



③ 矿床统计预测的基本方法(上)

武汉地质学院勘探教研室数学地质组

矿床统计预测的方法很多,哈里斯(1973)曾将矿产资源定量评价方法分为三大类:地质多元统计法、主观概率法及空间法。阿格特伯格(1974)则概括为单变量统计法及多元统计法两类,而将空间法归为单变量统计法之中。根据已有资料,这两类方法在矿产预测中均有广泛应用。

本文着重介绍几种常用的单变量统计预测法,下一讲则介绍地质多元统计分析在矿产预测中的应用。

用于矿产资源评价及预测的基础概率和单变量统计法主要有下面几种:

1. 区域矿床空间分布统计模型的研究

此法属于区域战略性评价。即研究单位面积内矿床(点)个数的频率分布,国外有人提出等面积单元内矿床(点)个数的分布服从波松分布,后来有人认为大范围内矿床通常不是彼此独立产出而是趋于成群出现,因而用负二项分布拟合更为适当。通过查明区域矿点空间分布统计模型,可以估计为发现一个矿点所至少需要调查的单元数(面积),以及评价该区找矿勘探工作是否充分等。同时,可用此资料估计地质条件类似的勘探新区每单位面积的矿点数,一个大区域总面积内的矿点总数,每单位面积的平均价值等,以大致评价不同地区的含矿远景。在宁芜盆地中段,我们以 2×2 公里网格将研究区划分为96个等面积单元(每单元4公里²)对铁矿点空间分布统计模型进行研究,计算结果见表2。表2资料说明,在信度 $\alpha=0.05$ 时铁矿点不服从波松分布而服从负二项分布,但降低信度 $\alpha=0.01$ 时可服从波松分布。波松分布是矿点随机式分布的基本统计模型。在宁芜中段因铁矿点主要受NNE向卜

表 2

单元内矿点数	观测频数	理 论 频 数		数 学 模 型
		波 松 分 布	负二项分布	
0	65	60.0	64	波 松 分 布 $P(x) = e^{-\lambda} \left(\frac{\lambda^x}{x!} \right)$
1	20	28.2	23	
2	8	6.6	7	
3	3	1.0	2	
Σ	96	95.8	96	负二项分布 $P(x) = \frac{(k+x-1)!}{x!(k-1)!} \cdot \frac{P^x}{q^{k+x}}$
X^2		4.32	1.097	
检 验		$X_{0.05}^2 = 3.84$	$X_{0.05}^2 = 2.71$	
(表列值)		$X_{0.01}^2 = 6.64$	$X_{0.01}^2 = 3.84$	

塘一黄梅山主断裂带控制具有丛集性而呈方向性分布,故用负二项分布有更好的拟合。同时还可做出以下判断,据表2资料按负二项分布计算的不含矿点的单元数理论值为64,略低于实际观测的65,而含1个矿点的理论单元数为23又略高于实际观测的20个单元。这说明,该区还有发现新矿点的潜力,但不太大。另外,本区找到至少含1个铁矿点的单元之累积概率为0.333,故为在本区找到一个铁矿点,需对大约3个单元即12公里²范围进行调查。

2. 资源勘探中的经济效果和风险分析及统计决策的制定

它是采用统计决策理论定量地估计大区域内矿产资源的潜力,评价风险和勘探策略。如在勘探某金属矿床时决策的制定如图1。因储量(A)、品位(B)、埋深(C)三个事件 可视为相

