

DOI: 10.13745/j.esf.sf.2024.5.13

基于大数据关联规则算法的卡林型金矿床元素富集规律及找矿方法研究

曹胜桃^{1,2}, 胡瑞忠^{1,2,*}, 周永章^{3,*}, 刘建中^{4,5}, 谭亲平¹, 高伟¹, 郑禄林⁵,
郑禄璟⁵, 宋威方⁵

1. 中国科学院地球化学研究所 矿床地球化学国家重点实验室, 贵州 贵阳 550081

2. 中国科学院大学 地球与行星科学学院, 北京 100049

3. 中山大学 地球环境与地球资源研究中心, 广东 广州 510275

4. 贵州省地质矿产勘查开发局, 贵州 贵阳 550004

5. 贵州大学, 贵州 贵阳 550025

CAO Shengtao^{1,2}, HU Ruizhong^{1,2,*}, ZHOU Yongzhang^{3,*}, LIU Jianzhong^{4,5},
TAN Qinqing¹, GAO Wei¹, ZHENG Lulin⁵, ZHENG Lujing⁵, SONG Weifang⁵

1. State Key Laboratory of Ore Deposit Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550081, China

2. College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

3. Center for Earth Environment & Resources, Sun Yat-sen University, Guangzhou 510275, China

4. Bureau of Geology and Mineral Exploration and Development, Guiyang 550004, China

5. Guizhou University, Guiyang 550025, China

CAO Shengtao, HU Ruizhong, ZHOU Yongzhang, et al. Element enrichment pattern and prospecting method for Carlin-type gold deposits based on big data association rule algorithm. *Earth Science Frontiers*, 2024, 31(4): 058-072

Abstract: Carlin-type gold deposits are an important reservoir of gold. Due to the gradual depletion of shallow surface gold resources, there is an urgent need for new prospecting methods to explore deep and hidden areas. The advent of the big data era has opened up new prospecting ideas. Association rule algorithm one of the most commonly used mining algorithms and can be used to effectively mine the inherent correlation between data items in large data sets. In this study, association rule mining is used to analyze the correlation between trace elements and gold mineralization in major Carlin-type gold deposits in the Yunnan-Guizhou-Guangxi “Golden Triangle” region. Combined with element migration and enrichment patterns, elemental anomaly combinations are extracted, and quantitative prospecting indicators are established. The elemental anomaly combinations are divided into elements with strong positive correlation and significantly enriched (As, Sb, Hg, Tl, Ag, W, Rb), indicating sulfidation and clayification (Rb); elements with strong positive correlation and slightly enriched (Zr, Th, Ta, Nb, Hf) or with strong negative correlation and strongly depleted (Li, Sr), indicating decarbonation; elements with strong positive correlation and slightly enriched (Sn, Zn, Ni, V, Co, Cu), likely reflecting their low contents in ore-forming fluids; and elements with weak correlation and not enriched (Cd, Pb, Ba, Bi, U, Mo)—these elements show no significant correlation with gold mineralization. The elemental anomaly combinations obtained by big data approach is consistent with previous understanding of the genesis of Au deposits, i. e., Au is mainly formed under

收稿日期: 2024-02-26; 修回日期: 2024-04-29

基金项目: 国家自然科学基金项目(41830432, U1812402); 国家重点基础研究发展计划“973”项目“华南大规模低温成矿作用(2014CB440900)”; 国家重点研发计划项目(2022YFF0801201); 贵州省卡林型金矿成矿与找矿科技创新人才团队建设项目(黔科合平台人才-CXTD[2021]007)

作者简介: 曹胜桃(1996—), 男, 博士研究生, 地球化学专业。E-mail: 719733205@qq.com

* 通信作者简介: 胡瑞忠(1958—), 男, 研究员, 中国科学院院士, 主要从事矿床学和矿床地球化学的研究工作。E-mail: huruizhong@vip.gyig.ac.cn
周永章(1963—), 男, 教授, 主要从事地球化学、大数据与数学地球科学等方面的研究工作。E-mail: zhouyz@mail.sysu.edu.cn

decarbonation and sulfidation processes accompanied by significant clayification, in which sulfidation is the main genetic mechanism. Through association rule mining, quantitative prospecting indicators are established: For sulfidation related elements (As, Hg, Sb, Tl, W, Ag, Rb), when the number of medium-high content elements in samples $\geq 1, 2, 3, 4$, or 5 , the corresponding Au contents $\geq 4.5 \times 10^{-9}, 47.0 \times 10^{-9}, 150 \times 10^{-9}, 500 \times 10^{-9}$, or 1000×10^{-9} ; when the number of high-content elements $\geq 1, 2$, or 3 , the corresponding Au contents $\geq 150 \times 10^{-9}, 500 \times 10^{-9}$, or 1000×10^{-9} ; during prospecting, both indicators should be used to ensure efficient delineation of ore bodies, without outcrops. For decarbonation related elements (Zr, Th, Ta, Nb, Hf), decarbonation is indicated when elemental content anomaly occurs at any two of the elements in samples. The method developed in this study for establishing quantitative prospecting indicators based on association rule algorithms should provide new ideas for other types of mineral deposits.

Keywords: geological big data; association rule algorithm; Carlin-type gold deposit; element enrichment law; factors of control; prospecting indicators

摘要:大数据时代的到来,为卡林型金矿床开拓了新的找矿思路。本研究应用关联规则算法,挖掘滇黔桂“金三角”卡林型金矿床内微量元素与金矿化海量数据之间的关联性,提取元素异常组合,分析控制因素,定量构建找矿标志。结果显示矿床内元素异常组合分为4组:(1)强正关联显著富集元素(As、Sb、Hg、Tl、Ag、W和Rb),显示硫化和黏土化作用;(2)较强正关联略富集元素I(Zr、Th、Ta、Nb和Hf)和强负关联强迁出元素(Li和Sr),显示去碳酸盐化作用;(3)较强正关联略富集元素II(Sn、Zn、Ni、V、Co和Cu),显示硫化作用;(4)弱关联基本无富集元素(Cd、Pb、Ba、Bi、U和Mo),与成矿无显著关联。从大数据角度获取的元素异常组合,与学界关于Au主要在去碳酸盐化、硫化和黏土化条件下形成的认识一致。通过关联规则算法分别对与硫化和去碳酸盐化相关的元素建立定量找矿标志。硫化找矿标志:样品中As、Hg、Sb、Tl、W、Ag和Rb等元素内中高含量项数 $\geq 1, \geq 2, \geq 3, \geq 4$ 和 ≥ 5 时,对应的Au矿化分别为 $\geq 4.5 \times 10^{-9}, \geq 47.0 \times 10^{-9}, \geq 150 \times 10^{-9}, \geq 500 \times 10^{-9}$ 和 $\geq 1000 \times 10^{-9}$;样品内中高含量项数 $\geq 1, \geq 2$ 和 ≥ 3 时,对应的Au矿化分别为 $\geq 150 \times 10^{-9}, \geq 500 \times 10^{-9}$ 和 $\geq 1000 \times 10^{-9}$;找矿过程中两组指标配合使用,确保不漏矿,高效圈矿。去碳酸盐化找矿标志:样品中Zr、Th、Ta、Nb和Hf含量任意两项出现正异常,认为样品经历过去碳酸盐化作用。定量识别的硫化和去碳酸盐化找矿标志可望在卡林型金矿找矿预测中发挥重要作用。本研究基于关联规则算法分析矿床元素富集规律、控制因素和定量构建找矿标志的方法,也可为其他类型矿床开展类似研究提供新思路。

关键词:地质大数据;关联规则;卡林型金矿;元素富集规律;控制因素;找矿标志

中图分类号:P628.4;P618.51 **文献标志码:**A **文章编号:**1005-2321(2024)04-0058-15

0 引言

美国内华达州和中国右江盆地是全球公认的卡林型金矿富集区^[1-4]。内华达州已探明卡林型金矿的金储量近8000t,是世界上的第二大金矿集区,年产量占全球总量的6%,且勘探深度超过1000m^[3]。右江盆地已探明卡林型金矿的金储量近1000t,是仅次于内华达地区的第二大卡林型金矿床集中分布区,勘探深度多不足500m,显示出巨大的找矿潜力^[4-8]。

前人对右江盆地卡林型金矿床开展了许多研究,并获得许多有意义的认识及海量的数据^[1-10]。然而,右江盆地浅部卡林型金矿资源大量开采,导致矿山资源出现枯竭危机,迫使找矿方向转向深部和盲区。在卡林型金矿找矿勘查过程中,除用Au含量进行找矿

和圈矿外,一直缺少其他元素作为定量指标。

大数据和人工智能时代的到来为新一轮找矿提供了新的突破思路,大数据正在改变地质学家的思维方式。其中,关联规则算法是大数据挖掘的基础算法,也是最活跃的研究方法之一,在许多领域找到应用场景,甚至进行成矿预测^[11-14]。

本研究以右江盆地内主要的卡林型金矿床为研究对象,在前人获得的海量矿床微量元素数据基础上,通过关联规则算法,定量挖掘元素之间、元素与Au矿化之间的关联关系,并结合矿床内元素的迁移富集规律分析,提取元素异常组合,分析控制因素,建立定量找矿指标,服务于区域上深部和盲区的找矿。

