

基于XGBoost模型的西藏山南金锑矿成矿预测

张毅明^{1,2}, 汤宇磊^{3*}, 韩聪², 邓俊峰², 黄勇²

1. 成都理工大学地球与行星科学学院, 四川 成都 610059;
2. 中国地质调查局军民融合地质调查中心, 四川 成都 610036;
3. 中国地质调查局地球物理调查中心, 河北 廊坊 065000



摘要: 西藏山南地处特提斯喜马拉雅东段, 是我国金、锑矿床集中发育的地区之一。其中, 措美县—隆子县一带已发现金、锑矿床(点)60余个。为了进一步发掘潜在矿床(点), 从而实现找矿突破, 可采用机器学习方法对潜在矿床(点)进行预测。以西藏山南措美县—隆子县15个1:5万矿调标准图幅范围作为研究区, 融合地质、地球化学、地球物理和地形数据等, 为该区金、锑矿构建了2个XGBoost预测模型, 模型精度分别为0.8807和0.9031。利用机器学习方法获得该区金、锑矿成矿预测数据集, 圈定了金、锑矿成矿远景区及找矿靶区。研究表明, XGBoost模型在西藏山南金锑矿成矿预测方面表现良好, 根据预测结果圈定金找矿靶区27个, 锑找矿靶区11个。研究成果可为该区域金锑矿找矿工作部署提供参考。

关键词: 金锑矿; 机器学习; 成矿预测; XGBoost模型; 措美县—隆子县; 西藏

中图分类号: P618.51 文献标志码: A 文章编号: 1005-2518(2026)03-0628-14 DOI: 10.11872/j.issn.1005-2518.2026.03.256

引用格式: 张毅明, 汤宇磊, 韩聪, 等. 基于XGBoost模型的西藏山南金锑矿成矿预测[J]. 黄金科学技术, 2026, 34(3): 628-641.
ZHANG Yiming, TANG Yulei, HAN Cong, et al. Mineralization Prediction of Shannan Gold and Antimony Deposits in Xizang Based on XGBoost Model[J]. Gold Science and Technology, 2026, 34(3): 628-641.

近年来, 随着地质大数据(陈建平等, 2017; 周永章等, 2018a; Li et al, 2021)和机器学习技术(周永章等, 2018b; 马瑶等, 2021)的快速发展, 作为数据驱动的智能分析技术, 机器学习成功地将成矿预测从定性描述阶段转向定位定量研究阶段(袁峰等, 2014)。其核心优势在于能够从地质、地球化学和遥感等多源异构数据中挖掘非线性关联规律, 突破传统单一手段仅能够反映成矿环境局部特征的局限性。机器学习不仅能够通过集成处理大量空间数据识别隐式矿化异常信息, 还能够构建多要素协同的成矿预测模型, 显著提升成矿预测的精度和

效率(周永章等, 2017)。众多学者先后通过机器学习算法验证了不同地区二维找矿预测模型(Li et al, 2020; 张振杰等, 2021; 欧阳渊等, 2023; 张必敏等, 2025)和不同矿区隐伏矿体三维预测模型的有效性(袁峰等, 2019; 陈进等, 2020; 姜渝明, 2023; 姜渝明等, 2025)。

在技术实现层面, 模型通过数据驱动、知识驱动或混合驱动模式, 对包含矿化标签及特征变量的样本数据集进行训练, 建立非线性分类器来量化成矿概率。如: 基于数据驱动的极端梯度提升模型, 采用增量式训练和特征分裂点精细调整等方式, 在

收稿日期: 2025-08-24; 修订日期: 2026-03-01

基金项目: 国家科技重大专项课题“北喜马拉雅东段锡锑多金属成矿模型与找矿预测”(编号: 2025ZD1006304)、自然资源部科技课题“藏南措美地区锑多金属勘查技术组合与示范”(编号: ZKKJ202427-03-04)、中国地质调查局地质调查项目“西藏措美—哲古地区区块优选调查评价”(编号: DD20240382)、中国地质调查局地质调查项目“西藏沿边抵边地区地质矿产数据集成与更新”(编号: DD20243090)和西藏自治区科技计划项目“喜马拉雅大型岩金矿床成矿作用及勘查示范”(编号: XZ202401YD0006-07)联合资助

作者简介: 张毅明(1996—), 男, 甘肃武威人, 工程师, 从事资源勘查与成矿预测、资源与环境遥感研究工作。E-mail: zhangyiming@mail.cgs.gov.cn

***通信作者:** 汤宇磊(1990—), 男, 安徽合肥人, 高级工程师, 从事资源环境遥感与机器学习应用研究工作。E-mail: tangyl_env@outlook.com

© Editorial Department of Gold Science and Technology (CC BY-NC-ND)

不断迭代中优化变量选择和权重参数调整,具有训练速度快且预测效果好的优势;基于地质知识图谱和大语言模型的预测系统,通过机器阅读与语义推理,实现具备因果解释能力的预测过程;基于知识、数据与算法的混合驱动人机交互模型,使人工智能预测实现“数据—知识—决策”的高阶应用(成秋明,2025;左仁广,2025)。

西藏山南地处特提斯喜马拉雅构造域东段,属于藏南金锑矿成矿带,是我国金、锑矿床发育较为集中的区域之一,尤其是措美县—隆子县一带。近10年来,研究人员在该区开展了1:5万区域地质矿产调查项目,发现60余个金、锑矿床和矿(化)点,并采用信息量算法进行成矿预测。信息量算法的优势在于解释性强且计算高效,适合小样本数据下的特征筛选和地质标志提取,但无法处理大数据和理解复杂成矿规律。由于矿床(点)与赋矿地层、控矿构造、岩浆活动、变质作用、地球化学异常、地球物理异常具有一定的成因联系,通过集成这些成矿有利信息进行机器学习模型训练,挖掘其与成矿潜在的时空耦合关系(聂虹等,2018;左仁广,2019,2021),能够实现潜在矿床(点)的成矿概率预测(Zuo et al, 2020; Xiong et al, 2018)。本文选择西藏山南措美县—隆子县15个1:5万矿调标准图幅范围作为研究区,融合地质、地球化学、地球物理和地形等数据,基于XGBoost算法构建金、锑矿预测模型,利用机器学习方法获得该区金、锑矿成矿预测数据集,并据此圈定金、锑矿找矿靶区,进一步指导找矿工作部署。

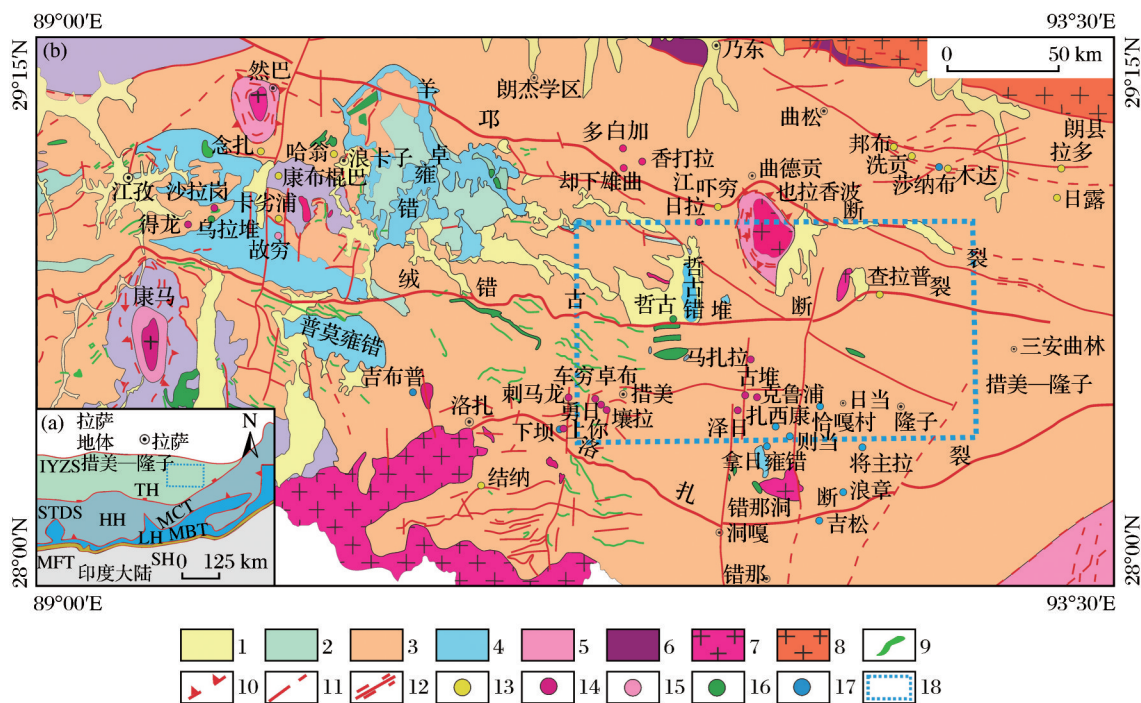
1 成矿地质背景

研究区主要位于西藏山南市措美县—隆子县,包括哲古、卡珠、邛多江、宗许、民那、松多、让宗、穷科当、俗坡下、雪莎区、卡果、甲坞、日当、隆子县和列麦幅共15个1:5万标准图幅,地跨琼结、乃东、曲松、措美、错那和隆子6个区县,覆盖面积约为6 780 km²,地理坐标为91°30′00″E~92°45′00″E, 28°20′00″N~28°50′00″N。区内广泛分布有金、锑和铅锌锑银多金属矿床(点),通过整理历年区矿调项目资料,梳理区内锑矿床(点)26个,岩金矿床(点)21个。其中,金、锑矿床主要类型为热液型。

