

doi:10.3799/dqkx.2012.123

# 加权证据权模型的应用与对比

张道军<sup>1,2</sup>, 成秋明<sup>1</sup>, 左仁广<sup>1</sup>, 王书旺<sup>1,2</sup>

1. 中国地质大学地质过程与矿产资源国家重点实验室, 湖北武汉 430074

2. 中国地质大学资源学院, 湖北武汉 430074

**摘要:** 证据权方法是目前最常用的信息综合方法之一, 广泛应用于矿产资源定量预测与评价。然而, 它要求变量间相互独立, 地质上很难满足这一条件。如何削弱条件不独立对证据权预测结果的影响, 已成为当前数学地球科学研究的热点。解决该问题的途径之一是对传统证据权模型进行校正, 比如采取加权的方法对原证据权模型计算的证据权重进行修正, 以便消除非条件独立性的影响。对近期提出的多种加权证据权模型进行了系统的对比研究, 基于同样的应用实例和实验方案, 对不同方法的应用效果进行了比较, 结果表明, 各种加权证据权模型均可不同程度地削弱证据图层条件不独立性的影响, 其中, 基于逻辑回归的加权证据权模型优于其他加权方法。

**关键词:** 矿产资源勘查与评价; 证据权方法; 条件独立性; 数据综合; 成矿预测; 矿床。

**中图分类号:** P628

**文章编号:** 1000-2383(2012)06-1160-09

**收稿日期:** 2012-07-10

## Application and Comparison of Weighted Weights of Evidence Models

ZHANG Dao-jun<sup>1,2</sup>, CHENG Qiu-ming<sup>1</sup>, ZUO Ren-guang<sup>1</sup>, WANG Shu-wang<sup>1,2</sup>

1. State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China

2. Faculty of Earth Resources, China University of Geosciences, Wuhan 430074, China

**Abstract:** Weights of evidence (WofE) is one of the most widely used methods in prediction and evaluation of mineral resources. It requires the independence of geological variables. However, it is almost impossible to meet the requirement among geological variables. The study of reducing the influence of the independence of variables has become a hot topic. One way to solve the problem is to modify the traditional WofE model such as modifying the weights of each evidential layer by calculating correction factor. In this study, some weighted WofE models proposed in recent years are systematically studied, i. e., the same data of case and the same test program are used to compare these models. The results illustrate that weighted WofE models can reduce the influence of dependence among variables to some extent, and the weighted WofE model based on logistic regression performs better than others.

**Key words:** mineral resources exploration; weights of evidence method; condition independence; data integration; ore prospecting; ore deposits.

证据权模型是一种基于贝叶斯定理的模型, 最初应用于医学领域专家系统 GLADYS (Spiegelhalter and Knill-Jones, 1984)。20世纪80年代末, 加拿大数学地质学家 F. P. Agterberg 和 G. F. Bonham-Carter 等将该方法引入到矿产资源定量预测与评价 (Agterberg, 1989; Agterberg *et al.*, 1990)。由于证据权模型形式直观、透明, 建模过程易于解释, 符合

地质思维, 并能很好地与 GIS 空间分析功能相结合, 且易于编程实现 (Agterberg, 2011), 很多软件都实现了该方法, 使得该方法成为最流行的矿产资源预测与评价方法之一。除此之外, 该方法还应用于滑坡敏感性评价 (Neuhäuser and Terhorst, 2007; 王志旺等, 2007; Cervi *et al.*, 2010; Netra *et al.*, 2010) 和生态环境 (Romero-Calcerrada and Luque, 2006;

**基金项目:** “覆盖区矿产综合预测”计划项目 (Nos. 1212011085468, 1212011085466); 国家自然科学基金 (No. 41002118); 中央高校基本科研业务费专项资金 (Nos. CUG120116, CUG120501); 国土资源部公益性行业科研专项 (No. 201211022); 地质过程与矿产资源国家重点实验室科技部专项经费资助 (No. MSFGPMR201203)。