1 研究区地质背景

右江盆地又被称为滇黔桂“金三角”,位于扬

子陆块西南缘(图 1), 其北东以紫云-垭都断裂、北西以弥勒-师宗断裂、南东以凭祥-南宁断裂和南西以红河断裂为界, 总面积约 $18.5 \times 10^4 \text{ km}^2$ [15]。区内矿地层和岩石丰富, 表现为几乎所有地层和岩石均可容矿, 主要容矿地层为二叠系和三叠系地层, 容矿岩石包括灰岩、白云岩、粉砂岩、黏土岩、硅质岩、辉绿岩、玄武岩和凝灰岩等, 产出代表性的矿床有水银洞、烂泥沟、泥堡、板其、丫他、紫

母凶、戈塘和架底等大型-超大型金矿床 [7-8]。前人研究显示 [1-8], 金矿体主要产于钙质碎屑岩系建造的断裂带和刚性碳酸盐岩系与钙泥质碎屑岩系之间的构造蚀变体 (SBT) 中, 围岩富含 Fe 碳酸盐矿物, 成矿主要与脱碳酸盐化、硅化、泥化和硫化等低温热液蚀变相关, 金呈微细粒不可见形式赋存于黄铁矿的含砷环带内, Au 通常与 As-Sb-Hg-Tl 等元素共/伴生。

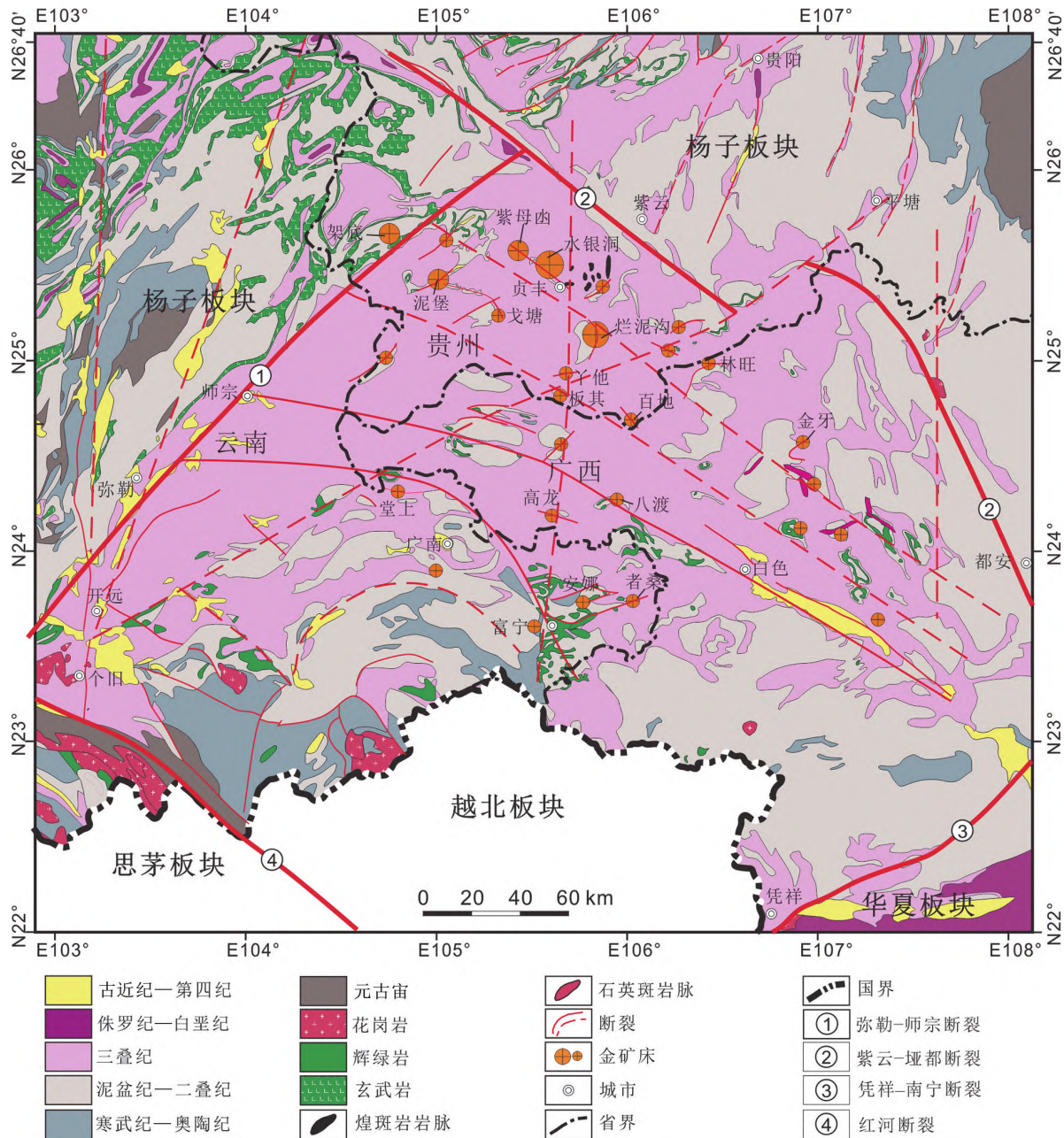


图 1 滇黔桂“金三角”区域地质图
(据文献[6]修改)

Fig. 1 Geologic map of the Yunnan-Guizhou-Guangxi “Golden Triangle” region, southwestern China. Modified after [6].

2 研究方法

关联规则算法是重要的大数据挖掘算法^[11]。它可以从大数据的项集之间发现有意义的、频繁出现的模式、关联和相关性。

本研究采用的 Apriori 算法主要包括两个步骤^[16]：一是寻找频繁项集；二是生成强关联规则(图 2)。基本工具包括支持度 (support)、置信度 (confidence) 和提升度 (lift)。其中,支持度表示 A 与 B 同时发生或同时出现的概率,即 $\text{support}(A \rightarrow B) = P(A \cap B)$ 。置信度是指 A 发生的条件下 B 发生的概率,即 $\text{confidence}(A \rightarrow B) = P(B|A)$ 。提升度是反映规则是否有增益,表示 A 发生的条件下 B 发生的概率与 B 总体发生的概率之比,即 $\text{lift}(A \rightarrow B) = P(B|A)/P(B)$ 。

当 $A \rightarrow B$ 支持度大于设定的最小支持度时,则 $A \rightarrow B$ 为频繁项集。在满足最小支持度的基础上,若 $A \rightarrow B$ 置信度大于最小置信度,则 $A \rightarrow B$ 为强关联规则。获取强关联规则后,利用提升度判断强关联规则是否有效。若 $A \rightarrow B$ 提升度 < 1 ,规则负相关,无增益。若 $A \rightarrow B$ 提升度 > 1 ,规则正相关,有增益^[11]。

3 数据预处理

3.1 数据清洗

本研究共收集 1 756 件全岩微量元素数据,数

据来源于知网文献数据(主要来源于水银洞、烂泥沟、泥堡、板其、丫他、紫母山、戈塘和架底等大型—超大型金矿床)和矿山勘查数据(主要来源于水银洞和烂泥沟金矿床)。分析项包含 Au、Ag、As、Ba、Be、Bi、Cd、Co、Cr、Cs、Cu、Ga、Ge、Hf、Hg、In、Li、Mn、Mo、Nb、Ni、Pb、Rb、Sb、Sc、Sn、Sr、Ta、Th、Tl、U、V、W、Y、Zn 和 Zr 等 36 个元素。

数据中元素种类缺失较多的样品直接剔除,元素含量显示高于或低于检测限,用对应检测限代替。元素对应的样品数介于 107 ~ 1 627,集中分布于 635 ~ 1 030。Be、Cr、Cs、Ga、Ge、In、Mn、Sc 和 Y 等元素对应的样品数目为 107 ~ 522,且除 Cr 的数量为 522 外,其余元素数目均低于 381,这部分元素暂不对其进行关联分析,将其从数据库中剔除。以上步骤,使缺失值锐减,Au 元素无缺失值,其他元素的缺失值仅占小部分,用空值代替。主要对数据量大的 Au、Ag、As、Ba、Bi、Cd、Co、Cu、Hf、Hg、Li、Mo、Nb、Ni、Pb、Rb、Sb、Sn、Sr、Ta、Th、Tl、U、V、W、Zn 和 Zr 等 27 个元素进行关联分析研究。在以 Au 为目标进行关联规则算法过程中,去掉缺失值,最大限度利用数据。数据清洗后剩余样品 1 627 件(表 1)。

3.2 数据变换

采取 k 均值(k-means)聚类算法实现数据结构化转换。为使聚类数据量分布均衡,将高异常值抽出,剩余元素分别采用 k-means 聚类成 3 类,对应元素含量低、中和高。Au 以边界品位 $1\ 000 \times 10^{-9}$ 为

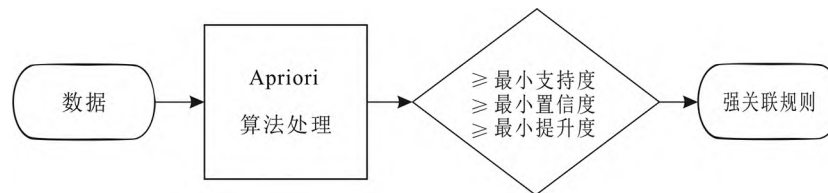


图 2 关联规则算法流程示意图

Fig. 2 Schematic flow of association rule algorithm

表 1 清洗后的数据(部分)

Table 1 Data after cleaning (partial data)

