

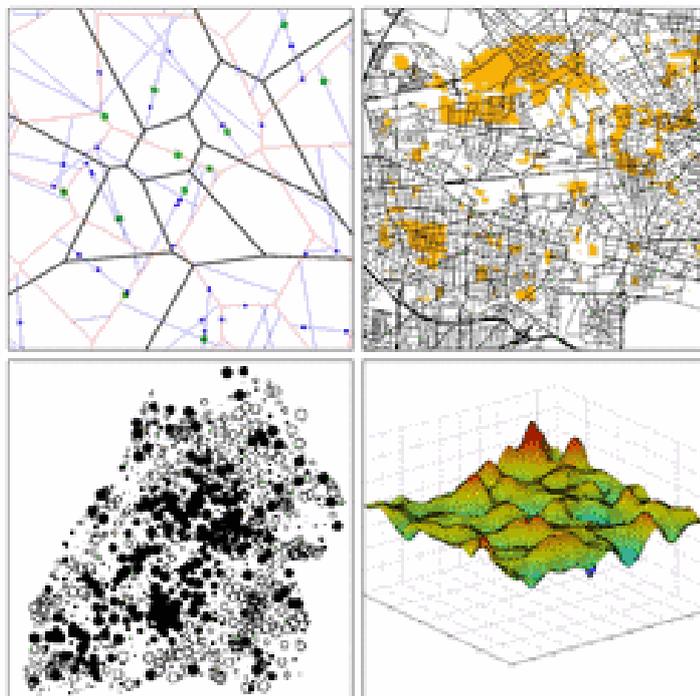


МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ – МСХА
имени К.А. ТИМИРЯЗЕВА
(ФГОУ ВПО РГАУ - МСХА имени К.А. Тимирязева)

ГЕОСТАТИСТИКА В ПОЧВОВЕДЕНИИ И ЭКОЛОГИИ

Мешалкина Ю.Л., Васенев И. И., Кузякова И.Ф.,
Романенков В.А.

Интерактивный курс



Москва 2010



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ – МСХА
имени К.А. ТИМИРЯЗЕВА
(ФГОУ ВПО РГАУ - МСХА имени К.А. Тимирязева)

**Мешалкина Ю.Л., Васенев И. И., Кузякова И.Ф.,
Романенков В.А.**

ГЕОСТАТИСТИКА В ПОЧВОВЕДЕНИИ И ЭКОЛОГИИ

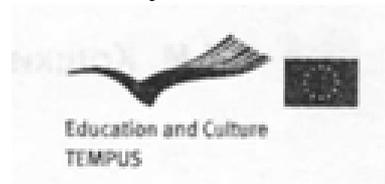
**Интерактивный курс
для подготовки магистров
по направлению «Агрохимия и агропочвоведение»
профилю «Агроэкология»
программе «Агроэкологический менеджмент и
инжиниринг»**

Москва 2010

УДК 519 : 631.58
ББК 26.1 + 41.41
А 26

Авторы: канд. с.-х. наук, доц. Ю. Л. Мешалкина, д-р биол. наук, проф. И. И. Васенев, канд. биол. наук Кузякова И.Ф., канд. биол. наук Романенков В.А.

*Издание осуществлено при поддержке
программы Темпус, грант Европейской
Комиссии CD_JER~26168-2005*



Рецензенты:

д-р биол. наук, проф., чл.-корр. РАН С.А. Шоба (МГУ имени М.В. Ломоносова),
д-р биол. наук, проф. М.А. Мазиров (РГАУ-МСХА имени К.А. Тимирязева)

Мешалкина Ю.Л., Васенев И. И., Кузякова И.Ф., Романенков В.А.

Геоинформационные системы в почвоведении и экологии. Интерактивный курс – М.: РГАУ-МСХА, 2010. 95 с.: илл. (Учебники и учебные пособия для студентов высших учебных заведений).

В учебном пособии в интерактивной форме представлены современные представления о теории и методах геостатистики и их применении в почвоведении и экологии для решения актуальных экологических и агроэкологических проблем. Основное внимание уделено вопросам построения семивариограмм и расчета их основных характеристик подбору моделей. Рассматривается понятие кригинга и дается описание некоторых его разновидностей. Описаны задачи, которые может решать геостатистика, при ее применении в экологии и почвоведении. Данное учебное пособие знакомит с возможностями использования методов геостатистики в почвоведении и экологии, а также содержит необходимый объем сведений, который позволит специалистам, работающим в области агроэкологии, экологии, почвоведения, агрохимии и земледелия, грамотно интерпретировать результаты геостатистических расчетов, выполняемых с помощью стандартных математических пакетов и специализированных статистических программ.

Для студентов агроэкологических и агрономических специальностей, обучающихся по программам магистратуры, а также аспирантов, преподавателей, научных сотрудников, работников сельского хозяйства и природоохранных учреждений, интересующихся вопросами математического и геоинформационного обеспечения их научной и повседневной деятельности.

Данный материал опубликован при поддержке Европейского Союза. Содержание публикации является предметом ответственности авторов и не отражает точку зрения Европейского Союза.

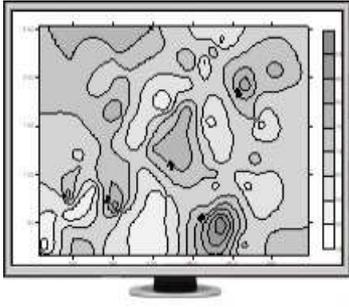
© Мешалкина Ю.Л., Васенев И. И., Кузякова И.Ф., Романенков В.А., 2010
© РГАУ-МСХА имени К.А. Тимирязева, 2010

Предисловие

В учебном пособии систематизированы современные представления о методах геостатистики, используемых в почвоведении и экологии. Основное внимание уделено вопросам построения семивариограмм и расчета их основных характеристик подбора моделей. Рассматривается понятие кригинга и дается описание некоторых его разновидностей. Описаны задачи, которые может решать геостатистика в области агроэкологии, почвоведения и земледелия. Содержится необходимый объем сведений, который позволит специалистам, работающим в области агроэкологии, экологии, почвоведения, агрохимии и земледелия, грамотно интерпретировать результаты геостатистических расчетов, выполняемых с помощью стандартных математических пакетов и специализированных статистических программ.

Материал представлен в интерактивной форме. Модули построены по общему плану. Теоретическому материалу предшествуют методические и организационные указания, ключевые слова, основная и дополнительная литература. В тексте обращается внимание на наиболее важные положения путем постановки вопросов и выделение материала восклицательным знаком. В конце каждого раздела имеются контрольные вопросы и задания, выполнение которых позволяет активизировать и закрепить полученные знания. Завершает учебное пособие толковый словарь основных терминов и понятий, а также англо-русский словарь профильной терминологии.

Наряду с отечественными и зарубежными публикациями последних лет, использованы результаты научной работы и преподавания на кафедре экологии РГАУ-МСХА имени К.А. Тимирязева, кафедре земледелия МГУ имени М.В. Ломоносова, в курсе геостатистики Университета Хоэнхайм (Германия) и лаборатории Географической сети опытов ВНИИА имени Д.Н. Прянишникова.



Модуль 1. Что такое геостатистика и ее отличие от классической статистики

Вы будете изучать

- Предпосылки использования геостатистики в почвоведении и экологии.
- Предмет и метод геостатистики и области ее применения
- Как описывается пространственная вариабельность

Цели модуля

- Дать определение геостатистике
- Показать отличия геостатистики от классической статистики
- Рассказать о составляющих пространственной вариабельности

После изучения модуля вы сможете

- Отличать геостатистику от классической статистики
- Знать определение геостатистики
- Иметь представление о пространственном варьировании.



Основная литература

1. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия, 2004. – 320 с.
2. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Дмитриев Е.А. Теоретические и методологические проблемы почвоведения. – М.: ГЕОС, 2001. – 374 с.
3. Журкин И. Г., Шайтура С. В. Геоинформационные системы. - М.: Кудиц-Пресс, 2008. – 272 с.
4. Козловский Ф.И. Теория и методы изучения почвенного покрова. – М.: ГЕОС, 2003 – 536 с.
5. Масштабные эффекты при исследовании почв. – М. Изд-во МГУ. 2001. – 400 с.
6. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
7. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85, no 2-3.



Ключевые слова:

Геостатистика, парадигма,
закономерная составляющая,
области применения геостатистики,
пространственная неоднородность,
истинная и наблюдаемая вариабельность.

ВВЕДЕНИЕ

Современная геостатистика – это быстро развивающаяся область прикладной статистики с большим набором методов, моделей для анализа, обработки и представления пространственной информации. Опора на геостатистический анализ позволяет значительно повысить уровень надежности и качество решений, принимаемых при использовании пространственно-распределенной информации. Спектр применения геостатистики – от традиционного использования в области геохимии ландшафта и оптимизации добычи природных ископаемых до современных приложений в социологии, экономике, экологии, медицине, финансах, эпидемиологии и т.д.

В последние годы геостатистические методы активно используются в почвоведении и экологии – при изучении структуры почвенного покрова, пространственной организации почвы и ее трансформации во времени, закономерностей пространственного варьирования агроэкологического качества почв, их химического состава и гидрофизических параметров, микроклиматических и микробиологических особенностей, уровня загрязнения и устойчивости к деградации и загрязнению.

1.1. ОПРЕДЕЛЕНИЕ ТЕРМИНА «ГЕОСТАТИСТИКА»

Последние 30 лет в почвоведении и экологии усилился интерес к изучению пространственно распределенных данных. Это связано с тем, что проблемы экологии вышли на первый план и перед человечеством встали задачи построения прогнозов в рамках Земного шара или континента, например: мирового урожая зерновых или глобального изменения климата.

Одновременно появились методические и технические условия для массового проведения детальных пространственных исследований:

- 1) Был разработан математический аппарат, обслуживающий работу с пространственными переменными;

2) Появились компьютеры с высоким быстродействием, позволяющие строить высокочувствительные геоинформационные системы;

3) Созданы широкодоступные системы глобального позиционирования (GPS);

4) Разработаны новые приборы и методы, позволяющие проводить многоплановый и разномасштабный мониторинг земной поверхности.

Таким образом, появилась аналитическая база, позволяющая получать большие массивы количественных данных, имеющих точные привязки на местности. Это привело к тому, что стали широко развиваться направления науки, связанные с анализом и интерпретацией пространственных данных.

В частности, в почвоведении оформилась новая область, именуемая педометрикой. Это слово образовано от греческих корней «pedos» (почва) и «metron» (измерение). Роберт Вебстер (Webster, 1994) предложил определить педометрику как "почвоведение в условиях неопределенности". В почвоведении геостатистика рассматривается как составная часть педометрики.



Что такое геостатистика?

Под термином «**геостатистика**» или «**пространственная статистика**» понимают методы статистического анализа пространственно распределенных данных. Часть понятий в геостатистику привнесено из теории аппроксимации. Именно поэтому добавлением "ГЕО" эти методы отделены от статистических .



Геостатистика включает методы, позволяющие проводить интерполяцию данных, рассчитывать оптимальные схемы пробоотбора, распространять точечные модели массопереноса на некоторую территорию и так далее. Отличие геостатистики от обычной статистики состоит в том, что в ней учитываются географические координаты точек опробования.



Что характерно для классической статистики?

При анализе результатов методами классической статистики исходят из отсутствия пространственной зависимости между значениями изучаемых показателей в точках опробования. Все данные рассматриваются как равноценные и взаимно независимые. Базовым понятием статистики является постулат, что статистическое испытание может быть повторено (хотя бы теоретически) бесконечное число раз.

Для уменьшения влияния природного варьирования на среднее измеряемых свойств применяют ряд известных приемов:

- ✓ увеличение объема выборки,
- ✓ отбор смешанных образцов,
- ✓ анализ средней пробы,
- ✓ закладку эксперимента в нескольких повторностях;
- ✓ рендомизированное расположение опытных делянок,
- ✓ организацию нескольких повторений опыта.

Несмотря на это, наличие неучтенных закономерностей варьирования почвенного покрова приводит подчас к противоречивым результатам полевых опытов даже на почвах одного типа.



В чем отличие геостатистической теории?



В отличие от классической статистики парадигма геостатистических теоретических построений рассматривает почвенный покров как единственную реализацию некоторой случайной функции. Эта случайная функция может быть изучена в силу наличия внутренних связей.

Пространственное варьирование при этом может быть разложено на составляющие, одна из которых зависит только от расстояния между точками

и не зависит от их координат. Существование закономерной составляющей природного варьирования почвенных свойств подтверждается наличием структурной организации почвенного покрова.

Количественные показатели почвенных свойств, как и большинства природных сред, в пределах небольших территорий часто ближе по значениям, чем для сравнительно удаленных друг от друга участков. Это выявляется как при мелкомасштабном, так и при крупномасштабном картировании почвенных свойств и отображается на картограммах в наличии ареалов, характеризующихся близкими значениями показателей.

Представление о неоднородности как пространственной организации объекта давно применяется в геологии (Матерон, 1968; Рац, 1973; Родионов, 1965). В российском почвоведении впервые Ф. И. Козловский (1970) предложил пространственное варьирование почвенных свойств подразделить на закономерную составляющую, обладающую определенной пространственной структурой и случайную составляющую.

В соответствии с этим подходом, однородность объекта по какому-либо признаку была определена (Розанов, 1983) как независимость признака от пространственных координат в пределах изучаемой области.

Применению геостатистики в почвоведении посвящены классическая работа Роберта Вебстера и Маргарет Оливер "Статистические методы в почвоведении и оценке природных ресурсов" (Webster, Oliver, 1990), работа П. Гуверта "Геостатистика в почвоведении: положение дел и перспективы" (Goovaerts, 1999), а также выпуски журнала *Geoderma* № 62 за 1994 год и № 97 за 2000 год.

В последнее время и в российском почвоведении стали все чаще использоваться методы геостатистики для решения задач, связанных с описанием пространственной изменчивости почвенных свойств (Гумматов и др., 1992; Дубин, 1990; Романенков, Кузякова, 2000; В.П. Самсоновой с соавт. (1999), В.А. Романенкова (1997), И.Ф. Кузяковой с соавторами (1997 и 2001).

1.2. ПОНЯТИЕ О ПРОСТРАНСТВЕННОЙ ВАРИАБЕЛЬНОСТИ



В чем состоит описание пространственной variability почв?

Вопрос о варьировании почвенных свойств в пространстве занимает особое место в почвоведении и экологии и вызывает бурные дискуссии. Суммируя предложения исследователей, участок можно назвать однородным, если он покрыт почвой, относящейся к одному и тому же типу, подтипу, роду, виду и разновидности почв (Козловский, Роде, 1976), свойства которой одинаково статистически распределены (Корнблум, 1975), а также их показатели не зависят от пространственных координат (Рац, 1973; Розанов, 1983).

Если эти условия не соблюдаются, участок будет неоднородным. Однако, из-за аналитических ошибок или недостаточного количества исходной информации, закономерности почвенной неоднородности могут быть потеряны (Матерон, 1968).

Пространственная неоднородность почвенных свойств и условий среды является основой эко- и биоразнообразия. Было показано, что пространственная неоднородность физико-химических свойств и почвенно-физических процессов является фундаментальной необходимостью образования структуры и осуществления почвой ее функций на каждом иерархическом уровне (Шеин, Милановский, 2001).

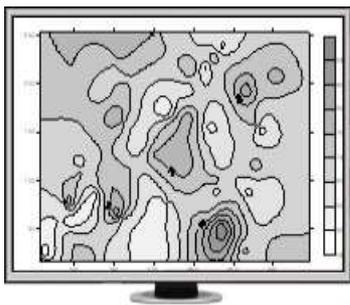
Матероном (1968) были сформулированы положения об «истинной» и наблюдаемой пространственной variability. Истинной variability соответствует почвенная неоднородность, а наблюдаемой – результат ее изучения. Е. А. Дмитриев (1993) предложил различать понятия «неоднородность» как собственное свойство почвы и «пространственная variability» как результат взаимодействия человека и природного объекта – почвы.

Таким образом, исследование и описание пространственной вариабельности свойств почвы представляет собой переход от объективного, существующего отдельно от исследователя явления к его модели. Модели только в некоторой степени отражают характеристики явления, поскольку модель отклоняется от истины, что обусловлено ошибками измерений и ограниченностью числа точек наблюдения, по которым оцениваются параметры модели, а также взглядами и научной школой исследователя. Из-за этого практически невозможно получить полное представление об истинной неоднородности почв. Изучение закономерностей из-за низкой точности и фрагментарности знаний может дать неотличимые от случайных результаты. В этом случае мы получаем модель почвы как случайно неоднородного объекта.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Каковы предпосылки развития геостатистики в экологии и почвоведения?**
- 2. Какие три основных постулата характерны для классической статистики?**
- 3. Для исследования каких задач наиболее активно используются методы классической статистики?**
- 4. Что такое геостатистика?**
- 5. Какая теория лежит в основе геостатистики? Кто ее разработал?**
- 6. Что такое пространственная переменная?**
- 7. Что может выступать в качестве аргумента для пространственной переменной?**



Модуль 2. Гипотетические модели пространственного варьирования почвенных тел

Вы будете изучать

- Концептуальные модели пространственного варьирования
- Представление пространственного варьирования методом хороплет
- Тренды
- Пространственную автокорреляцию

Цели модуля

- Дать представление о моделях, которые используются для описания пространственного варьирования
- Показать преимущества и недостатки метода хороплет
- Обсудить отличия трендов разного порядка
- Понять новое понятие – пространственную автокорреляцию

После изучения модуля Вы сможете

- Узнать основные закономерности моделирования пространственного варьирования
- Знать отличия между глобальным трендом и трендов других порядков
- Подходить обоснованно к выбору модели для пространственных данных в ГИС



Основная литература

1. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия, 2004. – 320 с.

2. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Дмитриев Е.А. Теоретические и методологические проблемы почвоведения. – М.: ГЕОС, 2001. – 374 с.
3. Журкин И. Г., Шайтура С. В. Геоинформационные системы. - М.: Кудиц-Пресс, 2008. – 272 с.
4. Козловский Ф.И. Теория и методы изучения почвенного покрова. – М.: ГЕОС, 2003 – 536 с.
5. Масштабные эффекты при исследовании почв. – М. Изд-во МГУ. 2001. – 400 с.
6. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
7. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85, no 2-3.



Ключевые слова

Модель, хороплеты, тренды,
пространственная автокорреляция,
отсутствие зависимостей, случайная составляющая,
глобальный тренд

ВВЕДЕНИЕ

Модель (в науке) – это описание объекта (предмета, процесса или явления) на каком-либо формализованном языке, составленное с целью изучения его свойств. Такое описание особенно полезно в случаях, когда исследование самого объекта затруднено или физически невозможно.

! □ Процесс построения и исследования модели называется моделированием. Таким образом, модель выступает как своеобразный инструмент для познания, который исследователь ставит между собой и объектом и с помощью которого изучает интересующий его объект.

2.1. ХОРОПЛЕТЫ

! □ Хороплеты – это однородные, согласно некоторым принятым критериям, участки территории. Эта самая простая модель пространственного изменения почвенных и экологических переменных.

В почвоведении и экологии при картографировании обычно предполагается, что естественные почвенные или растительные ареалы очерчены четкими границами, а внутри они являются однородными по своим свойствам. При этом считается, что изменение почвенных свойств происходит на границах скачкообразно (рис. 1 А и Б).

Значение почвенного признака (случайной переменной) Z в точке x задается по формуле:

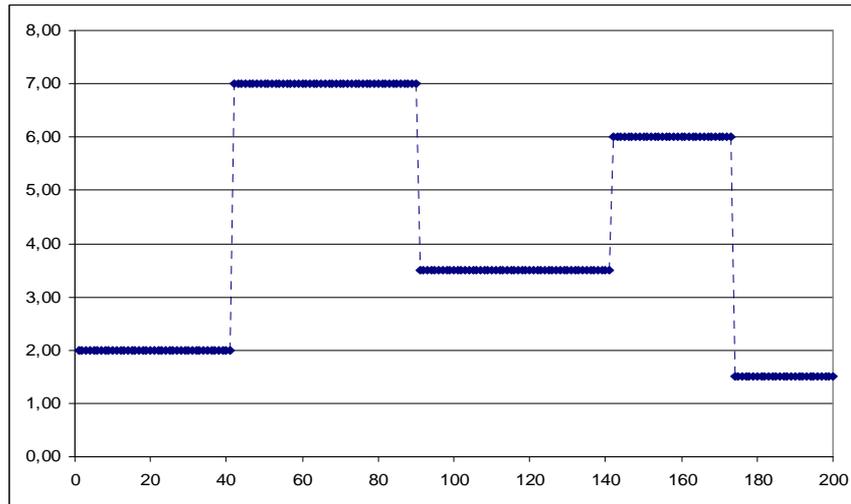
$$Z = b_0 + b_j + \varepsilon ,$$

где b_0 - общее среднее почвенного признака для всей изучаемой территории;

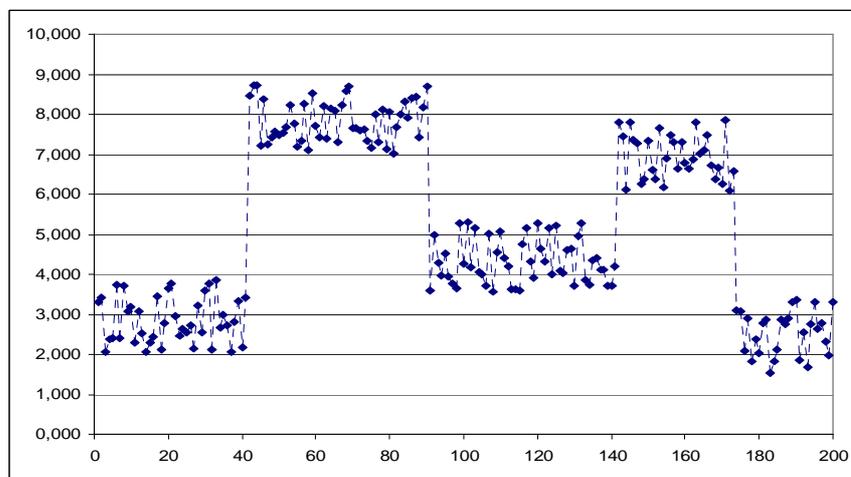
b_j - разница между средним внутри области j , в которой расположена точка x , и общим средним; сумма всех b_j равно нулю;

ε - независимая, нормально распределенная случайная ошибка, имеющая нулевое среднее.