研究区位于青藏高原南部的特提斯喜马拉雅构造域东段,地处印度河—雅鲁藏布江缝合带和藏南拆离系主拆离面之间[图1(a)]。三叠纪以后,区域构造活动强烈,先后经历了特提斯洋盆扩张、消减与闭合、喜马拉雅陆块与冈底斯陆块碰撞造山以及剪切走滑与伸展拆离作用,形成了复杂的区域大地构造格局。区内地层属于喜马拉雅地层区的康马—隆子地层分区,地层发育较完整,仅缺失晚古生代地层(图2)。其中,前寒武系亚堆扎拉岩群出露于北部也拉香波变质核杂岩内核;早古生代曲德贡岩组出露于也拉香波变质核杂岩外侧;三叠—侏罗纪新特提斯洋扩张形成一套含成矿元素的碳硅泥质岩系,为区内大面积出露的主要地层,后期受构造、岩浆和变质作用叠加改造,成为区内主要的赋矿地层。

研究区与成矿相关的构造主要为断裂和穹窿构造。古近纪渐新世同碰撞阶段特提斯喜马拉雅形成大规模近EW向拆离断层,以邛多江、古堆和洛扎拆离断层带为代表,区内金多金属矿床主要受该断层控制,典型矿床有邦卓玛金矿(姜元林,2019)和马扎拉锑矿(董富权,2012;邓舟,2020)。新近纪中新世在碰撞后伸展背景下形成近SN向高角度正断层,产生堑垒式构造和断陷盆地,以洞嘎、勒金康桑和下坝断裂带为代表[图1(b)],区内锑矿床主要受该断裂带控制,典型矿床有车穷卓布锑矿和勇日锑矿。另外,铅锌银锑多金属矿床主要受控于近EW向拆离断层与近SN向正断层的交会部位(张刚阳,2012),典型矿床有扎西康铅锌锑多金属矿、桑日则铅锌锑矿、柯月锑铅锌矿、索月锑铅锌矿和则当锑铅锌多金属矿等矿床。此外,研究表明区内穹窿构造与金锑的成矿作用具有成因联系,伸展拆离导致地壳重熔形成也拉香波穹窿内的淡色花岗岩,其多阶段的岩浆活动为金属成矿提供了成矿物质和动力来源,导致变质核杂岩体内部及周边形成大量金、金锑矿床(点)(张刚阳,2012),典型矿床有晒日金矿和达拉金矿。

研究区岩浆侵入和火山活动较为频繁,分布有早白垩—始新世中基性侵入岩、渐新—中新世酸性侵入岩以及晚三叠世、中侏罗世和晚侏罗—早白垩世中基性火山岩。晚三叠—早白垩世火山活动和早白垩世岩浆侵位带来了深部成矿物质,为初始



MFT-主前缘逆冲断裂;MBT-主边界逆冲断裂;MCT-主中央逆冲断裂;STDS-藏南拆离系;IYZS-印度—雅鲁藏布江缝合带;SH-次喜马拉雅;LH-低喜马拉雅;HH-高喜马拉雅;TH-特提斯喜马拉雅;1.第四系;2.上白垩统;3.上三叠统一白垩统;4.奥陶系—下三叠统;5.前寒武系;6.罗布莎群变形橄榄岩;7.新生代花岗岩;8.白垩纪花岗岩;9.白垩纪中基性脉岩;10.拆离断层;11.断裂构造;12.韧性剪切带;13.金矿;14.锑矿;15.金银矿;16.锑金矿;17.锑多金属矿;18.措美—隆子地区

图1 喜马拉雅地体构造格架简图(a)及特提斯喜马拉雅区域地质图(b)

Fig. 1 Simplified tectonic framework map of the Himalayan Terrane (a) and regional geological map of the Tethyan Himalaya (b)

矿源层的富集提供了物质来源,古近纪和新近纪中酸性岩浆侵位为成矿提供了物源、热源及动力。喜马拉雅期强烈的构造作用及岩浆活动,引发了广泛的区域变质、动力变质和接触变质作用。除第四系松散沉积物未受明显变质影响之外,区内其余地层岩石均经历了不同程度的变质改造(姜元林等, 2024)。其中,区域变质作用形成以板岩、片岩和变质砂岩为主的岩石组合,属于绿片岩相。总之,地层、构造、岩浆—热液活动和变质作用共同参与了区内金锑成矿过程(黄小东, 2011)。

2 数据与方法

2.1 数据来源与预处理

(1)数据来源

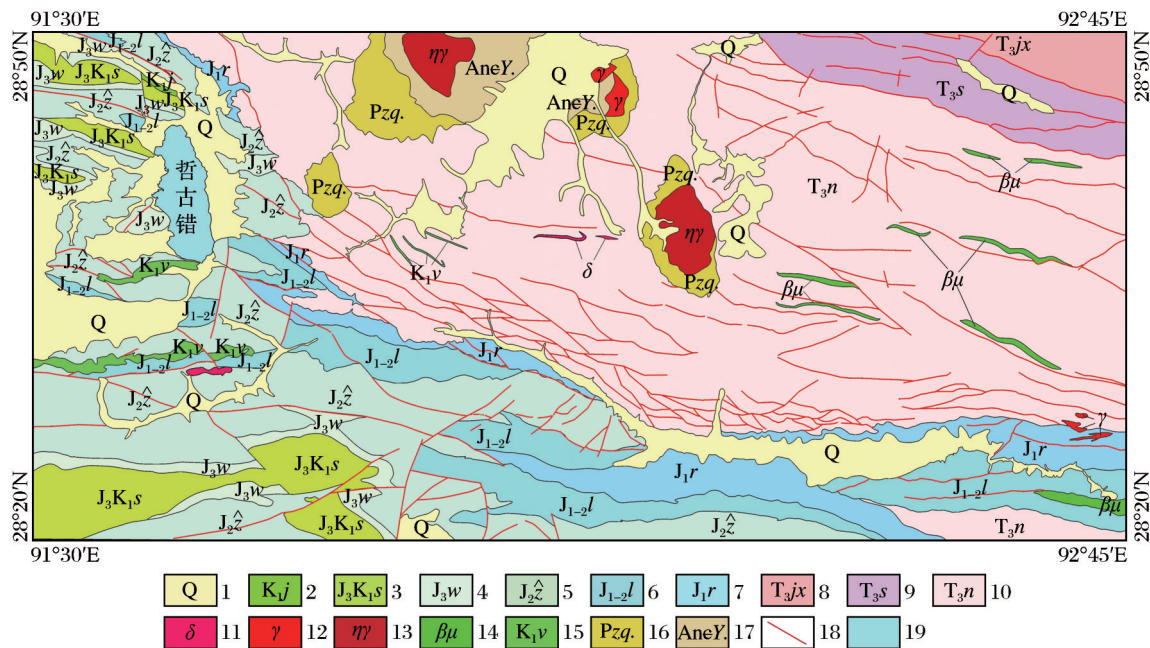
本文地质与地球化学数据来自中国地质调查局,主要包括哲古幅等15个1:5万图幅的地层、构

造、岩浆岩、变质岩数据和地球化学水系沉积物样品分析测试数据(除宗许幅之外),包含26 655件样品的Sb、Au、S、Pb、Zn、Cu、Ag、Hg和Bi共9个元素近240 000条测试结果,区内已发现的21个岩金和26个锑矿床(点),包括矿种、坐标和规模等数据。剩余重力异常数据来自西藏自治区地质调查院, SRTM地形数据来自美国太空总署和美国国家测绘局。数据来源统计如表1所示。

(2)变量提取

特征变量提取是机器学习建模预测的关键步骤,是筛选提取针对潜在矿床(点)具有指示意义的多源异构数据的重要手段。本文选取与研究区内金、锑矿床成矿密切相关的地质、构造、岩浆岩、变质岩数据、水系沉积物样品数据、剩余重力异常数据和SRTM地形数据作为特征变量参与建模(图3)。