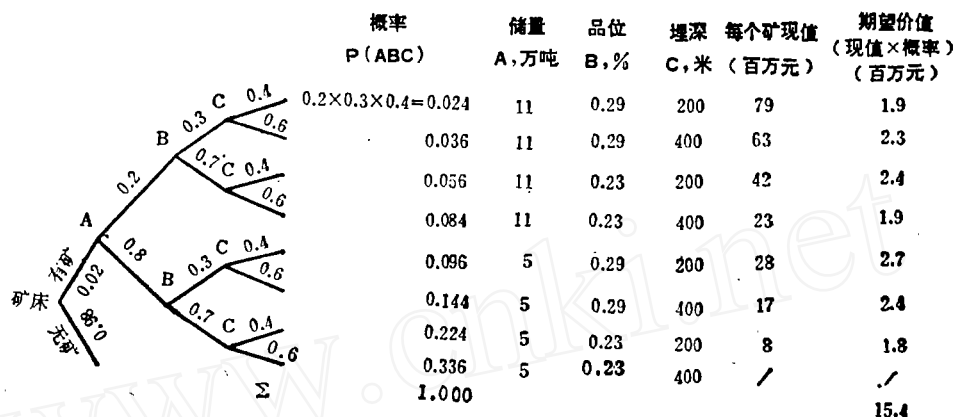


图1

互独立事件,因此可用 $P(ABC) = P(A) \cdot P(B) \cdot P(C)$ 来表示找到该种情况矿床的概率,图1中第一分支代表该区找到储量高品位高埋藏浅的矿床的概率即 $P(A_1, B_1, C_1) = 0.2 \times 0.3 \times 0.4 = 0.024$ 。设该区有矿的概率为0.02,无矿概率为0.98,则找到矿的期望价值(总期望价值×有矿概率)为 $15.4 \times 0.02 = 0.31$ 百万元。

决策意见:通过期望价值和勘探成本的对比,考虑开采是否合算,上例若勘探成本为0.18百万元,则 $0.31 - 0.18 = 0.13$ 百万元,故有盈利,可投入勘探。

当然,资本主义国家多数矿山是属于资本家所有,找矿和勘探的决策和布置都从营利出发。在我们社会主义制度下,矿山的开发首先决定于国民经济发展对矿产的需要程度及可能条件,但也应该考虑经济核算。

上述两方法的优点在于关键变量和不确定性相互作用中引入定量的概念,能比较粗略地估计矿产远景。缺点是纯粹从统计理论角度做出的判断,脱离了地质背景。此外,矿点空间分布统计模型还与所划分单元面积大小有关,划分不合理会影响统计结果。因而只能起参考作用。

3. 矿化、地质因素和标志的相关分析

通过分析矿化与各种控矿地质条件及找矿标志的相关联系,查明这些因素与标志最有利成矿的数值范围,从而找出多种有利因素及标志重合地段,即为要查明之成矿远景或预测地段。在这方面,秩相关分析是一种行之有效的方法,所谓“秩”,就是把一个变量的几个观测数据从小到大(或由大到小)排成序列,每个数据排的位次数就叫该数据的“秩”,然后计算变量的秩之间的相关系数。它实质是计算一个观测序列与一规则上升(或下降)的整数序列之间的相关性。现以宁芜盆地中段铁矿预测为例说明秩相关分析方法的应用。

根据该区地质构造条件及成矿理论分析,选择了该区有利控矿地质因素及找矿标志为:闪长玢岩出露面积、区内主断裂带与矿化的空间关系、围岩岩性组合及围岩蚀变组合、磁异

常值强度梯度及面积，并分别进行取值，如闪长玢岩出露面积比、单元中心至主断裂带距离、岩性及围岩蚀变组合特征值（相对熵值）以及磁异常指数及面积加权平均磁异常值。

在获得上述各项观测值后，再用秩相关分析查明各因素及标志与铁矿化的相关程度。秩相关系数 ρ 按下式求得：

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

式中 d 为所对比之两序列（矿化分级序列和某因素或标志分级序列）级序之差值，

n 为序列中划分之等级数。

ρ 值介于1与-1之间，代表不同的相关性质和程度。由于分级数一般较少，故以0.8或0.9做为相关显著性之准绳（若相邻级别 ρ 值差别不大，有可能是由于分级界限选择不当，这时需改变分级界限，再行计算）。

表3为闪长玢岩出露面积比与铁矿化程度的秩相关分析结果：

表 3

矿化分级（单元内矿点数）	单元内闪长玢岩出露面积比分级（%）					
	0	0~10	10~20	20~30	30~40	>40
0 ④**	① 60(92.23%)*	④ 4(6.15%)	② 1(1.53%)	④ 0	② 0	③④ 0
1 ③	③ 4(19.1%)	② 11(52.4%)	① 1(4.75%)	③ 2(9.5%)	① 2(9.5%)	② 1(4.75%)
2 ②	② 3(42.9%)	③ 2(28.6%)	④③ 0	② 1(14.3%)	③ 0	① 1(14.3%)
3 ①	④ 0	① 2(66.8%)	③④ 0	① 1(33.2%)	④ 0	④③ 0
ρ 值	-0.8	+0.8	-0.8, -0.6	+1	-0.8	-0.2, +0.4

