

Метод ландшафтно-ориентированной интерполяции геохимических полей

В.Б.Вагнер

23 января 1991

Одним из самых распространенных инструментов в геохимических исследованиях являются карты геохимических полей. Такие карты как правило строятся в виде карт изоконцентрат. Интерполяция изоконцентрат — процесс достаточно трудоемкий и требующий хорошей обеспеченности данными.

До сих пор не существовало алгоритма автоматизированной интерполяции, который бы позволял получать результаты, не уступающие ручной интерполяции при тех же исходных данных. Это связано с тем, что, проводя изолинии вручную, исследователь обычно учитывает ландшафтную дифференциацию территории, а автоматизированные методы интерполяции, достаточно широко сейчас применяющиеся, как правило не используют никаких данных кроме плановых координат опорных точек и значений интерполируемой величины в них.

Тем не менее эти методы с успехом применяются при некоторых видах геохимического картографирования. Их применимость зависит от характера имеющихся данных. Существует два основных подхода к сбору данных для карт изоконцентрат.

Первый — это *отбор проб по регулярной сети*. С последующей интерполяцией изолиний одним из широко распространенных методов: линейной интерполяцией, с использованием сплайнов, методом крайгинга [5] и др. Этот подход широко применяется в геолого-геохимической практике [6], но при ландшафтно-геохимических исследованиях он как правило не позволяет получить достаточно хорошие результаты, особенно в условиях территорий со сложной ландшафтной структурой.

Принципиально метод регулярной сети позволяет построить сколь угодно точную карту, но на практике для этого требуется количество, превосходящее возможности исследователя, так как при детализации сетки количество проб растет обратно пропорционально ее шагу.

Второй подход можно назвать *типологическим*. Точки опробования при этом выбираются таким образом, чтобы охарактеризовать каждый ландшафтный или почвенный выдел и при интерполяции изолиний их стараются проводить параллельно границам контуров. При ручном построении карт изолиний этот метод является основным.

Все другие методы построения карт геохимических полей — карты качественного фона, в легенду которых кроме чисто качественной информации



Рис. 1: Интерполяция в условиях двух контуров сложной формы

включены и данные об уровнях концентраций в соответствующих геохимических ландшафтах, картограммы и др. [3] рассчитаны на ту же методику отбора проб, что и типологическая интерполяция.

Н.Л.Беручашвили [2] считает, что картограммы отражают реальную ситуацию лучше, чем карты изолиний именно потому, что при их построении используется информация о ландшафтной структуре изучаемой территории, не учитываемая при прямой автоматизированной интерполяции.

По мнению автора, оптимальными являются карты изолиний, построенные по типологическому методу, так как, в отличие от картограмм они позволяют показать внутриландшафтные различия, и формируют в сознании исследователя образ плавно изменяющейся величины, более соответствующий реальности чем дискретные “ступеньки”.

Автором разработан метод автоматизированной типологической интерполяции данных, названный нами методом *ландшафтно-ориентированной интерполяции*. Сущность этого метода заключается в том, что существует ряд факторов, отраженных на ландшафтной карте и влияющих на распределение картографируемой величины. В основу метода положена интерполяционная формула Гаусса[1]. Известный недостаток этой формулы, заключающийся в том, что при расчете каждого интерполируемого значения используются все точки опробования, превращается в достоинство, так как геометрически удаленные точки могут оказаться качественно близкими и их влияние на интерполируемое значение может быть достаточно велико.

Значение интерполируемой величины в точке интерполяции определяется средним взвешенным по всем точкам наблюдений. Вес точки наблюдений зависит как от геометрического расстояния от нее до точки интерполяции, так и от расстояния в пространстве некоторой классификации, учитывающей факторы дифференциации. Как и во всех методах, где требуется одновременно учитывать качественные и количественные признаки, здесь возникает проблема соотношения весов тех и других. Кроме того большинство существующих ландшафтно-геохимических классификаций не рассчитано на использование в подобных целях и в них не существует методов количественного определения классификационного расстояния.

Наиболее удобным способом описания качественной информации, для кото-



Рис. 2: Интерполяция в условиях пятнистой ландшафтной структуры

рой проблема весовых коэффициентов имеет достаточно простое решение, является система признаков, каждый из которых может либо наблюдаться, либо не наблюдаться в соответствующей точке. При составлении почвенно-геохимических карт такими признаками могут например быть наличие карбонатного горизонта в почвах, вараженность подзолистого процесса и др.