编号	$w_B/10^{-9}$				$w_B/10^{-6}$				
	Au	Ag	As	Ba	Co	Cu	Hf	...	Zr
1	17.00	0.09	296.50	26.60	0.88	3.34	0.99	...	38.75
2	501.00	1.81	709.00	253.00	3.57	14.10	15.60	...	605.00
3	186.00	1.39	1 151.00	179.00	9.47	21.50	13.90	...	482.00
4	52.00	0.93	792.00	112.00	19.60	43.20	8.82	...	328.00
5	105.00	1.46	964.00	193.00	9.04	24.80	13.70	...	494.00
6	57.00	0.86	737.00	63.80	7.88	8.68	6.96	...	270.00
...
1 627	0.20	0.03	7.72	28.67	0.67	13.84	0.30	...	10.14

高异常界限值^[17]。Au 通过 k-means 聚类, 含量被分成了 $0 \sim 150 \times 10^{-9}$ (低含量)、 $(150 \sim 500) \times 10^{-9}$ (中含量) 和 $(500 \sim 1\ 000) \times 10^{-9}$ (高含量) (边界值略有调整)。通常根据 3 倍上地壳中 Au 的平均含量 (约 1.5×10^{-9})^[18] 判断成矿是否发生异常。本研究中 Au 低含量类显著高, 表明发生了矿化。因此, Au 的含量分为 $[0, 4.5) \times 10^{-9}$ 、 $[4.5, 150) \times 10^{-9}$ 、 $[150, 514) \times 10^{-9}$ 、 $[514, 1\ 000) \times 10^{-9}$ 和 $[1\ 000, 74\ 900] \times 10^{-9}$, 对应样品数为 841、142、192、100 和 352, 分别代表母岩发生的矿化程度为未发生矿化 (围岩)、低矿化、中矿化、高矿化和矿体 (表 2)。

其他元素采用箱型图将高异常值抽出, 剩余数据采用 k-means 聚类成 3 类, 对应元素含量低、中和高, 并将高异常值与高含量部分合并为高含量类。

结果显示, 元素单类样品数量均大于 94 件, 且各类占比平衡, 有利于关联规则算法分析。

4 元素与金矿化关联规则分析

4.1 以 Au 为目标的关联规则

从样品中随机抽 1 件, Au^1 、 Au^2 、 Au^3 、 Au^4 和 Au^5 的概率分别是 27.1%、33.3%、11.8%、6.15% 和 21.6%。以 Au 为目标, 其余 26 个元素作为输入项进行关联规则分析。预设最小支持度为 0.5%, 最小置信度为 70%, 最大前项为 6 项, 最小提升度 (增益) 设为 1。结果获 553 898 条强关联规则。5 类强关联规则中支持度与置信度之和最大的部分规则见表 3, 以规则 1 和规则 10 为例分析。

规则 1: $Sn^2-Rb^1-Tl^1-As^1 \rightarrow Au^1$, 支持度为 1.11%, 置信度为 100%。表示样品中元素含量出现 Sn^2 、 Rb^1 、 Tl^1 和 As^1 时, 样品矿化程度为 Au^1 的概率为 100%, 为围岩。反之, 样品中元素含量均不出现 Sn^2 、 Rb^1 、 Tl^1 和 As^1 时, 支持度为 98.89%, 样品矿化程度为 Au^{2-5} 的概率为 100%, 为矿化部位。

规则 10: $Tl^3-Hg^3-Ag^3-Sb^3-Ni^2 \rightarrow Au^5$, 支持度为 2.95%, 置信度为 100%。表示样品中元素含量出现 Tl^3 、 Hg^3 、 Ag^3 、 Sb^3 和 Ni^2 时, 样品矿化程度为 Au^5 的概率为 100%, 为矿石。

4.2 矿化与元素含量

Au^1 为围岩 (或弱矿化围岩), Au^{2-5} ($Au^2 \sim Au^5$) 为矿化部位。以 Au^1 和 Au^{2-5} 为目标, 其余 26 个元素为输入项进行关联规则分析 (高含量类和中含量类合并为中高含量类)。最小支持度为 2%, 最小置信度为 27.1%, 最大前项为 10 项 (理论可以是 26 项, 但超过 10 项的规则少), 最小提升度设为 1, 获取了 1 021 595 条强关联规则, 难直观获取元素异常组合。最小置信度为 27.1% 和最小提升度设为 1, 确保已生成的规则有增益, 同时不会遗漏以 Au^1 为后项有增益的强关联规则。在最小支持度、最小置信度和最小提升度限制下, 与目标项关联性越强的项在生成的强关联规则中出现的次数越高。关联规则算法相当于放大镜, 随着生成强规则前项数增加, 与目标关联程度强的项出现的次数会以几何倍数递增。反之, 与后项关联程度弱的项出现的次数不变或略微增加。

以 Au^1 为后项的强关联规则内中高含量类 (X_2) 出现的次数相较于低含量类 (X_1) 显示 (图 3a) 情况: As、Sb、Hg、Tl、Ag 和 W 显著低; Rb、Zr、Th、Ta、Nb、Hf、Sn、V、Zn、Ni、Co、Cd、Pb、Ba、U 和 Li 接近; Cu 和 Mo 略高; Bi 和 Sr 显著高。反映围岩与 As、Sb、Hg、Tl、Ag 和 W 低含量类 (X_1) 关联强, 与 Bi 和 Sr 中高含量类 (X_2) 关联强, 与 Rb、Zr、Th、Ta、Nb、Hf、Sn、V、Zn、Ni、Co、Pb、Ba、U、Li、Cu 和 Mo 低含量类 (X_1) 和中高含量类 (X_2) 关联程度无显著差异。

以 Au^{2-5} 为后项的强关联规则显示 (图 3b) 矿化部位与 As、Sb、Hg、Tl、Ag、W、Rb、Zr、Th、Ta、Nb、Hf、Sn、V、Zn、Ni、Co、Cd 和 Bi 等中高含量类 (Y_2) 关联强, 与 Li 和 Sr 低含量类 (Y_1) 关联强, 与 Ba、Pb 和 U 低含量类 (Y_1) 和中高含量类 (Y_2) 关联程度无显著差异。

4.3 矿化关联度

以上判别了围岩和矿化部位与 26 个元素含量类之间的相对关联强弱, 仍需量化 Au 与其他元素整体关联程度。通过获取矿化部位元素中高含量类出现次数 (Y_2) 与低含量类出现次数 (Y_1) 的比值 (Y_2/Y_1)、与围岩中高含量类出现次数 (X_2) 与低含量类出现次数 (X_1) 的比值 (X_2/X_1) 相比判断, 令比值为关联系数 a , 即

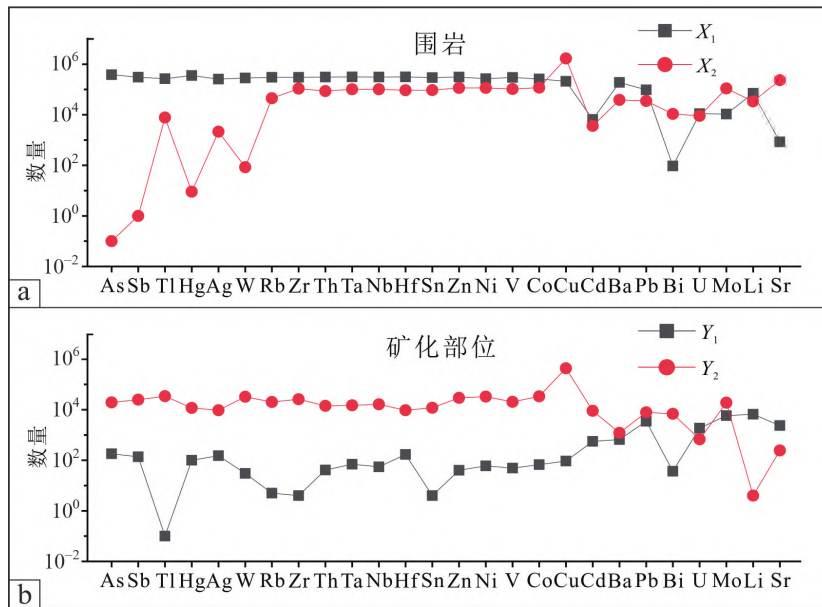
表 2 元素数据聚类特征
Table 2 Elemental data clustering features