* 前一数字为单元个数；

括弧内百分数为该单元数在该横行总单元数中的百分比，即在相同矿化级别条件下具不同闪长玢岩出露面积比的单元出现概率（频率估计）。

** 右上角圆圈内数字表示它在该列中的等级顺序号，即有不同矿点数（不同矿化级别）的单元在该标志区间内出现的概率大小顺序。而秩相关分析即考查该概率等级顺序与矿化等级顺序的相关性。

表3中闪长玢岩出露面积比为20~30%时， ρ 值为1，即与矿化程度呈完全正相关。面积比为0~10%时， ρ 值为+0.8，也有较好的正相关，且含有矿点的单元数最多(15个)。总的看来，岩体出露面积比<30%对铁矿化是有利的。单元中较小的岩体出露面积比可能代表岩体的局部突出部位，也可能是大岩体的边缘接触带或岩体被剥蚀深度不大，这些都应是寻找本类铁矿的有利地段。

还可以根据单元中心点所标定的观测值，编制各种因素或标志的等值线图，或编制各种因素及标志观测值的“上升序列曲线”，与各单元已知矿化情况相对照，找出最有利的成矿及找矿的数量范围。

例如，通过对宁芜中段各控矿因素及找矿标志统计分析，查明了各因素及标志最有利于形成或寻找铁矿的数量范围（可称为找矿的“统计标志”）如下：

- ①单元内闪长玢岩出露面积比 20~30%
- ②单元中心距北北东向主干断裂距离 0~1公里
- ③岩性组合特征值（Hr值） 40~60%
- ④围岩蚀变组合特征值（Hr值） >40%
- ⑤磁异常指数对数值 >1.5
- ⑥面积加权平均磁异常值 >1200Y

查明各因素及标志的有利数量范围后，再统计整个研究区内各单元中有利因素及标志的

总和数。在该研究区,目前已发现之工业铁矿床均位于具有三个或三个以上有利因素及标志的单元中。因此,有利因素总和数 ≥ 3 时为本区铁矿成矿远景区,凡有利因素总和 ≥ 3 而目前尚未发现铁矿床之单元,即做为铁矿之统计预测区。也可将各因素及标志观测值等值线图相重叠,多种有利数量范围叠加处即为成矿远景区。

综上所述,秩相关分析法可以具体查明各种地质因素与矿化关系。某因素与矿化具有简单线性相关时, ρ 值将随该因素级序之增加而递增或递减。这表明此种因素与矿化呈正相关或负相关。但当某因素与矿化不呈线性相关时,则可查明这种最有利于成矿的数量范围,同时,该方法是一种非参数性统计分析方法,不要求变量服从正态分布或其他特定分布,故对地质数据不作特殊要求(原始数据无需进行变换或处理)。此外,该法简便易行效果好,是主要优点。但这种统计预测基本上属于半定量性质,须对各因素分别进行考查,然后加以综合分析。在这当中,将各种因素和标志等量齐观,未分主次。再者,由于对原始数据不能做过细分级,分析结果的精度受一定影响,这是不足之处。应用时,须注意标志的合理分级,当相邻区间 ρ 值相差无几时,表明级距可以增大;当区间内观测值过少时, ρ 值的确定具有很大的随机性,结果只能供参考。

4. 成矿地质条件及找矿标志的信息量计算

某种地质因素、找矿标志对指导找矿的作用,可通过计算这些因素和标志所提供的找矿信息量进行评价。再根据各因素和标志在每个单元中所提供的信息量总和,评价找矿远景区并进行预测。具体做法分为两个步骤:

首先,计算各因素和标志对找矿的信息量(计算公式及其意义见第二讲)。

其次,计算在每个单元中各标志信息量的总和,其大小反映了相对找矿远景,单元信息总量越大越利于找矿。

信息量算法不仅可以考查单个变量与矿化的关系,还可以同时考查多个标志对找矿的信息量,其公式分两种情况:

当标志间有相依关系时,设 $A_1, A_2, \dots, A_n$ 为 n 个标志, X 为有矿床存在的事件,则

$$\begin{aligned} I_{A_1, A_2, \dots, A_n \rightarrow X} &= \log \frac{P(A_1, A_2, \dots, A_n | X)}{P(A_1, A_2, \dots, A_n)} \\ &= \log \frac{P(A_1 | X)}{P(A_1)} + \log \frac{P(A_2 | A_1 X)}{P(A_2 | A_1)} + \log \frac{P(A_3 | A_1 A_2 X)}{P(A_3 | A_1 A_2)} + \dots \\ &\quad \dots + \log \frac{P(A_n | A_1, A_2, \dots, A_{n-1} X)}{P(A_n | A_1 A_2 \dots A_{n-1})} \end{aligned}$$

当标志间相互独立时,则

$$\begin{aligned} I_{A_1, A_2, \dots, A_n \rightarrow X} &= \log \frac{P(A_1 | X)}{P(A_1)} + \log \frac{P(A_2 | A_1 X)}{P(A_2)} + \dots \\ &\quad \dots + \log \frac{P(A_n | A_1, A_2, \dots, A_{n-1} X)}{P(A_n)} \end{aligned}$$

应用这一方法,可以定量地评价各控矿地质条件和找矿标志的找矿意义,但在应用中需注意连续型变量状态的划分要尽可能反映与矿化联系的本质,划分不当就提供不了找矿信息。这要通过地质分析和多方案试验。好的状态划分应是使各不同状态信息量的差距有足够大,而且越大越好。

5. 主观概率法

由D.P.哈里斯1973年提出,他的基本出发点是认为地质因素和矿化之间只能是一种条件概率关系。因为我们不知道所有与成矿有关的地质因素,即使知道也不可能全部被观

测, 故漏失变量产生误差, 致使地质因素和矿化不是由单一解的确定性函数相联系。在实际工作中这种影响表现为: 一个地区表面上和含矿床地区相似, 但对 D 个地质因素 ($\hat{G}_1, \hat{G}_2, \dots, \hat{G}_D$) 进行考查时则可能是少矿或根本无矿。因而一个地区的矿床以有限的地质因素 $\hat{G}_i$ ($i = 1, 2, \dots, D$) 为条件时, 只能看成是一个随机变量:

$$P(M \leq \text{Mol} G_1, G_2, \dots, G_D) = F(\text{Mol} \hat{G}_1, \hat{G}_2, \dots, \hat{G}_D)$$

式中, Mo 、 M 为矿床值, 如矿床数 $1, 2, \dots$

G_i 为地质因素 ($i = 1, 2, \dots, D$);

$P(M \leq \text{Mol} G_i)$ 为在 G_i 条件下, 找到 $\leq \text{Mo}$ 矿床值的概率;

$F(\text{Mol} G_i)$ 为分布函数。

为了简化模型, 哈里斯规定: ①要预测的矿床数简化为两种情况, 即 $M \geq 1$ 及 $M = 0$, 分别表示有矿和无矿; ②要求首先建立一个地质概念模型, 它必须尽可能地反映矿床成因的地质理论。因而假定地质变量的存在和观测有一个自然顺序或符合于探矿过程中资料收集的次序。即当一个低序的变量 G_{i-1} 尚未被观测时, 变量 G_i 不可能观测。如对夕卡岩矿床, 有灰岩 G_{i-1} 、闪长岩 G_{i-1} , 才可能有接触带 G_i 。并规定都取二态变量 (有和无)。

计算方法是用贝西定理计算相应的条件概率, 结合找矿贝西定理定义如下:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})}$$

式中, A 为有矿, $\bar{A}$ 为无矿, A 与 $\bar{A}$ 构成一互不相容的完备事件组, B 为地质条件。故 $P(A|B)$ 为在 B 存在的条件下有矿 (A) 的概率。

据此, 矿产预测中的条件概率模型为:

$$P(M \geq 1 | G_1, G_2, \dots, G_D) =$$

$$= \frac{P(G_1, G_2, \dots, G_D | M \geq 1) \cdot P(M \geq 1)}{P(G_1, G_2, \dots, G_D | M \geq 1) \cdot P(M \geq 1) + P(G_1, G_2, \dots, G_D | M = 0) \cdot P(M = 0)}$$

其中 $P(G_1, G_2, \dots, G_D | M \geq 1) = P(G_D | G_{D-1}, G_{D-2}, \dots, G_1, M \geq 1) \cdot P(G_{D-1} | G_{D-2}, G_{D-3}, \dots, G_1, M \geq 1) \cdot \dots \cdot P(G_1 | M \geq 1)$