Информация на ландшафтных, почвенных и им подобных картах представлена обычно не в виде таких признаков, а с использованием классификаций, имеющих более двух вариантов на каждом из уровней, но выразить эту информацию в форме “присутствует или отсутствует в данном контуре” (тот или иной тип растительности, почвообразовательный или геоморфологический процесс) — задача совершенно тривиальная.

Если же исходными данными для ландшафтно-ориентированной интерполяции служит карта, оцифрованная в формате какой-либо ГИС, то преобразование ее легенды в таблицу признаков для каждого контура достаточно легко можно выполнить автоматически, либо средствами самой ГИС, либо с помощью любой СУБД.

Для таких систем признаков существует возможность расчета весовых коэффициентов методом наименьших квадратов.

Интерполированное значение в этом случае вычисляется формуле:

$$C_x = \frac{\sum_{i=1}^n C_i (a_0 + \sum_{j=1}^n a_j \lambda_{xij}) \phi(x_x, y_x, x_i, y_i)}{\sum_{i=1}^n (a_0 + \sum_{j=1}^n a_j \lambda_{xij}) \phi(x_x, y_x, x_i, y_i)}$$

где $\phi(x_x, y_x, x_i, y_i)$ — весовая функция, зависящая от расстояния между точками, a_j — весовые коэффициенты признаков и a_0 — коэффициент фона, учитывающийся независимо от признаков, λ_{xij} принимает значение 1 если j -тый признак в данных точках имеет одинаковые значения и 0 в противном случае.

Весовые коэффициенты в том числе и коэффициент фона рассчитываются



Рис. 3: Интерполяция с учетом нескольких признаков

путем минимизации суммы

$$\sum_{i=1}^n (C_i^{\text{ИЗМ}} - C_i^{\text{ИНТ}})^2$$

где $C_i^{\text{ИЗМ}}$ — концентрации, измеренные в i -той точке опробования, а $C_i^{\text{ИНТ}}$ — концентрация, интерполированная в этой точке с использованием данных по всем точкам опробования кроме нее самой.

Особый случай возникает, если существуют два признака, значения которых совпадают во всех точках наблюдений, но не совпадают на всей территории на которую строится карта. Для этих признаков можно рассчитать только один общий коэффициент. Такие признаки мы будем называть *сцепленными*. При этом невозможно определить долю влияния каждого из признаков, что приводит к неопределенности при интерполяции. В этом случае оптимальным решением, видимо, является присвоение каждому из признаков равной доли совокупного коэффициента

Значения рассчитанных весовых коэффициентов имеют содержательный смысл, так как коэффициент тем больше, чем больше влияние данного признака на изменение картографируемой величины.

Проскольку значения всех признаков изменяются в одном и том же диапазоне $[0 \dots 1]$ для анализа относительной значимости признаков можно не вводить специальных статистик, подобных коэффициенту корреляции, а сравнивать сами значения весовых коэффициентов. Поэтому весовые коэффициенты, вычисляемые по данному методу, названы нами *коэффициентами влияния*.

Возможности метода демонстрируются рядом примеров (рис 1–4), иллюстрирующих типичные физико-географические ситуации. Для всех примеров приведена карта ландшафтных контуров (а), по которой определялись используемые признаки дифференциации с нанесенными на нее "точками опробования", в которых подписаны значения концентраций, карта изоконцентрат, построенная методом прямой интерполяции (б) и карта, построенная методом ландшафтно-ориентированной интерполяции (в).

Первая группа примеров иллюстрирует интерполяцию в условиях наличия

Таблица 1: Значения и коэффициенты влияния признаков в примере 3 (рис 3).