**作者简介:** 张道军 (1985—), 男, 博士研究生, 主要研究方向为数学地质及国土资源信息化。E-mail: cugzdj@gmail.com

张生元等,2006; Cho *et al.*, 2008; Romero-Calcerada *et al.*, 2010; Gorney *et al.*, 2011)等领域. 使用证据权的前提是证据图层间具有相互独立性,然而,矿产资源预测与评价使用的地质变量间很难满足这一条件,使得数据综合的结果(后验概率)有偏差. 如何克服证据图层间的不独立或削弱证据图层间不独立对预测的影响,已成为矿产资源定量预测与评价研究的热点. 如 Bonham-Carter *et al.* (1989)、Agterberg (1992)、Bonham-Carter (1994)、Agterberg and Cheng (2002)等先后提出了条件独立性检验方法,主要包括成对检验、综合检验、新综合检验等; Bonham-Carter (1994)提出采用主成分分析技术削弱变量间相关关系; Cheng (2008)提出了校正后验概率的方法; Journal (2002)从如何解决数据冗余角度提出了 tau 模型,该模型与后验 Logit 形式上是一致的,可认为是广义形式,为后续模型的改进提供了基础. 依据 tau 模型,可从如下 2 个方面对证据权进行改进:第一,对普通证据权所计算的正负权重( $W^+$ 和 $W^-$ )进行修正,不考虑证据图层进入顺序,这类改进的模型称为加权证据权模型;第二,按一定的顺序纳入证据图层进行计算,后进入图层权重的计算依赖于之前进入图层与预测矿床(点)的空间关系. 如 Journal (2002)提出了用 Simple kriging(S-K)方程组求解证据图层的权重; Krishnan *et al.* (2005)提出了 2 种计算 tau 模型的方法:一种是基于条件相关性的顺序拟合方法,一种是基于随机指示克立格的方法; Polyakova and Journal (2007)提出了新的 nu 模型,给出了一种新的 tau 模型的表达形式;张生元等(2009)提出了秩相关系数加权方案; Deng (2009)提出校正证据权模型,该模型试图通过校正相关系数矩阵,来弱化证据图层间的重叠信息; Agterberg (2011)提出了一种修正证据权模型,该模型将证据权模型和逻辑回归方法结合,使得后验概率结果是无偏估计. 除了上述基于加权修正证据权模型的方法外,采取了完全不同的解决方案 (Cheng, 2011; 成秋明, 2012),提出了增强证据权模型(BoostWofE),本质上属于逐步证据权模型的一种. 本文将主要对加权证据权模型进行对比研究.

## 1 普通证据权模型

设有  $n$  个二值证据图层,  $B_i^*$  表示第  $i$  个证据图层, 0 表示证据图层不存在(此时 \* 为一), 1 表示证据图层存在(此时 \* 为+);  $D^*$  表示矿床(点), 1 为

含矿单元(此时 \* 为+), 0 为不含矿单元(此时 \* 为一).

第  $i$  个证据图层 \* 状态下的证据权重可表示为:

$$W_i^* = \ln \frac{P(B_i^* | D^+)}{P(B_i^* | D^-)}. \quad (1)$$

若  $n$  个证据图层满足条件独立性,则可用下式得到综合后的证据权重:

$$\ln(O(D^+ | B_1^* B_2^* \cdots B_n^*)) = W_0 + \sum_{i=1}^n W_i^*, \quad (2)$$

其中:  $O$  代表概率的比率;  $W_0 = \ln \frac{P(D^+)}{P(D^-)}$ , 为先验概率几率的对数.

通过以下公式,可计算出后验概率 (Bonham-Carter *et al.*, 1988, 1989; Agterberg, 1989; Agterberg *et al.*, 1990).

$$P(D^+ | B_1^* B_2^* \cdots B_n^*) = \frac{e^{\ln(O(D^+ | B_1^* B_2^* \cdots B_n^*))}}{1 + e^{\ln(O(D^+ | B_1^* B_2^* \cdots B_n^*))}}. \quad (3)$$

## 2 加权证据权模型

为了克服证据图层间条件不独立性的影响, Journal (2002)提出了 tau 模型,该模型可通过后验 Logit 模型转化成普通证据权模型的形式:

$$\ln(O(D^+ | B_1^* B_2^* \cdots B_n^*)) = W_0 + \sum_{i=1}^n T_i W_i^*, \quad (4)$$

其中:  $T_i = \frac{\ln \frac{P(B_i^* | B_1^* B_2^* \cdots B_{i-1}^* D^+)}{P(B_i^* | B_1^* B_2^* \cdots B_{i-1}^* D^-)}}{\ln \frac{P(B_i^* | D^+)}{P(B_i^* | D^-)}}$ , 当

$i=1$  时,  $T_1=1$ .

此式为 tau 模型的表达式. 这里隐含了证据图层进入模型顺序的问题,不同的图层进入顺序会得到不同的  $T_i$ ,但公式(4)的计算结果却与图层进入顺序无关. 精确计算  $T_i$  实际上对预测意义不大,因为那样就变成了完全的后验 Logit 模型,所有图层组合的后验概率都是确定的,没有任何估计. 公式(4)的意义在于它的形式,当  $T_i=1$  ( $i$  取 1 到  $n$ ) 时,它就是普通证据权模型,因此普通证据权可以看作是 tau 模型的特例. 用不同的方法对  $T_i$  进行估计,可得到不同的改进证据权模型. 如果  $T_i$  的估计结果不随图层进入顺序的变化而变化,就可以看作加权

证据权模型,接下来介绍 4 种常用加权证据权模型。

## 2.1 S-K 方程组加权方案

Journal(2002) 建议用 S-K 方程组求解  $T_i$ , 如下式:

$$\sum_{j=1}^n T_j \cdot \gamma(B_i, B_j) = \gamma(B_i, D), \quad (5)$$

其中:  $i, j$  表示图层编号, 可取 1 到  $n$ ;  $n$  表示图层数;  $\gamma$  表示距离. 由于  $i$  取值为 1 到  $n$ , 由公式(5)可构成  $n$  元一次方程组. 如果能有效定义  $\gamma(B_i, B_j)$ , 便能计算出该方程组的唯一解  $(T_1, T_2, \dots, T_n)$ . 对  $\gamma(B_i, B_j)$  的不同定义, 将得到不同的加权系数.  $\gamma(B_i, B_j)$  可定义为一种距离或广义的变异函数. 由于变量(证据图层)的个数一般远小于样本(预测单元)的数量, 因此无法求得马氏距离, 本研究将采用欧氏距离计算  $\gamma(B_i, B_j)$ .