例: 假定考查灰岩、闪长岩和矿的关系, 地质变量为二态变量, 用 G_{ij} 表示, i 为变量号 ($i = 1$ 为灰岩, $i = 2$ 为闪长岩) j 为该变量存在与否 ($j = 1$ 为存在, $j = 2$ 为不存在), 用如下枝状图表示 (图 2)。设要考查无灰岩有闪长岩条件下, 即枝状图中的 G_{12} , G_{21} , 至

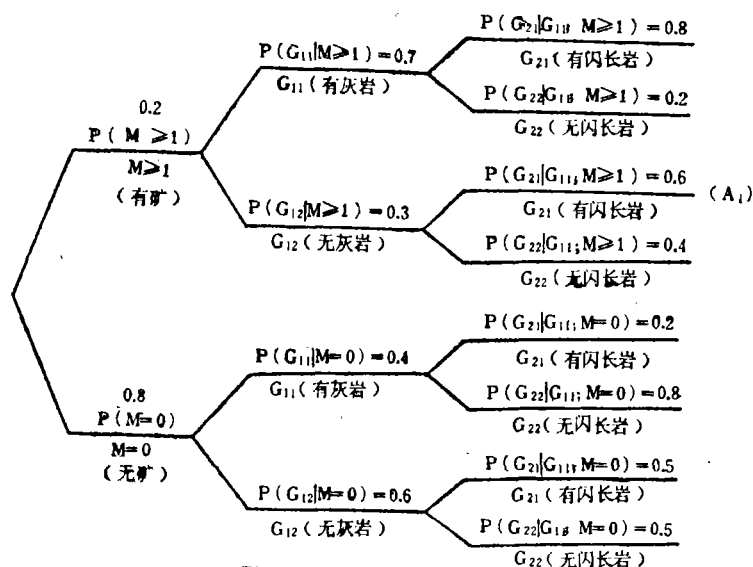


图 2

少发现一个矿床的概率 $P(M \geq 1 | G_{12}, G_{21})$ 如图 (A_1) 支:

$$\begin{aligned} P(M \geq 1 | G_{12}, G_{21}) &= \frac{P(G_{21} | G_{12}, M \geq 1) \cdot P(G_{12} | M \geq 1) \cdot P(M \geq 1)}{P(G_{21} | G_{12}, M \geq 1) \cdot P(G_{12} | M \geq 1) \cdot P(M \geq 1) + P(G_{12} | G_{12}, M = 0) \cdot P(G_{12} | M = 0) \cdot P(M = 0)} \\ &= \frac{0.6 \times 0.3 \times 0.2}{0.6 \times 0.3 \times 0.2 + 0.8 \times 0.6 \times 0.5} = 0.13 \end{aligned}$$

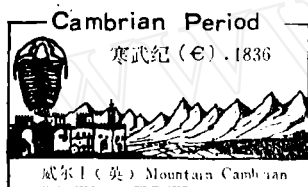
枝状图上各种条件的概率值, 哈里斯采用地质工作者的主观定量化意见。而一个地质工作者能反映全部矿床成因的不同观点是不可能的, 因此需要发动更多的地质工作者对此做出“响应”, 再计算各地质工作者所赋与概率(意见)的加权平均数。如

$$P(G_{11} | M \geq 1) = \frac{P_A(G_{11} | M \geq 1) \cdot P_A(M \geq 1) + P_B(G_{11} | M \geq 1) \cdot P_B(M \geq 1)}{P_A(M \geq 1) + P_B(M \geq 1)}$$

其中 P_A 、 P_B 分别为A、B两地质工作者所赋予的概率, 而权为边缘概率。

这就是哈里斯等“主观概率法”的实质。这个方法的优点, 哈里斯等人认为在于它能提供变量之间的相关信息, 而且可用于不同地质信息水平, 且随探矿过程中资料的不断增加可以进行概率复查。另外, 它不需要有一个“控制区”, 即为建立相应数学模型而选择的已知区。

应用上述模型先验概率 $P(M \geq 1)$ 及 $P(M = 0)$ 的确定是关键, 若其与实际情况差距太大可能造成较大失误。含矿与不含矿先验概率的确定, 有时也需依据类似控制区的资料进行统计。枝状图上各概率最好也根据已知区统计确定, 然后再对相似区进行预测。