А.



Б.



В.

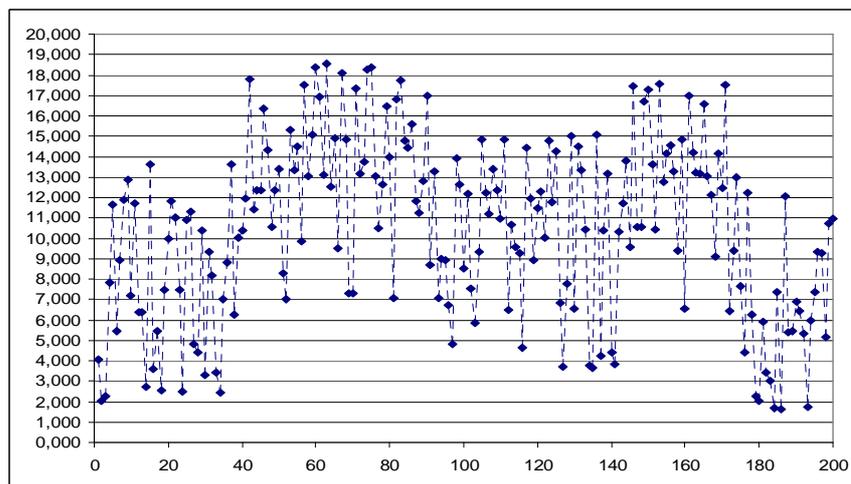


Рис. 1. Хороплеты:

- А. -идеальное варьирование между единицами с четкими границами;**
- Б. -варьирование внутри классификационных единиц меньше, чем между классификационными единицами;**
- В. -варьирование на небольших расстояниях перекрывает все другие изменения**

К сожалению, данная модель пространственного изменения почвенных переменных не всегда хорошо работает, поскольку в природе изменения редко происходят на четко различимых границах. Кроме того, границы, проведенные по каким-либо критериям, могут неверно отражать изменения других переменных. А пересмотр принципов выделения однородных ареалов может дать совершенно другие карты.



В чем состоит преимущество метода хороплет?

Преимущество такой модели состоит в том, что она позволяет генерализовать большое количество точечных наблюдений. До того, как компьютеры стали широко доступны для обработки большого числа комплексных пространственных данных, классификация и картирование хороплет было единственным подходом к проблеме картирования пространственного изменения комплексных экологических явлений (рис. 2).

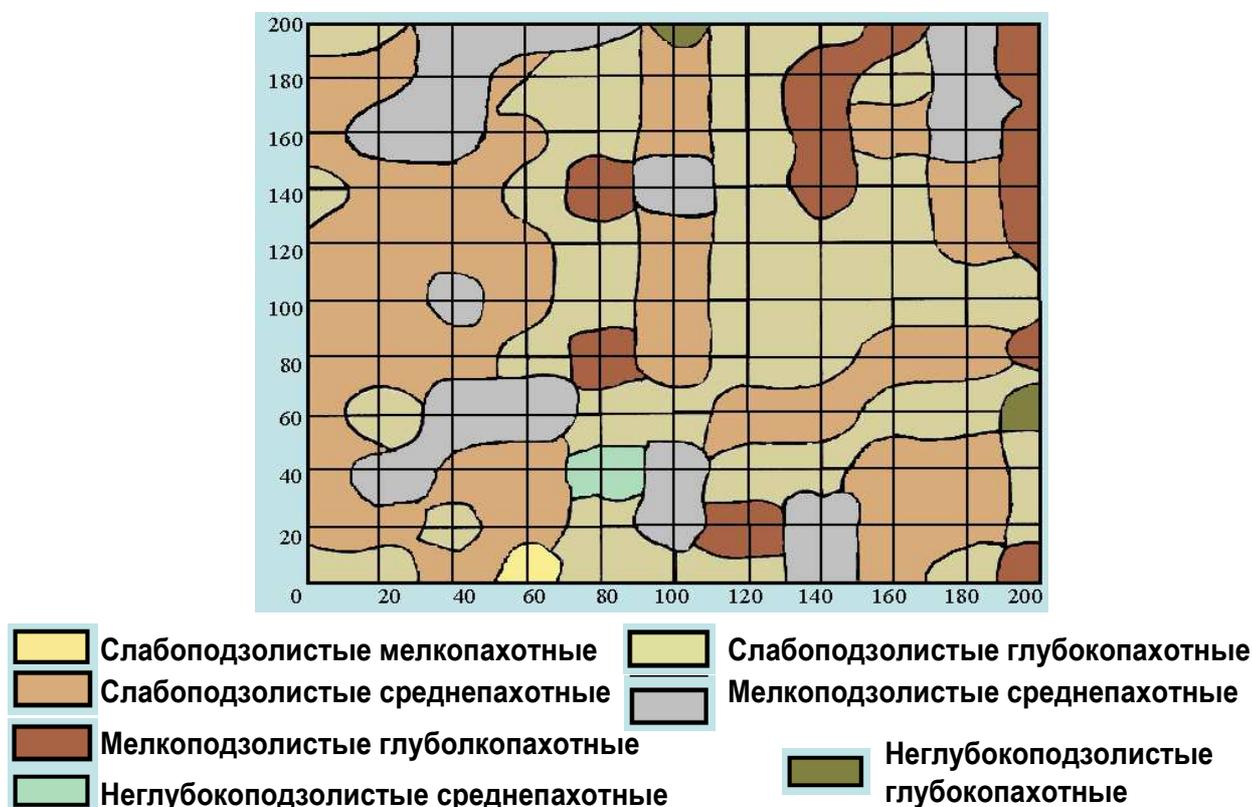


Рис. 2. Пример почвенной карты участка площадью 4 га – реализации метода хороплет.

2.2. ТРЕНДЫ



Что собой представляют тренды?



Тренды - это пространственные модели, описывающие постепенные изменения почвенного свойства. Трендовые поверхности, в основном, используются для описания постепенных крупномасштабных изменений. Полученные в результате линия (рис. 3), поверхность или объем описывает крупномасштабное пространственное варьирование Z (детерминистическая часть модели).

Случайная (мелкомасштабная) компонента представлена отклонением от регрессии (случайная составляющая модели):

$$Z = b_0 + b_1 X + \varepsilon$$

где ошибка ε считается независимой и нормально распределенной;

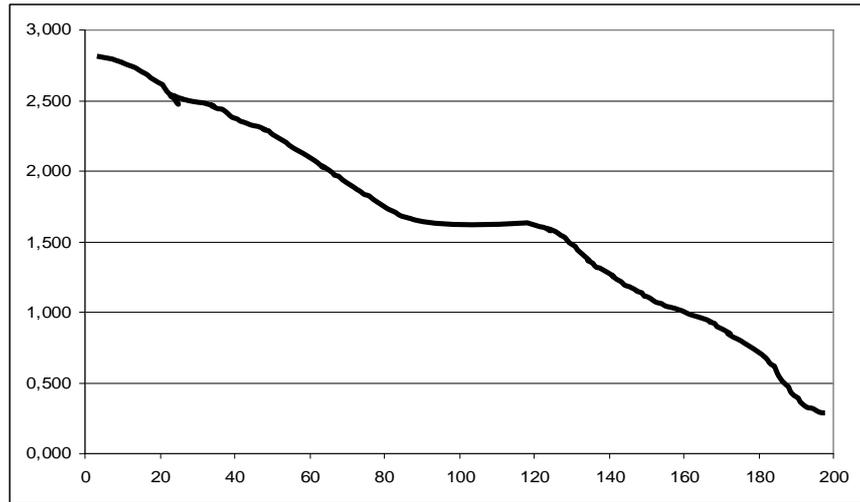
b_0 и b_1 этой модели могут быть оценены методом наименьших квадратов.

Если Z не является линейно возрастающей или убывающей, то изменения могут быть описаны уравнения более высокого порядка, например, второго. На рис. 4 представлены тренды 1-го, 2-го и 3-го порядка. Путем увеличения порядка уравнения можно точно описать любую сложную кривую.

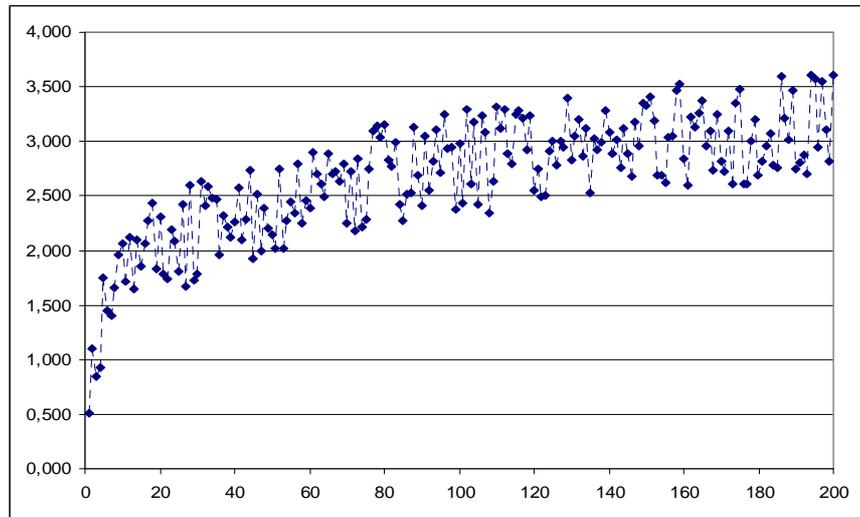
Подбор поверхности тренда не имеет большого смысла до тех пор, пока тренду не может быть дано физическое объяснение. Когда число данных невелико, экстремальные значения удаленных точек данных могут оказывать чрезмерно большое влияние и, в результате, давать неверные оценки, и существенно изменить поверхность.

Модели регрессии предполагают, что все отклонения от регрессии (остатки) нормально распределены и независимы. На практике это не всегда так, что может быть видно при картировании остатков, которые часто обнаруживают пространственную связь, образуя кластеры. Поверхности очень чувствительны к краевым эффектам.

А



Б.



В.

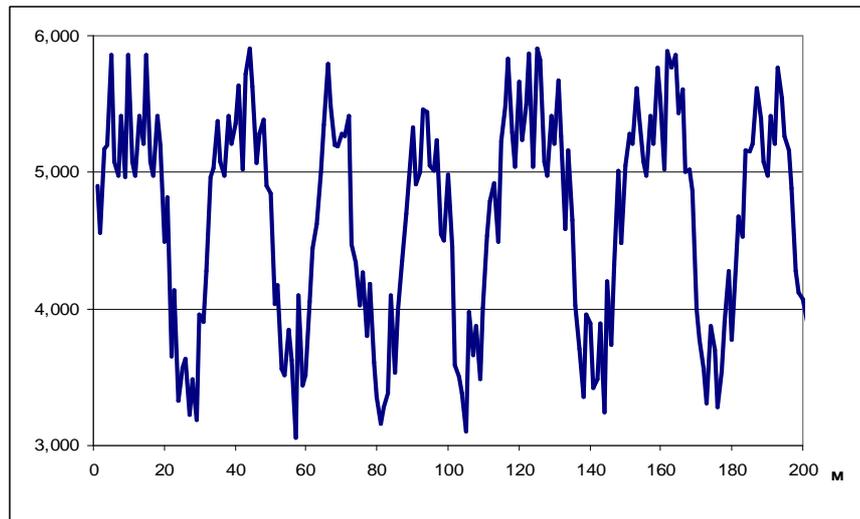


Рис. 3. А. - непрерывное сглаженное изменение, характерное для некоторых ландшафтных форм; Б.- непрерывное сглаженное изменение с локальным шумом, такое изменение характерно, например, для грунтовых вод. В. - простейший случай наличия пространственных зависимостей – такие периодические изменения часто образуются в результате антропогенного воздействия, например, вспашки

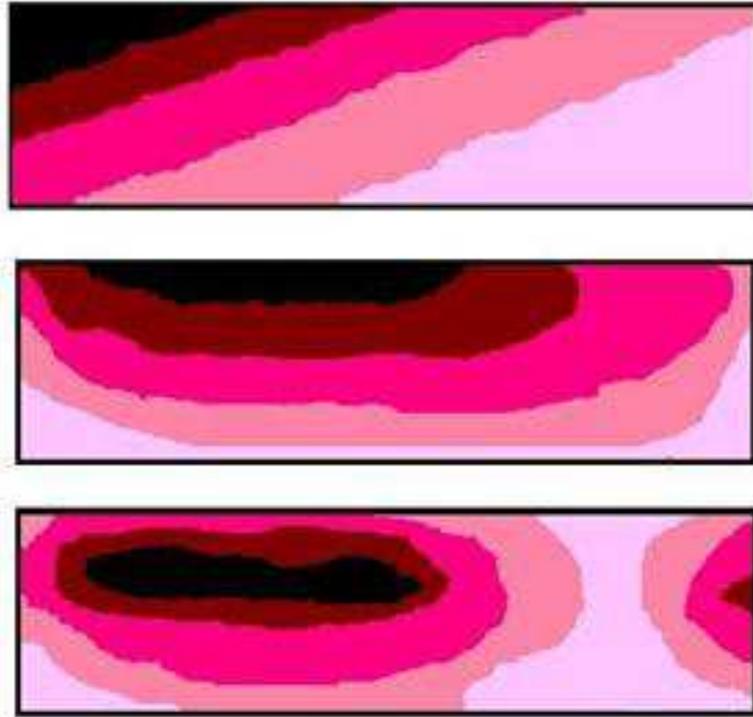


Рис. 4. Визуализация трендов разных порядков: от 1-го до 3-го.

2.3. ПРОСТРАНСТВЕННАЯ АВТОКОРРЕЛЯЦИЯ

Часто мелкомасштабные изменения почвенного свойства существенно коррелируют друг с другом – по крайней мере, в масштабе сделанных наблюдений. Такой характер пространственной зависимости приводит нас к модели, которая состоит из трех компонентов.

Первый - это средняя величина переменной внутри изучаемой территории. Второй вид варьирования – пространственно коррелируемые постепенные изменения. Третий компонент – некоррелируемые, случайные изменения, которые могут быть вызваны аналитическими ошибками, неточностями модели или пространственными изменениями на малых расстояниях:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon''$$

где $m(x)$ - член, описывающий общее среднее;

$\varepsilon'(x)$ - описывает пространственно коррелируемые, случайные изменения;

ε'' - остаток, пространственно некоррелируемый шум, имеющий нормальное распределение с нулевым средним.

Самым важным допущением является то, что статистические свойства $\varepsilon'(x)$ одни и те же – как в пределах всей изучаемой области, так и каждой из ее отдельных частей. Исследованию данной модели и будут посвящены дальнейшие лекции курса.



Какая модель является простейшей?

Простейшей моделью является случай, когда отсутствуют пространственные зависимости. Эта же модель наблюдается в том случае, когда случайная составляющая столь велика, что не позволяет выделить закономерную составляющую (рис. 5).

Значение свойства $Z(x)$ может быть описана следующим образом:

$$Z(x) = m(x) + \varepsilon$$

где $m(x)$ - член, описывающий общее среднее;

ε - остаток, пространственно некоррелируемый шум, имеющий нормальное распределение с нулевым средним.

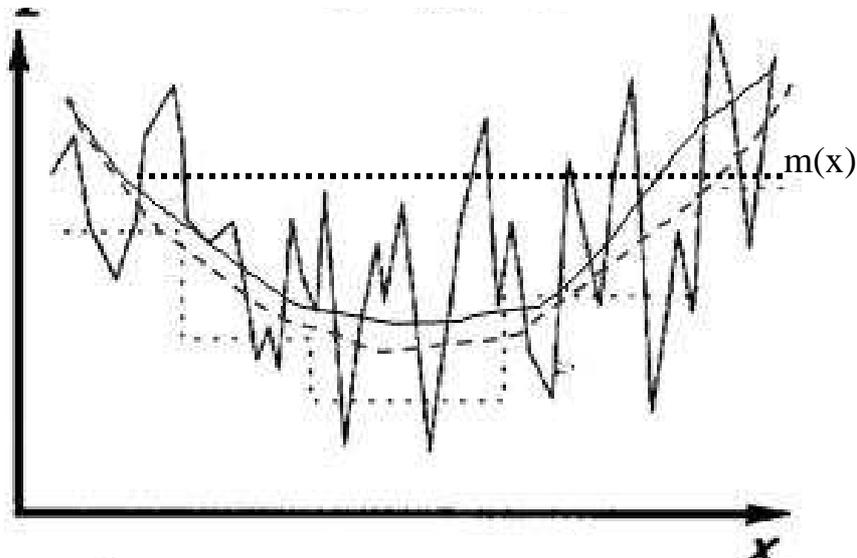


Рис. 5. Иллюстрация отсутствия пространственных зависимостей

2.4. ПРОСТЕЙШИЕ СПОСОБЫ ОТОБРАЖЕНИЯ ПРОСТРАНСТВЕННОГО ВАРЬИРОВАНИЯ СВОЙСТВ

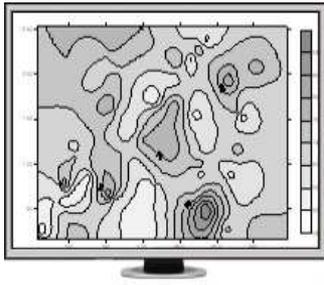
Если поле опробовано по регулярной или случайной сетке с достаточным числом точек, то о картине пространственного варьирования можно судить по процентному содержанию точек или долевого участию площади контуров, соответствующих каждому уровню обеспеченности.

В качестве наглядного представления пространственного варьирования используются квантили. Точки опробования отмечают на карте условными знаками, соответствующими четырем интервалам: 1) от минимума до нижнего квартиля (25%), 2) от нижнего квартиля до медианы (50% квантиль), 3) от медианы до верхнего квартиля (75%), 4) от верхнего квартиля до максимума.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Что такое хороплеты?**
- 2. Приведите пример пространственных изменений, которые могут быть описаны методом хороплет?**
- 3. Какие преимущества дает использование хороплет при моделировании пространственных зависимостей?**
- 4. Какие недостатки метода хороплет?**
- 5. В каких случаях используется модель трендов? Приведите пример.**
- 6. Чем отличаются тренды 1-го, 2-го и 3-го порядка?**
- 7. Какие недостатки проявляются при использовании трендов высокого порядка?**
- 8. В чем проявляется пространственная автокорреляция?**
- 9. Как может выглядеть функция, описывающая периодические зависимости?**
- 10. Приведите пример наблюдаемой периодической зависимости из экологии или почвоведения?**
- 11. Когда используется модель «отсутствие пространственных зависимостей» и имеет ли смысл ее использовать?**



Модуль 3. Концепция регионализированной переменной

Вы будете изучать

- Теорию регионализированной переменной
- Понятие стационарности 1-го и 2-го порядка
- Внутреннюю гипотезу

Цели модуля

- Дать представление о теории, на которой базируется геостатистика
- Показать истоки понятия «семивариограммы»

После изучения модуля Вы сможете

- Понимать отличия между случайной переменной и случайной функцией
- Знать допущения, на которых базируется теория регионализированной переменной



Основная литература

1. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия, 2004. – 320 с.
2. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.

2. Дмитриев Е.А. Теоретические и методологические проблемы почвоведения. – М.: ГЕОС, 2001. – 374 с.
3. Масштабные эффекты при исследовании почв. – М. Изд-во МГУ. 2001. – 400 с.
4. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
5. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85, no 2-3.
6. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML // Pedometron. Issue 18. July 2005.
7. McBratney A.B., Webster R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates // J.Soil Sci. 1986. V. 37. P. 617-639.
8. McBratney, A.B., Mendonca Santos, M.L., Minasny B. On digital soil mapping // Geoderma 117, 2003



Ключевые слова

Пространственные координаты,
случайная функция и случайная переменная,
теория регионализованной переменной,
стационарность 1-го порядка,
стационарность 2-го порядка,
внутренняя гипотеза

ВВЕДЕНИЕ

В геостатистике любой участок расценивается как состоящий из бесчисленного числа точек x_i , $i = 1, 2, \dots, \infty$. В каждой точке x почвенное свойство рассматривается как **случайная переменная** $Z(x)$, которая может принимать множество значений. В более привычных терминах классической статистики можно сказать, что в качестве случайной переменной в точке x можно рассматривать все значения генеральной совокупности любого из почвенных показателей изучаемой территории (величину рН, содержание $C_{орг}$, тяжелых металлов и т.д.)



Как можно описать распределение случайной переменной?

Случайная переменная подчинена закону распределения, который устанавливает связь между ее возможными значениями и соответствующими им вероятностями. Переменная в точке x имеет распределение со своим средним, дисперсией и моментами более высокого порядка, и ее настоящее значение $z(x)$ представляет собой случайную точку на графике данного распределения.

Значение, которое случайная переменная принимает в конкретном эксперименте, называется **реализацией случайной переменной** (или выборкой).

3.1. СЛУЧАЙНАЯ ФУНКЦИЯ

Переменные в других точках могут иметь конечное число фиксированных состояний, и настоящее значение в любой точке соответствует одному из этих состояний случайным образом. Количественное описание варьирования включает оценку характеристик того, что мы считаем случайной функцией.

Случайная функция – это набор случайных переменных $Z(x)$, характеризующих изменение показателя в некоторой области в зависимости от аргумента x .



Что может выступать в качестве аргумента?



