Записки Ленинградского ордена Ленина, ордена Октябрьской Революции
и ордена Трудового Красного Знамени горного института им. Г.В.Плеханова,
1983, т. 98

УДК 550.8:519.2

К ОЦЕНКЕ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ ПОИСКОВЫХ МЕТОДОВ

А.Г.МАРЧЕНКО

В сообщениях об опыте применения того или иного поисково-оценочного метода нередко встречаются фразы типа: «результаты, полученные данным методом, подтвердились в стольки-то процентах случаев". Составить впечатление об эффективности использования метода на основании этих сведений невозможно, так как неясно, о каких рекомендациях, полученных данным методом, идет речь - положительных или отрицательных, насколько экономически оправдано применение метода и сколь существенная информация добавляется к тому, что было известно до его применения.

Для общей оценки эффективности применения метода при поисках по результатам проверки прогнозов можно использовать два несложных для расчетов пути, один из которых базируется на оценках вероятности допустить ту или иную ошибку, а второй - на основополагающих понятиях теории информации[x].

Рассмотрим их на примере [4], который выгодно отличается полнотой численных характеристик и потому удобен для расчетов. В статье П.Н.Селезнева и А.В.Палагина речь идет о результатах применения в Комсомольском оловорудном районе метода глубинного геохимического прогноза, основанного «на наличии корреляции между разницей в уровне содержаний фтора в околорудных породах в жильном выполнении, с одной стороны, и величиной нижележащих запасов олова, с другой. Метод был проверен на 24 зонах, по которым опробовано 72 пересечения. Рекомендации выданы по 57 пересечениям. Прогнозы проверены по 40 пересечениям. В 20 случаях установлена сходимость результатов: из них в 17 - отрицательная оценка более глубоких горизонтов, а в 3 положительная. В остальных 20 случаях установлены значительные расхождения прогнозных запасов с фактическими: не подтвердились 16 положительных рекомендаций, а в 4 случаях, наоборот, при отрицательных прогнозах разведкой на глубине установлены промышленные рудные тела" [4, с.124].

I. Оценки ошибок, надежности и экономического риска

Под событиями x_1 и x_2 будем понимать состояние объекта - соответственно рудный и безрудный, а под событиями y_1 и y_2 - соответственно по-

[x] Автор благодарен проф. Е.М.Квятковскому, который был инициатором рассмотрения данного вопроса и обсуждал с автором выводы, содержащиеся в статье.

ложительные и отрицательные рекомендации по результатам геохимического прогноза. Рассмотрим те объекты (т.е. пересечения), по которым прогнозы проверены: количество рудных объектов $n(x_1) = 7$; количество безрудных объектов $n(x_2) = 33$; общее количество объектов $n = 40$.

Считая частости тех или иных событий оценками соответствующих вероятностей, заполним следующую таблицу:

Рекомендации Y	Состояние объекта		$p(y_i)$
	Рудный x_1	Безрудный x_2	
Положительные y_1	3/40	16/40	19/40
Отрицательные y_2	4/40	17/40	21/40
$p(x_i)$	7/40	33/40	1

Учитывая, что в практике геохимических и геофизических поисков за нулевую гипотезу обычно принимается фоновое (безрудное) распределение, примем следующие обозначения ошибок [3]: α - ошибка I-го рода (вероятность того, что безрудный объект признается рудным); β - ошибка 2-го рода (вероятность того, что рудному объекту дана отрицательная оценка).

Вероятности правильно оценить объект как безрудный и как рудный соответственно обозначаются $(I - \alpha)$ и $(I - \beta)$. Оценки указанных вероятностей будут:

$$\alpha = \frac{n(x_2, y_1)}{n(x_2)} = \frac{16}{33} \approx 0{,}48; \qquad I - \alpha = 0{,}52;$$

$$\beta = \frac{n(x_1, y_2)}{n(x_1)} = \frac{4}{7} = 0{,}57; \qquad I - \beta = 0{,}43.$$

Таким образом, вероятность правильно распознать рудный объект оценивается в 43 %, а безрудный - в 52 %. Поскольку эти оценки близки к 50 %, эффективность применения метода следует считать неудовлетворительной как в плане распознавания рудных пересечений, так и в плане отбраковки безрудных.

Заметим, что 50-процентные оценки соответствуют наибольшей случайности выводов. Например, если бы при оценке каждого из пересечений мы бросали монетку по принципу „орел"-„решка", мы получили бы примерно такие же оценки ошибок.

Оценка полной вероятности ошибки:

$$P_{\alpha,\beta} = \alpha\, p(x_2) + \beta\, p(x_1) = \frac{16}{33} \cdot \frac{33}{40} + \frac{4}{7} \cdot \frac{7}{40} = 0{,}5$$

(это можно заметить сразу, так как из 40 проверенных пересечений в 20 установлена сходимость результатов).

Полная вероятность не ошибиться, называемая обычно надежностью, оценивается следующим образом:

$$\gamma = I - P_{\alpha,\beta} = 0,5.$$

Надежность изменяется от 0 до I, причем $\gamma = I$, если выводы всегда верны, $\gamma = 0,5$ соответствует наибольшей неопределенности выводов, а $\gamma = 0$ означает, что выводы всегда неверны.