第一,地层数据。为了清楚地识别不同时代地层对成矿的贡献,将全区地层划分为三叠系、侏罗



1.第四系松散堆积物;2.甲不拉组:钙质页岩与灰岩互层;3.桑秀组:变质石英砂岩、变砂质泥岩夹变质石英砂岩;4.维美组:页岩与变质石英砂岩互层;5.遮拉组:变泥质粉砂岩、变粉砂质泥岩夹变质石英砂岩;6.陆热组:变钙质泥岩、变粉砂质泥岩夹变质泥灰岩;7.日当组:板岩与变质泥灰岩或变质砂岩互层;8.江雄组:变质长石石英砂岩、泥质板岩和千枚岩;9.宋热组:变质长石石英砂岩和变质石英砂岩;10.涅如组:变粉砂泥质板岩夹变质长石石英杂砂岩和变质长石石英杂砂岩;11.闪长岩;12.花岗岩;13.二长花岗岩;14.辉绿岩;15.辉长岩;16.曲德贡岩组:变质砂板岩;17.前寒武系亚堆拉岩群:混合质变质岩、混合岩和混合花岗岩;18.断裂;19.哲古错

图2 措美—隆子地区地质简图

Fig. 2 Simplified geological map of the Cuomei-Longzi area

表1 数据来源统计

Table 1 Statistics of data sources

数据	来源	格式	比例尺/分辨率
地层数据	中国地质调查局	MPJ,shp	1:5万
构造数据	中国地质调查局	MPJ,shp	1:5万
岩浆岩数据	中国地质调查局	MPJ,shp	1:5万
变质岩数据	中国地质调查局	MPJ,shp	1:5万
矿床(点)数据	中国地质调查局	Xlsx,shp	1:5万
水系沉积物数据	中国地质调查局	Xlsx,shp	1:5万
剩余重力异常数据	西藏地质调查院	MPJ,shp	1:50万
SRTM地形数据	美国航空航天局	tif	30 m

系及第四系3类数据参与建模。白垩纪地层分布面积积极小,几乎对模型不产生影响,因此忽略不计。第二,岩浆岩数据。区内成规模分布的岩浆岩主要为渐新—中新世中—中酸性侵入岩和白垩世中基性侵入岩,将二者分别提取为单独数据层参与建模。第三,变质岩数据。区内变质岩主要为分布于变质核杂岩内的亚堆扎拉岩群和曲德贡岩组变质结晶基底。第四,构造数据。由于断裂是区内主要

控矿因素,为了更好地识别控矿断裂并探讨其对成矿的贡献,区分近EW、近SN、NW和NE向4种不同方向提取区内断裂,形成线状要素。第五,化探数据。研究区具有明显的Au、Sb、As、Pb、Zn和Ag等单元素异常,以及Au-As-Sb-Hg、Ag-Pb-Zn-Mn-Cd-U-Be-Bi等元素组合异常,异常区内各元素套合好,异常规模较大,具有明显的浓集中心和浓度分带,因此提取9种主要致矿异常元素(Au、Sb、As、Pb、Zn、Ag、Cu、Hg和Bi)水系沉积物测量数据。第六,物探数据。研究区重力异常总体反映深部为由南向北倾斜的地幔缓坡区,具体表现为由南向北快速下降的等值线密集带,等值线呈近EW向展布,叠加局部重力高异常。第七,地形数据。热液矿床常发育硅化蚀变,岩体抗风化能力强,易形成正地形。发育泥化或碳酸盐化蚀变的破碎岩层易形成带状负地形,是矿化的间接标志(张刚阳,2012)。

(3) 样本构建

将所有已发现矿床、矿点和矿化点作为正样

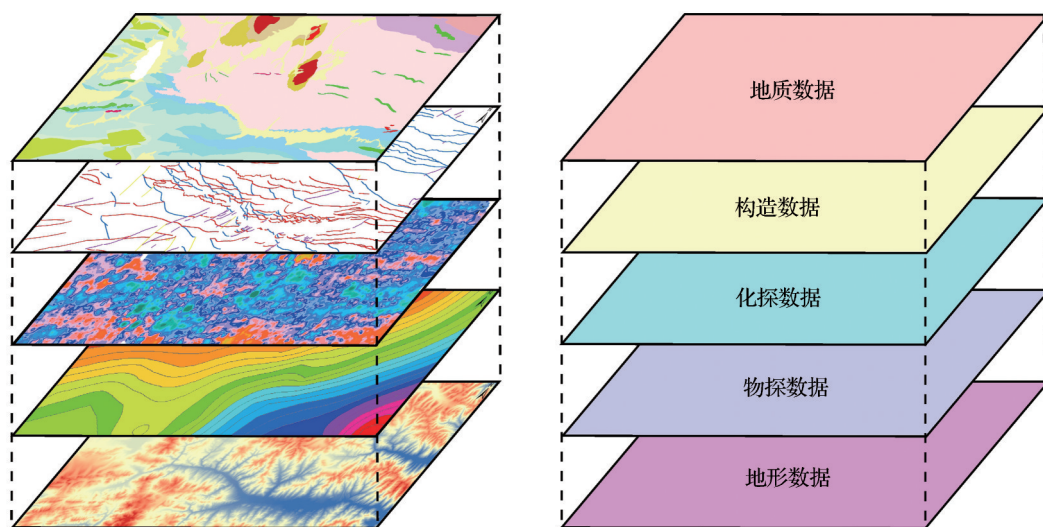


图3 变量提取示意图

Fig. 3 Schematic diagram of variable extraction

本,岩金矿床(点)共21个,锑矿床(点)共26个,随机选取30%的正样本作为模型的验证集,正样本(点)基本信息见表2。在已知矿床(点)、控矿构造和地球化学组合异常等成矿有利区域外,随机均匀生成26个点作为无矿负样本(Zuo et al,2020)。

(4)预处理

由于区内构造规模多延伸数千米至数十千米,每0.25 km²范围内至少有一个水系沉积物采样点,因此,运用网格法(Laffan et al,2005)将研究区按照251×101的网度划分为25 351个网格单元,每个网格单元为边长0.5 km的正方形,空间分辨率为0.5×0.5 km²,兼顾建模预测计算效率的同时,提供足够

空间分辨率的预测结果,满足找矿靶区的圈定需求。建立缓冲区是成矿预测中常用的数据增强方法,结合前人经验,按成矿有利区影响范围分别为超大、大、中、小型矿床、矿点和矿化点建立5 km、4 km、3 km、2 km、1 km和0.5 km缓冲区,同时为断裂建立0.5 km缓冲区。另外,将有效地球化学数据进行协同克里金插值,将30 m分辨率地形数据重采样生成0.5 km分辨率数据。所有变量网格格点值作为模型训练集,构建一个地质—地球化学—地球物理—地形协变量基础数据库,所有数据储存于PostgreSQL数据库,通过R、Python等语言进行处理分析。

表2 正样本基本信息

Table 2 Basic information of positive samples

矿种	规模	矿床、矿点和矿化点名称
金矿	大型矿床	查拉普金矿
	中型矿床	马扎拉金矿、姐纳各普金矿
	矿点	泽日金锑多金属矿、古堆乡金锑矿、扭麦拉金矿、错嘎金矿、邦卓玛金多金属矿点、宁拉金矿点、那穷金锑矿点、卓本日金多金属矿点、赛龙寺金矿点、卓日金多金属矿点
	矿化点	日拉金矿、哲古错金矿、当布金矿、果荣岩金矿、晒日岩金矿、达拉金矿点、多巴金矿点、那嘎迪金矿化点
锑矿	中型矿床	民那锑矿、柯月锑铅矿
	小型矿床	错姆其村锑矿
	矿点	沾玛锑金矿、哲古错锑金矿、纳鲁锑矿、拉定锑矿、姜仓锑金矿、龙中日锑矿、每金扎洛锑矿、克鲁浦锑矿、鼓多锑矿、龙中日锑矿、甲钨锑矿、泽日锑矿、则当锑铅锌多金属矿、象日锑矿点、普容朗锑矿、恰嘎锑矿点、古堆锑矿点、恰嘎锑矿点、克鲁浦锑矿点、龙中日锑矿点
	矿化点	嘎乌日锑矿、索月锑铅锌多金属矿、玉堆锑金矿

2.2 模型构建

(1) 技术方法

本文采用XGBoost算法构建成矿预测模型。XGBoost是一种基于Boosting集成框架的机器学习算法,其核心思想是通过逐次迭代构建弱学习器CART(Classification and Regression Tree)决策树以优化模型预测性能(Yin et al, 2022),其计算公式为

$$y = \sum_{k=1}^k f_k(x_i), f_k(x_i) \in F \quad (1)$$

式中: y 为模型预测值; i 为样本序号; x_i 为第 i 个样本的特征向量; F 为所有树的集合; k 为树的总数;

f_k 为经过第 k 轮迭代所得的第 k 棵树。

与传统梯度提升决策树GBDT(Gradient Boosting Decision Tree)相比,XGBoost引入二阶泰勒展开对损失函数进行精细化近似,并融合正则化项(L1/L2)以约束模型复杂度,从而有效平衡拟合能力和泛化性能。算法通过迭代优化特征权重,能够有效提取各类控矿因素之间的深层耦合关系,如构造交会部位、地球化学异常组合和地球物理异常区域等关键找矿信息,从而生成高置信度且高分辨率的成矿预测数据集(图4)。