## 2.2 Deng 加权方案

Deng(2009) 提出了一种校正证据权模型, 认为普通证据权采用的证据权重, 实际上包含了其他图层的一些信息, 亦即单个图层权重受到其他图层权重影响, 可通过系数矩阵进行校正. 实质上也可以看作是一种加权证据权模型. 然而 Schaeben and Boogart(2011) 通过特例证明 Deng 的模型存在自相矛盾的地方, 不能推广到所有情况. Schaeben(2012) 指出, 其推导存在一定问题. 目前的研究表明, Deng 的错误可能在于用几何平均代替算术平均. 如果是这样, 则表明 Deng 的推导虽然错了, 但是仍然可能在一定程度上起到削弱条件不独立性的影响, 因为几何平均在一定程度上近似等于算术平均. 有关 Deng 方法原理部分仍然需要进一步讨论和研究, 本文只从应用角度考察该方法的有效性.

## 2.3 逻辑回归加权方案

Agterberg(2011) 将证据权模型和逻辑回归方法结合, 提出了一种修正证据权模型. 该方法和逻辑回归一样, 可以得到后验概率的无偏估计(即预测事件发生数等于实际事件发生数), 同时还保留了传统证据权的形式及优点.

逻辑回归加权方案的实施分如下 3 步: (1) 计算各证据图层的正负权重  $(W_i^+, W_i^-)$ ; (2) 用  $W_i^+$ 、 $W_i^-$  代替初始的“1”、“0”值, 进行逻辑回归建模, 进而得到每个证据图层的偏回归系数  $\beta_i'$ ; (3) 将  $\beta_i'$  作修正系数, 对  $W_i^+$ 、 $W_i^-$  进行修正, 得到修正的证据权重, 见如下公式:

$$\begin{aligned} W_{Ni}^+ &= \beta_i' W_i^+ \\ W_{Ni}^- &= \beta_i' W_i^-, \end{aligned} \quad (6)$$

其中:  $\beta_i'$  是逻辑回归表达式  $\text{Logit}'(D | B_1 * B_2 * \dots * B_n) = \sum_{i=1}^n \beta_i' x_i + a'$  中的偏回归系数,  $x_i$  的取值为  $W_i^+$  或  $W_i^-$ ,  $W_{Ni}^*$  表示修正后的证据权重.

可以证明, Agterberg(2011) 模型与逻辑回归模型是等价的, 并且 Agterberg(2011) 模型中的  $\beta_i'$  就等于该证据图层对应的偏回归系数与回归系数的比值.

## 2.4 秩相关系数加权方案

如果第  $i$  个证据图层(简称图层  $i$ ) 与其他  $n-1$  个证据图层空间相关性很大, 则图层  $i$  与其他  $n-1$  个证据图层按照普通证据权进行综合后所得到的后验概率图层空间相关性很强, 反之亦然, 这就是秩相关系数加权方案的原理.

张生元等(2009) 在 Journal 等人的研究基础上, 给出如下秩相关系数加权算法(以计算第  $i$  个证据图层加权系数为例):

- (1) 以除第  $i$  个证据变量之外的  $n-1$  个证据变量进行普通证据权建模, 并得到后验概率排序;
- (2) 以第  $i$  个证据变量单独进行普通证据权建模, 得到此状态下的后验概率排序;
- (3) 在前 2 步基础上计算秩相关系数  $r_i$ , 并通过  $T_i = 1 - r_i$  计算加权系数;
- (4) 证据权重加权综合, 并计算后验概率.

该模型在计算秩相关系数时是针对 2 个变量进行的, 一个是二值证据图层变量(排序值分别是 1 和 2), 另一个是其他图层的综合(排序值为 1 到  $2^{n-1}$ ,  $n$  表示图层数), 这便具有很大的不稳定性, 因此本文对计算秩相关系数的方法进行了改进, 见下一小节.

## 2.5 改进的秩相关系数加权方案

改进的秩相关系数加权方案原理同上, 只是在秩相关系数求取方法上有所不同, 算法如下:

- (1) 将除第  $i$  个证据图层之外的其他证据图层与目标图层一起进行普通证据权建模, 则可以得到  $k$  ( $k \leq 2^{n-1}$ ) 个唯一值区域, 理想状况下  $k = 2^{n-1}$ . 对这些区域按照后验概率由高到低进行排序并编号;
- (2) 将第  $i$  个证据图层与步骤(1)得到的每个唯一值区域进行叠加, 可以将  $k$  个唯一值区域分为分别与  $B_j^+$  与  $B_j^-$  相交的两部分, 其面积分别记为  $Q_{j-1}$  和  $Q_{j-0}$ , 其中  $j = 1 \sim k$ . 对所有  $Q_{j-0} > 0$  的区域按照  $Q_{j-1}/Q_{j-0}$  由高到低进行排序并编号;
- (3) 计算  $k$  个唯一值区域在步骤(1)、(2)两种排序下的秩相关系数  $r_i$ , 从而得到第  $i$  个证据图层的