“寒武”(Cambria)是英国西南部威尔士的古称。它是英国地质学家赛得威克(Sedgwick)1835~1836年在威尔士寒武山系描述了含三叶虫及其它化石的古老沉积岩系而创立的名称。以后, 在欧洲、美洲和亚洲一些国家相继发现有类似地层存在, 其界限为1888年在伦敦召开的第四次国际地质会议所肯定, “寒武系”则为第八次国际地质会议所承认。

寒武纪距今约5~6亿年。从那时开始, 海生动物大量繁盛。目前全世界寒武纪地层中发现的生物化石有2500多种, 以节肢动物中的三叶虫最多(占60%)、腕足动物次之(占30%)、其它生物, 如: 蠕虫类、软体动物、古杯海绵和古植物孢子等亦有发现。几乎所有无脊椎动物的主要门类, 都在寒武纪地层中存在。这说明地球上动物界的一次大发展。由于三叶虫种类多、演化快, 成为划分寒武纪地层的重要依据。

当时气候在地球不同地区有所差异, 北半球主要是较干燥或温暖, 而南半球则较温湿或寒冷, 甚至在一些

地区曾发生过冰川作用。我国自寒武纪开始, 海浸广泛, 主要来自北方和西南, 其范围比震旦纪有了明显的扩展, 但大部分地区处于浅海区。从世界范围看, 由于地壳发展不平衡, 因此, 在较活动的海槽区, 沉积物以碎屑岩为主, 并伴有强烈的海底火山喷发, 沉积厚度常达数千至数万米; 而在活动性小的海盆, 沉积物以碳酸盐岩和泥质或砂质岩类为主, 厚度只有一、二千米。

我国寒武纪地层发育完整。北方地区为浅海相沉积, 以碳酸盐岩类和泥质岩类为主。下统主要为石灰质砾岩、角砾状灰岩及砂页岩, 局部夹冰碛层及磷矿层; 中、上统以鲕状灰岩、竹叶状灰岩为标志, 以紫红色、绿色页岩为特征。标准剖面以鲁西泰安地区和燕山地区为代表。南方地区以长江三峡及川滇地区为代表, 其下统发育最好, 成为世界上研究和解决早寒武世问题的重要地区之一, 主要由灰岩、白云岩组成, 次为砂页岩, 常夹含磷炭质页岩、硅质岩层, 部份地区夹海底喷发火山岩; 中、上统主要是灰岩、白云质灰岩及白云岩。

关于寒武系的底界问题, 国内外仍在争论, 一种意见以最古三叶虫出现为界; 另一种意见以小壳动物群化石作为底界。国际上成立了一个界线委员会, 委员会主席采取举手表决, 后一种意见占了上风。近年来在三峡、峨眉山、滇东一带, 在寒武系下

统含三叶虫化石层位以下发现许多小壳动物化石(如单板、似牙形虫、节壳类等)50余属、100多种。故其底界普遍下移2~11米。目前我国寒武系划为8个统、9个组、21个化石带。

寒武纪大部分地区地壳运动比较缓和, 主要表现为较大地块的整体由缓慢沉降到隆起, 岩浆活动较微弱, 区域变质作用不普遍, 地层褶皱也较平缓, 断裂规模巨大, 但为数少。相当寒武纪的岩浆活动属加里东早期(同位素年龄值: 4.90~5.20亿年)。北方主要分布在祁连山、阿尔金山, 其次是昆仑山, 在辽东、大兴安岭、秦岭、贺兰山、内蒙白云鄂博、北山、天山等地亦见零星出露。岩石类型复杂, 从超基性、基性岩到花岗岩类均很发育。南方加里东早期侵入岩很少见。近年南岭发现混合花岗岩类, 有人认为属加里东早一中期。

世界范围寒武纪地层蕴藏的矿产主要有: 磷矿、锰矿、铁矿、铅土矿、岩盐、石膏、油页岩等。我国寒武系底部是重要含磷层位之一。值得引起特别注意的是, 近年来世界各地报道的寒武系底部的含矿黑色页岩(炭质页岩), 在我国亦广泛分布, 常含有Cu、Mo、Ni、Cd、U、V、K、P等多种有用元素。还伴生有锰、铁、铜、镍、钼、盐类、石煤、稀土、稀有元素等矿产; 内生金属矿产有铬、镍、铂、铜、铁、汞、铅锌等。(莫杰供稿)