元素	Au					Ba					Co					Hg									
	Au ¹	Au ²	Au ³	Au ⁴	Au ⁵	Ba ¹	Ba ²	Ba ³	Co ¹	Co ²	Co ³	Hg ¹	Hg ²	Hg ³	Hg ⁴	Hg ⁵									
分类	Au ¹	Au ²	Au ³	Au ⁴	Au ⁵	Ba ¹	Ba ²	Ba ³	Co ¹	Co ²	Co ³	Hg ¹	Hg ²	Hg ³	Hg ⁴	Hg ⁵									
含量范围	[0, 4.5)	[4.5, 150)	[150, 500)	[500, 1 000)	[1 000, 74 900)	[0, 169)	[169, 350)	[350, 8 822]	[0, 9.11)	[9.11, 21.6)	[21.6, 169]	[0, 2.12)	[2.12, 5.84)	[5.84, 5 940]	[2.12, 5.84)	[5.84, 5 940]									
样品数/个	441	542	192	100	352	417	219	144	456	810	358	940	265	304	265	304									
元素	Ag					Bi					Cu					Li									
分类	Ag ¹	Ag ²	Ag ³	Ag ⁴	Ag ⁵	Bi ¹	Bi ²	Bi ³	Cu ¹	Cu ²	Cu ³	Li ¹	Li ²	Li ³	Li ⁴	Li ⁵									
含量范围	[0, 0.075)	[0.075, 0.164)	[0.164, 15.0)	[0, 0.189)	[0.189, 0.360)	[0.360, 14.2)	[0.360, 14.2)	[0.360, 14.2)	[0, 22.1)	[22.1, 46.9)	[46.9, 1 100]	[0, 14.8)	[14.8, 35.3)	[35.3, 1 154]	[14.8, 35.3)	[35.3, 1 154]									
样品数/个	810	360	348	267	470	254	254	254	534	696	365	382	138	135	382	138									
元素	As					Cd					Hf					Mo									
分类	As ¹	As ²	As ³	As ⁴	As ⁵	Cd ¹	Cd ²	Cd ³	Hf ¹	Hf ²	Hf ³	Mo ¹	Mo ²	Mo ³	Mo ⁴	Mo ⁵									
含量范围	[0, 668)	[668, 1 939)	[1 939, 200 900)	[0, 0.168)	[0.168, 0.28)	[0.280, 18.0)	[0.280, 18.0)	[0.280, 18.0)	[0, 4.66)	[4.66, 10.2)	[10.2, 37.3)	[0, 0.699)	[0.699, 1.50)	[1.50, 122]	[0, 0.699)	[0.699, 1.50)									
样品数/个	1 008	292	326	463	369	283	283	283	365	189	94	854	390	328	854	390									
元素	Nb					Rb					Sr					Tl					W				
分类	Nb ¹	Nb ²	Nb ³	Nb ⁴	Nb ⁵	Rb ¹	Rb ²	Rb ³	Sr ¹	Sr ²	Sr ³	Tl ¹	Tl ²	Tl ³	Tl ⁴	Tl ⁵									
含量范围	[0, 20.1)	[20.1, 47.3)	[47.3, 181]	[0, 42.7)	[42.7, 98.8)	[98.8, 285]	[98.8, 285]	[98.8, 285]	[0, 337)	[337, 807)	[807, 7 984]	[0, 0.636)	[0.636, 1.36)	[1.36, 164]	[0, 2.23)	[2.23, 4.63)									
样品数/个	308	176	145	315	285	109	285	109	361	244	149	543	624	317	884	334									
元素	Ni					Sb					Ta					U					Zn				
分类	Ni ¹	Ni ²	Ni ³	Ni ⁴	Ni ⁵	Sb ¹	Sb ²	Sb ³	Ta ¹	Ta ²	Ta ³	U ¹	U ²	U ³	Zn ¹	Zn ²									
含量范围	[0, 18.4)	[18.4, 37.9)	[37.9, 4 122]	[0, 8.51)	[8.51, 24.1)	[24.1, 328 200]	[24.1, 328 200]	[24.1, 328 200]	[0, 1.4)	[1.4, 3.28)	[3.28, 11.5)	[0, 1.43)	[1.43, 2.68)	[2.68, 31.3)	[0, 49)	[49, 102]									
样品数/个	435	840	344	911	323	355	323	355	325	166	132	347	204	171	427	827									
元素	Pb					Sn					Th					V					Zr				
分类	Pb ¹	Pb ²	Pb ³	Pb ⁴	Pb ⁵	Sn ¹	Sn ²	Sn ³	Th ¹	Th ²	Th ³	V ¹	V ²	V ³	Zr ¹	Zr ²									
含量范围	[0, 9.72)	[9.72, 19.2)	[19.2, 535]	[0, 1.92)	[1.92, 3.38)	[3.38, 12.0)	[3.38, 12.0)	[3.38, 12.0)	[0, 4.47)	[4.47, 9.69)	[9.69, 35.5)	[0, 96.8)	[96.8, 216)	[216, 558]	[0, 127)	[127, 341]									
样品数/个	531	630	434	314	198	140	198	140	339	282	97	347	214	212	259	261									

注: Au 元素含量单位为 10^{-9} , 其余元素含量单位为 10^{-6} 。

表3 部分强关联规则
Table 3 Examples of strong association rules

编号	规则	实例数量/个	支持度/%	置信度/%	增益
1	$\text{Sn}^2\text{-Rb}^1\text{-Tl}^1\text{-As}^1 \rightarrow \text{Au}^1$	18	1.11	100	3.69
2	$\text{Zr}^2\text{-Tl}^1\text{-Pb}^2\text{-W}^1\text{-As}^1 \rightarrow \text{Au}^1$	21	1.29	95.2	3.51
3	$\text{Cd}^2\text{-Mo}^2\text{-Tl}^2\text{-Cu}^2\text{-W}^1\text{-Sb}^1 \rightarrow \text{Au}^2$	19	1.17	100.0	3.00
4	$\text{Hg}^2\text{-Tl}^2\text{-Cu}^2\text{-W}^1\text{-Sb}^1 \rightarrow \text{Au}^2$	23	1.41	95.7	2.87
5	$\text{Rb}^2\text{-Sn}^1\text{-Ba}^1\text{-Hg}^1 \rightarrow \text{Au}^3$	17	1.04	70.6	5.98
6	$\text{Sn}^1\text{-Cd}^1\text{-Pb}^1\text{-Ag}^1\text{-As}^1 \rightarrow \text{Au}^3$	17	1.04	70.6	5.98
7	$\text{Rb}^3\text{-Th}^2\text{-Ni}^3\text{-Sr}^1\text{-Ba}^1 \rightarrow \text{Au}^4$	9	0.55	77.8	12.7
8	$\text{Rb}^3\text{-As}^3\text{-Th}^2\text{-Ni}^3\text{-Sr}^1\text{-Ba}^1 \rightarrow \text{Au}^4$	9	0.55	77.8	12.7
9	$\text{Hg}^3\text{-As}^3\text{-Tl}^3\text{-Ag}^3\text{-Ni}^2 \rightarrow \text{Au}^5$	48	2.95	100	4.62
10	$\text{Tl}^3\text{-Hg}^3\text{-Ag}^3\text{-Sb}^3\text{-Ni}^2 \rightarrow \text{Au}^5$	48	2.95	100	4.62



X_1 和 X_2 分别代表以 Au^1 (围岩)为后项的强关联规则中各元素低含量类和中高含量类出现的次数; Y_1 和 Y_2 分别代表以 Au^{2-5} (矿化部位)为后项的强关联规则中各元素低含量类和中高含量类出现的次数。

图3 强关联规则中元素在围岩和矿化部位出现的频数

Fig. 3 Number of occurrences of low- (squares) and high-content (dots) elements in wall rock (a) and mineralization site (b) under strong association rules

$$\frac{\frac{Y_2}{Y_1}}{\frac{X_2}{X_1}} = a \dots\dots\dots(\text{关联系数})$$

$$\frac{\frac{Y_2}{Y_1}}{\frac{X_2}{X_1}} > 1 \dots\dots\dots(\text{正关联})$$

$$\frac{\frac{Y_2}{Y_1}}{\frac{X_2}{X_1}} = 1 \dots\dots\dots(\text{无关联})$$

$$\frac{\frac{Y_2}{Y_1}}{\frac{X_2}{X_1}} < 1 \dots\dots\dots(\text{负关联})$$

其中 X_1 、 X_2 、 Y_1 、 Y_2 均大于或等于 0, 当 X_1 、 X_2 、 Y_1 、 Y_2 为 0 时(关联规则参数设置中尽可能使 X_1 、 X_2 、 Y_1 、 Y_2 不为 0), 令其为 0.1, 细小的变化不影响整体关联强度判断、突显关联强弱, 故 $a > 0$ 。若元素的 X_1 与 X_2 、或 Y_1 与 Y_2 值接近或均较小, 则表示围岩或矿化部位内元素低含量类和中高含量类的比例接近, 元素与 Au 可能有一定关联, 但并非特别强。在此基础上, 通过 a 的大小判断各元素与 Au 的关联程度。

当元素的关系系数 a 大于 1 时, 元素与 Au 含量呈正关联。 a 越大, 元素与成矿关联程度越强, 元素越富集。

当元素的关系系数 a 等于或接近 1 时, 元素与 Au 含量无关联, 元素矿化作用过程中含量变化

不大。

当关联系数 a 小于 1 时,元素与 Au 含量负关联。 a 越小,负关联程度越强,元素矿化作用过程中越可能流失。

根据关联系数 a 的值(表 4、图 4),将其他元素与成矿(Au)关联度划分为 4 组:强正关联组,元素包括 As、Sb、Tl、Hg、W、Rb、Zr、Sn 和 Ag, a 为 7 563~414 269 855;较强正关联组,元素包括 Zn、Ni、Th、V、Co、Nb、Ta、Cu 和 Hf, a 为 191~1 964;弱关联组,元素包括 Cd、Ba、Pb、Bi、U 和 Mo, a 为 0.31~29.3;强负关联组,元素包括 Li 和 Sr, a 为 0.001 3~0.000 38。

4.4 元素异常组合

通过矿化部位元素平均含量(Z_2)与围岩元素平均含量(Z_1)的比值(Z_2/Z_1),判断元素整体迁移富集规律(表 4、图 3)。

Au 在围岩的平均含量为 2.26×10^{-9} ,矿化部位平均含量为 $1 836.63 \times 10^{-9}$,矿化部位平均含量是围岩的 813 倍。其他元素富集程度分为 4 类(表 5):显著富集元素 As、Sb、Ag、W、Hg、Tl、Rb 和 Mo,矿化后平均含量是矿化前的 1.51~178 倍;略富集元素 V、Zn、Zr、Ni、Co、Cu、Bi、Nb、Hf、Ta、Ba、

Th、Sn、U 和 Pb,矿化后平均含量是矿化前的 1.06~1.27 倍;略迁出元素 Cd,矿化后平均含量是矿化前的 0.97 倍;强迁出元素 Li 和 Sr,矿化后平均含量是矿化前的 0.45~0.50 倍。