В качестве аргумента в геостатистике могут рассматриваться пространственные координаты:

- ✓ одна переменная (x), когда речь идет о пробоотборе вдоль линии,
- ✓ две переменные (x, y), при пробоотборе из траншеи по глубинам или площадном опробовании с фиксированной глубины (рис. 6),
- ✓ три переменные (x, y, z), например, при пробоотборе из разрезов с некоторой территории.

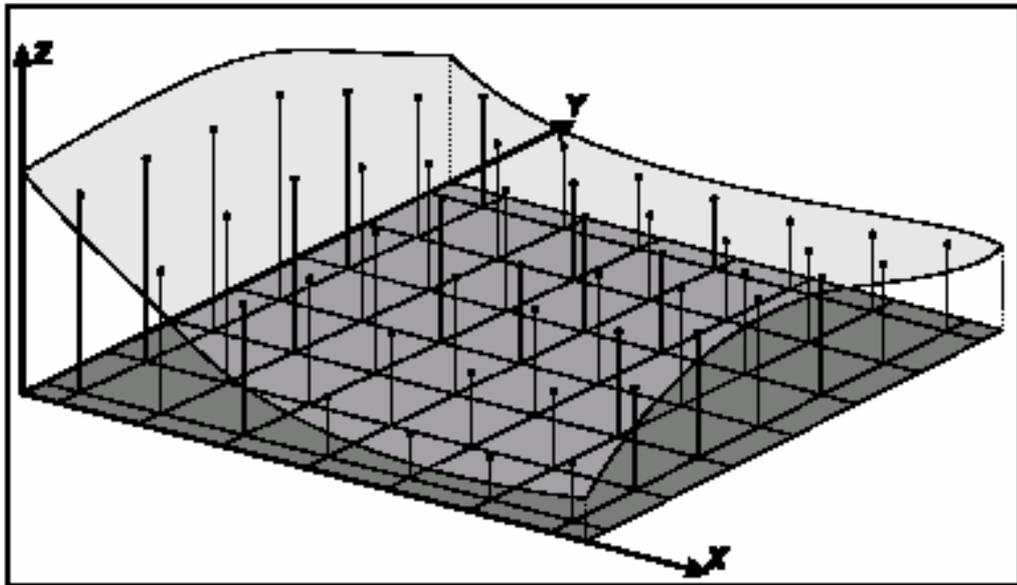


Рис. 6. Пространственная функция $Z(x)$ зависит от пространственных координат x – долготы и y – широты (по Gilberto и др., 2010)

При решении других задач аргументы могут быть и другими, например, при изучении сезонной динамики тех же почвенных свойств аргументом будет являться время. Вообще, под понятием пространство в геостатистике понимается как физическое пространство: недра Земли, почвы, атмосфера, подземные воды, так и абстрактные пространства (в математическом понимании).

В результате эксперимента, например – проведения почвенной съемки, случайная функция принимает вполне конкретный вид, называемый **реализацией** случайной функции.

Случайные функции изучаются одной из фундаментальных отраслей математики – теории случайных функций (Вентцель, 1969). Развитие этой теории применительно к пространственным объектам получило название теории регионализированной переменной.

3.2. ТЕОРИЯ РЕГИОнализованной переменной. СТАЦИОНАРНОСТЬ И ВНУТРЕННЯЯ ГИПОТЕЗА

Теория регионализированных переменных” (ТРП) была разработана Ж. Матероном (1968) на базе эмпирических зависимостей, которые применял южноафриканский геолог D.G. Krige для прогноза запаса золота в россыпях.



Что собой представляет регионализированная переменная?



В соответствии с ТРП регионализированная переменная, т.е. переменная, распределенная в пространстве, помимо случайной, "неорганизованной" компоненты варьирования, имеет закономерную составляющую варьирования, обладающую определенной пространственной структурой.

ТРП ориентирована на решение определенного типа математических задач, а именно описание структурных особенностей пространственного варьирования природных сред и создание на этой базе оптимального аппарата для интерполяции данных и интерпретации полученных результатов.

Базовым понятием теории случайных функций является понятие **стационарности**. **Стационарность 1-го порядка** выражается в том, что какую часть бы территории мы не обследовали, математическое ожидание и дисперсия являются величинами постоянными. Однако в реальной жизни такое условие практически никогда не выполняется.

Поэтому теория случайных функций основывается на **понятии стационарности 2 порядка (или слабой стационарности)**. Для характеристики случайной функции необходимо иметь данные о достаточно большом количестве ее реализаций. В реальных почвенных исследованиях обычно располагают результатами лишь одной почвенной съемки, что, казалось бы, не дает возможности охарактеризовать распределения случайной функции. Однако, эту проблему можно решить при наличии стационарности 2-го порядка в распределении изучаемого свойства.



Что такое стационарность 2-го порядка?

Под стационарностью второго порядка понимают:

1) постоянство значения математического ожидания (которое можно оценить по значению арифметического среднего) на всей изучаемой территории:

$$E \{z(x)\} = \mu = const \quad (1)$$

На практике это означает, что какую бы часть территории мы не рассматривали, локальное среднее на ней будет таким же как и генеральное среднее. Это означает также отсутствие тренда, т.е. выраженного изменения арифметического среднего в определенном направлении на изучаемой территории;

2) постоянство пространственной ковариации $C(h)$ для каждой пары случайных величин $\{z(x); z(x+h)\}$, ее зависимость только от расстояния между точками и независимость от положения точек в пространстве:

$$C(h) = E\{[z(x) - \mu] * [z(x+h)] - \mu\} \quad (2)$$

Применительно к почвенным исследованиям это означает, что вне зависимости от расположения точек опробования в пределах изучаемой территории, среднее произведение отклонений почвенных параметров от матема-

тического ожидания (или арифметического среднего) в точках, разделенных шагом h , есть величина постоянная¹.

Наличие стационарности второго порядка можно рассматривать как повторяемость функции в пространстве, поэтому становится возможным для ее характеристики заменить совокупность ее реализаций всего одной реализацией, описанной для достаточно большой (представительной) территории. Дисперсия случайной функции $z(x)$ в этом случае конечна и равна

$$C(0) = E\{[z(x) - \mu]^2\} \quad (3)$$

При несоблюдении условия стационарности второго порядка в распределении изучаемой переменной (например, при наличии тренда) в ряде случаев, можно применить несложные математические преобразования, приводящие к стационарности. Простейшим из таких преобразований является снятие тренда N -го порядка и последующий анализ остатков, для которых условие стационарности соблюдено. Кроме того, разработаны другие способы ухода от нестационарности.



Зависит ли стационарность от масштаба?



Понятие стационарности должно быть соотнесено с масштабом наблюдений. Например, некоторая область с локальным трендом может входить как небольшая часть в регион, который в целом может рассматриваться как стационарный (рис. 7).

Во многих случаях неоправданно ожидать, что среднее значение будет одинаково для всего изучаемого региона. В таких случаях ковариация не может быть определена, поскольку не существует значения для μ , чтобы подставить его в уравнение (2). Матерон (1968) отметил подобную ситуацию и предложил менее строгое утверждение, называемое внутренней гипотезой.

¹ Здесь рассматривается простейший случай так называемой изотропной реализации случайной функции.

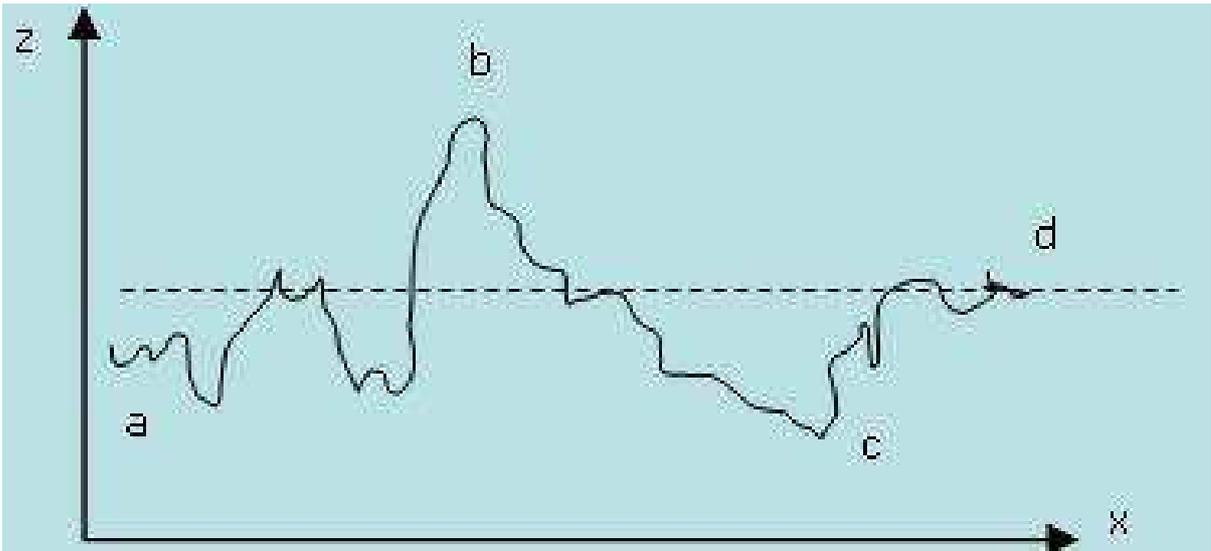


Рис. 7. Понятие стационарности выполняется на участке [a,d] и не выполняется для его части [b,c], где наблюдается тренд (по Анализ..., 1999)



Что такое внутренняя гипотеза?

Требования внутренней гипотезы еще менее строгие, чем стационарность 2-го порядка:

1. Ожидаемое различие между 2 точками x и $x+h$. Находящиеся на расстоянии вектора h , составляет нуль

$$E[Z(x) - Z(x+h)] = 0 \quad (4)$$

2. Делается предположение, что дисперсия точек, находящихся на расстоянии h зависит только от расстояния между точками и, следовательно,

$$E\{[Z(x) - Z(x+h)]^2\} = E\{[\varepsilon'(x) - \varepsilon'(x+h)]^2\} = 2\gamma(h) \quad (4a)$$

Так же как и ковариация, эта функция $\gamma(h)$, зависит только от расстояния между точками и не зависит от координат точек.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Чем случайная функция отличается от случайной переменной?**
- 2. Что такое ТРП?**
- 3. Что может выступать в качестве аргумента для $Z(x)$?**
- 4. Какие предположения формулируются относительно $Z(x)$ в ТРП?**
- 5. Для некоторой территории, в ходе исследований было показано, что для ряда свойств выполняется условие стационарности. Можно ли постулировать, что в ходе последующих разномасштабных исследований стационарность будет сохраняться?**
- 6. Что такое внутренняя гипотеза?**
- 7. Что такое реализация случайной переменной?**
- 8. Что такое реализация случайной функции?**



Модуль 4. Семивариограмма

Вы будете изучать

- Определение семивариограммы
- Пример расчета простейшей семивариограммы.
- Основные характеристики семивариограмм: ранг, порог, "наггет-эффект".
- Понятие об анизотропии
- Основные модели семивариограммы

Цели модуля

- Дать представление о том, что такое семивариограмма
- Показать, как анализируются семивариограммы
- Продемонстрировать каким образом происходит подбор модели

После изучения модуля Вы сможете

- Рассчитывать семивариограмму по исходным данным
- Осуществлять подбор модели



Основная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Масштабные эффекты при исследовании почв. – М.: МГУ. 2001. – 400 с.
3. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Кузякова И.Ф., Кузяков Я.В. Влияние микрорельефа на пространственное варьирование содержания гумуса в дерново-подзолистой почве длительного полевого опыта // Почвоведение. 1997. № 7.
2. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001; N 11.
3. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
4. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85, no 2-3.
5. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML // Pedometron. Issue 18. July 2005.
6. McBratney, A.B., Mendonca Santos, M.L., Minasny B. On digital soil mapping // Geoderma 117, 2003.



Ключевые слова

Семивариограмма,
транзитивность, шаг (лаг),
«нагетт»-эффект, ранг, порог,
расчет изотропной и анизотропной семивариограммы,
линейная, сферическая, экспоненциальная и Гауссова модели,
подбор модели, критерии подбора.

ВВЕДЕНИЕ

Величину $\gamma(h)$, получаемую в уравнении (4а) Матерон назвал «semivariance» или при дословном переводе «полудисперсия», подразумевая тем самым, что на каждую точку приходится половина от полученного значения математического ожидания квадрата разности значений. Однако по своей природе $\gamma(h)$ является дисперсией, а не ее половиной.

Эта же функция была ранее определена Колмогоровым (1939), который назвал ее «структурной» функцией. Как функция $\gamma(h)$ представляет из себя вариограмму и в русском переводе ее обычно называют «семивариограммой» или просто «вариограммой» (термины «полудисперсия» и «полувариограмма» нам кажутся не совсем удачными).

4.1. ОПРЕДЕЛЕНИЕ СЕМИВАРИОГРАММЫ

Помимо математического среднего и выборочной дисперсии, семивариограмма относится к основным статистическим показателям, используемым в геостатистике.



Что такое семивариограмма?

Семивариограммой $\gamma(h)$ называется величина:

$$\gamma(h) = 1/2E \{ [Z(x_i) - Z(x + h)]^2 \}, \quad (4b)$$

или, если вариограмма рассчитывается на основании данных измерений в отдельных точках:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x + h)]^2, \quad (5)$$

где $N(h)$ – количество пар точек, удаленных друг от друга на расстояние h , x – пространственная координата.

Как следует из приведенной выше формулы, функция $\gamma(h)$ от расстояния h характеризует зависимость среднего квадрата разности значений изучаемого показателя от расстояния между точками опробования, в которых

эти значения измерены. Вариограмма является основной характеристикой пространственной структуры варьирования природных объектов.



Что следует из постоянства пространственной ковариации?

Из постоянства пространственной ковариации для каждой пары случайных величин, находящихся друг от друга на расстоянии h , следует, что значения вариограммы не зависят от положения точек в пространстве, а зависят только от расстояния h между ними. Из (2), (3) и (4) следует, что

$$\gamma(h) = C(0) - C(h) \quad (6)$$

Данная зависимость связывает между собой вариограмму $\gamma(h)$ и ковариационную функцию $C(h)$, показывая, что первая по своей природе является дисперсией и составляет лишь некоторую часть от общей дисперсии $C(0)$ для показателей, измеренных на расстоянии h друг от друга. Если же пространственная ковариация отсутствует, то $\gamma(h)$ становится равной дисперсии $C(0)^2$.

Именно из этого предположения исходит классическая статистика, в которой все данные изучаемой совокупности рассматриваются как равноценные и взаимонезависимые. Такое утверждение верно для многих задач, решаемых почвоведением, когда точки обследования находятся относительно далеко друг от друга и не испытывают взаимовлияния.

4.2. ПРИМЕР РАСЧЕТА ПРОСТЕЙШЕЙ ВАРИОГРАММЫ

Приведем пример расчетов простейшего случая расчета семивариограммы. На рис. 8 приведены данные содержания подвижного фосфора в образцах, отобранных вдоль прямой с севера на юг с шагом 1 метр на поле 2-го севооборота УОПС "Чашниково" (дерново-подзолистые среднесуглинистые окультуренные почвы). Рассмотрим все точки, находящиеся на расстоянии 1 метра. Таких пар будет 8, то есть $N(h=1)$ равно 8.

² Отсюда следует, что термин «полудисперсия» не совсем оправдан и вносит путаницу.

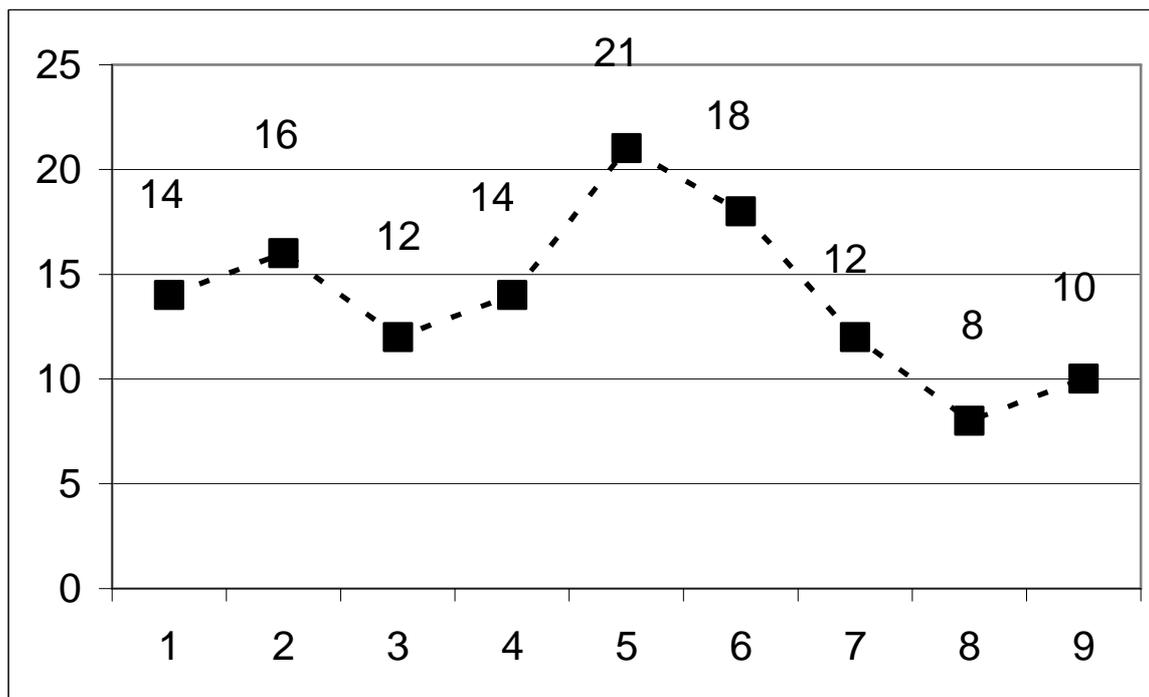


Рис. 8. Изменение содержания подвижного фосфора (мг-экв/ 100 г п) в образцах, отобранных вдоль прямой с севера на юг с шагом 1 метр на поле 2-го севооборота УОПС "Чашниково"

Построим для них функцию $\gamma(h)$ (табл. 1).

$$\gamma(1) = \frac{1}{2^{N(h=1)}} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x + 1)]^2 = \frac{1}{2^{*8}} \sum_{i=1}^8 (14-16)^2 + (16-12)^2 + (12-14)^2 + (14-21)^2 + (21-18)^2 + (18-12)^2 + (12-8)^2 + (8-10)^2 =$$

$$\frac{1}{2^{*8}} \sum_{i=1}^8 4 + 16 + 4 + 49 + 9 + 36 + 16 + 4 = 138 / 16 = 8,6$$

Рассмотрим все точки, находящиеся на расстоянии 2 метра. Таких пар будет 7, то есть $N(h=2)$ равно 7. Построим для них функцию $\gamma(h)$ (табл. 1).

$$\gamma(2) = \frac{1}{2^{N(h=2)}} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x + 1)]^2 = \frac{1}{2^{*7}} \sum_{i=1}^7 4 + 4 + 81 + 16 + 81 + 100 + 4 = 290 / 14 = 20,7$$

Рассмотрим все точки, находящиеся на расстоянии 3 метра. Таких пар будет 6. Построим для них функцию $\gamma(h)$ (табл. 1).

$$\gamma(3) = 298 / 12 = 24,8$$

Таблица 1. Пример расчета семивариограммы по шагам.

| Координаты точек | Содержание фосфора, мг-экв/100 г п | Шаг 1 | | Шаг 2 | | Шаг 3 | | Шаг 4 | |
|------------------|------------------------------------|-----------------------|-------------------|-----------------------|-------------------|-----------------------|-------------------|-----------------------|-------------------|
| | | $Z(x_i - Z(x_{j+1}))$ | То же, в квадрате | $Z(x_i - Z(x_{j+2}))$ | То же, в квадрате | $Z(x_i - Z(x_{j+3}))$ | То же, в квадрате | $Z(x_i - Z(x_{j+4}))$ | То же, в квадрате |
| 1 | 14 | -2 | 4 | 2 | 4 | 0 | 0 | -7 | 49 |
| 2 | 16 | 4 | 16 | 2 | 4 | -5 | 25 | -2 | 4 |
| 3 | 12 | -2 | 4 | -9 | 81 | -6 | 36 | 0 | 0 |
| 4 | 14 | -7 | 49 | -4 | 16 | 2 | 4 | 6 | 36 |
| 5 | 21 | 3 | 9 | 9 | 81 | 13 | 169 | 11 | 121 |
| 6 | 18 | 6 | 36 | 10 | 100 | 8 | 64 | | |
| 7 | 12 | 4 | 16 | 2 | 4 | | | | |
| 8 | 8 | -2 | 4 | | | | | | |
| 9 | 10 | | | | | | | | |
| Число пар точек | | | 8 | | 7 | | 6 | | 5 |
| Сумма | | | 138 | | 290 | | 298 | | 210 |
| Функция гамма | | | 8,6 | | 20,7 | | 24,8 | | 21 |

Рассмотрим все точки, находящиеся на расстоянии 4 метра. Таких пар будет 5. Построим для них функцию $\gamma(h)$ (табл. 1).

$$\gamma(4) = 210 / 10 = 21,0$$

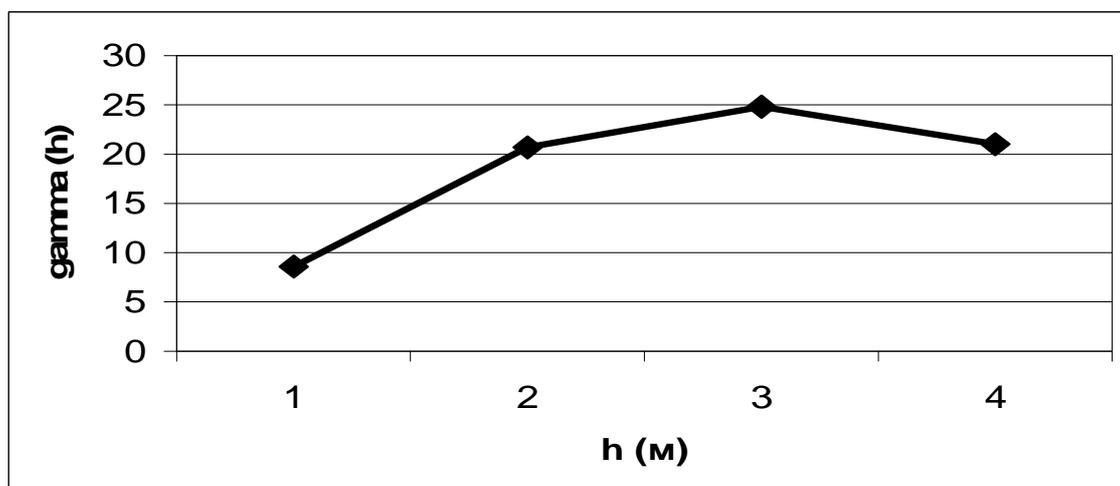


Рис. 9. Вариограмма, построенная по результатам расчетов в таблице 1

Построим по полученным значениям график. По оси абсцисс отложим расстояние между точками, а по оси ординат соответствующие им значения функции $\gamma(h)$ (рис.9).