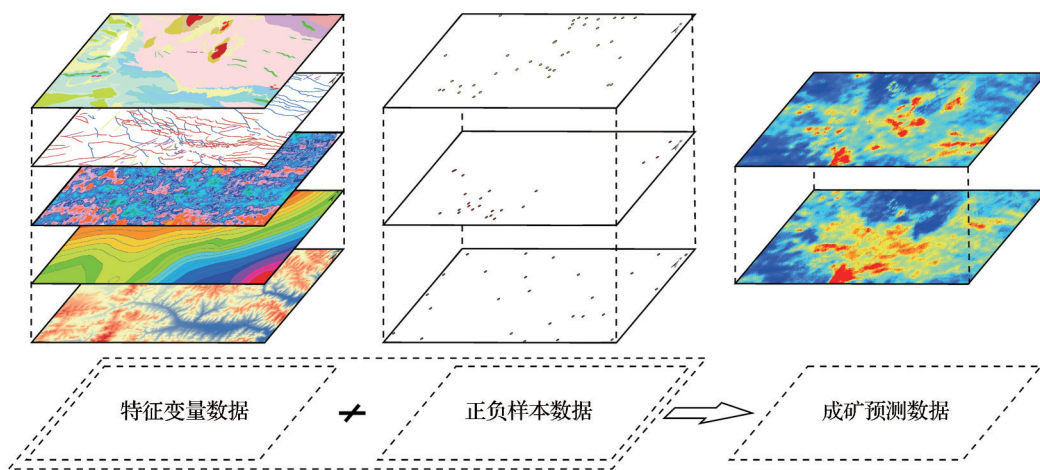


图4 技术方法示意图

Fig. 4 Schematic diagram of technical methods

(2) 技术路线

首先,搜集研究区地质、地球化学、地球物理和地形数据,建立已知矿床(点)正样本和无矿负样本,对地质和地球化学数据进行增强处理。其次,将所有数据进行预处理,如插值、重采样、对数变换和建立缓冲区等,利用网格法划分研究区并将各类数据叠加匹配至全区网格。然后,划分样本训练集、测试集和验证集,利用样本数据和特征变量建立XGBoost模型,调参优化使其预测性能达到最优。最后,开展模型精度评估,计算变量相对重要性,并输出成矿预测结果,技术路线如图5所示。

3 结果与讨论

3.1 模型评估

金矿成矿系统复杂且控矿因素较多,多期次、

多阶段矿化特征明显,因此设置较大的最大深度以捕捉高阶规律,搭配260个估计器控制过拟合。锑矿成矿控制因素较单一,数据噪声更少,采用较小的最大深度规避噪声,搭配310个估计器提升拟合精度。其余参数统一设置,具体如下:学习率为0.05,以平衡收敛和稳定性;样本权重为0.67,解决样本不平衡;特征采样比例为0.7,提升泛化性,以适配成矿预测共性需求。金、锑成矿预测模型相关参数如表3所示。

(1) mIoU 均值

本研究利用mIoU(Mean Intersection over Union, mIoU)均值进行模型评估,该指标是预测模型常用的评估手段。它是真实值和预测值2个集合的交集和并集之比的平均值。通过对每个子模型在验证集上的性能加权平均输出权重,对每个像元成矿概率的加权平均值阈值化,若某个像元成矿概率的

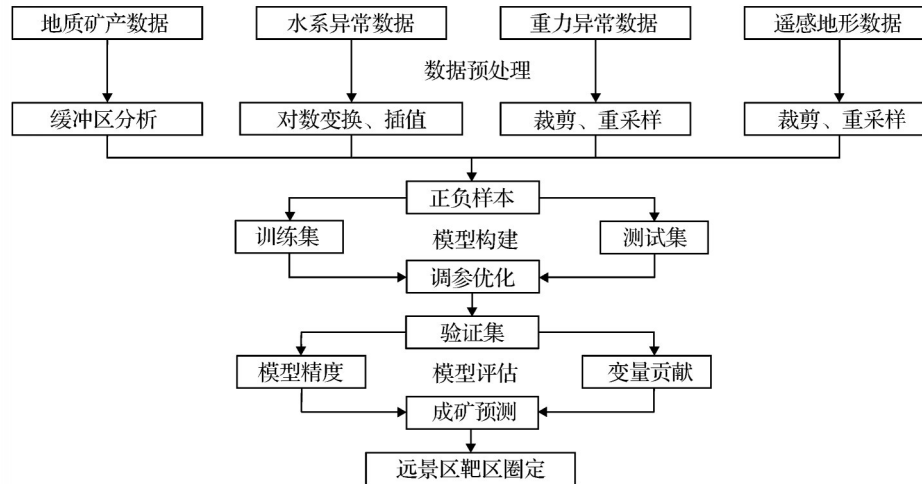


图5 技术路线图

Fig. 5 Diagram of technical route

表3 预测模型参数

Table 3 Parameters of prediction models

超参数	参数设置区域	金矿模型参数	锑矿模型参数
最大深度 (max_depth)	[3,4,5,6,7]	4	3
学习率 (learning_rate)	[0.01,0.05,0.1,0.2]	0.05	0.05
估计器数量 (n_estimators)	100~500	260	310
样本权重 (scale_pos_weight)	[0.5,0.67,1]	0.67	0.67
特征采样比例 (colsample_bytree)	0.6~0.9	0.7	0.7

加权平均值大于70%,则该像元被认为具有成矿潜力,像元成矿概率的加权平均值计算公式为

$$P(c)=\sum_{i=1}^{10}P_i(c)mIoU_i \quad (2)$$

式中: c 为待预测的类; $P_i(c)$ 为每个子模型的输出概率; $mIoU_i$ 为验证集上计算的第 i 个子模型的性能 $mIoU$, $mIoU$ 的计算公式为

$$mIoU=\frac{1}{n}\sum\frac{TP}{TP+FP+FN} \quad (3)$$

式中: TP (True Positive)表示正确预测为正类的像素数; FP (False Positive)表示错误预测为正类的像素数; FN (False Negative)表示错误预测为负类的像素数; n 为类别总数。

$mIoU$ 能够精准量化表征模型预测值与真实值(矿床、矿点和矿化点)在像素维度的重叠水平,可直接映射分割边界的精准度与类别归属的正确性。若 $mIoU$ 值为0~1,且越接近1,表示成矿高概率区与已知矿床(点)所在网格的重叠度越高,空间定位越准确,模型的预测性能越好。经加权平均计算,

$mIoU(Au)=0.8807$, $mIoU(Sb)=0.9031$,表明金、锑矿成矿预测模型具备良好的预测性能。

(2)ROC曲线

受试者操作特征曲线(Receiver Operating Characteristic Curve, ROC)是二分类模型性能评估的核心工具,其通过动态阈值调整揭示分类器的判别能力,广泛应用于成矿预测模型评价领域。在成矿预测二分类任务中,ROC曲线直观地展示了模型在不同分类阈值下,对“矿床(点)”与“地质背景”辨别能力的平衡关系。计算公式分别为

$$FPR=\frac{FP}{FP+FN} \quad (4)$$

$$TPR=\frac{TP}{TP+FN} \quad (5)$$

其中, TP 、 FP 和 FN 的含义与式(3)相同。模型通过概率排序生成阈值序列,从最高阈值(全样本判为负类)逐步降低至最低阈值(全样本判为正类),逐点计算对应的 FPR 与 TPR 并连接成曲线。以曲线下方面积(Area Under the Curve, AUC)作为量化指标,

将性能评估简化为单一数值,数值越接近1,代表分类精度越高。经计算, $AUC(Au)=0.915$, $AUC(Sb)=0.923$,表明金、锑矿成矿预测模型具有良好的预测性能。金、锑矿预测模型ROC曲线如图6和图7所示。