加权系数  $1-r_i$ ;

(4)对其他证据图层重复步骤(1)~(3),即可得到所有证据图层的加权系数。

### 3 实例研究

这里使用的研究数据为真实数据,来自 Agterberg *et al.* (1993)、Agterberg and Cheng(2002)和 Agterberg(2011)的东太平洋海岭火山通道预测研究,实验数据见表 1。

研究的目标图层为海底火山通道,该图层共有 13 个训练样点,5 个预测要素,见表 1,从第 2 列至第 6 列依次为:(1)海底玄武岩相对年龄;(2)海底高程;(3)距新形成玄武岩接触带的距离;(4)玄武岩类型;(5)距海底裂隙的距离。所有这 5 个图层均为二值化变量,阈值采用 Agterberg(2011)中方案。另外,表 1 的第 1 列为唯一值单元的编号,第 7 列和第 8 列分别为该唯一值单元中含火山通道的栅格数和总栅格数。

研究区面积约 3.99 km<sup>2</sup>,划分为 39 851 个单元格,单元格大小为 10 m×10 m,其中有火山通道的单元 13 个,先验概率为 0.032 6%。

研究分别采用普通证据权、S-K 方法加权证据权、本文改进的秩相关系数加权证据权、Deng(2009)方法加权证据权以及逻辑回归方法加权证据权这 5 种模型对表 1 的数据进行证据权建模,各方法加权系数及权重计算情况见表 2 和表 3,最终的后验概率分布见表 4。

虽然从原理上说,4 种加权方法均能减弱条件不独立性对预测结果的影响,但各方法的具体表现有差异。从表 2 看,各种加权方法加权系数数值大小差异较大,特别是 S-K 方法系数值偏小;从各方案加权系数相对大小看,逻辑回归加权方法与 Deng(2009)加权方法最为接近,而秩相关系数加权方法介于这 2 种方法与 S-K 方法之间。具体来说,S-K 方法的结果表明“裂隙距离”与其他图层信息重叠最多,其次是“高程”图层,秩相关系数加权方法的结果表明“年龄”和“高程”与其他图层相关性强,而逻辑回归加权和 Deng 加权方法“意见一致”,认为“岩石类型”和“高程”需要削弱影响。此外,除 S-K 之外的方法的结果均表明“裂隙距离”与其他图层在火山通道事件上条件相关性最小,其次是“接触距离”图层。

由表 3 可以明显看到,加权证据权模型对普通证据权的原始权重具有调节作用,原始权重被加大

表 1 实验数据(据 Agterberg *et al.*, 1993)

Table 1 Test data

编号	年龄	高程	接触距离	岩石类型	裂隙距离	通道数	面积
1	0	0	0	0	1	0	10 052
2	0	0	0	0	0	1	3 363
3	0	0	0	1	1	0	3 268
4	0	0	0	1	0	0	1 074
5	0	1	0	0	1	0	5 455
6	0	1	0	0	0	0	25
7	0	0	1	0	1	0	3 482
8	0	1	0	1	1	0	2 518
9	0	0	1	0	0	0	1 474
10	0	1	0	1	0	0	1 371
11	1	0	0	0	1	0	5
12	1	0	0	0	0	0	705
13	0	0	1	1	1	0	5
14	0	0	1	1	0	0	744
15	1	0	0	1	1	0	422
16	1	0	0	1	0	0	58
17	0	1	1	0	1	0	12
18	0	1	1	0	0	0	179
19	1	1	0	0	1	2	1 766
20	1	1	0	0	0	0	119
21	0	1	1	1	1	1	1 055
22	0	1	1	1	0	0	33
23	1	0	1	0	1	0	10
24	1	1	0	1	1	0	146
25	1	0	1	0	0	1	623
26	1	1	0	1	0	0	145
27	1	0	1	1	1	2	504
28	1	0	1	1	0	0	1
29	1	1	1	0	1	2	317
30	1	1	1	0	0	1	277
31	1	1	1	1	1	3	348
32	1	1	1	1	0	0	295

或缩小,但是各种方法所得到的最终权重大小排序却很接近,这主要是因为这些加权方法均是在普通证据权基础上对权重进行调整,其“基调”已由普通证据权模型确定好,具体来看,S-K 和秩相关系数加权方法与普通证据权完全一致,而逻辑回归和 Deng(2009)加权方法与普通证据权略有不同。

根据后验概率可以计算出预测火山通道数,普通证据权、S-K 加权、秩相关系数加权、逻辑回归加权及 Deng(2009)加权方法的结果分别为:24.8,8.4,15.4,13.0 和 16.8 个,而实际发生数为 13 个,加权方法得到的预测数值都比普通证据权更接近真实值,可见在实际运用中,这些加权方法都在一定程度上克服了证据图层间条件不独立的影响;逻辑回归加权得到的预测事件数与实际发生数是一样的,因为该方法与逻辑回归等价,而逻辑回归是无偏估计。

