknown deposits. Presence of undiscovered deposits in the study area would imply underestimation of all posterior probabilities P_f by approximately the same factor if $P_f \ll 1$.

ACKNOWLEDGMENTS

We are indebted to Graeme Bonham-Carter for critical review of the manuscript.

REFERENCES

Agterberg, F. P., 1992, Combining indicator patterns in weights of evidence modeling for resource evaluation: Nonrenewable Resources, v. 1, no. 1, p. 35–50.

Agterberg, F. P., and Bonham-Carter, G. F., 1999, Logistic regression and weights of evidence modeling in mineral exploration, *in Proc. 28th Intern. Symp.* Computer Applications in the Mineral Industries: Golden, CO, p. 483–490.

Agterberg, F. P., Bonham-Carter, G. F., and Wright, D. F., 1990, Statistical pattern integration for mineral exploration, *in* Gaál, G., and Merriam, D. F., eds., Computer Applications in Resource Estimation: Pergamon, Oxford, p. 1–21.

Agterberg, F. P., Bonham-Carter, G. F., Cheng, Q. M., and Wright, D. F., 1993, Weights of evidence modeling and weighted logistic regression for mineral potential mapping, *in* Davis, J. C., and Herzfeld, U. C., eds., Computers in Geology—25 Years of Progress: Oxford Univ. Press, New York, p. 13–32.

Ballard, R. D., Francheteau, J., Juteau, T., Rangan, C., and Norwark, W., 1981, East Pacific Rise at 21°N: The oceanic, tectonic and hydrothermal processes of the central axis: Earth and Planetary Sci. Lett., v. 55, no. 1, p. 1–10.

Bishop, M. M., Feinberg, S. E., and Holland, P. W., 1975, Discrete multivariate analysis: theory and practice: MIT Press, Cambridge, MA, 587 p.

Boleneus, D. E., Raines, G. L., Causey, J. D., Bookstrom, A. A., Frost, T. P., and Hyndman, P. C., 2001, Assessment method for epithermal gold deposits in northeast Washington State using weights-of-evidence GIS modeling: U.S. Geol. Survey Open-File Rept. 01–501.

Bonham-Carter, G. F., 1994, Geographic Information Systems for geoscientists: Pergamon, Oxford, 398 p.

Carranza, E. J. M., and Hale, M., 2000, Geologically constrained probabilistic mapping of gold potential, Baguio District, Philippines: Natural Resources Research, v. 9, no. 3, p. 237–253.

Carranza, E. J. M., and Hale, M., 2002, Where are porphyry copper deposits spatially localized? A case study in Berguet Province, Philippines: Natural Resources Research, v. 11, no. 1, p. 45–60.

Cheng, Q. M., and Agterberg, F. P., 1999, Fuzzy weights of evidence method and its application in mineral potential mapping: Natural Resources Research, v. 8, no. 1, p. 27–36.

Harris, J. R., Wilkinson, L., Heather, K., Fumeron, S., Bernier, M. A., and Ayer, J., 2001, Application of GIS processing techniques for producing mineral prospectivity maps—a case study: Mesothermal Au in the Swayze Greenstone Belt, Ontario, Canada: Natural Resources Research, v. 10, no. 2, p. 91–124.

Kemp, L., Bonham-Carter, G. F., and Raines, G. L., 1999, Arc-Wofe: Arcview extension for weights of evidence mapping: http://gis.nrcan.gc.ca/software/arcview/wofe.

Mihalasky, M. J., and Bonham-Carter, G. F., 2001, Lithodiversity and spatial association within metallic mineral sites, Great Basin of Nevada: Natural Resources Research, v. 10, no. 3, p. 209–225.

Pan, G., and Harris, D. P., 2000, Information synthesis for mineral exploration: Oxford Univ. Press, New York, 461 p.

Raines, G. L., 1999, Evaluation of weights of evidence to predict epithermal gold deposits in the Great Basin of the Western United States: Natural Resources Research, v. 8, no. 4, p. 257–276.

Scott, M., and Dimitrakopoulos, R., 2001, Quantitative analysis of mineral resources for strategic planning: Implications for Australian Geological surveys: Natural Resources Research, v. 10, no. 3, p. 159–177.

Singer, D. A., and Kouda, R., 1999, A comparison of weights-of-evidence methods and probabilistic neural networks: Natural Resources Research, v. 8, no. 4, p. 287–298.

Venkataraman, G., Madhavan, B. B., Ratha, D. S., Antany, J. P., Goyal, R. S., Sanglani, S., and Roy, S. S., 2000, Spatial modeling for base-metal mineral exploration through integration of geological data sets: Natural Resources Research, v. 9, no. 1, p. 27–42.