Однако полная вероятность ошибки, так же как и полная надежность, имеет для геолога меньше смысла, чем оценки ошибок и надежностей двух типов в отдельности. Например, если бы мы все объекты без их изучения признавали безрудными, то в силу того, что безрудные объекты встречаются чаще рудных, мы получили бы высокую надежность выводов ($\alpha = 0$, $\beta = I$):

$$\gamma' = 1 - \alpha p(x_2) - \beta p(x_1) = I - I \cdot \frac{7}{40} = 0,825 > 0,5.$$

Но очевидно, что говорить о высокой надежности бездействия не имеет смысла. Бесспорно, что в зависимости от задач и условий проведения поисково-разведочных работ, мы придаем ошибкам α и β разный удельный вес: в одних случаях стремимся не пропустить рудный объект, в других – в первую очередь уменьшить число ошибочных положительных рекомендаций. Можно учесть так называемые цены ошибок I-го и 2-го рода C_α и C_β, а также величину риска R, измеряемые в конкретном денежном выражении [I, 3]:

$$R = C_\alpha \alpha p(x_2) + C_\beta \beta p(x_1).$$

Например, если разбуривание пересечения на глубину требует затрат около $I0^4$ руб., а обнаружение на глубине рудного объекта позволит при его разработке получить эффект в $I0^5$ руб., то риск применения метода составит ($C_\alpha = I0^4$ руб., $C_\beta = I0^5$ руб.):

$$R = I0^4 \cdot \frac{I6}{33} \cdot \frac{33}{40} + I0^5 \cdot \frac{4}{7} \cdot \frac{7}{40} = I4 \text{ тыс.руб.}$$

(цифры условные).

Если бы мы все объекты проверяли разбуриванием на глубину, как рудные, то риск составил бы ($\beta = 0$, $\alpha = I$):

$$R'' = I0^4 \cdot I \cdot \frac{33}{40} = 8,25 \text{ тыс.руб.}$$

Итак, если бы соотношение цен ошибок было таким, как мы приняли, то пришлось бы признать нецелесообразным применение метода, поскольку он увеличивает риск по сравнению с тривиальным случаем. Из конкурирующих методов прогноза целесообразно выбирать тот, который обладает наименьшим риском.

Существуют и другие статистико-игровые способы принятия решений при той или иной стратегии поисков [I]. Однако цель настоящей статьи – проил-

люстрировать возможность простых и в то же время выразительных оценок.

2. Оценка общей информативности метода.

Оценить общую информативность метода позволяет аппарат теории информации.

Исходная система – совокупность рудных и безрудных объектов. Наши сведения о ней обладают неопределенностью, так как до применения метода прогноза мы не можем сказать, встретится ли нам рудный или безрудный объект. Метод тем более информативен, чем больше он уменьшает исходную неопределенность.

Пусть $H(X)$ – энтропия исходной системы (величина, характеризующая степень ее неопределенности), $H(Y)$ – энтропия сообщения о системе, т.е. примененного метода прогноза.

Информация о системе X, содержащаяся в сообщении Y, выражается формулой [2]:

$$I_{Y-X} = H(X) + H(Y) - H(X,Y).$$

Если теперь отнести количество информации I_{Y-X} к исходной энтропии $H(X)$ и выразить эту величину в процентах, то получим характеристику информативности сообщения, т.е. примененного метода:

$$J' = \frac{I_{Y-X}}{H(X)} 100\ \%.$$

Эта величина может изменяться от нуля (в случае, когда сообщение не содержит информации об изучаемой системе) до 100 % (в случае, когда сообщение позволяет однозначно определить состояние системы).

Для расчета энтропии в выбранном примере подставим данные таблицы в соответствующие формулы [2]:

$$H(X) = -\sum_{i=1}^{2} p(x_i) \log_2 p(x_i) = 0{,}669;$$

$$H(Y) = -\sum_{i=1}^{2} p(y_i) \log_2 p(y_i) = 0{,}998;$$

$$H(X,Y) = -\sum_{i=1}^{2}\sum_{j=1}^{2} p(x_i, y_j) \log_2 p(x_i, y_j) = 1{,}666;$$

$$I_{Y-X} = 0{,}669 + 0{,}998 - 1{,}666 = 0{,}001;$$

$$J' = \frac{0{,}001}{0{,}669} \cdot 100\ \% = 0{,}15\ \%.$$

До сих пор мы не обсудили тот факт, что не всегда удается дать прогноз по всем изучаемым объектам. Так, в рассматриваемом примере опробовано

72 пересечения, а рекомендации даны по 57. Естественно поэтому ввести коэффициент применимости метода $k_{пр}$ ($0 \leqslant k_{пр} \leqslant 1$) и учитывать его при расчете общей информативности метода:

$$J = k_{пр} \frac{I_{Y-X}}{H(X)} \cdot 100\ \% \qquad (0 \leqslant J \leqslant 100\ \%).$$

В рассматриваемом примере

$$J = \frac{57}{72} \cdot 0{,}15 = 0{,}12\ \%.$$

ЛИТЕРАТУРА

1. Бугаец А.Н., Дуденко Л.Н. Математические методы при прогнозировании месторождений полезных ископаемых. Л., Недра, 1976.

2. Вентцель Е.С. Теория вероятностей. М., Наука, 1969.

3. Никитин А.А. Статистические методы выделения геофизических аномалий. М., Недра, 1979.

4. Селезнев П.Н., Палагин А.В. Геохимическая изученность, результаты и задачи геохимических поисков на восточном участке зоны БАМа. – В кн.: Повышение эффективности геохимических методов поисков в зоне БАМ. Новосибирск, Наука, 1981.