结合矿床内元素与成矿的关联性,元素异常组合划分为:强正关联显著富集元素(As、Sb、Hg、Tl、Ag、W 和 Rb);较强正关联略富集元素(Zr、Th、Ta、Nb、Hf、Sn、Zn、Ni、V、Co 和 Cu);强负关联强迁出元素(Li 和 Sr)和弱关联基本无富集元素(Cd、Pb、Ba、Bi、U 和 Mo)。

5 讨论

5.1 控制因素

以往研究显示,卡林型金矿床的成矿过程为成矿流体交代蚀变含 Fe 碳酸盐,释放 Fe^{2+} 进入流体使 Au-HS 络合物失稳,导致含砷黄铁矿($Fe(As, S)_2$)沉淀并卸载 Au,从而在有利赋矿空间的岩石中形成金矿体,得出金主要以“不可见金”形式赋存于含砷黄铁矿和毒砂中,而去碳酸盐化和硫化作用是主要成矿机制,伴随大量的黏土化^[1,4,6]。

结合元素与 Au 矿化的关联程度和元素的迁移

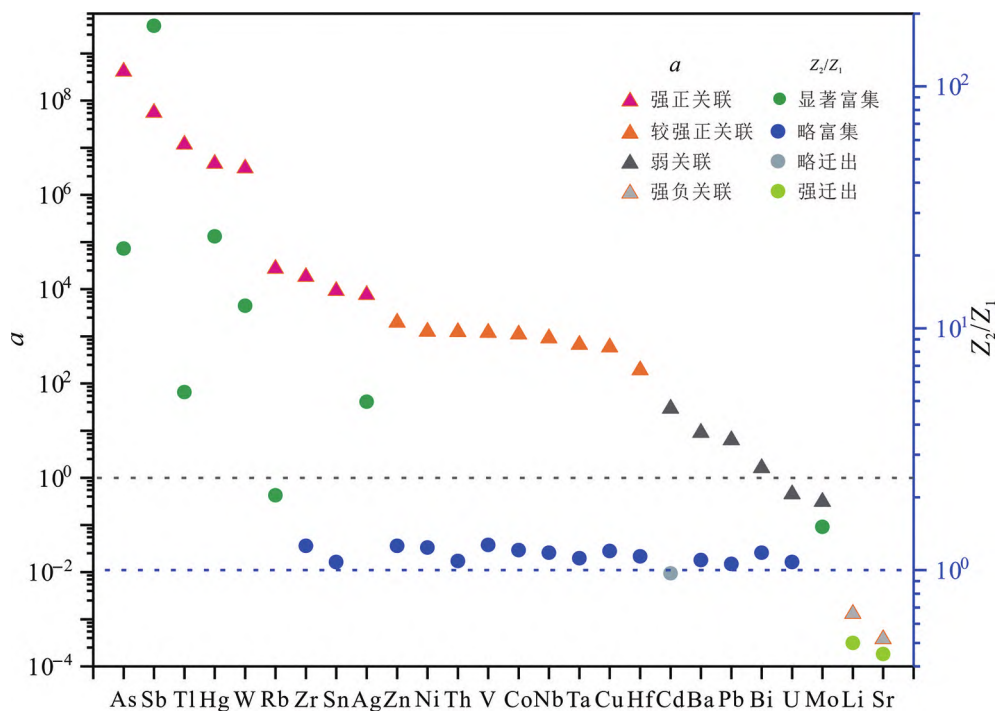


图 4 元素与 Au 的关联度及元素迁移富集特征

Fig. 4 Correlation coefficients between selected elements and Au mineralization (triangles) and elemental migration enrichment characteristics (dots)

表 4 元素关联系数与富集特征
Table 4 Elemental correlation coefficients and enrichment characteristics

指标参数	强正关联显著富集元素 (受硫化控制)				强正关联略富集元素 I (受碳酸盐控制)				强负关联强迁出元素 (受去碳酸盐控制)					
	As	Sb	Tl	Hg	W	Ag	Rb (受黏土化控制)	Zr	Th	Nb	Ta	Hf	Li	Sr
X ₁	385.298	307.759	267.595	356.365	284.726	256.715	302.612	303.658	310.478	311.954	313.724	315.367	70.882	848
X ₂	0	1	7.803	9	84	2.134	44.899	107.606	86.456	101.413	101.374	93.157	33.608	233.632
Y ₁	181	138	0	101	30	151	5	4	41	55	70	169	6.592	2.374
Y ₂	19.461	24.893	34.319	11.705	32.406	9.493	20.155	25.896	14.169	16.328	15.005	9.523	4	246
Z ₁	120.6	5.21	0.52	1.36	1.53	0.06	31.63	185.34	5.1	24.51	1.7	4.82	49.88	842.17
Z ₂	2.575.23	929.4	2.84	32.65	19.05	0.28	64.52	232.91	5.55	28.88	1.92	5.49	24.87	381.89
a	414.269.855	55.514.817	11.769.310	4.588.836	3.661.441	7.563	27.168	18.269	1.241	9.13	6.63	1.91	0.001.3	0.000.38
Z ₂ /Z ₁	21.35	178	5.44	24	12.4	4.97	2.04	1.26	1.09	1.18	1.12	1.14	0.5	0.45

指标参数	弱关联基本无富集元素 (受与成矿无显著关联控制)											
	Sn	Zn	Ni	V	Co	Cu	Cd	Ba	Pb	Bi	U	Mo
X ₁	293.550	309.505	266.853	298.543	260.850	210.042	6.546	190.041	96.862	94	11.044	10.659
X ₂	94.237	115.688	115.161	104.260	118.756	1.681.376	3.601	38.696	34.840	10.770	9.011	109.718
Y ₁	4	40	60	49	67	93	560	661	3.487	37	1.856	5.852
Y ₂	11.893	29.368	32.724	20.416	33.537	437.998	9.037	1.215	7.910	6.835	6.77	18.956
Z ₁	2.22	64.2	26.94	116.18	14.43	35.34	0.45	267.91	15.45	0.26	2.13	1.24
Z ₂	2.39	80.68	33.38	147.1	17.51	42.3	0.44	293.85	16.46	0.31	2.3	1.87
a	9.262	1.964	1.264	1.193	1.099	588	29.3	9.03	6.31	1.61	0.45	0.31
Z ₂ /Z ₁	1.08	1.26	1.24	1.27	1.21	1.2	0.97	1.1	1.06	1.18	1.08	1.51

注:Z₁和Z₂分别为成矿前后元素在岩石内的平均含量,单位为10⁻⁶。Z₂/Z₁指矿化作用后的富集程度,当Z₂/Z₁>1,为富集元素,Z₂/Z₁=1为稳定元素,Z₂/Z₁<1,为迁出元素。

富集规律可以发现,强正关联显著富集元素(As、Sb、Hg、Tl、Ag 和 W)在矿床中以硫化物或赋存在硫化物的形式存在,尤其是含砷黄铁矿^[1,4,6,19-20],表明硫化是主要成矿机制。 Li^+ 与 Mg^{2+} 和 Fe^{2+} 的半径相近, Sr^{2+} 与 Ca^{2+} 和 Mg^{2+} 具有相似的性质,均常以类质同象的形式存在^[21-22],且主要赋存于卡林型金矿床碳酸盐岩内(尤其是Fe白云石),强负关联强迁出元素(Li和Sr)指示在成矿过程中去碳酸盐化作用强烈。Rb离子半径(0.113 nm)与K(0.099 nm)相近,且Rb主要分散在含K的黏土矿物中(如白云母、黑云母和伊利石),同时黏土矿物对Rb有很强的吸附作用^[23],强正关联显著富集元素(Rb)表明矿化过程中黏土化作用明显。高场强元素Zr、Th、Ta、Nb和Hf受自身离子结构约束(离子电价较高,半径较小,具有较高离子场强),一般较稳定,不易受变质、蚀变和风化作用等的影响^[24]。卡林型金矿中Zr、Th、Ta、Nb和Hf整体略微富集,与Au矿化有较强正关联,主要是去碳酸盐化作用导致岩石体积减小所致。在矿化前后岩石内高场强元素总量几乎不变,去碳酸盐化作用导致岩石体积减小,因此含量略显增加,且与成矿呈较强正关联。Sn、Zn、Ni、V、Co和Cu与成矿关联较强,略富集,这些元素在中低温矿床中主要以硫化物或赋存在硫化物的形式存在^[25],本研究中的富集规律可能是这些元素在成矿流体中含量低所致。其余元素(Cd、Pb、Ba、Bi、U和Mo)整体关联程度和富集量均较低,表明成矿过程对其影响小。

综合以上,矿床内元素异常组合可分为4组:(1)强正关联显著富集元素(As、Sb、Hg、Tl、Ag、W和Rb),显示硫化和黏土化(Rb)作用;(2)较强正关联略富集元素I(Zr、Th、Ta、Nb和Hf)和强负关联强迁出元素(Li和Sr),显示去碳酸盐化作用;(3)较强正关联略富集元素II(Sn、Zn、Ni、V、Co和Cu),显示硫化作用;(4)弱关联基本无富集元素(Cd、Pb、Ba、Bi、U和Mo),与成矿无显著关联。从大数据角度获取的元素异常组合,与学界^[1,4-6]关于Au主要在去碳酸盐化、硫化、黏土化条件下形成的认识一致。