4.3. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ СЕМИВАРИОГРАММ: РАНГ, ПОРОГ, "НАГГЕТ-ЭФФЕКТ"

! По характеру пространственной зависимости от шага опробования выделяют следующие «крайние» типы семивариограмм:

а) значения семивариограммы колеблются на одном уровне, что говорит об отсутствии пространственной корреляции (чистый «наггет-эффект»);

б) значение дисперсии между точками опробования линейно повышается с увеличением расстояния между ними (тренд);

в) дисперсия постепенно увеличивается, а затем выходит на плато (**транзитивная семивариограмма**).

Максимальное значение функции $\gamma(h)$, если семивариограммы транзитивна, называют **порогом**, величина которого должна примерно соответствовать дисперсии случайной величины (рис. 10).

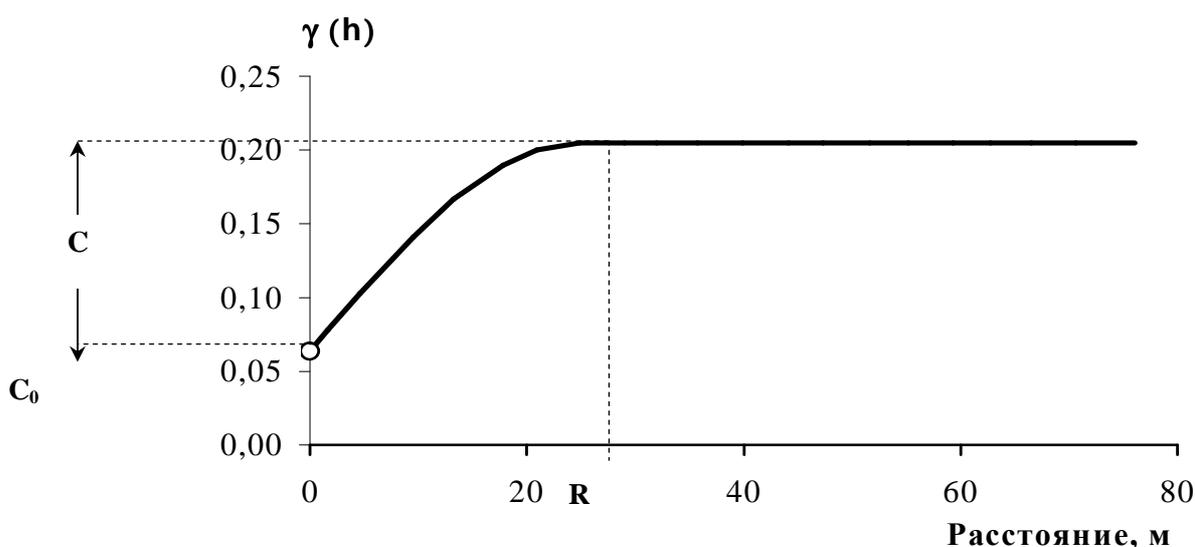


Рис. 10. Транзитивная семивариограмма (транзитивная модель).
 C_0 - «наггет»-эффект, R – ранг, C_0+C – порог



Что такое ранг?

Расстояние, при котором функция $\gamma(h)$ достигает постоянного значения, называется диапазоном скоррелированности или **рангом**. В пределах диапазона скоррелированности существует зависимость значений показателей между собой, т.е. $\gamma(h) < C(0)$, при расстояниях между точками опробования, превышающих диапазон скоррелированности, значения показателей действительно становятся взаимонезависимыми и равноценными, как того требует классическая статистика ($\gamma(h) = C(0)$).

Ранг определяет размеры однородных структур, выделяемых в почвенном покрове. Когда расстояние между точками (h) стремится к нулю, значение $\gamma(h)$ приближается к некоторому значению, называемому «наггет-эффект» (или «эффект самородка») и обозначается C_0 ³. В него входит варьирование случайной величины на расстояниях меньших, чем шаг, использованный в модели, а также варьирование, связанное с аналитическими ошибками. Разница между порогом и «наггет-эффектом» обозначается C .

❗ Порог и ранг наблюдаются не всегда, а только в случае транзитивной вариограммы. В этом случае интересно рассчитать отношение порог/наггет-эффект. Оно показывает, какая часть варьирования является пространственно зависимой, а какая случайной. Если значение близко к 1 и менее, то преобладает непространственное варьирование.



Как соотносятся выборочное значение дисперсии и порог?

Выборочная дисперсия значений изучаемых показателей оказывается, как правило, несколько ниже пороговой дисперсии. Это объясняется следующим. Выборочная дисперсия рассчитывается на основании значений всех

³ Значение модели семивариограммы в точке 0 равно нулю ($\gamma(h=0)=0$).

точек опробования изучаемой территории.

В реальном эксперименте территория обладает конечными размерами, и точки в пределах диапазона скоррелированности обычно вносят в значение выборочной дисперсии существенный вклад, понижая его относительно пороговой дисперсии.

! При расчете идеальной вариограммы такого ограничения территории не существует, поэтому значение дисперсии всей совокупности точек стремится к значению пороговой дисперсии.

4.4.ИЗОТРОПНАЯ ВАРИОГРАММА И ПРИНЦИП ЕЕ РАСЧЕТА В СОВРЕМЕННЫХ КОМПЬЮТЕРНЫХ ПРОГРАММАХ



Какие среды называют изотропными?

! Изотропными называют природные среды, свойства которых одинаковы во всех направлениях. В геостатистике под **изотропной семивариограммой** понимают вариограмму, характеризующую структуру варьирования показателя на всей территории опробования, независимо от направления.

Для построения изотропной вариограммы рассматривают все возможные пары точек опробования. Для каждой пары рассчитывают квадрат разности значений показателя в этих двух точках, пары объединяют в группы по удалению точек в парах друг от друга, а затем для каждой группы рассчитывают половину среднего квадрата разности значений показателя, т.е. функции $\gamma(h)$.

Для построения вариограммы используют значения функции $\gamma(h)$ при расстояниях, кратных единичному расстоянию h , называемому *шагом* (или *лагом*). Шаг может соответствовать наименьшему расстоянию между точками опробования, либо выбираться произвольно, исходя из условий конкретной задачи.

Таким образом, получают средние значения функции $\gamma(h)$ при расстояниях между точками пары, кратных шагу, т.е. при h , равных 1, 2, 3 и т.д. шагам. В реальности почти невозможно так организовать отбор образцов, чтобы расстояния между точками всегда оказывались точно кратными шагу. Поэтому расстояние между точками меньшее или равное единичному шагу приравнивают к шагу; расстояние в пределах от 1 до 2 шагов приравнивают к двум шагам и т.д.

В некоторых программах реализована другая схема: расстояние в пределах от 0,5 до 1,5 шагов приравнивают к одному шагу, расстояние от 1,51 до 2,5 приравнивают к 2 шагам и так далее.

На рис. 11 приведен принцип группировки точек для расчета изотропной вариограммы. Для рассмотрения всех пар точек с участием точки O проведены окружности радиусами, кратными шагу, с центром в этой точке. Точки A и B будут считаться удаленными от точки O на расстояние одного шага; точки C, D, E, F, G, H, I – на расстояние двух шагов и т.д.

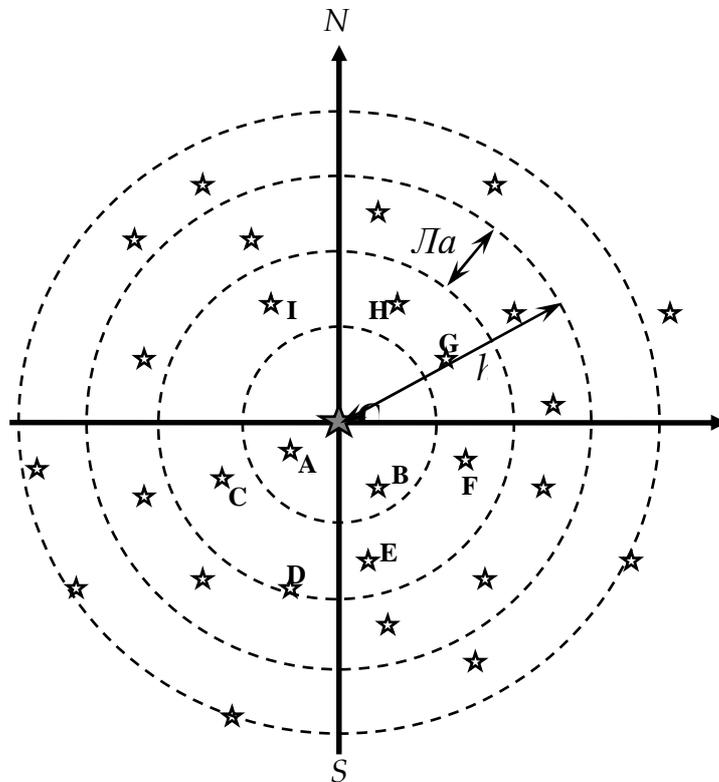


Рис. 11. Принцип расчета изотропной семивариограммы.

O, A, B, \dots, I - точки опробования, h - расстояние между точками опробования, кратное единичному шагу

Если провести такие же окружности с радиусами, кратными шагу, с центрами и во всех других точках, то получим все группы пар точек для расчета изотропной вариограммы.

4.5. РАСЧЕТ АНИЗОТРОПНОЙ ВАРИОГРАММЫ

Пространственная структура варьирования показателей не всегда одинакова по всем направлениям в пространстве. Для определения варьирования показателя по выбранным направлениям в пространстве рассчитывают *анизотропные вариограммы*. Пример расчета анизотропной вариограммы представлен на рис.12.

Направление в пространстве, для которого хотят рассчитать вариограмму, задается значением угла φ , который вектор, задающий направление, образует с осью X . (В некоторых программах за точку отсчета принимается ось Y).

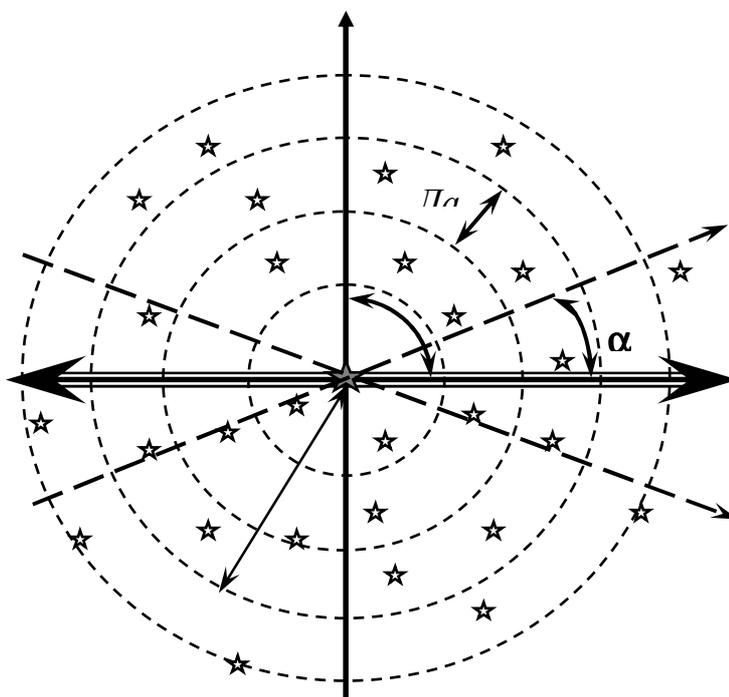


Рис. 12. Принцип расчета анизотропной вариограммы

α - угол толерантности, φ - значение угла между выбранным направлением в пространстве, для которого рассчитывается полувариограмма, и осью X , h - расстояние между точками опробования, кратное единичному шагу

От этого вектора в обе стороны откладывается так называемый *угол толерантности* α , определяющий тот пространственный сектор, в пределах которого рассматриваются точки для расчета вариограммы.

Таким образом, точки объединяют в группы не только по расстоянию между ними, но и по направлению вектора, соединяющего точки в паре. При угле толерантности, равном $\pm 90^\circ$ анизотропная вариограмма превращается в изотропную.

4.6. НАИБОЛЕЕ РАСПРОСТРАНЕННЫЕ ФОРМЫ (МОДЕЛИ) ВАРИОГРАММ И ФАКТОРЫ, ВЛИЯЮЩИЕ НА ТОЧНОСТЬ МОДЕЛИ

Семивариограмма рассматривается как графическое представление некоторой функции, поэтому можно подобрать модель, которой она описывается (аппроксимируется). Выделяют несколько типов вариограмм, характеризующих различные виды пространственной структуры варьирования показателей.

Основные типы вариограмм, а также формулы для их описания, представлены на рис. 13. Все модели приведены с «наггет-эффектом», так как на практике практически не встречается вариограмм, начинающихся из нуля.

Пространственная корреляция данных существует на расстояниях между точками, не превышающих ранг (диапазон скоррелированности), обозначенный в таблице как a . Далее семивариограмма выходит на плато (C_0+C), где достигает порогового значения.

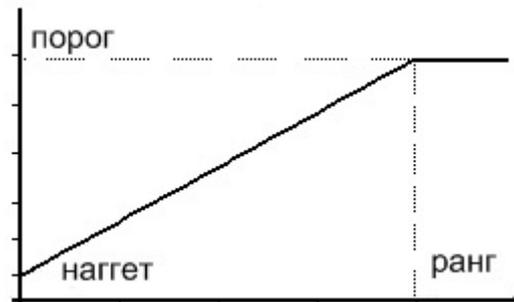
Для транзитивных семивариограмм обычно используют три основные стандартные модели: сферическую, экспоненциальную и модель Гаусса. Гауссова и экспоненциальная модели отличаются от сферической формой кривой на расстоянии меньших ранга. Линейная модель в настоящее время практически не используется, так как доказана ее неэффективность.

Модель/Формула/Параметры

График

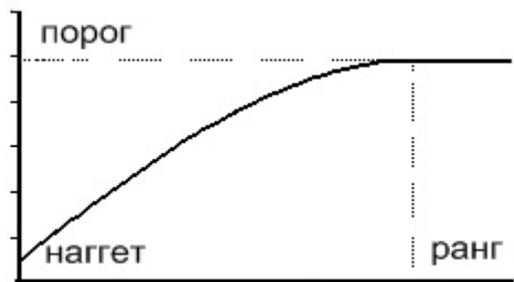
1. Линейная с выходом на плато

$$\chi(h) = \begin{cases} Co + C \cdot \frac{h}{a} & h < a \\ Co + C & h \geq a \end{cases}$$



2. Сферическая

$$\chi(h) = \begin{cases} Co + C \cdot \left[1,5 \frac{h}{a} - 0,5 \left(\frac{h}{a} \right)^3 \right], & 0 < h \leq a \\ Co + C, & h > a \end{cases}$$

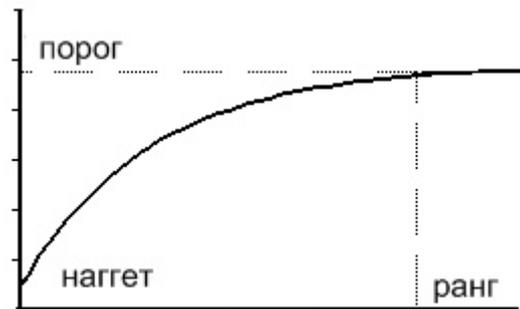


3. Экспоненциальная

$$\chi(h) = Co + C \cdot \left[1 - \exp\left(-\frac{h}{r}\right) \right]$$

$$a = 3 \cdot r; \chi(a) = Co + 0,95C$$

r - вспомогательный параметр.



4. Гауссова

$$\chi(h) = Co + C \cdot \left[1 - \exp\left(-\left(\frac{h}{r}\right)^2\right) \right]$$

$$a = 3 \cdot r; \chi(a) = Co + 0,95C$$

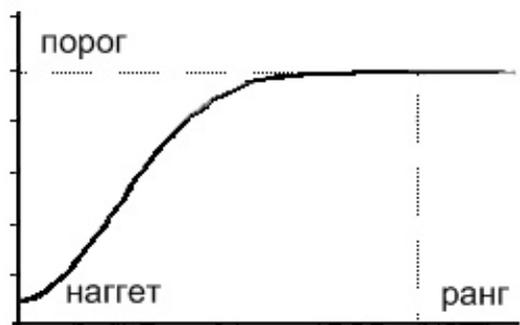


Рис. 13. Основные модели семивариограмм. *a* – ранг, $Co + C$ - порог



Как подбираются оптимальные параметры модели?



Для подбора оптимальных параметров модели используют такие методы, как вычисление минимальной суммы квадратов отклонений, максимального коэффициента детерминации, оценку максимального правдоподобия; наиболее распространенным методом является расчет взвешенной минимальной суммы квадратов отклонений.

В связи с математической неоднозначностью оптимизации экспериментальных данных по этим параметрам во многих программах подгон модели производится пользователем визуально (например, программы GEOSTAT, GEOEAS, Vesper).

Отношение суммы квадратов отклонений, связанных со случайностью, к общей сумме квадратов отклонений (ssd/sst) часто используется в качестве критерия качества подбора модели. Чем меньше отношение ssd/sst , тем лучше модель описывает данные.



Параметры модели семивариограмм используются для пространственной интерполяции данных методом кригинга (о чем будет рассказано далее), поэтому точность семивариограммы обуславливает точность интерполяции данных (Brus, Gruijter, 1997).



Что влияет на точность модели семивариограммы?

Факторы, влияющие на точность модели семивариограммы, были подробно изучены почвоведом-геостатистиком.

Для точной оценки параметров модели семивариограммы важным фактором является количество точек отбора. R. Webster и M. A. Oliver (1992) получили доверительный интервал для семивариограмм с помощью моделирования больших полей значений свойства и вычисления наблюдаемых семивариограмм и квартилей для них. Они сделали вывод, что выборка для по-

строения семивариограмм должна быть не менее 150 точек, а 225 точек дают хорошую точность.

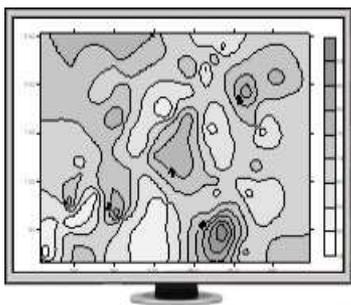
По данным публикаций можно назвать также такие факторы, как схема расположения точек отбора образцов. Для точной оценки параметров модели семивариограмм точки отбора должны располагаться на небольших расстояниях друг от друга.

Стоимость получения большого количества исходных данных является причиной недостаточно повсеместного использования методов геостатистики в почвенных исследованиях. Однако, Т. М. Burgess и R. Webster, (1980) показали, что семивариограммы могут быть рассчитаны по данным трансект, полученным на предварительной стадии почвенного исследования территории.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Что такое семивариограмма? Как еще называют эту функцию?**
- 2. Что такое шаг (или лаг)?**
- 3. Какие характеристики выделяют для транзитивной семивариограммы и что они описывают?**
- 4. Чем анизотропная семивариограмма отличается от изотропной?**
- 5. Зачем расчетную семивариограмму аппроксимируют некоторой функцией?**
- 6. Что влияет на точность модели семивариограммы?**
- 7. Как производят подгонку модели, по каким критериям?**



Модуль 5. Кригинг и кокригинг

Вы будете изучать

- Что такое кригинг?
- Проведение простейших расчетов
- Различные виды кригинга
- Кокригинг

Цели модуля

- Дать представление о кригинге как об интерполяционном методе
- Познакомить с примером расчетов
- Объяснить, чем отличаются разные виды кригинга и кокригинг

После изучения модуля Вы сможете

- Узнать об особенностях кригинга и кокригинга
- Выбрать вид кригинга, подходящий к данным конкретного исследования



Основная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Масштабные эффекты при исследовании почв. – М.: МГУ. 2001. – 400 с.
3. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Каневский М. Ф., Демьянов В. В., Савельева Е. А., Чернов С. Ю., Тимонин В. А. Кригинг и базовые модели геостатистики // В сб. Проблемы окружающей среды и природных ресурсов. М., 1999. № 11.
2. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001; N 11.
3. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
4. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85, no 2-3.
5. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML //Pedomatron. Issue 18. July 2005.



Ключевые слова

Кригинг,
универсальный кригинг,
разделительный кригинг,
точный и блочный кригинг,
ковариограмма,
кокригинг

ВВЕДЕНИЕ

Термин «кригинг» объединяет целое семейство алгоритмов пространственной регрессии. Важным преимуществом кригинга по сравнению с другими методами интерполяции, например – методом обратных расстояний, является наличие ошибки предсказания, то есть вычисляется не только точечная, но и интервальная оценка для предсказанных значений.

Другим преимуществом является учет особенностей пространственного варьирования (радиуса корреляции, анизотропии), отражающееся в модели семивариограммы. Многие виды кригинга позволяют включать априорную информацию в процесс предсказания.