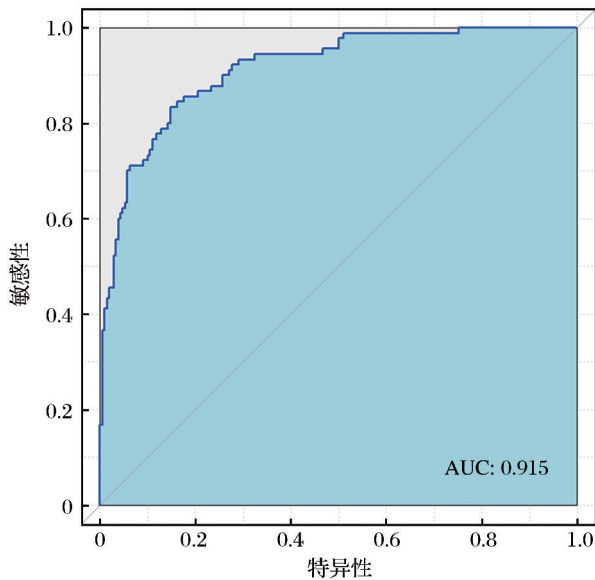


图6 金矿模型ROC曲线

Fig. 6 ROC curve of gold deposit model

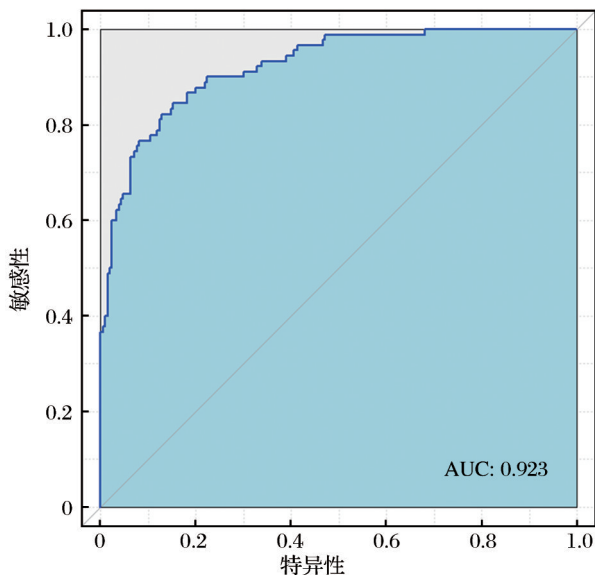


图7 锑矿模型ROC曲线

Fig. 7 ROC curve of antimony deposit model

3.2 变量相对重要性

变量相对重要性是衡量模型变量贡献程度的评判依据。在对数据进行回归分类的同时,计算各个变量的权重指数,评估各变量在预测中所起的作用

(Sun et al, 2019; 蒋海昆等, 2023)。在每次实验中,将包含特定变量的子模型全部去除,使用剩余的子模型组成新的加权平均集成模型。本研究利用参与建模的地层(三叠系T、侏罗系J、第四系Q)、岩浆岩(中酸性IA、中基性IB)、变质岩(MR)、构造(近东西GS-EW、近南北GS-SN、北西GS-NW、北东GS-NE和穹窿GS-Dome)、不同元素水系沉积物含量、剩余重力异常(RGA)和地形数据(DEM)作为变量,通过逐个置换变量并量化预测误差,输出各个变量的贡献(图8和图9)。

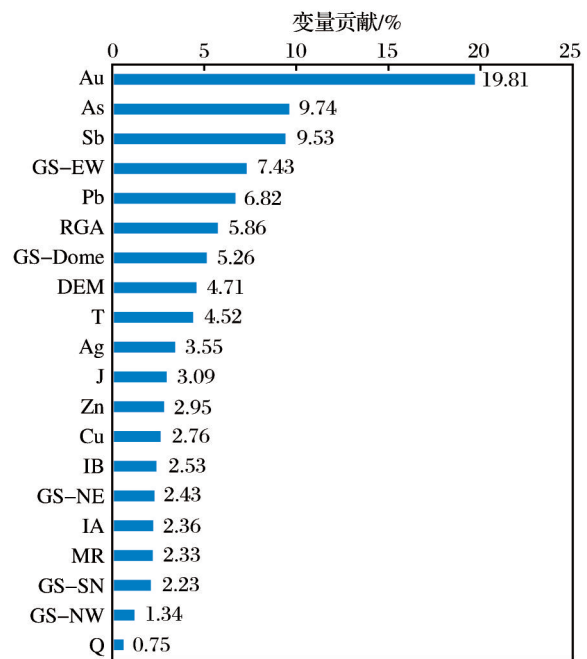


图8 金矿模型变量贡献

Fig. 8 Variable contribution of gold deposit model

在水系沉积物方面,Au、As、Sb和Pb元素是金矿成矿预测模型相对重要性较高的变量(>5%),分别占模型权重的19.81%、9.74%、9.53%和6.82%,贡献近半。其均为强亲硫元素,在成矿过程中优先与硫结合形成硫化物或硫盐矿物。在金锑成矿的中低温热液环境中,四者常以硫氢络合物形式迁移,相同的络合形式让它们能随热液同步迁移并富集沉淀,间接揭示了金与毒砂、辉锑矿、方铅矿和黄铁矿等金属硫化物可能存在共伴生关系。Sb、As、Hg、Bi和Ag元素是锑矿成矿预测模型相对重要性较高的变量(>5%),分别占模型权重的17.64%、11.22%、8.73%、6.24%和5.48%,贡献近半,这些元

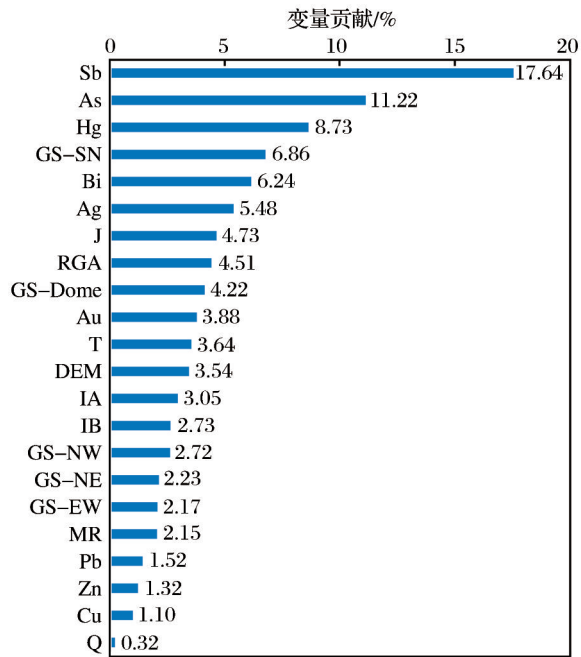


图9 锑矿模型变量贡献

Fig. 9 Variable contribution of antimony deposit model

素在成矿过程中同样具有高度类似的地球化学行为,亲硫性强、络合物一致,且常发生类质同象替代,间接揭示了锑矿形成于造山带或断裂带的浅成低温热液环境, $\text{Bi} \rightarrow \text{Ag} \rightarrow \text{Sb} \rightarrow \text{Hg}$ 依次记录了降温减压的流体演化过程。

在地层方面,三叠系对金矿模型的重要性高于侏罗系,表明三叠系与金矿相关性更强,侏罗系对锑矿模型的重要性高于三叠系,表明侏罗系与锑矿相关性更强,这种关系与区域矿床赋矿地层分布规

律一致;在构造方面,金矿模型近EW向断裂相对重要性占比为7.43%,穹窿构造占比为5.26%,表明金矿与近EW向断裂和穹窿具有较强的相关性。锑矿模型近SN向断裂占比为6.86%,表明锑矿与近SN向断裂具有较强的相关性。其余各向断裂重要性没有明显差异;在岩浆岩方面,金矿模型中生代中基性岩浆岩相对重要性略高于新生代中酸性岩浆岩,锑矿模型反之,与区域上金矿与中生代近EW向构造—中基性岩浆作用联系密切。锑矿与新生代近SN向构造—中酸性岩浆作用联系密切的认识一致;在变质岩方面,前寒武系和寒武系变质结晶基底对2个模型的贡献均不高,可能由于其并未直接参与成矿,且与金锑矿床不具有明显的空间相关性。剩余重力异常和地形的相对重要性较高,且金矿模型明显高于锑矿,表明二者分布与金矿的空间相关性高于锑矿。

3.3 成矿预测结果

基于XGBoost模型的金、锑矿成矿预测结果如图10和图11所示,成矿概率如图12和图13所示。根据成矿预测结果,该区金成矿预测高值区主要分布于中部、南部及东南部,如象日—古堆、那穷—赛龙寺、查拉普—错嘎、马扎拉—拉定、克鲁浦—扎西康、恰嘎和列麦等地区。锑成矿预测高值区主要分布于中部和南部,如象日—古堆、那穷—赛龙寺、克鲁浦—扎西康和恰嘎等地区,这些地区多为构造—岩浆活动耦合作用区,地球化学组合异常明显,多分布中小型金锑矿床(点)。