Контур	Признаки			Значения концентраций	
	Лес	Подзо- листые почвы	Коэф- фициент фона	в точках опробования	средние по контур
1	+	+	+	12.0 8.5	10.25
2	+	-	+	6.0 5.5	5.75
3	-	+	+		
4	-	-	+	3.0 4.0	3.5
Коэффициент влияния	1.279	1.365	0.558		

двух типов контуров сложной геометрической формы — долины в условиях однородного водораздела (рис 1) и системы заболоченных западин на плоской поверхности (рис 2). В этих примерах был использован только один признак — принадлежность точки к фоновому контуру (водораздельной поверхности). Вторая группа примеров показывает применение метода в условиях более сложной ландшафтной ситуации, когда для описания факторов дифференциации требуется использование нескольких признаков. Эти примеры демонстрируют возможность экстраполяции данных на основе качественной информации. На рис. 3 показана ситуация в которой все признаки независимы и для каждого рассчитан отдельный коэффициент влияния, а на рис 4 демонстрируется описанное выше явление сцепления признаков.

В долинно-водораздельной (рис 1) ситуации картографируемая территория в одном месте пересечена ландшафтно-геохимическим профилем с точками опробования.

Очевидно, что в этих условиях одного профиля будет недостаточно для того, чтобы отразить реальную структуру поля концентраций с помощью прямой интерполяции. Убедиться в этом позволяет рис. 1б. В данной ситуации потребуется сетка не менее чем из четырех профилей, для того чтобы адекватно изобразить поле концентраций этим методом. Метод ландшафтно-ориентированной интерполяции (рис 1в) позволяет выделить основную особенность поля концентраций — значимые различия между концентрациями в долине и концентрациями на плакоре, даже по одному профилю. При разработке данного примера мы предполагали, что вся дифференциация между долиной и плакором объясняется единственным использованным нами признаком. Если такой уверенности нет, при реальных работах рекомендуется использовать в аналогичной ситуации не менее 10–15 точек, чтобы с помощью анализа коэффициентов влияния убедиться, что значимость данного признака дифференциации действительно столь велика.

При использовании данного метода требуемое количество точек практически не зависит от сложности формы контуров, и определяется в основном количеством используемых признаков дифференциации и сложностью ландшафтной структуры территории. Поэтому в случае, если бы рассматриваемая в данном примере долина имела большее количество притоков, преимущества нашего ме-



Рис. 4: Сцепление признаков

тогда перед методом регулярной сетки были бы более заметными.

Рис 2 изображает не менее частую ситуацию — большое количество сильно контрастных (заболоченных или засоленных) западин на однородной поверхности, причем данные опробования имеются только для некоторых из них. При прямой интерполяции (рис 2б) пятна, не охваченные точками опробования на карте не отобразились, а общий фон получил значение среднее между западинами и основной поверхностью, в результате чего точки опробования на последней изображены как положительные аномалии. Метод ландшафтно-ориентированной интерполяции (рис 2в) и в этом случае дал результаты, для получения которых методом регулярной сетки потребовалось бы в несколько раз большее количество проб.

Следующие примеры (рис 3 и 4) показывают интерполяцию и экстраполяцию с использованием системы качественных признаков. В этих случаях опробованием охвачены не все качественно различные контура, но имеющаяся качественная информация позволяет делать более менее обоснованные предположения о значениях концентраций в этих контурах. Для анализа этих примеров важна информация о том, в какие признаки наблюдаются в каждом из контуров. Эти сведения, так же как и значения коэффициентов влияния, приведены в таблицах 1 и 2.

В примере с независимыми признаками (рис 3) имеются четыре контура и два признака (первый — наличие лесной растительности и второй — наличие подзолистых почв (табл 1)). Пробы отбирались

в обоих контурах с „лесом” (контура 1 и 2) и в контуре 4, в котором отсутствуют оба вышеуказанных признака.

Из значений „концентраций” ясно, что оба признака способствуют увеличению значений картографируемой величины, причем второй сильнее (в агроландшафте на серых лесных почвах значения 3.0 и 4.0, в лесом ландшафте на них же 6.0 и 5.5 и в лесном ландшафте на подзолистых почвах 8.0 и 12.0) если предположить, что влияние обоих признаков складывается и признаки независимы друг от друга, то после несложных арифметических операций со средними арифметическими по контурам получим, что первый признак „увеличивает” концентрацию на 2.25, а второй — на 4.5, и, следовательно среднее значение в кон-

Таблица 2: Значения и коэффициенты влияния признаков в примере 4 (рис 4)

Контур	Признаки			Коэффициент фона
	песчаные почвы	сосновый лес	гидроморфность	
1	+	+	+	+
2	+	+	-	+
3	-	-	-	+
4	-	-	+	+
5	+	-	+	+
Коэффициент влияния	7.052 ¹	7.052 ¹	9.328	4.982

^aСцепленные признаки

туре 3 должно составлять 8. Такая простая модель, конечно же неправомерна. Но качественный вывод, что значения концентраций в контуре 3 должны лежать между значениями в контурах 1 и 2 верен.