Применение каждого из этих методов зависит от целей исследования, от того в каком пространственном масштабе оно проводится и от требуемой точности предсказания (McBratney и др., 2003).

5.1. ЧТО ТАКОЕ КРИГИНГ?

Термин кригинг происходит от имени ученого D. Krige, в честь которого Матерон (1968) и назвал базовые геостатистические модели. Оценка методом кригинга обладает минимальной дисперсией и дает интервальную оценку в каждой точке интерполяции. Важным свойством кригинга является его невысокая чувствительность точности метода к изменению объема выборки (Laslet и др., 1987).



В чем состоит основной принцип крикинга?



Принцип кригинга состоит в определении весов значений переменной в окрестных некоторой точках для оценки значения переменной в искомой точке или области. При этом вес вклада каждой точки определяется по вариограмме.

Весы подбираются таким образом, чтобы оценка искомого значения была несмещенной, а дисперсия оценки – минимальной. Чем ближе конкрет-

ная точка опробования к точке или области, для которых нужно определить значение признака, тем больший вклад ее значение вносит в искомое. Понятно, что для оценки имеет смысл привлекать только те точки, которые лежат на расстояниях, не превышающих диапазона скоррелированности.

Для кригинга обычно используют не экспериментальные вариограммы, а теоретические модели, которыми эти экспериментальные вариограммы описываются.

Первые фундаментальные работы по применению кригинга в почвоведении относятся к началу 80-тых годов (Burgess и Webster, 1980). Так называемый "обычный" кригинг (Ordinary kriging) нашел широкое распространение во многих приложениях почвоведения: рекультивации (Samara, Singh, 1990), почвенной классификации (Odeh и др., 1992; Burrough и др., 1992), изучении почвенного засоления (Bourgault и др., 1995), изучении загрязнения (Hendricks Frassen и др., 1997).

Система уравнений простейшего кригинга выглядит следующим образом:

$$\left\{ \begin{array}{l} \sum_{j=1}^N \lambda_{j0} \gamma(x_i - x_j) + \mu = \gamma(x_i - x_0), \quad i = 1, \dots, N \\ \sum_{i=1}^N \lambda_{i0} = 1 \end{array} \right.$$

где x_0 – точка, где значение функции неизвестно и нужно его предсказать;
 $x_i, i = 1, 2, \dots, N$ – точки, где был проведен пробоотбор и измерено значение признака, всего таких точек N ;

μ – генеральное среднее;

λ_i – веса, по которым рассчитывается вклад значений признака в точках пробоотбора в значение признака в искомой точке, эти веса рассчитываются исходя из расстояний между точками пробоотбора и точкой x_0 , используя семи-вариограмму.

5.2. ПРИМЕР ПРОСТЕЙШИХ РАСЧЕТОВ

Цель – предсказать значение y_0 в т. x_0 , где не было наблюдения. Для этого воспользуемся формулами:

$$\mathbf{t} = \sum_{i=1}^n \lambda_i \mathbf{y}_i = \boldsymbol{\mu} + \mathbf{g}'_0 \mathbf{G}^{-1} (\mathbf{y} - \boldsymbol{\mu} \mathbf{1}_n)$$

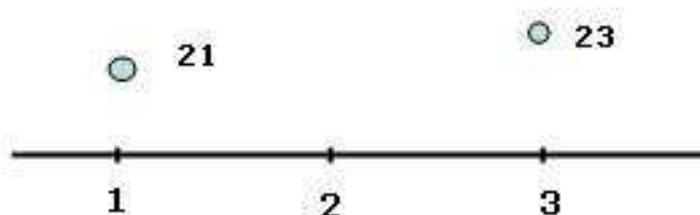
$$\boldsymbol{\mu} = (\mathbf{1}'_n \mathbf{G}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}'_n \mathbf{G}^{-1} \mathbf{y}$$

где \mathbf{t} – оценка для значения y_0 ; $\boldsymbol{\mu}$ – вектор средних

$\mathbf{1}_n$ – единичный вектор, длиной n ; знак ' означает транспонирование;

\mathbf{G} – матрица значений, полученная по семивариограмме для точек пробоотбора; -1 означает, что матрица обратная;

\mathbf{g}_0 – вектор значений функции семивариограммы, для расстояний от точки x_0 до точек пробоотбора; \mathbf{y} – вектор значений признака в точках пробоотбора.



Экспоненциальная модель: $\gamma(h) = 1 + 3(1 - \exp(-h/0,5))$

| h | 0,01 | 0,12 | 0,25 | 0,5 | 0,75 | 1 | 1,5 | 2 | 2,5 | 3 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| gamma | 1,059 | 1,640 | 2,180 | 2,896 | 3,331 | 3,594 | 3,851 | 3,945 | 3,980 | 3,993 |

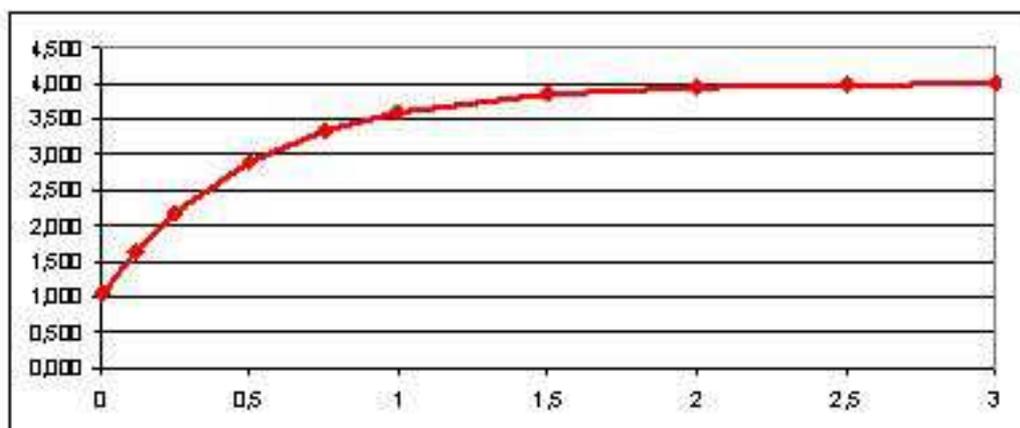


Рис. 14. Иллюстрация для примера расчетов кригинга

На рис. 14 даны исходные условия. Пусть имеется две точки, расположенные на прямой в точках 1 и 3. Значения признака в этих точках соответственно 21 и 23. Нужно предсказать значение признака в точке 2, расположенной на одной прямой с точками 1 и 3.

Известно, что пространственные зависимости описываются экспоненциальной моделью семивариограммы, которая приведена на рисунке. Там также рассчитаны значения функции $\gamma(h)$ для разных расстояний.

Запишем исходные условия в векторной форме:

$$y = \begin{pmatrix} 21 \\ 23 \end{pmatrix}, \mathbf{1}^2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{G} = \begin{pmatrix} 0 & 3,945 \\ 3,945 & 0 \end{pmatrix}, \mathbf{G}^{-1} = \begin{pmatrix} 0 & 0,253 \\ 0,253 & 0 \end{pmatrix}$$

$$\mathbf{G}\mathbf{G}^{-1} = \begin{pmatrix} 3,945 * 0,253 = 1 & 0 \\ 0 & 3,945 * 0,253 = 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Рассчитаем, значение μ

$$\mu = (\mathbf{1}'_2 \mathbf{G}^{-1} \mathbf{1}_2)^{-1} \mathbf{1}'_2 \mathbf{G}^{-1} \mathbf{y} =$$

$$= \left[(1 \ 1) \begin{pmatrix} 0 & 0,253 \\ 0,253 & 0 \end{pmatrix} * \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]^{-1} (1 \ 1) \begin{pmatrix} 0 & 0,253 \\ 0,253 & 0 \end{pmatrix} * \begin{pmatrix} 21 \\ 23 \end{pmatrix} = 22$$

Задача состоит в том, чтобы предсказать значение в т.2

Так как точка 2 находится на расстоянии 1 от точек отбора 1 и 3, то

$$g_0 = \begin{pmatrix} 3,594 \\ 3,594 \end{pmatrix}, \mu = 22$$

Теперь, произведем расчеты:

$$\mathbf{t}_2 = \mu + \mathbf{g}'_0 \mathbf{G}^{-1} (\mathbf{y} - \mu \mathbf{1}_n) =$$

$$= 22 + (3,594 \ 3,594) * \begin{pmatrix} 0 & 0,253 \\ 0,253 & 0 \end{pmatrix} * \left(\begin{pmatrix} 21 \\ 23 \end{pmatrix} - 22 * \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) =$$

$$= 22 + (0,91 \ 0,91) * \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 22$$

Ответ: в точке 2 значение признака будет 22.

5.3. ТОЧЕЧНЫЙ И БЛОЧНЫЙ КРИГИНГ

В зависимости от размера площади, для которой нужно оценить значение показателя, кригинг подразделяют на **точечный** и **блочный**. При точечном кригинге определяют значение показателя в отдельно взятой точке пространства.

При совпадении оцениваемой точки с одной из точек измерений, значение, даваемое кригингом, в точности совпадает с измеренным значением. Блочный кригинг используется, когда требуется оценить интегральное значение функции в некоторой области.



Что такое блочный кригинг?



Блочный кригинг, например, применяется при оценки запасов природных ископаемых или при оценки загрязняющего вещества на определенной площади. Этот метод осуществляется следующим образом: вся площадь разбивается на блоки. В каждом блоке производится несколько точечных оценок в разных местах блока. Затем берется среднее значение по блоку. Результатом является предсказанное локальное среднее значение.



Что такое точечный кригинг?



Точечный кригинг обычно рассматривают как предельный случай блочного, при котором размер блока соответствует размерам почвенной пробы. Блочный кригинг является незаменимым методом, когда требуется объединить данные, полученные из разных источников, производящих измерения в разных масштабах.

Например, с помощью блочного кригинга можно соединить в одной модели данные, полученные со спутникового снимка, размер пикселя кото-

рого составляет несколько десятков метров и данные, полученные в ходе наземных исследований, когда пробыв были отобраны с площадки в 1 кв. м.

Было показано преимущество блочного кригинга над точечным. Так, И. Ф. Кузякова и др. (2001) использовали метод блочного кригинга, чтобы повысить тесноту связей между показателями за счет частичного устранения высокочастотной компоненты варьирования. Т. М. Burgess и R. Webster (1980), используя блочный кригинг, получили карту с меньшей дисперсией, с меньшим «наггет-эффектом» и с более сглаженными контурами по сравнению с картой, построенной точечным кригингом.

5.4. ВИДЫ КРИГИНГА

Ограничения применения **обычного** кригинга связаны с наличием трендов, то есть направленных изменений локального среднего в рамках исследуемой территории. Преодолеть эту трудность позволяет модель универсального кригинга, представляющая собой комбинацию модели многомерной линейной регрессии и обычного кригинга. Этот метод относится к, так называемым, методам «нестационарной геостатистики»..



Что такое универсальный кригинг?



Универсальный кригинг (иногда его называют глобальным) предполагает, что в данных имеется доминирующая тенденция (например, господствующий ветер), и ее можно смоделировать с помощью полиномиальной функции.

Этот полином вычитается из исходных значений, и семивариограмма моделируется по остаткам. Когда к остаткам подобрана модель, перед расчетом интерполяции полином суммируется с полученными данными, чтобы получился осмысленный результат.



Что такое разделительный кригинг?

! **Разделительный** кригинг (его иногда называют дизъюнктивным, используя кальку с английского термина) позволяет применять стандартные статистические расчеты к переменным, распределение которых существенно отличается от нормального (т.е. стандартные формы преобразования исходной переменной: логарифмирование или извлечение квадратного корня, не позволяют получить распределения, близкого к нормальному), но подчиняется условию стационарности второго порядка.

При разделительном кригинге совокупность значений исходной переменной **разделяют** на ряд нормально распределенных подсовкупностей, для которых затем применяют интерполяцию методом кригинга.

Разделительный кригинг используют, в частности, для расчета вероятности превышения изучаемой переменной определенных пороговых значений - например, величины ПДК, уровня кислотности или щелочности и т.д.

Таким образом может проводится выявление площадей, требующих неотложных консервационных и мелиоративных мероприятий или регулярного мониторинга независимо от распределения значений показателя, что обуславливает преимущества данного метода по сравнению со стандартным кригингом.

Многомерные подходы, такие как **факториальный кригинг**, могут использоваться для исследования того, как взаимосвязи между переменными изменяются в зависимости от масштаба пространственного рассмотрения и, тем самым, дают нам ключ к пониманию физических процессов в зависимости от масштаба (Goovaerts, 1999).

Факториальный кригинг был предложен Матероном в 1982 году как один из вариантов кригинга (Matheron, 1982), целью которого является оценка и картографирование нескольких составляющих пространственного варьирования, выявляемых по экспериментальной семивариограмме. Во мно-

гих случаях эти пространственные составляющие обусловлены реальными процессами, связанными с формированием почвенного покрова.

Гуверт использовал факториальный кригинг для отделения локальной составляющей варьирования свойств почв и растений внутри поля от варьирования, обусловленного принадлежностью участков к разным классификационным разностям (Goovaerts, 1994). Вебстер (1994) использовал аналогичные подходы для картографирования концентраций тяжелых металлов, пространственное распределение которых связано с антропогенным воздействием, с одной стороны, и подчиняется региональным изменениям, обусловленным геологическими особенностями местности.

5.5. ПРОСТРАНСТВЕННАЯ ЗАВИСИМОСТЬ МЕЖДУ ПЕРЕМЕННЫМИ (КОВАРИОГРАММА) И КОКРИГИНГ

Достаточно часто пространственное распределение одного почвенного свойства определяется распределением другого свойства, либо распределение двух или нескольких свойств обусловлено действием одного и того же фактора варьирования. В качестве примера первого случая можно привести зависимость содержания гумуса от содержания тонкой гранулометрической фракции, с которой гумус обычно связан; в качестве второго – миграцию растворимых Ca^{2+} и Mg^{2+} по склону.

В этом случае говорят о *корегionalизации* переменных. Для оценки пространственной связи между переменными рассчитывают кроссвариограмму (или ковариограмму), которая по сути является произведением двух вариограмм. Если принять, что в каждой точке рисунка 1 измерены значения двух почвенных свойств, то значение ковариограммы при расстоянии между точками h рассчитывается как произведение вариограммы для всех точек одного показателя, находящихся друг от друга на расстоянии h , и вариограммы всех точек другого показателя, разделенных тем же расстоянием h .

Описание пространственного варьирования с помощью коррелограмм часто приводит к получению новой информации о связях почвенных свойств и окружающей среды, а также почвенного покрова и ландшафта. Так, например, МакБратни и Вебстер (1981), исследуя участок 0,4 га, выявили, что периодичность на коррелограмме урожая пшеницы обусловлена ранней вспашкой. Робертсон и др. описал пятнистости по количеству доступного азота размером 20-40 м, которые определили структуру растительных сообществ.



Что такое кокригинг?



Кокригингом называется многомерное развитие кригинга, когда в процесс предсказания почвенного свойства включается вторая переменная, которая легче измеряется или по которой имеются многочисленные данные (Wackernagel, 1995). Это в какой-то мере аналог методов регрессии в традиционной статистике.

Основная идея кокригинга заключается в интерполяции значений одной переменной по значениям другой, корегионализированной с ней. Кокригинг применяют с целью индикации одного свойства по другому, т.е. в тех случаях, когда определение одного свойства значительно проще, чем определение другого (например, из-за высокой стоимости анализов последнего).

Например, если имеются многочисленные данные по гранулометрическому составу и лишь немного измерений влажности на том же участке, то недостающие данные по влажности могут быть получены расчетным путем на основании данных о гранулометрическом составе.

Преимущество использования кокригинга по сравнению с кригингом тем выше, чем больше корреляция между переменными и чем больше объем вторичной информации, на основании которой предсказываются почвенные свойства.

5.6. УЧЕТ ВАРЬИРОВАНИЯ ВО ВРЕМЕНИ



На какие группы могут разделяться характеристики окружающей среды по типу их временной изменчивости?

Обычно ученые разделяют все характеристики окружающей среды на две большие группы: консервативные свойства и режимы.



Консервативные свойства, например, почвенные, такие как морфологические, физические, минералогические свойства и др., характеризуются значительным пространственным варьированием и практически не изменяются во времени или, по крайней мере, характерное время их изменений больше периода наблюдений.

Варьирование температурного и газового режимов, изменения жидкой фазы почв и др. обладает сильно выраженной временной составляющей. Поэтому традиционно для описания консервативных почвенных свойств применяют геостатистические методы, а для режимов используют анализ временных рядов.

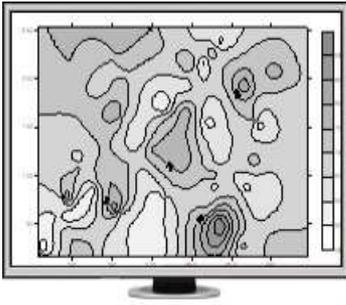
В геостатистике варьирование во времени может трактоваться аналогично варьированию в пространстве (Stein и др., 1997) Варьирование во времени качественно отличается тем, что может быть определено только тем значением, которое было в данной точке в предыдущий момент. Кроме того, варьирование во времени может иметь четко детерминированную составляющую: сезонную (суточную, годовую).

Для решения этих задач в геостатистике развиваются специальные методы (Hoosbeek, 1998) показано, что модель, рассматривающая время как еще одну координату и состоящая из стохастической и динамической составляющих лучше, чем модель, являющаяся просто суммой пространственной и временной составляющих.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Что такое кригинг?**
- 2. В чем состоит основной принцип крикинга?**
- 3. На основании каких данных делается предсказание значения в точке, где не было наблюдения?**
- 4. Для решения каких задач используется блочный кригинг?**
- 5. Что такое универсальный кригинг?**
- 6. Что такое разделительный кригинг?**
- 7. Кем и зачем был предложен факториальный кригинг?**
- 8. Какие еще виды крикинга Вы знаете? Чем они отличаются?**
- 9. В каких случаях используется кокригинг?**
- 10. Как в геостатистике учитывается фактор времени?**



Модуль 6. Системы пробоотбора, ошибка интерполяции и оценка точности картограмм

Вы будете изучать:

- Влияние схемы пробоотбора на получаемую картограмму
- Источники ошибок интерполяции
- Методы проверки точности картограмм
- Способы сравнения интерполяционных карт

Цель модуля:

- Показать роль пробоотбора в конечном результате.
- Рассказать о современных методах проверки точности картограмм
- Описать несколько процедур, используемых для сравнения картограмм

После изучения модуля вы сможете:

- Подобрать для конкретного проекта процедуру проверки точности построенной картограммы
- Сравнить количественными методами интерполяционные карты



Основная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Масштабные эффекты при исследовании почв. – М.: МГУ. 2001. – 400 с.
3. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Каневский М. Ф., Демьянов В. В., Савельева Е. А., Чернов С. Ю., Тимонин В. А. Кригинг и базовые модели геостатистики // В сб. Проблемы окружающей среды и природных ресурсов. М., 1999. № 11.
2. Козловский Ф.И. Теория и методы изучения почвенного покрова. - М.: ГЕОС, 2003 - 536с.
3. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001. N 11.
4. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
5. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85. no 2-3.
6. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML //Pedomtron. Issue 18. July 2005.



Ключевые слова.

Схема пробоотбора,
стратифицированный случайный отбор,
метод наибольшей беспристрастности,
рандомизированный отбор,
бутстреп тестирование, джекнайф,
ошибка интерполяции.

ВВЕДЕНИЕ

Полевые почвенные или экологические исследования никогда не бывают исчерпывающими. Кроме того, такие исследования являются дорогостоящими и требуют наличия в своем составе высококвалифицированных специалистов, количество которых ограничено. Поэтому всегда в ходе исследования необходимо ответить на вопросы:

- ✓ Какое количество точек пробоотбора является достаточным для решения поставленной задачи?
- ✓ Как расположить эти точки?
- ✓ Насколько достоверны результаты интерполяции и экстраполяции при данном пробоотборе?, то есть какова точность получаемых результатов.

6.1. РОЛЬ СХЕМЫ ПРОБООТБОРА

В зарубежной литературе довольно много методологических работ посвящено изучению количества почвенных проб и их расположению в пространстве. Основным фактором, влияющим на точность и стоимость обзорных почвенных карт, является схема расположения точек при отборе образцов.

Большое количество точек отбора обеспечивает отчетливую картину пространственного варьирования изучаемого свойства, что требует существенных затрат. Дешевле отобрать и проанализировать малое число образцов, но в этом случае возможна потеря части информации о пространственном распределении изучаемого свойства. В связи с этим в литературе предложены различные способы и схемы отбора почвенных проб в зависимости от задач исследования.

Предлагается отбирать образцы в точках, расположенных случайно, так называемый «рандомизированный» отбор. Такая процедура, названная авто-

рами «вероятностным» отбором, исключает смещение и субъективность оценок (Brus, Gruijter, 1997).

Вероятностный отбор образцов может быть проведен с помощью различных процедур, таких как «простой случайный» отбор, «стратифицированный» отбор, «кластерный» отбор и другие (Cochran, 1977, Saerndal и др., 1992).