表 2 各加权证据权模型加权系数

Table 2 Weighted coefficients of different kinds of weighted WofE model

证据图层	S-K 加权	秩相关系数加权	逻辑回归加权	* Deng 加权 $W^+$	* Deng 加权 $W^-$
年龄	0.44	0.76	0.84	0.82	0.80
高程	0.14	0.82	0.53	0.67	0.76
接触距离	0.27	0.94	0.84	0.90	0.93
岩石类型	0.21	0.91	0.00	0.71	0.45
裂隙距离	0.05	1.11	5.07	6.67	2.61

注: \* Deng 的加权方案下,  $W^+$  和  $W^-$  加权系数不同。

表 3 各种证据权模型的权重

Table 3 Weights of different kinds of WofE model

图层	原始权重		S-K 加权		秩相关系数加权		逻辑回归加权		Deng 加权	
	$W^+$	$W^-$	$W^+$	$W^-$	$W^+$	$W^-$	$W^+$	$W^-$	$W^+$	$W^-$
年龄	1.77	-1.72	0.78	-0.76	1.34	-1.30	1.49	-1.44	1.45	-1.37
高程	0.67	-0.74	0.09	-0.10	0.55	-0.61	0.36	-0.39	0.45	-0.56
接触距离	1.19	-1.20	0.33	-0.33	1.11	-1.12	1.00	-1.01	1.07	-1.12
岩石类型	0.43	-0.26	0.09	-0.05	0.39	-0.24	0.00	0.00	0.30	-0.12
裂隙距离	0.04	-0.13	0.00	-0.01	0.05	-0.15	0.22	-0.67	0.29	-0.34

表 4 各证据权模型预测后验概率

Table 4 Posterior probabilities using different kinds of WofE model

编号	普通证据权	S-K 加权	秩相关系数加权	Deng 加权	逻辑回归加权
1	7.00E-06	9.40E-05	1.31E-05	1.83E-05	2.36E-05
2	1.30E-05	1.09E-04	2.45E-05	2.79E-05	2.36E-05
3	1.10E-05	1.08E-04	2.02E-05	1.49E-05	9.76E-06
4	2.80E-05	1.15E-04	4.17E-05	5.04E-05	5.01E-05
5	2.30E-05	1.14E-04	3.44E-05	2.69E-05	2.07E-05
6	7.30E-05	1.82E-04	1.22E-04	1.63E-04	1.77E-04
7	5.60E-05	1.33E-04	7.80E-05	7.67E-05	5.01E-05
8	6.20E-05	1.80E-04	1.00E-04	8.68E-05	7.30E-05
9	4.70E-05	1.32E-04	6.43E-05	4.09E-05	2.07E-05
10	2.21E-04	4.38E-04	1.82E-04	3.08E-04	4.40E-04
11	1.86E-04	4.35E-04	1.51E-04	1.64E-04	1.82E-04
12	1.46E-04	2.10E-04	2.28E-04	2.48E-04	1.77E-04
13	1.23E-04	2.08E-04	1.88E-04	1.32E-04	7.29E-05
14	4.41E-04	5.06E-04	3.42E-04	4.68E-04	4.39E-04
15	3.70E-04	5.02E-04	2.82E-04	2.49E-04	1.82E-04
16	3.03E-04	2.21E-04	3.89E-04	4.48E-04	3.74E-04
17	2.55E-04	2.19E-04	3.20E-04	2.39E-04	1.55E-04
18	7.67E-04	5.29E-04	4.80E-04	4.51E-04	3.85E-04
19	5.07E-04	2.53E-04	6.00E-04	3.63E-04	1.55E-04
20	2.40E-03	8.44E-04	1.70E-03	2.73E-03	3.28E-03
21	1.82E-03	6.15E-04	1.09E-03	1.29E-03	9.31E-04
22	1.53E-03	6.10E-04	8.99E-04	6.86E-04	3.85E-04
23	4.01E-03	9.66E-04	2.62E-03	2.21E-03	1.36E-03
24	1.64E-02	1.17E-03	8.32E-03	6.07E-03	2.87E-03
25	6.00E-06	9.40E-05	1.08E-05	9.76E-06	9.77E-06
26	6.04E-04	2.55E-04	7.27E-04	6.82E-04	3.74E-04
27	2.02E-03	8.37E-04	1.40E-03	1.46E-03	1.36E-03
28	8.27E-03	1.02E-03	4.46E-03	3.99E-03	2.87E-03
29	9.12E-04	5.33E-04	5.82E-04	8.46E-04	9.31E-04
30	4.77E-03	9.73E-04	3.18E-03	4.15E-03	3.28E-03
31	9.83E-03	1.03E-03	5.40E-03	7.47E-03	6.92E-03
32	1.94E-02	1.18E-03	1.01E-02	1.13E-02	6.92E-03