5.2 找矿标志

强正关联显著富集元素(As、Sb、Hg、Tl、Ag、W和Rb)主要代表硫化作用。强负关联强迁出元素(Sr和Li)代表去碳酸盐化作用,较强正关联略富集元素II(Zr、Th、Ta、Nb和Hf)主要是去碳酸盐化作用引起体积减小,导致含量略增,两类均直接反映去

碳酸盐化作用。将两个控制因素内的元素以Au为目标,分别建立硫化和去碳酸盐化找矿标志。

5.2.1 硫化找矿标志

为判断矿化程度与其他元素含量之间的关系,Au含量设置5种组合:以 4.5×10^{-9} 为界(Au^1 和 Au^{2-5})、以 47×10^{-9} 为界、以 150×10^{-9} 为界、 150×10^{-9} 为界和以 $1\ 000 \times 10^{-9}$ 为界。As、Sb、Hg、Tl、Ag、W和Rb等元素含量设置2种组合:以中低含量界限为界,分析元素低含量类与中高含量类与不同程度Au矿化之间的特征;以中高含量界限为界,分析元素低中含量类与高含量类与不同程度Au矿化之间的特征。中低含量界限和中高含量界限见表2。Au元素的5种组合与其他元素的2种组合相互匹配,共10种组合(1-I、2-I、3-I、4-I、5-I、1-II、2-II、3-II、4-II和5-II)。以Au为目标,As、Sb、Hg、Tl、Ag、W和Rb作为输入进行关联规则分析。利用关联规则算法建立找矿模型,需考虑每种规则。设置最小支持度为0.01,最小置信度为0.01,最小提升度为0,最大前项数设为7。

结果显示:As、Sb、Hg、Tl、Ag、W和Rb元素内中高含量项或高含量项相互组合生成关联规则各有128条,每条规则内中高含量项或高含量出现项数的范围为0~7,对应规则数为1、7、21、35、35、21、7和1,对应支持度逐渐减小。同时,置信度和增益均逐渐增大,且变化趋势一致,因此主要分析置信度与矿化程度之间的特征,如图5所示。

1-I:中高含量项数为1,置信度为82.7%~98.2%,平均为91.6%,往后随中高含量项数对应的置信度迅速收敛。样品内元素中高含量项数 ≥ 1 时,样品中Au含量 $\geq 4.5 \times 10^{-9}$ (图5a)。

2-I:中高含量项数为2,置信度为76.9%~95.3%,平均为89.7%,往后随中高含量项数对应的置信度迅速收敛。样品内元素中高含量项数 ≥ 2 时,样品中Au含量 $\geq 47.0 \times 10^{-9}$ (图5b)。

3-I:中高含量数为3时,置信度大,为77.4%~94.4%,平均为86.3%,往后随中高含量项数对应的置信度迅速收敛。样品内中高含量项数 ≥ 3 时,样品中Au含量 $\geq 150 \times 10^{-9}$ (图5c)。

4-I:中高含量数为4时,置信度为59.9%~86.5%,平均为73.7%,往后随中高含量项数对应的置信度有所提高,但幅度不大。样品内中高含量项数 ≥ 4 时,Au含量 $\geq 500 \times 10^{-9}$ (图5d)。

5-I:中高含量数为5时,置信度为48.1%~

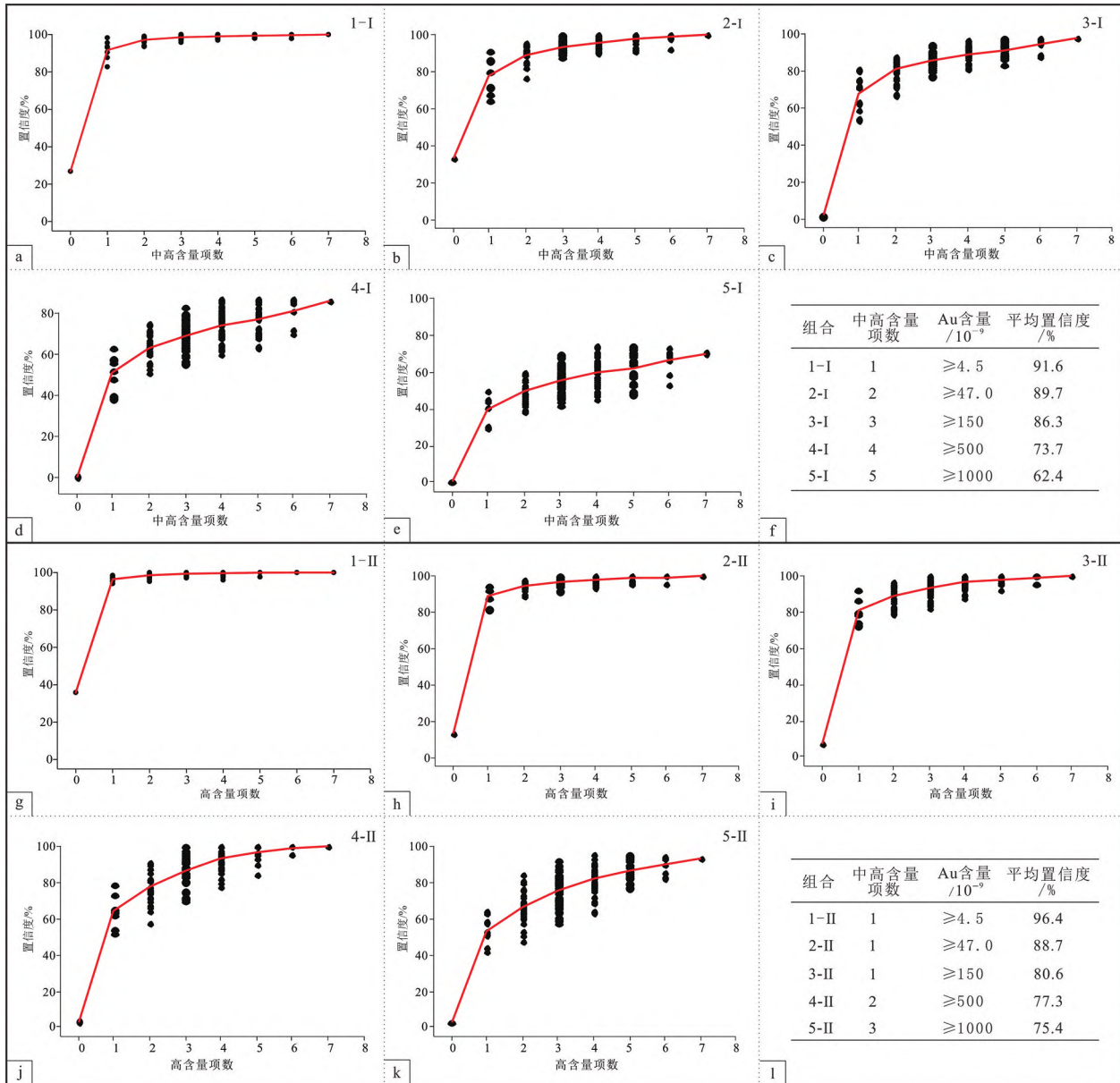


图5 硫化找矿标志分析图

Fig. 5 Data analysis for establishing sulfidation-related Au prospecting indicators

74.2%, 平均为 62.4%, 此时能够确保规则置信度整体上大于 60%。表明样品内元素中高含量项数 ≥ 5 时, Au 含量 $\geq 1000 \times 10^{-9}$ (图 5e)。

1-II: 高含量项数为 1, 置信度为 94.0% ~ 98.5%, 平均为 96.4%。往后随高含量项数增加置信度迅速收敛。故样品内元素高含量项数 ≥ 1 时, Au 含量 $\geq 4.5 \times 10^{-9}$ (图 5g)。

2-II: 高含量项数为 1, 置信度为 81.9% ~ 95.1%, 平均为 88.7%。往后随高含量项数增加置信度迅速收敛。故样品内元素高含量项数 ≥ 1 时,

Au 含量 $\geq 47.0 \times 10^{-9}$ (图 5h)。

3-II: 高含量数为 1 时, 置信度为 72.6% ~ 92.1%, 平均为 80.6% (相比 1-II 和 2-II 内高含量项数为 1 时的置信度略低), 往后随高含量项数增加置信度迅速收敛。故样品内元素高含量项数 ≥ 1 时, Au 含量 $\geq 150 \times 10^{-9}$ (图 5i)。

4-II: 高含量数为 2 时, 置信度为 57.7% ~ 91.2%, 平均为 77.3%, 大部分规则置信度均大于 60.0%, 往后随高含量项数增加置信度迅速收敛。故样品内元素高含量项数 ≥ 2 时, Au 含量 $\geq 500 \times$

10^{-9} (图 5j)。

5-II:高含量项数为 3,置信度为 57.6%~92.0%,平均为 75.4%,往后随高含量项数增加置信度迅速收敛。故认为样品内元素高含量项数 ≥ 3 时,Au 含量 $\geq 1\ 000\times 10^{-9}$ (图 5k)。

综合以上,在卡林型金矿床内根据 As、Hg、Sb、Tl、W、Ag 和 Rb 元素含量差异进行找矿和圈矿。当样品中 As、Hg、Sb、Tl、W、Ag 和 Rb 等元素内中高含量项数 ≥ 1 、 ≥ 2 、 ≥ 3 、 ≥ 4 和 ≥ 5 时,样品发生矿化,对应的 Au 含量分别为 $\geq 4.5\times 10^{-9}$ 、 $\geq 47.0\times 10^{-9}$ 、 $\geq 150\times 10^{-9}$ 、 $\geq 500\times 10^{-9}$ 和 $\geq 1\ 000\times 10^{-9}$ (图 5f)。当样品中 As、Hg、Sb、Tl、W、Ag 和 Rb 等元素内中高含量项数 ≥ 1 、 ≥ 2 和 ≥ 3 时,样品发生矿化,对应的 Au 含量分别为 $\geq 150\times 10^{-9}$ 、 $\geq 500\times 10^{-9}$ 和 $\geq 1\ 000\times 10^{-9}$ (图 5l)。因此,可根据 As、Hg、Sb、Tl、W、Ag 和 Rb 的中高含量项数和高含量项数判断卡林型金矿床矿化程度及找矿圈矿。