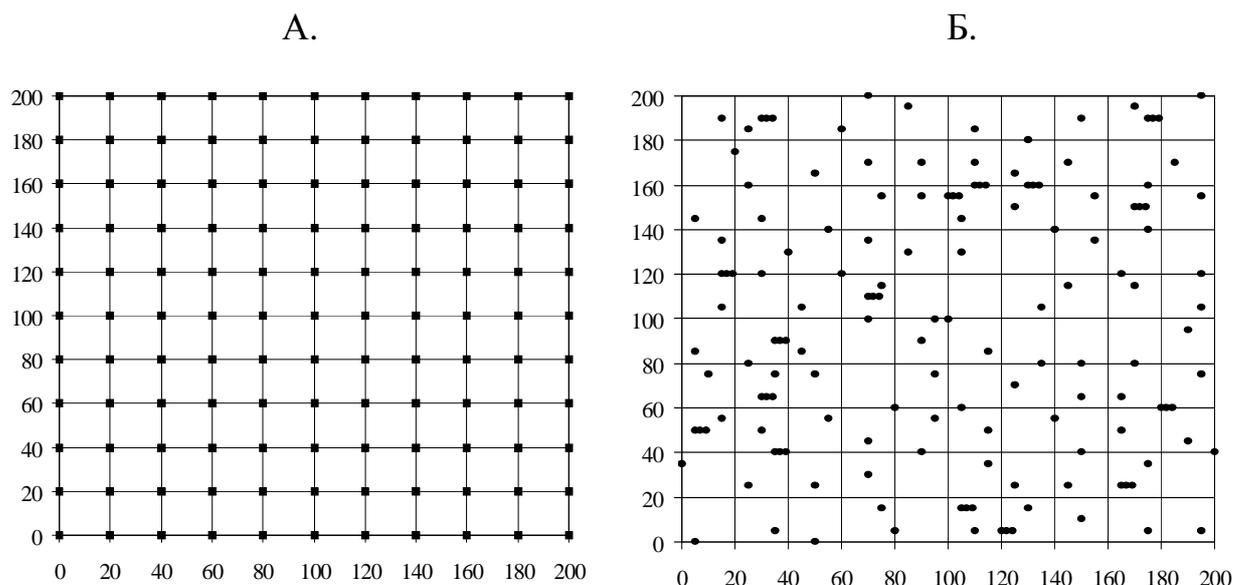


Рис. 15. Два способа пробоотбора одного и того же участка: А - систематический отбор; Б - стратифицированный случайный отбор

Точность результатов зависит от равномерности распространения точек отбора по полю. Систематический отбор в узлах регулярной сетки обеспечивает равномерность отбора проб. Равномерность охвата при случайном отборе достигается разбиением поля на блоки (страты) равной площади и выбором в каждом блоке 1 точки случайным образом. Такой отбор называется «стратифицированным случайным» отбором (Brus и др., 1999).

При исследовании содержания гумуса в пахотном слое пахотных дерново-подзолистых почв на ключевых участках образцы отбирались, используя два способа опробования: стратифицированный случайный отбор и систематический отбор (рис. 15).



Что собой представляет стратифицированный случайный отбор?

Стратифицированный случайный отбор состоит в следующем: все поле разбивается на квадраты одинакового размера (в данном случае 20 м на 20 м). В каждом квадрате случайным образом выбрана точка. Всего 100 точек. Дополнительно внутри 16-ти случайных квадратов выбраны еще по три точки. Причем первая точка выбрана случайно, а 2-я и 3-я на расстоянии 2 и 4 м от первой (рис. 15, Б).

Такая схема обеспечивает случайность координат точек и, в то же время, равномерно распределяет точки по всей площади. Преимуществом этой схемы по сравнению с систематическим отбором (отборе образцов в узлах регулярной сетки 20 м на 20 м) является наличие пар точек на расстоянии, меньших размера ячейки сетки отбора (20 м), а также увеличение числа точек на расстояниях, меньших 50 м.

Это хорошо видно на рис. 16, где показана зависимость числа пар точек от расстояния между ними для обоих способов отбора.

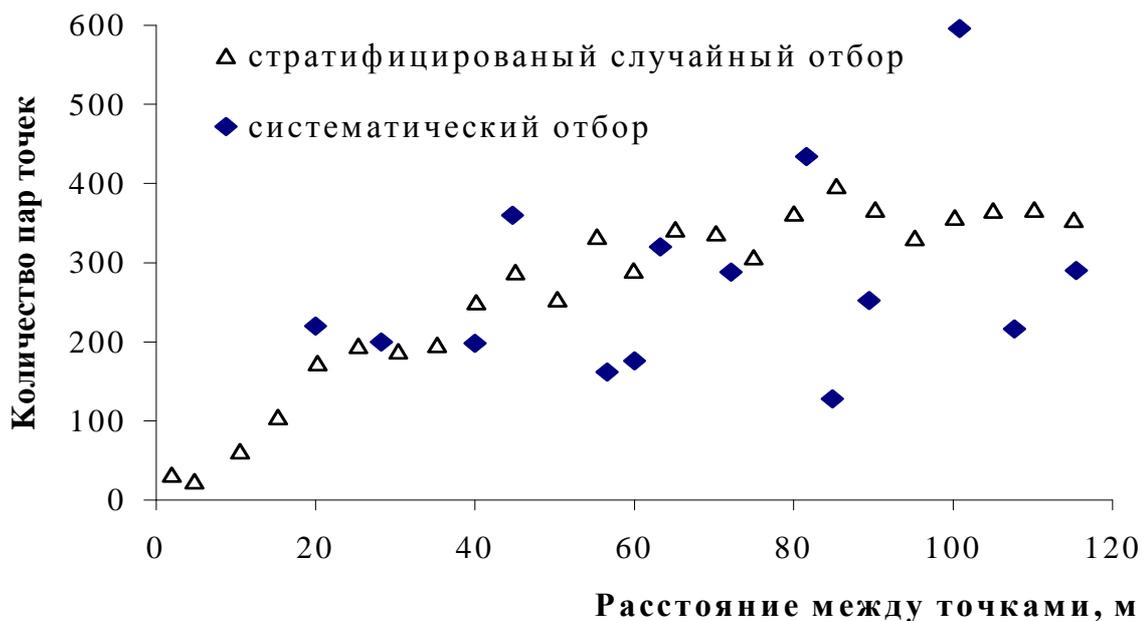


Рис. 16. Количество пар точек в зависимости от расстояния между ними

В работе J. W. van Groenigen с соавторами (1999) разработан способ выбора схемы отбора проб для минимизации дисперсией кригинга. Авторы показали, что выбор радиуса влияния (то есть числа ближайших соседних точек к той точке, в которой проводится интерполяция) при кригинге мало влияет на конфигурацию оптимальной схемы пробоотбора в условиях изотропии.

В российском почвоведении также имеются работы о количестве и расположении точек опробования при исследовании варьирования агрохимических показателей. Было показано, что для построения агрохимических карт опытного поля площадью 22 га необходим отбор не менее 70 индивидуальных образцов.

Карты, построенные на основе меньшего количества точек, в меньшей степени учитывают роль микрорельефа в воспроизведении пространственной неоднородности (Романенков, Кузякова, 2000). З.А. Прохорова и А.С. Фрид (1993) показали, что сетка отбора $3 \times 4 \text{ м}^2$ вскрывает варьирование агрохимических показателей дерново-подзолистой почвы внутри ЭПА, а шаг в 7 м является оптимальным интервалом для обнаружения различий между ЭПА.

6.2. ОШИБКА ИНТЕРПОЛЯЦИИ

К достоинствам кригинга следует отнести и то, что этот метод кроме несмещенной оценки с минимальной дисперсией дает также и величину этой дисперсии, а также - невысокую чувствительность точности кригинга к изменению объема выборки.

Webster и Oliver (1992) показали, что на точность предсказания при проведении интерполяции кригинг-методом влияют следующие источники неопределенности:

- почвенный покров является единственной доступной в природе реализация случайной функции;
- семивариограмма, построенная по выборочным данным, зависит от пробоотбора и аналитической ошибки;

- выбранная для описания семивариограммы модель и точность оценивания ее параметров также влияют на конечный результат.

6.3. ПРОВЕРКА ТОЧНОСТИ КАРТОГРАММ

Проверка точности карт обычно проводится по одному из следующих методов: по методу бутстреп методу, джекнайф или по методу наибольшей беспристрастности. Использование данных методов позволяет заменить полевые исследования имитационным моделированием.



Что такое бутстреп?



Бутстреп – это статистический метод для оценки распределения случайной величины, когда из исходной выборки отбираются достаточное число раз подвыборки с заменой (то есть подвыборки каждый раз возвращаются в исходную выборку). Чаще всего формируются подвыборки, составляющие 99%, 95% или 90% от исходной выборки.

В результате такой процедуры получают ошибку (или доверительный интервал) для параметров генеральной совокупности (среднего, медианы, коэффициента корреляции или коэффициента регрессии). Эта процедура также используется для проверки гипотез. Она может также использоваться для построения гипотез.

На рис. 17 показаны результаты бутстреп-тестирования, использованного для исследования картограмм распределения гумуса в пахотном горизонте на участке 4 га на пахотных дерново-подзолистых почвах (Московская область). Опираясь на результаты бутстреп тестирования можно утверждать, что случайно-стратифицированная схема опробования дает более устойчивые результаты, по сравнению с регулярной сеткой и эти результаты остаются устойчивыми при уменьшении выборок до 10%.

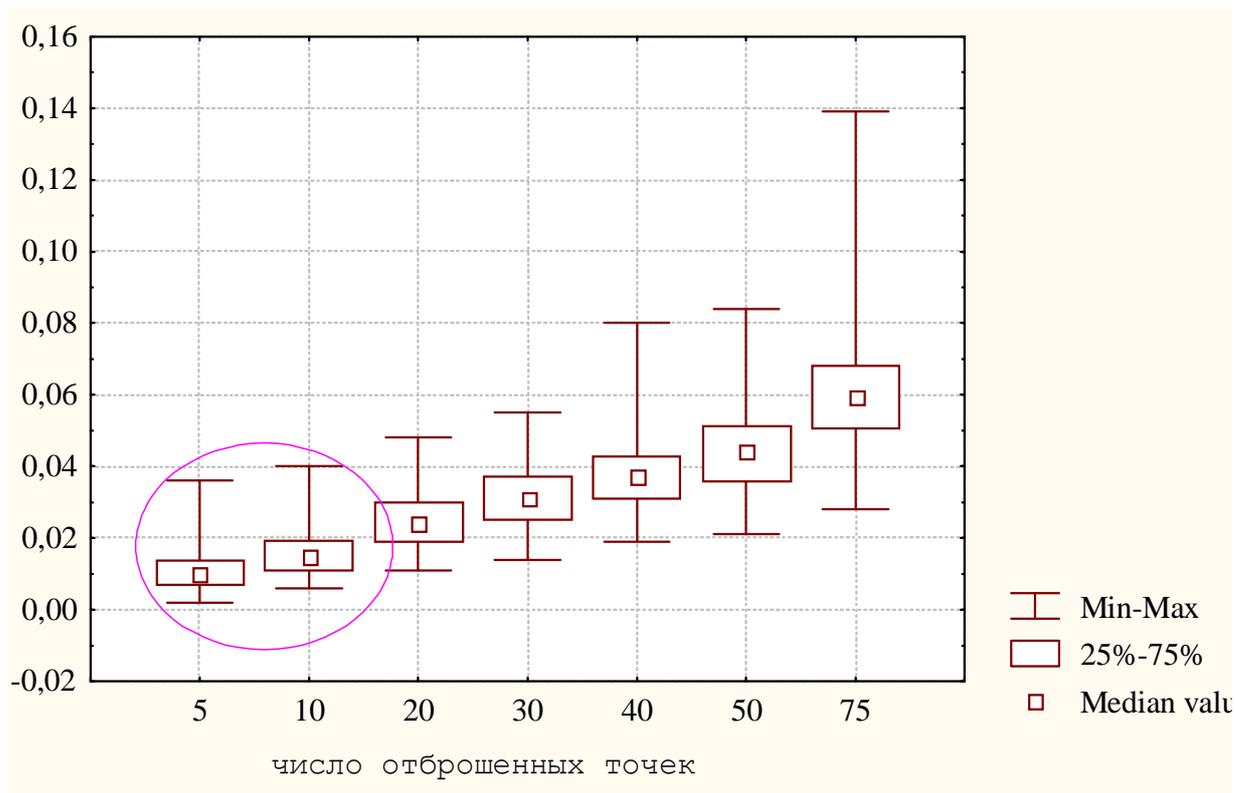


Рис. 17. Результаты бутстреп- тестирования (пояснения в тексте). По оси абсцисс показано число отбрасываемых точек, а по оси ординат – модуль невязки между картограммами, рассчитанными по сетке 1м²

Также бутстреп-тестирование показало, что более важным является объем выборки нежели реализация схемы опробования (идея схемы та же, но другая реализация).



Что собой представляет джекнайф?



Джекнайф – аналогичный бутстрепу статистический метод.

Основная идея метода джекнайфа заключается в систематическом перерасчете нужной статистики (среднего, медианы, коэффициента корреляции или коэффициента регрессии), выбрасывая из выборки случайным образом по одному наблюдению.

Таким образом, может быть получена несмещенная оценка статистики и ее ошибка. Оба метода оценки изменчивости статистика от изменчивости этой статистики между проб, а не из параметрических предположений.

Процедура джекнайф носит менее общий характер по сравнению с бутстрепом, Однако джекнайф проще применять для сложных схем пробоотбора, таких, например, как многоступенчатый отбор с различными весами. Джекнайф и бутстреп часто приводят к одинаковым результатам. Однако бутстреп может давать немного разные результаты для повторностях на одних и тех же данных, тогда как джекнайф дает точно такой же результат каждый раз (при условии, что подмножества выбираются из одной и той же выборки).



Как действует метод наибольшей беспристрастности?



Метод наибольшей беспристрастности – статистический метод для проверки прогнозных моделей. Подмножество данных делится на два поднабора: «обучающую» выборку и выборку, по которой производится проверка. Мерой точности прогноза считают среднюю оценку, полученную по результатам оценки каждого значения «проверяющей» выборки.

Один из вариантов метода наибольшей беспристрастности является аналогом метода джекнайф и состоит в том, что проверяющая выборка состоит из одного значения, и процедура проводится множество раз.

Другой вариант проверки, разделяет данные на k подмножеств; каждое из которых по очереди удаляется и используется для проверки. Такая процедура позволяет избежать псевдо-точности, когда происходит завышение точности результатов.

Еще в 70-х годах из-за большого объема вычислений эти три описанные процедуры были немислимы. Это машинно-ориентированные методы, существенно зависящие от развития вычислительной техники и внешних систем ЭВМ. Одним из важнейших достижений было само осознание воз-

возможности формулировать задачу имитации как задачу планирования эксперимента.

6.4. СРАВНЕНИЕ ИНТЕРПОЛЯЦИОННЫХ КАРТ

Часто требуется сравнить две или несколько карт разных лет, например, чтобы оценить потери гумуса или изменения местоположения контуров, сопоставить изоплеты влажности с мощностью второго гумусового горизонта, сделать пространственную оценку зон с повышенным значением (Губер, Архангельская, 2001).

И.Ф. Кузякова и Я.В. Кузяков (1997) сопоставили карты содержания органического углерода с топографической картой и, рассчитанной на ее основе, картой временных водно-эрозионных потоков, и сделали вывод о том, что существенное варьирование органического углерода обусловлено влиянием микрорельефа как фактора микроэрозии.

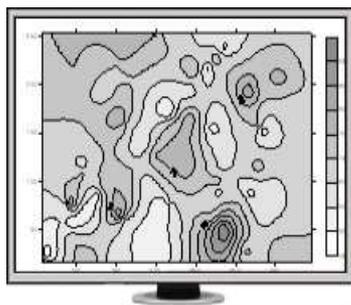
При сравнении карт и картограмм в большинстве своем используется **визуальное** сопоставление, которое носит качественный характер и субъективно. Одним из предложений количественного критерия для сравнения картограмм является работа Е.В. Жаровой и В.П. Самсоновой (2001), где они использовали таблицы сопряженности при сравнении карт для оценки пространственной устойчивости во времени контуров по рН и содержанию фосфора и калия.

Другим количественным методом сравнения картограмм является расчет коэффициента корреляции для двух матриц интерполированных значений по сетке с минимальным для данного исследования шагом (Готра, Мешалкина, 2003).



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Что собой представляет стратифицированный случайный отбор?**
- 2. Какой фактор влияет на стоимость и точность картограмм?**
- 3. Какие факторы влияют на точность предсказания методом кригинга?**
- 4. Что такое бутстреп-тестирование?**
- 5. Для чего используется джекнайф?**
- 6. Какими способами можно провести сравнение интерполяционных карт?**
- 7. Как действует метод наибольшей беспристрастности?**
- 8. Какой метод тестирования используется для прогнозных моделей?**



Модуль 7. Задачи, решаемые геостатистикой

Вы будете изучать

- Анализ пространственного варьирования
- Задачу интерполяции
- Прогнозирование
- Обоснование схем пробоотбора
- Исследование моделей

Цели модуля

- Познакомить с классификацией задач, решаемых геостатистикой

После изучения модуля Вы сможете

- Подобрать схему решения конкретной практической задачи, используя геостатистику.



Основная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Масштабные эффекты при исследовании почв. – М.: МГУ. 2001. – 400 с.
3. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия. 2004.
4. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Каневский М. Ф., Демьянов В. В., Савельева Е. А., Чернов С. Ю., Тимонин В.А. Кригинг и базовые модели геостатистики // В сб. Проблемы окружающей среды и природных ресурсов. М., 1999. № 11.
2. Козловский Ф.И. Теория и методы изучения почвенного покрова. - М.: ГЕОС, 2003 - 536с.
3. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001. N 11.
4. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
5. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85. no 2-3.
6. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML //Pedometron. Issue 18. July 2005.



Ключевые слова

Тип данных,
интерполяция,
прогнозирование,
исследование моделей

ВВЕДЕНИЕ

Ключом к повышению экономической эффективности является использование и обработка получаемой информации, позволяющие, например, выбрать лучший вид посевов для конкретного поля, или провести расчет количества удобрений и других вносимых в почву в течение года веществ, необходимого для получения максимального урожая.

Анализ данных по сельскохозяйственным предприятиям (фермерским хозяйствам, колхозам, совхозам...) может помочь в принятии верных решений и, за счет этого, получить дополнительный доход и снизить влияние неблагоприятных природных условий.

Природный ландшафт состоит из множества взаимосвязанных частей, и возможности ГИС и пространственного моделирования позволяют воссоздать этот сложный комплекс в виде менее сложной структуры компьютерной (модельной) среды. Слои данных с информацией о каком-то специфическом объекте или свойстве как раз и являются результатом подобного упрощения.

Тематические слои могут включать гидрологические свойства, характеристики почв, данные по уклонам местности и другие параметры, которые фермер или эколог рассматривают в качестве значимых при общем управлении земельной собственностью. После создания подобных слоев данных путь к повышению эффективности производства открывает пространственный анализ накопленной информации.

7.1. ОПИСАНИЕ И АНАЛИЗ ПРОСТРАНСТВЕННОГО ВАРИИРОВАНИЯ

Первая из задач, решаемых геостатистикой - это, естественно, описание и анализ пространственного варьирования, то есть количественное описание того, как свойство варьирует в данном контуре, районе, конкретном поле и т.д. Этот анализ обычно состоит в выявлении трендов, оценке параметров вариограммы, в том числе, оценке наггет-дисперсии.

Например, для пахотной серой лесной почвы со вторым гумусовых горизонтом (район г. Пущино Московской области) Н.Г. Гумматовым с соавторами (1992) для ряда физико-химических показателей была выявлена колебательная составляющая в варьировании почвенных свойств с периодом 30-50 м, что было объяснено в дальнейших исследованиях существованием древней блочно-полигональной структуры.

На рис. 18 приведены семивариограммы для подвижных фосфора (А) и калия (Б), а также суммы обменных оснований для аллювиальных луговых почв Быковского расширения (Московская область). Эти семивариограммы носят транзитивный характер, следовательно, для этих свойств отмечаются пространственные зависимости.

Для суммы обменных оснований «наггет-эффект» равен нулю следовательно варьирование на расстояниях, меньших шага, и аналитическое варьирование для этого свойства пренебрежимо мало. Для фосфора и калия составляет соответственно 63% и 20% от значения порога. Ранг для этих свойств составляет 220-270 м.

7.2. ИНТЕРПОЛЯЦИЯ

На втором месте стоит задача интерполяция или построения оптимального "предиктора" (в переводе "предсказателя"). Такая задача возникает, если необходимо предсказать значение в точке, где не было наблюдения; оценить точность данного предсказания; для некоторого участка вычислить среднее значение свойства, например, среднее количество нитратов, которое переходит в среднем в воды с данного участка.

Примерами может служить построение картограмм распределения веществ (в том числе, загрязнения) вдоль трансекты, на поле, слоя гумусового горизонта (рис. 19), трехмерного пласта полезного ископаемого.