表 5 各证据权模型预测后验概率排序

Table 5 The rank of posterior probabilities using different kinds of WofE model

编号	普通证据权	S-K 加权	秩相关系数加权	Deng 加权	逻辑回归加权
1	31	31	31	30	27
2	29	29	29	28	28
3	30	30	30	31	32
4	27	27	27	26	25
5	28	28	28	29	29
6	23	23	23	22	19
7	25	25	25	25	26
8	24	24	24	24	23
9	26	26	26	27	30
10	19	15	21	17	11
11	20	16	22	21	17
12	21	21	19	19	20
13	22	22	20	23	24
14	15	13	16	13	12
15	16	14	18	18	18
16	17	19	15	15	15
17	18	20	17	20	21
18	12	12	14	14	13
19	14	18	12	16	22
20	7	7	7	6	3
21	9	9	9	9	10
22	10	10	10	11	14
23	6	6	6	7	8
24	2	2	2	3	6
25	32	31	32	32	31
26	13	17	11	12	16
27	8	8	8	8	7
28	4	4	4	5	5
29	11	11	13	10	9
30	5	5	5	4	4
31	3	3	3	2	1
32	1	1	1	1	2

表 6 各证据权模型之间秩相关系数

Table 6 Rank correlation coefficients between posterior probabilities obtained from different WofE models

	普通证据权	S-K 加权	秩相关系数加权	Deng 加权	逻辑回归加权
不加权	1.00				
S-K 加权	0.99	1.00			
秩相关系数加权	0.99	0.96	1.00		
Deng 加权	0.99	0.98	0.98	1.00	
逻辑回归加权	0.94	0.96	0.92	0.97	1.00

各种证据权方法得到的后验概率由高到低排序见表 5,它们之间的秩相关系数见表 6.

从表 5 和表 6 可以看出,各种方法对后验概率的相对大小预测总体保持一致,各方法所得到的后验概率排序相关系数均在 0.9 以上;在各种加权证据权模型中,逻辑回归加权模型与普通证据权差别

最大(相关系数为 0.94),其余 4 种加权模型与普通证据权差别较小(相关系数均为 0.99);逻辑回归加权可以达到无偏估计,而 Deng 的加权模型与逻辑回归加权模型最为接近.

为了检验预测结果与已知训练样本之间吻合程度的好坏,分别按照各加权方案的预测后验概率由

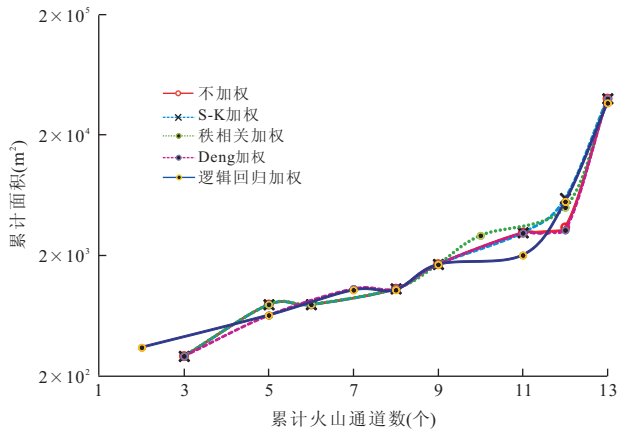


图 1 各证据权模型预测效果比较

Fig. 1 Plot of the number of events versus the area

高到低进行排序,统计累计火山通道数量和所需面积大小,并将其绘制在以实际火山通道累计数为横坐标、对应累计面积为纵坐标的二维图上(图 1)。预测成功同样多的火山通道数,所需累计面积越小,表明预测成本越低,因此图 1 中靠近底端的曲线所对应加权方法是相对较优的。据此,图 1 按横坐标从左至右可以大致分为 4 个阶段:第 1 阶段,累计火山通道数 $\leq 6$ ,Deng 加权方案占优;第 2 阶段,累计火山通道数 $> 6$  而 $< 8$ ,普通证据权、S-K 加权方案及秩相关系数加权方案占优;第 3 阶段,累计火山通道数 $\geq 8$  而 $\leq 11$ ,逻辑回归加权方案占优;第 4 阶段,累计火山通道数 $> 11$  且 $< 13$ ,Deng 加权方案占优。虽然总体上来说各种方法预测结果相差不大(图 1 中各曲线呈胶着状态),但相对来说,逻辑回归加权方案和 Deng 加权方案的预测效率较高,这也从实施效果上证明了 Deng(2009)加权模型的有效性;而 S-K 加权方案和秩相关系数加权方案与普通证据权表现非常接近,可见从最终预测效果的角度看,这 2 种加权方法改进不大。

## 4 结论

证据权模型以其建模简单、形式直观和容易实现等优点成为矿产资源预测与评价的重要方法之一。然而,该方法要求预测变量间相互独立,这一条件很难在地质上满足。因此,很多学者试图从不同角度对证据权模型进行改进,本文对比分析了多种加权证据权模型,得到如下认识:

(1)虽然各种加权方案的计算都表明各证据图层之间存在不同程度的条件相关性,但是最终的预

测结果差异并不大,各区域的后验概率排序具有很高的相关性(表 6);特别是对“最有利”区域的预测,各种方法下概率最高的区域变化很小(表 5)。因此,普通证据权方法仍然不失为一种简单而有效的靶区圈定方法。

(2)加权证据权可以消除或削弱由于证据图层条件不独立引起的后验概率偏差。从预测效果看,基于逻辑回归的加权方法表现更优,具有更高的预测效率。

本文从应用角度研究比较了各种加权证据权模型的原理及预测效果,以期对证据权模型的使用者在选用加权方案时有所帮助,必须指出的是:(1)更加科学和客观的评价需要大量实例和数据,而本文只用了一个案例;(2)对预测方法的评价,除了要考察其与已知矿点的吻合性,更应该评价其“外推能力”,即方法对训练样本以外事件的预测效果,这需要更多的已知信息。后续的研究将在这些方面加强;(3)由于篇幅所限,文中仅仅探讨加权证据权的几种模型,今后还需全面系统地对比其他新的方法,如增强证据权方法。

## References

- Agterberg, F. P., 1989. Computer programs for mineral exploration. *Science*, 245(4913): 76–81. doi: 10.1126/science.245.4913.76
- Agterberg, F. P., 1992. Combining indicator patterns in weights of evidence modeling for resource evaluation. *Nonrenewal Resources*, 1(1): 39–50. doi: 10.1007/BF01782111
- Agterberg, F. P., 2011. A modified weights-of-evidence method for regional mineral resource estimation. *Natural Resources Research*, 20(2): 95–101. doi: 10.1007/s11053-011-9138-0
- Agterberg, F. P., Cheng, Q. M., 2002. Conditional independence test for weights-of-evidence modeling. *Natural Resources Research*, 11(4): 249–55. doi: 10.1023/A:1021193827501
- Agterberg, F. P., Bonham-Carter, G. F., Cheng, Q. M., et al., 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. In: Davis, J. C., Herzfeld, U. C., eds., *Computers in geology—25 years of progress*. Oxford University Press, New York, 13–32.
- Agterberg, F. P., Bonham-Carter, G. F., Wright, D. F., 1990. Statistical pattern integration for mineral exploration. In: Gaál, G., Merriam, D. F., eds., *Computer ap-*