中高含量项数与高含量项数取并集或交集,整体均明显表现出含矿率提高、矿丢失率较低、正确率高且稳定(表 5)。但取并集或取交集时,相对存在各自的优缺点。取并集时,矿丢失率低、保存率高,但含矿率和正确率略低。取交集时,含矿率和正确率高,但矿丢失率相对取交集时较高,尤其是 Au 含量 $\geq 1\ 000\times 10^{-9}$ 时的模型组合。若仅采用中高含量项数或高含量项数判别,则找矿效果处于并集和交集之间。整体上,模型取并集,确保不漏矿。模型取交集,有效获取圈靶区,节约成本。找矿过程中,并集与交集配合使用,尽可能确保不漏矿、高效圈矿。

5.2.2 去碳酸盐化找矿标志

强负关联强迁出元素(Sr 和 Li)代表去碳酸盐

化作用,较强正关联略富集元素(Zr、Th、Ta、Nb 和 Hf)主要是去碳酸盐化作用引起体积减小,导致含量略增,两类元素均直接反映去碳酸盐化作用。去碳酸盐化是成矿过程的中间环节,并不能直接指示 Au 矿化,但可作为找矿标志。即发生去碳酸盐化部位不一定成矿,但成矿部位整体上均发生去碳酸盐化,去碳酸盐化远比金矿化范围广。因此,在区域上卡林型金矿的找矿过程中,通过寻找去碳酸盐化作用,更有可能找到金矿,去碳酸盐化作用是寻找深部和盲区矿床的重要标志。因此,通过 Sr 和 Li 负异常和高场强元素 Zr、Th、Ta、Nb 和 Hf 的正异常判断岩石是否发生去碳酸盐化。

元素的迁移富集规律结合元素的直方分布图显示(图 6):Au 的异常边界(10×10^{-9}),与预设的 4.5×10^{-9} 接近;As 的直方分布图与 Au 相似,As 的异常边界值为 120.6×10^{-6} 。Zr(100×10^{-6})、Th (2.51×10^{-6})、Ta(0.63×10^{-6})、Nb(12.6×10^{-6})和 Hf(2.51×10^{-6})的异常边界值明显低于围岩平均含量,Sr($1\ 258\times 10^{-6}$)和 Li(100×10^{-6})的异常边界值明显高于围岩平均含量,表明部分围岩已发生了去碳酸盐化作用。

以 Au 为目标(Au^1 与 Au^{2-5}),Sr、Li、Zr、Th、Ta、Nb 和 Hf(均采用异常界限进行二分类)为输入进行关联规则分析。在以 Au^{2-5} 为后项的规则中,Sr、Li、Zr、Th、Ta、Nb 和 Hf 出现异常项数 1~7 的置信度为 69.4%~79.3%,平均为 73.5%,表明这些元素有协同变化趋势,且在成矿过程中发生去碳酸盐化作用部位的 73.5%会发现矿化($Au\geq 4.5\times 10^{-9}$)。去碳酸盐化范围明显大于 Au 矿化范围,通过追索去碳酸盐化作用痕迹开展找矿。鉴于 Sr 和 Li 受背景岩性干扰较大,优选采用不活动元素 Zr、Th、

表 5 硫化找矿模型验证结果

Table 5 Validation results for sulfidation-related Au prospecting indicators

模型	Au 矿化	中高含量项数	高含量项数	样品特征			结果			
				总样品	含矿率/%	矿丢失率/%	矿保存率/%	含矿率/%	准确率/%	正确数
并集	≥ 4.5	≥ 1		1 627	73.0	8.6	91.4	82.4	79.5	1 293
	≥ 47	≥ 2		1 627	40.0	4.3	95.3	60.8	74.8	1 202
	≥ 150	≥ 3	≥ 1	1 627	39.6	5.6	94.4	67.8	80.0	1 302
	≥ 500	≥ 4	≥ 2	1 627	27.8	9.7	90.3	58.8	79.7	1 297
	$\geq 1\ 000$	≥ 5	≥ 3	1 627	21.6	18.2	81.8	52.6	80.1	1 303
交集	≥ 150	≥ 3	≥ 1	1 627	39.6	19.9	80.1	78.9	83.7	1 361
	≥ 500	≥ 4	≥ 2	1 627	27.8	27.7	72.3	68.8	83.2	1 354
	$\geq 1\ 000$	≥ 5	≥ 3	1 627	21.6	44.3	55.7	64.1	83.7	1 361

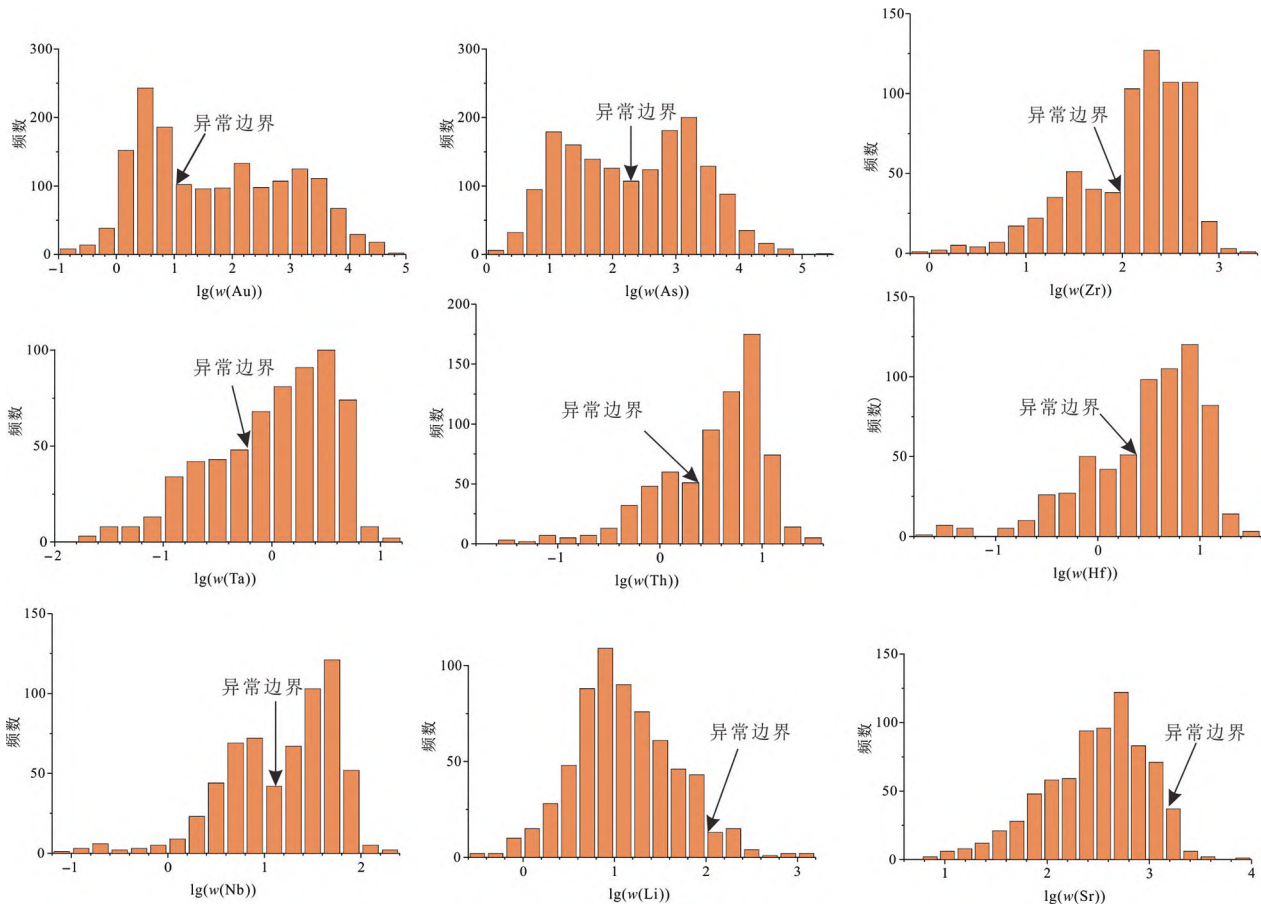


图6 去碳酸盐化相关元素直方图

Fig. 6 Histograms of elemental contents for decarbonation-related elements

Ta、Nb 和 Hf 作为找矿标志。Zr、Th、Ta、Nb 和 Hf 有相似的变化趋势, 为确保找矿标志的稳定性, 当样品中 Zr、Th、Ta、Nb 和 Hf 含量任意两项出现正异常时, 认为样品经历过碳酸盐化作用。

5.3 方法及意义

关联规则算法是大数据挖掘的基础算法, 也是最活跃的研究方法之一^[11]。关联规则算法首先由 Agrawal 等^[16]提出, 指从事务数据库、关系数据库和其他信息存储中的大量数据的项集之间发现有趣的、频繁出现的模式、关联和相关性, 从而提高对数据的理解和预测能力。本研究以 Au 为目标, 其余 26 个元素作为输入项进行关联规则分析, 生成的 1~10 条规则能够判断样品矿化程度 (表 3)。规则不仅能从与成矿关系密切的同/伴生元素 (As、Hg、Sb、Tl 和 Ag) 获取成矿信息, 还能从其他以往未引起注意的微量元素 (如 Zr、Sr、Rb、Ba、Cu、Sn、Ni 和 W 等) 中提取成矿信息, 表明关联规则算法能够挖掘矿床内潜在而不易察觉的元素