基于机器学习的成矿预测结果,通常用于圈

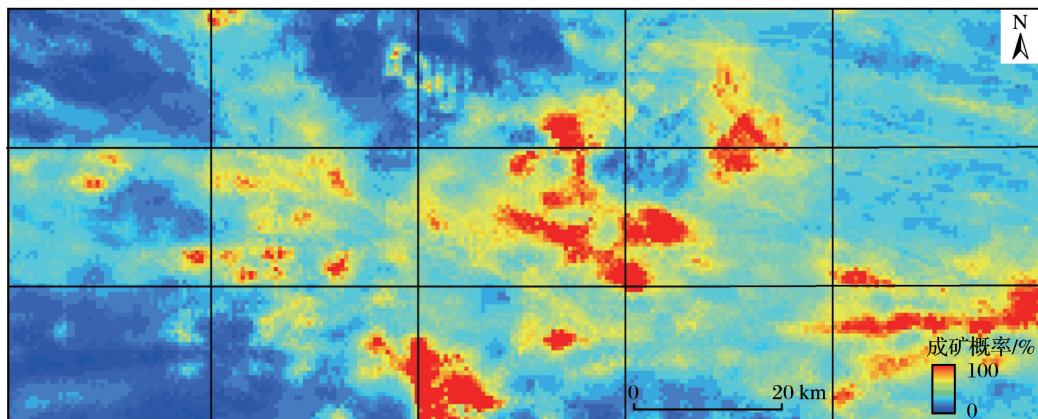


图10 金矿成矿预测结果

Fig. 10 Metallogenetic prediction results of gold deposits

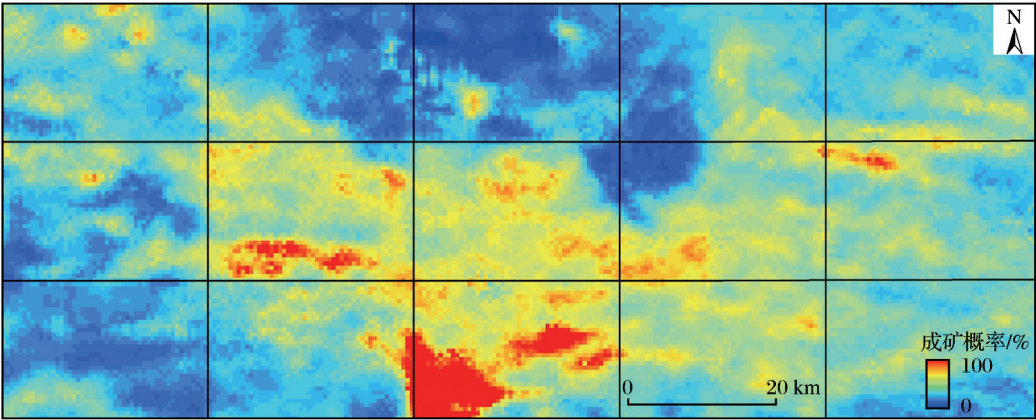


图 11 锑矿成矿预测结果
Fig. 11 Metallogenic prediction results of antimony deposits

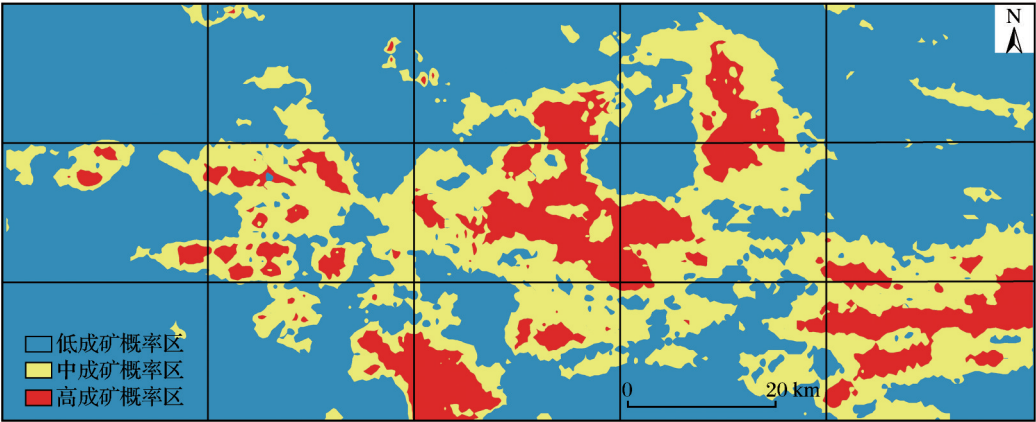


图 12 金矿成矿预测概率
Fig. 12 Metallogenic prediction probability of gold deposits

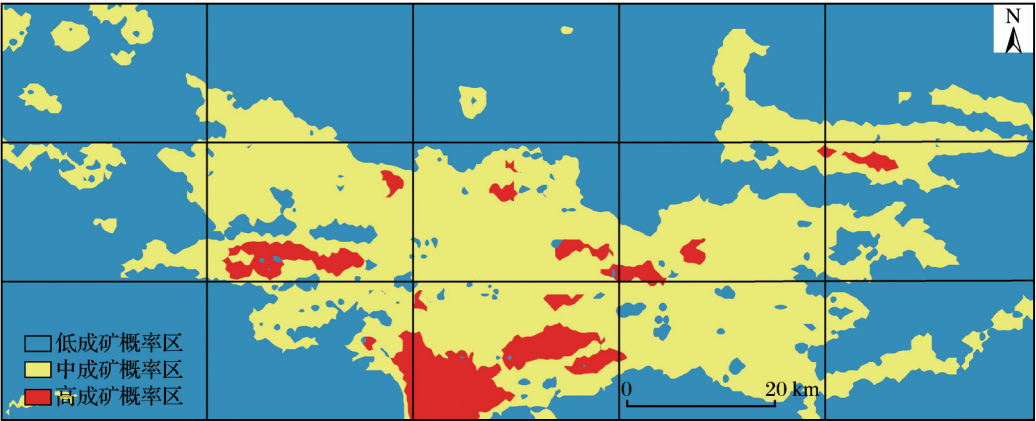


图 13 锑矿成矿预测概率
Fig. 13 Metallogenic prediction probability of antimony deposits

定成矿预测远景区和靶区(刘艳鹏等,2020)。依据中国地质调查局《固体矿产地质调查技术要求 (1:50 000)》(2023 版)分类标准,将预测靶区划分为 A 类、B 类和 C 类 3 类(表 4),不同类别代表不同

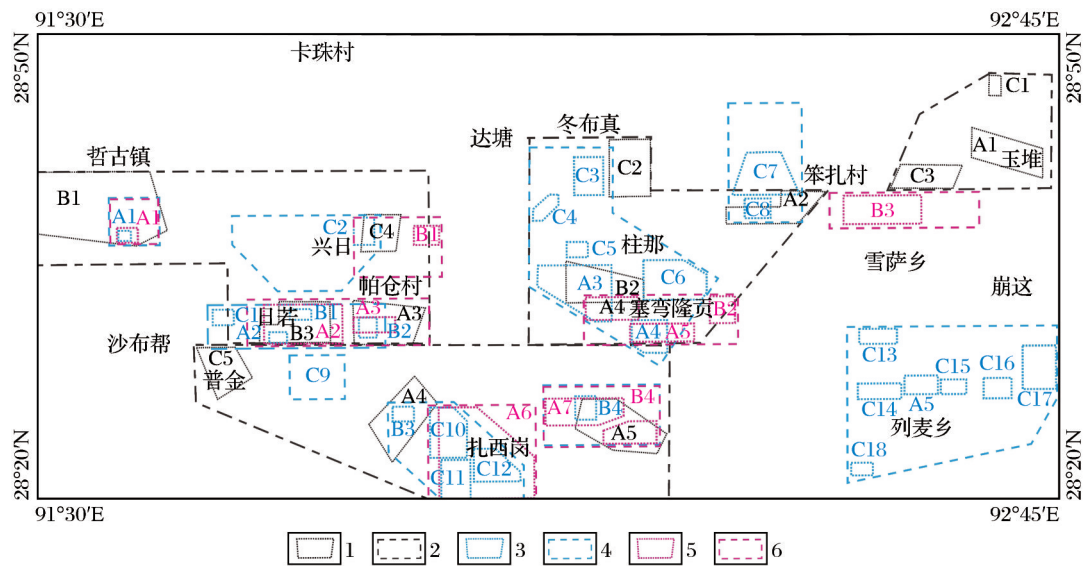
成矿预测潜力区域。根据模型预测结果及找矿靶区圈定标准,圈定金找矿靶区 27 个,其中 A 类金矿找矿靶区 5 个,B 类金矿找矿靶区 4 个,C 类金矿找矿靶区 18 个。圈定锑找矿靶区 11 个,其中 A 类锑找矿靶区 7 个,B 类锑找矿靶区 4 个。

利用地球化学数据开展成矿预测是传统预测

(张宇等,2022)和机器学习预测(李苍柏等,2020)的常用方法之一,在措美—隆子地区开展机器学习成矿预测的同时,基于地球化学水系沉积物数据运用分子区多背景圈定化探异常的方法进行传统成矿预测及靶区优选,2 种方法的预测结果对比如图 14 所示。