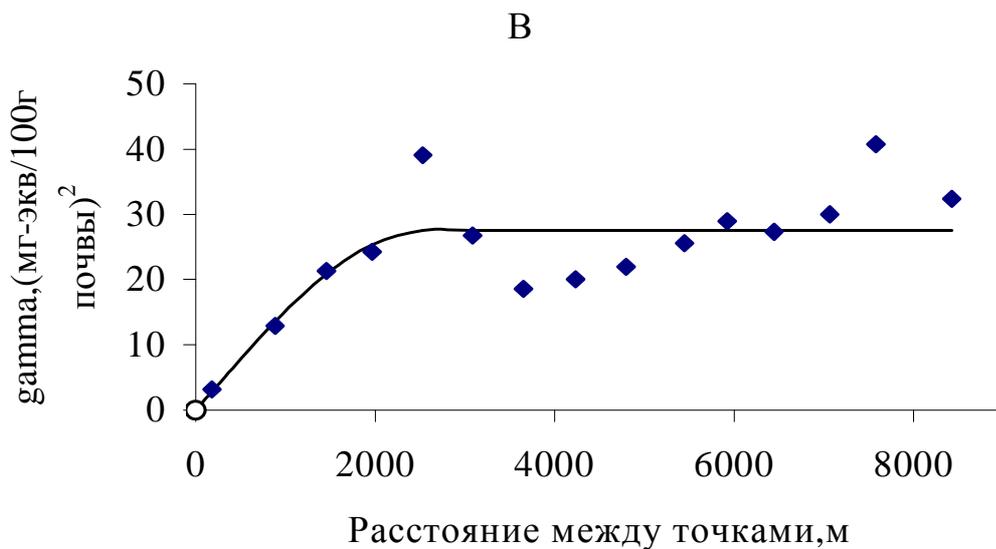
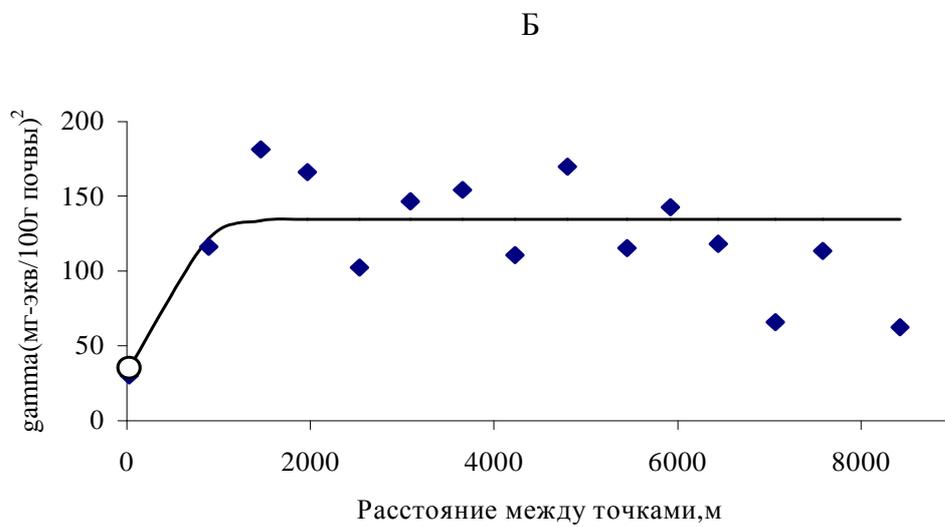
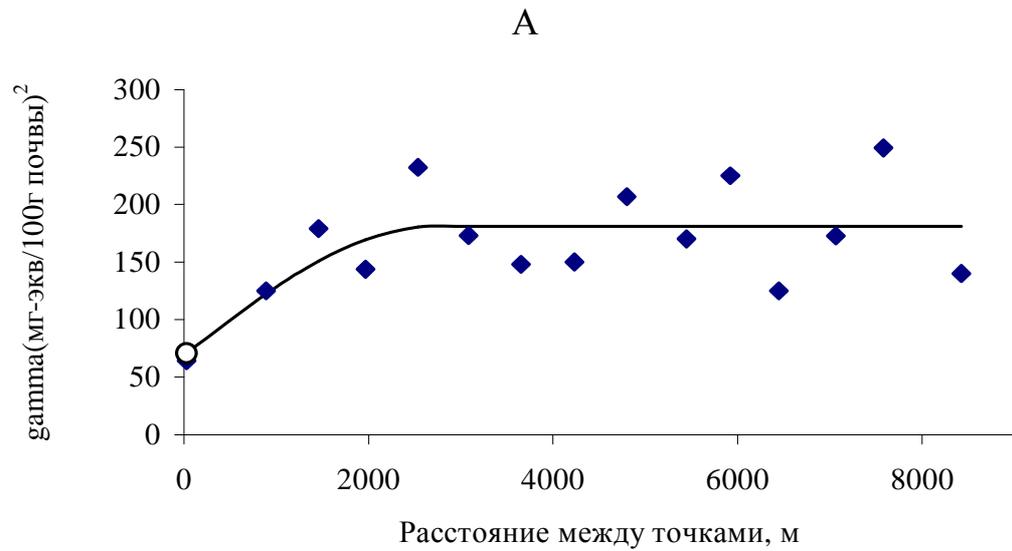


Рис. 18. Семивариограммы для подвижных фосфора (А), калия (Б) и суммы обменных оснований (В) для аллювиальных луговых почв Быковского расширения (Московская область)

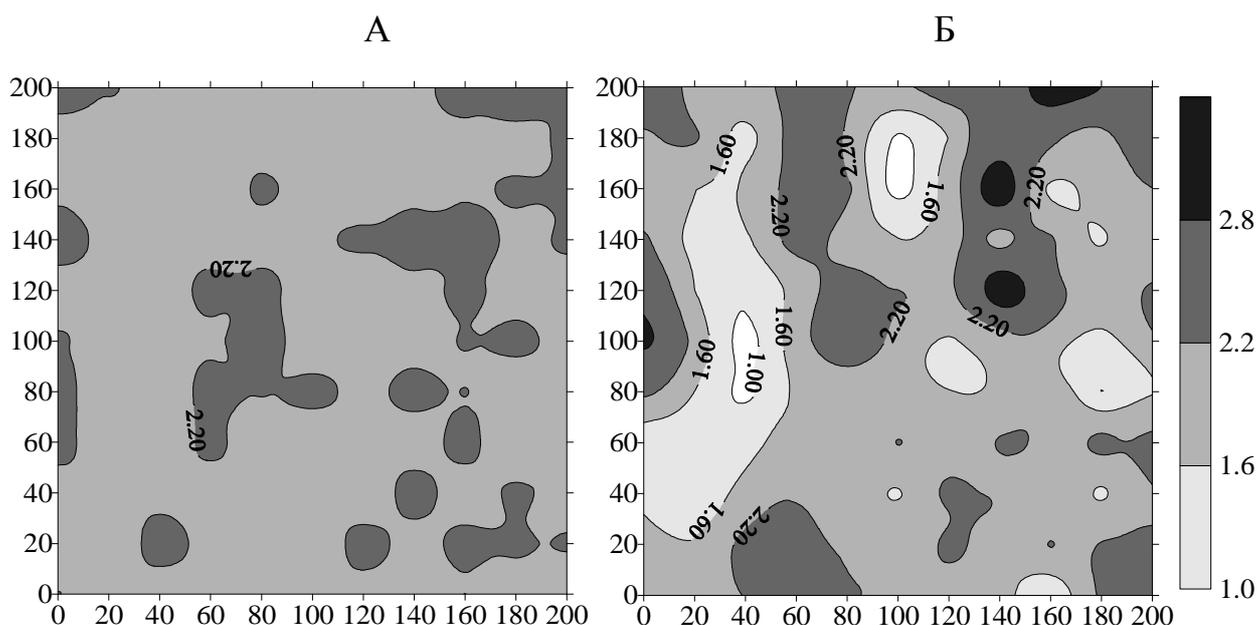


Рис. 19. Контурные карты содержания гумуса в верхней (а) и нижней (б) части пахотного слоя на участке 0,4 га для пахотных дерново-подзолистых почв (Московская область)

7.3. ПРОГНОЗИРОВАНИЕ

Третьей можно считать задачу прогнозирования. К этой задаче можно, например, отнести расчет вероятности превышения пороговых значений на данной территории для данного загрязнителя.

Другим примером будет решение вопроса: можно ли утверждать, что территория загрязнена с вероятностью 98%? С практической точки зрения такие карты вероятности могут быть использованы для принятия решений: например, для создания плана загрязненных зон или выделения областей с высоким потенциальным плодородием.

7.4. ОБОСНОВАНИЕ СХЕМ ПРОБОТБОРА

Четвертая задача связана с обоснованием схем пробоотбора. В ее рамках рассматриваются такие вопросы, как выбор достаточного числа наблюдений, расположение точек друг относительно друга в пространстве.

7.5. ИССЛЕДОВАНИЕ МОДЕЛЕЙ

Пятая группа задач может быть названа "исследование моделей". В современном почвоведении нашли большое распространение модели водо- и массопереноса. Для работы любой модели нужны данные. Входные данные должны быть четко описаны и корректно собраны, а выходные данные не могут быть точнее, чем данные на входе и/или чем качество модели.

Геостатистика решает в этом случае следующие вопросы:

1. Настройка параметров модели, то есть параметров, которые входят в модель, должны быть оценены для конкретного участка земли.
2. Пространственная интерполяция результатов исследований из точек наблюдения на весь участок.
3. Оптимизация модели. Это может быть расчет оптимального количества удобрений (нитратов, пестицидов) – посредством системного анализа 2-х противоположных целевых функций: увеличение урожая, с одной стороны, и сокращение вымывания удобрения (нитратов, пестицида) в грунтовые воды. На каком-то этапе увеличение нормы внесения увеличивает урожай и доход. С другой стороны, может возрасти количество удобрений (нитратов, пестицидов), вымываемых в грунтовые воды. Следовательно, возможен штраф, т.е. убыток.
4. Анализ чувствительности, то есть изучение того, как варьирование одного из входных параметров более влияет на результат.
5. Анализ пространственного варьирования результатов.
6. Анализ пространственно-временных отношений для данной модели.
7. Оценка необходимого числа наблюдений для получения вывод.

К этой же задаче относится и стохастическое моделирование - одно из самых активных областей исследования в геостатистике.



Что собой представляет стохастическое моделирование?



Стохастическое моделирование состоит в том, что генерируются многочисленные реализации, все из которых соответствуют данным и воспроизводят разновидности рисунков пространственной зависимости, или другие важные с точки зрения рассматриваемой проблемы статистики.

Некий сценарий (например, стратегия землепользования или рекультивация земель) может применяться к набору реализаций, позволяя оценить неопределенность отклика (эффективность рекультивации, почвенное плодородие, урожайность).

Все эти проблемы концентрируются при создании геоинформационных систем. Современные ГИС комбинируют информацию трех уровней: карты, модели и базы данных, содержащие подробные сведения о конкретных точках пространства. Составляющие каждого уровня реализованы в разных масштабах (поля, хозяйства, района, области, региона и так далее).

Особое место занимают случаи, когда рисунок пространственного варьирования видоизменяется во времени. Примером может служить моделирование загрязнения разового выброса. Оно начинается в момент времени t_0 , нарастает, достигает пика, затем снижается.

Промоделировать эти изменения, предсказать концентрации в определенных точках пространства в конкретный момент времени t - эти задачи могут быть решены в настоящее время как совокупность пространств, каждое из которых отвечает отмеченному в наблюдениях моменту времени.



В геостатистике варьирование во времени может трактоваться аналогично варьированию в пространстве. Но оно качественно отличается тем, что может быть определено только тем значением, которое было в этой точке в предыдущий момент.

Кроме того, варьирование во времени может иметь четко детерминированную составляющую: сезонную (суточную, годовую). Для решения этих задач в геостатистике развиваются специальные методы.



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Какие задачи в области почвоведения и экологии можно решить, используя методы геостатистики?**
- 2. Какие этапы анализа выполняются при исследовании моделей?**
- 3. Что понимается под термином «прогнозирование»?**
- 4. Зачем нужно обосновывать схемы пробоотбора?**
- 5. Что собой представляет стохастическое моделирование?**
- 6. Каким образом выполняется анализ пространственного варьирования?**
- 7. Как может исследоваться геостатистическими методами временная динамика почв и экосистем?**



Модуль 8. Сравнение кригинга с другими средствами интерполяции. Доступные программные средства.

Вы будете изучать:

- Сравнение кригинга с другими методами интерполяции
- Геостатистические компьютерные программы

Цели модуля:

- Сравнение кригинга другими методами интерполяции
- Дать краткий обзор существующих геостатистических компьютерных программ

После изучения модуля Вы сможете:

- Ориентироваться в компьютерных программах, связанных в той или иной степени с геостатистикой



Основная литература

1. Геоинформатика. Кн. 1 (под ред. В.С. Тикунова). – М.: Издательский центр «Академия», 2008. – 384 с.
2. Масштабные эффекты при исследовании почв. – М.: МГУ. 2001. – 400 с.
3. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия. 2004.
4. Самсонова В.П. Пространственная изменчивость почвенных свойств: на примере дерново-подзолистых почв. – М.: Изд-во ЛКИ, 2008. – 160 с.



Дополнительная литература

1. Каневский М. Ф., Демьянов В. В., Савельева Е. А., Чернов С. Ю., Тимонин В.А. Кригинг и базовые модели геостатистики // В сб. Проблемы окружающей среды и природных ресурсов. М., 1999. № 11.
2. Козловский Ф.И. Теория и методы изучения почвенного покрова. - М.: ГЕОС, 2003 - 536с.
3. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001. N 11.
4. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma. 1999, V. 89.
5. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // Geoderma. 1998, V. 85. no 2-3.
6. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML //Pedometron. Issue 18. July 2005.



Ключевые слова:

Скользящее среднее,
метод обратных расстояний,
сплайны,

ВВЕДЕНИЕ

Бурное развитие геостатистики за последние 20 лет напрямую связано с активным развитием, широким распространением и применением популярных пакетов ГИС и специализированных геостатистических программ.

8.1. СРАВНЕНИЯ КРИГИНГА С ДРУГИМИ МЕТОДАМИ ИНТЕРПОЛЯЦИИ

Помимо семейства методов кригинга существуют различные методы интерполяции: скользящее среднее, метод обратных расстояний, сплайны и т.д. С точки зрения теории приближений ни одному из них нельзя отдать предпочтение, так как все они (при соответствующем выборе параметров) имеют одинаковый порядок сходимости.

Рядом исследователей (Brus и др., 1996; Utset и др., 2000) было проведено практическое сравнение методов интерполяции **почвенных** свойств в ненаблюдаемые точки. Рассматривались методы: общего среднего, скользящего среднего, ближайшего соседа, обратных расстояний, сплайн-функция Лапласа и обычный точечный кригинг, а также комбинации из них.

Авторы показали, что различие между методами было сравнительно небольшое. В целом, не было найдено статистически значимого влияния метода на точность интерполяции. С.А. Gotway с соавторами (1996) и G.M. Laslett с соавторами (1987), сравнивая качество кригинга и метода обратных расстояний для построения карт распределения почвенных свойств, отметили, что оба метода дают хорошие результаты при построении карт. Это подтвердили и наши результаты (рис. 20).

Таким образом, можно предположить, что влияние методов интерполяции на точность интерполяции отдельных почвенных свойств несущественное. Однако, в случае редко расположенных опорных точек метод сплайн-интерполяции или кригинга дает лучшие результаты. Метод кригинга имеет

несомненное преимущество в том, что позволяет оценить точность предсказания в каждой точке.

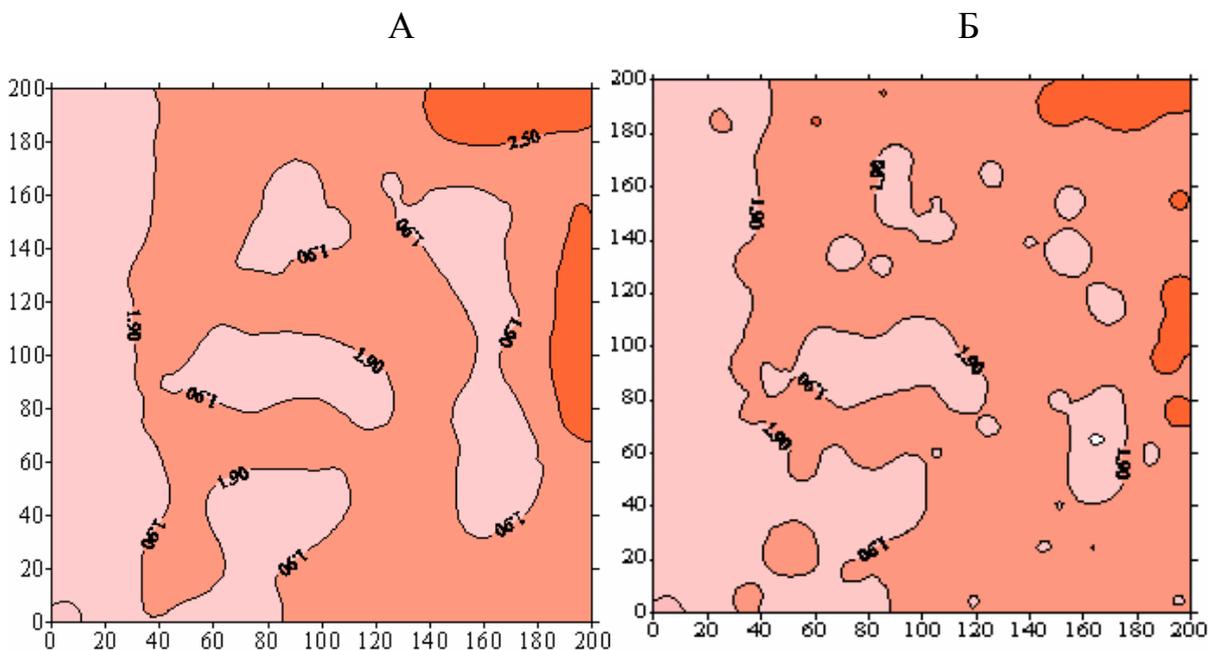


Рис. 20. Картограммы распределения гумуса на поле многолетних трав (с дерново-подзолистыми почвами), построенные по одним и тем же данным, но разными методами: А - методом кригинга, Б – методом обратных расстояний.

8.2. ГЕОСТАТИСТИЧЕСКИЕ КОМПЬЮТЕРНЫЕ ПРОГРАММЫ

Компьютерные программы, используемые для гестатистических расчетов, существенно отличаются друг от друга по количеству промежуточных расчетов и по наглядности представления результатов и удобству для пользователей. В настоящее время почти каждый коммерческий ГИС-пакет содержит геостатистический модуль.

Например, дополнительный модуль **Geostatistical Analyst** семейства программного обеспечения ArcGIS (фирма ESRI). Помимо развитых алгоритмов интерполяции, Geostatistical Analyst позволяет прогнозировать и рассчитывать статистическую вероятность наличия определенных свойств, явлений и объектов в пределах заданной области, отображать полученные результаты в виде интерактивно связанных между собой графиков и карт.

Этот модуль предоставляет полный набор инструментов пространственного анализа и обеспечивает решение двух задач: исследовательский анализ пространственных данных и построение поверхностей (реальных и прогнозных). В пакете реализованы разнообразные методы интерполяции (обратных взвешенных расстояний, выявления и оценки трендов, радиальной, подбораглобальных и локальных полиномов) и традиционные геостатистические методы (простой кригинг, универсальный кригинг, индикаторный кригинг, вероятностный кригинг, разделительный кригинг) для создания различных поверхностей.

Мощный картографический пакет для ученых и инженеров **Surfer** (фирма Golden Software Inc) – трехмерная программа вычерчивания поверхности карт, которая выполняется в среде Microsoft Windows. Она позволяет чертить контур, поверхность, каркас, вектор, изображение, заштрихованную область и карты, где отмечены исходные точки. Начиная с версии 6.0 и выше возможно для интерполяций задавать параметры вариограмм и использовать разные виды кригинга.

Программа **Vesper 1.0** Австралийского центра точного земледелия распространялась через Интернет бесплатно. В настоящее время продвинутые версии данной программы требуют регистрации и символической оплаты (50 \$ США). Программа позволяет подобрать модель семивариограммы и провести интерполяции различными методами кригинга.

Программы **GEOEAS** (Geostatistical Environmental Assessment Software – Геостатистическое программное обеспечение для оценивания рисков окружающей среде) и **GEOPACK** (Geostatistics foe Waste Management - Геостатистика для управления отходами) связанные друг с другом пакеты программ с дружественным интерфейсом. Включают различные виды кригинга и кокригинга. Распространяются бесплатно. Основная проблема, что они были разработаны под MSDOS. Краткий обзор программ, которые можно скачать бесплатно, приводится в таблице 2.

Таблица 2. Адреса в интернете, по которым можно получить геостатистические программы и обширную информацию по ним.

| Название программы (фирма) | Краткая характеристика | Адрес: http://www. |
|--|---|---|
| GS+ (Gamma Design Software) | DOS/Windows, вариограммы (изотр., анизотр.), кригинг (блочн., точечн.), джек-найф, удобная графика | http://www.gammadesign.com |
| GridstatPro (Applied Computer Engineering) | Windows/Unix, 3D, вариограммы (изотр., анизотр.), вариограммные карты | http://www.ace.inter.net |
| GViz (Correlations Company) | Windows, вариограммы (изотр., анизотр.), кригинг (блочн., точечн.), интерполяция методом ближайшего соседа, джек-найф, удобная графика | http://www.correlations.com |
| Isatis (Geovariances) | | http://www.geovariances.fr |
| Variowin | | http://www-sst.unil.ch/ |
| GSLIB (Geostatistical Software LIBrary) | Набор из 39 программных модулей, бесплатная версия под DOS и под Linux. Интерфейс не дружелюбный. Нужны навыки работы с программами на Фортране. Есть ограничения по объему выборки | http://www.staios.com/Quick/gslib.html |
| Geostatistical Toolbox (Roland Froidevaux) | | http://www-sst.unil.ch/ |
| Geo-EAS | | http://triton.cms.udel.edu/~vinton/gis_gip/packages/geoeas.html |
| Uncert (Department of Geology and Geological Engineering, Colorado School of Mines) | | http://uncert.mines.edu |

GSLIB – акроним для Geostatistical Software LIBrary (Библиотека геостатистических программ), разрабатываемых в Стэнфордском Университете последние 20 лет. Она содержит 39 модулей, охватывающей многие геостатистические приложения: моделирование 1, 2 и 3D, джекнайф и метод наибольшей беспристрастности, кокригинг, индикаторный кригинг, стохастическое моделирование. Поставляется с примерами данных и их интерпретаций. Первая версия GSLIB была написана Clayton Deutsch и André Journel еще в 1990 - 1992 гг. (www.gslib.com).



КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ

- 1. Какие Вы знаете программные продукты, с помощью которых можно проводить геостатистические расчеты?**
- 2. Что можно сказать о сравнении кригинга с другими методами интерполяции?**

Словарь терминов

Атрибутивная информация – информация, характеризующая качественные и количественные характеристики пространственного объекта, например таксационные описания для выделов, поквартальные итоги для кварталов, ГУЛ для лесхозов или лесничеств.

Бутстреп – это статистический метод для оценки распределения случайной величины, когда из исходной выборки отбираются достаточное число раз подвыборки с заменой (то есть подвыборки каждый раз возвращаются в исходную выборку). Используется для оценки точности карт.