富集规律。然而, 对大量数据 (具有众多属性) 进行关联规则分析, 虽有最小支持度、最小置信度和最小增益约束, 但生成的强关联规则数量庞大, 通常具有置信度高、支持度低的特点, 导致获取的强关联规则虽然有用, 但适用范围小, 规则挑选困难。

本研究从矿床围岩与矿化角度基于关联规则算法提出了关联系数, 反映目标 (Au) 与其他元素的整体关联程度, 能够较好地识别事物的特征和规律。同时, 关联系数也能反映元素的迁移富集规律: 关联系数明显大于 1, 代表富集; 小于 1, 代表贫化; 关联系数接近 1, 代表含量变化不大。关联系数区别于常规的相关性系数, 相关性系数是研究变量之间线性相关程度的量, 反映两两变量之间相关关系密切程度的统计指标, 而关联系数反映整体中元素之间的关联程度。通过关联规则算法分析矿床内元素与 Au 矿化之间的关联性, 结合元素整体富集规律获取元素异常组合, 能够很好地从大数据角度

解释卡林型金元素富集规律的控制因素,即 Au 主要在去碳酸盐化和硫化作用下形成,伴随大量黏土化。选取与成因关联性强、富集或迁出显著的元素进行关联规则分析,分别建立起了卡林型金矿床的硫化和去碳酸盐化找矿标志。此时,生成的规则少,置信度高,包含所有可能规则,具有较好的实用价值。

通过定量识别的硫化和去碳酸盐化找矿标志可望在卡林型金矿找矿预测中发挥重要作用。此外,本研究基于关联规则算法分析矿床元素富集规律、控制因素和定量构建找矿标志的方法,也可为其他类型矿床开展类似研究提供新思路。

6 结论

(1)卡林型金矿床蚀变围岩元素异常组合分为4组:强正关联显著富集元素(As, Sb, Hg, Tl, Ag, W 和 Rb),显示硫化和黏土化(Rb)作用;较强正关联略富集元素 I(Zr, Th, Ta, Nb 和 Hf)和强负关联强迁出元素(Li 和 Sr),显示去碳酸盐化作用;较强正关联略富集元素 II(Sn, Zn, Ni, V, Co 和 Cu),显示硫化作用;弱关联基本无富集元素(Cd, Pb, Ba, Bi, U 和 Mo),与成矿无显著关联。

(2)从大数据角度获取的元素异常组合,与学界关于 Au 主要在去碳酸盐化、硫化、黏土化条件下形成的认识一致。

(3)通过关联规则算法分别对硫化和去碳酸盐化相关的元素建立了定量找矿标志。硫化找矿标志:样品中 As, Hg, Sb, Tl, W, Ag 和 Rb 等元素的中高含量项数 ≥ 1 、 ≥ 2 、 ≥ 3 、 ≥ 4 和 ≥ 5 时,对应的 Au 矿化分别为 $\geq 4.5 \times 10^{-9}$ 、 $\geq 47.0 \times 10^{-9}$ 、 $\geq 150 \times 10^{-9}$ 、 $\geq 500 \times 10^{-9}$ 和 $\geq 1\,000 \times 10^{-9}$;其中,高含量项数 ≥ 1 、 ≥ 2 和 ≥ 3 时,对应的 Au 矿化分别为 $\geq 150 \times 10^{-9}$ 、 $\geq 500 \times 10^{-9}$ 和 $\geq 1\,000 \times 10^{-9}$ 。去碳酸盐化找矿标志:样品中 Zr, Th, Ta, Nb 和 Hf 等元素含量任意两项出现正异常,表明样品经历过去碳酸盐化作用。

(4)定量识别的硫化和去碳酸盐化找矿标志可望在卡林型金矿找矿预测中发挥重要作用。此外,本研究基于关联规则算法分析矿床元素富集规律、控制因素和定量构建找矿标志的方法,也可为其他类型矿床开展类似研究提供新思路。

感谢审稿专家和编辑对本论文提出宝贵的修改意见。

参考文献

- [1] HU R Z, SU W C, BI X W, et al. Geology and geochemistry of Carlin-Type gold deposits in China[J]. Mineralium Deposita, 2002, 37(3): 378-392.
- [2] CLINE J S, MUNTEAN J L, GU X X, et al. A comparison of Carlin-Type gold deposits: Guizhou Province, golden triangle, Southwest China, and northern Nevada, USA[J]. Earth Science Frontiers, 2013, 20(1): 1-18.
- [3] MUNTEAN J L, CLINE J S. Introduction diversity of Carlin-style gold deposits[M]//MUNTEAN J L. Diversity of Carlin-Style gold deposits. Littleton, Colorado: Society of Economic Geologists, 2018.
- [4] XIE Z J, XIA Y, CLINE J S, et al. Are there carlin-type gold deposits in China? A comparison of the Guizhou, China, deposits with Nevada, USA, deposits[M]//MUNTEAN J L. Diversity in Carlin-Style gold deposits. Littleton, Colorado: Society of Economic Geologists, 2018.
- [5] HU R Z, FU S L, HUANG Y, et al. The giant South China Mesozoic low-temperature metallogenic domain: reviews and a new geodynamic model[J]. Journal of Asian Earth Sciences, 2017, 137: 9-34.
- [6] SU W C, DONG W D, ZHANG X C, et al. Carlin-Type gold deposits in the Dian-Qian-Gui "golden triangle" of Southwest China[M]//MUNTEAN J L. Diversity in Carlin-Style gold deposits. Littleton, Colorado: Society of Economic Geologists, 2018.
- [7] 刘建中,王泽鹏,杨成富,等.中国滇黔桂及周邻区卡林型金矿构造蚀变体判别指标及其意义[J].黄金科学技术, 2022, 30(4): 532-539.
- [8] 刘建中,王泽鹏,宋威方,等.滇黔桂地区卡林型金矿多层次构造滑脱成矿系统构建和找矿实践[J].地质论评, 2023, 69(2): 513-525.
- [9] 胡瑞忠,陈伟,毕献武,等.扬子克拉通前寒武纪基底对中生代大面积低温成矿的制约[J].地学前缘, 2020, 27(2): 137-150.
- [10] 胡瑞忠,高伟,付山岭,等.华南中生代陆内成矿作用[J].地学前缘, 2024, 31(1): 226-238.
- [11] 周永章,张良均,张奥多,等.地球科学大数据挖掘与机器学习[M].广州:中山大学出版社, 2018.
- [12] 周永章,左仁广,刘刚,等.数学地球科学跨越发展的十年:大数据、人工智能算法正在改变地质学[J].矿物岩石地球化学通报, 2021, 40(3): 556-573, 777.
- [13] 刘心怡,周永章.关联规则算法在粤西庞西垌地区元素异常组合研究中的应用[J].地学前缘, 2019, 26(4): 125-130.
- [14] WANG X L, CAO S T, TAN Q P, et al. Exploration vectors and indicators extracted by factor analysis and association rule algorithms at the Lintan Carlin-Type gold deposit, Youjiang Basin, China[J]. Minerals, 2024, 14(5): 492.

- <https://doi.org/10.3390/min14050492>.
- [15] 中国地质调查局,中国地质科学院矿产资源研究所,中国地质调查局成都地质调查中心.中国重要成矿区带成矿特征、资源潜力和选区部署:南盘江—右江成矿区[M].北京:中国原子能出版社,2015.
- [16] AGRAWAL R, IMIELINSKI T, SWAMI A. Mining association rules between sets of items in large databases[J]. ACM SIGMOD Record, 1993, 22(2): 207-216.
- [17] 翟裕生,姚书振,蔡克勤.矿床学[M].3版.北京:地质出版社,2011.
- [18] RUDNICK R L, GAO S. Composition of the continental crust[M]//GOLDFARB R J. Treatise on geochemistry. Amsterdam: Elsevier, 2003; 1-64.
- [19] WEI D T, XIA Y, GREGORY D D, et al. Multistage pyrites in the Nibao disseminated gold deposit, southwestern Guizhou Province, China: insights into the origin of Au from textures, *in situ* trace elements, and sulfur isotope analyses[J]. Ore Geology Reviews, 2020, 122: 103446.
- [20] ZHENG L J, TAN Q P, ZUO Y J, et al. Two hydrothermal events associated with Au mineralization in the Youjiang Basin, southwestern China [J]. Ore Geology Reviews, 2022, 144: 104816.
- [21] HU H Y, CHAN L H, ZHANG L B, et al. Lithium and its isotopes in major world rivers: implications for weathering and the oceanic budget[J]. Geochimica et Cosmochimica Acta, 1998, 62(12): 2039-2051.
- [22] 张宏飞,高山.地球化学[M].北京:地质出版社,2012.
- [23] ZHANG Z Y, MA J L, WANG Z B, et al. Rubidium isotope fractionation during chemical weathering of granite [J]. Geochimica et Cosmochimica Acta, 2021, 313: 99-115.
- [24] 赵振华.微量元素地球化学原理[M].2版.北京:科学出版社,2016.
- [25] 叶天竺,韦昌山,王玉往,等.勘查区找矿预测理论与方法:各论[M].北京:地质出版社,2017.