表 4 预测靶区圈定依据
Table 4 Delineation basis of prediction target area

靶区类别	圈定依据
A 类靶区	模型成矿预测值高,高值区域分布明显,区内已发现矿床,资源潜力大的地区
B 类靶区	模型成矿预测值较高,高值区域分布较明显,区内或附近已发现矿点,资源潜力较大的地区
C 类靶区	模型成矿预测值相对较高,有一定高值区域分布,区内或附近有矿化点,具有一定资源潜力的地区



1.地球化学方法圈定找矿预测区;2.地球化学方法圈定找矿靶区;3.机器学习方法圈定金找矿预测区;4.机器学习方法圈定金找矿靶区;5.机器学习方法圈定锑找矿预测区;6.机器学习方法圈定锑找矿靶区

图 14 基于 2 种方法的远景区及靶区圈定结果

Fig. 12 Delineation results of prospective area and target area based on two methods

2 种方法圈定的预测远景区在研究区中西部高度重叠,且部分靶区也有一定的重叠。机器学习方法未能预测出玉堆远景区,地球化学方法未能预测出雪萨乡北部远景区,表明 2 种方法均具有局限性。对通过地球化学方法得到的预测结果,以化探异常组合为依据,圈定范围会远大于实际成矿区,为后期异常查证带来不便。对通过机器学习算法得到的预测结果以边长 500 m 像元为最小单元,圈定范围较为精细。在远景区内优选成矿概率较大的区域作为找矿靶区,预测定位定量,从而更加精

确地反映局部区域的成矿潜力。

4 结论

(1)利用机器学习技术构建的 XGBoost 成矿预测模型具备适应成矿地质变量分布不均衡性的特性,同时兼具处理局部区域地质变量数据缺失的预测能力,能够为区域找矿预测提供一定指示。研究表明,XGBoost 模型在成矿预测中表现良好,机器学习方法适用于西藏山南金锑矿成矿预测,为藏南金

锑成矿带的区域找矿预测提供了一个新视角。

(2)*mIoU*值能够量化表征模型预测值与真实值在像元级别的重叠水平,反映预测模型的准确性。根据模型评估结果,金成矿预测模型*mIoU*值为0.8807,锑成矿预测模型*mIoU*值为0.9031,显示XGBoost模型在金锑矿成矿预测方面的精度均较高。

(3)在水系沉积物方面,Au、As、Sb和Pb是金矿模型相对重要性较高的变量,Sb、As、Hg、Bi和Ag是锑矿模型相对重要性较高的变量;在地层方面,三叠系与金矿相关性更强,侏罗系与锑矿相关性更强;在构造方面,金矿模型与近EW向断裂、穹窿相关,锑矿模型与近SN向断裂相关;在岩浆岩方面,金矿模型中生代中基性岩浆岩相对重要性略高于新生代中酸性岩浆岩,锑矿模型反之;在变质岩方面,变质结晶基底对2个模型的贡献均不高;金矿模型剩余重力异常和地形的相对重要性明显高于锑矿。

(4)根据成矿预测结果圈定金找矿靶区27个,锑找矿靶区11个。通过与地球化学方法预测结果进行对比可知,机器学习方法具有较好的预测效果,靶区圈定结果更精细,研究成果可为该地区金锑矿找矿突破工作部署提供参考。

参考文献 (References):

- Laffan S W, Nielsen O M, Silcock H, et al, 2005. Sparse Grids: a new predictive modelling method for the analysis of geographic data [J]. *International Journal of Geographical Information Science*, 19(3): 267-292.
- Li S, Chen J P, Liu C, et al, 2021. Mineral prospectivity prediction via convolutional neural networks based on geological big data [J]. *Journal of Earth Science*, 32(2): 327-347.
- Li S, Chen J P, Xiang J, 2020. Applications of deep convolutional neural networks in prospecting prediction based on two-dimensional geological big data [J]. *Neural Computing and Applications*, 32(7): 2037-2053.
- Sun T, Chen F, Zhong L, et al, 2019. GIS-based mineral prospectivity mapping using machine learning methods: a case study from Tongling ore district, eastern China [J]. *Ore Geology Reviews*, 109: 26-49.
- Xiong Y H, Zuo R G, Carranza E J M, 2018. Mapping mineral prospectivity through big data analytics and a deep learning algorithm [J]. *Ore Geology Reviews*, 102: 811-817.
- Yin J N, Li N, 2022. Ensemble learning models with a Bayesian

optimization algorithm for mineral prospectivity mapping [J]. *Ore Geology Reviews*, 145: 104916.

Zuo R G, Wang Z Y, 2020. Effects of random negative training samples on mineral prospectivity mapping [J]. *Natural Resources Research*, 29(6): 3443-3455.

陈建平, 李靖, 谢帅, 等, 2017. 中国地质大数据研究现状 [J]. *地质学刊*, 41(3): 353-366. Chen Jianping, Li Jing, Xie Shuai, et al, 2017. China geological big data research status [J]. *Journal of Geology*, 41(3): 353-366.

陈进, 毛先成, 刘占坤, 等, 2020. 基于随机森林算法的大尹格庄金矿床三维成矿预测 [J]. *大地构造与成矿学*, 44(2): 231-241. Chen Jin, Mao Xiancheng, Liu Zhankun, et al, 2020. Three-dimensional metallogenic prediction based on random forest classification algorithm for the Dayingezhuang gold deposit [J]. *Geotectonica et Metallogenia*, 44(2): 231-241.

成秋明, 2025. 面向人类智能与人工智能融合的矿产资源预测新范式 [J]. *地学前缘*, 32(4): 1-19. Cheng Qiuming, 2025. A new paradigm for mineral resource prediction based on human intelligence-artificial intelligence integration [J]. *Earth Science Frontiers*, 32(4): 1-19.

邓舟, 2020. 藏南马扎拉金锑矿矿床成因研究 [M]. 成都: 成都理工大学. Deng Zhou, 2020. Study on genesis of the Mazhala Au-Sb deposit in south Tibet [M]. Chengdu: Chengdu University of Technology.

董富权, 2012. 西藏古堆—隆子地区断裂构造分形特征及其地质意义 [J]. *黄金科学技术*, 20(6): 41-45. Dong Fuquan, 2012. Fractal characteristics of fractures and its geological significance in Gudui-Longzi region in southern Tibet [J]. *Gold Science and Technology*, 20(6): 41-45.

黄小东, 2011. 藏南拆离系江孜—隆子金—锑成矿带成矿规律与找矿方向研究 [M]. 成都: 成都理工大学. Huang Xiaodong, 2011. Study on metallogenic regularity and ore-prospecting direction of Gyantse-Lhunze gold-antimony metallogenic belt in the South Tibetan detachment system [M]. Chengdu: Chengdu University of Technology.

蒋海昆, 王锦红, 2023. 适用于机器学习的地震序列类型判定特征重要性讨论 [J]. *地震研究*, 46(2): 155-172. Jiang Haikun, Wang Jinhong, 2023. Discussion on the importance of the features for the judgement of earthquake sequence types applicable to machine learning [J]. *Journal of Seismological Research*, 46(2): 155-172.

李苍柏, 肖克炎, 李楠, 等, 2020. 支持向量机、随机森林和人工神经网络机器学习算法在地球化学异常信息提取中的对比研究 [J]. *地球学报*, 41(2): 309-319. Li Cangbai,