- plications in resource estimation prediction and assessment of metals and Petroleum. Pergamon Press, New York, 1—12.
- Bonham-Carter, G. F. , 1994, Geographic information systems for geoscientists: modeling with GIS. Oxford, Pergamon, 398.
- Bonham-Carter, G. F. , Agterberg, F. P. , Wright, D. F. , 1988. Integration of geological datasets for gold exploration in Nova Scotia. *Photogrammetric Engineering & Remote Sensing*, 54(11): 1585—1592.
- Bonham-Carter, G. F. , Agterberg, F. P. , Wright, D. F. , 1989. Weights of evidence modelling: a new approach to mapping mineral potential. In: Agterberg, F. P. , Bonham-Carter, G. F. , eds. , Statistical applications in the earth sciences. Canadian Government Publishing Centre, Ottawa, 171—183.
- Cervi, F. , Berti, M. , Borgatti, L. , et al. , 2010. Comparing predictive capability of statistical and deterministic methods for landslide susceptibility mapping: a case study in the northern Apennines (Reggio Emilia Province, Italy). *Landslides*, 7(4): 433—444. doi: 10. 1007/s10346—010—0207—y
- Cheng, Q. M. , 2008. Non-linear theory and power-law models for information integration and mineral resources quantitative assessments. *Mathematical Geosciences*, 40 (5): 503—532. doi: 10. 1007/s11004—008—9172—6
- Cheng, Q. M. , 2011. Integration of AdaBoost and weights of evidence model for mineral potential probabilistic mapping. Session MG3 “Geo-process Modeling” at IAMG 2011 Annual Conference , Salzburg, Austria.
- Cheng, Q. M. , 2012. Application of a newly developed boost weights of evidence model (BoostWofE) for mineral resources quantitative assessments. *Journal of Jilin University: Earth Science Edition* (in press).
- Cho, S. H. , Poudyal, N. C. , Roberts, R. K. , 2008. Spatial analysis of the amenity value of green open space. *Ecological Economics*, 66(2—3): 403—416. doi:10. 1016/j. ecolecon. 2007. 10. 012
- Deng, M. F. , 2009. A conditional dependence adjusted weights of evidence model. *Natural Resources Research*, 18(4): 249—258. doi: 10. 1007/s11053—009—9101—5
- Gorney, R. M. , Ferris, D. R. , Ward, A. D. , et al. , 2011. Assessing channel-forming characteristics of an impacted headwater stream in Ohio, USA. *Ecological Engineering*, 37(3): 418—430. doi:10. 1016/j. ecoleng. 2010. 11. 013
- Journal, A. G. , 2002. Combining knowledge from diverse sources: an alternative to traditional conditional independence hypotheses. *Mathematical Geosciences*, 34(5): 573—596. doi: 10. 1023/A: 1016047012594
- Krishnan, S. , Boucher, A. , Journal, A. G. , 2005. Evaluating information redundancy through the Tau model. *Quantitative Geology and geostatistics*, 14(5): 1037—1046. doi: 10. 1007/978—1—4020—3610—1\_108
- Netra, R. , Regmi, N. R. , Giardino, J. R. , et al. , 2010. Modeling susceptibility to landslides using the weight of evidence approach: western Colorado, USA. *Geomorphology*, 115(1—2): 172—187. doi:10. 1016/j. geomorph. 2009. 10. 002
- Neuhäuser, B. , Terhorst, B. , 2007. Landslide susceptibility assessment using “weights-of-evidence” applied to a study area at the Jurassic escarpment (SW-Germany). *Geomorphology*, 86 (1—2): 12—24. doi: 10. 1016/j. geomorph. 2006. 08. 002
- Polyakova, E. I. , Journal, A. G. , 2007. The Nu expression for probabilistic data integration. *Mathematical Geology*, 39 (8): 715—733. doi: 10. 1007/s11004—007—9117—5
- Romero-Calcerrada, R. , Barrio-Parra, F. , Millington, J. D. A. , et al. , 2010. Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid (central Spain). *Ecological Modelling*, 221 (1) : 34—45. doi: 10. 1016/j. ecolmodel. 2009. 08. 008
- Romero-Calcerrada, R. , Luque, S. , 2006. Habitat quality assessment using weights-of-evidence based GIS modeling: the case of *picooides tridactylus* as species indicator of the biodiversity value of the finnish forest. *Ecological Modelling*, 196(1—2): 62—76. doi: 10. 1016/j. ecolmodel. 2006. 02. 017
- Schaeben, H. , 2012. Comparison of mathematical methods of potential modeling. *Mathematical Geosciences*, 44 (1): 101—129. doi: 10. 1007/s11004—011—9373—2
- Schaeben, H. , Boogart, K. G. V. D. , 2011. Comment on “a conditional dependence adjusted weights of evidence model” by Minfeng Deng in natural resources research 18(2009), 249—258. *Natural Resources Research*, 20 (4): 401—406. doi: 10. 1007/s11053—011—9146—0
- Spiegelhalter, D. J. , Knill-Jones, R. P. , 1984. Statistical and knowledge-based approaches to clinical decision-support systems, with an application in gastroenterology. *Journal of the Royal Statistical Society (Ser. A)*, 147(1): 35—77. doi: 10. 2307/2981737
- Wang, Z. W. , Li, D. Y. , Wang, X. G. , 2007. Zonation of landslide hazards based on weights of evidence model. *Chinese Journal of Geotechnical Engineering*, 29 (8): 1268—1273 (in Chinese with English abstract).
- Zhang, S. Y. , Cheng, Q. M. , Zhang, S. P. , et al. , 2009. Weighted weights of evidence and stepwise weights of evidence and their application in Sn-Cu mineral potential



mapping in Gejiu, Yunnan Province, China. *Earth Science—Journal of China University of Geosciences*, 34 (2): 281–286(in Chinese with English abstract).

Zhang, S. Y., Wu, Q., Cheng, Q. M., et al., 2006. Weights of evidence method based on fuzzy training layer and its application in desertification assessment. *Earth Science—Journal of China University of Geosciences*, 31 (3): 389–393(in Chinese with English abstract).

定量评价中的应用. 吉林大学学报: 地球科学版(待刊).

王志旺, 李端有, 王湘桂, 2007. 证据权法在滑坡危险度区划研究中的应用. 岩土工程学报, 29(8): 1268–1273.

张生元, 成秋明, 张素萍, 等, 2009. 加权证据权模型和逐步证据权模型及其在个旧锡铜矿产资源预测中的应用. 地球科学——中国地质大学学报, 34(2): 281–286.

张生元, 武强, 成秋明, 等, 2006. 基于模糊预测对象的证据权方法及其在土地沙漠化评价中的应用. 地球科学——中国地质大学学报, 31(3): 389–393.

附中文参考文献

成秋明, 2012. 增强证据权(BoostWofE)新方法在矿产资源

\*\*\*\*\*

《地球科学——中国地质大学学报》

2013 年 第 38 卷 第 1 期 要目预告

中国天山冰川积雪中 NO <sup>-</sup> 与 NH <sub>4</sub> <sup>+</sup> 的分布特征及其环境意义 .....	张晓宇等
南海神狐海域水合物钻探区钙质超微化石生物地层与沉积速率 .....	周 洋等
周口店岩体矿物学、年代学、地球化学特征及其岩浆起源与演化 .....	陶继东等
LGM 热带西太平洋硅藻席古生产力定量评估 .....	常凤鸣等
宁镇地区安基山侵入岩 SHRIMP 锆石 U—Pb 年龄其地质意义 .....	陆建培等
临南洼陷沙三段异常地层压力及其演化特征 .....	李纯泉等
珠江口盆地惠州凹陷古近系烃源岩特征及资源预测 .....	曹婷婷等
塔里木盆地库车坳陷致密砂岩气藏成因类型 .....	孟庆宽等