Аналогичный вывод можно сделать и из вычисляемых по нашему методу из значений коэффициентов влияния (табл 1), хотя здесь разница довольно невелика (1.37 против 1.28).

Карта, построенная методом ландшафтно-ориентированной интерполяции (рис 3в), выглядит в полном соответствии с этими предположениями. Изолиния 6.5 проходит практически по границе контура 2, (лес на серых лесных почвах) и захватывает большую часть контура 3, (агрландшафт на подзолистых), показывая, что почвенный фактор в данном случае влияет на поле концентраций сильнее чем фактор растительности, что мы и показали выше.

Ситуация в последнем примере (рис 3, табл 2) значительно более сложная. Здесь имеются три признака, два из которых (песчаный механический состав почв и сосновые леса) либо вместе присутствуют, либо оба отсутствуют во всех точках опробования. Это пример описанного выше явления *сцепления признаков*.

Поскольку разделить влияние признаков 1 и 2 не удалось, контур 5, в котором наблюдается песчаные почвы под другой растительностью, получил значения, промежуточные между контурами 1 и 4, между которыми он расположен, так как эти признаки получили равные коэффициенты влияния.

Строго говоря, это может быть и неверно. Если реальный коэффициент влияния признака 1 (следует помнить, что вычисляется только оценка методом наименьших квадратов суммы коэффициентов при признаках 1 и 2) много больше коэффициента влияния признака 2, то значения интерполируемой величины в этом контуре будет ближе к значениям в контуре 1 и наоборот.

Примеры, показанные на рисунках 1–4 демонстрируют возможности метода, но метод имеет и ряд ограничений.

Во-первых, если в совокупности выбранных признаков нет признаков, отражающих факторы дифференциации, которые не всегда известны а priori, каче-

ство карты изолиний будет не лучше, чем при прямой интерполяции.

Во-вторых, число точек опробования должно не менее чем в 2–3 раза превосходить число признаков, так как в противном случае метод неустойчив.

В-третьих, вычислительная сложность метода существенно выше, чем у метода прямой интерполяции, и его использование уже при нескольких десятках точек опробования требует заметных затрат машинного времени. Несколько компенсирует возросшую вычислительную сложность то, что вместе с построением карты изолиний вычисляются коэффициенты влияния, анализ которых позволяет сделать выводы, для получения которых другим путем потребовалось бы использование достаточно сложных статистических методов, например дисперсионного анализа.

Разработанный метод ландшафтно-ориентированной интерполяции позволяет использовать при построении карт изоконцентрат не только данные о результатах анализов, но и информацию, содержащуюся в тематических — ландшафтных, почвенных и других картах.

При правильном выборе системы признаков дифференциации он позволяет не только интерполировать, но и экстраполировать данные на природно-территориальные комплексы, по которым нет данных опробования.

Область применения описанного метода, видимо, не ограничивается геохимией ландшафтов. По-видимому нет причин, которые препятствовали бы его использованию, например при составлении микроклиматических карт или карт продуктивности растительности.

Список литературы

- [1] Агиштейн М.Э., Мигдал А.А. Как увидеть невидимое? - В кн. Эксперимент на дисплее. М.:Наука, 1989.
- [2] Беручашвили Н.Л. Этология ландшафта и картографирование состояний природной среды. Тбилиси, 1989
- [3] Гаврилова И.П. Ландшафтно-геохимическое картографирование. М.:МГУ, 1985
- [4] Гедымин А.В. О ландшафтно-геохимических картах. - В сб.: География почв и геохимия ландшафтов. М.:МГУ,1967
- [5] Кнорринг Л.Д., Деч В.А. Геологу о математике. М.,1989
- [6] Принципы и методика геохимических исследований при прогнозировании и поисках рудных месторождений. Л.:Недра, 1979.