- Xiao Keyan, Li Nan, et al, 2020. A comparative study of support vector machine, random forest and artificial neural network machine learning algorithms in geochemical anomaly information extraction [J]. *Acta Geoscientica Sinica*, 41(2): 309-319.
- 刘艳鹏, 朱立新, 周永章, 2020. 大数据挖掘与智能预测找矿靶区实验研究: 卷积神经网络模型的应用 [J]. *大地构造与成矿学*, 44(2): 192-202.
- Liu Yanpeng, Zhu Lixin, Zhou Yongzhang, 2020. Experimental research on big data mining and intelligent prediction of prospecting target area: application of convolutional neural network model [J]. *Geotectonica et Metallogenia*, 44(2): 192-202.
- 娄渝明, 2023. 基于地质—地球化学信息的西藏雄村矿集区地表和深部成矿预测研究 [M]. 成都: 成都理工大学.
- Lou Yuming, 2023. Study on surface and deep metallogenic prediction based on geological-geochemical information in Xiongcu district, Tibet [M]. Chengdu: Chengdu University of Technology.
- 娄渝明, 康旭, 赖渊平, 等, 2025. 隐式建模和机器学习算法在西藏巨龙斑岩型铜钼矿床三维成矿预测中的应用研究 [J]. *地学前缘*, 32(5): 440-455.
- Lou Yuming, Kang Xu, Lai Yuanping, et al, 2025. Application of implicit modeling and machine learning algorithm to 3D metallogenic prediction of the Julong porphyry copper-molybdenum deposit, Xizang [J]. *Earth Science Frontiers*, 32(5): 440-455.
- 娄元林, 陈武, 杨桃, 2019. 西藏隆子县邦卓玛金矿床成矿模式与找矿模型 [J]. *地质通报*, 38(增刊1): 449-461.
- Lou Yuanlin, Chen Wu, Yang Tao, 2019. Metallogenic model and prospecting pattern of the Bangzhuoma gold deposit in Longzi County, Tibet [J]. *Geological Bulletin of China*, 38(Suppl.1): 449-461.
- 娄元林, 钱建利, 朱志平, 等, 2024. 物化遥综合找矿方法在西藏隆子县拉九地区的应用 [J]. *黄金科学技术*, 32(2): 241-257.
- Lou Yuanlin, Qian Jianli, Zhu Zhiping, et al, 2024. Application of integrated geophysical, geochemical and remote sensing prospecting methods in Lajiu area, Longzi County, Tibet [J]. *Gold Science and Technology*, 32(2): 241-257.
- 马瑶, 赵江南, 2021. 机器学习方法在矿产资源定量预测应用研究进展 [J]. *地质科技通报*, 40(1): 132-141.
- Ma Yao, Zhao Jiangnan, 2021. Advances in the application of machine learning methods in mineral prospectivity mapping [J]. *Bulletin of Geological Science and Technology*, 40(1): 132-141.
- 聂虹, 朱月琴, 常力恒, 等, 2018. 数据驱动下的矿产预测模型构建方法研究 [J]. *中国矿业*, 27(9): 82-87.
- Nie Hong, Zhu Yueqin, Chang Liheng, et al, 2018. Research on construction method of data-driven minerals prediction model [J]. *China Mining Magazine*, 27(9): 82-87.
- 欧阳渊, 刘洪, 李光明, 等, 2023. 基于随机森林算法的找矿预测: 以冈底斯成矿带西段斑岩—浅成低温热液型铜多金属矿为例 [J]. *中国地质*, 50(2): 303-330.
- Ouyang Yuan, Liu Hong, Li Guangming, et al, 2023. Mineral search prediction based on Random Forest algorithm: a case study on porphyry-epithermal copper polymetallic deposits in the western Gangdise metallogenic belt [J]. *Geology in China*, 50(2): 303-330.
- 袁峰, 李晓晖, 张明明, 等, 2014. 隐伏矿体三维综合信息成矿预测方法 [J]. *地质学报*, 88(4): 630-643.
- Yuan Feng, Li Xiaohui, Zhang Mingming, et al, 2014. Three dimension prospectivity modelling based on integrated geoinformation for prediction of buried orebodies [J]. *Acta Geologica Sinica*, 88(4): 630-643.
- 袁峰, 张明明, 李晓晖, 等, 2019. 成矿预测: 从二维到三维 [J]. *岩石学报*, 35(12): 3863-3874.
- Yuan Feng, Zhang Mingming, Li Xiaohui, et al, 2019. Prospectivity modeling: from two-dimension to three-dimension [J]. *Acta Petrologica Sinica*, 35(12): 3863-3874.
- 张必敏, 王学求, 周建, 等, 2025. 老挝铜资源成矿规律与基于机器学习的远景预测 [J]. *地学前缘*, 32(1): 61-77.
- Zhang Bimin, Wang Xueqiu, Zhou Jian, et al, 2025. Copper mineralization pattern and machine learning-based copper prospectivity prediction in Laos [J]. *Earth Science Frontiers*, 32(1): 61-77.
- 张刚阳, 2012. 藏南金锑多金属成矿带成矿模式与找矿前景研究 [M]. 武汉: 中国地质大学(武汉).
- Zhang Gangyang, 2012. Metallogenic model and prospecting potential in Southern Tibet Au-Sb polymetallic belt [M]. Wuhan: China University of Geosciences.
- 张宇, 魏俊浩, 石文杰, 等, 2022. 藏南康马县布主金(锑)矿土壤地球化学异常信息提取及成矿预测 [J]. *黄金科学技术*, 30(1): 1-18.
- Zhang Yu, Wei Junhao, Shi Wenjie, et al, 2022. Soil geochemical anomaly information extraction and metallogenic prediction of the Buzhu gold (antimony) deposit in the Kangma County, south Tibet [J]. *Gold Science and Technology*, 30(1): 1-18.
- 张振杰, 成秋明, 杨玠, 等, 2021. 机器学习与成矿预测: 以闽西南铁多金属矿预测为例 [J]. *地学前缘*, 28(3): 221-235.
- Zhang Zhenjie, Cheng Qiuming, Yang Jie, et al, 2021. Machine learning for mineral prospectivity: a case study of

- iron-polymetallic mineral prospectivity in southwestern Fujian[J]. *Earth Science Frontiers*, 28(3): 221-235.
- 周永章, 陈烁, 张旗, 等, 2018a. 大数据与数学地球科学研究进展: 大数据与数学地球科学专题代序[J]. *岩石学报*, 34(2): 255-263. Zhou Yongzhang, Chen Shuo, Zhang Qi, et al, 2018a. Advances and prospects of big data and mathematical geoscience [J]. *Acta Petrologica Sinica*, 34(2): 255-263.
- 周永章, 黎培兴, 王树功, 等, 2017. 矿床大数据及智能矿床模型研究背景与进展[J]. *矿物岩石地球化学通报*, 36(2): 327-331, 344. Zhou Yongzhang, Li Peixing, Wang Shugong, et al, 2017. Research progress on big data and intelligent modelling of mineral deposits [J]. *Bulletin of Mineralogy, Petrology and Geochemistry*, 36(2): 327-331, 344.
- 周永章, 王俊, 左仁广, 等, 2018b. 地质领域机器学习、深度学习及实现语言[J]. *岩石学报*, 34(11): 3173-3178. Zhou Yongzhang, Wang Jun, Zuo Renguang, et al, 2018b. Machine learning, deep learning and Python language in field of geology [J]. *Acta Petrologica Sinica*, 34(11): 3173-3178.
- 左仁广, 2019. 基于深度学习的深层次矿化信息挖掘与集成[J]. *矿物岩石地球化学通报*, 38(1): 53-60. Zuo Renguang, 2019. Deep learning-based mining and integration of deep-level mineralization information [J]. *Bulletin of Mineralogy, Petrology and Geochemistry*, 38(1): 53-60.
- 左仁广, 2021. 基于数据科学的矿产资源定量预测的理论与方法探索[J]. *地学前缘*, 28(3): 49-55. Zuo Renguang, 2021. Data science-based theory and method of quantitative prediction of mineral resources [J]. *Earth Science Frontiers*, 28(3): 49-55.
- 左仁广, 2025. 智能矿产预测的技术挑战与解决方案[J]. *中国科学: 地球科学*, 55(9): 3104-3119. Zuo Renguang, 2025. Key technology for intelligent mineral prospectivity mapping: challenges and solutions [J]. *Scientia Sinica (Terrae)*, 55(9): 3104-3119.

Mineralization Prediction of Shannan Gold and Antimony Deposits in Xizang Based on XGBoost Model

ZHANG Yiming^{1,2}, TANG Yulei³, HAN Cong², DENG Junfeng², HUANG Yong²

1. College of Earth and Planetary Sciences, Chengdu University of Technology, Chengdu 610059, Sichuan, China;

2. Civil-Military Integration Center of China Geological Survey, Chengdu 610036, Sichuan, China;

3. Geophysical Survey Center of China Geological Survey, Langfang 065000, Hebei, China

Abstract: The Shannan region in Xizang, situated within the eastern segment of the Tethyan Himalaya, is a significant area for the concentrated development of gold and antimony deposits in China. Over 60 gold and antimony deposits (occurrences) have been identified within the zone extending from Cuomei County to Longzi County. To enhance the identification of potential deposits (occurrences) and achieve breakthroughs in prospecting, machine learning methodologies can be employed to predict these potential deposits. This study focuses on the area covered by 15 standard 1:50 000 geological survey sheets in Cuomei County and Longzi County, Shannan, Xizang. It integrates geological, geochemical, geophysical, topographical, and other relevant datasets to construct two XGBoost prediction models for gold and antimony deposits in the region, achieving model accuracies of 0.8807 and 0.9031, respectively. A dataset for predicting mineralization of gold and antimony deposits in the study area has been developed using machine learning techniques. This dataset facilitates the identification of prospective mineralization zones and prospecting targets for gold and antimony deposits. The findings demonstrate that the XGBoost model exhibits satisfactory performance in predicting the mineralization of gold and antimony deposits in Shannan, Xizang. Based on these predictive outcomes, 27 gold prospecting targets and 11 antimony prospecting targets have been identified. These research findings offer valuable insights for guiding gold and antimony prospecting efforts in this region.

Key words: gold and antimony deposits; machine learning; metallogenic prediction; XGBoost model; Cuomei-Longzi County; Xizang