Географические координаты – географическая широта и географическая долгота на земной сфере как аналоги геодезической широты и геодезической долготы в системе геодезических координат на земном эллипсоиде.

Геоданные (синонимы: пространственные данные, географические данные, геопропространственные данные) – данные об объектах и явлениях окружающей среды, требующие представления в координатно-временной форме, то есть для которых указаны координаты на земной поверхности и зафиксировано время, когда явление имело место.

Геостатистика – это совокупность прикладных статистических методов, которое исследует распределение объектов, явлений и процессов в географическом пространстве.

Джекнайф – это статистический метод для оценки распределения случайной величины, суть которого состоит в систематическом перерасчете нужной статистики (среднего, медианы, коэффициента корреляции или коэффициента регрессии), выбрасывая из выборки случайным образом по одному наблюдению.

Изотропная семивариограмма характеризует структуру варьирования показателя на всей территории опробования, независимо от направления.

Интерполяция, интерполирование – восстановление значений функции на заданном интервале по известным ее значениям в конечном множестве точек, принадлежащих этому интервалу.

Кригинг – интерполяционный метод, в котором значение в расчетной точке получается путем взвешенного усреднения по начальным (заданным) значениям в точках, находящихся в некоторой окрестности от расчетной точки. Вес вклада каждой точки определяется по вариограмме. Веса подбираются таким образом, чтобы оценка искомого значения была несмещенной, а дисперсия оценки – минимальной.

Кокригинг – многомерное развитие кригинга, когда в процесс предсказания почвенного свойства включается вторая переменная, которая легче измеряется или по которой имеются многочисленные данные.

Метаданные – это данные о данных, они включают сведения о составе данных, содержании, статусе, происхождении, местонахождении, качестве, форматах и формах представления, условиях доступа, приобретения и использования, авторских, имущественных и смежных с ними правах на данные и др.

Метод наибольшей беспристрастности – статистический метод для проверки прогнозных моделей. Подмножество данных делится на два подбора: «обучающую» выборку и выборку, по которой производится проверка. Мерой точности прогноза считают среднюю оценку, полученную по результатам оценки каждого значения «проверяющей» выборки.

Модель – описание объекта (предмета, процесса или явления) на каком-либо формализованном языке, составленное с целью изучения его свойств.

«Наггет-эффект» (или «эффект самородка») – значение к которому приближается функция семивариограммы, когда шаг (h) стремится к нулю. В него входит варьирование случайной величины на расстояниях меньших, чем шаг, использованный в модели, а также варьирование, связанное с аналитическими ошибками.

Порог – величина которую достигает транзитивная семивариограмма, когда выходит на «плато»

Пространственная информация – информация об объекте, для которого известно его положение в пространстве – его координаты (для точечного объекта), координаты точек поворота линий (для линейных объектов) или координаты образующих полигон границ (для площадных границ).

Прямоугольные координаты (декартовы) – прямоугольные координаты на плоскости или пространстве с взаимно-перпендикулярными осями.

Ранг – расстояние, при котором функция семивариограммы достигает постоянного значения (порога).

Регионализованный переменная – переменная, распределенная в пространстве, помимо случайной, "неорганизованной" компоненты варьирования, имеет закономерную составляющую варьирования, обладающую определенной пространственной структурой.

Регрессия – зависимость усредненных значений какой-либо величины, называемой зависимой или откликом, от другой величины или нескольких величин, называемых независимыми.

Семивариограмма, вариограмма характеризует зависимость среднего квадрата разности значений изучаемого показателя от расстояния между точками опробования, в которых эти значения измерены. Это инструмент для

оценки степени пространственной корреляции, имеющейся в данных, и основа для проведения кригинга.

Система глобального позиционирования – спутниковая навигационная система, состоящая из работающих в единой сети 24 спутников, находящихся на 6 орбитах высотой около 17 000 км над поверхностью Земли. Спутники постоянно движутся со скоростью около 3 км/сек, совершая два полных оборота вокруг планеты менее, чем за 24 часа. Позволяет в любом месте Земли, почти при любой погоде, также в космическом пространстве на расстоянии до 100 км от поверхности Земли, определить местоположение и скорость объектов.

Универсальный кригинг – комбинация модели многомерной линейной регрессии и обычного кригинга. Этот метод относится к, так называемым, методам «нестационарной геостатистики».

Факториальный кригинг – один из методов семейства кригинга, суть которого состоит в разложении семивариограммы на составляющие, связанные с разным масштабом (так называемые, факторы). Предсказание моделируется как сумма составляющих, соответствующих выделенным масштабам.

Цифровое моделирование рельефа - это система количественных методов моделирования и анализа земной поверхности и взаимосвязей между рельефом и другими естественными и антропогенными компонентами ландшафта.

Англо-русский словарь

| Английский термин | Перевод термина на русский язык |
|---|--|
| Block kriging | Блочный кригинг |
| Bootstrap | Бутстреп |
| Cartesian coordinates | Прямоугольные координаты (декартовы) |
| Cokriging | Кокригинг |
| Cross-validation method | Метод наибольшей беспристрастности |
| DEM – Digital Elevation Model | Цифровая модель рельефа |
| Disjunctive kriging | Разделительный кригинг |
| Intrinsic hypothesis | Внутренняя гипотеза |
| Jack-Knife | Джекнаиф (метод) |
| DTM- Digital Terrain Model | Модель рельефа местности, включающая морфометрические характеристики рельефа |
| Geographic(al) coordinates | Географические координаты |
| Geostatistics | Геостатистика |
| Global Positioning System - GPS или NAVSTAR | Система глобального позиционирования |
| Grid coordinates | Прямоугольные координаты (декартовы) |
| Feasibility study | Предпроектные исследования |

| | |
|--|---|
| Interpolation | Интерполяция, интерполирование |
| Kriging | Кригинг |
| Maximum likelihood estimation | Оценка максимального правдоподобия |
| Metaanalysis | Мета-анализ (метаанализ) |
| Multiple representation, multiscale representation | Множественные, или полимасштабные представления пространственных объектов |
| Multi-stage sampling | Многоступенчатый отбор |
| Nugget-effect | «наггет-эффект» (эффект самородка) |
| Random function | Случайная функция |
| Random variable | Случайная величина (переменная) |
| Range | Ранг |
| Regression | Регрессия |
| Rectangular coordinates | Прямоугольные координаты (декартовы) |
| Semivariogram | Семивариограмма, вариограмма |
| Sensativity analysis | Анализ чувствительности |
| Sill | Порог |
| Trend | Тренд |
| Universal kriging | Универсальный кригинг |
| Variability | Варьирование |
| Variance | Дисперсия |



СПИСОК ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ

1. Анализ данных в экологии сообществ и ландшафтов. Пер. с англ. под ред. А.Н. Гельфана, Н.М. Новиковой, М.Б. Шадринной. М.: РАСХН, 1999. - 306 с.
2. Васенев И.И. Почвенные сукцессии. М.: ЛКИ. 2008. 400 с.
3. Васенёв И.И., Руднев Н.И., Хахулин В.Г. Областные базы данных и геоинформационные системы // Информационно-справочные системы по оптимизации землепользования в условиях ЦЧЗ. Курск. 2002. С. 48-66.
4. Васенев И.И., Васенева Э.Г. Информационно-методическое обеспечение прецизионных систем земледелия в черноземной зоне // Информационно-справочные системы по оптимизации землепользования в условиях ЦЧЗ. Курск, 2002. С. 91–102.
5. Вентцель Е.С. Теория вероятностей. М.: Наука, 1969. - 383 с.
6. Возможности современных и будущих фундаментальных исследований в почвоведении. – М.: ГЕОС, 2000. - 131 с.
7. Готра О.Н., Мешалкина Ю.Л. Описание распределения гумуса в пределах одного поля на примере пахотного слоя дерново-подзолистой почвы // Вестник Московского университета, 2003, № 4. С. 3-8.
8. Гумматов Н.Г., Жиромский С.В., Мироненко Е.В., Пачепский Я.А., Щербиков Р.А. Геостатистический анализ пространственной изменчивости водоудерживающей способности серой лесной почвы // Почвоведение. 1992. № 6. С. 52-62.
9. Дмитриев Е.А. Теоретические и методологические проблемы почвоведения.- М.: ГЕОС. 2001. – 374 с.
10. Дэвис Дж.С. Статистический анализ данных в геологии. М.: Недра, 1990.
11. Жарова Е. В., Самсонова В. П. Устойчивость пространственных структур агрохимических свойств пахотной дерново-подзолистой почвы // Масштабные эффекты при исследовании почв. М.:МГУ, 2001. С. 229-234
12. Иванникова Л.А., Мироненко Е.В. Теория регионализированных переменных при исследовании пространственной вариабельности показателей агрохимических свойств почв //Почвоведение. 1988. № 5. С. 113-121.
13. Каневский М. Ф., Демьянов В. В., Савельева Е. А., Чернов С. Ю., Тимошин В. А. Кригинг и базовые модели геостатистики // Проблемы окружающей среды и природных ресурсов, М.,1999. № 11.

14. Козловский Ф.И. Почвенный индивидуум и методы его определения // Закономерности пространственного варьирования свойств почв и информационно-статистические методы их изучения. М.: Наука, 1970. С. 42-59.
15. Козловский Ф.И. Теория и методы изучения почвенного покрова. - М.: ГЕОС, 2003. – 536 с.
16. Кузякова И.Ф., Кузяков Я.В. Влияние микрорельефа на пространственное варьирование содержания гумуса в дерново-подзолистой почве длительного полевого опыта // Почвоведение. 1997. № 7.
17. Кузякова И.Ф.; Романенков В.А.; Кузяков Я.В. Применение метода геостатистики при обработке результатов почвенных и агрохимических исследований // Почвоведение, 2001. № 11.
18. Масштабные эффекты при исследовании почв.- М.: МГУ. 2001.
19. Матерон Ж. Основы прикладной геостатистики. - М.: Мир, 1968.
20. Пузаченко Ю.Г. Математические методы в экологических и географических исследованиях. – М.: Академия. 2004.
21. Самсонова В.П. Пространственная изменчивость почвенных свойств: На примере дерново-подзолистых почв. – М.: ЛКИ, 2008. – 160 с.
22. Самсонова В.П., Мешалкина Ю.Л., Дмитриев Е.А. Структуры пространственной вариабельности агрохимических свойств пахотной дерново-подзолистой почвы// Почвоведение. 1999ю №11.С.1359-1366
23. Bouma, J. The role of quantitative approaches in soil science when interacting with stakeholders // Geoderma 78, 1997.
24. Burgess T.M., Webster R. Optimal interpolation and isotithmic mapping of soil properties. II. Block kriging // J. Soil Sci. 1980. V.31. P. 333 - 341.
25. Burgess T.M., Webster R. Optimal interpolation and isotithmic mapping of soil properties. The semivariogram and punctual kriging // J.Soil Sci. 1980. V. 31. 315 - 331.
26. Burgess T.M., Webster R., McBratney A.B. Optimal interpolation and isotithmic mapping of soil properties. IV. Sampling strategy // J.Soil Sci. 1981. V.32. P. 643 - 659.
27. Cressie N. Statistics for spatial data.. N.Y. : John Wiley & Sons , 1991. 900 p.
28. Deutsch C.V., Journel A.G.1992. GSLIB. Geostatistical Software Library and User's Guide. Oxford University Press. 340p

29. Dobermann, A., Goovaerts P., Neue H.U. Scale-dependent correlations among soil properties in two tropical lowland rice fields// Soil Science Society of America Journal 1997, 61(5).
30. Englund E. and Sparks A., GEOstatistical Environmental Assessment Software. Environmental Protection Agency EPA-EMSL, Las Vegas USA
31. Gilberto C., Monteiro A.M., Fucks S.D., Carvalho M. Spatial Analysis and GIS: A Primer [Электронный ресурс] – Режим доступа: http://www.dpi.inpe.br/gilberto/tutorials/spatial_analysis/spatial_analysis_primer.pdf - Англ. – Редакция от 31.01. 2010.
32. Goovaerts, P. Geostatistics in soil science: state-of-the-art and perspectives // Geoderma 89, 1999.
33. Goovaerts, P. Study of relationships between two sets of variables using multivariate geostatistics // Geoderma 62, 1994.
34. Hoosbeek M. R. Incorporating scale into spatio-temporal variability : applications to soil quality and yield data //Geoderma. 1998, Vol. 85, no 2-3.
35. Isaaks,E.H. and R.M. Srivastava.1989.Applied geostatistics. Oxford Univ. Press.,New York, 1991.
36. Journel A.G., Huijbrects Ch.I. Mining geostatistics. L.: Academic Press, 1978. 498 p.
37. Krige D.G. A statistical approach to some basic mine problems on the Witwatersrand // J.Chem. Metall. Min.Soc.S. Afr. 1951. V.52. P.119-139.
38. Krige D.G. On the departure of ore values distributions from log-normal model in South African gold mines // J.South Afr. Inst. Min. Metall. 1960. V. 61. P. 231-233.
39. Lark, R.M., Webster, R. Separating random and deterministic spatial components in soil data with REML // Pedometron. Issue 18. July 2005.
40. Matheron G. Les variables regionalisees et leur estimation. Paris: Masson, 1965, 212 p.
41. Matheron, G. Pour une analyse Krigeante des donees regionalisees// Raport 732. 1982. Centre de Geostatistique, Fontainebleau.
42. McBratney A.B., Webster R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates // J.Soil Sci. 1986. V. 37. P. 617-639.

43. McBratney, A.B., Mendonca Santos, M.L., Minasny B. On digital soil mapping // *Geoderma* 117, 2003.
44. McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M. An overview of pedometric techniques for use in soil survey // *Geoderma* 97, 2000.
45. McBratney, A.B., Webster R. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. II. Program and examples // *Computers & Geosciences* 1981. N 7.
46. Oliver M.A., Webster R. How geostatistics can help you // *Soil use and management*. 1991. № 4. P. 206-217.
47. Oliver, M.A., Webster, R. Combining nested and linear sampling for determining the scale and form of spatial variation of regionalized variables // *Geographical Analysis* 1986. N 18.
48. Stein A., Brouwer J., Bouma J. Methods for comparing spatial variability patterns of millet yield and soil data // *Soil Sci. Soc. Am. J.* 1997. V.61.
49. Van Meirvenne, M., Cockx, L., Vitharana, U. Pedometrics in transition: from too few to too many data? // *Pedometrics 2005. Frontiers in Pedometrics. Bi-annual Meeting of Commission 1.5 Pedometrics, Division 1 of the International Union of Soil Science. September 12-14 2005. Naples Beach Hotel & Golf Club. Naples, Florida. USA.* 81.
50. Vasenev I.I. , Bojko O.S., Angombe S.T. Geoinformation methodical support for agroecological optimization of precision farming at the Chernozem zone // *Izvestia TSHA*. 2009. Spec. Vol. P. 41-49.
51. Webster R. The development of pedometrics // *Geoderma*.1994.V.62. P. 1-15.
52. Webster R., Oliver M.A. *Spatial dependence in statistical methods in soil and land resource survey*. Oxford Univ.Press, 1990. 307 p
53. Webster, R. Is soil variation random? // *Geoderma*, Vol. 97, 2000.
54. Webster,R., Atteia O., Dubois J.P. Coregionalization of trace metals in the soil in Swiss Jura// *Eur. J. Soil Sci.* 1994. N 45.

СОДЕРЖАНИЕ:



| | |
|--|-----------|
| МОДУЛЬ 1. ЧТО ТАКОЕ ГЕОСТАТИСТИКА И ЕЕ ОТЛИЧИЕ ОТ КЛАССИЧЕСКОЙ СТАТИСТИКИ..... | 2 |
| ВВЕДЕНИЕ | 4 |
| 1.1. ОПРЕДЕЛЕНИЕ ТЕРМИНА «ГЕОСТАТИСТИКА» | 4 |
| 1.2. ПОНЯТИЕ О ПРОСТРАНСТВЕННОЙ ВАРИАБЕЛЬНОСТИ | 8 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 9 |
| МОДУЛЬ 2. ГИПОТЕТИЧЕСКИЕ МОДЕЛИ ПРОСТРАНСТВЕННОГО ВАРИИРОВАНИЯ ПОЧВЕННЫХ ТЕЛ..... | 10 |
| ВВЕДЕНИЕ | 12 |
| 2.1. ХОРОПЛЕТЫ | 12 |
| 2.2. ТРЕНДЫ | 15 |
| 2.3. ПРОСТРАНСТВЕННАЯ АВТОКОРРЕЛЯЦИЯ | 17 |
| 2.4. ПРОСТЕЙШИЕ СПОСОБЫ ОТОБРАЖЕНИЯ ПРОСТРАНСТВЕННОГО ВАРИИРОВАНИЯ СВОЙСТВ..... | 19 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 19 |
| МОДУЛЬ 3. КОНЦЕПЦИЯ РЕГИОНАЛИЗИРОВАННОЙ ПЕРЕМЕННОЙ | 20 |
| ВВЕДЕНИЕ | 22 |
| 3.1. СЛУЧАЙНАЯ ФУНКЦИЯ..... | 22 |
| 3.2. ТЕОРИЯ РЕГИОНАЛИЗИРОВАННОЙ ПЕРЕМЕННОЙ. СТАЦИОНАРНОСТЬ И ВНУТРЕННЯЯ ГИПОТЕЗА..... | 24 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 28 |
| МОДУЛЬ 4. СЕМИВАРИОГРАММА | 29 |
| ВВЕДЕНИЕ | 31 |
| 4.1. ОПРЕДЕЛЕНИЕ СЕМИВАРИОГРАММЫ | 31 |
| 4.2. ПРИМЕР РАСЧЕТА ПРОСТЕЙШЕЙ ВАРИОГРАММЫ..... | 32 |
| 4.3. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ СЕМИВАРИОГРАММ: РАНГ, ПОРОГ, "НАГГЕТ-ЭФФЕКТ" | 35 |
| 4.4. ИЗОТРОПНАЯ ВАРИОГРАММА И ПРИНЦИП ЕЕ РАСЧЕТА В СОВРЕМЕННЫХ КОМПЬЮТЕРНЫХ ПРОГРАММАХ | 37 |
| 4.5. РАСЧЕТ АНИЗОТРОПНОЙ ВАРИОГРАММЫ..... | 39 |
| 4.6. НАИБОЛЕЕ РАСПРОСТРАНЕННЫЕ ФОРМЫ (МОДЕЛИ) ВАРИОГРАММ И ФАКТОРЫ, ВЛИЯЮЩИЕ НА ТОЧНОСТЬ МОДЕЛИ..... | 40 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 43 |

| | |
|---|-----------|
| МОДУЛЬ 5. КРИГИНГ И КОКРИГИНГ | 44 |
| ВВЕДЕНИЕ | 46 |
| 5.1. ЧТО ТАКОЕ КРИГИНГ?..... | 46 |
| 5.2. ПРИМЕР ПРОСТЕЙШИХ РАСЧЕТОВ..... | 48 |
| 5.3. ТОЧЕЧНЫЙ И БЛОЧНЫЙ КРИГИНГ | 50 |
| 5.4. ВИДЫ КРИГИНГА | 51 |
| 5.5. ПРОСТРАНСТВЕННАЯ ЗАВИСИМОСТЬ МЕЖДУ ПЕРЕМЕННЫМИ (КОВАРИОГРАММА) И КОКРИГИНГ | 53 |
| 5.6. УЧЕТ ВАРЬРОВАНИЯ ВО ВРЕМЕНИ | 55 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 56 |
| МОДУЛЬ 6. СИСТЕМЫ ПРОБООТБОРА, ОШИБКА ИНТЕРПОЛЯЦИИ И ОЦЕНКА ТОЧНОСТИ КАРТОГРАММ | 57 |
| ВВЕДЕНИЕ | 59 |
| 6.1. РОЛЬ СХЕМЫ ПРОБООТБОРА | 59 |
| 6.2. ОШИБКА ИНТЕРПОЛЯЦИИ..... | 62 |
| 6.3. ПРОВЕРКА ТОЧНОСТИ КАРТОГРАММ | 63 |
| 6.4. СРАВНЕНИЕ ИНТЕРПОЛЯЦИОННЫХ КАРТ | 66 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 67 |
| МОДУЛЬ 7. ЗАДАЧИ, РЕШАЕМЫЕ ГЕОСТАТИСТИКОЙ..... | 68 |
| ВВЕДЕНИЕ | 70 |
| 7.1. ОПИСАНИЕ И АНАЛИЗ ПРОСТРАНСТВЕННОГО ВАРЬИРОВАНИЯ | 70 |
| 7.2. ИНТЕРПОЛЯЦИЯ | 71 |
| 7.3. ПРОГНОЗИРОВАНИЕ..... | 73 |
| 7.4. ОБОСНОВАНИЕ СХЕМ ПРОБООТБОРА | 73 |
| 7.5. ИССЛЕДОВАНИЕ МОДЕЛЕЙ | 74 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 76 |
| МОДУЛЬ 8. СРАВНЕНИЕ КРИГИНГА С ДРУГИМИ СРЕДСТВАМИ ИНТЕРПОЛЯЦИИ. ДОСТУПНЫЕ ПРОГРАММНЫЕ СРЕДСТВА.... | 77 |
| ВВЕДЕНИЕ | 79 |
| 8.1. СРАВНЕНИЯ КРИГИНГА С ДРУГИМИ МЕТОДАМИ ИНТЕРПОЛЯЦИИ..... | 79 |
| 8.2. ГЕОСТАТИСТИЧЕСКИЕ КОМПЬЮТЕРНЫЕ ПРОГРАММЫ..... | 80 |
| КОНТРОЛЬНЫЕ ВОПРОСЫ И ЗАДАНИЯ | 83 |
| СЛОВАРЬ ТЕРМИНОВ..... | 84 |
| АНГЛО-РУССКИЙ СЛОВАРЬ..... | 88 |
| СПИСОК ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ..... | 90 |
| СОДЕРЖАНИЕ:..... | 94 |